THÈSE DE DOCTORAT
DE L'UNIVERSITÉ SORBONNE UNIVERSITÉ

Spécialité : Astrophysique

École doctorale n°127: Astronomie et Astrophysique d'Île-de-France

réalisée

à l'Institut d'Astrophysique de Paris

sous la direction de Yohan DUBOIS & Christophe PICHON

présentée par

Corentin CADIOU

pour obtenir le grade de :

DOCTEUR DE L'UNIVERSITÉ SORBONNE UNIVERSITÉ

Sujet de la thèse :

L'impact des grandes structures de l'Univers sur la formation des halos de matière noire et leur galaxie

soutenue le 26 Septembre 2019

devant le jury composé de :

Avishai DEKEL	Rapporteur
Simon WHITE	Rapporteur
Oliver HAHN	Examinateur
Dmitri POGOSYAN	Examinateur
Benoit SEMELIN	Examinateur
Adrianne SLYZ	Examinatrice
Christophe PICHON	Directeur de thèse
Yohan DUBOIS	Co-encadrant de thèse

Acknowledgements

Abstract

Résumé

À grande échelle, il est frappant de voir que la distribution anisotrope de la matière forme un large réseau de vides délimités par des murs qui, avec les filaments présents à leurs intersections, tissent la toile cosmique. La matière qui doit former plus tard les halos de matière noire et leurs galaxies afflue vers les nœuds compacts se situant à l'intersection des filaments et garde dans ce processus une empreinte/le souvenir de la toile cosmique.

Dans cette thèse, je développe une extension contrainte de la théorie de l'excursion dans son approximation dite "du franchissement vers le haut" pour prédire la masse, le taux d'accrétion et le temps de formation des halos de matière noire au voisinage des proto-filaments (qui sont identifiés comme des points-selles du potentiel). Les points-selles sont utilisés comme un référentiel local dans lequel l'évolution des propriétés physiques et morphologiques des galaxies est quantifiée aux grandes échelles / à grande échelle. À masse fixe, le modèle prédit que le taux d'accrétion et le temps de formation varient avec l'orientation et la distance au point-selle, confirmant que le biais d'assemblage est sensible aux forces de marées de la toile cosmique. Les halos peu massifs, s'étant formés tôt et "affamés" sont regroupés le long de l'axe principal des filaments, tandis que les halos plus massifs, plus jeunes sont répartis autour des nœuds. Les différents gradients observés pour différentes quantités, tels que la masse typique et le taux d'accrétion, ont pour origine l'anisotropie du point-selle et leur dépendance distincte aux moyennes et aux variances du champ. Pour les faibles décalages vers le rouge, ce modèle prédit qu'à masse fixe il y a un excès de galaxies rouges dans des directions préférentielles, comme l'ont montré des relevés spectroscopiques (GAMA) et photométriques (COSMOS), mais aussi les simulations hydrodynamiques (Horizon-AGN).

J'ai également calculé les taux de fusions par analyse multi-échelle des conditions initiales pour prédire l'assemblage anisotrope des halos et comprendre son impact sur la formation des galaxie. Outre les fusions de halos, j'ai aussi pris en compte les fusions de murs et de filaments qui ont un effet sur l'accrétion galactique et j'ai calculé leur propriétés à un et deux points en fonction du temps cosmique. J'ai établi le lien entre les taux de fusion et la connectivité. J'ai ensuite exploité ce lien pour estimer l'effet des structures à grande échelle sur le biais d'assemblage. Cette théorie décrit l'anisotropie de la toile cosmique, qui est un élément important pour décrire conjointement l'évolution de la physique et de la dynamique des galaxies dans leur environnement, en particulier les alignements intrinsèques ou les diversités morphologiques.

Afin d'étudier l'accrétion cosmique à de plus faibles échelles, j'ai implémenté une nouvelle méthode de particules traceuses dans le code à raffinement de grille adaptatif Ramses. Cette méthode est basée sur un échantillonnage de Monte-Carlo et est capable de reconstruire la trajectoire lagrangienne du gaz et son retraitement. Je démontre que la distribution spatiale des particules traceuses reproduit précisément celle du gaz, et je propose une extension capable de suivre tout le cycle des baryons dans leurs échanges avec les étoiles et les trous noirs. Cette approche est particulièrement adaptée aux problèmes astrophysiques qui requièrent simultanément d'avoir une résolution efficace des chocs avec un solveur de Godounov et de suivre l'histoire lagrangienne des baryons. Je l'utilise ensuite dans plusieurs simulations zoomées pour étudier l'acquisition du moment angulaire par les galaxies via leur accrétion bi-modale pour les hauts décalages vers le rouge. J'y observe que l'amplitude et l'orientation du moment angulaire du gaz froid sont conservées jusque dans le halo interne où le moment angulaire contribue à l'augmentation de la rotation des galaxies. Les couples de pressions sont plus importants en amplitude mais, de par leur turbulence, ils sont incohérents et sont dominés globalement par les couples gravitationnels. Les couples de la matière noire dominent dans le halo externe, ceux des étoiles dominent dans le disque.

Abstract

The strikingly anisotropic large-scale distribution of matter is made of an extended network of voids delimited by sheets, with filaments at their intersection which together form the cosmic web. Matter that will later form dark matter halos and their galaxies flows towards compact nodes at filaments' intersections and in the process, retains the imprint of the cosmic web.

In this thesis, I predict the mass, accretion rate, and formation time of dark matter halos near proto-filaments (identified as saddle points of the potential) using a conditional version of the excursion set theory in its so-called up-crossing approximation. The (filament-type) saddle points provides a local frame in which to quantify the induced physical and morphological evolution of statistical properties of galaxies on large scales. The model predicts that at fixed mass, mass accretion rate and formation time vary with orientation and distance from the saddle, demonstrating that assembly bias is indeed susceptible to the tides imposed by the cosmic web. Starved, early-forming halos of smaller mass lie preferentially along the main axis of filaments, while more massive and younger halos are found closer to the nodes. Distinct gradients for distinct quantities such as typical mass and accretion rate occur because the saddle condition is anisotropic, and because the statistics of these observables depend on both the means and their covariances. The signature of this model corresponds at low redshift to an excess of reddened galaxies at fixed mass along preferred directions, as recently reported in spectroscopic (GAMA) and photometric (COSMOS) surveys and in hydrodynamical simulations (Horizon-AGN).

I also compute the rate of merger events in the multi-scale initial conditions to forecast special events driving the anisotropic assembly of dark matter halos and understand their impact on galaxy formation. Beyond halo mergers, I consider all sets of mergers, including wall and filament mergers, as they impact the geometry of galactic infall. Their one- and two-points statistics are computed as a function of cosmic time. I establish the relation between merger rates and connectivity, which is then used to assess the impact the large scale structures on assembly bias. The anisotropy of the cosmic web, as encoded in this theory, is a significant ingredient to describe jointly the physics and dynamics of galaxies in their environment, e.g. in the context of intrinsic alignments or morphological diversity.

In order to explore the impact of cosmic infall on smaller scales I implemented a novel tracer particles algorithm in the Eulerian adaptive mesh refinement code Ramses. The tracer particles are based on a Monte Carlo approach and keep tracks of where fluid elements originate, so as to follow their Lagrangian trajectories and re-processing history. I show that they reproduce the gas distribution very accurately and I extend them to also trace the stars and black holes through the full cycle of baryons. These tracer particles are ideal to study complex astrophysical systems where both efficiency of shock-capturing Godunov schemes and a Lagrangian follow-up of the fluid are required simultaneously, in particular in cold flows. Thanks to this accurate tracer particle algorithm, the acquisition and loss of angular momentum of both cold and hot accretion flows onto galaxies at high redshift can be studied reliably. I find that the amplitude and orientation of the specific angular momentum of the cold gas is preserved down to the inner halo where the angular momentum contributes to the spin-up of galaxies, while for the hot gas it is lost at larger radii. Pressure torques, stronger in magnitude than gravitational torques are, however, spatially incoherent, which leads them to have no significant impact on the redistribution of angular momentum of the accretion flows. Gravitational torques, which dominate globally, are the main driver of the loss of angular momentum of the accretion flows in those halos, with dark matter gravitational torques dominating in the outer halo and stellar gravitational torques dominating in the disk.

Contents

1. Introduction

1.1   From a single galaxy to an expanding Universe

From the 18th century, most of the observed objects in the sky were thought to be located in a single entity — the Milky Way — whose borders were the borders of the Universe. Setting apart planets and stars, these objects were broadly classified as nebulæ, from the Latin word for cloud or fog, as they resemble diffuse clouds in sky. In 1771, Charles Messier published his "Catalogue des Nébuleuses et des Amas d'Étoiles" (Catalogue of Nebulæ and Star Clusters). Charles Messier first interest was in comets, but in order to observe them, he had to be able to distinguish moving objects from fixed objects in the sky, such as stars, star clusters and nebulæ. This led him to systematically compile a list of the objects in the sky that were impairing his observations. This catalogue, known as the Messier Catalogue is still today one of the most popular catalogues among amateur astronomers.

With the advent of better observations and the systematic classification of the objects, astronomers started distinguishing star clusters from diffuse nebulæ from spiral nebulæ. During the 18th and 19th centuries, many philosophers and mathematicians (E. Swedenborg, P.L. Maupertuis, T. Wright) speculated that the Milky Way is itself a "spiral nebula", made of a flattened disk of stars and that the spiral nebulæ are its analogues, but reside outside the Milky Way, while others argued that the spiral nebulæ were part of the Milky Way. This questioned not only the location of the Milky way and the spiral nebulæ in space, but also their relative sizes. However, observational evidences were missing to rule out any of the two models and it was not until the 20th century that it was finally shown that these nebulæ live outside of the Milky Way. One of the first proofs of the extra-galactic nature of the nebulæ can be attributed to Vesto Slipher. In 1912, he made spectrographic observations of the brightest spiral nebulæ ; all of them showed significant Doppler shifts, suggesting that the nebulæ are receding at velocities of hundreds to thousands of kilometres per seconds, much greater than the relative velocities of the stars of the Milky way. In 1917, observations of supernovæ in the Great Andromeda Nebula (now called the Andromeda galaxy) revealed that the supernovæ were 10 magnitudes fainter than supernovæ in the Milky Way, suggesting that they were much further away than the ones observed in the Milky Way. Using conservative assumptions, Shapley and Curtis, 1921 estimated that the Andromeda
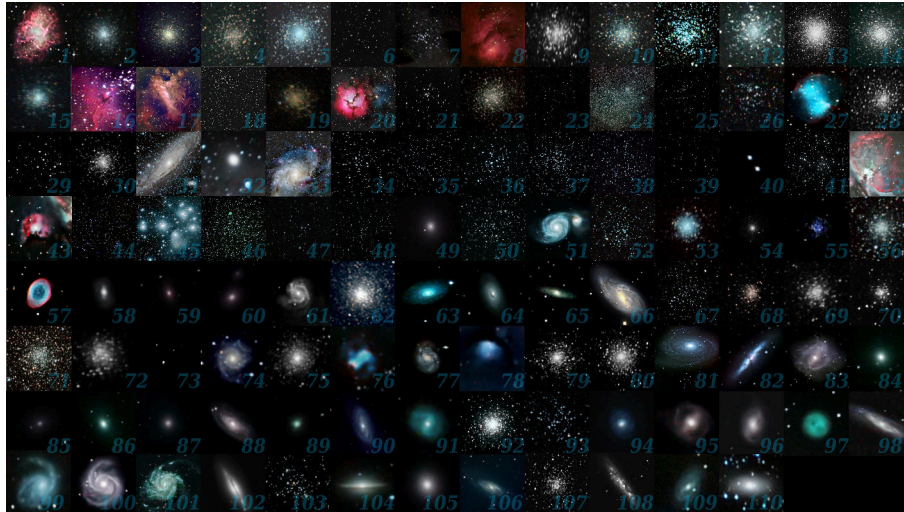
Figure 1.1.1: The 110 objects of the Messier catalogue, taken and compiled by an amateur astronomer. Credits: Michael A. Phillips.
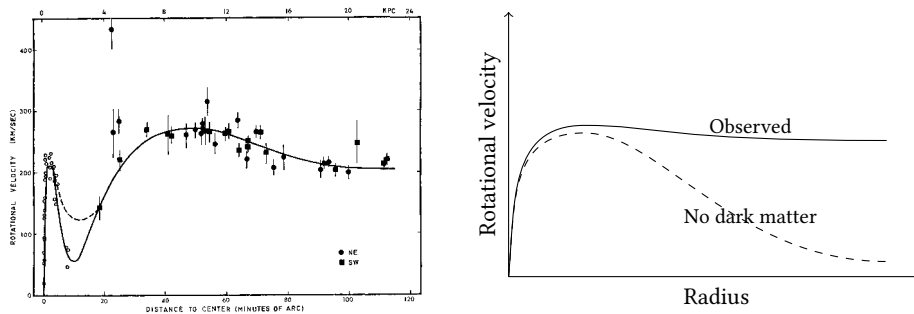
Figure 1.1.2: *Left:* Rotational profile of the Andromeda galaxy from Rubin and Ford, 1970. *Right:* Scheme of the rotational velocity profile. The expected rotational velocity profile with no dark matter is shown as a dashed line, the observed rotational profile is shown as a solid line. Without dark matter, the rotation profile decreases after some radii while the addition of dark matter makes the profile flatter at large radii, in agreement with observations.

Nebula could not be any closer than $20\,000\,\mathrm{ly}$, but still $7\,000\,\mathrm{ly}$ off the plane of the Milky way. This was further confirmed by the distance estimations of the nebulæ by Edwin Hubble, which definitely showed that nebulæ were too distant to be part of the Milky Way. We now call spiral (and elliptical) nebulæ "galaxies" from the greek words γάλα ("milk") and ξίας ("way").

In 1929, Hubble was able to show that galaxies were receding at increasing velocities with increasing distance, so that galaxies further away are receding faster, a relation now known as Hubble-Lemaître's law. The law states that the receding velocity is proportional to the distance times the Hubble constant H. Even though the measurements were largely inaccurate — Hubble's measurements gave $H = 500\,\mathrm{km/s/Mpc}$ while modern estimates are around $70\,\mathrm{km/s/Mpc}$ — the result showed that the Universe is not only made of multiple galaxies, but it is also expanding, paving the way to modern cosmology. At this point, the basic building blocks of the Universe were broadly found: the Universe is made of a multitude of individual galaxies, each of which has millions to several hundred billions stars, and sizes of the order of $10\,\mathrm{kly}$ to $100\,\mathrm{kly}$, while the distances between galaxies are of the order of the Mly.

The next step in our current understanding of the structure of galaxies and cosmology appeared

in the course of the 20th century. The discoveries of the distances and sizes of galaxies was shortly followed by estimations of their mass. In order to do so, a simple way is to use the laws of gravity: objects orbiting massive objects have smaller periods according to Kepler's laws. By measuring the velocities of objects gravitating in or around galaxies, one can infer their *gravitational* mass. Since galaxies are made of stars, one can also estimate the stellar mass from the galaxies' apparent luminosity. The ratio of gravitational mass required to explain the observed velocities to the observed stellar mass, known as the "mass-to-light" ratio was initially thought to be one. However, observations in the 1920s showed that the mass required to explain the motion of stars in the neighbourhood of the Sun is much larger than the observed one. Later in the 1930s, observations of the motion of galaxies in the Coma cluster led to the same conclusion that the gravitational mass should be much larger than the observed ones. These evidences were later confirmed when Rubin and Ford, 1970 showed that most of the mass of galaxies is not in stars. This was shown by measuring the rotational velocity of stars in the Andromeda galaxy. If the bulk of the mass of the galaxy was due to its stars, then the rotational velocity should increase from the centre to a radius of $10\,000\,\mathrm{ly}$, reaching a maximum of $\sim 200\,\mathrm{km/s}$ before decreasing. The observations however showed that the rotation curve rises as a function of radius before reaching a plateau at a radius of $10\,000\,\mathrm{ly}$ at about $250\,\mathrm{km/s}$, as shown on figure 1.1.2, left panel. This discovery, followed by multiple similar results, all confirmed that most of the mass in galaxies — and similarly in galaxy clusters — is hidden and is invisible. This matter that interacts *via* gravity but cannot be seen is now called Dark Matter (DM). In order to match observations, galaxies should be embedded in an extended DM halo, so that the decrease of stellar density with increasing radius is mitigated by the extended dark matter halo in the outskirts of the galaxy. This is schematically illustrated on figure 1.1.2, right panel. In addition to a correct distance ladder, astronomers now had access to a mass scale, albeit imprecise: in addition to the billion of stars that make galaxies, an extended and massive halo of dark matter surrounds each galaxy.

At about the same time other evidences for dark matter emerged with the discovery of the Cosmic Microwave Background (CMB) by Penzias and Wilson, 1965. This electromagnetic emission, emitted at the infancy of the Universe, shows that the Universe started in a quasi-homogeneous hot and dense state, with tiny density fluctuations of the order of the 10^{-4} to 10^{-5}. In a model missing dark matter, these initial density fluctuations would be too small for gravitational collapse to have time to pull matter together and form the observed large-scale structures of the Universe. This is a consequence of the interaction of baryonic matter with radiation: up to the emission of the CMB, gravitational collapse was prevented due to the radiative pressure of photons scattering from atoms to atoms. Dark matter provides a solution to the problem because it does not interact with light. Therefore, its density perturbations can grow first and create a potential well into which baryonic structures will later collapse. In addition to the CMB observations, several other observations such as gravitational lensing by galaxy clusters or the temperature distribution of hot gas in galaxies and clusters all pointed towards dark matter.

These discoveries, in conjunction with the development of general relativity led to the emergence of the standard model of cosmology, the Λ Cold Dark Matter (ΛCDM) model. The ΛCDM describes the evolution of the Universe after the CMB and is made of the following building bricks

1. the CMB is described by a Gaussian random field with known statistical properties,
2. the Universe is homogeneous and isotropic with no spatial curvature,
3. the Universe contains dark energy (Λ term), cold dark matter (CDM) in addition to ordinary matter and radiation.

According to the ΛCDM model, the Universe started from a hot dense state some $14\,\mathrm{Gyr}$ ago and has been expanding since then, as measured by Hubble, 1929 and is now in accelerated expansion, as a result of the non-null Λ term in Einstein's equations. About $85\,\%$ of the current matter of the Universe is DM, the remaining $15\,\%$ being ordinary baryonic matter (gas, stars,
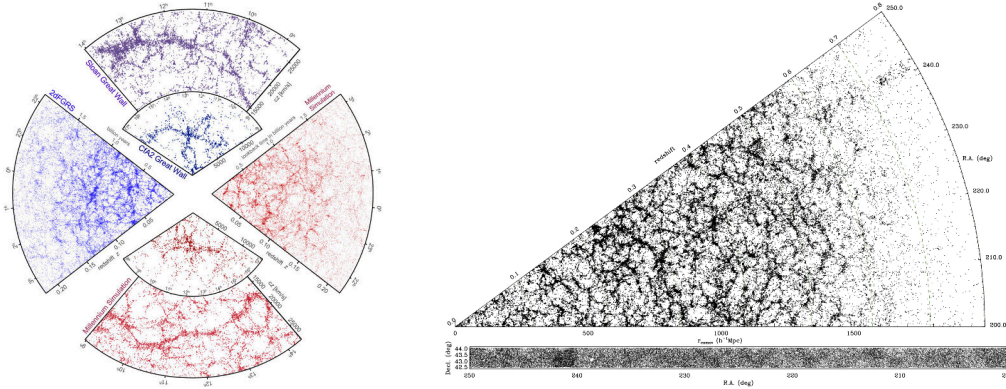
Figure 1.2.1: (*Left:*) The galaxy distribution obtained from spectroscopic redshift surveys (blue) and from mock catalogues (red) constructed from cosmological simulations. From Springel et al., 2006. (*Right:*) Catalogue of the spectroscopic HectoMAP survey in the local Universe (Hwang et al., 2016). The cosmic web made of large voids, filaments and dense nodes is clearly visible.

etc.). All this matter only adds up to 30% of the total energy density of the current Universe, the remaining 70% being dark energy. The success of the ΛCDM is well illustrated by the advent of the "precision cosmology" era, in which the parameters of the model can be fitted to observations down to percent levels using a variety of measurements, from CMB observations (Bennett et al., 2013; Planck Collaboration, 2018a), baryonic acoustic oscillations (e.g. Eisenstein et al., 2005; Moresco et al., 2016; Alam et al., 2017), type Ia supernovæ (e.g. Riess et al., 1998; Perlmutter et al., 1999; Abbott et al., 2019), weak lensing, cluster abundances (see e.g. Weinberg et al., 2013, and references therein for a detailed review).

1.2   The large scale structure of the Universe

Since the assumption of homogeneity clearly breaks down on small scales, as revealed by the presence of galaxies or stars, there must be certain homogeneities present at a certain time in history of the Universe. The homogeneities can be traced back in time to the CMB, but also to much larger scales, as can be seen in galaxy surveys that have revealed the existence of superstructures (cluster of galaxies, super-clusters and filaments and walls connecting them) on scales up to a few tens to hundreds of Mpc[1], as can be seen on figure 1.2.1. While each of these structures, from galaxies to super clusters or filaments, is unique in its morphology and mass, their overall statistical properties are homogeneous: the probability of any configuration is independent on the spatial location. Recent surveys, like the Sloan Digital Sky Survey (Abazajian et al., 2003), the 2MASS redshift survey (Huchra et al., 2012) or HectoMap (Hwang et al., 2016) have improved significantly our knowledge of the galaxy distribution showing with no doubt that galaxies form a complex web-like network on large scales made of voids, walls and filaments that interconnect with clusters of galaxies. This pattern is known as the *cosmic web*.

Due to the laws of gravity, the initial tiny fluctuations evolved into large and complex anisotropic structures that shape the current Universe. At scales of up to a few tens of Mpc, large under-dense regions called voids are found (Pan et al., 2012). Put together, the voids form a foam-like structure where each bubble is bound by denser walls or pancakes, sometimes called Zel'dovich pancakes (Zeldovich, 1970). The initial motion of particles can be well approximated in their linear regime by a rectilinear trajectory where the direction is set by the initial tides.

---

[1]The pc length unit is commonly used in astronomy, where $1\,\mathrm{pc} \approx 3.08 \times 10^{16}\,\mathrm{m} \approx 3.3\,\mathrm{ly}$.

Similar to parallel light rays bent by a disturbed water surface, the particles will travel until they form caustics. The first caustics to form, resulting from the collapse of matter along one direction, are bi-dimensional in nature. Following this first collapse, a second direction may collapse to form secondary caustics, resulting in uni-dimensional filamentary structures. Finally, filamentary structures may also collapse to form "knots" or nodes of the cosmic web.

As the cosmic web builds up, dense and spheroidal regions will undergo spherical collapse, resulting in the formation of dark matter halos. These primordial halos will later merge to form larger halos that in turn will also merge. This continuous accretion and successive merger scenario is often referred to as hierarchical formation. In classical models, galaxies form in (sub-)halos (Kauffmann et al., 1993) themselves located in the cosmic web. The distribution of galaxies in the Universe follows that of the large-scale structures as most of them are found in walls, filaments and nodes of the cosmic web. Recent developments have also shown that not only does the spatial distribution of galaxies relate to the cosmic web, but also some of their properties such as the orientation of their spin or their colour.

1.3 From the cosmic web to galaxy formation

According to the classical galaxy formation paradigm, gas falling on a proto-galaxy heats up to the Virial temperature of its host halo when crossing the virial radius (Rees and Ostriker, 1977; Silk, 1977). In this scenario, the gas acquires the same angular momentum distribution as dark matter before turning around and flowing towards the proto-galaxy, which has been confirmed by hydrodynamic simulations that do not describe cooling (e.g. van den Bosch et al., 2002). This process of angular momentum acquisition, at the core of the understanding of the formation of disk galaxies at high redshift, is well explained by the Tidal Torque Theory (TTT, Peebles, 1969; Doroshkevich, 1973; S. D. M. White, 1984). It predicts that the angular momentum of the dark matter increases under the effect of the gravitational torques of the cosmic web before dark matter decouples from the expansion of the Universe. In the classic scenario, the gas undergoes the same tidal field before decoupling but loses all dynamical and causal connection with the large scale structures at the Virial radius. Following this idea, classical models of galaxy formation typically ignore any explicit coupling of the baryons to their large scale environment, so that galaxy properties are directly inherited from the mass history of their host halo or some quantities averaged over all angles.

It has been established that the clustering of dark matter halos, as measured by halo bias, not only depends on halo mass but also on other halo properties such as formation time, concentration, spin and ellipticity (Gao et al., 2005; Wechsler et al., 2006; Gao and S. D. M. White, 2007; Hahn et al., 2007). This problem, commonly referred to as the "assembly-bias problem" can be rephrased as follows: the clustering of dark matter halos and their properties are correlated, beyond a mere mass and density relation. On large scales, systematic galactic surveys (de Lapparent et al., 1986; Geller and Huchra, 1989; Colless et al., 2001; Tegmark et al., 2004; Abazajian et al., 2003) have revealed that the Universe is structured around voids, sheets, filaments and knots that form the cosmic web. Using a different approach, a growing number of evidence have since showed that some halo and galaxy properties present distinct features at different locations in the cosmic web. As presented in Kraljic et al., 2018, void galaxies are found to be less massive, bluer and more compact than galaxies outside of voids (Rojas et al., 2004; Beygu et al., 2016); galaxies infalling into clusters along filaments show signs boosted star formation rate even before becoming part of the clusters while those infalling from the voids do not (Porter et al., 2008; Martínez et al., 2016); Kleiner et al., 2017 find systematically higher HI fractions for massive galaxies near filaments compared to field population, interpreted as evidence for a more efficient cold gas accretion from the intergalactic medium. A small but significant trend in the distribution of galaxy properties within filaments was reported in the spectroscopic surveys VIPERS (Malavasi et al., 2017) and

GAMA (Kraljic et al., 2018) and with photometric redshifts in the COSMOS field (Laigle et al., 2018). When corrected for large-scale density effects, these studies find significant mass and type segregations, where the most massive or quiescent galaxies are closer to filaments than less massive or active galaxies, emphasizing that large-scale cosmic flows play a role in shaping galaxy properties. On the other hand, other works reported that the most important driver of galaxy properties is stellar mass, as opposed to environment (Robotham et al., 2013; Alpaslan et al., 2015; Alpaslan et al., 2016), while the environment may impact the efficiency of galaxy formation (Guo et al., 2015; Eardley et al., 2015).

On large scales, the Tidal Torque Theory (TTT) naturally connects the distribution of matter to the angular momentum of halos (see section 2.1.6, e.g. Lee and Pen, 2001; Hirata and Seljak, 2004) in its recently revisited, conditioned formulation (Codis et al., 2015), with low-mass galaxies being preferentially aligned with filament's direction while more massive ones have their spin perpendicular to it. While it is far from obvious that the alignment of halo spin implies that the galactic spin are also aligned (Tenneti et al., 2015; Chisari et al., 2017, e.g.), the effect has also been confirmed for galaxies in numerical simulations (Dubois et al., 2014; Welker et al., 2014; Martizzi et al., 2019) and recently observationally (e.g. Trujillo et al., 2006; Lee and Erdogdu, 2007; Paz et al., 2008; Tempel et al., 2013; Tempel and Libeskind, 2013; Pahwa et al., 2016, see also for B. J. T. Jones et al., 2010; Cervantes-Sodi et al., 2010; Andrae and Jahnke, 2011; Goh et al., 2019 for contradictory results).

Classical models have proven quite successful in explaining many observed properties of galaxies, *via* the so-called halo model (see Cooray and R. Sheth, 2002, for a detailed review), in particular against isotropic statistics such as the two-point correlation function, yet they fail to capture some galactic properties, such as spin alignments, which are specifically driven by scale-coupling to the cosmic web (Codis et al., 2015), nor do they fully take into account how a given galaxy is gravitationally sensitive to the larger scales anisotropies. Indeed, when gas cooling is accounted for, it has been shown that a substantial part of the baryon mass and angular momentum is acquired *via* cold filamentary flows (Birnboim and Dekel, 2003; Ocvirk et al., 2008; Dekel et al., 2009; Kereš et al., 2009), feeding the galaxy in a highly anisotropic way. Unlike shock-heated gas, cold flows are able to penetrate halos to reach their innermost regions, feeding galaxies with pristine fuel for star formation. A three-dimensional visualization of galactic formation processes at intermediate scales (made possible by the joint use of tracer particles for the cold gas phase, and well-resolved zoom simulations) reveals that these gaseous flows stem from the cosmic web. In fact, the spatial distribution of caustics (the geometric location of the dark matter shell crossing and the isothermal shock of cold gas) provides us with direct information on the dynamical state of the gas likely to be accreted on the proto-galaxy: in this scenario, the gas first flows towards the caustics created by the dark matter to form wall-like structures, in which galaxies are embedded (Danovich et al., 2012). The gas then radiatively cools and looses its velocity component in the direction perpendicular to the walls to condense at the centre of dark matter filaments found at the intersection of walls. In the process, the gas retains a net transverse motion that sets the direction and amplitude of its angular momentum which will later be fed coherently into growing proto-galaxies. Doing so, it retains its angular momentum — and hence its causal connection to the cosmic web — until it reaches the innermost part of the galaxy (Pichon et al., 2011; Danovich et al., 2015), providing a unique testbed to assess the effect of the cosmic web on the formation of galaxies.

With the advent of large spectroscopic surveys (GAMA, Driver et al., 2011; VIPERS, Guzzo et al., 2014) and cosmological simulations (Illustris, Vogelsberger et al., 2014; Horizon-AGN Dubois et al., 2014; Dubois et al., 2016; Eagle, Schaye et al., 2015; Massive-Black II, Khandai et al., 2015), astronomers can now explore time modulations of the galactic properties with statistically meaningful data, but also their spatial modulations in the frame of the cosmic web (e.g. Alpaslan

et al., 2016; Malavasi et al., 2017; Laigle et al., 2018; Kraljic et al., 2019). There is now a dire need for both new methods and models to understand the coupling between the anisotropic cosmic web and the baryonic physics of galaxies. In particular one needs to build new estimators to quantify the spatial modulation of galaxy properties beyond isotropic two-point correlation functions on top of the classical halo model. As the effect of the cosmic web is expected to be second-order (after mass and local density dependence), new estimators and models have to be built that take into account explicitly the anisotropy of the cosmic web to prevent the signal from being lost when averaging over all possible angles.

The aim of this dissertation is to provide such estimators and models, with a novel framework devoted to the study of the effect of anisotropic features on the formation of dark matter halos and their galaxies. The approach followed in my work is two-fold: I study the effect of the cosmic web on large and small scales on the assembly of dark matter halos across cosmic time using conditional excursion set and critical set theory and use numerical simulations to unveil how these effects impact galaxy formation.

Chapter 2 describes the context on which this dissertation is based, presenting the different models and tools used in the course of my work. Chapter 3 presents an extension of the excursion set theory and predicts the accretion rate, formation time and typical mass of dark matter halos as a function of their environment. Chapter 4 presents a framework based on the peak theory to quantity the environmental effects acting on halo formation. In particular, it aims to provide a comprehensive description of the major events relevant to the assembly of galaxies. Chapter 5 presents a new numerical scheme able to accurately track the cosmic accretion in mesh-based hydrodynamical simulations. Chapter 6 presents results obtained from a numerical study of how angular momentum is acquired from the cosmic web and transported towards galaxies *via* cold flows. Chapter 7 wraps things up and discusses perspectives.

Disclaimer

The results presented in chapter 3 have led to a publication in MNRAS (Musso, Cadiou et al., 2018). I have produced all the results of the paper in collaboration with M. Musso, with contributions from C. Pichon, S. Codis, K. Kraljic and Y. Dubois.

The results presented in chapter 4 have been obtained in collaboration with C. Pichon and S. Codis, with contributions from D. Pogosyan, Y. Dubois and M. Musso.

The results presented in chapter 5 have led to a publication in A&A (Cadiou et al., 2019). I have produced all the results of the paper, with contributions from Y. Dubois and C. Pichon.

I have produced all the results of chapter 6, with contributions from Y. Dubois and C. Pichon.

I have read and contributed to to all publications presented in appendix B, albeit not as the main author. More specifically, I contributed to the theoretical sections of Kraljic et al., 2018; Kraljic et al., 2019. I contributed to the numerical setup of Beckmann et al., submitted by providing the tracer particle code.

To the best of my knowledge, all the results presented in the dissertation are original.

2. Context

In this chapter, I will present the different models used throughout my work. In section 2.1, I will present the cosmological context in which galaxy form, before describing the processes involved in galaxy formation in section 2.2.

2.1 Cosmology

In this section, I detail the different models that describe the initial conditions of the Universe in section 2.1.1. In section 2.1.2, I detail how these initial fluctuations grow to form the large-scale structure of the Universe and dark matter halos. In sections 2.1.3 to 2.1.5, I present models suited to study the statistical properties of dark matter halos and detail how they then acquire angular momentum in in section 2.1.6. Finally, in section 2.1.7, I provide some tools to describe the initial conditions of the Universe.
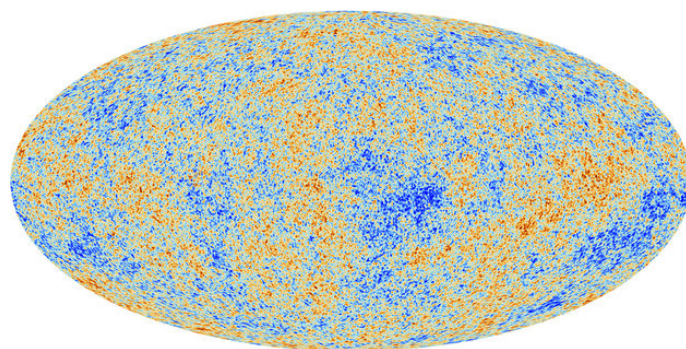
Figure 2.1.1: Map of the CMB as observed by the Planck satellite in 2013. Credit: ESA, Planck Collaboration.

2.1.1 Properties of the initial conditions of the Universe

In the standard model of cosmology, the time and space frame of the Universe was created at the Big Bang. Between $10^{-36}\,\mathrm{s}$ and $10^{-33}\,\mathrm{s} \sim 10^{-32}\,\mathrm{s}$, the Universe experienced a phase of exponential growth known as inflation. Quantum fluctuations in the microscopic scales were quickly expanded to cosmological scales to yield a flat, statistically homogeneous and isotropic Universe. After the end of inflation, the hot initial plasma cooled until light and matter decoupled at about $3\,000\,\mathrm{K}$, a moment known as the Last Scattering Surface (LSS). Photons emitted from the LSS were able to travel freely through space and experienced only the expansion of the Universe since then. Today, they can be observed in the microwave range of the light spectrum — as they have been emitted at $z \sim 1100$ — and form the well-studied CMB illustrated on figure 2.1.1.

The CMB is as-of-today the best example of a black body spectrum with a temperature of $2.726\,\mathrm{K}$. It is characterised by very small temperature fluctuations of about $0.0013\,\mathrm{K}$. These fluctuations can theoretically have two forms. Isocurvature perturbations have the property that the total energy density is constant in space so that the sum of the fractional variation of each component compensates exactly. An increase of $1\,\%$ of any component is compensated by a net *decrease* of $1\,\%$ of other components. Cosmic strings are commonly associated to isocurvature perturbations.

On the other hand, for adiabatic perturbations, the fractional variation of each component of the matter (baryons, photons, DM, neutrinos) are the same so that an excess of $1\,\%$ of photons results in an excess of $1\,\%$ of baryons. This is the model favoured by cosmic inflation. In the following, we will assume that the initial perturbations are adiabatic perturbations. Under this assumption, regions that are hotter are also denser ones so that the CMB is therefore also an observation of the density fluctuations of the Universe at $z \sim 1100$. The fluctuations in the initial density field are very well described by a homogeneity Gaussian Random Field. The mathematical properties of such fields are described in section 2.1.1.1.

The evolution of the Universe after inflation is well described by the ΛCDM model, as already mentioned in the introduction. The ΛCDM is made of cold DM and a cosmological constant, Λ, entering Einstein's equation, resulting in an expanding Universe. At low redshifts (later times), the expansion becomes accelerated once the density of the Universe is Λ-dominated. The model is described by six parameters: the baryon density Ω_b, the dark matter density Ω_c, the age of the Universe t_0, the spectral index n_s, the normalization of the amplitude of the primordial fluctuations Δ_R^2 and the reionisation optical depth. From these parameters, one can derive the Hubble constant H_0, the total matter density $\Omega_\mathrm{m} = \Omega_\mathrm{c} + \Omega_\mathrm{b}$, the root-mean-square of the field linearly evolved at $z = 0$ and smoothed with a Top-Hat filter of size $8\,\mathrm{Mpc}/h$, σ_8. The values of the six parameters are now measured from observations of the CMB (WMAP, Planck Collaboration, 2018a), as well as many other independent observations (see Weinberg et al., 2013, for a detailed review). The best-fit values from CMB observations are reported in Table 2.1. They show that today's Universe is in accelerated expansion, with 70% of the energy density in the form of dark energy (Λ term), 25% as DM and only 5% as baryonic matter while other particles — such as photons or neutrinos — make a negligible contribution.

In the following, we will particularly focus on the density contrast δ

$$\delta \equiv \frac{\rho - \bar{\rho}}{\bar{\rho}}. \tag{2.1}$$

Here $\bar{\rho}$ is the mean density of the Universe and ρ is the local density. This field is well represented by a Gaussian random with zero mean.

Table 2.1: Planck collaboration best-fit cosmological parameters. See Planck Collaboration, 2018a for more details.

Parameter	Comment	Value	Unit
n_s	Scalar spectral index	0.9667(40)	
H_0	Hubble constant	67.74(46)	$\mathrm{km\,s^{-1}\,Mpc^{-1}}$
Ω_b	Baryon density	0.0486(10)	
Ω_c	Dark matter density	0.2589(57)	
Ω_m	Matter density	0.3089(62)	
Ω_Λ	Dark energy density	0.6911(62)	
σ_8	R.m.s. of the matter fluctuation	0.8159(86)	
z_*	Redshift at decoupling	1089.90(23)	

2.1.1.1 Gaussian Random Field

Since the initial conditions of the Universe[1] are very well described by a Gaussian random field, it is worth providing a mathematical description of their structure and properties. While a more in-depth and mathematical description of random fields and their geometry is provided in Adler and Taylor, 2007, let us provide some basic definitions and properties.

A random variable $X \in \mathbb{R}$ has a Gaussian distribution (or normal distribution) with mean μ and variance σ if its PDF reads

$$P(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right). \tag{2.2}$$

This definition can easily be generalized to d dimensions: a random vector $\boldsymbol{X} \in \mathbb{R}^d$ has a multivariate Gaussian distribution with mean $\boldsymbol{\mu} \in \mathbb{R}^d$ and variance $\boldsymbol{\Sigma} \in M(\mathbb{R}^d)$, where $\boldsymbol{\Sigma}$ is a positive-definite matrix, if its PDF reads

$$P(\boldsymbol{X}) = \frac{1}{\sqrt{(2\pi)^d \det\boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^\mathrm{T}\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right). \tag{2.3}$$

The element i,j of the covariance matrix is the covariance of the ith element of $\boldsymbol{X}$ with its j-th element. It can be formally written as

$$\Sigma_{ij} = \langle (X_i - \mu_i)(X_j - \mu_j)\rangle, \tag{2.4}$$

where the brackets indicate an ensemble average. For a function $F : \mathbb{R}^d \to \mathbb{R}$, the ensemble average over the ensemble of possible realizations Ω is defined as

$$\langle F \rangle \equiv \int_\Omega \mathrm{d}^d X' \, F(\boldsymbol{X'}) P(\boldsymbol{X'}). \tag{2.5}$$

In the following, brackets symbols are expectation, integrated over all possible realisations. Using ergodicity in an isotropic and homogeneous field, this is equivalent to averaging over space.

Using multivariate Gaussian distributions, we can also define a discrete Gaussian random field. Let X be a discrete field defined at positions $\{\boldsymbol{r_i}\}_{i=1,\ldots,d}$. The field is said to be a Gaussian random field if the vector $\boldsymbol{X} = \{X(\boldsymbol{r_i})\}_{i=1,\ldots,d}$ is distributed following a multivariate Gaussian distribution. In cosmology, it is very common to use the two-point correlation function instead of

---

[1] In the following of the manuscript, I will call "the initial conditions" the initial conditions in the matter dominated Universe, which are set by the measurements of the CMB.

the covariance. For a given pair of point $\boldsymbol{r}_i, \boldsymbol{r}_j'$, the two-point correlation function of a field is defined as

$$\xi(\boldsymbol{r}_i, \boldsymbol{r}_j') \equiv \left\langle (X(\boldsymbol{r}_i) - \mu(\boldsymbol{r}_i))(X(\boldsymbol{r}_j') - \mu(\boldsymbol{r}_j')) \right\rangle, \tag{2.6}$$

where $\mu(\boldsymbol{r}) = \langle X(\boldsymbol{r}) \rangle$ is the mean of the field (it is a field itself). This is the generalization of equation (2.4) to a discrete field.[2] Since a multivariate Gaussian distribution is described uniquely by its mean and covariance, a discrete Gaussian random field is entirely described by its mean and two-point correlation function.

We are now in a position to define a (continuous) Gaussian random field. Let X by a field in d dimensions. The field is a Gaussian random field if for a given position $\boldsymbol{r}$ there exists μ, σ such that

$$X(\boldsymbol{r}) \sim \mathcal{N}(\mu, \sigma), \tag{2.7}$$

and the covariance of the field at any pair of point $\boldsymbol{r}, \boldsymbol{r}'$ is given by the two-point correlation function

$$\xi(\boldsymbol{r}, \boldsymbol{r}') = \left\langle (X(\boldsymbol{r}) - \mu(\boldsymbol{r}))(X(\boldsymbol{r}') - \mu(\boldsymbol{r}')) \right\rangle. \tag{2.8}$$

The standard model of cosmology further assumes that the Universe is statistically homogeneous and isotropic. Mathematically, a Gaussian random field is homogeneous and isotropic if its correlation functions verifies

$$\xi(\boldsymbol{r}, \boldsymbol{r}') = \xi(\left\| \boldsymbol{r} - \boldsymbol{r}' \right\|) = \xi(d), \tag{2.9}$$

where $d = \left\| \boldsymbol{r} - \boldsymbol{r}' \right\|$ and its mean is a constant

$$\mu(\boldsymbol{r}) = \mu_0. \tag{2.10}$$

As a consequence, the statistical properties of the field are invariant by translation and rotation.

2.1.1.2 Generating Gaussian random fields

Because the initial conditions of the Universe are well described by a Gaussian random field, the generation of Gaussian random field is the first step of any numerical simulation that aims at simulating the Universe from its birth. Let us describe how one can draw values from a multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. Let $\mathbf{A}$ be a real matrix such that $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \boldsymbol{\Sigma}$. Let $\boldsymbol{Z} = \{z_1, \ldots, z_N\}$ be a set of N independent normal variables[3] and $\boldsymbol{\mu}$ an N-dimensional real vector. It can be shown that

$$\boldsymbol{X} = \boldsymbol{\mu} + \mathbf{A}\boldsymbol{Z} \tag{2.11}$$

is distributed following a multivariate Gaussian variable with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

2.1.1.3 Power spectrum

In a statistically homogeneous Universe, it is convenient to represent random the field δ by its Fourier components using the following convention

$$\delta(\boldsymbol{k}) = \int \mathrm{d}^3\boldsymbol{r}\, \delta(\boldsymbol{r})e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}, \qquad \delta(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\boldsymbol{k}\, \delta(\boldsymbol{k})e^{i\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{2.12}$$

The power spectrum $P(k)$ of the field is the expectation value

$$\left\langle \delta(\boldsymbol{k})\delta^\star(\boldsymbol{k}') \right\rangle \equiv P(\boldsymbol{k})(2\pi)^3\delta_{\mathrm{D}}\left(\boldsymbol{k} - \boldsymbol{k}'\right). \tag{2.13}$$

---

[2]Note that this definition is not specific to Gaussian random fields.
[3]Normal variables can be generated using the Box-Muller method or the ziggurat algorithm.
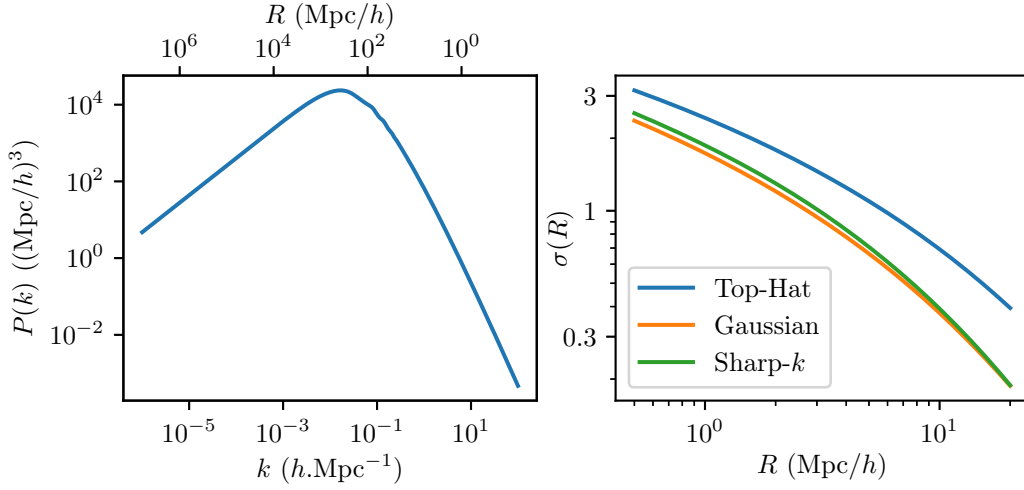
Figure 2.1.2: (*Left:*) Plot of the matter power spectrum as a function of the wave number assuming a Planck 2018 (Planck Collaboration, 2018a) cosmology. At scales involved in galaxy formation ($0.1\,\mathrm{Mpc}/h$ to $10\,\mathrm{Mpc}/h$), the spectrum resembles a power-law with slope ~ -2. (*Right:*) Standard deviation of the field smoothed with different filters as labelled.

Here the superscript * stands for the complex conjugate, which makes $P(k)$ positive definite. The Dirac delta is a consequence of translational invariance (homogeneity). Otherwise, the ensemble average $\langle\delta(\boldsymbol{k})\delta(\boldsymbol{k}')\rangle$ would acquire a phase factor when $\boldsymbol{r} \to \boldsymbol{r} + \boldsymbol{\Delta r}$. If the field is real $\delta^\star(\boldsymbol{k}) = \delta(-\boldsymbol{k})$, and we obtain

$$\langle\delta(\boldsymbol{k})\delta(\boldsymbol{k}')\rangle = P(\boldsymbol{k})(2\pi)^3\delta_{\mathrm{D}}\big(\boldsymbol{k} + \boldsymbol{k}'\big)\,. \tag{2.14}$$

Requesting further rotational invariance implies that the power spectrum depends only on $k = \|\boldsymbol{k}\|$, *i.e.*

$$\langle\delta(\boldsymbol{k})\delta^\star(\boldsymbol{k}')\rangle = P(k)(2\pi)^3\delta_{\mathrm{D}}\big(\boldsymbol{k} - \boldsymbol{k}'\big)\,. \tag{2.15}$$

Taking the Fourier transform of equation (2.13), we obtain the relation between the two point correlation function and the power spectrum

$$\xi(\boldsymbol{r}) = \frac{1}{(2\pi)^3}\int_{\boldsymbol{k}}\mathrm{d}^3\boldsymbol{k}\,P(k)e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\,. \tag{2.16}$$

For a statistically homogeneous and isotropic Gaussian random field in three dimensions this can be rewritten in a more compact way as

$$\xi(r) = \frac{1}{2\pi^2}\int_0^\infty\mathrm{d}k\,k^2P(k)j_1(kr), \tag{2.17}$$

where j_1 is the second spherical Bessel function. For the sake of completeness, note that in one and two dimensions we have

$$\xi^{1\mathrm{D}}(r) = \frac{1}{\pi}\int_0^\infty\mathrm{d}k\,P(k)j_0(kr), \tag{2.18}$$

$$\xi^{2\mathrm{D}}(r) = \frac{1}{2\pi}\int_0^\infty\mathrm{d}k\,kP(k)J_0(kr), \tag{2.19}$$

Table 2.2: Correspondence of linear operators between their real-space and Fourier-space representations. Note that real space filters are convolution operators while their Fourier representation is a multiplication.

Operator	Real space	Fourier space
Gradient	$\boldsymbol{\nabla}$	$i\boldsymbol{k}$
Laplacian	∇^2	$-k^2$
Spatial shift	$\delta(\boldsymbol{r}) \mapsto \delta(\boldsymbol{r} + \boldsymbol{\Delta r})$	$e^{i\boldsymbol{k}\cdot\boldsymbol{\Delta r}}$
Gaussian filter	$W_{\mathrm{G}}(r) \equiv \dfrac{1}{(2\pi)^{3/2}R^3}e^{-r^2/2R^2}$	$\tilde{W}_{\mathrm{G}}(k) = e^{-(kR)^2/2}$
Top-Hat filter	$W_{\mathrm{TH}}(r) \equiv \dfrac{3}{4\pi R^3}\vartheta_{\mathrm{H}}\left(1 - \dfrac{r}{R}\right)$	$\tilde{W}_{\mathrm{TH}}(k) = \dfrac{3j_1(kR)}{kR}$
Sharp k filter	$W_{\mathrm{sharp}}(r) \equiv \dfrac{3}{4\pi R^3}\dfrac{3j_1(r/R)}{r/R}$	$\tilde{W}_{\mathrm{sharp}}(k) = \vartheta_{\mathrm{H}}(1 - kR)$
Derivative of TH	$W'_{\mathrm{TH}} = \partial_R[W_{\mathrm{TH}}]$	$-\dfrac{3}{R}j_2(kR)$
Derivative of Gaussian	$W'_{\mathrm{G}}(r) = -\dfrac{1}{2R^2}W_{\mathrm{G}}(r)$	$-\dfrac{k^2}{2}\tilde{W}_{\mathrm{G}}(k)$

where J_0 is the first Bessel function and j_0 the first spherical Bessel function. The ΛCDM power spectrum depends on the properties of inflation and of the early Universe. It feature notably a peak at $\sim 100\,\mathrm{Mpc}/h$ and then decreases roughly as a power-law with index $n_{\mathrm{s}} \sim 1 - 2$, as shown on figure 2.1.2, left panel.

Generating Gaussian Random Field

Let us describe a method to generate Gaussian random fields on a periodic grid with a given power spectrum $P(k)$. The first naive approach works as follows

1. generate a white noise field from a Gaussian distribution with zero mean and unit variance

$$\mu(\boldsymbol{r}) \sim \mathcal{N}(0, 1),$$

2. compute $\mu(\boldsymbol{k})$, the Fourier transform of $\mu(\boldsymbol{r})$ (using e.g. the Fast Fourier Transform (FFT) algorithm),
3. compute the Fourier field

$$\delta(\boldsymbol{k}) = \sqrt{P(k)}\mu(\boldsymbol{k}),$$

4. obtain the real-space overdensity $\delta(\boldsymbol{r})$ using an inverse Fourier transform of $\tilde{\delta}(\boldsymbol{k})$.

While the naive approach is easy to implement, it has a number of issues for small boxes (see e.g. Pen, 1997). Indeed, the power spectrum is assumed to be spherically symmetric, an assumption that does not hold on a finite rectangular box. One way to partially solve the problem is due to Hahn and Abel, 2011, where they suggested sampling the power spectrum in real-space instead of Fourier space, so that the periodicity of the box is correctly accounted for. This is the approach used in the Music software to generate initial conditions for cosmological simulations.

2.1.1.4 Correlation of the field and linear operators

In the context of excursion set theory and peak theory, discussed in sections 2.1.3 and 2.1.5, one needs to be able to compute correlation functions of the field and its derivatives with respect to space or smoothing scale. The motivations to compute these correlation functions will be discussed in further details in sections 2.1.3 and 2.1.5 for excursion set theory and peak patch, while the importance of the smoothing operation is discussed in section 2.1.2.2. One of the properties of Gaussian random field is that any linear combination of a Gaussian random field is

itself a Gaussian random field. Stated differently, any linear transformation of a Gaussian random field is a Gaussian random field itself so that, in general, any linear operator will conserve the Gaussian property of a field. Following the lines of section 2.1.1.3, let us define the two-point cross-correlation function between two Gaussian random fields δ_1 and δ_2

$$\xi_{\delta_1,\delta_2}(\boldsymbol{r}) \equiv \frac{1}{(2\pi)^6} \int \mathrm{d}^3\boldsymbol{k} \int \mathrm{d}^3\boldsymbol{k}' \left\langle \delta_1(\boldsymbol{k})\delta_2^\star(\boldsymbol{k}') \right\rangle e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}. \tag{2.20}$$

The results of section 2.1.1.3 can be recovered setting $\delta_1 = \delta_2$.

Let $\mathcal{F}$ be a linear operator. We define its Fourier representation $\tilde{\mathcal{F}}$ as

$$\tilde{\mathcal{F}}[\delta](\boldsymbol{k}) = \int \mathrm{d}^3\boldsymbol{r}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \mathcal{F}[\delta](\boldsymbol{r}). \tag{2.21}$$

The Fourier representation of convolution filters, such as the Top-Hat filter and the Gaussian filter, become simple multiplications in Fourier space. Noting that any derivation operators ($\boldsymbol{\nabla}, \nabla^2, \dots$) can be written as a convolution with the relevant distribution, their representation in Fourier also become a simple multiplication, where the multiplication factor does not depend on the underlying field. Some common operators and their Fourier representations are given in Table 2.2. This means that $\tilde{\mathcal{F}}[\delta](\boldsymbol{k}) = \tilde{\mathcal{F}}(\boldsymbol{k})\delta(\boldsymbol{k})$. In the following, we will restrain ourselves to operators that can be written as multiplications in Fourier space[4]. Using this formalism, we can compute any correlation function between two operators applied to a Gaussian random field. The correlation between $\mathcal{F}_1[\delta]$ and $\mathcal{F}_2[\delta]$, where $\mathcal{F}_1$ and $\mathcal{F}_2$ are linear operators, reads

$$\begin{aligned}
\xi_{\mathcal{F}_1\mathcal{F}_2} &= \langle \mathcal{F}_1[\delta]\mathcal{F}_2[\delta] \rangle \\
&= \frac{1}{(2\pi)^6} \int \mathrm{d}^3\boldsymbol{k} \int \mathrm{d}^3\boldsymbol{k}'\, \tilde{\mathcal{F}}_1(\boldsymbol{k})\tilde{\mathcal{F}}_2^\star(\boldsymbol{k}') \left\langle \delta(\boldsymbol{k})\delta^\star(\boldsymbol{k}') \right\rangle \\
&= \frac{1}{(2\pi)^6} \int \mathrm{d}^3\boldsymbol{k} \int \mathrm{d}^3\boldsymbol{k}'\, \tilde{\mathcal{F}}_1(\boldsymbol{k})\tilde{\mathcal{F}}_2^\star(\boldsymbol{k}')(2\pi)^3 P(k)\delta_\mathrm{D}\big(\boldsymbol{k}-\boldsymbol{k}'\big) \\
&= \frac{1}{(2\pi)^3} \int \mathrm{d}^3\boldsymbol{k}\, P(k)\tilde{\mathcal{F}}_1(\boldsymbol{k})\tilde{\mathcal{F}}_2^\star(\boldsymbol{k}).
\end{aligned} \tag{2.22}$$

Note that in general the fields returned by the operators may not be invariant under rotation or even translation, even when the underlying field is itself isotropic or homogeneous. This is for example the case when considering the gradient of the field in a given direction of space. As an example, let us compute the correlation function between the field smoothed by a Gaussian filter at scale R_1 and the field smoothed by a Gaussian filter at scale R_2 at a separation $\boldsymbol{r}$[5] using equation (2.22)

$$\begin{aligned}
\xi_{R_1,R_2}(\boldsymbol{r}) &= \langle (W_{\mathrm{G},R_1} * \delta)(0) \times (W_{\mathrm{G},R_2} * \delta)(\boldsymbol{r}) \rangle \\
&= \frac{1}{(2\pi)^3} \int \mathrm{d}^3\boldsymbol{k}\, P(k)\tilde{W}_{\mathrm{G},R_1}(k)\tilde{W}_{\mathrm{G},R_2}(k)e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \\
&= \frac{1}{(2\pi)^3} \int \mathrm{d}^3\boldsymbol{k}\, P(k)\exp\left(-\frac{k^2(R_1^2+R_2^2)}{2} - i\boldsymbol{k}\cdot\boldsymbol{r}\right).
\end{aligned} \tag{2.23}$$

2.1.2 Formation of the structures of the Universe

In this section, I detail the models that describe the formation of the structures of the Universe. In section 2.1.2.1, I present how initial tiny fluctuations grow in the linear regime. In section 2.1.2.2, I detail the analytical solution of the spherical collapse that will be at the base of our understanding of the formation of dark matter halos. In section 2.1.2.3, I detail how the first structures emerge from the displacement of matter in the Zel'dovich approximation.

---

[4]This includes any operator that can be written as a convolution operation with a distribution, in particular convolution and linear differential operators.

[5]The separation can be interpreted as a shift operator applied to the field.

2.1.2.1   Linear perturbations

Since the initial conditions of the Universe are given by tiny fluctuations of the density field around its mean value, it is expected that the initial evolution can be described in the linear regime, expressed as perturbations of the density contrast δ. In the linear regime, the variance of the density contrast increases as matter departs under-dense regions to reach over-dense regions. In order to study this regime, let us first restrict ourselves to a pressure-less fluid approach that describe DM in the ΛCDM model. In this model, the DM is assumed to start from a state where the velocity field has no velocity dispersion (the velocity field is single-valued). Linear perturbations aim at providing a description of the density contrast up to the moment where multiple particles with different velocities are found at the same location, a moment known as shell-crossing. It turns out that this regime provides a good description of the first steps of the formation of the large scale structures of the Universe.

The equations of interest are the continuity, Euler (with no pressure) and Poisson equations

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \rho \boldsymbol{v} = 0, \tag{2.24}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \boldsymbol{\nabla})\boldsymbol{v} = -\boldsymbol{\nabla}\Phi, \tag{2.25}$$

$$\nabla^2 \Phi = 4\pi G \rho. \tag{2.26}$$

Equation (2.25) can be obtained from the Vlasov-Poisson equation, assuming that a single velocity is found at each location. Here spatial derivatives have been done in *proper* units. We can compute the comoving position *via* $\boldsymbol{r} = a(t)\boldsymbol{x}$. In these variables, the peculiar velocity $\boldsymbol{u}$ is the sum of the Hubble flow and the comoving velocity $\boldsymbol{v}$

$$\boldsymbol{u} \equiv \dot{a}(t)\boldsymbol{x} + \boldsymbol{v}, \quad \boldsymbol{v} \equiv a\dot{\boldsymbol{x}}. \tag{2.27}$$

Under the change of variable $\boldsymbol{r} \to \boldsymbol{x}$, equations (2.24)–(2.26) can be changed using the following transformation for the time derivative and gradient operators

$$\boldsymbol{\nabla} \to \frac{\boldsymbol{\nabla}_{\boldsymbol{x}}}{a}, \quad \frac{\partial}{\partial t} \to \frac{\partial}{\partial t} - \frac{\dot{a}}{a}\boldsymbol{x} \cdot \boldsymbol{\nabla}_{\boldsymbol{x}}. \tag{2.28}$$

Recalling that $\delta(\boldsymbol{r}, t) = (\rho(\boldsymbol{r}, t) - \bar{\rho}(t))/\bar{\rho}(t)$, this yields in comoving coordinates

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\boldsymbol{\nabla}_{\boldsymbol{x}} \cdot [(1 + \delta)\boldsymbol{v}] = 0, \tag{2.29}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} + \frac{1}{a}(\boldsymbol{v} \cdot \boldsymbol{\nabla}_{\boldsymbol{x}})\boldsymbol{v} = -\frac{\boldsymbol{\nabla}_{\boldsymbol{x}}\Phi}{a}, \tag{2.30}$$

$$\nabla_{\boldsymbol{x}}^2 \Psi = 4\pi G \bar{\rho} a^2 \delta, \tag{2.31}$$

$$\text{with } \Psi \equiv \Phi + a\ddot{a}x^2/2. \tag{2.32}$$

These equations can then be linearised at first order in δ, $\boldsymbol{v}$ and Ψ

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\boldsymbol{\nabla}_{\boldsymbol{x}} \cdot \boldsymbol{v} = 0, \tag{2.33}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} = -\frac{\boldsymbol{\nabla}_{\boldsymbol{x}}\Phi}{a}, \tag{2.34}$$

$$\nabla_{\boldsymbol{x}}^2 \Psi = 4\pi G \bar{\rho} a^2 \delta. \tag{2.35}$$

We then derive equation (2.33) w.r.t. t and use equations (2.34) and (2.35) to finally get the second order partial differential equation

$$\frac{\partial^2 \delta}{\partial t^2} + \frac{2\dot{a}}{a}\frac{\partial \delta}{\partial t} - 4\pi G \bar{\rho}\delta = 0. \tag{2.36}$$

In order to move forward, we can use the Fourier representation of the overdensity $\delta(\boldsymbol{k})$. This yields a second order ordinary differential equation

$$\frac{\mathrm{d}^2\delta(\boldsymbol{k},t)}{\mathrm{d}t^2} + \frac{2\dot{a}}{a}\frac{\mathrm{d}\delta(\boldsymbol{k},t)}{\mathrm{d}t} - 4\pi G\bar{\rho}\delta(\boldsymbol{k},t) = 0. \tag{2.37}$$

We can immediately see that equation (2.37) does not have any scale dependence: all modes grow (or decay) at the same rate. The perturbations evolve as a function of time only and can formally be written as

$$\delta(\boldsymbol{k},t) = A(\boldsymbol{k})D_+(t) + B(\boldsymbol{k})D_-(t), \tag{2.38}$$

where $D_+(t)$ is a growing mode and $D_-(t)$ is a decaying mode and A and B are constants of time. D_+ is usually normalized to its value at $t(z=0)$ so that $D_+(z=0)=1$. The decaying mode can be directly expressed as a function of the Hubble constant

$$D_-(t) \propto H(t), \tag{2.39}$$

and the growing mode is obtained using

$$D_+(t) \propto H(t)\int_0^t \frac{\mathrm{d}t'}{a^2(t')H^2(t')} \propto H(z)\int_z^\infty \frac{1+z'}{E^3(z')}\,\mathrm{d}z', \tag{2.40}$$

where

$$E(z) \equiv \frac{H(z)}{H_0} = \sqrt{\Omega_{\Lambda,0} + (1-\Omega_0)(1+z)^2 + \Omega_{\mathrm{m},0}(1+z)^3 + \Omega_{\mathrm{r},0}(1+z)^4}. \tag{2.41}$$

Ω_0 is equal to one in a flat Universe, while $\Omega_{\Lambda,0}, \Omega_{\mathrm{m},0}$ and $\Omega_{\mathrm{r},0}$ are the current Λ, matter and radiation density. They are linked to their value at redshift z by

$$\Omega_\Lambda(z) = \frac{\Omega_{\Lambda,0}}{E^2(z)}, \quad \Omega_\mathrm{m}(z) = \frac{\Omega_{\mathrm{m},0}(1+z)^3}{E^2(z)}, \quad \Omega_\mathrm{r}(z) = \frac{\Omega_{\mathrm{r},0}(1+z)^4}{E^2(z)}. \tag{2.42}$$

In the matter dominated era or in an Einstein de-Sitter (EdS) universe[6], the growing mode scales as

$$D_+(z) \sim t^{2/3} \sim \frac{1}{1+z}. \tag{2.43}$$

In the general case in a ΛCDM cosmology, their is no explicit formula for D_+ but equation (2.40) can be integrated numerically, as shown on figure 2.1.3. Alternatively, a good approximation is due to Carroll et al., 1992

$$D_+(z) \propto \frac{\Omega_\mathrm{m}(z)}{1+z}\left[\Omega_\mathrm{m}^{4/7}(z) - \Omega_\Lambda(z) + \left(1 + \frac{\Omega_\mathrm{m}(z)}{2}\right)\left(1 + \frac{\Omega_\Lambda(z)}{70}\right)\right]^{-1}. \tag{2.44}$$

This approximation holds for a close Universe with non-null matter density and a Λ contribution.

In the remainder of the dissertation, the linear growth factor will be noted $D(z) \equiv D_+(z)$ for the sake of simplicity, while the decaying mode will be neglected.

---

[6]An EdS Universe is a flat, matter-only Universe with no cosmological constant. It is a good approximation to our Universe after the radiation-dominated era $z < 300$ and before the Λ-dominated era $z > 2$.
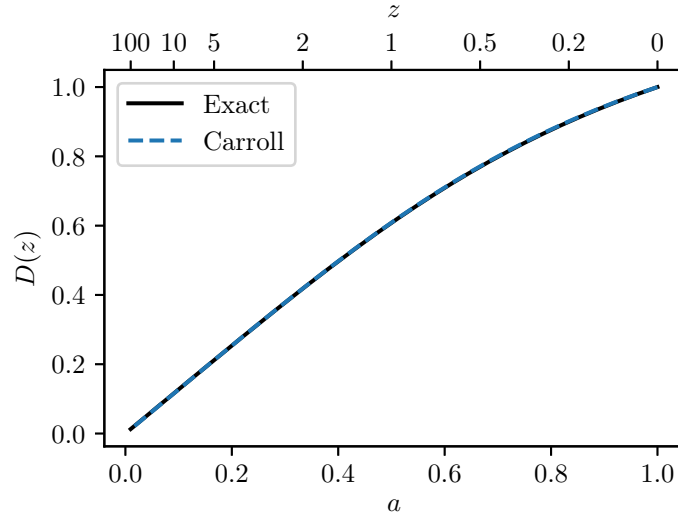
Figure 2.1.3: Linear matter growth factor in a Planck 2018 cosmology (Planck Collaboration, 2018a) using the exact expression of D_+ (equation (2.40), solid black) and the fit by Carroll et al. (equation (2.44), dashed blue).

2.1.2.2 Spherical collapse

Let us consider a region of the Universe with uniform initial density ρ_i and radius R_i. For the sake of simplicity, we will assume to be in an EdS Universe, but similar results can be found including a cosmological constant (see e.g. Lacey and Cole, 1993; Lahav et al., 1991). Following section 2.1.2.1, we assume that there is no shell-crossing. We therefore assume that collapse will happen in concentric spheres, with the outermost spheres collapsing in a time larger or equal to the collapse time of the inner spheres. Under the assumption, the total mass in a sphere is constant. Let $\delta_i = (\rho_i - \bar{\rho}_{m,i})/\bar{\rho}_{m,i}$ be the initial overdensity of the region w.r.t. the cosmic mean matter density $\bar{\rho}_{m,i}$. The total mass in the region is given by $M = (4\pi/3)R_i^3 \bar{\rho}_i(1 + \delta_i)$. The region evolves under the action of gravity following

$$\frac{\mathrm{d}^2 R}{\mathrm{d}t^2} = -\frac{GM}{R^2} = -\frac{H_i^2 R_i^3}{2R^2}(1 + \delta_i), \tag{2.45}$$

where H_i is the initial Hubble rate. It is worth noting that at fixed initial overdensity, equation (2.45) is scale invariant: the evolution of the sphere depends on the initial density only. Let us now integrate equation (2.45) over time to get the specific energy equation[7]

$$\frac{1}{2}\dot{R}^2 - \frac{H_i^2 R_i^3}{2R}(1 + \delta_i) = E. \tag{2.46}$$

If $E > 0$, the solution is unbound and the radius will grow forever. If $E < 0$, the solution is bound and the radius will eventually collapse to $R \to 0$. At early times, the bulk velocity is due to the Hubble flow $\dot{R}_i \approx H_i R_i$ so that the total energy reads

$$E = -\frac{H_i^2 R_i^2}{2}\delta_i. \tag{2.47}$$

The energy is negative for overdensities $\delta_i > 0$ and positive otherwise. This shows that in the spherical collapse model, any initial overdensity will eventually collapse. Let us now assume that

---

[7]Since the mass of the sphere is assumed to be constant, this specific energy is conserved.
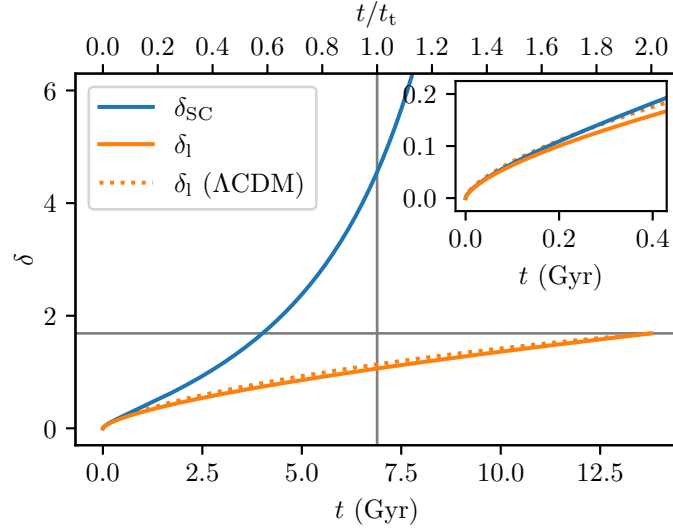
Figure 2.1.4: Plot of the evolution of spherical collapse overdensity δ_{SC} (blue) and linear overdensity δ_l (orange) in a EdS Universe and in a ΛCDM Universe (dotted orange). The spherical collapse solution diverges in a time $t = t_e = 2t_t$ (here $t_e = t_H \approx 13.8\,\text{Gyr}$). At this time, the linear overdensity has a value $\delta_c \approx 1.69$ (horizontal solid grey line). At turnaround (vertical solid grey line) the spherical-collapse overdensity is $\delta_{SC} \approx 4.55$ and the linear overdensity $\delta_l \approx 1.06$. At early times (inset), the spherical collapse model and the linear evolution coincide with $\delta(t) \sim (1+z)^{-1}$.

$\delta_i > 0$ to derive the evolution of the spherical region. Under this assumption, the solution of equation (2.45) can be written parametrically

$$r = A(1 - \cos\theta), \qquad t = B(\theta - \sin\theta). \tag{2.48}$$

Here A and B are set by the initial conditions

$$A = -\frac{GM}{E}, \qquad B^2 = -\frac{(GM)^2}{E^3}. \tag{2.49}$$

The evolution for a spherical region collapsing in a Hubble time ($r(t_H) = 0$) is shown on figure 2.1.4. Using the conservation of energy and equations (2.46) and (2.47), we can compute the turnaround radius — or maximum radius — R_t for which the radial velocity vanishes

$$R_t = \frac{(1 + \delta_i)}{\delta_i} R_i, \quad t_t = \frac{\pi}{2} H_i \frac{1 + \delta_i}{\delta_i^{3/2}}. \tag{2.50}$$

After turnaround, the region will start contracting until $R(t_f) = 0$ with $t_f = 2t_t$. For small initial overdensities, turnaround time and radius scale like

$$R_t \sim \frac{R_i}{\delta_i}, \quad t_t \sim \frac{\pi}{2} \frac{H_i}{\delta_i^{3/2}}. \tag{2.51}$$

This shows that small overdensities have large turnaround radii, since collapse time is inversely proportional to the initial overdensity. This regime is the one of interest assuming cosmological initial conditions, as the observation of the CMB gives us $|\delta_i| \sim 10^{-3}$ at $z \sim 1000$.

In practice, the region will not collapse to a single point. As the region collapses, the effect of the initial (random) velocity dispersion will become non-negligible so that the assumption of a perfectly spherical collapse will break. Instead of converting all the gravitational energy to kinetic energy, both term will eventually reach equilibrium. This process, known as "virialization" will relax to the state where the Virial theorem is verified, *i.e.*

$$2K + U = 0, \tag{2.52}$$

where $K = \dot{R}^2/2$ is the specific kinetic energy and $U = -GM/R$ is the specific gravitational energy. Combining equations (2.47) and (2.52) and introducing the Virial radius R_{vir} we find that

$$R_{\mathrm{vir}} = \frac{R_{\mathrm{t}}}{2}. \tag{2.53}$$

After virialization, the radius of the region will be half the radius at turnaround and the density is eight time the density at turnaround. It can be shown that the overdensity at the time of virialization is

$$1 + \delta(t_{\mathrm{vir}}) = 18\pi^2 \approx 178. \tag{2.54}$$

This overdensity is frequently used in numerical simulation to define the radius of dark matter halos and is written Δ_{178} or quite frequently Δ_{200} when using a value of 200. The corresponding linear density contrast at the time of virialization, which defines the critical density δ_{c} is

$$\delta_{\mathrm{c}} \equiv \delta_{\mathrm{l}}(t_{\mathrm{vir}}) = \frac{3}{5} \left(\frac{3}{4}\right)^{2/3} (\theta_{\mathrm{vir}} - \sin\theta_{\mathrm{vir}})^{2/3} = \frac{3}{5} \left(\frac{3\pi}{2}\right)^{2/3} \approx 1.6865. \tag{2.55}$$

This critical density is of interest, as it provides a way to find regions that will collapse non-linearly following the spherical models using the linear overdensity field: any region with their linear density $\delta > \delta_{\mathrm{c}}$ should be considered as collapsed and virialized. This will be further discussed in section 2.1.3

2.1.2.3 Zel'dovich approximation

An interesting approach to understand the genesis of the cosmic web is to adopt a Lagrangian view dual to the Eulerian description used in section 2.1.1. Instead of expression quantities at fixed comoving coordinates (Eulerian view), one can indeed write the cosmic fields as a function of the initial position $\boldsymbol{q}$. This initial position is related to the comoving coordinate $\boldsymbol{x}$ at time t by a displacement term

$$\boldsymbol{x}(\boldsymbol{q}, t) = \boldsymbol{q} + \boldsymbol{\psi}(\boldsymbol{q}, t), \tag{2.56}$$

where $\boldsymbol{\psi}(\boldsymbol{q}, t)$ is the displacement field. Starting from a homogeneous initial density field, the local density at time t then reads

$$\rho(\boldsymbol{q}, t) = \frac{\bar{\rho}(\boldsymbol{q})}{J}, \tag{2.57}$$

where J is the Jacobian of the Eulerian-to-Lagrangian transformation $J = |\mathrm{d}\boldsymbol{x}/\mathrm{d}\boldsymbol{q}|$ given by

$$J = \left| \delta_{ij}^{\mathrm{K}} + \frac{\partial\psi_i}{\partial q_j} \right|, \tag{2.58}$$

with δ_{ij}^{K} the Kronecker delta. Lagrangian Perturbation Theory finds a perturbative solution for the displacement field,

$$\boldsymbol{\psi}(\boldsymbol{q}, t) = \boldsymbol{\psi}^{(1)}(\boldsymbol{q}, t) + \boldsymbol{\psi}^{(2)}(\boldsymbol{q}, t) + \cdots. \tag{2.59}$$

The Zel'dovich approximation is the first-order approximation to equation (2.59), which reads

$$\boldsymbol{\psi}(\boldsymbol{q}, t) \approx \boldsymbol{\psi}^{(1)}(\boldsymbol{q}, t) = \int \frac{\mathrm{d}k}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\boldsymbol{q}} \frac{i\boldsymbol{k}}{k^2} \delta_{\boldsymbol{k}}(t). \tag{2.60}$$

Applying the same formalism as in section 2.1.2.1, we find that the displacement field has a growing and a decaying mode. Keeping only the growing mode, we can express equation (2.60) in terms of the linear matter growth function

$$\boldsymbol{\psi}(\boldsymbol{q}, t) \approx \boldsymbol{\psi}^+(\boldsymbol{q}) D_+(t). \tag{2.61}$$

Plugging equation (2.61) into the Jacobian of equation (2.58) we find that the density reads

$$\rho(\boldsymbol{q}, t) = \frac{\bar{\rho}}{|(1 - D_+(t)\lambda_1)(1 - D_+(t)\lambda_2)(1 - D_+(t)\lambda_3)|}, \tag{2.62}$$

with $\lambda_1 \geq \lambda_2 \geq \lambda_3$ the eigenvalues of $-\partial\psi_i^+/\partial q_j$. Note that the deformation tensor $-\partial\psi_i^+/\partial q_j$ is equal to the hessian of the gravitational potential — the tidal tensor — up to a time-dependant factor $(4\pi G\rho(t)a^2(t)D(t))^{-1}$ (Porciani et al., 2002). The Zel'dovich displacement is therefore a rectilinear trajectory that moves particles along the direction of the initial force that converts the three-dimensional sphere in $\boldsymbol{q}$-space to a flattened ellipsoid in the real $\boldsymbol{r}$-space, see figure 2.1.5. Multiple studies have shown that the Zel'dovich approximation holds up to the mildly non-linear regime of structure formation (e.g. M. White, 2014) and describes well the anisotropic collapse of matter that shapes the cosmic web. Indeed, equation (2.62) suggests that for $D_+(t)\lambda_1 \to 1$, the density diverges resulting in the formation of a caustic. The approximation clearly does not hold any more for particles that shell crossed, but it still provides a good approximation for particles that surround the shell-crossed region. In addition, the Zel'dovich approximation gives us a physical understanding of the next likely direction(s) of collapse. If $\lambda_2 \geq 0$, the region contracts in the corresponding direction, eventually leading to the formation of a filamentary structure. Finally, if $\lambda_3 \geq 0$, the region will also contract along the third direction, leading to the formation of a node of the cosmic web. While the details of the secondary and third collapse are not well predicted by the Zel'dovich approximation, various models have been designed to overcome this shortcoming, such as the adhesion model (Kofman et al., 1992) or more recently the origami model (Neyrinck, 2014).

2.1.3 The excursion set theory

The excursion set approach, originally formulated by Press and Schechter, 1974, assumes that virialized halos form from spherical regions whose initial mean density equals some critical value. The distribution of late-time halos can thus be inferred from the simpler Gaussian statistics of their Lagrangian progenitors. The approach implicitly assumes approximate spherical symmetry (but not homogeneity), and uses spherical collapse, as presented in section 2.1.2.2, to establish a mapping between the initial mean density of a patch and the time at which it recollapses under its own gravity.

According to this model, a sphere of initial radius R shrinks to zero volume at redshift z if its initial mean overdensity δ equals $\delta_c D(z_{\mathrm{in}})/D(z)$, where $D(z)$ is the growth rate of linear matter perturbations, z_{in} the initial redshift, and $\delta_c = 1.686$ for an Einstein–de Sitter universe, or equivalently, if its mean overdensity linearly evolved to $z = 0$ equals $\delta_c/D(z)$, regardless of the initial size. If so, thanks to mass conservation, this spherical patch will form a halo of mass $M = (4\pi/3)R^3\bar{\rho}$ (where $\bar{\rho}$ is the comoving background density). The redshift z is assumed to be a proxy for its virialization time.

Bond et al., 1991 added to this framework the requirement that the mean overdensity in all larger spheres must be lower than δ_c, for outer shells to collapse at a later time. This condition ensures that the infall of shells is hierarchical, and the selected patch is not crushed in a bigger volume that collapses faster (the so-called *cloud-in-cloud* problem). The number density of halos of a given mass at a given redshift is thus related to the volume contained in the largest spheres whose mean overdensity $\delta \equiv \delta(R)$ crosses δ_c. The dependence of the critical value δ_c on departures
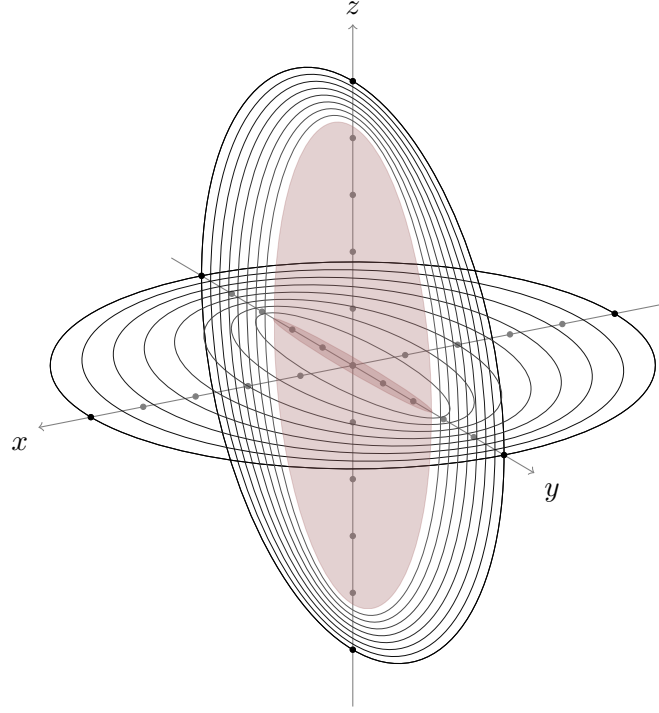
Figure 2.1.5: Scheme of the triaxial collapse under the Zel'dovich approximation, showing the collapse of an initial spherical shell as consecutive ellipsoids. The last shell is represented in shaded red and resembles a flattened spheroid (a Zel'dovich pancake) in the yz plane, the first axis to collapse is the x axis, the second y and the last z.

from spherical collapse induced by initial tides was studied by Bond and Myers, 1996, and later by R. K. Sheth et al., 2001, who approximated it as a scale-dependent barrier.

As the variation of $\delta(R)$ with scale resembles random diffusion, it is convenient to parametrize it with the variance

$$\sigma^2(R) \equiv \mathrm{Var}(\delta(R)) = \int \mathrm{d}k \, \frac{k^2 P(k)}{2\pi^2} \tilde{W}_{\mathrm{TH}}^2(kR) \tag{2.63}$$

of the stochastic process, smoothed with a real-space Top-Hat filter $\tilde{W}_{\mathrm{TH}}$, rather than with R or M; see section 2.1.1.4 for the details of the filter and its Fourier representation. In equation (2.63), $P(k)$ is the underlying power spectrum. The three quantities σ, R and M are in practice interchangeable. The mass fraction in halos of mass M at z is

$$\frac{M}{\bar{\rho}} \frac{\mathrm{d}n}{\mathrm{d}M} = \left| \frac{\mathrm{d}\sigma}{\mathrm{d}M} \right| f(\sigma) \,, \tag{2.64}$$

where $\mathrm{d}n/\mathrm{d}M$ is the number density of halos per unit mass (*i.e.* the mass function) and $f(\sigma)$ — often called the halo multiplicity — is the probability distribution of the first-crossing scale of the random walks, that is of the smallest σ (largest R) for which

$$\delta(R, \boldsymbol{r}) \equiv \frac{1}{(2\pi)^3} \int \mathrm{d}^3 k \, \delta_{\mathrm{m}}(\boldsymbol{k}) \tilde{W}_{\mathrm{TH}}(kR) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \frac{\delta_{\mathrm{c}}}{D(z)} \,, \tag{2.65}$$

where δ_{m} is the (unsmoothed) matter density. The first-crossing requirement avoids double counting and guarantees that $f(\sigma)$ is a well-behaved probability distribution, and the resulting mass fraction is correctly normalized.
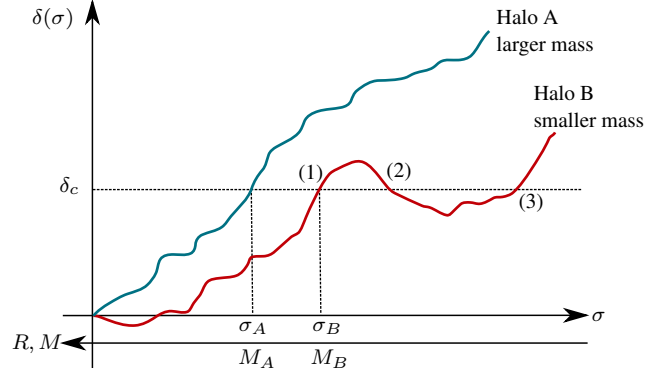
Figure 2.1.6: Pictorial description of the first-crossing and upcrossing conditions to infer the halo mass from the excursion set trajectory. The first-crossing condition on σ assigns at most one halo to each trajectory, with mass $M(\sigma)$. Upcrossing may instead assign several masses to the same trajectory (that is, to the same spatial location), thus over-counting halos. Trajectory B in the figure has a first crossing (upwards) at scale σ_B (1), a downcrossing (2) and second upcrossing (3), but the correct mass is only given by σ_B. However, the correlation of each step with the previous ones makes turns in small intervals of σ exponentially unlikely: at small σ most trajectories will thus look like trajectory A. Thanks to the correlation between steps at different scales, for small σ (large M) simply discarding downcrossings is a very good approximation.

The first-crossing probability, $f(\sigma)\Delta\sigma$, is the fraction of walks that cross the threshold between $\sigma - \Delta\sigma$ and σ for the first time. Considering discretized trajectories with a large number of steps $\sigma_1, \ldots, \sigma_N$ of width $\Delta\sigma \equiv \sigma_i - \sigma_{i-1}$ (corresponding to concentric spheres of radii $R_1, \cdots, R_N$), the first-crossing probability is the joint probability that $\delta_N > \delta_c$ and $\delta_i < \delta_c$ for $i < N$, with $\delta_i \equiv \delta(\sigma_i)$ and $\sigma_N = \sigma = N\Delta\sigma$. Hence, the distribution $f(\sigma)$ is formally defined as the limit

$$f(\sigma) \equiv \lim_{\Delta\sigma \to 0} \frac{1}{\Delta\sigma} \left\langle \vartheta_{\mathrm{H}}(\delta_N - \delta_c) \prod_i^{N-1} \vartheta_{\mathrm{H}}(\delta_c - \delta_i) \right\rangle, \tag{2.66}$$

where $\vartheta_{\mathrm{H}}(x)$ is Heaviside's step function, and the expectation value is evaluated with the multivariate distribution $p(\delta_i, \ldots, \delta_N)$. This definition discards crossings for which $\delta_i > \delta_c$ for any $i < N$, since $\vartheta_{\mathrm{H}}(\delta_c - \delta_i) = 0$, assigning at most one crossing (the first) to each trajectory. For instance, in figure 2.1.6, trajectory B would not be assigned the crossing marked with (3), since the trajectory lies above threshold between (1) and (2). Since taking the mean implies integrating over all trajectories weighed by their probability, $f(\sigma)$ can be interpreted as a path integral over all allowed trajectories with fixed boundary conditions $\delta(0) = 0$ and $\delta(\sigma) = \delta_c$ (Maggiore and Riotto, 2010).

In practice, computing $f(\sigma)$ becomes difficult if the steps of the random walks are correlated, as is the case for real-space Top-Hat filtering with a ΛCDM power spectrum, and for most realistic filters and cosmologies. For this reason, more easily tractable but less physically motivated sharp cutoffs in Fourier space have often been preferred, for which the correlation matrix of the steps becomes diagonal, treating the correlations as perturbations (Maggiore and Riotto, 2010; Corasaniti and Achitouv, 2011). The upcrossing approximation described below can instead be considered as the opposite limit, in which the steps are assumed to be strongly correlated (as is the case for a realistic power spectrum and filter). This approximation is equivalent to constraining only the last two steps of equation (2.66), marginalizing over the first $N - 2$.

The upcrossing approximation to $f(\sigma)$.

Indeed, Musso and R. K. Sheth, 2012 noticed that for small enough σ (*i.e.* for large enough masses), the first-crossing constraint may be relaxed into the milder condition

$$\delta' \equiv \frac{\mathrm{d}\delta}{\mathrm{d}\sigma} > 0 \, ; \tag{2.67}$$

that is, trajectories simply need to reach the threshold with positive slope (or with slope larger than the threshold's if δ_c depends on scale). This upcrossing condition may assign several halos of different masses to the same spatial location. For this reason, while first-crossing provides a well-defined probability distribution for σ (e.g. with unit normalization), upcrossing does not. However, since the first-crossing is necessarily upwards, and down-crossings are discarded, the error introduced in $f(\sigma)$ by this approximation comes from trajectories with two or more turns. Musso and R. K. Sheth, 2012 showed that these trajectories are exponentially unlikely if σ is small enough when the steps are correlated. The first-crossing and upcrossing conditions to infer the halo mass from excursion sets are sketched in figure 2.1.6: while the trajectory A would be (correctly) assigned to a single halo, the second upcrossing of trajectory B in the figure would be counted as a valid event by the approximation, and the trajectory would (wrongly) be assigned to two halos. The probability of this event is non-negligible only if σ is large. This is further illustrated on figure 2.1.7. The figure presents bundles of random trajectories drawn using a ΛCDM power spectrum constrained to a crossing (up or down) at a given radius. For small values of σ most of the trajectories that are first-crossing are also upcrossing. The fraction drops significantly for larger values of σ.

Returning to equation (2.66), expanding δ_{N-1} around δ_N gives

$$\vartheta_{\mathrm{H}}(\delta_c - \delta_{N-1}) \simeq \vartheta_{\mathrm{H}}(\delta_c - \delta_N) + \delta_{\mathrm{D}}(\delta_c - \delta)\, \delta' \Delta\sigma \, , \tag{2.68}$$

where the crossing scale σ, giving the halo's final mass M, is defined implicitly in equation (2.65), as the solution of the equation $\delta(\sigma) = \delta_c/D$[8]. The assumption that this upcrossing is first-crossing allows us to marginalize over the first $N-2$ variables in equation (2.66) without restrictions. The first term has no common integration support with $\vartheta_{\mathrm{H}}(\delta_N - \delta_c)$, and only the second one – containing the Jacobian $(\delta' - \delta'_c)$ – contributes to the expectation value (throughout the text, a prime will denote the derivative $\mathrm{d}/\mathrm{d}\sigma$). Adopting for convenience the normalized walk height $\nu \equiv \delta/\sigma$, for which $\langle \nu^2 \rangle = 1$, the corresponding density of solutions in σ-space obeys

$$\left| \nu' - \nu'_c \right| \delta_{\mathrm{D}}(\nu - \nu_c) = (|\delta'|/\sigma)\, \delta_{\mathrm{D}}(\nu - \nu_c) \, , \tag{2.69}$$

where $\nu_c \equiv \delta_c/(\sigma D)$ is the rescaled threshold. The probability of upcrossing at σ in equation (2.66) is therefore simply the expectation value of this expression,

$$f_{\mathrm{up}}(\sigma) \equiv p_{\mathrm{G}}(\nu = \nu_c) \int_0^\infty \mathrm{d}\delta' \, \delta' p_{\mathrm{G}}(\delta'|\nu_c) \, , \tag{2.70}$$

where the integral runs over $\delta' > 0$ because of the upcrossing condition (2.67). Usually, one sets $D = 1$ at $z = 0$ for simplicity so that $\nu_c = \delta_c/\sigma$. For Gaussian initial conditions[9], the conditional distribution $p_{\mathrm{G}}(\delta'|\nu_c)$ is a Gaussian with mean ν_c and variance $1/\Gamma^2$, where

$$\Gamma^2 = \frac{1}{\langle \delta'^2 \rangle - 1} = \frac{\gamma^2}{1 - \gamma^2} = \frac{1}{\sigma^2 \langle \nu'^2 \rangle} \, , \tag{2.71}$$

---

[8]A careful calculation shows that the step function should be asymmetric, so that $\vartheta_{\mathrm{H}}(\delta - \delta_c) = 1$ when $\delta = \delta_c$ instead of $1/2$.

[9]No conceptual complications arise in dealing with a non-Gaussian distribution, which is nonetheless beyond the scope of this dissertation.

and $\gamma^2 = \langle\delta'\delta\rangle^2 / \langle\delta'^2\rangle \langle\delta^2\rangle$ is the cross-correlation coefficient between the density and its slope[10]. Thanks to this factorization, integrating equation (2.70) over δ' yields the fully analytical expression

$$f_{\mathrm{up}}(\sigma) = p_{\mathrm{G}}(\nu_{\mathrm{c}})\frac{\mu}{\sigma}F(X)\,, \tag{2.72}$$

where p_{G} is a Gaussian with mean $\langle\nu\rangle = 0$ and variance $\mathrm{Var}(\nu) = 1$. For a constant barrier, the parameters μ and X are defined as

$$\mu \equiv \langle\delta'\,|\,\nu_{\mathrm{c}}\rangle = \nu_{\mathrm{c}}\,, \quad \text{and} \quad X \equiv \frac{\mu}{\sqrt{\mathrm{Var}(\delta'\,|\,\nu_{\mathrm{c}})}} = \Gamma\nu_{\mathrm{c}}\,, \tag{2.73}$$

with

$$F(x) \equiv \int_0^\infty \mathrm{d}y\, \frac{y}{x}\, \frac{e^{-(y-x)^2/2}}{\sqrt{2\pi}} = \frac{1+\mathrm{erf}(x/\sqrt{2})}{2} + \frac{e^{-x^2/2}}{x\sqrt{2\pi}}\,, \tag{2.74}$$

which is a function that tends to 1 very fast as $x \to \infty$, with correction decaying like $e^{-x^2/2}/x^3$. It departs from one by $\sim 8\%$ for a typical $\Gamma\nu_{\mathrm{c}} \sim 1$. Equation (2.72) can be written explicitly as

$$f_{\mathrm{up}}(\sigma) = \frac{\nu_{\mathrm{c}}e^{-\nu_{\mathrm{c}}^2/2}}{\sigma\sqrt{2\pi}}F(\Gamma\nu_{\mathrm{c}})\,, \tag{2.75}$$

where the first factor in the r.h.s. of equation (2.75) is the result of Press and Schechter, 1974, ignoring the factor of 2 they introduced by hand to fix the normalization. For correlated steps, their non-normalized result reproduces well the large-mass tail of $f(\sigma)$ (which is automatically normalized to unit and requires to correcting factor), but it is too low for intermediate and small masses. The upcrossing probability $f_{\mathrm{up}}(\sigma)$ also reduces to this result in the large mass limit, when $\Gamma\nu_{\mathrm{c}} \gg 1$ and $F(\Gamma\nu_{\mathrm{c}}) \simeq 1$. However, for correlated steps $f_{\mathrm{up}}(\sigma)$ is a very good approximation of $f(\sigma)$ on a larger mass range. For a ΛCDM power spectrum, the agreement is good for halo masses as small as $10^{12}\,\mathrm{M}_\odot\,h^{-1}$, well below the peak of the distribution. The deviation from the strongly correlated regime is parametrized by $\Gamma\nu_{\mathrm{c}}$, which involves a combination of mass and correlation strength: the approximation is accurate for large masses (small σ and large ν_{c}) or strong correlations (large Γ). Although Γ mildly depends on σ, fixing $\Gamma^2 \sim 1/3$ (or $\gamma \sim 1/2$) can be theoretically motivated (Musso and R. K. Sheth, 2014a) and mimics well its actual value for real-space Top-Hat filtering in ΛCDM on galactic scales. The limit of uncorrelated steps ($\Gamma = 0$), whose exact solution is twice the result of Press and Schechter, 1974, is pathological in this framework, with f_{up} becoming infinite. More refined approximation methods can be implemented in order to interpolate smoothly between the two regimes (Musso and R. K. Sheth, 2014b).

From equation (2.72), a characteristic mass $M_\star$ can be defined by requesting that the argument of the Gaussian be equal to one, *i.e.* $\nu_{\mathrm{c}} = 1$ or $\sigma(M_\star) = \delta_{\mathrm{c}}/D$. This defines $M_\star$ implicitly *via* equation (2.63) for an arbitrary cosmology. This quantity is particularly useful because $f_{\mathrm{up}}(\sigma)$ does not have well-defined moments (in fact, even its integral over σ diverges). This is a common feature of first passage problems (Redner, 2001), not a problem of the upcrossing approximation: even when the first-crossing condition can be treated exactly, and $f(\sigma)$ is normalized – it is a distribution function –, its moments still diverge. Therefore, given that the mean $\langle M\rangle$ of the resulting mass distribution cannot be computed, $M_\star$ provides a useful estimate of a characteristic halo mass. In chapter 3, I will revisit this subject to imposing larger tides. We will see that since the process remains Gaussian, it boils down to shifting the mean and the covariances.

---

[10] recalling that $\langle\delta'\delta\rangle = \sigma$ so that $\gamma^2 = 1/\langle\delta'^2\rangle$.
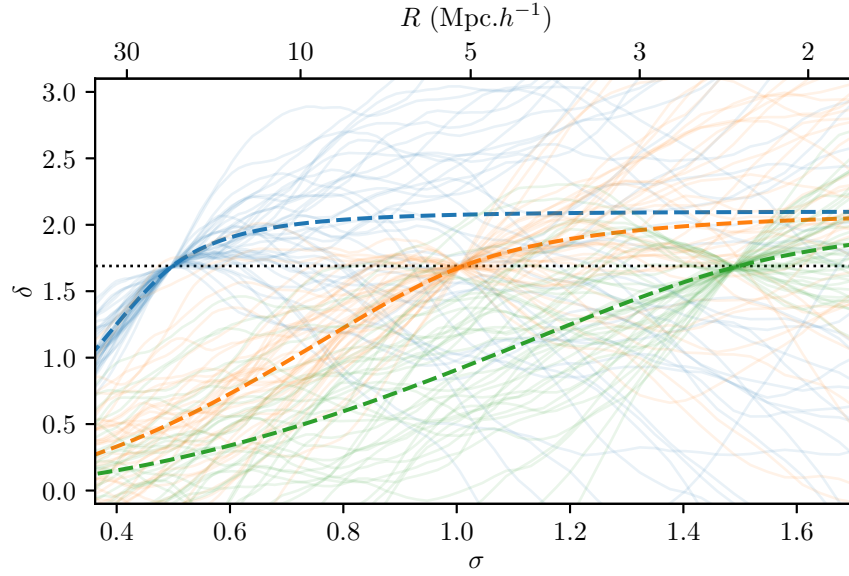
Figure 2.1.7: Excursion set trajectories constrained to $\delta = \delta_c$ (dotted line) at $\sigma = 0.5$ (blue bundle), $\sigma = 1$ (orange bundle) and $\sigma = 1.5$ (green bundle), dashed lines show mean trajectories. 93% (resp. 74%, 64%) of the first-crossing trajectories at $\sigma = 0.5$ (resp. $\sigma = 1$, $\sigma = 1.5$) are upcrossing.

2.1.5 The peak patch theory

The peak patch theory as introduced by Bond and Myers, 1996 aims at providing a more comprehensive description of the formation of dark matter halos. It is built as a combination of the excursion set theory (detailed in section 2.1.3), spherical collapse (detailed in section 2.1.2.2) and the Zel'dovich approximation (detailed in section 2.1.2.3). The theory aims to reproduce the mass distribution of dark matter halos using smoothing operations in the initial Lagrangian field, but also to predict the spatial distrbution in Eulerian space, using the Zel'dovich flow as an estimation of the displacement of the structures from their initial Lagrangian position.

The fundamental quantity in peak theory is the set of local maxima of the density field; therefore, peaks define a point process. Since the evolved density field is highly nonlinear, the peak constraint is generally applied to the initial (Lagrangian) Gaussian density field, with the assumption that the most prominent peaks should be in one-to-one correspondence with luminous galaxies or massive halos in the Universe. The theory is based on the study of the peaks of the initial density fields, which can be derived using the Kac-Rice formula (Kac, 1943; Rice, 1945). For a Gaussian random field δ, let $\{q_1, q_2, \ldots, q_p, \ldots\}$ be the Lagrangian position of point-particles such as centres of halos in some volume. The comoving Lagrangian density $n_g(q)$ of these point-particles is formally written as a sum of Dirac distributions

$$n(q) = \sum_p \delta_D^{(d)}(q - q_p).$$

(2.76)

In order to derive the number density in terms of the properties of the field, let us introduce the following variables

$$x \equiv \frac{\delta}{\sigma}, \quad x_i \equiv \frac{\partial_i \delta}{\sigma_1}, \quad x_{ij} \equiv \frac{\partial_i \partial_j \delta}{\sigma_2},$$

(2.77)

and σ_i are defined in section 2.1.7.1. Here we are using a different naming convention compared to the original paper Bardeen et al., 1986. For reference, their result can be obtained using the

following substitutions $x = \nu$, $x_i = \eta_i/\sigma_1$ and $x_{ij} = \xi_{ij}/\sigma_2$. Here we implicitly assumed the field to be smoothed at some scale R with any filter for which σ_2 is finite. Filters that verifies this property are notably the sharp-k filter and the Gaussian filter, while the physically motivated Top-Hat filter does not. Indeed, at large k, $W_{\mathrm{TH}}(kR) \sim -3\cos(kR)/kR$ so that the integrand of σ_2 becomes $\sim k^2 P(k)\cos^2(kR)$ which, for any power spectrum decaying more slowly than k^{-2}, has a UV-divergence.

In the case of critical points (maxima, saddle points and minima) the number density $n_{\mathrm{cp}}(\boldsymbol{q})$ can be entirely expressed in terms of x, x_i and x_{ij}. Without loss of generality, let us suppose that a critical point exists at the origin and let us derive the expression of the number density in its vicinity. Since the gradient at the critical points is null by definition, the gradient at a position $\boldsymbol{q}$ can be expressed using a Taylor expansion

$$x_i(\boldsymbol{q}) = \underbrace{x_i(0)}_{0} + \frac{\sigma_2}{\sigma_1}q_i x_{ij}(0). \tag{2.78}$$

This expression can then be plugged back in equation (2.76), provided that x_{ij} is invertible[11],

$$n_{\mathrm{cp}}(\boldsymbol{q}) = \delta_{\mathrm{D}}^{(d)}\left(\frac{\sigma_2}{\sigma_1}(x_{ij})^{-1}x_i\right), \tag{2.79}$$

where d is the number of dimensions. All the terms except for x_i can be taken out of the Dirac distribution so that the number density becomes

$$n_{\mathrm{cp}}(\boldsymbol{q}) = \left(\frac{\sigma_1}{\sigma_2}\right)^d |x_{ij}|\delta_{\mathrm{D}}^{(d)}(x_i). \tag{2.80}$$

In order to get the number density of a given kind of critical point, equation (2.80) need to be extended to take into account the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_d$ of the hessian of the field. The number density of maxima is given by

$$n_{\mathrm{max}}(\boldsymbol{q}) = \frac{|x_{ij}|}{R_*^d}\delta_{\mathrm{D}}^{(d)}(x_i)\prod_{j\leq d}\vartheta_{\mathrm{H}}(-\lambda_j). \tag{2.81}$$

Here R_* is the typical distance between extrema, see section 2.1.7.1. In more general terms, one can define the kind of a critical point by the sign of its (sorted) eigenvalues, also named the "signature". In three dimensions, maxima have a signature $---$, filament-type saddle points $+--$, wall-type saddle points $++-$ and minima $+++$. Noting k the number of negative eigenvalues, the number density reads

$$n_k(\boldsymbol{q}) = \frac{|x_{ij}|}{R_*^d}\delta_{\mathrm{D}}^{(d)}(x_i)\prod_{j\leq k}\vartheta_{\mathrm{H}}(-\lambda_j)\prod_{j>k}\vartheta_{\mathrm{H}}(\lambda_j). \tag{2.82}$$

The mean number density can be exactly calculated for a Gaussian field in two and three dimensions. In two dimensions, the mean number densities are

$$\langle n_{\mathrm{max}}\rangle = \langle n_{\mathrm{min}}\rangle = \frac{1}{8\sqrt{3}\pi R_*^2}, \tag{2.83}$$

$$\langle n_{\mathrm{sad}}\rangle = \frac{1}{4\sqrt{3}\pi R_*^2}. \tag{2.84}$$

---

[11]The extension to the case where x_{ij} is not invertible is provided in chapter 4.
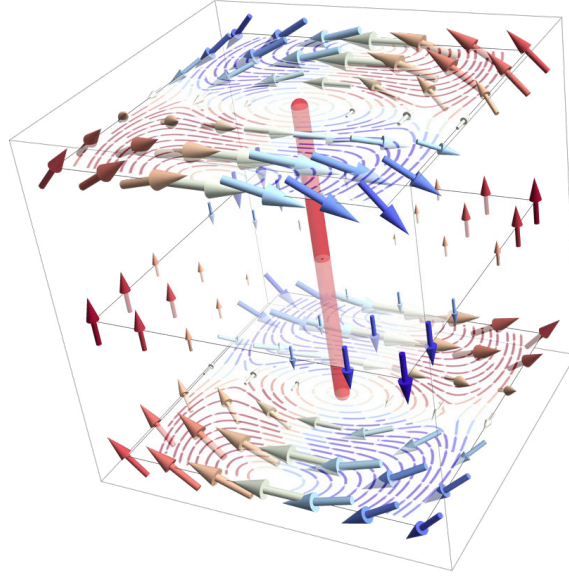
Figure 2.1.8: 3D representation of the mean angular momentum of halos (arrows) in the vicinity of a flattened filamentary structure (red cylinder) computed from first principle using conditional tidal torque theory. Close to the filament saddle point, the spins are aligned with the axis of the filament. Close to the nodes (at both ends of the filament), the spins become perpendicular to the axis of the filament and "rotate" around its axis.

In three dimensions, the mean number densities are

$$\langle n_{\mathrm{max}} \rangle = \langle n_{\mathrm{min}} \rangle = \frac{29\sqrt{15} - 18\sqrt{10}}{1800\pi^2 R_*^3}, \tag{2.85}$$

$$\langle n_{\mathrm{sadf}} \rangle = \langle n_{\mathrm{sadw}} \rangle = \frac{29\sqrt{15} + 18\sqrt{10}}{1800\pi^2 R_*^3}. \tag{2.86}$$

The "localized" number density, *i.e.* the number density at fixed height, can be easily derived from there

$$n_k(\boldsymbol{q}, \nu_0) = n_k(\boldsymbol{q})\delta_{\mathrm{D}}(x - \nu_0). \tag{2.87}$$

2.1.6   Tidal torque theory

The tidal torque theory has been developed to address the problem of the halo and galaxy angular momentum acquisition. In this model, proto-halo and proto-galaxies acquire their angular momentum by tidal torquing coming from the surrounding matter distribution (Hoyle, 1949; Peebles, 1969; Doroshkevich, 1970; S. D. M. White, 1984; Catelan and Theuns, 1996; Crittenden et al., 2001; Schäfer, 2009). Given a proto-halo that will later collapse, TTT provides an estimate of the growth of the angular momentum about the centre of mass, to the lowest non-vanishing order in perturbation theory. To do so, TTT links the evolution of the angular momentum (defined below) to the misalignment of the inertia tensor, which describes the spatial distribution of matter in the proto-halo, and the tidal tensor, which describes the tides from the larger scale environment. The upshot of the theory is that gravitational torques act to realign the inertia tensor of matter with the tidal tensor at larger scales, resulting in a net torque.

In general, the angular momentum $\boldsymbol{L}(t)$ of a rotating volume V, with velocity $\boldsymbol{v}(\boldsymbol{r}, t)$ and

density $\rho(\boldsymbol{r}, t)$ with respect to its centre of mass is

$$\boldsymbol{L}(t) \equiv \int_V \mathrm{d}^3 r\, \boldsymbol{r} \times \boldsymbol{v} \rho(\boldsymbol{r}, t). \tag{2.88}$$

Here I have implicitly assumed that the centre of mass is at the origin for the sake of simplicity and that the mean velocity is 0. Let me now assume an initial proto-halo of volume V_L in Lagrangian space, with mean density ρ_0. I have shown in section 2.1.2.3 that in the mildly non-linear regime, the time evolution of the particles can be described as a function of the displacement field using equation (2.56) and equation (2.61)

$$\dot{\boldsymbol{x}} = \dot{D}_+(t) \boldsymbol{\nabla} \psi. \tag{2.89}$$

One can then rewrite equation (2.88) for the proto-halo in terms of the peculiar velocity of its particles, *i.e.*

$$\boldsymbol{L}(t) = a^5 \rho_0 \int_{V_L} \mathrm{d}^3 q\, \boldsymbol{q} \times \dot{x} \approx a^5 \rho_0 D_+ \int_{V_L} \mathrm{d}^3 q\, \boldsymbol{q} \times \boldsymbol{\nabla} \psi(\boldsymbol{q}). \tag{2.90}$$

Let me further assume that the displacement field varies slowly in the proto-halo, so that

$$\boldsymbol{\nabla} \psi(\boldsymbol{q}) \approx \boldsymbol{\nabla} \psi(0) + \boldsymbol{q} \boldsymbol{\nabla} \boldsymbol{\nabla} \psi(0), \tag{2.91}$$

so that the displacement field can be expressed as function of the tidal shear $\boldsymbol{\nabla} \boldsymbol{\nabla} \psi$ tensor at the centre of mass. The expression of the angular momentum of the volume can be further simplified introducing the inertia tensor $\mathbf{I}$ (the quadrupole moment of the mass distribution) in Lagrangian coordinates

$$I_{ij} \approx \rho_0 a^3 \int_{L_V} \mathrm{d}^3 q\, \boldsymbol{q}_i \boldsymbol{q}_j. \tag{2.92}$$

For an initally uniform density, one can use the inertia tensor to describe the mass distribution as an ellipse whose axes are the eigenvectors of $\mathbf{I}$ and semi-axes are the square root of the eigenvalues. In the end, the i-th component of the angular momentum can be expressed as a function of time and the initial inertia tensor and tidal shear tensor

$$L_i(t) = a^2 \dot{D}(t) \epsilon_{ijk} I_{jl} \psi_{lk}. \tag{2.93}$$

Here I have used the fully anti-symmetric Levi-Civita tensor ϵ_{ijk} and the tidal shear tensor ψ_{ij}. Equation (2.93) shows that the angular momentum initially grows as $a^2(t)\dot{D}(t)$ which is $\sim t$ for a EdS Universe (Porciani et al., 2002). In addition, only the traceless parts of the inertia and the tidal shear enter equation (2.93), as the trace describes the compression (or expansion) of the proto-halo. Equation (2.93) also shows that $L(t)$ is null if the volume is spherical (so that the inertia tensor is symmetric) or is bounded by an equipotential surface (so that the tidal shear is symmetric). In the frame of the eigenvalues (t_1, t_2, t_3) of the tidal shear, equation (2.93) simply reads $L_i \propto (t_j - t_k) I_{jk}$, where i, j, k are cyclic permutations of 1, 2 and 3 (Porciani et al., 2002), e.g. $L_1 \propto (t_2 - t_3) I_{23}$. Tidal torquing is effective until the moment of the turnaround in the spherical collapse picture, because the collapse dramatically reduces the lever arms. After the collapse, the halo conserves the angular momentum it has accumulated until the turnaround.

More recently, Codis et al., 2012 suggested an extension of the TTT, coined "constrained TTT". The theory relies on the study of the primordial field, constrained to a large-scale filamentary structure, in which the tidal tensor as well as the inertia tensor become functions of space. Since the tidal tensor probes larger scales than the inertia tensor, the effect of large-scale structures act differentially on their typical orientations: the former tends to be aligned towards large overdensities (typically a node of the cosmic web), while the former is aligned to the local most massive structure (typically the nearest filament). On average, the upshot of the theory is that the

typical orientation of the angular momentum of the proto-halo, which measures the misalignment of the two tensors, is modulated by the cosmic web: the angular momentum is aligned with the filaments for small proto-halos, found close to the filament saddle point, and become perpendicular for larger ones, found close to the nodes (the spin rotates around the filament), as shown on figure 2.1.8. This is a typical example of the impact of the cosmic web on a galactic property.

This approach, where the initial conditions are constrained to take into account large-scale filamentary structures, will be further explored in the following of the dissertation.

2.1.7   Description of Gaussian random fields

In this section, I provide some useful variables that can be used to describe a Gaussian random field. In section 2.1.7.1, I define the variance of the field and its derivatives, as well as spectral parameters that encode the cross-correlation of the field and its derivatives. In section 2.1.7.2, I show how one can match smoothing scale when smoothing a field with different filters, in particular going from Gaussian filtering to Top-Hat filtering.

2.1.7.1   Spectral parameters of the field

When deriving quantities from the initial density field, it is of interest to quantify the variance of the field as they are natural scales for the rarity of events. This is usually done in terms of the generalized variance of the field and its derivatives and anti-derivatives

$$\sigma_i^2(R) = \frac{1}{2\pi^2} \int \mathrm{d}k \, k^2 P(k) k^{2i} W^2(kR) \,, \tag{2.94}$$

so that

$$\sigma_0^2 = \langle \delta^2 \rangle, \quad \sigma_1^2 = \langle (\nabla_i \delta)^2 \rangle, \quad \sigma_2^2 = \langle (\Delta \delta)^2 \rangle, \quad \sigma_3^2 = \langle (\Delta \nabla_i \delta)^2 \rangle, \tag{2.95}$$

where the gradient $\nabla_i \delta$ can be evaluated in any arbitrary direction. The evolution of $\sigma_0(R)$ for different filters is shown on figure 2.1.2, right panel.

Following closely Pogosyan et al., 2009b, let us introduce the characteristic scales of the field

$$R_0 = \frac{\sigma_0}{\sigma_1}, \qquad R_* = \frac{\sigma_1}{\sigma_2}, \qquad \tilde{R} = \frac{\sigma_2}{\sigma_3} \,. \tag{2.96}$$

These scales are ordered as $R_0 \geq R_* \geq \tilde{R}$. The first two have well-known meanings of typical separation between roots of the field R_0 and mean distance between extrema, R_* (Bardeen et al., 1986) and the third one, $\tilde{R}$ is, by analogy, the typical distance between inflection points. This also gives a motivation for the ordering of the scales: in 1D, there is at least one extrema between each root of a function and there is at least one inflection point between each root, so that the distance between consecutive roots is larger than the typical distance between peaks. As shown in section 2.1.5, the scales enter naturally the expressions of the number density of peaks (R_*) and anticipating the results to come, we can expect $\tilde{R}$ to enter any number density requiring the knowledge of the field and its third derivative, as will be the case in this chapter.

Let us define a set of spectral parameters that depend on the shape of the underlying power spectrum. Out of these three scales two dimensionless ratios may be constructed that are intrinsic parameters of the theory

$$\gamma \equiv \frac{R_*}{R_0} = \frac{\sigma_1^2}{\sigma_0 \sigma_2}, \qquad \tilde{\gamma} \equiv \frac{\tilde{R}}{R_*} = \frac{\sigma_2^2}{\sigma_1 \sigma_3} \,. \tag{2.97}$$

From the geometrical point of view γ specifies how frequently one encounters a maximum between two zero crossings of the field, while $\tilde{\gamma}$ describes, on average, how many inflection points are between two extrema. Using the results of section 2.1.1.4, one can rewrite σ_1^2 in terms of

$\langle\delta\nabla^2\delta\rangle$ so that γ and $\tilde{\gamma}$ are the cross-correlation coefficients between the field and its derivatives at the same point

$$\gamma = -\frac{\langle\delta\Delta\delta\rangle}{\sigma_0\sigma_2}, \quad \tilde{\gamma} = -\frac{\langle\boldsymbol{\nabla}\delta \cdot \Delta\boldsymbol{\nabla}\delta\rangle}{\sigma_1\sigma_3}. \tag{2.98}$$

These scales and scale ratios fully specify the correlations between the field and its derivative at the same point. For power-law spectra with Gaussian smoothing at the scale R, $R_0 = R\sqrt{2/(n+3)}$, $R_* = R\sqrt{2/(n+5)}$ and $\tilde{R} = R\sqrt{2/(n+7)}$ while $\gamma = \sqrt{(n+3)/(n+5)}$ and $\tilde{\gamma} = \sqrt{(n+5)/(n+7)}$. Note that the definition of equation (2.97) is not the same as the definition used in the excursion set theory (see e.g. equation (2.71)). In the following of the dissertation, unless stated otherwise, we will use the definition of equation (2.97). For the sake of completeness, let us remind here the definition of γ entering the excursion set theory, which we will distinguish from the definition above using the subscript "ES"

$$\gamma_{\mathrm{ES}}^2 = \frac{\langle\delta\delta'\rangle^2}{\langle\delta^2\rangle\langle\delta'^2\rangle}. \tag{2.99}$$

For power-law spectra with an index $n_{\mathrm{s}} < -1$ smoothed with a Top-Hat filter, $\gamma_{\mathrm{ES}} = (n_{\mathrm{s}} + 1)(n_{\mathrm{s}} + 3)/n_{\mathrm{s}}(n_{\mathrm{s}} + 5)$.

2.1.7.2 Matching smoothing scales

Figure 2.1.2, right panel shows the variance of the field smoothed with different filters. It shows that at the same scale, the Gaussian filter has a smaller variance than the Top-Hat filter. In order to study the same level of non-linearity, one has to establish a mapping between the smoothing scale. Using the definition of the variance of the field

$$\sigma^2(R) = \int_0^\infty \mathrm{d}k\, \frac{k^2 P(k)}{2\pi^2} W(kR)^2, \tag{2.100}$$

we have for a Top-Hat filter with a power-law power spectrum with spectral index n

$$\sigma_{\mathrm{TH}}^2(R) = 9 \times 2^{n-1}\frac{(n+1)R^{-n-3}\sin\left(\frac{n\pi}{2}\right)\Gamma(n-1)}{\pi^2(n-3)}, \tag{2.101}$$

while for a Gaussian filter it is

$$\sigma_{\mathrm{G}}^2(R) = \frac{R^{-n-3}\Gamma\left(\frac{n+3}{2}\right)}{4\pi^2}. \tag{2.102}$$

The field smoothed by a Top-Hat filter at scale R and a Gaussian filter at scale (R/α) have the same level of non-linearity if $\sigma_{\mathrm{TH}}(R) = \sigma_{\mathrm{G}}(R/\alpha)$, i.e.

$$\alpha(n)^{n+3} = 9 \times 2^{2-n}\frac{\sin\left(\frac{n\pi}{2}\right)\Gamma(n-1)}{(n-3)\Gamma\left(\frac{n+1}{2}\right)}. \tag{2.103}$$

For example $\alpha(-2) = \frac{6\sqrt{\pi}}{5} \approx 2.12$. For a ΛCDM power spectrum α becomes a weak function of the smoothing scale, for example $\alpha(R = 0.5\,\mathrm{Mpc}/h) = 2.14$ while $\alpha(R = 8\,\mathrm{Mpc}/h) = 2.12$. The evolution of α with the smoothing scale is shown on figure 2.1.9, which shows that at scales involved in galaxy formation $\alpha \approx 2.1$.
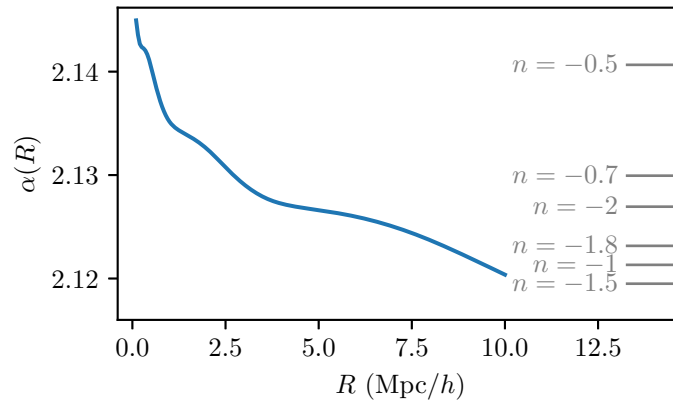
Figure 2.1.9: Ratio of the Gaussian smoothing scale to the Top-Hat smoothing scale to get the same level of non-linearity for a ΛCDM power spectrum $\sigma_{\mathrm{TH}}(R) = \sigma_{\mathrm{G}}(R/\alpha)$. Horizontal gray lines show the value of α for some power-law power spectra (which does not depend on R).

2.2   Galaxy formation

On large scales, the effect of baryonic processes is very small and baryons simply follow the dark matter dynamics. However, on small-to-intermediate scales, baryonic physics cannot be neglected and should be taken into account. This is particularly challenging, as the physics driving the evolution of the baryons is made of non-linear and highly coupled equations. To make things worse, the gravitational force and turbulence couple different scales together. A pragmatic approach to the problem of galaxy formation is to write numerical codes that simulate all relevant physical processes. In the simulated *in silico* Universe, one can then study the formation of galaxies to better understand observations and constrain their models.

In practice, the problem of galaxy formation could be rephrased as a Cauchy problem, where the initial conditions are set to a Gaussian random field according to the ΛCDM model, while coupled partial differential equations describe the interactions *via* the four fundamental forces. However, the different scales at play make a numerical treatment particularly difficult: the formation of a galaxy depends on its large scale environment on Mpc scales, its dynamical evolution is on kpc scales, while the evolution of its stars and its central Supermassive Black Hole (SMBH) act on sub-pc scales and these scales are coupled *via* the gravitational force.

The challenge for cosmological numerical simulations is then two-fold. First, since the resolution of numerical simulations is finite, effective models should be build to account for the unresolved physics. Second, numerical simulations should be able to capture processes at very different scales. Section 2.2.2 provides a description of the different physical processes involved in galaxy formation while focusing particularly on their implementation in the code RAMSES. Section 2.2.3 presents the set of equations solved and the numerical methods involved in their resolution.

2.2.1   Classical model of galaxy formation

In the classical model of galaxy formation, galaxies grow by the accretion of gas at the centre of the potential well of DM halos. The gas is initially distributed uniformly and traces the DM distribution (on scales larger than the Jeans length). Following the evolution of proto-halos, the gas first expands with the Hubble flow until turn-around. Let me first describe the physical state of the gas in the halo, before discussing implication on the mode of accretion.

Let me assume a cloud of monoatomic gas of mass M_{gas} in the potential of the DM halo, with mass M_{vir} and virial radius R_{vir}. If one assumes that the gas is in equilibrium, the virial theorem reads

$$2K + U + \Sigma = 0, \qquad (2.104)$$

where K is the kinetic (thermal) energy of the gas, U is the gravitational potential energy and Σ is the work of the external pressure forces. For an isothermal monoatomic gas ($\gamma = 5/3$) and assuming that the external pressure vanishes, we have

$$K = \frac{3 M_{\mathrm{gas}} k_{\mathrm{B}} T}{2 \mu \mathrm{m}_{\mathrm{p}}}, \qquad U = -\frac{3 G M_{\mathrm{gas}} M}{5r}, \qquad (2.105)$$

where μ is the mean molecular weight of the gas and m_{p} the proton mass. Here, I have assumed that the cloud has a radius r and spherical symmetry. If one introduces the circular velocity $V_c^2 = GM/r$, the temperature of the gas then reads

$$T = \frac{\mu \mathrm{m}_{\mathrm{p}}}{5 k_{\mathrm{B}}} V_c^2 = 2.4 \times 10^5 \mu \left(\frac{V_c}{100\,\mathrm{km\,s^{-1}}} \right)^2 \mathrm{K}, \qquad (2.106)$$

which defines the virial temperature. At the centre of the halo where the density is higher, gas slowly cools down from the inside-out and therefore looses its pressure support. This enables further gravitational collapse, which, in turn, leads to star formation.

Upon its entry in the halo, the accreted gas will encounter the hot halo gas at temperatures of the order of $2 \times 10^5\,\mathrm{K}$. If the accreted gas is cooler than the virial temperature, as is expected for primordial gas, a general expectation is that an accretion shock will form (see e.g. Binney, 1977; Bertschinger, 1985; Tozzi and Norman, 2001; Benson, 2010, and references therein), with a general conclusion that the shock occurs at a radius comparable to or slightly larger than the virial radius and the accreted gas will be shock-heated to the virial temperature.

On the contrary, if the cooling time is short compared to the dynamical time, the gas is able to flow into the centre of the halo without heating (S. D. M. White and Frenk, 1991). Based on 3D numerical simulations, it was confirmed that a significant fraction of the gas in low-mass galaxies has never been shock heated (see e.g. Kereš et al., 2005; Ocvirk et al., 2008; Kereš et al., 2009; Nelson et al., 2013) and reaches the galaxy through cold flows. Using an analytic treatment, Birnboim and Dekel, 2003 showed that if the cooling times are sufficiently short in the post-shock region, the shock looses its pressure support and becomes unstable, shrinking to smaller radii. This is expected for small mass halos $M < 10^{12}\,\mathrm{M}_\odot$, but also for more massive ones at $z > 2$. These cold flows have since been identified as a robust prediction of the ΛCDM model, consistently reproduced in different numerical codes with different subgrid models (Stewart et al., 2017).

2.2.2 Baryonic processes

This section provides an overview of the different phenomenon at play in galaxy formation. It is particularly focused on their numerical implementation and especially in Ramses. Section 2.2.2.1 describes how gas is cooled and heated. Section 2.2.2.2 describes how stars form and release energy as supernovæ. Section 2.2.3 details the different methods used to account for the cosmological context.

2.2.2.1 Gas cooling and heating

Following the results of section 2.1.2.2 halos are virialized structures that cannot collapse much further as their kinetic energy balances out their gravitational energy. In order to form galaxies at their centre, the gas need to be able to collapse further. This can only happen if the gas can get rid of its thermal energy, which happens mainly *via* cooling. In order to understand galaxy

formation, one need to compare the different timescales. The first timescale at play here is the cooling timescale

$$t_{\text{cool}} = \frac{E}{\dot{E}}. \tag{2.107}$$

The timescale associated with the expansion of the Universe is the Hubble timescale

$$t_{\text{H}} \sim H(z)^{-1}. \tag{2.108}$$

The timescale associated with the monolithic collapse of a pressure-less fluid is the free-fall, or dynamical, time

$$t_{\text{dyn}} \sim (G\rho)^{-1/2}. \tag{2.109}$$

There are then three scenarios. If $t_{\text{cool}} > t_{\text{H}}$, the Universe expands faster than the gas cool and no significant collapse can take place. If $t_{\text{dyn}} < t_{\text{cool}} < t_{\text{H}}$, the system evolves quasi-statically but the gas cannot cool efficiently to form galaxies. In practice, a succession of cooling followed by adiabatic contraction can happen at constant Jeans mass, but this does not lead to an efficient gravitational collapse. Finally if $t_{\text{cool}} < t_{\text{dyn}}$, the extra energy of the gas is quickly radiated away and gravitational collapse can happen. In this case, the loss of pressure following a temperature decrease is rapid so that the Jeans mass drops without giving the system a change to re-adjust its density. The drop in Jeans mass can lead to smaller structures being able to collapse gravitationally and hence to fragmentation. The precise study of the cooling and heating processes is therefore at the core of our understanding of galaxy formation. The main paths to cool the gas are the Compton cooling and radiative cooling.

Cooling processes

Compton cooling happens when a low-energy photon passes through an ionised thermal gas. In the process, photons and electrons exchange energy due to Compton scattering so that electrons lose energy to the radiation field, causing the gas to cool. It turns out that the change in the energy density of the radiation u_γ can be expressed as

$$\frac{\mathrm{d}u_\gamma}{\mathrm{d}t} = \frac{4k_b}{m_e c}\sigma_{\text{T}} n_e u_\gamma (T_e - T_\gamma), \tag{2.110}$$

where σ_{T} is the Thomson scattering cross-section, n_e is the electron number density, T_e is the electron fluid temperature and T_γ is the temperature of the radiation. In the case of cosmology, the photons come from the CMB and we have $T_e \gg T_\gamma$. Thus, we have a net gain of energy in the photons, and hence a net loss of energy in the electrons, which in turn will induce a net loss in the gas. This process is known as inverse Compton scattering. Using the fact that $u_\gamma = aT_\gamma^4$, where a is the expansion factor, the cooling rate per unit volume becomes

$$\mathcal{C}_{\text{Comp}} = \frac{4k_B T_e}{m_e c}\sigma_{\text{T}} n_e a T_\gamma^4 \propto n_e T_e (1+z)^4. \tag{2.111}$$

The cooling is therefore most efficient at high-redshift, but after reionisation since it requires free electrons to interact with.

$$t_{\text{cool,Comp}} \approx 2.3 \times 10^{12}(1+z)^{-4}\text{yr}, \tag{2.112}$$

which equals the Hubble time at about $z \sim 6$. After reionisation and before $z \sim 6$, the gas can cool efficiently using inverse Compton scattering.

Radiative cooling is a two-body radiative process that happens when a pair of atoms lose energy as a result of their interaction. The main processes of radiative cooling are listed in Table 2.3. The type of the interaction depends on the physical state of the electrons involved (free

Table 2.3: Various radiative transitions of importance in forming galaxies.

Type	Reaction	Name
Free-Free	$e^- + X^+ \rightarrow \gamma + e^- + X^+$	*Bremsstahlung*
Free-Bound	$e^- + X^+ \rightarrow X + \gamma$	Recombination
Bound-Free	$e^- + X \rightarrow 2e^- + X^+$	Collisional ionisation
Bound-Bound	$e^- + X \rightarrow e^- + X^*$	Collisional excitation

or bound). At high temperatures $T \gtrsim 10^7 \, \text{K}$[12] the dominating process in the fully ionised gas is *bremsstahlung*. The process is due to successive interactions between electrons and ions which bends the trajectories of the electrons, resulting in the emission of a radiation. The cooling rate per unit volume, assuming a charge number of unity and $n_i \sim n_e$, valid for a completely ionised hydrogen gas, is

$$\mathcal{C}_{\text{ff}} \approx 1.4 \times 10^{-23} \left(\frac{T}{10^8 \, \text{K}} \right)^{1/2} \left(\frac{n_e}{1 \, \text{cm}^{-3}} \right)^2 \text{erg s}^{-1} \, \text{cm}^{-3}. \tag{2.113}$$

This process gives us the behaviour of the cooling rate at high temperature $\mathcal{C}_{\text{ff}} \propto \sqrt{T}$.

At lower temperatures, several other processes become important. The first is collisional ionisation, in which atoms become ionised by collisions with other atoms. In the process, part of the kinetic energy of the atoms is used to ionise the electron. The second is recombination, in which an electron recombines with an ion, emitting a photon. The third is collisional excitation, in which atoms are first excited by collisions with electrons and then emit a photon on their transition to the ground state. These efficiency of the three processes depends strongly on the temperature as well as the chemical composition of the gas.

At temperatures below $10^4 \, \text{K}$, most of the electrons have recombined and cooling due to collisional excitation drops quickly. At this temperature, cooling is still possible, albeit smaller, e.g. by exciting the rovibrational levels of molecules. For metal enriched gas, CII and OI fine structure transitions contribute to the cooling (see e.g. Wolfire et al., 2003).

In practice, cooling is numerically treated using the cooling function

$$\Lambda \equiv \frac{\mathcal{C}}{n_{\text{H}}^2}, \tag{2.114}$$

where $\mathcal{C}$ is the total cooling rate (including all the mentioned processes) per unit volume and n_{H} is the number density of hydrogen atoms. The cooling function is usually derived in the collisional ionisation equilibrium limit, assuming that the relaxation times are fast enough. The cooling function also depends on the metallicity of the gas. Figure 2.2.1, left panel, shows the cooling function of a $Z = 0.02 Z_\odot$ gas. Most notably, the first peak of the cooling function is due to collisions involving H atoms, while the second peak is due to He and metals collisions and depends on the exact composition of the gas ; figure 2.2.1, right panel, shows the contribution of the different chemical species to the cooling function for a $Z = Z_\odot$ plasma. It is worth mentioning that some codes now compute out-of-equilibrium cooling rates for H and He, such as Grackle or KROME (Grassi et al., 2014; Smith et al., 2017).

Heating processes

In addition to the different cooling mechanism, an atom can also be ionised by absorbing a photon, a process called photoionisation. The presence of a radiation field can change the

---

[12]The temperature depends notably on the metallicity of the gas.
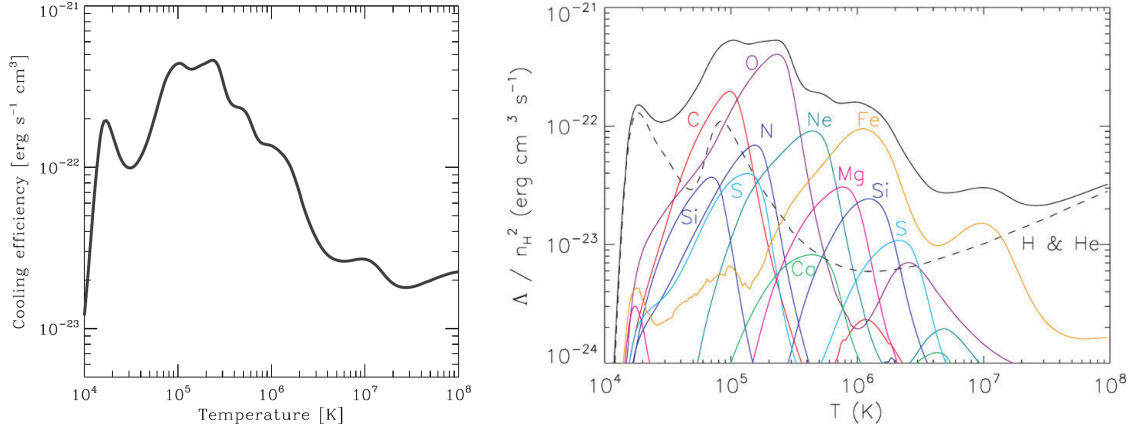
Figure 2.2.1: (*Left:*) Cooling function in the collisional ionisation equilibrium limit for a $Z = 0.02Z_\odot$ plasma. From Gnedin et al., 2015. (*Right:*) Cooling function for a $Z = Z_\odot$ plasma, indicating the contributions the each chemical species to the cooling function. From Wiersma et al., 2009. The dominant contribution to the cooling function at low temperature are H and He atoms. At $T \sim 10^5$ K, different metals dominate the cooling function, depending on the chemical composition of the gas. At large temperature, the cooling is dominated by Bremsstahlung.
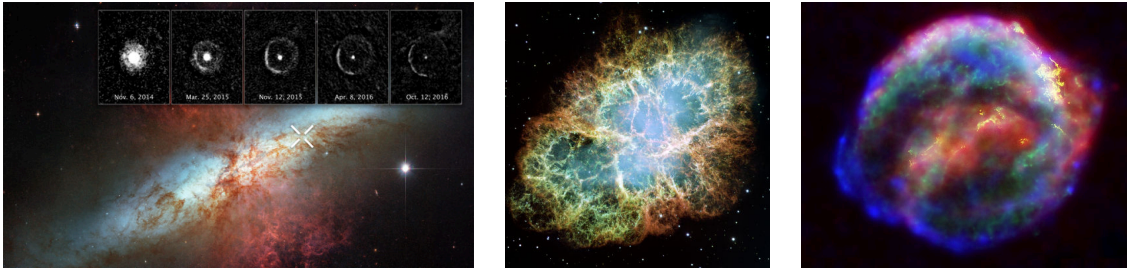
Figure 2.2.2: *Left:* HST image of M82 showing "light echoes" around a supernova bubble (from Yang et al., 2017) *Centre:* HST image of the crab nebula (remnant of SN1054). *Right:* Composite image of Kepler's supernova remnant (CXO, HST and Spitzer Space Telescope). Credits: NASA, ESA, JHU.

population of ions, which in turn can have an impact on the cooling rate of the gas. It can also heat the gas *via* photoionisation heating: an ionising photon is absorbed by an electron, part of the energy is used to ionise the electron and the surplus is transferred as kinetic energy. The photoionisation heating rate per unit volume is expected to be proportional to the intensity of the radiation field. Since the process is based on the ionisation of a an electron, it is most efficient at low temperatures where the gas is not fully ionised. In the presence of a UV radiation background of $J(\nu) = 10^{-22} \, (\nu_H/\nu) \, \mathrm{erg\,s^{-1}\,cm^{-2}\,sr^{-1}\,Hz^{-1}}$, a gas in ionisation equilibrium has a photoionisation heating that balances the cooling at temperature $T \lesssim 10^4 - 10^5$ K, depending on the gas density.

2.2.2.2 Stellar models

To understand the non-linear problem of galaxy formation and evolution, theorists use cosmological simulations of DM, describing the flow and collapse of baryonic star-forming gas either with directly coupled hydrodynamics or semi-analytic models. Strong feedback in galaxies is a vital ingredient in any model of galaxy evolution that comes even close to reproducing basic observables, such as the star formation history of the Universe, the stellar mass function of

galaxies, the Kennicutt–Schmidt relation, rotational velocities and outflows (e.g. Vogelsberger et al., 2013; Dubois et al., 2014; Hopkins et al., 2014; Schaye et al., 2015; Somerville and Davé, 2015).

Indeed, naive arguments would predict that star formation consumes stars over a few free-fall times, effectively depleting an entire galaxy in a few million years. Observations on the other side show that the process of star formation should be rather slow and inefficient. In addition, observations show the ubiquitous presence of large, massive outflows around galaxies (e.g. Cecil et al., 2001) of hundreds of km/s, that each release about 10^{51} erg. The origin of these outflows can be traced back to bubbles expanding around supernova remnants (see for example figure 2.2.2). Each exploding supernova releases large amounts of energy that are able to drive large-scale shocks, pushing gas outwards and leaving the shocked region heated and ionised.

Any simulation aimed at reproducing galaxies as we observe them must therefore include stars, but also track their evolution and their explosion. As of today, numerical simulations still struggle to consistently track the formation, evolution and destruction of stars.

A full treatment of the formation of stars would indeed require to resolve the physics at play in star formation: the collapse of molecular clouds into proto-stars, the ignition of their internal fusion, the accurate tracking of the winds, and eventually their explosion into supernovæ, etc. Some codes exist that do follow these processes (e.g. MESA, Paxton et al., 2011), yet they require a simulation of their own for each individual star. While these codes provide useful information to understand the evolution of small populations of star, they clearly cannot be scaled up to galaxy or, even worse, cosmological simulations.

Since it is unpractical to follow all stars that make up a galaxy in a large cosmological simulation, one should relax the goal to track each of them individually and track them as small populations instead. This is the current approach of most cosmological simulations (e.g. Hopkins et al., 2014; Dubois et al., 2016). Simulations are populated with "stellar particles" that represent a star population with a coherent formation time. The rate of star formation is usually given by a Kennicutt law (Kennicutt, 1998). This law links the local properties of the gas (local density and local free-fall time) to the star formation rate $\dot{\rho}_\star$

$$\dot{\rho}_\star = \varepsilon \frac{\rho}{t_{\text{ff}}}, \tag{2.115}$$

where ε is the star formation efficiency, which is usually set to a few percent in order to agree with observations, t_{ff} is the free-fall time and ρ the gas density. This law is the three-dimensional counterpart of the Schmidt law (Schmidt, 1959) that links the surface brightness of a galaxy to the observed surface density. More refined models have also been built in which the star formation efficiency becomes a function of additional properties of the gas, for example of their gravo-turbulent properties (Kimm et al., 2017; Trebitsch et al., 2017) following the results of Federrath and Klessen, 2012 which showed the role of turbulence in driving up or down the star formation efficiency. The stellar population is assumed to be sampled by its Initial Mass Function (IMF). Various models exist (Salpeter, 1955; Kroupa, 2001; Chabrier, 2003) that mostly differ on the low and high mass ends of the IMF, which will in turn have an impact on the feedback caused by the supernovæ. Indeed, IMF that are top-heavy have more stars on the massive end and will produce more supernovæ, boosting the efficiency of the stellar feedback.

After a few million years, the most massive stars start exploding into supernovæ. Doing so, they yield back metals as well as inject energy in the interstellar medium. Multiple models have been proposed to track how and where energy and momentum is fed back to the gas, as well as the total quantity returned (the yield). Let us briefly detail those used in Ramses as well as provide some hints of their pros and cons, following the lines of Rosdahl et al., 2017.

In the first supernova feedback models (Katz, 1992), all the energy was released as thermal energy in the gas surrounding the stars. This however had little effect on the star formation rate,
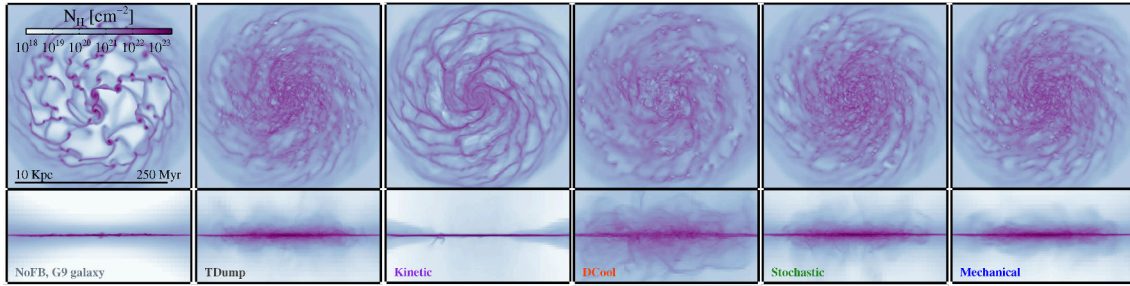
Figure 2.2.3: Effect of the different feedback models, from left to right: no feedback, direct thermal dumping, kinetic feedback, delayed cooling, stochastic feedback and mechanical feedback. From Rosdahl et al., 2017. Effective feedback leads to smoother, thick discs with larger outflows.

which leads to the so-called overcooling problem. In this model, the energy is diluted into a large amount of gas which in turn heats up a little bit. Because the energy has been spread over a large volume compared to the physical size of the supernova bubble, the gas is able to radiatively cool quickly. In practice, the cooling is so fast that the energy is usually radiated away in a few timesteps, and all the injected energy is lost before a significant fraction has been converted to e.g. kinetic energy. While the cooling of the gas is physically motivated (see section 2.2.2.1), the issue of this model is that the energy is spread into too large a volume, resulting in an overestimation of the cooling rates.

In order to solve the overcooling problem, different sub-grid models have been built, each of which aimed at reducing the amount of energy loss by (over-estimated) radiative cooling, which can be gathered into four classes.

In *kinetic feedback models*, a fraction of the supernova energy is directly injected as momentum in the gas (Navarro and S. D. M. White, 1993; Springel and Hernquist, 2003; Dubois and Teyssier, 2008). In *delayed cooling models*, radiative cooling is temporarily disabled in the cell containing the supernova remnant (Gerritsen, 1997; Stinson et al., 2013; Teyssier et al., 2013). In *stochastic feedback models*, the supernova energy is spread over time and space into fewer but more energetic explosions (Dalla Vecchia and Schaye, 2012; Rosdahl et al., 2017). Finally, *multiphase models* track the different phases of the gas (hot and cold), resulting in a more efficient feedback. A physically motivated approach to the problem would be to have different models for different grid resolutions and different states of the surrounding medium. This is the approach followed by the *mechanical feedback model* (Kimm and Cen, 2014; Kimm et al., 2015). The effect of these models is illustrated on figure 2.2.3, which presents a comparison of the different feedback models on the disc of an idealized galaxy. The study of the impact of the feedback models on galaxy formation is an active domain of research (e.g. Rosdahl et al., 2017; Kimm et al., 2017; Nelson et al., 2019).

2.2.3 Numerical simulations

While the formation of the large-scale structures of dark matter halos can be studied to some extent from first principles, as detailed in section 2.1, the complex baryonic physics involved in galaxy formation make the task much more complex on smaller scales. This is usually dealt with numerical simulations. The intrinsic multi-scale nature of the phenomenon involved in galaxy formation is however challenging to any numerical treatment, as was already underlined in the previous section, as very different scales are coupled. For example, sub-kpc scales involved in galaxy formation are coupled to the large-scale hydrodynamical evolution of the gas by powerful feedback events, which may disrupt the gas at hundreds of kpc. This in turn will impact the inflow of gas and couple back to feedback.

Let us illustrate this scale-coupling problem with some back-of-the-envelope calculation to estimate the number of resolution element required to resolve galaxies and the cosmic web at the same time. In order to accurately capture the evolution of a galaxy in its environment, the size of the simulated Universe should be at least an order of magnitude larger than the maximum distance travelled by the particles ending up in the central galaxy. Assuming that the galaxy is a Milky-Way like progenitor, its initial Lagrangian patch has a size of the order of a few Mpc so that the box size should at least be a few tens of Mpc. At the same time, in order to resolve the scale-height of a disk galaxies $h \sim 1\,\mathrm{kpc}$ accurately, cell sizes should be at least an order of magnitude smaller which sets the resolution to about $50\,\mathrm{pc}$. On a regular lattice, the number of cells we would therefore be

$$N_{\mathrm{cell}} \approx \left(\frac{50\,\mathrm{Mpc}}{50\,\mathrm{pc}} \right)^3 = 10^{18}\,\mathrm{cell}. \tag{2.116}$$

In a very simple simulation that only stores the physical state of a monoatomic neutral gas (density, pressure, velocities) in double-precision floats, each cell would require $20\,\mathrm{o}$ of storage. In total, storing the state of the gas on the full grid would therefore require $2 \times 10^{19}\,\mathrm{o} = 20\,\mathrm{Eo}$. For the sake of comparison, setting each cell to 0 would require at least $200\,\mathrm{yr}$ on a $3\,\mathrm{Ghz}$ single processor[13]. On the fastest currently-available super computer[14], using all the $2\,000\,000$ cores at the same time, it would still take more than a day for a single update. If we make a very conservative assumption that each cell is updated once per timestep and that a timestep corresponds to $1\,\mathrm{Myr}$ in the simulation, running a cosmological simulation on the fastest super computer for $14\,\mathrm{Gyr}$ would take $60\,\mathrm{yr}$ to complete. This approach is obviously not practical so that alternative approaches have been devised.

I first present the set of equations that numerical simulations have to solve in section 2.2.3.1. Section 2.2.3.2 presents the two approaches used in astrophysics to solve the scale-separation problem, focusing in particular on finite volume methods. Section 2.2.3.3 details how the hydrodynamical equations are solved in finite volume methods. Section 2.2.3.4 details how Poisson's equation is solved on a grid. Finally, section 2.2.3.5 presents the modifications required to take into account cosmological expansion, how the initial conditions are set and also presents the different state-of-the-art cosmological simulations at the time of the writing of this dissertation.

2.2.3.1 Hydrodynamical equations

In the context of cosmological astrophysical simulations, the scales considered are much larger than the mean free path λ

$$\lambda \sim \frac{\mathrm{m_p}}{\sigma \rho} = 7.4 \times 10^{-5} \left(\frac{\rho}{1\,\mathrm{m_p/cm^3}} \right)^{-1} \mathrm{pc}, \tag{2.117}$$

where $\mathrm{m_p} \approx 1.67 \times 10^{-27}\,\mathrm{kg}$ is the proton mass, $\sigma = 3.5 \times 10^{-20}\,\mathrm{m^2}$ is the Hydrogen-Hydrogen collisional cross section and ρ is the gas density. The mean free-path is below $1\,\mathrm{pc}$ as long as the density is larger than $1 \times 10^{-4}\,\mathrm{m_p/cm^3}$. As of today, no cosmological simulation reaches sub-parsec resolutions in regions with such low densities so that the equation describing the gas can be well approximated in the fluid limit.

Assuming that the gas is described by a pressure p, a density ρ, a velocity $\boldsymbol{v}$ and a specific

---

[13]Assuming that the processor can update memory once per cycle.
[14]DOE/SC/Oak Ridge National Laboratory, United States. Data from top500.org.

internal energy $\mathcal{E}$, the evolution of the gas is described by

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) = 0, \tag{2.118}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \boldsymbol{\nabla}) \, \boldsymbol{v} = - \left(\boldsymbol{\nabla} \Phi + \frac{\boldsymbol{\nabla} p}{\rho} \right), \tag{2.119}$$

$$\frac{\partial}{\partial t} \left[\rho \left(\frac{v^2}{2} + \mathcal{E} \right) \right] + \boldsymbol{\nabla} \cdot \left[\rho \left(\frac{v^2}{2} + \frac{P}{\rho} + \mathcal{E} \right) \boldsymbol{v} \right] - \rho \boldsymbol{v} \cdot \boldsymbol{\nabla} \Phi = \mathcal{H} - \mathcal{C}, \tag{2.120}$$

$$\nabla^2 \Phi = 4\pi G \rho_{\text{tot}}. \tag{2.121}$$

$\mathcal{H}, \mathcal{C}$ are the heating and cooling rates per unit volume, as described in section 2.2.2.1, ρ_{tot} is the total density accounting for the fluid, DM, stars and SMBHs. Φ is the gravitational potential. For an ideal gas with adiabatic index γ, the system of equation is closed by the perfect gas Equation of State (EoS)

$$P = \rho(\gamma - 1)\mathcal{E}. \tag{2.122}$$

These equations correspons to the equation of conservation of mass (equation (2.118)), linear momentum (equation (2.119)) and internal energy (equation (2.120)). The system is closed by the Poisson equation (equation (2.121)) that connects the density to the Newtonian gravitational potential.

The heating term is due to photoionisation where an atom of fluid is ionised by a photon. Ionising photons can originate from the UV background (see e.g. Haardt and Madau, 1996), stars or Active Galactic Nucleis (AGNs).

Equations (2.118)–(2.121) do not have an exact solution in the general case, so that a numerical treatment is required.

### 2.2.3.2   Finite-mass and finite-volume methods

Multiple methods have been developed to solve the set of equations (2.118)–(2.121), but in the context of cosmological simulations two main methods emerged that provide a practical solution to the scale-separation problem. They can be grouped in two main categories.

SPH simulations

Smooth Particle Hydrodynamics (SPH) simulations are based on a mass discretization of the fluid. The fluid is described as a set of fixed-mass macro-particle, whose interactions are described by the Lagrangian version of equations (2.118)–(2.120). In order to solve equation (2.121), the total density is interpolated on a grid, the equation is solved. Finally, the potential — or its derivatives — is interpolated at the particles' location. More details are provided in section 2.2.3.4. Each fluid particle has a variable "smoothing length" that depends on the density of the fluid *via* $\rho \sim m/r^3$, where m is the mass of the fluid particle. The exact normalisation depends on the choice of a kernel. The obvious advantage of this approach is to provide an accurate description of the Lagrangian evolution of the gas, while Eulerian quantities can be approximated by projecting particles onto an arbitrary mesh. This last step can easily be done in post-processing. This is the approach used in GADGET (Springel, 2005), Gasoline (Wadsley et al., 2004), Gizmo (Hopkins, 2015). In its simplest form (all particles have the same mass), the scale-separation problem is addressed by adapting the smoothing-length to match the local density.

AMR simulations

Adaptive Mesh Refinement (AMR) simulations are finite-volume methods. The evolution of the gas is described in an Eulerian framework. Equations (2.118)–(2.120) are solved on a fixed arbitrary grid. In order to capture the multi-scale evolution of the gas relevant to astrophysical phenomenon, the grid is adaptively refined following arbitrary criteria. Commonly used criteria are the following

(a) Patch-based AMR (ᴇɴᴢᴏ).   **(b)** Octree AMR (Rᴀᴍsᴇs).
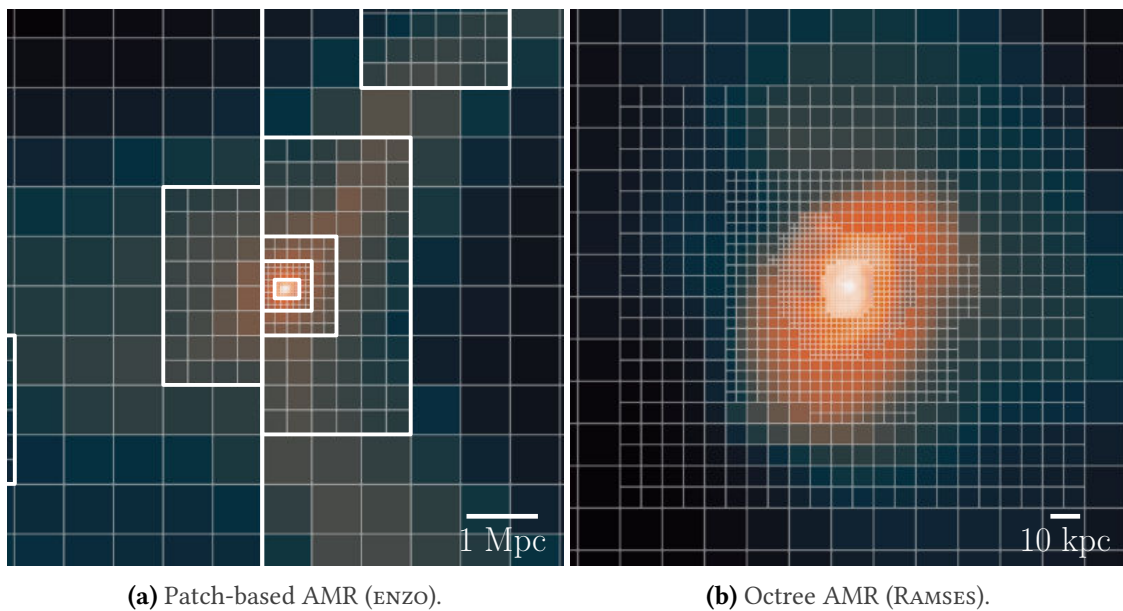
Figure 2.2.4: Plot of cosmological simulations with cell boundaries annotated in thin white lines. (a): patch-based codes divide space using a set of nested grids (thick white rectangles) made of an arbitrary number of cells in each dimension. (b): octree codes divide space using nested octs made of 8 cells, allowing a finer control of the grid structure.

(a) Domain decomposition of the unit-square   **(b)** Buffer zones
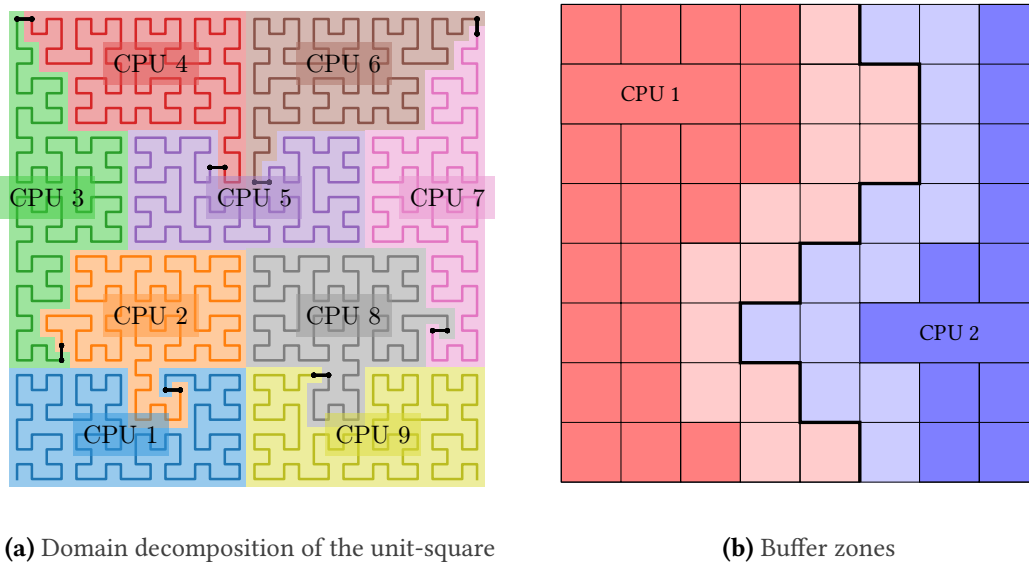
Figure 2.2.5: (a) Domain decomposition of the unit-square for a 32^2 grid over nine domains using the Hilbert space-filling curve, shown as the continuous line. (b) Buffer zones built at the interface between domains. The thick black line marks the boundary of the spatial decomposition between CPU 1 and CPU 2. CPU 1 owns all the red cells while CPU 2 owns the blue ones. There is a one-cell-thick buffer zone outside of each domain that a CPU can access (CPU 1 has access to light blue cells and CPU 2 has access to light red cells). This ensures that each cell has access to all its $26 = 27 - 1$ direct neighbours (for example to compute spatial gradients).

- Semi-Lagrangian criterion: a cell is refined if its mass exceeds a fixed mass $M_{\text{threshold}}$. This is commonly used in cosmological simulations in order to have cells of similar masses (but of different sizes), so that overdense regions (e.g. galaxies) are more refined than under-dense regions (e.g. cosmological voids).
- Jeans criterion: a cell is refined if its size exceeds the Jeans length $\lambda_{\text{J}} = c_s\sqrt{\pi/G\rho}$, where $c_s^2 = \gamma P/\rho$ is the local sound speed. This is commonly used to resolve the gravitational collapse (e.g. the collapse of molecular clouds in star forming regions).
- Pressure or density gradients: a cell is refined if the pressure or density gradient exceed some fraction of the quantity itself. This is commonly used to resolve shock fronts (e.g. Supernova (SN) blasts).

In addition, AMR codes can further be split into patch-based codes and octree codes. Patch-based codes (e.g. Enzo (Bryan et al., 2014), see figure 2.2.4a) use a nested hierarchy of rectangular patches of increasing resolution. The building blocks of the computational grid are therefore rectangular patches of various sizes, whose positions and aspects ratio are optimised with respect to flow geometry, speed and memory constraints in order to represent regions of increasing resolutions. Space is divided in nested rectangular patches (thick white lines) made of cells of fixed resolution.

Octree codes' building blocks are octs (e.g. Art (Kravtsov et al., 1997), Ramses (Teyssier, 2002)). An oct is an 2x2x2 set of cells, where each of the 8 cells is either a leaf cell or is itself an oct, as illustrated in figure 2.2.4b. The resulting grid follows complex flow geometry more closely, at the price of a data management which is more complicated than patch-based AMR. High density regions are followed by fine cells (as in the centre of the plot), while less dense regions have coarser cells.

These two strategies enable the code to partition space. In order to compute the time evolution of the hydrodynamical quantities, the codes then have to solve the so-called Riemann problem at cell boundaries, as described in section 2.2.3.3.

In order to increase the computation power wielded by numerical simulations, most of the numerical codes are now parallelized to run on multiple cores at the same time. They now routinely run on hundreds or even thousands of cores. In this context numerical codes have to be optimised to best balance the computation weight between each computation domain, while trying to minimise the number of communications. This is further complicated by the fact that for AMR codes, the grid is non-uniform so that there are no obvious space decomposition that will balance the cells evenly between all domains.

In order to solve this issue, AMR codes usually use space-filling curves. Space filling curves are bijective functions from 1d space to the 3d unitary cube, providing a unique index to each cell in the simulation. In addition, they should also conserve locality so that two cells that are close should have a close index. Using such a space-filling curve, the load balancing problem becomes a simple problem of sharing a set of N cells evenly between M domains. In Ramses the space-filling curve used is the Hilbert curve, as illustrated on figure 2.2.5a. Each computing unit has access to the list of the indexes on the Hilbert curve that separate the different domains, represented as black dashed lines on figure 2.2.5a, so that it can easily compute to which domain each cell belongs. This method is also a very efficient way to encode the volumetric partition of space into $M\log_2(N)$ bits. For example in a simulation with 20 levels of refinement and 4096 processors, the information about the spatial partitioning can be encoded optimally on $\log_2((2^{20})^3) \times 4096 \approx 250\,\text{kbit} = 31\,\text{kio}$.

Once the space has been decomposed between M domains, boundary regions are constructed at the interface between contiguous domains. The thickness of the boundary region depends on the order of spatial derivatives involved in the evolution equations. In cases where the hydrodynamical solver is using first-order finite differences, as is the case with Ramses, a 1-oct-thick layer is built

at the interface between each domain, as illustrated on figure 2.2.5b.

2.2.3.3 The Riemann problem and Godunov solvers

Let us consider the Riemann problem with initial left and right values values $U = U_1$ for $x \leq 0$ and $U = U_r$ for $x > 0$. The state vector U follows a conservation equation

$$U_t + F(U)_x = 0, \tag{2.123}$$

where F is the flux vector and subscripts indicate partial derivatives relative to the variable (e.g. $U_t = \partial U / \partial t$). Introducing the Jacobian matrix

$$\mathbf{A} = \frac{\partial F}{\partial U},$$

we can rewrite equation (2.123) into its conservative form

$$U_t + \mathbf{A}(U)U_x = 0. \tag{2.124}$$

The Riemann problem

The Riemann problem is the initial value problem of equation (2.124) with piece-wise initial conditions. In general, the equation does not accept an analytical solution, so that one need to design a numerical solver. Let us now focus Euler equation in 1D. U can be written using the conservative formulation of equations (2.118)–(2.120) with an ideal gas EoS

$$U = \begin{pmatrix} \rho \\ \rho u \\ \rho \left(\frac{u^2}{2} + \mathcal{E} \right) \end{pmatrix} \quad \text{and} \quad \mathbf{A}(U) = \begin{pmatrix} 0 & 1 & 0 \\ -\frac{1}{2}(\gamma - 3)u^2 & (3 - \gamma)u & \gamma - 1 \\ \frac{1}{2}(\gamma - 2)u^3 - \frac{a^2 u}{\gamma - 1} & \frac{3 - 2\gamma}{2}u^2 + \frac{a^2}{\gamma - 1} & \gamma u \end{pmatrix}, \tag{2.125}$$

where we have used the sound speed $a = \sqrt{\frac{\gamma p}{\rho}}$. We can also use the flux vector to have

$$F(U) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{pmatrix}, \tag{2.126}$$

where we have used the total energy E per unit volume

$$E = \rho \left(\frac{1}{2}u^2 + \mathcal{E} \right). \tag{2.127}$$

The different waves propagating at the interface are described by the eigenvalues and eigenvectors of the Jacobian $\mathbf{A}$. In the 1D case, the eigenvalues are $\lambda_1 = u - a$, $\lambda_2 = u$, $\lambda_3 = u + a$ describing three waves propagating downstream, with the stream and upstream. The wave associated with λ_2 is a contact discontinuity characterized with a constant pressure and velocity. The two waves associated with λ_1 and λ_3 are either rarefaction waves (smooth) or shock waves (discontinuities). Rarefaction waves are characterized by a smooth change of ρ, u and p across the front. On the contrary, shock waves are characterized by a jump of ρ, u and p which are described by the Rankine–Hugoniot jump condition. Let us introduce the Mach number

$$\mathcal{M} \equiv \frac{u}{a} = \sqrt{\frac{\rho u^2}{\gamma p}}. \tag{2.128}$$
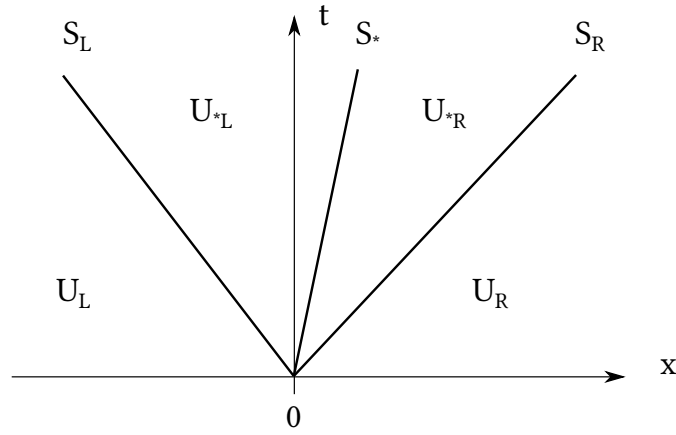
Figure 2.2.6: HLLC approximate Riemann solver. Solution in the star region consist of two constant states (U_{*L} and U_{*R}) separated from each other by a middle wave (S_*). The left and right state are separated from the star region by two waves (S_L and S_R). Adapted from Toro, 2009.

The Mach number is the ratio of the velocity to the sound speed. It can also be interpreted as the ratio of ram-pressure ρu^2 to thermal pressure. Denoting the pre- and post-shock regions with subscript 1 and 2, the Rankine-Hugoniot jump conditions read

$$\frac{\rho_2}{\rho_1} = \frac{u_1}{u_2} = \frac{(\gamma + 1)\mathcal{M}_1^2}{(\gamma - 1)\mathcal{M}_1^2 + 2}, \tag{2.129}$$

$$\frac{p_2}{p_1} = \frac{2\gamma\mathcal{M}_1^2 - (\gamma - 1)}{\gamma + 1}. \tag{2.130}$$

These conditions also imply the temperature jump

$$\frac{T_2}{T_1} = \frac{\left[(\gamma - 1)\mathcal{M}_1^2 + 2\right]\left[2\gamma\mathcal{M}_1^2 - (\gamma - 1)\right]}{(\gamma + 1)^2\mathcal{M}_1^2}. \tag{2.131}$$

Further details can be found in Toro, 2009, p.87–91.

Riemann solvers

The task of Riemann solvers is to solve the Riemann problem. One such solver is the HLLC solver (Toro et al., 1994) that is itself an extension of the HLL solver (Harten et al., 1983). The solver approximates the evolution of the contact discontinuity with the three waves described above (rarefaction, entropy and shock waves) that separate four states, as illustrated on figure 2.2.6. The flux is then computed using the conservation equations and the properties of the different contact discontinuities at the interface between each state, see Toro, 2009, chapter 10 for more details. The HLLC solver is frequently used in astrophysical setups, as it is very stable albeit quite diffusive.

Godunov solvers

Godunov solvers is a class of conservative numerical schemes first described by Godunov, 1959. It is based on a three-step algorithm[15]

1. *Reconstruction:* the value of the state variable is interpolated at cell faces using the values at the centre of the cells with a *slope limiter*. The slope limiter prevents spurious oscillations

---
[15]The first step is actually due to van Leer, 1984.

to appear in the solution.[16]

2. *Evolution:* the Riemann problem is solved using the interpolated values at the faces. This is the physical step. This step can be solved using either an exact Riemann solver or an approximate one (e.g. the HLLC solver described in the previous paragraph).

3. *Averaging:* the value of the state variable is updated using the flux of the Riemann solver.

While the original Godunov scheme was first order in space and time, higher order methods have since been introduced (e.g. MUSCL-Hancock (van Leer, 1984), PLM (Colella, 1985)). In the end the Godunov scheme outputs a flux that can readily be used to update the state vector. For MUSCL schemes, the state vector is updated using

$$U_i^{n+1} = U_i^n + \frac{\Delta t}{\Delta x} \left(F_{i-1/2}^{n+1/2} - F_{i+1/2}^{n+1/2} \right). \tag{2.132}$$

Here i is the index of the cell and n is the timestep ; the flux F is computed at half time steps $(n + 1/2)$ at cell boundaries.

2.2.3.4 Poisson solvers

In addition to fully solve the equations of hydrodynamics, special care should be taken for the gravitational force. Indeed, the gravitational force has infinite range so that each cell and particle in the simulation are gravitationaly coupled to any other. There are multiple methods to compute the gravitational force, either by solving Poisson equation or using so-called "direct methods". The former is based on Poisson equation (2.121) while the latter uses the force equation

$$F_i = -\sum_{j \wedge i} \frac{G m_i m_j}{\|r\|^3} r. \tag{2.133}$$

Direct methods

Direct methods compute for each particle (or for each gas cell) the force due to all other particles (or cells). While this method gives exact results, it suffers from performance issue since it scales as $\mathcal{O}(N^2)$, where N is the number of massive elements (particles, cells, …).

PM methods

In Particle Mesh (PM) methods, one first solves the Poisson equation to compute the potential, then uses the gradient of the potential to compute the gravitational force. The DM, star, black hole and gas density are projected onto a common mesh[17]. The Poisson equation is then solved on the grid and the gravitational acceleration is computed using the gradient of the potential $F = -\nabla \Phi$. There are two notable techniques to solve the Poisson equation: the multi-grid approach and the conjugate gradient approach.
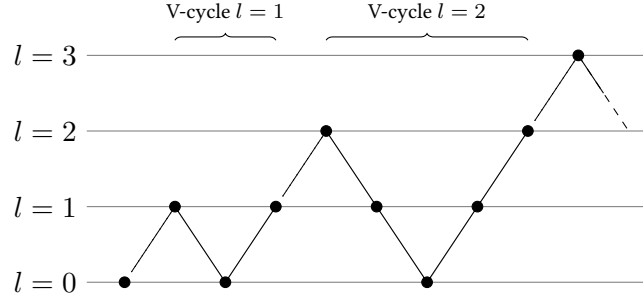
In the multi-grid approach, the Poisson equation is solved iteratively using a succession of "V cycles". This is illustrated in figure 2.2.7. At first, an exact solution is found at the coarsest level $l = 0$. Then the algorithm goes to the next finer level $l = 1$. An approximate solution is found using the coarser solution, then corrected using the information at coarser levels. The algorithm repeats itself until $l = l_{max}$.

The conjugate gradient method is a general method to find the solution of a linear problem

$$\mathbf{A} \cdot \boldsymbol{X} = \boldsymbol{B}.$$

---

[16]In practice, the slope limiter reduces the order of the scheme to 1 around discontinuities but increases the order of the method in smooth regions.

[17]The mesh is usually the AMR grid.

Figure 2.2.7: Illustration of the multigrid method. The multigrid algorithm starts at level $l = 0$ where it solves the Poisson equation exactly. It then proceeds to level 1, does a V-cycle, proceeds to level 2, does a V-cycle, ...

On a discrete grid, one can reformulate the Poisson equation into a simple linear problem, presented here for the one-dimensional case, with

$$
\mathbf{A} = \frac{1}{2\Delta x}
\begin{pmatrix}
-2 & 1 & 0 & \cdots & 0 \\
1 & -2 & 1 & \ddots & \vdots \\
0 & 1 & -2 & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & 1 \\
0 & \cdots & 0 & 1 & -2
\end{pmatrix},
\tag{2.134}
$$

$\boldsymbol{X} = \{\phi_i\}$ and $\boldsymbol{B} = 4\pi G\{\rho_{i,\text{tot}}\}$. The exact solution is found in N iterations for an AMR grid with N cells. In practice the error on the solution decreases with each step so that a simple convergence criterion is generally used to stop the iteration earlier, typically after a few hundreds iterations. Any iterative method can be used in place of the conjugate gradient method to solve the linear set of equations, for example the Gauss-Seidel method (which is the default method in RAMSES).

#### 2.2.3.5   Cosmological simulations

In this section, we detail the modifications to numerical code that are usually implemented to account for the cosmological expansion of the Universe and the initial conditions. Here, we focus especially on simulations with box sizes large enough to capture the large-scale structures of the Universe ($\gtrsim 50\,\mathrm{Mpc}/h$), with enough resolution to capture galactic scales ($\Delta x_{\min} \lesssim 1\,\mathrm{kpc}$) and that include at least star formation, SN feedback, SMBH formation, AGN feedback heating and cooling of the gas, gravity and dark matter, the whole simulation being evolved in an expanding Universe.

Accounting for the cosmological expansion

In RAMSES, cosmology is accounted for by the following change of "super-comoving" variable

$$
\mathrm{d}\tilde{t} = H_0 \frac{\mathrm{d}t}{a^2}, \quad \tilde{x} = \frac{1}{a}\frac{x}{L}, \quad \tilde{\rho} = a^3 \frac{\rho}{\Omega_\mathrm{m}\rho_\mathrm{c}}, \quad \tilde{p} = a^5 \frac{p}{\Omega_\mathrm{m}\rho_\mathrm{c}H_0^2 L^2}, \quad \tilde{\boldsymbol{v}} = a\frac{\boldsymbol{v}}{H_0 L}.
\tag{2.135}
$$

These variables have been introduced by Martel and Shapiro, 1998. Here H_0 is the Hubble constant, Ω_m is the matter density, L is the box size and ρ_c is the critical density. In these
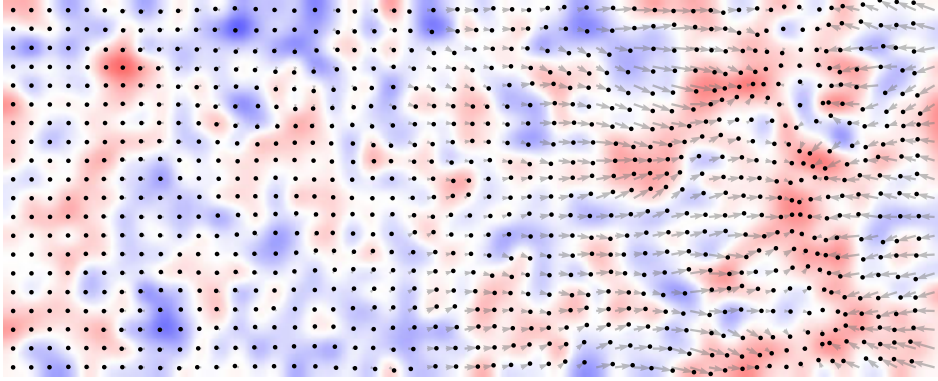
Figure 2.2.8: Location of particles following the Zel'dovich flow. Arrows indicate the direction of motion. Time increases linearly going from left ($t = 0$) to right (first shell-crossing). The background shows the initial overdensity field (red is overdense). These positions provide the initial conditions to numerical simulations.

variables, equations (2.118)–(2.121) become

$$\frac{\partial \tilde{\rho}}{\partial \tilde{t}} + \tilde{\boldsymbol{\nabla}} \cdot (\tilde{\rho} \tilde{\boldsymbol{v}}) = 0, \tag{2.136}$$

$$\frac{\partial (\tilde{\rho} \tilde{\boldsymbol{v}})}{\partial \tilde{t}} + \tilde{\boldsymbol{\nabla}} \cdot (\tilde{\rho} \tilde{\boldsymbol{v}} \times \tilde{\boldsymbol{v}}) = -\left(\tilde{\boldsymbol{\nabla}} \tilde{p} + \tilde{\rho} \tilde{\boldsymbol{\nabla}} \tilde{\phi} \right), \tag{2.137}$$

$$\frac{\partial (\tilde{\rho} \tilde{e})}{\partial \tilde{t}} + \tilde{\boldsymbol{\nabla}} \cdot \left(\tilde{\rho} \tilde{\boldsymbol{v}} \left[\tilde{e} + \frac{\tilde{p}}{\tilde{\rho}} \right] \right) = -\tilde{\rho} \tilde{\boldsymbol{v}} \cdot \tilde{\boldsymbol{\nabla}} \tilde{\phi}, \tag{2.138}$$

$$\tilde{\nabla}^2 \tilde{\phi} = \frac{3}{2} a \Omega_{\mathrm{m}} (\tilde{\rho} - 1). \tag{2.139}$$

Note that an extra term has to be added to equation (2.120) if $\gamma \neq 5/3$. In order to write equation (2.139), we have used the transformation from Eulerian potential Φ to peculiar potential ϕ

$$\Phi = \frac{2\pi G \bar{\rho} r^2}{3} + \frac{2\pi G \bar{\rho}_\Lambda r^2}{3} + \phi. \tag{2.140}$$

The peculiar potential is equal to the Eulerian potential in non-cosmological cases.

Initial conditions

In order to provide an *ab initio* scenario of the evolution of the Universe, cosmological simulations are usually started at high redshifts. The initial conditions can be computed using a random Gaussian random field and DM particles are then set on a regular grid and moved according to the Zel'dovich approximation (see section 2.1.2.3), as illustrated on figure 2.2.8. The approximation is used to fast-forward time to $z \sim 100$. At this time, the Zel'dovich approximation still provides very accurate results, yet the density contrast is high enough for the approximate Poisson solver to be able to solve the Poisson equation.

State-of-the-art cosmological simulations

Let us briefly introduce the different cosmological simulations. For the sake of comparison only simulations with sizes comparable to $\sim 100\,\mathrm{Mpc}/h$ are mentioned here. They are presented in Table 2.4; they usually have mass resolutions of the order of $1 \times 10^6\,\mathrm{M}_\odot$ for SPH and moving mesh simulations and spatial resolutions of $1\,\mathrm{kpc}$ for AMR simulations. In addition to different

Table 2.4: Summary of recent state-of-the art cosmological simulations. Only simulations including gas and with box sizes of the order of $100\,\mathrm{Mpc}$ have been included. The resolution is mentioned in mass for SPH simulations and in spatial resolution for AMR simulations. [1]: This simulation is a re-zoom simulation of the Horizon-AGN simulation.

Simulation name	Box size	Resolution	Method	Reference
FIRE m13	$60\,\mathrm{Mpc}$	$3 \times 10^5\,\mathrm{M_\odot}$	SPH	Hopkins et al., 2014
MassiveBlack II	$100\,\mathrm{Mpc}/h$	$2 \times 10^6\,\mathrm{M_\odot}$	SPH	Khandai et al., 2015
Illustris-1	$100\,\mathrm{Mpc}$	$1 \times 10^6\,\mathrm{M_\odot}$	Moving mesh	Vogelsberger et al., 2014
Illustris-TNG100	$100\,\mathrm{Mpc}$	$1 \times 10^6\,\mathrm{M_\odot}$	Moving mesh	Springel et al., 2018
Horizon-AGN	$100\,\mathrm{Mpc}/h$	$1\,\mathrm{kpc}$	AMR	Dubois et al., 2016
New Horizon[1]	$100\,\mathrm{Mpc}/h$	$30\,\mathrm{pc}$	AMR	Park et al., 2019

resolutions and hydrodynamical solvers, they have also very different feedback recipes. Comprehensive comparisons of the different simulation techniques and physical recipes is the topic of ongoing research (see e.g. the Aquila comparison project Scannapieco et al., 2012, the AGORA comparison project Kim et al., 2013; Kim et al., 2016).

In section 2.1, I have presented the cosmological context in which galaxies form. In section 2.1.1, I have presented the current evidences that the Universe is organised at large-scales in the so-called "cosmic-web". In section 2.1.2, I have presented the models that describe the formation of the large-scale structures of the Universe, and in particular the cosmic web. The remaining of the dissertation will in particular focus on the link between the cosmic web and the formation of dark halos and their galaxies. In order to do so, I have presented the different tools I have used in my work. From a theoretical perspective, one can predict properties of dark matter halos from first principle using the excursion set theory, as presented in sections 2.1.3 and 2.1.4 and the peak-patch theory, as presented section 2.1.5. These tools enable us to compute the properties of dark matter halos, yet they fail at predicting the fate of baryons in galaxies, which is usually understood in the classical model of galaxy formation presented in section 2.2.1. One way to study the evolution of baryons is to rely on hydrodynamical numerical simulations, which I presented in section 2.2. The different models used in the cosmological simulations I have used in my dissertation are detailed in section 2.2.2, while a more technical description focused on the numerical methods is provided in section 2.2.3.

2.3 State-of-the-art and synopsis
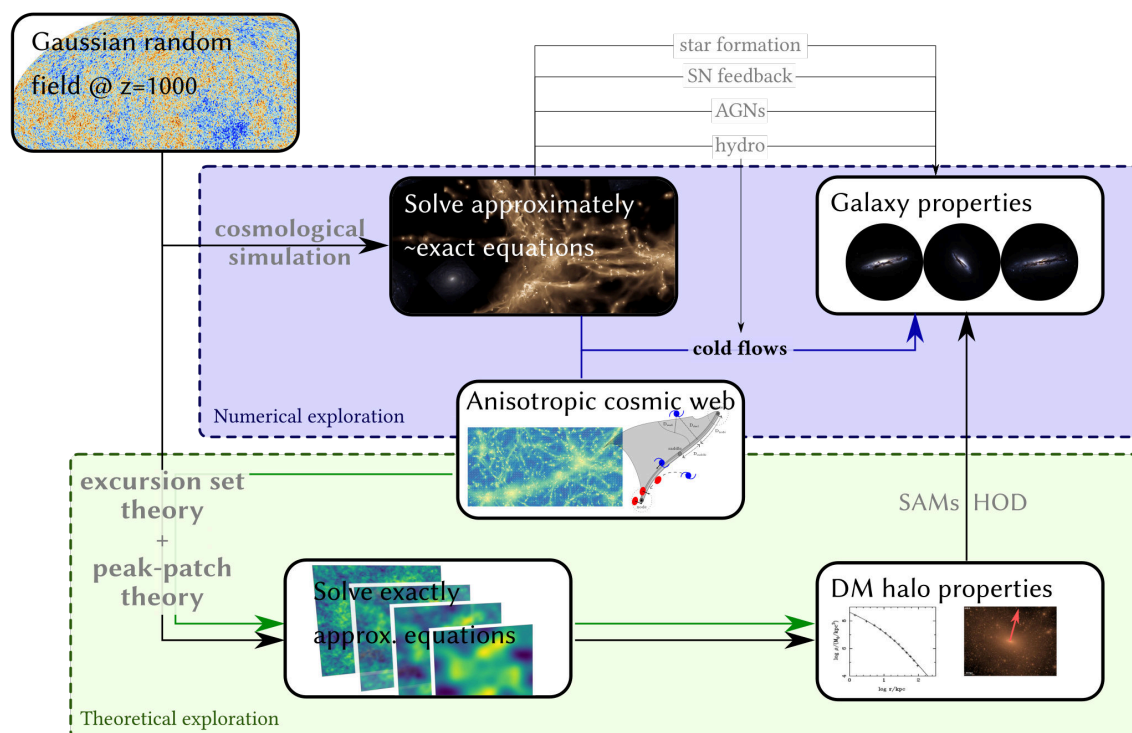
Figure 2.3.1: Sketch of the approaches used in this thesis to study the formation of dark matter halos and galaxies. The different processes relevant to the formation of dark matter halos and their galaxies are shown in gray. The two axes developed in the thesis are shown in the green area (theoretical exploration, chapters 3 and 4) and blue area (numerical exploration, chapter 6). Along the former, I developed extensions to the excursion set and peak theory that explicitly include the anisotropic effect of the cosmic web. This was used to predict environmental-driven effects on the properties of DM halos. Along the latter, I developed numerical methods to accurately follow the cosmic accretion of gas in simulations and applied it to understand how angular momentum is transported by cold flows on high-redshift galaxies.

One of the success of the ΛCDM model is its ability to predict a significant number of properties of DM halos and their galaxies. In the classical model of galaxy formation, galaxy form out of the condensation of the gas in the potential well of their host halo. As such, galaxy properties are usually understood as a result of the halo mass — which sets the amount of gas available and the internal kinematics — and the local density — which regulates gas accretion and pair interactions. The classical analytical and semi-analytical models intrinsically suppose that halo properties, and as a consequence, galaxy properties are only influenced by their local environment *via* the local density, with some extensions probing also the local tidal environment. These models have proven successful at predicting many galactic properties, such as their spatial clustering or their mass function.

In the context of assembly bias, many extensions of the halo model have been suggested to understand the modulation effects of the cosmic web in terms of local properties. In particular, it has been suggested that the local tidal field may explain part of the assembly bias signal (e.g. Hahn et al., 2009; Ludlow et al., 2014) when formulated in terms of the formation time. Tidal forces induce a shear flow in the vicinity of small halos that flow along filaments of the cosmic

web. One of the outcomes is that the accretion rate of small halos is decreased by neighbouring structures, so that small halos growing in dense environments are not able to accrete mass. As a consequence, these halos appear older resulting in a differential biasing as a function of formation time. Similarly, Paranjape et al., 2018 suggested that the effect of halo concentration on the bias is well explained by a local quantification of the local tidal anisotropy. All these models are typically extensions of the halo model with new halo-centred probes of the larger-scale environment.

Another possible approach, which is the one followed in this dissertation, is to relax the halo-centric assumption and work in the frame that sets the large scale environment: the cosmic web. Indeed, due to the statistical properties of the initial conditions of the Universe, the different scales involved in galaxy formation and the formation of the cosmic web are coupled statistically. In particular, large-scale structures such as large filaments have an impact on the statistical properties of the field out of which halos grow, which has the effect of biasing halo assembly. One can argue that the assembly signal can be explained simply *via* this biasing effect of the cosmic web: the cosmic web is responsible for driving the typical assembly history at fixed halo mass and local density.

This approach has already proven successful at providing a theoretical explanation to the spin-alignment problem (Codis et al., 2015). This framework has since been used to show that, in hydrodynamical simulations, the cosmic web has also an effect on the assembly of galaxies. In Kraljic et al., 2018; Kraljic et al., 2019 (appendices B.1 and B.2), we reported that the specific star formation rate and the velocity-to-velocity-dispersion ratio both present significant modulations along the filaments, highlighting that, indeed, filamentary structures can be used as a metric to parametrize the assembly of dark matter halos and galaxies therein. Using an extension of the excursion set theory (Bond et al., 1991; Lacey and Cole, 1993; Mo and S. D. M. White, 1996), I show in chapter 3 (Musso, Cadiou et al., 2018) that the cosmic web, and in particular large scale filaments, biases the formation of dark matter halos. In this dissertation, I also argue that the assembly bias problem stated in these terms can also provide a valuable understanding of how halos grow, but also how their galaxy forms.

Although a number of evidences are pointing towards an effect of the cosmic web on galaxy formation, the detailed physics that couples them is still poorly understood. One of the issues lays in the description of the cosmic web itself, so that different methods may lead to different effects on galaxy formation. One key parameter to further study the effect of the cosmic web is then the question of its description, the challenge residing in its continuous and multi-scale nature. Many methods have been developed to tackle this issue and provide a local frame in which galaxy properties can be studied (Bond et al., 1996; Sousbie et al., 2008 and Libeskind et al., 2018 for a review). In chapter 4, I highlight a process entering galaxy and dark halo formation, namely the coalescence of critical points of the cosmic web as a function of cosmic time. I present theoretical predictions that account for it in a compact way and provide theoretical predictions of the evolution of the cosmic web in the Lagrangian space of the initial conditions (based on the idea of Hanami, 2001) and link them to the connectivity of the cosmic web (Codis et al., 2018).

The complex coupling of the different processes involved in galaxy formation (star formation, feedback, gas cooling, hydrodynamics) render theoretical predictions particularly complex. One way around is to rely on numerical simulations that model these processes experimentally to study galaxy formation. This has been shown to reproduce well the spatial clustering of galaxies (Springel et al., 2006), but also their properties such as morphology, colour or sizes (e.g. Vogelsberger et al., 2014; Dubois et al., 2016; Schaye et al., 2015 and Scannapieco et al., 2012; Kim et al., 2016 for a comparison of the predictions). One of the strong predictions of the numerical codes and their physical models is the presence of cold gas, that flows along dark matter filaments that form the cosmic web (Stewart et al., 2017), so that one can now rely on hydrodynamical numerical simulations to simulate galaxy formation in their cosmological context. These flows

have been shown to be a main channel driving angular-momentum rich material down to the inner regions of the galaxies (Kimm et al., 2011; Stewart et al., 2013; Danovich et al., 2015; Tillson et al., 2015; Bullock et al., 2016), which are key to understand the emergence of the disk structure of galaxies. One of the key to understand the dynamics of cold flows is to understand how their history differs from that of the hot-accreted gas, and how it impacts the angular momentum acquisition of the central galaxy, by studying the Lagrangian history of the gas. Studies based on Godunov solvers, albeit very accurate at capturing hydrodynamical shocks, only provide the Eulerian history of the gas. In order to get the Lagrangian history of the gas, codes have been equipped with tracer particles. In chapter 5, I present a new tracer particle implementation for the code RAMSES based on a Monte Carlo approach. This implementation significantly improves over previous implementations and enables us to study accurately the Lagrangian history of the baryons through their hydrodynamical evolution and their recycling in stars and AGNs. Using the new tracer particle, I present in chapter 6 an analysis of the Lagrangian evolution of the angular momentum of the gas as it flows into galaxies at high redshift, so as to better understand how galaxies get their spin. The evolution of the magnitude and orientation of the angular momentum is computed for the cold- and the hot-accreted gas. I decompose the forces between stellar gravitational forces, dark matter gravitational forces and pressure forces to assess which component dominates where at different locations.

3. The impact of the large scale cosmic web on assembly bias

Outline

3.1 Introduction

Galaxies form and evolve within a complex network, the so-called cosmic web (Bond et al., 1996, see section 2.1.2), made of filaments embedded in sheet-like walls, surrounded by large voids and intersecting at clusters of galaxies (Jõeveer et al., 1978). Halo masses are highly dependent on their large-scale surrounding, as elegantly explained by the theory of biased clustering (Kaiser, 1984a; Efstathiou et al., 1988), such that high mass objects are preferentially found in over-dense regions near nodes (Bond and Myers, 1996; Pogosyan et al., 1996). The importance of interactions with the larger scale environment in driving their evolution has indeed recently emerged as a central tenet of halo formation theory.

It has been established that the clustering of dark matter halos, as measured by halo bias, not only depends on halo mass but also on other halo properties such as formation time, concentration, spin and ellipticity (Gao et al., 2005; Wechsler et al., 2006; Gao and S. D. M. White, 2007; Hahn et al., 2007). This effect, commonly referred to as "assembly bias" can be rephrased as follows: the clustering of dark matter halos and their properties are correlated, beyond a mere mass and density relation. Using a different approach, a growing number of evidence from simulations (Welker et al., 2014; Kraljic et al., 2018; Kraljic et al., 2019; Martizzi et al., 2019) and observations (Porter et al., 2008; Kleiner et al., 2017; Malavasi et al., 2017, e.g.) have since showed that some

halo and galaxy properties present distinct features at different locations in the cosmic web. One striking example is spin-alignments which have been measured for DM halos (e.g. Codis et al., 2012; Dubois et al., 2014) and galaxies (e.g. Tempel et al., 2013; Welker et al., 2014; Chisari et al., 2017), but also the colour segregation of galaxies (Rojas et al., 2004; Martínez et al., 2016; Beygu et al., 2016; Laigle et al., 2018; Kraljic et al., 2018; Kraljic et al., 2019, e.g.).

As a filament is formally the field line that joins two maxima of the density field through a filament-type saddle point (Pogosyan et al., 2009a), studying the expected properties of galaxies and halos in the vicinity of filament-type saddle points is a sensible choice. Indeed, TTT (Peebles, 1969; Schäfer, 2009) was recently revisited (Codis et al., 2015, see section 2.1.6) in the context of such anisotropic environments, biased by the presence of a filament within a wall, which is most efficiently represented by this point process of filament-type saddles. It predicts the alignment of the angular momentum distribution of the forming galaxies with the filament's direction, and perpendicular orientation for massive population. Since spin plays an important role in the physical and morphological properties of galaxies, a signature is also expected in the properties of galaxies as a function of the longitudinal and transverse distance to this saddle.

Most of the previous theoretical work on the impact of the anisotropy of the environment on galactic assembly history focused on dark matter halos. At a given mass, halos that are sufficiently far away from the potential wells grow by accreting their surrounding matter, leading to a correlation between the instantaneous accretion rate and the density of their environment (e.g. Zentner, 2007). On the other hand, halos close to the potential wells are expected to stall and stop to grow earlier, as their mass inflow is dynamically quenched by anisotropic tides generated in their vicinity (e.g. Dalal et al., 2008; Hahn et al., 2009; Ludlow et al., 2014; Borzyszkowski et al., 2017). Individual properties of dark matter halos, such as their mass, formation time or accretion, are thus expected to be affected by the exact position of halos within the large-scale anisotropic cosmic web.

These works underlined the role of the shear strength (a scalar quantity constructed out of the traceless shear tensor which does not correlate with the local density), measured on the same scale as the halo (Castorina et al., 2016; Paranjape et al., 2018). As tidal forces act against gravitational collapse, the shear strength encodes the delay induced by the dynamical quenching due to the environment. This has been justified as a phenomenological explanation of the scale-dependent scatter in the initial overdensity of proto-halos measured in simulations (Ludlow et al., 2014; R. K. Sheth et al., 2013) or as a theoretical consequence of the coupling between the shear and the inertia tensor which tends to slow down collapse (Bond and Myers, 1996; R. K. Sheth et al., 2001; Del Popolo et al., 2001).

The purpose of this paper is to address the question of the environmental quenching of halos. In particular, is the cosmic web responsible for the environmental quenching of halos? What effect does it have on different variables entering the assembly of dark matter halos? In collaboration with M. Musso, we extended the excursion set theory to account for the large-scale modulations induced by a filament-type saddle point. From this, we computed the mass function and the accretion rate and formation times at fixed final mass.

The results presented here were published in Musso, Cadiou et al., 2018.

3.2 "How does the cosmic web impact assembly bias?"

How does the cosmic web impact assembly bias?

M. Musso,[1,2,3]★ C. Cadiou,[1]★ C. Pichon,[1,4] S. Codis,[5] K. Kraljic[6] and Y. Dubois[1]

[1]*Institut d'Astrophysique de Paris, CNRS and UPMC, UMR 7095, 98 bis Boulevard Arago, F-75014 Paris, France*
[2]*Institut de Physique Théorique, Université Paris Saclay and CEA, CNRS, F-91191 Gif-sur-Yvette, France*
[3]*East African Institute for Fundamental Research (ICTP-EAIFR), KIST2 Building, Nyarugenge Campus, University of Rwanda, Kigali, Rwanda*
[4]*Korea Institute of Advanced Studies (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul, 02455, Republic of Korea*
[5]*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada*
[6]*Aix Marseille Université, CNRS, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France*

ABSTRACT

The mass, accretion rate, and formation time of dark matter haloes near protofilaments (identified as saddle points of the potential) are analytically predicted using a conditional version of the excursion set approach in its so-called upcrossing approximation. The model predicts that at fixed mass, mass accretion rate and formation time vary with orientation and distance from the saddle, demonstrating that assembly bias is indeed influenced by the tides imposed by the cosmic web. Starved, early-forming haloes of smaller mass lie preferentially along the main axis of filaments, while more massive and younger haloes are found closer to the nodes. Distinct gradients for distinct tracers such as typical mass and accretion rate occur because the saddle condition is anisotropic, and because the statistics of these observables depend on both the conditional means and their covariances. The theory is extended to other critical points of the potential field. The response of the mass function to variations of the matter density field (the so-called large-scale bias) is computed, and its trend with accretion rate is shown to invert along the filament. The signature of this model should correspond at low redshift to an excess of reddened galactic hosts at fixed mass along preferred directions, as recently reported in spectroscopic and photometric surveys and in hydrodynamical simulations. The anisotropy of the cosmic web emerges therefore as a significant ingredient to describe jointly the dynamics and physics of galaxies, e.g. in the context of intrinsic alignments or morphological diversity.

Key words: galaxies: evolution – galaxies: formation – galaxies: kinematics and dynamics – large-scale structure of Universe – cosmology: theory.

1 INTRODUCTION

The standard paradigm of galaxy formation primarily assigns galactic properties to their host halo mass. While this assumption has proven to be very successful, more precise theoretical and observational considerations suggest other hidden variables must be taken into account.

The mass–density relation (Oemler 1974), established observationally 40 yr ago, was explained (Kaiser 1984; Efstathiou et al. 1988) via the impact of the long-wavelength density modes of the dark matter (DM) field, allowing the proto-halo to pass earlier the critical threshold of collapse (Bond et al. 1991). This biases the mass function in the vicinity of the large-scale structure: the abundance of massive haloes is enhanced in overdense regions.

Numerical simulations have shown that denser environments display a population of smaller, older, highly concentrated 'stalled' haloes, which have stopped accreting and whose relationship with the environment is in many ways the opposite of that of large-mass actively accreting haloes that dominate their surroundings. This is the so-called assembly bias (e.g. Sheth & Tormen 2004; Gao, Springel & White 2005; Wechsler et al. 2006; Dalal et al. 2008; Paranjape & Padmanabhan 2017; Lazeyras, Musso & Schmidt 2017). More recently, Alonso, Eardley & Peacock (2015), Tramonte et al. (2017) and von Braun-Bates et al. (2017) have investigated the differential properties of haloes with respect to loci in the cosmic web. As they focused their attention to variations of the mass function, they also found them to vary mostly with the underlying density. Paranjape, Hahn & Sheth (2017) have shown that haloes in nodes and filaments behave as two distinct populations when a suitable variable based on the shear strength on a scale of the order of the halo's turnaround radius is considered.

In observations, galactic conformity (Weinmann et al. 2006) relates quenching of centrals to the quenching of their satellite galaxies. It has been detected for low- and high-mass satellite galaxies up to high redshift ($z \sim 2.5$, Kawinwanichakij et al. 2016) and

fairly large separation (4 Mpc, Kauffmann et al. 2013). Recently, colour and type gradients driven specifically by the anisotropic geometry of the filamentary network have also been found in simulations (Laigle et al. 2017; Kraljic et al. 2018) using the Horizon-AGN simulation (Dubois et al. 2014), and observations using SDSS (Yan, Fan & White 2013; Martínez, Muriel & Coenda 2016; Poudel et al. 2017; Chen et al. 2017), GAMA (Alpaslan et al. 2016; Kraljic et al. 2018) and, at higher redshift, VIPERS (Malavasi et al. 2017) and COSMOS (Laigle et al. 2017). This suggests that some galactic properties do not only depend on halo mass and density alone: the co-evolution of conformal galaxies is likely to be connected to their evolution within the same large-scale anisotropic tidal field.

An improved model for galaxy evolution should explicitly integrate the diversity of the geometry of the environment on multiple scales and the position of galaxies within this landscape to quantify the impact of its anisotropy on galactic mass assembly history. From a theoretical perspective, at a given mass, if the halo is sufficiently far from competing potential wells, it can grow by accretion from its neighbourhood. It is therefore natural to expect, at fixed mass, a strong correlation between the accretion rate of haloes and the density of their environment (Zentner 2007; Musso & Sheth 2014b). Conversely, if this halo lies in the vicinity of a more massive structure, it may stop growing earlier and stall because its expected feeding will in fact recede towards the source of anisotropic tide (e.g. Dalal et al. 2008; Hahn et al. 2009; Ludlow, Borzyszkowski & Porciani 2014; Wang et al. 2011).

Most of the work carried out so far has focused on the role of the shear strength (a scalar quantity constructed out of the traceless shear tensor which does not correlate with the local density) measured on the same scale of the halo: as tidal forces act against collapse, the strength of the tide will modify the relationship of the halo with its large-scale density environments, and induce distinct mass assembly histories by dynamically quenching mass inflow (Hahn et al. 2009; Castorina et al. 2016; Borzyszkowski et al. 2016). Such local shear strength should be added, possibly in the form of a modified collapse model that accounts for tidal deformations, so as to capture e.g. the effect of a central on its satellites' accretion rate. This modified collapse model has been motivated in the literature on various grounds, e.g. as a phenomenological explanation of the scale-dependent scatter in the initial overdensity of proto-haloes measured in simulations (Ludlow et al. 2014; Sheth, Chan & Scoccimarro 2013) or as a theoretical consequence of the coupling between the shear and the inertia tensor which tends to slow down collapse (Bond & Myers 1996; Sheth, Mo & Tormen 2001; Del Popolo, Ercan & Gambera 2001). Notwithstanding, the position within the large-scale anisotropic cosmic web also directly conditions the local statistics, even without a modification of the collapse model, and affects different observables (mass, accretion rate, etc.) differently.

The purpose of this paper is to provide a mathematical understanding of how assembly bias is indeed partially *driven* by the anisotropy of large-scale tides imprinted in the so-called cosmic web. To do so, the formalism of excursion sets will be adapted to study the formation of structures in the vicinity of saddle points as a proxy for filaments of the cosmic web. Specifically, various tracers of galactic assembly will be computed conditional to the presence of such anisotropic large-scale structure. This will allow us to understand why haloes of a given mass and local density stall near saddles or nodes, an effect which is not captured by the density–mass relation, as it is driven solely from the traceless part of the tide tensor. This should have a clear signature in terms of the

distinctions between contours of constant typical halo mass versus those of constant accretion rate, which may in turn explain the distinct mass and colour gradients recently detected in the above-mentioned surveys.

The structure of this paper is the following. Section 2 presents a motivation for extended excursion set theory as a mean to compute tracers of assembly bias. Section 3 presents the unconstrained expectations for the mass accretion rate and half-mass. Section 4 investigates the same statistics subject to a saddle point of the potential and computes the induced map of shifted mass, accretion rate, and half-mass time. It relies on the strong symmetry between the unconditional and conditional statistics. Section 5 provides a compact alternative to the previous two sections for the less theoretically inclined reader and presents directly the joint conditional and marginal probabilities of upcrossings explicitly as a function of mass and accretion rate. Section 6 reframes our results in the context of the theory of bias as the response of the mass function to variations of the matter density field. Section 7 wraps up and discusses perspectives. Appendix A sums up the definitions and conventions used in the text. Appendix B tests these predictions on realizations of Gaussian random fields (GRFs). Appendix C investigates the conditional statistics subject to the other critical points of the field. Appendix D presents the probability distribution function (PDF) of the eigenvalues at the saddle. Appendix E presents the covariance matrix of the relevant variables to the PDFs. Appendix F presents the relevant joint statistics of the field and its derivatives (spatial and with respect to filtering) and the corresponding conditional statistics of interest. Appendix G presents the generalization of the results for a generic barrier. Appendix H speculates about galactic colours.

2 BASICS OF THE EXCURSION SET APPROACH

The excursion set approach, originally formulated by Press & Schechter (1974), assumes that virialized haloes form from spherical regions whose initial mean density equals some critical value. The distribution of late-time haloes can thus be inferred from the simpler Gaussian statistics of their Lagrangian progenitors. The approach implicitly assumes approximate spherical symmetry (but not homogeneity), and uses spherical collapse to establish a mapping between the initial mean density of a patch and the time at which it recollapses under its own gravity.

According to this model, a sphere of initial radius R shrinks to zero volume at redshift z if its initial mean overdensity δ equals $\delta_c D(z_{in})/D(z)$, where $D(z)$ is the growth rate of linear matter perturbations, z_{in} the initial redshift, and $\delta_c = 1.686$ for an Einstein–de Sitter universe, or equivalently, if its mean overdensity linearly evolved to $z = 0$ equals $\delta_c/D(z)$, regardless of the initial size. If so, thanks to mass conservation, this spherical patch will form a halo of mass $M = (4\pi/3)R^3\bar{\rho}$ (where $\bar{\rho}$ is the comoving background density) . The redshift z is assumed to be a proxy for its virialization time.

Bond et al. (1991) added to this framework the requirement that the mean overdensity in all larger spheres must be lower than δ_c, for outer shells to collapse at a later time. This condition ensures that the infall of shells is hierarchical, and the selected patch is not crushed in a bigger volume that collapses faster (the so-called *cloud-in-cloud* problem). The number density of haloes of a given mass at a given redshift is thus related to the volume contained in the largest spheres whose mean overdensity $\delta \equiv \delta(R)$ crosses δ_c. The dependence of the critical value δ_c on departures from spherical collapse induced by initial tides was studied by Bond & Myers (1996), and later

by Sheth et al. (2001), who approximated it as a scale-dependent barrier. This will be further discussed in Section 7.2.

As the variation of $\delta(R)$ with scale resembles random diffusion, it is convenient to parametrize it with the variance

$$\sigma^2(R) \equiv \text{Var}(\delta(R)) = \int \mathrm{d}k \frac{k^2 P(k)}{2\pi^2} W^2(kR) \tag{1}$$

of the stochastic process, smoothed with a real-space Top-Hat filter W,[1] rather than with R or M. In equation (1), $P(k)$ is the underlying power spectrum. The three quantities σ, R, and M are in practice interchangeable. The mass fraction in haloes of mass M at z is

$$\frac{M}{\bar{\rho}} \frac{\mathrm{d}n}{\mathrm{d}M} = \left| \frac{\mathrm{d}\sigma}{\mathrm{d}M} \right| f(\sigma), \tag{2}$$

where $\mathrm{d}n/\mathrm{d}M$ is the number density of haloes per unit mass (i.e. the mass function) and $f(\sigma)$ – often called the halo multiplicity – is the probability distribution of the first-crossing scale of the random walks, that is of the smallest σ (largest R) for which

$$\delta(R, \mathbf{r}) \equiv \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \delta_\mathrm{m}(\mathbf{k}) W(kR) e^{i\mathbf{k} \cdot \mathbf{r}} = \frac{\delta_\mathrm{c}}{D(z)}, \tag{3}$$

where δ_m is the (unsmoothed) matter density. The first-crossing requirement avoids double counting and guarantees that $f(\sigma)$ is a well-behaved probability distribution, and the resulting mass fraction is correctly normalized. In equation (3), the linear growth factor, $D(z)$, is defined as a function of redshift via

$$D(z) = \frac{H(a)}{H_0} \int_0^a \frac{\mathrm{d}a}{\sqrt{\Omega_\mathrm{m}/a + \Omega_\Lambda a^2}}, \text{ with } a = \frac{1}{1+z}. \tag{4}$$

At early time, $D(z)$ scales like $1/(1 + z)$. Here, $H(a) = H_0\sqrt{\Omega_\mathrm{m}/a + \Omega_\Lambda a^2}$ is the Hubble constant.

The first-crossing probability, $f(\sigma)\Delta\sigma$, is the fraction of walks that cross the threshold between $\sigma - \Delta\sigma$ and σ for the first time. Considering discretized trajectories with a large number of steps $\sigma_1, \ldots, \sigma_N$ of width $\Delta\sigma \equiv \sigma_i - \sigma_{i-1}$ (corresponding to concentric spheres of radii $R_1, \ldots, R_N$), the first-crossing probability is the joint probability that $\delta_N > \delta_\mathrm{c}$ and $\delta_i < \delta_\mathrm{c}$ for $i < N$, with $\delta_i \equiv \delta(\sigma_i)$ and $\sigma_N = \sigma = N\Delta\sigma$. Hence, the distribution $f(\sigma)$ is formally defined as the limit

$$f(\sigma) \equiv \lim_{\Delta\sigma \to 0} \frac{1}{\Delta\sigma} \langle \vartheta(\delta_N - \delta_\mathrm{c}) \prod_i^{N-1} \vartheta(\delta_\mathrm{c} - \delta_i) \rangle, \tag{5}$$

where $\vartheta(x)$ is Heaviside's step function, and the expectation value is evaluated with the multivariate distribution $p(\delta_1, \ldots, \delta_N)$. This definition discards crossings for which $\delta_i > \delta_\mathrm{c}$ for any $i < N$, since $\vartheta(\delta_\mathrm{c} - \delta_i) = 0$, assigning at most one crossing (the first) to each trajectory. For instance, in Fig. 1, trajectory B would not be assigned the crossing marked with (3), since the trajectory lies above threshold between (1) and (2). Since taking the mean implies integrating over all trajectories weighed by their probability, $f(\sigma)$ can be interpreted as a path integral over all allowed trajectories with fixed boundary conditions $\delta(0) = 0$ and $\delta(\sigma) = \delta_\mathrm{c}$ (Maggiore & Riotto 2010).

In practice, computing $f(\sigma)$ becomes difficult if the steps of the random walks are correlated, as is the case for real-space Top-Hat filtering with a Λ cold dark matter (ΛCDM) power spectrum, and for most realistic filters and cosmologies. For this reason, more easily tractable but less physically motivated sharp cut-offs in Fourier

---
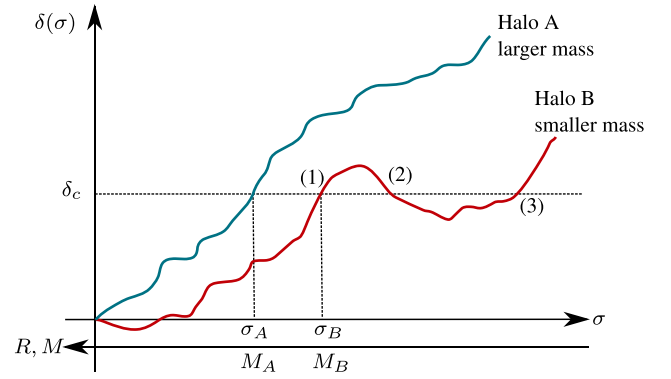
Figure 1. Pictorial description of the first-crossing and upcrossing conditions to infer the halo mass from the excursion set trajectory. The first-crossing condition on σ assigns at most one halo to each trajectory, with mass $M(\sigma)$. Upcrossing may instead assign several masses to the same trajectory (that is, to the same spatial location), thus overcounting haloes. Trajectory B in the figure has a first crossing (upwards) at scale σ_B (1), a downcrossing (2), and second upcrossing (3), but the correct mass is only given by σ_B. However, the correlation of each step with the previous ones makes turns in small intervals of σ exponentially unlikely: at small σ most trajectories will thus look like trajectory A. Thanks to the correlation between steps at different scales, for small σ (large M) simply discarding downcrossings is a very good approximation.

space have been often preferred, for which the correlation matrix of the steps becomes diagonal, treating the correlations as perturbations (Maggiore & Riotto 2010; Corasaniti & Achitouv 2011). The upcrossing approximation described below can instead be considered as the opposite limit, in which the steps are assumed to be strongly correlated (as is the case for a realistic power spectrum and filter). This approximation is equivalent to constraining only the last two steps of equation (5), marginalizing over the first $N - 2$.

2.1 The upcrossing approximation to $f(\sigma)$.

Indeed, Musso & Sheth (2012) noticed that for small enough σ (i.e. for large enough masses), the first-crossing constraint may be relaxed into the milder condition

$$\delta' \equiv \frac{\mathrm{d}\delta}{\mathrm{d}\sigma} > 0; \tag{6}$$

that is, trajectories simply need to reach the threshold with positive slope (or with slope larger than the threshold's if δ_c depends on scale). This upcrossing condition may assign several haloes of different masses to the same spatial location. For this reason, while first crossing provides a well-defined probability distribution for σ (e.g. with unit normalization), upcrossing does not. However, since the first crossing is necessarily upwards, and downcrossings are discarded, the error introduced in $f(\sigma)$ by this approximation comes from trajectories with two or more turns. Musso & Sheth (2012) showed that these trajectories are exponentially unlikely if σ is small enough when the steps are correlated. The first-crossing and upcrossing conditions to infer the halo mass from excursion sets are sketched in Fig. 1: while the trajectory A would be (correctly) assigned to a single halo, the second upcrossing of trajectory B in the figure would be counted as a valid event by the approximation, and the trajectory would (wrongly) be assigned to two haloes. The probability of this event is non-negligible only if σ is large.

Returning to equation (5), expanding δ_{N-1} around δ_N gives

$$\vartheta(\delta_\mathrm{c} - \delta_{N-1}) \simeq \vartheta(\delta_\mathrm{c} - \delta_N) + \delta_\mathrm{D}(\delta_\mathrm{c} - \delta)\delta'\Delta\sigma, \tag{7}$$

where the crossing scale σ, giving the halo's final mass M, is defined implicitly in equation (3), as the solution of the equation $\delta(\sigma) = \delta_c/D$.[2] The assumption that this upcrossing is first crossing allows us to marginalize over the first $N-2$ variables in equation (5) without restrictions. The first term has no common integration support with $\vartheta(\delta_N - \delta_c)$, and only the second one – containing the Jacobian $(\delta' - \delta'_c)$ – contributes to the expectation value (throughout the text, a prime will denote the derivative $d/d\sigma$). Adopting for convenience the normalized walk height $\nu \equiv \delta/\sigma$, for which $\langle \nu^2 \rangle = 1$, the corresponding density of solutions in σ-space obeys

$$\left| \nu' - \nu'_c \right| \delta_D(\nu - \nu_c) = (|\delta'|/\sigma)\, \delta_D(\nu - \nu_c)\,, \tag{8}$$

where $\nu_c \equiv \delta_c/(\sigma D)$ is the rescaled threshold. The probability of upcrossing at σ in equation (5) is therefore simply the expectation value of this expression,

$$f_{up}(\sigma) \equiv p_G(\nu = \nu_c) \int_0^\infty d\delta'\, \delta'\, p_G(\delta'|\nu_c)\,, \tag{9}$$

where the integral runs over $\delta' > 0$ because of the upcrossing condition (6). Usually, one sets $D = 1$ at $z = 0$ for simplicity so that $\nu_c = \delta_c/\sigma$. For Gaussian initial conditions,[3] the conditional distribution $p_G(\delta'|\nu_c)$ is a Gaussian with mean ν_c and variance $1/\Gamma^2$, where

$$\Gamma^2 = \frac{1}{\langle \delta'^2 \rangle - 1} = \frac{\gamma^2}{1 - \gamma^2} = \frac{1}{\sigma^2 \langle \nu'^2 \rangle}\,, \tag{10}$$

and $\gamma^2 = \langle \delta'\delta \rangle^2/\langle \delta'^2 \rangle \langle \delta^2 \rangle$ is the cross-correlation coefficient between the density and its slope.[4] Thanks to this factorization, integrating equation (9) over δ' yields the fully analytical expression

$$f_{up}(\sigma) = p_G(\nu_c)\frac{\mu}{\sigma}F(X)\,, \tag{11}$$

where p_G is a Gaussian with mean $\langle \nu \rangle = 0$ and variance $\text{Var}(\nu) = 1$. For a constant barrier (see Appendix G for the generalization to a non-constant case), the parameters μ and X are defined as

$$\mu \equiv \langle \delta'|\nu_c \rangle = \nu_c\,, \quad \text{and} \quad X \equiv \frac{\mu}{\sqrt{\text{Var}(\delta'|\nu_c)}} = \Gamma\nu_c\,, \tag{12}$$

with

$$F(x) \equiv \int_0^\infty dy\, \frac{y}{x}\frac{e^{-(y-x)^2/2}}{\sqrt{2\pi}} = \frac{1 + \text{erf}(x/\sqrt{2})}{2} + \frac{e^{-x^2/2}}{x\sqrt{2\pi}}\,, \tag{13}$$

which is a function that tends to 1 very fast as $x \to \infty$, with correction decaying like $\exp(-x^2/2)/x^3$. It departs from one by ~ 8 per cent for a typical $\Gamma\nu_c \sim 1$. Equation (11) can be written explicitly as

$$f_{up}(\sigma) = \frac{\nu_c e^{-\nu_c^2/2}}{\sigma\sqrt{2\pi}}F(\Gamma\nu_c)\,, \tag{14}$$

where the first factor in the right-hand side (RHS) of equation (14) is the result of Press & Schechter (1974), ignoring the factor of 2, they introduced by hand to fix the normalization. For correlated steps, their non-normalized result reproduces well the large-mass tail of $f(\sigma)$ (which is automatically normalized to unit and requires to correcting factor), but it is too low for intermediate and small masses. The upcrossing probability $f_{up}(\sigma)$ also reduces to this result

in the large-mass limit, when $\Gamma\nu_c \gg 1$ and $F(\Gamma\nu_c) \simeq 1$. However, for correlated steps $f_{up}(\sigma)$ is a very good approximation of $f(\sigma)$ on a larger mass range. For a ΛCDM power spectrum, the agreement is good for halo masses as small as $10^{12}\,M_\odot\,h^{-1}$, well below the peak of the distribution. The deviation from the strongly correlated regime is parametrized by $\Gamma\nu_c$, which involves a combination of mass and correlation strength: the approximation is accurate for large masses (small σ and large ν_c) or strong correlations (large Γ). Although Γ mildly depends on σ, fixing $\Gamma^2 \sim 1/3$ (or $\gamma \sim 1/2$) can be theoretically motivated (Musso & Sheth 2014c) and mimics well its actual value for real-space Top-Hat filtering in ΛCDM on galactic scales. The limit of uncorrelated steps ($\Gamma = 0$), whose exact solution is twice the result of Press & Schechter (1974), is pathological in this framework, with f_{up} becoming infinite. More refined approximation methods can be implemented in order to interpolate smoothly between the two regimes (Musso & Sheth 2014a).

From equation (11), a characteristic mass $M_\star$ can be defined by requesting that the argument of the Gaussian be equal to one, i.e. $\nu_c = 1$ or $\sigma(M_\star) = \delta_c/D$. This defines $M_\star$ implicitly via equation (1) for an arbitrary cosmology. This quantity is particularly useful because $f_{up}(\sigma)$ does not have well-defined moments (in fact, even its integral over σ diverges). This is a common feature of first passage problems (Redner 2001), not a problem of the upcrossing approximation: even when the first-crossing condition can be treated exactly, and $f(\sigma)$ is normalized – it is a distribution function –, its moments still diverge. Therefore, given that the mean $\langle M \rangle$ of the resulting mass distribution cannot be computed, $M_\star$ provides a useful estimate of a characteristic halo mass.

2.2 Joint and conditional upcrossing probability

The purpose of this paper is to recompute excursion set predictions such as equation (11) in the presence of additional conditions imposed on the excursions. Adding conditions (like the presence of a saddle at some finite distance) will have an impact not only on the mass function of DM haloes, but also on other quantities such as their assembly time and accretion rate.

Let us present in full generality how the upcrossing probability is modified by such supplementary conditions. When, besides $\delta(\sigma) = \delta_c$ and the upcrossing condition, a set of N linear[5] functional constraints $\{\mathcal{F}_1[\delta], \ldots, \mathcal{F}_N[\delta]\} = \{v_1, \ldots, v_N\}$ on the density field is enforced, the additional constraints modify the joint distribution of ν and ν'. The conditional upcrossing probability may be obtained by replacing $p(\nu, \nu')$ with $p(\nu, \nu'|\{v\})$ in equation (9). For a Gaussian process, when the functional constraints do not involve δ', this replacement yields after integration over the slope

$$f_{up}(\sigma, \{v\}) = p_G(\nu_c, \{v\})\frac{\mu_v}{\sigma}F(X_v)\,, \tag{15}$$

where $p_G(\nu_c|\{v\})$ is a Gaussian with mean $\langle \nu|\{v\} \rangle$ and variance $\text{Var}(\nu|\{v\})$, while μv and Xv are defined as

$$\mu_v \equiv \langle \delta'|\nu_c, \{v\} \rangle\,, \quad X_v \equiv \frac{\mu_v}{\sqrt{\text{Var}(\delta'|\nu_c, \{v\})}}\,, \tag{16}$$

and $\langle \delta'|\nu_c, \{v\} \rangle$ and $\text{Var}(\delta'|\nu_c, \{v\})$ are the mean and variance of the conditional distribution, $p_G(\delta'|\nu_c, \{v\})$ given by equations (F10) and (F11) and evaluated at $\delta = \delta_c$, while F is given by equation (13). Equation (15) is formally the conditional counterpart to equation

---

[2] A careful calculation shows that the step function should be asymmetric, so that $\vartheta(\delta - \delta_c) = 1$ when $\delta = \delta_c$ instead of 1/2.

[3] No conceptual complications arise in dealing with a non-Gaussian distribution, which is none the less beyond the scope of this paper.

[4] Recalling that $\langle \delta'\delta \rangle = \sigma$ so that $\gamma^2 = 1/\langle \delta'^2 \rangle$.

[5] Indeed the saddle condition below imposes linear constraints on the contrast and the potential, since the saddle's height and curvature are fixed.

Table 1. List of variables for the three different probabilities studied in the text (upcrossing, accretion rate given upcrossing, and formation time given upcrossing), conditioned or not to the presence of the saddle point, split by whether they relate to the height of the excursion set trajectory or its slope. Variables like μ and X always appear as $\mu F(X)$ and describe the mean slope of the upcrossing trajectories given the different conditions (presence of the saddle and/or height ν_f of the trajectory at formation). The unconditional case has $\mu = \nu_c$ and $X = \Gamma \nu_c$. The remaining variables appear as arguments of a Gaussian, and are used to define the typical values $\sigma_\star$, $\alpha_\star$, and $D_\star$ of the excursion set variables σ, α, and D_f. The height-related variables describe the probability of reaching the collapse threshold ν_c (unconditional or given the saddle), or the formation threshold ν_f given ν_c (with or without saddle). The slope-related ones describe the probability of having at upcrossing the slope corresponding to a given accretion rate. See also Table A1.

| | Without saddle | | With saddle | |
	Height	Slope	Height	Slope
Upcrossing (σ)	ν_c	μ, X	$\nu_{c,\mathcal{S}}$	$\mu_\mathcal{S}, X_\mathcal{S}$
Accretion (α)		Y_α		$Y_{\alpha,\mathcal{S}}$
Formation (D_f)	$\nu_{f,c}$	μ_f, X_f	$\nu_{f,c,\mathcal{S}}$	$\mu_{f,\mathcal{S}}, X_{f,\mathcal{S}}$

(11), while incorporating extra constraints corresponding to e.g. the large-scale Fourier modes of the cosmic web.

The brute force calculation of the conditional means and variances entering equation (15) can rapidly become tedious. To speed up the process, and gain further insight, one can write the conditional statistics of δ' in terms of those of δ and their derivatives. This is done explicitly in Appendix F1, which allows us to write explicitly the conditional probability of upcrossing at σ given $\{v\}$, obtained by dividing equation (15) by $p(\{v\})$, as

$$f_{\rm up}(\sigma|\{v\}) = -\nu'_{c,v} \frac{e^{-\nu^2_{c,v}/2}}{\sqrt{2\pi}} F\left(-\frac{\nu'_{c,v}}{\sqrt{{\rm Var}\left(\nu'_v\right)}}\right),\qquad (17)$$

given

$$\nu_{c,v} \equiv \frac{\delta_c - \langle \delta | \{v\} \rangle}{\sqrt{{\rm Var}\left(\delta | \{v\}\right)}}, \quad \text{and} \quad \nu'_{c,v} \equiv \frac{{\rm d}\nu_{c,v}}{{\rm d}\sigma}, \qquad (18)$$

where these conditionals and variances can be expressed explicitly in terms of the constraint via equations (F8)–(F11). Equation (17) is therefore also formally equivalent to equation (14), upon replacing $\nu_c \to \nu_{c,v}$ and $\langle \nu'^2 \rangle \to \langle \nu'^2_v \rangle$ to account for the constraint. Remarkably, the conditional probability $f_{\rm up}(\sigma|\{v\})$ is thus simply expressed as an unconditional upcrossing probability for the effective unit variance process obtained from the conditional density.

The above-sketched formal procedure will be applied to practical constraints in the next section. For convenience and consistency, Table 1 lists all the variables that are introduced in the following sections, for the combinations of the various constraints (on the slope at crossing, on the height of the trajectory at $\sigma(M/2)$, and on the presence of a saddle) that will be imposed.

3 ACCRETION RATE AND FORMATION TIME

Let us first present the tracers of galactic assembly when there is no large-scale saddle. Specifically, this section will consider the DM mass accretion rate and formation redshift. It will compute the joint PDFs, the corresponding marginals, typical scales, and expectations. Its main results are the derivation of the conditional probability of the accretion rate – equation (25) – and formation time – equation (36) – for haloes of a given mass. The emphasis will be on derivation in the language of excursion set. The reader
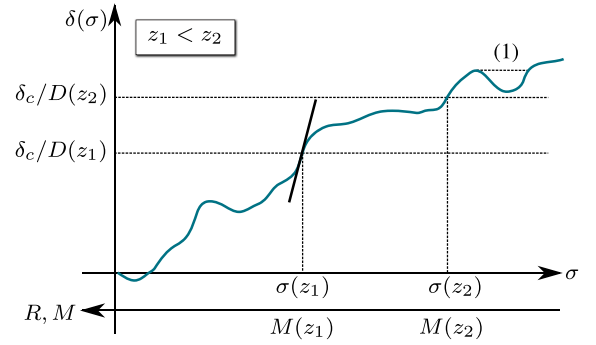
Figure 2. Pictorial representation of the procedure to infer accretion rates from excursion sets. As the redshift z grows, the barrier $\delta_c/D(z)$ becomes higher and the first-crossing scale $\sigma(z)$ moves to the right, towards smaller masses. This procedure reconstructs the entire mass accretion history $M(z)$ from the first-crossing history $\sigma(D)$. As the two redshifts z_1 and z_2 in figure get close to each other, the difference between $\sigma(z_1)$ and $\sigma(z_2)$ is completely fixed by the slope of the trajectory. This deterministic relation connects the excursion set slope to the halo's instantaneous mass accretion rate. Finite jumps of the first-crossing σ after a downturn [where the inverse function $\sigma(\delta)$ becomes multivalued, as in (1)] cannot describe smooth accretion and are traditionally associated with large mergers.

only concerned with statistical predictions in terms of quantities of direct astrophysical interest may skip to Section 5.

Following Lacey & Cole (1993), the entire mass accretion history of the halo is encoded in the portion of the excursion set trajectory after the first crossing: solving the implicit equation (3) at all z enables to reconstruct $M(z)$. As the barrier $\delta_c/D(z)$ decreases with time (since $D(z)$ grows as z decreases), the first-crossing scale moves towards smaller values (larger masses), thereby describing the accretion of mass on to the halo. Clearly, since $\delta(\sigma)$ is not monotonic, $M(z)$ is not a continuous function. Finite jumps of the first-crossing scale, corresponding to portions for which σ is not a global maximum of the interval $[0, \sigma]$, can be interpreted as mergers (see trajectory B in Fig. 1, or the portion marked with (1) in Fig. 2). In the upcrossing approximation, the constraint $\delta'(\sigma) > 0$ discards the downward part of each jump.

3.1 Accretion rate

In the language of excursion sets, finding the mass accretion history is equivalent to reconstructing the function $\sigma(D)$ [where D was defined in equation (4)]: because the barrier grows as D decreases with z, the crossing scale σ moves towards larger values (smaller masses). Differentiating both sides of equation (3) with respect to z gives

$$\alpha \equiv -\frac{D}{\sigma} \frac{{\rm d}\sigma}{{\rm d}D} = \frac{\delta_c}{\sigma \delta'} = \frac{\nu_c}{\sigma(\nu' - \nu'_c)}, \qquad (19)$$

where α measures the fractional change of the first-crossing scale $\sigma(M)$ with $D(z)$, and is related to the instantaneous relative mass accretion rate by

$$\frac{1}{M} \frac{{\rm d}M}{{\rm d}z} \equiv \frac{\dot{M}}{M} = \alpha \frac{{\rm d}\log D}{{\rm d}z}\left(-\frac{{\rm d}\log M}{{\rm d}\log \sigma}\right). \qquad (20)$$

The upcrossing condition implies that $\alpha > 0$: excursion set haloes can only increase their mass, since ${\rm d}\log M/{\rm d}\log \sigma < 0$.

A pictorial representation of this procedure is given in Fig. 2. Equation (19) gives a relation between the accretion rate of the final haloes and the Lagrangian slope of the excursion set trajectories,

which is statistically meaningful in the framework of excursion sets with correlated steps (because the slope then has finite variance). Note that α scales both like the inverse of the slope δ' and the logarithmic rate of change of σ with D. It also essentially scales like the relative accretion rate, $\dot{M}/M$ since in equation (20) $d\log D/dz$ is simply a time-dependent scaling, while on galactic scales, $(n \sim 2)$, $d\log M/d\log \sigma \sim -6$ (see also Section 5 and Appendix E for the generic formula).

Fixing the accretion rate establishes a local bidimensional mapping between $\{\nu, \nu'\}$, or $\{\delta, \delta'\}$, and $\{\sigma, \alpha\}$, defined as the solutions of the bidimensional constraint

$$\mathcal{C} \equiv \{\nu(\sigma) - \nu_c, \nu'(\sigma) - \nu'_c - \nu_c/\sigma\alpha\} = \mathbf{0}. \tag{21}$$

The density of points in the (σ, α) space satisfying the constraint is

$$|\det(\partial\mathcal{C}/\partial\{\sigma, \alpha\})| \, \delta_D^{(2)}(\mathcal{C}). \tag{22}$$

Since $\partial(\nu - \nu_c)/\partial\alpha = 0$, the determinant in equation (22) is simply $|(\nu' - \nu'_c)(\nu_c/\sigma\alpha^2)| = \nu_c^2/\sigma^2\alpha^3$, and is no longer a stochastic variable. Taking the expectation value of equation (22) gives

$$\begin{aligned}
f_{\mathrm{up}}(\sigma, \alpha) &= \frac{\nu_c^2}{\sigma^2\alpha^3} p_G(\nu_c, \nu'_c + \nu_c/\sigma\alpha), \\
&= \frac{\Gamma\nu_c^2}{\sigma\alpha^3} \frac{e^{-\nu_c^2/2}}{\sqrt{2\pi}} \frac{e^{-Y_\alpha^2/2}}{\sqrt{2\pi}},
\end{aligned} \tag{23}$$

with [using the conditional mean $\mu = \nu_c$ from equation (12)]

$$Y_\alpha \equiv \frac{\nu_c/\alpha - \mu}{\sqrt{\mathrm{Var}(\delta'|\nu_c)}} = \Gamma(\sigma\nu'_c + \nu_c/\alpha), \tag{24}$$

which is the joint probability of upcrossing at σ with accretion rate α.[6] This can be formally recovered setting $\langle\delta'|\nu_c, \alpha\rangle = \nu_c/\alpha$ and $\mathrm{Var}(\delta'|\nu_c, \alpha) \to 0$ in equation (16) (because the constraint fixes δ' completely), which gives $F(X_\alpha) = 1$ as needed.

The conditional probability of having accretion rate α given upcrossing at σ can be obtained taking the ratio of equations (23) and (14), which gives

$$f_{\mathrm{up}}(\alpha|\sigma) = \frac{\Gamma\nu_c}{\alpha^3} \frac{e^{-Y_\alpha^2/2}}{\sqrt{2\pi} \, F(\Gamma\nu_c)}, \tag{25}$$

and represents the main result of this subsection. The exact form of $f_{\mathrm{up}}(\alpha|\sigma)$ from equation (25), as σ changes is shown in Fig. 3. This conditional probability has a well-defined mean value, which reads

$$\langle\alpha|\sigma\rangle = \int_0^\infty d\alpha \, \alpha f_{\mathrm{up}}(\alpha|\sigma) = \frac{1 + \mathrm{erf}(\Gamma\nu_c/\sqrt{2})}{2F(\Gamma\nu_c)}; \tag{26}$$

however, the second moment $\langle\alpha^2|\sigma\rangle$ and all higher order statistics are ill defined. The nth moment is in fact proportional to the expectation value of $(1/\delta')^{n-1}$ (over positive slopes and given ν_c), which is divergent. Equation (25) shows that very small values of α (corresponding to very steep slopes) are exponentially unlikely, and very large ones (shallow slopes) are suppressed as a power law. Unlike $f_{\mathrm{up}}(\sigma)$, the conditional distribution $f_{\mathrm{up}}(\alpha|\sigma)$ is a well-defined normalized PDF. However, it is still an approximation to the exact PDF, as it assumes that the distribution of the slopes at first crossing is a (conditional) Gaussian. This assumption is accurate for steep slopes, but overestimates the shallow-slope tail, for which the exact first-crossing condition would impose a boundary condition $p_G(\delta' = 0|\delta_c) = 0$. The higher moments of the exact conditional

---

[6] As expected, marginalizing equation (23) over $\alpha > 0$ gives back equation (11), upon setting $\Gamma\nu_c/\alpha = x$.
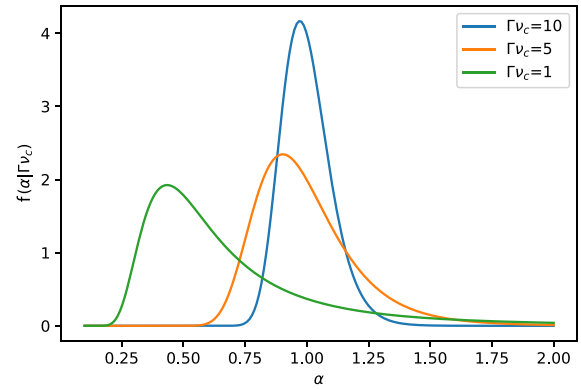
Figure 3. Plot of the conditional PDF $f_{\mathrm{up}}(\alpha|\sigma)$ of the accretion rate for values of σ corresponding to $\Gamma\nu_c = 10$, 5, and 1. As the mass gets smaller, so does $\Gamma\nu_c$ and the conditional PDF moves towards smaller accretion rates α. Therefore, haloes of smaller mass tend to accrete less.

distribution of accretion rates should be convergent. However, even if this was not the case, let us stress that these divergences would not represent a pathology of excursion sets, but are instead a rather common feature of first-passage statistics in a cosmological context.

Regardless of convergence issues, it remains true that the estimate (26) of the mean $\langle\alpha|\sigma\rangle$ gets a significant contribution from the less accurate side of the distribution. One may therefore look for other more informative quantities. In analogy with $M_\star$, defined as the value of M for which $\nu_c = 1$, one can define the characteristic accretion rate $\alpha_\star$ as the value for which Y_α, the argument of the Gaussian in equation (25), equals one

$$\alpha_\star(\sigma) = \frac{\Gamma\nu_c}{1 + \Gamma\nu_c}. \tag{27}$$

For the above-mentioned typical value, it follows that $\alpha_\star(M_\star) = (\sqrt{3} - 1)/2 \approx 1/3$. Another useful quantity is the most likely value of the accretion rate, corresponding to the maximum $\alpha_{\max}$ of $f_{\mathrm{up}}(\alpha|\sigma)$. Requesting the derivative of the PDF to vanish, one gets

$$\alpha_{\max}(\sigma) = \frac{(\Gamma\nu_c)^2}{6} \left[\sqrt{1 + \frac{12}{(\Gamma\nu_c)^2}} - 1\right]. \tag{28}$$

All three quantities $\langle\alpha|\sigma\rangle$, $\alpha_\star$, and $\alpha_{\max}$ tend to 1 in the large-mass limit, and decrease for smaller masses. They thus contain some equivalent information on the position of the bulk of the conditional PDF of α at given mass. Hence, haloes of smaller mass accrete less on average.

3.2 Halo formation time

The formation time is conventionally defined as the redshift z_f at which a halo has assembled half of its mass. It is thus related to the height of the excursion set trajectory at the scale $\sigma_{1/2} \equiv \sigma(M/2)$ corresponding to the radius $R_{1/2} = R/2^{1/3}$. As the barrier $\delta_c/D(z)$ grows with z, and the first-crossing scale moves to the right towards higher values of σ, z_f is the redshift at which $\sigma_{1/2}$ becomes the first-crossing scale for that trajectory, if it exists. That is, neglecting for the time being the presence of finite jumps in the first-crossing scale (interpreted as mergers), one simply needs to solve for z_f the implicit relation $\delta(\sigma_{1/2}) = \delta_c/D(z_f)$, which makes z_f a stochastic variable. As described in Fig. 4, trajectories with the same upcrossing scale σ but different heights at $\sigma_{1/2}$ describe different formation times: a
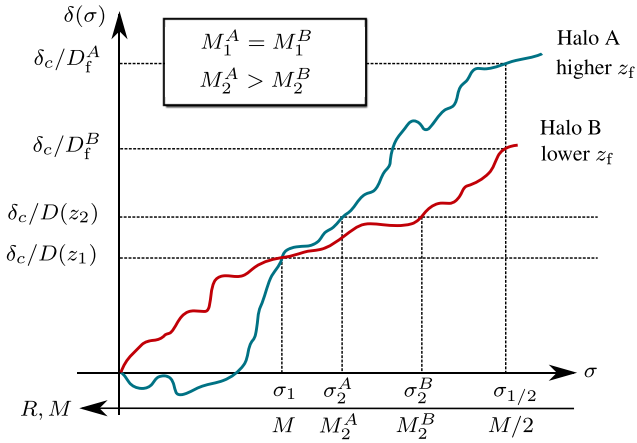
Figure 4. Pictorial representation of the interplay between accretion rate and formation time as inferred from excursion sets. Two haloes A and B upcross the threshold $\delta_c/D(z_1)$ at the same scale σ. At redshift z_1, they have therefore the same mass. Halo A has a steeper slope than halo B, and has thus a lower accretion rate. At a slightly larger redshift z_2, halo A crosses the higher threshold $\delta_c/D(z_2)$ at a lower σ, and its mass is thus larger than halo B's: halo A assembles its mass earlier, consistent with its lower accretion at z_1. At the half-mass scale $\sigma_{1/2} = \sigma(M/2)$, the trajectory of halo A is higher: its threshold δ_c/D_f has a value of D_f lower than halo B's at the same $\sigma_{1/2}$. Halo A has thus assembled half of its mass at a redshift z_f higher than halo B.

higher $\delta_{1/2}$ corresponds to a smaller $D(z_f)$ and thus to a halo with larger z_f, which assembled half of its mass earlier.

In the language of excursion sets, it is convenient to work with $D_f \equiv D(z_f)$ rather than with z_f. In terms of unit variance variables, haloes with formation time D_f correspond to trajectories satisfying

$$\nu_{1/2} \equiv \frac{\delta(\sigma_{1/2})}{\sigma_{1/2}} = \frac{\delta_c}{\sigma_{1/2}D_f} \equiv \nu_f \, , \tag{29}$$

where $\nu_{1/2}$ is the Gaussian variable at $\sigma_{1/2}$ and ν_f is the threshold at D_f. This constraint at $\sigma_{1/2}$ imposes a second condition on the trajectory after $\nu = \nu_c$, which selected the crossing scale σ. One then needs to transform the bidimensional constraint

$$\tilde{\mathcal{C}} \equiv \{\nu - \nu_c, \nu_{1/2} - \nu_f\} = \mathbf{0} \tag{30}$$

on $\{\nu, \nu_{1/2}\}$ into one for $\{\sigma, D_f\}$, which gives

$$\left| \det \left(\partial \tilde{\mathcal{C}} / \partial\{\sigma, D_f\} \right) \right| \delta_D^{(2)}(\tilde{\mathcal{C}}) = \left| \nu' - \nu_c' \right| \frac{\nu_f}{D_f} \delta_D^{(2)}(\tilde{\mathcal{C}}) , \tag{31}$$

thanks to the fact that $\partial(\nu_c - \nu)/\partial D_f = 0$.

The joint probability of upcrossing at σ having formation time D_f, denoted $f_{\mathrm{up}}(\sigma, D_f)$, is defined as the expectation value of equation (31) with the condition $\nu' > \nu_c'$. That is,

$$f_{\mathrm{up}}(\sigma, D_f) \equiv \frac{\nu_f}{D_f} \int_{\nu_c'}^{\infty} \mathrm{d}\nu' (\nu' - \nu_c') \, p_G(\nu_c, \nu', \nu_f) ,$$
$$= \frac{\nu_f}{D_f} p_G(\nu_c, \nu_f) \frac{\mu_f}{\sigma} F(X_f) , \tag{32}$$

where the second equality follows from setting $\{\nu\} = \nu_f$ in the general expression (15), while μ_f and X_f are given by

$$\mu_f(D_f) \equiv \langle \delta' | \nu_c, \nu_f \rangle , \quad X_f(D_f) \equiv \frac{\mu_f(D_f)}{\sqrt{\mathrm{Var}\left(\delta' | \nu_c, \nu_f\right)}} , \tag{33}$$

as specified by equation (16). The conditional mean $\langle \delta' | \nu_c, \nu_f \rangle$ and variance $\mathrm{Var}\left(\delta' | \nu_c, \nu_f\right)$ are computed in equations (F21) and (F22),

which give

$$\mu_f(D_f) = \frac{\omega' \delta_c}{\sigma_{1/2} D_f} + \frac{\sigma - \omega' \omega}{\sigma^2 - \omega^2} \left(\delta_c - \frac{\omega \delta_c}{\sigma_{1/2} D_f} \right) , \tag{34}$$

$$X_f(D_f) = \mu_f(D_f) \Bigg/ \left[\langle \delta'^2 \rangle - \omega'^2 - \frac{(\sigma - \omega' \omega)^2}{\sigma^2 - \omega^2} \right]^{1/2} , \tag{35}$$

where $\omega = \langle \delta \nu_{1/2} \rangle$ and $\omega' = \langle \delta' \nu_{1/2} \rangle$ are given by equations (E14) and (E15), respectively.

The conditional probability of D_f given upcrossing at σ – the main result of this subsection – is obtained dividing equation (32) by equation (11)

$$f_{\mathrm{up}}(D_f | \sigma) = \frac{\nu_f}{D_f} p_G(\nu_f | \nu_c) \frac{\mu_f F(X_f)}{\nu_c F(X)} ,$$
$$= \frac{\left(\delta_c / \sigma_{1/2} D_f^2 \right) e^{-\nu_{f,c}^2/2}}{\sqrt{2\pi(1 - \langle \nu \nu_{1/2} \rangle^2)}} \frac{\mu_f F(X_f)}{\nu_c F(X)} , \tag{36}$$

where $(\nu_f/D_f) p_G(\nu_f | \nu_c) = p(D_f | \nu_c)$, not surprisingly, is the conditional probability of the (non-Gaussian) variable D_f given ν_c, and

$$\nu_{f,c} \equiv \frac{\nu_f - \langle \nu \nu_{1/2} \rangle \nu_c}{\sqrt{1 - \langle \nu \nu_{1/2} \rangle^2}} = \frac{\delta_c}{\sigma_{1/2}} \frac{1/D_f - \langle \delta \delta_{1/2} \rangle / \sigma^2}{\sqrt{1 - \langle \nu \nu_{1/2} \rangle^2}} . \tag{37}$$

Recall also that $X = \Gamma \nu_c$. The conditional probability $f_{\mathrm{up}}(D_f | \sigma)$ depends on D_f directly, through $\nu_{f,c}$ and through μ_f (which appears also in X_f). As both $\nu_{f,c}$ and μ_f are proportional to $1/D_f$ in the small-D_f limit, equation (36) scales like $e^{-\nu_{f,c}^2/2}/D_f^3$. Hence, $f_{\mathrm{up}}(D_f | \sigma)$ is exponentially suppressed for small D_f, that is for large formation redshift z_f: it is exponentially unlikely for a halo to assemble half of its mass at very high redshift.

Like in the previous section, the Gaussian cut-off in equation (36) enables to define a characteristic value $D_\star(\sigma)$ of the formation time, below which $f_{\mathrm{up}}(D_f | \sigma)$ is exponentially suppressed, by requesting that $\nu_{f,c} = 1$. This definition corresponds to

$$D_\star(\sigma) = \frac{\delta_c / \sigma_{1/2}}{\langle \nu \nu_{1/2} \rangle \nu_c + \sqrt{1 - \langle \nu \nu_{1/2} \rangle^2}} , \tag{38}$$

which can then be solved for the typical formation redshift $z_\star$. Similarly, one may define the most likely formation time $D_{\max}$ by finding the value of D_f that maximizes equation (36). Because its expression is rather involved and not much more informative than $D_\star$, it is not reported here.

Expanding $D_\star$ in powers of $\Delta \sigma_{1/2} \equiv \sigma_{1/2} - \sigma$ (even though $\Delta \sigma_{1/2}/\sigma \simeq -(1/2)\mathrm{d}\log\sigma/\mathrm{d}\log M$ may not be small, in which case this expansion may just give a qualitative indication), one gets

$$D_\star \simeq 1 - \frac{\Delta \sigma_{1/2}}{\sigma} \left(1 + \frac{\sqrt{\langle \delta'^2 \rangle - 1}}{\nu_c} \right) \simeq 1 - \frac{1}{\alpha_\star} \frac{\Delta \sigma_{1/2}}{\sigma} , \tag{39}$$

confirming the intuitive relation between accretion rate and formation time. Haloes with smaller accretion rates today must have formed earlier, in order for their final mass to be the same. To derive this expression, $\langle \delta \delta_{1/2} \rangle$ was expanded up to second order in $\Delta\sigma$, using $\langle \delta \delta' \rangle = \sigma$ and $\langle \delta \delta'' \rangle = 1 - \langle \delta'^2 \rangle = \Gamma^{-2}$. Let us stress that, strictly speaking, the conditional probability $f_{\mathrm{up}}(D_f | \sigma)$ is not a well-defined probability distribution. For instance, just like $f_{\mathrm{up}}(\sigma)$, equation (36) is not normalized to unity when integrated over $0 < D_f < D$. This is an artefact introduced by the upcrossing approximation to the first-crossing problem, because equation (29) does not require trajectories to reach δ_c/D_f for the first time. As D_f gets close to D, most trajectories reaching δ_c/D_f do so with negative slope, or after one or more crossings, which leads to overcounting. For $D_f = D$,

trajectories that first crossed δ_c/D_f at σ cannot first cross again at $\sigma_{1/2}$, since $\sigma_{1/2} - \sigma$ remains finite: the true distribution should then have $f(D_f|\sigma) = 0$. This is clearly not the case for $f_{up}(D_f|\sigma)$. In spite of these shortcomings, equation (36) approximates well the true conditional PDF for $D_f \ll D_\star$, and the characteristic time $D_\star$ still provides a useful parametrization of the height of the tail.

A better approximation than equation (36) may be obtained by imposing an upcrossing condition at $\sigma_{1/2}$ as well

$$\frac{\delta_c}{D_f^2} \int_0^\infty \mathrm{d}\delta'\, \delta' \int_0^\infty \mathrm{d}\delta'_{1/2}\, p_G(\delta_c, \delta', \delta_c/D_f, \delta'_{1/2}). \tag{40}$$

Notice the absence in this expression of the Jacobian factor $\delta'_{1/2}$: this is because the constraint at $\sigma_{1/2}$ is not differentiated with respect to $\sigma_{1/2}$, but only with respect to D_f. This reformulation, which unfortunately does not admit a simple analytical expression, would improve the approximation for values of D_f closer to $D_\star$, but it would still not yield a formally well-defined PDF. Furthermore, the mean $\langle D_f|\sigma\rangle$ and all higher moments would still be infinite: these divergences are in fact a common feature of first passage statistics, which typically involve the inverse of Gaussian variables. For all these reasons, this calculation is not pursued further.

This section has formalized analytical predictions for accretion rates and formation times from the excursion set approach with correlated steps. It confirmed the tight correlation between the two quantities, according to which at fixed mass, early-forming haloes must have small accretion rates today. Because the focus is here on accounting for the presence of a saddle of the potential at finite distance, for simplicity and in order to isolate this effect we have restricted our analysis to the case of a constant threshold δ_c. More sophisticated models (e.g. scale-dependent barriers involving other stochastic variables that account for deviations from spherical collapse) could however be accommodated without extra conceptual effort (see Appendix G).

4 HALO STATISTICS NEAR SADDLES

Let us now quantify how the presence of a saddle of the large-scale gravitational potential affects the formation of haloes in its proximity. To do so, let us study the tracers introduced in the previous section (the distributions of upcrossing scale, accretion rate, and formation time) using conditional probabilities. The condition we enforce is that the upcrossing point (the centre of the excursion set trajectories) lies at a finite distance $\boldsymbol{r}$ from the saddle point. The focus will be on (filament-type) saddles of the potential that describe local configurations of the peculiar acceleration with two spatial directions of inflow (increasing potential) and one of outflow (decreasing potential). See Appendix C for other critical points. These initial regions will evolve into filaments (at least in the Zel'dovich approximation), where particles accumulate out of the neighbouring voids from two directions, and the saddle points filament centres, where the gravitational attraction of the two nodes balances out. A schematic representation of this configuration is given in Fig. 5.

The saddles are identified as points with null gradient of the gravitational potential, smoothed on a sphere of radius R_S (which is assumed to be larger than the halo's scale R). This condition guarantees that the mean peculiar acceleration of the sphere, which at first order is also the acceleration of its centre of mass, vanishes. That is, the null condition (for $i = 1, \ldots, 3$)

$$g_i \equiv \frac{1}{R_\star} \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \frac{i k_i}{k^2} \delta_m(\boldsymbol{k}) \frac{W(kR_S)}{\sigma_S} = 0, \tag{41}$$
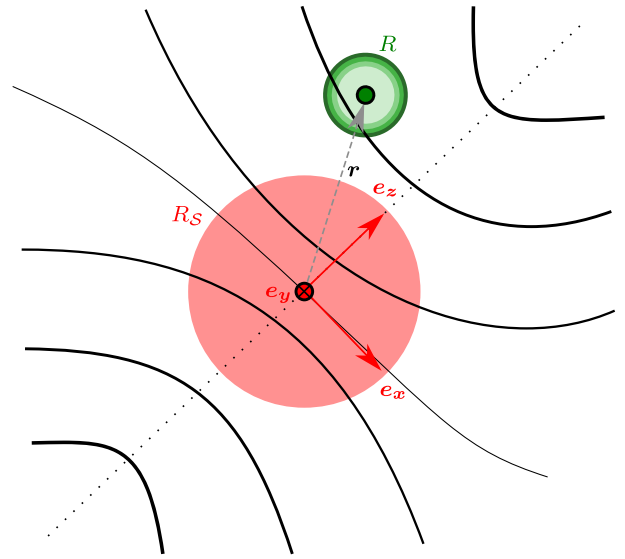
Figure 5. Illustration of the conditional excursion set smoothing on a few infinitesimally close scales around R (in green) at finite distance $\boldsymbol{r}$ from a saddle point of the gravitational potential smoothed on scale $R_S \gg R$ (in red). The eigenvectors $\boldsymbol{e}_x$ and $\boldsymbol{e}_z$ of the tidal tensor at the saddle give the directions of steepest increase and decrease of the potential, corresponding to maximum inflow and outflow, respectively. The region is compressed along $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$ and stretched along $\boldsymbol{e}_z$, thus creating a filament. The solid lines are isocontours of the mean density, the thickest the densest. The dotted line indicates a ridge of mean density (the filament), parallel to $\boldsymbol{e}_z$ near the saddle.

where $\sigma_S \equiv \sigma(R_S)$, is imposed on the mean gradient of the potential smoothed with a Top-Hat filter on scale R_S. This mean acceleration is normalized in such a way that $\langle g_i g_j\rangle = \delta_{ij}/3$ by introducing the characteristic length-scale[7]

$$R_\star^2 \equiv \int \mathrm{d}k\, \frac{P(k)}{2\pi^2} \frac{W^2(kR_S)}{\sigma_S^2}. \tag{42}$$

Having null peculiar acceleration, the patch sits at the equilibrium point of the attractions of what will become the two nodes at the end of the filament.[8]

The configuration of the large-scale potential is locally described by the rank 2 tensor

$$q_{ij} \equiv \frac{1}{\sigma_S} \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \frac{k_i k_j}{k^2} \delta_m(\boldsymbol{k}) W(kR_S), \tag{43}$$

which represents the Hessian of the perturbed potential smoothed on scale R_S, normalized so that $\langle \mathrm{tr}^2(q)\rangle = 1$. This tensor is the opposite of the so-called strain or deformation tensor. The peculiar gravitational acceleration at the surface of the sphere is proportional to $-q_{ij}r_j$. Thus, the trace $\mathrm{tr}(q) = \nu_S$ of q_{ij} describes the

---

[7] This scale is similar, but not equivalent, to the scale often defined in peak theory. Calling σ_i^2 the variance of the density field filtered with $k^{2i}W(kR)$, the $R_\star$ defined here is σ_{-1}/σ_0, while the peak theory scale is $\sqrt{3}\sigma_1/\sigma_2$.

[8] The mean gravitational acceleration g_i includes an unobservable infinite wavelength mode, which should in principle be removed. A way to circumvent the problem would be to multiply $W(kR_S)$ by a high-pass filter on some large-scale R_0 to remove modes with $k \lesssim 1/R_0$. Because g_i is set to 0, it does not introduce any anisotropy, but simply affects the radial dependence of the conditional statistics through its covariance $\langle g_i g_j\rangle$, which however is not very sensitive to long wavelengths. For this reason, this minor complication is ignored.

average infall (or expansion, if negative) acceleration of the three axes with respect to the background, while the anisotropic shear is given by the traceless part $\bar{q}_{ij} \equiv q_{ij} - \delta_{ij}\nu_{\mathcal{S}}/3$, which deforms the region by slowing down or accelerating each axis. By construction, $\langle \nu_{\mathcal{S}}\bar{q}_{ij} \rangle = 0$.

For the initial spherical patch to evolve into a filament, the eigenvalues q_i of q_{ij} must obey $q_1 < 0 < q_2 < q_3$ (see also Fig. D1). In this configuration, the Zel'dovich flow of the patch has one expanding direction and two infalling ones. The non-linear evolution is unlikely to revert this behaviour, and the spherical region will end up in a filament (Zel'dovich 1970; Bond, Kofman & Pogosyan 1996). There is no clear consensus on what the initial density of a protofilament should be for the structure to form at $z = 0$ (see however Shen et al. 2006). The value $\nu_{\mathcal{S}} = 1.2$ was chosen here, corresponding to a mean density of 0.8 within a sphere of $R_{\mathcal{S}} = 10 \, \mathrm{Mpc}h^{-1}$, which is about one standard deviation higher than the mean value for saddle points of this type (see Appendix D for details), and thus corresponds to a filament slightly more massive than the average (or to an average filament that has not completely collapsed yet). The qualitative results presented in this paper do not depend on the exact value of $\nu_{\mathcal{S}}$ (even though they obviously do at the quantitative level).

4.1 Expected impact of saddle tides

The mean and covariance of δ and δ' at $\mathbf{r}$ are modified by the presence of the saddle at the origin. The zero mean density field is replaced by $\delta - \langle \delta | \mathcal{S} \rangle$, where (using Einstein's convention as usual)

$$\langle \delta | \mathcal{S} \rangle = \langle \delta | \mathcal{S} \rangle \langle \delta \nu_{\mathcal{S}} \rangle \nu_{\mathcal{S}} + 3\langle \delta g_i \rangle g_i + \frac{15}{2} \langle \delta \bar{q}_{ij} \rangle \bar{q}_{ij} , \qquad (44)$$

where the correlation functions are evaluated at finite separation. Here, $\mathcal{S}$ stands for a filament-type saddle condition of zero gradient and two positive eigenvalues of the tidal tensor, see Fig. 5. The slope δ' is replaced by the derivative of this whole expression with respect to σ, which gives $\delta' - \langle \delta' | \mathcal{S} \rangle$, since the correlation functions of δ' with the saddle quantities correspond to the derivatives of the δ correlations. These modified height and slope no longer correlate with any saddle quantity. Thus, the abundance of the various tracers at $\mathbf{r}$ can be inferred from standard excursion sets of this effective density field. The building blocks of this effective excursion set problem – the variance of the field and of its slope, height, and slope of the effective barrier – are derived in full in Appendix F. The main text of this section discusses how the saddle condition affects the upcrossing statistics, and the excursion set proxies for accretion rate and formation time.

For geometrical reasons, since statistical isotropy is broken only by the separation vector, any angular dependence of the correlation functions may arise only as r_i or $r_i r_j$. Let us thus write equation (44) as

$$\langle \delta | \mathcal{S} \rangle = \xi_{00}\nu_{\mathcal{S}} + 3\xi_{11}\frac{r}{R_\star}\hat{r}_i g_i - 5\xi_{20}\frac{3\hat{r}_i\bar{q}_{ij}\hat{r}_j}{2} , \qquad (45)$$

where $\hat{r}_i \equiv r_i/r$ and the correlation functions $\xi_{\alpha\beta}(r, R, R_{\mathcal{S}})$ – whose exact form is given in equation (E11) – depend only on the radial separation $r = |\mathbf{r}|$ and the two smoothing scales, and have positive sign. Notice the presence of a minus sign in the shear term. In the frame of the saddle, oriented with the $\hat{z}$-axis in the direction of outflow,

$$\mathcal{Q} \equiv \hat{r}_i\bar{q}_{ij}\hat{r}_j = \bar{q}_3 \sin^2\theta \cos^2\phi + \bar{q}_2 \sin^2\theta \sin^2\phi + \bar{q}_1 \cos^2\theta , \quad (46)$$

where θ and ϕ are the usual cylindrical coordinates in the frame of the eigenvectors $(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1)$ of $\bar{q}_{ij}$ with eigenvalues $\bar{q}_3 > \bar{q}_2 > \bar{q}_1$.
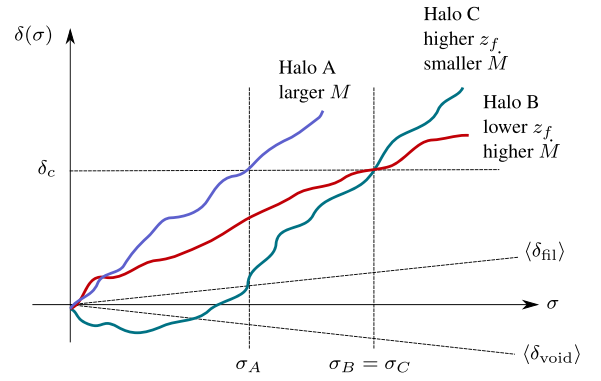
Figure 6. Pictorial representation of the effect of the presence of saddle point on the excursion set trajectories at a finite distance from it. Haloes A and B lie in the direction of the filament ($\mathcal{Q} \equiv \hat{r}_i\bar{q}_{ij}\hat{r}_j < 0$), where the mean density is higher than the average density. Halo C lies in the direction orthogonal to it ($\mathcal{Q} > 0$), where the mean density is lower. Haloes in the filament are likely to cross the collapsing threshold earlier, like halo A, than haloes in the voids. They thus tend to have larger mass. At fixed crossing scale $\sigma_{\mathrm{B}} = \sigma_{\mathrm{C}}$, haloes in the filament are likely to cross with shallower slopes, like halo B, than halo in the voids. At their half-mass scale $\sigma_{1/2} > \sigma_{\mathrm{A}}$, their trajectories tend to be lower. Hence, at fixed mass, haloes in the filaments tend to have larger accretion rates and to assemble half of their mass later. Conversely, haloes in the voids assemble their mass earlier, and then stop accreting.

When setting $g_i = 0$, an angular dependence can only appear as a functional dependence on $\mathcal{Q}(\hat{\mathbf{r}}) = \hat{r}_i\bar{q}_{ij}\hat{r}_j$. That is, a dependence on the direction $\hat{\mathbf{r}}$ with respect to the eigenvectors of the shear $\bar{q}_{ij}$. As shown by equation (45), a negative value of $\mathcal{Q}$ corresponds to a higher mean density, which makes it easier for δ to reach δ_{c} and for haloes to form. At fixed distance from the saddle point, halo formation is thus enhanced in the outflow direction with respect to the inflow direction: haloes are naturally more clustered in the filament than in the voids. Moreover, excursion set trajectories with a lower mean will tend to cross the barrier with steeper slopes than those crossing at the same scale but with a higher mean, and will reach higher densities at smaller scales. Hence, haloes of the same mass that form in the voids will form earlier and have a lower accretion rates. These trends are shown in Fig. 6.

To understand the radial dependence, one may expand equation (45) for small r away from the saddle, obtaining

$$\langle \delta | \mathcal{S} \rangle \simeq \langle \delta \nu_{\mathcal{S}} \rangle_{r=0}\nu_{\mathcal{S}} + \langle \delta \nabla^2 \nu_{\mathcal{S}} \rangle_{r=0}\frac{r^2}{2}\hat{r}_i q_{ij}\hat{r}_j ; \qquad (47)$$

whether the mean density increases or decreases with r depends on the sign of the eigenvalues, i.e. the curvatures of the saddle, of the full $\mathbf{q}$ defined in equation (43). Since $\langle \delta \nabla^2 \nu_{\mathcal{S}} \rangle < 0$, the mean density grows quadratically with r if $\hat{r}_i q_{ij}\hat{r}_j < 0$, and decreases otherwise. One thus expects the saddle point to be a maximum of halo number density, accretion rate, and formation time in the two directions perpendicular to the filament, and a minimum in the direction parallel to it (corresponding to the negative eigenvalue q_1).

4.2 Conditional halo counts

The conditional distribution of the upcrossing scale σ at finite distance $\mathbf{r}$ from a saddle point of the potential can be evaluated following the generic procedure described in Section 2.2, fixing

$$\{v_I\} = \{\nu_{\mathcal{S}}, 0, -\sqrt{5}(3\mathcal{Q}/2)\} \equiv \mathcal{S}(\mathbf{r}) \qquad (48)$$

as the constraint. With this replacement, equation (15) divided by $p_G(S)$ gives

$$f_{\mathrm{up}}(\sigma;\boldsymbol{r}) = \frac{\mathrm{e}^{-\nu_{c,S}^2/2}}{\sqrt{2\pi\mathrm{Var}(\delta|S)}}\,\mu_S\,F(X_S)\,, \qquad (49)$$

which is the sought conditional distribution, with

$$\mu_S(\boldsymbol{r}) \equiv \langle\delta'|\nu_c,S\rangle\,, \quad X_S(\boldsymbol{r}) \equiv \frac{\mu_S(\boldsymbol{r})}{\sqrt{\mathrm{Var}(\delta'|\nu_c,S)}}\,, \qquad (50)$$

as in equation (16). The effective threshold $\nu_{c,S}$ given the saddle condition is obtained replacing the generic constraint ν with S in equation (18).

The explicit calculation of the conditional quantities needed to compute $\nu_{c,S}$, μ_S, and X_S is carried out in Appendix F. The results of Appendix F2 [namely, equation (F13)] give

$$\nu_{c,S}(\boldsymbol{r}) \equiv \frac{\delta_c - \langle\delta|S\rangle}{\sqrt{\mathrm{Var}(\delta|S)}} = \frac{\delta_c - \xi_{00}\nu_S + \frac{15}{2}\xi_{20}\mathcal{Q}(\hat{r})}{\sqrt{\sigma^2 - \xi^2}}\,, \qquad (51)$$

consistently with equation (45), where

$$\xi^2(r) \equiv \xi_{00}^2(r) + 3\xi_{11}^2(r)r^2/R_\star^2 + 5\xi_{20}^2(r)\,. \qquad (52)$$

The effective slope parameters, obtained by replacing equations (F10) and (F11) into equation (50), are

$$\mu_S(\boldsymbol{r}) = \xi_I'\mathcal{S}_I + \frac{\sigma - \xi_I'\xi_I}{\sqrt{\sigma^2 - \xi^2}}\nu_{c,S}(\boldsymbol{r})\,, \qquad (53)$$

$$X_S(\boldsymbol{r}) = \mu_S(\boldsymbol{r}) \Big/ \left[\langle\delta'^2\rangle - \xi'^2 - \frac{(\sigma - \xi_I'\xi_I)^2}{\sigma^2 - \xi^2}\right]^{1/2}\,, \qquad (54)$$

in terms of the vectors

$$\xi(r) \equiv \{\xi_{00}(r),\,\sqrt{3}\xi_{11}(r)r/R_\star,\,\sqrt{5}\xi_{20}(r)\}\,, \qquad (55)$$

$$\xi'(r) \equiv \{\xi_{00}'(r),\,\sqrt{3}\xi_{11}'(r)r/R_\star,\,\sqrt{5}\xi_{20}'(r)\}\,. \qquad (56)$$

The correlation functions $\xi_{\alpha\beta}(r, R, R_S)$ and their derivatives $\xi_{\alpha\beta}' = \mathrm{d}\xi_{\alpha\beta}/\mathrm{d}\sigma$ are given in equations (E11) and (E12), respectively. Note that throughout the text, $\xi_{\alpha\beta}$ or $\xi_{\alpha\beta}(r)$ will be used as a shorthand for $\xi_{\alpha\beta}(r, R, R_S)$.

Equation (49), the main result of this subsection, is the conditional counterpart of equation (11), and is formally identical to it upon replacing ν_c, ν_c', and X with $\nu_{c,S}(\boldsymbol{r})$, $\nu_{c,S}'(\boldsymbol{r}) = -\mu_S(\boldsymbol{r})/\sqrt{\sigma^2 - \xi^2}$ and $X_S(\boldsymbol{r})$. The position-dependent threshold $\nu_{c,S}(\boldsymbol{r})$ and the slope parameter $\mu_S(\boldsymbol{r})$, given by equations (51) and (53), respectively, contain anisotropic terms proportional to $\mathcal{Q}$ These terms account for all the angular dependence of $f_{\mathrm{up}}(\sigma;\boldsymbol{r})$. In the large-mass regime, as $\{\xi_I'\} \simeq 0$, $X_S \simeq \nu_{c,S}/(1 - \xi^2) \gg 1$ and $F(X_S) \simeq 1$. The most relevant anisotropic contribution is thus the angular modulation of $\nu_{c,S}$, which raises or lowers the exponential tail of $f_{\mathrm{up}}(\sigma;\boldsymbol{r})$ along or perpendicular to the filament. Upcrossing, and hence halo formation, will be most likely in the direction that makes the threshold $\nu_{c,S}$ smallest, as this makes it easier for the stochastic process to reach it.

In analogy to the unconditional case, when a characteristic mass scale could be defined for which $\sigma = \delta_c$, equation (49) suggests to define the characteristic mass scale $\sigma_\star = \sigma(M_\star)$ for haloes near the saddle as the one for which $\nu_{c,S} = 1$ in equation (51). In the language of excursion sets, this request naturally sets the scale

$$\sigma_\star^2(\boldsymbol{r}) \equiv \left(\delta_c - \xi_{00}\nu_S + \frac{15}{2}\xi_{20}\mathcal{Q}\right)^2 + \xi^2(r)\,. \qquad (57)$$

This is now an implicit equation for $\sigma_\star$, because the RHS has a residual dependence on $\sigma_\star$ through $\xi_{\alpha\beta}(r, R(\sigma_\star), R_S)$, as shown in
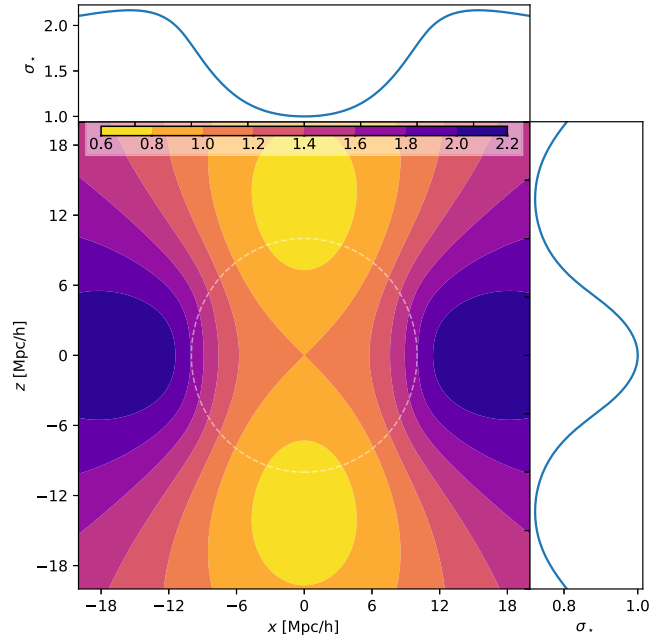
Figure 7. Isocontours in the x–z plane of the typical upcrossing scale $\sigma_\star$ around a saddle point [at $(0, 0)$]. The saddle point is defined using the values of Table D1. The profiles in the direction of the filament (z-direction) and of the void (x-direction) are plotted on the sides. The smoothing scale is $R = 1\,\mathrm{Mpc}\,h^{-1}$. They are obtained by solving equation (57) for $\sigma_\star$ at each point, with a ΛCDM power spectrum, and normalized to the value at the saddle point. In the filament, haloes form at a smaller σ (higher mass) and conversely in the void.

Appendix E. This equation can be solved numerically for $\sigma_\star$ and then for $M_\star$.

The angular dependence of $\sigma_\star(\boldsymbol{r})$ is entirely due to $\xi_{20}\mathcal{Q}$. Since the pre-factor of $\mathcal{Q} \equiv \hat{r}_i q_{ij} \hat{r}_j$ is positive, $\sigma_\star(\boldsymbol{r})$ will be smallest when $\boldsymbol{r}$ aligns with the eigenvector with the smallest eigenvalue, and $\mathcal{Q}$ is most negative. This happens when $\theta = 0$ in equation (46): that is, in the direction of positive outflow, along which a filament will form. Thus, in filaments haloes tend to be more massive than field haloes. The full radial and angular dependence of the characteristic mass scale $\sigma_\star$ is shown in Fig. 7.

4.3 Conditional accretion rate

The abundance of haloes of given mass and accretion rate at distance $\boldsymbol{r}$ from a saddle is obtained by replacing the probability distribution $p_G(\nu_c, \nu_c' + \nu_c/\sigma\alpha)$ in equation (23) with its conditional counterpart given the saddle constraint. As shown by equation (F12), this conditional distribution is equal to the distribution of the effective independent variables $\tilde{\nu}$ and $\delta' - \langle\delta'|\nu_c, S\rangle$ introduced in Section 2.2, times a Jacobian factor of $\sigma/(1 - \xi^2/\sigma^2)$. Furthermore, the relation (19) giving the excursion set slope in terms of the accretion rate reads in these new variables

$$\delta' - \langle\delta'|\nu_c, S\rangle = \frac{\nu_c}{\alpha} - \mu_S\,. \qquad (58)$$

Putting these two ingredients together, equation (23) becomes

$$f_{\mathrm{up}}(\sigma, \alpha;\boldsymbol{r}) = \frac{\nu_c^2}{\sigma^2\alpha^3}p_G(\nu_c, \nu_c' + \nu_c/\sigma\alpha|S)\,,$$

$$= \frac{\nu_c^2}{\alpha^3}\frac{\mathrm{e}^{-\left(\nu_{c,S}^2 + Y_{\alpha,S}^2\right)/2}}{2\pi\sqrt{(\sigma^2 - \xi^2)\mathrm{Var}(\delta'|\nu_c, S)}}\,, \qquad (59)$$

where $\mathrm{Var}\left(\delta'|\nu_c, \mathcal{S}\right)$ is given by equation (F17) and

$$Y_{\alpha,\mathcal{S}}(\boldsymbol{r}) \equiv \frac{\nu_c/\alpha - \mu_{\mathcal{S}}(\boldsymbol{r})}{\sqrt{\mathrm{Var}\left(\delta'|\nu_c, \mathcal{S}\right)}} \,, \tag{60}$$

with $\mu_{\mathcal{S}}(\boldsymbol{r})$ given by equation (53). Again, like equation (23), this result could be obtained by taking $\langle\delta'|\nu_c, \alpha, \mathcal{S}\rangle = \nu_c/\alpha$ and the limit $\mathrm{Var}\left(\delta'|\nu_c, \alpha, \mathcal{S}\right) \to 0$ in equation (16), which would give $F(X_{\alpha,\mathcal{S}}) = 1$.

To investigate the anisotropy of the accretion rate for haloes of the same mass, one needs the conditional probability of α given upcrossing at σ, that is the ratio of equations (59) and (49). This conditional probability reads

$$f_{\mathrm{up}}(\alpha|\sigma; \boldsymbol{r}) = \frac{\nu_c e^{-Y_{\alpha,\mathcal{S}}^2/2}}{\alpha^3 \sqrt{2\pi\mathrm{Var}\left(\delta'|\nu_c, \mathcal{S}\right)}} \frac{\nu_c}{\mu_{\mathcal{S}} F(X_{\mathcal{S}})} \,, \tag{61}$$

with $\mu_{\mathcal{S}}(\boldsymbol{r})$ and $X_{\mathcal{S}}(\boldsymbol{r})$ given by equations (53) and (54), respectively. The second fraction in this expression is thus a normalization factor that does not depend on α, and which tends to 1 when $\nu_c \gg 1$ in the large-mass limit. Equation (61) is the main result of this subsection. It depends on the angular position $\hat{\boldsymbol{r}}$ through the terms $\xi_{20}' \mathcal{Q}$ and $\xi_{20} \mathcal{Q}$ contained in $\mu_{\mathcal{S}}(\boldsymbol{r})$, and thus also in $Y_{\alpha\mathcal{S}}$ and $X_{\mathcal{S}}$. The angular dependence is now weighted by two different functions $\xi_{20}(r)$ and $\xi_{20}'(r)$, whose relative amplitude matters to determine the overall effect.

To understand the angular variation of the exponential tail of this distribution, let us focus on how $Y_\alpha(\boldsymbol{r})$ depends on $\hat{\boldsymbol{r}}$. That is, on the anisotropic part of $-\mu_{\mathcal{S}}(\boldsymbol{r})$. In the large-mass limit, when $\sigma\xi'_{\alpha\beta}(r) \ll \xi_{\alpha\beta}(r)$, equation (53) tells us that the anisotropic part of $Y_\alpha(\boldsymbol{r})$ is proportional to $-\xi_{20}\mathcal{Q}$, with a proportionality factor that is always positive and $\mathcal{O}(1)$. Thus, the modulation has the opposite sign of the anisotropic part of $\nu_{c,\mathcal{S}}$, given in equation (51): for trajectories with the same upcrossing scale, the probability of having a given accretion rate is lowest in the direction of the eigenvector of $\bar{q}_{ij}$ with the lowest (most negative) eigenvalue, for which Y_α is largest. That is, for haloes with the same mass, the probability of having a given accretion rate is lowest along the ridge of the potential saddle, which will become the filament.

The typical accretion rate $\alpha_\star$ of the excursion set haloes described by the distribution (61) corresponds to the condition $Y_{\alpha_\star,\mathcal{S}} = 1$. This definition transforms equation (27) into

$$\alpha_\star(\sigma, \boldsymbol{r}) \equiv \frac{\nu_c}{\sqrt{\mathrm{Var}\left(\delta'|\nu_c, \mathcal{S}\right)} + \mu_{\mathcal{S}}(\boldsymbol{r})} \,, \tag{62}$$

where $\mathrm{Var}\left(\delta'|\nu_c, \mathcal{S}\right)$ and $\mu_{\mathcal{S}}(\boldsymbol{r})$ are given by equations (F17) and (53). In the limit of small anisotropy, the angular variation of the typical accretion rate is

$$\Delta\alpha_\star(\sigma, \boldsymbol{r}) = \frac{\alpha_\star^2|_{\bar{q}=0}}{\nu_c} \frac{15}{2} \left[\xi_{20}' - \frac{\sigma - \xi_I'\xi_I}{\sigma^2 - \xi^2} \xi_{20} \right] \hat{r}_i \bar{q}_{ij} \hat{r}_j \,, \tag{63}$$

where $\alpha_\star|_{\bar{q}=0}$ – the value of $\alpha_\star(\sigma, \boldsymbol{r})$ when $\bar{q}_{ij} = 0$ – is function of r but not of the angles. Therefore, at a fixed distance r from the saddle, haloes that form in the direction of the filament tend to have higher accretion rates than haloes with the same mass that form in the orthogonal direction. The full dependence of the characteristic accretion rate $\alpha_\star$ for haloes of the same mass on the position with respect to the saddle point of the potential is shown in Fig. 8. The figure shows that the saddle point is a local minimum of the accretion rate along the direction connecting two regions with high density of final objects, which is two peaks of the final halo density field. This is consistent with the result that the accretion of haloes in filaments is suppressed by the effect of the tidal forces (as shown by, e.g.

Hahn et al. 2009; Borzyszkowski et al. 2016). The threshold $\delta \lesssim \delta_c$ is reached at smaller σ in filaments than in void, hence the slope is smaller at upcrossing. It is shown schematically in the top panel of Fig. B3. A verification with a constrained random field is shown in the bottom panel of Fig. B3. The details of the method used are given in Appendix B.

One can also evaluate the mean of the conditional distribution (61) following equation (26), integrating $\alpha f_{\mathrm{up}}(\alpha|\sigma, \mathcal{S})$ over the range of positive α. This conditional mean value is

$$\langle\alpha|\sigma\rangle(\boldsymbol{r}) = \frac{\nu_c}{\mu_{\mathcal{S}}(\boldsymbol{r})} \frac{1 + \mathrm{erf}(X_{\mathcal{S}}(\boldsymbol{r})/\sqrt{2})}{2F(X_{\mathcal{S}}(\boldsymbol{r}))} \,; \tag{64}$$

in the large-mass regime, where $X_{\mathcal{S}} \gg 1$ and the whole second fraction tends to 1, the position-dependent conditional mean $\langle\alpha|\sigma\rangle(\boldsymbol{r})$ is essentially the same as $\alpha_\star(\boldsymbol{r})$ defined in equation (62). As for $f_{\mathrm{up}}(\alpha|\sigma)$, all higher order moments are ill defined. One can also find useful information in the most likely accretion rate

$$\alpha_{\max}(\sigma, \boldsymbol{r}) = \frac{\nu_c^2}{6\mathrm{Var}\left(\delta'|\nu_c, \mathcal{S}\right)} \left[\sqrt{1 + \frac{12}{X_{\mathcal{S}}^2(\boldsymbol{r})}} - 1 \right] \,, \tag{65}$$

which generalizes equation (28) to the presence of a saddle point at distance $\boldsymbol{r}$. The same conclusion holds here namely the most likely accretion rate increases from voids to saddles and saddles to nodes. The following only considers maps of $\alpha_\star(\sigma, \boldsymbol{r})$, since the information encoded in $\alpha_{\max}(\sigma, \boldsymbol{r})$ and $\langle\alpha|\sigma\rangle(\boldsymbol{r})$ is somewhat redundant.

4.4 Conditional formation time

The formation time in the vicinity of a saddle is obtained by fixing the saddle parameters $\mathcal{S} = \{\nu_{\mathcal{S}}, \hat{r}_i g_i, \hat{r}_i \bar{q}_{ij} \hat{r}_j\}$, with $g_i = 0$, besides $\nu = \nu_c$ and $\nu_{1/2} = \nu_f$. A five-dimensional constraint on the Gaussian variables must now be dealt with, and mapped into $\{\sigma, D_f, \mathcal{S}\}$. Since the mapping of the saddle parameters is the identity, the Jacobian of the transformation still gives $|\nu' - \nu_c'|\nu_f/D_f$, like in Section 3.2 (where there was no saddle constraint). The formalism outlined in Section 2.2 still applies: the joint probability of upcrossing at σ with formation time D_f given the saddle is obtained replacing $\{\nu\}$ with $\{\nu_f, \mathcal{S}\}$ in equation (16), multiplying by the Jacobian ν_f/D_f and dividing by the probability $p_G(\mathcal{S})$ of the saddle. The result is

$$f_{\mathrm{up}}(\sigma, D_f; \boldsymbol{r}) = \frac{\nu_f}{D_f} p_G(\nu_c, \nu_f|\mathcal{S}) \frac{\mu_{f,\mathcal{S}}}{\sigma} F(X_{f,\mathcal{S}}) \tag{66}$$
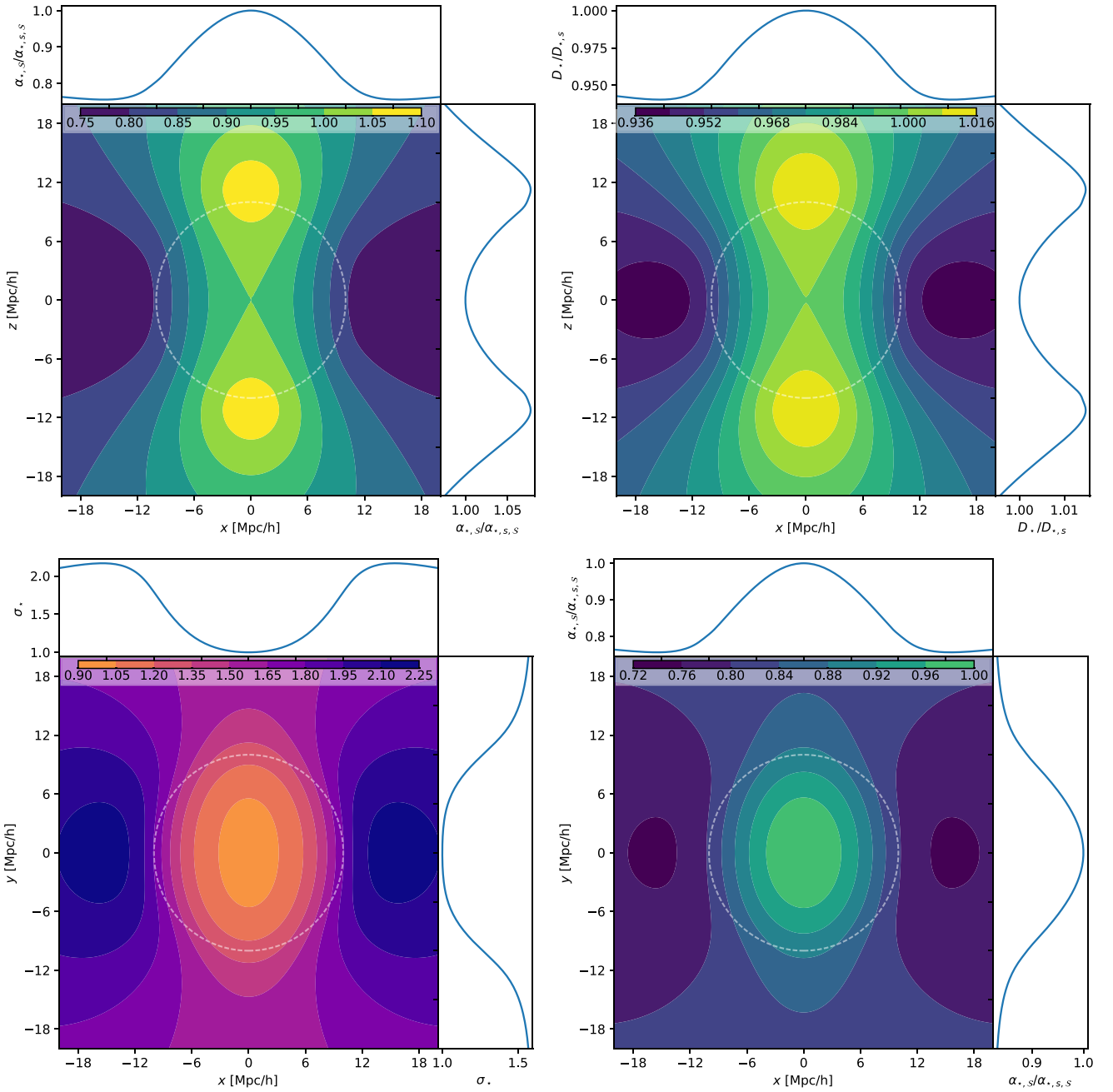
which extends equation (32) by including the presence of a saddle point of the potential at distance $\boldsymbol{r}$, with

$$\mu_{f,\mathcal{S}} \equiv \langle\delta'|\nu_f, \nu_c, \mathcal{S}\rangle \,, \quad X_{f,\mathcal{S}} \equiv \frac{\mu_{f,\mathcal{S}}}{\sqrt{\mathrm{Var}\left(\delta'|\nu_f, \nu_c, \mathcal{S}\right)}} \,. \tag{67}$$

The conditional mean and variance of δ' given $\{\nu_f, \nu_c, \mathcal{S}\}$ are explicitly computed in Appendix F4, equations (F30) and (F31).

The conditional probability of the formation time D_f given σ at a distance $\boldsymbol{r}$ from the saddle follows dividing equation (66) by $f_{\mathrm{up}}(\sigma|\boldsymbol{r})$, given by equation (49). This ratio – which is the main result of this section – gives

$$\begin{aligned} f_{\mathrm{up}}(D_f|\sigma; \boldsymbol{r}) &= \frac{\nu_f}{D_f} p_G(\nu_f|\nu_c, \mathcal{S}) \frac{\mu_{f,\mathcal{S}}}{\mu_{\mathcal{S}}} \frac{F(X_{f,\mathcal{S}})}{F(X_{\mathcal{S}})} \,, \\ &= \frac{(\delta_c/D_f^2)e^{-\nu_{f,c,\mathcal{S}}^2/2}}{\sqrt{2\pi\mathrm{Var}\left(\delta_{1/2}|\nu_c, \mathcal{S}\right)}} \frac{\mu_{f,\mathcal{S}}}{\mu_{\mathcal{S}}} \frac{F(X_{f,\mathcal{S}})}{F(X_{\mathcal{S}})} \,. \end{aligned} \tag{68}$$

Figure 8. Isocontours in the x–z plane of the typical accretion rate $\alpha_\star$ (upper left) and formation time $D_\star$ (upper right) around a saddle point [at (0, 0)] and in the x–y plane of the characteristic upcrossing scale $\sigma_\star$ (lower left) and typical accretion rate (lower right). The saddle point is defined using the values of Table D1. The profiles going through the saddle point in the x–z (upper panels) and x–y (lower panels) planes are plotted on the sides. The smoothing scale is $R = 1\,\mathrm{Mpc}\,h^{-1}$. They were obtained with a ΛCDM power spectrum, and normalized to the value at the saddle point. Since the filament has higher mean density, excursion set trajectories upcrossing at a given σ have shallower slopes. Hence, typical haloes are more massive in filaments and at fixed mass, haloes forming in the filament have larger accretion rates at $z = 0$ and form later. The same hierarchy exists between the two perpendicular directions.

Equation (68) provides the counterpart of equation (36) near a saddle point, in terms of the effective threshold

$$\nu_{\mathrm{f,c},\mathcal{S}}(D_{\mathrm{f}}, \boldsymbol{r}) \equiv \frac{\delta_{\mathrm{c}}/D_{\mathrm{f}} - \langle \delta_{1/2} | \nu_{\mathrm{c}}, \mathcal{S} \rangle}{\sqrt{\mathrm{Var}\left(\delta_{1/2} | \nu_{\mathrm{c}}, \mathcal{S}\right)}} \,, \tag{69}$$

with

$$\langle \delta_{1/2} | \nu_{\mathrm{c}}, \mathcal{S} \rangle = \xi_{1/2} \cdot \mathcal{S} + \frac{\langle \delta \delta_{1/2} \rangle - \xi \cdot \xi_{1/2}}{\sigma^2 - \xi^2} (\delta_{\mathrm{c}} - \xi \cdot \mathcal{S}) \,, \tag{70}$$

$$\mathrm{Var}\left(\delta_{1/2} | \nu_{\mathrm{c}}, \mathcal{S}\right) = \sigma_{1/2}^2 - \xi_{1/2}^2 - \frac{(\langle \delta \delta_{1/2} \rangle - \xi \cdot \xi_{1/2})^2}{\sigma^2 - \xi^2} \,. \tag{71}$$

It also depends on the effective upcrossing parameters $\mu_{\mathcal{S}}(\boldsymbol{r})$ and $X_{\mathcal{S}}(\boldsymbol{r})$, given in equations (50)–(53). The explicit forms of the functions $\mu_{\mathrm{f},\mathcal{S}}(D_{\mathrm{f}}, \boldsymbol{r})$ and $X_{\mathrm{f},\mathcal{S}}(D_{\mathrm{f}}, \boldsymbol{r})$ are reported in Appendix F4 for convenience [equations (F33) and (F34)].

Note that in equation (68), $f_{\mathrm{up}}(D_{\mathrm{f}} | \sigma; \boldsymbol{r})$ depends on D_{f} also through $\nu_{\mathrm{f,c},\mathcal{S}}$ and $\mu_{\mathrm{f},\mathcal{S}}$. For early formation times ($D_{\mathrm{f}} \ll 1$), the

conditional mean $\langle \delta' | \nu_{\mathrm{f}}, \nu_{\mathrm{c}}, \mathcal{S} \rangle$ becomes large, since the trajectory must reach a very high value at $\sigma_{1/2}$. Hence, $\mu_{\mathrm{f},\mathcal{S}}(D_{\mathrm{f}}, \boldsymbol{r}) \propto 1/D_{\mathrm{f}}$. In this limit, the last ratio in equation (68) above tends to 1, and $f_{\mathrm{up}}(D_{\mathrm{f}}|\sigma; \boldsymbol{r}) \propto (1/D_{\mathrm{f}}^3)\exp(-\nu_{\mathrm{f},\mathrm{c},\mathcal{S}}^2/2)$, with a proportionality constant that does not depend on the angle. Then, the probability decays exponentially for small D_{f} as $\nu_{\mathrm{f},\mathrm{c},\mathcal{S}}$ grows. The typical formation time $D_\star = D(z_\star)$ can be defined as that value for which $\nu_{\mathrm{f},\mathrm{c},\mathcal{S}} = 1$ and this exponential cut-off stops being effective, that is

$$D_\star(\boldsymbol{r}, \sigma) \equiv \frac{\delta_{\mathrm{c}}}{\sqrt{\mathrm{Var}\left(\delta_{1/2} | \nu_{\mathrm{c}}, \mathcal{S}\right)} + \langle \delta_{1/2} | \nu_{\mathrm{c}}, \mathcal{S} \rangle}, \tag{72}$$

which provides the anisotropic generalization of the expression given in equation (38). The explicit expression for the conditional mean $\langle \delta_{1/2} | \nu_{\mathrm{c}}, \mathcal{S} \rangle$ and variance $\mathrm{Var}\left(\delta_{1/2} | \nu_{\mathrm{c}}, \mathcal{S}\right)$ are given by equations (70) and (71), respectively.

As the angular variation of $\langle \delta_{1/2} | \nu_{\mathrm{c}}, \mathcal{S} \rangle$ is approximately

$$\frac{15}{2} \Delta\sigma_{1/2} \xi_{20}(r) \mathcal{Q}(\hat{\boldsymbol{r}}), \tag{73}$$

where $\mathcal{Q}(\hat{\boldsymbol{r}}) \equiv \hat{r}_i \bar{q}_{ij} \hat{r}_j$, $\Delta\sigma_{1/2} = \sigma_{1/2} - \sigma > 0$, the formation time $D_\star$ is larger when $\boldsymbol{r}$ is aligned with the eigenvector with the most negative eigenvalue, corresponding to the direction of the filament. One has in fact

$$\Delta D_\star(\boldsymbol{r}, \sigma) = -\frac{D_\star^2|_{\bar{q}=0}}{\delta_{\mathrm{c}}} \frac{15}{2} \Delta\sigma_{1/2} \xi_{20}(r) \mathcal{Q}(\hat{\boldsymbol{r}}), \tag{74}$$

where $D_\star$ depends only on the radial distance r, which shows that at a fixed distance from the saddle point, haloes in the direction of the filament tend to form later (larger $D_\star$). The saddle point is thus a minimum of the half-mass time $D_\star$ along the direction of the filament, that is a maximum of $z_\star$: haloes that form at the saddle point assemble most of their mass the earliest. Fig. 8 displays a cross-section of a map of $D_\star$ in the frame of the saddle.

5 ASTROPHYSICAL REFORMULATION

The joint and conditional PDFs derived in Sections 2–4 were expressed in terms of variables (σ, α, and D_{f}) that are best suited for the excursion set theory. Now, for the sake of connecting to observations and gathering a wider audience, let us write explicitly what the main results of those sections – equations (14), (25), and (36), and their constrained counterparts (49), (61), and (68) – imply in terms of astrophysically relevant quantities like the distribution of mass, accretion rate, and formation time of DM haloes.

5.1 Unconditional halo statistics

The upcrossing approximation provides an accurate analytical solution of the random walk problem formulated in the Extended Press–Schechter model, for a Top-Hat filter in real space and a realistic power spectrum. In this framework, the mass fraction in haloes of mass M is

$$\frac{M}{\bar{\rho}} \frac{\mathrm{d}n}{\mathrm{d}M} = \left| \frac{\mathrm{d}\sigma}{\mathrm{d}M} \right| f_{\mathrm{up}}(\sigma(M)), \tag{75}$$

with $f_{\mathrm{up}}(\sigma)$ given by equation (14) and is a function of mass via equation (1). For instance, for a power-law power spectrum $P(k) \propto k^{-n}$ with index $n = 2$ one has $M/M_\star = (\sigma/\sigma_\star)^{-6}$. The general power-law result $M \propto \sigma^{6/(n-3)}$ follows from equation (E17).

The excursion set approach also establishes a natural relation between the accretion rate of the halo and the slope of the trajectory at barrier crossing. One can thus predict the joint statistics of σ and

of the excursion set proxy $\alpha \equiv \nu_{\mathrm{c}}/[\mathrm{d}(\delta - \delta_{\mathrm{c}})/\mathrm{d}\sigma]$ for the accretion rate. In order to get the joint mass fraction in haloes of mass M and accretion rate $\dot{M}$, one needs to introduce the Jacobian of the mapping from (σ, α) to (M, $\dot{M}$). Since $\sigma(M)$ does not depend on α, this Jacobian has the simple factorized form $|\mathrm{d}\sigma/\mathrm{d}M||\mathrm{d}\alpha/\mathrm{d}\dot{M}|$. Since $\mathrm{d}\alpha/\mathrm{d}\dot{M} = \alpha/\dot{M}$ from equation (20), one can write the joint analogue of equation (75) as

$$\frac{M\dot{M}}{\bar{\rho}} \frac{\mathrm{d}^2 n}{\mathrm{d}M\mathrm{d}\dot{M}} = \left| \frac{\mathrm{d}\log\sigma}{\mathrm{d}M} \right| \sigma\alpha f_{\mathrm{up}}(\sigma, \alpha), \tag{76}$$

where $f_{\mathrm{up}}(\sigma, \alpha)$ is now given by equation (23), whereas $\sigma(M)$ and $\alpha(M, \dot{M})$ are functions of M and $\dot{M}$ via equations (1) and (20), respectively. From the ratio of equations (76) and (75), the expected mean density of haloes of given mass and accretion rate can be reformulated as

$$\dot{M} \frac{\mathrm{d}^2 n}{\mathrm{d}M\mathrm{d}\dot{M}} = \alpha f_{\mathrm{up}}(\alpha|\sigma) \frac{\mathrm{d}n}{\mathrm{d}M}, \tag{77}$$

where $f_{\mathrm{up}}(\alpha|\sigma)$ is given by equation (25). This expression relates analytically the number density of haloes binned by mass and accretion rate to the usual mass function.

Similarly, the joint mass fraction of haloes of mass M and formation time z_{f} (defined as the redshift at which the halo has assembled half of its mass) can be inferred from the joint statistics of σ and $D_{\mathrm{f}} \equiv \delta_{\mathrm{c}}/\delta(\sigma_{1/2})$, where $\sigma_{1/2} \equiv \sigma(M/2)$ is the scale containing half of the initial volume. The redshift dependence of the growth function $D(z)$ is defined by equation (4). Hence, the mass fraction in haloes of given mass M and formation time z_{f} is

$$\frac{M}{\bar{\rho}} \frac{\mathrm{d}^2 n}{\mathrm{d}M\mathrm{d}z_{\mathrm{f}}} = \frac{\mathrm{d}\sigma}{\mathrm{d}M} \frac{\mathrm{d}D_{\mathrm{f}}}{\mathrm{d}z_{\mathrm{f}}} f_{\mathrm{up}}(\sigma, D_{\mathrm{f}}), \tag{78}$$

and its conditional is

$$\frac{\mathrm{d}^2 n}{\mathrm{d}M\mathrm{d}z_{\mathrm{f}}} = \frac{\mathrm{d}D_{\mathrm{f}}}{\mathrm{d}z_{\mathrm{f}}} f_{\mathrm{up}}(D_{\mathrm{f}}|\sigma) \frac{\mathrm{d}n}{\mathrm{d}M}, \tag{79}$$

where the joint and conditional distributions of D_{f} and σ are given by equations (32) and (36), respectively.

Interestingly, while the excursion set mass function is subject to the limitation of upcrossing theory, the conditional statistics of accretion rate, or formation redshift, at given mass should be considerably more accurate. This is because the main shortcoming of excursion sets is the lack of a prescription for where to centre in space each set of concentric spheres giving a trajectory. These spheres are placed at random locations, whereas they should insist on the centre of the protohalo. However, choosing a better theoretical model (e.g. the theory of peaks) to set correctly the location of the excursion set trajectories would not dramatically modify the conditional statistics. Changing the model would modify the function $F(x)$, defined in equation (13), that modulates each PDF. In conditional statistics, only ratios of this function appear, which are rather model independent, whereas the probability of the constraint does not appear. The relevant part for our analysis – the exponential cut-off of each conditional distribution given the constraint – would not change. Hence, even though equation (75) does not provide a good mass function $\mathrm{d}n/\mathrm{d}M$, one may argue that the relations (77) and (79) are still accurate in providing the joint abundance statistics of mass and accretion rate, or mass and formation redshift, once a better model – or even a numerical fit – is used to infer $\mathrm{d}n/\mathrm{d}M$.

5.2 Halo statistics in filamentary environments

In the tide of a saddle of given height and curvature, equations (75), (76), and (78) remain formally unchanged, except for the
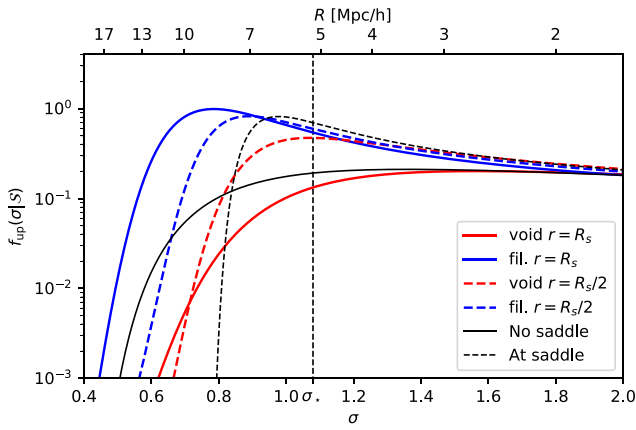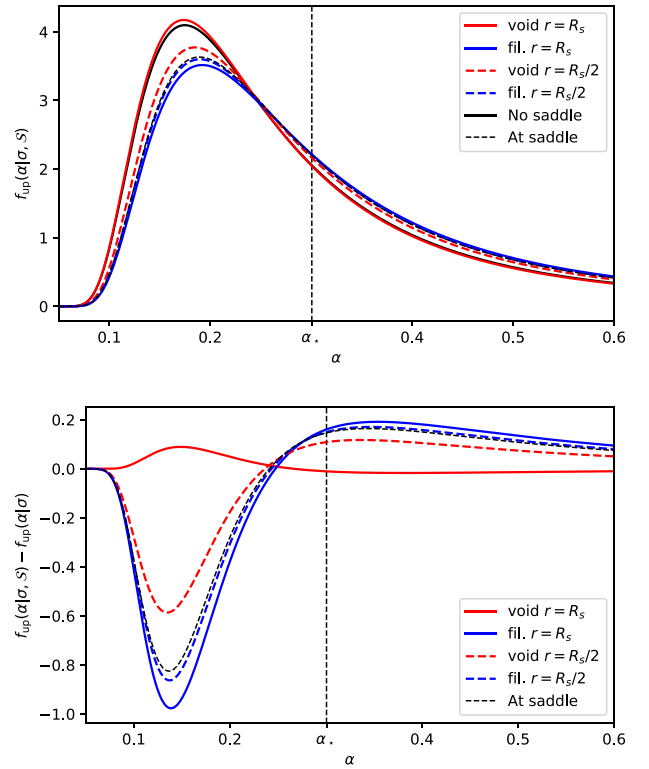
Figure 9. PDF of σ at upcrossing given the saddle point in the x (void, in red) and z (filament, in blue) directions at distance $r = 10$ Mpc h^{-1} (solid lines) and $r = 5$ Mpc h^{-1} (dashed lines). The saddle point is defined using the values of Table D1. The PDF without the saddle point is shown in black and at the saddle point in dashed black. The value of $\sigma_\star$ at the saddle point is shown by the vertical dashed line. In the filament, the PDF is boosted for small values of σ: there are more massive haloes in the filament. The opposite trend is seen in the void.

replacement of $f_{up}(\sigma)$, $f_{up}(\sigma, \alpha)$, and $f_{up}(\sigma, D_f)$ by their position-dependent counterparts $f_{up}(\sigma; \boldsymbol{r})$, $f_{up}(\sigma, \alpha; \boldsymbol{r})$, and $f_{up}(\sigma, D_f; \boldsymbol{r})$ conditioned to the presence of a saddle, given by equations (49), (59), and (66), respectively. Similarly, in equations (77) and (79), one should substitute the distribution $f_{up}(\alpha|\sigma)$ and $f_{up}(D_f|\sigma)$ by their conditional counterparts $f_{up}(\alpha|\sigma; \boldsymbol{r})$ and $f_{up}(D_f|\sigma; \boldsymbol{r})$ of accretion rate and formation time at fixed halo mass, given by equations (61) and (68).

These functions depend on the mass M, accretion rate $\dot{M}$, and formation time z_f of the halo through $\sigma(M)$, $\alpha(M, \dot{M})$, and $D_f(z_f)$, as before. However, conditioning on $\mathcal{S}$ introduces a further dependence on the geometry of the environment (the height $\nu_\mathcal{S}$ of the saddle and its anisotropic shear $\bar{q}_{ij}$) and on the position $\boldsymbol{r}$ of the halo with respect to the saddle point. This dependence arises because the saddle-point condition modifies the mean and variance of the stochastic process (δ, δ') – the height and slope of the excursion set trajectories – in a position-dependent way, making it more or less likely to form haloes of given mass and assembly history within the environment set by $\mathcal{S}$. The mean becomes anisotropic through $\mathcal{Q} = \hat{r}_i \bar{q}_{ij} \hat{r}_j$, and both mean and variance acquire radial dependence through the correlation functions $\xi_{\alpha\beta}$ and $\xi'_{\alpha\beta}$, defined in equation (E12), which depend on r, $R_\mathcal{S}$, and R [the variance remains isotropic because the variance of $\bar{q}_{ij}$ is still isotropic, see e.g. equation (71) and Appendix E].

The relevant conditional distributions are displayed in Figs 9–11. The plots show that haloes in the outflowing direction (in which the filament will form) tend to be more massive, with larger accretion rates and forming later than haloes at the same distance from the saddle point, but located in the infalling direction (which will become a void). This trend strengthens as the distance from the centre increases. The saddle point is thus a minimum of the expected mass and accretion rate of haloes, and a maximum of formation redshift, as one moves along the filament. The opposite is true as one moves perpendicularly to it. This behaviour is consistent with the expectation that filamentary haloes have on average lower mass and accretion rate, and tend to form earlier, than haloes in peaks.

Figure 10. PDF of α at upcrossing given the smoothing scale and the saddle point in the x (void, in red) and z (filament, in blue) directions at distance $r = 10$ Mpc h^{-1} (solid lines) and $r = 5$ Mpc h^{-1} (dashed lines) (upper panel) compared to the PDF without the saddle point (lower panel). The saddle point is defined using the values of Table D1. The PDF with no saddle point is shown in solid black and the PDF at the saddle point in dashed black. In the filament, the PDF is boosted at its high end: haloes accrete more. The opposite trend is seen in the void.
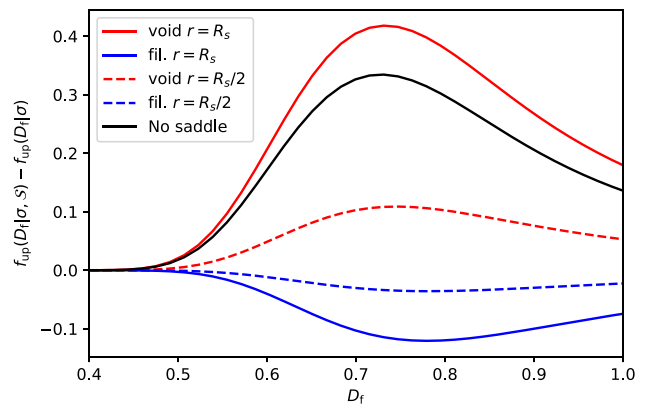
Figure 11. PDF of D_f at upcrossing given the smoothing scale and the saddle point in the x (void, in red) and z (filament, in blue) directions at distance $r = 10$ Mpc h^{-1} (solid lines) and $r = 5$ Mpc h^{-1} (dashed lines) and without saddle point (black) compared to the PDF at the saddle point. The saddle point is defined using the values of Table D1. In the filament, the PDF is boosted at the late formation end: haloes form later. The opposite trend is seen in the void.
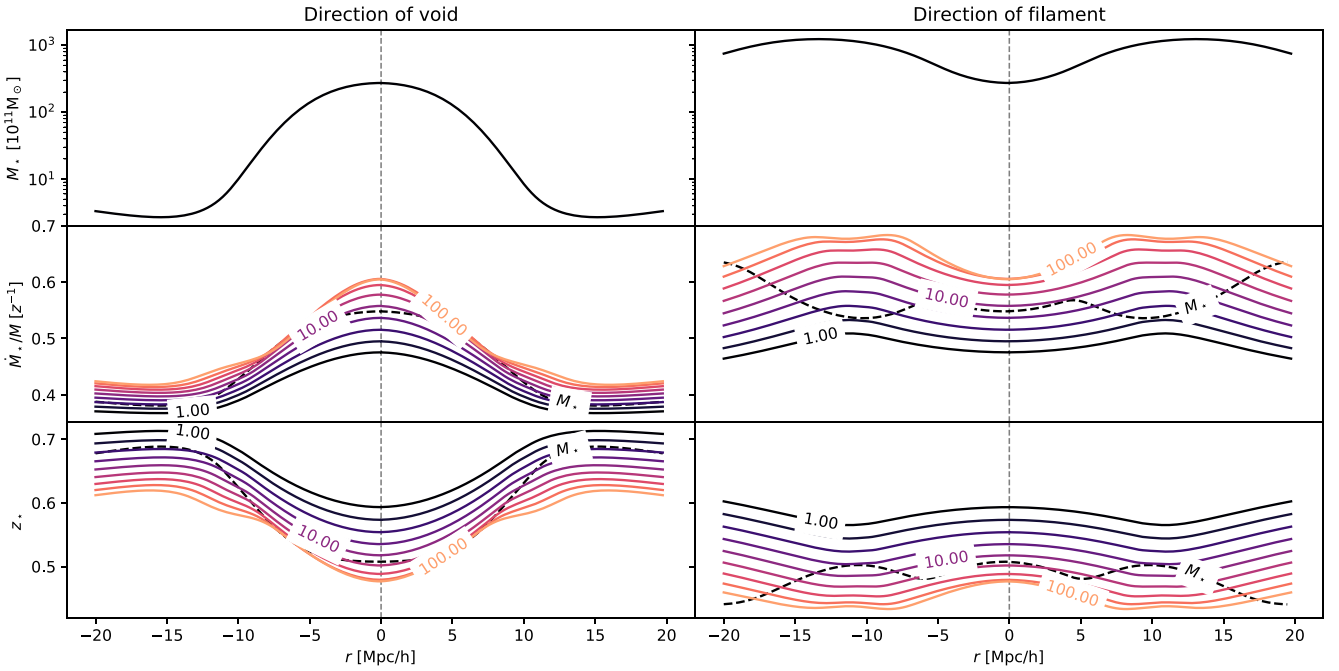
Figure 12. Top: plot of the typical mass $M_\star$, middle: the typical specific accretion rates $\dot{M}/M$, and bottom: the formation redshifts $z_\star$ for different masses as a function of the distance to the saddle point, left: in the direction of the void and right: in the direction of the filament. The colour of each line encodes the smoothing scale (hence the mass), from dark to light $M = 10^{11}\ M_\odot\ h^{-1}$ ($R = 0.8$ Mpc h^{-1}) to $M = 10^{13}\ M_\odot\ h^{-1}$ ($R = 3.7$ Mpc h^{-1}) logarithmically spaced; the dashed line is evaluated at $M = M_\star$. Labels are given in unit of $10^{11}\ M_\odot\ h^{-1}$. The saddle point has been defined using the values given in Table D1. More massive haloes accrete more and form later than less massive ones. At the typical mass, the space variation of the specific accretion rate and the formation redshift is smaller in the direction of the filament than in the direction of the void.

To better quantify these trends let us define the tidally modified characteristic quantities

$$M_\star(\boldsymbol{r}) = M(\sigma_\star(\boldsymbol{r})),\tag{80}$$

$$\dot{M}_\star(\boldsymbol{r}, M) = -\frac{\mathrm{d}\log D}{\mathrm{d}z}\frac{\mathrm{d}M}{\mathrm{d}\log\sigma}\alpha_\star(\boldsymbol{r}, \sigma),\tag{81}$$

$$z_\star(\boldsymbol{r}, M) = z(D_\star(\boldsymbol{r})) \simeq 1/D_\star(\boldsymbol{r}, \sigma) - 1,\tag{82}$$

giving the typical mass and the accretion rate and formation time at given mass as a function of the position with respect to the centre of the saddle.

The last approximation holds for haloes that assemble half of their mass before $z \sim 2$, since at early times $D \simeq (1 + z)^{-1}$. These typical quantities are known functions of the position-dependent typical values of the excursion set parameters $\sigma_\star(\boldsymbol{r})$, $\alpha_\star(\boldsymbol{r}, \sigma)$, and $D_\star(\boldsymbol{r}, \sigma)$ given by equations (57), (62), and (72), respectively. They generalize the corresponding characteristic quantities obtained without conditioning on the saddle, given by $\sigma_\star = \delta_c$, and by the functions $\alpha_\star(\sigma)$ and $D_\star(\sigma)$ defined in equations (27) and (38).

Taylor expanding equation (57) in the anisotropy gives the first-order angular variation of $M_\star$ at fixed distance r from the saddle

$$\Delta M_\star(\boldsymbol{r}) = -\frac{15}{2}\frac{\delta_c\,\xi_{20}(r)}{|(\mathrm{d}\sigma/\mathrm{d}M)_{M_\star}|}\mathcal{Q}(\hat{\boldsymbol{r}}),\tag{83}$$

where $\xi_{20}(r)$ is the radial part of the shear-height correlation function at finite separation. Since ξ_{20} is positive, this variation is largest when $\boldsymbol{r}$ is parallel to the eigenvector with the smallest eigenvalue. That is, in the direction of positive outflow (with negative $\mathcal{Q} = \hat{r}_i\bar{q}_{ij}\hat{r}_j$), along which a filament will form. Thus, in filaments haloes tend to be more massive, and haloes of large mass are more likely.

The full dependence of the characteristic mass $M_\star$ as a function of the position with respect to the saddle point of the potential is shown in Fig. 12.

Similarly, like equations (63) and (74) for $\alpha_\star$ and $D_\star$, the first-order angular variations of $\dot{M}_\star$ and $z_\star$ are

$$\Delta\dot{M}_\star(\boldsymbol{r}, M) = -\frac{\mathrm{d}\log D}{\mathrm{d}z}\frac{\mathrm{d}M}{\mathrm{d}\log\sigma}\frac{\alpha_\star^2|_{\bar{q}=0}}{\nu_c}$$
$$\times\frac{15}{2}\left[\xi_{20}' - \frac{\sigma - \xi_I'\xi_I}{\sigma^2 - \xi^2}\xi_{20}\right]\mathcal{Q}(\hat{\boldsymbol{r}}),\tag{84}$$

$$\Delta z_\star(\boldsymbol{r}, M) = \left|\frac{\mathrm{d}z}{\mathrm{d}D}\right|\frac{D_\star^2|_{\bar{q}=0}}{\delta_c}\frac{15}{2}\left|\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right|\frac{M}{2}\xi_{20}(r)\mathcal{Q}(\hat{\boldsymbol{r}}).\tag{85}$$

These results confirm that in the direction of the filament, haloes have on average larger mass accretion rates and smaller formation redshifts than haloes of the same mass that form at the same distance from the saddle point, but in the direction perpendicular to it. The space variation becomes larger with growing halo mass and fixed R_S, as shown in Fig. 12, because the correlations become stronger as the difference between the two scales gets smaller. Conversely, for smaller masses haloes have on average smaller accretion rates (like in the unconditional case, see Fig. 3) and later formation times, but also less prominent space variations.

Note that two estimators of delayed mass assembly, $\Delta\dot{M}_\star$ and $\Delta z_\star$ do not rely on the same property of the excursion set trajectory and do not lead to the same physical interpretation. In particular, when extending the implication of delayed mass assembly to galaxies and their induced feedback, one should distinguish between the instantaneous accretion rate, and the integrated half-mass time as

they trace different components of the excursion hence different epochs.

5.3 Expected differences between the isocontours

In order to investigate whether the assembly bias generated by the cosmic web and described in this work is purely an effect due to the local density (itself driven by the presence of the filament), this section studies the difference between the isocontours of the local density field and any other statistics (mass accretion rate for instance). The latter will be shown not to follow exactly the isodensity surfaces, but to intersect each other. This misalignment may only appear if spherical symmetry is broken (all isocontours would otherwise be spherical). However, it also shows that halo properties do not depend only on the local density, indicating that the role of the anisotropy of the nearby filament in the formation of structures goes beyond the simple creation of an anisotropic density field.

The normals to the level surfaces of $\dot{M}_\star(\boldsymbol{r}, M)$, $M_\star(\boldsymbol{r})$, $z_\star(\boldsymbol{r}, M)$, and $\langle\rho\rangle(\boldsymbol{r}) \equiv \bar{\rho}(1 + \langle\delta|\mathcal{S}\rangle)$ scale like the gradients of these functions. First note that any mixed product (or determinant) such as $\nabla\dot{M}_\star \cdot (\nabla M_\star \times \nabla\langle\rho\rangle)$ will be null by symmetry; i.e. all gradients are coplanar. This happens because the present theory focuses on scalar quantities (mediated, in our case, by the excursion set density and slope). In this context, all fields vary as a function of only two variables, r and $\mathcal{Q} = \hat{r}_i\bar{q}_{ij}\hat{r}_j$, hence the gradients of the fields will all lie in the plane of the gradients of r and $\mathcal{Q}$.[9] Ultimately, if one focuses on a given spherically symmetric peak, then $\mathcal{Q}$ vanishes, so all gradients are proportional to each other and radial. Let us now quantify the misalignments between two normals within that plane. In spherical coordinates, the Nabla operator reads

$$\nabla = \left(\frac{\partial}{\partial r}, \frac{1}{r}\frac{\partial}{\partial\theta}, \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right) \equiv \left(\frac{\partial}{\partial r}, \frac{1}{r}\tilde{\nabla}\right), \qquad (86)$$

so that for instance

$$\nabla\dot{M}_\star \propto \left(\frac{\partial\dot{M}_\star}{\partial r}, \frac{1}{r}\frac{\partial\dot{M}_\star}{\partial\mathcal{Q}}\tilde{\nabla}\mathcal{Q}\right),$$

where equation (46) implies that

$$\tilde{\nabla}\mathcal{Q} = \begin{pmatrix} \sin 2\theta\left(\bar{q}_3\cos^2\phi + \bar{q}_2\sin^2\phi - \bar{q}_1\right) \\ \sin\theta(\bar{q}_2 - \bar{q}_3)\sin 2\phi \end{pmatrix}. \qquad (87)$$

Hence, for instance the cross product $\nabla M_\star \times \nabla\dot{M}_\star$ reads

$$\left(\frac{\partial\dot{M}_\star}{\partial r}\frac{\partial M_\star}{\partial\mathcal{Q}} - \frac{\partial\dot{M}_\star}{\partial\mathcal{Q}}\frac{\partial M_\star}{\partial r}\right)\tilde{\nabla}\mathcal{Q}. \qquad (88)$$

It follows that the two normals are not aligned, since the pre-factor in equation (88) does not vanish: the fields are explicit distinct and independent functions of both r and $\mathcal{Q}$. The origin of the misalignment lies in the relative amplitude of the radial and 'polar' derivatives (with respect to $\mathcal{Q}$) of the field. For instance, even at linear order in the anisotropy, since $\Delta\dot{M}_\star$ in equation (84) has a radial dependence in ξ'_{20} as a pre-factor to $\mathcal{Q}$, whereas $M_\star$ has only ξ_{20} as a pre-factor in equation (83), the bracket in equation (88) will involve the Wronskian $\xi'_{20}\partial\xi_{20}/\partial r - \xi_{20}\partial\xi'_{20}/\partial r$ which is non-zero

because ξ_{20} and its derivative with respect to filtering are linearly independent. This misalignment does not hold for $M_\star$ and $\langle\rho\rangle$ at linear order, since $\Delta M_\star$ (equation 83) and $\langle\rho\rangle$ (equation 45) are proportional in this limit. Yet it does arises when accounting for the fact that the contribution to the conditional variance in $M_\star$ also depends additively on $\xi^2(r)$ in equation (57) [with $\xi^2(r)$ given by equation (52) as a function of the finite separation correlation functions $\xi_{\alpha\beta}$ computed in equation (E12) for a given underlying power spectrum]. Indeed, one should keep in mind that the saddle condition not only shifts the mean of the observables but also changes their variances. Since the critical 'star' observables ($M_\star$, $z_\star$, etc.) involve rarity, hence ratio of the shifted means to their variances (e.g. entering equation 60), both impact the corresponding normals. It is therefore a clear specific prediction of conditional excursion set theory relying on upcrossing that the level sets of density, mass density, and accretion rates are distinct.

Physically, the distinct contours could correspond to an excess of bluer or reddened galactic hosts at fixed mass along preferred directions depending on how feedback translate inflow into colour as a function of redshift. Indeed feedback from active galactic nuclei (AGNs), triggered during merger events, regulates gas inflows (Dubois et al. 2016), which in turn impacts star formation: when it is active, at intermediate and low redshift, it may reverse the naive expectation (see Appendix H). This would be in agreement with the recent excess transverse gradients (at fixed mass and density) measured both in cosmological hydrodynamical simulation Horizon-AGN (Dubois et al. 2014) and those observed in spectroscopic (e.g. VIPERS or GAMA, Malavasi et al. 2017; Kraljic et al. 2018) and photometric (e.g. COSMOS, Laigle et al. 2017) surveys: bluer central galaxies at high redshifts when AGN feedback is not efficient and redder central galaxies at lower redshift.

Our predictions are formulated in the initial conditions. However, one should take into account a Zel'dovich boost to get the observable contours of the quantities derived in the paper. Regions that will collapse into a filament are expected to have a convergent Zel'dovich flow in the plane perpendicular to the filament and a diverging flow in the filament's direction. As such, the contours of the different quantities will be advected along with the flow and will become more and more parallel along the filament. This effect is clearly seen in Fig. 13 which shows the contours of both the typical density and the accretion rate[10] (bottom panel) after the Zel'dovich boost (having chosen the amplitude of the boost corresponding to the formation of the filamentary structure). The contours are compressed towards the filament and become more and more parallel. Hence, the stronger the non-linearity, the more parallel the contours. This is consistent with the findings of Kraljic et al. (2018), whose colour and (stellar) mass gradients follow the underlying mean density, when the density is averaged on sufficiently small scales.

6 ASSEMBLY BIAS

The bias of DM haloes (see Desjacques, Jeong & Schmidt 2016, for a recent review) encodes the response of the mass function to variations of the matter density field. In particular, the Lagrangian bias function b_1 describes the linear response to variations of the initial matter density field. For Gaussian initial conditions, the

---

[9] In order to break this degeneracy, one would need to look at the statistics of higher spin quantities. For instance, the angular momentum of the halo would depend on the spin-one coupling $\varepsilon_{ijk}\hat{r}_j\bar{q}_{kl}\hat{r}_l$, with ε_{ijk} the totally antisymmetric tensor (see e.g. Codis, Pichon & Pogosyan 2015), or to consider a barrier that depends on the local shear at $\boldsymbol{r}$ filtered on scale R (e.g. Castorina et al. 2016), like e.g. $\delta_c + \beta\sigma\bar{q}_{ij}(\boldsymbol{r}, R)\bar{q}_{ij}(\boldsymbol{r}, R)$ with some constant β.

[10] Interactive versions can be found online https://cphyc.github.io/research/assembly/with_boost.html and https://cphyc.github.io/research/assembly/no_boost.html.
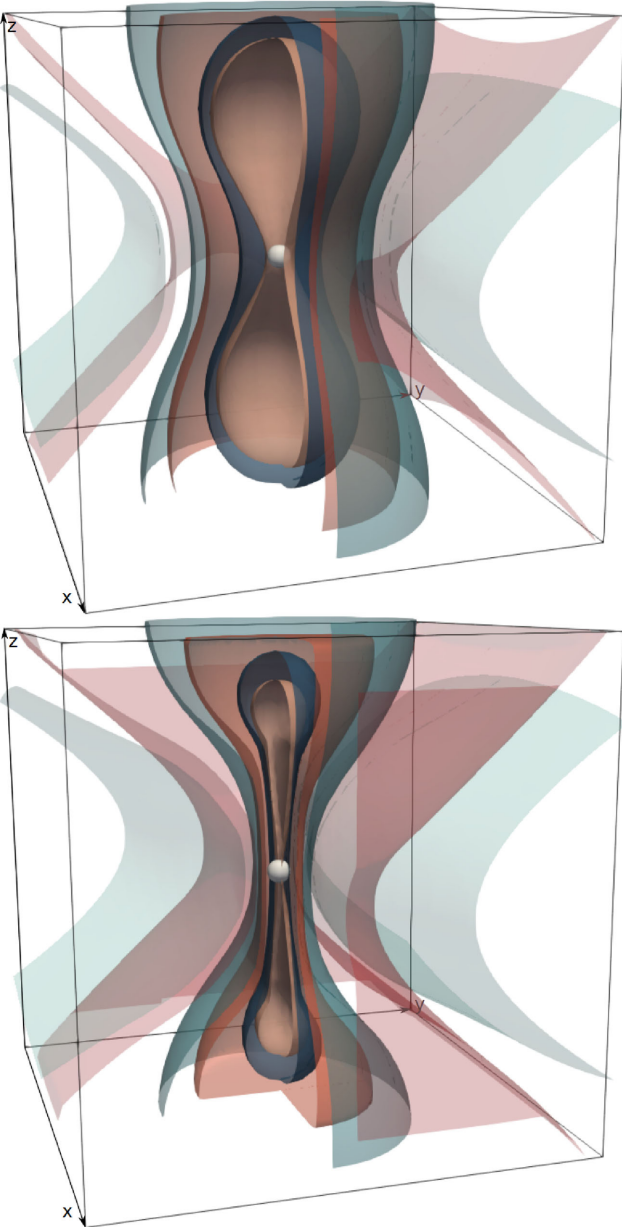
Figure 13. Level surfaces of the typical density $\rho_\star$ (light to dark blue) and of the accretion rate $\alpha_\star$ (light to dark red) with no Zel'dovich boost (upper panel and with a Zel'dovich boost (lower panel). The saddle is represented by a ball. Once boosted, the structure of the filament in the z-direction is clearly seen and the isocontours align one with each other.

correlation of the halo overdensity with an infinite wavelength matter overdensity δ_0 is then (Fry & Gaztanaga 1993),

$$\langle \delta_0 \delta_{\rm h}(\boldsymbol{r}, M) \rangle = \int {\rm d}\boldsymbol{r}_1 \langle \delta_0 \delta_{\rm m}(\boldsymbol{r}_1) \rangle b_1(\boldsymbol{r}, \boldsymbol{r}_1, M), \qquad (89)$$

where formally $b_1(\boldsymbol{r}, \boldsymbol{r}_1, M) \equiv \langle \partial[\delta_{\rm h}(\boldsymbol{r}, M)]/\partial[\delta_{\rm m}(\boldsymbol{r}_1)] \rangle$ is the expectation value of the functional derivative of the local halo over-density with respect to the (unsmoothed) matter density field $\delta_{\rm m}(\boldsymbol{r})$ (Bernardeau, Crocce & Scoccimarro 2008). In the standard setup, because of translational invariance (which does not hold here), it is only a function of the separation $|\boldsymbol{r} - \boldsymbol{r}_1|$.

The dependence of the halo field on the matter density field can be parametrized with a potentially infinite number of variables

constructed in terms of the matter density field, evaluated at the same point. With a simple chain rule applied to the functional derivative, equation (89) can be written as the sum of the cross-correlation of δ_0 with each variable, times the expectation value of the ordinary partial derivative of the halo point process with respect to the same variable. The latter are the so-called bias coefficients, and are mathematically equivalent to ordinary partial derivatives of the mass function with respect to the expectation value of each variable.

The most important of these variables is usually assumed to be the density $\delta(\boldsymbol{r}, R)$ filtered on the mass scale of the haloes, which mediates the response to the variation of an infinite wavelength mode of the density field, the so-called large-scale bias. Because the smoothed density correlates with the $k = 0$ mode of the density field, this returns the peak-background split bias. Its bias coefficient is also equal to (minus) the derivative with respect to $\delta_{\rm c}$.

Excursion sets make the ansatz that the next variable that matters is the slope $\delta'(\boldsymbol{r}, R)$ (Musso, Paranjape & Sheth 2012). In the simplest excursion set models with correlated steps and a constant density threshold, trajectories crossing $\delta_{\rm c}$ with steeper slopes have a lower mean density on larger scales (Zentner 2007). They are thus unavoidably associated with less strongly clustered haloes. This prediction is in agreement with N-body simulations for large-mass haloes, but the trend is known to invert for smaller masses (Sheth & Tormen 2004; Gao et al. 2005; Wechsler et al. 2006; Dalal et al. 2008). Although more sophisticated models are certainly needed in order to account for the dynamics of gravitational collapse, we will see that the presence of a saddle point contributes to explaining this inversion.

None of the concepts outlined above changes in the presence of a saddle point: the bias coefficients are derivatives of ${\rm d}n/{\rm d}M$, that is of the upcrossing probability through equation (75). Because we are interested in the bias of the joint saddle-halo system, we must differentiate the joint probability $f_{\rm up}(\sigma; \boldsymbol{r})p(\mathcal{S})$, rather than just $f_{\rm up}(\sigma; \boldsymbol{r})$, and divide by the same afterwards. Of course, the result picks up a dependence on the position within the frame of the saddle. The relevant uncorrelated variables are $\delta - \langle \delta | \mathcal{S} \rangle$, $\delta' - \langle \delta' | \nu, \mathcal{S} \rangle$, $\nu_{\mathcal{S}}$, $\hat{r}_i g_i = 0$, and $\mathcal{Q} = \hat{r}_i \bar{q}_{ij} \hat{r}_j$. Differentiating equation (49), the bias coefficients of the halo are

$$b_{10}(M; \boldsymbol{r}) \equiv \frac{\partial \log \left[f_{\rm up}(\sigma; \boldsymbol{r}) \right]}{\partial \langle \delta | \mathcal{S} \rangle} = \frac{\delta_{\rm c} - \xi_I \mathcal{S}_I}{\sigma^2 - \xi^2}, \qquad (90)$$

$$b_{01}(M; \boldsymbol{r}) \equiv \frac{\partial \log \left[f_{\rm up}(\sigma; \boldsymbol{r}) \right]}{\partial \langle \delta' | \nu_{\rm c}, \mathcal{S} \rangle} = \frac{1 + {\rm erf}(X_{\mathcal{S}}(\boldsymbol{r})/\sqrt{2})}{2\mu_{\mathcal{S}}(\boldsymbol{r})F(X_{\mathcal{S}}(\boldsymbol{r}))}, \qquad (91)$$

which without saddle reduce to (a linear combination of) those defined by Musso et al. (2012). The coefficients of the saddle are

$$b_{100}^{(\mathcal{S})} \equiv -\frac{\partial}{\partial \delta_{\rm s}} \log p_{\rm G}(\mathcal{S}) = \frac{\nu_{\mathcal{S}}}{\sigma_{\mathcal{S}}}, \qquad (92)$$

$$b_{010}^{(\mathcal{S})} \equiv -\frac{\partial}{\partial (\hat{r}_i g_i)} \log p_{\rm G}(\mathcal{S}) \bigg|_{g_i = 0} = 0, \qquad (93)$$

$$b_{001}^{(\mathcal{S})} \equiv -\frac{\partial}{\partial \mathcal{Q}} \log p_{\rm G}(\mathcal{S}) = \frac{15}{2} \frac{3\mathcal{Q}}{2}. \qquad (94)$$

A constant δ_0 does not correlate with $\bar{q}_{ij}$, since there is no zero mode of the anisotropy. One can see this explicitly by noting that $\xi_{20}(R_0, R_{\mathcal{S}}, r) \to 0$ as $R_0 \to \infty$. The only coefficients that survive in the cross-correlation with δ_0 are thus b_{10}, b_{01} and $b_{100}^{(\mathcal{S})}$, so that equation (89) becomes

$$\langle \delta_0 \delta_h(\boldsymbol{r}, M) \rangle = b_{100}^{(\mathcal{S})} \langle \delta_0 \delta_{\rm s} \rangle + b_{10} {\rm Cov}\,(\delta_0, \delta | \mathcal{S})$$
$$+ b_{01} {\rm Cov}\,\left(\delta_0, \delta' | \nu_{\rm c}, \mathcal{S}\right). \qquad (95)$$

Similarly, in this limit δ_0 does not correlate with g_i either, while $\langle \delta_0 \delta \rangle$ becomes independent of R. Thus, $\langle \delta_0 \delta \rangle \simeq \langle \delta_0 \delta_s \rangle$ and $\langle \delta_0 \delta' \rangle \simeq 0$. Hence,

$$\frac{\langle \delta_0 \delta_h \rangle}{\langle \delta_0 \nu_{\mathcal{S}} \rangle} \simeq \nu_{\mathcal{S}} + \frac{\delta_c - \xi_I \mathcal{S}_I}{\sigma^2 - \xi^2}(\sigma_s - \xi_{00})$$
$$- b_{01}\left[\xi'_{00} + \frac{\sigma - \xi'_I \xi_I}{\sigma^2 - \xi^2}(\sigma_s - \xi_{00}) \right]. \quad (96)$$

Setting $\nu_{\mathcal{S}} = \xi_{\alpha\beta} = \xi'_{\alpha\beta} = 0$ recovers Musso et al.'s (2012) results.

The anisotropic effect of the saddle is easier to understand looking at the sign of the terms in the round and square brackets, corresponding to Cov $(\delta_0, \delta | \mathcal{S})$ and $-\text{Cov}(\delta_0, \delta' | \nu_c, \mathcal{S})$ respectively. One can check that for $R = 1$ Mpc h^{-1} and $R_{\mathcal{S}} = 10~$Mpc h^{-1} both terms are negative near $r = 0$, but become positive at $r \simeq 0.75 R_{\mathcal{S}}$. This separation marks an inversion of the trend of the bias with $\nu_{c,\mathcal{S}}$, the parameter measuring how rare haloes are given the saddle environment. Far from the saddle, haloes with higher $\nu_{c,\mathcal{S}}$ are *more biased*, which recovers the standard behaviour since $\nu_{c,\mathcal{S}} \to \nu_c$ as $r \to \infty$. However, as $r/R_{\mathcal{S}} \lesssim 0.75$, the trend inverts and haloes with higher $\nu_{c,\mathcal{S}}$ become *less biased*. Therefore, one expects that at fixed mass and distance from the saddle-point haloes in the direction of the filament are less biased far from the saddle, but become more biased near the saddle point. The upper panel of Fig. 14, displaying the exact result of equation (96), confirms these trends and their inversion at $r \simeq 0.75 R_{\mathcal{S}}$. The height of the curves at $r = 0$ depends on the chosen value for $\nu_{\mathcal{S}}$, but the inversion at $r \simeq 0.75 R_{\mathcal{S}}$ and the behaviour at large r do not. Fig. 14 also shows that a saddle point of the potential need not be a saddle point of the bias (in the present case, it is in fact a maximum).

The inversion can be interpreted in terms of excursion sets. Near the saddle, fixing $\nu_{\mathcal{S}}$ at $\mathbf{r} = 0$ puts a constraint on the trajectories at $\mathbf{r}$ that becomes more and more stringent as the separation gets small. At $r = 0$, the value of the trajectory at $R_{\mathcal{S}}$ is completely fixed. Therefore, trajectories constrained to have the same height at both $R_{\mathcal{S}}$ and R, but lower $\langle \delta | \mathcal{S} \rangle$ at R, will tend to drift towards lower values between $R_{\mathcal{S}}$ and R, and thus towards higher values for $R_0 \gg R_{\mathcal{S}}$. This effect vanishes far enough from the saddle point, since the constraint on the density at $R_{\mathcal{S}}$ becomes looser as the conditional variance grows. Hence, trajectories with lower $\langle \delta | \mathcal{S} \rangle$ at R will remain lower all the way to R_0. Note however that interpreting these trends in terms of clustering is not straightforward, because the variations happen on a scale $R_{\mathcal{S}} \ll R_0$ (they are thus an explicit source of scale-dependent bias). The most appropriate way to understand the variations of clustering strength is looking at the position dependence of dn/dM, which is predicted explicitly through $f_{\text{up}}(\sigma; \mathbf{r})$ in equation (49).

When one bins haloes also by mass and accretion rate, the bias is given by the response of the mass function at fixed accretion rate. That is, to get the bias coefficients one should now differentiate the joint probability $f_{\text{up}}(\sigma, \alpha; \mathbf{r}) p_G(\mathcal{S})$ with respect to mean values of the different variables, with $f_{\text{up}}(\sigma, \alpha; \mathbf{r})$ given by equation (59). The only bias coefficient that changes is b_{01}, the derivative with respect to $\langle \delta' | \nu_c, \mathcal{S} \rangle$, which becomes

$$b_{01}(M, \dot{M}, \mathbf{r}) \equiv \frac{\partial \log \left[f_{\text{up}}(\sigma, \alpha; \mathbf{r}) \right]}{\partial \langle \delta' | \nu_c, \mathcal{S} \rangle} = \frac{\nu_c/\alpha - \mu_{\mathcal{S}}(\mathbf{r})}{\text{Var}(\delta' | \nu_c, \mathcal{S})}, \quad (97)$$

with α defined by equation (20). Inserting this expression in equation (96), returns the predicted large-scale bias at fixed accretion rate. Notice that in this simple model, the coefficient multiplying the $1/\alpha$ term is purely radial. The asymptotic behaviour of the bias at small accretion rates will then always be divergent and isotropic,
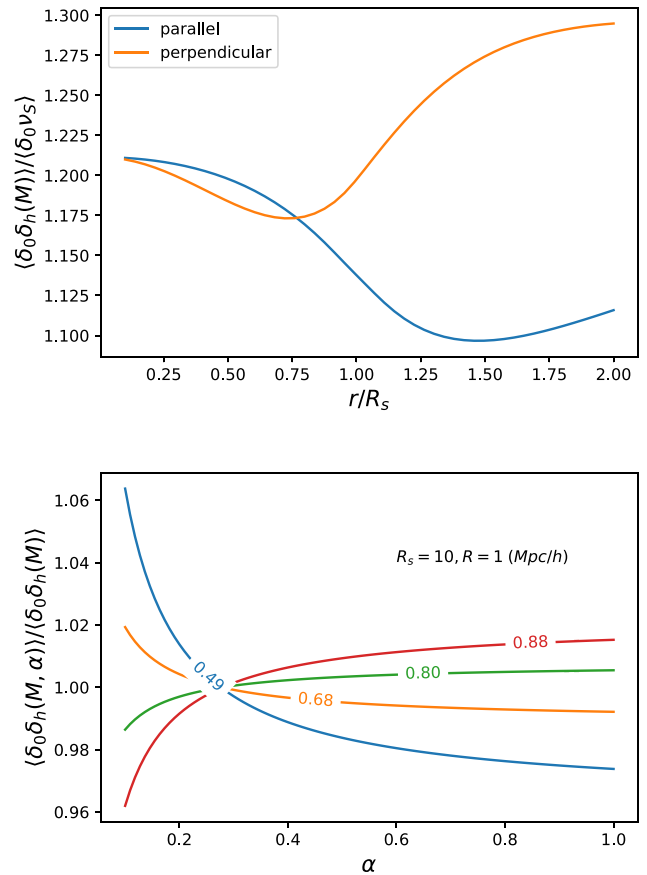
Figure 14. Upper panel: large-scale Lagrangian bias as a function of the distance from the saddle point, along the filament and perpendicularly to it, for haloes of mass $M = 2.0 \times 10^{11}$ M$_\odot$ h^{-1} ($R = 1$ Mpc h^{-1}). Haloes in the perpendicular direction are less biased at small separation, but the trend inverts at $r/R_{\mathcal{S}} \simeq 0.75$. Lower panel: bias as a function of accretion rate, for different values of the separation $r/R_{\mathcal{S}}$ in the direction of the filament. For haloes closer to the centre, bias decreases with accretion rate, but the trend inverts at $r/R_{\mathcal{S}} \simeq 0.75$. In the perpendicular direction, the effect is 30 per cent smaller, but the relative amplitudes and the inversion point do not change appreciably. As discussed in the main text, both inversions depend on the fact that $\delta - \langle \delta | \mathcal{S} \rangle$ and δ_0 correlate at large distance from the saddle, but they anticorrelate at small separation.

with a sign depending on that of the square bracket in equation (96). If this term is positive, the bias decreases as α gets smaller, and vice versa. Clearly, the value of α for which the divergent behaviour becomes dominant depends on the size of all the other terms, and is therefore anisotropic.

As one can see from Fig. 14, the sign of the small-α divergence depends on the distance from the saddle point. It is negative for $r \gtrsim 0.75 R_{\mathcal{S}}$, but it reverses closer to the centre. This effect is again a consequence of the constraint on the excursion set trajectories at $R_{\mathcal{S}}$. Trajectories with steeper slopes at R will sink to lower values between $R_{\mathcal{S}}$ and R, then turn upwards to pass through $\delta(R_{\mathcal{S}})$, and reach higher values for $R_0 \gg R_{\mathcal{S}}$. The haloes they are associated with are thus *more biased*. This trend is represented in Fig. 15. This inversion effect is lost as the separation increases, and the constraint on the density at $R_{\mathcal{S}}$ becomes loose, and trajectories that reach R with steeper slopes are likely to have low (or even negative) values at very large scales. These haloes are thus *less biased*, or even antibiased.
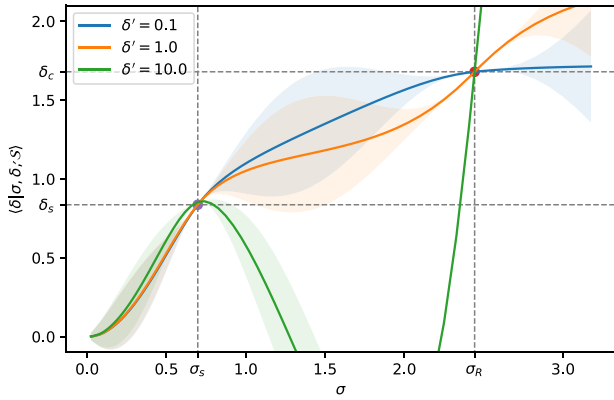
Figure 15. Plot of the mean of density given the saddle point, the upcrossing condition and the slope at *R* for different slopes. The saddle point was defined using the values of Table D1. The details of the calculation are provided in Appendix B. For steep slopes (small accretion rate), the mean of the density overshoots at small σ, resulting in a larger bias.

It follows that the bias of haloes far from structures grows with accretion rate (the usual behaviour expected from excursion sets), while the trend inverts for haloes near the centre of the filament. Because typical mass of haloes also depends on the position along the filament, with haloes towards the nodes being more massive, the different curves of Fig. 14 correlate with haloes of different mass. This effect explains why low-mass haloes with small accretion rate (or early formation time, or high concentration) are more biased, when measuring halo bias as a function of mass and accretion rate (or formation time or concentration, which strictly correlate with accretion rate), without knowledge of the position in the cosmic web. Conversely, the high-mass ones are less biased (Sheth & Tormen 2004; Gao et al. 2005; Wechsler et al. 2006; Dalal et al. 2008; Faltenbacher & White 2010; Paranjape & Padmanabhan 2017). It is also intriguing to compare this result with the measurements by Lazeyras et al. (2017, , namely their fig. 7) which show the same trends (although their masses are not small enough to clearly see the inversion).

Note in closing that the conditional bias theory presented here does not capture changes in accretion rate and formation time presented in Sections 4.3 and 4.4.

7 CONCLUSION AND DISCUSSION

7.1 Conclusion

With the advent of modern surveys, assembly bias has become the focus of renewed interest as a process which could explain some of the diversity of galactic morphology and clustering at fixed mass. It is also investigated as a mean to mitigate intrinsic alignments in weak-lensing survey such as *Euclid* or LSST. Both observations and simulations have hinted that the large-scale anisotropy of the cosmic web could be responsible for stalling and quenching. This paper investigated this aspect in Lagrangian space within the framework of excursion set theory. As a measure of infall, we computed quantities related to the slope of the contrast conditioned to the relative position of the collapsing halo with respect to a critical point of the large-scale field. We focused here on mass accretion rate and half-mass redshift and found that their expectation vary with the orientation and distance from saddle points, demonstrating that assembly bias

is indeed influenced by the geometry of the tides imposed by the cosmic web.

More specifically, we derived the Press–Schechter typical mass, typical accretion rate, and formation time of dark haloes in the vicinity of cosmic saddles by means of an extension of excursion set theory accounting for the effect of their large-scale tides. Our principal findings are the following: we have computed the (i) *up-crossing PDF* for halo mass, accretion rate, and formation time; they are given by equations (14), (23), and (32), and their constrained-by-saddles counterparts equations (49), (61), and (68). These PDFs allowed us to identify the (ii) *typical halo mass*, and *typical accretion rate and formation time at given mass* as functions of the position within the frame of the saddle via equations (83)–(85). All quantities are expressed as a function of the geometry of the saddle for an arbitrary cosmology encoded in the underlying power spectrum via the correlations $\xi_{\alpha\beta}$ and $\xi'_{\alpha\beta}$ given by equations (E11) and (E12). In turn, this has allowed us to compute and explain the corresponding (iii) *distinct gradients* for the three typical quantities and for the local mean density (Section 5.3). The misalignment of the gradients, defined as the normals to the their isosurfaces, arises because the saddle condition is anisotropic and because it does not only shift the local mean density and the mean density profile (the excursion set slope) but also their variances, affecting different observables in different way. Finally, we have presented (iv) *an extension of classical large-scale bias theory* to account for the saddle (Section 6).

Our simple conditional excursion set model subject to filamentary tides makes intuitive predictions in agreement with the trends found in *N*-body simulations: haloes in filaments are less massive than haloes in nodes, and at equal mass they have earlier formation times and smaller accretion rates today. The same hierarchy exists for haloes in walls with respect to filaments. For the configuration we examined, the effect is stronger as one moves perpendicularly to the filament. The typical mass changes by a factor of 5 along the filament, and by two orders of magnitude perpendicularly. The relative variation of accretion rates and formation times is of about 5–10 per cent along the filament, and of about 20–30 per cent in the perpendicular direction, for haloes of 10^{11} M$_\odot$ h^{-1}. Furthermore, our model predicts that at fixed halo mass, the trend of the large-scale bias with accretion rate depends on the distance from the centre of the filament. Far from the centre, the large-scale bias grows with accretion rate (which is the naive expectation from excursion sets), while near the centre the trend inverts and haloes with smaller accretion rates become more biased. Since haloes near the centre are also on average less massive, this effect should contribute to explaining why the trend of bias with accretion rate (or formation time) inverts at masses much smaller than the typical mass.

These findings conflict with the simplistic assumption that the properties of galaxies of a given mass are uniquely determined by the density of the environment. The presence of distinct space gradients for the different typical quantities is also part and parcel of the conditional excursion set theory, simply because the statistics of the excursion set proxies for halo mass, accretion rate, and formation time (the first-crossing scale and slope, and the height at the scale corresponding to $M/2$) are different functions of the position with respect to the saddle point. They have thus different level surfaces. At the technical level, the contours depend on the presence of the conditional variance of $\delta(\mathbf{r})$, besides its conditional mean, and of the correlation functions of $\delta'(\mathbf{r})$. At finite separation, the traceless shear of the large-scale environment modifies in an anisotropic way the statistics of the local mean density $\delta(\mathbf{r})$ (and of its derivative $\delta'(\mathbf{r})$ with respect to scale). The variations are modulated by $\mathcal{Q} = \hat{r}_i \bar{q}_{ij} \hat{r}_j$,
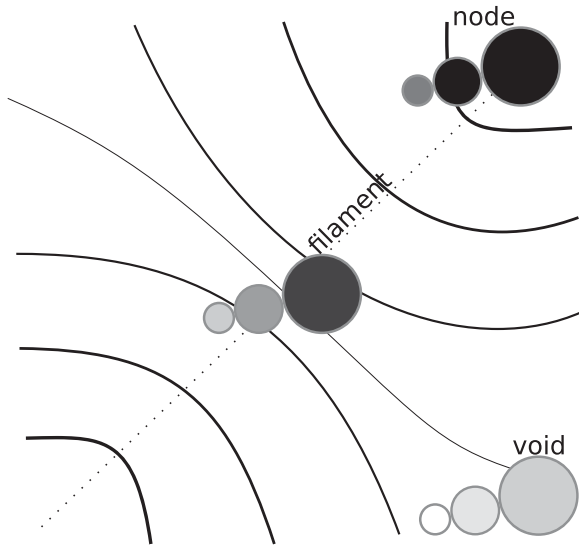
Figure 16. Scheme of the intensity of the accretion rate at different locations near a filament-type saddle for different final halo masses. The darkness of the colour encodes the intensity of the accretion rate (darker is more accretion). At fixed mass, the accretion rate increases from voids to saddle points and from saddle points to nodes (along dotted line which marks the filament's direction). At a given location, the accretion rate increases with mass.

i.e. the relative orientation of the separation vector in the frame set by the tidal tensor of the saddle. This angular modulation enters different quantities with different radial weights, which results in different angular variations of the local statistics of density, mass, and accretion rate/formation time. It provides a supplementary vector space, $\tilde{\nabla}\mathcal{Q}$, beyond the radial direction over which to project the gradients, whose statistical weight depend on each specific observable. These quantities have thus different isosurfaces from each other and from the local mean density, a genuine signature of the traceless part of the tidal tensor. The qualitative differences in terms of mass accretion rate and galactic colour is sketched in Fig. 16.

7.2 Discussion and perspectives

In contrast to the findings of Alonso et al. (2015), Tramonte et al. (2017), and von Braun-Bates et al. (2017), we focused our attention on variations of *mass accretion rates* with respect to the cosmic web rather than mass functions. We have found that, even in a very simple model like excursion sets, halo properties are indeed affected by the anisotropic tides of the environment (involving the traceless part of the tidal tensor), and not just by its density (involving the trace of the tidal tensor). This effect cannot be explained by a simple rescaling of the local mean density (the average density in a sphere of radius of the order of the Lagrangian radius, centred around the halo). Our predictions are in qualitative agreement with the observational results of Kraljic et al. (2018), who detect a misalignment between the isocontours of mass, secondary halo property (type/colour in their case), and local mean density averaged on sufficiently large scales. This misalignment tends to disappear as the scale of the smoothing becomes small, and the signal is increasingly driven by the density alone: this can be interpreted as a consequence of the dynamical stretching of all contours as the filament forms.

Although the excursion set approach is rather crude, and additional constraints (e.g. peaks) would be needed to pinpoint the exact location of halo formation in the initial conditions, we argued that

the effect we are investigating does not strongly depend on the presence of these additional constraints. The underlying reason is that the extra constraints usually involve vector or tensor quantities evaluated *at the same location r as the excursion set sphere*, which do not directly correlate with the scalars considered here (they only do so through their correlation with the saddle point). They may add polynomial corrections to the conditional distributions, but will not strongly affect the exponential cut-offs on which we built our analysis. Our formalism may thus not predict exactly whether a halo will form (hence, the mass function), but it can soundly describe the secondary properties and the assembly bias of haloes that actually form. A more careful treatment would change our results only at the quantitative level. For this reason, we chose to prefer the simplicity of the simple excursion set approach. Furthermore, in order to describe the cosmic web, we focused on saddle points of the initial gravitational potential, rather than of the density field, as these are more suitable to trace the *dynamical* impact of filamentary structures in connection to the spherical collapse model.

The present Lagrangian formalism only aims at describing the behaviour of the central galaxy: it cannot claim to capture the strongly non-linear process of dynamical friction of subclumps within dark haloes, nor strong deviations from spherical collapse. We refer to Hahn et al. (2009) which captures the effect on satellite galaxies, and to Ludlow et al. (2014), Castorina et al. (2016), and Borzyszkowski et al. (2016) which study the effect of the local shear on haloes forming in filamentary structures. Incorporating these effects would require adopting a threshold for collapse that depends on the local shear, as discussed in the Introduction. Such a barrier would not pose a conceptual problem to our treatment;[11] technically, however it requires two extra integrations (over the amplitude of the local shear and its derivative with respect to scale), and cannot be done analytically. The shear-dependent part of the critical density (and its derivative) would correlate with the shear of the saddle at $r = 0$, and introduce an additional anisotropic effect on top of the change of mean values and variances of density and slope we accounted for. Evaluating this effect will be the topic of future investigation.

Our analysis demonstrated that the large-scale tidal field alone can induce specific accretion gradients, distinct from mass and density ones. One would now like to translate those distinct DM gradients into colour and specific star formation rate (SFR) gradients. At high redshift, the stronger the accretion, the bluer the central galaxy. Conversely at low redshift, one can expect that the stronger the accretion, the stronger the AGN feedback, the stronger the quenching of the central. Should this scaling hold true, the net effect in terms of gradients would be that colour gradients differ from mass and density ones. The transition between these two regimes (and in general, the inclusion of baryonic effects) is beyond the scope of this paper, but see Appendix H for a brief discussion.

Beyond the DM-driven processes described in this paper, different explanations have been recently put forward to explain filamentary colour gradients. On the one hand, it has been argued (Aragon-Calvo, Neyrinck & Silk 2016) that the large-scale turbulent flow within filaments may explain the environment dependence in observed physical properties. Conversely, the vorticity of gas inflow within filaments (Laigle et al. 2015) may be prevalent in feeding galactic discs coherently (Pichon et al. 2011; Stewart et al. 2011). Both processes will have distinct signatures in terms of the efficiency and stochasticity of star formation. A mixture of both may

---

[11] The details of the impact on the present derivation are given in Appendix G.

in fact be taking place, given that the kinematic of the large-scale flow is neither strictly coherent nor fully turbulent. Yet, even if ram-pressure stripping in filaments operate as efficiently as in clusters, it will remain that the anisotropy of the tides will also impact the consistency of angular momentum advection, which is deemed important at least for early-type galaxies. The amplitude of thermodynamical processes depends on the equation of state of the gas and on the amplitude of feedback which are not fully calibrated today. Recall that shock heating, AGN and stellar feedback are driven by cold gas infall, which in turn is set by gravity (as the dominant dynamical force). Since gravity has a direct effect through its tides, unless one can convincingly argue that its direct impact is negligible on galactic scales, it should be taken into account.

Codis et al. (2015), following a formally related route, investigated the orientation of the spin of dark haloes in relation to their position with respect to the saddle points of the (density) cosmic web (see also Wang & Kang 2018, for a slightly different approach). Together with their predictions on spin orientation, this work could be extended to model galaxy colours based on both spin and mass accretion. It could also guide models aiming at mitigating the effect of intrinsic alignments (Joachimi et al. 2011) impacting weak-lensing studies, while relying on colour gradients. More generally, galactic evolution as captured by semi-analytical models will undoubtedly gain from a joint description of involving both mass and spin acquisition as relevant dynamical ingredients. Indeed, it has been recently shown in hydrodynamical simulation (e.g. Zavala et al. 2016) that the assembly of the inner DM halo and its history of specific angular momentum loss is correlated to the morphology of galaxies today. One should attempt to explain the observed diversity at a given mass driven by anisotropic large-scale tides, which will impact gas inflow towards galaxies, hence their properties. An improved model for galaxy properties should eventually explicitly integrate the geometry of the large environment (following, e.g. Hanami 2001) and quantify the impact of its anisotropy on galactic mass assembly history.

Thanks to significant observational, numerical, and theoretical advances, the subtle connection between the cosmic web and galactic evolution is on the verge of being understood.

ACKNOWLEDGEMENTS

REFERENCES

Alonso D., Eardley E., Peacock J. A., 2015, MNRAS, 447, 2683
Alpaslan M. et al., 2016, MNRAS, 457, 2287
Aragon-Calvo M. A., Neyrinck M. C., Silk J., 2016, preprint (arXiv:1607.07881)
Bernardeau F., Crocce M., Scoccimarro R., 2008, Phys. Rev. D, 78, 103521
Bond J. R., Myers S. T., 1996, ApJS, 103, 1
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Borzyszkowski M., Porciani C., Romano-Diaz E., Garaldi E., 2016, MNRAS, 469, 594
Castorina E., Paranjape A., Hahn O., Sheth R. K., 2016, preprint (arXiv:1611.0361)
Chen Y.-C. et al., 2017, MNRAS, 466, 1880
Codis S., Pichon C., Pogosyan D., 2015, MNRAS, 452, 3369
Corasaniti P. S., Achitouv I., 2011, Phys. Rev. D, 84, 023009
Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12
Del Popolo A., Ercan E. N., Gambera M., 2001, Balt. Astron., 10, 629
Desjacques V., Jeong D., Schmidt F., 2016, preprint (arXiv:1611.09787)
Doroshkevich A. G., 1970, Astrophysics, 6, 320
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2010, MNRAS, 409, 985
Dubois Y., Pichon C., Devriendt J., Silk J., Haehnelt M., Kimm T., Slyz A., 2013, MNRAS, 428, 2885
Dubois Y. et al., 2014, MNRAS, 444, 1453
Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, MNRAS, 463, 3948
Efstathiou G., Frenk C. S., White S. D. M., Davis M., 1988, MNRAS, 235, 715
Faltenbacher A., White S. D. M., 2010, ApJ, 708, 469
Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447
Gao L., Springel V., White S. D. M., 2005, MNRAS, 363, L66
Gradshteyn I. S., Ryzhik I. M., 2007, Table of Integrals, Series, and Products, Seventh edn. Elsevier/Academic Press, Amsterdam
Hahn O., Porciani C., Dekel A., Carollo C. M., 2009, MNRAS, 398, 1742
Hanami H., 2001, MNRAS, 327, 721
Joachimi B., Mandelbaum R., Abdalla F. B., Bridle S. L., 2011, A&A, 527, A26
Kaiser N., 1984, ApJ, 284, L9
Kauffmann G., Li C., Zhang W., Weinmann S., 2013, MNRAS, 430, 1447
Kawinwanichakij L. et al., 2016, ApJ, 817, 9
Kraljic K. et al., 2018, MNRAS, 474, 547
Lacey C. G., Cole S., 1993, MNRAS, 262, 627
Laigle C. et al., 2015, MNRAS, 446, 2744
Laigle C. et al., 2017, MNRAS, 474, 5437
Lazeyras T., Musso M., Schmidt F., 2017, J. Cosmol. Astropart. Phys., 3, 059
Ludlow A. D., Borzyszkowski M., Porciani C., 2014, MNRAS, 445, 4110
Maggiore M., Riotto A., 2010, ApJ, 711, 907
Malavasi N. et al., 2017, MNRAS, 465, 3817
Martínez H. J., Muriel H., Coenda V., 2016, MNRAS, 455, 127
Musso M., Sheth R. K., 2012, MNRAS, 423, L102
Musso M., Sheth R. K., 2014a, MNRAS, 438, 2683
Musso M., Sheth R. K., 2014b, MNRAS, 443, 1601
Musso M., Sheth R. K., 2014c, MNRAS, 443, 1601
Musso M., Paranjape A., Sheth R. K., 2012, MNRAS, 427, 3145
Oemler A., Jr, 1974, ApJ, 194, 1
Paranjape A., Padmanabhan N., 2017, MNRAS, 468, 2984
Paranjape A., Hahn O., Sheth R. K., 2017, preprint (arXiv:1706.09906)
Pichon C., Pogosyan D., Kimm T., Slyz A., Devriendt J., Dubois Y., 2011, MNRAS, 1739
Pogosyan D., Bond J. R., Kofman L., Wadsley J., 1998, in Colombi S., Mellier Y., Raban B., eds, Wide Field Surveys in Cosmology. Editions Frontieres, Dreux, p. 61
Poudel A., Heinämäki P., Tempel E., Einasto M., Lietzen H., Nurmi P., 2017, A&A, 597, A86
Press W. H., Schechter P., 1974, ApJ, 187, 425
Redner S., 2001, A Guide to First-Passage Processes. Cambridge University Press, Cambridge
Shen J., Abel T., Mo H. J., Sheth R. K., 2006, ApJ, 645, 783
Sheth R. K., Tormen G., 2004, MNRAS, 350, 1385
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Sheth R. K., Chan K. C., Scoccimarro R., 2013, Phys. Rev. D, 87, 083002
Sousbie T., Pichon C., Colombi S., Pogosyan D., 2008, MNRAS, 383, 1655
Stewart K. R., Kaufmann T., Bullock J. S., Barton E. J., Maller A. H., Diemand J., Wadsley J., 2011, ApJ, 738, 39
Tramonte D., Rubino-Martin J. A., Betancort-Rijo J., Dalla Vecchia C., 2017, MNRAS, 467, 3424

von Braun-Bates F., Winther H. A., Alonso D., Devriendt J., 2017, J. Cosmol. Astropart. Phys., 3, 012

Wang P., Kang X., 2018, MNRAS, 473, 1562

Wang J. et al., 2011, MNRAS, 413, 1373

Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, ApJ, 652, 71

Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, MNRAS, 366, 2

Yan H., Fan Z., White S. D. M., 2013, MNRAS, 430, 3432

Zavala J. et al., 2016, MNRAS, 460, 4466

Zel'dovich Y. B., 1970, A&A, 5, 84

Zentner A. R., 2007, Int. J. Mod. Phys. D, 16, 763

APPENDIX A: DEFINITIONS AND NOTATIONS

Table A1 presents all the definitions introduced in the paper. Table 1 gives also the motivation behind the choice of variables. The following conventions is used throughout:

(i) unless stated otherwise, all the quantities evaluated at (halo) scale R have their dependence on R omitted (e.g. $\sigma = \sigma(R)$);

(ii) the quantities that have a radial dependence are evaluated at a distance r when the radius is omitted. Sometimes, the full form is used to emphasize the dependence on this variable;

(iii) unless stated otherwise, the quantities are evaluated at $z = 0$ and $D(z) = 1$ (e.g. $\delta_c = 1.686$);

(iv) a prime denotes a derivative with respect to σ of the excursion set (e.g. $\delta' = \mathrm{d}\delta/\mathrm{d}\sigma$);

(v) variables carrying a hat have unit norm (e.g. $|\hat{\boldsymbol{r}}| = 1$), matrices carrying an overbar are traceless (e.g. $\mathrm{tr}(\bar{q}_{ij}) = 0$);

(vi) the *Einstein*'s convention on repeated indexes is used throughout, except in Appendix F2.

APPENDIX B: VALIDATION WITH GRFS

Let us first compare the prediction of Section 4 to statistics derived from realization of GRF, while imposing a saddle-point condition. The values used at the saddle point are reported in Table D1. We further imposed the saddle point's eigenframe to coincide with the x, y, z frame, which in practice has been done by imposing $\bar{q}_{ij}$ to be diagonal. We have used two different methods to validate our results, by generating random density cubes (Appendix B1) and by computing the statistics of a constrained field (Appendix B2).

B1 Validation for $\sigma_\star$

The procedure is the following: (i) 4000 cubes of size $(128)^3$ and width $L_{\mathrm{box}} = 200$ Mpc h^{-1} centred on a saddle point were generated following a ΛCDM power spectrum; (ii) each cube has been smoothed using a Top-Hat filter at 25 different scales ranging from 0.5 to 20 Mpc h^{-1}; (iii) for each point of each cube, the *first-crossing* point σ_{first} was computed; and (iv) the 4000 realizations were stacked to get a distribution of σ_{first} and to compute the median value. It is worth noting that the value of $\Gamma(\sigma(R))$ in the GRF is not the same as in theory. This is a well-known effect (see e.g. Sousbie et al. 2008) that arise on small scales due to the finite resolution of the grid and on large scale because of the finite size of the box. The Γ measured in a GRF is correct at scales verifying $\Delta L \lesssim R \ll L_{\mathrm{box}}$, where ΔL is the grid spacing. In our case, the largest smoothing scale is 20 Mpc $h^{-1} = L_{\mathrm{box}}/10$. However, the smallest scale is comparable to the grid spacing. To attenuate the effect of finite resolution, we have measured $\Gamma(\sigma(R))$ in the GRF and used its value to compute the theoretical cumulative distribution function (CDF). The results of the measured CDF F_{first} and

Table A1. Summary of the variables used throughout the paper.

Variable	Definition	Comment
$\bar{\rho}_{\mathrm{m}}$	$(2.8 \times 10^{11}\ h^2 \mathrm{M}_\odot/\mathrm{Mpc}^3) \times \Omega_{\mathrm{M}}$	Uniform matter background density
$R, M, M_\star$	$M = 4/3\pi R^3 \bar{\rho}_{\mathrm{m}}$	Smoothing scale, mass, and typical mass
δ_{m}	$(\rho_{\mathrm{m}} - \bar{\rho}_{\mathrm{m}})/\bar{\rho}_{\mathrm{m}}$	Linear matter overdensity
$W(x)$	$3j_1(x)/x$	Real-space Top-Hat filter (Fourier representation)
δ	$\int \dfrac{\mathrm{d}^3 k}{(2\pi)^3} \delta_{\mathrm{m}}(\boldsymbol{k}) W(kR) \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}$	Linear matter overdensity smoothed at scale R, position $\boldsymbol{r}$
σ^2	$\mathrm{Var}(\delta)$	Variance of the overdensity at scale R
ν	δ/σ	Rescaled overdensity
$\delta_{\mathrm{c}}, \nu_{\mathrm{c}}$	$1.68, \delta_{\mathrm{c}}/\sigma$	Critical overdensity
δ', ν'	$\mathrm{d}\delta/\mathrm{d}\sigma, \mathrm{d}\nu/\mathrm{d}\sigma$	Slope of the E.S. trajectories
Γ^{-2}	$\mathrm{Var}(\delta') - 1 = \langle(\sigma\nu')^2\rangle = \mathrm{Var}\left(\delta'\vert\nu\right)$	Conditional variance of δ' at fixed ν
$R_{\mathcal{S}}, \sigma_{\mathcal{S}}$	$\sigma_{\mathcal{S}} = \sigma(R_{\mathcal{S}})$	Smoothing scale used at the saddle point
$R_\star^2$	$(42) \int \mathrm{d}k \dfrac{P(k)}{2\pi^2} \dfrac{W^2(kR_{\mathcal{S}})}{\sigma_{\mathcal{S}}^2}$.	Characteristic length-scale of the saddle (squared)
$g_i, q_{ij}, \nu_{\mathcal{S}}$	(41) and (43)	Mean acceleration, tidal tensor, and overdensity at saddle (see Table D1 for their value)
$\bar{q}_{ij}, \mathcal{Q}$	$\bar{q}_{ij} = q_{ij} - \nu_{\mathcal{S}}\delta_{ij}/3, \hat{r}_i \bar{q}_{ij} \hat{r}_j$	Traceless tidal tensor and anisotropy ellipsoidal-hyperbolic coordinate
$\xi_{\alpha\beta}, \xi'_{\alpha\beta}$	(E11) and (E12); $\xi'_{\alpha\beta} = \mathrm{d}\xi_{\alpha\beta}/\mathrm{d}\sigma$	Two-point correlation functions at separation r and scales $R, R_{\mathcal{S}}$
$\alpha, \alpha_\star$	$\nu_{\mathrm{c}}/[\sigma(\nu' - \nu'_{\mathrm{c}})]$; (27) and (62)	Accretion rate and typical accretion rate
$R_{1/2}, \sigma_{1/2}$	$R/2^{1/3}, \sigma(R_{1/2})$	Half-mass radius and variance
$\delta_{1/2}, \nu_{1/2}$	$\delta(\sigma_{1/2}), \delta_{1/2}/\sigma_{1/2}$	Overdensity at half-mass
$D_{\mathrm{f}}, D_\star$	$\delta_{\mathrm{c}}/\delta_{1/2}$; (38) and (72)	Formation time and typical formation time
ν_{f}	$\delta_c/(\sigma_{1/2}D_{\mathrm{f}})$	Density threshold at formation time
ω, ω'	(E14) and (E15); $\omega' = \mathrm{d}\omega/\mathrm{d}\sigma$	Zero-distance correlation functions between scales R and $R_{1/2}$
Ω, Ω'	(F27) and (F32); $\Omega' = \mathrm{d}\Omega/\mathrm{d}\sigma$	Zero-distance conditional covariance between scales R and $R_{1/2}$ given the saddle point
δ_0	$\delta(R_0 \gg R)$	Large-scale overdensity
δ_{h}		Local halo number density contrast

Figure B1. Theoretical CDF of σ at upcrossing (bold lines) and numerical CDF (steps) at first-crossing at four locations around the saddle point (the distances are in Mpc h^{-1} in the x (void) and z (filament) directions). The CDF have been normalized to share the same 50˜ per cent quantile (the horizontal line). See the text for the details of the normalization.
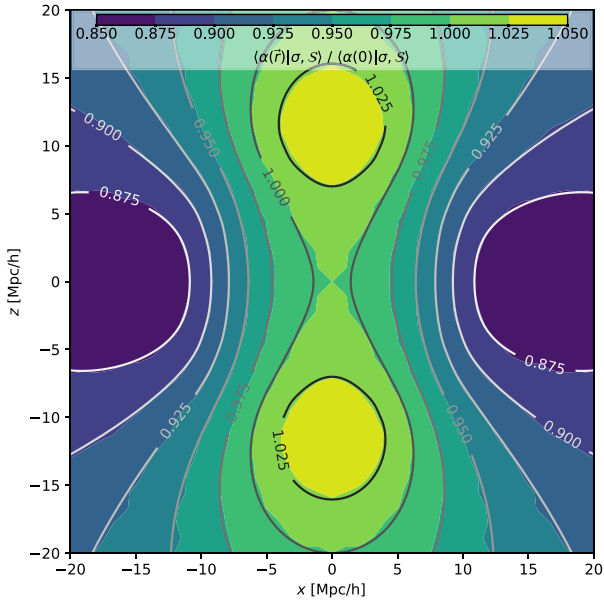
Figure B2. Mean value of α using a numerical method (purple to yellow) versus its theoretical value (grey contours). Both are normalized by the theoretical value at the saddle point.

theoretical CDF F_{up} (with the measured Γ) at four different positions are shown on Fig. B1. The measured CDFs have been normalized so that $F_{first}^{-1}(0.5) = F_{up}^{-1}(0.5)$: we impose that the CDF match at the 'median' (defined as the σ such that $F(\sigma) = 0.5^{12}$). As shown on Fig. B2, the abscissa of the peak of the PDF in the direction of the void is around $\sigma \approx 2.7$. As $\sigma(R_{min}) \approx 3$, it means that in the direction of the void, the PDF is only sampled up to its peak. The experimental CDF at such location is hence only probing less than 50˜ per cent of the distribution and the median is not reached. In this case, we are normalizing the experimental CDF to have the same

---

[12] This definition matches the classical one for distributions that have a normalized CDF, which is not true for F_{up}.

value at the largest σ as the theoretical CDF. As shown on Fig. B1, the experimental and theoretical CDFs start diverging at $F \gtrsim 0.5$. At larger σ, the upcrossing approximation used in the theory breaks as more and more trajectories cross multiple time the barrier (they are counted once for the first crossing and multiple times for upcrossing). The orange and blue lines, in the direction of the filament show this clearly as they diverge one from each other at large σ. As $\sigma_\star$ is a measure of the location of the peak of the PDF (which is where the CDF is the steepest), it is sufficient that the experimental and theoretical CDF match up to their flat end to have the same $\sigma_\star$ values.

B2 Validation for $\alpha_\star$ using constrained fields

A second check was implemented on the accretion rate as follows: (i) for each location, the covariance matrix of ν, δ', ν_S, $\bar{q}_{ij}$, g_i was computed at finite distance. These quantities all have a null mean; (ii) the covariance matrix and the mean of ν, δ' conditioned to the value at the saddle point was computed using the values of Table D1; (iii) the variance and mean of ν, δ' were computed given $\nu = \nu_c$ and the saddle point; and (iv) a sample of 10^6 points were then drawn from the distribution of $\delta' > 0$ (upcrossing). (v) The values of $\alpha \propto 1/\delta'$ were computed to obtain a sample of α. Each draw was weighted by $1/\alpha$ (the Jacobian of the transform from δ' to α). Finally, the numerical value of $\langle \alpha | \sigma, S \rangle$ was estimated from the samples and compared with the theoretical value. The results are shown on Fig. B2 and are found to be in very good agreement.

We computed Fig. B3 by following steps (i)–(iii) at 10 Mpc h^{-1} in the direction of the filament (blue) and of the void (orange) and plotting the mean and standard deviation of δ given the saddle and the threshold. Fig. 15 was computed by following steps (i)–(iii) at the saddle point ($r = 0$). An extra constrain on the value of δ' was then added to compute the different curves.

APPENDIX C: OTHER CRITICAL POINTS

For the sake of generality, let us discuss here the conditional excursion set expectations in the vicinity of other critical points of the potential. At the technical level, all the formulae we derived in Section 4 depend on the eigenvalues of q_{ij} with no a priori assumption on their sign. The expressions will thus remain formally the same, with all information about the environment being channelled through the values of ν_S and $\hat{r}_i \bar{q}_{ij} \hat{r}_j$. For instance, the typical quantities $M_\star$, $\dot{M}_\star$, and $z_\star$ parametrizing the PDFs of interest will be defined in exactly the same way as in equations (80)–(82). However, their level curves will have different profiles in different environments.

As physical intuition suggests, and equation (47) explicitly shows, the dependence of the various halo statistics on the distance from the stationary point (whether the probability of a given halo property increases or decreases with separation) is encoded in the signs of the eigenvalues q_i of q_{ij}. Besides filaments (having two positive eigenvalues), one may thus be interested in wall-type saddles (one positive eigenvalue), maxima (all negative), and minima (all positive), corresponding to voids and nodes, respectively. In general, $q_1 + q_2 + q_3 = \nu_S$ parametrizes the mean variation with distance (averaged over the angles), whereas the traceless shear $\bar{q}_{ij}$ is responsible for the angular variation at fixed distance.

In all cases, however, for a given direction $M_\star$, $\dot{M}_\star$, and $-z_\star$ will either all increase (if $r_i q_{ij} r_j < 0$) or all decrease (if $r_i q_{ij} r_j > 0$).
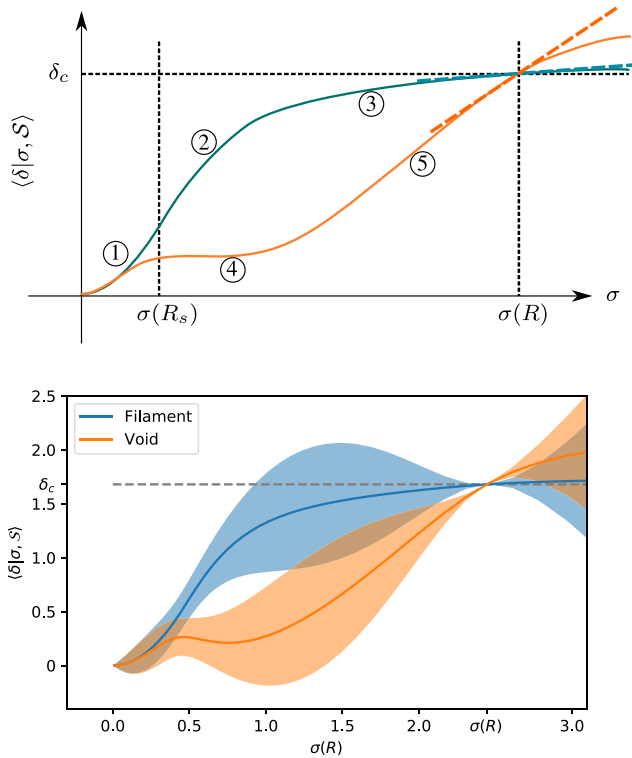
Figure B3. Top: scheme of the mean value of the density in the direction of a filament (red) and void (blue) close to a saddle point smoothed at $\sigma = \sigma_S$ with the constrain that $\delta(\sigma(R)) = \delta_c$. (1) The value of the density imposed at the saddle point forces both mean densities to increase. (2) In the direction of the filament, a large-scale overdensity, the mean density at a given point increases quickly, but (3) the constrain $\delta(\sigma) = \delta_c$ prevents any further increase at $\sigma \lesssim \sigma(R)$, hence the slope δ' is small at upcrossing. (4) In the direction of the void, a large-scale underdensity, the mean density at a given point cannot increase with σ. (5) At $\sigma \lesssim \sigma(R)$, the upcrossing constrain forces a sharp increase of the density to reach $\delta(R) = \delta_c$, hence the slope is high at upcrossing. Bottom: a validation using constrained GRF at a distance of $10 \, \mathrm{Mpc} \, h^{-1}$ in the direction of the filament (blue) and of the void (orange). See the text for the details.

Their increase will be fastest (or their decrease slowest) in the direction of $\bar{q}_3$, the least negative eigenvalue, and slowest in that of $\bar{q}_1$. The rationale of this behaviour will always be that an increase of the conditional mean density will make it easier for excursion set trajectories to reach the threshold. Upcrossing will happen preferentially at smaller σ, corresponding to the formation of haloes of bigger mass. At fixed mass (fixed crossing scale σ), the crossing will happen preferentially with shallower slopes, corresponding to higher accretion rates and more recent formation (i.e. assembly of half-mass).

C1 Walls

A wall will form in correspondence of a saddle point of the potential filtered on scale R_S, for which $q_1 < q_2 < 0 < q_3$. This combination of eigenvalue signs generates collapse in one spatial direction and expansion in the other two. As argued, a saddle point of the potential induces a saddle point of the opposite type in $M_\star$, $\dot{M}_\star$, and $-z_\star$, which will increase along two space directions following the increase of the mean density, and decrease along one. Since for walls (like for filaments), the value of ν_S is likely to be smaller than $\sqrt{\mathrm{tr}(\bar{q}^2)}$, they will tend to have an angular modulation larger
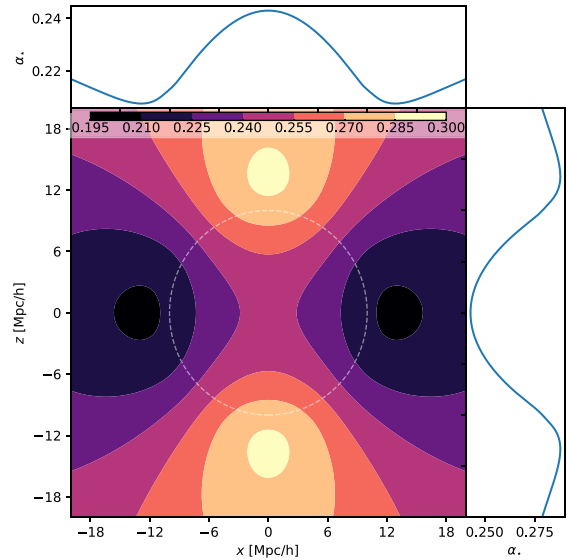
Figure C1. Isocontours in the x–z plane of the typical accretion rate $\alpha_\star$ around a wall-type saddle point [at (0, 0)]. The saddle point is defined using the values of Table D1. The profiles in the main direction of the wall (z-direction) and of the void (x-direction) are plotted on the sides. The smoothing scale is $R = 1 \, \mathrm{Mpc} \, h^{-1}$. The typical accretion rate is computed using a ΛCDM power spectrum. Similarly to what happens in filaments, haloes accrete more in the direction of the wall than in the direction of the void.

than the radial angle-averaged variation. Walls are thus likely to be highly anisotropic configurations also of the accretion rate and of the formation time. This is illustrated for example in Fig. C1 for the accretion rate. On average, ν_S will be smaller for a wall-type saddle (which has two negative eigenvalues) than for a filament-type one. Thus, haloes in walls tend to be less massive, and at fixed mass, they tend to have smaller accretion rates and earlier assembly times.

C2 Voids

A void will eventually form (although not necessarily by $z = 0$) when $r = 0$ is a local maximum of the potential filtered on scale R_S (from which matter flows away), for which $q_1 < q_2 < q_3 < 0$. The centre of the void is a minimum of $M_\star$, $\dot{M}_\star$, and $-z_\star$. All these quantities will gradually increase with the separation. As $|\nu_S|$ may be large (in particular for a large, early-forming void), halo statistics in voids may not show a large anisotropy relative to their radial variation. However, because voids have the most negative ν_S, they are the environment with the least massive haloes, the smallest accretion rates and the earliest formation times (at fixed mass).

C3 Nodes

Nodes form out of local minima of the gravitational potential, for which $0 < q_1 < q_2 < q_3$ (corresponding to three directions of infall). The centre of the node is thus a maximum of $M_\star$, $\dot{M}_\star$, and $-z_\star$, all of which decrease with radial separation. Like voids, large early-forming nodes (whose density ν_S must reach ν_c when σ_S is very small) are relatively less anisotropic, since the relative amplitude of the angular variation induced by $\bar{q}_{ij}$ is likely to be small compared to the radial variation. Since ν_S is the largest for nodes, they host the most massive haloes, and at fixed mass, those with the largest accretion rates and the latest formation times.

APPENDIX D: PDF OF SADDLES

This section presents the distribution of the eigenvalues of the anisotropic (i.e traceless) part of the tidal tensor at critical points of the potential field. By definition, a critical point is such that $g_i = 0$ and its kind is given by the signature (the signs of the eigenvalues of the hessian of the potential, q_{ij}): $+ + +$ for a peak, $- + +$ for a filament-type saddle point, $- - +$ for a wall-type saddle point, and $- - -$ for a void. Because the anisotropic tidal tensor reads $\bar{q}_{ij} = q_{ij} - \delta_{ij} \nu_{\mathcal{S}}/3$, the type of the critical point is then given by the number of eigenvalues of $\bar{q}_{ij}$ above $-\nu_{\mathcal{S}}/3$.

The distribution of the eigenvalues of the (normalized) tidal tensor denoted $q_1 < q_2 < q_3$ is described by the Doroshkevich formula (Doroshkevich [1970]; Pogosyan et al. [1998])

$$p(q_i) = \frac{675\sqrt{5}}{8\pi} \exp\left[\frac{15}{2}I_2 - 3I_1^2\right] (q_3 - q_1)(q_3 - q_2)(q_2 - q_1), \tag{D1}$$

where $\{I_n\}$ denotes the rotational invariants which define the characteristic polynomial of q_{ij}, namely its trace $I_1 = q_1 + q_2 + q_3$, trace of the comatrix $I_2 = q_1 q_2 + q_2 q_3 + q_1 q_3$, and determinant $I_3 = q_1 q_2 q_3$. Subject to a filament-type saddle-point constraint, this PDF becomes

$$p(q_i | -++) = \frac{540\sqrt{5\pi}}{29\sqrt{2} + 12\sqrt{3}} q_1 q_2 q_3 \vartheta(q_2) \vartheta(-q_1) p(q_i), \tag{D2}$$

after imposing the condition of a saddle $|\det q_{ij}| \delta_D(g_i) \vartheta(q_2) \vartheta(-q_1)$ for which as the acceleration is decoupled from the tidal tensor, only the condition on the sign of the eigenvalues and the determinant contribute. From this PDF, it is straightforward to compute the distribution of saddles of heights $\nu_{\mathcal{S}} = q_1 + q_2 + q_3$

$$p(\nu_{\mathcal{S}} | -++) = p^+(\nu_{\mathcal{S}}) \vartheta(\nu_{\mathcal{S}}) + p^-(\nu_{\mathcal{S}}) \vartheta(-\nu_{\mathcal{S}}), \tag{D3}$$

with

$$p^+(\nu_{\mathcal{S}})$$
$$= \frac{5\sqrt{10\pi} e^{-\frac{\nu_{\mathcal{S}}^2}{2}} \left(3\nu_{\mathcal{S}} - \nu_s^3\right) \mathrm{Erfc}\left(\frac{\sqrt{5}\nu_{\mathcal{S}}}{2\sqrt{2}}\right) + e^{-\frac{9\nu_{\mathcal{S}}^2}{8}} \left(32 + 155\nu_s^2\right)}{\left(29\sqrt{2} + 12\sqrt{3}\right)\sqrt{\pi}},$$

$$p^-(\nu_{\mathcal{S}})$$
$$= \frac{5\sqrt{10\pi} e^{-\frac{\nu_{\mathcal{S}}^2}{2}} \left(3\nu_{\mathcal{S}} - \nu_s^3\right) \mathrm{Erfc}\left(\frac{-\sqrt{5}\nu_{\mathcal{S}}}{\sqrt{2}}\right) + e^{-3\nu_{\mathcal{S}}^2} \left(32 - 10\nu_{\mathcal{S}}^2\right)}{(29\sqrt{2} + 12\sqrt{3})\sqrt{\pi}}.$$

In particular, the height of filament-type saddles has mean and standard deviation given by

$$\langle \nu_{\mathcal{S}} | -++ \rangle = 250 \left(3(29\sqrt{2} + 12\sqrt{3})\sqrt{\pi}\right)^{-1} \approx 0.76,$$

$$\mathrm{Std}\left(\nu_{\mathcal{S}} | -++\right) = \frac{\sqrt{696\sqrt{6} + 75\pi(10 - 3\sqrt{6}) - 2114}}{15\sqrt{\pi}} \approx 0.55.$$

For other types of critical points, a similar calculation can be done. As expected, the heights of wall-type saddle points follow the same distribution as $-\nu_{\mathcal{S}}$. Peak and void heights have mean $\pm\sqrt{2114 + 696\sqrt{6}}/15\sqrt{\pi} \approx \pm 2.3$ and standard deviation $\sqrt{75\pi(10 + 3\sqrt{6}) - (2114 + 696\sqrt{6})}/15\sqrt{\pi} \approx 0.62$.

This work picks a typical value for the filament-type saddle at roughly 1σ from the mean $\nu_{\mathcal{S}} = 1.2$. For wall-type saddles, $\nu_{\mathcal{S}} = 0$ is chosen. The distribution of eigenvalues of the anisotropic tidal

tensor $\bar{q}_i$ for a filament-type saddle point with a given positive[13] height can then be easily obtained from equation (D2)

$$p(\bar{q}_1 | \nu_{\mathcal{S}}) = \frac{15(3\bar{q}_1 + \nu_{\mathcal{S}}) \left[a_1 e^{-\frac{4\nu_{\mathcal{S}}^2}{3} + \frac{5}{2}\bar{q}_1 \nu_{\mathcal{S}} - \frac{15\bar{q}_1^2}{2}} - a_2 e^{-\frac{\nu_{\mathcal{S}}^2}{2} - \frac{45\bar{q}_1^2}{8}}\right]}{16(29\sqrt{2} + 12\sqrt{3})\sqrt{\pi}\mathcal{P}^+(\nu_{\mathcal{S}})},$$

where $\bar{q}_1 < -\nu_{\mathcal{S}}/3$ and a_1 and a_2 are two polynomials of $\bar{q}_1$ and $\nu_{\mathcal{S}}$ given by

$$a_1(\bar{q}_1, \nu_{\mathcal{S}}) = 32\left[5|\nu_{\mathcal{S}} - 6\bar{q}_1|(3\bar{q}_1 + \nu_{\mathcal{S}}) + 12\right],$$

and

$$a_2 = 6075\bar{q}_1^4 - 8100\bar{q}_1^3 \nu_{\mathcal{S}} + 900\bar{q}_1^2 \left(3\nu_{\mathcal{S}}^2 - 4\right) + 480\bar{q}_1 \nu_{\mathcal{S}}$$
$$- 160\nu_{\mathcal{S}}^2 + 384.$$

Similarly, the PDF of the intermediate and major eigenvalues are, respectively, given by

$$p(\bar{q}_2 | \nu_{\mathcal{S}}) = \frac{15(3\bar{q}_2 + \nu_{\mathcal{S}}) a_1 e^{-\frac{11}{12}\nu_{\mathcal{S}}^2 + \frac{5}{4}\bar{q}_2 \nu_{\mathcal{S}} - 15\bar{q}_2^2 - \frac{5}{12}(\nu_{\mathcal{S}} + 3\bar{q}_2)|\nu_{\mathcal{S}} - 6\bar{q}_2|}}{16(29\sqrt{2} + 12\sqrt{3})\sqrt{\pi}\mathcal{P}^+(\nu_{\mathcal{S}})}$$

where $\bar{q}_2 > -\nu_{\mathcal{S}}/3$ and $a_1 = a_1(\bar{q}_2, \nu_{\mathcal{S}})$, and

$$p(\bar{q}_3 | \nu_{\mathcal{S}}) = \frac{15(3\bar{q}_3 + \nu_{\mathcal{S}}) \left[a_1 e^{-\frac{\nu_{\mathcal{S}}^2}{2} - \frac{45\bar{q}_3^2}{2}} + \bar{a}_1 e^{\frac{4\nu_{\mathcal{S}}^2}{3} + \frac{5}{2}\bar{q}_3 \nu_{\mathcal{S}} - \frac{15\bar{q}_3^2}{2}}\right]}{16(29\sqrt{2} + 12\sqrt{3})\sqrt{\pi}\mathcal{P}^+(\nu_{\mathcal{S}})}$$

where $\bar{q}_3 > \nu_{\mathcal{S}}/6$, having defined $a_1 = a_1(\bar{q}_3, \nu_{\mathcal{S}})$ and $\bar{a}_1(\bar{q}_3, \nu_{\mathcal{S}}) = -a_1(-\bar{q}_3, -\nu_{\mathcal{S}})$. Similar expressions can be obtained for wall-type saddles (together with peaks and voids). The top panel of Fig. [D1] shows the distribution of eigenvalues for a filament-type saddle point of height $\nu_{\mathcal{S}} = 1.2$ and the bottom panel shows the distribution for a wall-type saddle point of height $\nu_{\mathcal{S}} = 0$. Typical values of $\bar{q}_{ij}$ were selected to correspond roughly to the maximum of the above-mentioned distributions of $\bar{q}_1, \bar{q}_2, \bar{q}_3$ and are reported in Table [D1]. Note that all the results obtained in this section are independent of the power spectrum. The only assumption is that the density is a GRF.

APPENDIX E: COVARIANCE MATRICES

Let us present here the covariance matrix of all variables introduced in the main text. The density δ and slope δ' are evaluated at position $\boldsymbol{r}$ and smoothed on the halo scale R, the half-mass density $\delta_{1/2}$ is also evaluated at the halo position $\boldsymbol{r}$ but smoothed on $R_{1/2} = 2^{-1/3}R$, while the saddle rareness $\nu_{\mathcal{S}}$, acceleration g_i, and detraced tidal tensor $\bar{q}_{ij}$ are evaluated at the origin and smoothed on a scale $R_{\mathcal{S}} \gg R$. The correlation matrix of $X \equiv \left\{\delta, \delta', \nu_{1/2}, \nu_{\mathcal{S}}, g_i, \bar{q}_{ij}\right\}$, a vector with 12 Gaussian components, is

$$\boldsymbol{C} = \begin{pmatrix} \sigma^2 & \sigma & \omega & \boldsymbol{C}_{14} & \boldsymbol{C}_{15} & \boldsymbol{C}_{16} \\ \sigma & \langle\delta'^2\rangle & \omega' & \boldsymbol{C}_{24} & \boldsymbol{C}_{25} & \boldsymbol{C}_{26} \\ \omega & \omega' & \sigma_{1/2}^2 & \boldsymbol{C}_{34} & \boldsymbol{C}_{35} & \boldsymbol{C}_{36} \\ \boldsymbol{C}_{14} & \boldsymbol{C}_{24} & \boldsymbol{C}_{34} & 1 & 0 & 0 \\ \boldsymbol{C}_{15}^T & \boldsymbol{C}_{25}^T & \boldsymbol{C}_{35}^T & 0 & \boldsymbol{C}_{55} & 0 \\ \boldsymbol{C}_{16}^T & \boldsymbol{C}_{26}^T & \boldsymbol{C}_{36}^T & 0 & 0 & \boldsymbol{C}_{66} \end{pmatrix}, \tag{E1}$$

with $\omega = \langle\delta\nu_{1/2}\rangle$, $\omega' = \langle\delta'\nu_{1/2}\rangle$, and

$$\boldsymbol{C}_{14} = \langle\delta\nu_{\mathcal{S}}\rangle = \xi_{00}, \quad \boldsymbol{C}_{15} = \langle\delta g_i\rangle = \frac{r_i}{R_\star}\xi_{11}, \tag{E2}$$

$$\boldsymbol{C}_{16} = \langle\delta\bar{q}_{ij}\rangle = \left(\frac{\delta_{ij}}{3} - \hat{r}_i \hat{r}_j\right)\xi_{20}, \tag{E3}$$

[13] A similar expression can be obtained for negative heights.
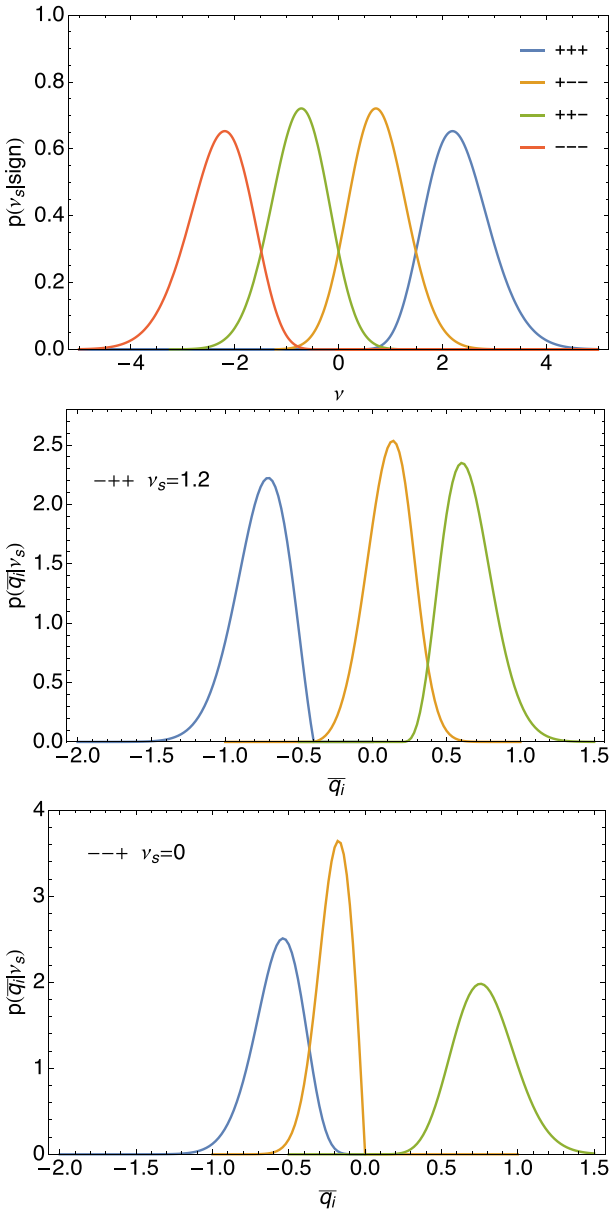
Figure D1. Top panel: distribution of heights of critical points of various signatures (peaks, filament-type saddles, wall-type saddles, and voids) for GRF with any power spectrum. Middle panel: PDF of the eigenvalues, $\bar{q}_1$ (blue), $\bar{q}_2$ (yellow), and $\bar{q}_3$ (green), of the anisotropic tidal tensor given a filament-type constraint at $\nu_{\mathcal{S}} = 1.2$. Bottom panel: same as middle panel for a wall-type constraint at $\nu_{\mathcal{S}} = 0$.

Table D1. Eigenvalues $\bar{q}_i = q_i - \nu_{\mathcal{S}}/3$ of the traceless tidal tensor $\bar{q}_{ij}$, height $\nu_{\mathcal{S}}$, and smoothing scale used to define the saddle points. See Appendix D for details.

	Traceless tide			Height	Scale	Saddle type
Quantity	$\bar{q}_1$	$\bar{q}_2$	$\bar{q}_3$	$\nu_{\mathcal{S}}$	$R_{\mathcal{S}}$	
Value	-0.7	0.1	0.6	1.2	$10\,\mathrm{Mpc}\,h^{-1}$	Filament-type
Value	-0.6	-0.2	0.8	0	$10\,\mathrm{Mpc}\,h^{-1}$	Wall-type

$$C_{24} = \langle \delta' \nu_{\mathcal{S}} \rangle = \xi'_{00}, \quad C_{25} = \langle \delta' g_i \rangle = \frac{r_i}{R_\star} \xi'_{11}, \tag{E4}$$

$$C_{26} = \langle \delta' \bar{q}_{ij} \rangle = \left(\frac{\delta_{ij}}{3} - \hat{r}_i \hat{r}_j \right) \xi'_{20}, \tag{E5}$$

$$C_{34} = \langle \nu_{1/2} \nu_{\mathcal{S}} \rangle = \frac{\xi^{(1/2)}_{00}}{\sigma_{1/2}}, \quad C_{35} = \langle \delta_{1/2} g_i \rangle = \frac{r_i}{R_\star} \frac{\xi^{(1/2)}_{11}}{\sigma_{1/2}}, \tag{E6}$$

$$C_{36} = \langle \delta_{1/2} \bar{q}_{ij} \rangle = \left(\frac{\delta_{ij}}{3} - \hat{r}_i \hat{r}_j \right) \frac{\xi^{(1/2)}_{20}}{\sigma_{1/2}}, \tag{E7}$$

$$C_{55} = \langle g_i g_j \rangle = \frac{\delta_{ij}}{3}, \quad C_{66} = \langle \bar{q}_{ij} \bar{q}_{kl} \rangle = \frac{2 P_{ij,kl}}{15}. \tag{E8}$$

Hence, C_{14}, C_{24}, and C_{34} are scalars, C_{15}, C_{25}, and C_{35} are three vectors, C_{16}, C_{26}, and C_{36} are 3×3 traceless matrices (or five vectors in the space of symmetric traceless matrices), C_{55} is a 3×3 matrix, and C_{66} is a 5×5 matrix. The matrix C_{66} involves

$$P_{ij,kl} \equiv \frac{\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}}{2} - \frac{\delta_{ij}\delta_{kl}}{3}, \tag{E9}$$

a projector that removes the trace and the antisymmetric part from a matrix. Since $P_{ij,ab} P_{ab,mn} = P_{ij,mn}$ and so $P^{-1}_{ij,mn} = P_{ij,mn}$, it acts as the identity in the space of symmetric traceless matrices. $P_{ij,kl}$ can be written in its matrix form by numbering the pairs $\{(1, 1), (2, 2), (1, 2), (1, 3), (2, 3)\}$ from 1 to 5, the dimensionality of the space, resulting in a 5×5 matrix. The element $(3, 3)$ has been dropped because it is linearly linked to $(1, 1)$ and $(2, 2)$. The explicit value of C_{66} is therefore

$$C_{66} = \frac{1}{45} \begin{pmatrix} 4 & -2 & 0 & 0 & 0 \\ -2 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}. \tag{E10}$$

The finite separation correlation functions $\xi_{\alpha\beta}(r, R, R_{\mathcal{S}})$ and $\xi'_{\alpha\beta}(r, R, R_{\mathcal{S}})$ are defined as

$$\xi_{\alpha\beta} \equiv \int \mathrm{d}k \, \frac{k^2 P(k)}{2\pi^2} W(kR) \frac{W(kR_{\mathcal{S}})}{\sigma_{\mathcal{S}}} \frac{j_\alpha(kr)}{(kr)^\beta}, \tag{E11}$$

$$\xi'_{\alpha\beta} \equiv \int \mathrm{d}k \, \frac{k^2 P(k)}{2\pi^2} W'(kR) \frac{W(kR_{\mathcal{S}})}{\sigma_{\mathcal{S}}} \frac{j_\alpha(kr)}{(kr)^\beta}, \tag{E12}$$

where $W'(kR) = [\mathrm{d}W(kR)/\mathrm{d}R]/(\mathrm{d}\sigma/\mathrm{d}R)$. Similarly, the correlation functions at the two different mass scales M and $M/2$ are

$$\xi^{(1/2)}_{\alpha\beta} \equiv \xi_{\alpha\beta}(r, R_{1/2}, R_{\mathcal{S}}), \tag{E13}$$

where $R_{1/2} \equiv R/2^{1/3}$. At null separation $(r = 0)$, it yields

$$\omega = \frac{\langle \delta \delta_{1/2} \rangle}{\sigma_{1/2}} = \int \mathrm{d}k \, \frac{k^2 P(k)}{2\pi^2} W(kR) \frac{W(kR_{1/2})}{\sigma_{1/2}}, \tag{E14}$$

$$\omega' = \frac{\langle \delta' \delta_{1/2} \rangle}{\sigma_{1/2}} = \int \mathrm{d}k \, \frac{k^2 P(k)}{2\pi^2} W'(kR) \frac{W(kR_{1/2})}{\sigma_{1/2}}. \tag{E15}$$

Recall that for a Top-Hat filter, one has

$$W(kR) = \frac{3 j_1(kR)}{kR} \quad \text{and} \quad W'(kR) = \frac{3 j_2(kR)}{R|\mathrm{d}\sigma/\mathrm{d}R|}, \tag{E16}$$

and notice that $W'(kR)$ is suppressed by a factor of $k^2 R^2$ with respect to $W(kR)/\sigma$ when $k \ll 1/R$. In fact, in this limit $j_n(kR) \sim (kR)^n/(2n + 1)!!$. Hence, the action of $\mathrm{d}/\mathrm{d}\sigma$ is proportional to that of $R^2 \nabla^2$, and $\sigma \xi'_{\alpha\beta} \propto R^2 \nabla^2 \xi_{\alpha\beta} \sim (R/R_{\mathcal{S}})^2 \xi_{\alpha\beta}$. It follows that for $R \ll R_{\mathcal{S}}$ one has $\sigma \xi'_{\alpha\beta} \ll \xi_{\alpha\beta}$. In presence of a strong hierarchy of scales, the terms containing $\xi'_{\alpha\beta}$ are negligible (see Fig. E1).
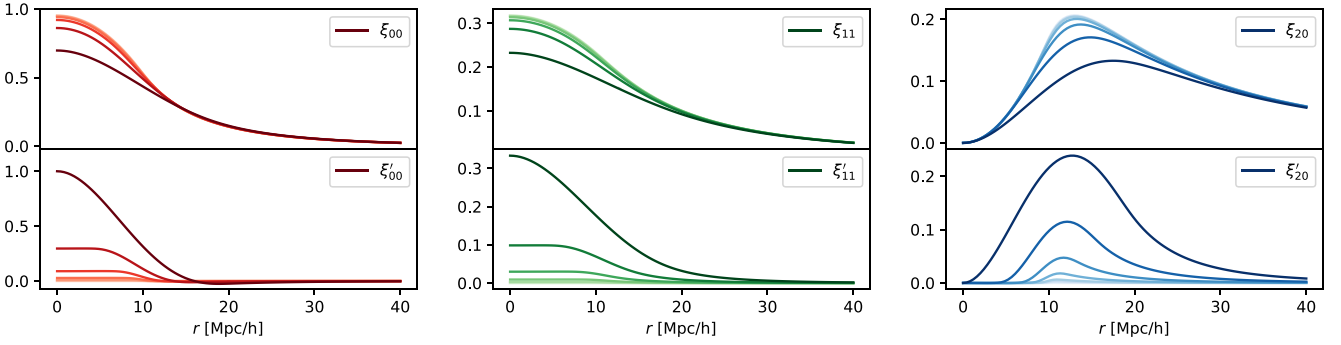
Figure E1. Plot as a function of r of the correlation functions defined in equation (E12). From left to right on the top row ξ_{00}, ξ_{11}, and ξ_{20}. The bottom row shows the same quantities derived with respect to σ. The correlation functions are evaluated at $R_S = 10^{-}$Mpc h^{-1} for different values of R logarithmically spaced between 10^{-1} Mpc h^{-1} (light colour) and 10 Mpc h^{-1} (dark colours) with a ΛCDM power spectrum and plotted as a function of the distance r.

For a scale invariant power spectrum $P(k) = A(k/k_0)^{-n}$, $\xi_{\alpha\beta}$ and $\xi'_{\alpha\beta}$ have an analytical expression that depends on the relation between r, R_S, and R. For example, when $R_S > r + R$:

$$\frac{\xi_{\alpha\beta}(r, R, R_S)}{\sigma_S}$$
$$= B F_4 \left(\frac{\alpha - \beta - n}{2}, \frac{3 + \alpha - \beta - n}{2}; \frac{5}{2}, \alpha + \frac{3}{2}; \frac{R^2}{R_S^2}, \frac{r^2}{R_S^2} \right)$$

and

$$\xi'_{\alpha\beta}(r, R, R_S) = \frac{2(\alpha - \beta - n + 3)(n - \alpha + \beta)}{5(n - 3)} \left(\frac{R}{R_S} \right)^{\frac{7-n}{2}} B$$
$$\times F_4 \left(\frac{2 + \alpha - \beta - n}{2}, \frac{5 + \alpha - \beta - n}{2}; \frac{7}{2}, \alpha + \frac{3}{2}; \frac{R^2}{R_S^2}, \frac{r^2}{R_S^2} \right),$$

where F_4 is the Appell Hypergeometric function of the fourth kind (Gradshteyn & Ryzhik 2007, p. 677),[14] while

$$B = -\left(\frac{r}{R_S} \right)^{\alpha - \beta}$$
$$\times \frac{\pi(n + 3) \csc\left(\frac{n\pi}{2} \right) \Gamma\left(\frac{3 + \alpha - \beta - n}{2} \right)}{2^{\beta + 2n + 2} 3(n - 1) \Gamma\left(\frac{3 + 2\alpha}{2} \right) \Gamma(-n - 1) \Gamma\left(\frac{n - \alpha + \beta + 2}{2} \right)}$$

and

$$\sigma^2(R) = \sigma_8^2 \left(\frac{R}{R_8} \right)^{n-3}, \quad \frac{d\log\sigma^2}{d\log R} = n - 3, \quad (E17)$$

where $R_8 = 8$ Mpc h^{-1} and $\sigma_8 = \sigma(R_8)$ are normalization factors. For the same power-law power spectrum, setting $\alpha = 1 + n$ and $\beta = R_{1/2}/R = 2^{-1/3}$, ω and ω' defined in equations (E14) and (E15) have the analytical expressions

$$\frac{\omega}{\sigma} = \frac{(1 + \beta)^\alpha \left(\beta^2 - \alpha\beta + 1 \right) - (1 - \beta)^\alpha \left(\beta^2 + \alpha\beta + 1 \right)}{2^\alpha (2 - \alpha) \beta^{\frac{\alpha + 2}{2}}}, \quad (E18)$$

and

$$\omega' = \frac{\left(3\beta^3 + \beta n^2 + 3\beta^2 n + n \right)(1 - \beta)^n}{2^n \beta^{\frac{n+3}{2}}(n - 3)(n - 1)}$$
$$+ \frac{\left(3\beta^3 + \beta n^2 - 3\beta^2 n - n \right)(1 + \beta)^n}{2^n \beta^{\frac{n+3}{2}}(n - 3)(n - 1)}. \quad (E19)$$

[14] http://mathworld.wolfram.com/AppellHypergeometricFunction.html

APPENDIX F: CONDITIONAL STATISTICS

The goal of this section is to derive explicitly the conditional statistics needed in the paper. Assuming that the underlying density field obeys Gaussian statistics, the PDF of the 12-dimensional vector $\boldsymbol{X} \equiv \left\{ \delta(\boldsymbol{r}), \delta'(\boldsymbol{r}), \nu_{1/2}(\boldsymbol{r}), \nu_S, g_i, \bar{q}_{ij} \right\}$ already defined in Appendix E involves inverting the 12×12 covariance matrix $\boldsymbol{C} \equiv \langle \boldsymbol{X} \cdot \boldsymbol{X}^{\mathrm{T}} \rangle$, given by equation (E1). Since however the focus here is on conditioning heights and slopes, which are scalar quantities, their correlation with the saddle is the correlation with the three unit-variance Gaussian components

$$\mathcal{S}(\hat{\boldsymbol{r}}) \equiv \{ \nu_S, \sqrt{3}\hat{r}_i g_i r / R_\star, -\sqrt{5}(3\hat{r}_i \bar{q}_{ij} \hat{r}_j / 2) \}. \quad (F1)$$

Hence, the six-dimensional vector $\tilde{\boldsymbol{X}} \equiv \{ \delta(\boldsymbol{r}), \delta'(\boldsymbol{r}), \nu_{1/2}(\boldsymbol{r}), \mathcal{S} \}$ is sufficient, and has a 6×6 covariance matrix given by

$$\tilde{\boldsymbol{C}}(r) = \begin{pmatrix} \sigma^2 & \sigma & \omega & \xi(r) \\ \sigma & \langle \delta'^2 \rangle & \omega' & \xi'(r) \\ \omega & \omega' & \sigma_{1/2}^2 & \xi_{1/2}(r) \\ \xi^T(r) & \xi'^T(r) & \xi_{1/2}^T(r) & \mathbb{1}_{3\times 3} \end{pmatrix}, \quad (F2)$$

where

$$\xi(r) \equiv \left\{ \xi_{00}, \sqrt{3}\xi_{11} r / R_\star, \sqrt{5}\xi_{20} \right\},$$
$$\xi'(r) \equiv \left\{ \xi'_{00}, \sqrt{3}\xi'_{11} r / R_\star, \sqrt{5}\xi'_{20} \right\},$$
$$\xi_{1/2}(r) \equiv \left\{ \xi_{00}^{(1/2)}, \sqrt{3}r/R_\star \xi_{10}^{(1/2)}, \sqrt{5}\xi_{20}^{(1/2)} \right\} / \sigma_{1/2}. \quad (F3)$$

The PDF of $\tilde{\boldsymbol{X}}$ is the six-variate Gaussian

$$p_{\mathrm{G}}(\tilde{\boldsymbol{X}}) = \frac{1}{(2\pi)^3 \sqrt{\det\tilde{\boldsymbol{C}}}} \exp\left(-\frac{1}{2} \tilde{\boldsymbol{X}} \cdot \tilde{\boldsymbol{C}}^{-1} \cdot \tilde{\boldsymbol{X}} \right), \quad (F4)$$

so that in each case, the task is to invert the appropriate section of the covariance matrix $\tilde{\boldsymbol{C}} \equiv \langle \tilde{\boldsymbol{X}} \cdot \tilde{\boldsymbol{X}}^{\mathrm{T}} \rangle$, marginalizing over the variables that are not involved.

F1 The general conditional case

To speed up the computation of conditional statistics, rather than doing a brute force block inversion of $\tilde{\boldsymbol{C}}$, it is best to use the decorrelated variables

$$\nu_v \equiv \frac{\delta - \langle \delta | \{v\} \rangle}{\sqrt{\mathrm{Var}\left(\delta | \{v\}\right)}}, \quad \text{and} \quad \nu'_v \equiv \frac{d\nu_v}{d\sigma}, \quad (F5)$$

where the possible $\{v\}$ considered in this work are $\nu_{1/2}$, $\mathcal{S}$ or $\{\nu_{1/2}, \mathcal{S}\}$. By construction, ν_v and ν'_v are uncorrelated, because ν_v has unit variance. Furthermore, if each ν_I is independent of σ (as

it will be the case in the following), ν'_v does not correlate with the constraint either, since $\langle \nu'_v \nu_I \rangle = \langle \nu_v \nu_I \rangle' = 0$. Then, being a linear combination of δ', ν, and $\{v\}$ that does not correlate with ν nor ν_I, ν'_v must be proportional to $\delta' - \langle \delta' | \nu, \{v\} \rangle$ (the only such linear combination by definition), and $\langle \nu'^2_v \rangle$ to $\mathrm{Var}\left(\delta' | \nu, \{v\}\right)$. That is,

$$\langle \delta' | \nu, \{v\} \rangle = \delta' - \sqrt{\mathrm{Var}\left(\delta | \{v\}\right)}\, \nu'_v,$$

$$= \langle \delta' | \{v\} \rangle + \frac{[\mathrm{Var}\left(\delta | \{v\}\right)]'}{2\mathrm{Var}\left(\delta | \{v\}\right)}\left(\delta - \langle \delta | \{v\} \rangle\right), \tag{F6}$$

$$\mathrm{Var}\left(\delta' | \nu, \{v\}\right) = \mathrm{Var}\left(\delta | \{v\}\right) \langle \nu'^2_v \rangle,$$

$$= \mathrm{Var}\left(\delta' | \{v\}\right) - \frac{[\mathrm{Var}\left(\delta | \{v\}\right)]'^2}{4\mathrm{Var}\left(\delta | \{v\}\right)}, \tag{F7}$$

providing the conditional statistics of δ' given ν and $\{v\}$ in terms of those of δ and δ' given $\{v\}$ alone. Since $[\mathrm{Var}\left(\delta | \{v\}\right)]' = 2\mathrm{Cov}\left(\delta, \delta' | \{v\}\right)$, these formulae reduce to the standard results for constrained Gaussian variables, but taking derivatives makes their calculation easier.

To compute ν_v and ν'_v explicitly, one needs to insert (using Einstein's convention on repeated indices)

$$\langle \delta | \{v\} \rangle = \psi_I C_{IJ}^{-1} v_J, \tag{F8}$$

$$\mathrm{Var}\left(\delta | \{v\}\right) = \sigma^2 - \psi_I C_{IJ}^{-1} \psi_J, \tag{F9}$$

in equation (F5), where $C_{IJ} \equiv \langle v_I v_J \rangle$ is the covariance matrix of the constraint, and $\psi_I \equiv \langle \delta v_I \rangle$ is the mixed covariance. The conditional statistics obtained from equations (F6) and (F7) are then

$$\langle \delta' | \nu, \{v\} \rangle = \psi'_I C_{IJ}^{-1} v_J + \frac{\sigma - \psi'_I C_{IJ}^{-1} \psi_J}{\sqrt{\sigma^2 - \psi_I C_{IJ}^{-1} \psi_J}} \nu_v, \tag{F10}$$

$$\mathrm{Var}\left(\delta' | \nu, \{v\}\right) = \langle \delta'^2 \rangle - \psi'_I C_{IJ}^{-1} \psi'_J - \frac{(\sigma - \psi'_I C_{IJ}^{-1} \psi_J)^2}{\sigma^2 - \psi_I C_{IJ}^{-1} \psi_J}, \tag{F11}$$

[where ν_v is given by equation (F5)] from which one can evaluate equations (15) and (16), after setting $\delta = \delta_c$. Since $\langle \delta' | \nu_c \rangle = \nu_c$ and $\mathrm{Var}\left(\delta' | \nu_c\right) = 1/\Gamma^2$, equation (11) is recovered in the unconstrained case. For later convenience, let us also note that the conditional probability of ν and ν' given the constraint $\{v\}$ is

$$p_G(\nu, \nu' | \{v\}) = \sigma \frac{p_G(\nu_v)\, p_G(\delta' - \langle \delta' | \nu_c, \{v\} \rangle)}{\sqrt{1 - \psi_I C_{IJ}^{-1} \psi_J / \sigma^2}}, \tag{F12}$$

since by construction ν_v and $\delta' - \langle \delta' | \nu_c, \{v\} \rangle \propto \nu'_v$ are independent.

F2 Conditioning to the saddle

Equation (F8) and its derivative guarantee that conditioning on the values of $\mathcal{S}$ (that is, fixing the geometry of the saddle) returns

$$\langle \delta | \mathcal{S} \rangle = \xi \cdot \mathcal{S}, \quad \mathrm{Var}\left(\delta | \mathcal{S}\right) = \sigma^2 - \xi^2,$$

$$\langle \delta' | \mathcal{S} \rangle = \xi' \cdot \mathcal{S}, \quad \mathrm{Var}\left(\delta' | \mathcal{S}\right) = \langle \delta'^2 \rangle - \xi'^2,$$

$$\langle \nu_{1/2} | \mathcal{S} \rangle = \xi_{1/2} \cdot \mathcal{S}, \quad \mathrm{Var}\left(\nu_{1/2} | \mathcal{S}\right) = 1 - \xi_{1/2}^2. \tag{F13}$$

To make the equations less cluttered, here and in the following, scalar products of these vectors are denoted with a dot, rather than in *Einstein*'s notation. Equation (F13) effectively amounts to replacing in all unconditional expressions

$$\delta \rightarrow \delta - \xi \cdot \mathcal{S},$$

$$\delta' \rightarrow \delta' - \xi' \cdot \mathcal{S},$$

$$\nu_{1/2} \rightarrow \nu_{1/2} - \xi_{1/2} \cdot \mathcal{S}, \tag{F14}$$

reducing the problem to three zero-mean variables that no longer correlate with $\mathcal{S}$ (but still do with each other!). The covariance of δ, δ' and $\nu_{1/2}$ at fixed $\mathcal{S}$ reads

$$\mathrm{Cov}\left(\delta, \delta' | \mathcal{S}\right) = \sigma - \xi \cdot \xi',$$

$$\mathrm{Cov}\left(\delta, \nu_{1/2} | \mathcal{S}\right) = \omega - \xi \cdot \xi_{1/2},$$

$$\mathrm{Cov}\left(\delta', \nu_{1/2} | \mathcal{S}\right) = \omega' - \xi' \cdot \xi_{1/2}, \tag{F15}$$

with ω and its derivative ω' given by equations (E14) and (E15). The first equation in (F15) is one half the derivative of $\mathrm{Var}\left(\delta | \mathcal{S}\right)$ with respect to σ from equation (F13), consistently with taking the conditional expectation value of the relation $\delta \delta' = (1/2)\mathrm{d}\delta^2 / \mathrm{d}\sigma$. The third is the derivative of the second, since $\xi_{1/2}$ depends on $\sigma_{1/2}$ and not on σ (the relation between the two scales arising since $\sigma_{1/2} = \sigma(M/2)$ should be imposed after taking the derivative).

F3 Slope given height at distance *r* from the saddle

The saddle point being fixed, it can now be assumed that the excursion set point is at the critical overdensity $\nu = \nu_c$. The conditional mean and variance of the slope are then

$$\langle \delta' | \nu_c, \mathcal{S} \rangle = \langle \delta' | \mathcal{S} \rangle + \frac{\mathrm{Cov}\left(\delta', \delta | \mathcal{S}\right)}{\mathrm{Var}\left(\delta | \mathcal{S}\right)}\left(\delta_c - \langle \delta | \mathcal{S} \rangle\right)$$

$$= \xi' \cdot \mathcal{S} + \frac{\sigma - \xi \cdot \xi'}{\sigma^2 - \xi^2}\left(\delta_c - \xi \cdot \mathcal{S}\right), \tag{F16}$$

after using equations (F13) and (F15), and

$$\mathrm{Var}\left(\delta' | \nu_c, \mathcal{S}\right) = \mathrm{Var}\left(\delta' | \mathcal{S}\right) - \frac{\mathrm{Cov}\left(\delta', \nu | \mathcal{S}\right)^2}{\mathrm{Var}\left(\nu | \mathcal{S}\right)},$$

$$= \langle \delta'^2 \rangle - \xi'^2 - \frac{(\sigma - \xi \cdot \xi')^2}{\sigma^2 - \xi^2}, \tag{F17}$$

respectively. This result is equivalent to decorrelating the effective variables $\delta - \xi \cdot \mathcal{S}$ and $\delta' - \xi' \cdot \mathcal{S}$ introduced in equation (F14), whose covariance is in fact $\sigma - \xi' \cdot \xi$.

Equation (F16) contains an angle-dependent offset $\hat{r}_i \bar{q}_{ij} \hat{r}_j \xi_{20}$ and a density dependent one $\xi_{00} \nu_{\mathcal{S}}$, entering through $\mathcal{S}$. On the contrary, the conditional variance does not depend on the angle nor the height of the saddle. At large distance from the saddle, when $\xi = \xi' = 0$, equations (F16) and (F17) tend as expected to the unconditional mean ν_c and variance $1/\Gamma^2 = \langle \delta'^2 \rangle - 1$.

From equations (F16) and (F17), one can compute the effective upcrossing parameters presented in the main text

$$\mu_{\mathcal{S}}(\mathbf{r}) = \xi' \cdot \mathcal{S} + \frac{\sigma - \xi' \cdot \xi}{\sigma^2 - \xi^2}(\delta_c - \xi \cdot \mathcal{S}), \tag{F18}$$

$$X_{\mathcal{S}}(\mathbf{r}) = \mu_{\mathrm{S}}(\mathbf{r}) / \sqrt{\mathrm{Var}\left(\delta' | \nu_c, \mathcal{S}\right)}. \tag{F19}$$

F4 Upcrossing at σ with given formation time but no saddle

Recalling that $\omega = \langle \delta \delta_{1/2} \rangle / \sigma_{1/2}$ and $\omega' = \langle \delta' \delta_{1/2} \rangle / \sigma_{1/2}$, as defined by equations (E14) and (E15), the conditional statistics of δ and δ' given that $\nu_{1/2} = \nu_f$ are

$$\langle \delta | \nu_f \rangle = \omega \nu_f, \quad \mathrm{Var}\left(\delta | \nu_f\right) = \sigma^2 - \omega^2,$$

$$\langle \delta' | \nu_f \rangle = \omega' \nu_f, \quad \mathrm{Var}\left(\delta' | \nu_f\right) = \langle \delta'^2 \rangle - \omega'^2,$$

$$\mathrm{Cov}\left(\delta, \delta' | \nu_f\right) = \sigma - \omega \omega'. \tag{F20}$$

Hence, the conditional mean and variance of δ' given $\nu_c = \delta_c / \sigma$ and ν_f are

$$\langle \delta' | \nu_c, \nu_f \rangle = \omega' \nu_f + \frac{\sigma - \omega' \omega}{\sigma^2 - \omega^2}\left(\delta_c - \omega \nu_f\right), \tag{F21}$$

$$\mathrm{Var}\left(\delta'|\nu_c,\nu_f\right) = \langle \delta'^2\rangle - \omega'^2 - \frac{(\sigma - \omega'\omega)^2}{\sigma^2 - \omega^2}\,. \tag{F22}$$

which is equivalent to decorrelating the zero-mean effective variables $\delta - \omega\nu_f$ and $\delta' - \omega'\nu_f$, whose covariance is $\sigma - \omega'\omega$. From equations (F21) and (F22), one can compute the parameters of the effective upcrossing problem

$$\mu_f(D_f) = \langle \delta'|\nu_c, \nu_f\rangle\,, \tag{F23}$$

$$X_f(D_f) = \mu_f(D_f)/\sqrt{\mathrm{Var}\left(\delta'|\nu_c, \nu_f\right)}\,, \tag{F24}$$

introduced in Section 2.2.

F5 Upcrossing at σ given formation time and the saddle

Similarly, thanks to equations (F13) and (F15), the mean and covariance of $p_G(\nu|\nu_f, \mathcal{S})$ are

$$\langle \delta|\nu_f, \mathcal{S}\rangle = \langle \delta|\mathcal{S}\rangle + \frac{\mathrm{Cov}\left(\delta, \nu_{1/2}|\mathcal{S}\right)}{\mathrm{Var}\left(\nu_{1/2}|\mathcal{S}\right)}\left(\nu_f - \langle \nu_{1/2}|\mathcal{S}\rangle\right)\,,$$
$$= \xi \cdot \mathcal{S} + \Omega\,\nu_{f,\mathcal{S}}\,, \tag{F25}$$

$$\mathrm{Var}\left(\delta|\nu_c, \mathcal{S}\right) = \mathrm{Var}\left(\delta|\mathcal{S}\right) - \frac{\mathrm{Cov}\left(\delta, \nu_{1/2}|\mathcal{S}\right)^2}{\mathrm{Var}\left(\nu_{1/2}|\mathcal{S}\right)}\,,$$
$$= \sigma^2 - \xi^2 - \Omega^2\,, \tag{F26}$$

where [recalling that ξ has the dimensions of δ but $\xi_{1/2}$ has those of ν, see equation (F3)]

$$\nu_{f,\mathcal{S}} \equiv \frac{(\nu_f - \xi_{1/2} \cdot \mathcal{S})}{\sqrt{1 - \xi_{1/2}^2}}\,, \quad \Omega \equiv \frac{\omega - \xi \cdot \xi_{1/2}}{\sqrt{1 - \xi_{1/2}^2}}\,. \tag{F27}$$

As discussed in Appendix F1, the statistics of $p_G(\delta'|\nu_c, \nu_f, \mathcal{S})$ can be derived from those of $p_G(\delta|\nu_f, \mathcal{S})$ as follows:

$$\langle \delta'|\nu_c, \nu_f, \mathcal{S}\rangle = \langle \delta|\nu_f, \mathcal{S}\rangle' + \frac{\mathrm{Var}\left(\delta|\nu_f, \mathcal{S}\right)'}{2\mathrm{Var}\left(\delta|\nu_f, \mathcal{S}\right)}(\delta_c - \langle \delta|\nu_f, \mathcal{S}\rangle) \tag{F28}$$

thanks to the relations $\langle \delta|\nu_f, \mathcal{S}\rangle' = \langle \delta'|\nu_f, \mathcal{S}\rangle$ and $\mathrm{Var}\left(\delta|\nu_f, \mathcal{S}\right)' = 2\mathrm{Cov}\left(\delta\delta'|\nu_f, \mathcal{S}\right)$, and

$$\mathrm{Var}\left(\delta'|\nu_c, \nu_f, \mathcal{S}\right) = \mathrm{Var}\left(\delta'|\nu_f, \mathcal{S}\right) - \frac{[\mathrm{Var}\left(\delta|\nu_f, \mathcal{S}\right)']^2}{4\mathrm{Var}\left(\delta|\nu_f, \mathcal{S}\right)}\,. \tag{F29}$$

Hence, taking derivatives of equations (F25) and (F26) give

$$\langle \delta'|\nu_c, \nu_f, \mathcal{S}\rangle = \xi' \cdot \mathcal{S} + \Omega'\nu_{f,\mathcal{S}}$$
$$+ \frac{\sigma - \xi' \cdot \xi - \Omega'\Omega}{\sigma^2 - \xi^2 - \Omega^2}(\delta_c - \xi \cdot \mathcal{S} - \Omega\,\nu_{f,\mathcal{S}})\,, \tag{F30}$$

and

$$\mathrm{Var}\left(\delta'|\nu_c, \nu_f, \mathcal{S}\right) = \langle \delta'^2\rangle - \xi'^2 - \Omega'^2$$
$$- \frac{(\sigma - \xi' \cdot \xi - \Omega'\Omega)^2}{\sigma^2 - \xi^2 - \Omega^2}\,, \tag{F31}$$

where

$$\Omega' = \frac{\omega' - \xi' \cdot \xi_{1/2}}{\sqrt{1 - \xi_{1/2}^2}}\,, \tag{F32}$$

which can finally be used to compute the effective slope parameters

$$\mu_{f,\mathcal{S}}(D_f, \boldsymbol{r}) = \langle \delta'|\nu_c, \nu_f, \mathcal{S}\rangle\,, \tag{F33}$$

$$X_{f,\mathcal{S}}(D_f, \boldsymbol{r}) = \mu_{f,\mathcal{S}}(D_f, \boldsymbol{r})/\sqrt{\mathrm{Var}\left(\delta'|\nu_c, \nu_f, \mathcal{S}\right)}\,. \tag{F34}$$

APPENDIX G: GENERIC AND MOVING BARRIER

The results presented hereby hold for a constant barrier, however, one can easily recover the results for a non-constant one – where the upcrossing conditions becomes $\delta_c > \delta'_c$ – by replacing μ_v by $\mu_v - \delta'_c$ in the general formula of equations (15) and (16), yielding

$$\mu_v \equiv \langle \delta'|\nu_c, \{v\}\rangle - \delta'_c\,, \tag{G1}$$

and by taking into account contributions from δ'_c in ν'_c

$$\nu'_c = \frac{\delta'_c}{\sigma} - \frac{\delta_c}{\sigma^2}\,, \tag{G2}$$

and in the definition of accretion rate

$$\alpha = \frac{\delta_c}{\sigma(\delta' - \delta'_c)} \tag{G3}$$

in equation (19). In practical terms, dealing with a moving barrier simply amounts to replacing

$$\mu \to \langle \delta'|\nu_c\rangle - \delta'_c\,, \tag{G4}$$

$$\mu_f \to \langle \delta'|\nu_c, \nu_f\rangle - \delta'_c\,, \tag{G5}$$

$$\mu_{\mathcal{S}} \to \langle \delta'|\nu_c, \mathcal{S}\rangle - \delta'_c\,, \tag{G6}$$

$$\mu_{f,\mathcal{S}} \to \langle \delta'|\nu_c, \nu_f, \mathcal{S}\rangle - \delta'_c\,, \tag{G7}$$

in equations (12), (33), (50), and (67), which automatically affects also the corresponding X, X_f, $X_{\mathcal{S}}$, and $X_{f,\mathcal{S}}$, as well as Y_α and $Y_{\alpha,\mathcal{S}}$ in equations (24) and (60).

For instance, for a barrier of the type $\delta_c + \beta\sigma\bar{q}_{ij,R}\bar{q}_{ij,R}$ (Castorina et al. 2016), where $\bar{q}_{ij,R}$ is the traceless tidal tensor smoothed on scale R, and β is some constant, one would use

$$\delta'_c \to \beta(\bar{q}_{ij,R}\bar{q}_{ij,R} + 2\sigma\bar{q}'_{ij,R}\bar{q}_{ij,R})\,. \tag{G8}$$

More generally, barriers should involve $\{I_n\}$, the rotationally invariants of $\bar{q}_{ij,R}$ defined in Appendix D.

APPENDIX H: IMPLIED GALACTIC COLOURS

Let us in closing attempt to convert the position-dependent accretion rates, computed in the main text, in terms of colour modulo some reasonable assumption on the respective role of AGN and how star formation proceeds at low and high redshifts. Galaxy colours are proportional to the amount of recent star formation, which in turn is driven by the recently accreted gas from cosmic infall. One complication comes from the impact of feedback on heating the gas to be accreted on to galaxies. Cosmological hydrodynamical simulations, which include the feedback of supermassive black holes, suggest that, at intermediate and low redshift, mass accretion through mergers triggers AGN feedback in massive galaxies. This in turn heats up the circumgalactic medium and prevents subsequent smooth gas accretion from feeding central galaxies efficiently (e.g. Dubois et al. 2010), quenching star formation and reddening massive galaxies (hosted in haloes with mass of $10^{12}\,\mathrm{M}_\odot\,h^{-1}$ or more). Conversely, at higher redshift, cold flows are less impacted by galactic feedback and reach the centre of dark haloes unimpaired, so that matter infall translates into bluer galaxies (though it has been suggested that in massive haloes, the disruption of cold flows can be significant, Dubois et al. 2013). Fig. H1 sketches these ideas, while distinguishing low- and high-mass haloes. As argued in the main text, this
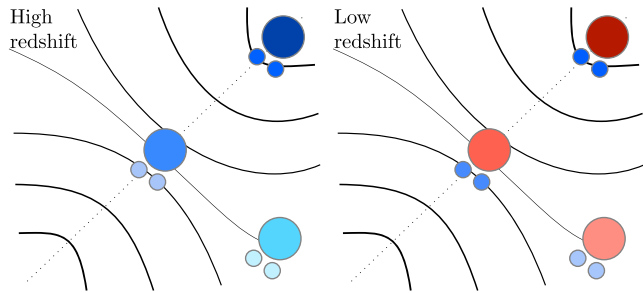
Figure H1. Scheme of the intensity of expected colour/SFR at different location near a filament-type saddle for different final halo mass. The displayed colour encodes galactic colour (or equivalently sSFR from high blue to low red). Massive galaxies in the filament (respectively, nodes) are expected to accrete more cold baryonic matter at high redshift and be bluer than less massive ones and than their counterparts in voids (respectively, filaments). At lower redshifts, AGN feedback is expected to quench cold gas accretion, thus reddening the massive ones – they are more likely to be central ones. The impact on lower mass satellite galaxies may also depend on the efficiency of processes such as starvation or ram-pressure stripping.

scenario remains speculative, if only because the impact of AGN feedback is still a fairly debated topic. For instance ram-pressure stripping on satellites plunging into clusters is known to induce reddening, but its efficiency within filaments is unclear. Fig. 16 encodes the robust result of the present investigation.

This paper has been typeset from a TEX/LATEX file prepared by the author.
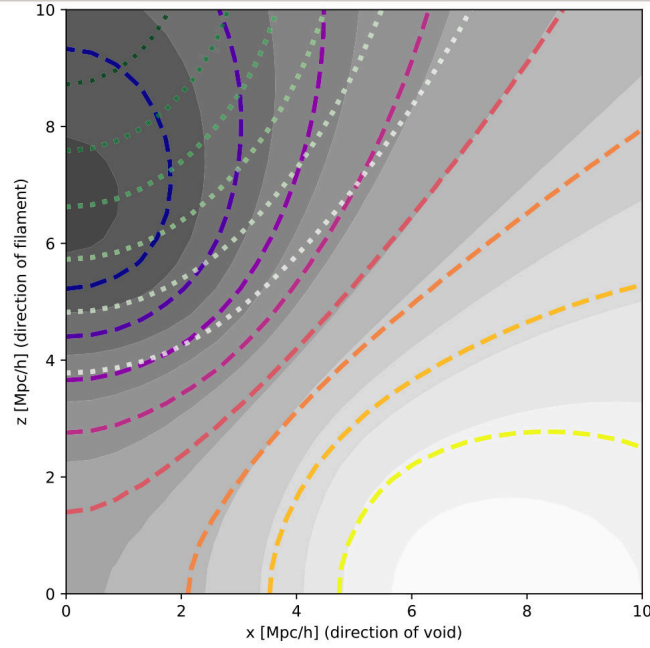
Figure 3.3.1: Isocontours of constant typical redshift $z = 0$ mean density (filled contours), mass (dotted lines) and accretion rate (dashed lines) in the frame of a filament (along the Oz axis) in Lagrangian space (initial conditions) from low (light colours) to high values (dark colours). The saddle is at coordinate $(0, 0)$ while the induced peak and void are at coordinates $(0, \pm 7)$ and $(\pm 8, 0)\mathrm{Mpc}/h$, respectively. The gradients of the three fields, are not parallel (the contours cross). The choice of scale sets the units on the x- and z-axis (chosen here to be $5\,\mathrm{Mpc}/h$, while the mass and accretion rates are computed for a local smoothing of $0.5\,\mathrm{Mpc}/h$). At lower redshift/smaller scales, one expects the non-linear convergence of the flow towards the filament to bring those contours together, as shown on figure 3.3.2.

3.3 Conclusion

Let me complement the conclusions of this article in the context of this dissertation and of subsequent works.

We have shown that the excursion set theory can be extended to take into account anisotropic effects induced by the cosmic web. This can be done by constraining the statistics entering the excursion to the presence of a proto-filament at a given location, which in turn spatially modulates the mean and the variance of the field, resulting in a biasing of the excursion. From this, one can show that different quantities derived from the properties of the excursion under the anisotropic constrain, such as the halo mass function and the accretion rate and formation time at fixed final mass, become distinct functions of the local mean and variance of the field, so that their modulation by the cosmic web is different.

The differential effect induced by the cosmic web can be illustrated by computing the isocontours of the different assembly variables, which can be shown to explicitly cross, as illustrated on figure 3.3.1. The figure shows that, for example, isodensity contours cross isocontours of accretion rate at fixed final mass. As a consequence, while most of the spatial variation of the accretion rate can be attributed to the modulation due to the local density surrounding a given halo, part of the variation is due to the tidal effect of the large-scale filamentary structure. The same conclusion can be drawn for the formation time and lead to the conclusion that the structure of the cosmic web, as encoded by the filament-type saddle point, drives part of the assembly bias signal. More
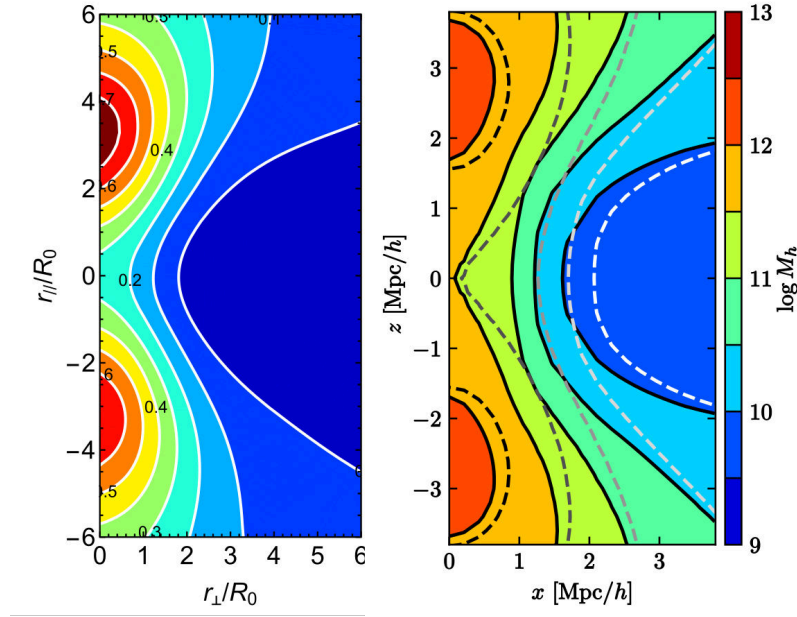
Figure 3.3.2: Typical mass measured in the Horizon-AGN simulation (left panel, from Kraljic et al., 2019) and predictions from constrained excursion set theory (right) along the axis of filaments (vertical axis). Compared to the prediction without Zel'dovich boost (dashed lines), the isomass contour lines after the boost are compressed in the direction of filament.

massive halos are found in the filament compared to the surrounding void and wall, while the most massive halos are found in nodes of the cosmic web. At fixed final mass, halos forming close to the saddle are stalled and formed early, whereas those forming close to the nodes formed later and accrete more. The same hierarchy is found between wall and filaments. Similar trends have been measured for galaxies in the GAMA spectroscopic survey (Driver et al., 2011) and the Horizon-AGN simulation in a paper I contributed to (Kraljic et al., 2018, see appendix B.1). Namely, it was shown that galaxies in filaments are more massive than their wall counterparts. In addition, galaxies also segregate by colour, with an excess of red passive galaxies close to the filament core than in the wall.

In a follow-up work (Kraljic et al., 2019, see appendix B.2), we measured the properties of virtual galaxies in the Horizon-AGN simulation in the frame of the cosmic web, reproducing the same maps as Musso, Cadiou et al., 2018, figure 8. In this work, I have shown that the results from the constrained excursion set theory can be qualitatively reproduced if one takes into account the mean Zel'dovich displacement (following the idea of the Bond and Myers, 1996), which has the effect of squeezing the isocontour lines in the direction perpendicular to the filament and stretch them in the direction parallel to the filament. This is for example illustrated on figure 3.3.2, which shows typical mass isocontours in the Horizon-AGN simulation (left panel) and the prediction from the constrained excursion set theory (right panel).

The constrained excursion set theory presented in this chapter enabled us to study the impact of the cosmic web on the formation of dark matter halos, yet more work is required to understand its impact on galaxy formation, as was presented in appendices B.1 and B.2. In chapter 4, I propose a new theoretical model that can be used to quantify the effect of the cosmic web on galaxy formation, by looking at special events that drive the evolution of galaxies, namely halo mergers and filament disconnections. These events impact how gas and baryons are acquired by forming galaxies, which then impact galaxy formation, and in particular disk formation. This is further

studied chapters 5 and 6, where I study the transport of angular momentum from the large scales down to the disk.

4. Forecasting special events driving the assembly of dark halos

Outline

4.1   Introduction

To what extent can today's properties of galaxies be predicted from the initial Gaussian random field from which they emerge? Within the paradigm of the spherical collapse, one can draw a relationship between the time of collapse of a given proto-halo given its over density, and between its mass and the scale at which its initial patch must be smoothed to pass a given threshold. As the halo grows in mass, it will explore larger and larger radii. In the extended Press-Schechter theory, this excursion is usually described in terms of the mean overdensity found at increasing radii, recovering the result that large overdensities collapse earlier in cosmic time and can be further refined to take into account non-spherical collapse (e.g. R. K. Sheth et al., 2001; Hahn et al., 2009), or the effect of gravitational clustering (Bond and Myers, 1996). In this sense, the fate of a given region is encoded in its initial conditions and is captured by the multi-scale properties of the corresponding Gaussian random field. Most of the aforementioned works have typically described proto-halos as peaks in the primordial field, effectively compressing the continuous density field into a set of points (peaks). In a more general way, the topology of the field can be described by the set of its critical points (peaks, saddle points and voids). In Hanami, 2001 it was suggested that the drift of these critical points, which draws the so-called skeleton tree, bears physical meaning, as it captures the variation of this topology with scale, hence cosmic time. In Manrique and Salvador-Sole, 1995; Manrique and Salvador-Sole, 1995; Hanami, 2001 the focus was on the coalescence of filament saddles with maxima which the authors called slopping saddles (as they are vanishing saddle points on the slope of peaks), and are proxy for halo merging events.

More generally here I will consider the coalescence of minima with wall-saddles and wall-saddles with filament-saddles corresponding respectively to the disappearance of a wall and a filament. It is the sequence and geometry of these special events in the Lagrangian patch and its vicinity that will later form a halo which will shape the fate of its host galaxy. Indeed, these coalescences impact the geometry of the cosmic web (in particular the filaments) which in turn defines preferred directions along which galaxies are fed with cold gas and acquire their spin. Merger events are also known to play an important role in triggering AGN feedback, which in turn impacts gas inflow and therefore galactic morphology. Hence, I will extend Hanami, 2001 by studying the clustering of these other merger events in the multi-scale landscape. The aim is to provide a compact description of the cosmic web in the initial conditions that is able to capture important events in the life of a galaxy, which includes its merger history, but also the merger
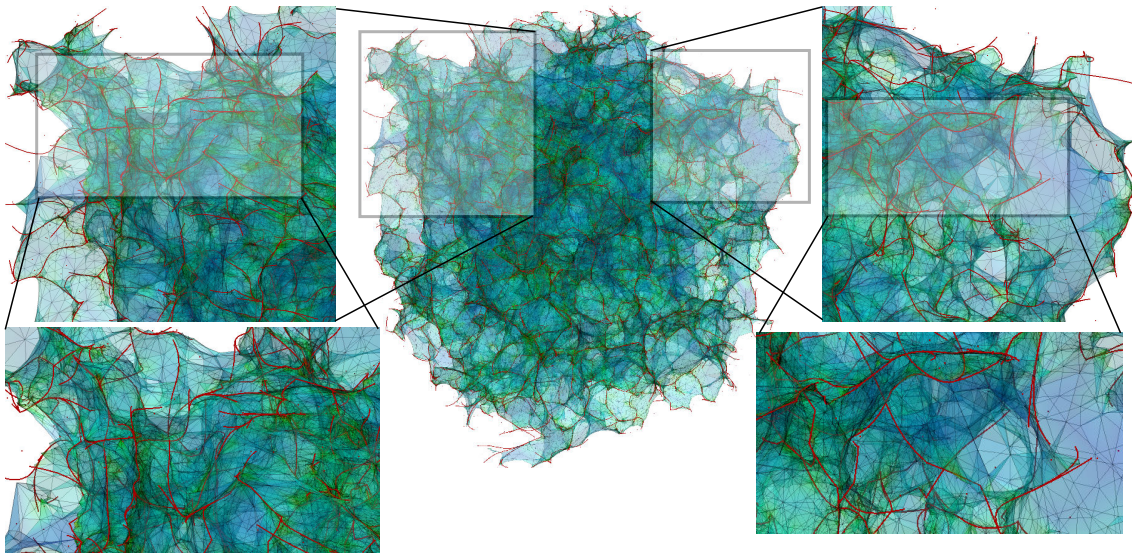
Figure 4.1.1: Snapshot and zooms of a hydrodynamical simulation showing filaments (in red) walls (in shades of blue to green) and peaks (at the nodes of the filamentary network) as traced by DISPERSE (Sousbie et al., 2011). The cosmic evolution of these large scale structures impacts the geometry of infall. As this simulation forms galaxies their properties reflect partially the corresponding tides and the funnelling of cold gas along the filamentary structure. Understanding when and how the topology of this network changes is therefore of great interest in this context.

history of the filaments feeding it and the merger history of the walls feeding its filaments. My motivations are many-fold:

i) Study the generalised history of accretion: what mergers happen when, at what frequency?
ii) Study the relation between different merger events, and their clustering in space and time,
iii) quantify the merger rates in a larger scale filamentary structure to study assembly bias.

In order to achieve these goals, I will present the general theory of the merger events, which I will refer to as "critical events". Section 4.2 provides a mathematical description of these events in the initial conditions and compute their one-point statistics (number counts). Section 4.3 predicts the clustering properties of these special events. Section 4.4 compares the predictions to realisations of Gaussian random fields and validates the theoretical formulas. Section 4.5 presents applications of the theory in the context of galaxy formation. Finally section 4.6 wraps up. Section 4.A presents the counts in arbitrary dimensions and illustrates them in up to 6D. Section 4.B explains how the critical events are measured in random field maps and cubes. Section 4.C presents the joint PDF of a Gaussian random field up to the third derivative of the field. Throughout the chapter, sections where the third form is used (we, us) were done in collaboration with S. Codis and C. Pichon.

4.2 Theory: one-point statistics

Let me consider the overdensity field $\delta = (\rho - \bar{\rho})/\bar{\rho}$ to be a homogeneous and isotropic Gaussian random field of zero mean, described by its power spectrum $P(k)$, as defined in section 2.1.1.3. In this section, I will focus on one-point statistics associated with merger rates. In section 4.2.1, I define the concept of critical events. In section 4.2.2, I present the number counts of critical events, counted together and by type (peak, filament and wall mergers). In section 4.2.3, I present the number counts as a function of the events' height. Section 4.2.4 sketches the corresponding

theory in two-dimensions, while section 4.2.5 presents its extension to non-Gaussian fields.

4.2.1   Critical events definition

When studying the time evolution of the density field, the spherical collapse model has shown that one can establish a mapping between collapse time and overdensity – high overdensity regions collapse earlier in the history of the Universe than underdense ones. At the same time, larger overdensities enclose more mass and will hence give birth to more massive structures. These relations mathematically read

$$\nu_{\rm c}(R) = \frac{\delta_{\rm c}}{\sigma(R)D(z)}, \quad M = \frac{4\pi}{3}\bar{\rho}R^3, \tag{4.1}$$

where R is the smoothing scale of the Top-Hat filter, $\sigma(R)$ is the variance of the field at that scale, $D(z)$ is the linear matter growth function at redshift z (see section 2.1.2.1), $\delta_{\rm c} = 1.69$ is the spherical collapse critical overdensity (see section 2.1.2.2) and $\bar{\rho}$ is the mean matter density of the Universe. Here I have introduced the density contrast $\nu \equiv \delta/\sigma(R)$, which is a zero-mean unit-variance Gaussian random field. The spherical collapse threshold can also be adapted to study the formation of voids (R. K. Sheth and van de Weygaert, 2004; Jennings et al., 2013) with $\delta_{\rm v} = -2.7$. From a theoretical perspective, the action of smoothing the density field δ enables to probe the time-evolution of spherical proto-halos by following the density evolution of peaks as smoothing scale increases. One caveat of using a Top-Hat filtering lays in the fact that the second derivative of the smoothed field has an infinite variance, so that one cannot study the statistics of its peaks and extrema. In the following of the work, I will make use of a Gaussian filtering instead, as it provides smooth fields[1]. In order to match the results of equation (4.1) with a Gaussian filter, one needs to establish a mapping of the smoothing scales between Top-Hat filtering and Gaussian filtering. This is usually achieved by matching the variance of the field $\sigma_{\rm G}(R/\alpha) = \sigma_{\rm TH}(R)$. At scales of a few Mpc/h, the scale ratio is of the order of $\alpha \approx 2.1$ for a ΛCDM power spectrum (see section 2.1.7.2) so that equation (4.1) becomes

$$M = \frac{4\pi}{3}\bar{\rho}(\alpha R)^3 \tag{4.2}$$

for a field smoothed by a Gaussian filter of radius R. This translates the fact that the variance of the field smoothed with a Gaussian filter at scale R is the same as the field smoothed with a Top-Hat filter at scale αR, so that at fixed smoothing scale, one can assign a larger mass to a peak found using Gaussian filtering compared to a region smoothed with a Top-Hat filtering.

Let me now define critical events associated to mergers. These events are defined in smoothing-position space and correspond to mergers of critical *points* (peaks, saddle points and minima). The slopping saddles defined in Hanami, 2001 are particular critical events that correspond to mergers between a peak and a saddle point. In this chapter, I will instead focus on all critical events as they are of interest to study the evolution of the geometry of the cosmic web. The formation and location of critical events is illustrated for a 1D field on figure 4.2.1: critical events are found at the tip of critical point lines and represent the disappearance of a critical point into a critical point of another kind (e.g. a maximum and a minimum in 1D, a maximum and a saddle point in 2 or 3D). They encode locations where the topology of the field is changed by removing a pair of critical points.

Let me emphasize here that critical points are a compact encoding of the proto-structures: each proto-filament has at its centre a filament-type saddle-point, while proto-walls have at their centre a wall-type saddle-point. Using an analogy with a mountainous landscape, one can describe a given mountain range by giving the set of its peaks and passes. In practice, this procedure has

---

[1]In practice, all the derivatives of the field have a well-behaving variance.

compressed the continuous information about the height of the mountains into a discrete set of critical points. A similar approach can be used to describe the skeleton of the cosmic web as a set of its critical events.

Let me illustrate the concept of critical events using the same analogy, the latter being restricted to 2D space, see figures 4.2.2 and 4.2.3. A mountainous landscape is made of peaks analogous to proto-halos. Each pair of neighbouring peaks is linked *via* a pass, analogous to a proto-filamentary structure. Following the ridge from one peak to another one is analogous to following a filamentary structure between two proto-halos. On each downhill side of a pass there are two valleys whose faces are analogous to proto-walls in the cosmic web while their depth (hence their geometry) is described by their lowest point. With the action of time, the mountains will erode until eventually no peak will subsist – this is analogous to the smoothing operation. In the process, a disappearing peak will see its height (the density) decrease with time. If the peak is not prominent enough, it will eventually be smoothed to the point where it no longer is a peak but a shoulder on another peak's slope. Just before the peak disappears, it is still linked to its neighbour *via* a pass. When the peak disappears so does the pass – indeed a pass is always located between *two* peaks ; when one disappears, so does the pass. This particular event is what I define as a critical *event*. It encodes the moment when two critical points (here a peak and a saddle point) annihilate. This can also be interpreted as the moment a peak disappears on the slope of its nearest neighbour – the two peaks merged and the most prominent subsisted. Critical events have hence a dual interpretation: in the initial Lagrangian space, critical points are found at the location where a critical event merges into a critical event of another kind (e.g. a peak with a filament saddle-point). In the Eulerian physical space, critical points spot the merger of two similar structures, for example two halos merging into a single one (squashing the filament in between them).

Since the primordial density field is a 3D field, the density landscape is made of peaks (proto-halos), saddle-points (proto-filaments and proto-walls) and minima (proto-voids). Critical events record the merger of peaks into proto-filaments (PF critical events), of proto-filaments into proto-walls (FW critical events) and of proto-walls into proto-voids (WV critical events).

Using the duality discussed above, they also encode halo mergers (PF critical events), filament mergers (FW critical events) and wall mergers (WV critical events). This is illustrated on figure 4.2.4. PF critical events (top panel) encode the merger of two halos separated by a filament. After the merger, the most prominent peak subsists, while the other proto-halo and the proto-filament have annihilated. FW critical events (centre panel) encode the merger of two filaments separated by a wall. After the merger, the most prominent filament subsists, while the other proto-filament and the proto-wall have annihilated. WV critical events (bottom panel) encode the merger of two walls separated by a void. After the merger, the most prominent wall subsists, while the other proto-wall and the proto-void have annihilated.

4.2.2 3D critical events number counts

In this section, I will present the derivation of the number count of critical events in smoothing-position space in 3D. In section 4.2.2.1, I present how one can express the critical event constraint as a function of the local properties of the field and its derivatives. I then express the condition in the frame of the Hessian of the field in section 4.2.2.2 where it takes a simpler expression. In section 4.2.2.3, I extend the previous formula to distinguish between different critical event types (halo mergers, filament mergers, wall mergers). In the following of the section, I will use the quantities defined in section 2.1.7.1, namely σ_i, γ, $\tilde{\gamma}$ which were defined as

$$\sigma_i^2(R) = \frac{1}{2\pi^2} \int dk \, k^2 P(k) k^{2i} W^2(kR), \quad \gamma = \frac{\sigma_1^2}{\sigma_0 \sigma_2}, \quad \tilde{\gamma} = \frac{\sigma_2^2}{\sigma_1 \sigma_3}. \tag{4.3}$$

Here W is a Gaussian filter $W(x) = \exp\left(-k^2 R^2/2\right)$.
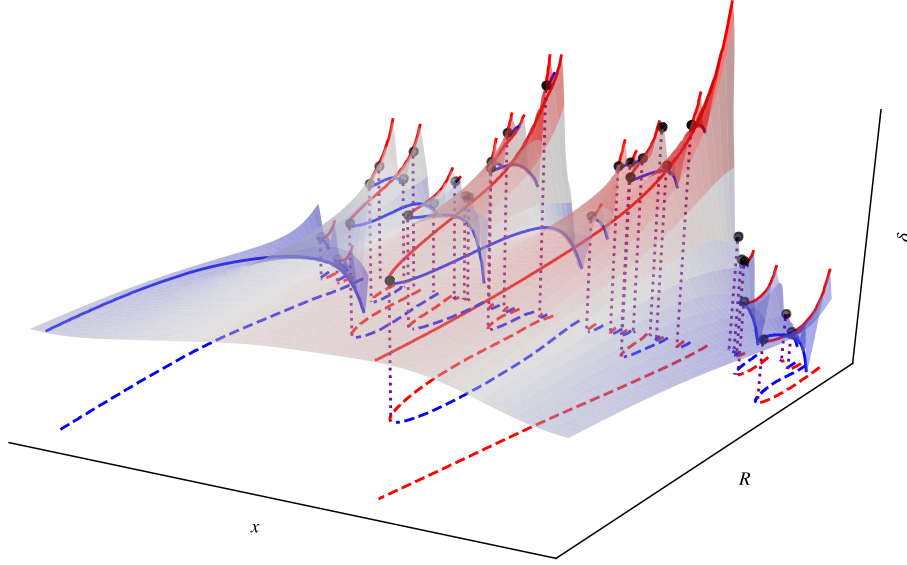
Figure 4.2.1: 2D "landscape" of a 1D field smoothed at a scale R in smoothing-position space. Here R is the smoothing scale, while δ is the density smoothed at the given scale. Solid lines indicate maxima (red) and minima (blue). Critical point lines end at critical events (black dots). The projections of the critical point lines are shown as red and blue dashed lines, while vertical dotted purple lines indicate the projection of critical events to illustrate that critical events are found at the location where two critical points merge.

4.2.2.1   General formulation

Following Hanami, 2001, the number density of critical events in smoothing-position space is given by

$$\frac{\partial^4 \mathcal{N}}{\partial r^3 \partial R} \equiv \langle \delta_{\mathrm{D}}^{(3)}(\boldsymbol{r} - \boldsymbol{r}_0)\, \delta_{\mathrm{D}}(R - R_0) \rangle \,, \tag{4.4}$$

where $\boldsymbol{r}_0$ is the position of a critical event (i.e. a critical point with a degenerate direction) in real space and R_0 its associated smoothing scale. Following the definition of section 4.2.1, critical events are found at the smoothing-position location where two critical points of different types (maximum, saddle points or minimum) merge. The nature of a critical point (occurring where $\boldsymbol{\nabla}\delta = 0$) is characterised by its index, that is to say the number of negative eigenvalues of the density Hessian matrix at this point. Critical *events* can then be defined as critical points for which one of the eigenvalues vanishes, which is also equivalent to having a vanishing determinant. By definition, only critical points whose indices differ by one can merge (peak–filament type saddle point, filament–wall type saddles, wall type saddle–void) so that only one eigenvalue vanishes[2].

Let me therefore first define the determinant of the Hessian $d(\delta) \equiv \det(\boldsymbol{\nabla}\boldsymbol{\nabla}\delta) = \lambda_1 \lambda_2 \lambda_3$, $\lambda_1 \leq \lambda_2 \leq \lambda_3$ being the ordered eigenvalues of the Hessian matrix $\boldsymbol{\nabla}\boldsymbol{\nabla}\delta$. In the following, I will use ∂_R to denote derivatives with respect to scale R. Since critical events are found where $d = 0$ and $\boldsymbol{\nabla}\delta = 0$, let me rewrite equation (4.4) in terms of the properties of the field, using the coordinate transformation from $\boldsymbol{r}, R$ to $\boldsymbol{\nabla}\delta, d$. This involves the 4D Jacobian of the

---

[2]The event where two eigenvalues vanish has a null probability.
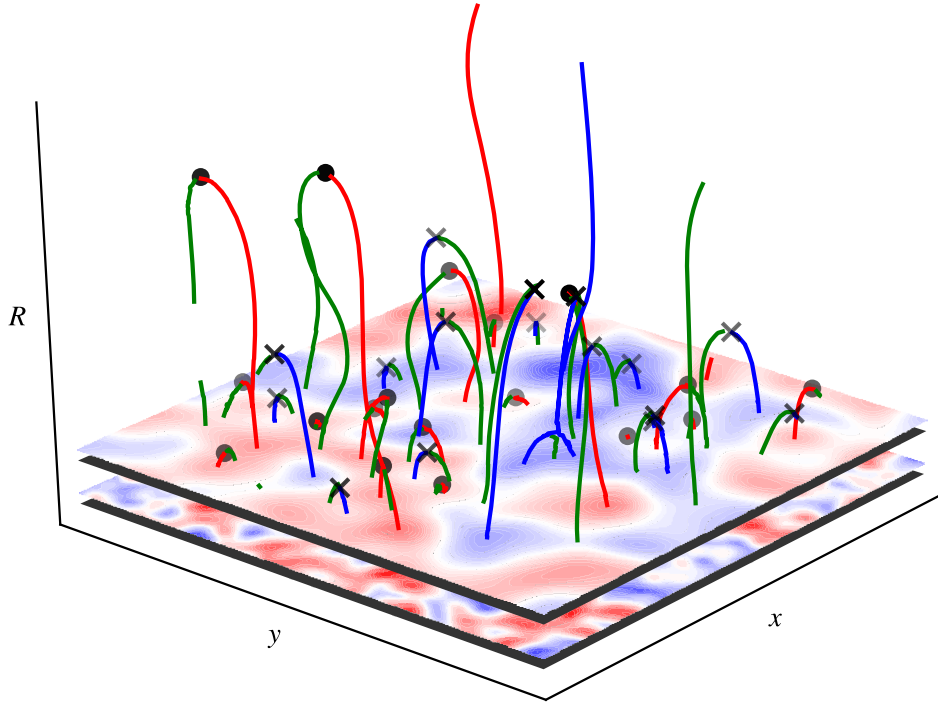
Figure 4.2.2: 3D "landscape" of a 2D field smoothed at a scale R in smoothing-position space. The density field (blue to red map) is smoothed at increasing R. For each scale, the critical points (red lines: peaks, green lines: saddle points, blue lines: minima) are found. At the tip of each branch a critical event is found (●: peak-saddle critical events, ×: saddle-minima). Lines near the boundaries have been hidden for the sake of clarity.

transformation[3]

$$J(d, \boldsymbol{\nabla}\delta) = \begin{vmatrix} \partial_R d & \boldsymbol{\nabla} d \\ \partial_R \boldsymbol{\nabla}\delta^T & \boldsymbol{\nabla}\boldsymbol{\nabla}\delta \end{vmatrix} = \begin{vmatrix} \partial_R d & \boldsymbol{\nabla} d \\ -R\boldsymbol{\nabla}\nabla^2\delta^T & \boldsymbol{\nabla}\boldsymbol{\nabla}\delta \end{vmatrix},$$

(4.5)

using the fact that for a Gaussian filter (see Table 2.2)

$$\partial_R \delta = -R\nabla^2\delta,$$

(4.6)

with ∇^2 the Laplacian operator. The fully covariant formulation of the number density of critical events is then

$$\frac{\partial^4 \mathcal{N}}{\partial r^3 \partial R} = \left\langle |J| \, \delta_{\mathrm{D}}^{(3)}(\boldsymbol{\nabla}\delta) \, \delta_{\mathrm{D}}(d) \right\rangle.$$

(4.7)

The expectation value in equation (4.7) can be evaluated using the joint distribution of the field and its successive derivatives up to third order, $P(x, x_i, x_{ij}, x_{ijk})$ which involves 20 variables, see section 4.C for the PDF for Gaussian random fields. One difficulty in evaluating equation (4.7) spans from $\delta_{\mathrm{D}}(d)$. In practice, it can for instance be dealt with numerically by 'broadening' the Dirac delta function: this method is used for validation and when considering two-point statistics in the section 4.3.1. Alternatively, one can go to the Hessian's eigenframe as described in the next section.

---

[3]Note that the determinant can be developed along the first line or the first column of the Jacobian matrix to find out – as shown by the simplifications in the next section – that the final result in our case does not depend on $\partial_R d$, thanks to the zero determinant constraint $\det \boldsymbol{\nabla}\boldsymbol{\nabla}\delta = 0$.
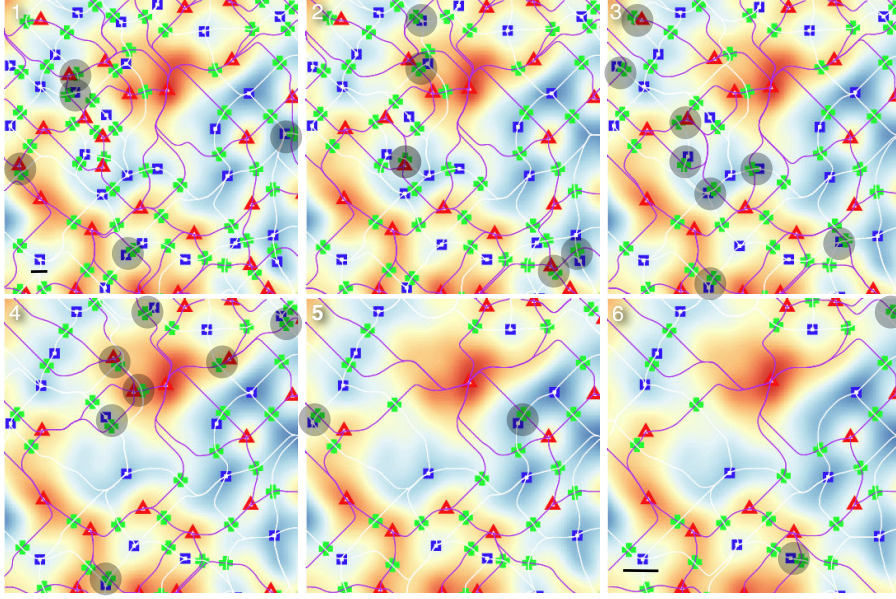
Figure 4.2.3: From left to right and top to bottom, a smoothing sequence of a Gaussian random field, whose density is colour coded from blue to red as a function of height (analogous to the slices shown on figure 4.2.2). The skeleton tracing the ridges is shown in purple, while the anti-skeleton tracing the trough is shown in white. The saddles shown as green crosses lay at the intersection. The Maxima are shown as red triangles while the minima as blue squares. As one smooths the field, these critical points drift towards each other along the skeletons, until they vanish in pairs. The upcoming coalescence are identified with grey circles. Note that as saddle points vanish, the two corresponding skeletons do too. Note also that the direction of coalescence is typically set by the skeleton's just before coalescence. In this two dimensional example, the ratio of peak+saddle to void+saddle event is one. The black segment in the bottom left of the first and last image represents the amount of smoothing. This chapter is concerned with studying the one and two-point statistics of these grey circles. Note that these events are indeed proxy for mergers of the peaks of the underlying field: for instance, between snapshot 3 and 5 the central four peaks have merged into one. Similarly, between 1 and 4 the central four voids have merged into one. I provide an interactive tool to follow such events in 2D and 3D.

4.2.2.2   Expression in the frame of the Hessian

The Jacobian is by construction invariant under rotation, so one can rewrite it in the frame of the eigenvalues of the Hessian (which will be denoted with tildas) without loss of generality. Developing d into $\sigma_2^3 \tilde{x}_{11} \tilde{x}_{22} \tilde{x}_{33}$ and assuming (arbitrarily) that direction 3 is the degenerate one, the Jacobian can be rewritten as follows

$$\frac{J(d,\delta)}{\sigma_1 \sigma_2^4 \sigma_3} = |\tilde{x}_{11}\tilde{x}_{22}| \begin{vmatrix} \partial_R \tilde{x}_{33} & \tilde{x}_{33i} \\ \partial_R \tilde{x}_i & \tilde{x}_{ij} \end{vmatrix}, \tag{4.8}$$

$$= |\tilde{x}_{11}\tilde{x}_{22}| \begin{vmatrix} \partial_R \tilde{x}_{33} & \tilde{x}_{133} & \tilde{x}_{233} & \tilde{x}_{333} \\ \partial_R \tilde{x}_1 & \tilde{x}_{11} & 0 & 0 \\ \partial_R \tilde{x}_2 & 0 & \tilde{x}_{22} & 0 \\ \partial_R \tilde{x}_3 & 0 & 0 & 0 \end{vmatrix}, \tag{4.9}$$

$$= |\tilde{x}_{11}\tilde{x}_{22}|^2 |\partial_R \tilde{x}_3||\tilde{x}_{333}|, \tag{4.10}$$
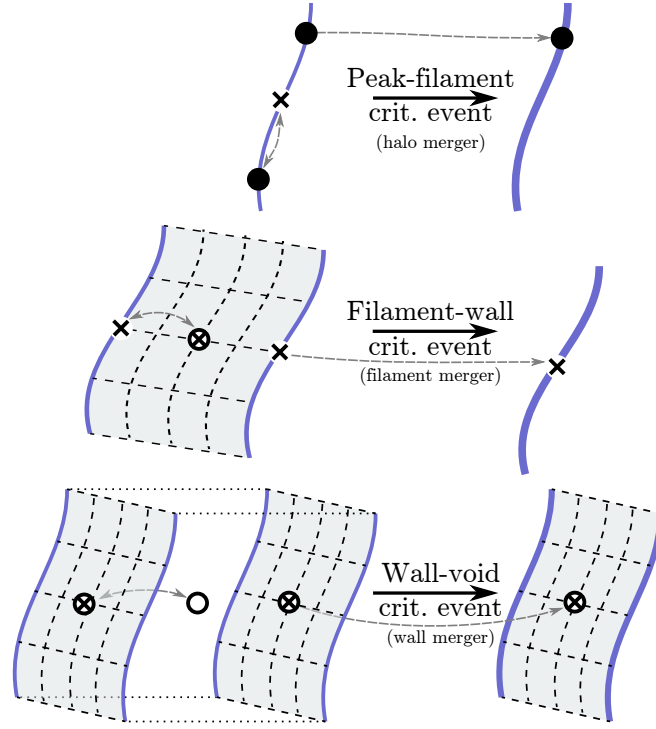
Figure 4.2.4: Illustration of critical events in a 3D random fields and their physical meaning. ● symbols are peaks, × symbols are filament-type saddle points (filament centres), ⊗ symbols are wall-type saddle points (wall centres) and ○ symbols are minima (void centres). *Top:* Peak-filament critical events encode the merger of two halos and the disappearance of their shared filament. After the merger, only one peak subsists and the filament disappears. *Middle:* Filament-wall critical events encode the merger of two filaments and the disappearance of their shared wall. After the merger, only one filament subsists. *Bottom:* Wall-void critical events encode the merger of two walls and the disappearance of their joint void (surrounded by the two walls and the dotted lines). After the merger, only one wall-type saddle-point subsists and the void has disappeared. Halo mergers are encoded by peak-filament critical events, filament mergers. Alternatively, one could have chosen to describe these events as resp. filament, wall and void disappearances.

where the factorisation with $|\tilde{x}_{11}\tilde{x}_{22}|$ along the first line in equation (4.8) is a consequence of $\tilde{x}_{33}$ being zero – which also nulls the last component of equation (4.9). Using equation (4.6) again to re-express the derivative w.r.t. smoothing in terms of the Laplacian of the field, one can rewrite the number density of critical events using the typical scales of equation (2.96) as [4]

$$\frac{\partial n}{\partial R} = \frac{2\pi^2 R}{\tilde{R}^2 R_*^3} \left\langle |\sum_i \tilde{x}_{3ii}| |\tilde{x}_{333}| \delta_D^{(3)}(\tilde{x}_i) |\tilde{x}_{11}\tilde{x}_{22}| \delta_D(\tilde{x}_{33}) \right\rangle , \qquad (4.11)$$

where I introduced $n = \partial^3 \mathcal{N}/\partial r^3$ the volume density of critical events (that does not depend on the spatial location $\boldsymbol{r}$ as the field is assumed to be stationary). Let me stress that the distribution of the fields expressed in the frame of the Hessian matrix differs from the original ones. The statistics of x and x_i and x_{ijk} are left unchanged and I therefore drop the tildes for the field and its first and third derivatives . However, going from cartesian coordinates to the Hessian eigenframe

---

[4]One factor of $|\tilde{x}_{11}\tilde{x}_{22}|$ drops between equation (4.10) and (4.11) because of the Dirac of d in equation (4.7).

modifies the distribution of the second derivatives that were chosen here to be ordered (such that the Doroshkevich formula is recovered)

$$\tilde{P}(\tilde{x}_{11}, \tilde{x}_{22}, \tilde{x}_{33}) = 2\pi^2(\tilde{x}_{33} - \tilde{x}_{22})(\tilde{x}_{22} - \tilde{x}_{11})(\tilde{x}_{33} - \tilde{x}_{11})\times$$
$$P(x_{11} = \tilde{x}_{11}, x_{22} = \tilde{x}_{22}, x_{33} = \tilde{x}_{33}, x_{12} = 0, x_{23} = 0, x_{13} = 0),$$

where $\tilde{x}_{11} < \tilde{x}_{22} < \tilde{x}_{33}$ are distributed according to $\tilde{P}$ and fields in cartesian coordinates follow the distribution $\mathcal{P}$. Note that the factor $2\pi^2$ is due to the integration over the Euler angles. Equation (4.11) therefore introduces a jacobian $2\pi^2|x_{11}x_{22}(x_{11} - x_{22})|$, as x_{33} is null, when going from the Hessian eigenframe to cartesian coordinates and the differential number count of critical events becomes

$$\frac{\partial n}{\partial R} = \frac{2\pi^2 R}{\tilde{R}^2 R_*^3} \left\langle |\sum_i x_{3ii}||x_{333}|\delta_{\mathrm{D}}^{(3)}(x_i)\, |x_{11}x_{22}|^2|x_{11} - x_{22}|\delta_{\mathrm{D}}(x_{33})\, \delta_{\mathrm{D}}^{(3)}(x_{i\neq k}) \right\rangle, \qquad (4.12)$$

where $\delta_{\mathrm{D}}^{(3)}(x_{i\neq k})$ must be understood as a product of Dirac delta functions of all the off-diagonal components of the Hessian matrix. Here R_* and $\tilde{R}$ are the typical inter critical point separation and inter inflection point separation introduced in equation (2.96), section 2.1.7.1. The novelty of equation (4.12) w.r.t. the classical BBKS formula is the weight $|\sum_i x_{3ii}||x_{333}|$ which requires the knowledge of the statistics of the 3rd order derivative of the field. The expectations in equation (4.12) can be evaluated with the joint statistics of the field and its successive derivatives, $P(x_{113}, x_{223}, x_{333}, x_{11}, x_{22})$ which now only involves 5 variables. Interestingly, because the dominant contribution to the expectation value of $\langle |\sum_i x_{3ii}||x_{333}|\rangle$ comes from $\langle x_{333}^2\rangle$ with very good accuracy (at the percent level), equation (4.12) is very well approximated by

$$\frac{\partial n}{\partial R} \approx \frac{2\pi^2 R}{\tilde{R}^2 R_*^3} \left\langle x_{333}^2\delta_{\mathrm{D}}^{(3)}(x_i)\, |x_{11}x_{22}|^2|x_{11} - x_{22}|\delta_{\mathrm{D}}(x_{33})\, \delta_{\mathrm{D}}^{(3)}(x_{i\neq k}) \right\rangle. \qquad (4.13)$$

Note that this equation closely resembles the equation giving the flux of critical lines per unit surface presented in Pogosyan et al., 1998, up to the delta function on the third eigenvalue in the present context. This is in fact expected since I require here that along the filament's direction the curvature should be flat, whereas they marginalised over all possible longitudinal curvature. The similarity reflects the fact that critical points essentially slide along critical lines as one smooths the field, see figure 4.2.3. In some sense the 3D event count can be approximatively recast into a 1D event count along the ridges. The expectation involves the product of the transverse curvatures because the larger those curvatures the larger the flux of such lines per unit transverse surface.

4.2.2.3 Gaussian number density of critical events per type

The aforementioned formalism makes no assumption on the type of the merging critical points. While the coalescence of peaks and filaments (PF critical events, the slopping saddles of Hanami, 2001) are clearly central to the theory of mass assembly, the filament-saddle to wall-saddle (FW critical events) and wall-saddle to minima coalescence (WV critical events) also impact the topology of galactic infall, as they destroy filaments, walls, voids within the surrounding cosmic web.

Let me therefore compute the number density of critical events of each type of mergers ($\mathcal{P} \equiv$ PF, $\mathcal{F} \equiv$ FW and $\mathcal{W} \equiv$ WV). Using the fact that for Gaussian random fields, equation (4.12) can

be split into odd- and even-derivative terms, one can write

$$
\frac{\partial n_j}{\partial R} = \frac{2\pi^2 R}{\tilde{R}^2 R_*^3} \overbrace{\left\langle \left| \sum_i x_{jii} \right| |x_{jjj}| \delta_{\mathrm{D}}^{(3)}(x_i) \right\rangle}^{C_{\mathrm{odd}}} \times
$$

$$
\underbrace{\left\langle \vartheta_{\mathrm{H}}(x_{33} - x_{22})\, \vartheta_{\mathrm{H}}(x_{22} - x_{11})\, \delta_{\mathrm{D}}(x_{jj})\, \delta_{\mathrm{D}}^{(3)}(x_{k \neq l}) \left| \sum_{kl} \frac{\varepsilon^{jkl}}{2} x_{kk}^2 x_{ll}^2 (x_{kk} - x_{ll}) \right| \right\rangle}_{C_{j,\mathrm{even}}}
$$

(4.14)

where ε is the completely antisymmetric Levi-Civita tensor, ϑ the Heaviside function, and $j = 1, 2, 3$ for peak ($\mathcal{P}$), filament ($\mathcal{F}$) and wall ($\mathcal{W}$) mergers respectively. Note that equation (4.14) for a given value of j is essentially the same as equation (4.12), modulo a choice of null eigenvalue and the requirement that the eigenvalues are sorted. In 3D, C_{odd} and $C_{j,\mathrm{even}}$ have analytical expressions given by

$$
C_{2,\mathrm{even}} = \langle \lambda_1 \lambda_3 \delta_{\mathrm{D}}(\lambda_2) \rangle = \frac{2}{\sqrt{15\pi}},
$$

$$
C_{1,\mathrm{even}} = C_{3,\mathrm{even}} = \langle \lambda_1 \lambda_2 \delta_{\mathrm{D}}(\lambda_3) \rangle = \frac{29 - 6\sqrt{6}}{18\sqrt{10\pi}},
$$

(4.15)

and

$$
C_{\mathrm{odd}} = \frac{\sqrt{27}(1 - \tilde{\gamma}^2)}{\sqrt{50\pi^5}} \left(\frac{2}{\sqrt{21(1 - \tilde{\gamma}^2)}} + \tan^{-1} \frac{\sqrt{21(1 - \tilde{\gamma}^2)}}{2} \right),
$$

(4.16)

which can also be computed in arbitrary dimensions as shown in section 4.A. From this I can compute the ratio of peak to filament mergers $r_{\mathcal{P}/\mathcal{F}} = C_{2,\mathrm{even}}/C_{1,\mathrm{even}}$. Interestingly, the event ratio is independent of the spectral index of the field and is given by

$$
r_{\mathcal{P}/\mathcal{F}} = \frac{24\sqrt{3}}{29\sqrt{2} - 12\sqrt{3}} \approx 2.05508 \approx \frac{37}{18},
$$

(4.17)

which is nothing but the ratio between the mean number of wall-type saddles and peaks minus 1, a relationship which is valid in arbitrary dimension, as shown in section 4.A.4. This equation shows that there are twice more filament disappearing in filament merger events ($\mathcal{F}$ events) than in halo merger events ($\mathcal{P}$ events). Similarly, I can compute $r_{\mathcal{F}/\mathcal{W}}$ to deduce that there are twice more walls disappearing due to filament mergers ($\mathcal{F}$ events) than due to void mergers ($\mathcal{W}$ events). Section 4.A also presents these ratios in dimension 4 to 6.

4.2.3 3D differential event counts of a given height

Introducing $\delta_{\mathrm{D}}(x - \nu)$ in the expectation of equation (4.14) allows me to write the density of critical events as a function of height, hence make the distinction between mergers of important critical points and less significant ones.

For Gaussian random fields, the field only correlates with its even derivatives (second in this case). Imposing the height of the critical events considered here therefore only modifies the term $C_{j,\mathrm{even}}$ while C_{odd} is left unchanged, following

$$
C_{j,\mathrm{even}}(\nu) = \left\langle \vartheta_{\mathrm{H}}(x_{33} - x_{22})\, \vartheta_{\mathrm{H}}(x_{22} - x_{11})\, \delta_{\mathrm{D}}(x_{jj}) \delta_{\mathrm{D}}^{(3)}(x_{k \neq l})\, \delta_{\mathrm{D}}(x - \nu) \right.
$$

(4.18)

$$
\left. \times \left| \sum_{kl} \frac{\varepsilon^{jkl}}{2} x_{kk}^2 x_{ll}^2 (x_{kk} - x_{ll}) \right| \right\rangle.
$$

Interestingly, $C_{j,\text{even}}(\nu)$ appears to have an analytical expression once rotational invariants are used to evaluate the expectations. Following the formalism described first in Pogosyan et al., 2009b, we introduce the variables

$$J_1 = I_1, \quad J_2 = I_1^2 - 3I_2, \tag{4.19}$$

$$J_3 = \frac{27}{2}I_3 - \frac{9}{2}I_1 I_2 + I_1^3, \quad \zeta = \frac{x + \gamma J_1}{\sqrt{1 - \gamma^2}}, \tag{4.20}$$

that are linear combinations of the density field x and rotational invariants of its second derivatives namely the trace $I_1 = \text{tr}\,\mathbf{H} = \lambda_1 + \lambda_2 + \lambda_3$, minor $I_2 = 1/2((\text{tr}\,\mathbf{H})^2 - \text{tr}\,\mathbf{H} \cdot \mathbf{H}) = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1$ and determinant $I_3 = \det \mathbf{H} = \lambda_1\lambda_2\lambda_3$ of the Hessian matrix $\mathbf{H} = (x_{ij})$. The distribution of these variables is given by

$$P(\zeta, J_1, J_2, J_3) = \frac{25\sqrt{10\pi}}{24\pi^2} \exp\left(-\frac{1}{2}\zeta^2 - \frac{1}{2}J_1^2 - \frac{5}{2}J_2\right), \tag{4.21}$$

where J_3 is uniformly distributed between $-J_2^{3/2}$ and $J_2^{3/2}$ and J_2 is positive. Using these rotational invariants, one can rewrite equation (4.18) for each type of critical event

$$C_{1,\text{even}}(\nu) = \left\langle |I_2|\delta_{\text{D}}(x - \nu)\,\delta_{\text{D}}(I_3)\,\mathcal{B}(-2J_2^{1/2} < J_1 < -J_2^{1/2})\right\rangle,$$

$$C_{2,\text{even}}(\nu) = \left\langle |I_2|\delta_{\text{D}}(x - \nu)\,\delta_{\text{D}}(I_3)\,\mathcal{B}(-J_2^{1/2} < J_1 < J_2^{1/2})\right\rangle, \tag{4.22}$$

$$C_{3,\text{even}}(\nu) = \left\langle |I_2|\delta_{\text{D}}(x - \nu)\,\delta_{\text{D}}(I_3)\,\mathcal{B}(J_2^{1/2} < J_1 < 2J_2^{1/2})\right\rangle = C_{1,\text{even}},$$

with

$$\delta_{\text{D}}(I_3) = \frac{27}{2}\delta_{\text{D}}\left(J_3 - \frac{3J_1 J_2 - J_1^3}{2}\right), \tag{4.23}$$

$$\delta_{\text{D}}(x - \nu) = \frac{1}{\sqrt{1 - \gamma^2}}\delta_{\text{D}}\left(\zeta - \frac{\nu + \gamma J_1}{\sqrt{1 - \gamma^2}}\right), \tag{4.24}$$

and the condition that the determinant is null due to λ_j being zero is enforced by restricted the range of J_1 according to the Boolean specified in equations (4.22). Eventually, the integration in equation (4.22) can be done symbolically and an analytical expression for $C_{j,\text{even}}(\nu)$ follows

$$C_{1,\text{even}}(\nu) = \sum_{i=1,6,9} c_{1,i}\,\exp\left(-\frac{\nu^2}{2\left(1 - \gamma^2/i\right)}\right), \tag{4.25}$$

$$C_{2,\text{even}}(\nu) = c_{2,6}\,\exp\left(-\frac{\nu^2}{2(1 - 5\gamma^2/6)}\right), \tag{4.26}$$

with

$$c_{1,1} = \frac{3\sqrt{\frac{5}{2}}\gamma\sqrt{1 - \gamma^2}\,\nu\left(275\gamma^4 + 30\gamma^2\left(2\nu^2 - 23\right) + 351\right)}{\pi^{3/2}\left(9 - 5\gamma^2\right)^4},$$

$$c_{1,6} = -\frac{\text{erf}\left(\frac{\gamma\nu}{\sqrt{2}\sqrt{5\gamma^4 - 11\gamma^2 + 6}}\right) + 1}{\sqrt{5}\pi\sqrt{6 - 5\gamma^2}},$$

$$c_{2,6} = \frac{2}{\pi\sqrt{30 - 25\gamma^2}},$$

$$c_{1,9} = \frac{\text{erf}\left(\frac{\sqrt{2}\gamma\nu}{\sqrt{5\gamma^4 - 14\gamma^2 + 9}}\right) + 1}{4\pi\sqrt{5}\left(9 - 5\gamma^2\right)^{5/2}}\left(\frac{3600\gamma^4\nu^4}{(9 - 5\gamma^2)^2} + \frac{120\gamma^2\left(27 - 35\gamma^2\right)\nu^2}{9 - 5\gamma^2} + 575\gamma^4 - 1230\gamma^2 + 783\right).$$
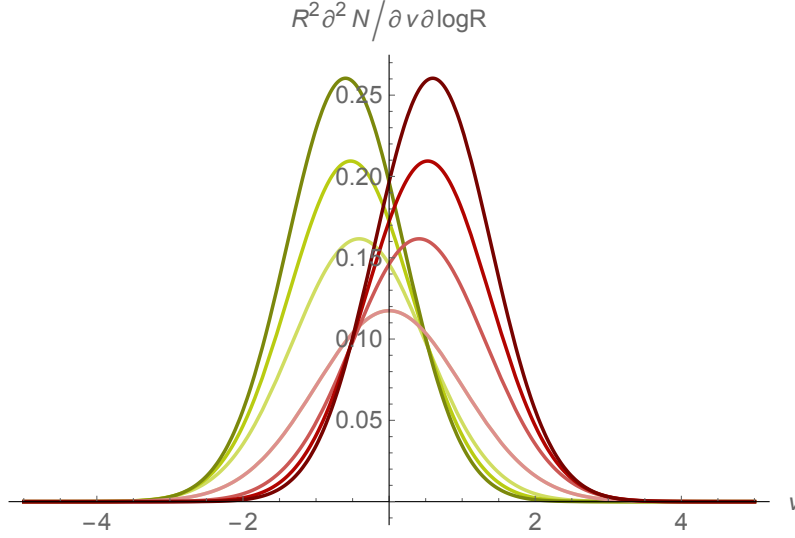
Figure 4.2.5: The PDF of critical events of the various types $(\mathcal{P}, \mathcal{F})$ in 2D for $n_s = -2, -3/2, -1, -1/2$ from light to dark. Note that the dominant change with spectral index is in the amplitude which scales like $1/\tilde{R}^2/R_\star^d$. The rest of the shape variation comes from the weaker γ and $\tilde{\gamma}$ dependence of C_{odd} and C_{even}.

The resulting counts of critical events as a function of their height ν is plotted in figure 4.4.1 for different values of the spectral index n_s. Note that $\partial^2 n/\partial R \partial \nu$ scales like $1/R^4$ but is also a function of R *via* the spectral parameters γ and $\tilde{\gamma}$.

4.2.4   2D event counts and differential counts

Since the formalism is very similar, let me also briefly present the analogues of equation (4.14) for 2D fields. It reads

$$\frac{\partial^2 n}{\partial R \partial \nu} = \frac{2\pi R}{\tilde{R}^2 R_\star^2} \langle |x_{211} + x_{222}||x_{222}|\delta_{\mathrm{D}}(x_1)\,\delta_{\mathrm{D}}(x_2)\rangle \times \tag{4.27}$$

$$\langle \vartheta_{\mathrm{H}}(x_{22} - x_{11})\,\delta_{\mathrm{D}}(x_{22})\delta_{\mathrm{D}}(x_{12})\,\delta_{\mathrm{D}}(x - \nu)|x_{11} - x_{22}|\rangle \,,$$

which after some algebra, given the knowledge of the 2D PDF given in section 4.C, yields for the peak merger rate

$$\frac{\partial^2 n}{\partial R \partial \nu} = \frac{R C_{\mathrm{odd}}}{\tilde{R}^2 R_\star^2} \left[\frac{4\gamma\nu\sqrt{1-\gamma^2}}{(3-2\gamma^2)^2} \exp\left(-\frac{\nu^2}{2(1-\gamma^2)}\right) \right. \tag{4.28}$$

$$\left. + \frac{\sqrt{8\pi}(2\gamma^4 + \gamma^2(\nu^2 - 5) + 3)}{(3 - 2\gamma^2)^{5/2}} \mathrm{erfc}\left(\frac{-\gamma\nu}{\sqrt{4\gamma^4 - 10\gamma^2 + 6}}\right) \exp\left(-\frac{3\nu^2}{6 - 4\gamma^2}\right) \right],$$

with

$$C_{\mathrm{odd}} = \frac{\hat{\gamma} + 3\hat{\gamma}^2 \tan^{-1}(3\hat{\gamma})}{4\pi^2} \,, \text{ given } \hat{\gamma} = \sqrt{1 - \tilde{\gamma}^2} \,.$$

The wall merger rate is obtained by swapping ν to $-\nu$ in this expression. The two rates are plotted in figure 4.2.5 and validated against Gaussian random fields in figure 4.4.2. The counts, $\partial n/\partial R = 2C_{\mathrm{odd}}R/(3\sqrt{3}\tilde{R}^2 R_\star^2)$ follows by integration over ν .

Section 4.A also presents differential counts in dimension 4 to 6, together with asymptotic expressions in the large dimension limit for the integrated count ratios. As expected, for any

dimension the number counts per unit log-volume is logarithmically scale invariant (up to the slow variation in the spectral parameters), i.e. $R^d\, \partial^2 n^d / \partial \log R \partial \nu$ is a function of γ, $\tilde{\gamma}$ and ν only.

4.2.5   Beyond Gaussian statistics

Let us finally compute the one-point statistics for close to Gaussian fields. The Edgeworth expansion joint statistics of the field at x, $P(x, x_i, x_{ij}, x_{ijk})$, involving the hierarchy of cumulants obeys

$$P(\boldsymbol{x}) = P_{\mathrm{G}}(\boldsymbol{x}) \left(1 + \sum_{k=3}^{\infty} \sigma^{k-2} \frac{\langle \mathbf{H}_k(\boldsymbol{x}) \rangle}{\sigma^{2k-2}} \cdot \mathbf{H}_k(\boldsymbol{x}) \right), \tag{4.29}$$

where $\mathbf{H}_k$ is a vector of orthogonal polynomials with respect to the kernel P_{G} obeying $\mathbf{H}_k = (-1)^k\, \partial^k P_{\mathrm{G}} / \partial \boldsymbol{x}^k\, / P_{\mathrm{G}}$ while at three order in perturbation theory (Bernardeau et al., 2002), $\langle \mathbf{H}_k(\boldsymbol{x}) \rangle / \sigma^{2k-2}$ is independent of the variance $\sigma^2(z)$ below $k = 6$. Equation (4.29) is in practice an expansion of the Gaussian PDF in the mildly non-linear regime where $\sigma(z) \ll 1$, so that the model is particularly accurate at large scales and at early times. Cumulants such as $\langle x_1^2 x_{111} \rangle$ entering equation (4.29) could in the context of a given cosmological model involve a parametrisation of modified gravity (via e.g. a parametrisation of $F_2(\boldsymbol{k}_1, \boldsymbol{k}_2)$), and/or primordial non-gaussianities (via e.g. f_{NL}). From this expansion, or relying on the connection between event ratio and connectivity discussed in section 4.A.5, we can for instance compute the non-Gaussian correction to the ratio of critical events, defined in equation (4.17) as

$$\frac{r_{\mathcal{P}/\mathcal{F}}}{r_{\mathcal{P}/\mathcal{F},\mathrm{G}}} = 1 + c_r \left(8 \langle J_1^3 \rangle - 10 \langle J_1 J_2 \rangle - 21 \langle q^2 J_1 \rangle \right) + \mathcal{O}(\sigma^2). \tag{4.30}$$

where $c_r = \left(29\sqrt{2} + 12\sqrt{3} \right) / 210 / \sqrt{\pi}$, while $\sigma_1^2 q^2 = |\nabla \rho|^2$ the modulus square of the gradient, J_1 and J_2 are defined in equation (4.20) *via* the trace and minor of the Hessian. These extended skewness parameters are isotropic moments of the underlying bispectrum which, when gravity drives the evolution, scale with σ at three order in perturbation theory (e.g. $\langle J_1^3 \rangle / \sigma$ is independent of σ). The correction to one entering equation (4.30) is negative (approximately equal to $-\sigma(1/7 - \log(L)/5)$ for a ΛCDM spectrum smoothed over L Mpc$/h$), suggesting that gravitational clustering reduces the relative number of peak mergers compared to filament mergers: filaments disconnect. When astronomers constrain the equation of state of dark energy using the cosmic evolution of voids disappearance they effectively measure σ in equation (4.30). Conversely, for primordial non Gaussianities, the extended skewness parameters must be updated accordingly (see Gay et al., 2012; Codis et al., 2013). For instance, $\langle J_1 q^2 \rangle = \langle J_1 q^2 \rangle_{\mathrm{grav}} - 2f_{\mathrm{NL}} \sqrt{1 + f_{\mathrm{NL}}^2} / (1 + 4 f_{\mathrm{NL}}^2)$.

Since the computation of the expectation (4.14) with the Edgeworth expansion (4.29) is beyond the scope of this dissertation, let us investigate an alternative proxy for the event rate. Figure 4.2.6 makes use of the perturbative prediction of Gay et al., 2012 to first order in σ for the gravitationally-driven non-gaussian differential extrema counts to compute the product of such counts as a proxy for the events, namely $\mathcal{P}(\nu) \propto P(\nu) \times F(\nu)$, $\mathcal{F}(\nu) \propto F(\nu) \times W(\nu)$, and $\mathcal{W}(\nu) \propto W(\nu) \times V(\nu)$. This Ansatz is reasonable, since for a merger to occur, two critical points of the same height must exist beforehand. We use the Gaussian PDF as a reference, to recalibrate the relative amplitude of the filament to peak merger counts. Since Gay et al., 2012 provide fits to the critical PDFs as a function of σ, it is straightforward to compute their product.

From figure 4.2.6, we see that gravitational clustering shifts the peak event counts to lower contrast, as it should. This is confirmed in simulation in figure 4.5.6. Less trivially, the filament merger rates also shift towards negative contrasts. From these PDFs one can re-compute the cosmic evolution of the ratio of critical events: it scales like $r_{\mathcal{P}/\mathcal{F}} = 7/34(1 - \sigma/7)$ (for $n = -1$), in good agreement with equation (4.30), suggesting that this approximation indeed captures the
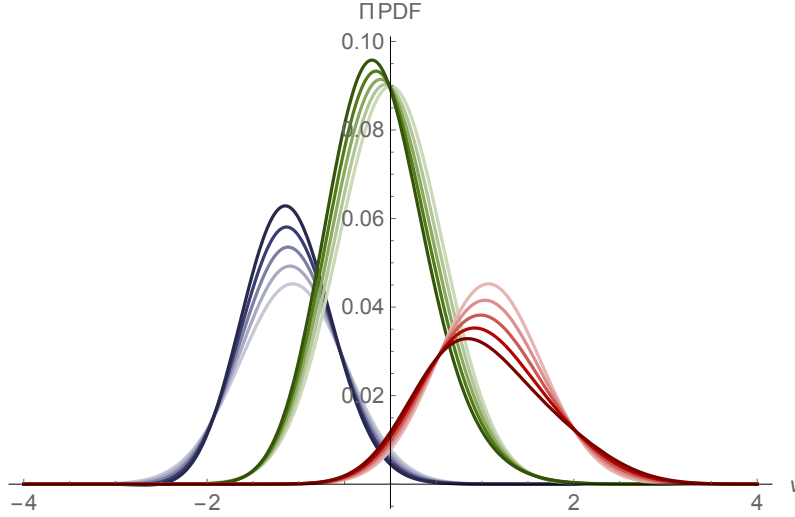
Figure 4.2.6: Predicted cosmic evolution of the product of extrema counts as a proxy for the event counts ($\mathcal{W}$ in blue, $\mathcal{F}$ in green and $\mathcal{P}$ in red) for the variances $\sigma = 0, 0.04,$ $0.08, 0.12, 0.16$ (from light to dark) and an underlying scale invariant power spectra of index $n = -1$. The $\mathcal{F}$ counts have been rescaled by a constant $205/332$ factor to better match the actual counts. The predicted trend with σ are in qualitative agreement with the measured counts presented in figure 4.5.6.

main features of gravitational clustering. This provides a physical understanding of the evolution of the one-point distribution of the critical events in the mildly non-linear regime.

4.3 Theory: two-point statistics

In the previous section, I have presented the concept of critical events (section 4.2.2.2) and derived their number counts counted together and by type (section 4.2.2.3), and by height (section 4.2.3). I have also presented how these results can be transposed in two dimensions (section 4.2.4 and eventually in d dimensions, see section 4.A.3). The formalism has also been extended in the mildly non-linear regime section 4.2.5.

Let me now present a method to compute the two-point statistics of critical events. Such statistics are of interest, for example to study the cosmic evolution of the connectivity of peaks, or to understand how large scale tides bias mass accretion (the so-called assembly bias). Section 4.3.1 presents the two-point statistics of merger events in 3D, while section 4.3.2 provides analytical approximations while assuming mergers occur along a straight filament. Section 4.3.3 computes the conditional merger rates subject to larger scale tides.

4.3.1 Clustering of critical events in R, r space

One cannot generally assume that the orientation of the two critical events are aligned w.r.t. the separation vector, so the covariant condition for critical event of type $j \in \{\mathcal{P}, \mathcal{F}, \mathcal{W}\}$, cond_j, is given by the argument of the expectation in equation (4.7) multiplied by requirement on the sign of the two non-zero eigenvalues. For instance

$$\mathrm{cond}_{\mathcal{P}}(\boldsymbol{x}) = |J| \delta_{\mathrm{D}}^{(3)}(x_i)\, \delta_{\mathrm{D}}(d) \times \vartheta_{\mathrm{H}}(-\mathrm{tr}(x_{ik}))\vartheta_{\mathrm{H}}\big(\mathrm{tr}^2(x_{ik}) - \mathrm{tr}(x_{il}x_{lk})\big)\,,$$

where the two Heaviside conditions ensure that the trace is negative and the minor positive so that the two eigenvalues are negative. From the joint two-point count of critical events, I can

define the relative clustering of critical events of kind i, j smoothed at scales (R_x, R_y) and located at positions $(\boldsymbol{r}_x, \boldsymbol{r}_y)$, $\xi_{ij}(\boldsymbol{s})$ as

$$1 + \xi_{ij}(\boldsymbol{s}) = \frac{\langle \mathrm{cond}_i(\boldsymbol{x}) \quad \mathrm{cond}_j(\boldsymbol{y}) \rangle}{\langle \mathrm{cond}_i(\boldsymbol{x}) \rangle \langle \mathrm{cond}_j(\boldsymbol{x}) \rangle} \,, \tag{4.31}$$

with

$$\boldsymbol{s} \equiv \sqrt{2} \left(\frac{\boldsymbol{r}_x - \boldsymbol{r}_y}{\sqrt{R_x^2 + R_y^2}} \right) \,, \tag{4.32}$$

the event separation between $\boldsymbol{x}(\boldsymbol{0})$ and $\boldsymbol{y}(\boldsymbol{s})$. Note that this definition of the separation includes the dependence of the correlation functions to the smoothing scale, as the product of two Gaussian kernels with scales R_x, R_y is equivalent to smoothing at a single scale $R = \sqrt{(R_x^2 + R_y^2)/2}$. The definition of equation (4.32) provides a natural distance ladder when comparing events at two different scales. Evaluating the expectation in equation (4.31) requires full knowledge of the joint statistics of the field $P(x, x_i, x_{ij}, x_{ijk}, y, y_i, y_{ij}, y_{ijk})$ (involving 40 variables, see section 4.C.2).

We rely on Monte-Carlo methods in MATHEMATICA in order to evaluate numerically equation (4.31). Namely, we draw random numbers from the conditional probability that $\boldsymbol{x}$ and $\boldsymbol{y}$ satisfy the joint PDF, subject to the condition that $x_k = 0, y_k = 0, x = \nu_1$ and $y = \nu_2$. For each draw $(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)})$ depending on the type of critical event hence the sign of $\mathrm{tr}(x_{ij})$ and $\mathrm{tr}^2(x_{ik}) - \mathrm{tr}(x_{ik}x_{kj})$ we drop or keep the sample; if it is kept, we evaluate $|J(\boldsymbol{x})| \delta_{\mathrm{D}}^{(\epsilon)}(d(\boldsymbol{x})) |J(\boldsymbol{y})| \delta_{\mathrm{D}}^{(\epsilon)}(d(\boldsymbol{y}))$ where $\delta_{\mathrm{D}}^{(\epsilon)}$ is a normalized Gaussian of width ϵ, which in the limit of $\epsilon \to 0$ would correspond to a Dirac function imposing here that the two determinants are zero. Eventually

$$\langle \mathrm{cond}_i(\boldsymbol{x}) \mathrm{cond}_j(\boldsymbol{y}) \rangle \approx \frac{P_m(x = \nu_1, y = \nu_2, x_l = y_l = 0)}{N} \times$$
$$\sum_{k \in \mathcal{S}_{ij}} \left| J(\boldsymbol{x}^{(k)}) \right| \delta_{\mathrm{D}}^{(\epsilon)} \left(d(\boldsymbol{x}^{(k)}) \right) \left| J(\boldsymbol{y}^{(k)}) \right| \delta_{\mathrm{D}}^{(\epsilon)} \left(d(\boldsymbol{y}^{(k)}) \right), \quad (4.33)$$

where N is the total number of draws, P_m the marginal probability for the field values and its gradients, and $\mathcal{S}_{ij}$ is the subset of the indices of draws satisfying the constraints i, j on the Hessians. The same procedure can be applied to evaluate the denominator. Equation (4.31) then yields an estimation of $\xi_{ij}(s, \nu_1, \nu_2)$. This algorithm is embarrassingly parallel.

This is illustrated in figure 4.3.1a which shows the auto-correlation of peak merger $\xi_{\mathcal{PP}}$ on the one hand, and the cross correlation of peak and filament merger $\xi_{\mathcal{PF}}$ on the other at fixed merger height, as labelled. Here we used $\epsilon = 0.1$. Note that because equation (4.31) is a ratio, the prefactors in the counts involving scales all cancel out. Similar results are presented in 2D on figure 4.3.1b.

4.3.2 Correlation of peak merger along filament

Let us briefly present the two-point statistics of high density peak mergers while assuming for simplicity that the mergers occur along the same (straight) filament (discussed in section 4.2.2), as it is instructive and simpler. In this approximation we can resort to one dimensional statistics. In the high density limit, we may drop the Heaviside constraint on the sign of the eigenvalues since all high density critical points tend to be automatically maxima. Then the (1D) correlation function of peak mergers, $1 + \xi_{\nu_1 \nu_2}(s)$ of height ν_1 and ν_2 becomes

$$\frac{\langle \delta_{\mathrm{D}}(x - \nu_1) \, x_{111}^2 \delta_{\mathrm{D}}(x_1) \, \delta_{\mathrm{D}}(x_{11}) \quad \delta_{\mathrm{D}}(y - \nu_2) \, y_{111}^2 \delta_{\mathrm{D}}(y_1) \, \delta_{\mathrm{D}}(y_{11}) \rangle}{\langle \delta_{\mathrm{D}}(x - \nu_1) \, x_{111}^2 \delta_{\mathrm{D}}(x_1) \, \delta_{\mathrm{D}}(x_{11}) \rangle \langle \delta_{\mathrm{D}}(y - \nu_2) \, y_{111}^2 \delta_{\mathrm{D}}(y_1) \, \delta_{\mathrm{D}}(y_{11}) \rangle}$$

(a) 3D

(b) 2D
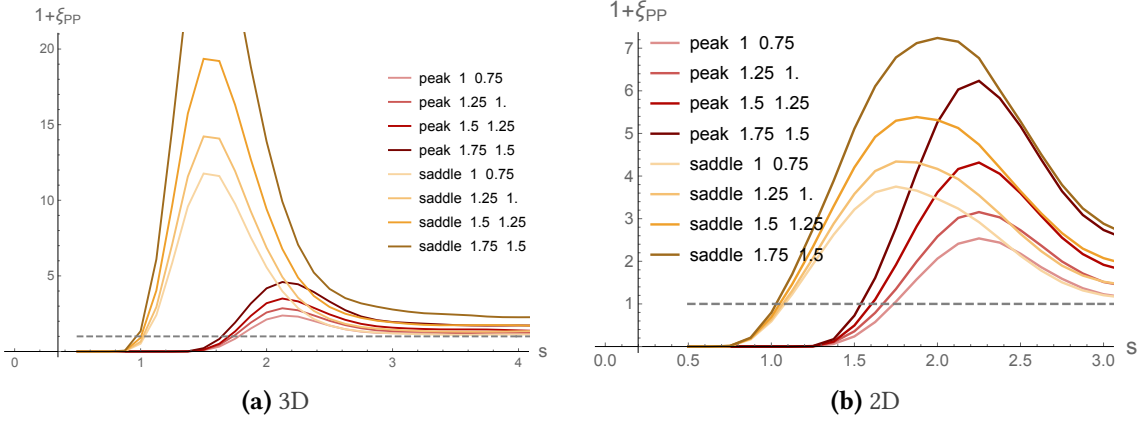
Figure 4.3.1: (a): The auto-correlation of peak merger $\xi_{\mathcal{PP}}$ (in shades of red, as labelled in terms of the height of the two critical points) and the cross correlation of peak and filament merger $\xi_{\mathcal{PF}}$ (in shades of yellow, as labelled) as a function of separation s. As expected, the saddle mergers are clustered closer to the higher peak compared to the peak mergers. (b): The two-point correlation of events in 2D fields with scale invariant power spectra of index $n_{\mathrm{s}} = -1$

where the expectation is over the Gaussian PDF whose covariance for the field $(x, x_1, x_{11}, x_{111}, y, y_1, y_{11}, y_{111})$ obeys

$$
\begin{pmatrix}
1 & 0 & -\gamma & 0 & \gamma_{00} & \gamma_{01} & \gamma_{02} & \gamma_{03} \\
0 & 1 & 0 & -\tilde{\gamma} & \gamma_{01} & \gamma_{11} & \gamma_{12} & \gamma_{13} \\
-\gamma & 0 & 1 & 0 & \gamma_{02} & \gamma_{12} & \gamma_{22} & \gamma_{23} \\
0 & -\tilde{\gamma} & 0 & 1 & \gamma_{03} & \gamma_{13} & \gamma_{23} & \gamma_{33} \\
\gamma_{00} & \gamma_{01} & \gamma_{02} & \gamma_{03} & 1 & 0 & -\gamma & 0 \\
\gamma_{01} & \gamma_{11} & \gamma_{12} & \gamma_{13} & 0 & 1 & 0 & -\tilde{\gamma} \\
\gamma_{02} & \gamma_{12} & \gamma_{22} & \gamma_{23} & -\gamma & 0 & 1 & 0 \\
\gamma_{03} & \gamma_{13} & \gamma_{23} & \gamma_{33} & 0 & -\tilde{\gamma} & 0 & 1
\end{pmatrix} ,
\tag{4.34}
$$

where for instance $\gamma_{02}(s) = \langle x(0) y_{22}(s) \rangle$, which can be computed using the formalism presented in section 2.1.1.4. The dominant contribution in the large threshold $\nu_1, \nu_2 \gg 1$, large separation $s \gg 1$ regime reads

$$
\Delta \xi^0_{\nu_1 \nu_2}(s) = \frac{\nu_1 \nu_2 \left(\gamma_{00}(s) + \gamma \left(2\gamma_{02}(s) + \gamma \gamma_{22}(s) \right) \right)}{(1 - \gamma^2)^2} ,
\tag{4.35}
$$

which as expected scales like the underlying correlation, $\gamma_{00}(s)$, boosted by the bias factor $\nu_1 \nu_2$ (Kaiser, 1984b). In that limit, the next order correction to the correlation function involving the third derivative of the field reads

$$
\Delta \xi^1_{\nu_1 \nu_2}(s) = \frac{2 \left(\tilde{\gamma}^2 \gamma_{11}(s) + 2\tilde{\gamma} \gamma_{13}(s) + \gamma_{33}(s) \right)^2}{(1 - \tilde{\gamma}^2)^2} ,
\tag{4.36}
$$

where $\tilde{\gamma}$-weighted linear combination of the autocorrelation of $\nabla \Delta \delta$ and the cross correlation of $\nabla \Delta \delta$ and $\nabla \delta$ appear, evaluated at events separated by s. The assumption of successive mergers of peaks occurring along a straight filament is of course very idealised, and prevents us from considering cross correlations between peak mergers and e.g. filament mergers.

4.3.3   Conditional merger rates in vicinity of larger tides

In the context of galaxy formation, it is of interest to quantify conditional merger rates computed subject to tides imposed by the large scale structure to explain geographically the origin of assembly bias. To do so one must compute the conditional event counts, subject to a given large scale critical point at some distance s from the running point x. The critical point can be e.g. a peak of a given geometry and height, if one is concerned with the impact of clusters on mergers trees of dark halos in their vicinity (Hahn et al., 2009; Ramakrishnan et al., 2019), or it could be a saddle point, as a proxy for a larger scale filament, when studying how halos growth stalls in such vicinity (Borzyszkowski et al., 2017; Musso et al., 2018). In turn this involves the joint expectation

$$\langle \mathrm{cond}_j(\boldsymbol{x})\, \delta_\mathrm{D}(y_i) | \det y_{ij} | \rangle. \tag{4.37}$$

Here cond_j is defined as in equation (4.31), namely it is the argument of equation (4.7) for a critical event of kind j

Evaluating equation (4.37) requires the full knowledge of the joint statistics of the field at $\boldsymbol{r}_x$ and $\boldsymbol{r}_y$, $\mathcal{P}(x, x_i, x_{ij}, x_{ijk}, y, y_i, y_{ij})$ (involving 30 variables). The correlations of the PDF involves the covariance of the field and its derivatives computed at two smoothing scales, R_x and R_y. We can then marginalise over all variables, subject to e.g. imposing the height, ν_c and shape, μ_i^c of the large scale critical mode:

$$\langle \mathrm{cond}(\boldsymbol{x})\delta_\mathrm{D}(y_i) | \det y_{ij} | \delta_\mathrm{D}(x-\nu)\delta_\mathrm{D}(y-\nu_c)\vartheta_\mathrm{H}(-\lambda_i)\, \delta_\mathrm{D}(\mu_i - \mu_i^c) \rangle$$

where λ_i are the eigenvalues of x_{ij} and μ_i are the eigenvalues of y_{ij}. The conditions imposed by the mergers and the properties of the peaks and large scale environment reduces the number of integrals from 30 to 21. Section 4.B.3 describes how to sample conditional event counts using constrained realisation of Gaussian random fields.

4.4   Measurements for Gaussian random fields

In the previous sections, I have provided the one-point statistics of critical events section 4.2 and their two-point statistics section 4.3.

Let me now validate the theory while counting critical events within realisations of Gaussian random fields. Section 4.4.1 details the procedure followed to generate Gaussian random fields. Section 4.4.2 measures the one-point statistics and compares them to predictions, while section 4.4.3 measures the two-point statistics.

4.4.1   Method

For each power-law power spectrum with spectral index $n_\mathrm{s} = -2, -1.5, -1, -0.5$, I have generated 200 gaussian random fields. I have also generated 200 gaussian random fields with a ΛCDM power spectrum using mpgrafic (Prunet et al., 2008) in a Planck cosmology (Planck Collaboration, 2018a) generated using the Eisenstein and Hu, 1999 fitting formula. Each realisation will henceforth be called a "cube". Each cube has a size of 256^3 pixels and a physical extent of $100\,\mathrm{Mpc}/h$.[5] Each cube has been smoothed using a Gaussian filter with scale ranging from $1\,\mathrm{Mpc}/h$ to $20\,\mathrm{Mpc}/h$ ($2.56\,\mathrm{px}$ to $51.2\,\mathrm{px}$). The smoothing operation were operated in Fourier space, assuming periodic boundary conditions. At each scale, all critical points are detected (minima, saddle points and extrema) using the method detailed in section 4.B.1 while the critical events have been detected using the method detailed in section 4.B.2.

Additionally, I have generated 200 2048^2 cubes with a power-law power spectrum with spectral index $n_\mathrm{s} = -1$ and a physical box size of $1\,000\,\mathrm{Mpc}/h$ which I smoothed with a Gaussian filter with scale ranging from $1\,\mathrm{Mpc}/h$ to $20\,\mathrm{Mpc}/h$.

---

[5]The box size is only relevant in the ΛCDM case, as the power-law cases are scale invariant.

4.4.2 Critical events counts

In this section I present the number density of critical event measured in cubes with a power-law power spectrum and compare the theoretical predictions of section 4.2.3 to measurements in cubes.

I first measured the ratio of the number of critical events of different kind, which is found to be $r_{\mathcal{F}/\mathcal{P}} = r_{\mathcal{F}/\mathcal{W}} \approx 2.1$, regardless of the smoothing scale or the underlying power spectrum. This excess of about 2% in the ratio originates to an over-detection of saddle points with respect to local extrema. Theory predicts this ratio to be $N_{\mathrm{saddle}}/N_{\mathrm{peak}} \approx 3.055$ in 3D (see e.g. Codis et al., 2018, equation 2) while the measured value is 3.1. In the following of the chapter, I have corrected the excess number density of $\mathcal{F}, \mathcal{W}$ critical events by applying a correction factor to their number counts.

Let me now proceed to the number count at fixed density. Figure 4.4.1 shows the PDF of the critical events as a function of their height for different power-law spectra ($n_{\mathrm{s}} = -2, -1.5, -1, -0.5, \Lambda\mathrm{CDM}$). The critical events have been selected at scale $2.35 \, \mathrm{Mpc}/h \leq R \leq 3.01 \, \mathrm{Mpc}/h$ ($6.0 \, \mathrm{px} \leq R \leq 7.7 \, \mathrm{px}$). The lower boundary ensures that the critical points are well separated[6]. The upper boundary is fixed so that the smoothed cubes have consistent effective spectral parameters $\gamma_{\mathrm{eff}}(R)$ and $\tilde{\gamma}_{\mathrm{eff}}(R)$. Indeed, the cubes have scale-dependant spectral parameters induced by the finiteness of the box and the discreteness of the grid (see e.g. Gay, 2011, figure 5.1). Errorbars have been estimated using a bootstrap method ran on 400 subsamples each made of 50 randomly chosen cubes. Solid lines show the result of a fit of the theoretical formula to the cube data with free parameters $\hat{\gamma}, \hat{\tilde{\gamma}}$.

The effective spectral index $\hat{n}_{\mathrm{s}}$ is fixed using $\gamma = \sqrt{(n_{\mathrm{s}}+3)/(n_{\mathrm{s}}+5)}$. The measured values of γ and $\tilde{\gamma}$ are consistent with the effective values measured directly in the cubes using equation (2.98). For example with $n_{\mathrm{s}} = -2$, the values measured in the cubes are $\gamma_{\mathrm{eff}} = 0.62 \pm 0.02, \tilde{\gamma}_{\mathrm{eff}} = 0.72 \pm 0.01$ ($n_{\mathrm{s,eff}} = -1.75 \pm 0.13$) using equation (2.98). The mean values have been estimated with a sample of 100 cubes and the errors are the standard deviations of the sample. The fitting procedure on the PDF of the critical events yields $\hat{\gamma} = 0.621 \pm 0.002, \hat{\tilde{\gamma}} = 0.737 \pm 0.004$ ($\hat{n}_{\mathrm{s}} = -1.74 \pm 0.02$). The relative difference between theory and measurements, presented on the upper panel of figure 4.4.1, show no systematic deviation of the measurements and is within a few percents in the region where most of the events are.

In order to further test the theoretical prediction, I have proceeded to the same analysis in the 2D case. The results are presented on figure 4.4.2 and show that the agreement between theory and measurements is of the order of the percent. Once again, no systematic deviation of the measurements is noted. The results in 2 and 3D confirm the analytical formula derived in section 4.2.3 and illustrate the accuracy of the detection algorithm presented in section 4.B. Interestingly, since the algorithm has been designed to make no assumption on the number of dimensions, it is expected to work as well in d dimensions.

4.4.3 Two-point statistics

Let me now estimate the two-point statistics of critical events using the critical events from the cubes presented above. For any two subsets A and B of critical events, one can estimate their correlation function numerically using

$$\xi_{AB}(s) = \frac{\langle AB \rangle}{f\sqrt{\langle AR_A \rangle \langle BR_B \rangle}}, \tag{4.38}$$

where R_A and R_B are uniformly distributed random points with $1/f$ times the number of points as A and B respectively. I have additionally checked that common estimators, such as the Landy-

---

[6]Critical points are typically separated by $R_* \gtrsim 0.6R$ (for $n_{\mathrm{s}} < 0$), so $R = 6 \, \mathrm{px}$ gives a typical separation of $3.6 \, \mathrm{px}$ between critical points, which is larger than the number of points used to infer the curvature.
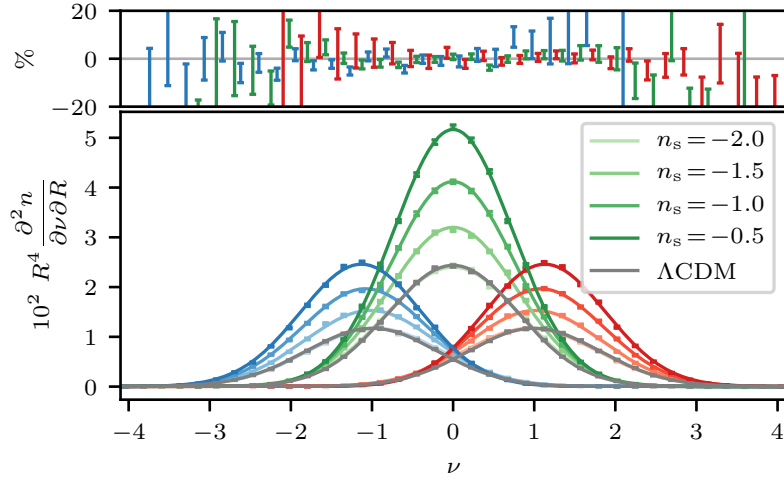
Figure 4.4.1: PDF of the critical events as a function of height in a scale invariant GRF as labelled. The left bundle corresponds to void mergers, the middle bundle to filaments mergers and the right bundle to peak mergers. The solid curve corresponds to the theory while the error bars correspond to the error on the mean extracted from 160 simulations. The grey lines are the results obtained for a ΛCDM power spectrum initially smoothed over a scale of $2.5\,\mathrm{Mpc}/h$. The top panel shows the residuals for $n_{\mathrm{s}} = -2$. The detection algorithm is still accurate in 3D.
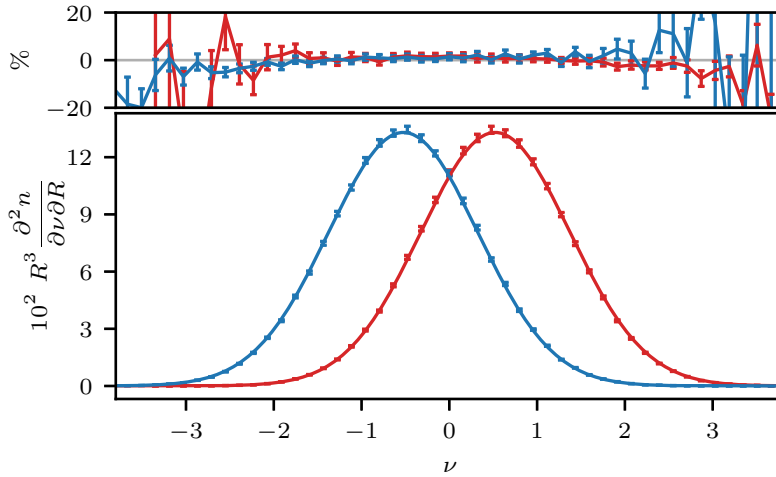
Figure 4.4.2: PDF of the critical events as a function of height in a scale invariant GRF in 2D with spectral index $n_{\mathrm{s}} = -1$. The left curve corresponds to filament mergers and the right curve to peak mergers. The solid curve correspond to the theory while the error bars correspond to the error on the mean extracted from 200 simulations. The top panel shows the residuals. The agreement between the analytic prediction and the measurements reflects the accuracy of the algorithm presented in section 4.B in identifying critical events.

(a) 2D **(b)** 3D
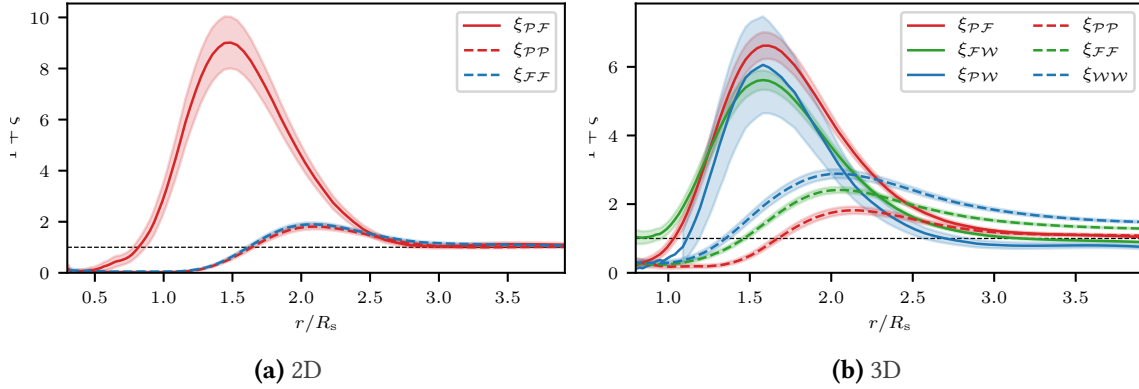
Figure 4.4.3: (a): Correlation functions between critical events $\mathcal{P}, \mathcal{F}$ in 2D at fixed smoothing scale. (b): Correlation functions between critical events $\mathcal{P}, \mathcal{F}, \mathcal{W}$ in 3D at fixed smoothing scale. Pairs of critical events have been selected at $\nu = 0.7$ and $\nu = 1.0$. The correlation function of halo-merger with filament-merger, $\xi_{\mathcal{P}\mathcal{F}}$, peaks at $r \sim 1.5R$ while the halo-merger autocorrelation functions $\xi_{\mathcal{F}\mathcal{F}}$ peaks at $r \sim 2R$. This shows that halo-mergers are more likely to be followed by filament-mergers. The data have been filtered using a Savgol filter. Errorbars have been estimated assuming a Poisson noise on the sample.

Szalay estimator yield similar results. This is further discussed in section 4.B.4, which shows that both estimators yield similar results at scales of interest to our analysis ($s = r/R \gtrsim 1$). For each cube in the simulation, I then select all critical events in a thick slice of smoothing scales ($\Delta R/R = 0.3$). The critical events are then split in two subsamples, the first is selected at an overdensity $\nu = 1$ with kind j and the second at $\nu = 0.7$ with kind k ($j, k \in \{\mathcal{P}, \mathcal{F}, \mathcal{W}\}$). The correlation functions are then given by the number of pairs at distance $s = r/R$ in all cubes using equation (4.38). The pair counting was done using a dual-tree algorithm, as described in Moore et al., 2001[7].

Figure 4.4.3 shows the measured correlation functions in 2D for a power law power spectrum with spectral index $n_s = -1$ (top panel) and in 3D with a ΛCDM power spectrum smoothed at scales between 1 and $20 \, \mathrm{Mpc}/h$ (bottom panel). In both cases the $\mathcal{P}\mathcal{F}$ cross-correlation function (peak merger to filament merger correlation) peaks at $r \approx 1.5R$ while the $\mathcal{P}\mathcal{P}$ auto-correlation function (peak merger autocorrelation) peaks at $r \approx 2.5R$. This indicates that each halo merger is more likely to be followed by a filament merger compared to another halo merger. Interestingly, peak mergers are also more likely to be followed by void mergers. Indeed, a halo merger induces a topological defect, as it leads to a resulting over-connected halo. The defect is quickly corrected by a filament merger, decreasing the local connectivity of the halo back towards the cosmic average. Doing so another topological defect appears as a void becomes under-connected as one of its walls disappeared. This last defect is then corrected by a last void merger that makes the under-connected void disappear. On average, critical events happen so that the global ratio of peak-to-filament, filament-to-walls and wall-to-void stays constant as smoothing increases, so that the global connectivity is preserved. The link between critical events and global connectivity of the cosmic web is further discussed in section 4.5.2.

---

[7]See the scipy doc for more information.

4.5   Applications to galaxy formation and discussions

I showed how one can derive the one-point and two-points statistics of critical events in sections 4.2 and 4.3. I have then successfully compared the predictions to Gaussian random fields and provided the two-point correlations functions in section 4.4.1.

The scope of application of the present formalism is obviously very wide. Rather than attempting to cover it all, I will present a few examples here, while a more thorough investigation is left for future work.

In a cosmic framework, section 4.5.1 will first translate the one-point statistics presented in the previous section into halo and void merger rates as a function of mass and redshift. Section 4.5.2 explains how mergers of filaments need to match that of peaks in order to preserve the connectivity of peaks. Section 4.5.3 explains how conditional merger counts in the vicinity of a filament explains how the environment drives assembly bias. Section 4.5.4 illustrates how the theoretical predictions compare to results from N-body simulations and galaxy catalogues. It also shows how wall merger rates can be used to yield constraints on modified gravity or primordial non gaussianities.

4.5.1   Merger rates in M, z space

The predictions in the initial Lagrangian space bear theoretical interest, yet they do not translate easily to measurable quantities. In this section, let me show how one can map the predictions to observable quantities, and in particular merger rates in M, z space. It is straightforward to change variable from R to M (recalling that $M = \frac{4}{3}\pi\bar{\rho}(\alpha R)^3$ from equation (4.2)) and from ν to z using the spherical collapse condition (equations 4.1 and 4.2), so that for condition c (peak, void) one has[8]

$$\frac{\partial^2 n}{\partial \log M \partial z}\Big|_c = \frac{\partial^2 n}{\partial R \partial \nu}\Big|_c \frac{\partial R}{\partial \log M}\frac{\partial \nu}{\partial z}$$

$$= -\frac{\partial^2 n}{\partial R \partial \nu}\Big|_c \frac{\delta_c}{3\alpha D(z)^2}\frac{\mathrm{d}D}{\mathrm{d}z}\left(\frac{3M}{4\pi\bar{\rho}}\right)^{1/3}, \tag{4.39}$$

where $\alpha \approx 2.1$ and $\bar{\rho} \approx 2.8 \times 10^{11}\,h^2\mathrm{M}_\odot/\mathrm{Mpc}^3\,\Omega_\mathrm{M}$ (see e.g. Musso, Cadiou et al., 2018, Table A1). The same reasoning can be applied to get a similar relation for void mergers (or equivalently wall mergers) substituting δ_c by $|\delta_\mathrm{v}| = 2.7$ (see Jennings et al., 2013, equation 8). Note that this simple relation holds in principle for small enough voids only ($R \lesssim 3\,\mathrm{Mpc}/h$). A more detailed study will be provided in future works. From equations (4.14) and (4.39), I am now in a position to count how many (peak, void) mergers occur early or late in the accretion history of a certain mass or within some mass range, *via* straightforward integration.

Figure 4.5.1 shows the merger rate of peaks and voids as a function of the mass of non linearity. The cosmology-dependant terms of equation (4.39) ($D(z)$ and $\mathrm{d}D/\mathrm{d}z$) have been computed using the code Colossus (Diemer, 2018) in a Planck cosmology. With increasing time mergers of increasing size are able to happen, as the collapse barrier decreases. Note that the cutoff at large-scale is significantly faster that the Press-Schechter cutoff, as discussed bellow. In order to evaluate the number density of critical events, I have assumed a scale-dependant equivalent power-law power spectrum[9].

It should in principle be possible to generalize equation (4.39) for filament mergers, but this would require the knowledge of a relation between the initial overdensity and the mass of the filament or its length, as well as a collapse condition. In practice, this would likely result in

---

[8]Note that $\mathrm{d}D/\mathrm{d}z = -Df/(1+z)$ with $f \equiv \mathrm{d}\log D/\mathrm{d}\log a \sim \Omega_\mathrm{m}^{0.6}$.

[9]At each scale, the equivalent power-law power spectrum is given by the formula $n_\mathrm{s,eq} = -3-2\,\mathrm{d}\log\sigma/\mathrm{d}\log R$, where σ is computed using a ΛCDM power spectrum.
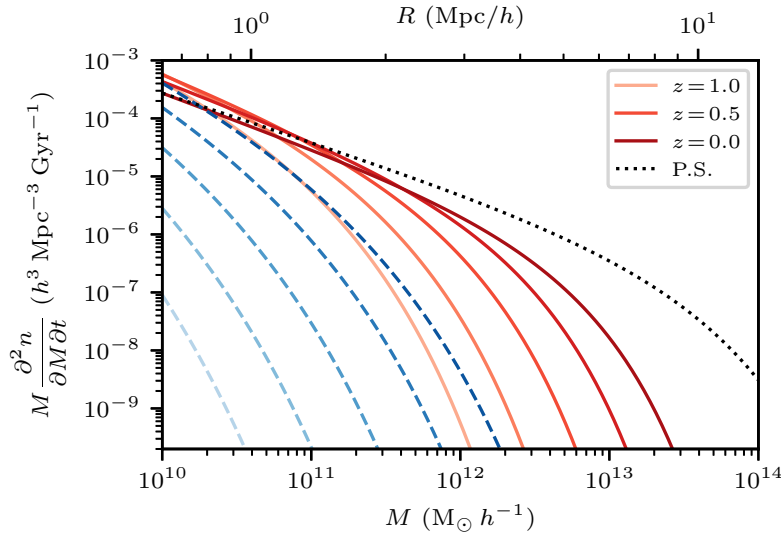
Figure 4.5.1: PDF of the halo merger rate (solid red lines) and the wall merger rate (dashed blue lines) as a function of redshift of formation (from dark to light at $z = 0, 0.25, 0.5, 0.75, 1$), see the text for details, up to a renormalisation. For the sake of clarity, only the redshifts $z = 0, 0.5, 1$ appear in the legend. For small masses the merger rate follows the Press-Schechter (Press and Schechter, 1974) halo mass function up to a renormalisation (black dotted line), while at larger masses the halo merger rate decays significantly faster. As expected, the transition mass increases with time. The same evolution is found for void mergers.

implementing a cylindrical collapse condition, while Pogosyan et al., 2009b suggested this could be achieved using a somehow smaller critical overdensity for filamentary collapse. The impact of our results on filament merger rates in M, z space will be done in a follow-up work.

4.5.2 Consistency with cosmic connectivity evolution

The properties of the initial random field was shown by Codis et al., 2018 to control to a large extent the connectivity of dark halos, as defined by the number of connected filaments (locally and globally) at a given cosmic time. The upshot of this work is that the packing of peaks (*i.e.* the "volume" they occupy, as imposed by their exclusion zone) and saddles implies that 3-4 filaments typically dominate locally. Interestingly, the rate of filament disappearing must match the peak merger rate, in order to preserve this number. Beyond numerology, this rate is important because filaments later feed coherently dark halos, so their lifespan matters to understand the balance between filamentary cold gas inflow (from subsisting filaments) and environmentally-driven disruptions (from filaments mergers).

In practice, one should distinguish the local and global connectivity (see Codis et al., 2018, for more details). Unfortunately, the link between global connectivity and merger rates that was discussed in the present work does not translate straightforwardly to the local connectivity. Our qualitative understanding of the critical structure of Gaussian random fields remains in close relation to packaging: each vicinity of a critical point, and with the same argument, of a critical event, must by continuity occupy a certain volume of space, as set by its eigenvalues, which puts constrains on the position of other points in the vicinity. Indeed, critical points are found where the gradient vanishes, with some local curvature, so that the field is quadratic in each eigenvector's direction. As a consequence, the gradient of the field is linear at non-null separation and cannot vanish, so that no other critical point can be found in the direct vicinity of another

critical point or event. At large separations the field decorrelates from its values at the critical point, so that another critical point event becomes likely. The same reasoning applies to critical event, except that the field has a third order behaviour along the ridge of the vanishing saddle point (it is an inflection point in that direction). The idea is that e.g. before connecting a given peak to a peak of a different height, the field must first go through a local minimum along the ridge, which distance is set by the 'width' of that peak. For events, the process of smoothing the field will impact both the local curvature but also the curvature of these other points. Hence, it is expected that smoothing jointly disconnects neighbouring peaks as mergers occur: the ridges are smoothed out because technically their saddle points vanish.

I can quantify this process *via* the two-point correlation functions of these events. From the auto- and cross-correlations of the $\mathcal{P}$ and $\mathcal{F}$ events presented in section 4.3, I can define the ratio of the separation at the maximum of these two correlations ($s_{ij} = \mathrm{argmax}_s \xi_{ij}(s)$) as a measure of the relative 'proximity' of the two events. Since this ratio $s_{\mathcal{PF}}/s_{\mathcal{PP}} \approx 3/4$ is smaller than one, it means that filament mergers are more clustered around halo mergers than halo mergers around halo mergers, so that the rate at which filaments disappear matches the merger rate and the typical number of filaments per halo remains constant through cosmic time. As a result of this spatial clustering, the most likely sequence happening is a $\mathcal{PFFP}$, as presented on the cartoon of figure 4.5.2 in 2D. This sequence conserves the connectivity of peaks, and is consistent with the relative rates of events. Figure 4.5.3 illustrates an analogous consistent $\mathcal{PF}^4\mathcal{P}$ sequence in 3D. Figure 4.5.4 shows how the local connectivity of 3 can also be preserved, as the weaker filaments typically lie off the main plane.

Finally, the clustering of filament disappearance impacts the connectivity of peaks as they merge as discussed in the next section (see figure 4.5.5, bottom right panel). This is a direct consequence of the clustering of events of the various types.

4.5.3   Assembly bias in the frame of filaments

Let me now make use of the merger statistics to study the impact of the large scale structures on assembly bias, following section 4.3.3. Previous works have highlighted the modulation effect induced by large-scale filamentary structure on the assembly of dark matter halos and their galaxies. Indeed, it is expected on theoretical ground that the typical accretion rate increases when going from saddle towards nodes Musso, Cadiou et al., 2018 Looking at galactic properties instead, Kraljic et al., 2019 showed that the galactic ratio of rotational-velocity-to-velocity-dispersion (v/σ) is also modulated as a function of the distance and orientation to the nearest filamentary structure. Using the framework developed in this work, I generate a suite of Gaussian random fields constrained to the presence of a proto-filament at the centre of the box, the exact generation procedure being described in section 4.B.3. The proto-filament is defined at a scale $R = 5 \, \mathrm{Mpc}/h$, is oriented along the z axis and lays in a wall in the yz plane. Using the set of constrained cubes, I compute the excess density of each kind of critical event with respect to the cosmic mean, at fixed smoothing scale (hence at fixed object mass) $2.5 \leq R \leq 5 \, \mathrm{Mpc}/h$. The results are shown on figure 4.5.5.

The peak merger rate is shown on the top left panel of figure 4.5.5. Going from the voids to the wall, from the wall to the filament and from the filament to the nearest node, the peak merger rate increases and the maximum peak merger rate is found at the location where a node is expected ($z \sim \pm 10 \, \mathrm{Mpc}/h$). At larger scales, the field becomes unconstrained so that the peak merger rate falls back to its cosmic mean. I reproduce here from first principle the results of Borzyszkowski et al., 2017, showing that halos close to the filament saddle are stalled compared to those in nodes: they do not undergo many mergers nor do they accrete much as the local tidal fields channels all the matter towards the two surrounding nodes, bypassing the centre of the filament. Quantitatively, halos forming at the centre of the filament are found to have a halo
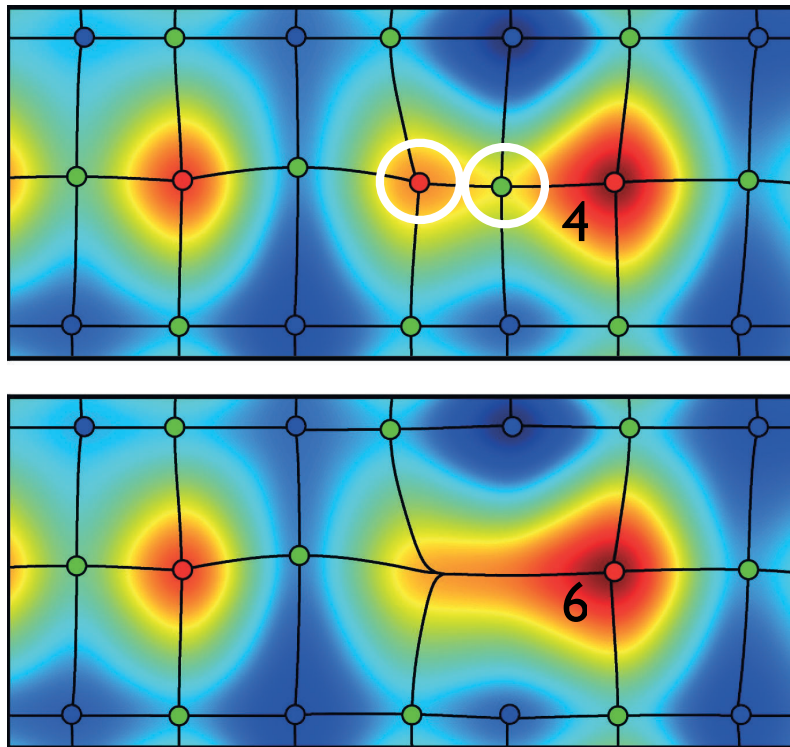
Figure 4.5.2: Snapshots of the density field at two smoothing scales (colour coded from blue, low density to red high density). The black line represents density ridges/trough connecting the red peaks, and the blue voids *via* the green saddle points. As the two low persistence pair of peaks (in white) merge the connectivity increases from 4 to 6 (as labeled). The fate of this connectivity now depends on the nature and location of the next merger events inspired from Sousbie et al., 2011.

merger rate close to the cosmic average, while those close to the nodes are expected to have 40% more mergers. Conversely, halos forming in a void next to a filamentary structure are expected to have a merger rate -20% smaller than the cosmic mean.

Filament merger rates act locally to decrease the connectivity of halos, as each filament merger will disconnect one filament from two halos. The top right panel of figure 4.5.5 shows that the merger rate is maximal along the wall and minimal along the filament. Going off the plane of the wall (x direction), the filament merger rate simply decreases towards the cosmic mean. Interestingly the filament merger rate is minimal in the nodes (-13%) and maximal in the wall ($+10\%$). As a consequence, halos forming close to a node have a larger halo merger rate but a smaller filament merger rate. This in turn will have an impact on the assembly of dark matter halos and their galaxies. In the wall where the filament merger rate is the highest, I expect filaments to merge faster than halos, resulting in halos with fewer connected filaments. This can be interpreted using the results of section 4.2.4. Indeed, in a cosmic wall, the geometry is locally 2D so that the theoretically expected connectivity becomes 4 instead of 6.

The bottom left panel of figure 4.5.5 shows that the wall merger rate is decreased in walls and even more strongly in filaments. The minimum wall merger rate is found at the location of the node with a rate -40% smaller than the cosmic mean. Conversely, the wall merger rate is enhanced in the two voids surrounding the wall with a rate 20% above the cosmic mean.

The evolution of the connectivity with cosmic environment is resumed by the bottom right panel of figure 4.5.5, which shows the ratio of halo mergers ($\mathcal{P}$ critical events) to filament mergers
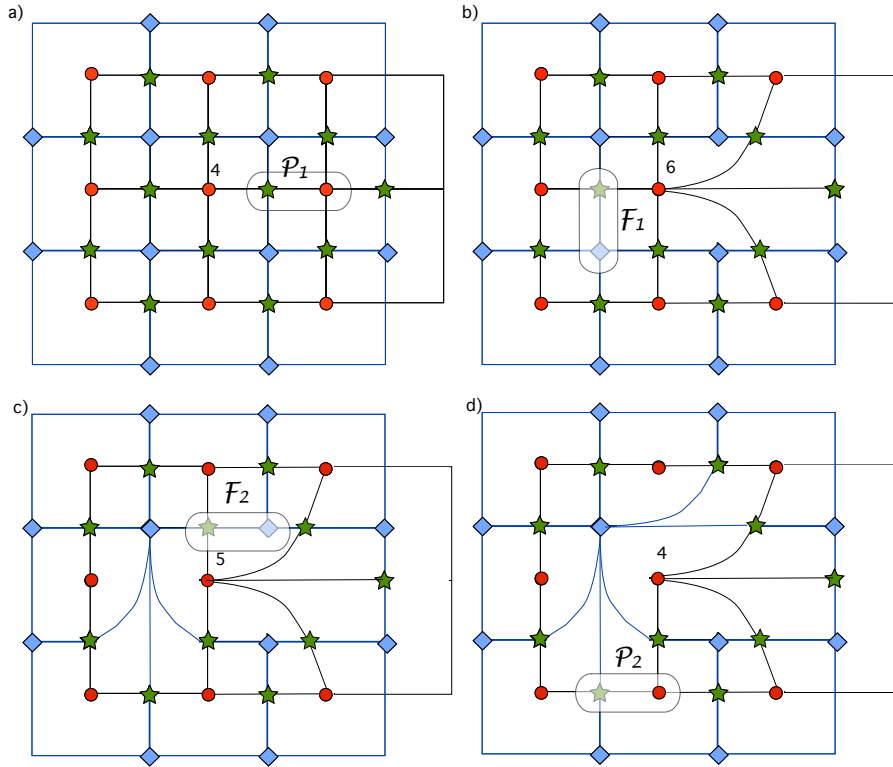
Figure 4.5.3: As labelled from a) to d) an abstraction of the merger sequence of a 2D 'cosmic crystal' impacting the connectivity of the central peak. Ridges are shown in black while troughs are shown in dark blue. The red circles represent the peaks, the green stars the saddles and the blue diamonds the voids. A $\mathcal{P}_1$ merger (highlighted in light grey) rises the mean connectivity of the central peak from 4 to 6, but the next two $\mathcal{F}_{1,2}$ mergers (highlighted in darker grey) lower it back to 4. The next $\mathcal{P}_2$ merger (panel d) will reduce the void's connectivity. A more realistic representation of this process is also visible on figure 4.2.3.

($\mathcal{F}$ critical events), for which the cosmic mean is 2.055 (see equation (4.17)). Small values of $r_{\mathcal{F}/\mathcal{P}}$ indicate that halos merge faster than their surrounding filaments, so that the connectivity increases as halos grow. On the contrary, large values of $r_{\mathcal{F}/\mathcal{P}}$ indicate that filaments merge faster than halos, so that the connectivity decreases as halos grow. The bottom right panel of figure 4.5.5 shows that in nodes, the ratio drops to about $r_{\mathcal{F}/\mathcal{P}} \approx 1.1$. On the contrary halos forming in voids are expected to have a ratio of about 2.4. I therefore expect that, at fixed final mass, halos forming next to a node will grow an increasing number of connected filaments[10]. The expected physical outcome of this process is that the streams feeding a galaxy growing next to a node will become more and more isotropic with increasing connectivity.

Assuming that an isotropic acquisition of matter leads to a smaller amount of angular momentum being transferred down to the disk, I propose that this effect prevents the formation of gaseous disks in the vicinity of nodes. Conversely, I expect that halos growing in the neighbouring voids see their filaments destroyed faster than they merge, so that the halo is likely to grow with steadier flows coming from a few filaments (see also Codis et al., 2015; Laigle et al., 2015, section 6.2.1, and 5 resp. for similar conclusions reached *via* the kinematic structure of large scale

---

[10]Conversely Codis et al., 2015 found that when averaged over all large scale structures, connectivity increases with mass.
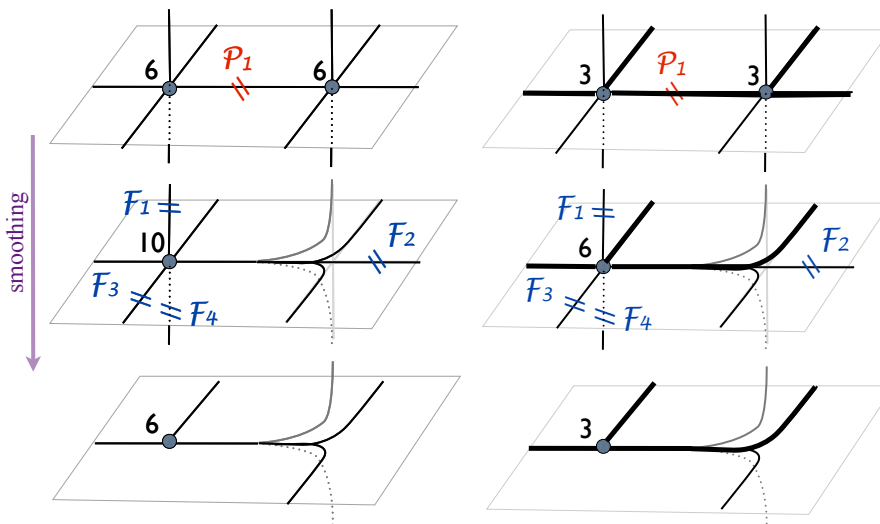
Figure 4.5.4: Following the cartoon shown in figure 4.5.3, the left panel shows a smoothing sequence (from top to bottom) which would preserve the connectivity of a 3D peak. It requires that each $\mathcal{P}$ merger should be followed by four $\mathcal{F}$ mergers in the vicinity. The right panel highlights how the multiplicity is preserved if one starts with 3 dominant co-planar filaments.

flows in filaments). The results presented here show a significant effect of the saddle point on the dynamical evolution of halos and their surrounding. At first order, the one-point statistics presented in section 4.2 show that the number density of halo mergers increases with increasing density, while the number density of filament merger decreases so that a more detailed analysis will be required to disentangle the effect from density from the modulation from the cosmic web. Following the arguments of Musso, Cadiou et al., 2018, one can however argue that the saddle point differentially impacts the statistics of halo mergers compared to filament mergers (as their two-points correlation functions are in principle different), so that an effect beyond density can be expected. This could be done by comparing the filament merger-to-halo merger ratio estimated from the mean local density alone to the maps presented in figure 4.5.5. These effects however require a more in-depth analysis and will need to be studied in future works.

4.5.4 Departures from gaussianity

With increasing time, non-linearities arising from gravitational collapse translate into departures from gaussianity (Bernardeau et al., 2002). This can formally be studies in the framework described in section 4.2.5. In particular, the PDFs will become skewed as the dynamics of gravitational collapse depends on the primordial overdensity.

Results in N-body simulations

Here, let me quantify the effect first on simulations, and then compare to the proxy of section 4.2.5 relying on known perturbative results. Figure 4.5.6 presents the redshift evolution of critical counts measured in 200 realisations of ΛCDM simulations in boxes of $500\,\mathrm{Mpc}/h$ involving 256^3 particles evolved using GADGET (Springel et al., 2001) sampled on a 256^3 grid smoothed with a Gaussian filter over $6\,\mathrm{Mpc}/h$. The algorithm described in section 4.B is used to identify and match the critical points. The qualitative similarity with the cosmic evolution of the measured event counts and the prediction shown in figure 4.2.6 is striking, strongly suggesting that indeed, the set of critical events in the initial condition do capture the upcoming cosmic evolution of the field.
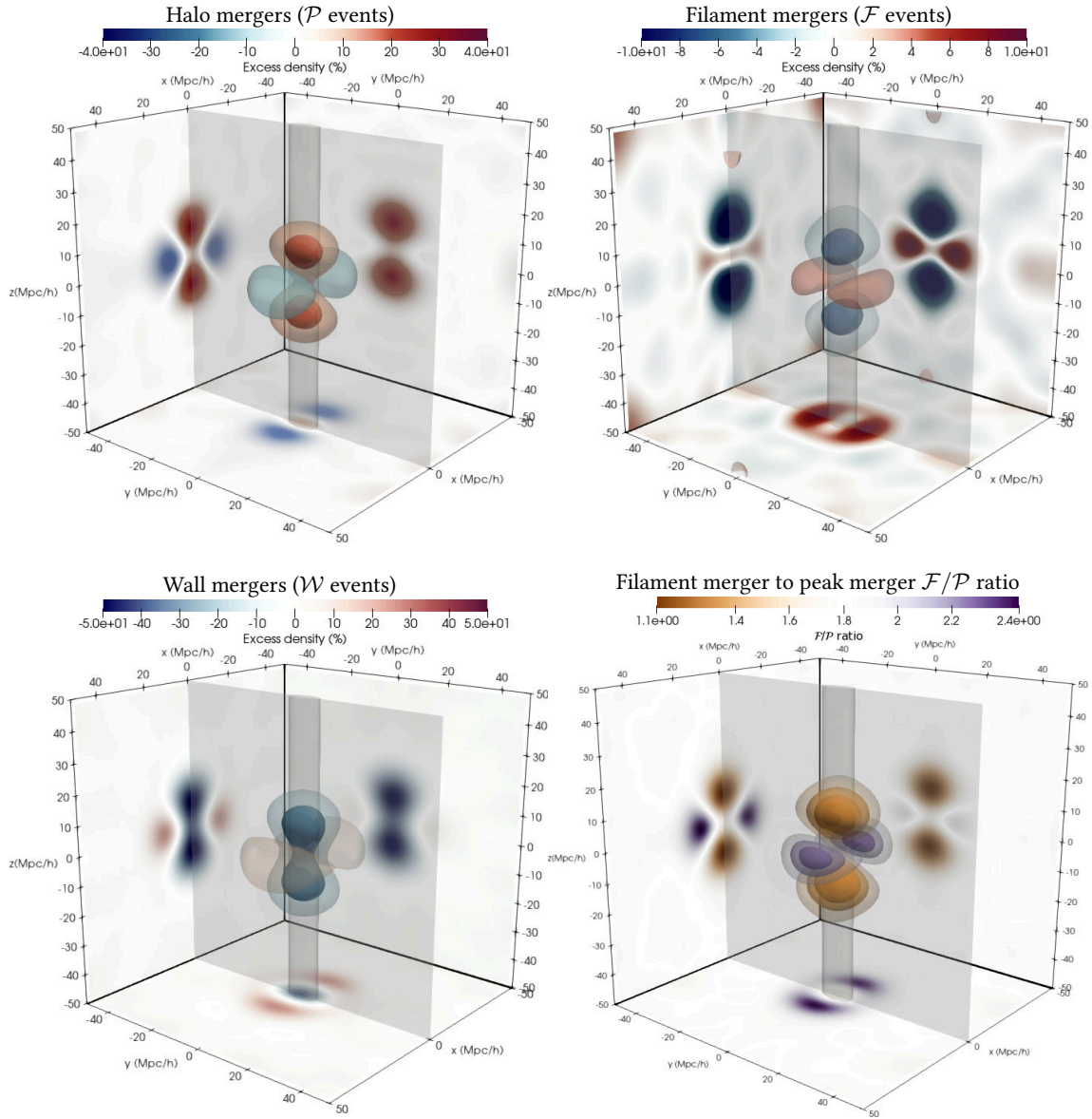
Figure 4.5.5: From left to right and top to bottom, peak-merger, filament-merger and wall-merger excess density around a large-scale proto-filament, illustrated by the vertical cylinder (z direction) and the wall in which it resides, illustrated by the grey plane (yz plane). The bottom right panel shows the local ratio of filament to peak mergers $r_{\mathcal{F}/\mathcal{P}}$. Each side of the cube shows a slice through the centre, shifted to the side of the plot for visualisation purposes. Red regions have an excess of critical events while blue regions have a deficit of critical events with respect to cosmic average. Interactive versions of these plots can be found online for the halo mergers, filament mergers, wall mergers and filament to peak merger ratio. Going from voids to wall, from wall to filament and from filament to the nearest node (along the z axis), the halo merger rate increases and the filament merger rate decreases. Halos in the filament are therefore stalled: they merge less than those in the nodes. At the same time, the filament merger rate decreases when going from the filament towards the node so that the mean connectivity, given by the ratio of halo merger to filament merger, is expected to increase.
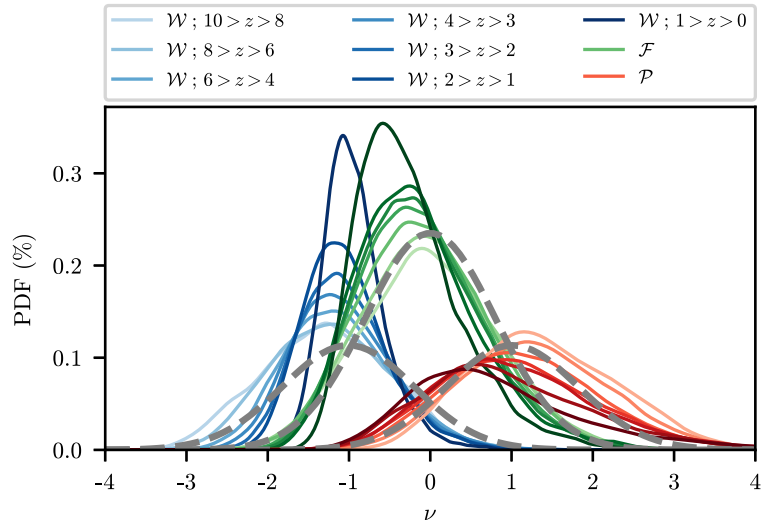
Figure 4.5.6: Critical events number count as a function of the rarity in dark-matter only simulations in different redshift bins as mentioned in the legend, with the same colours as figure 4.4.1. The curves have been normalised so that in each redshift bin, the integral of the three curves ($\mathcal{W}, \mathcal{P}, \mathcal{F}$) equals one. At high redshift, the merger rates resembles the Gaussian prediction (thick dashed grey lines, with an arbitrary normalisation). The skewness of the distributions increases with decreasing redshift as the field departs from gaussianity.

At high redshift, the Gaussian prediction is recovered. At lower redshift, the $\mathcal{P}$ and $\mathcal{F}$ counts shift towards lower contrast, but respectively decrease and increase in amplitude, while the $\mathcal{W}$ counts increase in amplitude. Since the first halos to merge are due to high σ peaks, it is expected that the low-z PDFs are biased towards low densities. Similarly, the mean density of filamentary structure decreases with increasing time, as the less dense filaments take more time to gravitationally form, so that the PDFs of the filament mergers shift to smaller densities at low z. The evolution of void structures with cosmological time is somehow symmetric to the evolution of peaks: early forming voids are the most underdense while late-time voids form out of less underdense regions. At fixed resolution, this results in a shift of the typical density of voids towards higher densities. Indeed, in the limit of infinite time, it is expected that the only voids found at a given size stem from $\nu = 0$, as any void with $\nu < 0$ will have had time to collapse earlier.

Comparison to galaxy catalogues

It is of interest to follow the position of all critical points (not just the maxima) explicitly as a function of true cosmic time in galaxy catalogue extracted from hydrodynamical simulations, so as to assess i) the impact of biasing involved in selecting specific tracers and ii) how non-linear clustering impacts the statistics. This was done by N. Cornuault and C. Pichon illustratively using 330 snapshots of galaxies extracted from HORIZON-AGN (shown on figure 4.1.1 at redshift zero with its set of walls and filaments), for which the critical points are derived using DISPERSE with a persistence threshold of $\sigma/100$. The algorithm described in section 4.B is used to match merging critical points as a function of redshift. The set of events are then binned as a function of log density for 4 redshift bins and shown on figure 4.5.7. Gravitational clustering has skewed the PDFs, but most dramatically galaxies poorly trace under dense regions, hence the number of wall mergers plummeted. While more work needs to be done in order to be conclusive, this
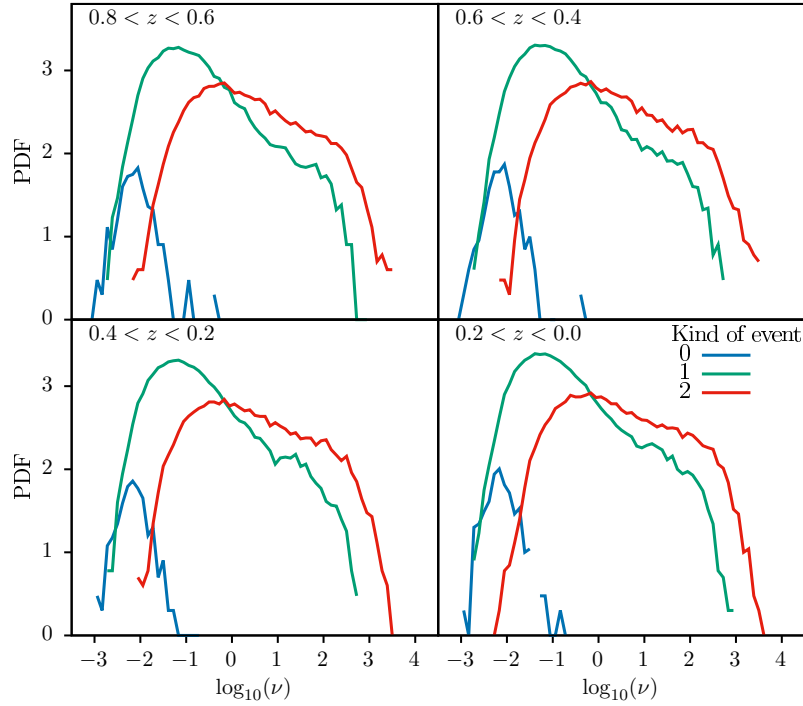
Figure 4.5.7: PDF of the critical events extracted from the galaxy catalogue of Horizon-AGN as a function of the log galaxy density for a range of cosmic time as labelled. When compared to figure 4.5.6, the PDFs of are strongly biased, with much fewer walls hence wall mergers detected. From Nicolas Cornuault, private communications.
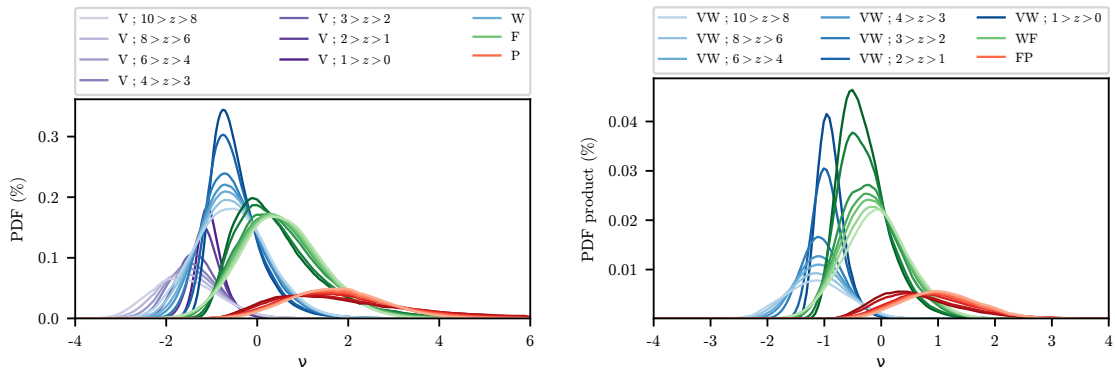
(a) Critical *point* PDF

(b) Critical *point* PDF product

Figure 4.5.8: (a): Critical point number counts as a function of the rarity in dark-matter only simulations in different redshift bins as mentioned in the legend. The curves have been normalised so that in each redshift bin, the integral of the four curves equals one. The purple bundle corresponds to voids, the blue one to walls, the green one to filaments and the red one to peaks. (b): Product of the PDFs. At large redshifts, the curves resemble the prediction of figure 4.2.6.

illustrates that the detection algorithms presented here can readily be applied to virtual catalogues in state-of-the-art simulations.

Void counts as a cosmological measurement

One particular application of these results is the study of the void number counts. Voids are very interesting laboratory both for galaxy evolution and cosmology. They represent primitive environments for galaxies, where density is low and matter flow is still relatively curl-free. Void galaxies are therefore interesting probes for galaxy formation (e.g. Lindner et al., 1996). Voids are also a tool of choice to probe the cosmology or to test theory of modified gravity (e.g. Gay et al., 2012; Lavaux and Wandelt, 2012; Cai et al., 2015) as a mean to constrain the equation of state of dark energy. In particular, these authors have used the cosmic evolution of the size and the number of voids as constrains on the linear matter growth function $D(z)$. In the present formalism void disappear as a function of cosmic time *via* mergers of walls, hence the one-point statistics of wall merger could be used as a cosmic probe. From equation (4.29) the cosmic evolution of the rate of void of volume $\mathcal{V}$ merging during time interval δz can be expanded to first order in σ *via* equation (4.39) as

$$\frac{\partial^2 n}{\partial \log \mathcal{V} \partial z} = \frac{\partial^2 n}{\partial \log \mathcal{V} \partial z}\Big|_{\mathrm{G}} + \sigma(z) \frac{\partial^2 n}{\partial \log \mathcal{V} \partial z}\Big|_{\mathrm{NG}}, \tag{4.40}$$

where the first term reflects cosmic evolution of the rate of void disappearance presented in section 4.5.1, while the second term is obtained by substituting $\partial^2 n/\partial R \partial \nu\big|_{\mathrm{G}}$ by $\partial^2 n/\partial R \partial \nu\big|_{\mathrm{NG}}$ into equation (4.39). As discussed in section 4.2.5, the scaling of these non-Gaussian corrections yield joint estimates for the cumulants (Codis et al., 2013), hence a measure of f_{NL} or a parametrisation of modified gravity.

4.6 Conclusion

As a proxy for cosmic evolution, I computed the rate of merging critical points as a function of smoothing scale from the primordial density field to forecast special events driving the assembly of dark halos and possibly galaxies. I considered all sets of critical points coalescence, including wall-saddle to filament-saddle and wall-saddle to minima, as they impact the topology of galactic infall, such as filament disconnection or void disappearance. The theory developed in this chapter, hereafter the "critical event theory", is central to the understanding of the effect of the cosmic web on the formation of galaxies, since their evolution is the result of their past history, which is usually encoded by their merger tree and the properties of their host halo. In this context, the critical event theory provides a way to encode not only the evolution of the halo hosting the galaxy *via* its merger tree, but also the evolution of its upcoming internal structure with time, which itself is responsible for driving the angular momentum acquisition, as will be seen in chapter 6. I argue that the theory can be seen as an extension to the classical halo model, where the properties of galaxies have an extra dependence on their "dressed" merger history which takes into account the merger history of their surrounding filaments and walls.

The scope of this theory is obviously very broad, but let me sum up here the results relevant to astrophysics and in particular to the problem of the assembly of galaxies.

- I studied critical events of all types, their clustering properties, and presented analytical formulae for the one-point statistics of these events in fields of dimensions up to 6, and also the two-point statistics.
- I have established the link between critical events and connectivity. This allows me to compute the connectivity of peaks and other critical events in arbitrary dimensions. Physically, I established the duality between the evolution of the cosmic web (critical events) and its topological features (connectivity).

- I provided a covariant formulation of the critical event theory which allowed me to also compute the two-point statistics for critical events. The two-point statistics show that halo mergers are typically followed by a filament merger, so that the connectivity is conserved.
- I have shown that the critical event theory can be further extended to take into account the early stages of non-linear gravitational evolution. This has then been compared positively to numerical simulations at high redshift. This extension also probes the non-Gaussianities that arise from primordial non-Gaussianities and can be used as a cosmological measurement.
- I have shown that halos forming in nodes grow by successive mergers, while their filaments do not merge, so that that their local connectivity increases. The trend is expected to reverse in voids, where filament mergers happen faster than halo mergers, resulting in halos with a small connectivity. This is likely to have an impact on galaxy formation, and in particular on angular momentum acquisition.

I have only touched on practical applications for the forecasting of special events in a multi-scale landscape. It should prove to be a fruitful field of research in astronomy and beyond in the future.

This work is part of an ongoing research effort and will lead to a publication in the near future.

4.A   Critical events in ND

For the sake of completeness and possible interest in other fields of research, let us present the one-point statistics of critical events in arbitrary dimension d.

4.A.1   Joint PDF of the field and its second derivatives

From Pogosyan et al., 2009b the probability of measuring the set of d eigenvalues of the d dimensional Hessian $\{\lambda_i\}$ and density ν obeys

$$\mathcal{V}_d \prod_{i \leq d} d\lambda_i \prod_{i < j} (\lambda_j - \lambda_i) \exp\left(-\frac{1}{2} Q_\gamma(\nu, \{\lambda_i\})\right), \tag{4.41}$$

where Q_γ is a quadratic form in λ_i and ν given by

$$Q_\gamma(\nu, \{\lambda_i\}) = \nu^2 + \frac{\left(\sum_i \lambda_i + \gamma\nu\right)^2}{(1 - \gamma^2)} + \mathcal{Q}_d(\{\lambda_i\}), \tag{4.42}$$

with

$$\mathcal{Q}_d(\{\lambda_i\}) = (d + 2) \left[\frac{1}{2}(d - 1) \sum_i \lambda_i^2 - \sum_{i \neq j} \lambda_i \lambda_j\right]. \tag{4.43}$$

In equation (4.41) $\mathcal{V}_d$ arises from the integration over the angles and is given by equation (4.56) below.

4.A.2   Joint PDF of the first and third derivatives

Here, we will look into the PDF of the first and third derivatives in d dimensions in order to compute the odd derivative term C_{odd} that enters critical event number counts in d dimensions.

First, let us note that the first derivatives are Gaussian distributed with individual variance $\langle x_i^2 \rangle = 1/d$ so that the probability for all first derivatives to be zero is

$$P(\nabla\rho = 0) = \left(\frac{d}{2\pi}\right)^{d/2}. \tag{4.44}$$

Now let us study the statistics of the third derivatives. By symmetry, one can note that

$$\left\langle \left(\sum_i x_{1ii} \right)^2 \right\rangle = \frac{1}{d}, \tag{4.45}$$

because the third derivatives are rescaled by σ_3, and

$$\langle x_{1jj}^2 \rangle = \langle x_{111} x_{1jj} \rangle = \frac{1}{5} \langle x_{111}^2 \rangle = 3 \langle x_{1jj} x_{1kk} \rangle \quad \forall j \neq k \neq 1.$$

Therefore,

$$\frac{1}{d} = \langle x_{111}^2 \rangle + (d-1) \langle x_{1jj}^2 \rangle + 2(d-1) \langle x_{111} x_{1jj} \rangle + (d-1)(d-2) \langle x_{1kk} x_{1jj} \rangle \quad \forall j \neq k \neq 1 \tag{4.46}$$

implies that $\langle x_{iii}^2 \rangle = 15/d(d+2)(d+4)$ and the full covariance matrix of the third derivatives is therefore now known. However, we are interested in statistics subject to a zero gradient constraint, in particular the three quantities of interest are (fixing d as the degenerate direction and assuming an implicit summation on the i indices)

$$\langle x_{ddd}^2 | x_d = 0 \rangle = \langle x_{ddd}^2 \rangle - \frac{\langle x_{dii} x_d \rangle^2}{\langle x_d^2 \rangle}, \tag{4.47}$$

$$\left\langle (x_{dii})^2 | x_d = 0 \right\rangle = \left\langle (x_{dii})^2 \right\rangle - \frac{\langle x_{ddd} x_d \rangle^2}{\langle x_d^2 \rangle}, \tag{4.48}$$

$$\langle x_{dii} x_{ddd} | x_d = 0 \rangle = \langle x_{dii} x_{ddd} \rangle - \frac{\langle x_d x_{ddd} \rangle \langle x_d x_{dii} \rangle}{\langle x_d^2 \rangle}, \tag{4.49}$$

which can easily be computed thanks to the additional relation $\langle x_{11}^2 \rangle = 3/d(d+2)$,

$$\langle x_{ddd}^2 | x_d = 0 \rangle = \frac{3}{d(d+2)} \left[\frac{5}{d+4} - \frac{3\tilde{\gamma}^2}{d+2} \right], \tag{4.50}$$

$$\left\langle (x_{dii})^2 | x_d = 0 \right\rangle = \frac{1 - \tilde{\gamma}^2}{d}, \tag{4.51}$$

$$\langle x_{dii} x_{ddd} | x_d = 0 \rangle = \frac{3}{d(d+2)} (1 - \tilde{\gamma}^2). \tag{4.52}$$

4.A.3   Critical event number counts in ND

It now follows that the critical event number counts of type j at height ν in dimension d read

$$\frac{\partial^2 n_j^d}{\partial R \partial \nu} = \frac{R \, \mathcal{V}_d \, C_{d,\text{odd}}}{\tilde{R}^2 \, R_*^d} \left\langle \delta_{\text{D}}(\lambda_j) \left| \prod_{i \neq j \leq d} \vartheta_{\text{H}}(\lambda_i - \lambda_j) \lambda_i \right| \right\rangle, \tag{4.53}$$

where this expectation is computed using the conditional expectations presented in the previous section. Equation (4.53) is a function of ν because of the correlation between ν and $\sum_i \lambda_i$ seen in equation (4.42). Recalling the formal analogy with the flux of critical lines per unit hyper surface,

$$\frac{\partial^2 n_{\mathcal{P}}^d}{\partial R \partial \nu} \overset{\gamma \nu \to \infty}{\sim} \frac{R}{\tilde{R}^2 \, R_*^d} \frac{\mathcal{V}_d \, C_{d,\text{odd}}}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} \nu^2 \right] \left(\frac{\nu}{R_0} \right)^{d-1},$$

in the large d large ν limit (Pogosyan et al., 2009b). The contribution from the odd part of the distribution function, $C_{d,\text{odd}}$ obeys

$$C_{d,\text{odd}} = \left\langle \left| \sum_i x_{jii} \right| |x_{jjj}| \delta_{\text{D}}^{(d)}(x_i) \right\rangle, \tag{4.54}$$
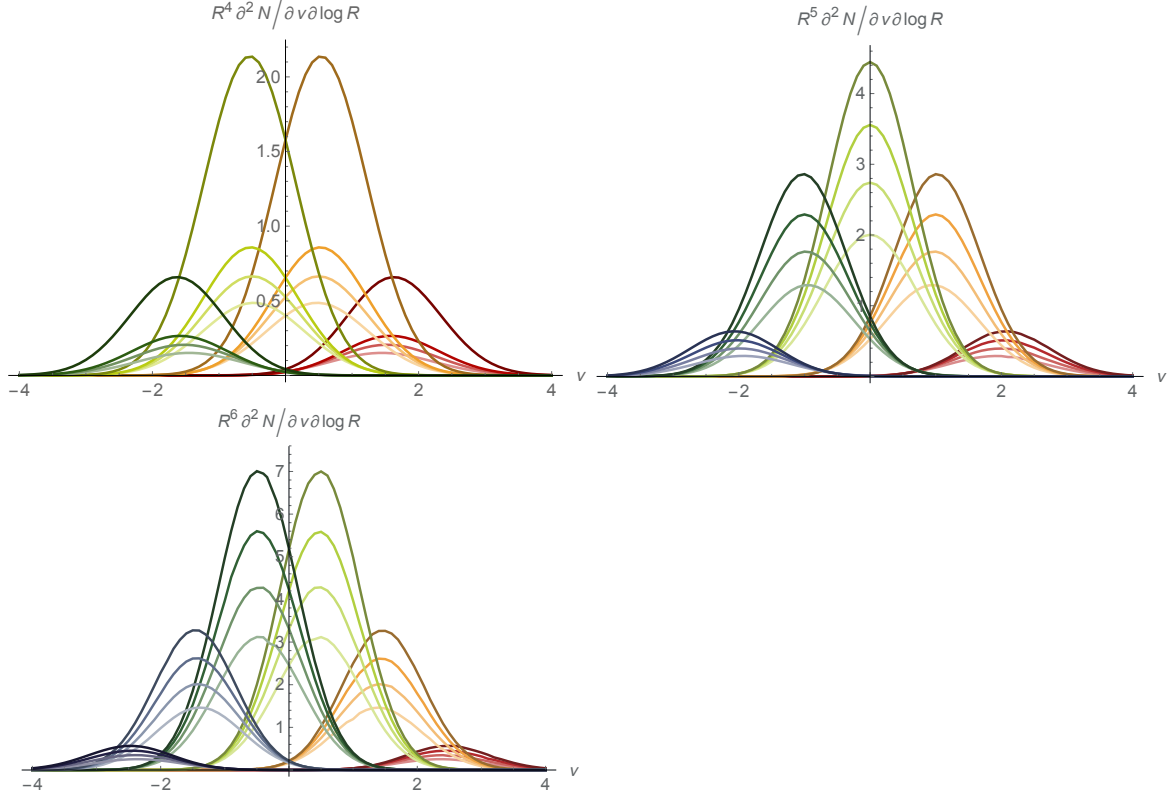
Figure 4.A.1: The PDF of critical events of the various types $(\mathcal{P}, \mathcal{F}, \mathcal{W}_1, \mathcal{W}_2)$ in 4D (*left*), in 5D (*right*) and 6D (*bottom*) for $n_{\rm s} = -2, -3/2, -1, -1/2$ from light to dark.

where the expectation in equation (4.54) should be computed with the odd derivative PDF given in section 4.A.2. After a bit of algebra,

$$
C_{d,\mathrm{odd}} = \left(\frac{d}{2\pi}\right)^{\frac{d}{2}} \left[\frac{2\sqrt{6}}{\pi} \sqrt{\frac{(d-1)\left(1-\tilde{\gamma}^2\right)}{d^2(d+2)^2(d+4)}} + \frac{6\left(1-\tilde{\gamma}^2\right)}{\pi d(d+2)} \tan^{-1}\left(\sqrt{\frac{3}{2}} \frac{\sqrt{d+4}\sqrt{1-\tilde{\gamma}^2}}{\sqrt{d-1}}\right) \right].
$$
(4.55)

Finally, the volume $\mathcal{V}_d$ of the hyper-wedge corresponding to the marginalisation over the orientation of the Hessian obeys

$$
\mathcal{V}_d = \frac{1}{2^{d-1}d!} \iint \mathrm{dSO}(d) = \frac{1}{2^{d-1}d!} \prod_{i=1}^{n-1} \mathrm{Vol}(S^i),
$$

$$
= \frac{1}{2^{d-1}d!} \prod_{i=1}^{n-1} \frac{2\pi^{(i+1)/2}}{\Gamma((i+1)/2)},
$$
(4.56)

where $\mathrm{Vol}(S^i)$ denotes the i-dimensional volume (i.e. surface area) of the unit i-sphere in $\mathbb{R}^{i+1}$, the factor $d!$ comes from not sorting the eigenvalues and the factor 2^{d-1} from not imposing their sign. It follows that $\mathcal{V}_2 = \pi/2$, $\mathcal{V}_3 = \pi^2/3$, $\mathcal{V}_4 = \pi^4/12$, $\mathcal{V}_5 = \pi^6/45$ and $\mathcal{V}_6 = \pi^9/540$. The PDFs of critical events in 4D, 5D and 6D are shown in figure 4.A.1. Note that the intermediate signature events dominate in number over the extreme ones, in accordance with the relative number of critical points.

4.A.4 Ratios of critical events

From equation (4.42), the integration over ν yields the marginal probability of $\{\lambda_i\}$

$$\mathcal{V}_d \prod_{i\leq d} d\lambda_i \prod_{i<j} (\lambda_j - \lambda_i) \exp\left(-\frac{1}{2}\mathcal{Q}_d(\{\lambda_i\}) - \frac{1}{2}\left(\sum_i \lambda_i\right)^2\right). \tag{4.57}$$

Finally, the d dimensional ratio of critical event of type j and k is simply given by

$$r_{j/k} = \frac{\left\langle \delta_{\mathrm{D}}(\lambda_j) \left| \prod_{i\neq j} \vartheta_{\mathrm{H}}(\lambda_i - \lambda_j)\,\lambda_i \right| \right\rangle}{\left\langle \delta_{\mathrm{D}}(\lambda_k) \left| \prod_{i\neq k} \vartheta_{\mathrm{H}}(\lambda_i - \lambda_k)\,\lambda_i \right| \right\rangle}, \tag{4.58}$$

where the PDF to evaluate this expectation is given by equation (4.57). Note that these counts correspond to the area below each curve shown in figure 4.A.1. In 3D, we recover the ratio presented in the main text. In 4D the ratio is analytic and reads

$$r_{\mathcal{F}/\mathcal{W}} = \frac{2(57 + 25\pi - 50\cot^{-1}(3))}{75\pi - 2(57 + 50\cot^{-1}(2))} \approx 3.17. \tag{4.59}$$

More generally,

$$
\begin{aligned}
d = 2: \quad & r_{\mathcal{F}/\mathcal{W}} = 1\,, \\
d = 3: \quad & r_{\mathcal{F}/\mathcal{P}} = 2.06\,, \\
d = 4: \quad & r_{\mathcal{F}/\mathcal{P}} = 3.17\,, \quad r_{\mathcal{W}/\mathcal{P}} = 3.17\,, \\
d = 5: \quad & r_{\mathcal{F}/\mathcal{P}} = 4.36\,, \quad r_{\mathcal{W}_1/\mathcal{P}} = 6.72\,, \quad r_{\mathcal{W}_2/\mathcal{P}} = 4.36\,, \\
d = 6: \quad & r_{\mathcal{F}/\mathcal{P}} = 5.67\,, \quad r_{\mathcal{W}_1/\mathcal{P}} = 11.97\,, \quad r_{\mathcal{W}_2/\mathcal{P}} = 11.97\,, \quad r_{\mathcal{W}_3/\mathcal{P}} = 5.67\,.
\end{aligned}
$$

Note that these ratios are pure numbers and do not depend on the detailed shape of the underlying powerspectrum (for Gaussian random fields).

4.A.5 Self-consistency links with critical points counts

These results can be used to derive the connectivity as defined in Codis et al., 2018. Indeed, let us formally write N_i the number density of critical point of kind i in d dimensions and $\mathcal{N}_i$ the number density of critical event of kind $i - i + 1$. The evolution of N_i is given by

$$\frac{\partial N_i}{\partial R} = -\begin{cases} \mathcal{N}_0 & \text{if } i = 0, \\ (\mathcal{N}_{i-1} + \mathcal{N}_i) & \text{if } 0 < i < d-1, \\ \mathcal{N}_{d-1} & \text{if } i = d-1. \end{cases} \tag{4.60}$$

For Gaussian random fields, the number density of critical point can be formally written as

$$N_i = \frac{1}{R_*^d} \underbrace{\left\langle \left| \prod_j \lambda_j \right| \right\rangle \left\langle \delta_{\mathrm{D}}^{(3)}(x_i) \right\rangle}_{C_i},$$

where the PDF to evaluate the left part of the r.h.s. is given by equation (4.57). Here C_i is a number common to all power spectra. The derivative of N_i with respect to the smoothing scale is then

$$\frac{\partial N_i}{\partial R} = -N_i \times d\frac{\mathrm{d}\log R_*}{\mathrm{d}R}. \tag{4.61}$$

Using equation (4.60) and equation (4.61) yields a simple relation between the number density of critical points and the number density of critical events

$$
N_i = \frac{1}{d \times \mathrm{d} \log R_*/\mathrm{d}R}
\begin{cases}
\mathcal{N}_0 & \text{if } i = 0, \\
(\mathcal{N}_{i-1} + \mathcal{N}_i) & \text{if } 0 < i < d-1, \\
\mathcal{N}_{d-1} & \text{if } i = d-1.
\end{cases}
$$

For Gaussian random fields, one has the property that $N_i = N_{d-i-1}$ and $\mathcal{N}_i = \mathcal{N}_{d-i-2}$. This follows from the fact that the field δ is invariant under sign change so that $-\delta$ has the same properties. This provides us with simple way to compute the ratio of critical events as a function of the ratio of the critical points. For any d, the ratio of filament to peak is connected to the ratio of $\mathcal{F}$ to $\mathcal{P}$ critical events

$$
\frac{N_1}{N_0} = \frac{\mathcal{N}_0 + \mathcal{N}_1}{\mathcal{N}_0} = 1 + \frac{\mathcal{N}_1}{\mathcal{N}_0} = 1 + r_{\mathcal{F}/\mathcal{P}}. \tag{4.62}
$$

As an example, let use derive the ratio of other critical points in dimensions up to 6D. For $d = 4$,

$$
\frac{N_1}{N_0} = \frac{N_2}{N_3} = 1 + r_{\mathcal{F}/\mathcal{P}} \approx 4.17,
$$
$$
\frac{N_2}{N_1} = \frac{\mathcal{N}_1 + \mathcal{N}_2}{\mathcal{N}_0 + \mathcal{N}_1} = \frac{\mathcal{N}_0 + \mathcal{N}_1}{\mathcal{N}_0 + \mathcal{N}_1} = 1.
$$

For $d = 5$,

$$
\frac{N_1}{N_0} = \frac{N_3}{N_4} = 1 + r_{\mathcal{F}/\mathcal{P}} \approx 5.36,
$$
$$
\frac{N_2}{N_1} = \frac{N_2}{N_3} = \frac{\mathcal{N}_1 + \mathcal{N}_2}{\mathcal{N}_0 + \mathcal{N}_1} = \frac{r_{\mathcal{F}/\mathcal{P}} + r_{\mathcal{W}_1/\mathcal{P}}}{1 + r_{\mathcal{F}/\mathcal{P}}} \approx 2.07.
$$

For $d = 6$,

$$
\frac{N_1}{N_0} = \frac{N_4}{N_5} = 1 + r_{\mathcal{F}/\mathcal{P}} \approx 6.67,
$$
$$
\frac{N_2}{N_1} = \frac{N_3}{N_4} = \frac{\mathcal{N}_1 + \mathcal{N}_2}{\mathcal{N}_0 + \mathcal{N}_1} = \frac{r_{\mathcal{F}/\mathcal{P}} + r_{\mathcal{W}_1/\mathcal{P}}}{1 + r_{\mathcal{F}/\mathcal{P}}} \approx 2.64,
$$
$$
\frac{N_3}{N_2} = 1.
$$

Given that Codis et al., 2018 provides an asymptotic limit for the connectivity, I can re-express it in terms of the ratio of critical events as

$$
\frac{N_1}{N_0} = \frac{N_{d-2}}{N_{d-1}} = 1 + r_{\mathcal{F}/\mathcal{P}} = d + \frac{1}{2}\left((2d-4)/7\right)^{7/4}, \tag{4.63}
$$

which in the large d limit, asymptotes to

$$
r_{\mathcal{F}/\mathcal{P}} \overset{d \to \infty}{\sim} \frac{1}{2}\left(\frac{2}{7}\right)^{7/4} d^{7/4} \approx \frac{1}{17} d^{7/4}. \tag{4.64}
$$
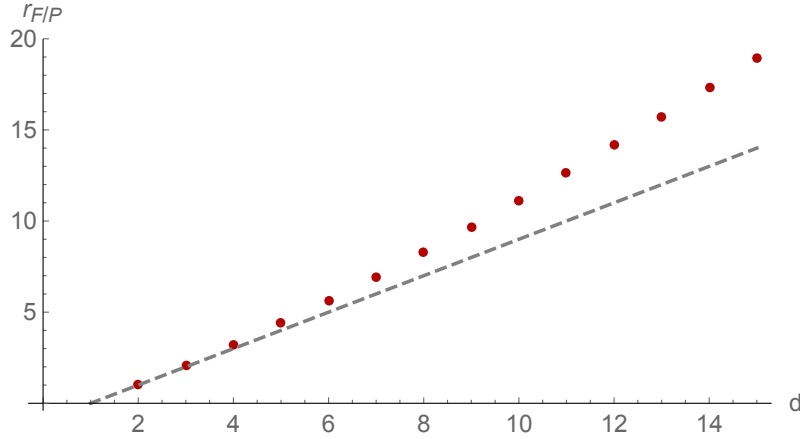
Figure 4.A.2: The ratio of peak to filament merger as a function of d. For reference, the first diagonal is shown as a dashed grey line as well. The ratio is approximately fitted as $d-1+((2d-4)/7)^{7/4}/2$ and shown as red dots. The dashed line is the identity.

4.A.6 Testing the link between critical pts and events counts

From equation (4.61) and because for a Gaussian filter, we have

$$\frac{\mathrm{d}\sigma_i^2}{\mathrm{d}R^2} = -\sigma_{i+1}^2,$$

one can easily derive

$$\frac{\partial N_i}{\partial R} = -N_i \times d\frac{R}{R_\star^2}\frac{1-\tilde{\gamma}^2}{\tilde{\gamma}^2} \tag{4.65}$$

which in $d = 3$ for peaks reads

$$-\frac{\partial N_0}{\partial R} = 3N_0\frac{R}{R_\star^2}\frac{1-\tilde{\gamma}^2}{\tilde{\gamma}^2} \tag{4.66}$$

$$= \frac{3R}{R_\star^3\tilde{R}^2}(1-\tilde{\gamma}^2)\frac{29\sqrt{15}-18\sqrt{10}}{1800\pi^2} \tag{4.67}$$

which happens to be equal to the differential number counts of 3D critical events (equation (4.14)) but only if C_{odd} is computed with the approximation in equation (4.13) that boils down to (using equation (4.50))

$$C_{\mathrm{odd}} \approx \frac{3(1-\tilde{\gamma}^2)}{d(d+2)}\left(\frac{d}{2\pi}\right)^{d/2}. \tag{4.68}$$

The discrepancy is however tiny, as a result of the transverse third derivatives being only very weakly correlated to the on-diagonal terms, *i.e.* $\langle x_{ijj}x_{iii}\rangle$ is small. More work will however be required to understand the origin of this disagreement.

4.B Algorithms

The source code of the implementation can be found online. It is based on Python and the Scipy stack (E. Jones et al., 2001).

4.B.1   Critical points detection

This section presents the algorithm used to find the extrema in a N-dimensional field. Let F, F_i and F_{ij} be a field evaluated on a grid, its derivative and its Hessian. For any point $\boldsymbol{x}$ on the grid, we have the following relation

$$F_j(\boldsymbol{x}) = F_j(\boldsymbol{x_c}) + (x_i - x_{c,i})F_{ij}(\boldsymbol{x}) + \mathcal{O}(\Delta x_i^2). \tag{4.69}$$

Critical points are found where $F_j' = 0$ by solving the linear system of equation

$$\Delta x_i F_{ij} = -F_j, \tag{4.70}$$

where $\boldsymbol{\Delta x} = \boldsymbol{x} - \boldsymbol{x}_c$. The algorithm works as follows:

1. Solve equation (4.70) for each cell on the grid. We then get a set of points $(\boldsymbol{x_c^i}, \boldsymbol{x^i})$, where the former is the cell centre and the latter the closest critical point.
2. Remove all critical points found at $|\boldsymbol{x_c^i}, \boldsymbol{x^i}|_\infty \geq \Delta x$, where Δx is the grid spacing.
3. For all critical point, compute the value of the Hessian by interpolating linearly from the $2N$ (4 in 2D, 6 in 3D) neighbouring cells.
4. Compute the eigenvalues of the Hessians and the type of the critical point (maximum, saddle point(s) or minimum).
5. Merge all critical points of the same kind closer than Δx. To do this, we first build a KD-Tree of the critical points and find all the pairs located at a distance $d_{ij} = |\boldsymbol{x^i} - \boldsymbol{x^j}|_\infty \leq \Delta x$. For each pair, we keep only the point that is the closest to its associated cell.

4.B.2   Critical events detection

The algorithm is based on the idea that each critical event has two predecessors at the previous smaller smoothing scale (two critical points). Conversely, each critical point has either a critical point successor of the *same kind* at the next (larger) smoothing scale or a critical event. Therefore, a way to detect critical events is to find critical points that do not have a successor. These points will be referred to as "heads" as they are the tip of a continuous line of critical points in the smoothing scale direction. Critical events are then found between pairs of heads of kind k and $k + 1$ (e.g. a peak and a filament).

Following this idea, the algorithm can be decomposed in two steps: compute the heads of each kind, than find pairs of heads to detect critical events. In the following of the section, let us call R_0 (resp. R_1) the smallest (resp. largest) scale at which the field is smoothed. Let $C_{R,k} = \{\boldsymbol{r}_i, R\}_{i=1,\dots,N}$ be the set of the N critical points of kind k at scale R. The whole detection algorithm reads

1: **procedure** FindCritEvents($C_{R,k}, \alpha$)
2:      $E \leftarrow \{\}$ ▷ All critical events
3:      **for** k **in** $1, \dots, d$ **do**   ▷ Find heads of critical points
4:          $H_k \leftarrow$ BuildHeads($k, \Delta \log R$)
5:      **end for**
6:      $R \leftarrow R_0$
7:      **while** $R \leq R_1$ **do**   ▷ Find pairs of heads (crit. events)
8:          $\Delta R \leftarrow R \times \Delta \log R$ ▷
9:          $E \leftarrow E +$ FindHeadPairs($H_1, \dots, H_d, R, \alpha \Delta R$)
10:         $R \leftarrow R + \Delta R$
11:     **end while**
12:     **return** E
13: **end procedure**

The parameter α controls how far heads can be in the smoothing scale direction, in units of $\log R$. A value of 1 looks for pairs of heads at the same scale, a value of 2 looks for pairs of heads at a scales $R, R + \Delta R$.

The first step (line 4) of the algorithm builds the set of heads H_k. It works as follows

```
 1: procedure BUILDHEADS(k, Δ log R)                          ▷ Build heads of kind k
 2:     H_k ← C_{R_1,k}                                        ▷ Initialize heads
 3:     P_k ← H_k                                              ▷ Initialize progenitors
 4:     R ← R_1
 5:     while R ≥ R_0 do
 6:         P'_k ← {}                                          ▷ Initialize new progenitors at R
 7:         for p, c, d in SORTEDPAIRS(P_k, C_{R,k}, R) do
 8:             if c ∉ P'_k then
 9:                 P'_k ← P'_k + {p, c}                       ▷ Found new progenitor
10:             end if
11:         end for
12:         P_k ← P'_k
13:         for c in C_{R,k} do                                ▷ Loop over crit. points
14:             if c ∉ P'_k then                               ▷ Keep only unpaired ones…
15:                 H_k ← H_k + {c}                            ▷ …and add them to heads
16:                 P_k ← P_k + {c}
17:             end if
18:         end for
19:         R ← R(1 − Δ log R)
20:     end while
21:     return H_k                            ▷ Heads are points with no successors at larger R
22: end procedure
```

Here, SortedPairs$(X, Y, R_{\max})$ returns (x, y, d), where x, y are points in X, Y and $d \le R_{\max}$ is their relative distance (in $(\boldsymbol{r}, R)$ space). The tuples are sorted by increasing distance. This can be efficiently implemented using a KD-tree with periodic boundary conditions. BuildHeads builds all heads by using a watershed approach. Starting from the largest smoothing scales, it finds and discards all critical events that are progenitors of a head at any larger scale. The remaining points have no successor (they are the progenitor of nothing) and are hence heads.

Once the heads have been computed, the second step of the algorithm pairs them (line 9)

```
 1: procedure FINDHEADPAIRS(H_1, …, H_d, R, ΔR)          ▷ Find pairs of heads (crit. events)
 2:     H_{R,k} ← {c ∈ H_k | R ≤ c.R < R + ΔR}               ▷ Keep heads at scale R
 3:     P ← {}                                               ▷ Head pair list
 4:     for k in 1, …, d − 1 do
 5:         P ← P+SORTEDPAIRS(H_{R,k}, H_{R,k+1}, R)
 6:         P ← P+SORTEDPAIRS(H_{R,k+1}, H_{R,k}, R)
 7:     end for
 8:     P ← SORTBYDISTANCE(P)
 9:     P' ← {}                                              ▷ Pairs with no double counts
10:     for c_1, c_2, d in P do
11:         if c_1 ∉ P' and c_2 ∉ P' then
12:             P' ← P' + {c_1, c_2}
13:         end if
14:     end for
15:     E ← {}                                               ▷ Critical events
16:     for c_1, c_2 in P' do
```

17:          $E \leftarrow E+$ CRITEVENTDATA(c_1, c_2)
18:     **end for**
19:     **return** E
20: **end procedure**

Lines 5-6 ensure that the detection method is invariant by permutation of $k \leftarrow d - k + 1$. CritEventData(c_1, c_2) computes the properties (position, kind, gradient, …) of the critical events given two critical points. FindHeadPairs works as follows. It first finds all pairs of heads separated by less than a smoothing scale. It then loops over all pairs (sorted by increasing distance) and greedily consumes heads. Each head can only be paired once, to its closest not-yet-paired head of either the previous or next kind. This prevents for example F critical points from being paired to a P and a W critical point, which would result in a double count. Note that this procedure may leave some heads unpaired (e.g. critical points at the largest smoothing scale do not merge but have no successor). In practice the unpaired heads typically account for less than a percent (0.5 % for $\Delta R = \alpha R \Delta \log R$ with $\alpha = 2$) of the total number of heads.

An alternative to the present algorithm could involve modifying DISPERSE to only retain the points of lowest persistence.

4.B.3 Generation algorithm

I have used CONSTRFIELD coupled with MPGRAFIC from Prunet et al., 2008 to generate constrained realisations of a Gaussian random field. I generate an unsmoothed Gaussian random field, constrained to have a filament-type saddle point of height $\delta = 1$ ($\nu = 1.17$) at smoothing scale $R = 5 \, \mathrm{Mpc}/h$. The eigenvalues of the Hessian are constrained to be $\{\lambda_1, \lambda_2, \lambda_3\} = \sigma_2\{-1/2, -1/2, -1\}$ with eigenvectors $\{\hat{x}, \hat{y}, \hat{z}\}$. Figure 4.B.1 shows the mean density profiles as well as one realisation. As expected, the density is locally entirely set by the constrain and has a parabola-like shape. At larger scales, the field decouples from the constrains resulting in large fluctuations around the mean value.

4.B.4 Comparison of two-point correlation function estimators

In the field of cosmology, some efforts (see Kerscher et al., 2000, and references therein) have been dedicated to build unbiased estimators of the two-point correlations. Indeed, such estimator are impacted by the size of the sample as well as finite volume effects if the catalog does not cover the entire sky. Because of periodic boundaries, I do not have problem with the size of the box. The estimator used in this work is

$$\xi_{AB} = \frac{\langle AB \rangle}{f \sqrt{\langle AR_A \rangle \langle BR_B \rangle}}, \tag{4.71}$$

where A, B are two catalogs and R_A, R_B are random samples with $1/f$ times more data than A, B respectively. I have compared it to the popular Landy-Szalay (LS) estimator (Landy and Szalay, 1993; Szapudi and Szalay, 1999)

$$\xi_{AB,\mathrm{LS}} = \frac{\langle (A - R_A/f)(B - R_B/f) \rangle}{\langle R_A R_B \rangle / f^2}. \tag{4.72}$$

The results are shown on figure 4.B.2. At large scales, both estimators converge to the expected value of one. However at small scales, the LS estimator is more noisy since at small scales, no pairs AB are found so that the estimator of equation (4.71) returns 0, while the LS estimator includes contributions from AR_B and BR_A pairs and returns a non-null, noise-dominated signal. Following a pragmatic approach I have used throughout all our analysis the estimator of equation (4.71).
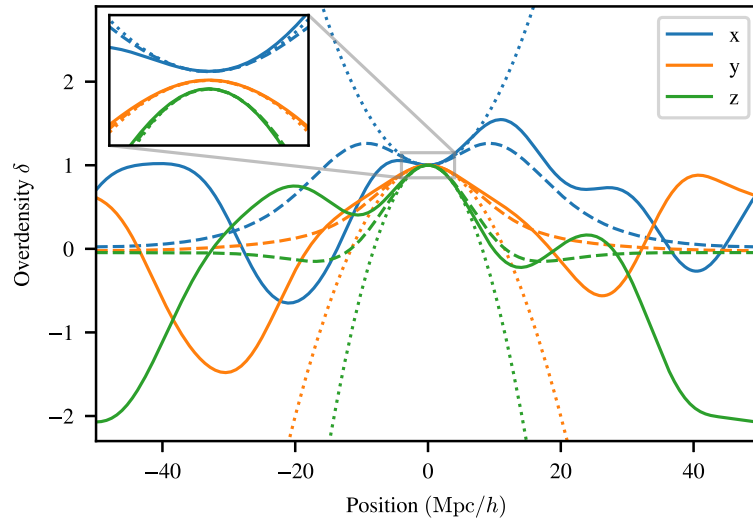
Figure 4.B.1: Density profile of a random field constrained to a density $\delta = 1$, null gradient and a Hessian with eigenvalues $\sigma_2/2, -\sigma/2, -\sigma$ in directions x, y, z at the centre of the box, assuming periodic boundary conditions. The expectation of the field is shown in dashed lines and the value of the field in one realisation is shown in solid lines. Dotted lines show the second order Taylor series of the field around the constrained point. The inset shows a zoom on the constrained zone. For the sake of clarity, each curve have been shifted by 0.02. At small distances from the constrain, the field resembles its mean and its Taylor expansion.
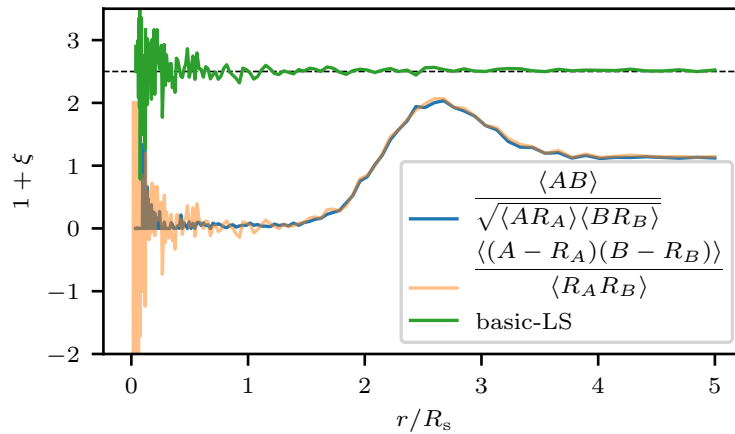
Figure 4.B.2: $\mathcal{PP}$ correlation function in the 2D case using the estimator of equation (4.71) (blue line) vs the Landy-Szalay estimator (light orange line). The difference (green line) has been shifted by 2.5 for visualisation purposes. The LS estimator yields a correlation function that is more noisy at small separations.

4.C   Joint PDFs

Let us present here the PDF of the field and its (up to 3rd) derivative which will allow us to compute the expectations involved in the main text.

4.C.1   One-point PDFs

Since the odd and even variables of Gaussian random fields do not correlate, let us write the joint PDF as $P_{\mathrm{G}} = P_0(x, x_{kl}) P_1(x_i, x_{ijk})$. The expression for $P_0(x, x_{kl})$ for the Gaussian field was first given by Bardeen et al., 1986. Introducing the variables

$$u \equiv -\Delta x = -(x_{11} + x_{22} + x_{33}), \tag{4.73}$$

$$w \equiv \frac{1}{2}(x_{11} - x_{33}), \tag{4.74}$$

$$v \equiv \frac{1}{2}(2x_{22} - x_{11} - x_{33}), \tag{4.75}$$

in place of diagonal elements of the Hessian (x_{11}, x_{22}, x_{33}) one finds that $u, v, w, x_{12}, x_{13}, x_{23}$ are uncorrelated. Importantly, the field, x is only correlated with u and

$$\langle xu \rangle = \gamma, \quad \langle xv \rangle = 0, \quad \langle xw \rangle = 0, \quad \langle xx_{kl} \rangle = 0, \; k \neq l,$$

where γ is the same quantity as in equation (2.97). The full expression of $P_0(x, x_{kl})$ is then

$$P_0(x, x_{kl}) = \frac{5^{1/2}15^2}{(2\pi)^{7/2}(1 - \gamma^2)^{1/2}} \exp\left(-\frac{1}{2}\left[Q_0 + Q_2\right]\right),$$

with the quadratic forms Q_0 and Q_2 given by

$$Q_0 = x^2 + \frac{(u - \gamma x)^2}{(1 - \gamma^2)}, \tag{4.76}$$

$$Q_2 = 5v^2 + 15(w^2 + x_{12}^2 + x_{13}^2 + x_{23}^2),$$

$$= \frac{15}{2}\, \overline{x}_{ab}\overline{x}_{ab}, \tag{4.77}$$

where the last identity is demonstrated in Pogosyan et al., 2009b and involves the detraced tensors:

$$\overline{t}_{ij} = t_{ij} - \frac{1}{3}t_{aa}\delta_{ij}, \tag{4.78}$$

$$\overline{t}_{ijk} = t_{ijk} - \frac{3}{5}\, t_{aa(j}\delta_{kl)}, \tag{4.79}$$

with an implicit summation over repeated indices and symmetrization between parenthesised indices (for instance: $t_{aa(j}\delta_{kl)} = [t_{aaj}\delta_{kl} + t_{aak}\delta_{lj} + t_{aal}\delta_{jk}]/3$ and so on). Equation (4.76) depends only on a single correlation parameter: γ. A similar procedure can be performed for the joint probability of the first and third derivatives of the fields, $P_1(x_i, x_{ijk})$ by defining the following nine parameters (see also Hanami, 2001)

$$u_i \equiv \nabla_i u, \quad v_i \equiv \frac{1}{2}\epsilon^{ijk}\nabla_i\left(\nabla_j\nabla_j - \nabla_k\nabla_k\right)x, \; \text{with } j < k,$$

$$w_i \equiv \sqrt{\frac{5}{12}}\nabla_i\left(\nabla_i\nabla_i - \frac{3}{5}\Delta\right)x, \tag{4.80}$$

and replacing the variables $(x_{i11}, x_{i22}, x_{i33})$ with (u_i, v_i, w_j). In that case, the only cross-correlations in the vector $(x_1, x_2, x_3, u_1, v_1, w_1, u_2, v_2, w_2, u_3, v_3, w_3, x_{123})$ which do not vanish are between the same components of the gradient and the gradient of the Laplacian of the field:

$$\langle x_i u_i \rangle \;\; = \;\; \tilde{\gamma}/3, \quad i = 1, 2, 3, \tag{4.81}$$

where $\tilde{\gamma}$ was defined in equation (2.97). This allows us to write:

$$P_1(x_i, x_{ijk}) = \frac{105^{7/2} 3^3 \exp\left(-\frac{1}{2}\left(Q_1 + Q_3\right)\right)}{(2\pi)^{13/2}(1 - \tilde{\gamma}^2)^{3/2}} \, , \tag{4.82}$$

with the quadratic forms:

$$Q_1 = 3 \sum_i \left(\frac{(u_i - \tilde{\gamma} x_i)^2}{(1 - \tilde{\gamma}^2)} + x_i^2\right) \, , \tag{4.83}$$

$$Q_3 = 105 \left(x_{123}^2 + \sum_{i=1}^3 (v_i^2 + w_i^2)\right) \, ,$$

$$= \frac{35}{2} \, \overline{x}_{ijk} \overline{x}_{ijk} \, . \tag{4.84}$$

4.C.2 Two-point PDFs

Calling $\boldsymbol{x} = (x, x_i, x_{ij}, x_{ijk})$ and $\boldsymbol{y} = (y, y_i, y_{ij}, y_{ijk})$, the Joint PDF reads

$$\mathcal{P}_2(\boldsymbol{x}, \boldsymbol{y}) = \frac{\exp\left[-\frac{1}{2}\begin{pmatrix}\boldsymbol{x}\\\boldsymbol{y}\end{pmatrix}^{\mathrm{T}} \cdot \mathbf{C}^{-1} \cdot \begin{pmatrix}\boldsymbol{x}\\\boldsymbol{y}\end{pmatrix}\right]}{\det|\mathbf{C}|^{1/2} (2\pi)^{15}} \, , \tag{4.85}$$

where $\mathbf{C}$ is the covariance matrix which depends on the separation vectors only because of homogeneity

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{\boldsymbol{xx}} & \mathbf{C}_{\boldsymbol{xy}} \\ \mathbf{C}_{\boldsymbol{xy}}^{\mathrm{T}} & \mathbf{C}_{\boldsymbol{yy}} \end{pmatrix} \, . \tag{4.86}$$

Note that $\boldsymbol{x}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{x}}^{-1} \cdot \boldsymbol{x}$ is given by $Q_0(x) + Q_2(x) + Q_1(x) + Q_3(x)$, where the Q_i are given by equations (4.76) and (4.84). The cross terms will involve correlations of all components of $\boldsymbol{x}$ and $\boldsymbol{y}$
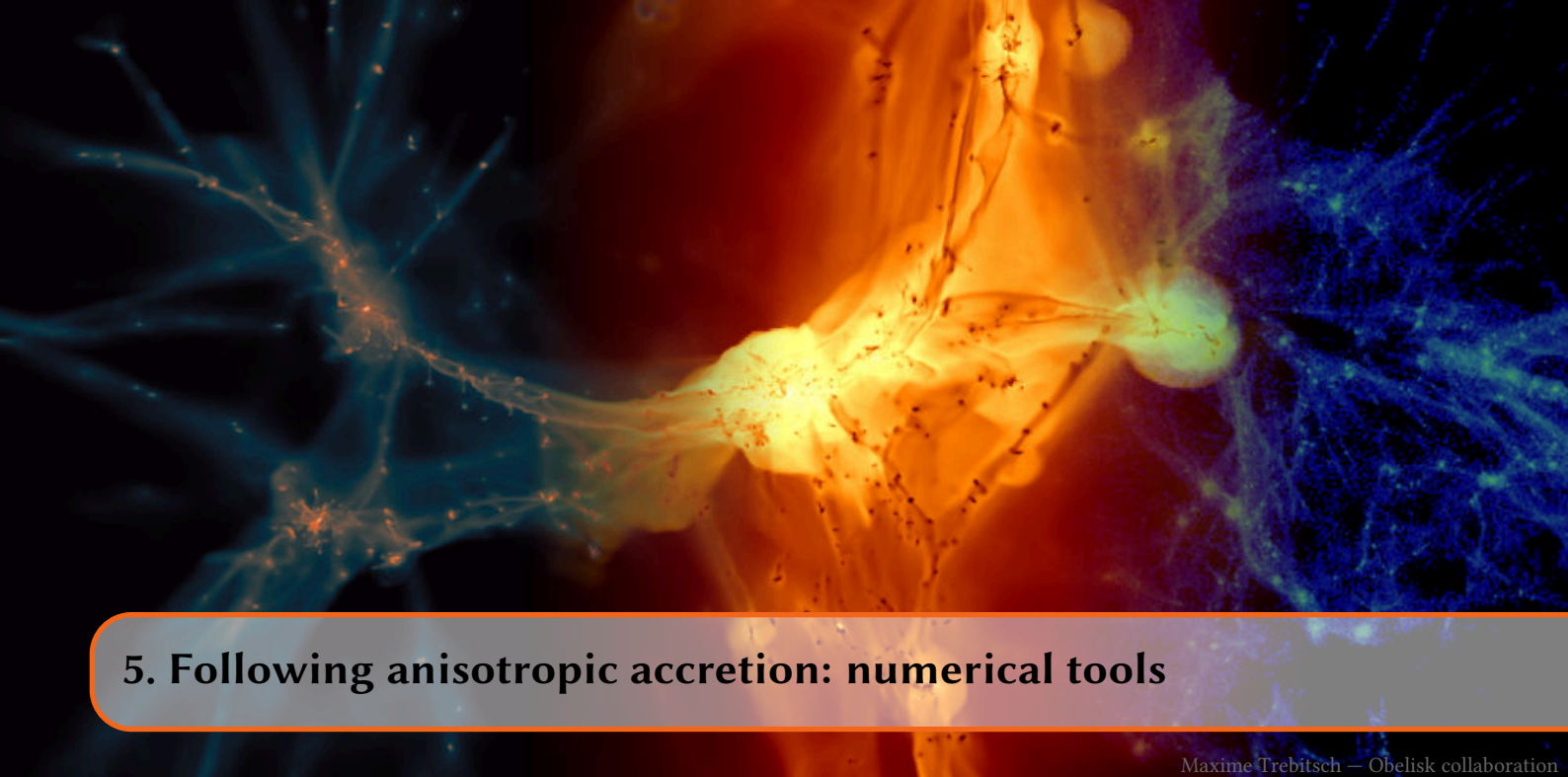
$$\mathbf{C}_{\boldsymbol{xy}} = \langle \boldsymbol{x} \cdot \boldsymbol{y}^{\mathrm{T}} \rangle \, . \tag{4.87}$$

The correlation length of the various components of $\mathbf{C}_{\boldsymbol{xy}}$ differ, as higher derivatives decorrelate faster, see figure 4.C.1. Note that the separations are measured in units of R, whereas the Q_i are independent of R.
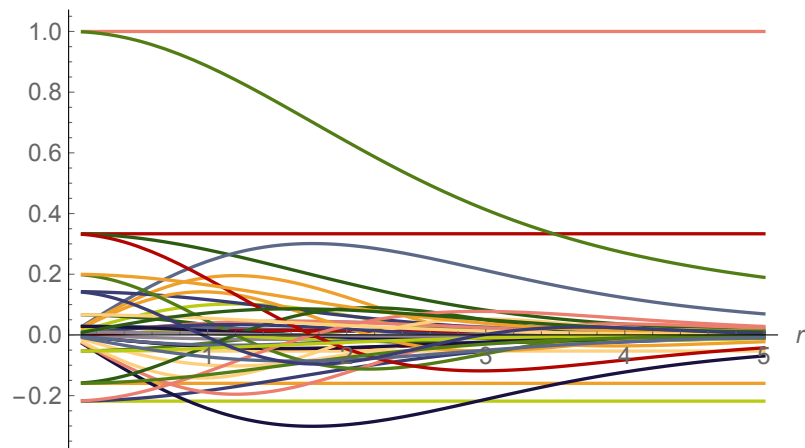
Figure 4.C.1: The correlation functions entering equation (4.87) for a scale invariant powerspectrum of index $n_{\mathrm{s}} = -3/2$.

5. Following anisotropic accretion: numerical tools

Maxime Trebitsch — Obelisk collaboration

Outline

5.1 Introduction

Astrophysical numerical codes usually solve the equation of hydrodynamics in two different approaches, following either the Lagrangian or Eulerian description of the gas. In the former approach known as SPH, the basic elements are macro-particles that represent a given mass of gas (Springel et al., 2001; Wadsley et al., 2004; Price et al., 2017). This approach very conveniently provides the Lagrangian history of the gas, as each particle carries a fixed amount of mass along with their thermodynamical quantities (temperature, density, velocity, metallicity). Another possible approach known as AMR is based on a Eulerian point of view. In these methods, the basic elements are finite volumes of gas (Teyssier, 2002; Bryan et al., 2014), spatially laid on a grid. One of the strength of AMR methods is their ability to control the spatial resolution in regions of interest by adapting the resolution. The most common approach is the so-called "pseudo-Lagrangian" refinement scheme, where regions containing a lot of mass are refined, effectively ending up with each volume containing a similar mass (see section 2.2.3). One can

however trigger refinement on arbitrary criterion, such as the gas vorticity, the local Jeans length or above a certain pressure gradient, so that the geometry of the flow can be captured by adapting the geometry of the grid. As AMR methods are built around Riemann solvers, they are also very good at capturing shocks in numerical simulations. They can therefore provide a very good solution to the problem of cosmic accretion, which requires to accurately resolve shocks around the virial radius. However, due to their Eulerian nature, AMR codes are unable to capture the Lagrangian evolution of the gas and do not provide the past thermodynamical history of a parcel of gas.

This caveat is particularly problematic when studying cosmic accretion, and in particular cold flows as their peculiar evolution is usually captured by their maximum temperature, as the gas that composes them never heated up above a given threshold (see section 6.2.3) This effectively selects the gas that crossed the virial radius without shocking, so that the definition of cold flows requires the knowledge of the past Lagrangian history of the gas. To overcome this issue, AMR codes have been equipped with "tracer" particles. Tracer particles are passively displaced with the gas flow and hence track its Lagrangian evolution. On their trajectory, they usually record instantaneous quantities, in particular the temperature of the gas and its density, but also the torques resulting from the gravitational interaction of the gas with the halo and the disk and from the pressure of the surrounding hot medium.

This chapter presents a technical description of the tools I developed in order to make Ramses suited to the study of cold flows, although the methods presented hereafter can also be applied to a variety of other problems. Section 5.2 presents a new tracer particle scheme developed for the code Ramses. I compare the implementation to previous ones and show that it largely improves the results. In particular, the new tracer particles are able to accurately reproduce the Eulerian distribution of the gas (so that they are trustable) and also provide the entire Lagrangian evolution of the gas (so that they provide a comprehensive history of the gas evolution). Section 5.3 presents the methods developed to compute the different torques acting on a parcel of fluid in post-processing. As will be shown in the next chapter, these methods can then be used to provide a detailed description of the evolution of the angular momentum of the cold flows.

5.2 "Accurate tracer particles of baryon dynamics in the adaptive mesh refinement code Ramses"

One of the requirements of tracer particles is that they should accurately reproduce the Eulerian distribution of the gas. In the naive approach, tracer particles are advected by the gas by interpolating the velocity. This is usually done with a cloud-in-cell interpolation (first order interpolation), where the value of the velocity is interpolated from the 8 closest cells. Such a velocity-based approach was implemented in Ramses (Dubois et al., 2012) and used to probe the link between cosmic gas infall and galactic gas feeding. This approach yields smooth Lagrangian trajectories, yet it falls short of reproducing the gas density distribution accurately in regions of converging flows, as I showed in Cadiou et al., 2019 (paper below). In addition, there is no natural way of taking into account transfers of baryons between the gas, stars and SMBHs which are particularly relevant in the context of galaxy formation in the inner regions surrounding the galaxy.

Using a different approach, Genel et al., 2013 suggested to instead sample mass fluxes *via* a Monte-Carlo method. In this approach, the mass flux between cells, which is readily computed by the Riemann solver of the code, is approximated by moving particles across cells interface: each particle jumps from cell i to cell j with probability

$$p_{ij} = \frac{\Delta M_{ij}}{M_i},$$

(5.1)

where ΔM_{ij} is the transferred mass (as computed by the Riemann solver) and M_i is the mass of the cell originally containing the particle. Since tracer particles reproduce mass fluxes, their evolution is fully consistent with that of the gas, up to a sampling noise.

Instead of providing smooth trajectories, Monte Carlo tracer particles provide a statistical sample whose mean accurately tracks the properties of baryons in the simulation and whose spatial distribution matches the Eulerian gas density. They are therefore perfectly suited to the problem of cold filamentary accretion. In the paper provided hereafter, I present the details of the implementation for gas-to-gas transfers. I then present how one can extend equation (5.1) to take into account any baryon transfers, providing a clear improvement over previous tracer particle implementations. I then show that my implementation is able to accurately reproduce the Eulerian distribution of the gas, while providing at the same time the full Lagrangian evolution of baryons in their journey in the gas, stars and SMBHs. As a proof of concept, the method is then applied to the problem of cold flows to recover the bimodal accretion mode observed in SPH simulations (e.g. Kereš et al., 2005).

The paper, published in A&A in Cadiou et al., 2019, is provided hereafter.

Astronomy
&
Astrophysics

Accurate tracer particles of baryon dynamics in the adaptive mesh refinement code Ramses

Corentin Cadiou[1], Yohan Dubois[1], and Christophe Pichon[1,2]

[1] Institut d'Astrophysique de Paris, CNRS & UPMC, UMR 7095, 98 bis Boulevard Arago, 75014 Paris, France
e-mail: corentin.cadiou@iap.fr
[2] Korea Institute of Advanced Studies (KIAS), 85 Hoegiro, Dongdaemun-gu 02455, Seoul, Republic of Korea

ABSTRACT

We present a new implementation of the tracer particles algorithm based on a Monte Carlo approach for the Eulerian adaptive mesh refinement code Ramses. The purpose of tracer particles is to keep track of where fluid elements originate in Eulerian mesh codes, so as to follow their Lagrangian trajectories and re-processing history. We provide a comparison to the more commonly used velocity-based tracer particles, and show that the Monte Carlo approach reproduces the gas distribution much more accurately. We present a detailed statistical analysis of the properties of the distribution of tracer particles in the gas and report that it follows a Poisson law. We extend these Monte Carlo gas tracer particles to tracer particles for the stars and black holes, so that they can exchange mass back and forth between themselves. With such a scheme, we can follow the full cycle of baryons, that is, from gas-forming stars to the release of mass back to the surrounding gas multiple times, or accretion of gas onto black holes. The overall impact on computation time is ∼3% per tracer per initial cell. As a proof of concept, we study an astrophysical science case – the dual accretion modes of galaxies at high redshifts –, which highlights how the scheme yields information hitherto unavailable. These tracer particles will allow us to study complex astrophysical systems where both efficiency of shock-capturing Godunov schemes and a Lagrangian follow-up of the fluid are required simultaneously.

Key words. hydrodynamics – methods: numerical – cosmology: theory – Galaxy: formation

1. Introduction

Many astrophysical problems of interest require us to solve equations of hydrodynamics on very different timescales and physical scales. Two main methods have been developed to solve these equations. On the one hand, one can study the motion of the gas by following the evolution of interacting particles. This Lagrangian approach is the one used by smooth particle hydrodynamics (SPH, e.g. Springel 2005; Wadsley et al. 2004; Price et al. 2018) codes. These codes sample the gas distribution using a set of fixed-mass macro-particles smoothed with a given kernel, and move particles accordingly. By construction, this approach provides the Lagrangian evolution of the gas. This property is also one of its shortcomings: low-density regions are populated by large particles and hence lack resolution. On the other hand, gas hydrodynamics can also be described on a grid, where gas distribution is sampled on finite volumes, and solved with efficient shock-capturing Godunov solvers. Adaptive mesh refinement (AMR, e.g. Kravtsov et al. 1997; Teyssier 2002; Springel 2010; Bryan et al. 2014) codes follow this approach and allow for a dynamical refinement of the mesh. Though quasi-Lagrangian refinement is most commonly adopted in situations addressing galaxy formation problems, super-Lagrangian resolutions can also be achieved by refining the grid based on gas quantities such as the Jeans length to follow gravitationaly unstable star-forming regions (Agertz et al. 2009), the vorticity to follow the seeding of turbulence (e.g. Iapichino & Niemeyer 2008), the relative variation of any hydro quantity (such as e.g. the ionised fraction of hydrogen; Rosdahl & Blaizot 2012), or using

a passive scalar to keep track of a particular gas phase (such as for jets, see, e.g. Bourne & Sijacki 2017), among others. While super-Lagrangian refinement provides a very flexible method to trigger refinement, it falls short of providing the Lagrangian history of the gas.

To overcome this issue, AMR codes have been equipped with "tracer" particles. Tracer particles are passively displaced with the gas flow, and hence track its Lagrangian evolution. Each tracer can also be used to record instantaneous quantities, such as the thermodynamical properties of the gas or any other property. Many astrophysical problems can can benefit greatly from this Lagrangian information. For example, when studying galaxy formation, the past Lagrangian history of the gas is crucial to understand how gas has been accreted and how it has been ejected in large-scale galactic outflows. Tracer particles can be used to study the density and temperature evolution of the gas (e.g. Nelson et al. 2013; Tillson et al. 2015) that will eventually form stars. For example, one could use tracer particles to study the temperature evolution of the gas as it falls onto galaxies, to study the number of dynamical times before it becomes star forming or to quantity the number of time gas is recycled in stars or sent in galactic fountains. Another problem that requires the use of tracer particles is the study of mixing. Particularly in turbulent environments, such as the interstellar or the intergalactic medium, the Lagrangian information provides information about, for example, mixing timescale (e.g. Federrath et al. 2008), the origin of turbulence (e.g. Vazza et al. 2011, 2012), or how it contributes to core buildup Mitchell et al. (2009). In addition to this, the past Lagrangian evolution of a parcel of fluid

can also impact the modelling itself (e.g. Federrath et al. 2008; Silvia et al. 2010).

In this paper we present a new implementation of tracer particles in the AMR RAMSES code (Teyssier 2002). This implementation is based on the one developed by Genel et al. (2013) for the moving mesh AREPO code (Springel 2010). It has been extended to track the full Lagrangian history of baryons in any phase, including their conversion from gas to stars, from stars back into the gas via supernova feedback, their interaction with feedback from black holes, and their accretion onto them. This Monte Carlo (MC) tracer particle implementation improves the previous implementation, velocity-advected tracers. With the velocity-based approach, tracer particles are moved based on the interpolated local values of the gas velocity field. While this yields qualitative results, it suffers from systematic effects: tracer particles over-condensate in regions of converging flows (Genel et al. 2013). Monte Carlo tracer particles follow a different idea. They are moved so that the tracer particle mass flux at each cell interface is statistically equal to that of the gas. Thanks to this property, the Eulerian distribution of tracers converge to that of the gas when the number of tracer particles goes to infinity. In addition to matching the gas distribution, the implementation of tracer particles here is also able to match the distribution of baryons in stars and in black holes.

The paper is structured as follows. Section 2 details the implemented algorithm. Section 3 presents tests and validations of the new implementation. In particular, Sect. 3.1 presents the results from idealised tests and Sect. 3.2 presents an analysis of the properties of tracers in a real astrophysical simulation. Using the same simulation, Sect. 3.3 illustrates the efficiency of the scheme applied to a specific science case – the bimodal accretion of gas onto galaxies at high redshift. Section 4 assesses the performance of the scheme. Section 5 provides a discussion of our results and our conclusions. Appendix A provides more details about the algorithm.

2. Implementation

The RAMSES code (Teyssier 2002) solves the full set of Euler equations by formulating the equations in terms of finite-volume, that is, by calculating fluxes at the interfaces of cells of the adaptive mesh. This is done by using a MUSCL-Hancock method with a second-order Godunov solver calculating the fluxes from linearly interpolated values at cell faces from the cell-centred values limited by a total-variation-diminishing scheme. Such a Eulerian-based method has proven efficient at capturing shock discontinuities and achieves efficient mixing of shear layers of gas; however, its main drawback is that it does not naturally provide the Lagrangian trajectories of gas elements.

To address this problem, it is possible to introduce the so-called tracer particles of the flow that should follow the flow lines of the gas. A naive approach to track the motion of the gas is to use the velocity of the gas itself, assign it to tracer particles, and move them accordingly. This is done with a cloud-in-cell interpolation of the velocity values of the overlapped cells where the volume of the cloud is that of the host cell, though the level of the interpolation is not particularly important (nearest grid point or triangular shape cloud; Federrath et al. 2008). Such a velocity-based approach was implemented in RAMSES (Dubois et al. 2012a) and used to probe the link between cosmic gas infall and galactic gas feeding, and its acquisition of angular momentum (Pichon et al. 2011; Dubois et al. 2012a; Tillson et al. 2015). While this approach yields smooth trajectories, it falls short of reproducing the gas density distribution

accurately in regions with strong convergence of the velocity field (Genel et al. 2013).

To address this shortcoming, we have implemented in RAMSES the MC approach of tracer particles introduced by Genel et al. (2013) for AREPO (Springel 2010). Instead of having proper velocities and positions, MC tracers are attached to individual cells and are allowed to "jump" from the centre of one cell to the centre of another according to the mass fluxes obtained through the Godunov solver.

We have generalised the MC method to track exchanges of baryons between gas, star particles, and supermassive black hole (SMBH) particles, and in the following we refer to them as "buckets". At each time step, tracers are allowed to jump from any bucket i to any bucket j with a probability (gas→gas, gas↔star, gas→black hole) of

$$p_{ij} = \begin{cases} \dfrac{\Delta M_{ij}}{M_i}, & \text{if } \Delta M_{ij} \geq 0, \\ 0, & \text{if } \Delta M_{ij} < 0, \end{cases} \qquad (1)$$

where ΔM_{ij} is the mass flowing from bucket i to bucket j between t and $t + \Delta t$ and M_i is the mass of the depleted bucket i at time t. This probability is also the fraction of baryons flowing from one bucket to another. If the initial Eulerian distributions of tracers and baryons are equal, then in the limit where the number of tracers becomes large, satisfying Eq. (1) is sufficient for the Eulerian distributions to remain equal at all times. Here is an outline of the proof. For any bucket i containing N_t tracers of equal mass m_t, let the total tracer mass read $M_t \equiv N_t m_t$. Because tracers are moved stochastically, the tracer mass flux $\Delta M_{t,ij}$ is a random variable. If at time t, $M_t = M_i$ (i.e. the Eulerian distributions are the same), then the expected tracer flux is $\mathrm{E}\left[\Delta M_{t,ij}\right] = N_t \times p_{ij} m_t = M_i p_{ij} = \Delta M_{ij}$. When the number of tracers becomes large, the tracer mass flux converges to the baryon flux, $\Delta M_{t,ij} \rightarrow \Delta M_{ij}$. The buckets have the same initial mass and are updated with the same mass fluxes, so they remain equal at the next time step, $t + \Delta t$. Therefore, if the initial Eulerian distributions are equal, by induction they remain equal at all times (in the limit of a large number of tracers)[1].

All the processes that are able to move tracers from bucket to bucket are summarised in Fig. 1. Tracers can move from one gas cell to another through gas dynamics, and the jet mode of AGN feedback from SMBHs, from gas to stars via star formation, from stars to gas via supernova (SN) feedback, and from gas to SMBHs via black hole accretion. Below, we present the different algorithms used for each of these processes.

2.1. Gas dynamics

The algorithm moving tracer particles from one gas cell to another is the following. For each level of refinement, all the unrefined leaf cells are iterated over. For each leaf cell i containing tracer particles, the total outgoing mass is computed as $\Delta M \equiv \sum_{j=1}^{2N_d} \max(\Delta M_{ij}, 0)$, where j runs over the index of the neighbouring cells, N_d is the number of dimensions, and ΔM_{ij} is the mass transferred between cell i and cell j in one time step and obtained from the Godunov flux of mass $F_{m,ij}$, that is, $\Delta M_{ij} = F_{m,ij} \Delta t$. We take

$$p_{\text{gas}} = \frac{\Delta M}{M_i}, \qquad (2)$$

---

[1] In general, any stochastic scheme for which the expected tracer flux equals that of the baryons is able to track the Eulerian distribution at all times.
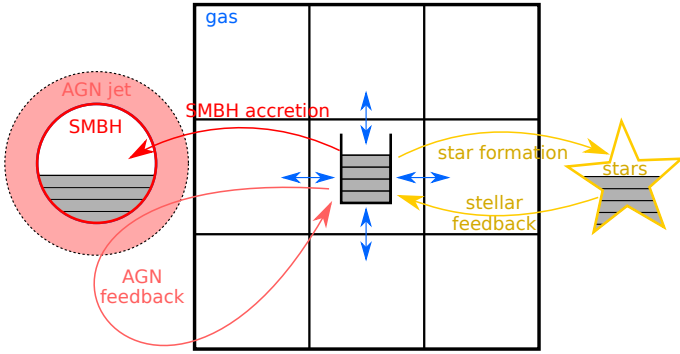
Fig. 1. Scheme of the different "buckets" that can hold tracer particles and the process that moves them around. The three buckets are gas cells, stars, and SMBHs. Arrows indicate outgoing mass fluxes between buckets and the physical process associated, and grey squares represent tracer particles. The jet mode feedback from AGNs (around SMBHs) is able to move gas tracer particles from the central cell to the surrounding cells. The particles have no spatial distribution within the buckets or any phase-space distribution. Tracer particles are exchanged probabilistically between buckets based on the mass fluxes. For example, for the gas, they are exchanged based on the mass fluxes at the boundary of the cells.

to be the probability of displacing a gas tracer particle from one cell to any other of its neighbouring cell, and

$$p_j = \max\left(\frac{\Delta M_{ij}}{\Delta M}, 0\right), \tag{3}$$

to be the probability of moving this tracer particle into cell j. For each tracer particle in cell i, a random number r is drawn from a uniform distribution between 0 and 1. If $r < p_{\mathrm{gas}}$, the tracer is selected. For each selected tracer, another random number r' is drawn. For each neighbouring cell j with a positive flux (such that $\Delta M_{ij} > 0$), if $r' < p_j$ the tracer particle is moved into cell j and the algorithm proceeds to the next particle; else, r' is decreased so that $r' \leftarrow r' - p_j$ and the algorithm proceeds to the next neighbouring cell. Because the sum of all the p_j is 1, this procedure will assign the tracer to a single cell.

When a cell of mass M_0 is refined between two time steps, all its tracers are distributed randomly to one of the eight newly created cells, the probability for a tracer particle to be attached to the new cell i being $p = M_i/M_0$ (refined density can be interpolated from neighbouring values or equally distributed amongst new cells). Conversely when a cell is derefined all its tracers are attached to the parent cell.

2.2. Star formation

This part of the algorithm involves moving tracers from the gas phase into star particles, and moving the star-tracer particles along with their star particles.

We first recall that the star formation process in Ramses is usually modelled by a Schmidt law, where the star formation rate density is non-zero and

$$\frac{\mathrm{d}\rho_\star}{\mathrm{d}t} = \epsilon_\star \frac{\rho_g}{t_{\mathrm{ff}}}, \tag{4}$$

when $\rho_g > \rho_0$, and where ρ_g is the gas density, ρ_0 a gas density threshold, $t_{\mathrm{ff}} = (3\pi/(32G\rho_g))^{1/2}$ the gas free-fall time, and $\epsilon_\star$ the efficiency of star formation, which can be taken as an ad hoc constant, or as a function of the local gravo-turbulent properties

of the gas (Krumholz & McKee 2005; Hennebelle & Chabrier 2011; Padoan & Nordlund 2011). A single star particle made of $N_\star$ stars of mass resolution $M_{\star,0}$ is created, where $N_\star$ is drawn according to random Poisson process (Rasera & Teyssier 2006):

$$P_{\mathrm{sf}} = \frac{\lambda^{N_\star}}{N_\star!} \exp(-\lambda), \tag{5}$$

where P_{sf} is the probability of creating $N_\star$ particles from the gas (and accordingly removing $M_\star \equiv N_\star M_{\star,0}$ mass from the gas cell), and

$$\lambda = \frac{\rho_g \Delta x^3}{M_{\star,0}} \frac{\Delta t}{\epsilon_\star^{-1} t_{\mathrm{ff}}}. \tag{6}$$

Finally, the transfer of gas tracer particles to star-tracer particles at time of creation t of $M_\star$ is given by the probability

$$p_\star = \frac{M_\star}{M_i}. \tag{7}$$

In more details, we loop over all the gas tracer particles contained in the cell where the new star is created. For each of them, a random number r is drawn from a uniform distribution between 0 and 1. If $r < p_\star$, the gas tracer particle is turned into a star-tracer particle at the same position as that of the star particle (i.e. at the centre of the cell). The star-tracer particle is "attached" to the star particle by moving along with the star particle, which is done through a classic leap-frog integration of its motion. Therefore, at all time steps, the position of the tracer is updated to match the position of its star. The index of the star is also recorded on the tracer for convenience.

The implementation also comes with two alternatives to initialise the tracer particles. One can feed in a list of positions to the code; one tracer will be created at each location. Alternatively, we developed an initialisation scheme that takes as input the mass that each tracer particle represents, m_t. The scheme is called "in-place initialisation" as it is performed directly within the code: the scheme is called once at startup, after the AMR grid has been built. It loops over all cells, and for each of them computes the number of tracer particles to create. The expected number of tracers created in a cell of mass M_{cell} is $N = m_t/M_{\mathrm{cell}}$. Let us write $N_0 = \lfloor N \rfloor$. The scheme creates $N_0 \equiv \lfloor N \rfloor$ particles in the cell and then creates an additional one with probability $N - N_0$. In the end, the expected number of tracer particles created in the cell is N, meaning that on average each cell is populated with the correct number of tracer particles. In the following, unless stated otherwise, the tracer particle distribution is always initialised using the in-place method.

2.3. Supernova feedback

Let us describe the transfer of mass of a star particle to the gas according to type II SN explosions (referred to henceforth as SNII) and their associated tracer particles. This can be trivially extended to the more complete description of the evolution of stellar mass loss.

When a star particle sampling an initial mass function (IMF) of mass $M_\star$ explodes into type II SNe, it releases a mass $\eta_{\mathrm{SN}} M_\star$, where η_{SN} can be crudely approximated by the mass fraction of the IMF going SNII. The probability of releasing a star-tracer particle into the gas is $p_{\mathrm{SN}} = \eta_{\mathrm{SN}}$. For each star particle turning into SNe, we loop over all the star-tracer particles attached to it. For each of these, a random number r is drawn from a uniform
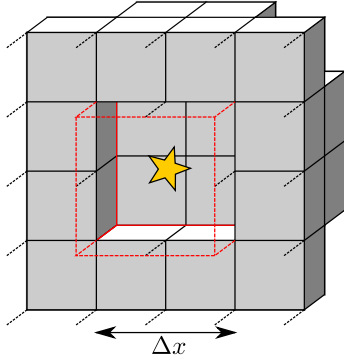
Fig. 2. Scheme of the 48 neighbouring virtual cells (only the 24 rear ones are shown) where mass and momentum are deposed during a SN event. The cell containing the SN has a size of Δx and is outlined in red.

distribution between 0 and 1. If $r < p_{\mathrm{SN}}$, the star-tracer particle is released in the gas and turned into a gas tracer particle. Otherwise, the tracer is left attached to the stellar remnant.

The transfer of star-tracer particles to the gas by SNII is described here for the so-called mechanical feedback model of (Kimm & Cen 2014; see also Kimm et al. 2015)[2]. In this model, the gas is released into the neighbouring cells. The mechanical feedback scheme is designed to overcome the consequences of radiative losses in SN bubbles due to the lack of resolution. Where the cooling time of the SN-heated gas is shorter than the hydrodynamical time step, the energy-conserving phase (with Sedov-Taylor solution), during which the momentum is growing by the pressure work of the bubble, cannot be captured properly, and leads to spurious energy and momentum loss. To circumvent this problem, Kimm & Cen (2014) introduced a model that correctly accounts for the momentum injection according to the Sedov-Taylor or snow-plough solution (Thornton et al. 1998), which depends on the cooling rate of the gas, or more precisely on the energy release, local gas density, and metallicity. The cell containing the exploding star particle is virtually represented by 8 cells refined by an additional level, which are equivalently surrounded by 48 such virtual neighbouring cells, as illustrated in Fig. 2 (Kimm & Cen 2014). The mass ejecta together with the mass of the swept-up gas of the central true cell is released evenly in all the virtual cells, and is attributed back accordingly to the true existing cells.

The tracer particles are interfaced with this feedback model as follows: For each released star to gas tracer particle, a random integer number $l \in [1, 48]$ is drawn uniformly. The star tracer is then moved to the centre of the lth virtual cell and attributed to the corresponding true existing cell as a new gas tracer particle.

2.4. SMBH formation and gas accretion

Supermassive black holes are modelled as sink particles that can form out of the dense star-forming gas, grow by accretion of gas, and coalesce following the implementation described in Dubois et al. (2012b).

---

[2] We have extended this implementation to i) simple thermal pulses of energy (with or without delayed cooling; Teyssier et al. 2013), where the mass is released to the central cell only, and ii) to the so-called kinetic model of (Dubois & Teyssier 2008; in its more recent form described in Rosdahl et al. 2017) where "debris" particles are replaced by a bubble injection region of energy, momentum, and mass according to the Sedov-Taylor solution.

A SMBH forms according to several user-defined criteria, typically above a given gas density threshold ρ_0 and outside an exclusion distance radius r_{ex} within which SMBH is artificially prevented if any other SMBH already exists (in order to prevent creation of multiple SMBHs within the same galaxy). When a SMBH forms with an initial seed mass $M_{\mathrm{SMBH},0}$, gas tracer particles in the cell of mass M_i containing the SMBH are attached to the SMBH and become SMBH tracer particles according to a probability

$$p_{\mathrm{SMBH}} = \frac{M_{\mathrm{SMBH},0}}{M_i}. \tag{8}$$

SMBHs can continuously accrete gas according to the Bondi–Hoyle–Littleton accretion rate, capped at Eddington with

$$\dot{M}_{\mathrm{SMBH}} = (1 - \varepsilon_{\mathrm{r}}) \, \dot{M}_{\mathrm{acc}} = (1 - \varepsilon_{\mathrm{r}}) \, \min(\dot{M}_{\mathrm{B}}, \dot{M}_{\mathrm{Edd}}), \tag{9}$$

$$\dot{M}_{\mathrm{B}} = \frac{4\pi\rho G^2 M_{\mathrm{SMBH}}^2}{(c_{\mathrm{s}}^2 + u^2)^{3/2}} \left(\frac{\rho}{\rho_{\mathrm{boost}}} \right)^\alpha, \tag{10}$$

$$\dot{M}_{\mathrm{Edd}} = \frac{4\pi G m_{\mathrm{p}} M_{\mathrm{SMBH}}}{\sigma_T \varepsilon_{\mathrm{r}} c}, \tag{11}$$

where $\dot{M}_{\mathrm{acc}}$, $\dot{M}_{\mathrm{SMBH}}$, $\dot{M}_{\mathrm{B}}$, and $\dot{M}_{\mathrm{Edd}}$ are the disc, SMBH, Bondi–Hoyle–Littleton, and Eddington accretion rates, respectively, m_{p} is the mass of a proton, G the gravitational constant, σ_T the Thomson cross-section, ε_{r} the radiative efficiency, c_{s} the speed sound, u the mean velocity of the gas with respect to the SMBH, and c the speed of light. ρ_{boost} and α are free parameters set, here, to $\rho_{\mathrm{boost}} = 8 m_{\mathrm{p}} \, \mathrm{cm}^{-3}$ and $\alpha = 2$ introduced to boost the accretion rate due to unresolved small-scale larger densities (Booth & Schaye 2009). The value of ε_{r} is either chosen as a constant value equal to 0.1, or, here, as a varying function of the spin of SMBH, whose evolution is governed by gas accretion and BH coalescence (see Dubois et al. 2014a,b, and Dubois et al., in prep., for details).

The mass taken from the gas cell in one time step is $\Delta M_{\mathrm{acc}} \equiv \Delta t \min(\dot{M}_{\mathrm{BH}}, \dot{M}_{\mathrm{Edd}})$. We note that $\Delta M_{\mathrm{acc}} > \dot{M}_{\mathrm{SMBH}} \Delta t$ as part of the accreted mass is radiated away due to relativistic effect (and lost to the simulation). Each gas tracer in the cell containing the SMBH at a given time is accreted into the black hole with a probability of

$$p_{\mathrm{acc}} = \frac{\Delta M_{\mathrm{acc}}}{M_i}. \tag{12}$$

Tracer particles also model SMBH merger events. All the tracer particles attached to the two parent SMBHs are moved to the newly formed SMBH. There is no mechanism to extract tracers from the SMBH (reflecting the fact that there is no way to extract matter from a BH). One should also note that the total mass of SMBH tracers is larger than the total mass of SMBHs, as part of the energy-mass has been radiated away during accretion (and tracers have a fixed mass).

2.5. AGN feedback

In Dubois et al. (2012b), there are two modes of AGN feedback: a quasar/heating mode and a radio/jet mode. The mode is selected based on the ratio of the Bondi–Hoyle–Littleton accretion rate to the Eddington accretion rate $\chi = \dot{M}_{\mathrm{B}}/\dot{M}_{\mathrm{Edd}}$. If $\chi < 0.01$, the AGN is in jet mode, and, otherwise, it is in quasar mode (Merloni & Heinz 2008).

In quasar mode, all the energy of the AGN proportional to $E_{\mathrm{AGN},Q} = \varepsilon_{\mathrm{f},Q} \varepsilon_{\mathrm{r}} \dot{M}_{\mathrm{acc}} c^2 \Delta t$ (the value $\varepsilon_{\mathrm{f},Q} = 0.15$ is calibrated to
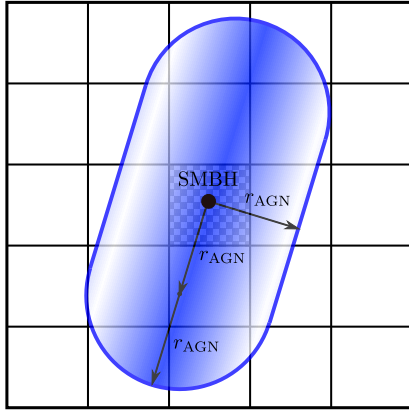
Fig. 3. Schematic representation of the jet model. Gas is transported from the central cell (hatched region) containing the SMBH (black dot) into the jet (blue shaded region). The radial profile of the jet is a Gaussian of scale r_{AGN}. The shape of the jet is a "capsule" (a cylinder capped with two half spheres).

match the BH-to-galaxy mass relation; Dubois et al. 2012b) is released as thermal energy in all cells within a sphere of size R_{AGN} and the mass of the gas is left untouched. This feedback mode has only an indirect effect on the gas mass distribution (and hence on tracer particles), turning some fraction of the released thermal energy into kinetic energy and launching a quasar-like wind.

In radio mode, a jet is launched from the AGN. The jet moves mass from the central cell only and spreads it into the jet and injects linear momentum, and energy. The released energy (and hence, momentum within the jet), as for the quasar mode, is proportional to the rest-mass accreted energy with an efficiency of $\varepsilon_{f,R}$, which is either taken as a constant value of 1 (to match the SMBH-to-galaxy mass relation; Dubois et al. 2012b) or a varying function of the spin of the SMBH following the results of magnetically arrested discs (MADs) from McKinney et al. 2012; see Dubois et al., in prep. for details). The jet is modelled by a "capsule" (a cylinder with spherical caps) of size r_{AGN}, as illustrated in Fig. 3. The radius of the jet r_{AGN} is usually set to a few times the cell resolution. The mass sent through the jet is proportional to the accreted mass onto the SMBH

$$\dot{M}_{jet} = f_{Load}\dot{M}_{SMBH}, \tag{13}$$

where f_{Load} is a mass-loading factor, usually 100. The mass transported by the jet is distributed to all the cells intersecting with the capsule. Each cell i receives a relative fraction ψ_i of the mass (and of the injected linear momentum)

$$\psi_i = \frac{\rho_i \int_{\mathcal{I}} e^{-r^2/2r_{AGN}^2} d^3V}{\sum_j \rho_j \int_{\mathcal{J}} e^{-r^2/2r_{AGN}^2} d^3V}, \tag{14}$$

where $\mathcal{I}$ (resp. $\mathcal{J}$) is the volume of the intersection between the AGN capsule and the cell i (resp. j) and ρ_i is the cell mean density. The radius r in Eq. (14) is the polar radius in the cylindrical frame centred on the AGN and aligned with its direction (it is the distance to the jet centre). This integral is computed approximately, using a numerical integration scheme.

The tracer particles are interfaced with the jet model as follows. Each gas tracer particle in the cell i containing the SMBH is moved into the jet volume with a probability of

$$p_{jet} = \frac{\dot{M}_{jet}\Delta t}{M_i}. \tag{15}$$

For each of these particles a random number r is drawn from a uniform distribution between 0 and 1. If $r < p_{jet}$, the tracer is selected and moved into the jet. The new position of the tracer (x, y, z) is drawn randomly, z being the coordinate in the direction of the jet; x and y are drawn from a normal distribution of variance r_{AGN} and z is drawn uniformly between $-2r_{AGN}$ and $2r_{AGN}$. The algorithm uses a draw-and-reject method until one position inside the capsule is found. We note that the gas tracer distribution (as given by Eq. (15)) is consistent with the distribution of the gas sent through the jet (as given by Eq. (14))[3].

More details about the algorithm are given in Appendix A.

3. Validations and tests

Let us now present various validation tests of the algorithm. Section 3.1 presents the results of idealised tests for gas-only tracer particles. Section 3.2 presents the results obtained from a cosmological zoom-in simulation of a galaxy with its SMBH at $z = 2$ and provides the details of the observed distribution of tracer particles. Unless stated otherwise, the gas hydrodynamics is solved with an adiabatic index of $\gamma = 5/3$ and the HLLC approximate Riemann solver (Toro 2009), applying the MinMod slope limiter on the linearly reconstructed states.

3.1. Idealised tests

In this section, we introduce different idealised tests to confirm that the evolution of the gas is correctly tracked by gas tracers. Section 3.1.1 presents a simple two-dimensional (2D) advection of an overdensity to quantify diffusion effects. Sections 3.1.2 and 3.1.3 present a Sedov–Taylor explosion and a Kelvin–Helmoltz instability and confirm that the gas tracers are able to accurately follow the motion of the gas for a strong shock case and a mixing layer of gas, respectively. Section 3.1.4 presents an idealised halo with an AGN at its centre to confirm that the gas tracers correctly track the evolution of the gas in AGN jets.

3.1.1. Uniform advection

In order to quantify the level of diffusion of MC tracers, we run a simulation similar to that run for Fig. 6 of Genel et al. (2013). The simulation is a region of 1 cm in size with a constant density of $1\,\mathrm{g\,cm^{-3}}$ and a velocity of $0.01\,\mathrm{cm\,s^{-1}}$. An overdensity of $14\,\mathrm{g\,cm^{-3}}$ is set at $0 < x < 0.05\,\mathrm{cm}$. The sound speed is $c_s = 1.3\,\mathrm{cm\,s^{-1}}$ in the under-dense region and $0.35\,\mathrm{cm\,s^{-1}}$ in the over-dense region. The simulation is performed on a uniform 2D 128^2 grid including 250 000 tracer particles, initially distributed in the same way as the gas. Due to the intrinsic numerical diffusion (advection error) of the hydrodynamical solver, the spatial extent of the overdensity increases as a function of time as it is advected away. This is illustrated in the central panel of Fig. 4. We note that the density profiles have each been shifted along their x coordinate for visualisation purposes and do not reflect their real absolute position (in fact the rightmost peak travelled 5 cm in 100 s). The top panel of Fig. 4 shows that, when rescaled by the expected noise level $\sigma \equiv 1/\sqrt{M_{cell}/m_t} = 1/\sqrt{N}$ (N is the expected number of tracer particles in the cell), the relative error between the gas tracers and the gas distributions shows

---

[3] In practice, the numerical evaluation of the integrals of Eq. (14) may lead to small yet undetected discrepancies between the gas tracer and the gas distributions.

Fig. 5. *Bottom panel*: radial profile at different times of a Sedov explosion (from blue to yellow) for the gas (solid lines) and the gas tracer (dots). The error bars are 2σ errors. *Top panel*: relative difference between the gas profile and the gas tracer profile. Data have been shifted by -0.25, -0.125, 0, 0.125 and 0.25 radius units respectively (from blue to yellow) so that one may easily distinguish the different data points. Details of the simulation are discussed in the text. The gas tracer particles are accurately advected with the gas.

Fig. 4. *Top panel (bottom):*: gas density profile (solid line) and gas density profile (plus symbols) at different times (reported in the legend). The profiles have been recentred and shifted horizontally by -0.12 cm, 0, 0.12 cm, and 0.24 cm for $t = 0, 1, 9$, and 100 s, respectively. *Top panel (top):*: relative difference between the gas and gas tracer density profiles in units of the expected noise level $\sigma = 1/\sqrt{M_{\text{cell}}/m_{\text{t}}}$. *Bottom panel*: evolution of the spatial extent of an advected overdensity as a function of time for the gas (dashed) and the gas tracer particles (dot symbols) for a high-resolution run (blue) and a low-resolution run (orange, see text for details). The difference shows no spatial dependence. The gas tracers diffuse exactly as the gas.

no spatial modulation. Their distributions are the same with an extra factor that is entirely due to sampling noise, which in turn depends only on the local cell mass and the (constant) tracer mass.

In more quantitative terms, let us compare the time evolution of the spatial extent of the gas tracer overdensity to that of the gas. We rerun the simulation on a 32^2 grid (low resolution) in addition to the previous run (high resolution). We compute the spatial extent by fitting a Gaussian function $\rho(x) = \rho_0 + H \exp(-(x - x_0)^2/(2\sigma_\rho^2))$ to the gas and gas tracer profiles, with free parameters ρ_0 the base density, H the amplitude of the overdensity, x_0 the position of the overdensity, and σ_ρ its spatial extent. The results are shown in the bottom panel of Fig. 4. As expected due to the numerical diffusion, the spatial extent of the overdensity increases as a function of time and the diffusion becomes larger when the resolution is decreased. In both cases, the Eulerian distribution of tracer particles is diffused exactly as much as the gas[4].

---

[4] This result complements that of Genel et al. (2013). Indeed we study here the diffusion of the Eulerian distribution of the tracer particles, while the original paper presents the Lagrangian diffusion of the tracer particles.

3.1.2. Sedov-Taylor explosion

We ran a classical Sedov-Taylor explosion in three dimensions and compare the gas density radial profile to the density profile of gas tracer particle. The simulation was performed on a coarse grid of 128^3, refined on the relative variation of the density and of the pressure: a new level is triggered when the local relative variation of one of these quantities is larger than 1% with up to two levels of refinement. The simulation was initialised with a uniform density and pressure of $1\,\text{g cm}^{-3}$ and $10^{-5}\,\text{dyne cm}^{-2}$, respectively, and an over-pressure in the central cell of the box of $6.7 \times 10^6\,\text{dyne cm}^{-2}$. $2\,900\,000$ tracers, statistically uniformly distributed initially in the box, hence, with around ~ 1.4 tracer per initial cell.

The evolution of the spherically averaged radial density profile of the gas and of the tracers is shown in Fig. 5. The tracer density has been computed by depositing the gas tracer mass in the nearest cell. The axes have been normalised so that the radius of the blast is one at the latest output. The error bars have been estimated assuming that the number of tracers per radial bin is given by a Poisson distribution. This assumption is discussed in more detail in Sect. 3.2.2.

At all stages of the blast, the tracer particles radial profile matches that of the gas at percent levels. This is more easily seen in the top panel of Fig. 5 where the relative difference between the gas tracer density and the gas density is plotted. The errors are all within a few percent and consistent with random fluctuations. As the explosion expands, the swept-up mass of gas in the shocked region increases. This is well tracked by the tracer distribution. Because the mass increases, the total number of tracer particles in the shock increases proportionally, causing the sample noise to decrease. In this particular test, the tracer distribution accurately reproduces that of the gas in the pre- (which is trivially that of the initial distribution) and post-shocked regions (shocked shell plus hot bubble interior). The noise level is a function of the number of tracer particles; its expected value is proportional to the total gas mass only.
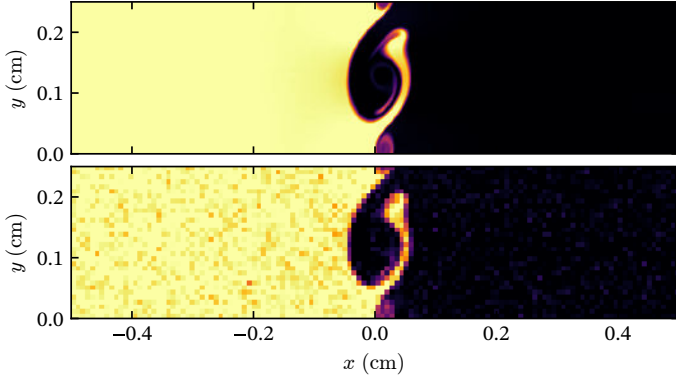
Fig. 6. Projection of the density (*top panel*) and of the gas tracer density (*bottom panel*) around a developing Kelvin–Helmoltz instability. To reduce the noise of the gas tracer projection, we have superposed the four projections of the forming rollers (each of size 0.25 cm). The gas tracer distribution resembles that of the gas with extra noise due to their stochastic nature.

The Sedov explosion is a reliable way of testing the ability of hydrodynamical codes to deal with shocks: more specifically it tests the ability of the code to capture the shock dynamics properly and also tests that the code resolves the shock interface with a few cells in a regime where the Mach number is largely above 1. Here, the gas tracer distribution has been shown to match that of the gas to a high degree of confidence, confirming that the gas tracers are correctly transported with the flow and are able to resolve shocks.

3.1.3. Kelvin–Helmholtz instability

We ran a classical Kelvin–Helmholtz (KH) instability in three dimensions to compare the gas density to the gas tracer density projected maps. The gas has an adiabatic index $\gamma = 7/5$[5]. The simulation is performed on a 128^3 grid with a physical size of 1 cm and a maximum level of refinement of 2^{10}. Cells are refined based on the relative variation of the density: a new level is triggered when the local relative variation of the density is larger than 1%. Only hydrodynamics is included. The instability is initialised with two regions of left and right density of $2 \, \mathrm{g \, cm^{-3}}$ and $1 \, \mathrm{g \, cm^{-3}}$, and of tangential velocity $u_{\mathrm{y,L}} = -1 \, \mathrm{cm \, s^{-1}}$ (resp. $u_{\mathrm{y,R}} = 1 \, \mathrm{cm \, s^{-1}}$). The instability was initially triggered by adding a small damped sinusoidal perturbation of the perpendicular velocity field $u_x = u_0 \cos(k(x - \lambda/2)) \exp(-k|x - x_0|)$, where $\lambda = 0.25$ cm, $k = 2\pi/\lambda$, $x_0 = 0.5$ cm and $v_0 = 0.1 \, \mathrm{cm \, s^{-1}}$. Here 2 900 000 gas tracers were initially distributed in the box, so that their Eulerian distribution matched that of the gas.

Figure 6 shows a projection of the gas density and of the tracer density at time $t = 0.3$ s, when the Kelvin–Helmholtz was already settled. The gas tracer distribution reproduces well the vortices found in the gas distribution, with extra noise due to the reduced number of tracer particles.

The largest k wave numbers of the perturbation are the first to grow following a KH growth timescale of $\tau_{\mathrm{KH}} = 2\pi \mathcal{R}^{1/2}/(|\Delta u|k)$, with $\pm \mathcal{R} = \rho_{\mathrm{R}}/\rho_{\mathrm{L}}$, and $\Delta u = u_{\mathrm{y,R}} - u_{\mathrm{y,L}}$. Therefore, as time proceeds, larger rollers develop in the shear interface between the two phases of gas, and hence, the mixing layer spreads further. We computed the evolution of the cross-section profile of the density at different times. The results are presented in Fig. 7. The phase-mixing region grows as a function of time and the growth

---

[5] This value is consistent with the adiabatic index of air at 20°.
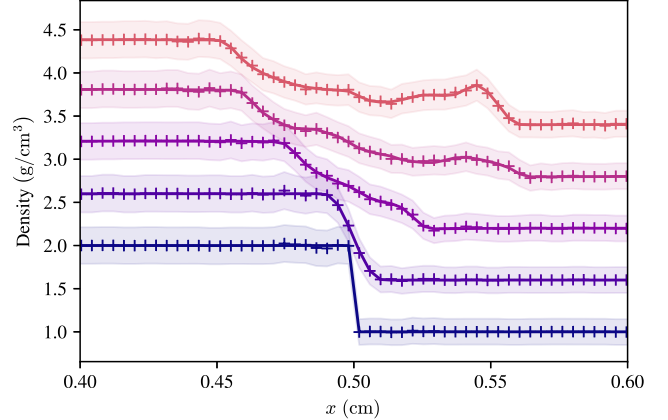
Fig. 7. Evolution of the cross-section of the gas density (solid lines) and the gas tracer density (symbols and shaded regions) for the Kelvin–Helmholtz instability at different times (from blue to red from the start to the end of the simulation at $t = 0.3$ s). The profiles have been shifted vertically (each by $0.6 \, \mathrm{g \, cm^{-3}}$) so that one may easily distinguish them from one another. The shaded regions are $\pm 5\sigma$, where σ has been estimated using a Poisson sampling noise. The gas tracers are accurately following the diffusion of the gas.

is correctly captured by the tracer particles that are able to track it within their intrinsic noise level. Therefore, the gas tracer particles are able to correctly capture the KH shear instability leading to mixing of two gas phases. Interestingly, the present algorithm does not lead to any relative diffusion between the gas and the tracers, as is illustrated quantitatively in Sect. 3.1.1.

3.1.4. AGN feedback

We subsequently tested the accuracy of the mass transfer for the jet mode of AGN feedback, which transfers part of the gas of the central cell to the surrounding cells within a "capsule" region (see Sect. 2.5 for details). We ran an idealised simulation of a halo with an AGN at its centre. The simulation is performed on a coarse grid of 128^3, refined according to a quasi-Lagrangian refinement criterion: a cell is refined/derefined wherever the mass resolution is above/below $1.4 \times 10^7 \, M_\odot$ up to a maximum level of refinement of 12. The box size is 1.2 Mpc, hence with a minimum cell size of 300 pc. The max level of refinement is also enforced in all the cells closer than $4\Delta x$ from the SMBH, where Δx is the minimum cell size. The gas distribution follows a NFW (Navarro et al. 1997) gas density profile, while the dark matter part follows a similar NFW profile modelled with a static gravitational profile (no back reaction of gas onto dark matter). The NFW profile has parameters $V_{200} = 200 \, \mathrm{km \, s^{-1}}$ (at 200 times the critical density of a $H_0 = 70 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$ Universe), a concentration of $c = 6.8$, and is 10% gas. The gas is initially put at rest and at hydrostatic equilibrium. A SMBH of mass $M_{\mathrm{SMBH,0}} = 3.5 \times 10^{10} \, M_\odot$[6] is set at the centre of the box and 10^6 tracers are set in the cell containing the black hole. We force the AGN to be in jet mode with a fixed direction in space and boost its efficiency so that all the tracer particles are sent into the jet in one time step. The radius and height of the jet is $r_{\mathrm{AGN}} = 50$ kpc. This value is much larger than usual values which are usually a few times the cell resolution (here typical values would be a

---

[6] We note that the SMBH mass is taken anomalously high for a typical halo mass of $M_{200} \simeq 3 \times 10^{12} \, M_\odot$. This is chosen simply to get a sufficient power of the jet through the Bondi accretion rate given the NFW distribution of gas.

few kiloparsecs). This is chosen so that the jet reaches cells at different levels of refinement and in other CPU domains. Within 50 kpc of the AGN, there are 1200, 24 000, 12 000, 13 000 and 8000 cells at levels 2^8 to 2^{12} (Δx from 5 kpc to 0.3 kpc) so that the tracer particles are deposited in regions of different refinement level. This region also covers 8 of the 16 CPU domains used. This controlled test enables us to check that the distribution of tracers sent through the jet matches the expected distribution, in the presence of deep refinement and parallelism.

Let us first present the theoretical probability distribution function as a function of the distance to the jet and along the jet. We then compare theoretical figures to those of the simulation. The marginal probability density function (PDF) in the direction of the jet $r_\parallel$ is given by

$$p(r_\parallel) = \frac{1}{A} \begin{cases} \sqrt{e} - e^{r_\parallel^2/2r_{AGN}^2}, & \text{if } |r_\parallel| < r_{AGN}, \\ \sqrt{e} - 1, & \text{if } r_{AGN} < |r_\parallel| < 2r_{AGN}, \end{cases} \quad (16)$$

where

$$A = 2\sqrt{e}r_{AGN}\left(2 + \sqrt{2}F\left(1/\sqrt{2}\right) - 1/\sqrt{e}\right). \quad (17)$$

Here F is Dawson's integral. The marginal PDF in the radial direction $r_\perp$ is

$$p(r_\perp) = \frac{r_\perp e^{-r_\perp^2/2r_{AGN}^2}\left(1 + \sqrt{1 - r_\perp^2/r_{AGN}^2}\right)}{r_{AGN}^2\left(2 - \sqrt{2}F\left(1/\sqrt{2}\right) - 1/\sqrt{e}\right)}. \quad (18)$$

The marginal PDF in the radial distribution is similar to a χ distribution with two degrees of freedom with an extra factor due to the two spherical caps: more particles are found close to the centre of the jet since the capsule is more extended close to its centre.

Figure 8 presents the results from the comparison of the simulation to the expected distribution. The distribution in the radial direction has been rescaled by a factor of two to span the same range as in the parallel direction. Theoretical curves (Eqs. (16) and (18)) are in very good agreement with the observed distributions, confirming that the algorithm is distributing tracer particles correctly in jets. In addition we have also run the same idealised simulation without forcing the AGN efficiency. We report that the tracer mass flux is equal to the gas mass flux. This confirms that the physical model of the jet is accurately sampled by the tracer particles interacting with it, both in terms of its mass and for its spatial distribution.

3.2. Astrophysical test

We have run a 50 cMpc/h-wide cosmological simulation down to $z = 2$ zoomed on a group of mass $1 \times 10^{13} M_\odot$ at $z = 0$, where the size of the zoom in the Lagrangian volume of initial conditions is chosen to encapsulate a volume of two times the virial radius of the halo at $z = 0$. We start with a coarse grid of 128^3 (level 7) and several nested grids with increasing levels of refinement up to level 11. The adopted cosmology has a total matter density of $\Omega_m = 0.3089$, a dark energy density of $\Omega_\Lambda = 0.6911$, a baryonic mass density of $\Omega_b = 0.0486$, a Hubble constant of $H_0 = 67.74\,\text{km s}^{-1}\,\text{Mpc}^{-1}$, a variance at 8 Mpc $\sigma_8 = 0.8159$, and a non-linear power spectrum index of $n_s = 0.9667$, compatible with a Planck 2015 cosmology (Planck Collaboration XIII 2016).

The simulation includes a metal-dependant tabulated gas-cooling function following Sutherland & Dopita (1993) allowing the gas to cool down to $T \sim 10^4\,\text{K}$ via Bremsstrahlung
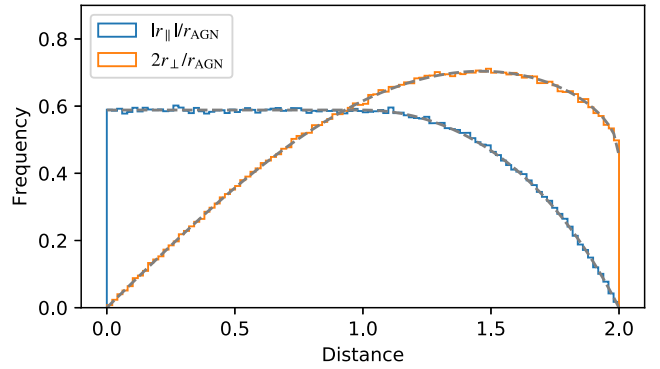
Fig. 8. Distribution of particles moved by a jet before any hydrodynamical time step has occurred. Shown is the parallel distribution marginalised over the plane of the jet (blue) and the radial distribution marginalised over the direction of the jet (orange) vs. the expected theoretical distributions from Eqs. (16) and (18) (dashed grey). The abscissa is in units of r_{AGN} in the parallel direction and in units of $r_{AGN}/2$ in the radial direction. The distribution of gas tracers sent into the jet perfectly matches the expected one.

radiation (effective until $T \sim 10^6\,\text{K}$), and via collisional and ionisation excitation followed by recombination (dominant for $10^4\,\text{K} \leq T \leq 10^6\,\text{K}$). The metallicity of the gas in the simulation is initialised to $Z_0 = 10^{-3} Z_\odot$ to allow further cooling below $10^4\,\text{K}$ down to $T_{min} = 10\,\text{K}$. Reionisation occurs at $z = 8.5$ using the Haardt & Madau (1996) model and gas self-shielding above $10^{-2}\,\text{m}_p\,\text{cm}^{-3}$. Star formation is allowed above a gas number density of $n_0 = 10\,\text{H cm}^{-3}$ according to the Schmidt law and with an efficiency ε_{ff} that depends on the gravo-turbulent properties of the gas (for details, see Kimm et al. 2017; Trebitsch et al. 2017). The main distinction of this turbulent star-formation recipe with the traditional star formation in RAMSES (Rasera & Teyssier 2006) is that the efficiency can approach and even exceed 100% (with $\varepsilon_{ff} > 1$ meaning that stars are formed faster than in a free-fall time). The stellar population is sampled with a Kroupa (2001) initial mass function, where $\eta_{SN} = 0.317$ and the yield (in terms of mass fraction released into metals) is 0.05. The stellar feedback model is the mechanical feedback model of Kimm et al. (2015) with a boost in momentum due to early UV pre-heating of the gas following Geen et al. (2015). The simulation also tracks the formation of SMBHs and the evolution of AGN feedback in jet mode (radio mode) and thermal mode (quasar mode) using the model of Dubois et al. (2012b). The jet is modelled in a self-consistent way by following the angular momentum of the accreted material and the spin of the black hole (Dubois et al. 2014b). The radiative efficiency and spin-up rate of the SMBH is then computed using the MAD results of McKinney et al. (2012).

We have a minimum roughly constant physical resolution of 35 pc (one additional maximum level of refinement at expansion factor 0.1, 0.2, and 0.4), a star particle mass resolution of $m_{\star,res} = 1.1 \times 10^4 M_\odot$, a dark matter (DM) particle mass resolution of $m_{DM,res} = 1.5 \times 10^6 M_\odot$, and gas mass resolution of $2.2 \times 10^5 M_\odot$ in the refined region. A cell is refined according to a quasi-Lagrangian criterion: if $\rho_{DM} + \rho_b/f_{b/DM} > 8m_{DM,res}/\Delta x^3$, where ρ_{DM} and ρ_b are respectively the DM and baryon density (including stars plus gas plus SMBHs), and where $f_{b/DM}$ is the cosmic mean baryon-to-DM mass ratio. The max level of refinement is also enforced in all cells closer than $4\Delta x$ from any SMBH, where Δx is the minimum cell size. We add tracer particles in the refined region with a fixed mass of $m_t = 2.0 \times 10^4 M_\odot$
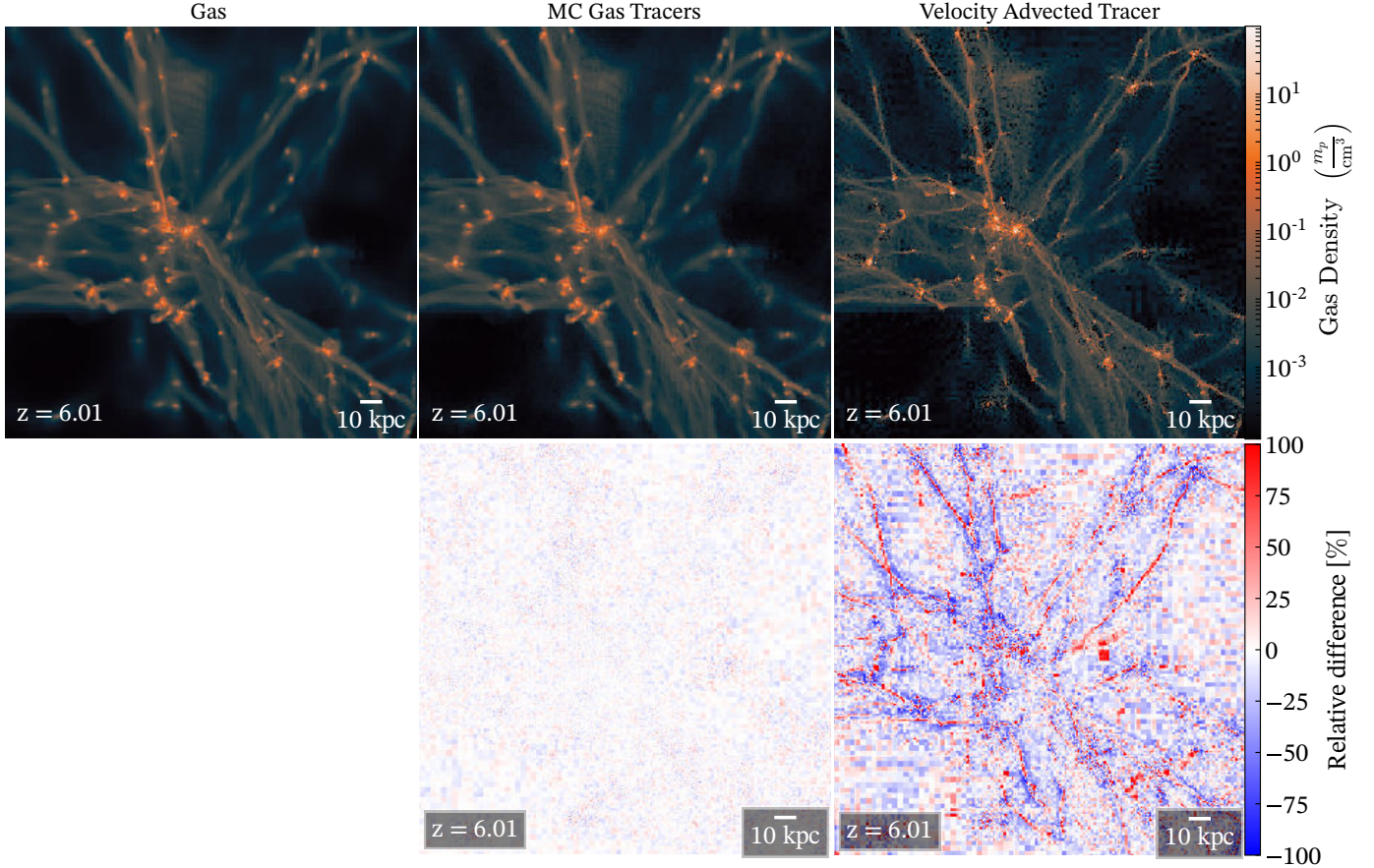
Fig. 9. *Top panels*: density weighted projection of the gas density in a cosmological simulation (*left*), of the velocity tracer distribution (*right*), and of the MC gas tracer distribution (*centre*). All the plots share the same colour map. *Bottom panels*: relative difference between the tracer and the gas. Velocity tracers accumulate in convergent regions (e.g. filaments, nodes). The MC gas tracer distribution reproduces more accurately that of the gas than velocity tracers.

($N_{tot} \approx 1.3 \times 10^8$ particles). There is on average 0.55 tracers per star and 22 per initial cell. Cells of size 35 pc and density 20 cm^{-3} contain on average one tracer per cell.

3.2.1. Velocity tracers versus Monte Carlo tracers

In addition to the above simulation, we ran the exact same one replacing each MC tracer with a velocity-advected tracer. This simulation was performed down to $z = 6$ and compared to the fiducial one. Both have a similar gas distribution, confirming that the tracer particles are indeed passive[7]. At this redshift, 99% of the baryons are still in the gas phase (0.72% in stars and 8×10^{-5}% in SMBHs), meaning that the comparison between MC tracers (that can be transferred into stars) and velocity tracers is fair when looking at cosmological scales. Since the velocity tracers have not been linked to star formation or SMBHs, we expect significant discrepancies within galaxies, where the gas-to-star ratio is much smaller.

The top panels of Fig. 9 show projections of the density-weighted density of gas (top left panel), of MC tracers (top-centre panel), and of velocity-advected tracers (top-right panel). The distribution of the MC tracers resembles that of the gas with extra noise due to sampling noise. All the prominent structures

in the gas are also present in the MC tracer distribution. On the other hand, the velocity tracer distribution is much sharper than that of the gas. The velocity tracers aggregate in converging flows (filaments and centres of galaxies) while MC tracers do not (they aggregate in high-mass regions, as expected). At such large scales, the origin of the discrepancy is an intrinsic issue of velocity tracers. This test shows that on a qualitative level, the MC tracers have a distribution that is in much better agreement with the gas distribution than the velocity advected tracers. The relative difference between the gas distribution and the tracer distribution is presented in the bottom panels of Fig. 9. The relative difference between the MC tracer density and the gas density (bottom central panel) is significantly smaller than the relative difference between the velocity advected tracer density and the gas density (bottom right panel). The latter is also much more biased: the velocity advected tracer density in convergent flows (e.g. filaments) can be up to an order of magnitude larger than the gas density, while in the vicinity of converging regions, the velocity advected tracer density is largely underestimated (e.g. around filaments). On the contrary, the MC tracer density is found to be in better agreement with the gas density and is not biased.

3.2.2. Gas tracers

As we have seen, velocity tracer particles are a less reliable tracer of the actual gas density compared to MC tracer particles, and this can already be seen on cosmological scales. Therefore, we

---

[7] They have however an indirect impact on stochastic processes such as star formation and SN feedback as they impact the random number generator (hence the outcome of these random processes will vary depending on how many and where the tracer particles are).
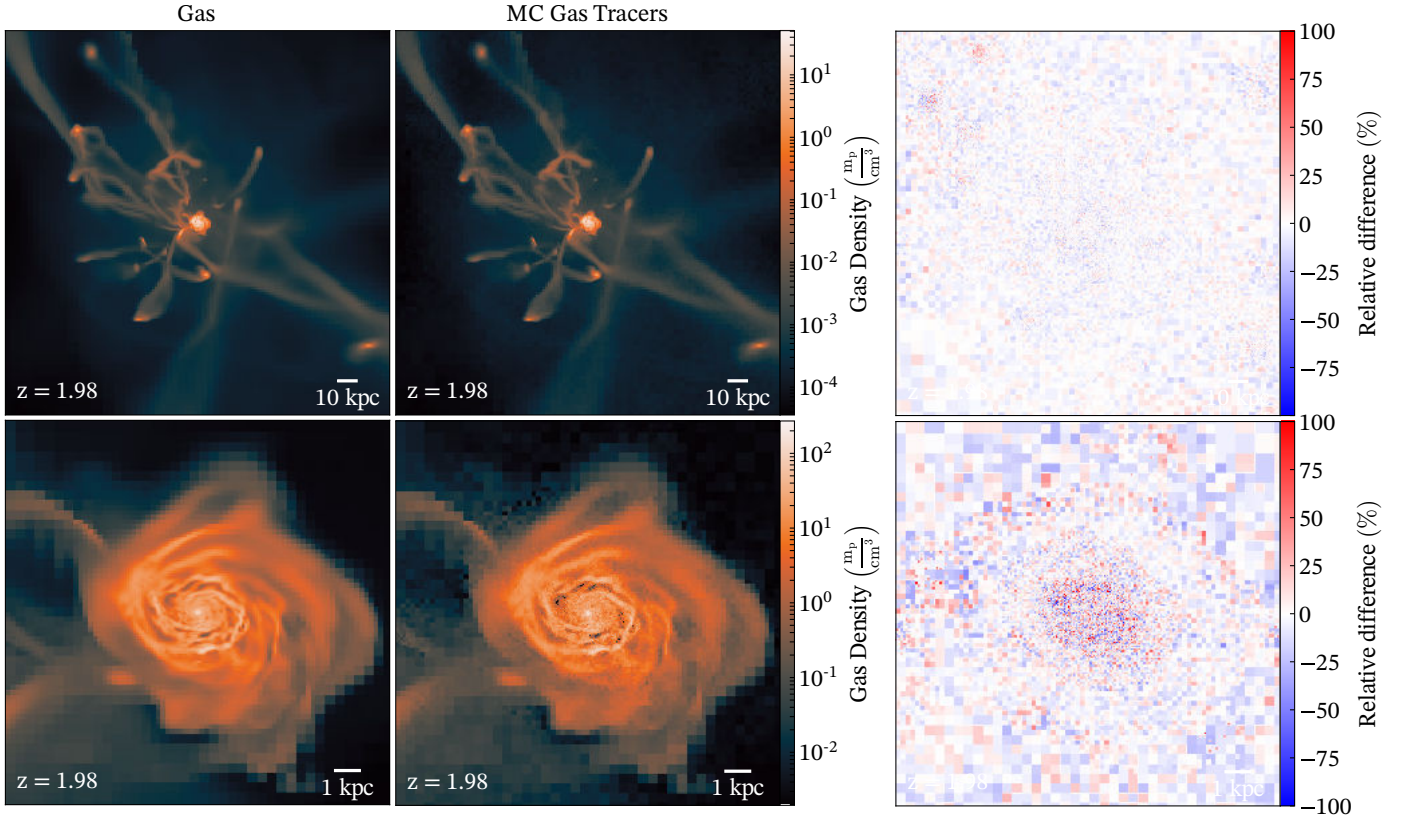
Fig. 10. Density-weighted projection of the gas density (*left panels*), of the gas tracer density (*centre panels*), and of their relative difference (*right panels*) along the *x* axis around the most massive galaxy of the cosmological simulation at *z* = 2. *Top panels*: large-scale structure of the gas; data have been selected within 200 kpc of the centre. *Bottom panels*: zoom on the central galaxy; data have been selected within 10 kpc of the centre of the galaxy. The MC tracer density is similar to that of the gas. The radial modulations are due to differences in cell mass at fixed cell resolution: massive cells (closer to the centre at fixed resolution) are best sampled by the MC tracers.

now continue to explore only the distribution of MC tracer particles with respect to the actual distribution of baryons. Figure 10 shows the density-weighted projected gas density and cloud-in-cell interpolated gas tracers around the zoomed galaxy of the simulation. Visual inspection reveals that the gas tracer distribution matches that of the gas with additional noise. All structures with a contrast above the noise level are reproduced by the gas tracers. More quantitatively, Fig. 11 shows the density of tracers versus the density of gas for the entire available range of gas densities (i.e. 9 orders of magnitude); the expected one-to-one relation is seen, with some scatter due to MC sampling noise.

More quantitative results can be obtained by computing the statistical properties of the gas tracer population. A cell of mass M_{cell} is expected to contain on average M_{cell}/m_t tracers. For a sample of cells of similar masses, we expect the mean number of tracers per cell to be $\lambda \equiv \langle M_{cell} \rangle / m_t$. The distribution of the number of tracers per cell in the simulation is shown in Fig. 12 for different cell-mass bins. Within a cell-mass bin, the number of tracers N_t can be seen to be very well approximated by a Poisson distribution with parameter λ

$$p_\lambda(N_t = k) = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{19}$$

To confirm this observation, we compared the mean number of tracers per cell to the expected number λ in the top panel of Fig. 12. For all cell masses, the mean number of tracer particles per cell is accurately described by its expected Poisson distribution. At large values of gas mass within a cell (right of the plot),

the scatter in the histogram count is due to the small number of massive cells in the simulation. Indeed, these cells can only be found in the most refined regions (otherwise they would be refined into smaller cells) where they also tend to be converted into stars.

In the following we assume that the gas tracer distribution is given by a Poisson distribution with parameter $\lambda = \langle M_{cell} \rangle / m_t$. This yields a simple rule of thumb to estimate the precision of the tracer scheme. The accuracy of the Eulerian distribution of the tracer can be written $1/\sqrt{\lambda} \sim \sqrt{m_t/M_{cell}}$.

3.2.3. Star formation and feedback

Figure 13 shows the integrated stellar mass and star-tracer mass around the zoomed galaxy of the cosmological simulation. Both distributions are visually in agreement and feature the same spatial distribution. At large radii where the star density is smaller than the gas density ($r \gtrsim 4$ kpc, see Fig. 14), the noise level of the star-tracer distribution is larger than that of the gas. This is due to the fact that small masses are poorly resolved by the MC tracers. Close to the galactic centre, the increasing star density induces a larger star-tracer density, and therefore, at fixed resolution, a smaller noise sampling. This is illustrated by the right panel of Fig. 13, where the centre of the plot shows smaller fluctuations than at large radii. More quantitative results are presented below.

We first present the analytical distribution of tracer particles for stars and for the number of tracers released in SN events, derived from first principles. When a star particle is formed, each tracer in the cell containing the newly created star
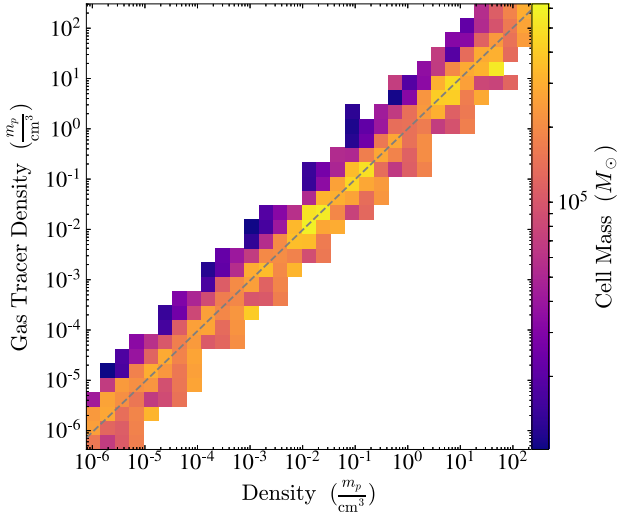
Fig. 11. Gas density vs. gas tracer density, colour coded by cell mass. The grey dashed line shows the one-to-one relation. The gas and gas tracer densities match on nine orders of magnitude.

Fig. 12. *Bottom panel*: distribution of the number of gas tracers for different cell-mass bins as observed in the simulation (solid lines) vs. a Poisson distribution with parameter $\lambda = \langle M_{\text{cell}} \rangle / m_{\text{t}}$ (dashed lines, reported in the legend). *Top panel*: relative difference between the observed mean number of tracer particles and the expected one, λ, as a function of λ. For all cells, the distribution of the number of gas tracers per cell is given by a Poisson distribution with parameter λ.

particle is attached to the star particle and has a probability of $p_\star \equiv M_{\star,0} / M_{\text{cell}}$ of becoming a "star tracer", where $M_{\star,0}$ is the mass of the newly created star particle[8]. Because $M_{\star,0} < M_{\text{cell}}$ – a star particle cannot be formed with more material than what is available – this probability is well defined: $0 < p_\star < 1$. When the heavy stars in a star particle go into SN, they yield $\eta M_\star$, and the mass of the corresponding star particle becomes $M_\star = (1 - \eta)M_{\star,0}$. The star tracers are then returned to the gas with a probability of η. Before the SNe explode, the distribution of tracers for an individual star particle is given by a binomial distribution with parameters N_i (the initial number of tracer in the cell where the star particle formed) and $p_\star$

$$p_{\text{form}}(N_i; N_f = k) = \binom{N_i}{k} p_\star^k (1 - p_\star)^{N_i - k}. \tag{20}$$

The number of tracer particles released in the SN event reads

$$p_{\text{SN}}(N_f; N = k) = \binom{N_f}{k} \eta^k (1 - \eta)^{N_f - k}, \tag{21}$$

where N_f is the number of star tracers in the star particle before the SN explosion. The number of tracers in the star particle after the SN has exploded is, thus, given by a binomial distribution of parameters N_i and $(1 - \eta)p_\star$,

$$p_\star^{\text{f}}(N_i; N = k) = \binom{N_i}{k} ((1 - \eta)p_\star)^k (1 - (1 - \eta)p_\star)^{N_i - k}. \tag{22}$$

In the limit where the N_i becomes large and $(1 - \eta)p_\star$ small, Eq. (22) converges mathematically to a Poisson distribution with parameter $N_i(1 - \eta)p_\star$.

Now, we compare the expected distribution of tracer particles to the measured one. Figure 15 presents the distribution of the number of tracer particles per star particle for different star particle mass bins. The number of star tracers per star particle can be seen to be well approximated by a Poisson distribution with parameter $\lambda = \langle M_\star \rangle / m_{\text{t}}$. There is a clear deviation at the tail of the distribution which displays an excess of probability.

This is however expected as when a star forms in a cell, a significant part of the cell mass is converted into the star, so that $p_\star \approx 1$. Because usually $(1 - \eta) \approx 0.9$, the product $p_\star(1 - \eta)$ is also of order unity. At the same time, cells where stars form have a typical mass of $10^4 M_\odot \sim m_{\text{t}}$, meaning that they contain only a few gas tracers at star formation. Therefore, we expect a significant deviation from a Poisson distribution, as the requirement for Eq. (22) to converge to a Poisson distribution is not met. This argument is reinforced by the fact that, compared to light stars (e.g. the blue curve of Fig. 15), the most massive stars have a more top-heavy distribution (e.g. the red curve) than a Poisson distribution. Indeed, these massive stars are relatively more massive than their parent cell, meaning that the parameter $p_\star$ is larger. In the simulation, star formation is only activated for cells above a given (fixed) density threshold. This is usually achieved at the maximum resolution, causing cells experiencing star formation to have typically the same mass, and therefore the same number of gas tracer particles, regardless of the mass of the forming stars. Consequently, the massive star particle distribution is indeed less Poissonian than that of the light stars, since their $p_\star$ is larger at fixed N_i. Figure 15 is in qualitative agreement with this.

3.2.4. SMBH evolution

Using our cosmological simulations, we have checked that the total mass of SMBH tracer particles ($M_{\text{t SMBH,tot}} = (3.5 \pm 0.3) \times 10^6 M_\odot$[9]) matches that of SMBH in the simulation ($M_{\text{SMBH,tot}}/(1 - \varepsilon_{\text{r}}) = 3.1 \times 10^6 M_\odot$) at the 10% level, up to an ε_{r} factor. This factor is due to the mass lost by the accreted material as it falls onto the black hole. This mass is radiated away and lost to the simulation. Because the tracer particles have a fixed mass in our implementation, they are unable to capture the mass energy that is radiated. However, one could store the value of ε_{r} at accretion time onto each tracer to be able to reconstruct the exact mass that the SMBH tracer represents.

---

[8] We note that in practice the star particles have a mass that is a multiple of the stellar mass resolution.

[9] The uncertainty has been estimated using a 1-σ Poissonian noise.
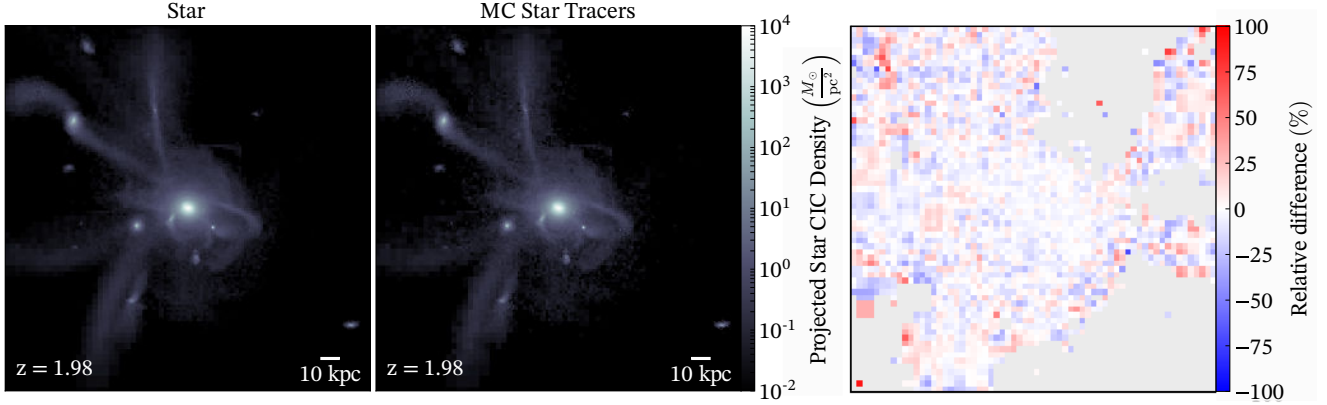
Fig. 13. Stellar surface density (*left panel*), star-tracer surface density (*centre panel*), and relative difference (*right panel*). The data are the same as in Fig. 10. In the difference map, regions where no stars are found are indicated in grey. The star and star-tracer distributions are in very good agreement; their difference shows no spatial dependence. The noise level is higher than in Fig. 10 at large radii where the star surface density is smaller than the gas surface density, hence the star mass distribution is less resolved than the gas.
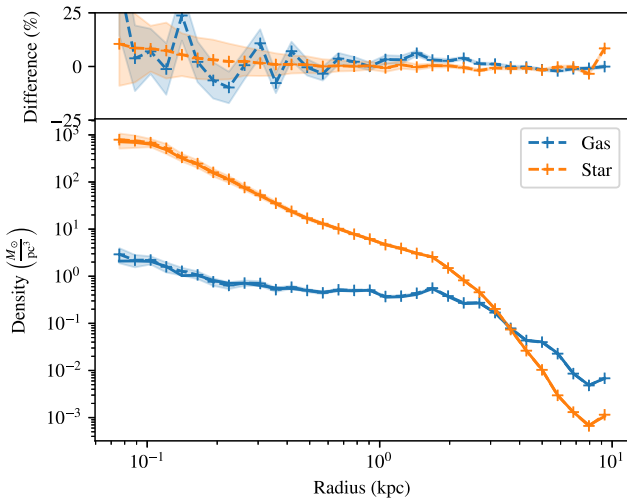
Fig. 14. *Bottom panel*: radial profile of the gas density (solid blue) and star density (solid orange) vs. the gas tracer density (blue cross) and the star-tracer density (orange cross). The error bars are given by a Poisson sampling noise. *Top panel*: relative difference between the baryon and the tracer profiles. The tracers match their baryon counterpart at a few percent level.

3.3. Bi-modal accretion at high redshift: a science case for tracer particles

Low-mass galaxies (embedded in halos $M_h \lesssim 10^{11} M_\odot$) exhibit a significant amount of "cold-mode" cosmological accretion made of cold gas streaming in narrow filaments with a temperature typically below $T_{max} \lesssim 10^5$ K (Birnboim & Dekel 2003; Kereš et al. 2005; Ocvirk et al. 2008; Nelson et al. 2013, 2016). A "hot-mode" phase made of gas that was shock heated before entering the virial radius ($T_{max} \sim 10^6$ K) appears in halos with higher mass. At early times ($z > 2.5$), the accretion is dominated by the cold mode. As time goes by, halos grow in mass so that an increasing fraction of the gas heats up before entering the halo. The outcome of this is a decrease of the relative importance of cold accretion compared to hot accretion. By $z \lesssim 2$, most of the accreted material comes from the diffuse hot phase. Hence, getting access to the Lagrangian history of the stars and of the star-forming gas is key to pinning down the origin of gas acquisition in galaxies.
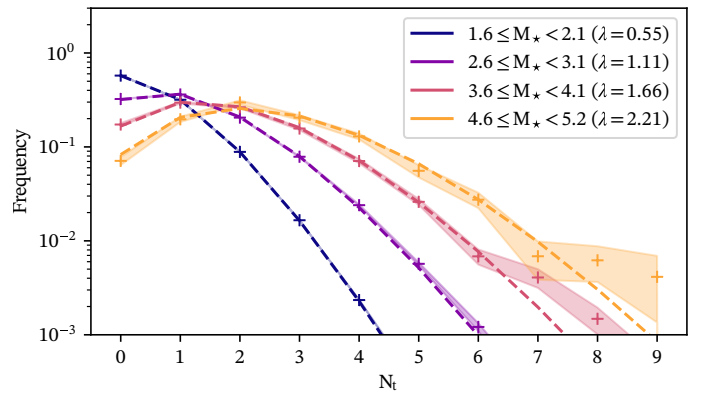
Fig. 15. Distribution of the number of star tracers per star for different star particle mass bins (in units of $10^4 M_\odot$) as observed in the simulation (symbols and shaded surfaces) vs. as given by a Poisson distribution with parameter $\lambda = \langle M_\star \rangle / m_t$ (dashed). The error bars have been estimated using a bootstrap method. For all stars, the distribution of the number of star tracers per star is approximated by a Poisson distribution with parameter λ.

We revisit this result using RAMSES and the MC tracer particles. Using the cosmological simulation of Sect. 3.2, we study the accretion of gas as a function of time around the central galaxy. We select all the gas tracers that end up in star particles (not the star-forming gas) at $z = 2$ and $r < 0.1 R_{vir}$. The halos were detected using the AdaptaHOP halo finder (Aubert et al. 2004). For the positioning of the centre of the DM halo, we start from the first AdaptaHOP guess of the centre (densest particle in the halo) and from a sphere the size of the virial radius of the halo; we use a shrinking sphere (Power et al. 2003) by recursively finding the centre of mass of the DM within a sphere 10% smaller than the previous iteration. We stop the search once the sphere has a size smaller than $\simeq 100$ pc and take the densest particle in the final region. Twenty neighbours are used to compute the local density. Only structures with a density greater than 80 times the average total matter density and with more than 200 particles are taken into account. The original AdaptaHOP finder is applied to the stellar distribution in order to identify galaxies with more than 200 particles. Their Lagrangian history is reconstructed in post-processing from the 132 equally spaced ($\Delta t = 25$ Myr) outputs, and the thermodynamical properties of
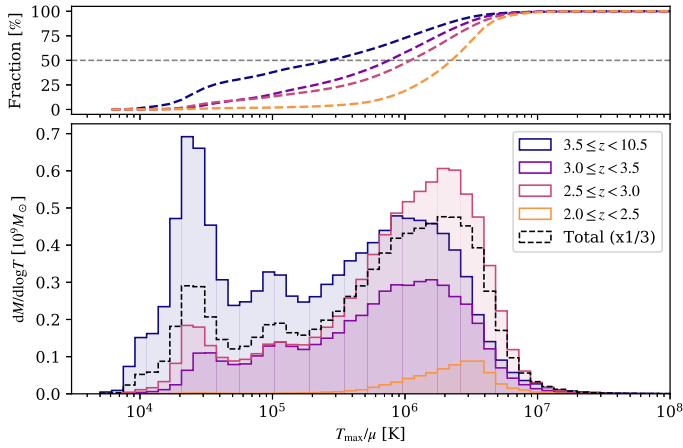
Fig. 16. *Bottom panel*: histogram of the maximum temperature of the gas accreted onto the central galaxy between different redshifts (from early accretion time in blue to late accretion time in yellow). *Top panel*: cumulative distribution of the gas temperature. Only the gas-forming stars within the virial radius are selected. The total distribution integrated over the total accretion time is shown with the black dashed line in the bottom panel. The total distribution has been rescaled by a factor of one third for visualisation. The halo has two modes of accretion: a cold and a hot mode. At high z the cold mode dominates and at low z the hot mode dominates.

the gas are extracted from the local gas cell value. For each tracer particle, the maximum temperature $T_{\max}$ reached before falling into the virial radius is recorded. The infall time is defined as the last inward crossing of the virial radius. The merger tree is computed following Tweed et al. (2009). The procedure only selects tracer particles falling onto the galaxy in the gas phase. This excludes gas tracers tracking gas that formed stars in satellite galaxies but includes gas from wet mergers. Figure 16 presents the temperature distribution of the accreted gas for different bins of infall time. At early times (blue lines, $z \gtrsim 3$) the accretion is bi-modal. About 50% of the gas is accreted via the cold mode, as shown in the top panel of Fig. 16. At later redshifts ($z \lesssim 2.5$), the accretion becomes dominated by the hot mode. The relative importance of the cold accretion decreases and the distribution become less and less bimodal, until it is eventually entirely dominated by the hot mode. This is in qualitative agreement with the findings of Kereš et al. (2005) though the exact quantitative amount of cold versus hot accreted gas relies significantly on i) the numerical scheme to model gas dynamics (Nelson et al. 2013) and ii) the modelled feedback processes (Dubois et al. 2013).

Caution should be taken here: contrary to what was done in the original study, only the accretion onto a single galaxy is investigated. In particular, our results are sensitive to the particular accretion and merger history of that galaxy, which impact the temperature distribution of the gas. In order to achieve a fairer comparison, one would have to run a full cosmological simulation and study the gas accretion of the full population within the box. While this would now technically be possible thanks to the new tracer algorithm, it is nonetheless well beyond the scope of this paper.

4. Performance

To quantify the performance of the tracer particles and their associated CPU overhead (defined as the excess of computation time required by the tracer particles), we restarted the simulation of
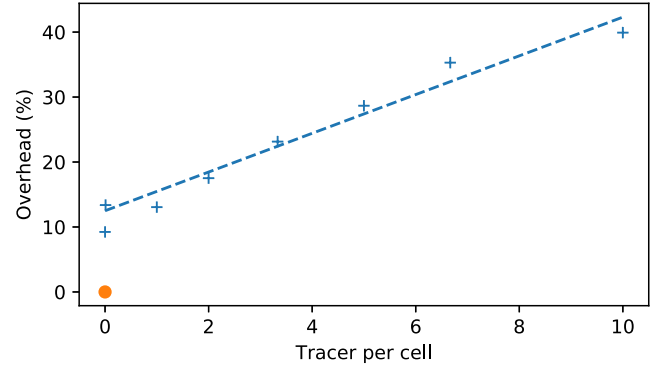
Fig. 17. Overhead as a function of the number of tracer particles per initial cell (symbols). The orange symbol is the simulation with the tracer deactivated. The data (excluding the run with the tracer deactivated) have been fitted with a linear function (dashed line). The estimated overhead (slope of the fit) is ~3% per tracer per initial cell with an extra constant of ~10%.

Table 1. Run time per coarse time step for the different runs.

Name	Absolute number	Tracer per cell	Run time (s)	Overhead (%)
t100	129325116	10	1310	39.9
t67	86214303	6.7	1270	35.3
t20	64656206	5	1210	28.7
t33	43104621	3.3	1160	23.1
t20	25861310	2	1100	17.5
t10	12929077	1	1060	13.1
t0.1	130250	0.01	1060	13.4
t0	0	0	1020	9.2
not	0	0	940	–

Notes. The run `notr` was performed with no tracer particles and with all the tracer particle routines deactivated. The column "Tracer per cell" is the number of tracer particles per initial cell in the zoomed region. The "Overhead" column contains the run-time overhead defined with respect to the `notr` run.

Sect. 3.2 at redshift $z = 2$, while varying the numbers of tracer particles to test the scaling of the algorithm. At restart, we decimate the tracer population to keep only 67, 50, 33, 20, 10, or 0.1% of the initial population (in the gas, star, and black holes). We also run a simulation with no tracer but all the tracer routines activated (`t0`) and a simulation with no tracer and the tracer routines deactivated (`notracer`). The parameters of the runs are presented in the first three columns of Table 1. The run time is defined as the total run time divided by the number of steps. The overhead is defined as the relative increase of the run time with respect to the run `not`. All the runs were stopped after two iterations of the coarse time step (about ~2000 s of run time, ~2.8 Myr of simulation time). The results are also plotted in Fig. 17.

By comparing the two runs `t0` and `notr`, we conclude that the tracer particle machinery adds a constant cost of about 10% to the computation. This is due to the fact that the tracer particles require the fluxes at the interface of each cell (six quantities per cell) to be stored, which then have to be communicated between CPUs. In addition, there are multiple loops that iterate over all the cells and all the particles (see Sect. 2 for more details). In principle, this could be optimised by setting tracer particles in their own linked list, but we exploited the particle machinery

available in Ramses, and treated tracer particles just like standard particles (star or DM) with respect to code structure. In the following, the computation overhead will be expressed in terms of the number of tracer per initial cell: $N_t/N_{cell,i}$, where N_t is the number of tracer particles and $N_{cell,i}$ is the number of initial (gas) cells.

The runs with tracers show that the total run time starts increasing with the number of tracer particles per cell[10] when this number becomes of the order of ~0.1 tracer per initial cell. Above this threshold, the run time scales roughly linearly with the number of tracer per initial cell. We have run the simulation on the Occigen supercomputer with 672 cores (28 nodes of 24 cores). Each node is made of two Intel Haswell 12-Core E5-2690 V3s[11] running at a clock frequency of 2.6 GHz. The nodes are wired together with a DDR Infiniband network (20 Gbit s^{-1}). The code was compiled with the Intel Fortran compiler version 17.0 and OpenMPI 2.0.2. In this setup the overhead is 3% per tracer per initial cell. For example the run t100 with 10 tracer per initial cell had a 40% overhead. Part of the overhead is due to the tracer particles themselves (moving, generating random numbers, etc.). Another part is due to the load balancing. Indeed, in this simulation, tracer particles are only found in the zoomed region, which is already the most CPU-intensive region. Our simulation can be seen as a worst-case scenario for the tracer particles. In general, let us write the conservative formula giving an estimate of the overhead induced by the tracer particles

$$\frac{\Delta t}{t} = 0.03 \left(\frac{N_t}{N_{cell,i}} \right) + 0.1, \tag{23}$$

where t is the run time and Δt the extra cost induced by the tracer particles. Here, N_t and $N_{cell,i}$ are the total number of tracer particles and the total number of initial cells, respectively.

5. Conclusions

We present a new implementation of tracer particles in the Ramses AMR code based on the Monte Carlo approach from Genel et al. (2013). It has been interfaced with the most common physical models used in cosmological simulations (star formation and stellar feedback, SMBH growth and AGN feedback). We have shown that the Lagrangian history of the gas is accurately reconstructed by testing the accuracy of the tracer distribution in an advection-dominated problem and in a diffusion-dominated problem. The gas tracer distribution matches that of the gas, even in complex situations that involve subgrid models. We have also provided a comparison of the new MC tracer particles to the previous velocity-based implementation and showed that the new version largely outperforms the accuracy of the previous one. We have made a detailed study of the distribution of tracer particles in a zoom-in cosmological simulation including state-of-the art subgrid model physics (cooling, star formation, SN feedback, SMBHs, and AGN feedback) and show that: (i) in each cell, the gas tracer distribution is given by a Poisson distribution with parameter $\lambda = M_{cell}/m_t$; and (ii) for each star, the number of star tracers can be approximated by a Poisson distribution with parameter $\lambda = M_\star/m_t$. The properties of the Poisson distribution give a simple rule to estimate the sampling noise of the tracer particle, as the noise can be represented by $1/\sqrt{\lambda}$. In turn this should allow users to quantify how many particles

are needed to reach their sought accuracy. We have also shown that the gas tracer particles sample exactly the intrinsic numerical diffusion of the Godunov solver. To highlight the assets of tracer particles in a realistic setting, they were implemented in the problem of cold flow accretion at high redshift. The known bi-modality in the temperature of gas was recovered.

The performance of the algorithm was explored. In a zoom-in full physics cosmological simulation, the run time grows roughly linearly with the number of tracer particles per cell. The overall impact on computation time is estimated to be ~3% per tracer per initial cell plus a constant computation time overhead of 10%, regardless of the number of tracer particles. These figures should serve as upper limits on the computation time. The performance of the scheme could be optimised by using two separate linked lists for the tracer particles and the other particles, as is done in Arepo (Genel et al. 2013). Implementing these possible improvements will be the subject of future studies. Presently, the performance is significantly lower than that reported in the original paper of Genel et al. (2013): in addition to using a specific linked list for the tracer particles, the moving mesh of Arepo reduces the number of tracer movements and mitigates the cost of each tracer.

In comparison to the original paper by Genel et al. (2013), we provide an additional detailed description of the statistical properties of the ensemble of tracer particles not only in the gas but also in stars and in AGN jets. We also studied how their distributions behave when complex sub-grid models are involved (star formation and feedback, AGN feedback, BH accretion) and checked that their distribution is in agreement with the baryon distribution.

This implementation provides an efficient method to accurately track the evolution of the Lagrangian history in the Eulerian code Ramses. It opens new perspectives to study how baryon flows interact in hydrodynamical simulations. For instance, tracer particles could be used to quantify the spatial and time evolution of the anisotropically accreted gas, its contribution to the spin of galaxies, and how these processes impact galactic morphology. Specifically, following Tillson et al. (2015), Danovich et al. (2015), and DeFelippis et al. (2017), one could address the following open questions: Where does the angular momentum go? Does it contribute to the spin-up of the galaxies or is it re-distributed before entering the disk? If it is, is it due to turbulent pressure, shock-heating or SN and AGN feedback?

References

Agertz, O., Lake, G., Teyssier, R., et al. 2009, MNRAS, 392, 294
Aubert, D., Pichon, C., & Colombi, S. 2004, MNRAS, 352, 376
Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349
Booth, C. M., & Schaye, J. 2009, MNRAS, 398, 53
Bourne, M. A., & Sijacki, D. 2017, MNRAS, 472, 4707
Bryan, G. L., Norman, M. L., O'Shea, B. W., et al. 2014, ApJS, 211, 19
Danovich, M., Dekel, A., Hahn, O., Ceverino, D., & Primack, J. 2015, MNRAS, 449, 2087
DeFelippis, D., Genel, S., Bryan, G. L., & Fall, S. M. 2017, ApJ, 841, 16
Dubois, Y., & Teyssier, R. 2008, A&A, 477, 79
Dubois, Y., Pichon, C., Haehnelt, M., et al. 2012a, MNRAS, 423, 3616

---

[10] We note that here the number of cells is the one in the refined regions, not the initial number of cells.

[11] See Intel-Xeon-Processor- E5-2690.

Dubois, Y., Devriendt, J., Slyz, A., & Teyssier, R. 2012b, MNRAS, 420, 2662
Dubois, Y., Pichon, C., Devriendt, J., et al. 2013, MNRAS, 428, 2885
Dubois, Y., Volonteri, M., & Silk, J. 2014a, MNRAS, 440, 1590
Dubois, Y., Volonteri, M., Silk, J., Devriendt, J., & Slyz, A. 2014b, MNRAS, 440, 2333
Federrath, C., Glover, S. C. O., Klessen, R. S., & Schmidt, W. 2008, Phys. Scr. Vol. T, 132, 014025
Geen, S., Rosdahl, J., Blaizot, J., Devriendt, J., & Slyz, A. 2015, MNRAS, 448, 3248
Genel, S., Vogelsberger, M., Nelson, D., et al. 2013, MNRAS, 435, 1426
Haardt, F., & Madau, P. 1996, ApJ, 461, 20
Hennebelle, P., & Chabrier, G. 2011, ApJ, 743, L29
Iapichino, L., & Niemeyer, J. C. 2008, MNRAS, 388, 1089
Kereš, D., Katz, N., Weinberg, D. H., & Dave, R. 2005, MNRAS, 363, 2
Kimm, T., & Cen, R. 2014, ApJ, 788, 121
Kimm, T., Cen, R., Devriendt, J., Dubois, Y., & Slyz, A. 2015, MNRAS, 451, 2900
Kimm, T., Katz, H., Haehnelt, M., et al. 2017, MNRAS, 466, 4826
Kravtsov, A. V., Klypin, A. A., & Khokhlov, A. M. 1997, ApJS, 111, 73
Kroupa, P. 2001, MNRAS, 322, 231
Krumholz, M. R., & McKee, C. F. 2005, ApJ, 630, 250
McKinney, J. C., Tchekhovskoy, A., & Blandford, R. D. 2012, MNRAS, 423, 3083
Merloni, A., & Heinz, S. 2008, MNRAS, 388, 1011
Mitchell, N. L., McCarthy, I. G., Bower, R. G., Theuns, T., & Crain, R. A. 2009, MNRAS, 395, 180
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Nelson, D., Vogelsberger, M., Genel, S., et al. 2013, MNRAS, 429, 3353
Nelson, D., Genel, S., Pillepich, A., et al. 2016, MNRAS, 460, 2881

Ocvirk, P., Pichon, C., & Teyssier, R. 2008, MNRAS, 390, 1326
Padoan, P., & Nordlund, Å. 2011, ApJ, 730, 40
Pichon, C., Pogosyan, D., Kimm, T., et al. 2011, MNRAS, 418, 2493
Planck Collaboration XIII. 2016, A&A, 594, A13
Power, C., Navarro, J. F., Jenkins, A., et al. 2003, MNRAS, 338, 14
Price, D. J., Wurster, J., Tricco, T. S., et al. 2018, PASA, 35, e031
Rasera, Y., & Teyssier, R. 2006, A&A, 445, 1
Rosdahl, J., & Blaizot, J. 2012, MNRAS, 423, 344
Rosdahl, J., Schaye, J., Dubois, Y., Kimm, T., & Teyssier, R. 2017, MNRAS, 466, 11
Silvia, D. W., Smith, B. D., & Shull, J. M. 2010, ApJ, 715, 1575
Springel, V. 2005, MNRAS, 364, 1105
Springel, V. 2010, MNRAS, 401, 791
Sutherland, R. S., & Dopita, M. A. 1993, ApJS, 88, 253
Teyssier, R. 2002, A&A, 385, 337
Teyssier, R., Pontzen, A., Dubois, Y., & Read, J. I. 2013, MNRAS, 429, 3068
Thornton, K., Gaudlitz, M., Janka, H.-T., & Steinmetz, M. 1998, ApJ, 500, 95
Tillson, H., Devriendt, J., Slyz, A., Miller, L., & Pichon, C. 2015, MNRAS, 449, 4363
Toro, E. F. 2009, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction (Springer)
Trebitsch, M., Blaizot, J., Rosdahl, J., Devriendt, J., & Slyz, A. 2017, MNRAS, 470, 224
Tweed, D., Devriendt, J., Blaizot, J., Colombi, S., & Slyz, A. 2009, A&A, 506, 647
Vazza, F., Brunetti, G., Gheller, C., Brunino, R., & Brüggen, M. 2011, A&A, 529, A17
Vazza, F., Roediger, E., & Brüggen, M. 2012, A&A, 544, A103
Wadsley, J. W., Stadel, J., & Quinn, T. 2004, New Astron., 9, 137

Appendix A: Tracer particle algorithm

Let us describe here the pseudo-code underlying the tracer particle algorithm. The corresponding FORTRAN code is available upon request.

A.1. Gas to gas cells

The main function in charge of moving tracers between gas cells is called TREATCELL. It takes as input the index of a cell and loops over all tracers in it. It requires all the (mass) fluxes to be stored. The pseudo code is the following.

```
function TREATCELL(i_cell)
    m_cell ← MASSOFCELL(i_cell)
    F_net ← 0
    for i_dir ← 1, 2N_dim do            ▷ Compute outgoing flux
5:      F ← GETFLUXINDIR(i_cell, i_dir)
        if F > 0 then
            F_net ← F_net + F
        end if
    end for
10: tracers ← GETTRACERPARTICLESINCELL(i_cell)
    p_out ← F_net/m_cell   ▷ Probability to move part. out of cell
    for j_part in tracers do          ▷ Loop on tracer particles
        r_1 ← DRAWUNIFORM(0, 1)
        if r_1 < p_out then
15:         r_2 ← DRAWUNIFORM(0, 1)
            for i_dir ← 1, 2N_dim do         ▷ Select a direction
                F ← GETFLUXINDIR(i_cell, i_dir)
                p = F/F_net
                if r_2 < p then       ▷ Move in direction i_dir
20:                 MOVEPARTICLE(i_cell, j_part, i_dir)
                    break
                else
                    r_2 ← r_2 − p
                end if
25:         end for
        end if
    end for
end function
```

This function requires the MOVEPARTICLE function, which is defined as follow

```
function MOVEPARTICLE(i_cell, i_part, i_dir)
    F_tot ← GETFLUXINDIR(i_cell, i_dir)
    neighbors ← GETCELLSONFACE(i_cell, i_dir)
    ī_dir ← GETOPPOSITEDIRECTION(i_dir)
5:  r ← DRAWUNIFORM(0, 1)
    for j_cell in neighbors do
        F ← − GETFLUXINDIR(j_cell, ī_dir)
        p ← F/F_tot
        if r < p then▷ Move particle to the centre of the cell
10:         SETPARTICLEATCENTER(i_part, j_cell)
            break
        else                        ▷ Proceed to next cell
            r ← r − p
        end if
15: end for
end function
```
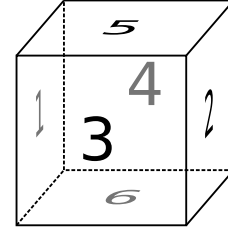

Fig. A.1. Cell faces numbering.

GETFLUXINDIR returns the mass that goes through the cell face in one timestep. Assuming that cell faces are numbered from 1 to 6 (left, right, top, bottom, front, rear, see Fig. A.1), GETOPPOSITEDIRECTION reads

```
function GETOPPOSITEDIRECTION(i_dir)
    mask ← [2, 1, 4, 3, 6, 5]
    return mask[i_dir]
end function
```

When looped over all cells, the algorithm treating all the tracers has complexity $O(N)$ where N is the total number of tracer particles and requires $O(N_{dim}N_{cell})$ memory to store the fluxes and $O(N)$ to store the tracer particles information.

A.2. AGN

Here we present how the volume of the jet is computed. We also present how the positions of the tracer particles in the jet are drawn. The function in charge of drawing position for the tracer particles in the jet is TRACER2JET

```
function TRACER2JET(j)
    loop
        c ← 2
        while c > 1 do
5:          a ← NORMALDISTRIBUTION(0, 1)
            b ← NORMALDISTRIBUTION(0, 1)
            c ← a² + b²
        end while
        x ← r_AGN × a
10:     y ← r_AGN × b
        h ← UNIFORM(−2r_AGN, 2r_AGN)
        r² ← x² + y²
        if |h| > r_AGN and (|h| − r_AGN)² + r² < r²_AGN then
            break
15:     else if |h| ≤ r_AGN then
            break
        end if
    end loop
            ▷ We now have a position in the frame of the jet.
20: u_z ← j/|j|
    u_x ← [j_y + j_z, −j_x + j_z, −j_x − j_y]
    u_x ← u_x/|u_x|
    u_y ← u_z ∧ u_x
    return x u_x + y u_y + h u_z
25: end function
```

5.3 Torque extraction

Most of the previous works (Danovich et al., 2015; Prieto et al., 2017) have studied the relative contribution of each torques to the angular momentum (AM) evolution of the cold gas focusing in particular on their magnitude, splitting the torques between the pressure and the gravitational torques. This section provides an improvement over these past works by computing the gravitational torques from each source (stars, DM and the gas) separately. I also lay down a general method to compute gradients in post-processing in AMR codes, which I then use to compute pressure gradient, and in particular, pressure torques. The precise computation of pressure gradients and the component-by-component decomposition of the gravitational torques will prove important in the context of the study of the evolution of cold flows in galaxies, as will be discussed in more depth in chapter 6, but let me first present here how one can rigorously compute them.

Section 5.3.1 details the method I developed to compute the gravitational torques from the different sources in numerical simulations. I show that I am able to extract the torques originating from the DM, stars, the gas. The method is then checked and I show that it provides percent-accurate results. Section 5.3.2 details the method I developed to compute the pressure torques in post-processing. The method can be used in general to compute any spatial derivative on an AMR grid that can be computed from the values of the direct neighbours, which includes the gradient, the divergence, the curl and the Laplacian operators.

5.3.1 Gravitational torques

In the vicinity of galaxies, the different massive sources (DM, stars, gas) all contribute to the total gravitational potential $\phi = \phi_{\mathrm{DM}} + \phi_\star + \phi_{\mathrm{gas}}$ *via* the Poisson equation

$$\nabla^2 \phi_i = 4\pi G \rho_i, \tag{5.2}$$

where ϕ_i and ρ_i are the gravitational potential and the density of the component i (DM, stars, gas). One can then compute the specific forces resulting from each potential $\boldsymbol{F}_i = -\boldsymbol{\nabla}\phi_i$ which can then be used to compute the specific torques at position $\boldsymbol{r}$

$$\boldsymbol{\tau}_i \equiv \boldsymbol{r} \times \boldsymbol{F}_i. \tag{5.3}$$

In order to extract the torques resulting from each gravitational source, I have modified the code RAMSES to extract in post-processing the specific forces due to the different matter components (DM, gas, stars). This was performed by stripping down RAMSES to keep only the Poisson solver, applied to the density of each individual component[1]. Since the resulting code is a simplified version of RAMSES, it can be run with exactly the same parameters as the original run, so that the results yielded are consistent (for example, the cosmology is the same). Using the numerical simulation detailed in section 5.2, I have computed the gravitational force of the stars, gas and dark matter that act on the gas for each output. For each component (star, gas and DM), I have also computed the rate of change of specific angular momentum (sAM) of the gas as

$$f_i = \frac{\boldsymbol{\tau}_i \cdot \boldsymbol{l}}{\|\boldsymbol{l}\|^2}, \tag{5.4}$$

where $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{v}$ is the gas sAM, both positions and velocities are evaluated for the gas in the frame of the central halo. Note that equation (5.4) yields a quantity that can be interpreted as the number of time torques are able to remove all the AM per unit time (it is a frequency). Equation (5.4) is therefore a measure of the inverse e-folding time along the Lagrangian trajectory of a particle. f_i is positive and large where torques are efficient at increasing the sAM and negative where

---

[1] The fiducial implementation solves the Poisson equation directly on the total matter density (gas + stars + DM).

(a) Error on evolution rate computation

(b) Error on acceleration computation
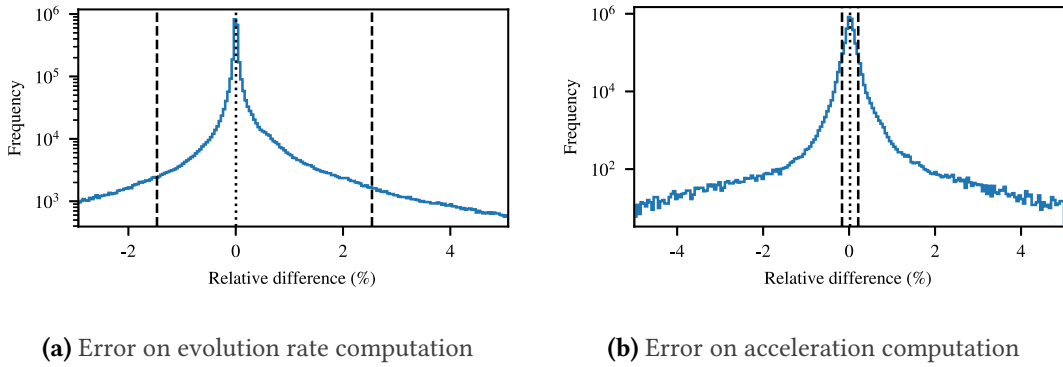
Figure 5.3.1: (a): Relative difference between the sum of the sAM evolution rate due to stars, DM and gas gravitational forces (as computed with the method presented in the text) and the rate due the total gravitational torques (as computed by RAMSES). (b): Same, but with the total gravitational accelerations. Vertical dashed line indicate 5 % and 95 % quantiles. The vertical dotted line indicates the median value. The two methods yield similar results within a few percents.

torques are efficient at decreasing the sAM. Figure 5.3.1a shows the relative difference between the sum of the evolution rates $f_\star + f_{DM} + f_{gas}$ extracted individually in post-processing and the total evolution rate $f = \boldsymbol{\tau} \cdot \boldsymbol{l}/\|\boldsymbol{l}\|^2$ computed on-the-fly by RAMSES. Figure 5.3.1b shows the relative difference between the gravitational accelerations computed using the two methods. The agreement is of the order of less than a percent in 90% of the cells. Note that a perfect agreement is not expected, as the potential from the SMBHs has been neglected in the post-processing method. In addition, RAMSES' Poisson solver has an intrinsic accuracy of 10^{-4}, consistent with the median error obtained in the gravitational accelerations ($0.02\,\%$). Overall, the agreement between the computed rates are within a few percents. The errors on the evolution rate are slightly larger, albeit still small, as a result of the division by l that skews the distribution and spreads assigns larger weights in regions where $\boldsymbol{l}$ is small. This confirms that the post-processing decomposition yields results consistent with the on-the-fly-computed gravitational field used internally to evolve the simulation.

5.3.2 Pressure gradient estimation

The precise capture of shocks is fundamental to most of the astrophysical codes. These shocks then result in strong, short-wavelengths gradients which are usually captured by a few cells in most AMR codes. While numerical codes routinely deal with strong gradients, most AMR post-processing tools either do not provide any utility to compute them (pynbody, Pontzen et al., 2013, pymses Guillet et al., 2013), or have gradient computing capacities that are not available for octtree-based AMR datasets, as is the case with RAMSES, (e.g. yt Turk et al., 2011). The approach usually followed is to project data on a fixed resolution grid, which is then used to compute gradients using a finite-difference scheme. Even though this approach yields sensible results at scales comparable to the (arbitrary) grid spacing, any information at finer scales is smoothed out while values at coarser levels have to be interpolated, eventually leading to spurious gradients. In the case of the study of accretion onto galaxies, the fixed-grid approach fails to provide a precise description of the gradients at play (pressure and potential gradients), as shocks may form anywhere in a large volume $\sim R_{\rm vir}^3$. In order to capture all shocks on a regular grid, one would then require the grid to be at least as well-resolved as the AMR grid, effectively requiring

$\sim (R_{\rm vir}/\Delta x)^3 \approx (100\,{\rm kpc}/30\,{\rm pc})^3 \approx 3 \times 10^{10}$ cells.[2] This is in practice too large to fit in memory, as it would require about $120\,{\rm Gio}$ of data as each pixel requires $4\,{\rm o}$ of data. In practice, it is much more efficient *and* consistent to directly work on the AMR structure dumped alongside the physical information of the simulation.

Using a tree search algorithm, as illustrated on figure 5.3.3, I have developed a post-processing tool that is able to compute finite difference gradients directly on the AMR grid. The binary search algorithm ensures that any given location is found in at most N steps, where N is the number of AMR levels in the simulation (typically between 10 and 20). To do so, I have extended the yt code (Turk et al., 2011) to enable computation of gradients for oct-based AMR datasets. The algorithm works as follows. (a) Loop over all octs in the tree. (b) Compute the positions of the $4^3 = 64$ virtual cells centred on the oct and extending in $\pm 2\Delta x$ in three directions, as illustrated on figure 5.3.2, left panel. (c) Get the value of interest at the centre of each virtual cell from the AMR grid. If the virtual cell exists on the grid or is contained in a coarser cell, the value on the grid is directly used. If the virtual cell contains leaf cells, the mean of these cells is used.[3] (d) Compute the gradient of the quantity using a centred finite-difference scheme on the 4^3 grid, as illustrated on figure 5.3.2, right panel. (e) Store the value of the gradient in the central 2^3 cells.

This approach aims to provide results as close as possible to the values used internally by Ramses. It is worth noting that this approach is exactly consistent with the internal approach of Ramses, except at the interface between different grid levels where a linear interpolation is used by Ramses, whereas our method uses a simple average. One way to check the consistency is to compare gradients computed by the post-processing tool to the ones computed internally by Ramses. This is for example done using the velocity divergence, as shown on figure 5.3.4. The figure shows that the post-processing method recovers the velocity divergence within a few percent, while most of the scatter is attributed to the fact that Ramses uses a linear interpolation at the interface between coarse and fine cells.

---

[2]Here I have used the resolution of the simulation presented in section 5.2 and the typical size of the virial radius of a $10^{12}\,{\rm M}_\odot$ halo at $z = 2$.

[3]Note that to be fully consistent with Ramses at fine-to-coarse boundaries, one should either use a linear interpolation with a total variation diminishing scheme (TVD) or a straight injection and use $1.5\Delta x$ distance in the gradient estimate.
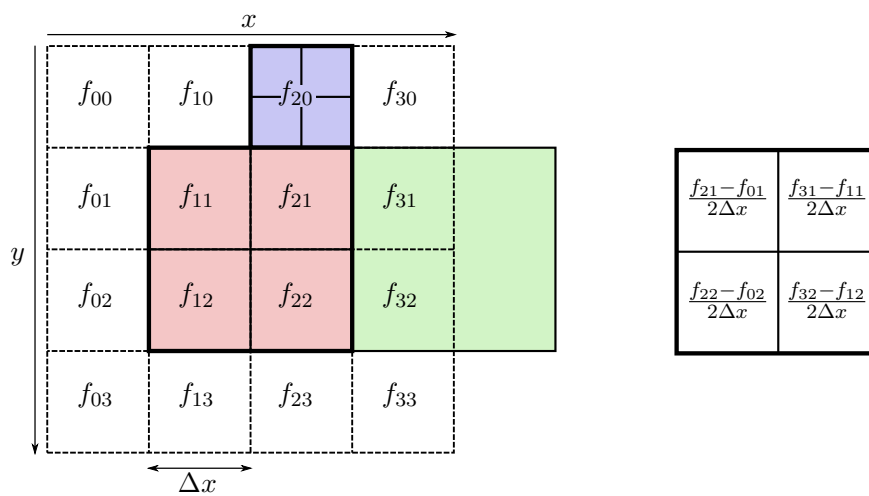
Figure 5.3.2: Scheme of the AMR structure used to estimate the gradient of a quantity f in the central oct (red). Octs are represented in thick lines, cells in thin lines and virtual cells in dashed lines. *Left panel:* The virtual cell values on a 4^3 grid are interpolated from the nearest cell in the AMR grid. If the nearest cell is at the same level, its value is directly used. If the cell is at a coarser level, its value is directly used (for example f_{31} and f_{32} have the value of the green cell). If the cell is refined, the mean of its children is used (for example f_{20} is the mean of all the *blue* cells). *Right panel:* Gradients are estimated using a first-order finite difference centred scheme on the 4^3 virtual cells.

(a) Binary search in 1D.

(b) Binary search in 2D.
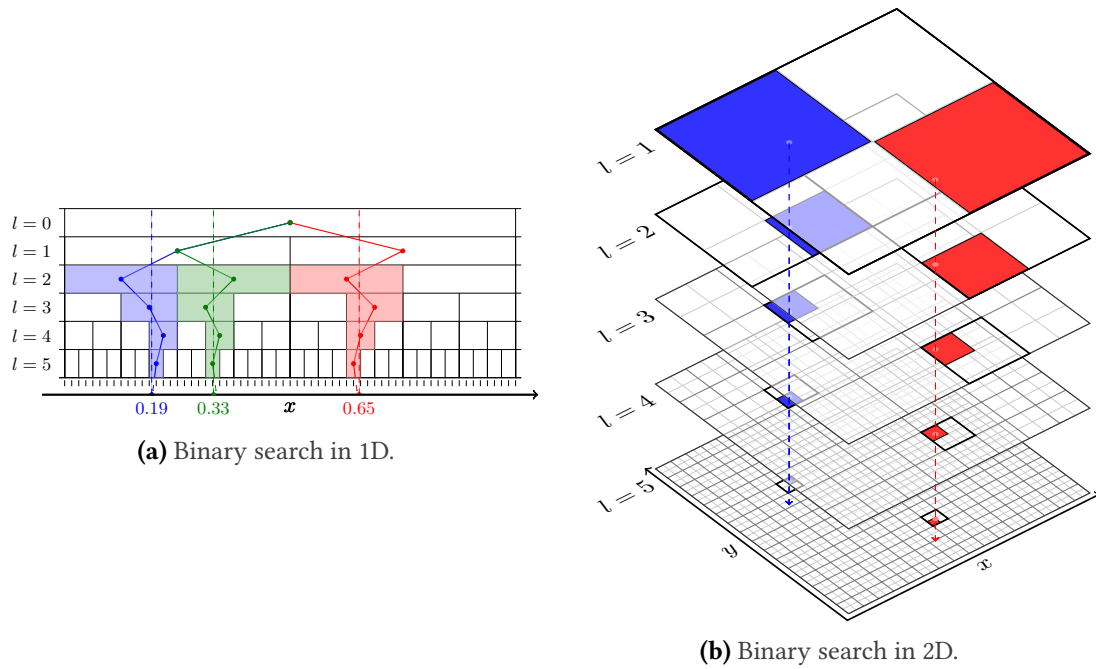
Figure 5.3.3: (a) Scheme of a binary search in an oct structure in 1D. The requested points are shown as red and blue dashed lines. The algorithm starts at the root level $l = 0$ and goes down the structure ; at each level, it picks the cell that contains the requested point. (b) A similar illustration in 2D, the algorithm works in the same way. At each level, it selects one of the four cells (red and blue squares) from the oct (thick line). The algorithm can be easily generalised to three or more dimensions. It is able to find any cell containing a given point in l_{max} iterations exactly. If the grid is sparse, as is the case for an AMR structure, l_{max} becomes an upper boundary.
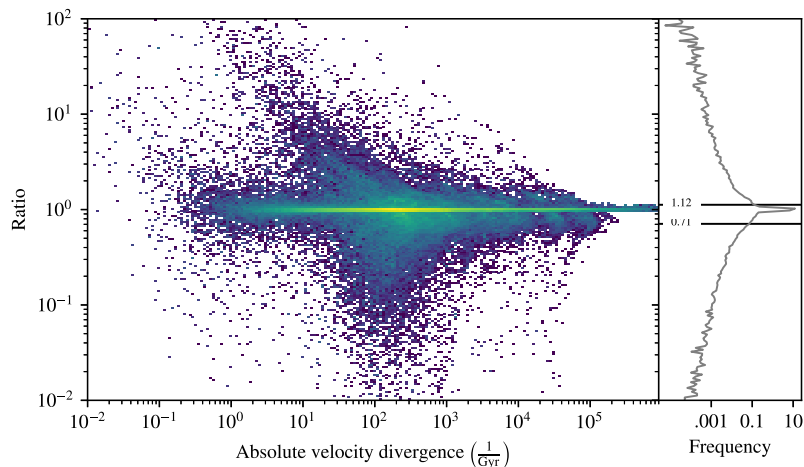
Figure 5.3.4: *Left panel:* Plot of the velocity divergence as computed by RAMSES vs. the ratio of the value computed in post-processing to RAMSES's one. *Right panel:* PDF of the ratio. 95 % of the distribution falls between the two horizontal lines. 95 % of the cells have a value between 0.71 and 1.12 times the value computed internally by RAMSES.

5.4 Conclusion

In section 5.2, I have presented a new tracer particle scheme. I have shown that it is able to accurately capture the Lagrangian evolution of the baryons in a full-featured cosmological simulation. In particular, I showed that the tracer particles have a spatial distribution close to that of the gas, which significantly improves over previous methods. I also showed that the tracer particles can be used to get the full Lagrangian evolution of the baryons as they move from gas to stars and into SMBHs. In section 5.3, I presented two numerical methods to extract the torques acting on a parcel of fluid. The methods presented have been shown to yield consistent results that are much improved compared to previous methods. The gravitational torque extraction method can be used to decompose the contribution to the gravitational torques due to each individual components (DM, stars and gas). I have also presented a method to compute pressure gradients on an AMR grid which I then applied to the computation of pressure torques. This approach provides results consistent with the internal values of the code.

The methods detailed in this chapter have a broad range of application. The tracer particle scheme has already been adopted by other researchers to study the formation and destruction of clumps in clusters (see appendix B.3) and to study gas flows in the circumgalactic medium (P. Mitchell, private communications). They are also used in the context of protoplanetary disk formation, where they have been modified to follow the Lagrangian evolution of dust grains (U. Lebreuilly, private communications). The gradient computation was for example used to compare the cosmic ray pressure gradient, thermal pressure gradient and gravitational acceleration in isolated simulations of dwarf galaxies, and concluded that cosmic ray pressure was the dominant acceleration mechanism above $1\,\mathrm{kpc}$ from the plane of the disc (G. Dashyan at al., submitted).

6. Angular momentum acquisition from the cosmic web

Maxime Trebitsch — Obelisk collaboration

Outline

6.1 Introduction

Figure 6.1.1: Sketch of the galaxy and its environment with the outer halo ($r > R_{\rm vir}/3$), the inner halo ($R_{\rm vir}/3 > r > R_{\rm vir}/10$) and the disk ($r < R_{\rm vir}/10$). The mode of accretion are the cold mode *via* cold flows (in blue) and the hot mode, characterized by a stable shock at the virial radius (in red). Cold flows may also shock and heat at a smaller radius to be determined. On large scales, cold flows are embedded in the cosmic web.

One of the successes of the ΛCDM model is its ability to reproduce the large-scale structure of the Universe observed in galaxy distribution (e.g. Springel et al., 2006). These structure form out of the initial tiny density fluctuations of the primordial density field and under the effect of gravitational forces, matter departs from underdense regions to flow through cosmic sheets into filamentary structures. Matter then flows from these filaments towards high-density peaks that will later become halos. In the process, matter acquires kinetic properties (e.g. vorticity Pichon and Bernardeau, 1999; Laigle et al., 2015) in its journey through voids, sheets and filaments of the cosmic web, which, in turn, affect the assembly of dark matter halos, as shown in chapter 4. Before shell crossing, baryons follow the same initial fate as DM and flow from underdense regions to sheets. Yet, as they flow in sheets, pressure forces prevents them from shell-crossing so that they lose their normal velocity component to the shock front, dissipating large-scale acquired kinetic energy into internal energy (eventually radiated away by gas cooling processes). Following potential wells created by dark matter, baryons then flow towards filamentary structures where they lose a second component of their velocity[1] and reach a dense-enough state to efficiently cool radiatively.

At first order, galaxy formation is affected by the mass of their dark matter halo host and the local environment, as encoded by the local density on sub-Mpc scales, as it is assumed that baryons have the same past accretion history as dark matter. These models have proven successful at explaining a number of observed trends, in particular against isotropic statistics, in the so-called halo model, yet they fail at explaining some effects such as spin alignments (Tempel and Libeskind,

---

[1]The component lost is in the direction perpendicular to the shock, which is in the plane of the wall and the filament.

2013; Codis et al., 2015; Dubois et al., 2014; Chisari et al., 2017), colour segregation (Laigle et al., 2018; Kraljic et al., 2018; Kraljic et al., 2019) or star formation rates (Malavasi et al., 2017; Kraljic et al., 2019). Indeed, galaxies form by converting their gas into stars and by successive mergers, which are in turn affected by the tides and large-scale modulations of the density field induced by the cosmic web. The detailed history of how the gas was acquired and how much AM it brought, as well as the origin of the mergers should in principle impact the formation of the galaxy. Since the physical processes involved in dark matter halo formation differ from the baryonic processes at the core of galaxy formation, one can expect that the cosmic web will have a different impact, if any, on the formation of galaxies and may explain the disparity of their properties in similar-looking dark matter halos.

In particular, at fixed halo mass and local density, properties of galaxies such as their colour or the kinematic structure vary with their location in the cosmic web. One key process in the differential evolution of galaxies is gas accretion. Indeed, at large redshifts it has been suggested that the accretion of gas is dominated by flows of cold gas funnelled from the large scales to galactic scales (Birnboim and Dekel, 2003; Dekel and Birnboim, 2006). This mode of accretion has then been confirmed in numerical simulations using different methods (Kereš et al., 2005; Dekel and Birnboim, 2006; Ocvirk et al., 2008; Nelson et al., 2013) as the source of a significant fraction of the baryonic mass but also AM (Pichon et al., 2011; Kimm et al., 2011; Stewart et al., 2013; Stewart et al., 2017) and it has been proposed that these flows may feed supermassive black holes (Di Matteo et al., 2012; Dubois et al., 2012), which in turn affect the cold inflow rates (Dubois et al., 2013). Using an extension of TTT (Peebles, 1969; Schaefer, 2009), Codis et al., 2015 showed that anisotropic environments, such as large-scale filamentary structures, biases the AM distribution to align it with the cosmic web. It is then expected that this gas will fall in galaxies *via* cold flows, feeding disks with angular-momentum rich gas that is itself aligned with the tides of the cosmic web.

Recent works have shown that the flows are subject to a variety of processes: they may fragment (Cornuault et al., 2018) or be disrupted by hydrodynamical instabilities (Mandelker et al., 2016; Mandelker et al., 2019), but they are also sensible to feedback events (Dubois et al., 2013). In this context, Danovich et al., 2015 showed that in numerical simulations, cold flows are nevertheless able to feed galaxies with angular-momentum rich material (as speculated by Pichon et al., 2011; Stewart et al., 2013). In this study, it was shown that the AM acquired outside the halo at $z = 2$ is transported down to the inner halo; the gas then settles in a ring surrounding the disk, where gravitational torques spin the gas down to the mean spin of the baryons. Another study, albeit at larger redshifts, found that the dominant force was pressure (Prieto et al., 2017). Since there is not much freedom on the final AM of the galaxies, as constrained by their radius, the excess AM brought by cold flows has to be redistributed somehow before it reaches the disk. The details of where this AM will end up are key to understand the AM distribution in galaxies, but also to understand to what extent their spin is aligned with the cosmic web. If the dominant forces acting on the AM are pressure forces, resulting from internal processes (SN winds, AGN feedback bubbles), then the spin of the galaxy would likely be a result of chaotic internal processes and would lose its connection to the cosmic web. Similarly, if the AM is lost into thermal energy (which is then radiated away) in shocks, the galactic spin would be a weak function of the large-scale AM induced by the cosmic web. On the contrary, if the dominant forces are gravitational forces, then the spin-down of the cold gas is likely to drive a spin-up of either the disk or the dark matter halo, which themselves are the result of their past AM accretion history. In this last scenario, the details of which part(s) of the halo or the disk interact exchange AM with the infalling material would constrain models aimed to understand the evolution of the spin of galaxies.

Historically, the study of cold accretion has been particularly challenging in numerical simulations. Early simulations using SPH methods largely over-estimated the fraction of gas accreted

cold (see e.g. Nelson et al., 2013, for a discussion on this particular issue) as a result of the difficulty to capture shocks using SPH. AMR simulations do not suffer from this caveat (Ocvirk et al., 2008), yet they fail at providing the Lagrangian history of the gas — in particular its past temperature — which is required to detect the cold-accreted gas. In order to circumvent this limitation, most simulations relied on velocity-advected tracer particles (Dubois et al., 2013; Tillson et al., 2015). However, this approach yields a very biased tracer distribution that fails at reproducing correctly the spatial distribution of gas in filaments: most tracer particles end up in convergent regions (centre of galaxies, centre of filaments) while divergent regions are under-sampled. In order to reproduce more accurately the gas distribution, Genel et al., 2013 suggested relying on a Monte-Carlo approach where tracer particle follow mass fluxes instead of being advected. Using this approach, I have showed in chapter 5 (Cadiou et al., 2019) that tracer particles are able to faithfully reproduce the gas distribution while providing the Lagrangian history of the gas, and in particular its past temperature and position.

In this chapter, I investigate the evolution of the AM of the cold and hot gas using cosmological simulations of group progenitors at $z > 2$. I provide a detailed study of the evolution of the AM of the cold and hot gas. In particular, this chapter aims at answering the question of which forces are responsible for the spin-down and realignment of the AM of the gas accreted in the two modes of accretion (hot and cold). Section 6.2 presents the numerical setup. Section 6.3 presents the AM evolution of the cold and hot gas. It follows the evolution of the magnitude and orientation of the AM and the different forces and torques at play in the different regions of the halos. It details the evolution of the magnitude and orientation of the AM and the different forces and torques at play in the different regions of the halos. Section 6.4, I present their implication on the distribution of AM in the galaxy and the inner halo. Finally, section 6.5 wraps things up and concludes.

In the following of this chapter, I will adopt the same naming conventions as Danovich et al., 2015. I will write $R_{\rm vir}$ the virial radius of a halo. The outer halo is defined as the region between $R_{\rm vir}$ and $R_{\rm vir}/3$. The inner halo is defined as the region between $R_{\rm vir}/3$ and $R_{\rm vir}/10$. The "disk" is the region at radius $r < R_{\rm vir}/10$ where the galaxy is found. This is sketched on figure 6.1.1.

6.2 Methods

In section 6.2.1, I establish the equations that link the AM evolution of the gas to the different torques. In section 6.2.2, I describe the simulations I ran. In section 6.2.3, I describe how I selected the cold gas being accreted on the halos in the simulations.

6.2.1 Equations

In the following, the position and velocities are computed in the frame of halo, as measured with the ADAPTAHOP halo finder (Tweed et al., 2009). Let me first derive the equation driving the evolution of the sAM of the gas,

$$\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{v}. \tag{6.1}$$

To do so, let us start from Euler's equation and the mass conservation equation

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) = 0, \tag{6.2}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \boldsymbol{\nabla})\boldsymbol{v} = -\frac{\boldsymbol{\nabla} P}{\rho} - \boldsymbol{\nabla}\phi. \tag{6.3}$$

Taking the derivative of equation (6.1) w.r.t. time, one gets that

$$\frac{\mathrm{d}\boldsymbol{l}}{\mathrm{d}t} = \boldsymbol{r} \times \left(\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \boldsymbol{\nabla})\boldsymbol{v}\right) + \left(\frac{\partial \boldsymbol{r}}{\partial t} + (\boldsymbol{v} \cdot \boldsymbol{\nabla})\boldsymbol{r}\right) \times \boldsymbol{v}. \tag{6.4}$$

After trivial algebra, the rightmost part of the right-hand side vanishes. Using equations (6.3) and (6.4), the Lagrangian time derivative of the sAM then reads

$$\frac{\mathrm{d}\boldsymbol{l}}{\mathrm{d}t} = \boldsymbol{\tau}_P + \boldsymbol{\tau}_\phi, \tag{6.5}$$

where $\boldsymbol{\tau}_P \equiv -\boldsymbol{r} \times \boldsymbol{\nabla}P/\rho$, $\boldsymbol{\tau}_\phi = -\boldsymbol{r} \times \boldsymbol{\nabla}\phi$ are the specific pressure and gravitational torques. Here P and ρ are the pressure and density of the gas and ϕ is the gravitational potential. The potential is defined using Poisson equation

$$\nabla^2\phi = 4\pi G\rho_{\mathrm{tot}}, \tag{6.6}$$

where ρ_{tot} is the total matter density (DM, stars, gas and SMBHs). Using the linearity of equation (6.6), the total potential can be written as the sum of the potential due to each component $\phi = \phi_{\mathrm{DM}} + \phi_\star + \phi_{\mathrm{gas}}$[2]. One can similarly decompose the gravitational torques into three different components $\boldsymbol{\tau}_\phi = \boldsymbol{\tau}_{\phi,\mathrm{DM}} + \boldsymbol{\tau}_{\phi,\star} + \boldsymbol{\tau}_{\phi,\mathrm{gas}}$. In the following, I will use this decomposition to assess which gravitational components contributes to the evolution of the sAM of the gas.

Following section 5.3, let me define the rate of change induced by each torque

$$f_i = \frac{\boldsymbol{\tau}_i \cdot \boldsymbol{l}}{\|\boldsymbol{l}\|^2}, \tag{6.7}$$

where i indicates the torque source (DM, star or gas gravitational torques or pressure torques). I have shown in the previous chapter that this quantity has the physical meaning of the inverse e-folding time of the sAM on a Lagrangian trajectory. Using equation (6.5) and after some algebra, one also gets that the total rate of change can be simply expressed as the variation of the sAM $f = \mathrm{d}\log l/\mathrm{d}t$. Let me emphasize here that this relation only holds for the total rate of change. From equation (6.7), one can also define a typical timescale associated with the torques as

$$t_{\tau,i} = \left|\frac{1}{f_i}\right|. \tag{6.8}$$

These timescales measure the typical time over which a given torque will significantly change the sAM of the gas. As equations (6.7) and (6.8) only use the contribution of the torque in the direction of the sAM, the timescale measures the typical time required to loose all the sAM, but not to reorient it. Note that one could compute similarly the realignment timescale by replacing the dot product of equation (6.7) with the norm of the cross product $f_{\perp,i} \propto |\boldsymbol{\tau}_i \times \boldsymbol{l}|$. In addition, the timescale and the rate of change are scalars that can have a negative value if the torques are spinning the gas down (anti-aligned with the sAM vector) while it has a positive value if they are spinning the gas up (aligned with the sAM vector).

6.2.2 Numerical simulation

I have run a suite of three $50\,\mathrm{cMpc/h}$-wide cosmological simulations, hereafter named S1, S2, S3. The three simulations contain 6 halos with $M \gtrsim 5 \times 10^{11}\,\mathrm{M}_\odot$[3], hereafter named A, B, C, D, E and F. Their properties are presented in Table 6.1. The size of the zoomed Lagrangian volume in the initial conditions is chosen to encapsulate twice the virial radius of the halo at $z = 2$. The simulation are started with a coarse grid of 128^3 (level 7) and several nested grids with increasing levels of refinement up to level 11. The adopted cosmology has a total matter density of $\Omega_m = 0.3089$, a dark energy density of $\Omega_\Lambda = 0.6911$, a baryonic mass density of $\Omega_\mathrm{b} = 0.0486$, a Hubble constant of $H_0 = 67.74\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, a variance at $8\,\mathrm{Mpc}$ $\sigma_8 = 0.8159$, and a non-linear power spectrum index of $n_s = 0.9667$, compatible with a Planck 2015 cosmology (Planck Collaboration, 2015).

---

[2]Here I neglect the contribution from SMBHs as it is negligible on galactic scales.
[3]Only pure halos in the zoomed-region have been selected.

Table 6.1: Properties of the halos at $z = 2$.

Name	Simulation	$M_{\mathrm{vir}}/10^{11}\,\mathrm{M}_\odot$	$M_\star/10^{10}\,\mathrm{M}_\odot$
A	S1	3.66	6.07
B	S2	7.82	9.20
C	S3	6.64	5.09
D	S1	7.29	4.18
E	S1	5.23	7.84
F	S3	4.63	3.49

The simulations include a metal-dependant tabulated gas cooling function following Sutherland and Dopita, 1993 allowing gas to cool down to $T \sim 10^4\,\mathrm{K}$ *via* Bremsstrahlung radiation (effective until $T \sim 10^6\,\mathrm{K}$), *via* collisional and ionisation excitation followed by recombination (dominant for $10^4\,\mathrm{K} \leq T \leq 10^6\,\mathrm{K}$) and *via* Compton cooling. The metallicity of the gas in the simulation is initialised to $Z_0 = 10^{-3}\,\mathrm{Z}_\odot$ to allow further cooling below $10^4\,\mathrm{K}$ down to $T_{\min} = 10\,\mathrm{K}$ (Rosen and Bregman, 1995). Reionisation occurs at $z = 8.5$ using the Haardt and Madau, 1996 model and gas self-shielding above $10^{-2}\,\mathrm{m_p\,cm^{-3}}$. Star formation is allowed above a gas number density of $n_0 = 10\,\mathrm{m_p\,cm^{-3}}$ and with efficiency ϵ_{ff} that depends on the gravoturbulent properties of the gas (for details, see Kimm et al., 2017; Trebitsch et al., 2017). The main distinction of this turbulent star-formation recipe with the traditional star formation in RAMSES (Rasera and Teyssier, 2006) is that the efficiency can approach and even exceed $100\,\%$ (with $\epsilon_{\mathrm{ff}} > 1$ meaning that stars are formed faster than in a free-fall time). The stellar population is sampled with a Kroupa, 2001 initial mass function, where $\eta_{\mathrm{SN}} = 0.317$ and the yield (in terms of mass fraction released into metals) is 0.05. The stellar feedback model is the mechanical feedback model of Kimm et al., 2015 with a boost in momentum due to early UV pre-heating of the gas following Geen et al., 2015. The simulation also tracks the formation of SMBHs and the evolution of AGN feedback in jet mode (radio mode) and thermal mode (quasar mode) using the model of Dubois et al., 2012. The jet is modelled self-consistently by following the AM of the accreted material and the spin of the black hole (Dubois et al., 2014). The radiative efficiency and spin-up rate of the SMBH is then computed using the MAD results of McKinney et al., 2012. SMBHs are created with a seed mass of $10^4\,\mathrm{M}_\odot$ for S1 and $10^5\,\mathrm{M}_\odot$ for S2 and S3.

The simulations have a roughly constant physical resolution of $35\,\mathrm{pc}$ (one additional maximum level of refinement at expansion factor 0.1 and 0.2), a star particle mass resolution of $m_{\star,\mathrm{res}} = 1.1 \times 10^4\,\mathrm{M}_\odot$, a dark matter (DM) particle mass resolution of $m_{\mathrm{DM,res}} = 1.5 \times 10^6\,\mathrm{M}_\odot$, and gas mass resolution of $2.2 \times 10^5\,\mathrm{M}_\odot$ in the refined region. A cell is refined according to a quasi-Lagrangian criterion: if $\rho_{\mathrm{DM}} + \rho_{\mathrm{b}}/f_{\mathrm{b/DM}} > 8 m_{\mathrm{DM,res}}/\Delta x^3$, where ρ_{DM}, and ρ_{b} are respectively the DM and baryon density (including stars plus gas plus SMBHs), and where $f_{\mathrm{b/DM}}$ is the universal baryon-to-DM mass ratio. The max level of refinement is also enforced up to 4 minimum cell size distance around all SMBHs. Tracer particles (Cadiou et al., 2019) are added in the refined region with a fixed mass of $m_{\mathrm{t}} = 2.0 \times 10^4\,\mathrm{M}_\odot$ ($N_{\mathrm{tot}} \approx 1.3 \times 10^8$ particles) and are detailed in more depth in the next section. The description of the tracer particle scheme is detailed in chapter 5. There is on average 0.55 tracer per star and 22 per initial gas resolution element. Cells of size $35\,\mathrm{pc}$ and density $20\,\mathrm{cm^{-3}}$ contain on average one tracer per cell.

6.2.3 Cold gas selection

The ratio of the total accreted mass with a maximum temperature below a given threshold $T_{\max}$ to the total gas mass — the cold fraction — is a widely reported quantity in the study of the cosmological gas accretion, dating back to Kereš et al., 2005. In this study, a temperature cut
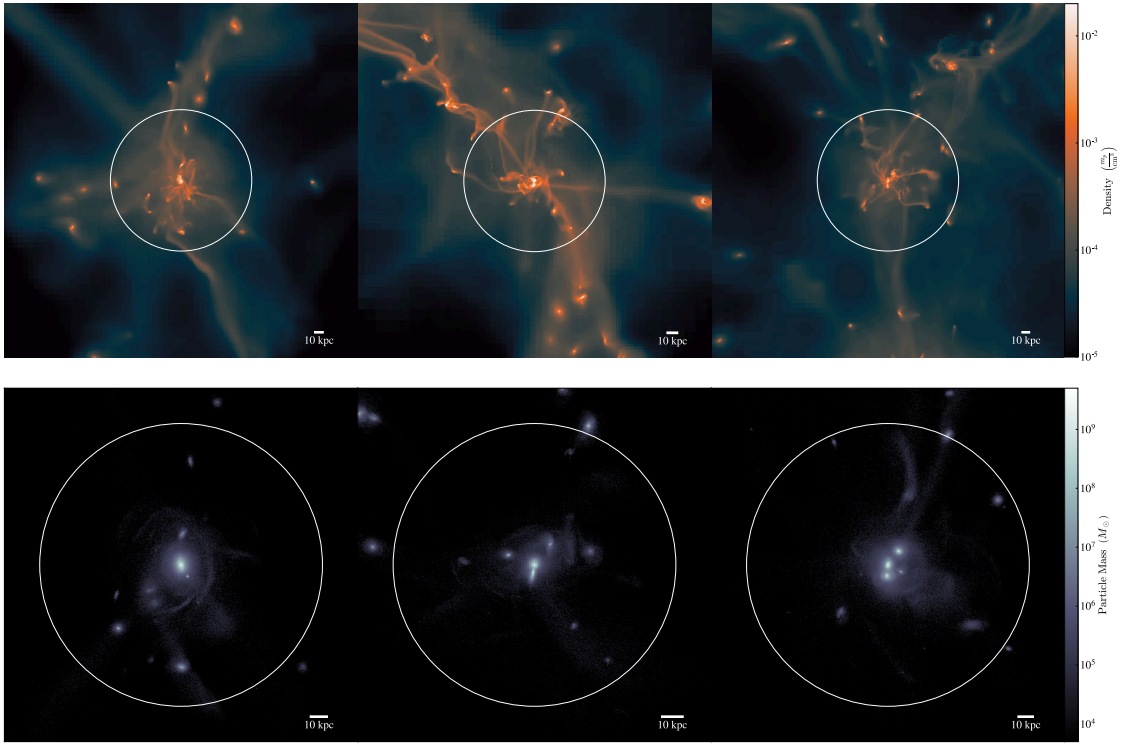
Figure 6.2.1: *Upper panel:* Projection of the gas density around the halos A (left), B (centre) and C (right) at $z = 2$. *Lower panel:* Line-of-sight integrated star density.

$T \lesssim T_{\mathrm{max}} = 2.5 \times 10^5 \,\mathrm{K}$ (see e.g. Nelson et al., 2013, for a discussion on the effect of the threshold) is used. In order to study the sAM evolution of the cold gas, I use the Lagrangian history of all the baryons (gas and star) that end up within $2R_{\mathrm{vir}}$ of the central galaxy. This ensemble of particle in the vicinity of the galaxy are then grouped in three sets.

1. the baryons that end up in the inner halo $r < 0.3R_{\mathrm{vir}}$ at the end of the simulation. I will refer to this subset as "baryons in the galaxy".
2. the baryons that never heated above the threshold temperature $T \leq T_{\mathrm{max}}$ from $1.5R_{\mathrm{vir}}$ to $0.3R_{\mathrm{vir}}$. I will refer to this subset as "cold baryons".
3. the baryons that were never accreted on a satellite galaxies. I will refer to this subset as "directly accreted baryons". This effectively selects gas whose first accretion is onto the main halo. In practice, this is done by excluding any tracer found at any time at less than a third of the virial radius of any halo other than the main one.

The distribution of the gas in halo A at $z = 2$ is shown on figure 6.2.2 where baryons in the galaxy are represented in the red ensemble, cold-accreted baryons in blue and directly accreted baryons in green. In the following of the chapter, the subset of interest is the intersection of the three ensembles: this is the gas that was accreted cold onto the galaxy, that end up in the inner halo at $z = 2$ and that was not accreted *via* mergers. In the remaining of the paper, I will refer to this subset as the "cold gas" while I will use "hot gas" to describe gas that was not accreted *via* mergers but which eventually heated up above the temperature threshold.

I have checked that the fraction presented on figure 6.2.2 are robust to changes of the threshold radius for first-accretion detection: using $R_{\mathrm{thresh}} = 0.5R_{\mathrm{vir}}$ instead of $0.3R_{\mathrm{vir}}$ only leads to percent differences. Indeed, most of the gas already within $0.5R_{\mathrm{vir}}$ of a halo is likely to later fall into the inner part of the galaxy.
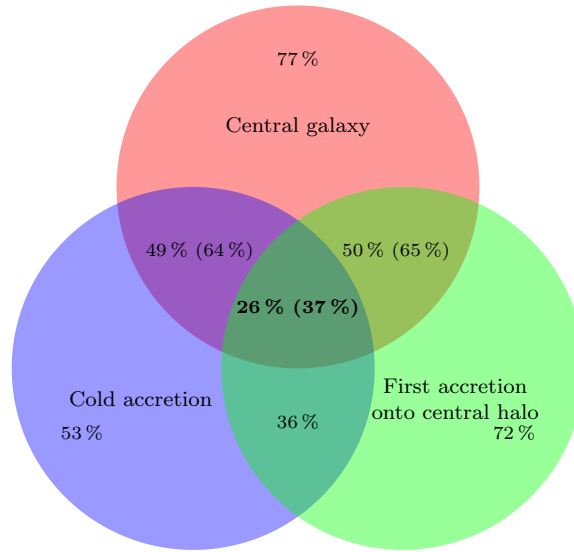
Figure 6.2.2: Venn diagram of the ensembles of tracer particles used to define the cold-accreted tracer particles. Direct cold-accreted tracer particles are the intersection of the tracer particles accreted cold between 1.5 and $0.5R_{\mathrm{vir}}$ (blue) that end up in the central galaxy at $z = 2$ (red) and that were first accreted onto the central halo (green). See the text for details on how each of these ensembles are defined. Percentages indicate the fraction in simulation A of all the particles within $2R_{\mathrm{vir}}$ found in each part of the diagram. Percentages within parenthesis indicate the fraction of tracer in the inner halo ($r < 0.3R_{\mathrm{vir}}$) found in each part of the diagram. Direct cold-accreted baryons represent 26% of the baryons that end up within $2R_{\mathrm{vir}}$ and 37% of the baryons within $0.3R_{\mathrm{vir}}$.

6.3   Results

I have now described the equations driving the AM evolution of the gas (section 6.2.1), the numerical setup (section 6.2.2) and described how the cold gas is selected (section 6.2.3). In this section, I detail the results obtained. In section 6.3.1, I present the differences between the sAM and the AM per unit volume, as used in Danovich et al., 2012. In section 6.3.2, I detail the dominant forces found in the different regions surrounding the galaxy. In sections 6.3.3 and 6.3.4, I describe the evolution of the AM magnitude and orientation respectively. Finally, in section 6.3.5, I describe which torques dominate the evolution of the AM of the gas.

6.3.1   Specific angular momentum vs. angular momentum per unit volume

Equation (6.5) differs from Eq. 9 of Danovich et al., 2015. Indeed, it is an equation on the *sAM* instead of the AM per unit volume. The rate of change of AM per volume includes a dependence to the cell volume *via* the velocity divergence, which is itself highly sensible to the compression and decompression of the gas. This is particularly important in astrophysical flows that are highly compressible. Contrary to what Danovich et al., 2015 reported, I find that the divergence term dominates over the gravitational and pressure terms. Inflowing gas typically moves at $100\,\mathrm{km/s}$ with typical variation scales of a few kpc. An order of magnitude of the divergence is then $\approx\ 100\,\mathrm{km\,s^{-1}}/1\,\mathrm{kpc}\ \approx\ 100\,\mathrm{Gyr^{-1}}$, with larger values found in shocked and highly compressed regions. These values are comparable or larger than pressure and gravitational torques, highlighting their importance in the study of the evolution of the AM per unit volume.

In the following of the dissertation, I will use the sAM, its evolution being described by

equation (6.5). I will hence *not* consider the divergence term in the study, as it does not enter the equation of evolution of the sAM. In addition to ignoring this term, following the Lagrangian evolution of the sAM has the advantage of interfacing naturally with tracer particles. Indeed, Lagrangian tracer particles have a fixed mass, so that their sAM is linked to their AM *via* a constant factor (their mass).

6.3.2 Dominant forces in the cold and hot phase

The different accretion mode for the cold and the hot phase of the gas leads to a spatial segregation of the cold phase into thin collimated filamentary structures, as shown on figure 6.2.1. In addition, their thermodynamical properties differ: the cold phase is made of a quite homogenous gas, so that the internal pressure gradients are weak. As a result, strong pressure gradients are found at their interface, as shown by Danovich et al., 2015. On the contrary the hot gas is less homogenous, so that pressure forces may be locally dominant. Figure 6.3.1 presents projected maps of the magnitude of the gravitational forces and pressure forces around one halo at $z = 2.7$ for the hot gas (top panel) and the cold gas (bottom panel). In addition, figures 6.A.1a and 6.A.1b in annex presents similar maps in the three directions (x, y and z) for the pressure and DM gravitational forces. In the hot gas, the two dominant forces are qualitatively DM gravitational forces and pressure forces, with stellar gravitational forces being important only in the inner halo. In the cold phase, pressure forces are significantly smaller, while gravitational forces are mostly unimpaired. In the inner halo, a notable "pressure-ring" is clearly visible in the cold gas, as shown in the bottom right panel of figure 6.3.1.

In order to better disentangle the different contributions to the dynamical evolution of the gas, one needs to distinguish the radial component of the forces — that is responsible for the infall of the gas — and the ortho-radial component — that is mostly responsible for the sAM variation. This is shown on figure 6.3.2 that presents radial profiles of the two components of each (specific) forces (pressure forces, gravitational forces) in one of the simulated halos. In the disk, the dominant forces in the radial and ortho-radial directions are stellar gravitational forces due to the disk. The forces are mostly radial, with their ortho-radial component one order of magnitude smaller than the radial one. In the inner halo, stellar gravitational forces become less dominant. The radial acceleration become DM-dominated, while the ortho-radial component is dominated by pressure torques. This is in particular the case for the hot gas, where ortho-radial pressure forces are one order of magnitude larger than DM gravitational forces. In the cold phase, the ortho-radial acceleration is due to both the DM and the pressure forces. The ortho-radial acceleration stays pressure-dominated in the hot phase up to a few Virial radii. Interestingly, both components of the gravitational forces have similar magnitudes in the cold and hot phase. I also notice that in the outer halo, the magnitude of both components of the pressure forces are comparable, indicating that pressure forces do not have a preferred direction. Here, I report that the "pressure-ring" corresponds to the rise of the ortho-radial pressure forces, as shown clearly in figure 6.3.1, right panel. This probably marks the transition between the free-falling cold flows and the circumgalactic medium and may have a significant impact on the structure of the accreted gas, and in particular on the cold flows.

6.3.3 The magnitude of the angular momentum

Before turn-around, gas acquires AM *via* torque with the cosmic web as explained by TTT (Hoyle, 1949; Peebles, 1969; S. D. M. White, 1984; Catelan and Theuns, 1996, see section 2.1.6). At these scales, the torque magnitudes are proportional to the mean density of the gas and DM component. Indeed, when the gas is far from the halo, the density ratio sourcing the gravitational torques is given $\Omega_{\rm b}/\Omega_{\rm DM} \approx 0.19$. As a consequence, a similar ratio is expected on the torque ratio, as shown on figure 6.3.3, which presents the evolution of the torques acting on the cold gas accreted
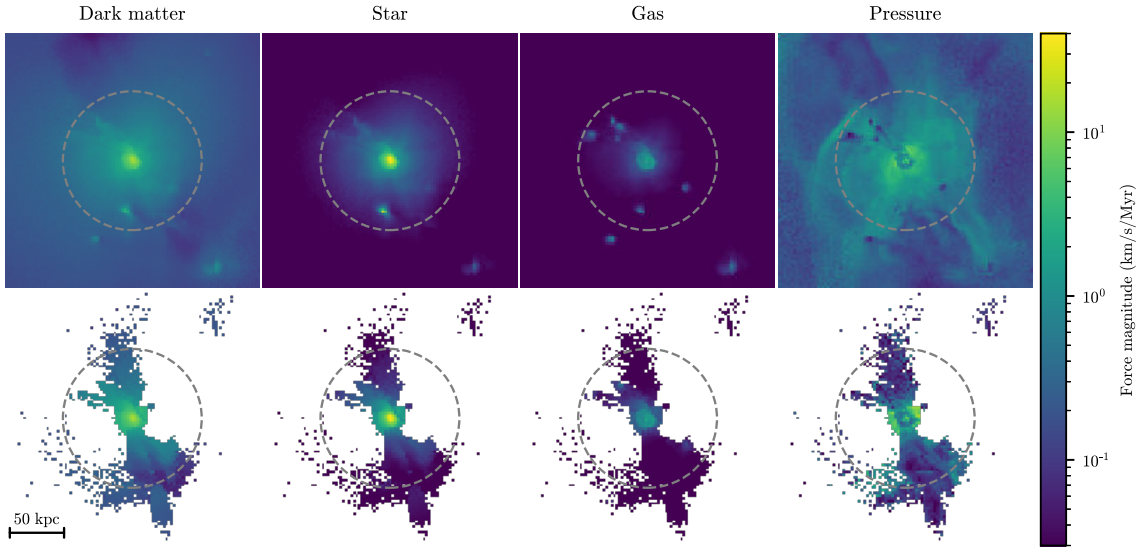
Figure 6.3.1: From left to right, mass-weighted projection of the magnitude of the DM gravitational forces, stellar gravitational forces, gas gravitational forces and gas pressure gradients, *top panel:* for all the gas but the cold one and *bottom panel:* only the cold gas in halo A at $z = 2.7$. In the hot phase outside the halo, gas pressure and DM gravitational forces have similar magnitudes. In the inner halo star and DM gravitational forces have a magnitude comparable to pressure forces. In the cold phase, the pressure forces are significantly weaker. The gas gravitational forces are negligible everywhere in both the cold and hot phases.
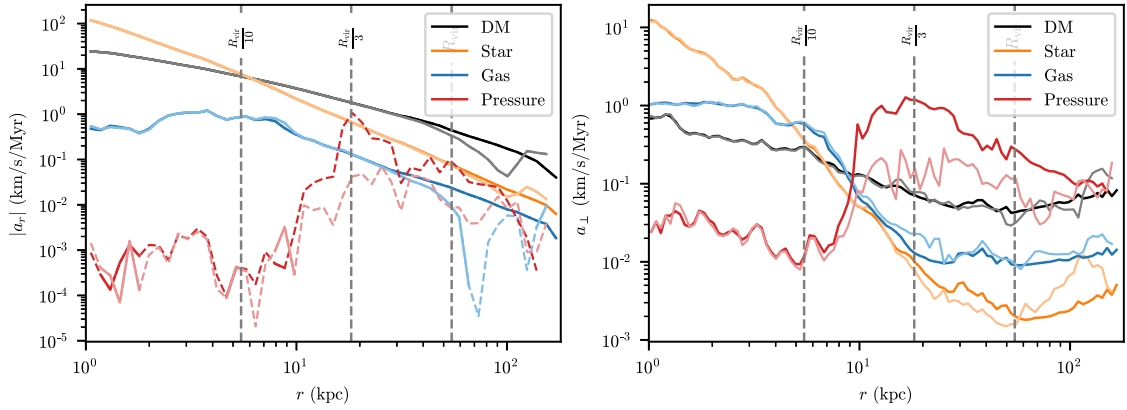
Figure 6.3.2: *Left:* Radial profile of the radial component and *right:* of the ortho-radial component of the different forces around halo A at $z = 2.7$: DM gravitational forces (black), stellar gravitational forces (orange), gas gravitational forces (blue) and pressure forces (red). Inward radial accelerations are shown as solid lines and outward accelerations are shown as dashed lines. Dark lines show the profiles for the hot gas and light lines for the cold gas. The virial radius $R_{\rm vir}$, $R_{\rm vir}/3$ and $R_{\rm vir}/10$ are shown as vertical dashed grey lines. Gravitational forces have a similar action on cold gas. The ortho-radial component of pressure forces is significantly smaller in the cold gas outside the inner halo.
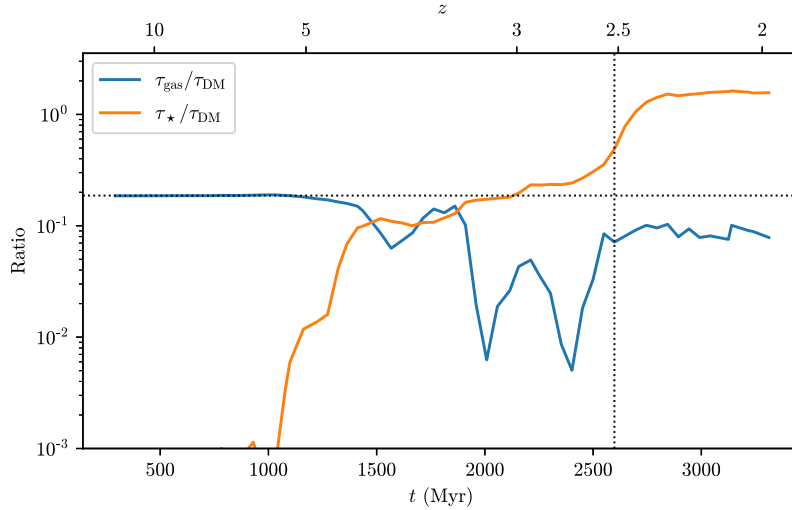
Figure 6.3.3: Evolution of the ratio of the gas gravitational torques to the DM gravitational torques (blue) and of the ratio of the stellar gravitational torques to the DM gravitational torques (orange) for gas crossing a $R_{\rm vir}/3$ at $z = 2.5$ in halo A. The ratio $\Omega_{\rm b}/\Omega_{\rm DM}$ (horizontal dotted line) corresponds to the initial gas-to-DM density ratio. Star torques become important in the inner halo $r \lesssim R_{\rm vir}/3$ (vertical dotted line).

in halo A at $z = 2.5$. As expected, the gas-to-DM torque ratio goes to $\Omega_{\rm b}/\Omega_{\rm DM}$ in the early times, when the gas is still far from the galaxy.

The sAM of the hot and cold gas follows a different path. In order to study how the sAM evolves, one can study the Lagrangian evolution of the sAM of all the gas accreted at the same time as a function of its radius, as shown on figure 6.3.4. The figure presents the Lagrangian evolution of the sAM as a function of radius for the cold (solid lines) and hot gas (dashed lines). In all halos, the sAM of the cold gas is conserved down to smaller radii, typically $r \sim R_{\rm vir}/3$ than in the hot gas.

For the hot gas, the virial shock is able to efficiently mix the pristine, freshly-accreted high-sAM gas with the gas already in the halo. In the process, most of the AM is either turned into thermal energy or transferred to the hot halo. This picture is consistent with the results of section 6.3.2 and figure 6.3.1, where I showed that the dominant forces in the outer halo and up to the outskirts of the halo in the hot gas are pressure forces.

The fate of cold gas is significantly different. On average in all our halos, the cold gas has a sAM ~ 3 times larger than the hot gas throughout its accretion in the outer halo down to the inner halo. The cold gas is mostly in free-fall (Rosdahl and Blaizot, 2012) up to the inner halo, where the cold gas shocks and the sAM quickly drops down to values comparable to the hot gas. While significant deviations are found from halo to halo, see the different panels of figure 6.3.4, the mean Lagrangian history of the sAM is clearly different between the cold and the hot gas. Our results, together with figure 6.3.1, suggest that the spin-down of the gas happens due to the interaction with the inner DM halo and the stellar disk.

6.3.4 The orientation of the angular momentum

So far, I have only described the evolution of the magnitude of the sAM of the gas. In practice, the evolution of the orientation of the sAM evolves slightly differently. In order to quantify the evolution of the sAM orientation, a relevant quantity is the relative angle between the sAM at
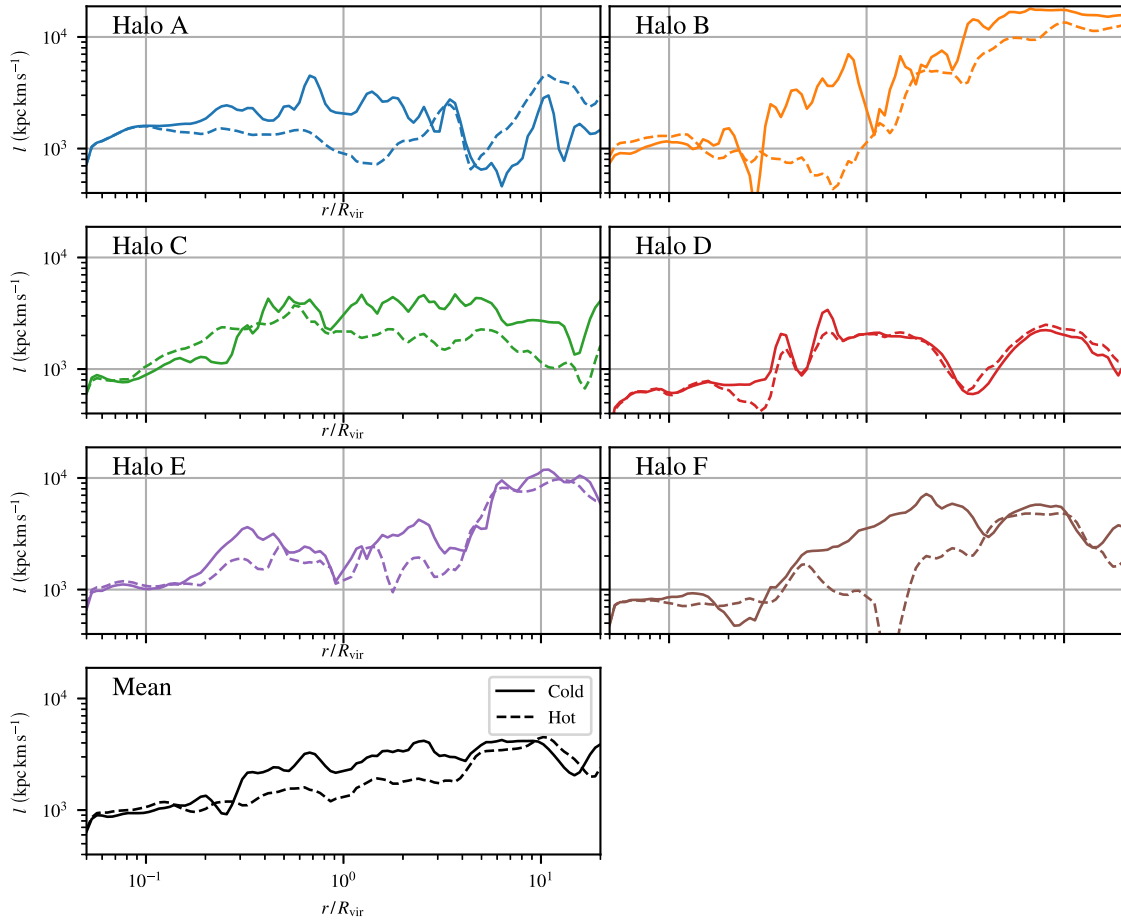
Figure 6.3.4: Evolution of the magnitudes of the mean sAM of the cold gas (solid lines) and of the hot gas (dashed lines) as a function of the distance to the halo centre for all halos. *Bottom right:* Mean value of the sAM averaged over all halos. The gas has been selected to cross the virial radius inward for the first time at $t = 2.2\,\mathrm{Gyr}$ ($z = 2.9$). In the outskirts of the halos ($r \sim 3R_{\mathrm{vir}}$), hot gas starts loosing sAM while cold gas conserves it down to the inner halo ($r \sim R_{\mathrm{vir}}/3$).

radius R_1, R_2, defined as

$$\cos\theta = \frac{\boldsymbol{l}(R_1) \cdot \boldsymbol{l}(R_2)}{\|\boldsymbol{l}(R_1)\|\|\boldsymbol{l}(R_2)\|}. \tag{6.9}$$

If the sAM orientation is conserved, equation (6.9) should have values close to one, whereas random reorientations yield values close to zeros. Values close to -1 are found in anti-aligned cases. The evolution of $\cos\theta$ is shown on figure 6.3.5, which presents the relative alignment of the sAM between its value at R_{vir} and its past value at $3R_{\mathrm{vir}}$ (left panel) and its value at the interface between the outer and inner halo ($0.3R_{\mathrm{vir}}$, centre panel) and between the inner halo and the disk ($0.1R_{\mathrm{vir}}$, right panel). The alignment angle is computed at crossing time ($r = R_{\mathrm{vir}}$) for all six halos. The sAM of the cold gas stays mostly aligned from $3R_{\mathrm{vir}}$ to $0.3R_{\mathrm{vir}}$ with typical misalignments of the order of $\pi/3$ ($\sim 60°$) or less. At its entry in the disk, most of the original orientation has been lost. I however report a weak yet non-null alignment. Before entering the halo, the evolution of the hot gas is similar to the cold gas: the orientation is conserved from $3R_{\mathrm{vir}}$ to R_{vir} but it becomes significantly less aligned between R_{vir} and $R_{\mathrm{vir}}/3$, where the misalignment is typically of the order of $2\pi/5$ ($\sim 70°$). I do not report any significant evolution of the sAM
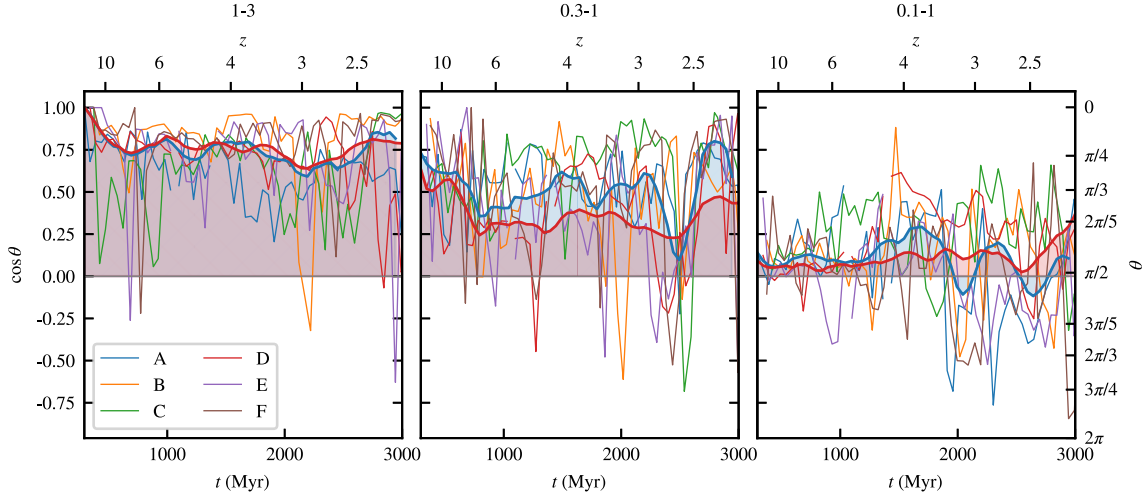
Figure 6.3.5: *Left:* Relative orientation of the sAM of the cold gas at R_{vir} compared to its value at $3R_{\mathrm{vir}}$ (*left*), $R_{\mathrm{vir}}/3$ (*middle*) and $R_{\mathrm{vir}}/10$ (*right*) for each halo (thin lines). The blue thick line shows the median value for the cold gas, smoothed over 11 consecutive outputs ($550\,\mathrm{Myr}$) using a fourth-order Savgol filter and the red thick line shows the median value for the hot gas smoothed in a similar way. In all simulations, the orientation of the sAM of the cold gas is conserved down to $\sim R_{\mathrm{vir}}/3$. Upon the entry in the disk, the sAM is reoriented and loses its connection to the large scale. The sAM of the hot gas start decoupling at larger radii.

orientation with redshift.

6.3.5 Dominant torques in the cold and hot phase

Figure 6.3.6 shows a 3D representation of the sAM, pressure torques and gravitational torques acting on the cold gas of halo A at $z = 3$. The figure illustrates that both sAM and gravitational torques have a coherent long-range spatial structure, as neighbouring vectors are aligned one with each other. On the contrary, pressure torques vary on scales similar or smaller than the size of filamentary structures, so that the net contribution of the pressure torque on a slab of filament cancels out, as neighbouring vectors point in opposite directions. Gravitational torques have a net (positive or negative) contribution thanks to their large-scale coherence, so that their effect adds up. More quantitatively, the coherence of the torques can be estimated by comparing the local torque value to the local torque standard deviation. This is similar to computing their "signal-to-noise" ratio, where the signal is the torque magnitude and the noise is its local deviation. Large values of this quantity are found in regions where torques have a coherent structures while small values are found in regions with no structure. In figure 6.3.7, I present mass-weighted projections of their signal-to-noise ratio, where the local standard deviation is computed using the 3^3 nearest cells in the cold gas. This illustrates that pressure torques have no spatial coherence, so that different locations of the cold flows may be either spun-up or spun down. On the contrary, large patches of the cold flows undergo coherent gravitational torques that can add up. Interestingly, gas gravitational torques seem to have more fluctuations than other gravitational torques, so that their net effect is small, even though they may contribute to the local force budget in the inner halo. In the disk, all torque sources lose their long-range spatial coherence and appear noisy. This signal-to-noise ratio for pressure torques is of the order of 10^{-3}, so that it is expected that the net contribution of pressure torques can be decreased by three order of magnitudes compared to their mean magnitude.
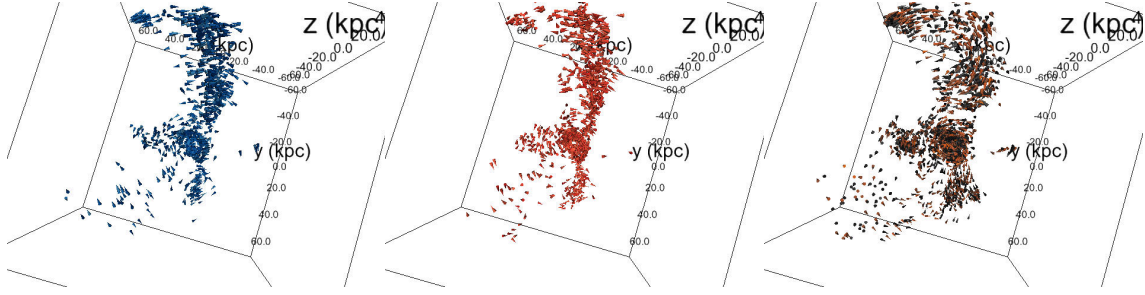
Figure 6.3.6: 3D representation of the sAM (left panel), pressure torques (central panel) and DM gravitational torques (right panel, black) and stellar gravitational torques (right panel, yellow) of the cold gas being accreted onto the central galaxy of halo A at $z = 2.7$. An interactive version can be found online. Pressure torques applied to the cold gas are mostly directed radially with respect to the filamentary structure, so that their net impact averages to zero. Gravitational torques are spatially coherent and contribute to a non-null net torque on the cold gas.

In order to go one step further, let me study the evolution of the cold gas by computing the contributions of the different torques to the spin-up or spin-down of the gas, projected on the axis of the mean sAM at a given radius. This is done on figure 6.3.8, which presents the Lagrangian evolution of the projection of the torques on the mean sAM at $5R_{\rm vir}$ (left panel), $R_{\rm vir}$ (centre panel) and $R_{\rm vir}/2$ (right panel) for halo B. The quantity plotted here is the projection of each torque on the mean sAM of the cold gas at a given radius

$$\tau_{\parallel,i,R_0}(t) \equiv \boldsymbol{\tau}_i(t) \cdot \frac{\sum_{\rm part} \boldsymbol{l}_i(r = R_0)}{\left|\sum_{\rm part} \boldsymbol{l}_i(r = R_0)\right|}, \tag{6.10}$$

where $R_0 = 5R_{\rm vir}, R_{\rm vir}, R_{\rm vir}/2$ respectively and i denotes the pressure torque or any of the gravitational torques. The cold gas has been selected to cross $r = R_{\rm vir}/3$ at $t = 2\,{\rm Gyr}$ ($z = 3.2$). The projected torque measures the propensity of the torque to spin the gas down if negative or spin the gas up if positive. In particular, large negative values contribute to remove the sAM acquired in the cosmic web. I have checked that the results presented are not sensitive to the radius at which the sAM has been measured, as long as it its measured in the outer halo or beyond. This is expected from figure 6.3.5, where I have showed that the sAM of the cold gas is well-aligned down to the inner halo, so that the orientation is conserved. I also report that using the mean sAM of the gas or the individual value of the sAM of each tracer particle lead to similar results (only the plot with the mean value is reported in this dissertation).

Figure 6.3.8 shows that, once averaged over the entire cold phase, pressure forces do not contribute significantly to the variation of the sAM of the gas. Indeed, I have shown on figure 6.3.7 that pressure forces are dominated by high-frequency spatial modulations, with a signal-to-noise ratio of the order of 10^{-3}. While the magnitude of the pressure forces are comparable to the DM gravitational forces, their net contribution to the torque budget is shown to be at least three order of magnitude smaller. As gas falls towards the galaxy, gravitational forces exert increasing torques resulting in a spin-down of the gas. In the inner halo down, torques become weakly aligned to the mean sAM of the gas at $R_{\rm vir}$, so that their projection can either contribute to the spin-up or spin-down in this specific frame, which is shown on figure 6.3.8 by rapid jumps from negative to positive values. Similar results can be found if one projects the torques on the axis of the AM vector of the galaxy at the end of the simulation, $L_\star(z = 2)$, as shown for halos A and B on figure 6.A.2. These plots also feature individual Lagrangian trajectories of the gas and illustrate that pressure torques spin the gas up as much as they spin it down. In contrast, gravitational
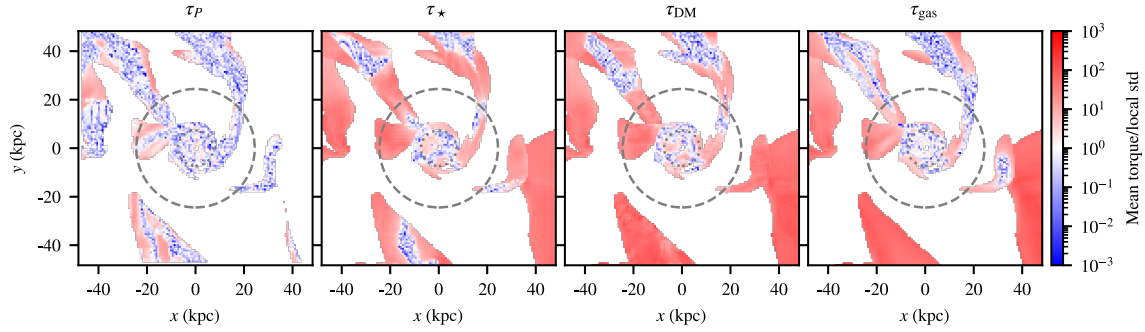
Figure 6.3.7: Mass-weighted projection of the ratio between the magnitude of the torques and the local standard deviation of the torques in halo A at $z = 2$ from left to right, for pressure torques, stellar gravitational torques, DM gravitational torques and gas torques. The local standard deviation is computed using the value of the torque in the 8 nearest cells. Blue regions indicate regions where torques are distributed randomly and red regions indicate where torques have a smooth and coherent distribution. The inner halo ($R_{\mathrm{vir}}/3$) is indicated by the grey dashed circle, while the dotted grey circle indicates $R_{\mathrm{vir}}/10$. In all regions, pressure torques have no spatial coherence on kpc scales. All gravitational sources have a much larger coherence scale, apart in a few regions in the filaments and in the disk.

torques are coherent over the Lagrangian evolution of the gas, so that their contribution adds up to spin the cold gas down. The bottom-right panels of figure 6.A.2 show the ratio of the DM gravitational torques to the stellar gravitational torques. As shown in figure 6.3.8, stellar gravitational torques are negligible in the outer halo but become dominant in the inner halo and in the disk.

The hierarchy between the different torques can in principle evolve with redshift. In order to study their relative importance, I have computed the total pressure torques, DM gravitational torques and stellar torques and compared the magnitude of each torques to the total torques from all sources $\boldsymbol{\tau}_{\mathrm{all}} = \boldsymbol{\tau}_{\mathrm{P}} + \boldsymbol{\tau}_{\mathrm{DM}} + \boldsymbol{\tau}_{\star}$. The ratio r is then defined as

$$r_i = \frac{\left|\sum_{\mathrm{particles}} \boldsymbol{\tau}_i\right|}{\left|\sum_{\mathrm{particles}} \boldsymbol{\tau}_{\mathrm{all}}\right|}. \tag{6.11}$$

Here i can be any of $\mathrm{P}, \mathrm{DM}, \star$ and sums run over all cold gas particles. The results are presented on figure 6.3.9, where torque ratios are presented as a function of the radial distance to the galaxy. Note that due to the definition, ratios can exceed one[4]. The figure shows that after a settling time of about $1\,\mathrm{Gyr}$ ($z = 5.7$), the ratios of each torque are constant at all radii, with the DM gravitational torques dominating in the outer halo and stellar gravitational torques dominating around the disk. Outside the halo at $2R_{\mathrm{vir}}$, the gravitational torques contribute to 90% of the total torques, while this fraction decreases to about 60% at $R_{\mathrm{vir}}/3$. In the disk, gravitational torques are dominated by stellar gravitational torques. In the six halos, the net contribution of pressure torques is negligible.

6.4 Discussion

At large radii, the evolution of the AM follows the tides imposed by the cosmic web, as explained by the TTT (e.g. Codis et al., 2012). The gas then flows on the forming galaxy *via* two different

---

[4]This can happen if two torques have similar magnitudes but opposite directions.
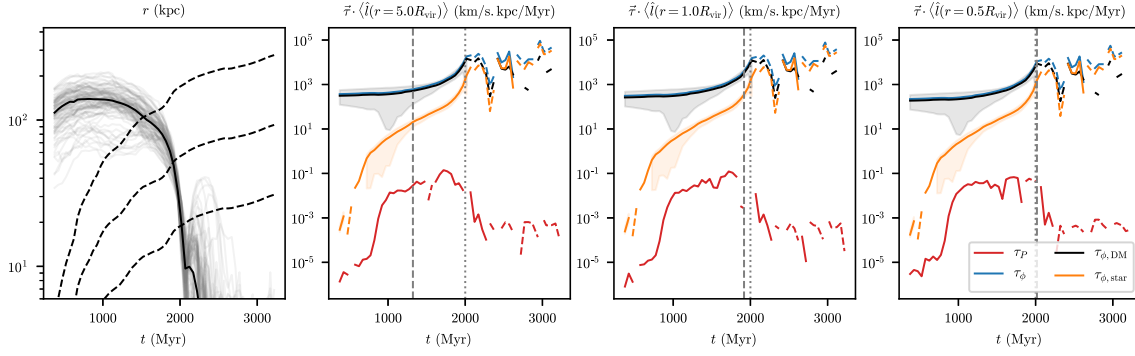
Figure 6.3.8: Evolution of the radius (*left panel*) and the specific torques projected on the direction of the mean sAM of the gas at $r = 5R_{\mathrm{vir}}$, $r = R_{\mathrm{vir}}$ and $r = R_{\mathrm{vir}}/2$ for halo B (from left to right). Solid lines indicate negative values (spin down) and dashed lines positive values (spin up). Particle are selected to cross $R_{\mathrm{vir}}/3$ at $t = 2\,\mathrm{Gyr}$ ($z = 3.2$) (vertical dotted lines). The mean time at which the sAM is measured is shown as vertical dashed lines. In all regions, pressure torques are negligible, as a result of averaging the projection of a randomly oriented field (the pressure torques) onto a fairly smooth field (the sAM).

channels: the hot and cold accretion, in particular for massive enough galaxies at $z \gtrsim 2$ (Birnboim and Dekel, 2003; Dekel and Birnboim, 2006; Pichon et al., 2011; Nelson et al., 2013). The predominance of one or the other channels of accretion can be used to understand the formation of disky galaxies and the internal evolution of the galaxy. Indeed, in cold flows that result from cold accretion, the gas is able to penetrate deep in the halo and can feed the galaxy with fresh gas, with a steady AM orientation (Pichon et al., 2011; Stewart et al., 2013; Danovich et al., 2015). In numerical simulations, it has been observed that cold gas has a higher AM at larger radii, as measured by their spin parameter (Kimm et al., 2011; Tillson et al., 2015; Danovich et al., 2015) which is up to one order of magnitude larger than that of the DM. In the inner halo and the disk however, the spin parameter of the cold-accreted gas is found to be only three times larger than that of the DM at the same location. The nature of the torques acting to reduce the AM of the gas is still debated today. While Danovich et al., 2015 argued that the dominant torques are gravitational torques regardless of the distance to the galaxy, Prieto et al., 2017 instead found that the dominant torques were pressure torques. In this work, I find that the pressure forces are dominant in the hot phase and are as important as the DM gravitational forces in the outer halo, in particular in the ortho-radial direction. In the inner halo, a transition occurs so that the dominant forces become stellar gravitational forces. I also report a significant "pressure ring" in the inner halo that may affect the kinematics of the infalling gas in which pressure forces dominate over all forces in both the cold- and hot-accreted gas. This pressure ring may impact the thermodynamical evolution of the gas, as well as to contribute to mixing the cold-accreted material to the hot gas, effectively blurring the line between hot- and cold-accreted material. The study of this pressure ring will be the topic of future work.

While pressure forces can act locally as the dominant forces, I report that their net contribution to the evolution of the cold gas is negligible. Indeed, pressure forces do not possess any structure over hundreds of parsecs, so that their individual contributions to the evolution of the cold gas cancel out. On the contrary, gravitational forces, that depend on the distribution of matter on larger scales, are able to coherently apply torques on the infalling material, resulting in most of the spin-down signal.

The net effect of the gravitational forces is reported to be a spin-down of the accreted gas, as
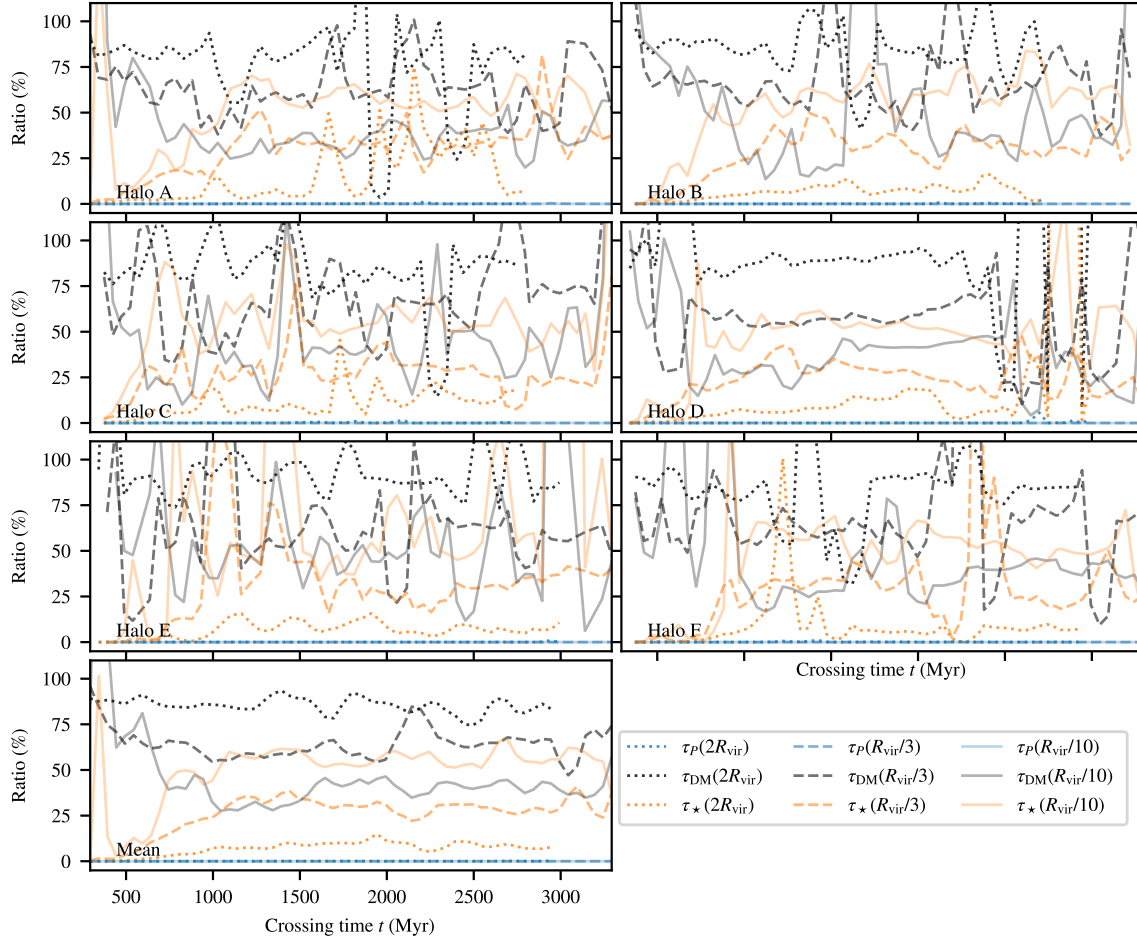
Figure 6.3.9: Absolute value of the torque ratios r_i measured in the cold gas (see text for details) as a function of time in different halos for different radial distance, as labelled. *Bottom left:* Mean value of the torque ratios, averaged over all six halos. After 1 Gyr, there is no average evolution of the torque ratios at any radius.

a result of dynamical friction and gravitational torques. Most of the spin-down of the hot gas happens before entering the halo and is due to DM torques. One possible reason is the following: under the effect of gas infall, the DM halo becomes slightly polarised which in turn creates a tidal field that will torque the hot gas down. Using the ortho-radial gravitational forces reported in figure 6.3.1, the typical angular momentum of the gas upon its entry in the halo ($\sim 10^4\,\mathrm{km/s\,kpc}$) would be depleted in a time $t_{\tau,\mathrm{DM}}(R = 100\,\mathrm{kpc}) \approx 10^4\,\mathrm{km/s\,kpc}/10^{-1}\,\mathrm{km/s/Myr} \times 80\,\mathrm{kpc} \approx 1\,250\,\mathrm{Gyr}$, which is about twice the free-fall time of the halo $t_{\mathrm{ff}} = 500\,\mathrm{Myr}$ at $z = 2$. If the hot gas lingers in the outskirts of the outer halo during two free-fall times, the DM gravitational torques are large enough to get rid of most of the angular momentum before accretion. In our simulations, hot gas takes on average $(1\,000 \pm 500)\,\mathrm{Myr}$ to fall from $3R_{\mathrm{vir}}(z = 2)$ to $R_{\mathrm{vir}}(z = 2)/3$ where $R_{\mathrm{vir}}(z = 2)$ is the final virial radius of the halo at $z = 2$.

Interestingly, I find that, even though most of the AM has been lost before entering the halo, the orientation of the AM of the hot gas is well-conserved between R_{vir} and $R_{\mathrm{vir}}/3$. This can be explained either by the fact that the spin of the halo, which has been reported to be well aligned with the first principal axis of the large scale tides (Danovich et al., 2012) do not reorient significantly the AM of hot gas (which is itself aligned with the first principal axis of the large scale tides), or that the infall of the hot gas coincides with the loss of most, but not all, of its

angular momentum. In this scenario, the hot gas starts infalling at the sweet spot where most of the angular momentum has been lost (so that the centrifugal force becomes negligible), but before all of it has been removed. This problem will however require a more detailed analysis.

As reported in Rosdahl and Blaizot, 2012, the trajectory of the cold gas is different and follows a mostly radial (with a non-null impact parameters) free-fall trajectory. In our simulation, the cold gas typically takes $(500 \pm 350)\,\mathrm{Myr}$ to go from $3R_{\mathrm{vir}}$ to $R_{\mathrm{vir}}/3$, so that the halo gravitational torques are not large enough to reduce the AM of the cold gas. As the cold gas plunges into the halo, the influence of the disk increases up to the point where torques become dominated by stars. I report here that the location where the disk torques become important coincides with the location where most of the AM of the cold gas has been lost. This may be an indication that the disk is actually responsible for the spin-down of the cold gas. If so, one would need to understand how the disk responds to an anisotropic accretion and more work is needed to understand in details the origin of the torques originating from the disk. One can then suggest that both the inner halo and the disk will then tend to be aligned to the mean orientation of the inflowing material in a similar way. This may explain why galactic spin is well aligned with the internal halo's, while being only mildly aligned with the global halo spin.

Cornuault et al., 2018 suggested that cold flows do not survive within the halo. They suggested that they instead fragment into clouds while their internal pressure increases. In the process, the kinetic energy of the gas is lowered as part of it is converted into turbulence and the gas mixes much more efficiently with the hot gas of the halo, effectively loosing the shielding effect usually assumed for cold flows. In this scenario, the pressure gradients observed in my simulations may contribute to efficiently mix the angular-momentum rich cold gas to the hot gas. This would likely result to a diffusion of the AM of the cold gas into the hot medium and increase the relative importance of pressure torques to the problem of the AM transport.

Using idealised simulations, Mandelker et al., 2016; Padnos et al., 2018; Mandelker et al., 2019 showed that cold flows may also be sensitive to the Kelvin-Helmholtz instability. In particular, they showed that thin-enough filaments are destroyed before reaching the galaxy. In this last case, the cold gas would effectively lose its angular momentum to the hot halo before interacting with the galaxy. Interestingly, these studies also suggested that cold flows may entrain the neighbouring hot gas as they fall in while slowing down the infall of the cold gas, which may result in an efficient mixing of the AM at the boundary of the cold flows. Berlok and Pfrommer, 2019 suggested that the mixing may be decreased if one considers magnetised flows with field lines parallel to the flow, as a result of a magnetic tension working against the Kelvin-Helmholtz instability.

Nelson et al., 2015 studied the effect of AGN feedback on cold accretion. They showed that feedback is able to significantly increase the infall time. If the delay is large enough, DM gravitational torques may have time to remove all AM from the cold gas before it enters the inner halo — as is already the case with hot gas in the simulations presented in this dissertation. In another study on the effect of AGN feedback on cold flows, Dubois et al., 2013 showed that at $z \geq 6$, the AGN activity in massive halo is able to prevent cold flows from reaching the disk and significantly decreases the cold gas mass in the inner halo. On large scales, AGN activity increases the curvature of filaments and decreases their length. The exact effect of AGN feedback will require further studies. It may lead to a revision of the results presented in this dissertation depending on their ability to disrupt the cold flow structures prior to accretion and delay the infall. One possible way to study this would require running numerical simulations with different AGN feedback strengths.

6.5   Conclusion
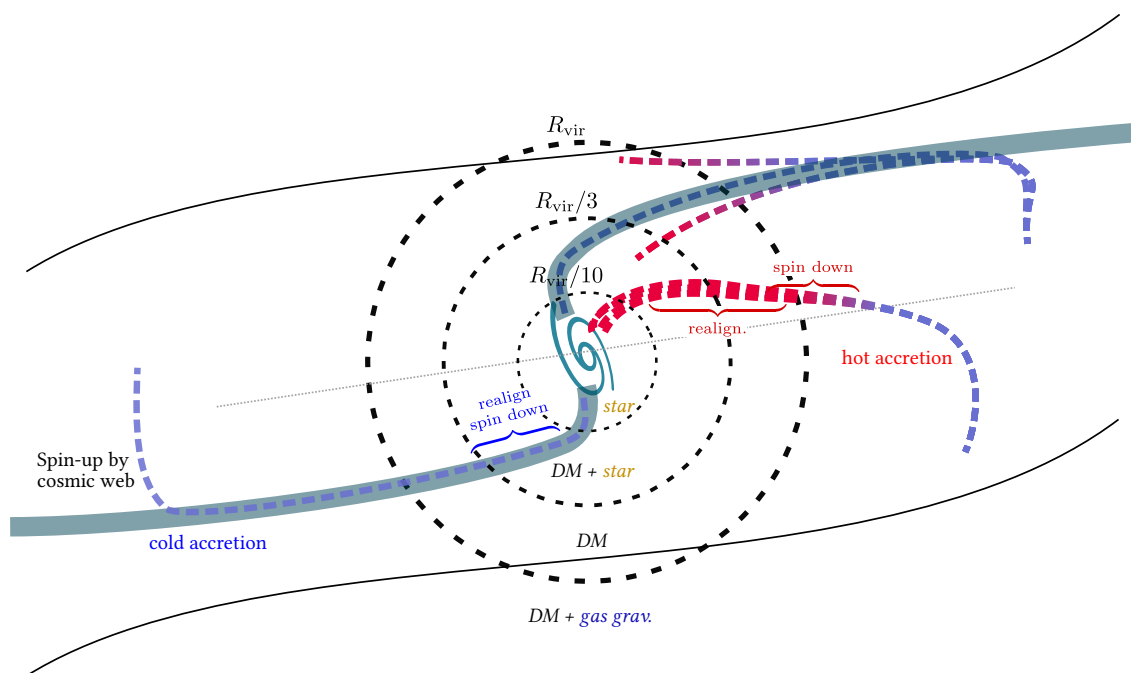
Figure 6.5.1: Sketch of the evolution of the AM at large z (not to scale). Hot gas (red dashed line) is spun-up by the cosmic web and loses most of its AM at the virial radius in the shock. Cold gas (blue dashed line) is spun-up by the cosmic web and retains its AM down to the inner halo. Between the inner halo and the disk, most of the AM is lost due to interactions with the DM halo and the disk.

Using a set of high-resolution zoom-in simulations, I have studied the evolution of the AM of gas accreted *via* the cold and the hot mode around six group progenitors at $z \gtrsim 2$. I also presented new numerical methods to extract the contributions of the different forces and torques (gravitational and pressure torques). My findings are the following:

1. the magnitude of the sAM of the cold gas is conserved down to the inner halo, the magnitude of the sAM of the hot gas is lost outside the halo,
2. the orientation of the sAM of the cold gas is conserved down to the inner halo,
3. the sAM of the hot gas is significantly less aligned to the large scale environment,
4. the dominant local forces in the cold gas are pressure forces and DM gravitational forces in the outer halo, and DM gravitational and stellar gravitational forces in the inner halo and the disk,
5. though, the pressure forces lack a spatial structure, so that their net contribution averages out in the cold gas,
6. therefore, the dominant torques in the cold cold gas are gravitational torques: DM gravitational torques dominate in the outer halo, stellar gravitational torques dominate in the disk.

The results on the major torques are sketched on figure 6.5.1. My findings indicate that the acquisition of the AM for group progenitors at $z \gtrsim 2$ is driven by the AM acquired at large-scale, consistent with the findings that the spin of galaxies is aligned with their environment. Most of the AM is able to flow down to the inner halo where gravitational torques redistribute it to the DM and the disk component, effectively transporting AM from the scales of the cosmic web to the scales involved in disk formation. These findings indicate that galaxy formation models aimed at

understanding AM acquisition should take into account the cold accretion mode, at least at high redshift. I have underlined that AM acquisition is dominated at large scales by the interaction with the cosmic web. In the halo, the evolution of the sAM of the cold flows is dominated by interactions with the inner halo and the disk. The sAM of the hot-accreted material is dominated by its interaction with the halo.

This work is part of an ongoing research effort. As a significant part of my work was devoted to the development of numerical methods, I have only touched some aspects of the complex angular momentum exchanges in the CGM. This fascinating topic will continue to motivate some of my research in the upcoming years.

6.A Additional material

In this section, I provide additional material. Figures 6.A.1a and 6.A.1b present mass-weighted projections of the magnitude of the pressure gradients and of the DM gravitational forces for halo A. They complement figure 6.3.1 by providing views in the xz and yz planes. In particular, the structure of the pressure ring is clearly visible in the pressure plots and not visible at all in the gravitational forces plots. Figures 6.A.2a and 6.A.2b present the evolution of the torques projected on the AM vector of the galaxy at $z = 2$ for halos A and B. The evolution of the torques projected on the axis of the galaxy are qualitatively similar for halos A and B.

(a) Pressure gradients

(b) DM gravitational forces
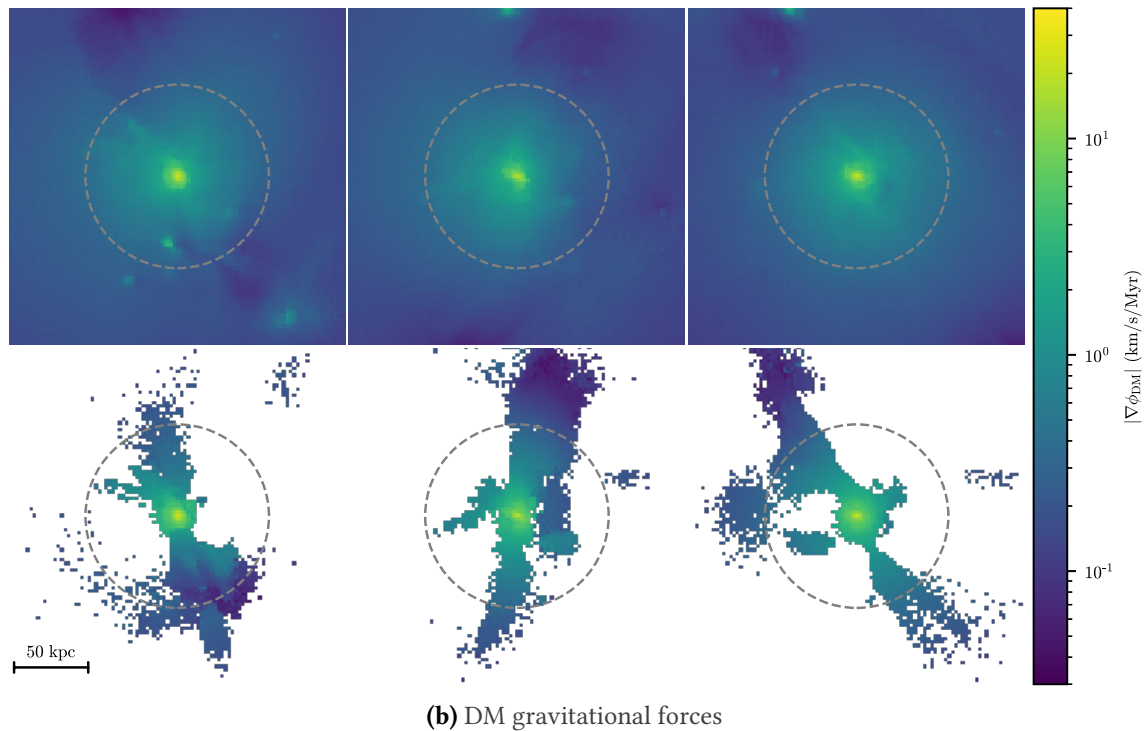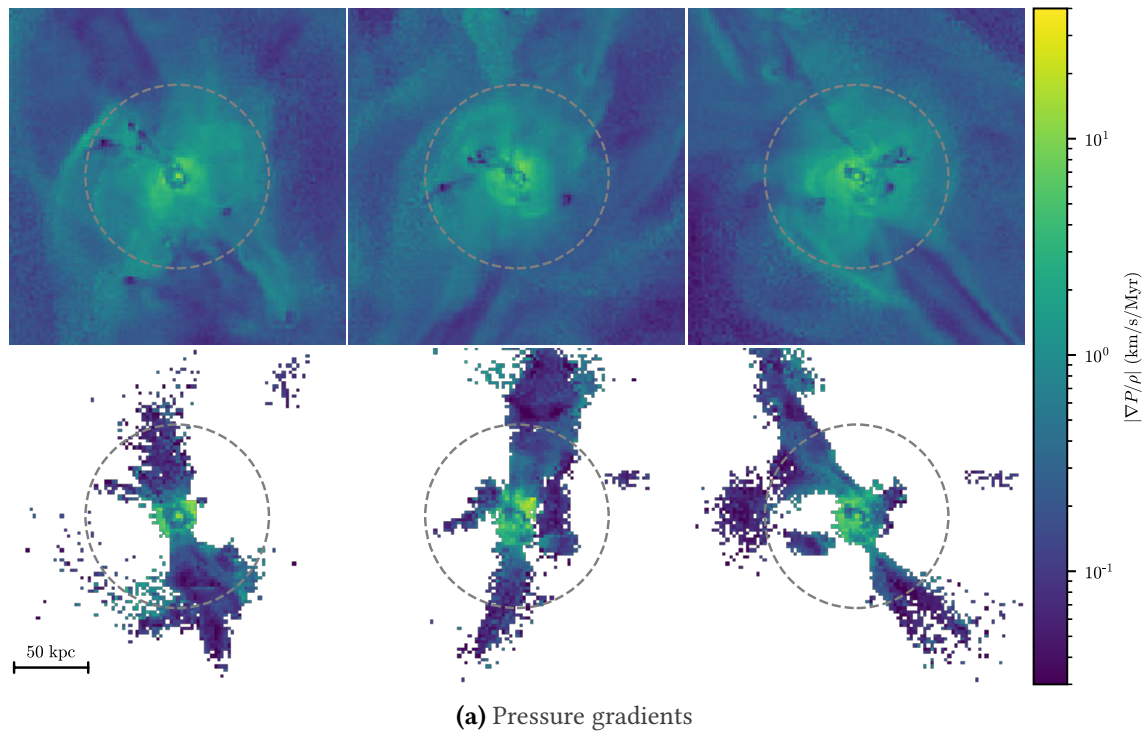
Figure 6.A.1: (a) Mass-weighted projections of the magnitude of pressure gradients and (b) of the magnitude of DM gravitational forces for the hot gas (top rows) and the cold gas (bottom rows) in halo A. From left to right in the xy, xz and yz plane. The grey circle is the virial radius of the halo. Pressure forces have a smaller magnitude in the cold gas than in the hot gas. DM gravitational forces have comparable magnitudes in the cold and hot gas.

(a) Halo A

(b) Halo B
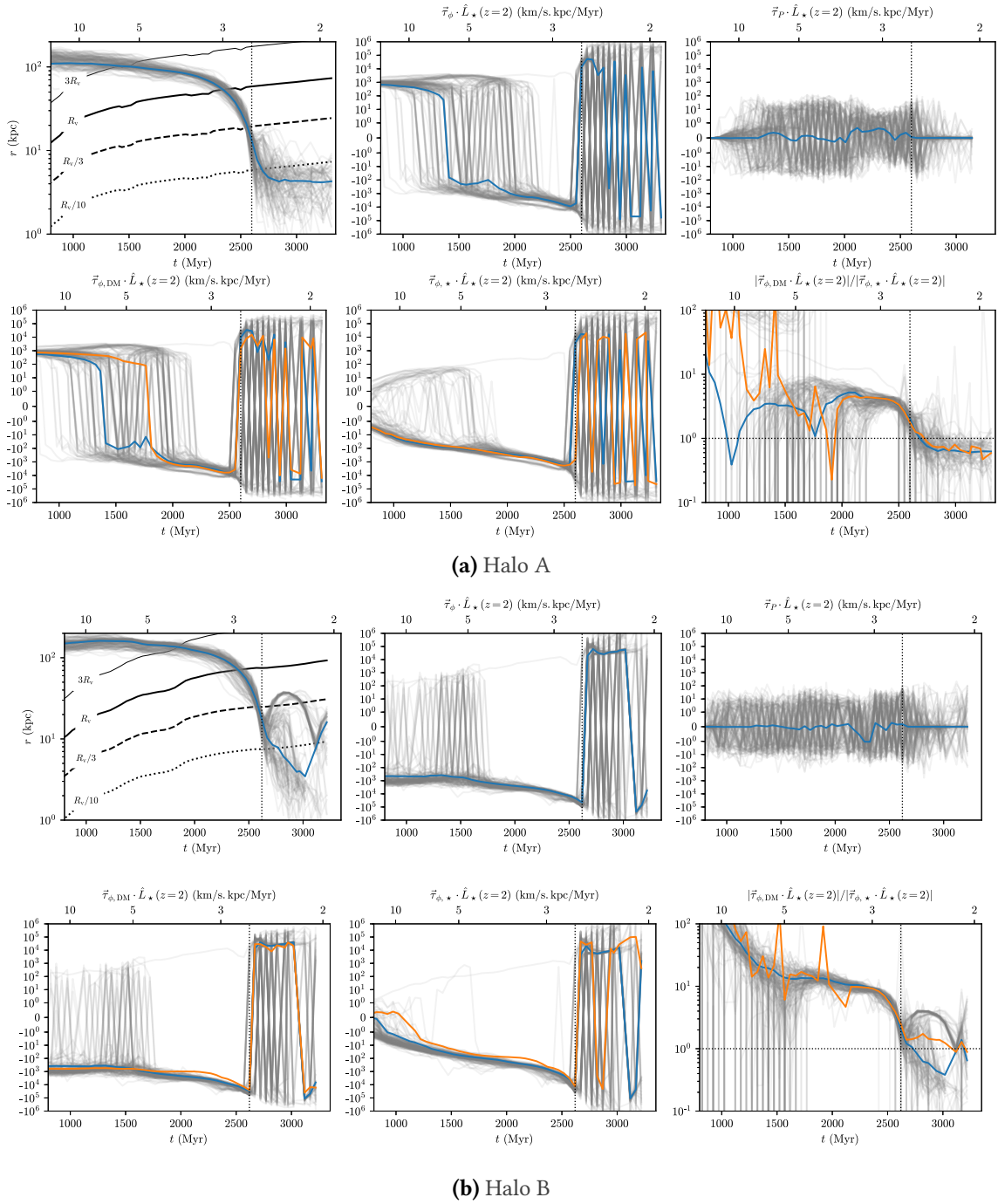
Figure 6.A.2: Trajectories (grey lines) of cold accreted gas particles with a first infall at $z = 2.5$ (vertical dotted line) in halo A (a) and B (b). *Upper left:* The radial distance to the galaxy, *upper centre:* The total gravitational torque, *upper right:* the pressure torques, *bottom left:* the DM gravitational torques, *bottom centre:* the stellar gravitational torques and *bottom right:* the ratio of the DM to stellar gravitational torques. All the torques are projected onto the normalised AM vector of the galaxy $\hat{L}_\star$. Median (blue) and mean (orange) values are shown as a function of time. DM is responsible to sAM acquisition at large radii. Stars and DM are responsible for the decrease of sAM at $\sim 2R_{\rm vir} > r > R_{\rm vir}/3$. In the inner halo, torques become dominated by stars.

7. Conclusion

7.1 Closing remarks

One of the key value of natural science lays in its predictive power. Hence, in the context of structure formation, a central question that theoretical cosmology must address is the following: to what extent can today's properties of galaxies be predicted from the initial Gaussian random field from which they emerge? More specifically, how can one encode the initial conditions in a compressed way to predict the fate of galaxies that will emerge from them?

Using an extension of the excursion set theory (Bond et al., 1991; Lacey and Cole, 1993; Mo and S. D. M. White, 1996), I have shown in chapter 4 (Musso, Cadiou et al., 2018) that the cosmic web, and in particular large scale filaments, biases the formation of dark matter halos. The formalism predicts that the variables entering the assembly history of the halo, namely the halo formation time and the accretion rate, are modulated by the cosmic web. As a result, at fixed final mass, halos forming close to nodes of the cosmic web are found to accrete more and have formed at later times, in agreement with N-body simulations for large-mass halos (R. K. Sheth and Tormen, 2004; Gao et al., 2005; Wechsler et al., 2006; Dalal et al., 2008). This effect complements other suggestions that the tides may be responsible for the assembly bias signal as it is purely geometric: as halos grow by accreting material, they also probe larger scales whose statistical structure is set by the cosmic web. I also argued that this provides a natural frame in which the assembly signal is simply a spatial modulation, or stated differently, different assembly histories are to be expected at different locations. This framework has since been used to show that, in hydrodynamical simulations, the cosmic web has also an effect on the assembly of galaxies. In Kraljic et al., 2018; Kraljic et al., 2019 (appendices B.1 and B.2), we reported that the specific star formation rate and the velocity-to-velocity-dispersion ratio both present significant modulations along filaments, highlighting that, indeed, filamentary structures can be used as a metric to parametrize the assembly of dark matter halos and galaxies therein.

In my dissertation, I also highlighted a process entering galaxy and dark matter halo formation, namely the coalescence of critical points of the cosmic web, whose theory was developed in chapter 4. I computed the rate of merger events as a function of smoothing scale from the initial cosmic landscape to forecast special events which impact the geometry of galactic infall,

and in particular filament disconnection. Using an extension of the theory to the mildly non-linear regime, I showed that one can connect our predictions to results obtained from N-body simulations. In particular, the formalism is able to detect halo merger events, but also filament- and wall-mergers in the Lagrangian initial conditions. I argued that these events, named "critical events", may be relevant parameters entering galaxy models, in particular to understand the evolution of galaxy properties that depend on the geometry of the accretion (such as their spin or their velocity-to-velocity-dispersion). This could readily be used to constrain further the assembly of galaxies by providing variables describing the evolution of the environment.

In the current understanding of galaxy formation, the evolution of the baryons is driven by the cosmic web on large scales, while at small scales complex interactions between the gas, stars and AGNs and the dark matter halo drive most of the physics. While the impact of the cosmic web on halo and galaxy formation can be studied to some extent from first principles as I have demonstrated in chapters 3 and 4, the complex baryonic physics at play make the task much more complex on smaller scales. I have presented a novel tracer particle scheme in chapter 5 (Cadiou et al., 2019) that is able to accurately trace the Lagrangian trajectories of gas elements in the adaptive mesh refinement code RAMSES. The method has then been applied on a suite of hydrodynamical cosmological simulations to study the angular momentum acquisition on $z > 2$ galaxies, presented in chapter 6. I have studied the formation of disk galaxies at large redshift and showed that the information acquired by the gas at large scales is transported to the inner regions of the halo and in the galaxy. In particular, cold flows are able to retain most of their angular momentum down to the inner halo. In the inner halo and around the disk, complex gravitational torques redistribute the angular momentum to the inner halo and the stellar component. I argue that this may lead to a good alignment of the inner halo and the galaxy, since their angular momentum is partially driven by their interaction with cold flows. This internal alignment is also expected to reflect the large-scale tidal field set by the cosmic web, as most of the anisotropic information is transported to the internal regions.

As a final conclusion, I have shown that the cosmic web is able to influence the assembly of dark matter halos. One can build theoretical models in which part of the assembly bias can simply be interpreted as a large-scale environment modulation, which cannot be parametrized easily in terms of the local properties of the field, both for dark matter halos and galaxies. I proposed a set of parameters, the critical events, that are suited to the compact description of the evolution of the cosmic web and argued that the geometry of the accretion onto galaxies *via* cold flows, and its evolution, can have a significant impact on the properties of galaxies, in particular against the ones sensitive to the anisotropy of the flows. This is in particularly highlighted by a numerical study that showed that the angular momentum of the gas, set by the cosmic web, is effectively transported down to the galaxy where complex interactions redistribute it. I suggest that in order to capture effects beyond mass and density relations, models of galaxy and halo formation should be augmented by parameters describing the non-local structure of the cosmic web at large-scales in terms of its critical points (nodes, filament and wall centres) but also in terms of their evolution, as described by critical events.

7.2   Perspectives and applications of my work

Thanks to the novel tools (constrained excursion set theory, critical event theory, Monte Carlo tracer particles) I have recently developed, our prospects of understanding the impact of the cosmic web on galaxy formation is hopefully brighter. It should enable us to soon produce new results in the context of galaxy formation and large scale structure hydrodynamics.

7.2.1 Extensions of the constrained excursion set theory

On the theoretical side, the predictions on assembly bias have proven quite successful at providing a physical understanding of the effect of the cosmic web on halo formation (Kraljic et al., 2018; Kraljic et al., 2019, see appendices B.1 and B.2). There is however room for further improvement to get a more refined model. One possible extension would require taking into account ellipsoidal collapse (Hahn and Paranjape, 2014; Ludlow et al., 2014; Ramakrishnan et al., 2019). As a first step, one could extend the work of chapter 4 with a barrier that depends on the *mean* tidal structure of the field. While this would not provide a definitive answer, one would recover the predictions of Musso et al., 2018 with an extra modulation coming from the mean tidal environment in the form of a shifted, spatially-modulated barrier $\delta_{\rm c} \to \delta_{\rm c}(\boldsymbol{r})$, where $\boldsymbol{r}$ is the distance to the nearest structure. This would likely have the consequence of delaying collapse in high-tide regions. Another possible approach to the problem of the tides on the collapse of halos is to rely on a Monte-Carlo sampling to explore the different excursion trajectories constrained to their large-scale environment. One could then seek the first-crossing given a barrier that is itself a random variable of the smoothing scale (using for example the shear strength, Hahn et al., 2009; Castorina et al., 2016; Borzyszkowski et al., 2017). In addition to explicitly taking into account ellipsoidal collapse and the effect of tides, this would also enable us to use the first-crossing condition instead of an approximation, so that the results could be extended to smaller masses. A detailed study would also require taking into account the Zel'dovich displacement (in the spirit of the peak patch theory, Bond and Myers, 1996), and in particular study the effect of a galaxy travelling from voids to sheets, from sheets to filaments, and along filaments to nodes, as significant relative velocities can be acquired by the travelling galaxy which may explain the presence of quenched halos at the geometrical centre of filaments (Borzyszkowski et al., 2017; Romano-Díaz et al., 2017; Garaldi et al., 2018).

7.2.2 Following dust formation using Monte-Carlo sampling methods

The Monte Carlo tracer particle, described in chapter 5, have obviously a broad range of applications. They have already been used to follow the formation and disruption of clumps in Coma-like clusters (Beckmann et al., submitted, appendix B.3). The problem of following the Lagrangian history of elements in a Eulerian framework should also find applications in coupled dust grain-gas models. In particular, bi-fluid models of strongly coupled dust and gas mixtures are difficult to integrate numerically. For Lagrangian-based methods they tend to produce spurious dust aggregates when the grains are accumulated below the resolution length of the gas (Ayliffe et al., 2012). In AMR codes, it was recently proposed to treat the dust-gas mixture as a single fluid with a diffusion term fluid instead (Lebreuilly et al., 2019), yielding accurate results for strongly coupled (small) grains. This method however requires one fluid per grain size bin, so that the treatment of a large range of grain size quickly becomes tedious. In order to circumvent this problem, one could modify the Monte Carlo approach of chapter 5 to follow individual dust grains with arbitrary grain size distribution. One would have to modify the transition probability involved in the tracer particle scheme (equation 5.1) to account for the relative drift of the dust grain with respect to the gas. This approach should prove particularly useful in simulations where different grain sizes can form, as the bi-fluid approach fails in the small grain regime (where dust is strongly coupled to the gas), while the dust-gas mixture approach fails in the large grain regime (where dust decouples from the gas).

7.2.3 Applications beyond cosmology

The analysis of chapter 4 was mostly restricted to (quasi) Gaussian random fields, because of their relevance in cosmology and also because in this context the theory can be developed in some details. However, any system involving random field whose (continuous) evolution is controlled

by one parameter could in principle be investigated with this framework in order to identify merger of ridges (though the specific role played by Gaussian smoothing would clearly generally not hold). For instance, critical events in dust maps (such as Meisner and Finkbeiner, 2014; Planck Collaboration, 2018b) could be used as an alternative statistics to quantify the properties of the underlying turbulence, a process which is known to display self similarities.

A wide range of important physical processes occur when rare events collide, hence boosting probabilities and passing thresholds, which in the context of this work corresponds to mergers of rare peaks (e.g. analysing dust map emission or disintegration events in Fermi maps). In this context, the process of interest is the appearance of pairs of critical points as one 'unsmooths' the field: this will corresponds to the generation of pairs of critical points. Following the results of section 4.A.3, the formalism could be extended to situations where the field whose evolution is investigated corresponds to probability distributions living in higher dimensions (or on more complex manifolds).

In the context of streaming of hierarchical images the set of critical events within a 2D image characterises its multi-scale topology. It would therefore be of interest to send beforehand a description of this set as a mean of prioritising which sub region of the image needs to be streamed first because the topology of its excursion (*i.e.* the local parsimonious representation of the image as iso-contours) has changed. This would allow the received image to acquire its most important higher resolution features first.

7.2.4   Critical events as input to Machine learning and Bayesian inference

The physics of galaxies is largely driven by non-linear processes (gas cooling, feedback, star formation). In order to capture these processes, we usually make use of numerical simulations that reproduce the known physics and can later be used to learn which processes are key to galaxy formation. In particular, I argue that some properties of galaxies, and in particular vector quantities (the spin, see e.g. Obuljen et al., 2019) or those sensible to the recent accretion history (v/σ, star formation rate) can be better understood if one takes into account the merger history of the halo *and* its environment. Indeed, there is a long tradition of relying on merger trees of dark halos extracted from simulations as a mean to predict the physical properties of galaxies (with so-called semi-analytical models, see e.g. Benson, 2010, and reference therein). One of the long term main motivations for the present work is to extend this strategy to the other two merger trees (filaments and walls), and to rely on modern segmentation techniques to identify which combination of events are most likely to lead to galaxies of a certain type to be produced in cosmological simulations. This strategy is likely to be efficient and rewarding, as the set of critical events is a very strong compression of the initial conditions, and because once the segmentation has been done, the subset of events which are in the past lightcone of a galaxy with a given tag have physical meaning. For instance, recent disconnection of filaments at a given smoothing scale are likely to impact gas infall at the redshift of corresponding smoothing scale, hence associated galaxy star formation and disc reformation. The set of critical events represents a useful effective topological compression of the initial conditions which will impact the upcoming 'dressed' mergers (*i.e.* the cosmic evolution of peaks *and* their filaments and walls). Note that the exact relative configuration of critical events in the smoothing-position space may be of relevance, and is not fully captured by the sole knowledge of the one and two-point statistics. Since the link between the galaxy formation and their properties is still poorly understood, a model agnostic approach can be used to study the effect of critical events on galaxy formation.

Predicting galactic properties using machine learning

In order to assess this, we can rely on machine learning techniques. Let me illustrate the strategy one could use using a catalogue of synthetic galaxies from a cosmological simulation. Let me assume that a set of virtual galaxies have been classified with a continuous parameter, e.g. based

on their morphology *via* their kinematic properties, v/σ. This ratio is computed from the 3D velocity distribution of stellar particles of each galaxy. In the frame of the angular momentum of that galaxy, the velocity is decomposed into cylindrical components v_r, v_θ, v_z, and the rotational velocity of a galaxy v is defined as the mean of v_θ of individual stars. The average velocity dispersion of the galaxy $\sigma^2 = (\sigma_r^2 + \sigma_\theta^2 + \sigma_z^2)/3$ is computed using the velocity dispersion of each velocity component σ_r, σ_θ and σ_z. This ratio allows me to separate rotation-dominated ($v/\sigma \gg 1$) from dispersion-dominated ($v/\sigma \ll 1$) galaxies. For each central galaxy identified in the simulation, one can identify their corresponding dark matter halo to trace the Lagrangian patch of dark matter particles back into the ICs. This defines a connected gravitational patch that contains all critical events causally connected to the final galaxy. Hence, the simulation provides me with a set of relations for k patches and three types of critical events $j \in [\mathcal{P}, \mathcal{F}, \mathcal{W}]$

$$\left(\{\Delta \boldsymbol{r}_{j,i}, R_{j,i}, \nu_{j,i}\}_{i \leq n_{j,k}} \right)_{j \in [\mathcal{P}, \mathcal{F}, \mathcal{W}]} \to (v/\sigma)_k \,, \qquad (7.1)$$

where $\Delta \boldsymbol{r}_{j,i}$ is the relative position *within the patch* of the critical event i of type j measured w.r.t. the centre of mass of the patch, $\nu_{j,i}$ is its contrast, and $R_{j,i}$ the corresponding smoothing scale, while v/σ_k is the velocity ratio of the patch k. Let me call $\mathcal{E}_k$ the l.h.s. of this relation. Standard machine learning tools (random tree forest, stochastic gradient descent), allows me to build a predictor, $P_r(\mathcal{E})$ from a subset of $(\mathcal{E}_k \to v/\sigma_k)_{k \leq K_{\text{train}}}$ drawn randomly from the full sample. From this training, one can do one of two things: i) use it as a predictor to associate $(v/\sigma)_k$ to other patches for which we computed their set of events, $\mathcal{E}_k$. ii) identify which features in this event set is responsible for the corresponding value of v/σ.

The former approach would be useful to find regions of interest in the initial conditions, therefore avoiding a costly try-and-error approach. For example, this could provide a likelihood of finding a galaxy with given morphology in the initial conditions, so that only regions of interest are resolved with high resolution. This is usually tackled by running larger than necessary simulations, in which only the regions of interest are kept *a posteriori*. While the approach suggested here would still have a chance of failing, it could significantly decrease the computation volume required to simulate a given configuration and could complement other approaches, such as genetically modified initial conditions (Roth et al., 2016; Rey and Pontzen, 2017). This could also provide useful insight to understand which halos host early-growing SMBHs. Indeed, it was recently proposed (Huang et al., 2019) that the structure of the initial conditions preconditions the early evolution of SMBHs in numerical simulations, while another study showed that the environment has an impact on AGN properties (Porqueres et al., 2018; Man et al., 2019). This problem is tightly coupled to the problem of feeding galaxies with fresh gas, which, at high-redshift, is linked to the orientation and structure of the local cosmic web *via* cold filamentary accretion. Here I argue that critical events are useful tools to study this class of problem, in particular at high redshifts.

The latter approach could be implemented over sets of simulations which implement different feedback recipes as a mean of disentangling the relative impact of environment and sub-grid physics on the evolution of galaxies

Critical events as input to Bayesian framework

Let me illustrate how one could use critical events to study the susceptibility of a set of galactic parameters to their cosmic web environment. Using a similar approach as described in the previous paragraph, the sample can be decomposed as a set of inputs and output variables $\boldsymbol{X}, \boldsymbol{Y}$, treated here as random variables

$$\boldsymbol{X}_j^{(0)} = \{M_j, \rho_j\} \,, \quad \boldsymbol{X}_j^{(1)} = \left\{ \{\Delta \boldsymbol{r}_{i,j}, R_{i,j}, \nu_{i,j}, k_{i,j}\}_{i=1,\dots,N_j} \right\}_j \,, \quad \boldsymbol{Y}_j = (v/\sigma)_j. \qquad (7.2)$$

where $\Delta \boldsymbol{r}_{j,i}$ is the relative position *within the patch*, $R_{i,j}$ is the smoothing scale, $\nu_{i,j}$ is the density contrast and $k_{i,j}$ is the kind ($k \in [\mathcal{P}, \mathcal{F}, \mathcal{W}]$) of the critical event i. Here I have split the inputs

between the variables commonly included in galaxy formation models, the mass of the halo M_j and the density ρ_j, which I named $\boldsymbol{X}_j^{(0)}$. The input variable $\boldsymbol{X}_j^{(1)}$ contains the "augmented" past history as encoded by the N_j critical events found in the Lagrangian patch of galaxy j. The output variable is chosen here to be v/σ, but it could be any other galactic property, such as the bulge mass or the SMBH mass. Using the concept of cross-entropy, one can then compute the information gain on the distribution of v/σ when adding to the classical mapping $\boldsymbol{X}^{(0)} \mapsto \boldsymbol{Y}$ the augmented merger tree $(\boldsymbol{X}^{(0)}, \boldsymbol{X}^{(1)}) \mapsto \boldsymbol{Y}$. The cross-entropy will tell us how many more bit of (Shannon) information is gained by adding the information from the merger history of the cosmic web in the gravitational patch of the galaxy. This can then be further extended by segmenting $\boldsymbol{X}^{(1)}$ into subsets containing only halo mergers, filament mergers and wall mergers to quantity which event better encodes the parameter v/σ.

When co-analysing the evolution of galactic properties with critical point mergers, one could relate the various (filament, wall) mergers to special events in terms of change in connectivity and feedback (e.g. the destruction of filaments by AGN activity, see Dubois et al., 2013). It could also be used to explore the relation between spin flip and with filaments or walls vanishing.

In this dissertation, I have provided new models and tools, which, in conjunction with my numerical work, should prove fruitful for research in galaxy and halo formation theory, astronomy and beyond.

A. Notations and conventions

Table A.1 presents the different notations and conventions used throughout the manuscript.

Table A.1: Conventions and notations used throughout the manuscript

Name	Definition
Vectors	$\boldsymbol{r}, \boldsymbol{X}, \boldsymbol{\Sigma}, \ldots$
Matrices	$\mathbf{A}, \mathbf{X}, \boldsymbol{\Sigma}, \ldots$
Unitary vector	$\hat{\boldsymbol{x}}, \hat{\boldsymbol{r}}, \ldots$
Probability density function	$p(x)$
Gaussian probability density function	$p_{\mathrm{G}}(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\!\left(-x^2/2\sigma^2\right)$
Spatial derivatives, Laplacian	$\partial_i, \nabla^2 \equiv \delta^{ij}\partial_i\partial_j$
Divergence, curl operator	$\boldsymbol{\nabla} \cdot f, \boldsymbol{\nabla} \times f$
Fourier transform	$f(\boldsymbol{k}) = \displaystyle\int \mathrm{d}^3\boldsymbol{r}\, f(\boldsymbol{r})e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$
Expectation	$\langle Q \rangle = \displaystyle\int \mathrm{d}x\, Q(x)p(x)$
2-point function	$\xi(r) = \langle \delta(0)\delta(r) \rangle$
Kronecker symbol	$\delta_{ij} = 1$ for $i = j$ and 0 otherwise
Dirac distribution	$\delta_{\mathrm{D}}(x)\,, \displaystyle\int \mathrm{d}x\, \delta_{\mathrm{D}}(x)\, f(x) = f(0)$
Heaviside step function	$\vartheta_{\mathrm{H}}(x) = 1$ for $x > 0$ and 0 otherwise
Linear matter power spectrum	$P(k)$
Linear density field smoothed on scale R	$\delta(R)$
Variance of linear density field on scale R	$\sigma^2(R)$
Critical density (spherical collapse)	$\delta_{\mathrm{c}} = 1.686$
Peak rarity on scale R	$\nu(R) = \delta(R)/\sigma(R)$

B. Contributed publications

Outline

In the context of my work, I contributed to the three papers presented in this chapter. The first two (Kraljic et al., 2018, appendix B.1 and Kraljic et al., 2019, appendix B.2) stemmed from my theoretical work presented in chapter 4. I contributed to third one (Beckmann et al., submitted, appendix B.3) by providing my tracer particle implementation presented in chapter 5.

B.1   "Galaxy evolution in the metric of the cosmic web" (published in MNRAS)

In this paper, the role of the cosmic web in shaping the properties of the cosmic web is explored. The skeleton of the cosmic web is extracted from the spatial distribution of galaxies in the Galaxy And Mass Assembly (GAMA) spectroscopic survey. As a comparison, the similar job is carried on the HORIZON-AGN simulation. The properties of galaxies are projected on the frame of the cosmic web, namely they are computed as function of the distance to the closest filament and the closest node.

The paper shows that the cosmic web induces a segregation of galaxies. In particular most massive galaxies are found close to filament centres and the trend subsists for star forming galaxies. At fixed mass, there are more passive galaxies, red galaxies and quenched galaxies close to filament centres than outside of it. The paper also shows that part of the segregation signal

cannot be interpreted as the effect of the local density only, showing that the cosmic web has an effect on galaxy properties beyond the mere mass and density relations.

Using the formalism developed in chapter 3, I have shown (section 7.2 of the paper) that one can provide a theoretical explanation to this segregation signal. Given a large-scale filamentary structure, one can compute mass, density and accretion rate isocontours to show that they are misaligned one with each other. Rephrasing, different variables entering the assembly of DM halos and their galaxies show distinct spatial dependences on the environment as set by cosmic web. For example, it is expected from first principles that at fixed halo mass, the density maximum is found closer to the filament centre than the DM accretion rate peak. As a result, the cosmic web systematically biases the different variables responsible for galactic properties, so that they become a function of their spatial location with respect to the cosmic web, on top of their mass and density dependence.

Galaxy evolution in the metric of the cosmic web

K. Kraljic,[1]⋆ S. Arnouts,[1] C. Pichon,[2,3] C. Laigle,[4] S. de la Torre,[1] D. Vibert,[1]
C. Cadiou,[2] Y. Dubois,[2] M. Treyer,[1] C. Schimd,[1] S. Codis,[5] V. de Lapparent,[2]
J. Devriendt,[4] H. S. Hwang,[6] D. Le Borgne,[2] N. Malavasi,[7] B. Milliard,[1] M. Musso,[2,8]
D. Pogosyan,[9] M. Alpaslan,[10] J. Bland-Hawthorn[11] and A. H. Wright[12]

[1]*Aix Marseille University, CNRS, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France*
[2]*Institut d'Astrophysique de Paris, UMR 7095 CNRS et Université Pierre et Marie Curie, 98bis Bd Arago, F-75014, Paris, France*
[3]*School of Physics, Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul, 02455, Republic of Korea*
[4]*Sub-department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, United Kingdom*
[5]*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada*
[6]*Quantum Universe Center, Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea*
[7]*Department of Physics and Astronomy, Purdue University, 525 Northwestern Avenue, West Lafayette, IN 47907, USA*
[8]*Institut de physique théorique, Université Paris Saclay and CEA, CNRS, 91191 Gif-sur-Yvette, France*
[9]*Department of Physics, University of Alberta, 412 Avadh Bhatia Physics Laboratory, Edmonton, AB T6G 2J1, Canada*
[10]*NASA Ames Research Center, Moffett Field, Mountain View, CA 94035, USA*
[11]*Sydney Institute for Astronomy, School of Physics, University of Sydney, NSW 2006, Australia*
[12]*Argelander-Institut für Astronomie (AIfA), Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*

ABSTRACT

The role of the cosmic web in shaping galaxy properties is investigated in the Galaxy And Mass Assembly (GAMA) spectroscopic survey in the redshift range $0.03 \leq z \leq 0.25$. The stellar mass, $u - r$ dust corrected colour and specific star formation rate (sSFR) of galaxies are analysed as a function of their distances to the 3D cosmic web features, such as nodes, filaments and walls, as reconstructed by DisPerSE. Significant mass and type/colour gradients are found for the whole population, with more massive and/or passive galaxies being located closer to the filament and wall than their less massive and/or star-forming counterparts. Mass segregation persists among the star-forming population alone. The red fraction of galaxies increases when closing in on nodes, and on filaments regardless of the distance to nodes. Similarly, the star-forming population reddens (or lowers its sSFR) at fixed mass when closing in on filament, implying that some quenching takes place. These trends are also found in the state-of-the-art hydrodynamical simulation HORIZON-AGN. These results suggest that on top of stellar mass and large-scale density, the traceless component of the tides from the anisotropic large-scale environment also shapes galactic properties. An extension of excursion theory accounting for filamentary tides provides a qualitative explanation in terms of anisotropic assembly bias: at a given mass, the accretion rate varies with the orientation and distance to filaments. It also explains the absence of type/colour gradients in the data on smaller, non-linear scales.

Key words: large-scale structure of Universe – cosmology: observations – galaxies: evolution – galaxies: high-redshift – galaxies: statistics.

1 INTRODUCTION

Within the Λ cold dark matter (ΛCDM) cosmological paradigm, structures in the present-day Universe arise from hierarchical clustering, with smaller dark matter haloes forming first and progressively merging into larger ones. Galaxies form by the cooling and condensation of baryons that settle in the centres of these haloes (White & Rees 1978) and their spin is predicted to be correlated with that of the halo generated from the tidal field torques at the moment of proto-halo collapse (tidal torque theory, TTT; e.g. Peebles 1969; Doroshkevich 1970; Efstathiou & Jones 1979; White 1984). However, dark matter haloes, and galaxies residing within them, are not isolated. They are part of a larger-scale pattern, dubbed the cosmic web (Jõeveer, Einasto & Tago 1978; Bond, Kofman & Pogosyan 1996), arising from the anisotropic collapse

⋆E-mail: katarina.kraljic@lam.fr

of the initial fluctuations of the matter density field under the effect of gravity across cosmic time (Zel'dovich 1970).

This web-like pattern, brought to light by systematic galaxy redshift surveys (e.g. De Lapparent, Geller & Huchra 1986; Geller & Huchra 1989; Colless et al. 2001; Tegmark et al. 2004), consists of large nearly-empty void regions surrounded by sheet-like walls framed by filaments which intersect at the location of clusters of galaxies. These are interpreted as the nodes, or high-density peaks of the large-scale structure pattern, containing a large fraction of the dark matter mass (Bond et al. 1996; Pogosyan et al. 1996). The baryonic gas follows the gravitational potential gradients imposed by the dark matter distribution, then shocks and winds up around multistream, vorticity-rich filaments (Codis et al. 2012; Hahn, Angulo & Abel 2015; Laigle et al. 2015). Filamentary flows, along specific directions dictated by the geometry of the cosmic web, advect angular momentum into the newly formed low mass galaxies with spins typically aligned with their neighbouring filaments (Pichon et al. 2011; Stewart et al. 2013). The next generation of galaxies forms through mergers as they drift along these filaments towards the nodes of the cosmic web with a post merger spin preferentially perpendicular to the filaments, having converted the orbital momentum into spin (e.g. Aubert, Pichon & Colombi 2004; Navarro, Abadi & Steinmetz 2004; Aragón-Calvo et al. 2007b; Codis et al. 2012; Libeskind et al. 2012; Trowland, Lewis & Bland-Hawthorn 2013; Aragon-Calvo & Yang 2014; Dubois et al. 2014; Welker et al. 2015).

Within the standard paradigm of hierarchical structure formation based on ΛCDM cosmology (Blumenthal et al. 1984; Davis et al. 1985), the imprint of the (*past*) large-scale environment on galaxy properties is therefore, to some degree, expected via galaxy mass assembly history. Intrinsic properties, such as the mass of a galaxy (and internal processes that are directly linked to its mass), are indeed shaped by its build-up process, which in turn is correlated with its *present* environment. For instance, more massive galaxies are found to reside preferentially in denser environments (e.g. Dressler 1980; Postman & Geller 1984; Kauffmann et al. 2004; Baldry et al. 2006). This mass-density relation can be explained through the biased mass function in the vicinity of the large-scale structure (LSS; Kaiser 1984; Efstathiou et al. 1988) where the enhanced density of the dark matter field allows the proto-halo to pass the critical threshold of collapse earlier (Bond et al. 1991) resulting in an overabundance of massive haloes in dense environments. However, what is still rightfully debated is whether the large-scale environment is also driving other observed trends such as morphology-density (e.g. Dressler 1980; Postman & Geller 1984; Dressler et al. 1997; Goto et al. 2003), colour-density (e.g. Blanton et al. 2003; Baldry et al. 2006; Bamford et al. 2009) or star formation-density (e.g. Hashimoto et al. 1998; Lewis et al. 2002; Kauffmann et al. 2004) relations, and galactic 'spin' properties, such as their angular momentum vector, their orientation or chirality (trailing versus leading arms).

On the one hand, there are evidences that the cosmic web affects galaxy properties. Void galaxies are found to be less massive, bluer, and more compact than galaxies outside of voids (e.g. Rojas et al. 2004; Beygu et al. 2016); galaxies infalling into clusters along filaments show signs of some physical mechanisms operating even before becoming part of these systems, that galaxies in the isotropic infalling regions do not (Porter et al. 2008; Martínez, Muriel & Coenda 2016); Kleiner et al. (2017) find systematically higher HI fractions for massive galaxies ($M_\star > 10^{11}\,M_\odot$) near filaments compared to the field population, interpreted as evidence for a more efficient cold gas accretion from the intergalactic

medium; Kuutma, Tamm & Tempel (2017) report an environmental transformation with a higher elliptical-to-spiral ratio when moving closer to filaments, interpreted as an increase in the merging rate or the cut-off of gas supplies near and inside filaments (see also Aragon-Calvo, Neyrinck & Silk 2016); Chen et al. (2017) detect a strong correlation of galaxy properties, such as colour, stellar mass, age and size, with the distance to filaments and clusters, highlighting their role beyond the environmental density effect, with red or high-mass galaxies and early-forming or large galaxies at fixed stellar mass having shorter distances to filaments and clusters than blue or low-mass and late-forming or small galaxies, and Tojeiro et al. (2017) interpret a steadily increasing stellar-to-halo mass ratio from voids to nodes for low mass haloes, with the reversal of the trend at the high-mass end, found for central galaxies in the GAMA survey (Driver et al. 2009, 2011), as an evidence for halo assembly bias being a function of geometric environment. At higher redshift, a small but significant trend in the distribution of galaxy properties within filaments was reported in the spectroscopic survey VIPERS ($z \simeq 0.7$; Malavasi et al. 2017) and with photometric redshifts ($0.5 < z < 0.9$) in the COSMOS field (with a 2D analysis; Laigle et al. 2017). Both studies find significant mass and type segregations, where the most massive or quiescent galaxies are closer to filaments than less massive or active galaxies, emphasizing that large-scale cosmic flows play a role in shaping galaxy properties.

On the other hand, Alpaslan et al. (2015) find in the GAMA data that the most important parameter driving galaxy properties is stellar mass as opposed to environment (see also Robotham et al. 2013). Similarly, while focusing on spiral galaxies alone, Alpaslan et al. (2016) do find variations in the star formation rate (SFR) distribution with large-scale environments, but they are identified as a secondary effect. Another quantity tracing different geometric environments that was found to vary is the luminosity function. However, while Guo, Tempel & Libeskind (2015) conclude that the filamentary environment may have a strong effect on the efficiency of galaxy formation (see also Benítez-Llambay et al. 2013), Eardley et al. (2015) argue that there is no evidence of a direct influence of the cosmic web as these variations can be entirely driven by the underlying *local* density dependence. These discrepancies are partially expected: the present state of galaxies must be impacted by the effect of the past environment, which in turn does correlate with the present environment, if mildly so, but these environmental effects must first be distinguished from mass-driven effects which typically dominate.

The TTT, naturally connecting the large-scale distribution of matter and the angular momentum of galactic haloes (e.g. Jones & Efstathiou 1979; Barnes & Efstathiou 1987; Heavens & Peacock 1988; Porciani, Dekel & Hoffman 2002a,b; Lee 2004), in its recently revisited, conditioned formulation (Codis, Pichon & Pogosyan 2015) predicts the angular momentum distribution of the forming galaxies relative to the cosmic web, which tend to first have their angular momentum aligned with the filament's direction while the spin orientation of massive galaxies is preferentially in the perpendicular direction. Despite the difficulty to model properly the halo-galaxy connection, due to the complexity, non-linearity and multiscale character of the involved processes, modern cosmological hydrodynamic simulations confirm such a mass-dependent angular momentum distribution of galaxies with respect to the cosmic web (Dubois et al. 2014; Welker et al. 2014, 2017). On galactic scales, the dynamical influence of the cosmic web is therefore traced by the distribution of angular momentum and orientation of galaxies, when measured relative to their embedding large-scale environment. The impact of such environment on the spins of galaxies has

only recently started to be observed (confirming the spin alignment for spirals and preferred perpendicular orientation for ellipticals; Trujillo et al. 2006; Lee & Erdogdu 2007; Paz et al. 2008; Tempel et al. 2013; Tempel & Libeskind 2013; Pahwa et al. 2016, but see also Jones, van de Weygaert & Aragón-Calvo 2010; Cervantes-Sodi, Hernandez & Park 2010; Andrae & Jahnke 2011, for contradictory results). What is less obvious is whether observed integrated scalar properties such as morphology or physical properties (SFR, type, metallicity, which depend not only on the mass but also on the past and present gas accretion) are also impacted.

Theoretical considerations alone suggest that local density as a sole and unique parameter (and consequently any isotropic definition of the environment based on density alone) is not sufficient to account for the effect of gravity on galactic scale (e.g. Mo, van den Bosch & White 2010) and therefore capture the environmental diversity in which galaxies form and evolve: one must also consider the relative past and present orientation of the tidal tensor with respect to directions pointing towards the larger-scale structure principal axes. At the simplest level, on large scales, gravity should be the dominant force. Its net cumulative impact is encoded in the tides operating on the host dark matter halo. Such tides may be decomposed into the trace of the tidal tensor, which equals the local density, and its traceless part, which applies distortion and rotation to the forming galaxy. The effect of the former on increasing scales has long been taken into account in standard galaxy formation scenarios (Kaiser 1984), while the effect of the latter has only recently received full attention (e.g. Codis et al. 2015). Beyond the above-discussed effect on angular momentum, other galaxy's properties could in principle be influenced by the large-scale traceless part of the tidal field, which modifies the accretion history of a halo depending on its location within the cosmic web. For instance, the tidal shear near saddles along the filaments feeding massive haloes is predicted to slow down the mass assembly of smaller haloes in their vicinity (Hahn et al. 2009; Borzyszkowski et al. 2017; Castorina et al. 2016). Bond & Myers (1996) integrated the effect of ellipsoidal collapse (via the shear amplitude), which may partially delay galaxy formation, in the Extended Press-Schechter (EPS) theory. Yet, in that formulation, the geometry of the delay imposed by the specific relative orientation of tides imposed by the large-scale structure is not accounted for, because time delays are ensemble-averaged over all possible geometries of the LSS. The anisotropy of the large-scale cosmic web – voids, walls, filaments, and nodes (which shape and orient the tidal tensor beyond its trace) – should therefore be taken into account explicitly, as it impacts mass assembly. Despite of the above-mentioned difficulty in properly describing the connection between galaxies and their host dark matter haloes, this anisotropy should have direct observational signatures in the differential properties of galaxies with respect to the cosmic web at fixed mass and local density. Quantifying these signatures is the topic of this paper. Extending EPS to account for the geometry of the tides beyond that encoded in the density of the field is the topic of the companion paper (Musso et al. 2017).

This paper explores the impact of the cosmic web on galaxy properties in the GAMA survey, using the Discrete Persistent Structure Extractor code (DisPerSE; Sousbie 2011; Sousbie, Pichon & Kawahara 2011) to characterize its 3D topological features, such as nodes, filaments and walls. GAMA is to date the best data set for this kind of study, given its unique spectroscopic combination of depth, area, target density and high completeness, as well as its broad multiwavelength coverage. Variations in stellar mass and colour, red fraction and star formation activity are investigated as a function of galaxy distances to these three features. The rest of the paper is or-

ganized as follows. Section 2 summarizes the data and describes the sample selection. The method used to reconstruct the cosmic web is presented in Section 3. Section 4 investigates the stellar-mass and type/colour segregation and the star formation activity of galaxies within the cosmic web. Section 5 shows how these results compare to those obtained in the HORIZON-AGN simulation (Dubois et al. 2014). Section 6 addresses the impact of the density on the measured gradients towards filaments and walls. Results are discussed in Section 7 jointly with predictions from Musso et al. (2017). Finally, Section 8 concludes. Additional details on the matching technique and the impact of the boundaries to the measured gradients are provided in Appendices A and B, respectively. Appendix C investigates the effect of smoothing scale on the found gradients, Appendix D briefly presents the HORIZON-AGN simulation, Appendix F provides tables of median gradients and a short summary of predicted gradient misalignments is presented in Appendix E.

Throughout the study, a flat ΛCDM cosmology with $H_0 = 67.5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.31$ and $\Omega_\Lambda = 0.69$ is adopted (Planck Collaboration XIII 2016). All statistical errors are computed by bootstrapping, such that the errors on a given statistical quantity correspond to the standard deviation of the distribution of that quantity re-computed in 100 random samples drawn from the parent sample with replacement. All magnitudes are quoted in the AB system, and by log we refer to the 10-based logarithm.

2 DATA AND DATA PRODUCTS

This section describes the observational data and derived products, namely the galaxy and group catalogues, that have been used in this work.

2.1 Galaxy catalogue

The analysis is based on the GAMA survey[1] (Driver et al. 2009, 2011; Hopkins et al. 2013; Liske et al. 2015), a joint European-Australian project combining multiwavelength photometry (UV to far-IR) from ground and space-based facilities and spectroscopy obtained at the Anglo-Australian Telescope (AAT, NSW, Australia) using the AAOmega spectrograph. GAMA provides spectra for galaxies across five regions, but this work only considers the three equatorial fields G9, G12 and G15 covering a total area of $180\,\mathrm{deg^2}$ ($12 \times 5\,\mathrm{deg^2}$ each), for which the spectroscopic completeness is >98 per cent down to a r-band apparent magnitude $m_r = 19.8$. The reader is referred to Wright et al. (2016) for a complete description of the spectro-photometric catalogue constructed using the LAMBDAR[2] code that was applied to the 21-band photometric data set from the GAMA Panchromatic Data Release (Driver et al. 2016), containing imaging spanning the far-UV to the far-IR.

The physical parameters for the galaxy sample such as the absolute magnitudes, extinction corrected rest-frame colours, stellar masses and specific star formation rate (sSFR) are derived using a grid of model spectral energy distributions (SED; Bruzual & Charlot 2003) and the SED fitting code LEPHARE[3] (Arnouts et al. 1999; Ilbert et al. 2006). The details used to derive these physical parameters are given in the companion paper (Treyer et al. in preparation).

---

[1] http://www.gama-survey.org/

[2] Lambda Adaptive Multi-Band Deblending Algorithm in *R*.

[3] http://cesam.lam.fr/lephare/lephare.html
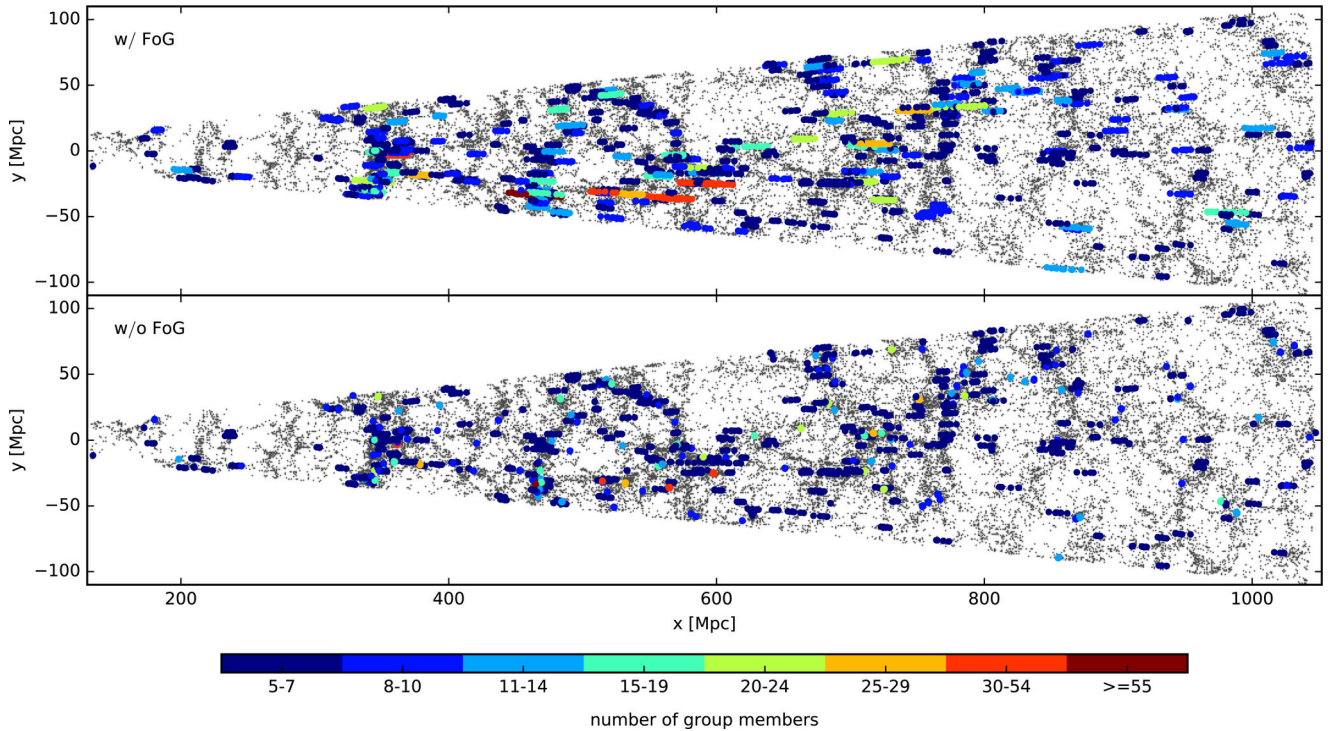
Figure 1. Spatial distribution of whole galaxy population with $m_r < 19.8$ in the GAMA field G12 in the redshift range $0.03 \le z \le 0.25$ (grey points). Overplotted are galaxy group members, colour coded by the size of their group. Only groups having five or more members are shown. The top and bottom panels illustrate the galaxy group members before and after correcting for the FoG effect, respectively.

The classification between the active (star-forming) and passive (quiescent) populations is based on a simple colour cut at $u - r = 1.8$ in the rest-frame extinction corrected $u - r$ versus r diagram that is used to separate the two populations. This colour cut is consistent with a cut in sSFR at $10^{-10.8} \, \mathrm{yr}^{-1}$ (see Treyer et al. in preparation). Hence, in what follows, the terms red (blue) and quiescent (star-forming) will be used interchangeably.

The analysis is restricted to the redshift range $0.03 \le z \le 0.25$, totalling 97 072 galaxies. This is motivated by the high galaxy sampling required to reliably reconstruct the cosmic web. Beyond $z \sim 0.25$, the galaxy number density drops substantially (to $2 \times 10^{-3} \, \mathrm{Mpc}^{-3}$ from $8 \times 10^{-3} \, \mathrm{Mpc}^{-3}$ at $z \le 0.25$, on average), while below $z \sim 0.03$, the small volume does not allow us to explore the large scales of the cosmic web.

The stellar mass completeness limits are defined for the passive and active galaxies as the mass above which 90 per cent of galaxies of a given type (blue/red) reside at a given redshift $z \pm 0.004$. This translates into mass completeness limits of $\log(M_\star/\mathrm{M}_\odot) = 9.92$ and $\log(M_\star/\mathrm{M}_\odot) = 10.46$ for the blue and red populations at $z \le 0.25$, respectively.

2.2 Group catalogue

Since the three-dimensional distribution of galaxies relies on the redshift-based measures of distances, it is affected by their peculiar velocities. In order to optimize the cosmic web reconstruction, one needs to take into account these redshift-space distortions. On large scales, these arise from the coherent motion of galaxies accompanying the growth of structure, causing its flattening along the line of sight, the so-called Kaiser effect (Kaiser 1987). On small scales, the so-called Fingers of God (FoG; Jackson 1972; Tully & Fisher 1978) effect, induced by the random motions of galaxies within virialized

haloes (groups and clusters) causes the apparent elongation of structures in redshift space, clearly visible in the galaxy distribution in the GAMA survey (Fig. 1, top panel). While the Kaiser effect tends to enhance the cosmic web by increasing the contrast of filaments and walls (e.g. Subba Rao et al. 2008; Shi et al. 2016), the FoG effect may lead to the identification of spurious filaments. Because the impact of the Kaiser effect is expected to be much less significant than that of the FoG (e.g. Subba Rao et al. 2008; Kuutma et al. 2017), for the purposes of this work, we do not attempt to correct for it and we focus on the compression of the FoG only. To do so, the galaxy groups are first constructed with the use of an anisotropic Friends-of-Friends (FoF) algorithm operating on the projected perpendicular and parallel separations of galaxies, that was calibrated and tested using the publicly available GAMA mock catalogues of Robotham et al. (2011) (see also Merson et al. 2013, for details of the mock catalogues construction). Details on the construction of the group catalogue and related analysis of group properties can be found in the companion paper (Treyer et al. in preparation). Next, the centre of each group is identified following Robotham et al. (2011) (see also Eke et al. 2004, for a different implementation). The method is based on an iterative approach: first, the centre of mass of the group (CoM) is computed; next its projected distance from the CoM is found iteratively for each galaxy in the group by rejecting the most distant galaxy. This process stops when only two galaxies remain and the most massive galaxy is then identified as the centre of the group. The advantage of this method, as shown in Robotham et al. (2011), is that the iteratively defined centre is less affected by interlopers than luminosity-weighted centre or the central identified as the most luminous group galaxy. The groups are then compressed radially so that the dispersions in transverse and radial directions are equal, making the galaxies in the groups isotropically distributed about their centres (see e.g. Tegmark et al. 2004). In practice, since

the elongated FoG effect affects mostly the largest groups, only groups with more than six members are compressed. Note that the precise correction of the FoG effect is not sought. What is needed for the purpose of this work is the elimination of these elongated structures that could be misidentified as filaments.

Fig. 1 displays the whole galaxy population and the identified FoF groups (coloured by their richness) in the GAMA field G12. The top and bottom panels show the groups before and after correcting for the FoG effect, respectively. For the sake of clarity, only groups having at least five members are shown. The visual inspection reveals that most of the groups are located within dense regions, often at the intersection of the apparently filamentary structures.

3 THE COSMIC WEB EXTRACTION

With the objective of exploring the impact of the LSS on the evolution of galaxy properties, one first needs to properly describe the main components of the cosmic web, namely the high-density peaks (nodes) which are connected by filaments, framing the sheet-like walls, themselves surrounding the void regions. Among the various methods developed over the years, two broad classes can be identified. One uses the geometrical information contained in the *local* gradient and the Hessian of the density or potential field (e.g. Novikov, Colombi & Doré 2006; Aragón-Calvo et al. 2007a,b; Hahn et al. 2007a,b; Sousbie et al. 2008a,b; Forero-Romero et al. 2009; Bond, Strauss & Cen 2010a,b), while the second exploits the topology and connectivity of the density field by using the watershed transform (Aragón-Calvo, van de Weygaert & Jones 2010) or Morse theory (e.g. Colombi, Pogosyan & Souradeep 2000; Sousbie et al. 2008a; Sousbie 2011). The theory for the former can be built in some details, (see e.g. Pogosyan et al. 2009), shedding some light on physical interpretation, while the latter avoids shortcomings of a second-order Taylor expansion of the field and provides a natural metric in which to compute distances to filaments. Within these broad categories, some algorithms deal with discrete data sets, while others require that the density field must be first estimated (possibly on multiple scales). An exhaustive description of several cosmic web extraction techniques and a comparison of their classification patterns as measured in simulations are presented in Libeskind et al. (2017). While this paper found some differences between the various algorithms, which should in principle be accounted for as modelling errors in this work, these differences remain small on the scales considered.

3.1 Cosmic web with disperse

This work uses the Discrete Persistent Structure Extractor (DisPerSE; see Sousbie et al. 2011, for illustrations in a cosmological context), a geometric three-dimensional ridge extractor dealing directly with discrete data sets, making it particularly well adapted for astrophysical applications. It allows for a scale and parameter-free coherent identification of the 3D structures of the cosmic web as dictated by the large-scale topology. For a detailed description of the DisPerSE algorithm and its underlying theory, the reader is referred to Sousbie (2011); its main features are summarized below.

DisPerSE is based on discrete Morse and persistence theories. The Delaunay tessellation is used to generate a simplicial complex, i.e. a triangulated space with a geometric assembly of cells, faces, edges and vertices mapping the whole volume. The Delaunay Tessellation Field Estimator (DTFE; Schaap & van de Weygaert 2000; Cautun & van de Weygaert 2011) allows for estimating the density

field at each vertex of the Delaunay complex. The Morse theory enables to extract from the density field the critical points, i.e. points with a vanishing (discrete) gradient of the density field (e.g. maxima, minima and saddle points). These critical points are connected via the field lines tangent to the gradient field in every point. They induce a geometrical segmentation of space, where all the field lines have the same origin and destination, known as the Morse complex. This segmentation defines distinct regions called ascending and descending k-manifolds.[4] The morphological components of the cosmic web are then identified from these manifolds: ascending 0-manifolds trace the voids, ascending 1-manifolds trace the walls and filaments correspond to the ascending 2-manifolds with their extremities plugged on to the maxima (peaks of the density field). In addition to its ability to work with sparsely sampled data sets while assuming nothing about the geometry or homogeneity of the survey, DisPerSE allows for the selection of retained structures on the basis of the significance of the topological connection between critical points. DisPerSE relies on persistent homology theory to pair critical points according to the birth and death of a topological feature in the excursion. The 'persistence' of a feature or its significance is assessed by the density contrast of the critical pair chosen to pass a certain signal-to-noise threshold. The noise level is defined relative to the RMS of persistence values obtained from random sets of points. This thresholding eliminates less significant critical pairs, allowing to simplify the Morse complex, retaining its most topologically robust features. Fig. 2 shows that filaments outskirt walls, themselves circumventing voids. The filaments are made of a set of connected segments and their end points are connected to the maxima, the peaks of the density field where most of clusters and large groups reside. Each wall is composed of the facets of tetrahedra from the Delaunay tessellation belonging to the same ascending 2-manifold. In this work, DisPerSE is run on the flux-limited GAMA data with a 3σ persistence threshold. Fig. 3 illustrates the filaments for the G12 field, overplotted on the density contrast of the underlying galaxy distribution, $1 + \delta$, where the local density is estimated using the DTFE density estimator. Even in this 2D projected visualization, one can see that filaments trace the ridges of the 3D density field connecting the density peaks between them.

3.2 Cosmic web metric

Having identified the major cosmic web features, let us now define a new metric to characterize the environment of a galaxy, which will be referred to as the 'cosmic web metric' and into which galaxies are projected. Fig. 4 gives a schematic view of this framework. Each galaxy is assigned the distance to its closest filament, D_{skel}. The impact point in the filament is then used to define the distances along the filament towards the node, D_{node} and towards the saddle point, D_{saddle}. Similarly, D_{wall} denotes the distance of the galaxy to its closest wall. In this work, only distances D_{node}, D_{skel} and D_{wall} are used. Other investigations of the environment in the vicinity of the saddle points are postponed to a forthcoming work.

The accuracy of the reconstruction of the cosmic web features is sensitive to the sampling of the data set. The lower the

---

[4] The index k refers to the critical point the field lines emanate from (ascending) or converge to (descending), and is defined as its number of negative eigenvalues of the Hessian: a minimum of the field has index 0, a maximum has index 3 and the two types of saddles have indices 1 and 2.
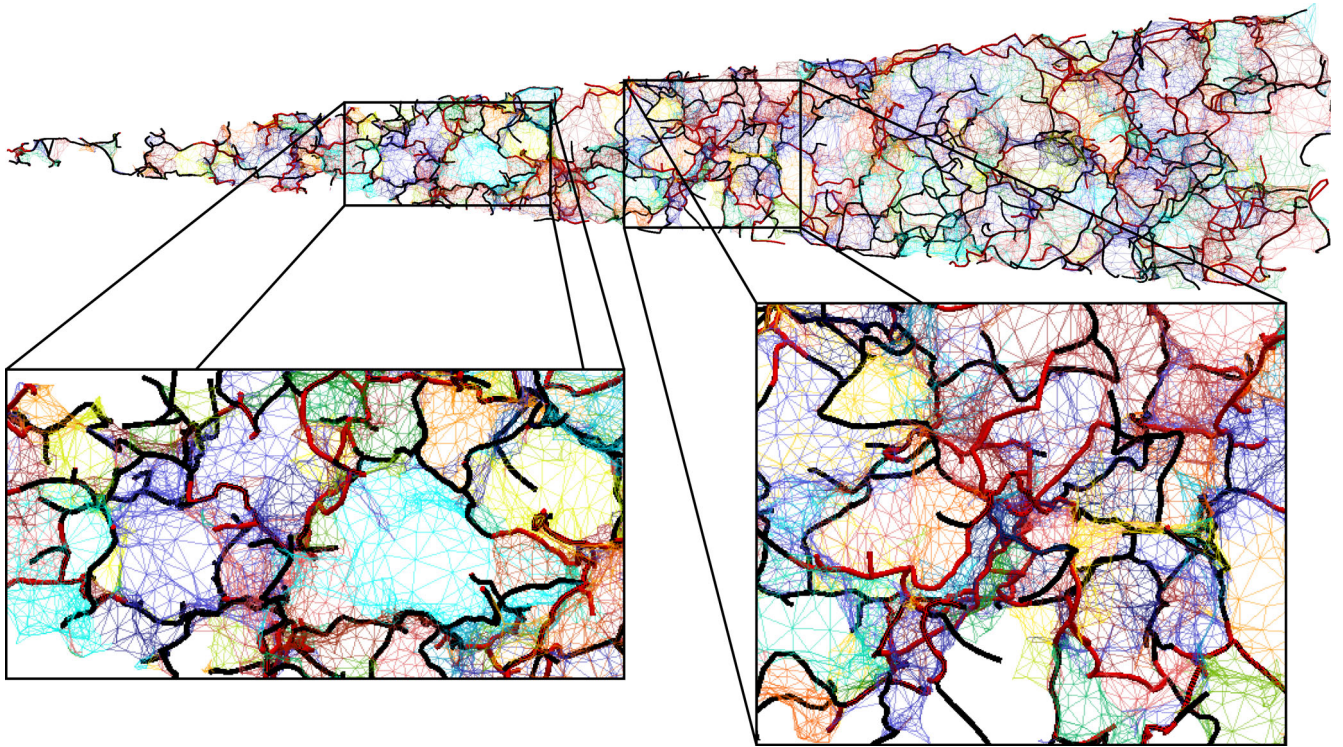
Figure 2. Illustration of the walls and filaments in the G12 field. For the sake of clarity and for the illustrative purposes, only the cosmic web features detected above a persistence threshold of 5σ are shown. Filaments are coloured in black, with the most persistent ones ($>6\sigma$) plotted in red, while walls are colour coded randomly. Note how DisPerSE is capable of recovering the important features of the underlying cosmic field by identifying its (topologically) most-robust features. In particular, it extracts filaments as a set of connected segments, which outskirt walls, themselves circumventing voids.
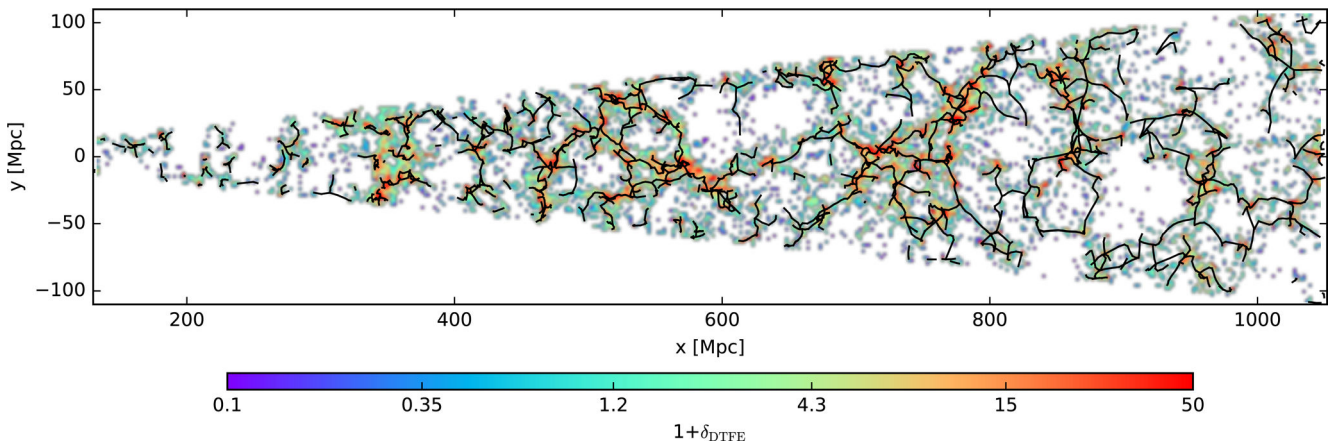
Figure 3. Illustration of the filamentary network (black lines) extracted with the DisPerSE code within the $\pm 1.2°$ of the central declination of the G12 field. The persistence threshold with which the filamentary network and the associated structures, used in this work and shown here, are extracted is 3σ. Also shown is the density contrast of the underlying galaxy distribution, measured with the small-scale adaptive DTFE estimator (see the text) and averaged over cells of $2.3 \times 2.3 \, \mathrm{Mpc}^2$ (white colour is used for empty cells). In spite of the projection effects, the visual inspection reveals that filaments follow the ridges of the density field which connect the peaks together.

sampling the larger the uncertainty on the location of the individual components of the cosmic web. To account for the variation of the sampling throughout the survey, unless stated differently, all the distances are normalized by the redshift-dependent mean inter-galaxy separation $\langle D_z \rangle$, defined as $\langle D_z \rangle \equiv n(z)^{-1/3}$, where $n(z)$ represents the number density of galaxies at a given redshift z. For the combined three fields of GAMA survey, $\langle D_z \rangle$ varies from 3.5 to 7.7 Mpc across the redshift range $0.03 \leq z \leq 0.25$, with a mean value of ~ 5.6 Mpc.

4 GALAXY PROPERTIES WITHIN THE COSMIC WEB

In this section, the dependence of various galaxy properties, such as stellar mass, $u - r$ colour, sSFR and type, with respect to their location within the cosmic web is analysed. First, the impact of the nodes, representing the largest density peaks, is investigated. Next, by excluding these regions, galaxy properties are studied within the intermediate density regions near the filaments. Finally, the analysis is extended to the walls.
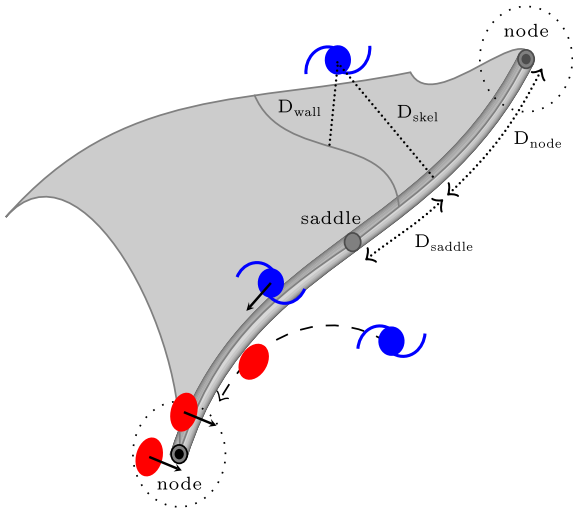
Figure 4. Schematic view of the 'cosmic web' metric in which the analysis is performed. The position of a galaxy within the cosmic web is parametrized by its distance to the closest filament, D_{skel}, and its distance to the closest wall, D_{wall}. D_{node} and D_{saddle} represent the distances from the impact point to the node and saddle along the corresponding filament, respectively.

4.1 The role of nodes via the red fractions

Let us start by analysing the combined impact of nodes and filaments on galaxies through the study of the red fractions. The red fraction, defined as the number of passive galaxies with respect to the entire population, is analysed as a function of the distance to the nearest filament, D_{skel} and the distance to its associated node, D_{node}.

This analysis is restricted to galaxies more massive than $\log(M_\star/\mathrm{M}_\odot) \geq 10.46$, as imposed by the mass limit completeness of the passive population (see Section 2). The stellar mass distributions of the passive and star-forming populations are not identical, with the passive galaxies dominating the high mass end. Therefore, to prevent biases in the measured gradients introduced by such differences, the mass-matched samples are used. The detailed description of the mass-matching technique can be found in Appendix A1.

In Fig. 5 the red fraction of galaxies is shown as a function of D_{skel} in three different bins of D_{node}. While the fraction of passive galaxies is found to increase with decreasing distances to both the filaments and nodes, the dominant effect is the distance to the nodes. At fixed D_{skel}, the fraction of passive galaxies sharply increases with decreasing distance to the nodes. Recalling that the mean inter-galaxy separation $\langle D_z \rangle \sim 5.6\,\mathrm{Mpc}$, a 20–30 per cent increase in the fraction of passive galaxies is observed from several Mpc away from the nodes to less than ~500 kpc. This behaviour is expected since the nodes represent the loci where most of the groups and clusters reside and reflect the well-known colour-density (e.g. Blanton et al. 2003; Baldry et al. 2006; Bamford et al. 2009) and star formation-density (e.g. Lewis et al. 2002; Kauffmann et al. 2004) relations. However, the gradual increase suggests that some physical processes already operate before the galaxies reach the virial radius of massive haloes. At fixed D_{node}, the fraction of passive galaxies increases with decreasing distance to filaments, but this increase is milder compared to that with respect to nodes: an increase of ~10 per cent is observed regardless of the distance to the node. These regions with intermediate densities appear to be a place where the transformation of galaxies takes place as emphasized in the next section.
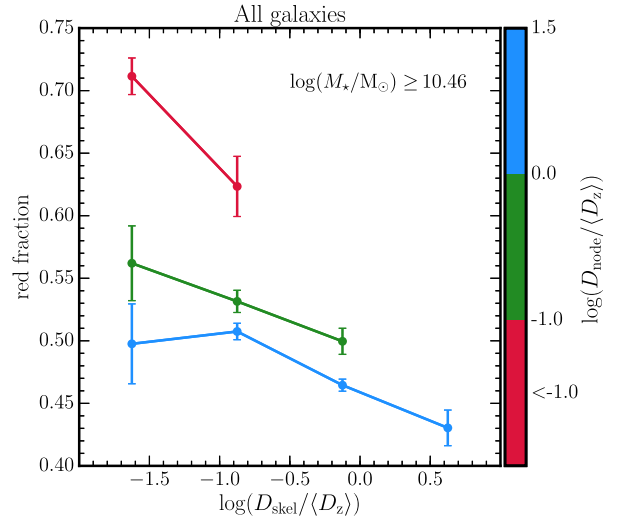
Figure 5. Red fraction of galaxies (the number of quiescent galaxies over the entire population) as a function of D_{skel} for three different bins of D_{node} as indicated by the colour. Both distances are normalized by the redshift-dependent mean inter-galaxy separation $\langle D_z \rangle$. Only galaxies with $\log(M_\star/\mathrm{M}_\odot) \geq 10.46$ are considered. Star-forming and quiescent populations are matched in mass (see Section 4.2.1). The error bars are calculated from 100 bootstrap samples. The fraction of red galaxies is found to increase with decreasing distances both to the closest filament D_{skel} and to the node of this D_{node}. Recalling that $\langle D_z \rangle \sim 5.6\,\mathrm{Mpc}$, the fraction of passive galaxies increases at given D_{skel} by ~20 per cent from several tens of Mpc away from the nodes (blue line) to less than ~0.5 Mpc (red line). At fixed D_{node}, the increase of the red fraction with decreasing distance to filaments is milder, of ~10 per cent, regardless of the distance to the node.

4.2 The role of filaments

In order to infer the role played by filaments alone in the transformation of galactic properties, the impact of nodes, the high-density regions has to be mitigated. By construction, nodes are at the intersection of filaments: they drive the well-known galaxy type-density as well as stellar mass-density relations. To account for this bias, Gay et al. (2010) and Malavasi et al. (2017) adopted a method where a given physical property or distance of each galaxy was down-weighted by its local density. Laigle et al. (2017) adopted a more stringent approach by rejecting all galaxies that are too close to the nodes. This method allows us to minimize the impact of nodes, avoiding the difficult-to-quantify uncertainty of the residual contribution of the density weighting scheme. We therefore adopt the latter approach. As shown in Appendix B1, this is achieved by rejecting all galaxies below a distance of 3.5 Mpc from a node.

4.2.1 Stellar mass gradients

Fig. 6 shows the normalized probability distribution functions (PDFs) of the distance to the nearest filament D_{skel} in three stellar mass bins for the entire population and star-forming galaxies alone (top left-hand and right-hand panels, respectively). The medians of the PDFs, shown by vertical lines, are listed together with the corresponding error bars in Table 1. The significance of the observed trends is assessed by computing the residuals between the distributions in units of σ (bottom panels), defined as $\Delta_{1-2}/\sqrt{\sigma_1^2 + \sigma_2^2}$, where Δ_{1-2} is the difference between the PDFs of the populations 1 and 2, and σ_1 and σ_2 are the corresponding standard deviations.

For the entire population (left-hand panels), differences between the PDFs of the three stellar mass bins are observed: the most
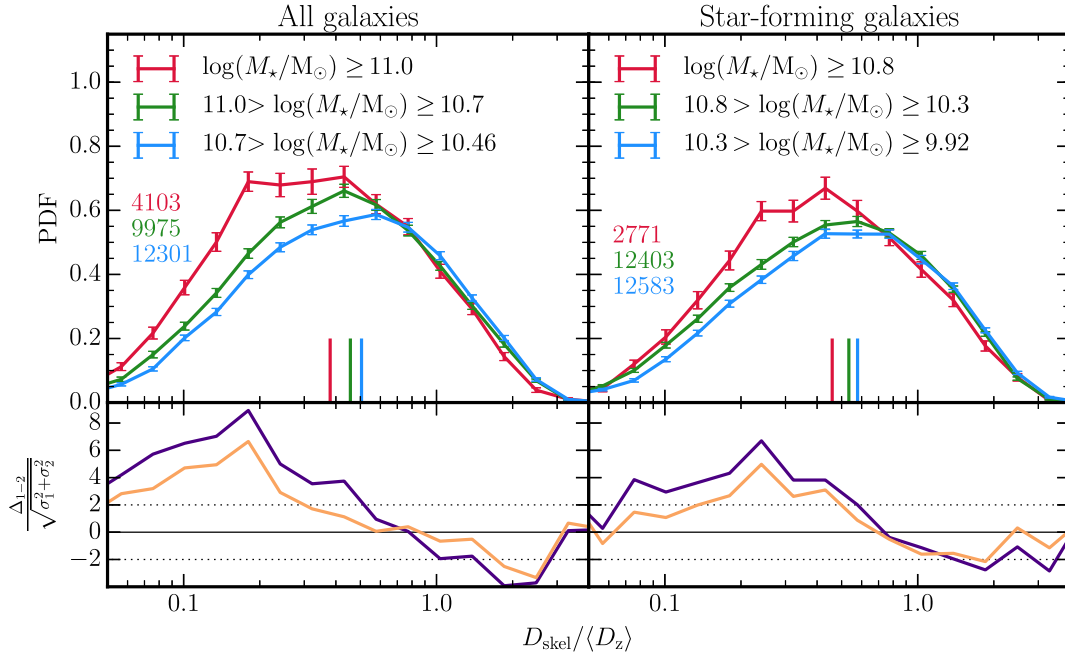
Figure 6. *Top row:* Differential distributions of the distances to the nearest filament, $D_{\rm skel}$ (normalized by $\langle D_z \rangle$, the redshift-dependent mean inter-galaxy separation) for the entire galaxy population (left-hand panel) and star-forming galaxies alone (right-hand panel) in three different stellar mass bins. Note that these bins are different for the two populations: this is due to the stellar mass completeness limit that is different (see Section 2). To highlight an effect specific to the filaments, the contribution of node is minimized (see the text for details). The vertical lines indicate the medians of the distributions and their values together with associated error bars are listed in Table 1. The numbers of galaxies in different considered bins are indicated in each panel. The error bars are calculated from 100 bootstrap samples. There is a mass segregation of galaxies with respect to filaments of the entire as well as star-forming population: more massive galaxies tend to be preferentially located closer to the filaments compared to their lower-mass counterparts. *Bottom row:* Residuals in units of σ between the two most extreme mass bins (purple line; $10.7 > \log{(M_\star/{\rm M}_\odot)} \geq 10.46$ and $\log{(M_\star/{\rm M}_\odot)} \geq 11.0$ on the left-hand panel and $10.3 > \log{(M_\star/{\rm M}_\odot)} \geq 9.92$ and $\log{(M_\star/{\rm M}_\odot)} \geq 10.8$ on the right-hand panel), and between the high and intermediate mass bins (orange solid line; $\log{(M_\star/{\rm M}_\odot)} \geq 11.0$ and $11.0 > \log{(M_\star/{\rm M}_\odot)} \geq 10.7$ on the left-hand panel and $\log{(M_\star/{\rm M}_\odot)} \geq 10.8$ and $10.8 > \log{(M_\star/{\rm M}_\odot)} \geq 10.3$ on the right-hand panel).

Table 1. Medians for the PDFs displayed in Figs 6–10.

Selection[a]	Bin	Median[b]	
		$D_{\rm skel}/\langle D_z \rangle$	$D_{\rm wall}/\langle D_z \rangle$
Mass[c]	All galaxies $\log{(M_\star/{\rm M}_\odot)} \geq 11$	0.379 ± 0.009	0.334 ± 0.005
	$11 > \log{(M_\star/{\rm M}_\odot)} \geq 10.7$	0.456 ± 0.007	0.381 ± 0.004
	$10.7 > \log{(M_\star/{\rm M}_\odot)} \geq 10.46$	0.505 ± 0.006	0.403 ± 0.004
	SF galaxies $\log{(M_\star/{\rm M}_\odot)} \geq 11$	0.459 ± 0.012	0.385 ± 0.011
	$11 > \log{(M_\star/{\rm M}_\odot)} \geq 10.4$	0.534 ± 0.007	0.429 ± 0.006
	$10.4 > \log{(M_\star/{\rm M}_\odot)} \geq 9.92$	0.578 ± 0.007	0.453 ± 0.007
Type[d]	SF versus passive[e] Star-forming	0.504 ± 0.008	0.411 ± 0.006
	Passive	0.462 ± 0.007	0.376 ± 0.006

Notes. [a]Panels of Figs 6–10.
[b]Medians of distributions as indicated in Figs 6–10 by vertical lines; errors represent half width at half-maximum of the bootstrap distribution, i.e. the distribution of medians from each of 100 bootstrap samples, fitted by a Gaussian curve.
[c]Figs 6 and 9.
[d]Figs 7 and 10.
[e]Only galaxies with stellar masses $\log{(M_\star/{\rm M}_\odot)} \geq 10.46$ are considered.

massive galaxies ($\log{(M_\star/{\rm M}_\odot)} \geq 11$) are located closer to the filaments than the intermediate population ($11 > \log{(M_\star/{\rm M}_\odot)} \geq 10.7$), while the population with the lowest stellar masses ($10.7 > \log{(M_\star/{\rm M}_\odot)} \geq 10.46$) is found furthest away from the filaments. The significances of the difference between the most massive and the two lowest stellar mass bins are shown in the bottom panel. Between the most extreme stellar mass bins (purple line), the difference exceeds 4σ close to the filament and 2σ at

larger distances. It is slightly less significant between the intermediate and lowest stella mass bins (orange line), but still in excess of 2σ close to the filament. The differences between the PDFs can be also quantified in terms of their medians, where the differences between the highest and lowest stellar mass bins are significant at an $\sim 10\sigma$ level (see Table 1). These results confirm previous claims of a mass segregation with respect to filaments, where the most massive galaxies are located near the core of the filaments, while
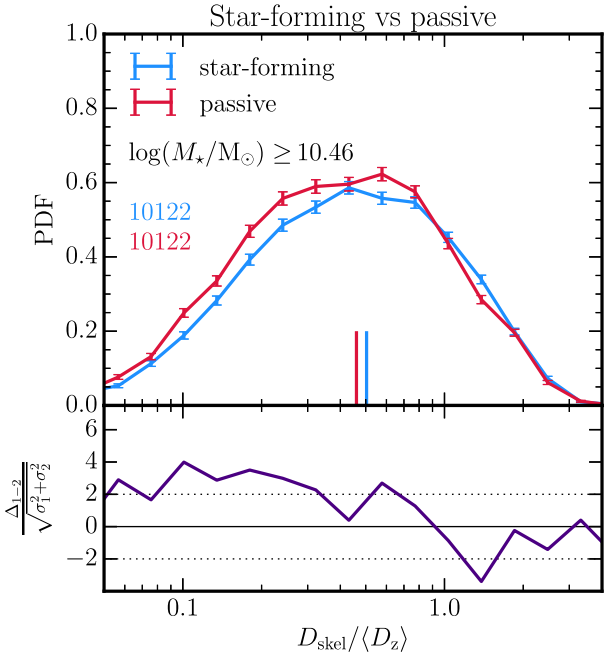
Figure 7. *Top:* Differential distributions of the distances to the nearest filament, D_{skel} (normalized by $\langle D_z \rangle$, the redshift-dependent mean inter-galaxy separation) for star-forming and quiescent galaxies that have been matched in mass (see the text for details). To highlight an effect specific to the filaments, the contribution of node is minimized (see the text for details). The vertical lines indicate the medians of the distributions and their values, together with associated error bars, are listed in Table 1. The numbers of galaxies in different considered bins are indicated in each panel. The error bars are calculated from 100 bootstrap samples. Galaxies are found to segregate, relative to filaments, according to their type: quiescent galaxies tend to be preferentially located closer to the filaments compared to their star-forming counterparts. *Bottom:* Residuals in units of σ between the star-forming and passive galaxies.

the less massive ones tend to reside preferentially on their outskirts (Laigle et al. 2017; Malavasi et al. 2017). As the impact of the nodes has been minimized, it is therefore established that this stellar mass gradient is driven by the filaments themselves and not by the densest regions of the cosmic web.

The mass segregation is also found among the star-forming population alone (right-hand panels), such that more massive star-forming galaxies tend to be closer to the geometric core of the filament than their less massive counterparts. Note that the mass bins for star-forming galaxies differ from mass bins used for the entire population. The completeness stellar mass limit allows us to decrease the lowest mass bin to $\log(M_\star/M_\odot) = 9.92$ when considering the star-forming galaxies alone (see Section 2). The significance of these stellar mass gradients between the extreme stellar mass bins exceeds 4σ near the filaments, while the difference of the medians reaches an $\sim 8\sigma$ level (see Table 1).

4.2.2 Type gradients

Let us now investigate the impact of the filamentary network on the type/colour of galaxies. To do so, galaxies are split by type between star-forming and passive galaxies based on the dust corrected $u - r$ colour as discussed in Section 2.1. As for the analysis of the red fraction (Section 4.1), the sample is restricted to galaxies with $\log(M_\star/M_\odot) \geq 10.46$ and the star-forming and passive populations are matched in stellar mass. Fig. 7 shows the PDFs of
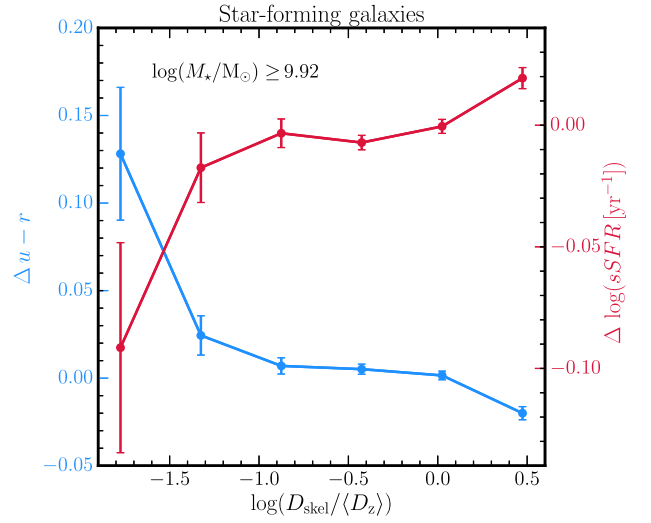
Figure 8. $u - r$ colour (blue line) and sSFR (red line) of star-forming galaxies as a function of D_{skel}. The y-axes indicate the amount by which $u - r$ colour and sSFR differ from the median values at given mass (see the text for details). Only galaxies with $\log(M_\star/M_\odot) \geq 9.92$ and far-away from nodes (at $D_{node} > 3.5$ Mpc) are considered. Star-forming galaxies tend to have higher $u - r$ colour (tend to be redder) and lower sSFR when they get closer to the filaments than their more distant counterparts.

the normalized distances D_{skel} within the mass-matched samples of star-forming and passive populations, which by construction have the same number of galaxies. Galaxies are found to segregate according to their type such that passive galaxies tend to reside in regions located closer to the core of filaments than their star-forming counterparts. The significance of the type gradients between the two populations exceeds 3σ near filaments while the difference between the medians reaches an $\sim 4\sigma$ level (see Table 1).

4.2.3 Star formation activity gradients

To explore whether the impact of filaments on the star formation activity of galaxies can be detected beyond the red fractions and type segregation reported above, the focus is now on the star-forming population alone through the study of their (dust corrected) $u - r$ colour and sSFR.

Both these quantities are known to evolve with stellar mass which itself varies within the cosmic web (see above). To remove this mass dependence, the offsets of $u - r$ colour and sSFR, $\Delta u - r$ and ΔsSFR, respectively, from the median values of all star-forming galaxies at a given mass are computed for each galaxy. Fig. 8 shows the medians of $\Delta u - r$ and ΔsSFR as a function of D_{skel}. Both quantities are found to carry the imprint of the large-scale environment. At large distances from the filaments ($D_{skel} \geq 5$ Mpc), star-forming galaxies are found to be more active than the average. At intermediate distances ($0.5 \leq D_{skel} \leq 5$ Mpc), star formation activity of star-forming galaxies does not seem to evolve with the distance to the filaments, while in the close vicinity of the filaments ($D_{skel} \leq 0.5$ Mpc), they show signs of a decrease in star formation efficiency (redder colour and lower sSFR). The significance of these results will be discussed in Section 7.

4.3 The role of walls in mass and type gradients

Let us now investigate the impact of walls on galaxy properties. Figs 9 and 10 show the PDFs of the distances to the closest wall D_{wall} for the same selections as in Figs 6 and 7, respectively. The distances

All galaxies

Star-forming galaxies
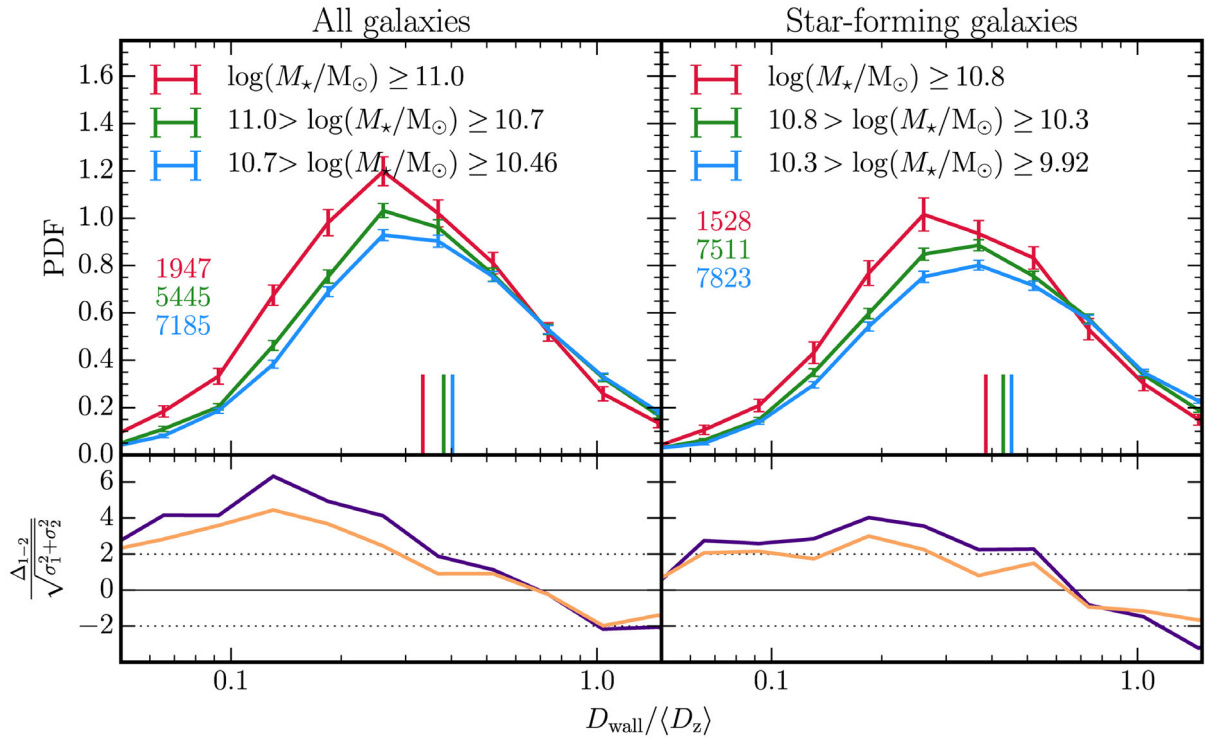
Figure 9. *Top row:* As in Fig. 6, but for the distances to the nearest wall, D_{wall}. To minimize the contribution of nodes and filaments to the measured signal, galaxies located closer to a node than 3.5 Mpc and closer to a filament than 2.5 Mpc are removed form the analysis. There is a mass segregation of galaxies with respect to walls of the entire as well as star-forming population: more massive galaxies tend to be preferentially located closer to the filaments compared to their lower-mass counterparts. *Bottom row:* Residuals in units of σ as in Fig. 6.

Star-forming vs passive
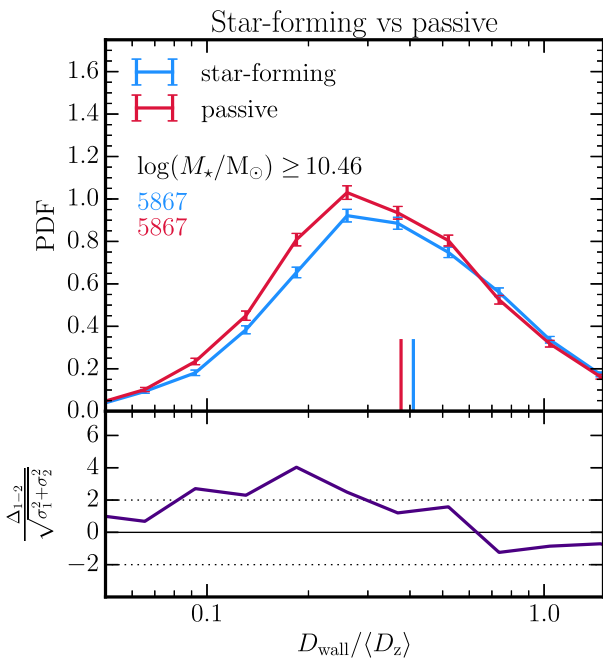
Figure 10. *Top row:* As in Fig. 7, but for the distances to the nearest wall, D_{wall}. To minimize the contribution of nodes and filaments to the measured signal, galaxies located closer to a node than 3.5 Mpc and closer to a filament than 2.5 Mpc are removed from the analysis. Galaxies are found to segregate, with respect to walls, according to their type: quiescent galaxies tend to be preferentially located closer to the walls compared to their star-forming counterparts. *Bottom row:* Residuals in units of σ as in Fig. 7.

are again normalized by the redshift-dependent mean inter-galaxy separation $\langle D_z \rangle$. The values of medians with corresponding error bars are listed in Table 1. As for filaments, one seeks signatures induced by a particular environment solely, walls in this case. Given that filaments are located at the intersections between walls, in addition to the contamination by nodes, which is of concern for filaments, one has to make sure that the contribution of filaments themselves is minimized as well. Following the method adopted in Section 4.2.1, Appendix B2 shows that this can be achieved by removing from the analysis galaxies having distances to the nodes smaller than 3.5 Mpc and distances to the closest filaments less than 2.5 Mpc.

The derived trends are qualitatively similar to those measured with respect to filaments. Massive galaxies are located closer to walls compared to their low-mass counterparts; star-forming galaxies preferentially reside in the outer regions of walls; and mass segregation is present also among star-forming population of galaxies with more massive star-forming galaxies having smaller distances to the walls than low-mass counterparts. Since these walls typically embed smaller-scale filaments, the net effect of transverse gradients perpendicular to these filaments should add up to transverse gradients perpendicular to walls.

The significance of the measured trends, in terms of the residuals between medians (see Table 1), is above 3σ for all considered gradients, slightly lower than for the gradients towards filaments. The deviations of $\sim 10\sigma$ and $\sim 5\sigma$ are detected between the highest and lowest stellar mass bins among the whole and star-forming population alone, respectively, while between the star-forming and passive galaxies it reaches $\sim 4\sigma$, as in the case of gradients towards filaments.

Table 2. Medians for the PDFs displayed in Fig. 11.

Selection[a]	Bin	Median[b]	
		D_{skel} (Mpc)	D_{wall} (Mpc)
Mass	$\log(M_\star/\text{M}_\odot) \geq 10.8$	1.34 ± 0.09	0.79 ± 0.04
	$10.8 > \log(M_\star/\text{M}_\odot) \geq 10.4$	1.73 ± 0.08	1.14 ± 0.03
	$10.4 > \log(M_\star/\text{M}_\odot) \geq 10$	1.97 ± 0.04	1.22 ± 0.02
sSFR[c]	$-10.8 > \log(sSFR/\text{yr}^{-1})$	1.46 ± 0.07	1.02 ± 0.03
	$-10.4 > \log(sSFR/\text{yr}^{-1}) \geq -10.8$	1.88 ± 0.06	1.18 ± 0.03
	$\log(sSFR/\text{yr}^{-1}) \geq -10.4$	2.0 ± 0.04	1.18 ± 0.02

Notes. [a]Panels of Fig. 11.
[b]Medians of distributions as indicated in Fig. 11 by vertical lines; errors are computed as in Table 1.
[c]Only galaxies with stellar masses $\log(M_\star/\text{M}_\odot) \geq 10$ are considered.

5 COMPARISON WITH THE HORIZON-AGN SIMULATION

In this section, a qualitative support for the results on the mass and star-formation activity segregation is provided via the analysis of the large-scale cosmological hydrodynamical simulation HORIZON-AGN (Dubois et al. 2014). Note that the main purpose of such an analysis is to provide a reference measurement of gradients in the context of a large-scale 'full physics' experiment. The construction of the GAMA-like mock catalogue is not performed because the geometry of HORIZON-AGN does not allow us to recover the entire GAMA volume and the flux-limited sample requires a precise modelling of fluxes in different bands.

A brief summary of some of the main features of the simulation can be found in Appendix D. Here, the results on the mass and sSFR gradients towards filaments and walls are presented. The HORIZON-AGN simulation is analysed at low redshift ($z \sim 0.1$), comparable to the mean redshift studied in this paper, and the same analysis is performed as in the GAMA data. The filamentary network and associated structures are extracted by running the DisPerSE code with the persistence threshold of 3σ.

Fig. 11 shows the mass (left-hand panels) and sSFR (right-hand panels) gradients towards filaments (figure a) and walls (figure b) as measured in the HORIZON-AGN simulation. The impact of the nodes and filaments on the measured signal is minimized by removing from the analysis galaxies that are closer to the node than 3.5 Mpc and closer to the filament than 1 Mpc. The detailed description of the method used to identify these cuts in distances can be found in Appendix B1. Consistently with the measurements in GAMA, galaxies in HORIZON-AGN are found to segregate by stellar mass, with more massive galaxies being preferentially closer to both the filaments and walls than their low-mass counterparts. Similarly, the presence of the sSFR gradient, whereby less star-forming galaxies tend to be closer to the cores of filaments and walls than their more star-forming counterparts, is in qualitative agreement with the type/colour gradients detected in the GAMA survey. Note that the three bins of sSFR are used to separate out the highly star-forming galaxies, with $\log(sSFR/\text{yr}^{-1}) \geq -10.4$, from passive ones, with $\log(sSFR/\text{yr}^{-1}) < -10.8$, in order to compare with the type gradients in the observations. In the simulation, sSFR is a more reliable parameter for type than for the colour.

The significance of the trends is measured, as previously, in terms of the residuals between medians (see Table 1). For the gradients towards filaments, the difference of $\gtrsim 6\sigma$ is found between the most extreme, both mass and sSFR, bins, while it drops to ~ 2–3σ between the intermediate and lowest bins. For the gradients towards walls, the deviation between the most extreme bins is ~ 10 and 4σ for mass and sSFR bins, respectively, while there is only a little to no

difference between intermediate and lowest stellar mass and sSFR bins, respectively. The gradients are slightly less significant than in the GAMA measurements, most likely due to the low numbers of galaxies per individual bins in HORIZON-AGN, but qualitatively similar as in GAMA.

6 THE RELATIVE IMPACT OF DENSITY

Let us now address the following questions: what is the specific role of the geometry of the large-scale environment in establishing mass and type/colour large-scale gradients? Are these gradients driven solely by density, or does the large-scale anisotropy of the cosmic web provide a specific signature?

A key ingredient in answering these questions is the choice of the scale at which the density is inferred. The properties of galaxies at a given redshift are naturally a signature of their past light-cone. This light-cone in turn correlates with the galaxy's environment: the larger the scale is, the longer the look-back time one must consider, the more integrated the net effect of this environment. This past environment accounts for the total accreted mass of the galaxy, but may also impact the geometry of the accretion history and more generally other galactic properties such as its star formation efficiency, its colour or its spin. At small scales, the density correlates with the most recent and stochastic processes, while going to larger scales allows taking the integrated hence smoother history of galaxies into account. Since this study is concerned about the statistical impact of the large-scale structure on galaxies, it is natural to consider scales large enough to average out local recent events they may have encountered, such as binary interactions, mergers and outflows. Therefore in the discussion below, the density is computed at the scale of 8 Mpc, the 'smallest' scale at which the effect of the anisotropic large-scale tides can be detected.

In practice, in order to try to disentangle the effect of density from that of the anisotropic large-scale tides, the following reshuffling method (e.g. Malavasi et al. 2017) is adopted. For mass gradients, 10 equipopulated density bins are constructed and in each of them the stellar masses of galaxies are randomly permuted. By construction, the underlying mass-density relation is preserved, but this procedure randomizes the relation between the stellar mass and the distance to the filament or the wall. For the type/colour gradients, in each of 10 equipopulated density bins, 10 equipopulated stellar mass bins are constructed. Within each of such bins, $u - r$ colour of galaxies are randomly permuted. Thus by construction, this preserves the underlying colour-(mass)-density relation, but breaks the relation between the colour/type and the distance to the particular environment, the filament or wall.
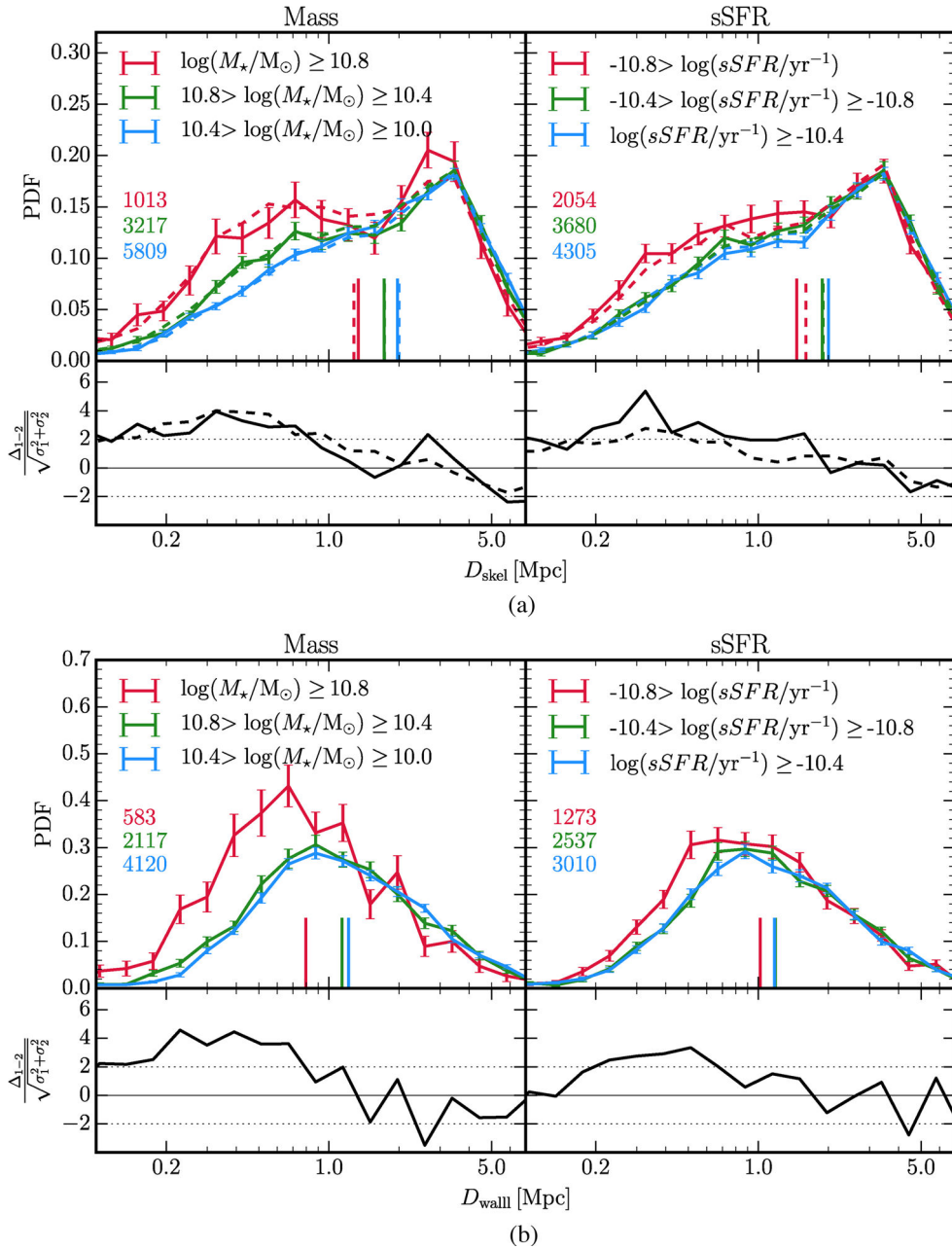
Figure 11. *Top rows:* Differential distributions of the distances as a function of stellar mass (left-hand panels) and sSFR (right-hand panels) for galaxies in HORIZON-AGN. To minimize the contribution of nodes and filaments to the measured signal, galaxies located closer to a node than 3.5 Mpc and closer to a filament than 1 Mpc are removed from the analysis. The vertical lines indicate the medians of the distributions (see Table 2 for the numerical values). Numbers of galaxies in different considered bins are indicated in each panel. There is mass and sSFR segregation of galaxies with respect to both filaments and walls: more massive and less star-forming galaxies tend to be preferentially located closer to the cores of filaments and walls compared to their lower-mass and more star-forming counterparts, respectively. These results are in qualitative agreement with the measurements in GAMA. *Bottom rows:* Residuals in units of σ between the two most extreme mass and sSFR bins, $\log(M_\star/M_\odot) \geq 10.8$ and $10.4 > \log(M_\star/M_\odot) \geq 10$ on the left-hand panel and $-10.8 > \log(sSFR/yr^{-1})$ and $\log(sSFR/yr^{-1}) \geq -10.4$ on the right-hand panel, respectively. (a) Differential distributions of the distances to the nearest filament, D_{skel}. (b) Differential distributions of the distances to the nearest wall, D_{wall}.

In order to account for the variation of the density through the survey, the density contrast, defined as $1 + \delta = n/n(z)$, where $n(z)$ corresponds to the mean redshift-dependent number density, is used in logarithmic bins. The number density n is computed in the Gaussian kernel and every time five reshuffled samples are constructed.

In Fig. 12(a), the mass and type gradients towards filaments, as measured in GAMA and previously shown in Figs 6 and 7, are compared with the outcome of the reshuffling technique. The original signal is found to be substantially reduced after the reshuffling of masses and colours of galaxies. For the mass gradients, the deviation between the most extreme bins before reshuffling exceeds 3σ, while after the reshuffling, the signal gets reduced, with typical deviations of $\sim 1\sigma$. The original signal for the type/colour gradients is weaker than in the case of the mass gradients, however it is similarly nearly cancelled out once the reshuffling method is
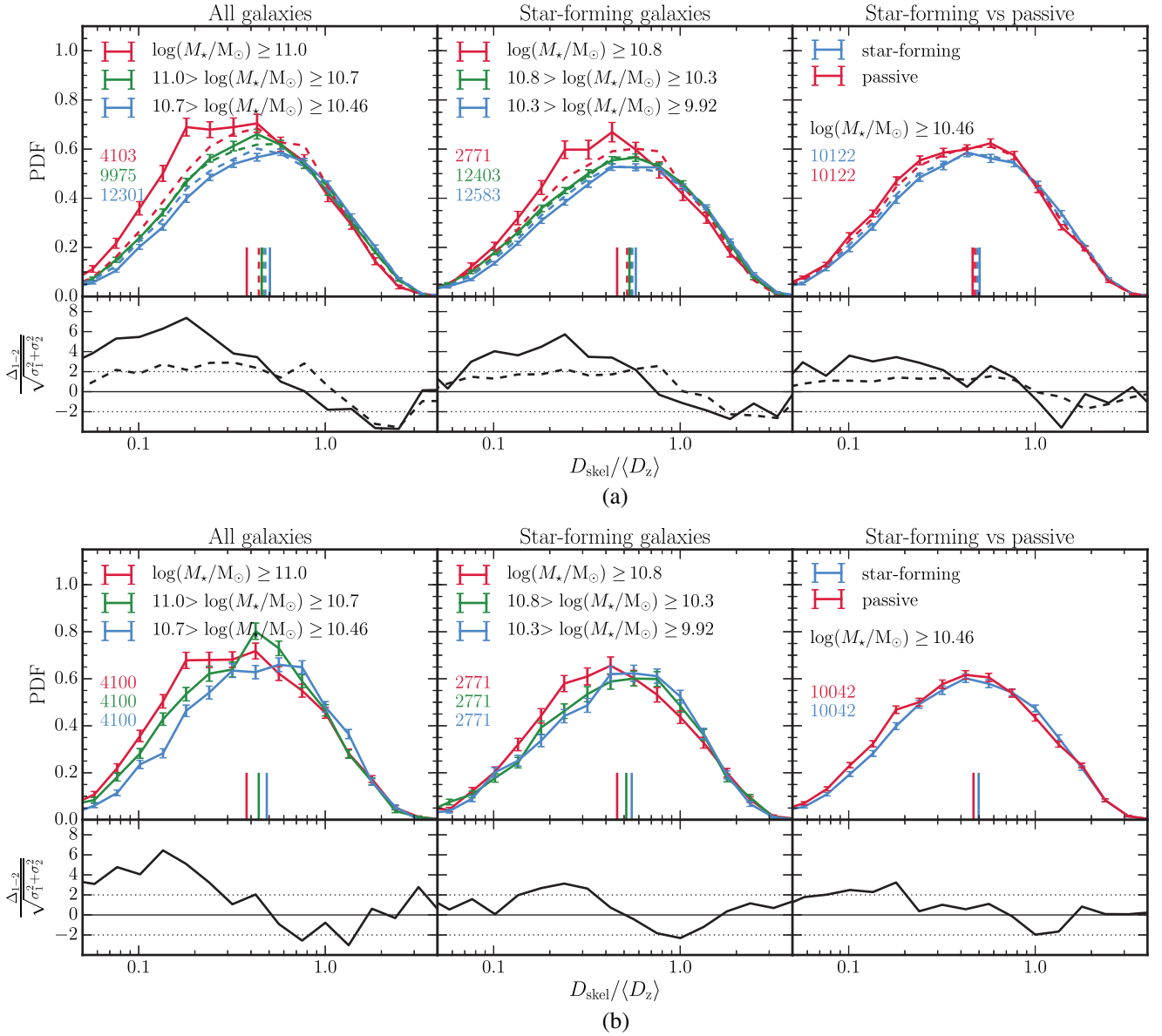
Figure 12. *Top rows:* Differential distributions of the normalized distances to the nearest filament, D_{skel} as a function of stellar mass of the entire galaxy population (left-hand panels), for star-forming galaxies only (middle panels) and as a function of galaxy's type (right-hand panels) with reshuffling (Figure a) and with density-matched samples (Figure b). In Figure (a), the distributions before applying the reshuffling method (solid lines) are compared to the results after the reshuffling (dashed lines). Figure (b) illustrates the distributions for the galaxy samples that are matched so that their density distributions are the same (see the text for details on the matching). The density estimators used in both the reshuffling and density matching corresponds to the (large-scale) density computed in the Gaussian kernel at the scale of 8 Mpc. As previously, the contribution of nodes to the measured signal is minimized. The numerical values of medians, shown as vertical lines, are listed in Table 3. The two methods yield qualitatively similar result: on the one hand when the large-scale density is used in reshuffling, the signal is reduced (dashed lines, Figure a) suggesting that the measured gradients (solid lines, Figure a) are not driven by the density at this scale, on the other hand, the gradients are measured within the samples that are matched in density at large scale. *Bottom rows:* Residuals in units of σ between the highest and lowest mass bins (left-hand and middle panels) and between the star-forming and passive galaxies (right-hand panels). (a) Reshuffling. (b) Density matching.

applied. The values of medians of the distributions after the reshuffling can be found in Table 3. Qualitatively similar behaviour is obtained for the gradients towards walls (not shown here). The analysis in HORIZON-AGN provides a qualitative support for these results. In Appendix D2, Fig. D1(a), the same reshuffling method is applied to simulated galaxies. The density used for this test is computed in the Gaussian kernel at 5 Mpc. This scale corresponds to the ~1.5× mean inter-galaxy separation in HORIZON-AGN, consistently with the GAMA data.

Alternatively, to assess the impact of the density on the measured gradients within the cosmic web, one may want to use density matching. The purpose of this method is to construct mass- and colour-density matched samples, whereby galaxies with different masses and/or colours have similar density distributions, in order to make sure that the measured properties are not driven by their differences (see Appendix A2 for details on the matching technique). As shown in Fig. 12(b), the main result on the density-matching technique leads to the same conclusions as the reshuffling method.

Table 3. Medians for the PDFs displayed in Fig. 12: large-scale density

Selection[a]	Bin	Median[b] $D_{skel}/\langle D_z \rangle$		
		Original[c]	reshuffling[d]	Matching[e]
Masses				
All galaxies	$\log(M_\star/M_\odot) \geq 11$	0.379 ± 0.009	0.441 ± 0.009	0.379 ± 0.01
	$11 > \log(M_\star/M_\odot) \geq 10.7$	0.456 ± 0.007	0.463 ± 0.006	0.44 ± 0.009
	$10.7 > \log(M_\star/M_\odot) \geq 10.46$	0.505 ± 0.007	0.475 ± 0.006	0.486 ± 0.01
SF galaxies	$\log(M_\star/M_\odot) \geq 11$	0.459 ± 0.01	0.541 ± 0.015	0.459 ± 0.011
	$11 > \log(M_\star/M_\odot) \geq 10.4$	0.534 ± 0.007	0.543 ± 0.007	0.514 ± 0.012
	$10.4 > \log(M_\star/M_\odot) \geq 9.92$	0.578 ± 0.007	0.552 ± 0.007	0.549 ± 0.012
Types				
SF versus passive[f]	Star-forming	0.503 ± 0.007	0.491 ± 0.007	0.498 ± 0.007
	Passive	0.462 ± 0.007	0.476 ± 0.007	0.467 ± 0.006

Notes. [a]Panels of Fig. 12.
[b]Medians of distributions as indicated in Fig. 12 by vertical lines; errors are computed as in Table 1.
[c]As in Table 1 for $D_{skel}/\langle D_z \rangle$.
[d]Reshuffling is done in bins of density computed at 8 Mpc (see the text for details).
[e]Medians for the density-matched sample, where the density considered is computed at 8 Mpc.
[f]Only galaxies with stellar masses $\log(M_\star/M_\odot) \geq 10.46$ are considered.

After matching galaxy populations in the large-scale density, mass and type gradients towards filaments and walls are still detected, suggesting that beyond the density, large-scale structures of the cosmic web do impact these galactic properties.

7 DISCUSSION

Let us first discuss the observational findings of the previous section in the framework of existing work (Section 7.1) and then focus on a recent extension of anisotropic excursion set which is developed in the companion paper (Section 7.2). The latter will allow us to explain why colour gradients prevail at fixed density.

7.1 Cosmic web metric: expected impact on galaxy evolution

In the current framework for galaxy formation, in which galaxies reside in extended dark matter haloes, it is quite natural to split the environment into the *local* environment, defined by the dark matter halo and the *global* large-scale anisotropic environment, encompassing the scale beyond the halo's virial radius. The anisotropy of the cosmic web is already a direct manifestation of the generic anisotropic nature of gravitational collapse on larger scales. It provides the embedding in which dark haloes and galaxies grow via accretion, which will act upon them via the combined effect of tides, the channeling of gas along preferred directions and angular momentum advection on to forming galaxies.

The observations and simulations presented in Sections 4, 5 and 6 provide a general support for this scenario. While rich clusters and massive groups are known to be environments which induce major galaxy transformations, the red fraction analysis presented in Section 4.1 (Fig. 5) reveals that the fraction of passive galaxies in the filaments starts to increase several Mpc away from the nodes and peaks in the nodes. This gradual increase suggests that some 'pre-processing' already happens before the galaxies reach the virial radius of massive haloes and fall into groups or clusters (e.g. Porter et al. 2008; Martínez et al. 2016). The above-mentioned morphological transformation of elliptical-to-spiral ratio when getting closer to the filaments (see also Kuutma et al. 2017) can be interpreted as the result of mergers transforming spirals into passive elliptical galaxies along the filaments when migrating towards nodes as suggested by theory and simulations (Codis et al. 2012; Dubois

et al. 2014). These findings show that filamentary regions, corresponding to intermediate densities, are important environments for galaxy transformation. This is also confirmed by the segregation found in Sections 4.2 (Figs 6 and 7). More massive and/or passive galaxies are found closer to the core of filaments than their less massive and/or star-forming counterparts. These differential mass gradients persist among the star-forming population alone. In addition to mass segregation, star-forming galaxies show a gradual evolution in their star formation activity (see Fig. 8). They are bluer than average at large distances from filaments ($D_{skel} \gtrsim 5$ Mpc), in a 'steady state' with no apparent evolution in star formation activity at intermediate distances ($0.5 \leq D_{skel} \leq 5$ Mpc) and they show signs of decreased star formation efficiency near the core of the filaments ($D_{skel} \lesssim 0.5$ Mpc). These results are in line with the picture where on the one hand more massive/passive galaxies lay in the core of filaments and merge while drifting towards the nodes of the cosmic web. On the other hand, the low mass/star-forming galaxies tend to be preferentially located in the outskirts of filaments, a vorticity-rich regions (Laigle et al. 2015), where galaxies acquire both their angular momentum (leading to a spin parallel to the filaments) and their stellar mass via essentially smooth accretion (Dubois et al. 2012b; Welker et al. 2017, also relying on HORIZON-AGN). The steady state of star-formation in these regions can reflect the right balance between the consumption and refuelling of the gas reservoir by the cold gas controlled by their surrounding filamentary structure (as shown by Codis et al. 2015, following Pichon et al. 2011, the outskirts of filaments are the loci of most efficient helicoidal infall of cold gas). This may not be true anymore when galaxies fall in the core of the filaments. The decline of star formation activity can, in part, be due to the higher merger rate but also due to a quenching process such as strangulation, where the supply of cold gas is halted (Peng, Maiolino & Cochrane 2015). It could also find its origin in the cosmic web detachment (Aragon-Calvo et al. 2016), where the turbulent regions inside filaments prevent galaxies to stay connected to their filamentary flows and thus to replenish their gas reservoir.

7.2 Link with excursion set theory

The distinct transverse gradients found for mass, density and type or colour may also be understood within the framework of conditional excursion set theory. Qualitatively, the spatial variation of the (traceless part of the) tidal tensor in the vicinity of filaments will

delay infall on to galaxies, which will impact differentially galactic colour (at fixed mass), provided accretion can be reasonably converted into star formation efficiency.

7.2.1 Connecting gradients to constrained excursion set

The companion paper (Musso et al. 2017) revisits excursion set theory subject to conditioning the excursion to the vicinity of a filament. In a nutshell, the main idea of excursion set theory is to compute the statistical properties of the initial (over)density – a stochastic variable – enclosed within spheres of radius R, the scale which, through the spherical collapse model, can be related to the final mass of the object (should the density within the sphere pass the threshold for collapse). Increasing the radius of the sphere provides us with a proxy for 'evolution' (larger sphere, larger mass, smaller variance, later formation time) *and* a measure of the impact of the environment (different sensitivity to tides for different, larger, spheres). The expectations associated with this stochastic variable can be re-computed subject to the tides imposed by larger scale structures, which are best captured by the geometry of a filament-saddle point, $\mathcal{S}$, providing the local natural 'metric' for a filament (Codis et al. 2015). These large-scale tides will induce distinct weighting in the conditional PDF($\delta, \partial_R \delta | \mathcal{S}$) for the overdensity δ, and its successive derivatives with respect to scale, $\partial_R \delta$ etc. (so as to focus on collapsed accreting regions). Indeed, the saddle will shift not only the mean expectation of the PDFs but also importantly their co-variances (see Musso et al. 2017, for details). The derived expected (dark matter) mean density $\rho(r, \theta, \phi)$, Press-Schechter mass $M(r, \theta, \phi)$ and typical accretion rate $\dot{M}(r, \theta, \phi)$ then become explicit *distinct* functions of distance r and relative orientation to the closest (oriented) saddle point. Within this model, it follows that the orientation of the mass, density and accretion rate gradients differ. The misalignment arises because the various fields weight differently the constrained tides, which will physically e.g. delay infall, and technically involve different moments of the aforementioned conditional PDF (see Appendix E for more quantitative information on contour misalignment). This is shown in Fig. 13, which displays a typical longitudinal cross-section of those three maps in the frame of the saddle (with the filament along the Oz axis) in Lagrangian space.[5]

This line of argument explains environmentally driven differential gradients, yet there is still a stretch to connect it to the observed gradients. While there is no obvious consensus on the detailed effect of large-scale (dark matter) accretion on to the colour or star formation of galaxies at fixed mass and density, one can expect that the stronger the accretion, the stronger the AGN feedback, the stronger the quenching. Should this (reasonable) scaling hold true, the net effect in terms of gradients would be that colour gradients differ from mass and density ones. This is qualitatively consistent with the findings of this paper.

---

[5] This companion paper does not capture the strongly non-linear process of dynamical friction of sub clumps within dark haloes, nor strong deviations from spherical collapse. We refer to Hahn et al. (2009), which captures the effect on satellite galaxies, and to Ludlow, Borzyszkowski & Porciani (2014), Castorina et al. (2016) and Borzyszkowski et al. (2017) which study the effect of the local shear on haloes forming in filamentary structures. This requires adopting a threshold for collapse that depends explicitly on the local shear. The shear-dependent part of the critical density (and its derivative) correlates with the shear of the saddle, and introduces an additional anisotropic effect on top of the change of mean values and variances of density and slope.
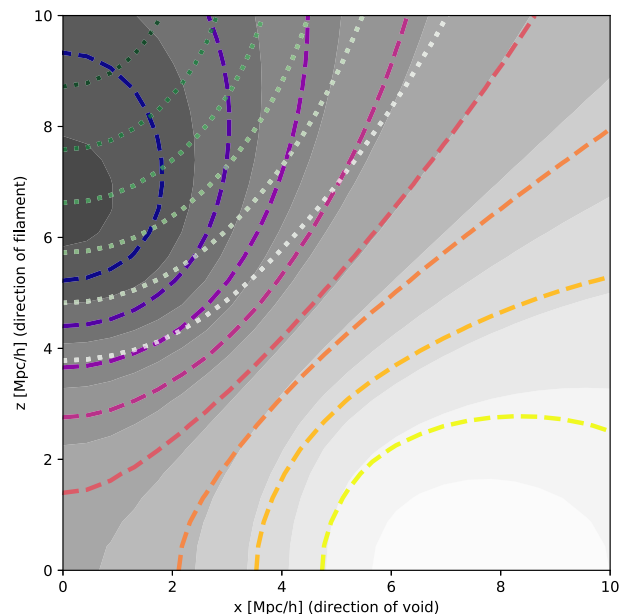
Figure 13. Isocontours of constant typical redshift $z = 0$ mean density (filled contours), mass (dotted lines) and accretion rate (dashed lines) in the frame of a filament (along the Oz axis) in Lagrangian space (initial conditions) from low (light colours) to high values (dark colours). The saddle is at coordinate (0,0) while the induced peak and void are at coordinates (0,$\pm$7) and ($\pm$8,0) h^{-1}Mpc, respectively. As argued in the main text, this figure shows that the contours, hence the gradients of the three fields, are not parallel (the contours cross). The choice of scale sets the units on the x- and z-axis (chosen here to be $5\,h^{-1}$Mpc, while the mass and accretion rates are computed for a local smoothing of $0.5\,h^{-1}$Mpc). At lower redshift/smaller scales, one expects the non-linear convergence of the flow towards the filament to bring those contours together, aligning the gradients (see Fig. 14).

7.2.2 Gradient alignments on smaller non-linear scales

The above-presented Lagrangian theory clearly applies only on sufficiently large scales so that dynamical evolution has not driven the large-scale flow too far from its initial configuration. On smaller scales, one would expect the same line of argument to operate in the frame set by the saddle smoothed on the corresponding scale, but with one extra caveat: the increased level of non-linearity will have compressed the local filament transversally and stretched it longitudinally, following the generic kinetic flow measured in N-body simulation (e.g. Sousbie et al. 2008a), or predicted at the level of the Zel'dovich approximation (Codis et al. 2015).

Consequently, the contours of constant dark matter density ρ, typical dark halo mass M and typical relative accretion rate $\dot{M}/M$ in the frame of the saddle shown in Fig. 13 will be driven more parallel to each other, hence the difference in the orientation of the density, mass and accretion gradient will become smaller and smaller as one considers smaller scales, and/or more non-linear dynamics (see Fig. 14). As colour gradient at fixed mass, and mass gradient at fixed density towards filaments originate from this initial misalignment, it should come as no surprise that as one probes smaller scales, such relative gradients disappear. When considering statistical expectations concerned with anisotropy (delayed accretion, acquisition of angular momentum, etc.), the net effect of past interactions should first be considered on the largest significant scale, beyond which the universe becomes isotropic. Conversely, the level of stochasticity should increase significantly on smaller scales, where one must
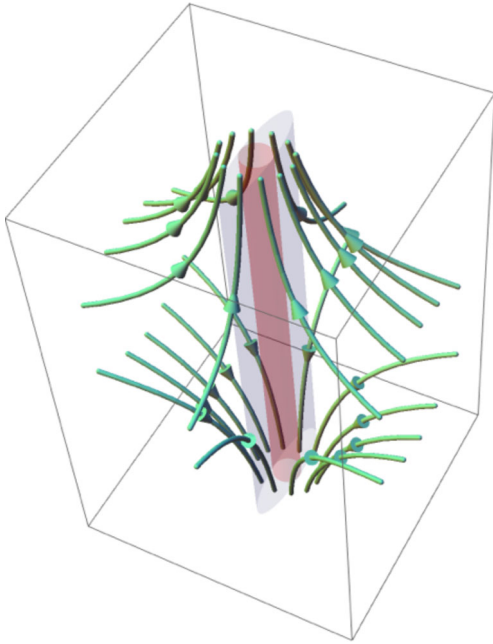
Figure 14. Illustration of the Zel'dovich flow (green arrows) in the vicinity of a filament (red cylinder) embedded in a wall (purple flattened cylinder), with filament saddle at the centre. The non-linear evolution operating more strongly on smaller scales will advect the contours presented in Fig. 13 along the green arrows, bringing them more parallel to each other. Consequently at these smaller scales, the mass and accretion gradients do not differ significantly from the density gradients. See Codis et al. (2015) and Musso et al. (2017) for more details.

account for, e.g. the configuration of the last merger event, or the last fly-by. Such a scenario is indeed supported by our findings in both GAMA and HORIZON-AGN, presented in Appendices C and D2, Figs C1 and D1, respectively, whereby the use of the small-scale density tracer does not allow us to disentangle between the effects of the local density and that of cosmic web, suggesting that at such scale, they are closely correlated through the small-scale processes.

7.2.3 Relationship to wall gradients

When measured relative to the walls, galaxy properties are found to exhibit the same trends as for filaments, in that more massive and/or quiescent galaxies are found closer to the walls than their low mass and/or star-forming counterparts. This result is again in qualitative agreement with the idea of walls being, together with the filaments, the large-scale interference patterns of primordial fluctuations capable of inducing anisotropic boost in overdensity together with the corresponding tides, and consequently imprinting their geometry in the measured properties of galaxies. The gradients measured for walls have the same origin as those inducing the differential gradients near the filament-type saddles, but are sourced by the geometry of the tides near the wall-type saddles (Codis et al. 2015, Appendix B). The main difference between the two saddles lies in the transverse curvatures, which is steeper for wall-type saddles than for filament-type saddles (when considering the mean, eigenvolume weighted, eigenvalues of the curvature tensor with the relevant signatures) leading to weaker differences between the different gradients when considering walls. This is consistent with the findings of Section 4.3.

In closing, note that the (resp. Eulerian and Lagrangian) interpretations presented in Sections 7.1 and 7.2 are complementary, but fall short in explaining in details the origin of quenching. Nevertheless, in view of both observation and theory, the cosmic web metric appears as a natural framework to understand galaxy formation beyond stellar mass and local density.

8 SUMMARY AND CONCLUSIONS

This paper studies the impact of the large-scale environment on the properties of galaxies, such as their stellar mass, dust corrected $u - r$ colour and sSFR. The discrete persistent structure extractor (DisPerSE) was used to identify the peaks, filaments and walls in the large-scale distribution of galaxies as captured by the GAMA survey. The principal findings are the following.

(i) *Mass segregation.* Galaxies are found to segregate by stellar mass, such that more massive galaxies are preferentially located closer to the cores of filaments than their lower mass counterparts. This mass segregation persists among the star-forming population. Similar mass gradients are seen with respect to walls in that galaxies with higher stellar mass tend to be found closer to the walls compared to galaxies with lower mass and persisting even when star-forming population of galaxies is considered alone.

(ii) *Type/colour segregation.* Galaxies are found to segregate by type/colour, with respect to both filaments and walls, such that passive galaxies are preferentially located closer to the cores of filaments or walls than their star-forming counterparts.

(iii) *Red fractions.* The fraction of passive galaxies increases with both decreasing distance to the filament and to the node, i.e. at fixed distance to the node, the relative number of passive galaxies (with respect to the entire population) increases as the distance to the filament decreases and similarly, at a given distance to the filament, this number increases with decreasing distance to the node.

(iv) *Star formation activity.* Star-forming galaxies are found to carry an imprint of large-scale environment as well. Their dust corrected $u - r$ and sSFR are found to be more enhanced and reduced, respectively, in the vicinity of the filaments compared to their outskirts.

(v) *Consistency with cosmological simulations.* All the found gradients are consistent with the analysis of the HORIZON-AGN 'full physics' hydrodynamical simulation. This agreement suggests that what drives the gradients is captured by the implemented physics.

(vi) *Connection to excursion set theory.* The origin of the distinct gradients can be qualitatively explained via conditional excursion set theory subject to filamentary tides (Musso et al. 2017).

This work has focused on filaments, nodes and in somewhat lesser details on walls. Similar observational results were recently reported at high redshift by using the cosmic web filamentary structures in the VIPERS spectroscopic survey (Malavasi et al. 2017) and while using projected filaments in photometric redshift slices in the COSMOS field (Laigle et al. 2017). These observations are of intrinsic interest as a signature of galactic assembly; they also comfort theoretical expectations which point towards distinct gradients for colour, mass and density with respect to the cosmic web. The tides of the large-scale environment play a significant specific role in the evolution of galaxies, and are imprinted in their integrated physical properties, which vary as a function of scale and distance to the different components of the cosmic web in a manner which is specific to each observable.

These observations motivate a theory which eventually should integrate the anisotropy of the cosmic web as an essential ingredient to (i) describe jointly the dynamics and physics of galaxies, (ii) explain galactic morphological diversity, and (iii) mitigate intrinsic alignment in upcoming lensing dark energy experiments, once a proper modelling of the mapping between galaxies and their haloes (allowing e.g. to convert the DM accretion rate into colour of galaxy) becomes available.

Future large-scale spectrographs on 8 metre class telescopes (MOONS;[6] Cirasuolo et al. 2014; Cirasuolo & MOONS Consortium 2016, PFS;[7] Sugai et al. 2015) or space missions (WFIRST;[8] Spergel et al. 2013, 2015, and Euclid;[9] Laureijs et al. 2011, the deep survey for the latter) will extend the current analysis at higher redshift ($z \geq 1$) with similar samplings, allowing us to explore the role of the environment near the peak of the cosmic star formation history, an epoch where the connectivity between the LSS and galaxies is expected to be even tighter, with ubiquitous cold streams. Tomography of the Lyman-α forest with PFS, MOONS, ELT-HARMONI (Thatte et al. 2010) tracing the intergalactic medium will make the study of the link between galaxies and this large-scale gas reservoir possible.

ACKNOWLEDGEMENTS

REFERENCES

Alpaslan M. et al., 2015, MNRAS, 451, 3249
Alpaslan M. et al., 2016, MNRAS, 457, 2287
Andrae R., Jahnke K., 2011, MNRAS, 418, 2014
Aragon-Calvo M. A., Yang L. F., 2014, MNRAS, 440, L46
Aragón-Calvo M. A., Jones B. J. T., van de Weygaert R., van der Hulst J. M., 2007a, A&A, 474, 315
Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., van der Hulst J. M., 2007b, ApJ, 655, L5

Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., 2010, MNRAS, 408, 2163
Aragon-Calvo M. A., Neyrinck M. C., Silk J., 2016, preprint (arXiv:1607.07881)
Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, MNRAS, 310, 540
Aubert D., Pichon C., Colombi S., 2004, MNRAS, 352, 376
Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469
Bamford S. P. et al., 2009, MNRAS, 393, 1324
Barnes J., Efstathiou G., 1987, ApJ, 319, 575
Benítez-Llambay A., Navarro J. F., Abadi M. G., Gottlöber S., Yepes G., Hoffman Y., Steinmetz M., 2013, ApJ, 763, L41
Beygu B., Kreckel K., van der Hulst J. M., Jarrett T. H., Peletier R., van de Weygaert R., van Gorkom J. H., Aragon-Calvo M. A., 2016, MNRAS, 458, 394
Blanton M. R. et al., 2003, ApJ, 594, 186
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
Bond J. R., Myers S. T., 1996, ApJS, 103, 1
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Bond N. A., Strauss M. A., Cen R., 2010a, MNRAS, 406, 1609
Bond N. A., Strauss M. A., Cen R., 2010b, MNRAS, 409, 156
Borzyszkowski M., Porciani C., Romano-Diaz E., Garaldi E., 2017, MNRAS, 469, 594
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Castorina E., Paranjape A., Hahn O., Sheth R. K., 2016, preprint (arXiv:1611.03619)
Cautun M. C., van de Weygaert R., 2011, Astrophysics Source Code Library, record ascl:1105.003
Cervantes-Sodi B., Hernandez X., Park C., 2010, MNRAS, 402, 1807
Chen Y.-C. et al., 2017, MNRAS, 466, 1880
Cirasuolo M. et al., 2014, Proc. SPIE, 9147, 91470N
Cirasuolo M., MOONS Consortium 2016, in Skillen I., Barcells M., Trager S., eds, ASP Conf. Ser. Vol. 507, Multi-Object Spectroscopy in the Next Decade: Big Questions, Large Surveys, and Wide Fields. Astron. Soc. Pac., San Francisco, p. 109
Codis S., Pichon C., Devriendt J., Slyz A., Pogosyan D., Dubois Y., Sousbie T., 2012, MNRAS, 427, 3320
Codis S., Pichon C., Pogosyan D., 2015, MNRAS, 452, 3369
Colless M., Dalton G., Maddox S., Sutherland W., Norberg P., Cole S., Bland-Hawthorn J., 2001, MNRAS, 328, 1039
Colombi S., Pogosyan D., Souradeep T., 2000, Phys. Rev. Lett., 85, 5515
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
De Lapparent V., Geller M. J., Huchra J. P., 1986, ApJ, 302, L1
Doroshkevich A. G., 1970, Astrofizika, 6, 581
Dressler A., 1980, ApJ, 236, 351
Dressler A. et al., 1997, ApJ, 490, 577
Driver S. P. et al., 2009, Astron. Geophys., 50, 5.12
Driver S. P., Hill D. T., Kelvin L. S., Robotham A. S. G., Liske J., Norberg P., Baldry I. K., 2011, MNRAS, 413, 971
Driver S. P. et al., 2016, MNRAS, 455, 3911
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2012a, MNRAS, 420, 2662
Dubois Y., Pichon C., Haehnelt M., Kimm T., Slyz A., Devriendt J., Pogosyan D., 2012b, MNRAS, 423, 3616
Dubois Y. et al., 2014, MNRAS, 444, 1453
Eardley E., Peacock J. A., McNaught-Roberts T., Heymans C., Norberg P., Alpaslan M., Baldry I., 2015, MNRAS, 448, 3665
Efstathiou G., Jones B. J. T., 1979, MNRAS, 186, 133
Efstathiou G., Frenk C. S., White S. D. M., Davis M., 1988, MNRAS, 235, 715
Eke V. R. et al., 2004, MNRAS, 348, 866
Forero-Romero J. E., Hoffman Y., Gottlöber S., Klypin A., Yepes G., 2009, MNRAS, 396, 1815
Gay C., Pichon C., Le Borgne D., Teyssier R., Sousbie T., Devriendt J., 2010, MNRAS, 404, 1801
Geller M. J., Huchra J. P., 1989, Science, 246, 897

---

[6] Multi-Object Optical and Near-infrared Spectrograph.
[7] Prime Focus Spectrograph; http://pfs.ipmu.jp/
[8] Wide-Field Infrared Survey Telescope; http://wfirst.gsfc.nasa.gov
[9] http://sci.esa.int/euclid/, http://www.euclid-ec.org

Goto T., Yamauchi C., Fujita Y., Okamura S., Sekiguchi M., Smail I., Bernardi M., Gomez P. L., 2003, MNRAS, 346, 601

Guo Q., Tempel E., Libeskind N. I., 2015, ApJ, 800, 112

Haardt F., Madau P., 1996, ApJ, 461, 20

Hahn O., Porciani C., Carollo C. M., Dekel A., 2007a, MNRAS, 375, 489

Hahn O., Carollo C. M., Porciani C., Dekel A., 2007b, MNRAS, 381, 41

Hahn O., Porciani C., Dekel A., Carollo C. M., 2009, MNRAS, 398, 1742

Hahn O., Angulo R. E., Abel T., 2015, MNRAS, 454, 3920

Hashimoto Y., Oemler Jr A., Lin H., Tucker D. L., 1998, ApJ, 499, 589

Heavens A., Peacock J., 1988, MNRAS, 232, 339

Hopkins A. M. et al., 2013, MNRAS, 430, 2047

Ilbert O. et al., 2006, A&A, 457, 841

Jackson J. C., 1972, MNRAS, 156, 1P

Jõeveer M., Einasto J., Tago E., 1978, MNRAS, 185, 357

Jones B. J. T., Efstathiou G., 1979, MNRAS, 189, 27

Jones B. J. T., van de Weygaert R., Aragón-Calvo M. A., 2010, MNRAS, 408, 897

Kaiser N., 1984, ApJ, 284, L9

Kaiser N., 1987, MNRAS, 227, 1

Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, MNRAS, 353, 713

Kleiner D., Pimbblet K. A., Jones D. H., Koribalski B. S., Serra P., 2017, MNRAS, 466, 4692

Komatsu E. et al., 2011, ApJS, 192, 18

Kuutma T., Tamm A., Tempel E., 2017, A&A, 600, L6

Laigle C. et al., 2015, MNRAS, 446, 2744

Laigle C. et al., 2017, preprint (arXiv:1702.08810)

Laureijs R. et al., 2011, preprint (arXiv:1110.3193)

Lee J., 2004, ApJ, 614, L1

Lee J., Erdogdu P., 2007, ApJ, 671, 1248

Lewis I. et al., 2002, MNRAS, 334, 673

Libeskind N. I., Hoffman Y., Knebe A., Steinmetz M., Gottlöber S., Metuki O., Yepes G., 2012, MNRAS, 421, L137

Libeskind N. I. et al., 2017, preprint (arXiv:1705.03021)

Liske J. et al., 2015, MNRAS, 452, 2087

Ludlow A. D., Borzyszkowski M., Porciani C., 2014, MNRAS, 445, 4110

Malavasi N. et al., 2017, MNRAS, 465, 3817

Martínez H. J., Muriel H., Coenda V., 2016, MNRAS, 455, 127

Merson A. I. et al., 2013, MNRAS, 429, 556

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution. Cambridge Univ. Press, Cmabridge

Musso M., Cadiou C., Pichon C., Codis S., Dubois Y., 2017, preprint (arXiv:1709.00834)

Navarro J. F., Abadi M. G., Steinmetz M., 2004, ApJ, 613, L41

Novikov D., Colombi S., Doré O., 2006, MNRAS, 366, 1201

Pahwa I. et al., 2016, MNRAS, 457, 695

Paz D. J., Stasyszyn F., Padilla N. D., 2008, MNRAS, 389, 1127

Peebles P. J. E., 1969, ApJ, 155, 393

Peng Y., Maiolino R., Cochrane R., 2015, Nature, 521, 192

Pichon C., Pogosyan D., Kimm T., Slyz A., Devriendt J., Dubois Y., 2011, MNRAS, 418, 2493

Planck Collaboration XIII, 2016, A&A, 594, A13

Pogosyan D., Bond J. R., Kofman L., Wadsley J., 1996, in American Astronomical Society Meeting Abstracts, p. 1289

Pogosyan D., Pichon C., Gay C., Prunet S., Cardoso J. F., Sousbie T., Colombi S., 2009, MNRAS, 396, 635

Porciani C., Dekel A., Hoffman Y., 2002a, MNRAS, 332, 325

Porciani C., Dekel A., Hoffman Y., 2002b, MNRAS, 332, 339

Porter S. C., Raychaudhury S., Pimbblet K. A., Drinkwater M. J., 2008, MNRAS, 388, 1152

Postman M., Geller M. J., 1984, ApJ, 281, 95

Robotham A. S. G. et al., 2011, MNRAS, 416, 2640

Robotham A. S. G. et al., 2013, MNRAS, 431, 167

Rojas R. R., Vogeley M. S., Hoyle F., Brinkmann J., 2004, ApJ, 617, 50

Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29

Schmidt M., 1959, ApJ, 129, 243

Shi F. et al., 2016, ApJ, 833, 241

Sousbie T., 2011, MNRAS, 414, 350

Sousbie T., Pichon C., Colombi S., Novikov D., Pogosyan D., 2008a, MNRAS, 383, 1655

Sousbie T., Pichon C., Courtois H., Colombi S., Novikov D., 2008b, ApJ, 672, L1

Sousbie T., Pichon C., Kawahara H., 2011, MNRAS, 414, 384

Spergel D. et al., 2013, preprint (arXiv:1305.5422)

Spergel D. et al., 2015, preprint (arXiv:1503.03757)

Stewart K. R., Brooks A. M., Bullock J. S., Maller A. H., Diemand J., Wadsley J., Moustakas L. A., 2013, ApJ, 769, 74

Subba Rao M. U., Aragón-Calvo M. A., Chen H. W., Quashnock J. M., Szalay A. S., York D. G., 2008, New J. Phys., 10, 125015

Sugai H. et al., 2015, J. Astron. Telescopes Instrum. Syst., 1, 035001

Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253

Tegmark M. et al., 2004, ApJ, 606, 702

Tempel E., Libeskind N. I., 2013, ApJ, 775, L42

Tempel E., Stoica R. S., Saar E., 2013, MNRAS, 428, 1827

Teyssier R., 2002, A&A, 385, 337

Thatte N. et al., 2010, in Proc. SPIE Conf. Ser. Vol. 77352I, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 77352I

Tojeiro R. et al., 2017, MNRAS, 470, 3720

Trowland H. E., Lewis G. F., Bland-Hawthorn J., 2013, ApJ, 762, 72

Trujillo I., Carretero C., Patiri S. G., 2006, ApJ, 640, L111

Tully R. B., Fisher J. R., 1978, in Longair M. S., Einasto J., eds, IAU Symp. Vol. 79, Large Scale Structures in the Universe. Springer, Dordrecht, p. 31

Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, A&A, 506, 647

Welker C., Devriendt J., Dubois Y., Pichon C., Peirani S., 2014, MNRAS, 445, L46

Welker C., Dubois Y., Pichon C., Devriendt J., Chisari E. N., 2015, preprint (arXiv:1512.00400)

Welker C., Dubois Y., Devriendt J., Pichon C., Kaviraj S., Peirani S., 2017, MNRAS, 465, 1241

White S. D. M., 1984, ApJ, 286, 38

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wright A. H. et al., 2016, MNRAS, 460, 765

Zel'dovich Y. B., 1970, A&A, 5, 84

APPENDIX A: MATCHING TECHNIQUE

A1 Mass matching

This Appendix provides details on the mass matching procedure. First the mass distributions of the two populations are cut so that they cover the same stellar mass range, i.e. they have the same minimum and maximum value of stellar mass. Then, in each stellar mass bin, the population with lower number of galaxies is taken as the reference sample and N_{match} samples of galaxies are extracted in the other population, such that their mass distribution is the same as the one of the reference sample. In practice, for each galaxy in the reference sample, the corresponding galaxy of the larger sample is sought among galaxies whose mass difference with respect to the reference mass is smaller than $\Delta M_\star$ in logarithmic space. If there is no galaxy in larger sample satisfying this condition, the galaxy of the reference sample is removed from the analysis. In each of N_{match} samples, every galaxy of the larger sample is considered only once, however repetitions are allowed across all samples. By construction, after applying this procedure, one ends up with N_{match} samples consisting of the same number of star-forming and passive galaxies and having very similar stellar mass distributions.

If not stated differently, 20 mass-matched samples are typically constructed using 10 equipopulated stellar mass bins for each and choosing a value of 0.1 for $\Delta M_\star$ parameter. Varying the values

of N_{match}, $\Delta M_\star$ and the number of stellar mass bins within the reasonable range does not alter our conclusions.

A2 Density matching

This Appendix provides details on the density matching procedure. First, let us describe how the mass-density matched samples are constructed. The galaxy sample is first divided into three logarithmic stellar mass bins for which the density matched samples are to be constructed. In each of the 10 equipopulated logarithmic overdensity $(1+\delta)$ bins, the reference sample is identified as that of the previously constructed stellar-mass subsamples with the lowest number of galaxies. Next, for each galaxy in the reference sample, a galaxy is randomly chosen from each of two stellar mass bins having the overdensity closest to the galaxy in the reference sample. In practice, the nominal absolute difference in the $\log(M_\star/M_\odot)$ values used to match galaxies is 0.1. If no suitable galaxy is found in at least one of the two stellar mass bins, the galaxy of the reference sample is removed from the analysis. This procedure is repeated 10 times, ending up with 10 samples of galaxies having the same overdensity distributions in three different stellar mass bins.

Similarly, to construct type-density matched samples, the entire galaxy sample is first divided into the subsamples of star-forming and passive galaxies. Then, in each of the 10 equipopulated logarithmic overdensity $(1+\delta)$ bins, the reference sample (sample of passive or star-forming galaxies) is identified as the one having the lowest number of galaxies. We continue by randomly choosing a galaxy from the larger sample with an overdensity and stellar mass close to that of the galaxy from the reference sample. In practice, we pair galaxies for which the distance in the two-parameter logarithmic space, defined by the stellar mass and the overdensity, is minimal and smaller than 0.1. The procedure is again repeated 10 times in order to construct 10 samples of star-forming and passive galaxies having their mass and density distributions close to each other.

APPENDIX B: THE IMPACT OF COSMIC BOUNDARIES

It was stated in Sections 4.2.1 and 4.3 that the measured gradients towards filaments (Figs 6 and 7) and walls (Figs 9 and 10) are not simply due to gradients towards nodes in the former and due to gradients towards nodes and filaments in the latter case. This Appendix presents the performed tests that allowed us to reach such a conclusion.

B1 Gradients towards filaments

Let us start by considering the gradients towards filaments. In order to probe these gradients without being substantially contaminated by the contribution from nodes, galaxies that are closer to nodes than 3.5 Mpc are removed from the analysis. The choice of this distance d_{min}^{node} is motivated by the compromise between eliminating the most of the gradient coming from nodes while keeping enough objects to have a statistically significant sample. Note that the distance of 3.5 Mpc is greater than the typical size of groups, which is ~ 1.5 Mpc in the redshift range considered in this work, measured as a median (or mean) projected group radius. The value of median (and mean) is insensitive to the definition of the group radius (see Robotham et al. 2011, for various definitions considered). In Fig. B1, the solid lines show the mass gradients towards filaments for the entire sample (left-hand panel) on the one hand and after excluding galaxies with distances to the node $D_{node} \leq 3.5$ Mpc (right-hand panel).

The contribution of nodes to mass gradients towards filaments is measured by randomizing distances to the filament, D_{skel}, in bins of distances to the node, D_{node}. By construction, gradients towards nodes are preserved. 20 samples are constructed in each of which this reshuffling method is applied in 20 equipopulated logarithmic bins. As shown by the dashed lines in Fig. B1 and values of medians listed in Table F1, the reshuffling cancels the gradients towards filaments for $d_{min}^{node} = 3.5$ Mpc.
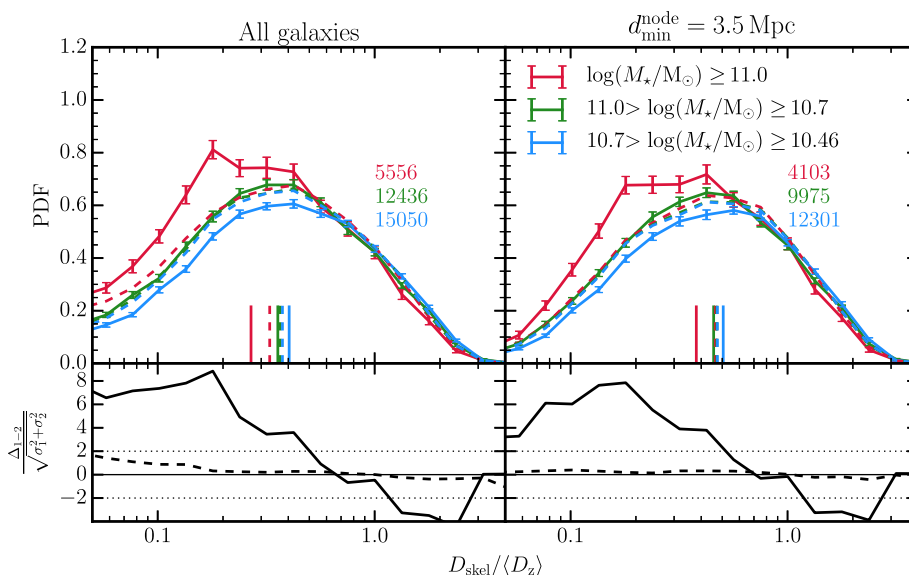
Figure B1. *Top row:* Differential distributions of the normalized distances to the nearest filament, D_{skel}. The solid lines show mass gradients for all galaxies (left-hand panel) and after removing galaxies with distances to the node smaller than 3.5 Mpc (right-hand panel). The dashed lines illustrate mass gradients after the reshuffling of D_{skel} of galaxies in bins of distances to the node D_{node}. The vertical lines indicate the medians of the distributions and their values, together with associated errors, are listed in Table F1. The reshuffling method cancels mass gradients towards filaments once galaxies at distances closer than 3.5 Mpc from nodes are removed. *Bottom row:* Residuals in units of σ between the two most extreme mass bins ($\log(M_\star/M_\odot) \geq 11.0$ and $10.7 > \log(M_\star/M_\odot) \geq 10.46$) before (solid lines) and after (dashed lines) the reshuffling.
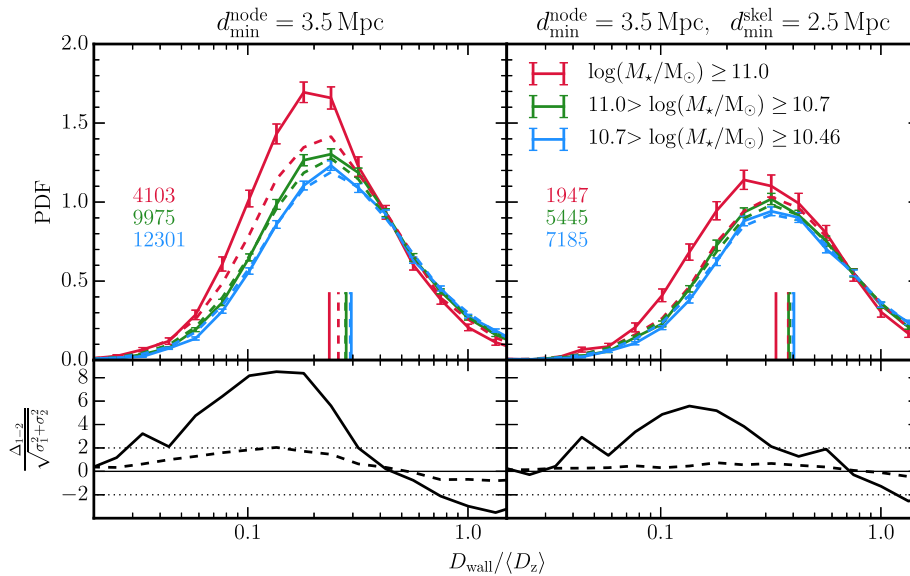
Figure B2. *Top row:* Differential distributions of the normalized distances to the nearest wall, D_{wall}. The solid lines show mass gradients after removing galaxies with distances to the node smaller than 3.5 Mpc (left) and after applying an additional criterion on the distance to the filament, such that galaxies with distances to the filament smaller than 2.5 Mpc (right) are removed. The dashed lines illustrate mass gradient after reshuffling of D_{skel} of galaxies in bins of distances to the node D_{node}. As shown on the right-hand panel, these are almost completely cancelled after removing sufficiently large regions around nodes and filaments. The vertical lines indicate the medians of the distributions and their values, together with associated errors, are listed in Table F2. *Bottom row:* Residuals are in units of σ as in Fig. B1.

In addition, following Laigle et al. (2017), it can be shown that in the regions sufficiently far away from nodes, gradients towards nodes and those towards filaments are independent. It was checked that the mass gradients towards nodes, present for the entire galaxy sample, are substantially reduced once galaxies for which distances to the node $D_{node} \leq 3.5$ Mpc are excluded. This time, the distances to the node, D_{node}, were randomized in bins of distances to the filament, D_{skel}, i.e. by construction, gradients towards filaments were preserved. Again, 20 samples were constructed using 20 equipopulated logarithmic bins. After reshuffling, weak gradients at the level of at most 1σ are still present, but note that additional increase in d_{min}^{node} does not reduce them further.

This analysis allows us to conclude that by removing from our sample galaxies that are closer to nodes than 3.5 Mpc, the impact of nodes to the measured gradients towards filaments is minimized, and even if weak gradients towards nodes still exist, these are independent of gradients towards filaments, i.e. gradients towards filaments and gradients towards nodes can be disentangled.

Let us finish this section with two remarks. First, note that distances to the node considered here are 3D euclidian distances. Curvilinear distances along the filaments could have been used instead (as illustrated in Fig. 4). This alternative choice of the distance does not alter our conclusions. Secondly, instead of using distances to the node D_{node}, one could have considered distances normalized by the redshift-dependent mean inter-galaxy separation, $D_{node}/\langle D_z \rangle$. These two approaches give consistent results not only qualitatively, but also quantitatively.

B2 Gradients towards walls

As with filaments, when measuring the gradients towards walls, one should investigate whether the gradient is not dominated by other component of the environments. As filaments are regions where walls intersect, these represent on top of nodes an additional source of contamination for the measured gradients towards walls.

Fig. B2 shows the mass gradients towards walls for the galaxy sample outside the zone of influence of nodes parametrized by $d_{min}^{node} = 3.5$ Mpc (left-hand panel) and after applying an additional criterion by excluding galaxies with distances to the closest filament $D_{skel} \leq d_{min}^{skel}$ with $d_{min}^{skel} = 2.5$ Mpc (right-hand panel). The contribution of filaments to the mass gradients towards walls is measured by randomizing distances to the wall, D_{wall}, in bins of distances to the filament, D_{skel}. By construction, the gradients towards filaments are preserved. Here 20 samples are constructed in each of which the reshuffling method is applied in 20 equipopulated logarithmic bins. As shown by the dashed lines in Fig. B2 and values of medians listed in Table F2, the reshuffling cancels the gradients towards walls for $d_{min}^{skel} = 2.5$ Mpc.

Following the method used in Appendix B1, it was verified (but not shown here) that the mass gradients towards filaments after randomization of the distances D_{skel} in bins of distances to the nearest wall D_{wall} are substantially reduced. Only a very weak mass gradient (at a 1σ level at most) is detected after randomization even for $d_{min}^{skel} = 2.5$ Mpc. Similarly to what was found in Section B1, increasing this parameter does not induce any substantial reduction of the gradient. Thus this distance was chosen as the limit for the exclusion region around filaments.

APPENDIX C: SMALL-SCALE DENSITY-COSMIC WEB RELATION

In this Appendix, the impact of the small-scale density estimator on the mass and type/colour gradients is presented. The density used here is DTFE, i.e. the density computed at the smallest possible scale.[10] As in Section 6, the two methods, the reshuffling and density-matching, are applied.

---

[10] There is no specific scale associated with the DTFE: it is a local adaptive method which determines the density at each point while preserving its multiscale character.
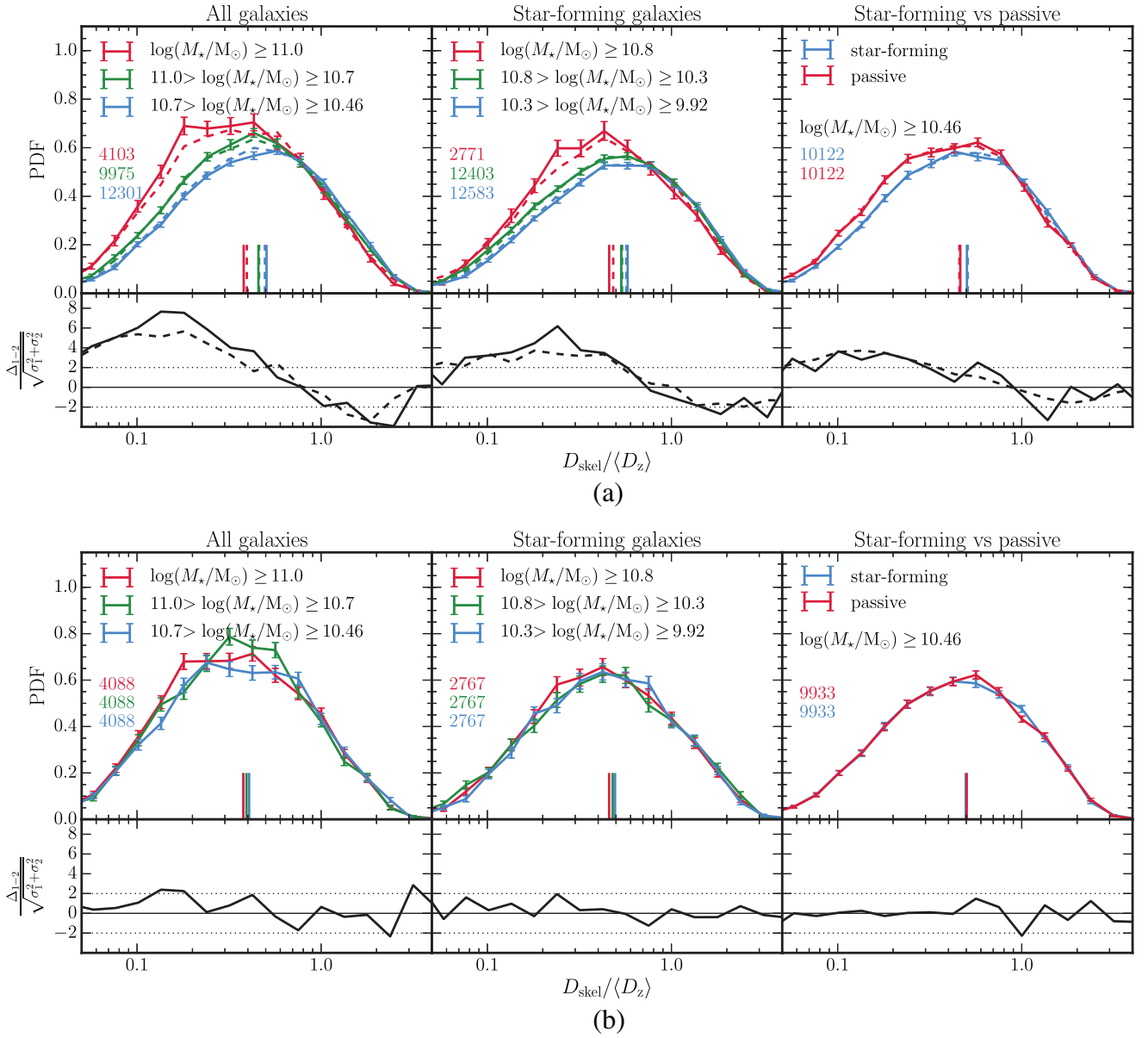
Figure C1. *Top rows:* As in Fig. 12, but using the DTFE density for both methods, reshuffling (Figure a) and density matching (Figure b). The numerical values of medians, shown as vertical lines, are listed in Table F3. When the small-scale density, DTFE in this case, is used in the reshuffling method, the randomized (dashed lines) and original signal (solid lines) are nearly identical. Similarly, all gradients are almost completely erased, as expected. *Bottom rows:* Residuals are in unit of σ as in Fig. 12. (a) Reshuffling. (b) Density matching.

Fig. C1 shows the differential distributions of the distances to the nearest filament, D_{skel} (normalized by $\langle D_z \rangle$), for the same selections as in Fig. 12. The contribution of the nodes to the measured signal is minimized, by removing from the analysis galaxies located closer to a node than 3.5 Mpc. Star-forming and passive galaxies have been matched in mass, as described in Appendix A1. The vertical lines indicate the medians of the distributions, whose values, together with the error bars, are listed in Table F3.

In Figure (a), the mass and type gradients are shown before (solid lines, as in 12) and after (dashed lines) applying the reshuffling of galaxies in the bins of overdensity $(1+\delta)$, where the number density corresponds to the DTFE density. The result conforms to the expectations. The reshuffling does not remove the observed mass

and type/colour gradients, i.e. the distributions before and after the reshuffling are almost identical, suggesting that at the small scale, traced by DTFE, the density and cosmic web are closely correlated through the small-scale processes.

Figure (b) illustrates the PDFs for samples that have been matched in overdensity $(1 + \delta)$, as described in Appendix A2, where the density considered is DTFE. The density-matching technique yields qualitatively similar result than the above used reshuffling in that almost no mass and type gradients are detected when galaxies matched in the DTFE density.

Qualitatively same results are obtained for both methods when applied to the measurements of gradients with respect to the walls (not shown).

APPENDIX D: THE HORIZON-AGN SIMULATION

This Appendix is dedicated to presenting the large-scale cosmological hydrodynamical simulation HORIZON-AGN (Dubois et al. 2014). First, some of the main features of the simulation are briefly summarized. The reshuffling method is then implemented on the simulation, as defined in Section 6, and shown to yield qualitatively similar results to those obtained in GAMA for both large- and small-scale density tracers.

D1 Simulation summary

The detailed description of the HORIZON-AGN simulation[11] can be found in Dubois et al. (2014), here only its brief summary is given. The cosmological parameters used in the simulation correspond to the ΛCDM cosmology with total matter density $\Omega_m = 0.272$, dark energy density $\Omega_\Lambda = 0.728$, amplitude of the matter power spectrum $\sigma_8 = 0.81$, baryon density $\Omega_b = 0.045$, Hubble constant $H_0 = 70.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $n_s = 0.967$ compatible with the WMAP-7 data (Komatsu et al. 2011).

The simulation was run with the Adaptive Mesh Refinement code RAMSES (Teyssier 2002) in a box of length $L_{box} = 100\,h^{-1}\,\mathrm{Mpc}$ containing 1024^3 dark matter (DM) particles, with a DM mass resolution of $M_{DM,res} = 8 \times 10^7\,\mathrm{M}_\odot$, and initial gas resolution of $M_{gas,res} = 1 \times 10^7\,\mathrm{M}_\odot$.

The collisionless DM and stellar components are evolved using a particle-mesh solver. The dynamics of the gaseous component are computed by solving Euler equations on the adaptive grid using a second-order unsplit Godunov scheme.

The refinement is done in a quasi-Lagrangian manner starting from the initial coarse grid down to $\Delta x = 1$ proper kpc (seven levels of refinement) as follows: each AMR cell is refined if the number of DM particles in a cell is more than 8, or if the total baryonic mass in a cell is eight times the initial DM mass resolution. This results in a typical number of 7×10^9 gas resolution elements (leaf cells) in the HORIZON-AGN simulation at $z = 0$.

Heating of the gas from a uniform UV background takes place after redshift $z_{reion} = 10$ following Haardt & Madau (1996). Gas is allowed to cool down to 10^4 K through H and He collisions with a contribution from metals using a Sutherland & Dopita (1993) model.

The conversion of gas into stars occurs in regions with gas density exceeding $\rho_0 = 0.1\,\mathrm{H\,cm^{-3}}$ following the Schmidt (1959) relation of the form $\dot{\rho}_* = \epsilon_* \rho_g / t_{ff}$, where $\dot{\rho}_*$ is the SFR mass density, ρ_g the gas mass density, $\epsilon_* = 0.02$ the constant star formation efficiency, and t_{ff} the local free-fall time of the gas.

Feedback from stellar winds, supernovae type Ia and type II are included into the simulation with mass, energy and metal release. HORIZON-AGN simulation takes also into account the formation of black holes (BHs) that can grow by gas accretion at a Bondi–Hoyle–Lyttleton rate capped at the Eddington accretion rate when they form a tight enough binary. The AGN feedback is a combination of two different modes (the so-called quasar and radio mode) in which BHs release energy in the form of heating or jet when the accretion rate is, respectively, above and below 1 per cent of Eddington, with efficiencies tuned to match the BH-galaxy scaling relations at $z = 0$ (see Dubois et al. 2012a, for details).

Galaxies are identified using the updated method (Tweed et al. 2009) of the AdaptaHOP halo finder (Aubert et al. 2004) directly operating on the distribution of stellar particles. Only galactic structures with a minimum of $N_{min} = 100$ stellar particles are considered, which typically selects objects with masses larger than $2 \times 10^8\,\mathrm{M}_\odot$.

D2 Density reshuffling

Let us finally present the impact of the reshuffling method, as defined in Section 6, and the choice of the density tracer in the HORIZON-AGN simulation.

Fig. D1 illustrates that the result of reshuffling depends on the scale at which the density is computed. As expected, when using the small-scale density tracer, such as e.g. the DTFE density (Figure a), both mass and sSFR gradients are almost unchanged, while on sufficiently large scales, the gradients tend to cancel out (Figure b). The numerical value of the scale at which this happens is ~5 Mpc. This is again in a qualitative agreement with the scale required in the GAMA survey, corresponding to the ~1.5× mean inter-galaxy separation.

---

[11] http://www.horizon-simulation.org
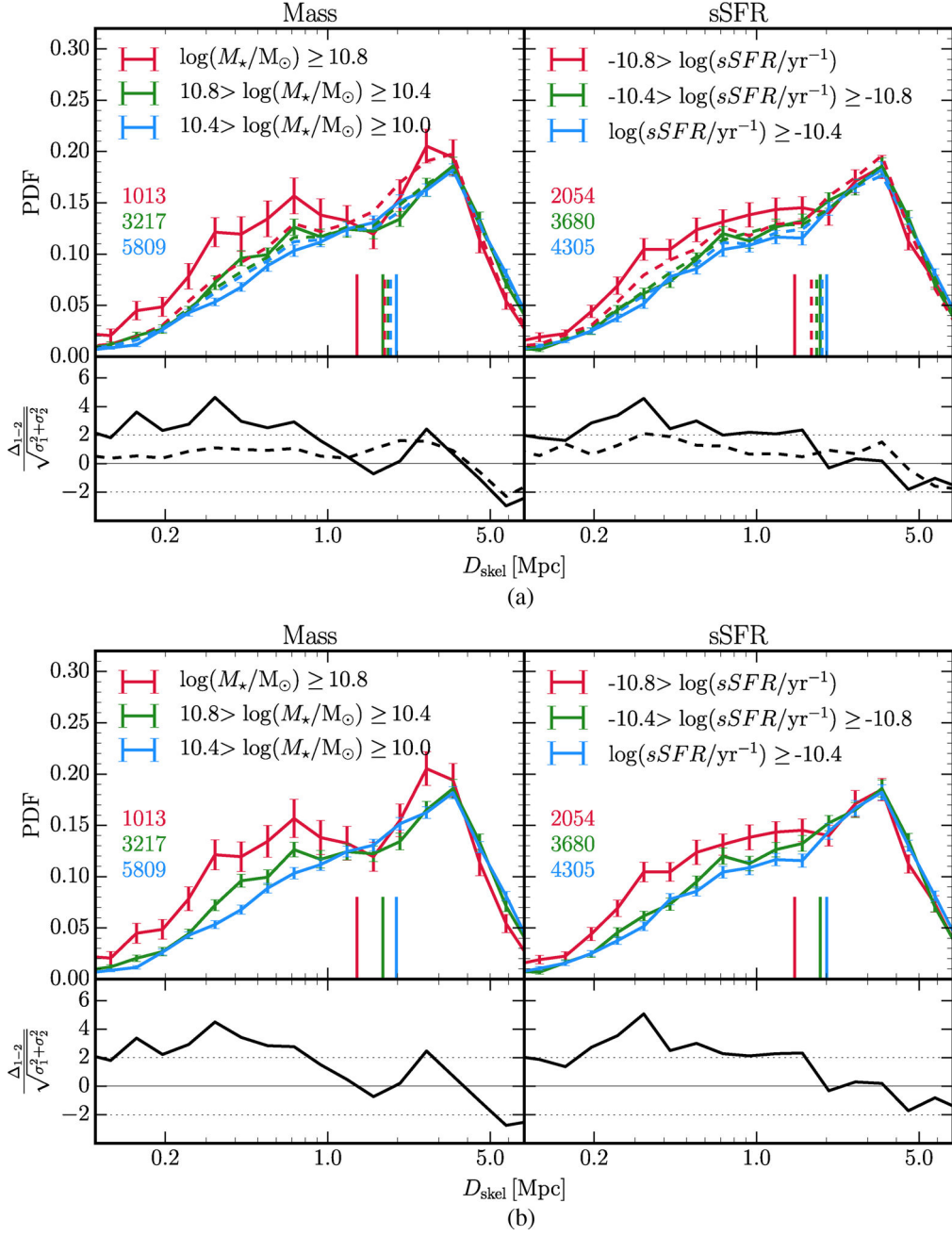
Figure D1. *Top rows:* As in Fig. 11 for the distances to the nearest filament, D_{skel}. The contribution of the nodes is minimized by removing galaxies located within 3.5 Mpc around them from the analysis. The dashed lines correspond to the distributions after the application of the reshuffling method using two different density tracers, a large (Figure a) and small-scale (Figure b) estimators. The numerical values of medians, shown as vertical lines, are listed in Table F4. In qualitative agreement with the results obtained with the observed data, in order to cancel the gradients, density at sufficiently large scale has to be considered. This corresponds to 5 Mpc in the HORIZON-AGN simulation, representing $\sim 1.5\times$ mean inter-galaxy separation, again in agreement with the value found in observations. *Bottom rows:* As in Fig. 11 before (solid lines) and after (dashed lines) the reshuffling. (a) Reshuffling using the density computed in the Gaussian kernel at the scale of 5 Mpc. (b) Reshuffling using the DTFE density.

APPENDIX E: GRADIENT MISALIGNMENTS

In the context of conditional excursion set theory subject to a saddle $\mathcal{S}$ at some finite distance (r, θ, ϕ) from a forming halo, let us consider the Hessian of the potential, $q_{ij} \equiv \partial^2 \psi / \partial r_i \partial r_j$, smoothed on the saddle scale $R_{\mathcal{S}}$ and normalized so that $\langle \mathrm{tr}^2(\boldsymbol{q}) \rangle = 1$. The anisotropic shear is given by the traceless part $\bar{q}_{ij} \equiv q_{ij} - \delta_{ij} \mathrm{tr}\, \boldsymbol{q}/3$, which deforms the region by slowing down or accelerating the

collapse along each axis. At finite separation, this traceless shear modifies in an anisotropic way the statistics of the smooth mean density (and of its derivative with respect to scale). The variations are modulated by $\mathcal{Q} = \sum_{i,j} \hat{r}_i \bar{q}_{ij} \hat{r}_j$, with $\hat{r}_i = r_i/r$, i.e. by the relative orientation of the separation vector, $\boldsymbol{r}$ in the frame set by the tidal tensor of the saddle. This extra degree of freedom, $\mathcal{Q}(\theta, \phi)$, provides a supplementary vector space, beyond the radial direction, over which to project the gradients, with statistical weight depending on each

specific observable (mass, accretion rate, etc.). These quantities have thus potentially different iso-surfaces from each other and from the local mean density, a genuine signature of the impact of the traceless part of the tidal tensor. Indeed, for each observable, the conditioning on $\mathcal{S}$ introduces a further dependence on the geometry of the environment (the height of the saddle and its anisotropic shear $\bar{q}_{ij}$) and on the position $\mathbf{r}$ of the halo with respect to the saddle point. This dependence arises because the saddle point condition modifies the mean and variance of the stochastic process $(\delta, \partial_R \delta)$ – the height and slope of the excursion set trajectories – in a position-dependent way, making it more or less likely to form haloes of given mass and assembly history within the environment set by $\mathcal{S}$. The expectation of the process becomes anisotropic through $\mathcal{Q}$, and both mean and variance acquire distinct radial dependence through the relevant correlation functions $\xi_{\alpha\beta}$ defined below in equation (E8).

For instance, considering the typical mass, $M_\star$ and accretion rate, $\dot{M}_\star$, at scale R, straightforward trigonometry shows that cross-product of their gradients reads

$$\left(\frac{\partial \dot{M}_\star}{\partial r} \frac{\partial M_\star}{\partial \mathcal{Q}} - \frac{\partial \dot{M}_\star}{\partial \mathcal{Q}} \frac{\partial M_\star}{\partial r} \right) \tilde{\nabla} \mathcal{Q} , \tag{E1}$$

where $\tilde{\nabla} = (\partial/\partial\theta, (1/\sin\theta)\partial/\partial\phi)$. The companion paper (Musso et al. 2017) shows that the Taylor expansion in the anisotropy for the angular variation, $\mathcal{Q}$, of $M_\star$ and $\dot{M}_\star$ at fixed distance r from the saddle scale like

$$\Delta M_\star \propto \xi_{20}(r) \mathcal{Q}(\theta, \phi) , \tag{E2}$$

and

$$\Delta \dot{M}_\star \propto \left[\xi'_{20}(r) - \frac{\sigma - \boldsymbol{\xi}' \cdot \boldsymbol{\xi}}{\sigma^2 - \boldsymbol{\xi} \cdot \boldsymbol{\xi}} \xi_{20}(r) \right] \mathcal{Q}(\theta, \phi) , \tag{E3}$$

in terms of the variance

$$\sigma^2(R) = \int dk \frac{k^2 P(k)}{2\pi^2} W^2(kR) , \tag{E4}$$

and the radius-dependent vectors

$$\boldsymbol{\xi}(r) \equiv \{\xi_{00}(r), \sqrt{3}\xi_{11}(r)r/R_\star, \sqrt{5}\xi_{20}(r)\} , \tag{E5}$$

$$\boldsymbol{\xi}'(r) \equiv \{\xi'_{00}(r), \sqrt{3}\xi'_{11}(r)r/R_\star, \sqrt{5}\xi'_{20}(r)\} , \tag{E6}$$

where

$$R_\star^2 \equiv \int dk \frac{P(k)}{2\pi^2} \frac{W^2(kR_\mathcal{S})}{\sigma_\mathcal{S}^2} , \tag{E7}$$

with $P(k)$ the underlying power spectrum, $W(k)$ the top hat filter in Fourier space, $\sigma_\mathcal{S} = \sigma(R_\mathcal{S})$, while the finite separation correlation functions, $\xi_{\alpha\beta}(r, R, R_\mathcal{S})$ and $\xi'_{\alpha\beta}(r, R, R_\mathcal{S})$ are defined as

$$\xi_{\alpha\beta} \equiv \int dk \frac{k^2 P(k)}{2\pi^2} W(kR) \frac{W(kR_\mathcal{S})}{\sigma_\mathcal{S}} \frac{j_\alpha(kr)}{(kr)^\beta} , \tag{E8}$$

$$\xi'_{\alpha\beta} \equiv \int dk \frac{k^2 P(k)}{2\pi^2} W'(kR) \frac{W(kR_\mathcal{S})}{\sigma_\mathcal{S}} \frac{j_\alpha(kr)}{(kr)^\beta} , \tag{E9}$$

where $j_\alpha(x)$ are the spherical Bessel functions of the first kind and prime denote derivate with respect to σ. Note that equation (E3) clearly highlights the shifted variance, $\sigma^2 - \boldsymbol{\xi} \cdot \boldsymbol{\xi}$, which contributes to the difference between $\Delta M_\star$ and $\Delta \dot{M}_\star$. From equation (E3), since the square bracket is not proportional to ξ_{20} as in equation (E2), it follows that the cross-product in equation (E1) is non-zero, which in turn implies that the contours of mass and accretion rate differ.

APPENDIX F: MEDIANS OF DISTRIBUTIONS

This Appendix gathers tables of medians with corresponding error bars used in previous sections.

Table F1. Medians of $D_{\text{skel}}/\langle D_z \rangle$ for Fig. B1.

Selection[a]	Mass bin	Median[b] $D_{\text{skel}}/\langle D_z \rangle$	
		Before reshuffling[c]	After reshuffling
All galaxies	$\log(M_\star/M_\odot) \geq 11$	0.27 ± 0.01	0.33 ± 0.02
	$11 > \log(M_\star/M_\odot) \geq 10.7$	0.36 ± 0.01	0.37 ± 0.01
	$10.7 > \log(M_\star/M_\odot) \geq 10.46$	0.40 ± 0.01	0.38 ± 0.01
$d_{\min}^{\text{node}} = 3.5\,\text{Mpc}$	$\log(M_\star/M_\odot) \geq 11$	0.38 ± 0.01	0.46 ± 0.02
	$11 > \log(M_\star/M_\odot) \geq 10.7$	0.46 ± 0.01	0.47 ± 0.01
	$10.7 > \log(M_\star/M_\odot) \geq 10.46$	0.51 ± 0.01	0.47 ± 0.01

[a]Panels of Fig. B1.
[b]Medians of distributions as indicated in Fig. B1 by vertical lines; errors are computed as in Table 1.
[c]Randomization of D_{skel} in bins of D_{node}.

Table F2. Medians of $D_{\text{wall}}/\langle D_z \rangle$ for Fig. B2.

Selection[a]	Mass bin	Median[b] $D_{\text{wall}}/\langle D_z \rangle$	
		Before reshuffling[c]	After reshuffling
$d_{\min}^{\text{node}} = 3.5\,\text{Mpc}$	$\log(M_\star/M_\odot) \geq 11$	0.234 ± 0.005	0.258 ± 0.011
	$11 > \log(M_\star/M_\odot) \geq 10.7$	0.279 ± 0.003	0.278 ± 0.005
	$10.7 > \log(M_\star/M_\odot) \geq 10.46$	0.295 ± 0.003	0.292 ± 0.004
$d_{\min}^{\text{node}} = 3.5\,\text{Mpc}, d_{\min}^{\text{skel}} = 2.5\,\text{Mpc}$	$\log(M_\star/M_\odot) \geq 11$	0.334 ± 0.007	0.379 ± 0.028
	$11 > \log(M_\star/M_\odot) \geq 10.7$	0.381 ± 0.004	0.386 ± 0.011
	$10.7 > \log(M_\star/M_\odot) \geq 10.46$	0.403 ± 0.004	0.398 ± 0.008

[a]Panels of Fig. B2.
[b]Medians of distributions as indicated in Fig. B2 by vertical lines; errors are computed as in Table 1.
[c]Randomization of D_{wall} in bins of D_{skel}.

Table F3. Medians for the PDFs displayed in Fig. C1: small-scale density

Selection[a]		Bin	Median[b] $D_{skel}/\langle D_z \rangle$		
			Original[c]	reshuffling[d]	Matching[e]
Masses	All galaxies	$\log(M_\star/M_\odot) \geq 11$	0.379 ± 0.009	0.397 ± 0.009	0.378 ± 0.01
		$11 > \log(M_\star/M_\odot) \geq 10.7$	0.456 ± 0.007	0.459 ± 0.006	0.393 ± 0.009
		$10.7 > \log(M_\star/M_\odot) \geq 10.46$	0.505 ± 0.006	0.495 ± 0.006	0.406 ± 0.008
	SF galaxies	$\log(M_\star/M_\odot) \geq 10.8$	0.459 ± 0.012	0.489 ± 0.013	0.458 ± 0.011
		$10.8 > \log(M_\star/M_\odot) \geq 10.3$	0.534 ± 0.007	0.541 ± 0.008	0.479 ± 0.01
		$10.3 > \log(M_\star/M_\odot) \geq 9.92$	0.578 ± 0.007	0.567 ± 0.007	0.494 ± 0.006
Types	SF versus passive[f]	Star-forming	0.504 ± 0.008	0.508 ± 0.007	0.495 ± 0.006
		Passive	0.462 ± 0.007	0.458 ± 0.007	0.504 ± 0.006

[a]Panels of Fig. C1.

[b]Medians of distributions as indicated in Fig. C1 by vertical lines; errors are computed as in Table 1.

[c]As in Table 1 for $D_{skel}/\langle D_z \rangle$.

[d]Reshuffling is done in bins of DTFE density (see the main text for more details).

[e]Medians for the density-matched sample, where the density considered is DTFE.

[f]Only galaxies with stellar masses $\log(M_\star/M_\odot) \geq 10.46$ are considered.

Table F4. Medians for the PDFs displayed in Fig. D1

Selection[a]	Bin	Median[b] D_{skel} [Mpc]		
		Original[c]	after reshuffling[d]	
			DTFE	G5Mpc
Mass	$\log(M_\star/M_\odot) \geq 10.8$	1.34 ± 0.09	1.26 ± 0.08	1.72 ± 0.1
	$10.8 > \log(M_\star/M_\odot) \geq 10.4$	1.73 ± 0.08	1.71 ± 0.06	1.82 ± 0.06
	$10.4 > \log(M_\star/M_\odot) \geq 10$	1.97 ± 0.04	2.0 ± 0.05	1.86 ± 0.04
sSFR	$-10.8 > \log(sSFR/\mathrm{yr})$	1.46 ± 0.07	1.61 ± 0.07	1.74 ± 0.08
	$-10.4 > \log(sSFR/\mathrm{yr}) \geq -10.8$	1.88 ± 0.06	1.89 ± 0.06	1.81 ± 0.06
	$\log(sSFR/\mathrm{yr}) \geq -10.4$	2.0 ± 0.04	1.9 ± 0.05	1.91 ± 0.06

[a]Panels of Fig. D1.

[b]Medians of distributions as indicated in Fig. D1 by vertical lines; errors are computed as in Table 1.

[c]As in Table 2 for D_{skel} (corresponding to the solid lines in Fig. D1).

[d]Reshuffling is done in the bins of the DTFE density and the density computed at the scale of 5 Mpc (corresponding to the dashed lines in Figures a and b, respectively).

This paper has been typeset from a TEX/LATEX file prepared by the author.

B.2   "Galaxies flowing in the oriented saddle frame of the cosmic web" (published in MNRAS)

In this follow-up of Kraljic et al., 2018, we have used the virtual galaxies of Horizon-AGN to further study the spatial distribution of galactic properties as a function of their location in the frame of the cosmic web. The properties of galaxies in the simulation have been plotted in the same frame as the one used throughout section 4.5.3. Namely, the galactic properties have been plotted as a function of their distance to the nearest filament centre and their angle with respect to the filament orientation. The key result of the paper is a confirmation that galactic properties are spatially modulated by the cosmic web. In particular, it has been found that after having removed the mean stellar mass, halo mass and density effects, maps of the properties of galaxies still show a modulation with respect to the cosmic web.

Theoretically, I showed (appendix F of the paper) that one can qualitatively recover the spatial signal observed in numerical simulations if one takes into account the displacement field induced by Zel'dovich boost (see section 2.1.2.3). Indeed, the mean flow around filamentary structures squeezes isocontours in the direction transverse to the filament and stretches them towards nodes. This result, already highlighted in Musso, Cadiou et al., 2018 has been shown to qualitatively reproduce the spatial distribution of halo mass in the simulation (see figures E1 and E2).

While theoretical predictions seem to be able to forecast the assembly of DM halos, significant improvements need to be made to relate this to the evolution of their host galaxies. This in particular relevant to the study of satellite galaxies which are influenced by the activity of their more massive neighbours. This work showed that the cosmic web provides a natural frame in which complex effects driving the formation of galaxies can be studied (AGN feedback, gas stripping). Following the results of chapter 3 and chapter 4, I suggest that the frame of the cosmic web will prove useful to understand assembly bias, but can also be used as an ingredient entering halo and galaxy models.

Galaxies flowing in the oriented saddle frame of the cosmic web

K. Kraljic,[1]★ C. Pichon,[1,2,3] Y. Dubois,[2] S. Codis,[2] C. Cadiou,[2] J. Devriendt,[4,5]
M. Musso,[6] C. Welker,[7] S. Arnouts,[8] H. S. Hwang [ID],[9] C. Laigle,[4] S. Peirani,[2,10]
A. Slyz,[4] M. Treyer [ID][8] and D. Vibert[8]

[1]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[2]*Institut d'Astrophysique de Paris, CNRS and Sorbonne Université, UMR 7095, 98 bis Boulevard Arago, F-75014 Paris, France*
[3]*School of Physics, Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea*
[4]*Sub-Department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, UK*
[5]*Observatoire de Lyon, UMR 5574, 9 avenue Charles André, Saint Genis Laval F-69561, France*
[6]*East African Institute for Fundamental Research (ICTP-EAIFR), KIST2 Building, Nyarugenge Campus, University of Rwanda, Kigali, Rwanda*
[7]*International Centre for Radio Astronomy Research and ASTRO 3D, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*
[8]*Laboratoire d'Astrophysique de Marseille, Aix Marseille Univ, CNRS, LAM, 38 Rue Frédéric Joliot Curie, F-13013 Marseille, France*
[9]*Korea Institute for Advanced Study, Quantum Universe Center, 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea*
[10]*Laboratoire Lagrange, UMR7293, Université de Nice Sophia Antipolis, CNRS, Observatoire de la Côte d'Azur, F-06300 Nice, France*

ABSTRACT

The strikingly anisotropic large-scale distribution of matter made of an extended network of voids delimited by sheets, themselves segmented by filaments, within which matter flows towards compact nodes where they intersect, imprints its geometry on the dynamics of cosmic flows, ultimately shaping the distribution of galaxies and the redshift evolution of their properties. The (filament-type) saddle points of this cosmic web provide a local frame in which to quantify the induced physical and morphological evolution of galaxies on large scales. The properties of virtual galaxies within the HORIZON-AGN simulation are stacked in such a frame. The iso-contours of the galactic number density, mass, specific star formation rate (sSFR), kinematics, and age are clearly aligned with the filament axis with steep gradients perpendicular to the filaments. A comparison to a simulation without feedback from active galactic nuclei (AGNs) illustrates its impact on quenching star formation of centrals away from the saddles. The redshift evolution of the properties of galaxies and their age distribution are consistent with the geometry of the bulk flow within that frame. They compare well with expectations from constrained Gaussian random fields and the scaling with the mass of non-linearity, modulo the redshift-dependent impact of feedback processes. Physical properties such as sSFR and kinematics seem *not* to depend *only* on mean halo mass and density: the residuals trace the geometry of the saddle, which could point to other environment-sensitive physical processes, such as spin advection, and AGN feedback at high mass.

Key words: methods: analytical – methods: numerical – galaxies: evolution – galaxies: formation – galaxies: interactions – galaxies: kinematics and dynamics.

1 INTRODUCTION

Galaxies form and evolve within a complex network, the so-called cosmic web (Bond, Kofman & Pogosyan 1996), made of filaments embedded in sheet-like walls, surrounded by large voids and intersecting at clusters of galaxies (Jõeveer, Einasto & Tago 1978). Do

the properties of galaxies, such as their morphology, retain a memory of these large-scale cosmic flows from which they emerge? The importance of interactions with the larger scale environment in driving their evolution has indeed recently emerged as central tenet of galaxy formation theory. Galactic masses are highly dependent on their large-scale surrounding, as elegantly explained by the theory of biased clustering (Kaiser 1984; Efstathiou et al. 1988), such that high-mass objects preferentially form in overdense environment near nodes (Bond & Myers 1996; Pogosyan et al. 1996).

★ E-mail: kat@roe.ac.uk

Conversely, what are the signatures of this environment away from the nodes of the cosmic web?

While galaxies grow in mass when forming stars from intense gas inflows at high redshift, they also acquire spin through tidal torques and mergers biased by these anisotropic larger scales (e.g. Aubert, Pichon & Colombi 2004; Peirani, Mohayaee & de Freitas Pacheco 2004; Navarro, Abadi & Steinmetz 2004; Aragón-Calvo et al. 2007; Codis et al. 2012; Libeskind et al. 2012; Stewart et al. 2013; Trowland, Lewis & Bland-Hawthorn 2013; Aragon-Calvo & Yang 2014, for dark matter (DM), and Pichon et al. 2011; Dubois et al. 2014; Welker et al. 2014, in hydrodynamical simulations). This should in turn have a significant impact on galaxy properties including morphology, colour, and star formation history of galaxies.

As a filament is formally the field line that joins two maxima of the density field through a filament-type saddle point[1] (Pogosyan et al. 2009), studying the expected properties of galaxies *in the vicinity of filament-type saddle points* is a sensible choice. Indeed, Tidal Torque Theory (Peebles 1969; Schaefer 2009) was recently revisited (Codis, Pichon & Pogosyan 2015b) in the context of such anisotropic environments, biased by the presence of a filament within a wall, which is most efficiently represented by this point process of filament-type saddles.[2] It predicts the alignment of the angular momentum distribution of the forming galaxies with the filament's direction, and perpendicular orientation for massive population. Since spin plays an important role in the physical and morphological properties of galaxies, a signature is also expected in the properties of galaxies as a function of the longitudinal and transverse distances to this saddle.

Most of the previous theoretical work on the impact of the anisotropy of the environment on galactic assembly history focused on DM haloes. In the emerging picture of halo assembly history, at a given mass, haloes that are sufficiently far from the potential wells of other haloes can grow by accretion from their neighbourhood, leading to a correlation between the accretion rate of haloes and the density of their environment (e.g. Zentner 2007). Haloes that are close to more massive structures are on the other hand expected to be stalled and their growth may stop earlier, as their mass inflow is dynamically quenched by anisotropic tides generated in their vicinity (e.g. Dalal et al. 2008; Hahn et al. 2009; Ludlow, Borzyszkowski & Porciani 2014; Borzyszkowski et al. 2017; Paranjape, Hahn & Sheth 2018a). Individual properties of DM haloes, such as their mass, formation time, or accretion, are thus expected to be affected by the exact position of haloes within the large-scale anisotropic cosmic web (e.g. Lazeyras, Musso & Schmidt 2017). Such expectations are complementary to the recent work of Musso et al. (2018) whose analytical prediction of the mass, accretion rate, and formation time of DM haloes near proto-filaments (identified as saddle points of the gravitational potential field) confirms that the anisotropy of the cosmic web is a significant ingredient to describe jointly the dynamics and physics of haloes. Their model predicts that at *fixed* mass, mass accretion rate, and formation time of haloes also vary with orientation and distance from the saddle.

Theoretical predictions on the impact of the anisotropic tides of the cosmic web on the specific properties of *galaxies* embedded in those haloes are hampered by the complexity of baryonic processes and the lack of knowledge of detailed physics driving them. Some attempts were recently made by Alam et al. (2018) and Paranjape, Hahn & Sheth (2018b) which compared the observed clustering and quenching properties of galaxies in the Sloan Digital Sky Survey (SDSS) with corresponding measurements in mock galaxy catalogues. These studies focused on whether the cosmic web leaves an imprint on the galaxy clustering beyond the effects of halo mass, by constructing mock catalogues using a halo occupation distribution in such a way that dependencies of galaxy properties on the tidal anisotropy and isotropic overdensity are driven by the underlying halo mass function across the cosmic web alone. As such prescription qualitatively reproduces the main observed trends, and quantitatively matches many of the observed results, they concluded that any additional direct effect of the large-scale tidal field on galaxy formation must be extremely weak.

In this work, the adopted approach is different in that it focuses directly on galaxies, their physical properties and redshift evolution as measured in the large-scale cosmological hydrodynamical simulation HORIZON-AGN (Dubois et al. 2014, 2016). The main purpose of this paper is to show how the 3D distribution of the physical properties of these synthetic galaxies reflects the (tidal) impact of the cosmic web on the assembly history of galaxies. It is partly motivated by recent studies carried in the VIPERS, GAMA, and COSMOS surveys (Malavasi et al. 2017; Kraljic et al. 2018; Laigle et al. 2018) which showed that the colour and specific star formation rate (sSFR) of galaxies are sensitive to their proximity to the cosmic web at fixed stellar mass and local density. This paper focuses specifically on the distribution of the galaxy properties stacked in the oriented frame of the filament on large (∼Mpc) scales. The natural choice of frame for stacking is defined by filament-type saddle points connecting two nodes by one filament (in contrast to nodes which are typically places where the connectivity of filaments is higher).

This paper is organized as follows. Section 2 shortly describes the simulation and the detection of filaments within. Section 3 presents the galactic maps near the saddle, focusing first on the transverse and longitudinal (azimuthally averaged) maps, and then their 3D counterparts, while Section 4 shows their redshift evolution. Section 5 relates our finding to the properties of weakly non-Gaussian random fields near saddles. Some observational implications of our work together with the comparison with theoretical predictions are discussed in Section 6. Section 7 wraps up.

Appendix A explores the robustness of our finding with respect to smoothing and choice of filament tracer, Appendix B discusses the redshift evolution of the geometry of filaments, Appendix C presents complementary 2D maps, and Appendix D quantifies the position-in-the-saddle frame efficiency of AGN feedback. Appendix E sketches the derivation of the theoretical results presented in the main text. Appendix F presents the geometry of the bulk galactic velocity flow in the frame of the saddle. Finally, Appendix G motivates statistically the mediation of mass and density maps over tides. Throughout this paper, by log, we refer to the 10-based logarithm and we loosely use $\log M$ as a short term for $\log(M/\mathrm{M}_\odot)$ and $\log \rho$ for $\log(\rho/\mathrm{M}_\odot\, h^{-2}\, \mathrm{Mpc}^3)$.

---

[1]Where the gradient of the density field is null and the density Hessian has two negative eigenvalues.

[2]The constrained misalignment between the tidal and the inertia tensors in the vicinity of filament-type saddles simply explains the distribution of spin directions and its mass dependent flip.

2 NUMERICAL METHODS

Let us briefly review the main numerical tools used in this work to study the properties of virtual galaxies within the frame of the saddle.

2.1 Hydrodynamical simulation

The details of the HORIZON-AGN simulation[3] can be found in Dubois et al. (2014), here, only brief description is provided. The simulation is performed with the adaptive mesh refinement code RAMSES (Teyssier 2002) using a box size of $100\,h^{-1}$ Mpc and adopting a Λ cold dark matter (ΛCDM) cosmology with total matter density $\Omega_m = 0.272$, dark energy density $\Omega_\Lambda = 0.728$, baryon density $\Omega_b = 0.045$, amplitude of the matter power spectrum $\sigma_8 = 0.81$, Hubble constant $H_0 = 70.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and $n_s = 0.967$ compatible with the *WMAP*-7 data (Komatsu 2011). The total volume contains 1024^3 DM particles, corresponding to a DM mass resolution of $M_{\mathrm{DM,res}} = 8 \times 10^7\,M_\odot$. The initial gas resolution is $M_{\mathrm{gas,res}} = 1 \times 10^7\,M_\odot$. The refinement of the initially coarse 1024^3 grid down to $\Delta x = 1$ proper kpc is triggered in a quasi-Lagrangian manner: if the total baryonic mass reaches 8 times the initial DM mass resolution, or the number of DM particles becomes greater than 8 in a cell, resulting in a typical number of 7×10^9 gas resolution elements (leaf cells) at redshift zero.

The gas heating from a uniform ultraviolet background that takes place after redshift $z_{\mathrm{reion}} = 10$ is modelled following Haardt & Madau (1996). Gas is allowed to cool down to 10^4 K through H and He collisions with a contribution from metals (Sutherland & Dopita 1993). Star formation follows a Schmidt relation in regions of gas number density above $n_0 = 0.1\,\mathrm{H\,cm^{-3}}$ and $\dot{\rho}_* = \epsilon_* \rho_g / t_{\mathrm{ff}}$, where $\dot{\rho}_*$ is the star formation rate mass density, ρ_g the gas mass density, $\varepsilon_* = 0.02$ the constant star formation efficiency, and t_{ff} the local free-fall time of the gas. Feedback from stellar winds, supernovae type Ia and type II are included into the simulation with mass, energy and metal release (see Kaviraj et al. 2017, for further details).

The HORIZON-AGN simulation includes the formation of black holes (BHs) that can grow by gas accretion at a Bondi-capped-at-Eddington rate and coalesce when they form a tight enough binary. Energy of BHs can be released in a heating or jet mode (respectively 'quasar' and 'radio' mode) when the accretion rate is respectively above and below one per cent of Eddington, with efficiencies tuned to match the BH–galaxy scaling relations at redshift zero (see Dubois et al. 2012, for further details).

In order to assess the impact of active galactic nuclei (AGNs) feedback on galaxy properties in the frame of saddle, this analysis also relies on the HORIZON-NOAGN simulation, which was performed with identical initial conditions and sub-grid modelling, but without BH formation, thus without AGN feedback (Dubois et al. 2016; Peirani et al. 2017).

2.2 Galaxy properties

The identification of galaxies is performed using the most massive sub-node method (Tweed et al. 2009) of the ADAPTAHOP halo finder (Aubert et al. 2004) operating on the distribution of star particles with the same parameters as in Dubois et al. (2014). Only structures with a minimum of $N_{\min} = 100$ star particles are considered, which typically selects objects with masses larger than $2 \times 10^8\,M_\odot$. For each redshift output analysed in this paper ($0.05 < \mathrm{redshift} < 2$) catalogues containing up to $\sim$350 000 haloes and $\sim$180 000 galaxies are produced.

For each galaxy, its V/σ, stellar rotation over dispersion, is extracted from the 3D distribution of velocities. This is meant to provide a kinematic proxy for morphology. The total angular momentum (spin) of stars is first computed in order to define a set of cylindrical spatial coordinates (r, θ, z), with the z-axis oriented along the spin of galaxy. The velocity of each individual star particle is decomposed into cylindrical components v_r, v_θ, v_z, and the rotational velocity of a galaxy is $V = \bar{v}_\theta$, the mean of v_θ of individual stars. The average velocity dispersion of the galaxy $\sigma^2 = (\sigma_r^2 + \sigma_\theta^2 + \sigma_z^2)/3$ is computed using the velocity dispersion of each velocity component σ_r, σ_θ, and σ_z.

2.3 Saddle frame identification

In order to quantify the position of galaxies relative to the cosmic web, a geometric 3D ridge extractor called DISPERSE[4] (Sousbie 2011; Sousbie, Pichon & Kawahara 2011) is run on the full volume gas density distribution over 512^3 cells with a 3σ persistence threshold. This density distribution is smoothed with a Gaussian kernel with smoothing length of 0.8 comoving Mpc h^{-1}. The orientation and distribution of galaxies can be measured relative to the direction of the closest filament's segment. In particular, the code identifies saddle points along those filaments. This is a costly method to identify saddle points, but it provides us with a local preferred polarity in the frame of the density Hessian (positively towards the larger of the two maxima). It was checked that the distributions presented below are relatively insensitive to the choice of smoothing length (see Appendix A). It was also checked there that these results do not show a strong dependence on the tracer (DM or gas density) used to compute the skeleton.

3 SADDLE STACKS IN 2D AND 3D

With the aim of studying the geometry of the galaxy distribution around filaments, stacking centred on the saddle points of filaments is applied. When stacking, two different strategies are explored. First, stacks are produced centred on the saddle, and physical properties of galaxies are binned as a function of transverse and longitudinal distances away from the skeleton. These properties are also stacked in 3D in the local frame set by the direction of the filament at the saddle and the 2D inertia tensor in the plane perpendicular to the filament. The former method avoids the flaring induced by the drift of the curved filaments away from the saddle, only associate one saddle to each galaxy and stacks azimuthally, while the latter one allows us to probe the transverse anisotropic geometry of filaments at the saddle.

3.1 Azimuthally averaged stellar mass and number density

Let us start by considering azimuthally averaged 2D maps in the frame defined by the saddle and its steepest ascent direction, and study the cross-sections of galactic number density and stellar mass in the vicinity of the saddle point. In order to infer the variation of galaxy properties beyond its stellar mass, stellar mass will be fixed by considering three bins, defined as low ($9.0 \leq \log M_\star \leq 9.05$), intermediate ($9.69 \leq \log M_\star \leq 9.75$), and high ($10.93 \leq \log M_\star \leq 11.99$) stellar mass bins. These bins correspond to the first, middle, and last 27-quantiles of the stellar mass distribution of all galaxies at a given redshift above the stellar mass

---

[3]See http://www.horizon-simulation.org

[4]The code DISPERSE, which stands for Discrete-Persistent-Structure-Extractor algorithm is publicly available at the following URL http://www.iap.fr/users/sousbie/disperse/.
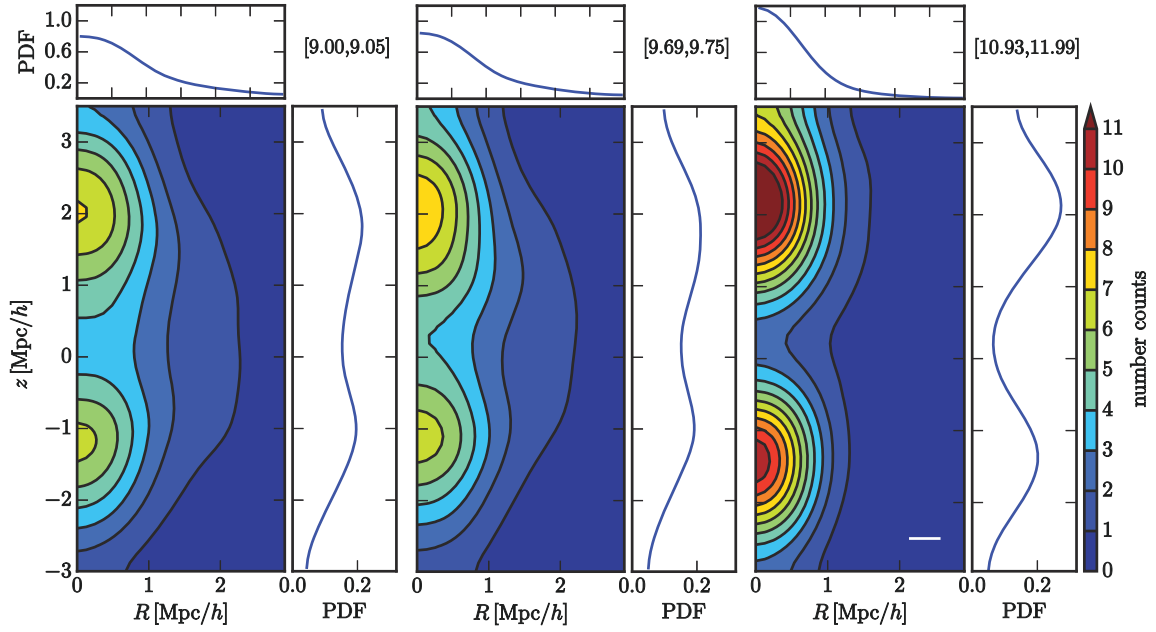
Figure 1. The galaxy number counts in low (left), intermediate (middle), and high (right) stellar mass bins (see the text for definition), as labelled (in square brackets), at redshift zero in the frame of the closest saddle. The vertical axis corresponds to the distance from the saddle along the skeleton, while the horizontal axis corresponds to the transverse direction to the skeleton. The upward direction is defined as the direction of the node with the highest density. The white horizontal line represents the smoothing length used in the analysis. The sub-panels on the top and the right of each panel show the marginalized 1D distributions of R and z, respectively. Note that the behaviour of the gradient of the number density of galaxies changes with stellar mass, noticeably in high-mass bin. As expected, high-mass galaxies are more tightly clustered near the filament axis and near nodes (right-hand panel) compared to their lower mass counterparts (middle and left-hand panels).

limit of 10^9 M$_\odot$. Each of such constructed stellar mass bin contains ∼3500 galaxies. The smoothing scale applied to the profiles is 0.4 Mpc h^{-1}.[5]

Figs 1 and 2 show the galactic number counts at low, intermediate, and high stellar mass, and mean stellar mass for all galaxies above the stellar mass limit, respectively, at redshift zero in the frame of the saddle. In that frame, the vertical axis corresponds to the distance from the saddle point along the skeleton, upwards towards the densest node, while the horizontal axis corresponds to the transverse direction. Note that the length of filaments is not constant, however its distribution is quite narrow with median length of ∼5.5 Mpc h^{-1} at redshift zero (see Appendix B, Fig. B1). Iso-contours clearly display a dependence both on the radial distance from the saddle point and the orientation with respect to the filament's direction. At fixed distance from the saddle point, the number of galaxies is enhanced in the direction of the filament, i.e. they are more clustered in the filaments than in the voids. The gradient of the number density of galaxies is also found to change with stellar mass. The high-mass galaxies are more tightly clustered near the filament axis and tend to be further away from saddles along the filament compared to their low-mass counterparts. Saddle points are, as expected, local minima of both galaxy number counts and stellar mass in the direction along the filament towards the nodes, and local maxima in the perpendicular direction. Thus, galaxies in filaments tend to be more massive than galaxies in voids and within filaments, while the stellar mass of galaxies increases with increasing distance from the saddle
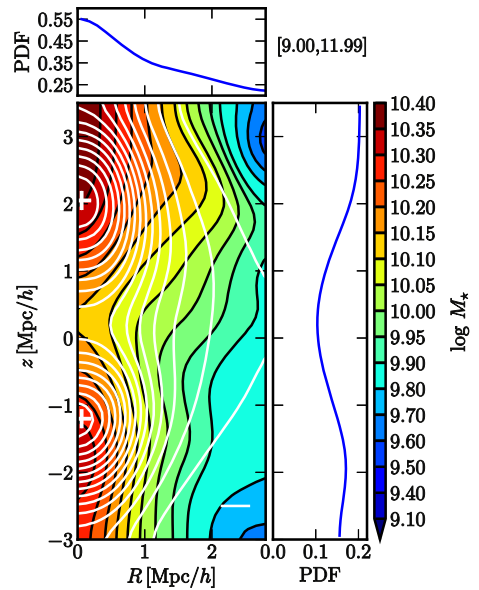
Figure 2. Mean stellar mass in the frame of the closest saddle for the entire galaxy population with masses in the range 10^9– 10^{12} M$_\odot$ at redshift zero. The white curves correspond to the contours of the galaxy number counts, while the white crosses represent the peaks in galactic number density on axis. More massive galaxies are further away from the saddle (respectively, closer to the saddle) than the low-mass population in longitudinal (respectively, transverse) direction.

point in the direction toward nodes. This effect is stronger in the direction perpendicular to the filament, where the relative variation of the mean stellar mass is about a factor of 2 higher compared to that along the filament.

---

[5]Changing the smoothing scale used to produce the maps to 0.2 and 0.8 Mpc h^{-1} leads to qualitatively similar conclusions. The smoothing impacts mostly the position of maxima in the transverse direction. At low values, these tend to be offset from the filament's axis because of the smoothing of the skeleton itself.
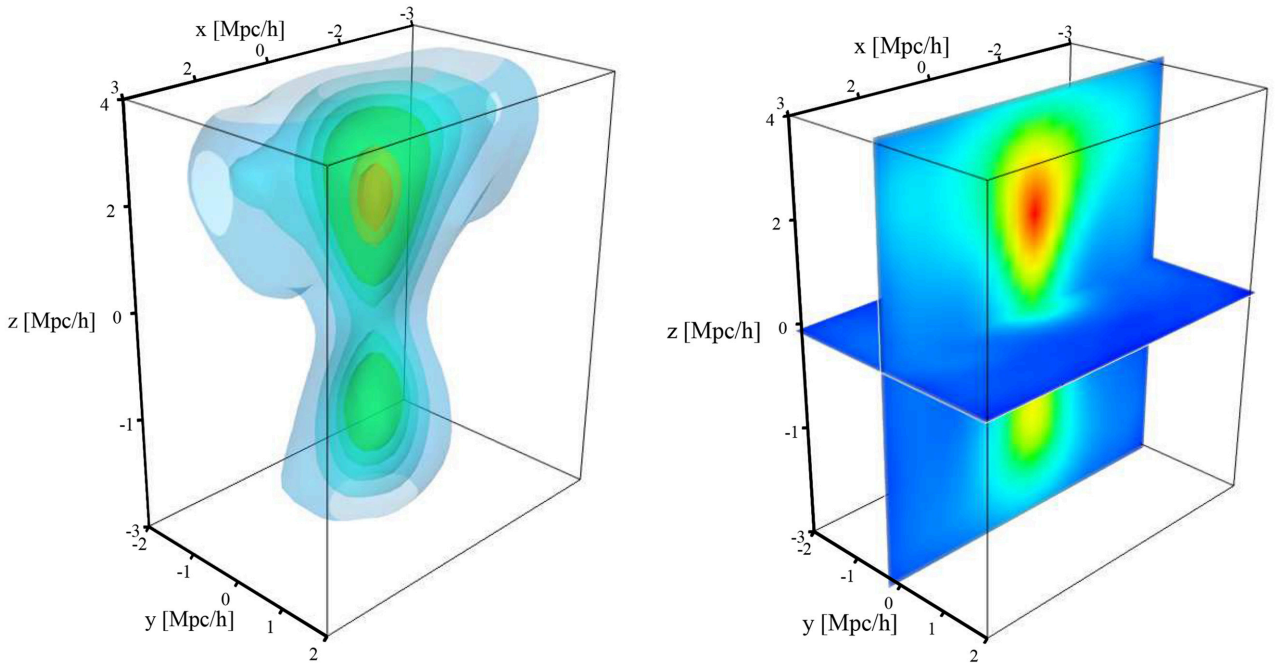
Figure 3. 3D structure of the neighbourhood of filaments at redshift zero. The galaxy number counts in the frame of the saddle for masses in the range 10^9–10^{12} M$_\odot$ (left) are shown together with two 2D cross-sections, longitudinal and transverse, of the filament at the saddle (right). The flattened flaring away from the saddle reflects the co-planarity of filamentary bifurcation within the wall. The top–bottom asymmetry reflects the orientation of the skeleton.

The mass gradients shown on Fig. 2 can be qualitatively understood within peak and excursion set theories (see Section 5 and Codis et al. 2015b; Musso et al. 2018).

3.2 3D stacks of stellar mass and number density

Let us now investigate the 3D structure of the neighbourhood of filaments by stacking galaxies relative to a 3D-oriented local reference frame, with its origin defined by the position of the saddle point and its axes defined as follows: the z-axis corresponds to the direction of the filament at the saddle, and the x- and y-axes represent major and minor principal axes of the inertia tensor in the plane perpendicular to the filament axis at the saddle point, respectively.[6]

In order to increase the signal-to-noise ratio, galaxies are stacked by flipping them with respect to the filament axis to produce longitudinal cross-sections, and with respect to both principal axes of the inertia tensor in the case of transverse cross sections.

The 3D distribution of galaxies in such defined frame is shown in Fig. 3 (left-hand panel) together with planes representing 2D cross-sections, longitudinal and transverse, as used in the analysis (right-hand panel). In practice, individual cross-sections are obtained by projecting galaxies within ± 1 and ± 0.75 Mpc h^{-1} from the plane passing through the saddle point for longitudinal and transverse cross-sections, respectively. Note the flaring near the nodes which arises because the typical saddle is flattened (the two negative

---

[6]In practice, the 2D inertia tensor is computed by considering galaxies within ± 1 (Mpc h^{-1}) around the saddle point and projected into the plane perpendicular to the filament and passing through the saddle. Note that changing the volume of the considered region within a factor of a few does not have a strong impact on orientation.

eigenvalues of the Hessian differ, while the corresponding eigenvectors are aligned when stacking), and the Hessian remains correlated away from the saddle. Correspondingly, the skeleton bifurcates within that plane (Pogosyan et al. 2009; Codis, Pogosyan & Pichon 2018). The top–bottom asymmetry reflects the fact that higher density contours are drawn near the more prominent peak (which is traced by the orientation of the skeleton).

As in the case of azimuthally averaged cross-sections, three stellar mass bins are defined as low ($9.0 \leq \log M_\star \leq 9.05$), intermediate ($9.7 \leq \log M_\star \leq 9.77$), and high ($10.96 \leq \log M_\star \leq 11.99$) stellar mass bins, containing $\sim 10\,000$ and ~ 1000 galaxies, for longitudinal and transverse cross-sections, respectively. The upward direction along z-axis corresponds to the direction of the node with highest density, and the smoothing scale applied to the profiles is 0.4 Mpc h^{-1}, as previously.

The cross-sections of galactic number counts, stellar mass, specific star formation rate, sSFR =SFR/$M_\star$, where SFR is computed over a time-scale of 50 Myr, V/σ, and age will be studied in the vicinity of the saddle. Figs 4 and 5 show the galaxy number counts in three different stellar mass bins, and mean stellar mass for all galaxies above the stellar mass limit, respectively, at redshift zero in the longitudinal (top panels) and transverse (bottom panels) planes in the frame of the saddle. Once again, iso-contours clearly depend on both the radial distance from the saddle and the orientation with respect to the filament's direction. Galaxies are found to be more clustered in filaments than in voids at all masses, i.e. at fixed distance from the saddle point, the number of galaxies is enhanced in the direction of the filament. What changes with stellar mass is the behaviour of the gradients with the most massive galaxies being more tightly clustered near the filament axis compared to their lower mass counterparts. As in the case of azimuthally averaged cross sections, mass gradients seen in Fig. 5 (left-hand panels) can be also understood in the context of con-

Figure 4. The galaxy number counts at redshift zero in the frame of the saddle for low (left), intermediate (middle), and high (right) stellar mass bins (see the text for definition), as labelled (in square brackets), in the longitudinal (top) and transverse (bottom) planes at the saddle. The vertical axis on top panels corresponds to the direction of the skeleton at the saddle (upwards toward the node with the highest density), while the horizontal axis corresponds to the major principal axis in the transverse direction. The sub-panels on the top and the right of each panel show the marginalized 1D distributions along respective axes. The white dashed contours represent the galaxy number counts with the horizontal axis corresponding to the minor principal axis in the transverse direction at the saddle. The black crosses represent the peaks in galactic density on axis and the white horizontal line represents the smoothing length used in the analysis. The projection is carried over ± 1 Mpc h^{-1} for the longitudinal slice and ± 0.75 Mpc h^{-1} away from the saddles transversally. The strength of the gradient of the galaxy number density changes with stellar mass. As expected, the high-mass galaxies are more tightly clustered near the filament axis and near nodes (right-hand panel) compared to their low-mass counterparts (left-hand and middle panels).

strained random field and excursion set theory, as discussed in Section 5.

Interestingly, the distribution of most massive galaxies around the saddle points in the transverse direction is axisymmetric up to the distance of ~ 1 Mpc h^{-1}, while the iso-contours of lower mass galaxies are more flattened (in the direction of x-axis corresponding to the major axis of the inertia tensor in the transverse cross section) and extended to larger distances from the saddle. This behaviour is a manifestation of the mass dependence of galaxy's connectivity: higher mass galaxies in denser environments are expected to be fed by numerous filaments, while lower mass galaxies are typically embedded in a single filament (Codis et al. 2018).

3.3 Longitudinal and transverse sSFR cross-sections

Let us now focus on sSFRs. Fig. 6 (top row) shows the mean stellar mass-weighted sSFR at redshift zero in HORIZON-AGN. Iso-contours display qualitatively similar behaviour in all stellar mass bins in the direction perpendicular to the filament, for which the saddle point represents the maximum of sSFR. In the direction along the filament, the behaviour is more complex: at high stellar mass sSFR increases with increasing distance from the saddle towards the

nodes, but while the maximum of sSFR overlaps with the position of the low-density node, it is located closer to the saddle in the direction of the densest node, as will be discussed below. The sSFR then decreases in this direction in the vicinity of the node and beyond. With decreasing stellar mass, the maximum of sSFR moves closer to the saddle point, until it overlaps with the saddle point for lowest stellar mass bin.

A general trend of decreasing sSFR with increasing stellar mass is clearly recovered, with most massive galaxies having their sSFR substantially reduced in particular in the vicinity of the densest node, where the average sSFR value can be up to 10 times lower compared to their low-mass counterparts. Indeed AGN feedback is an important ingredient for the formation of the more massive galaxies, suppressing star formation so as to reproduce the observed high end of the galaxy luminosity function. By comparing the iso-contours of mean stellar mass-weighted sSFR in HORIZON-AGN and HORIZON-NOAGN (bottom row of Fig. 6), the two main specific consequences of AGN feedback can be identified. First, and not surprisingly, when AGN feedback operates, the overall sSFR is reduced, mostly in the high stellar mass bin (the mean sSFR in the highest stellar mass bin changes by a factor of ~ 3, while in the lowest stellar mass bin, it remains ~ 1.15). Secondly, AGN
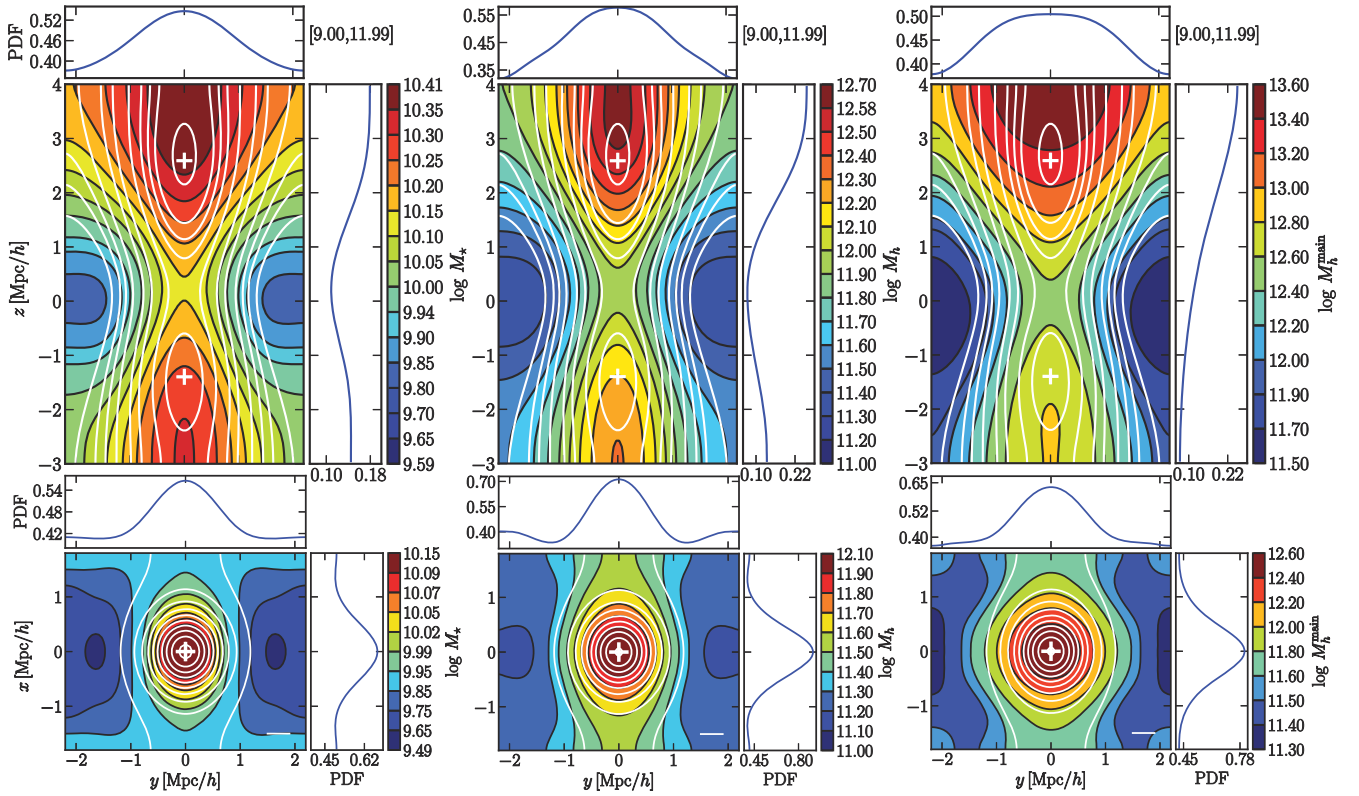
Figure 5. Mean galaxy stellar mass (left), sub-halo mass (middle), and host halo mass (right) in the frame of the saddle for masses in the range 10^9–10^{12} M$_\odot$ at redshift zero, in the longitudinal (top panels) and transverse (bottom panels) planes at the saddle. The vertical axis corresponds to the direction of the skeleton at the saddle (upwards toward the node with the highest density), while the horizontal axis corresponds to the major principal axis in the transverse direction. The white contours correspond to the galaxy number counts with the horizontal axis corresponding to the major principal axis in the transverse direction at the saddle. The white cross represents the peak in galactic density on axis. More massive galaxies are further away from the saddle than the low-mass population in the longitudinal direction, while they are closer to the saddle transversally. As expected, more massive galaxies are also residing in more massive haloes. Note in particular that the iso-contours of stellar mass are very similar to those of sub-halo mass, while the iso-contours of host halo mass, the shape of which differ from the two others, show much more resemblance to the iso-contours of density (as discussed in Section 6). The peak of maximum mass is further away from the saddle than the counts.

feedback modifies the shape of sSFR iso-contours. This effect is most prominent amongst most massive galaxies[7] in the vicinity of the densest node that represents the maximum of the sSFR in the direction along the filament from the saddle when AGN feedback is absent. A similar effect is seen at low and intermediate stellar mass, albeit less pronounced. Overall, the reduced star formation activity of galaxies due to AGN feedback in the densest environment translates into an offset of the maximum of the mean stellar mass-weighted sSFR away from the node. This clearly demonstrates the importance of the AGN feedback and its ability not only to reduce the star formation activity of individual objects, but also to modify their distribution on larger scales in the vicinity of high-density regions such as nodes, corresponding to galaxy groups and clusters,

consistently with our findings of AGN feedback being most efficient near nodes at high stellar mass (see Appendix D).

3.4 Centrals and satellite differential counts

In order to gain a better understanding of what processes regulate sSFR of galaxies in their anisotropic environment, galaxies are next split into centrals and satellites (respectively the most massive galaxy within 10 per cent the current virial radius of halo, or sub-halo). Making this separation is further motivated by more straightforward comparison with theoretical prediction of Musso et al. (2018, that is strictly applicable to central galaxies alone, as the effect of the large-scale tidal field on the low-mass objects is not accounted for). Fig. 7 shows stellar mass-weighted sSFR for centrals (top panels) and satellites (bottom panels) separately in both simulations, HORIZON-AGN (leftmost panels) and HORIZON-NOAGN (rightmost panels) in low and high stellar mass bins. Not surprisingly, the low-mass end is dominated by the population of satellites, while central galaxies dominate the highest stellar mass bins. What is more interesting is the distinct response of centrals and satellites in terms of their sSFR as a function of the exact position within the cosmic web (in both HORIZON-AGN and HORIZON-NOAGN) and more surprisingly, the distinct impact of AGN feedback on the sSFR of these two populations.

---

[7]Note that the highest stellar mass bin is not identical in the two simulations. This is due to the difference in the stellar mass distributions, such that at high stellar mass end, there are more galaxies in HORIZON-NOAGN than in HORIZON-AGN that also tend to be more massive (see also Beckmann et al. 2017). However, considering the same stellar mass bins does not impact our results. Another difference is in the halo-to-stellar mass relation, especially at the high-mass end. It was checked that the medians of halo masses in the highest stellar mass bin considered in this work are comparable in both simulations.

Figure 6. Mass-weighted sSFR in the frame of the saddle at redshift 0 for low (left), intermediate (middle), and high (right) stellar mass bins, as labelled, in the longitudinal and transverse planes at the saddle, in HORIZON-AGN (topmost panels) and HORIZON-NOAGN (bottommost panels). The vertical axis corresponds to the direction of the skeleton at the saddle (upwards toward the node with the highest density), while the horizontal axis corresponds to the major principal axis in the transverse direction at the saddle. The white contours and the white crosses correspond to the galaxy number counts and the peak in galactic density on axis, respectively. The saddle represents maximum of sSFR in transverse direction at all masses and regardless of the presence of the AGN feedback. What does change is the star formation activity in particular of the most massive galaxies, where AGN feedback substantially reduces the values of sSFR. Moreover, note that at high-mass end, the sSFR iso-contours are modified by AGN feedback in the vicinity of the densest node, such that in the longitudinal direction away from the saddle, the maximum of sSFR is offset from the densest peak. Overall, the sSFR iso-contours display a stellar mass dependence in the longitudinal direction in that at low-mass (respectively, high-mass) sSFR is maximum (respectively, minimum) at the saddle and it decreases (respectively, increases) in the direction towards the nodes.
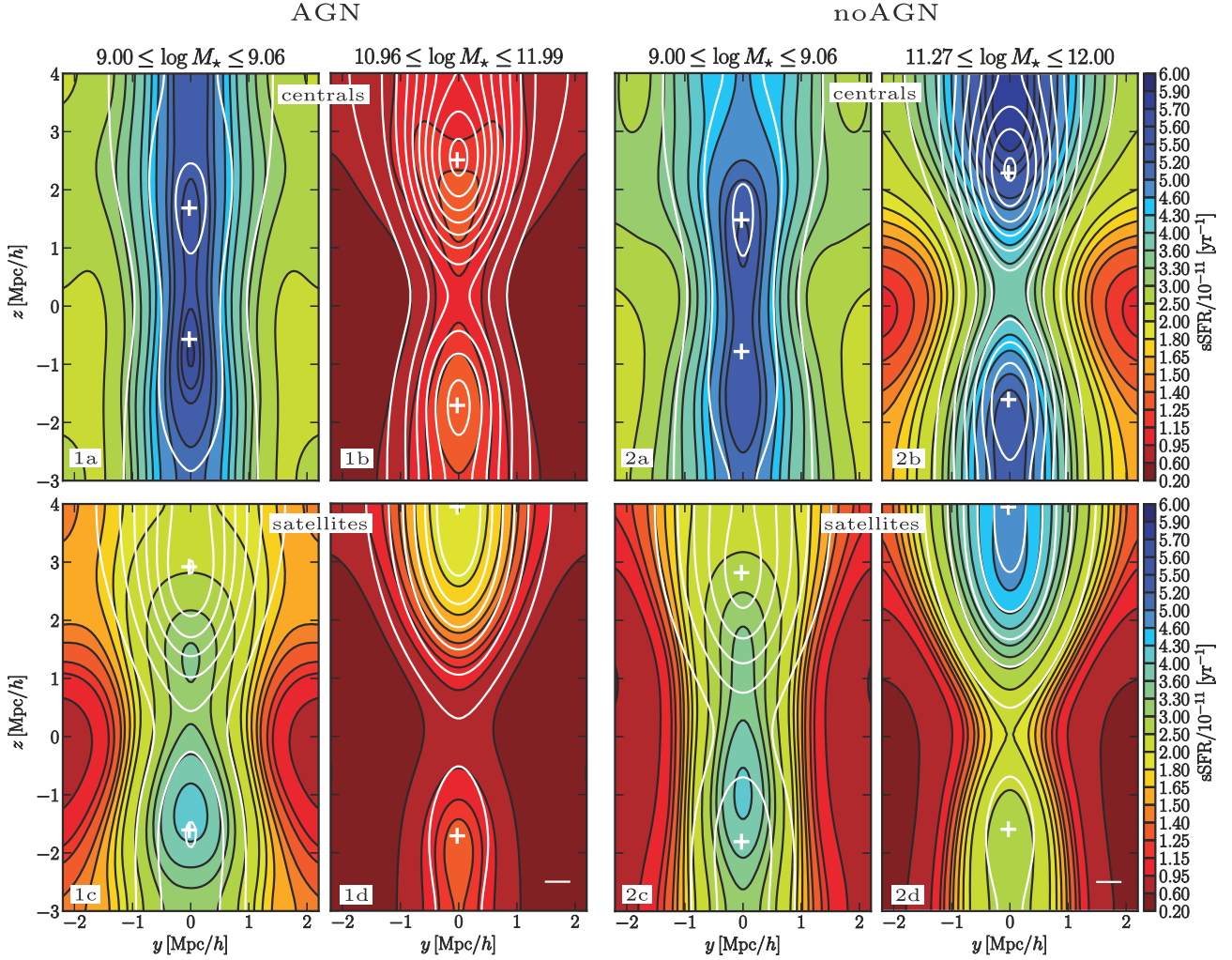
Figure 7. Mass-weighted sSFR in the frame of the saddle at redshift 0 for low and high stellar mass bins, as labelled, in the longitudinal and transverse planes at the saddle, shown for centrals (top row) and satellites (bottom row) separately with (left) and without (right) AGN feedback. The vertical axis corresponds to the direction of the skeleton at the saddle (upwards toward the node with the highest density), while the horizontal axis corresponds to the major principal axis in the transverse direction at the saddle. The white contours and the white crosses correspond to the galaxy number counts and the peak in galactic density on axis, respectively. AGN feedback has the strongest impact on high-mass centrals and in the vicinity of the densest node (compare panel 1b with 2b and with panel 1d), where it modifies the shape of the sSFR iso-contours as already noticed for the entire high-mass population (see Fig. 6). At low stellar mass, satellites are generally less star forming compared to centrals, but note also that the sSFR iso-contours of centrals and satellites are also different. For satellites, the maximum of sSFR is located between the saddle and the peak in the direction along the filament towards the densest node (compare panel 1a with panel 1c or panel 2a with panel 2c).

AGN feedback seems to have a stronger impact on centrals which are closer to the denser node (compare panel 1b with 2b and 1d with 2d). At high mass, AGN feedback quenches much less efficiently star formation in satellites than it does in centrals (compare panel 1b with 1d), where it distorts the shape of the sSFR iso-contours in the vicinity of the denser node. High-mass satellites seem to feel the impact of both the AGN feedback and environmental processes, in particular in dense regions, but less so than the centrals (compare panel 1d with 2d). A possible explanation for massive satellites being less affected by the AGN feedback (compared to centrals at the same stellar mass) could be the tidal influence of their main halo (Hahn et al. 2009) which reduces accretion and merger rate onto the satellite. As mergers trigger bursts of AGN activity, this induces less star formation.

At low stellar mass, as expected, AGN feedback does not seem to have a strong impact on the sSFR of both satellites and centrals

(compare panel 1a with 2a and 1c with 2c). At low stellar mass, sSFR iso-contours are different for satellites and centrals: (i) satellites have lower sSFR compared to centrals of the same mass, in the direction both perpendicular to the filament, and along the filament towards the nodes, and (ii) the shape of sSFR iso-contours is different for satellites and centrals, in particular in the vicinity of denser node, in that for satellites, the sSFR reaches its maximum before reaching the densest node in the direction along the filament (compare panel 1a with 1c or 2a with 2c). Presumably, satellite specific processes, such as e.g. strangulation, are driving this difference.[8] Note that the sSFR contours for massive centrals in the

[8] Strangulation (Larson, Tinsley & Caldwell 1980), together with mergers (Toomre & Toomre 1972), are traditionally considered as group-specific processes impacting star formation activity of satellites. Other environmental

HORIZON-NOAGN simulation (see panel 2b) are, as expected, in qualitative agreement with the DM accretion predicted by Musso et al. (2018, see Section 6 for a more detailed discussion).

3.5 Longitudinal and transverse kinematic/age cross-sections

Let us finally focus on the kinematics, quantified by the ratio of rotation to dispersion-dominated velocity, V/σ, and the age of galaxies in the frame of the saddle. The observational proxies of these quantities would be morphology and colour, respectively. Higher V/σ typically characterizes disc-dominated morphologies, while lower V/σ indicates the presence of a substantial bulge component. The age of galaxies corresponds to the mean ages that are given by the mass-weighted age of star particles belonging to each galaxy. Fig. 8 shows iso-contours of V/σ (top panels) and age (bottom panels) as a function of stellar mass at redshift zero. Again, the contours exhibit both radial and angular gradients with respect to the saddle point. At all stellar mass bins, galaxies tend to have higher V/σ in the vicinity of the saddle point that decreases in the orthogonal direction away from the saddle, while in the direction along the filament towards the nodes it first increases, reaches its maximum before getting to the densest node and decreases afterwards. This effect is strongest for highest mass galaxies. In terms of quantitative comparison of V/σ at different stellar mass, galaxies in the lowest stellar mass bin have the lowest V/σ, while intermediate-mass galaxies show the largest V/σ values. V/σ of the most massive galaxies is lower compared to intermediate stellar masses, but higher than at lowest stellar mass end. This can be explained by the presence of few massive disc-dominated galaxies present in the HORIZON-AGN simulation and higher fraction of ellipticals at low-mass end compared to observations. Indeed, as shown in Dubois et al. (2016), the maximum probability of finding discs in HORIZON-AGN is in the stellar mass range of 10^{10}–$10^{11} \, M_\odot$.

Similarly, age gradients display clear radial and angular dependence with respect to the saddle point at all stellar mass bins, however, with qualitatively different behaviour. In the transverse direction, saddle point is still maximum of the age at all stellar mass, while in the direction along the filament away from the saddle, age increases all the way beyond the node. Interestingly, in this aspect, age gradients are similar to stellar mass gradients with the oldest and most massive galaxies being located closer to the node in the direction of the filament, and in the vicinity of the filament in the orthogonal direction. This is consistent with the redshift evolution of the stacks as discussed now.

4 REDSHIFT EVOLUTION

Let us now examine the evolution of galaxy properties with redshift. When comparing different epochs one may either consider the fate of a given set of galaxies, or quantify the cosmic evolution of the galactic population as a whole.

Fig. 9 shows galaxy number counts in low (left-hand column), intermediate (middle column), and high (right-hand column) stellar mass bins at redshifts two (topmost rows) and one (bottommost

rows),[9] while Fig. 10 shows the mean stellar mass of the entire population above the mass limit at these redshifts, as indicated.[10] The corresponding redshift zero maps are shown on Figs 4 and 5, respectively.

At each redshift, more massive galaxies are more tightly clustered in the filaments than in the voids, and near the nodes than near the saddles. Part of this redshift evolution is simply due to the mass evolution of objects. In other words, one could fix the level of non-linearity by considering mass bins that evolve with redshift following the non-linear mass for instance and then consider the residual redshift evolution. This procedure would allow to focus on the same class of objects across redshifts.

On Fig. 9, one can follow the progenitors of a given class of objects by fixing the level of non-linearity which is equivalent to move approximatively along the diagonal (by adding Fig. 4), i.e. to focus on less massive objects at high redshift. As galaxies grow in mass, i.e. as non-linear gravitational clustering proceeds (the local dynamical clocks being set by inverse square root of the local density), they also become more concentrated towards the filaments and nodes (see Appendix E2). For instance, comparing the bottom right transverse cross-section at redshift one and zero (from Fig. 4), the vicinity of the saddle is less populated by massive objects as these have drifted towards the nodes. This redshift evolution is consistent with the global flow of galaxies first towards the filaments and then along them (as quantified kinematically in Appendix F), and with the fact that galaxies accumulate mass with cosmic time.

For a population as a whole, in the close vicinity of the saddle, the breadth of the filament broadens with cosmic time as shown in Fig. 11, comparing the filament's thickness for all galaxies above the stellar mass limit at redshifts two and zero. Specifically, the full width at half-maximum (FWHM) of the transverse galaxy number counts profiles was computed at different positions along the filament's direction. As argued in the next section, the measured increase of the filament's width with cosmic time is consistent with the theoretical expectations.

Finally, Fig. 12 shows the redshift evolution of stellar mass-weighted sSFR. Again, it is interesting to note that the global sSFR traces the level of non-linearity of the collapse of structures: at high-redshift, low-mass population (top left panel) has the highest sSFR, whereas the high-mass low-redshift population (bottom right panel) is the most quenched. This is also reflected in the position of maximum of sSFR, which drifts with cosmic time i.e. with the level of non-linearity of the field. The peak of sSFR seems to occur further from the denser nodes towards the saddle as a function of cosmic time. Hence, for the sSFR at least two processes compete: advection with the main flow and star formation activity which is impacted by the proximity to AGNs and the local dynamical time-scale (but see Section 6.6 below).

5 THEORETICAL PREDICTIONS

Let us briefly present the theoretical framework which will allow us to interpret the measurements presented in Sections 3 and 4. This will involve predictions for DM and halo density cross-sections in

---

quenching processes, mostly operating in clusters include galaxy harassment (Moore et al. 1996) or ram pressure stripping of gas (Gunn & Gott 1972). However, in this work, we are not attempting to address the processes impacting satellite population in particular.

[9]The skeleton and stellar mass bins are constructed as for redshift zero, see Sections 2.3 and 3.2, respectively. Consequently, the stellar mass bins are not identical at different redshifts, but they still contain comparable number of galaxies.

[10]Note that these cross-sections are in qualitative agreement with azimuthally averaged counterparts (see Appendix C).
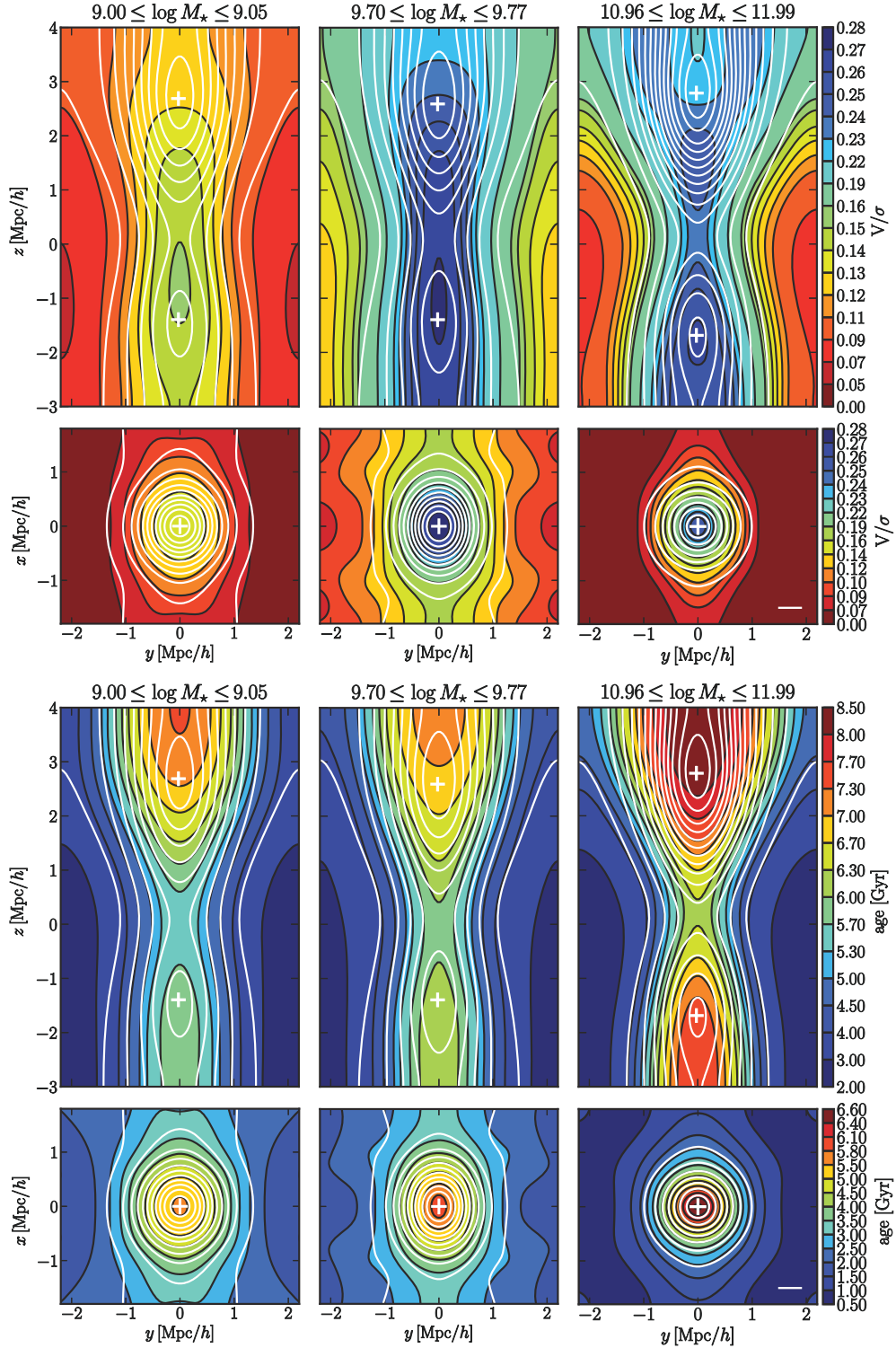
Figure 8. Stellar mass weighted V/σ (topmost panels) and age (bottommost panels) for low (left), intermediate (middle), and high (right) stellar mass bins as labelled, in the longitudinal and transverse planes at the saddle. The vertical axis corresponds to the direction of the skeleton at the saddle (upwards toward the node with the highest density), while the horizontal axis corresponds to the major principal axis in the transverse direction at the saddle. The white contours and the white crosses correspond to the galaxy number counts and the peak in galactic density on axis, respectively. The behaviour of the number density of galaxies changes with stellar mass for both physical properties, but much more dramatically for V/σ. The shape of iso-contours are qualitatively different, while the maximum of V/σ along the filament in the upward direction is located in between the saddle point and the density peak, age increases with increasing distance away from the saddle towards the nodes, and beyond. Transverse gradients are similar, both in terms of shape of iso-contours and in that the saddle point is maximum for both quantities in radial direction.
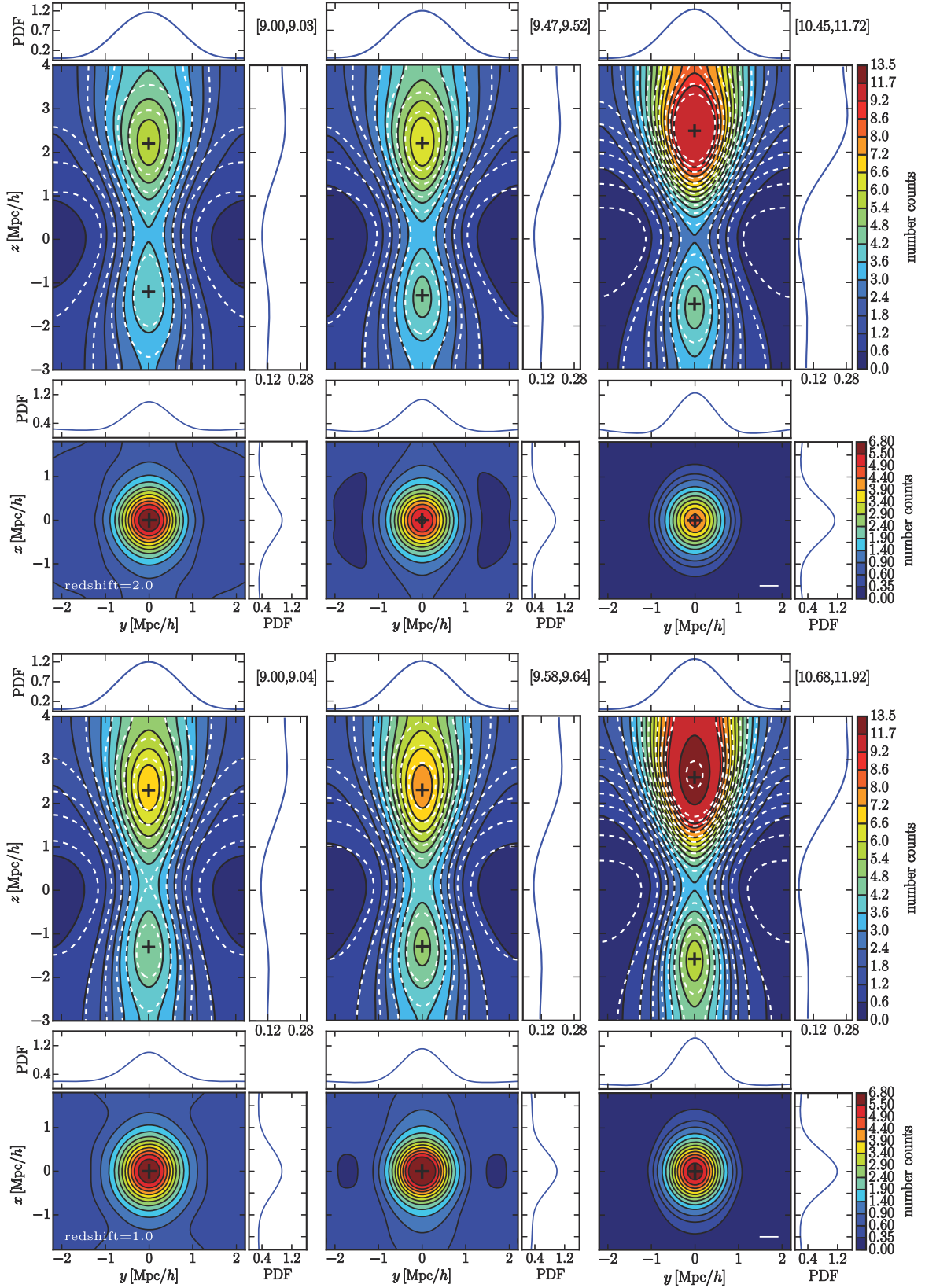
Figure 9. Redshift evolution of the galaxy number counts in the frame of the saddle, in the longitudinal and transverse planes at the saddle. Low (left-hand column), intermediate (middle column), and high (right-hand column) stellar mass bins are shown at redshifts 2 (topmost panels) and 1 (bottommost panels), respectively. The white dashed contours represent the galaxy number counts with the horizontal axis corresponding to the minor principal axis in the transverse direction at the saddle. The corresponding redshift zero maps are shown on Fig. 4. High-mass galaxies are more clustered near the filaments and nodes at all redshifts considered compared to their lower mass counterparts. Note that as galaxies grow in mass with time, they follow the global flow of matter, reflected by the increased distance between the saddle point and two respective nodes at lower redshift.
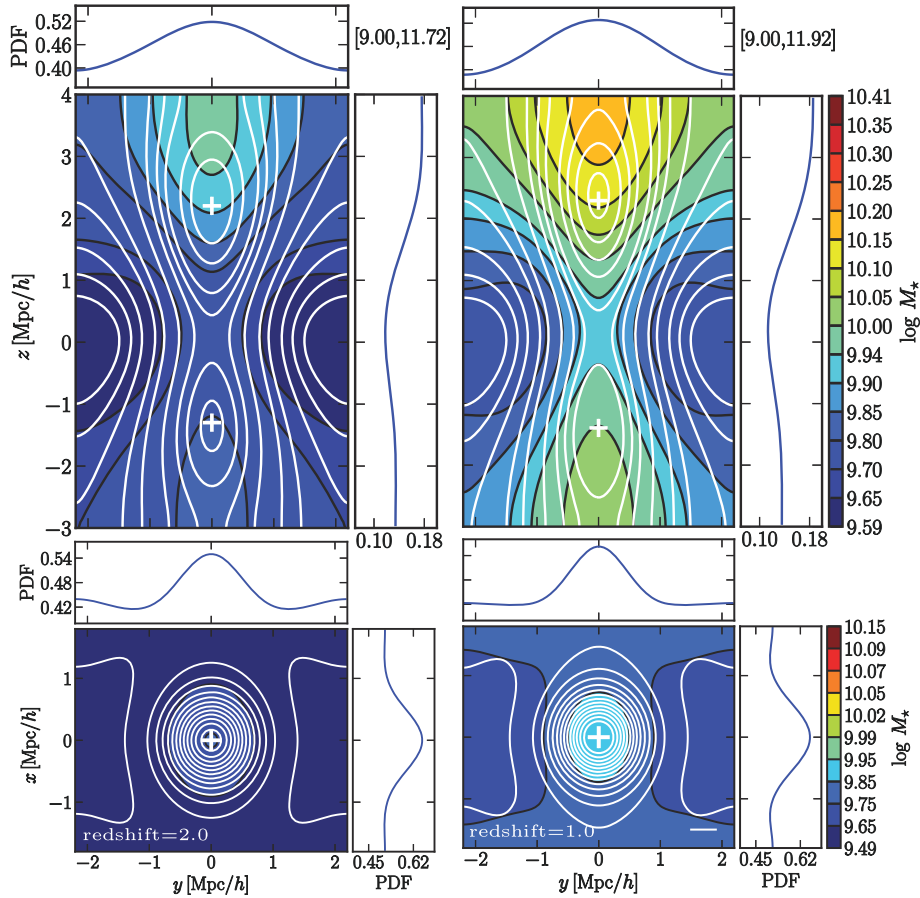
Figure 10. Mean galaxy stellar mass in the frame of the saddle, in the longitudinal and transverse planes at the saddle, for all masses in the range $10^{9.0}$–$10^{12.0}$ M$_\odot$ at redshifts two (left-hand column) and one (right-hand column). The redshift zero maps are shown on Fig. 5. At a given mass, the corresponding (coloured) contours get further away from the filament axis with cosmic time. Transverse cross-sections (bottom panels) of number counts (white contours) become more elongated with decreasing redshift, while longitudinally (top panels), they are further away from the saddle at lower redshift: the filaments become more elliptical and thicken with cosmic time (see Fig. 11 for quantitative estimate of this effect).

the frame of the saddle, and their expected non-linear evolution with cosmic time.

5.1 Constrained random fields

For Gaussian cosmological initial conditions, peak theory (Bardeen et al. 1986) can be adapted to predict the mean (total) matter density maps around saddles. Appendix E derives this mean initial matter distribution marginalized to the constraint of a saddle point of arbitrary geometry (height and curvatures) when the direction of the largest (positive) eigenvalue of the Hessian, i.e. the direction of the filament, is fixed together with its orientation. This last requirement is achieved by imposing that the coordinate of the gradient of the gravitational potential along the filament is always negative. The resulting oriented map of the density distribution around saddles is shown in Fig. 13 (left-hand and middle panels). As expected, more mass is found close to the filament axis and in the direction of the most attractive potential well (towards the top of the map). Fig. 13 (right-hand panel) also presents the expected mass distribution of DM haloes within the frame of the saddle when the Press–Schechter threshold for collapse is decreased by the mean density (following the prescription described in Codis et al. 2015b).

5.2 Expected redshift evolution

Different approaches can be used to incorporate the non-linear evolution in the theoretical predictions, e.g. by doing a Zel'dovich boost of the mean density map predicted from excursion set theory, or by incorporating the gravity induced non-Gaussianity of the distribution using a perturbative approach as sketched in Appendix E. Both predict that gravitational clustering distorts and enhances the contours of the matter density field within the frame of the saddle, with a scaling proportional to σ ($M_\star$,redshift), the mass- and redshift-dependent scale of non-linearity. The net effect will depend on what is held fixed while stacking. At fixed rareness, which is essentially achieved when focusing on the more massive objects, filaments will collapse with cosmic time and therefore get thinner and more concentrated (see Fig. E1). On the other hand, when the entire population of galaxies is considered at each redshift, filaments typically get thicker, because less rare and therefore less-connected and less-biased objects form at low redshift and dominate the population.

While the realm of these predictions is limited (in redshift and range of tracers), it none the less allows us to understand the trend at the level of gravity-driven processes, and highlight by contrast the contribution of AGN or stellar feedback. We refer to Codis et al.
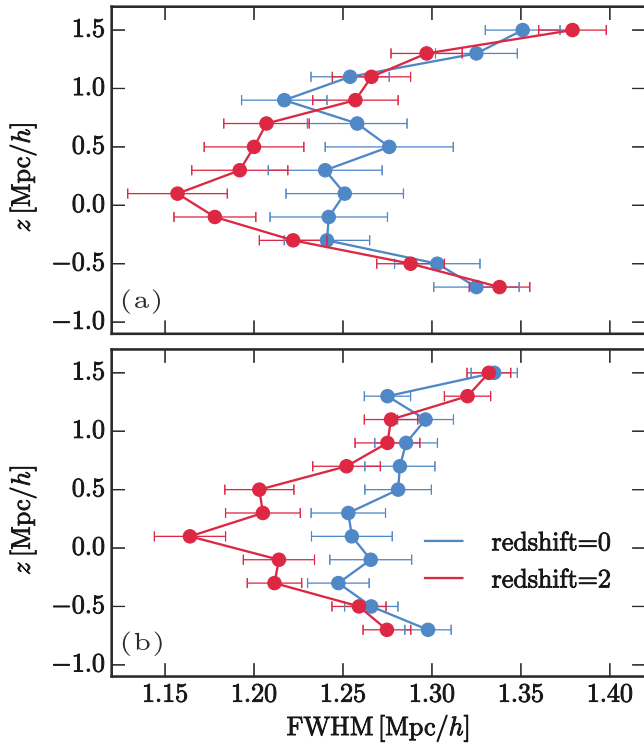
Figure 11. Thickness of the filaments, defined as the FWHM of the Gaussian fit of the transverse galaxy number counts profiles marginalized over *x*-(panel a) and *y*-axes (panel b) at different positions along the filament's direction (*z*-axis on the longitudinal cross-sections) at redshifts 2 and 0, in red and blue, respectively, for all galaxies in the mass range 10^9–10^{12} M$_\odot$. The transverse projections are carried over 0.2 Mpc h^{-1} longitudinally (along the *z*-axis). As previously, the upward direction along the *z*-axis corresponds to the direction of the skeleton at the saddle toward the node with the highest density. When considering the entire population of galaxies, the cross-sections of filaments in the vicinity of the saddle point (at $z = 0$ Mpc h^{-1}) grow with time. For the sake of clarity, only measurements at redshifts zero and two are shown, however, their redshift evolution is consistent throughout. Note also that the widths are computed in comoving coordinates: the growth at low redshifts is much stronger in physical coordinates. See also Appendix B (Fig. B2) for the thickness of the filaments and its redshift evolution at distances extending more faraway from saddle.

(2015b, their section 4) and Laigle et al. (2015, their section 5) for predictions for the expected angular momentum and vorticity distributions and their evolution in the frame of the saddle, which will prove useful when discussing V/σ maps (and less directly sSFR maps, which are sensitive to the recent accretion of cold gas).

6 INTERPRETATION AND DISCUSSION

Let us now discuss the findings of Sections 3 and 4 in the context of existing surveys and structure formation models (Section 5).

6.1 Complementary top-down approach to galaxy formation

Let us start by putting the adopted approach and the results of this work in the classical context of structure formation models. Traditionally, galaxy formation and evolution is studied in the hierarchical framework where galaxies are considered as evolving in (sub)-haloes possibly embedded in larger haloes (e.g. Kauffmann, White & Guiderdoni 1993; White 1996). Dynamically, this means that we can associate two typical time-scales

(or 'clocks') to each encapsulated environment. This approach is justified in the well-established bottom–up scenario of structure formation. One can address the impact of the isotropic environment on the scales of haloes, or equivalently the local density (i.e. the trace of the Hessian of the gravitational potential) while considering the merger tree history of individual haloes (and thus galaxies residing within).[11] Such scenario has proven quite successful in explaining many observed properties of galaxies, via the so-called halo model (Cooray & Sheth 2002) – in particular against isotropic statistics (e.g. two-point functions). In this classical view, the impact of the larger anisotropic scales set by the cosmic web is ignored because it is assumed that these scales do not couple back down to galactic scales. Yet this view fails to capture e.g. spin alignments which are specifically driven by scale coupling to the cosmic web (Codis et al. 2015a), nor does it fully take into account how the light-cone of a given galaxy is gravitationally sensitive to the larger scale anisotropies.

By contrast, Musso et al. (2018) recently investigated the impact of the large-scale anisotropic cosmic web on the assembly history of DM haloes within the framework of extended excursion set theory, accounting for the effect of its large-scale tides. They derived the typical halo mass, typical accretion rate, and formation time of DM haloes as a function of the geometry of the saddle. These quantities were predicted to vary with the orientation and distance from saddle points, such that haloes in filaments are less massive than haloes in nodes, so that at equal mass, they have earlier formation times and smaller accretion rates at redshift zero, the effect being stronger in the direction perpendicular to the filament. These findings suggest that on top of the mass and local mean density, the tides of the larger scale environment also impact haloes' properties through a third time-scale.

The approach adopted here follows up and assesses specifically the impact of this large-scale environment on *galaxy* properties, and in particular the top–down relevance of the imposed tides (captured by the traceless part of the Hessian of the gravitational potential) on galaxy assembly. In other words, the aim here is to identify properties of galaxies which are specific to their relative position within the saddle frame. To do so, the analysis is carried out at fixed stellar mass and quantified at additional fixed (sub)-halo mass and anisotropic density (through the analysis of stacked re-oriented residual maps, see below), instead of the conventional galaxy–halo–group mass isotropic perspective. This framework does not invalidates past results expressed in terms of group and halo masses – which remain the dominant effect impacting galaxy formation, but complements them at first- or second-order corrections.[12] Qualitatively, the aim is to understand the impact of the stretching and twisting imposed by those tides above and below the impact of the density. As shown in Section 5.2, it also provides as a bonus a good understanding of the bulk flows within that frame, which enlightens the geometry of filaments' iso-contours traced by galaxies at fixed mass or fixed cosmic age.

---

[11]The local density is indeed strongly correlated with the group halo mass, as can be seen by comparing e.g. Figs 5 and 15.

[12]In fact one could indeed alternatively extend the classical framework by adding the larger scale group distribution, i.e. the cosmic web traced by DM haloes as an extra 'hidden variable' driving galactic assembly. Below that scale, the statistics is isotropic, while beyond it one has to define how ensemble average should be carried. The frame of its saddles is chosen here as a proxy for this web so as to be able to stack galactic distributions while taking its effect into account.
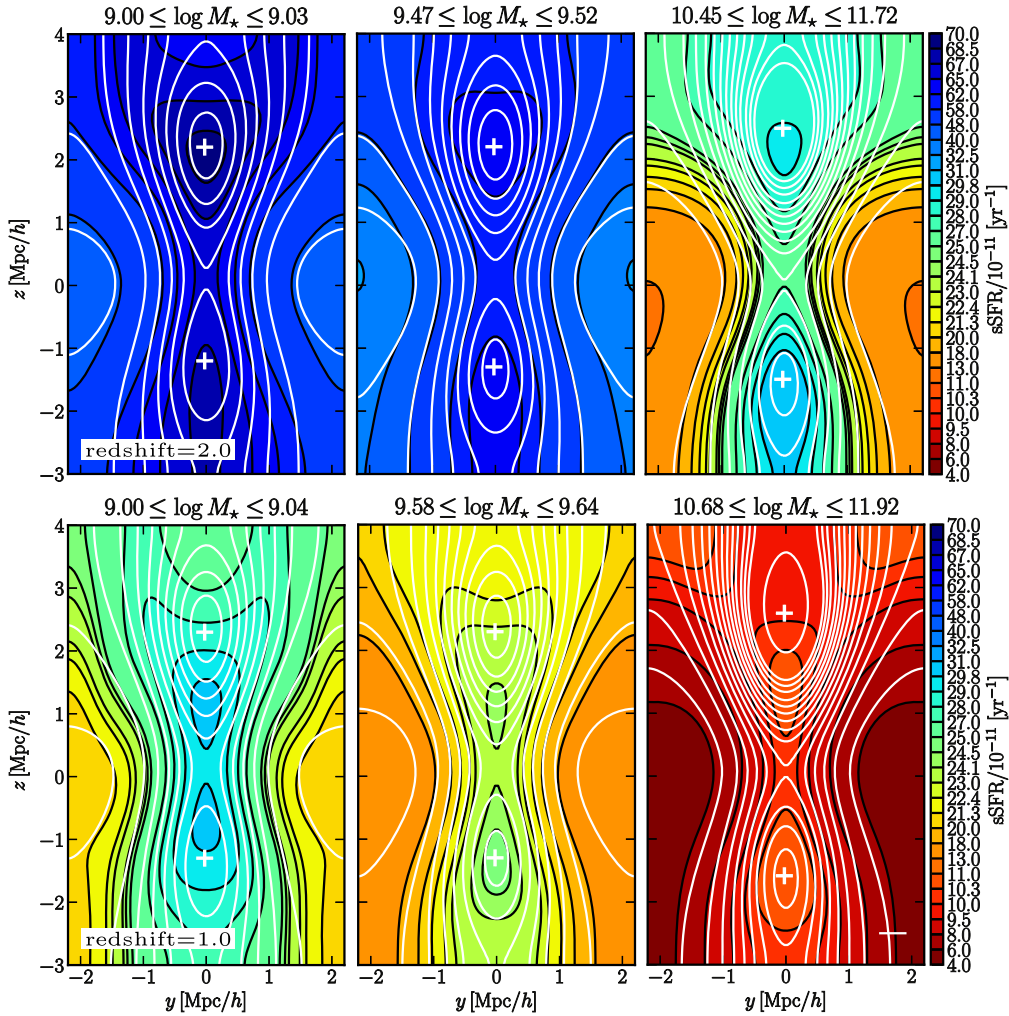
Figure 12. Mass-weighted sSFR in the frame of the saddle, in the longitudinal plane. Low (left-hand column), intermediate, (middle column), and high (right-hand column) stellar mass bins are shown at redshifts two (top), and one (bottom), respectively. The redshift zero map is shown on the top panel of Fig. 6. sSFR decreases with cosmic time at all stellar masses independently of the relative position with respect to the saddle. Interestingly, the peak of sSFR drifts away from the densest node as a function of cosmic time or increasing mass (i.e. level of non-linearity).

6.2 Observational signature for the impact of the cosmic web

The idea that galaxy properties, such as their stellar mass, colour, or sSFR are also driven specifically by the anisotropy of the cosmic web has only recently started to be explored in observations (e.g. Eardley et al. 2015; Alpaslan et al. 2016; Tojeiro et al. 2017). Stellar mass and colour or sSFR gradients have been reported at low (e.g. Chen et al. 2017), intermediate ($z \lesssim 0.25$; Kraljic et al. 2018), and higher redshifts ($z \sim 0.7$–0.9; Chen et al. 2017; Malavasi et al. 2017; Laigle et al. 2018), with more massive and/or less star-forming galaxies being found closer to the filaments compared to their lower mass and/or higher star-forming counterparts. The focus in this paper is on 2D and 3D cross-sections at fixed stellar mass, allowing to explore more complex geometric environment of the filamentary network. The (marginalized) 1D distributions (over distance along the filament) are in qualitative agreement with the above-mentioned observed stellar mass and colour or sSFR gradients with respect to filaments. Marginalizing over the distance perpendicular to the axis of the filaments yields gradients along the filament, such that at fixed orthogonal distance from the filament,

more massive and/or less star-forming galaxies are preferentially located in the vicinity of the node. Such a signature was found by Kraljic et al. (2018) in terms of red fractions, who reported the increasing fraction of passive galaxies with decreasing distances both to the filaments and nodes, with the dominant effect being the distance to the nodes. These gradients should now be measured in the 3D distribution of galaxies inferred from large galaxy redshift surveys, such as e.g. SDSS (York et al. 2000) or GAMA (Driver et al. 2009, 2011), providing a large statistical sample of galaxies and for which additional information about the properties of group haloes is available.

In terms of redshift evolution of sSFR, note that while at redshift above one the sSFR of galaxies increases in the direction along the filament away from the saddle and reaches its maximum near the node – in the region where the density is typically highest, this maximum is *shifted* away from the nodes towards the saddle at redshift one and below (top panels of Figs 6 and 12). Qualitatively similar behaviour, known as the reversal of the star formation–density relation at high redshift, was tentatively identified in observations (e.g. Elbaz et al. 2007; Cooper et al. 2008; Hwang et al. 2010, but see
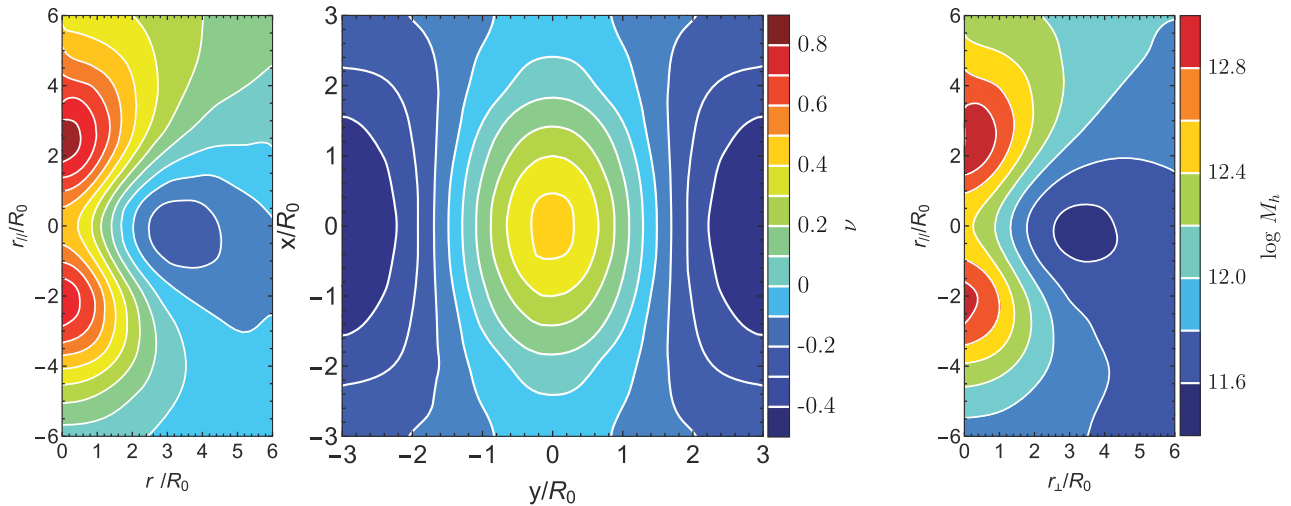
Figure 13. Mean predicted maps of the DM distribution around a saddle point of arbitrary shape and height. The direction of the filament is fixed to be along the vertical axis for left- and right-hand panels (perpendicular to the plane of the figure in the middle panel) and the top–bottom symmetry is broken by imposing that the most attractive peak is at the top. Left-hand panel: predicted distribution of the density fluctuation for Gaussian random fields ν (in units of the variance). Middle panel: corresponding transverse cross-section with the same colour coding. Right-hand panel: corresponding log of the non-linear mass when the threshold for collapse in the Press–Schechter mass is decreased by the mean density obtained on the left-hand panel. Its numerical counterpart measured in HORIZON-AGN is shown on Fig. E2 at low and high redshifts (see also Fig. E1 for a prediction).

e.g. Patel et al. 2009; Ziparo et al. 2014, for contradictory results).[13] Overall, our results suggest that in order to understand the complex behaviour of galaxies' properties, one may need to take into account the large-scale environment where tides are expected to play an important role, beyond that of density.

Note finally that a possible reason for the recent non-detection of Alam et al. (2018) and Paranjape et al. (2018b) with the SDSS resolution is that the ensemble average of the non-linearly evolved galactic properties predicted from angular-averaged fields does not differ by much from the ensemble and angular average of the non-linearly evolved galactic properties from anisotropic fields. To a good approximation, angular-averaging and dynamical non-linear evolution commute, which has of course been the basis of the success of the spherical collapse model.[14] One has to compute expectation in the frame of the filament to underline the differences, which is precisely the purpose of this paper.

6.3 Inferred age, mass, and counts statistics

The findings presented in this work, based on the analysis of galaxy-related gradients in the frame of saddle, are in qualitative agreement with the predictions of Musso et al. (2018) and those of Section 5: the iso-contours of studied galaxy properties show dependence on both the distance and orientation with respect to the saddle point of the cosmic web. Specifically, galaxies tend to be more massive closer to the filaments compared to voids, and inside filaments near nodes compared to saddles (Figs 1–5). Similarly and equivalently (given the duality between mass and cosmic evolution discussed in Appendix E2), Figs 9–11 show that as galaxies grow in mass, they

become more clustered near filaments and nodes with cosmic time, the width of the filaments narrows for a given mass bin, while the evolution of the entire population is consistent with broadening of the filaments, as expected from the theory of rare events (Bernardeau 1994). The number counts maxima are closer to the saddles than the stellar mass maxima as the former is dominated by the less-massive and more-common population, forming more evenly within the frame of the cosmic web, so that they have not had time to drift to the nodes. Consistently, older galaxies (Fig. 8) are preferentially located near the nodes of the comic web when comparing their distribution in the direction along the filaments, and in the vicinity of filaments in the perpendicular direction. These age gradients are seemingly at odds with the formation time of haloes predicted by Musso et al. (2018), where haloes that form at the saddle point assemble most of their mass the earliest. However, note that the formation time of haloes does not necessarily trace galactic age as inferred from the mean age of the stellar population. Indeed, our findings reflect the so-called downsizing (Cowie et al. 1996) of both galaxies and haloes (e.g. Neistein, van den Bosch & Dekel 2006; Tojeiro et al. 2017), such that oldest galaxies tend to be most massive, and galaxies in high-mass haloes are older (they formed their stars earlier).

Note finally that the theoretical predictions in Musso et al. (2018) are made at fixed halo mass, while the analysis presented so far in this work is performed at fixed stellar mass. However, the halo mass used in their study is physically closer to a sub-halo mass than a host halo mass,[15] and is therefore more strongly correlated with the stellar mass of galaxies which justifies further the qualitative comparison at this stage. As anticipated in Section 6.1, additional fixed sub-halo mass and density will be taken into account through the analysis of residuals (see Section 6.5).

---

[13]Elbaz et al. (2007) specifically found evidence of this reversal for massive galaxies, such that the sSFR increases with increasing galaxy density at redshift ∼one.

[14]This is in fact seen even at the level of the one-point function: one needs to invoke a moving barrier (Sheth & Tormen 2002), i.e. corrections to spherical collapse to match the measured mass function of dark haloes.

[15]The formalism adopted in Musso et al. (2018) does not capture the strongly non-linear processes operating on satellite galaxies.

6.4 The impact of AGN feedback

Relating the predicted specific accretion gradients of DM haloes to galaxies' observables requires some assumptions. One can in principle translate DM accretion gradients into sSFR gradients by considering the role of baryons in the accretion and feedback cycle. In the current framework of galaxy formation and evolution, galaxies acquire their gas by accretion from the large-scale cosmic web structure. The average growth rate of the baryonic component can be related to the cosmological growth rate of DM haloes, from which follows that higher star formation rate corresponds to higher DM accretion rate, providing that the SFR follows the gas supply rate. At high redshift, the vast majority of galaxies are believed to grow by acquiring gas from steady, narrow, and cold streams (e.g. Kereš et al. 2005; Ocvirk, Pichon & Teyssier 2008; Dekel et al. 2009). Using these arguments, it should follow that at high redshift, the stronger the accretion, the higher the sSFR of galaxy. Such a scenario is consistent with the gradients of the DM accretion rates found by Musso et al. (2018), where high-mass haloes that form in the direction of the filament tend to have higher accretion rates than haloes with the same mass that form in the orthogonal direction. This qualitatively agrees with the sSFR gradients in the frame of saddle at high redshift (Fig. 12) and in the simulation *without* AGN feedback (Fig. 6) at redshift zero, where galaxies with highest sSFR at fixed stellar mass tend to be located in the vicinity of the node in the direction along the filament, and near the saddle in the orthogonal direction.

In the presence of BHs, it is reasonable to expect at low redshift that the stronger the accretion, the stronger the AGN feedback, thus the stronger the quenching of star formation. This should result in an overall reduced sSFR, a behaviour that is indeed found when comparing the sSFR iso-contours between the HORIZON-AGN and HORIZON-NOAGN simulations. Interestingly, Figs 6, 7, and 12 also show that the shape of the sSFR iso-contours is modified in the presence of AGN feedback such that, at the high-mass end, galaxies with highest sSFR seem to be offset from the highest density nodes of the cosmic web (see also Appendix D which quantifies the difference of sSFR between HORIZON-AGN and HORIZON-NOAGN). Satellites are much less impacted by AGN feedback than centrals, and their sSFR is mostly affected by the environment of groups and clusters.

6.5 Evidence for other processes driving galaxy formation

Closer inspection specifically shows that the iso-contours of sSFR, V/σ (Figs 6 and 8) on the one hand, and stellar mass (Fig. 5) on the other differ from one another. This suggests that there may exist hidden processes driving galactic physics (beyond mass and local density).

Let us attempt to quantify their nature. Fig. 14 displays the host's halo mass (respectively, sub-halo's mass for satellites defined as the current virial mass of the sub-halo) in the frame of the saddle, in the longitudinal cross-section at redshift zero for different *stellar* mass bins (see also Fig. 5). Not surprisingly, galaxies with higher stellar mass are found to live in more massive DM haloes. These halo mass gradients are in agreement with Section 5's theoretical prediction and reflect what was already seen for the stellar mass gradients of the entire galaxy population, i.e. saddle points represent maxima of the halo mass in the direction perpendicular to the filament, while they are minima in the direction along the filament towards nodes independently of stellar mass. Note that in a given stellar mass bin, halo mass increases towards filaments and nodes, i.e. the $M_\star/M_h$ ratio is decreasing along those directions. Strikingly, there is little

change in the shape of these halo mass gradients when varying stellar mass. This is strongly indicative that stellar mass is at first order only a function of DM mass (at a given position within the cosmic web).[16] This is in sharp contrast with Fig. 6 (respectively, Fig. 8), which shows that the sSFR (respectively, V/σ) contours do vary significantly across stellar mass bins *and* have also distinct shapes compared to Fig. 14.

Besides halo mass, density is another obvious candidate for a variable that could drive the observed sSFR (respectively, V/σ) distributions in the frame of the saddle. Fig. 15 shows the density in the frame of the saddle, in the longitudinal cross-section at redshift zero for different *stellar* mass bins (see also Fig. 5). This density is computed on the scale of 0.8 Mpc h^{-1}, at which the skeleton was defined (and where the corresponding level of anisotropy was defined). Not surprisingly, galaxies with higher stellar mass are found to live in denser regions. These maps are again in agreement with Section 5's theoretical prediction and are qualitatively similar to what was already seen for the halo mass gradients, i.e. saddle points represent maxima of the density in the direction perpendicular to the filament, while they are minima in the direction along the filament towards nodes independently of stellar mass. As for halo mass, there is little change in the shape of these maps versus stellar mass. This in turn may indicate that there exist other position-dependent variables which impact sSFR (respectively, V/σ).

Let us attempt to quantify this effect by calibrating from the full simulation the mapping $s\hat{S}FR(M_h, \rho)$, defined as the median sSFR at given M_h and local density ρ (and $M_\star$ given that the mapping is defined in a given stellar mass bin), where ρ is computed on the scale of 0.8 Mpc h^{-1}. To do this, the median sSFR is computed in bins of $\log M_h$ and $\log \rho$ constructed adaptatively such that each of 10 equipopulated bins of M_h is further divided in eight equipopulated bins of $\log \rho$, in a given stellar mass bin. This median relation is then used in a 2D interpolation to obtain a relation $s\hat{S}FR(M_h, \rho)$ that can be applied to each galaxy (see Appendix G for details). Should the physical process driving star formation only depend on mean density and mass,[17] this operation would reproduce exactly Fig. 6. What is found instead is that at given stellar mass, there is a clear position-dependent discrepancy between the two, as shown in Fig. 16.[18] This figure displays the difference of the mean sSFR measured at the given position, and the mean sSFR estimated using the above-defined mapping, in highest stellar mass bin, normalized by the median sSFR (computed over the whole saddle region). This discrepancy is indicative that the impact of the saddle accounts for at least a fraction of the dispersion from the median sSFR$-M_h-\rho$

---

[16]A tight correlation between the stellar and halo mass of galaxies in the current framework of galaxy formation (Rees & Ostriker 1977; Fall & Efstathiou 1980) is expected based on abundance matching (e.g. Conroy & Wechsler 2009; Moster, Naab & White 2013; Rodríguez-Puebla et al. 2017) and confirmed with more direct measurements using e.g. satellite kinematics (e.g. van den Bosch et al. 2004; More et al. 2009) or weak lensing (e.g. Moster et al. 2010; Han et al. 2015; van Uitert et al. 2016).

[17]Note that we cannot rule out that position-dependent shape of the PDF of the distribution of sSFR, halo mass and density accounts for some residuals, as one would not expect the averaging and the mapping to fully commute, see Appendix G.

[18]We also computed maps of the density smoothed on 2 and 3 Mpc h^{-1}. This had little impact on the equivalent of Fig. 16, while significant residuals are found at the lower stellar mass bin, as expected since the smaller the mass the smaller the scale and the smaller the correlation with the field smoothed on larger (fixed) scale. This is consistent with the findings of Kraljic et al. (2018).
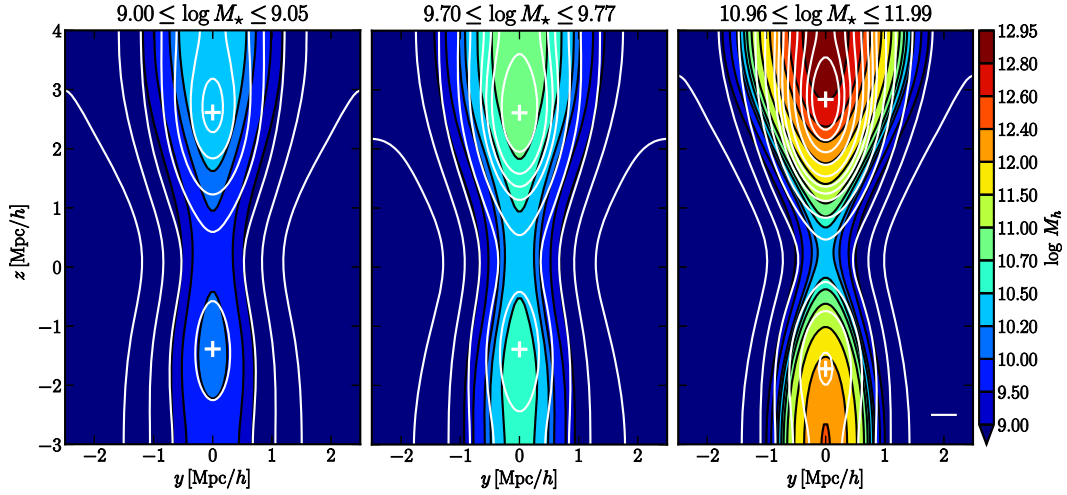
Figure 14. Stellar mass-weighted halo mass in the frame of the saddle at redshift zero for low (left), intermediate (middle), and high (right) stellar mass bins, as labelled, in the longitudinal plane at the saddle. The shape of iso-contours does not change dramatically with stellar mass and not surprisingly, galaxies with highest stellar masses live in most massive haloes. Note that low values for halo mass result from the smoothing of mean values in sparsely occupied regions.
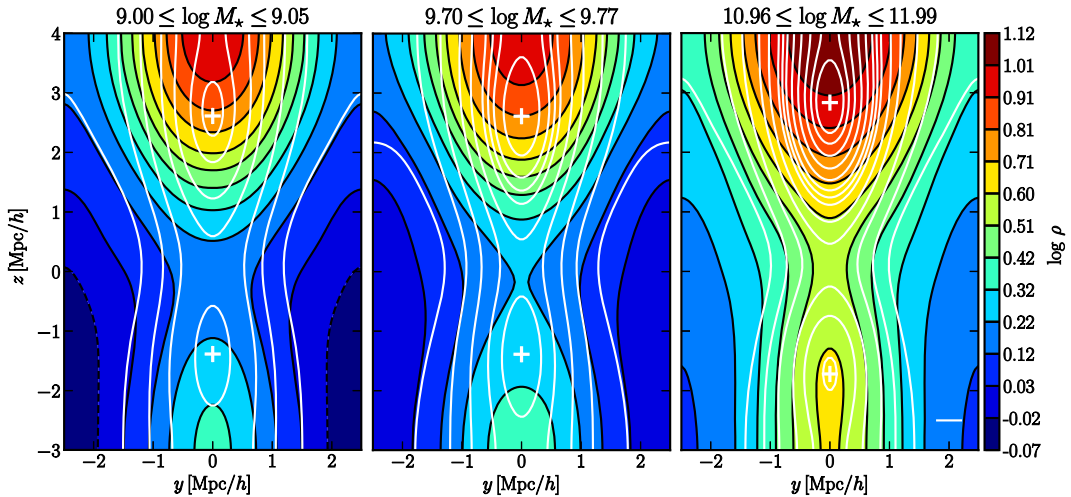
Figure 15. Stellar mass-weighted density in the frame of the saddle at redshift zero for low (left), intermediate (middle), and high (right) stellar mass bins, as labelled, in the longitudinal plane at the saddle. As for halo mass (see Fig. 14), the shape of iso-contours does not change dramatically with stellar mass and not surprisingly, galaxies with highest stellar masses live in densest regions.

relation (middle and right-hand panels), either because of the imposed local tides and/or because of the scatter in density imposed by this saddle (which might also be position-dependent). Interestingly, when the same transformation is applied to galactic age, no significant residuals are found (left-hand panel). This suggests that mean stellar mass and age, which are integrated quantities, do not seem to be very sensitive to anything but mean dark halo mass and mean density. Appendix G discusses in more details how to statistically disentangle mass, density, and tidal effects.

6.6 Is spin advection one of the residual processes?

In closing, let us speculate on the nature of the physical process which may be responsible for the residual scatter – having removed some of the effect of mean mass and local density, while relying on our saddle-centred stacks to identify processes that may be driven by anisotropy. As already mentioned, the (radial) distance to the node quenching from AGN feedback is an obvious candi-

date for the amplitude of the residual maps. Nevertheless, it had long been known that angular momentum stratification – undoubtedly built from anisotropic tides – is a key underlying property driving morphology of galaxies, which correlates with their star formation efficiency. Angular momentum acquisition is controlled by the large-scale tidal tensor, which imprints its torque along the galaxy's lightcone. The induced tides not only impact the assembly and accretion history of the host, but also the filamentary flow of cold gas connecting to the host, hence its coherent gas supply. It has recently been shown (Welker et al. 2015) following galaxies that the quadrupolar vorticity-rich large-scale filaments are indeed the loci where low- and intermediate-mass galaxies steadily acquire angular momentum via quasi-polar cold gas accretion, with their angular momentum aligned with the host filament (see Fig. 17 for the high-mass bin which has the most significant alignment signal at low redshift, and Laigle et al. 2015): galaxies are expected to accrete more efficiently cold gas when their angular momentum is aligned with the preferential direction of the gas infall, i.e. aligned
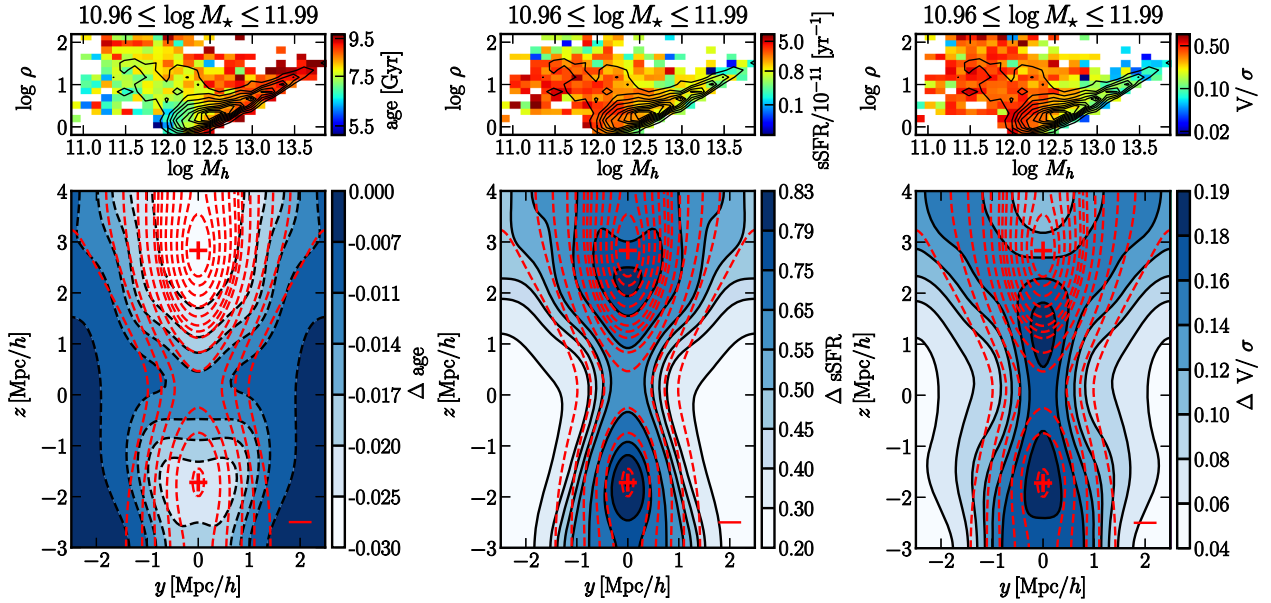
Figure 16. Age, sSFR, and V/σ residuals, from left to right, having removed the mean stellar mass, halo mass, and density effects, respectively by binning and considering the median mapping (see the text for details), in terms of fraction of the median values in the frame of the saddle at redshift zero for high stellar mass bin, in the longitudinal plane at the saddle. The red dashed contours and the red crosses correspond to the galaxy number counts and the peaks in galactic density on axis, respectively and the red horizontal lines represent the smoothing length used in the analysis. The sub-panels on the top show the distributions of the three parameters, age, sSFR, and V/σ (in colours), as a function of halo mass, M_h and local density ρ (computed on the scale of 0.8 Mpc h^{-1}, at which the skeleton and the corresponding level of anisotropy were defined), respectively, with number counts overplotted in black. Interestingly, the residuals for sSFR and V/σ display an excess at finite distance between the saddle and the nodes, which points towards the expected loci of maximum spin up and limited AGN quenching. Conversely, the age residuals are very small ($\lesssim 3$ per cent) relative to the values of residuals obtained for sSFR and V/σ, consistently with the observation that the age, halo mass, and local density gradients show many similarities (compare bottom panel of Figs 8, 14, and 15, respectively).

with the filament (Pichon et al. 2011; Stewart et al. 2011). This has typical local kinematic signatures in terms of (i) spin and (ii) vorticity orientation as predicted by Codis et al. (2015b), and as measured in HORIZON-AGN with respect to the direction of its closest filament (Fig. 17), and (iii) in terms of internal kinetic anisotropy in the velocity dispersion of dark haloes (Faltenbacher & White 2009). The V/σ of galaxies increases as they drift along the filament without significant merger, as they align themselves to the saddle's tides (Fig. 17).

The efficiency of star formation, as traced by sSFR, also depends on the infalling rate and impact parameter of the cold gas in the circumgalactic medium. Hence, one also expects star formation efficiency to be strongest wherever the alignment is tightest. The locus of this induced excess of star formation and/or V/σ should therefore have measurable signatures in observations when quantified in the metric of the filament (as discussed e.g. in Codis et al. 2015b, equation 40, in terms of loci of maximum cold gas advection at some finite distance from the saddle along the filament). There is a hint of such excess in the residuals shown in Fig. 16 in terms of both sSFR and V/σ (which should co-evolve). While quenching is also playing a significant position-dependent role for the high-mass population, its impact on the lower mass galaxies will be less significant. Fig. 18 shows indeed that for the lower mass bins, the residual maps peak significantly on axis, which supports the idea that the efficiency of angular momentum advection is a relevant process. This is worth emphasizing, given the above-given theoretical prejudices based on following galaxies in the flow (Welker et al. 2015), and on the orientation of galaxies traced by their spin's orientation distribution in the vicinity of the filament axis, predicted to exhibit a point-reflection symmetric structure (Codis et al. 2015b) as mea-

sured in Fig. 17. While this discussion is more speculative, recall in any case that most properties of the galactic population measured within the frame of saddles presented in the previous section – including redshift evolution and filament thickening/thinning – can be understood when accounting for their cosmic advection with the bulk flow along and transverse to the filament. The present study clearly highlighted that an improved model for galaxy properties should also explicitly integrate the diversity of the topology of the large environment on multiple scales (following, e.g. Hanami 2001) and quantify the impact of its anisotropy on galactic mass assembly history, and more generally on the kinematic history of galaxies. The details of how the kinematics impact star formation remains to be understood. The vorticity-rich kinematics of the large-scale flow is neither strictly coherent nor fully turbulent. Does the offset of merger and accretion rate imposed by the large-scale turbulent flow explain the residual environment dependence in observed physical properties (Aragon-Calvo, Neyrinck & Silk 2016), or is the helicity of gas inflow within filaments prevalent in feeding galactic discs coherently (Pichon et al.)?

7 CONCLUSIONS

This paper investigated the properties of virtual galaxies in the neighbourhood of filament-type saddle points of the cosmic web. These properties were measured within the frame set by the principal axes of the saddle in the HORIZON-AGN simulation. The impact of AGN feedback was assessed by comparing to results obtained in the HORIZON-NOAGN simulation. The principal findings are the following:
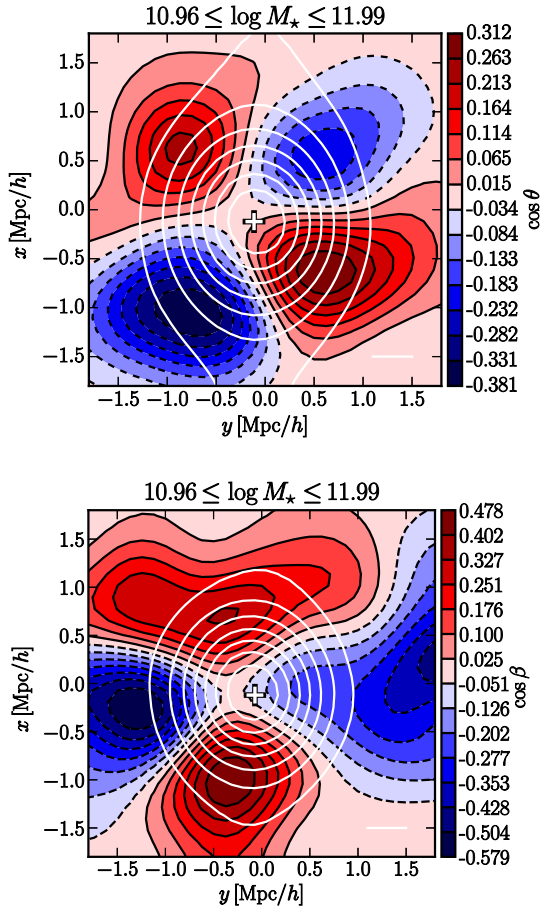
Figure 17. Stellar mass-weighted cosine of angle between the spin of the galaxy and the direction of the filament (top panel) and between the vorticity of the gas at the position of the galaxy and the direction of the filament (bottom panel) for highest stellar mass bin at redshift zero. Transverse cross-sections comprise galaxies with z-coordinate between 1 and 2.5 Mpc h^{-1} for the spin and between 0.5 and 1.8 Mpc h^{-1} for the vorticity, where the signal is most significant. The vorticity is computed on the gas distribution at the resolution and smoothing length used to define the skeleton, and interpolated at the position of galaxies. The map is normalized so that the integrated amplitude in each quadrant is preserved while smoothing. The white contours and the white crosses correspond to the galaxy number counts and the peaks in galactic density, respectively. Note the quadrupolar (point-reflection symmetric) structure of the spin's orientation distribution in the vicinity of the filament axis, in qualitative agreement with the prediction of Codis et al. (2015b) for DM. The distribution of the vorticity of the gas is also in qualitative agreement with this prediction and with the measurement of Laigle et al. (2015, their appendix A), which focused on cooling runs (without star formation). The tilt in the plane of symmetry of the vorticity map is likely to be driven by shot noise.

(i) The iso-contours of the galactic number density, mass, sSFR, V/σ, and age in the saddle's frame display a clear alignment with the filament axis and stronger gradients perpendicular to the filaments, quantifying the impact of the cosmic web in shaping galaxies.

(ii) High-mass galaxies are more clustered around filaments and within filaments around nodes compared to their low-mass counterparts. As expected, the filament's width of the whole galaxy population grows with cosmic time (as it becomes dominated by less rare galaxies). Conversely, at fixed mass, it decreases with cosmic time at the saddle.
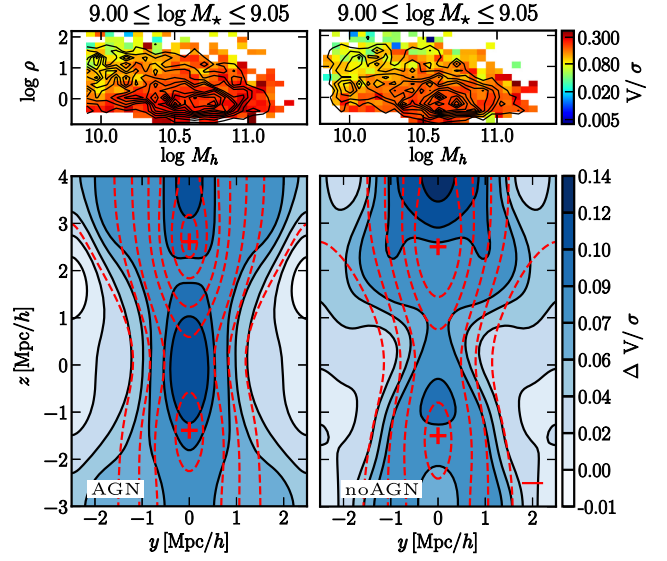
Figure 18. V/σ residuals in HORIZON-AGN (left-hand panel) and HORIZON-NOAGN (right-hand panel), having removed the mean stellar mass, halo mass, and density effects, respectively by binning and considering the median mapping (see the text), in the lowest stellar mass bin at redshift zero. The red dashed contours and the red crosses correspond to the galaxy number counts and the peaks in galactic density on axis, respectively and the red horizontal lines represent the smoothing length used in the analysis. The two maps are qualitatively similar, as expected for this mass bin where AGN feedback should not have a strong impact, and none the less, the amplitude of the map is ∼14 per cent, concentrated along the filament's axis.

(iii) In addition to reducing the overall sSFR of galaxies, AGN feedback also impacts the shape of the sSFR iso-contours, in particular for high-mass galaxies and in the vicinity of the nodes of the cosmic web. AGN feedback quenches centrals more efficiently than satellites. Satellite strangulation seems to occur within the filaments and nodes of the cosmic web.

(iv) While the dominant effect of the cosmic web on galaxy formation seems to be captured by the distance to cosmic nodes, the full 3D geometry of the web, in particular its saddle points, provides a natural oriented frame for stacking galaxies, showing significant effects of the environment beyond solely the distance to nodes. Hence, galaxies do retain a memory of the large-scale cosmic flows from which they emerged.

(v) The redshift evolution of the galactic counts and the age distribution of galaxies are consistent with a drift of the population towards the filaments and along them (see Appendix F). The cosmic evolution of the sSFR reflects both this drift and the triggering of quenching as centrals become massive enough to trigger AGN feedback near the peaks of the cosmic web. The geometry of the stacks and their cosmic evolution compare favourably to expectations for constrained Gaussian random fields in the weakly non-linear regime.

(vi) The maps of V/σ and sSFR (and their residuals) are consistent with the role played by feedback *and* angular momentum in shaping galaxies, beyond that played by mass and density, and its connection with the geometry of the cosmic web, as described by Codis et al. (2015b) and Laigle et al. (2015) (in a Lagrangian and an Eulerian framework, respectively). The point-reflection symmetric distribution of the orientation of the spin of galaxies and vorticity of the gas presented in this paper is also in agreement with this picture.

(vii) At high mass and low redshift, AGN feedback coupled with advection of galaxies along filaments induces some level of anisotropy in the distribution of galaxy properties (sSFR, V/σ, and age) which is partially degenerate with the effect of how angular momentum of galaxies is acquired from the large-scale vorticity of the anisotropic environment.

(viii) While sSFR responds to the saddle frame over and above what is expected from halo mass and local density, other indicators such as stellar age do not.

Overall, all distributions are consistent with the geometry of the flow in the vicinity of saddles, including quenching by AGN feedback, strangulation of satellites near the nodes, and possibly time delays induced by asymmetric tides on local *and* intermediate scales. They complement the findings of Kraljic et al. (2018), which also showed that galaxy properties occupy more than a 2D manifold (in physical parameter space such as age, sSFR, V/σ, etc.), but at the expense of not resolving the 3D distribution of fields in the frame of the saddle, which was the adopted strategy here.[19] This strategy allows us to *suggest* that one extra degree of freedom is the angular momentum acquired from the anisotropy of the cosmic web.

The signal-to-noise ratio in the counts is in the current analysis limited by the number of galaxies in the simulated box and by the choice of sampling the population in 3D. In order to e.g. probe the transverse asymmetry of saddles (reflecting the relative depth and distance to neighbouring voids and wall saddles), the present study could be followed up using simulations with better statistics so that the counts may be orientated with respect to the connecting walls and voids. A larger sample would also allow us to quantify the effect of non-linearities when constructing residual maps, as discussed in Appendix G. It would also be of interest to stack observationally measurable quantities such as colour or metallicity. These predictions could then be directly compared to observations from upcoming spectroscopic surveys such as *4MOST* (de Jong et al. 2012), DESI (DESI Collaboration et al. 2016), PFS (Takada et al. 2014), MSE (McConnachie et al. 2016), integral field spectroscopy such as MANGA (Bundy et al. 2015), SAMI (Croom et al. 2012), Hector (Bland-Hawthorn 2015) or in projection using photometric redshifts with DES (Rykoff et al. 2016), *Euclid* (Laureijs et al. 2011), *WFIRST* (Spergel et al. 2013), LSST (LSST Dark Energy Science Collaboration 2012), KiDs (de Jong et al. 2013), following the pioneer work of Laigle et al. (2018) in the COSMOS field. Connecting the present findings with work on spin orientation (Codis et al. 2015b) in the frame of the saddle may also prove useful to mitigate the effect of intrinsic alignment (e.g. Joachimi et al. 2011; Chisari et al. 2015). Investigating the distribution and survival of filaments on much smaller scales as they enter dark haloes is also of interest and will be the topic of future work (Ford et al. in preparation).

ACKNOWLEDGEMENTS

REFERENCES

Alam S., Zu Y., Peacock J. A., Mandelbaum R., 2018, preprint (arXiv: 1801.04878)
Alpaslan M. et al., 2016, MNRAS, 457, 2287
Aragon-Calvo M. A., Yang L. F., 2014, MNRAS, 440, L46
Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., van der Hulst J. M., 2007, ApJ, 655, L5
Aragon-Calvo M. A., Neyrinck M. C., Silk J., 2016, preprint (arXiv:1607.07881)
Aubert D., Pichon C., Colombi S., 2004, MNRAS, 352, 376
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Baron R. M., Kenny D. A., 1986, J. Pers. Soc. Psychol., 51, 1173
Beckmann R. S. et al., 2017, MNRAS, 472, 949
Bernardeau F., 1994, ApJ, 427, 51
Bland-Hawthorn J., 2015, in Ziegler B. L., Combes F., Dannerbauer H., Verdugo M., eds, IAU Symp. Vol. 309, Galaxies in 3D across the Universe, Kluwer, Dordrecht. p. 21
Bond J. R., Myers S. T., 1996, ApJS, 103, 1
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Borzyszkowski M., Porciani C., Romano-Díaz E., Garaldi E., 2017, MNRAS, 469, 594
Bundy K. et al., 2015, ApJ, 798, 7
Chen Y.-C. et al., 2017, MNRAS, 466, 1880
Chisari N. et al., 2015, MNRAS, 454, 2736
Codis S. et al., 2015a, MNRAS, 448, 3391
Codis S., Pichon C., Devriendt J., Slyz A., Pogosyan D., Dubois Y., Sousbie T., 2012, MNRAS, 427, 3320
Codis S., Pichon C., Pogosyan D., 2015b, MNRAS, 452, 3369
Codis S., Pogosyan D., Pichon C., 2018, MNRAS, 479, 973
Conroy C., Wechsler R. H., 2009, ApJ, 696, 620
Cooper M. C. et al., 2008, MNRAS, 383, 1058
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839
Croom S. M. et al., 2012, MNRAS, 421, 872
Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12
de Jong R. S. et al., 2012, Proc. SPIE, 8446, 84460T
de Jong J. T. A. et al., 2013, The Messenger, 154, 44
Dekel A. et al., 2009, Nature, 457, 451
DESI Collaboration et al., 2016, preprint (arXiv:1611.00036)
Driver S. P. et al., 2009, Astron. Geophys., 505.12
Driver S. P. et al., 2011, MNRAS, 413, 971
Dubois Y. et al., 2014, MNRAS, 444, 1453
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2012, MNRAS, 420, 2662
Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, MNRAS, 463, 3948
Eardley E., Peacock J. A., McNaught-Roberts T., Heymans C., Norberg P., Alpaslan M., Baldry I., 2015, MNRAS, 448, 3665
Efstathiou G., Frenk C. S., White S. D. M., Davis M., 1988, MNRAS, 235, 715
Elbaz D. et al., 2007, A&A, 468, 33
Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
Faltenbacher A., White S. D. M., 2009, ApJ, 708, 469
Gay C., Pichon C., Pogosyan D., 2012, Phys. Rev. D, 85, 023011
Gunn J. E., Gott J. R., III, 1972, ApJ, 176, 1

---

[19]Alternatively, one could stack in the theoretically motivated (Musso et al. 2018) 'natural' 2D frame of the saddle using radius, r, and 'angle', $\mathcal{Q} = \mathbf{r} \cdot \bar{\mathbf{H}} \cdot \mathbf{r}/r^2$, where $\bar{\mathbf{H}}$ is the tidal tensor.

Haardt F., Madau P., 1996, ApJ, 461, 20

Hahn O., Porciani C., Dekel A., Carollo C. M., 2009, MNRAS, 398, 1742

Han J. et al., 2015, MNRAS, 446, 1356

Hanami H., 2001, MNRAS, 327, 721

Hwang H. S., Elbaz D., Lee J. C., Jeong W.-S., Park C., Lee M. G., Lee H. M., 2010, A&A, 522, A33

Joachimi B., Mandelbaum R., Abdalla F. B., Bridle S. L., 2011, A&A, 527, A26

Jõeveer M., Einasto J., Tago E., 1978, MNRAS, 185, 357

Kac M., 1943, Bull. Am. Math. Soc., 49, 938

Kaiser N., 1984, ApJ, 284, L9

Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201

Kaviraj S. et al., 2017, MNRAS, 467, 4739

Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2

Komatsu E. et al., 2011, ApJS, 192, 18

Kraljic K. et al., 2018, MNRAS, 474, 547

Laigle C. et al., 2015, MNRAS, 446, 2744

Laigle C. et al., 2018, MNRAS, 474, 5437

Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692

Laureijs R. et al., 2011, preprint (arXiv:1110.3193)

Lazeyras T., Musso M., Schmidt F., 2017, J. Cosmol. Astropart. Phys., 3, 059

Libeskind N. I., Hoffman Y., Knebe A., Steinmetz M., Gottlöber S., Metuki O., Yepes G., 2012, MNRAS, 421, L137

LSST Dark Energy Science Collaboration, 2012, preprint (arXiv:1211.031 0)

Ludlow A. D., Borzyszkowski M., Porciani C., 2014, MNRAS, 445, 4110

Malavasi N. et al., 2017, MNRAS, 465, 3817

McConnachie A., et al., 2016, preprint (arXiv:1606.00043)

Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, Nature, 379, 613

More S., van den Bosch F. C., Cacciato M., Mo H. J., Yang X., Li R., 2009, MNRAS, 392, 801

Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903

Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121

Musso M., Cadiou C., Pichon C., Codis S., Kraljic K., Dubois Y., 2018, MNRAS, 476, 4877

Navarro J. F., Abadi M. G., Steinmetz M., 2004, ApJ, 613, L41

Neistein E., van den Bosch F. C., Dekel A., 2006, MNRAS, 372, 933

Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326

Paranjape A., Hahn O., Sheth R. K., 2018a, MNRAS, 476, 3631

Paranjape A., Hahn O., Sheth R. K., 2018b, MNRAS, 476, 5442

Patel S. G., Holden B. P., Kelson D. D., Illingworth G. D., Franx M., 2009, ApJ, 705, L67

Peebles P. J. E., 1969, ApJ, 155, 393

Peirani S. et al., 2017, MNRAS, 472, 2153

Peirani S., Mohayaee R., de Freitas Pacheco J. A., 2004, MNRAS, 348, 921

Pichon C., Pogosyan D., Kimm T., Slyz A., Devriendt J., Dubois Y., 2011, MNRAS, 418, 2493

Pogosyan D., Bond J. R., Kofman L., Wadsley J., 1996, in American Astronomical Society Meeting Abstracts. p. 1289

Pogosyan D., Pichon C., Gay C., Prunet S., Cardoso J. F., Sousbie T., Colombi S., 2009, MNRAS, 396, 635

Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541

Rice S. O., 1945, Bell System Tech. J., 25, 46

Rodríguez-Puebla A., Primack J. R., Avila-Reese V., Faber S. M., 2017, MNRAS, 470, 651

Rykoff E. S. et al., 2016, ApJS, 224, 1

Schaefer B. M., 2009, Int. J. Mod. Phys. D, 18, 173

Sheth R. K., Tormen G., 2002, MNRAS, 329, 61

Sobel M. E., 1982, in Leinhardt S., ed. Sociological Methodology. Jossey-Bass, San Francisco, p. 290

Sousbie T., 2011, MNRAS, 414, 350

Sousbie T., Pichon C., Colombi S., Novikov D., Pogosyan D., 2008, MNRAS, 383, 1655

Sousbie T., Pichon C., Kawahara H., 2011, MNRAS, 414, 384

Spergel D. et al., 2013, preprint (arXiv:1305.5422)

Stewart K. R., Kaufmann T., Bullock J. S., Barton E. J., Maller A. H., Diemand J., Wadsley J., 2011, ApJ, 738, 39

Stewart K. R., Brooks A. M., Bullock J. S., Maller A. H., Diemand J., Wadsley J., Moustakas L. A., 2013, ApJ, 769, 74

Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253

Takada M. et al., 2014, PASJ, 66, R1

Teyssier R., 2002, A&A, 385, 337

Tojeiro R. et al., 2017, MNRAS, 470, 3720

Toomre A., Toomre J., 1972, ApJ, 178, 623

Trowland H. E., Lewis G. F., Bland-Hawthorn J., 2013, ApJ, 762, 72

Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, A&A, 506, 647

van den Bosch F. C., Norberg P., Mo H. J., Yang X., 2004, MNRAS, 352, 1302

van Uitert E. et al., 2016, MNRAS, 459, 3251

Welker C., Devriendt J., Dubois Y., Pichon C., Peirani S., 2014, MNRAS, 445, L46

Welker C., Dubois Y., Pichon C., Devriendt J., Chisari E. N., 2015, preprint (arXiv:1512.00400)

White S. D. M., 1996, in Schaeffer R., Silk J., Spiro M., Zinn-Justin J., eds, Cosmology and Large Scale Structure. Elsevier Scientific Publishing Company, Amsterdam, p. 349

Wright S., 1934, Ann. Math. Stat., 5, 161

York D. G. et al., 2000, AJ, 120, 1579

Zentner A. R., 2007, Int. J. Mod. Phys. D, 16, 763

Ziparo F. et al., 2014, MNRAS, 437, 458

APPENDIX A: VALIDATION

Let us briefly study how the measured distributions presented in the main text are impacted by the smoothing length of the gas density distribution and the type of tracer used to extract the skeleton. Results are presented in the frame of the saddle using the curvilinear coordinates (see Section 3.1), but qualitatively similar conclusions are obtained for 3D distributions. Fig. A1 shows the galaxy number counts for the entire galaxy population with masses in the range $10^{9.0}$–$10^{12.0}$ $M_\odot$ at redshift zero, using, after rescaling, the same smoothing length as in the main text (left) and twice as big (right).
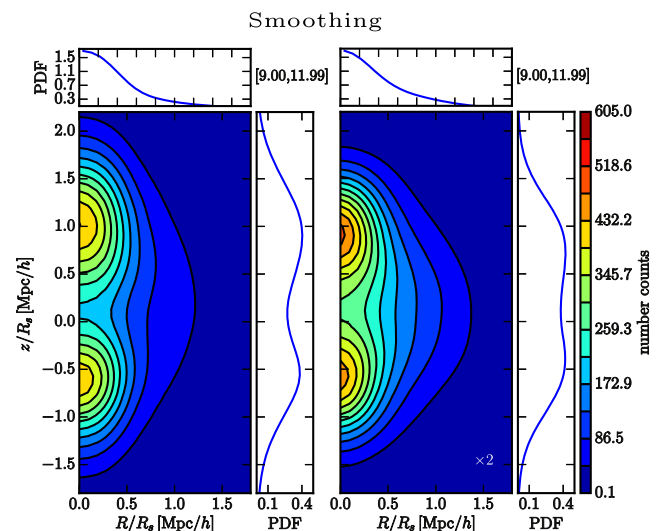
Figure A1. Galaxy number counts in the frame of the saddle (curvilinear coordinates) for all masses in the range $10^{9.0}$–$10^{12.0}$ $M_\odot$ at redshift zero (left) compared to the smoothing twice as big (right). Note that the R- and z-axes have been rescaled by the smoothing length. Similarity of the contours suggest that the measured distributions are relatively insensitive to the choice of the smoothing length.
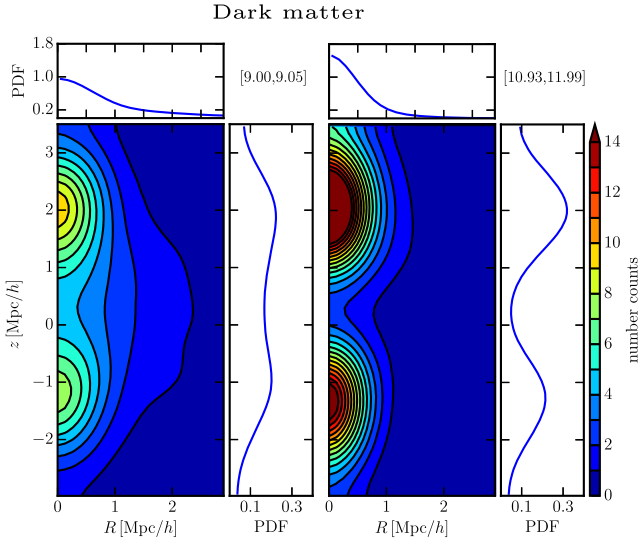
Figure A2. Galaxy number counts in the frame of the saddle (curvilinear coordinates) in low (left) and high (right) stellar mass bins at redshift zero, using the DM as a skeleton tracer. Similarity between these contours and those obtained using gas (see Fig. 1) suggest that the measured distributions are relatively insensitive to the choice of the tracer used to construct the skeleton.

Similarity of these iso-contours suggests that as expected, the measured distributions are relatively insensitive to the level of smoothing applied.

Fig. A2 shows the galaxy number counts in low (left) and high (right) stellar mass bins at redshift zero, using the DM particles as a tracer of the cosmic web. These iso-contours that should be compared with left- and right-hand panels of Fig. 1, suggest again only a weak dependence of results on the choice of the tracer (i.e. gas or DM). Note none the less that the skeleton built directly from galaxies using persistence is significantly different, as it becomes multiscale in nature. The corresponding complication is beyond the scope of this paper and will be explored elsewhere.

APPENDIX B: FILAMENTS' LENGTH AND WIDTH

Fig. B1 shows the probability distribution of the length of filaments at redshift two and zero. The length of filaments decreases with time, in agreement with the expected evolution of matter distribution in the ΛCDM universe with accelerated expansion at redshift $\lesssim 1$ and
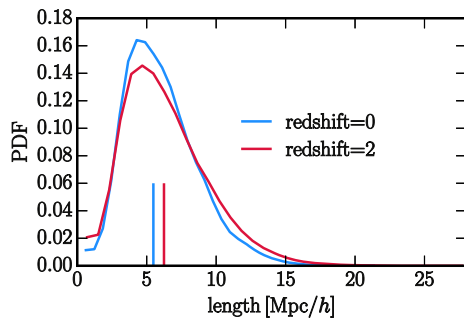
Figure B1. Probability distribution function of the length of filaments at redshift two (red) and redshift zero (blue). The vertical lines correspond to the medians of distributions.
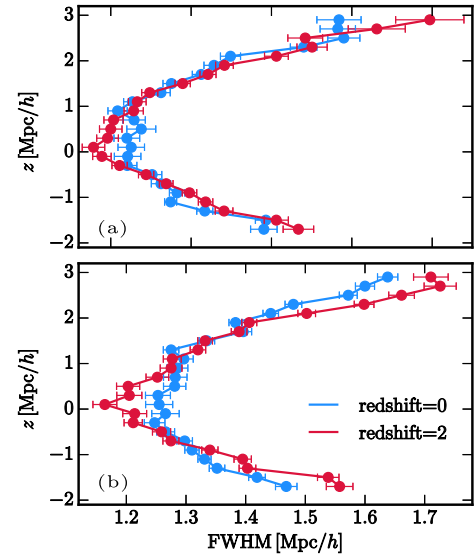
Figure B2. Thickness of the filaments, defined as the FWHM of the Gaussian fit of the transverse galaxy number counts profiles as in Fig. 11. When considering the entire population of galaxies, the cross-sections of filaments in the vicinity of the saddle point (at $z = 0$ Mpc h^{-1}) grow with time, while in the vicinity of nodes ($z \sim 2.5$ and -1.5 Mpc h^{-1} for highest and lowest density nodes, respectively), they get thinner.

as measured by Sousbie et al. (2008) for the DM. As universe expands, more low-mass objects form leading to the formation of filaments on smaller scales that eventually merge together while longer filaments are stretched. Because larger scale filaments are less numerous than filaments on small scales, the net result is a shift of the median length towards lower values at lower redshift.

Fig. B2 shows the thickness of the filaments as a function of the position in the direction along the filament and is complementary to Fig. 11 in that it extends to the vicinity of the nodes. Regions near the nodes (in both upper and lower directions from the saddle, corresponding to the nodes of highest and lowest densities, respectively) are getting thinner with time.

APPENDIX C: AZIMUTHALLY AVERAGED SECTIONS

All distributions presented in Section 3 – considering the stacks in 3D, and Section 4 presenting their redshift evolution – are in qualitative agreement with azimuthally averaged maps in 2D, adopting curvilinear coordinates as in Section 3.1. Let us here focus on redshift evolution alone. Fig. C1 shows the galaxy number counts and mean stellar mass in the frame of the saddle using curvilinear coordinates, at redshifts two and one, complementing Fig. 1. The redshift evolution of both number counts and mean stellar mass is in qualitative agreement with the results obtained when considering stacks in 3D (see Section 4) and consistent with the global flow of matter towards the filament first and along them afterwards (see Sousbie et al. 2008, for the DM flow).
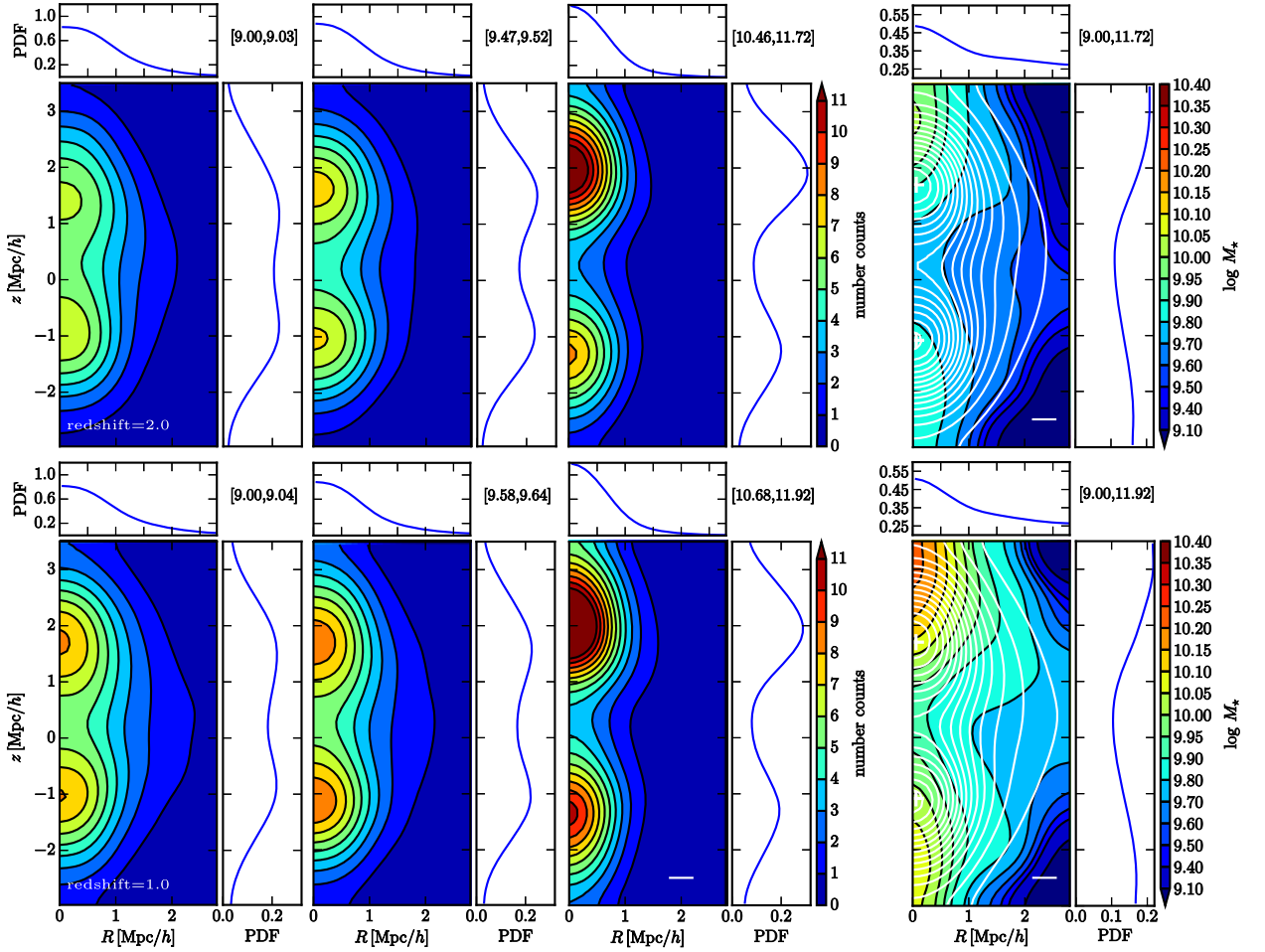
Figure C1. Left-hand and middle panels: redshift evolution of the galaxy number counts in low (left-hand column), intermediate (middle column), and high (right-hand column) stellar mass bins, in the frame of the saddle (curvilinear coordinates) at redshift two (top row) and one (bottom row), respectively. The white horizontal lines represent the smoothing length used in the analysis. Rightmost panels: redshift evolution of mean stellar mass iso-contours in the frame of the saddle (curvilinear coordinates) for the entire galaxy population at redshift two (top row) and one (bottom row), respectively. The white curves correspond to the contours of the galaxy number counts, while the white crosses represent the peaks in galactic number density on axis. Note how galaxies become more clustered towards filaments and nodes as they grow in mass with decreasing redshift, consistently with the global flow of matter within the cosmic web and in agreement with results considering the 3D distributions.

APPENDIX D: AGN QUENCHING EFFICIENCY

Fig. D1 shows the normalized difference of sSFR in the HORIZON-AGN and HORIZON-NOAGN simulations at redshift zero for highest stellar mass bin. This quantity allows to quantify where the quenching is most efficient. As expected, highest reduction of the sSFR is in the vicinity of the densest node.

APPENDIX E: THEORETICAL PREDICTIONS

Let us briefly predict from first principles the expected shape of the matter and halo distribution in the vicinity of a saddle and its cosmic evolution.

E1 Predictions for the mean constrained initial density field

The initial density field in which the cosmic web develops being Gaussian, the theory of constrained Gaussian random field provides a natural framework in which to compute the expectation of the matter distribution and typical halo mass within the frame set by the saddle point, as we do not expect the dynamics to be strongly non-linear on such scales. An important ingredient here is therefore to impose a filament-type saddle point constraint. Such a critical point form when the gradient of the density field is zero and is defined by its geometry, namely (i) its height ν defined as the density contrast divided by its rms $\sigma_0 = \sqrt{\langle \delta^2 \rangle}$ and (ii) its curvature by means of the three eigenvalues of the Hessian matrix of the density contrast rescaled again by their rms $\sigma_2 = \sqrt{\langle (\Delta \delta)^2 \rangle}$. For a filament-type saddle point, $\lambda_1 \geq 0 \geq \lambda_2 \geq \lambda_3$.

The so-called peak theory (Kac 1943; Rice 1945) then allows us to predict all statistical properties of critical points once the (supposedly Gaussian here) probability density function (PDF) of the field $\nu = \delta/\sigma_0$, its first $\nu_i = \delta_{,i}/\sigma_1$ and second derivatives $\nu_{ij} = \delta_{,ij}/\sigma_2$ is known. The saddle constraint reads

$$\mathcal{C}_{sad} = \frac{1}{R_\star^3} \lambda_1 \lambda_2 \lambda_3 \Theta_H(\lambda_1) \Theta_H(-\lambda_2) \delta_D(\nu_i), \tag{E1}$$
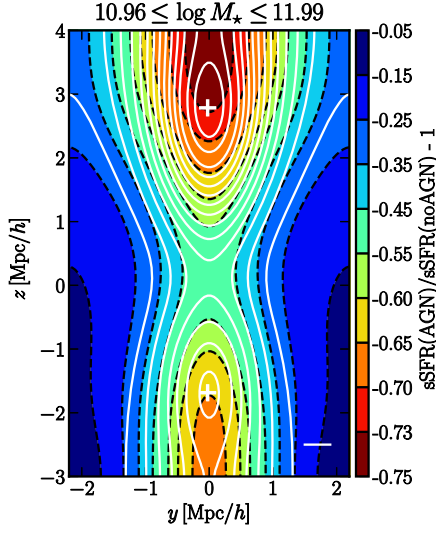
Figure D1. sSFR ratio in the HORIZON-AGN, sSFR (AGN), and HORIZON-NOAGN, sSFR (noAGN), simulations at redshift zero for highest stellar mass bin. Note that the highest impact of AGN feedback on sSFR of galaxies is, as expected, in the vicinity of the densest node. White contours represent galaxy number counts in the HORIZON-AGN simulation.

where the Dirac delta function ensures the gradient to be zero, the Heaviside Theta functions impose the sign of the eigenvalues, the Jacobian $\lambda_1\lambda_2\lambda_3 = \det v_{ij}$ accounts for the volume associated with a saddle point, and $R_\star = \sigma_2/\sigma_1$.

To predict the mean density map around a saddle point, one has to consider the joint statistics of (v, v_i, v_{ij}) together with the density field v' at a distance $\boldsymbol{r}$ from the saddle point. In addition, the symmetry along the axis of the filament ($i = 1$ here) will be broken by imposing the first axis to be oriented in the opposite direction from that of the gradient of the gravitational potential (i.e. towards the deepest potential well, the most attractive node). One therefore also has to consider Φ_1 the derivative of the gravitational potential along the first direction rescaled by its corresponding variance $\sigma_{-1} = \sqrt{\langle(\nabla\Phi)^2\rangle}$. Let us gather those 12 fields in a vector $\boldsymbol{X} = \{v', v, v_1, v_2, v_3, v_{11}, v_{22}, v_{33}, v_{12}, v_{13}, v_{23}, \Phi_1\}$ whose PDF can be written

$$\mathcal{P}(\boldsymbol{X}) = \frac{1}{\sqrt{\det|2\pi\mathbf{C}|}} \exp\left(-\frac{1}{2}\boldsymbol{X}^{\mathrm{T}} \cdot \mathbf{C}^{-1} \cdot \boldsymbol{X}\right), \qquad (E2)$$

where the covariance matrix $\mathbf{C} = \langle \boldsymbol{X} \cdot \boldsymbol{X}^{\mathrm{T}}\rangle$ depends on the separation vector $\boldsymbol{r}$ and the linear power spectrum $P_k(k)$ which can include a filter function on a given scale. In this work, a ΛCDM power spectrum is used (using the same values for the cosmological parameters as HORIZON-AGN) with a Gaussian filter defined in Fourier space by

$$W_{\mathrm{G}}(\boldsymbol{k}, L) = \frac{1}{(2\pi)^{3/2}} \exp\left(\frac{-k^2 L^2}{2}\right), \qquad (E3)$$

with $L = 0.8\,\mathrm{Mpc}\,h^{-1}$. One-point covariances do not depend on the separation but may depend on the spectral parameter $\gamma = \sigma_1^2/(\sigma_0\sigma_2)$. The variance of the density field is one by definition, $C_{11} = 1$, while the diagonal block corresponding to the saddle po-

sition, $\mathbf{C}_0 = (C_{\mathrm{ij}})_{i,j>1}$, reads

$$\mathbf{C}_0 = \begin{pmatrix}
1 & 0 & 0 & 0 & -\gamma/3 & -\gamma/3 & -\gamma/3 & 0 & 0 & 0 & 0 \\
0 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta/3 \\
0 & 0 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\gamma/3 & 0 & 0 & 0 & 1/5 & 1/15 & 1/15 & 0 & 0 & 0 & 0 \\
-\gamma/3 & 0 & 0 & 0 & 1/15 & 1/5 & 1/15 & 0 & 0 & 0 & 0 \\
-\gamma/3 & 0 & 0 & 0 & 1/15 & 1/15 & 1/5 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/15 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/15 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/15 & 0 \\
0 & -\beta/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3
\end{pmatrix},$$

with $\beta = \sigma_0^2/\sigma_{-1}\sigma_1$. The cross-correlations between v' and the fields at the position of the saddle are to be computed carefully as they depend on both the separation and the orientation of the separation vector in the frame of the Hessian described by the coordinates with indices $i = 1, 2$, and 3. They are explicit function of the shape of the power spectrum and are therefore computed numerically (the angle dependence is analytical, hence only the integration with respect to $k = |\boldsymbol{k}|$ requires a numerical integration). They read for j between 2 and 12

$$\langle v'X_j\rangle = \frac{\int \mathrm{d}^3\boldsymbol{k}\,P_k(k) \prod_{i=1}^{3}(-\iota k_i)^{\alpha_i}(\iota k)^{-2p}\exp(\iota\boldsymbol{k}\cdot\boldsymbol{r})}{\int \mathrm{d}^3\boldsymbol{k}\,P_k(k) \int \mathrm{d}^3\boldsymbol{k}\,P_k(k) \prod_{i=1}^{3}k_i^{2\alpha_i}}, \qquad (E4)$$

where $p = 1$ only for $j = 12$ (because of Poisson equation) and zero elsewhere and α_i counts the number of derivatives with respect to index i. Note that the mean density map around a saddle point of fixed height and curvatures with no symmetry breaking (i.e not imposing $\Phi_1 < 0$) is analytical and given by (Codis et al. 2015b)

$$\langle v'|\mathcal{S}\rangle = \frac{(\lambda_1+\lambda_2+\lambda_3)\left(\langle v'\mathrm{tr}\,v_{ij}\rangle + \gamma\langle v'v\rangle\right)}{1-\gamma^2}$$
$$+ \frac{v\left(\langle v'v\rangle + \gamma\langle v'\mathrm{tr}\,v_{ij}\rangle\right)}{1-\gamma^2} + \frac{45}{4}\left(\hat{\boldsymbol{r}}^{\mathrm{T}} \cdot \overline{\boldsymbol{H}} \cdot \hat{\boldsymbol{r}}\right)\langle v'(\hat{\boldsymbol{r}}^{\mathrm{T}} \cdot \overline{\boldsymbol{H}} \cdot \hat{\boldsymbol{r}})\rangle,$$

where $\overline{\boldsymbol{H}}$ is the *detraced* Hessian of the density and $\hat{\boldsymbol{r}} = \boldsymbol{r}/r$. However, here the goal is to compute this mean map around an arbitrary saddle (marginalizing over its height and curvatures) and with symmetry breaking. To do so, a Monte Carlo technique is implemented to compute the integrals of typically six dimensions with MATHEMATICA.

The mean map marginalized over the direction perpendicular to the filament is shown on the left-hand panel of Fig. 13. As expected, a filamentary ridge is predicted along the λ_1 direction with two nodes at about three smoothing lengths from the saddle. In the direction perpendicular to the filament, two voids are typically found on both sides of the saddle. In addition, Fig. 13 also shows the mean density in a plane perpendicular to the filament and containing the saddle point. As expected the filament cross-section is squashed in the direction of the wall (λ_2). This squashing will depend on the peak height and therefore on the mass of galaxies and haloes, namely the rarer objects will display a more spherical cross-section and vice versa. Note that for both plots, 10 million draws of the fields per point are drawn from a Gaussian distribution conditioned to having $v_i = v_{12} = v_{13} = v_{23} = 0$. All configurations with positive Φ_1 and wrong signs of the eigenvalues are thrown before computing the mean density v' in those configurations with weights $\lambda_1\lambda_2\lambda_3$ (because of the v_i condition) times $(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3)$ (because of the $v_{12} = v_{13} = v_{23}$ condition).

E2 Cosmic evolution of the dark matter maps

The above formalism is valid in the Gaussian initial conditions and can in principle be extended perturbatively to the subsequent weakly non-linear cosmic evolution. For the sake of simplicity, only the mean non-linear evolution of the density distribution around a saddle point of *fixed geometry* is described. Using a Gram–Charlier expansion (Gay, Pichon & Pogosyan 2012) for the joint distribution of the field and its derivative, the first non-Gaussian correction to the mean density map is found to be

$$\langle \delta(\boldsymbol{r}|\mathcal{S}) \rangle \sim \langle \delta(\boldsymbol{r}|\mathcal{S}) \rangle_{\mathrm{G}} + \sigma_0 \left[\sum_{ijk \leq 11} S_{\mathrm{ijk}} H_{\mathrm{ijk}}(\boldsymbol{r}) \right], \tag{E5}$$

where higher order terms $\mathcal{O}(\sigma^2)$ are neglected. In equation (E5), the $S_{\mathrm{ijk}} \equiv \langle X_i X_j X_k \rangle / \sigma$ coefficients generalize the so-called $S_3 \equiv \langle \delta^3 \rangle / \langle \delta^2 \rangle^2$ to expectations of cubic combinations of the field $\nu = \delta / \sigma_0$ and the components of its gradients, ν_{k}, and its Hessian, ν_{ij} (rescaled by their respective variance) evaluated at the running point $\boldsymbol{r}$ and at the saddle. In equation (E5), the function $H_{\mathrm{ijk}}(\boldsymbol{r})$ only involves known combinations of the Gaussian covariance matrix C_{ij} evaluated at separation $\boldsymbol{r}$ (so H_{ijk} is independent of redshift). Note importantly that at tree order, the S_{ijk} also do not depend on σ_0, so that the only (degenerate) dependence on cosmic time τ and smoothing scale L (over which the saddle is defined) is through $\sigma_0(L, \tau)$ in front of the square bracket of equation (E5). For the purpose of this paper, this equation therefore implies that gravitational clustering will distort and enhance the contours of DM density within the frame of the saddle, with a scaling proportional to σ_0.[20] Here, the considered scale L can also be related to the typical mass, $M_\star$ of the population considered so that the local clock becomes $\sigma(M_\star, \mathrm{redshift})$. Hence, equation (E5) simply predicts the observed mass and redshift scalings of the main text. In practice, computing the whole S_{ijk} suite takes us beyond the scope of this paper and will be investigated elsewhere.

Notwithstanding, as a first approximation, most of the effect is simply due to the density boost $\nu = \nu_\mathcal{S}$ at the location of the saddle. The corresponding non-Gaussian correction is simply given by σ_0 multiplied by

$$\frac{H_2(\nu_\mathcal{S})}{2} \frac{\xi(r)}{\sigma_0^2} (C_{12}(r) - S_3), \tag{E6}$$

with $H_2(x) = x^2 - 1$ the second Hermite polynomial, $C_{12} = \langle \rho^2(\boldsymbol{x}) \rho(\boldsymbol{x} + r) \rangle_{\mathrm{c}} / \sigma_0^2 \xi$ and again $S_3 = \langle \rho^3 \rangle_{\mathrm{c}} / \sigma_0^4$. Note that at tree order in perturbation theory, in the large separation limit, $C_{12}(r) - S_3 \to -34/21$. For a saddle point 1σ above the mean, $H_2(\nu_\mathcal{S}) > 0$, which means that the non-linear evolution tend to sharpen the density profile around the saddles (given that the height of the saddle, $\nu_\mathcal{S}$ is fixed here), as one would have expected.

Alternatively, excursion set theory (Musso et al. 2018) allows us to predict the typical mass distribution in the vicinity of a given saddle point (with fixed geometry) and as was done in that paper, the predicted profile can be displaced via a so-called Zel'dovich boost. This is shown in Fig. E1, which corresponds to a cross section through fig. 13 of Musso et al. (2018) where the length have been rescaled by a factor α and the masses by a factor α^3 to match the smoothing scale used in this paper and to account for differences arising from the use of a different filter (Gaussian versus Top-Hat). Using the same approach, it is also possible to compute the expected

---

[20]As such, the thickening of filaments provides us with a cosmological probe, though admittedly it might not be the most straightforward one!
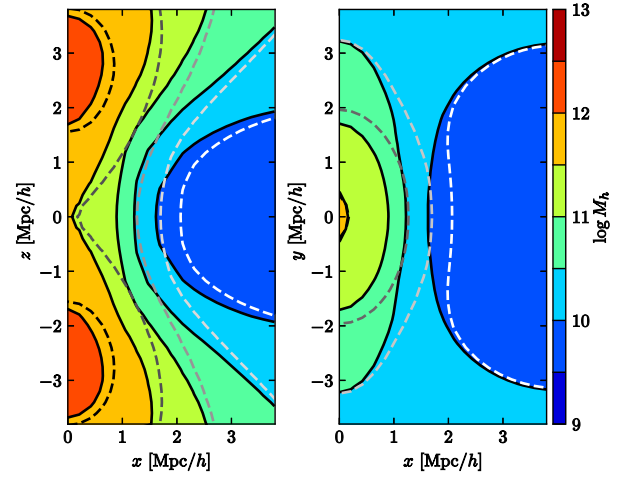
Figure E1. Coloured contours: predicted cross-section of the halo mass density after a Zel'dovich boost for a fixed geometry of the saddle in the plane of the wall and filament (left-hand panel) and perpendicular to the filament (right-hand panel). Dashed contours: the cross-section before the boost.
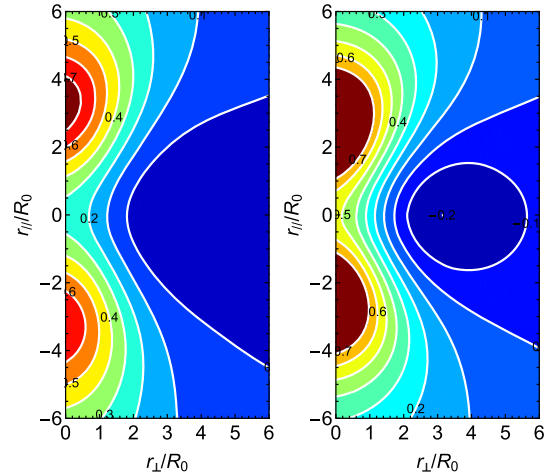
Figure E2. Left-hand panel: DM longitudinal cross-section measurements in HORIZON-AGN at redshift 0. Right-hand panel: same quantity at redshift 2. The corresponding prediction at high redshift is shown in Fig. 13 and the agreement is fairly good.

accretion rate of the DM halo. One then recovers fig. 12 of Musso et al. (2018) that is showing that the effect of the saddle point on the accretion rate decreases as the mass of the halo decreases. In the (simplistic) picture where DM accretion rate correlates with fresh gas accretion and specific star formation, one then qualitatively recovers the results of Fig. 6, where the effect of the cosmic web onto the sSFR decreases with the stellar mass. Indeed, as these two quantities (DM accretion rate and sSFR) only probe the recent accretion history of the halo, they are sensitive to *differential* effects induced by the saddle point which vanish at scales much smaller than that of the filament.

In order to compare this Lagrangian prediction to simulations, the mean total matter distribution was measured around saddles in the HORIZON-AGN simulation. The low-redshift measurement is shown on the left-hand panel of Fig. E2. Interestingly, the prediction for Gaussian random fields recovers the qualitative picture found in the HORIZON-AGN simulation in terms of the geometry of the

contours. As expected, the non-linear evolution (not captured by the Gaussian prediction) further contracts the filaments which become more concentrated. As one goes to higher redshifts (right-hand panel of Fig. E2), the contours clearly become closer to the Gaussian prediction.

APPENDIX F: KINETIC BULK FLOW NEAR SADDLE

Extending the result of Sousbie et al. (2008, which focused on DM), let us quantify the geometry of the bulk galactic velocity flow in the frame of the saddle. Fig. F1 displays the (normalized) velocity field of galaxies in the frame of the saddle while tracking (left-hand panel) or not (middle panel) the orientation of the saddle, and the PDF of the velocity's modulus and orientation (right-hand panel). The velocities of the left-hand and middle panels are computed as previously, i.e. as an average velocity for all galaxies in given 2D bin and smoothed over 0.3 Mpc h^{-1}. Note that no flipping with respect to the z-axis is applied here. As expected, when the frame is orientated towards the larger node (left-hand panel), the net flow is directed towards that node throughout that frame. Interestingly, note that the flow actually overshoots the peak of density in that frame, which is in fact expected, in so far that the velocity should, at the level of the Zel'dovich approximation, point towards the minimum of the potential, whose peak is typically further away from the saddle. When the orientation of the saddle is ignored (middle panel), one recovers a 'saddle-like' geometry for the flow, i.e. the saddle point locally repels the flow longitudinally but attracts it transversally. The right-hand panel is consistent with fig. 6 of Sousbie et al. (2008), but applies now to galaxies in HORIZON-AGN. The PDF velocity orientation and moduli present a tail of high velocities (at $\cos \delta < 0$), corresponding to galaxies converging transversally towards filaments.

The geometry of the flow displayed in Fig. F1, together with the distinct initial population distribution (and accretion history) for the progenitor of high- and low-mass galaxies allow us to understand their cosmic evolution presented in the main text. On top of this passive advection, Section 6.5 argues that the tides of the saddle may impact directly dark halo growth while shifting the conditional mean and covariances of the accretion rate, and galactic V/σ or sSFR while biasing spin (hence cold gas) acquisition.

APPENDIX G: STATISTICAL ORIGIN OF RESIDUALS

Let us finally discuss the statistical basis of the procedure described in the main text to study second-order effects beyond the mass and density and capture the origin of these hidden variables. When attempting to disentangle the specific role of tides from that of the local density and/or that of the dark halo mass, we are facing a statistical mediation problem (see e.g. Wright 1934; Sobel 1982; Baron & Kenny 1986), in that we aim to determine if the tidal tensor plays a specific role impacting the sSFR (or V/σ or age etc.) which is not already encoded in other quantities such as density and DM mass (which also vary away from the saddle, but typically with different maps in that frame). For the sake of being concrete, let us assume that the effect of the tides can be summed up by a scalar field α (e.g. the squared sum of the difference of the eigenvalues of the Tidal tensor, $\alpha = \sum (\lambda_i - \lambda_j)^2$, which quantifies the anisotropy of the collapse, or the net flux of advected angular momentum, etc.). Our purpose is to extract the map $\alpha(r)$ and check its structure relative to the saddle.

G1 Conditional mediation

Let us motivate the procedure used in Section 6.5 while relying on a statistical description of the random variables describing the various fields at some given position away from the saddle. Let us first assume for simplicity that the field $X = (\text{sSFR}, \delta \equiv \log \rho, m \equiv \log M_{\text{DH}}, \alpha)$ obeys a centred[21] joint Gaussian statistics:

$$\text{PDF(sSFR}, \delta, m, \alpha) = \frac{1}{\sqrt{\det(\mathbf{C}_0)}} \exp\left(-\frac{1}{2} X^T \cdot \mathbf{C}_0^{-1} \cdot X\right),$$

where $\mathbf{C}_0$ is the matrix of the covariance of the four fields, which we will also assume for now to be position independent (but see below). Note the change of variable to m and δ which are likely to behave more like Gaussian variables than M_{DM} and ρ.

Applying Bayes' theorem, we can compute the conditional PDF(sSFR$|\delta, m, \alpha$) = PDF(sSFR, δ, m, α)/PDF(δ, m, α), where PDF(δ, m, α) is the marginal (after integration over sSFR). From this conditional PDF, the expectation sSFR(r) $\equiv \langle$sSFR$|\delta, m, \alpha, r\rangle$ subject to the constraint of the three fields $\delta(r), m(r)$, and $\alpha(r)$ reads

$$\text{sSFR}(r) = \big(\langle\text{sSFR}\,\delta\rangle, \quad \langle\text{sSFR}\,m\rangle, \quad \langle\text{sSFR}\,\alpha\rangle\big)$$
$$\cdot \begin{pmatrix} \langle\delta^2\rangle & \langle\delta m\rangle & \langle\delta\alpha\rangle \\ \langle\delta m\rangle & \langle m^2\rangle & \langle\alpha m\rangle \\ \langle\delta\alpha\rangle & \langle\alpha m\rangle & \langle\alpha^2\rangle \end{pmatrix}^{-1} \cdot \begin{pmatrix} \delta(r) \\ m(r) \\ \alpha(r) \end{pmatrix} \equiv \sum_{i>1} \beta_i X_i,$$
$$\text{(G1)}$$

so that the conditional sSFR(r) is simply a linear combination of the three X_i maps, $\delta(r), m(r)$, and $\alpha(r)$ (with coefficients β_i involving the covariances).[22] Let us now take the statistical expectation of this equation *at a given pixel*. Subtracting the contribution of $\langle\delta(r)\rangle$ and $\langle m(r)\rangle$ from the measured $\langle$sSFR(r)$\rangle$ (while using a linear fit to the simulation to estimate the β_i since we do not know a priori what the covariances involving α might be[23]) and focusing on residuals provides a position-dependent estimate of the field $\langle\alpha(r)\rangle$. If its amplitude is statistically significant, its geometry may tell us if it is consistent with the nature of the mediating physical process, as discussed in the main text.

If we relaxed the assumption of Gaussian statistics, the conditionals derived from a Gram–Charlier expansion of the joint PDF (Gay et al. 2012) would lead (to leading order in non Gaussianity) to the mapping

$$\text{sSFR}(r) = \sum_{i>1} \beta_i X_i + \sum_{i,j>1} \beta_{ij} X_i X_j + \cdots \quad \text{(G2)}$$

where β_i (respectively β_{ij}) are functions of the second respectively, second- and third-order cumulants of the fields (such as $\langle\delta^2\alpha\rangle$ etc.). Once again we could subtract the (up to quadratic) fitted contribution of $\langle\delta(r)\rangle$ and $\langle m(r)\rangle$ from the measured $\langle$sSFR(r)$\rangle$ so as to fit the manifold of the X samples. Unfortunately, in this non-linear regime, the expectations would not compute any more, $\langle X_i X_j\rangle \neq \langle X_i\rangle\langle X_j\rangle$ and the residuals will also involve terms such as $\langle\delta\alpha\rangle$, $\langle m\alpha\rangle$, or $\langle\alpha^2\rangle$.

In practice though, the extracted relationships expressed in terms of $\delta = \log \rho$ and $m = \log M_{\text{DH}}$ do in fact look fairly linear, see

---

[21] The PDF is assumed to be centred on the mean value of the field averaged over the whole map

[22] This relationship could have also been obtained by principal component analysis in the extended X space: it would have led to the same sets of covariances as linear coefficients.

[23] Note that for an explicit choice of α, we could have extracted the covariances entering equation (G1) from the simulation and estimated the β_i accordingly.
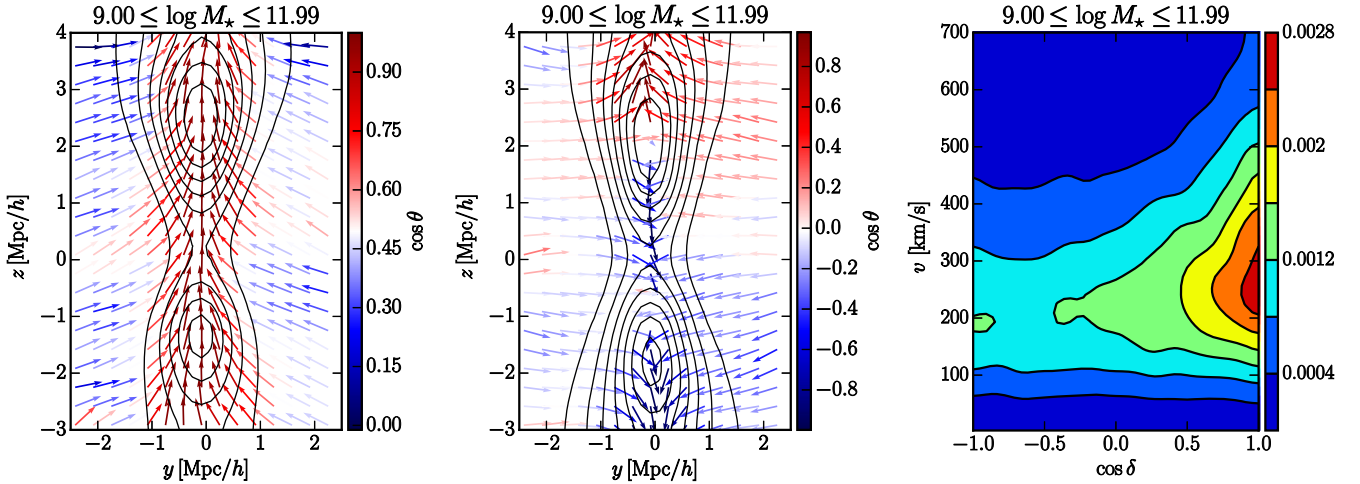
Figure F1. Left-hand and middle panels: velocity vector of galaxies in the frame of the saddle for masses in the range 10^9–10^{12} M$_\odot$ at redshift zero, in the longitudinal plane of the saddle, with the upward direction defined to be towards the node of highest density (left), and without imposing any condition on the direction of the filament at the saddle (middle). The arrows represent the unitary velocity vector in the $y - z$ plane, colour coded by the cosine of the angle between this vector and the $z-$axis. Note that as the modulus of the velocity near the saddle point is low, the direction of the velocity vector is not very well constrained there. Note also that the structure of the velocity field in the saddle frame is intrinsically complex due to the relative velocity of a saddle with respect to the nodes. The grey contours represent the galaxy number counts. Right-hand panel: PDF of the velocity field of galaxies within the cosmic web as a function of its modulus v, and the cosine of its angle δ with the closest filament. The excess of galaxies with $\cos \delta$ close to unity shows that the bulk of the population appears to be flowing along the filaments in the direction of nodes, i.e. the high-density regions.

e.g. the top panels of Fig. 16, which favours the assumption of Gaussianity, as was assumed in the main text. We also checked that relationships such as equation (G1) did not significantly vary with position within the saddle frame (by marginalizing over sub-regions within the frame). Finally, we used the median to extract the β_i coefficients, as it is a more robust estimator.

As a word of caution, it should nevertheless be stressed that since we are aiming to extract a secondary effect (beyond mass and density), the impact of departure from our assumptions may prove to be of the same order as the sought signal. Eventually, larger statistical samples may allow us to statistically disentangle more robustly the various processes. Note finally that carrying out the analysis at fixed *stellar* mass allows us to avoid the bimodality of some physical parameters which would clearly have broken the assumption of joint Gaussian statistics.

G2 Mediation of multiple causes

An alternative strategy to address the fact that more than one variable impact V/σ (and/or sSFR, age etc.) is to sample over narrow bins of stellar and dark halo mass, local density, and position r within the frame of the cosmic web, and estimate the full joint PDF. This is challenging for a sample of only 10^5 galaxies, hence can only be applied to relatively large bins in practice. We attempted to disentangle halo mass, density, and tidal effects by computing the residuals of V/σ (and/or sSFR, age) from median halo mass mapping in a given stellar mass and density bin. We found comparable residuals to those shown in Fig. 16. However, given the size of the bins we use, we cannot draw any definitive conclusions here. Simulations with more statistics should be able to address this difficult point in the future.

This paper has been typeset from a TeX/LaTeX file prepared by the author.

B.3 "Dense gas formation and destruction in a simulated Perseus-like galaxy cluster with spin-driven black hole feedback" (submitted to A&A)

One striking feature of nearby clusters is the extended filamentary Hα[1] emission found at their centre, yet their formation mechanism and the processes which shape their morphology remain unclear. In this study, the effect of AGN feedback on the evolution of clumps, defined as contiguous regions with $\rho > 1\,\mathrm{m_p\,cm^{-3}}$ and $T < 10^6\,\mathrm{K}$ is studied. It shows that the formation of clumps is directly influenced by the AGN jet and that, for low black hole spin, the clumps are able to reorient the spin axis (and hence the jet direction). The study makes use of the tracer particles described in chapter 5, which are key to provide the Lagrangian history of the gas found in clumps. In particular, they are used to estimate the condensation rate of gas into clumps by computing the mass of tracer particles that entered each individual clump between two snapshots. It is found that most of the condensation happens when clump are falling in, although a small fraction of the condensation happen for outflowing clumps. Two processes driving the fragmentation of clumps into smaller structures are found. (1) AGN feedback events are able to break large clumps into smaller ones, essentially by blowing out 3/4 of the gas. The surviving clumps are entrained and are ejected from the cluster centre. (2) At the top of their ballistic trajectories, clumps also fragment before the surviving ones fall back onto the cluster centre.

This study would have proven impossible with previous implementations of velocity-advected tracers, as their distribution is particularly inaccurate in regions where condensation occur (they gather in clump centres and stick there). In addition, the velocity-advected approach do not track the conversion of baryons into stars which may form in the clumps. While they provide somehow precise trajectories, velocity-advected tracers could not be used to compute e.g. condensation rate, which require tracer particles to accurately reproduce the Eulerian distribution of the gas and its time evolution (see Cadiou et al., 2019, chapter 5 for a discussion).

---

[1]Not to be confused with with cold filamentary accretion discussed in chapters 5 and 6.

Dense gas formation and destruction in a simulated Perseus-like galaxy cluster with spin-driven black hole feedback

R. S. Beckmann[1]⋆, Y. Dubois[1], P. Guillard[1], P. Salomé[2], V. Olivares[2], F. Polles[2], C. Cadiou[1], F. Combes[2,3], S. Hamer[4], M. D. Lehnert[1], and G. Pineau des Forets[2]

[1] Institut d'Astrophysique de Paris, CNRS UMR 7095, Sorbonne Université, 75014 Paris, France
[2] LERMA, Observatoire de Paris, PLS research Université, CNRS, Sorbonne Université, 75104 Paris, France
[3] Collège de France, 11 Place Marcelin Berthelot, 75005 Paris, France
[4] Department of Physics, University of Bath, Claverton Down, BA2 7AY, UK

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Context. Extended filamentary Hα emission nebulae are a striking feature of nearby galaxy clusters but the formation mechanism of the filaments, and the processes which shape their morphology remain unclear.
Aims. We conduct an investigation into the formation, evolution and destruction of dense gas in the center of a simulated, Perseus-like, cluster under the influence of a spin-driven jet. The jet is powered by the supermassive black hole located in the cluster's brightest cluster galaxy. We particularly study the role played by condensation of dense gas from the diffuse intracluster medium, and the impact of direct uplifting of existing dense gas by the jets, in determining the spatial distribution and kinematics of the dense gas.
Methods. We present a hydrodynamical simulation of an idealised Perseus-like cluster using the adaptive mesh refinement code RAMSES. Our simulation includes a supermassive black hole (SMBH) that self-consistently tracks its spin evolution via its local accretion, and in turn drives a large-scale jet whose direction is based on the black hole's spin evolution. The simulation also includes a live dark matter (DM) halo, a SMBH free to move in the DM potential, star formation and stellar feedback.
Results. We show that the formation and destruction of dense gas is closely linked to the SMBH's feedback cycle, and that its morphology is highly variable throughout the simulation. While extended filamentary structures readily condense from the hot intracluster medium, they are easily shattered into an overly clumpy distribution of gas during their interaction with the jet driven outflows. Condensation occurs predominantly onto infalling gas located 5 - 15 kpc from the center during quiescent phases of the central AGN, when the local ratio of the cooling time to free fall time falls below 20, i.e. when $t_{\rm cool}/t_{\rm ff} < 20$.
Conclusions. We find evidence for both condensation and uplifting of dense gas, but caution that purely hydrodynamical simulations struggle to effectively regulate the cluster cooling cycle and produce overly clumpy distributions of dense gas morphologies, compared to observation.

Key words. Galaxies: clusters: intracluster medium – Galaxies: jets – Galaxies: clusters: general – Methods: numerical – Hydrodynamics

1. Introduction

One of the most striking features of the nearby Perseus cluster, NGC1275, is the extended filamentary Hα emission nebula in its center (Lynds 1970; Heckman et al. 1989; Crawford & Fabian 1992; Conselice et al. 2001; Hatch et al. 2007; Fabian et al. 2008). Harbouring up to 5×10^{10} M$_\odot$ of cold gas (Salomé, P. et al. 2006), this emission nebula has a filamentary morphology, with individual filaments up to 40 kpc long and only 70 pc wide (Conselice et al. 2001; Fabian et al. 2016). Within the extended, filamentary Hα emission, dense clumps of molecular gas have been observed (Salomé, P. et al. 2006; Lim et al. 2012), and some filaments show signs of star formation (Fabian et al. 2008; Canning et al. 2010, 2014). Larger observational samples have shown that the Perseus cluster is not the only object to house such Hα emission nebulae, with many massive galaxy clusters showing similar features (Crawford & Fabian 1992; Heckman et al. 1989; McDonald et al. 2010, 2012; Olivares et al. 2019) in their center. Where does this gas come from, and what causes its characteristic filamentary morphology?

Finding cold gas in cluster centers is not unexpected. As cooling times in the intra-cluster medium (ICM) of massive galaxy clusters are short, a massive cooling flow of the order of $100 - 1000$ M$_\odot$ yr^{-1} is expected to develop in the cluster center (Fabian 1994). Observations of clusters, however, predict cooling flows of the order of only $10 - 100$ M$_\odot$ yr^{-1} (Heckman et al. 1989; Peterson & Fabian 2006), several orders of magnitude below the predicted cooling rates. Clusters must therefore contain a heating source which prevents overcooling and slows down star formation. Many clusters show evidence for extended jets powered by active galactic nuclei (AGN), which are inflating large cavities in the ICM whose power is sufficient to offset cooling (McNamara & Nulsen 2007; Rafferty et al. 2006; Fabian 2012). Via the self-regulation cycle, which consists of cold gas feeding the AGN, which in turn powers a jet, which then inflates cavities that heat the ICM, AGN are expected to play a decisive role in determining the cooling and star formation properties of the cluster (see McNamara & Nulsen 2007; Fabian 2012, for a review). This picture of self-regulation cycles from AGN jets is getting increasing support from hydrodynamical simulations both in an

⋆ ricarda.beckmann@iap.fr

idealised (Cattaneo & Teyssier 2007; Gaspari et al. 2011; Li & Bryan 2014a) and in a cosmological context (Dubois et al. 2010).

The cospatiality of the Hα emission nebula with the AGN jets and bubbles suggests that the AGN might not only control global cooling properties of the cluster but also be more directly responsible for the morphology of the existing dense gas (Salomé, P. et al. 2006; Russell et al. 2017; Vantyghem et al. 2017, 2018; McKinley et al. 2018; Tremblay et al. 2018). The often complex line-of-sight velocity field of the nebula in Perseus also suggests that this gas is not merely free-falling, or rotationally supported (McDonald et al. 2012; Gendron-Marsolais et al. 2018), but most likely interacts with the turbulence injected by the AGN jets and buoyantly rising bubbles (Fabian et al. 2003; Hatch et al. 2006; Revaz et al. 2008). However, with only line-of-sight velocity information, the three-dimensional velocity pattern of gas is difficult to ascertain.

Simulations by McCourt et al. (2012) and Sharma et al. (2012) showed that even for a globally thermally stable ICM (required to avoid overly strong cooling flows) dense gas can condense out of the hot ICM via local thermal instabilities when $t_{\rm cool}/t_{\rm ff} < 1 - 10$, with the critical ratio depending on the strength of the perturbations and the role of the AGN. Here, $t_{\rm cool}$ is the local cooling time

$$t_{\rm cool} = \frac{3}{2} \frac{nk_BT}{n_e n_i \Lambda}, \tag{1}$$

where n_i, n_e and n are the ion, electron and total number density respectively, T is the temperature and Λ the cooling rate. The free fall time is

$$t_{\rm ff} = \left(\frac{2r}{g}\right)^{\frac{1}{2}}, \tag{2}$$

where g is the local gravitational acceleration and r is the radius from the cluster center. It has been confirmed observationally that molecular gas is observed at the minima of $t_{\rm cool}/t_{\rm ff}$ profiles (Hogan et al. 2017; Pulido et al. 2018; Olivares et al. 2019), with some of these authors stressing that only $t_{\rm cool}$ determines condensation rates as the growth of linear perturbations is largely independent of the geometry of the gravitational potential (Choudhury & Sharma 2016). Simulations have shown that the turbulence injected by AGN feedback can cause the local thermal instabilities predicted by McCourt et al. (2012), but struggle to reproduce the observed morphologies, with dense gas having either a very clumpy morphology (Li & Bryan 2014b; Yang & Reynolds 2016a) or settling into a massive central disk (Gaspari et al. 2012; Li & Bryan 2014a,b; Prasad et al. 2015). While the latter has been observed in some clusters, such as in Hydra-A (Hamer et al. 2014; Olivares et al. 2019), only a small central disk is observed in Perseus (Nagai et al. 2019). The dense gas morphology therefore seems to sensitively trace the energy balanced in the ICM.

One feature of these clusters is that the cold gas is expected to rain down on the AGN in a cold and chaotic fashion (Gaspari et al. 2013; Voit & Donahue 2015; Voit et al. 2017), so the cold gas accreted by the black hole is expected to lack coherent angular momentum, which in turn could lead to a reorientation of the black hole spin axis over time. In this paper, we investigate the impact of this chaotic dense accretion on the formation of further gas by explicitly tracing the spin of the black hole, and using this black hole spin axis as the axis for the AGN driven jet (Dubois et al. 2014). In contrast to existing simulations, which rely on a fixed jet axis with pre-defined precession within a narrow jet cone (Li & Bryan 2014a; Yang & Reynolds 2016a; Ruszkowski

et al. 2017; Li et al. 2017; Prasad et al. 2018; Martizzi et al. 2019; Wang et al. 2019), the spin driven approach used here is able to inject turbulence over a larger volume of the cluster center, and to respond dynamically to the evolving dense gas morphology throughout the simulation.

In this paper, we will investigate the formation and time evolution of dense gas structures in a Perseus-like cluster under the influence of a spin-driven jet, with a particular focus on clump dynamics. The simulations are introduced in section 2. A general overview of results is given in section 3.1, the jet evolution is studied in Section 3.2 and the clump properties are investigated in Section 3.3. A detailed look at the role of uplifting in clump properties and dynamics is given in Section 3.4, and the impact of condensation is studied in Section 3.5. A discussion of results can be found in section 4, and conclusions are summarised in 5.

2. Simulation setup

This paper presents a set of hydrodynamical simulations of isolated galaxy clusters, produced with the adaptive mesh refinement code RAMSES (Teyssier 2002).

2.1. Technical details and refinement

For the simulations presented here, the Euler equations are solved with the second order MUSCL-Hancock scheme, which computes Godunov fluxes using an approximate HLLC Riemann solver and a MinMod total variation diminishing scheme to reconstruct the interpolated variables. The Courant factor is set to a value of 0.8.

The simulation is performed in box of size 8 Mpc with a root grid of 64^3, and then adaptively refined to a maximum resolution of 120 pc. Refinement proceeds according to several criteria. We use a quasi-Lagrangian criterion: when a cell contains a mass greater than $3.5 \times 10^9 \, {\rm M}_\odot$, it is refined (and it is derefined if it contains less than 0.125 this). We also use a Jeans-based criterion: a cell is refined until the local Jeans length is > 4 times the cell size. To refine regions of interest to this work, we also employ two additional refinement criteria. First, the cell containing the BH is forced to be refined at the maximum resolution. Second, a passive scalar variable is injected by the BH jet with a mass density $\rho_{\rm scalar}$ equals to that of the gas $\rho_{\rm gas}$, which is advected with the gas and marks regions affected by BH feedback. The scalar decays exponentially, with a decay time of $t_{\rm jet} = 10$ Myr to ensure that the scalar traces only recent AGN feedback events. After testing different decay times, we can confirm that the results do not sensitively depend on this value. To resolve the AGN bubbles, cells are allowed to be further refined when the scalar fraction exceeds $\rho_{\rm scalar}/\rho_{\rm gas} > 10^{-4}$, equivalent to 92 Myr since the last feedback event, and its relative variation from one cell to another exceeds 10^{-2}. The latter two refinement criteria ensure that the regions affected by AGN feedback, including the hot, low density bubbles which would de-refine under a purely Lagrangian refinement scheme, remain maximally refined over a reasonable duration of the jet propagation and mixing with hot ICM.

2.2. Initial conditions

The initial conditions for dark matter (DM) and gas consist of a cored Navarro-Frenck-White profile:

$$\rho_{\rm DM} = (1 - f_{\rm gas})\rho_s \frac{r_s^3}{(r + r_{\rm core})(r + r_s)^2}, \tag{3}$$

where $r_s = r_{200}/c$ is the scale radius, $r_{core} = 20$ kpc is the core radius, $\rho_s = \rho_c \delta_{200}$ is the density scaling of the profile, with ρ_c the critical density of the Universe. The rescaling factor $\delta_{200} = \frac{200\Omega_m}{3} \frac{c^3}{f(c)}$, where $f(c) = \ln(1+c) - c/(1+c)$, to the radius at which the average density of the profile is 200 times the mean density of the Universe. f_{gas} is the gas fraction of the halo, here taken to be 15 %. The halo has a concentration parameter $c = 6.8$, and a virial velocity $v_{200} = 1250$ km s^{-1}.

DM particles have a mass resolution of 3.7×10^8 M$_\odot$ and are distributed using the DICE code (Perret et al. 2014). The profile is truncated at a radius 2.2 Mpc, for a total DM halo mass of 3.4×10^{14} M$_\odot$. DM particle are live and able to move under gravity, allowing the DM potential to respond to the evolution of the cluster core throughout the simulation.

Gas is initiated in hydrostatic equilibrium assuming a gas fraction of 15 % distributed according to the profile of the DM (see Eq. 3), and then allowed to cool. As part of the initial conditions, turbulence is injected into the gas with a velocity dispersion of 15 km s^{-1}, but no rotation is added to the halo. This small initial velocity dispersion in the hot gas serves to break the symmetry of the initial conditions. Metallicity is initially set to 0.3 Z$_\odot$ throughout, and the BH sinkparticle is placed in the centre of the halo. No stars are added as part of the initial conditions. In order to avoid edge effects, the halo is placed in a sufficiently large box (8 Mpc on a side), and initiated with a gas density of 9.8×10^{-8} cm^{-3} outside of the truncation radius of the halo.

2.3. Cooling

The metal-dependant cooling of the gas is followed using the tabulated values of Sutherland & Dopita (1993) down to 10^4 K. The cooling function is extended below 10^4 K with the fitting functions from Rosen & Bregman (1995). Solar abundance ratios of the elements is assumed throughout independent of the overall metallicity.

2.4. Star formation and stellar feedback

Star formation proceeds according to a combined density and temperature criterion, with star formation permitted in cells with hydrogen number density of $n_H > 1$ H cm^{-3} and temperature $T < 10^4$ K. The mass resolution of stars is $n_H m_p \Delta x^3 / X_H = 5.6 \times 10^4$ M$_\odot$, where $X_H = 0.74$ is the fractional abandundance of hydrogen. The star formation rate density proceeds according to a Schmidt law $\dot\rho_* = \epsilon_* \rho / t_{ff}$, where ρ is the gas density, t_{ff} is the gas free-fall time, and $\epsilon_* = 0.1$ is the constant efficiency of star formation.

Stellar feedback is included in the form of type II supernovae only. We use the energy-momentum model of Kimm et al. (2015) with each stellar particle releasing an energy of $e_{*,SN} = m_* \eta_{SN} 10^{50}$ erg M$_\odot^{-1}$ at once after 10 Myr, where $\eta_{SN} = 0.2$ corresponding to the mass fraction of the initial mass function for stars ending up their life as type II supernovae, and m_* is the stellar particle mass. These explosions also enrich the gas with metals with a constant yield of 0.1. Metals are treated as a single species and are advected as a passive scalar.

2.5. Black hole accretion and feedback

AGN feedback from the central BH is followed using the model from Dubois et al. (2010) with several modifications that include the self-consistent evolution of the BH spin (Dubois et al.

2014) and the spin-dependent feedback efficiency (Dubois et al., in prep.).

A BH "sink" particle is placed at the centre of the halo as part of the initial conditions, with a mass of 3.4×10^8 M$_\odot$. The BH is then free to move across the grid throughout the simulation. To compensate for unresolved dynamical friction from the stars within the host galaxy, an analytic drag force is applied to the sink particle according to Pfister et al. (2019). We do not model the equally unresolved gas drag explicitly as the difficulty in measuring the relative velocity between the sink and the turbulent ISM introduces too many numerical artifacts in the black holes trajectory (see Beckmann et al. 2019). A particular worry is the black hole getting attached to its own feedback and ejected from the central galaxy, which we avoid here by not using a subgrid prescription for the gas drag.

The BH accretes according to the Bondi-Hoyle-Lyttleton accretion rate

$$\dot{M}_{BHL} = \frac{G^2 M_{BH}^2 \bar\rho}{(\bar{c_s}^2 + \bar{v}^2)^{3/2}} , \tag{4}$$

where $\bar\rho$, $\bar{c}_s$ and $\bar{v}$ are the mass weighted local average density, sound speed and relative velocity between the gas and the BH. All quantities are measured within a sphere with radius $4\Delta x_{min}$ centered on the instantaneous position of the BH, with the BH free to move across the grid. Δx_{min} is the size of the smallest resolution element of the simulation. Accretion is not limited to the Eddington accretion rate.

The AGN feedback is modelled with jets following the injection method from Dubois et al. (2010). At each feedback event, feedback energy

$$\dot{E}_{feed} = \eta_{MAD} \dot{M}_{BHL} c^2 \tag{5}$$

(where c is the speed of light) is injected as kinetic energy within all cells contained in a cylinder of radius 0.4 kpc and height 0.8 kpc. The cylinders is aligned with the BH spin axis. The efficiency η_{MAD} is a spin-dependent efficiency obtained from magnetically arrested disc (MAD) simulations from McKinney et al. (2012), which has a minimum at a spin of 0, and a maximum at a spin of 1. The BH spin-up rate is taken from the same simulations. The AGN jet is always taken to be aligned with the BH spin axis, and the conditions for BH-disc alignment in misaligned grid-scale gas angular momentum (with that of the BH spin) is obtained by Lens-Thirring considerations (see Dubois et al. 2014, for details). As the spin-axis changes self-consistently throughout the simulation, we do not need to add any explicit precession to the jet, as it naturally arises from the chaotic nature of the cold gas accretion onto the BH (Gaspari et al. 2013), which regularly changes the BH spin direction over time (see Section 3.2).

As mentioned in Section 2.1, a passive scalar of density $\rho_{scalar} = \rho_{gas}$ is injected within the feedback cylinder at each feedback event, where ρ_{gas} is the gas density. This scalar then decays exponentially with a decay time of 10 Myr, allowing cells recently affected by the AGN jet to be identified. Therefore, with the AGN passive scalar quantity, one can define an age for the gas that has been impacted by the AGN, with $t_{AGN} = -10 \ln(Y_{scalar})$ Myr, where $Y_{scalar} = \rho_{scalar}/\rho_{gas}$.

2.6. Tracer particles

To follow the dynamical history of gas in the simulation we employ Monte-Carlo based tracer particles from Cadiou et al. (2019). These tracer particles are a significant improvement over

classical "velocity"-based tracer particles, in particular in regions with strongly converging flows such as cold gas condensation and gas collapsing under self-gravity. We set up 2×10^8 tracer particles, with each particle tracing a gas mass of 4×10^5 M$_\odot$. They are initially distributed according to the gas density profile in the initial conditions, out to a radius of 200 kpc.

3. Results

3.1. Cluster evolution

As can be seen in Fig. 1, which shows the gas density, temperature, AGN age (see section 2.5) and the gas radial velocity at various times, the gas in the cluster develops a multi-phase structure with a complex morphology that evolves significantly over the course of the simulation.

The hot gas in the intra-cluster medium, which has temperatures in the range $0.09 - 113$ keV, cools down and condensates into dense clumps and filaments within the central 50 kpc of the cluster, with an average temperature of the dense gas of 4.0×10^{-4} keV. This dense gas falls towards the center where it feeds the central BH and thereby triggers the AGN jet, which, in return, interacts with existing dense gas and stirs turbulence into the hot gas, generating hot outflows with temperatures up to 113 keV and outflow velocities up to 3.5×10^4 km s^{-1}. As the radio jet is oriented along the SMBH spin axis, which in turn is updated according to the chaotic cold accretion onto the central BH (Gaspari et al. 2013; Voit et al. 2017), the jet continuously re-orients throughout the simulation (see Section 3.2 for details). As a result, the shapes of the jet relics indicated by the "AGN age" also change significantly over time.

More quantitatively, Fig. 2 shows that gas begins to cool after approximately 100 Myr, equivalent to the initial cooling time of gas in the cluster center as set by the initial conditions. Dense gas, for the remainder of the analysis, is defined to be gas with a maximum temperature of 10^6 K. By 139 Myr, the dense gas mass first exceeds 10^9 M$_\odot$, and the cluster enters a cyclic behaviour where dense gas repeatedly builds up to a total mass in excess of 2×10^{10} M$_\odot$ before being reduced to closer to 2×10^9 M$_\odot$.

We have split the evolution of the cluster into two regimes using the total dense gas mass. A cooling dominated regime, when the total dense gas mass of the cluster increases (marked with a grey background in Fig. 2), and a heating dominated regime, when the total dense gas mass of the cluster decreases. The regime of the cluster is evaluated using the smoothed derivative of the mass of dense gas $M_{\text{gas,dense}}$. The total dense gas mass in the cluster can be reduced in a number of different ways: dense gas can be consumed in star formation, accreted onto the BH or destroyed via hot winds or shocks driven by AGN feedback.

AGN activity (see second panel of Fig. 2) is highest during the heating-dominated phase, with maxima in dense gas followed by maxima in AGN activity within 50 Myr or less. These peaks in AGN activity destroy dense gas in the cluster, causing the AGN to enter a low feedback state until the dense gas mass has had time to build up again. Only a small fraction of the gas is directly accreted by the SMBH, as can be seen by the fact that the mass increase of the SMBH mass in the top panel of Fig. 2 is much smaller than the decrease in dense gas mass over the equivalent period of time.

As can be seen in the bottom panel of Fig. 2, the star formation rate varies strongly over time, following the general trends set by the total dense gas mass in the cluster. There are clear bursts of star formation in the cooling dominated time. This suggests that a significant amount of the dense gas is directly con-

sumed by star formation. At peaks of up to 1000 M$_\odot$/yr, the star formation rate of our simulated cluster is extremely high in comparison to observations, which for equivalent mass clusters report star formation rates in the range $1 - 100$ M$_\odot$/yr (O'Dea et al. 2008). The dense gas mass, by contrast, falls close to the $10^{10} - 10^{11}$ M$_\odot$ observed in Perseus (Bridges & Irwin 1998; Salomé, P. et al. 2006).

There are several possible reasons for this elevated SFR. One possibility is that dense gas is being converted too efficiently into stars. The star formation efficiency employed here, of $\epsilon_* = 0.1$, is on the upper end of observed values for massive clusters, which can range from as low as 0.01 to as high as 0.1 (Mittal et al. 2015; McDonald et al. 2018). Indeed, because most of the warm cooling gas is not settled in a galactic disk, the star formation efficiency is expected to be lower than in disks, because the pressure is lower due to the absence of the stellar potential (Shi et al. 2011; Salomé et al. 2016). We also use a comparatively simple density-based star formation recipe of the form $\dot{\rho}_* = \epsilon_* \rho / t_{\text{ff}}$, which does not take the effects of small-scale turbulence into account. In addition, being purely hydrodynamical, the simulation disregards effects such as magnetic fields and other non-thermal energy sources such as cosmic rays, which could heat the gas and provide an extra pressure support against collapse on small scales.

Alternatively, it is possible that gas in the cluster is cooling too efficiently, and that too much gas is transitioning from the hot, diffuse phase to the dense phase. The X-ray luminosity of the central 50 kpc of the simulated cluster are in the range of $1.2 - 5.3 \times 10^{45}$ erg s^{-1}, with the observed values for Perseus of 1.26×10^{45} s^{-1} (Ebeling et al. 1996) at the lower end of that range. While the initial conditions were chosen to reproduce observed profiles, the emitted X-ray luminosity increases due to the gas cooling in the cluster center.

In this simulation, we rely on equilibrium cooling with an initially uniform metallicity of 0.3 $Z_\odot$ everywhere, based on observations of the metallicity in the outskirts of Perseus by Werner et al. (2013). By 1 Gyr, the volume weighted metallicity in the central 50 kpc of the hot ICM has risen to 0.36 $Z_\odot$ due to stellar feedback. While this is higher than the initial value, it still falls below the value of 0.6 observed by Schmidt et al. (2002). One possibility is that equilibrium cooling assumed here overestimates the contribution of metal cooling at high temperatures. X-rays emitted by the AGN could dissociate metals in high temperature gas, reducing their contribution to cooling (Dubois et al. 2011; Agertz et al. 2013).

If radiative transfer and non-equilibrium processes were included, the hard X-rays emitted by the AGN would be able to photo-ionize some important metal coolants further so that their contribution to cooling is reduced (e.g. Gnedin & Hollon 2012; Segers et al. 2017). As metal line cooling is the dominant cooling channel for gas between $10^5 - 10^7$ K, shutting down metal cooling would hamper the transition of gas from the hot, diffuse to the dense phase. To test this theory, we ran a simulation using a cooling function in which the metal cooling function is modified by a kernel

$$f_{\text{cool}} = \exp\left[-\left(\frac{T}{10^4 \text{ K}} \right)^{10} \right], \qquad (6)$$

which effectively shuts off metal cooling for gas with temperatures above T $> 10^4$ K. As can be seen in the bottom panel of Fig. 2, while the initial cooling is delayed in comparison to the fiducial simulation, SFRs remain high even with truncated metal line cooling and the evolution of dense gas is qualitative indistinguishable between the two simulations. We therefore conclude
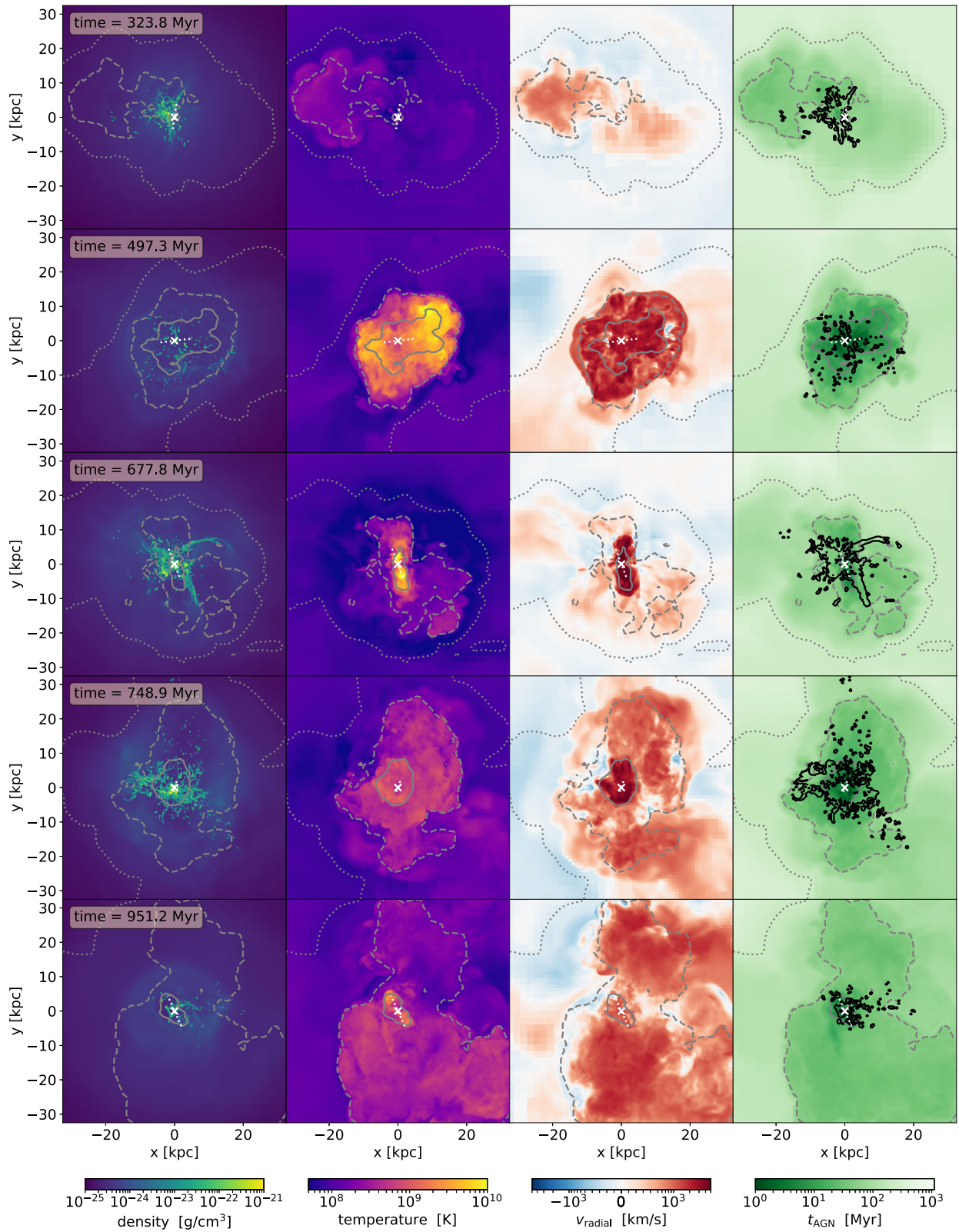
Fig. 1: Projections of (from left to right) density, temperature, radial velocity and the time since a cell has been affected by AGN feedback, t_{AGN}, at five different points in time. Radial velocity and temperature are weighted by t_{AGN}. Radial velocity is measured in 3D space with the SMBH at the origin, and negative velocities are inflowing. Contours are based on the plot of t_{AGN}, and are drawn at 10 (solid), 50 (dashed) and 200 (dotted) Myr. The location of the SMBH is marked by a white cross, and black contours in the right hand column denote the outline of dense gas structures. The white dotted line lies along the instantaneous jet axis, which is plotted to be exactly 10 kpc long in 3D space. The shorter it appears, the more it is aligned with the line of sight of the image (here taken to be the z-axis of the box at all times).
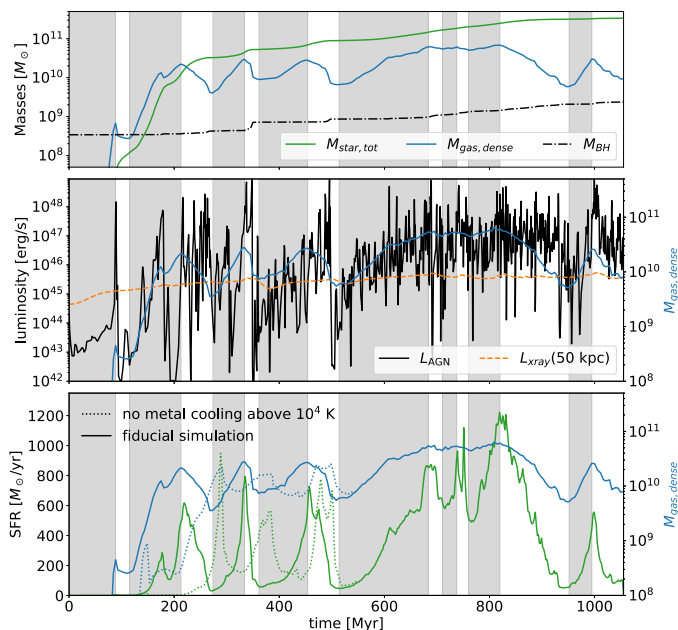
Fig. 2: Time evolution of cluster properties [top panel] including stellar mass M_{star}, BH mass M_{BH} and gas mass M_{gas}. The middle panel shows AGN luminosity, the X-ray luminosity of hot gas within 50 kpc of the cluster center, and the dense mass again for comparison. The bottom panel shows the star formation rate, as well as the dense gas mass again for comparison, for both the fiducial simulation and for a companion simulation without metal cooling for gas with $T > 10^4$ K (see text for details). Dense gas is defined to be gas with a temperature at or below $T_{dense} = 10^6$ K, hot, diffuse gas with a temperature above that. White and grey background colours show the heating and cooling dominated regimes of the fiducial simulation.

that metal line cooling is not the root cause of the over-cooling reported here. It is more likely that the over-cooling occurs due the absence of non-thermal energies from cosmic rays, which are expected to be able to offset as much as 60 % of the thermal cooling in a cluster environment (Pfrommer 2013; Jacob & Pfrommer 2017a,b; Ruszkowski et al. 2017), while only contributing on the percent level to the overall pressure (Reimer et al. 2004; Brown et al. 2011). Due to the large reservoir of heat in cluster outskirts, thermal conductivity in massive clusters can also be an efficient process to bring balance back to the hot cooling gas in the center of clusters (Narayan & Medvedev 2001; Ruszkowski & Oh 2010; Yang & Reynolds 2016b; Kannan et al. 2017). These avenue of investigation will be explored in future work.

3.2. Jet evolution and turbulence in the cluster

One important difference between the work presented here, and previous works on the subject (Li & Bryan 2014a,b; Yang & Reynolds 2016a; Ruszkowski et al. 2017; Li et al. 2017; Cielo et al. 2018; Martizzi et al. 2019; Wang et al. 2019) is that our jet axis is not fixed throughout the simulation, nor do we add explicit precession. Instead, the spin evolution of the BH not only determines the feedback energy but also, crucially, the direction of the jet, as the jet axis is taken to be aligned with the BH spin axis, and the BH spin axis is continuously updated according to the angular momentum of accreted gas.
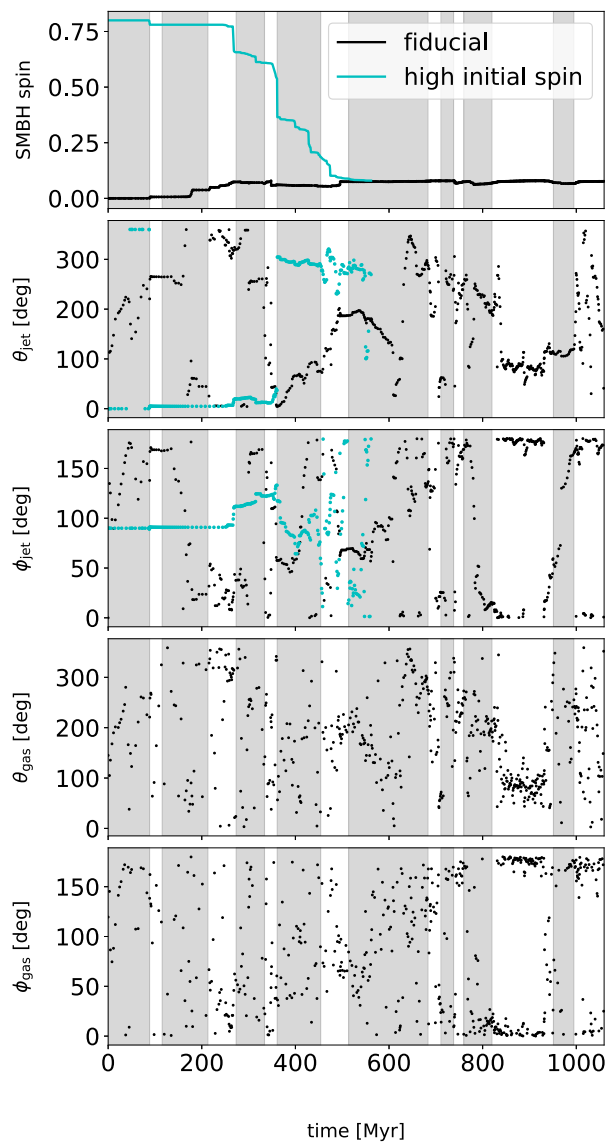
Fig. 3: Spin evolution of the SMBH, showing the spin magnitude (top panel), and the two angles defining the SMBH axis (second + third panel) and the angular momentum of the accreted gas at that particular timestep (bottom two panels). The angles are measured in the box frame, and are defined to be the same as in polar coordinates, where θ_{jet} is measured in the x-y plane of the box (shown in Fig. 1) and ϕ_{jet} is the angle with the z-axis (the line of sight in Fig. 1). Angles are measured in the range $0 \leq \theta < 360°$ and $0 \leq \phi < 180°$. Discontinuous jumps from just below the upper end of the range, to just above the lower end of the range, or vice versa, are due to the cyclic nature of the coordinate system. The top three panels show both the fiducial simulation, and a second, identical simulation initiated with a higher spin value. White and grey background colours show the heating and cooling dominated regimes of the fiducial simulation.

Fig. 3 shows that the direction of the jet explores the full parameter space of the simulation, repeatedly traversing the full range of both polar and azimuthal angles ($0 \leq \theta_{jet} < 360°$ and $0 \leq \phi_{jet} < 180°$). This is a consequence of the chaotic angular momentum accreted by the SMBH. As can be seen in the bottom two panels of Fig 3, the angular momentum of the accreted gas varies extremely rapidly, both in θ and in ϕ, as clumps rain
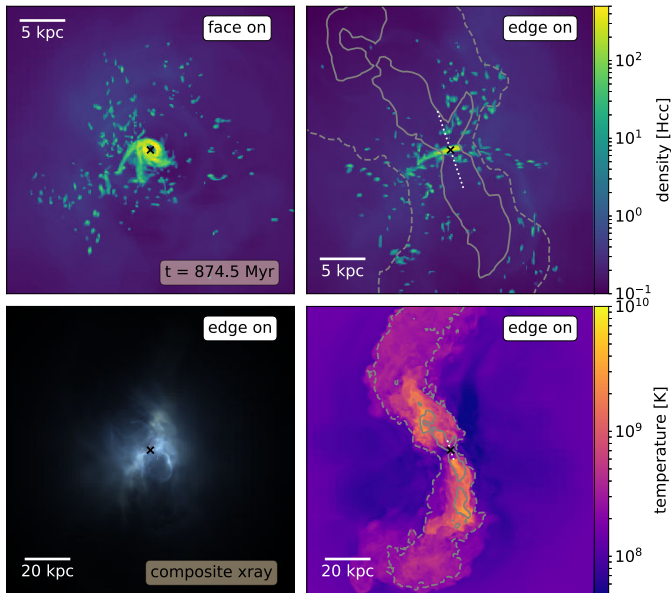
Fig. 4: Projection plots at $t = 874.5$ Myr, showing the central gas disc in the cluster: Top row - density projections of the cluster center along two different lines of sight, Bottom left: composite x-ray image, using the same x-ray bins as in Figure 5, Bottom right: jetscalar weighted temperature projection. Contours mark $t_{AGN} = 10$ and 50 Myr (solid, dashed). The SMBH location is marked by a black cross, and the jet direction is shown by a white dotted line in the right hand panels only.

down on the BH from all directions. As the BH spin evolution is a continuous measure, it varies more slowly than the angular momentum of the accreted gas. The only time both the gas angular momentum and the BH spin direction settle occurs in the period $t = 820-950$ Myr, when both the BH spin and the angular momentum have $\theta \approx 90°$ and ϕ close to zero (the apparent large gap in ϕ at this time is a feature of the coordinate system chosen. $\phi = 2°$ to $\phi = 178°$ only represents a rotation of $4°$, as both $0°$ and $180°$ are aligned with the z-axis of the box). At this time a rotating central gas disc forms around the SMBH, as can be seen in Fig. 4,which drives jet bubbles out to more than 70 kpc from the cluster center. Dense clumps continue to exist at larger radii, but are preferentially found outside the region recently affected by AGN feedback (see solid grey contours on the image).

Our jets self-consistently produces a three-dimensional distribution of fat feedback bubbles seen in Fig. 5, without the need for adding an ad hoc precession or reorientation of the jet (as done in e.g. Li & Bryan 2014a,b; Yang & Reynolds 2016a; Ruszkowski et al. 2017; Li et al. 2017; Cielo et al. 2018; Martizzi et al. 2019; Wang et al. 2019). Firstly, the jet reorientation due to spin helps to self-regulate the cooling flow in clusters (Cielo et al. 2018) by more uniformly redistributing the energy in the hot gas as long as the reorientation is moderate (i.e. not too close to mimicking isotropic energy input, see Gaspari et al. 2012). Secondly, the reorienting jet has important consequences for the distribution of turbulence in the cluster center, as over time a much larger volume is directly affected by the AGN jet. However, the bubbles shown here are less round and more broken up than observed X-ray cavities in clusters. This is due to the fact that in the absence of viscosity and magnetic fields, strong Rayleigh-Taylor and Kelvin-Helmholtz instabilities at the bub-

ble surface break up bubbles prematurely and shorten their overall lifetime (Ogiya et al. 2018).

While we explicitly track the spin evolution of the BH, as described in Section 2.5, the magnitude of the BH spin remains small throughout, as can be seen in Fig. 3, with a maximum spin parameter of 0.08. This is partially a consequence of the model chosen, as the MAD jet model always preferentially reduces the spin of the BH. This low spin value in turn has consequences for the jet direction, as the jet axis is aligned with the BH spin axis. Due to the low spin value of the BH, the chaotic angular momentum of accreted gas (see bottom two panels of Fig. 3), driven by the chaotic infall of the clumps, is able to significantly realign the spin axis throughout the simulation.

As can be seen in Eq. 5, the feedback energy of the BH is determined by the feedback efficiency η_{MAD}, which in turn is determined by the BH spin. Due to the consistently low spin-values, the simulation presented here has an average luminosity-weighted feedback efficiency of only 0.046.

To test the consequences of a higher initial spin value of the BH, we ran a companion simulation to our fiducial simulation. The only difference between the two was that the companion simulation had an initial SMBH spin value of 0.8. As can be seen in the top panel of Fig. 3, the SMBH spin persistently decreases over the course of the simulation, until it converges with the fiducial simulation after ~ 500 Myr. While the spin is high, the jet changes direction very slowly in comparison to the fiducial simulation, as the high angular momentum of the rapidly spinning BH makes reorientation more difficult. Once the spin has dropped below 0.4, the jet direction changes more rapidly and the two simulations become statistically indistinguishable. The bubbles remain comparatively fat even in the absence of precession. This is due to the fact that our jets are very light and hot, and therefore over-pressurized in comparison to the background medium. While injected bimodally, the bubbles quickly expand outwards into the surrounding medium. We note that the absence of magnetic fields, whose wound-up helical structure along the jet is expected to keep it confined over kpc scales (see Pudritz et al. 2012, for a review), will have contributed to the fatness of the bubbles. We therefore postpone a comparison between bubble structures in a high spin and a low spin case to future, magnetised simulations.

3.3. Dense gas structures

3.3.1. Quantifying clump morphology

As can be seen visually in Fig. 1, the dense gas in the cluster center can be found in clumps of a wide range of sizes and shapes. A clump is defined here to be a connected volume of space, for which all cells have a minimum density of 1 H cm^{-3} and a maximum temperature of 10^6 K. All properties are measured by summing over all cells contained within a given clump. Tracer particles are associated with a particular clump if they are contained within the clump volume at the point of measurement.

To quantify this parameter space, we measured the physical extent of individual clumps using the following methodology:

1. Find the center of mass for each clump by summing over all cells contained within the clump, treating each cell as a point mass located at the cell center.
2. Calculate the clumps mass-weighted reduced inertia tensor using

$$I_{i,j} = \sum_{n=1} \frac{m_n x_{n,i} x_{n,j}}{R_n^2} \qquad (7)$$
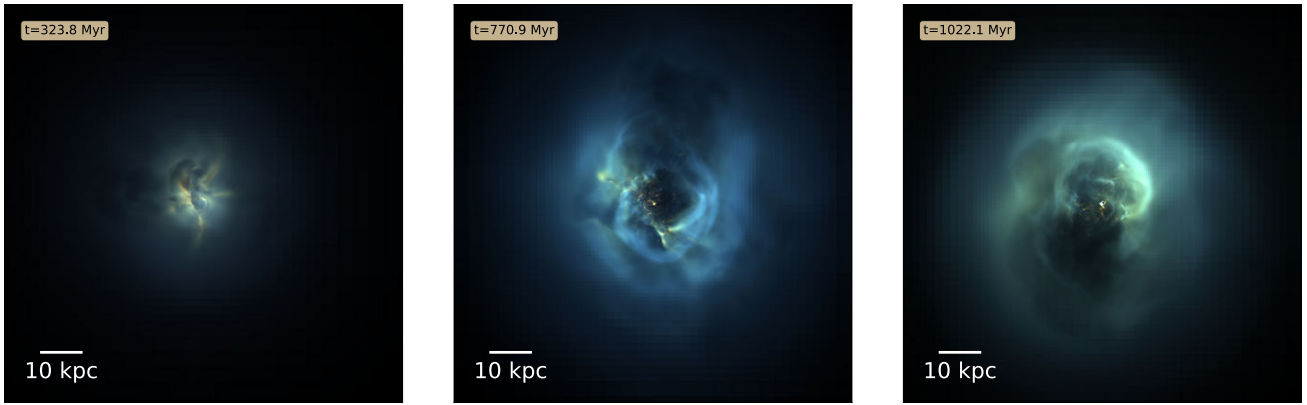
Fig. 5: Synthetic composite X-ray images of the cluster center, with 0.3-1.2 keV in red, 1.2-2 keV in green and 2-7 keV in blue, to match the image of Perseus in Fabian et al. (2006), towards the beginning, middle and end of the simulation. Each channels is scaled to highlight fainter features. Each image is 100 kpc across.

by summing over all cells n contained within a clump, where $x_{n,i}$ is the ith coordinate of the nth cell within the clump, measured in the center of mass frame of the clump. R_n is the nths cells distance from said center of mass, and m_n is its gas mass.

3. Calculate the physical extent of the major axis r_{maj} by finding the largest distance between any two cell centers contained within the clump. To this value, Δx_{min} is added to extrapolate from the cell centers to the cell edges contained within the clumps.

4. Find the axis vectors and axis length ratios using the eigenvalues and eigenvectors of the inertia tensor from Eq. 7.

5. Calculate the median and minor axis length, r_{med} and r_{min} respectively, using the axis length ratios from the previous step, and the length of the major axis, r_{maj}.

6. Calculate the volume of the ellipse defined by the three axes:

$$V_{ellipse} = \frac{4}{3}\pi r_{maj} r_{med} r_{min}. \tag{8}$$

7. Calculate the volume filling fraction f_V, which is defined to be the ratio of the volume defined by the axis vectors, $V_{ellipse}$ in Eq. 8, and the sum of the cell volumes contained within the clump:

$$f_V = \frac{V_{ellipse}}{\sum_n V_n} \tag{9}$$

where V_n is the volume of the nth cell contained in the clump. For solid, round clumps well described by an ellipse, f_V will have a value close to unity. For clumps with a complex morphology, such as bent filaments and three-dimensional networks of filaments and clumps, the volume fraction will be low as the axis vectors used to describe the ellipse mark the total physical extent of the clump along a given axis vector in 3D space, and said ellipse will therefore contain many cells outside the clump.

For further analysis, we split the population of clumps into three categories depending on the length of their major axis relative to the mean major axis of the whole sample, $\bar{r}_{maj} = 1.54$ kpc, and the samples standard deviation $\sigma_{maj} = 1.42$ kpc:

1. **small clumps** have a major axis $r_{maj} < \bar{r}_{maj} = 1.54$ kpc.

2. **big clumps** have a major axis length in the range $\bar{r}_{maj} = 1.54 < r_{maj} < \bar{r}_{maj} + \sigma_{maj} = 2.96$ kpc.

3. **filaments** have $r_{maj} > \bar{r}_{maj} + \sigma_{maj} = 2.96$ kpc.

Some example decompositions according to these criteria can be seen in Fig. 6.

3.3.2. Clump properties

A variety of bulk clump properties versus axis length are shown in Fig. 7, for the stacked sample of clumps of the whole simulation. As can be seen in column (a), the distribution of major axis lengths ranges from the resolution limit of the simulation to very large, extended objects that have major axes of the order 10 kpc or more. The stacked sample shown here, which contains all objects from all snapshots at all points in time of the simulation, contains 37897 small clumps (87.4 %), 4283 big clumps (9.9 %) and 1153 filaments (2.7 %).

As expected, smaller clumps contain less gas mass (Fig. 7, column (b)), with a minimum gas mass for the current resolution of 5×10^5 M$_\odot$, and an average value of 1.8×10^7 M$_\odot$ for small clumps and 1.1×10^8 M$_\odot$ for large clumps. The population of filaments is much more massive, with an average gas mass of 2.5×10^9 M$_\odot$. Structures with a mass above 10^9 M$_\odot$ are all classified as filamentary. This lower mass limit for gas clumps is determined by the resolution. As we tested with a companion simulation, in which we reduced Δx_{min} to 30 pc, i.e. a factor 4 smaller than in the fiducial simulation. With this improved resolution, the gas structures fragment further into even smaller clumps, with a new minimum mass of 2.2×10^3 M$_\odot$, and a new minimum axis length still approaching the resolution limit. This suggests that the shattering into smaller structures is by no means complete, and with even more resolution, the clumps would continue to break apart, as in the "cloudlet" model by McCourt et al. (2018). However, larger, filamentary structures continued to exist even in the higher resolution simulation.

In terms of shape, smaller clumps have higher values of f_V, so they are indeed much more compact (column (c), Fig 7). Values of $f_V > 1$ can occur for compact objects when the axis length for the median and minor axis are under-estimated in comparison to the true extent of the clump which happens mainly for clumps with less than 20 cells. However, the volume of the ellipse used to fit the clump never exceeds that of the sum of the cells contained in the clump by more than 40 %. More extended
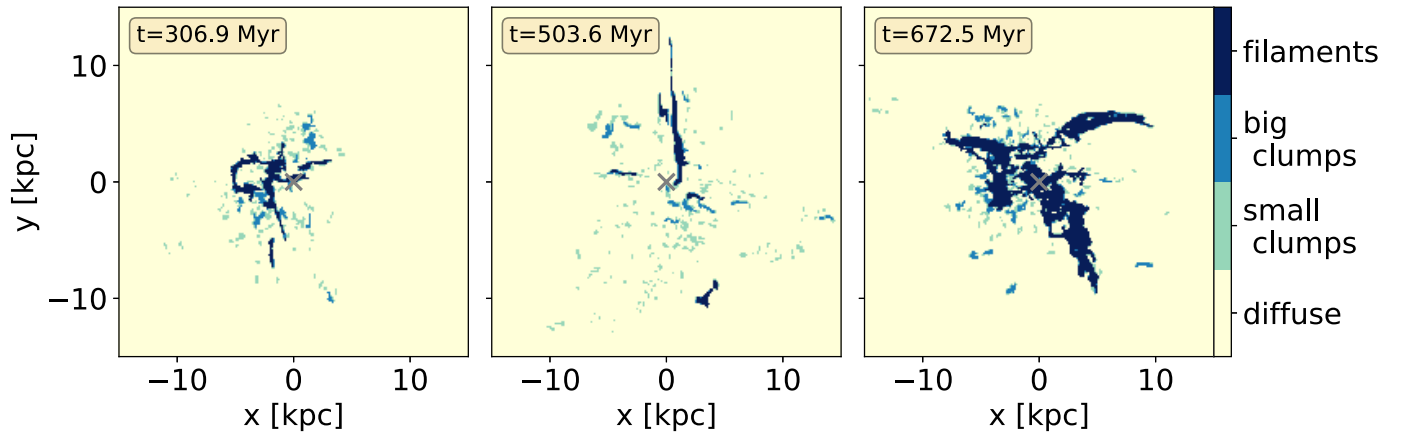
Fig. 6: Example projections of the decomposition of structures into small clumps, big clumps and filamentary structures at three different points in time. Structures are considered to be distinct when not connected in 3D space. From left to right, the plots contain 5, 5 and 9 distinct filaments respectively, partially superimposed due to projection effects.
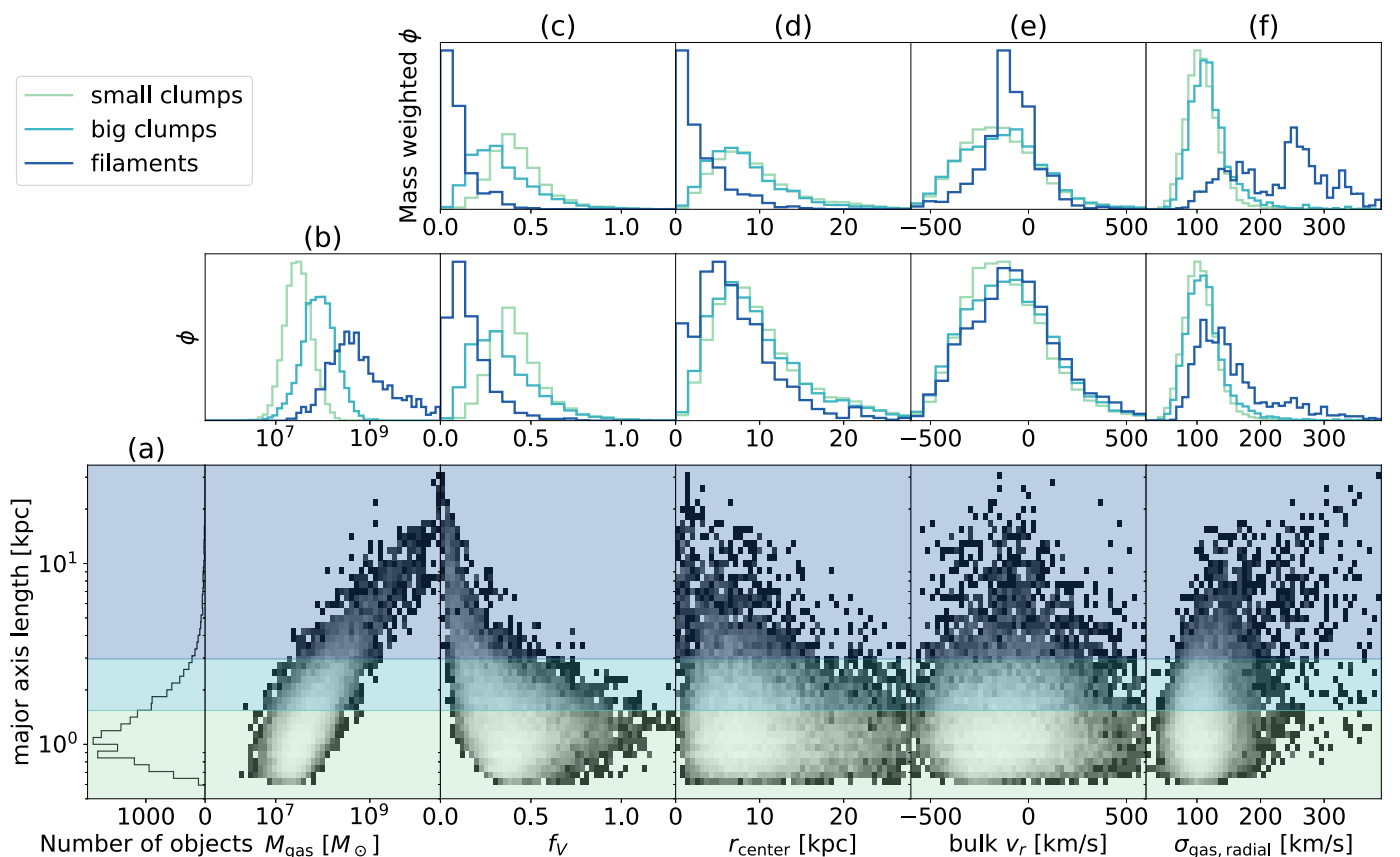
Fig. 7: Clump properties for the whole sample (bottom row) and split into the three structure categories (top two rows). From left to right: clump gas mass M_{gas}, volume ratio f_V, distance between the clump center of mass and the cluster center r_{center}, bulk velocity v_r and gas velocity dispersion within the clump $\sigma_{gas,radial}$. The probability distributions ϕ in the top row is mass weighted, while the one in the row below is unweighted.

objects have f_V far below unity, which is an indicator of complex morphology. The most clumpy filament produced in this simulation still has $f_V < 0.7$ so large clumpy structures do not form at any point of the simulation.

Small and big clumps have a similar radial distribution (column (d), Fig. 7) and are preferentially found between $3 - 10$ kpc from the cluster center. Filaments, on the other hand, include both a subsample found at large radii, and a sample of particularly extended structures in the cluster center, an example of which can be seen in the right hand panel of Fig. 6. This suggests that gas structures merge into larger objects as they reach the cluster center, consistent with a model in which small structures rain down onto a central massive gas structure. This structure can take the form of a massive gas disk, as for example seen

in Li & Bryan (2014a) and briefly also in the simulation presented here (see Fig. 4), or in the form of an extended but not rotationally-supported object such as the one in the right hand panel of Fig. 6, or the gas structures seen in the first, third and fourth snapshot of Fig. 1.

In velocity space, all three populations are similarly distributed (column (e), Fig. 7), with no discernible difference in the unweighted probability distribution of small and medium clumps, as well as filaments. The mass-weighted distribution in the top row shows that all three categories of structures are preferentially infalling (i.e. have $v_r < 0$). The time-stacked sample of the simulation has an unweighted mean radial velocity of $75\,\mathrm{km\,s^{-1}}$, with a full width half max of $198\,\mathrm{km\,s^{-1}}$, where radial velocity is measured in 3D space with the SMBH at the origin. Negative values denote gas falling towards the SMBH. These values are comparable to observed bulk velocities of $100\,\mathrm{km\,s^{-1}}$ but are at the upper end of observed velocity widths of $100 - 218\,\mathrm{km\,s^{-1}}$ for molecular gas in Perseus (Salomé, P. et al. 2008; Hitomi Collaboration 2016; Gendron-Marsolais et al. 2018). By comparison, they fall easily within the range of observed velocity widths for warm ionised gas in massive clusters (Hamer et al. 2016). We note that, in contrast to the observational values, the full width-half max calculated here is calculated across the entire time-stacked sample, not just along the line of sight. While the mean and dispersion values show good agreement with observations, the sample of clumps presented here has an overall larger velocity range than found in cold-gas maps of nearby clusters, which report velocity values across the map in the range of $350\,\mathrm{km\,s^{-1}}$ at most (Olivares et al. 2019; Gendron-Marsolais et al. 2018).

The velocity dispersion $\sigma_{\mathrm{gas,radial}}$ is defined to be the velocity dispersion of the radial velocities of all resolution elements within an individual clump. It therefore quantifies the range of velocities found within an individual object. Clumpy structures, both small and big, have a low velocity dispersion (column (f), Fig 7), i.e. a small range of radial velocities, with an average value of just $90\,\mathrm{km\,s^{-1}}$. The bulk of the filaments, despite major axis lengths of 10 kpc or more, have radial velocity dispersion of less than $200\,\mathrm{km\,s^{-1}}$ but there is a small population of high-velocity dispersion objects with $\sigma_{\mathrm{gas}} > 200\,\mathrm{km\,s^{-1}}$, which is preferentially populated by filaments: They make up 28 % of the high dispersion objects versus only 2.7 % of the total sample.

Dynamically, the clumps are therefore a surprisingly uniform population, despite more than 2 orders of magnitude in size difference, and more than 4 orders of magnitude in mass range. Gas properties across all three populations are also similar, with a temperature range of $10 - 10^6$ K (the latter being the cut-off temperature for the definition of a dense gas structure in this paper), and densities in the range of $1 - 10^5\,\mathrm{H\,cm^{-3}}$. The bulk of the gas has a temperature around 10^4 K and a density of $10 - 10^3\,\mathrm{H\,cm^{-3}}$. This is not to say that all objects have the same properties at a given point in time, but that all types of objects can be found at all points in phase space at some point throughout the simulation.

The morphology and distribution of objects can vary strongly on a 5 Myr timescale, as can be seen in Fig. 8. Overall, the number of structures at all points in time is dominated by small clumps, who are always the most abundant and make up 87.4 % of the time-integrated sample. During some parts of the cooling-dominated phase, they also contain the bulk of the dense gas mass, such as around 400 Myr and at $500 - 550$ Myr. The rest of the time, the bulk of dense gas mass can be found in filaments, despite the fact that they only make up 2.7 % of the overall sample by number. Big clumps contain dense gas mass on the order
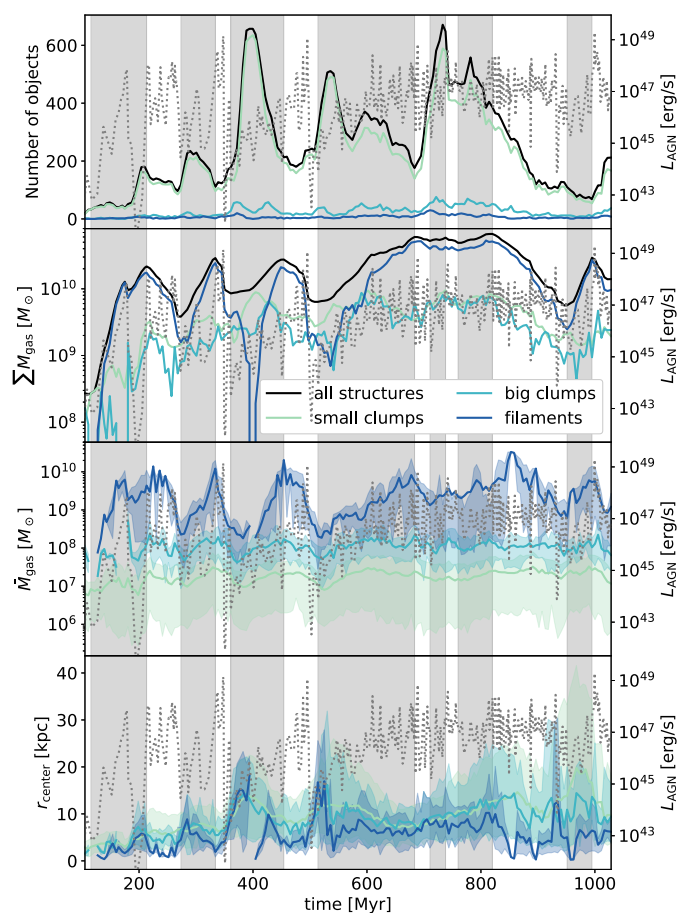
Fig. 8: Time evolution of (from top to bottom) the number, total dense gas mass, mean gas mass and mean distance to the cluster center for the three structure categories. In the bottom two panels, solid lines show the mean and shaded regions the range from the 10th to the 90th percentile of the distribution. The dashed grey line in all plots shows the AGN luminosity for comparison.

of that contained in the small clumps, but represent 9.9 % of the total number of objects.

From Fig. 8, strong bursts of AGN feedback are followed by a strong increase in the number of small clumps, as well as an equally strong drop in both the total mass of gas contained in filaments (second panel) and the average mass of gas per filament (third panel). At the same time, the average radial distance between the cluster center and a clumps center of mass increases (bottom panel). While the bulk of clumps can usually be found within the central 20 kpc of the cluster, strong AGN outbursts produce clumps at much larger radii, up to 50 kpc from the location of the cluster center. This suggests a scenario where large objects are being shattered into smaller clumps during their interaction with strong AGN jets, and highlights the importance of the AGN jet not just for slowing down cooling onto the cluster center but also for the morphology and kinematics of the existing dense gas structures. The details of this interaction will be explored further in the next section.

3.4. Uplifting

Uplifting has been used to explain the non-structures velocity profiles observed in nearby clusters (Pulido et al. 2018; Gendron-Marsolais et al. 2018). When talking about uplifting

dense gas in clusters, two different mechanisms need to be distinguished. On the one hand, there is the entrainment of existing dense gas by the AGN driven outflows, which turns previously infalling dense gas into outflowing dense gas, which will be discussed in this section. Alternatively, outflowing dense gas could form via condensation, where the lower-entropy gas from the hot atmosphere is compressed and lifted to larger radii by the AGN jet, where it cools while still outflowing[1]. The latter scenario in particular has been advocated as a source of the non-structured velocity profiles seen in nearby clusters, and will be discussed in Section 3.5.

The impact one interaction between the AGN jet and the dense gas in the cluster center, namely the outburst at $320 - 400$ Myr, is shown visually in Fig. 9: at $t = 323$ Myr (top left), the dense gas is predominantly infalling and contained in radially oriented filaments. At this point in time, the filaments contain $M_{\rm gas,filaments} = 1.7 \times 10^{10} M_\odot$, i.e. 76 % of the total dense gas mass, with an average gas mass per filament of $\bar{M}_{\rm gas,filaments} = 1.47 \times 10^8 \, M_\odot$. As the AGN outburst commences, fed by this infalling dense gas ($t = 338 - 356$ Myr, middle and top right panel), the filaments are broken into small and medium size clumps, and their velocity turns from infalling to outflowing. By 371 Myr (bottom left), gas is predominately outflowing, and the total mass budget of 8.8×10^{10} M$_\odot$ is evenly split between small clumps, medium clumps and filaments. The filaments that continue to exist are much less massive, with an average mass of just $\bar{M}_{\rm gas,filaments} = 4.2 \times 10^7$ M$_\odot$.

By $t = 388$ Myr, the gas has reached its largest radial extent for this episode and is beginning to fall back onto the cluster center in the form of a shower of small, distinct clumps. From 371.9 Myr to 388.5 Myr, the total gas only increases by 5 %, from 8.8×10^{10} M$_\odot$ to 9.3×10^{10} M$_\odot$, but the total number of objects triples as objects continue to break apart, from 244 at 371.9 Myr to 651 individual objects by 388.5 Myr. By this point, small clumps dominate the population, as they represent 94 % of objects and contain 64 % of the total gas mass, with a further 27 % contained in big clumps.

The timeseries of the number of different objects in the top panel of Fig. 8 shows that this behaviour is generic for the cluster presented here. Following a strong feedback outburst, the number of small objects spikes, while the total gas mass and the average mass per filament decrease strongly. At the same time, the average distance for objects of all categories increases as they are ejected from the cluster, with the outermost small clumps being found as far as 40 kpc or more from the cluster center.

Looking directly at the number of inflowing and outflowing objects, as shown in Fig. 10, strong AGN feedback bursts are followed by a spike in the number of outflowing objects, as larger, filamentary structures are entrained and broken up by the hot winds of AGN feedback and lifted to larger radii. As gas is evacuated from the cluster center the AGN turns off. The entrained clumps then decelerate under gravity and fall back onto the cluster center. During this process, they shatter into even smaller components so the number of individual objects continues to increase even after the AGN has become quiescent again. As the small clumps fall back onto the cluster center, they coalesce and trigger another strong outburst of AGN feedback, which repeats the cycle. The results presented in this paper are similar to work by Yang & Reynolds (2016a), who presented evidence for ex-

isting dense gas to be redistributed by the AGN jet. Contrary to their work, the dense gas in the simulations presented here is not indestructible. In our simulations, only 25 % of the dense gas survives its interaction with the hot jet. It gets entrained by the AGN driven outflows and lifted to large radii. We note that, with a temperature cut of 10^6 K, the gas discussed here is equivalent to the ionised dense gas seen in observation, not to the molecular gas. We expect that if we were able to adequately distinguish between ionised warm gas and molecular cold gas, the molecular gas would be much more difficult to uplift by the AGN jet.

This is surprising in the context of work by Klein et al. (1994), who showed that for adiabatic cold structures in hot winds, the drag timescale $t_{\rm drag} \approx \chi r_{\rm clump}/v_{\rm wind}$ is always longer than the clump crushing timescale $t_{\rm cc} \approx \chi^{1/2} r_{\rm clump}/v_{\rm wind}$, where χ is the density contrast between wind and cold clump, $r_{\rm clump}$ is the clump radius and $v_{\rm wind}$ is the relative velocity. It should therefore be impossible to accelerate cold clumps with a hot wind. However, recent work by Gronke & Oh (2018) shows that radiative cooling can replenish the cold clump mass from the hot gas during uplifting and thereby dramatically increase the clump lifetime. Under these assumptions, clumps with radii larger than $r_{\rm clump} > v_{\rm wind} t_{\rm cool,mixing}/\chi$, where $t_{\rm cool,mixing}$ is the cooling time in the mixing layer surrounding the cold clumps, should survive the uplifting process, as cooling from the hot to the cold phase replenishes gas faster than cold gas from the clumps is being evaporated. For the simulation presented here, the maximum outflow velocities in the vicinity of clumps is of the order 10^4 km s^{-1}, the cooling time in the mixing layer around clumps is of the order 0.1 Myr and the density contrast $\chi \approx 10^4$. Therefore, clumps with a minimum value of $r_{\rm clump} \approx 1$ pc should survive their interaction with the hot wind, which is much smaller than the smallest cell size of 120 pc. While poorly resolved clumps most likely lack this mixing layer, and are therefore destroyed during the jet interaction, well-resolved cold clouds would be expected to survive their interaction with the hot outflows and become entrained without being destroyed, as shown in Fig. 9. These results are also in agreement with work by Armillotta et al. (2017), who show that the bulk of cold gas in clouds with radii above 250 pc survives being accelerated by a hot wind for 200 Myr. It is however likely that the 25 % of dense gas that survives the interaction in our simulations is an overestimate, as work by Sparre et al. (2019) showed that more highly resolved clouds shatter more efficiently during their interaction with hot winds and therefore have shorter overall lifetimes than less resolved clouds.

In comparison to the observed velocity maps for Hα emitting gas in Perseus by Gendron-Marsolais et al. (2018), the velocity maps shown in Fig. 9 are much more episodic, with clumps either predominantly infalling or outflowing. However, as can be seen in Fig. 10, this particular AGN outburst is unusual in that it is the only time within the simulation when the number of infalling clumps falls almost to zero due to the AGN outburst. It was chosen here as it illustrates uplifting by AGN feedback particularly cleanly but we note that at all other points in time, dense gas can be observed to be inflowing and outflowing at the same time, due to the directionality of the jet and the limited width of the jet cone.

It is also important to remember that the observed velocities are line-of-sight velocities, while Fig. 9 shows radial velocities. As can be seen visually in Fig. 11, which shows both radial velocities (top row) and line-of-sight velocities (bottom row) for an inflowing dominated (left column), an outflow dominated (middle panel) and a mixed (right column) point in time, the line-of-sight velocities appear less ordered than the radial ve-

---

[1] For an adiabatically compressed gas, the net effect is an enhanced cooling rate since adiabatic compression leads to a pressure ratio of $\mathcal{R}_\rho^\gamma$, where $\mathcal{R}_\rho$ is the gas density ratio, and, hence, an enhancement of the Bremsstrahlung cooling rate of $\mathcal{R}_\rho^{2+(\gamma-1)/2} = \mathcal{R}_\rho^{7/3}$ for $\gamma = 5/3$.
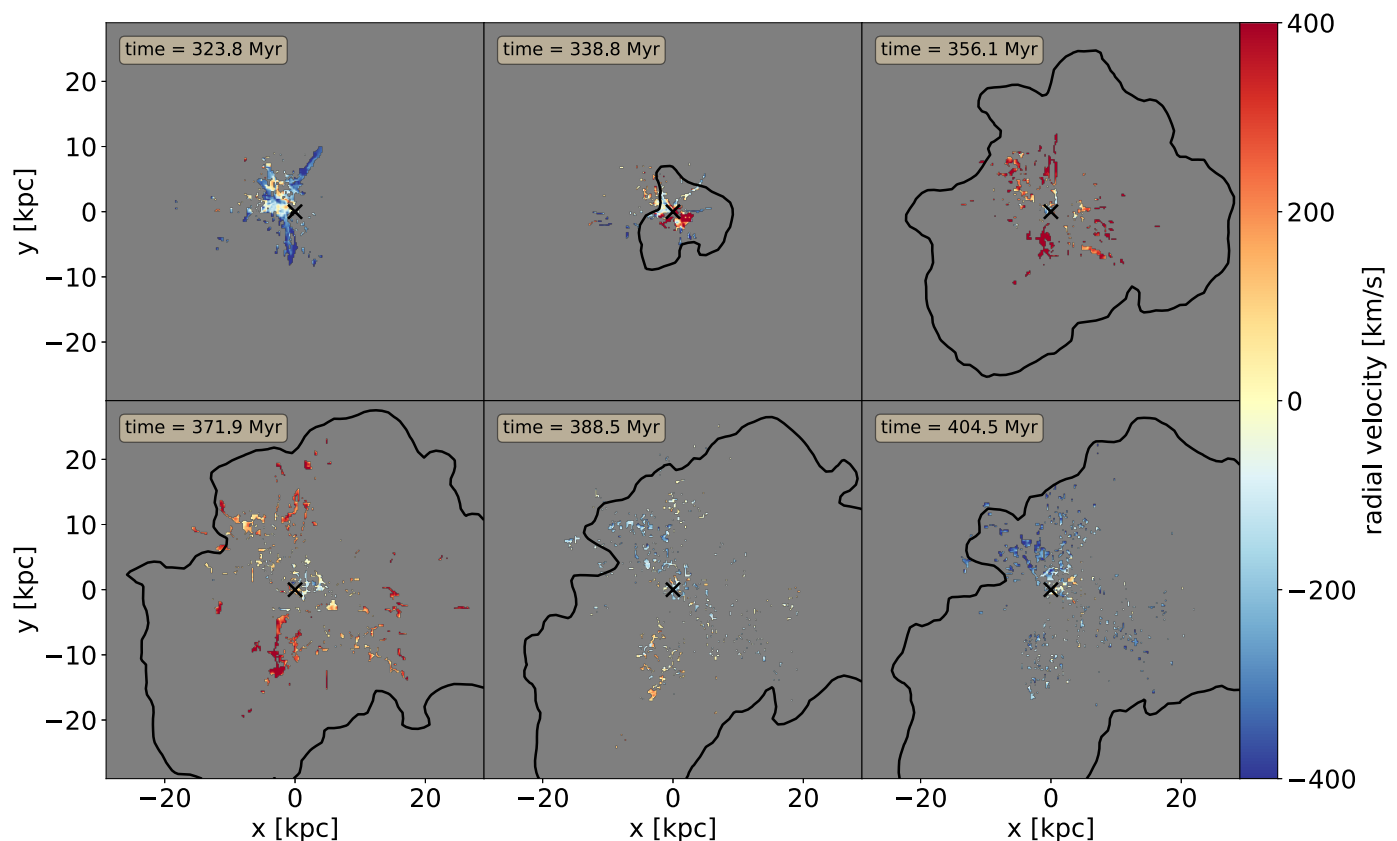
(a) Visual time evolution of one episode of AGN feedback that starts around $t = 350$ Myr. Only the dense gas is plotted. The colourmap shows the radial velocity of the gas, with negative values denoting infall, with the background color set to grey for clarity. The location of the BH is marked by a cross, and the contours show the extent of the AGN feedback bubbles produced by the feedback event that starts at $t = 323$ Myr.
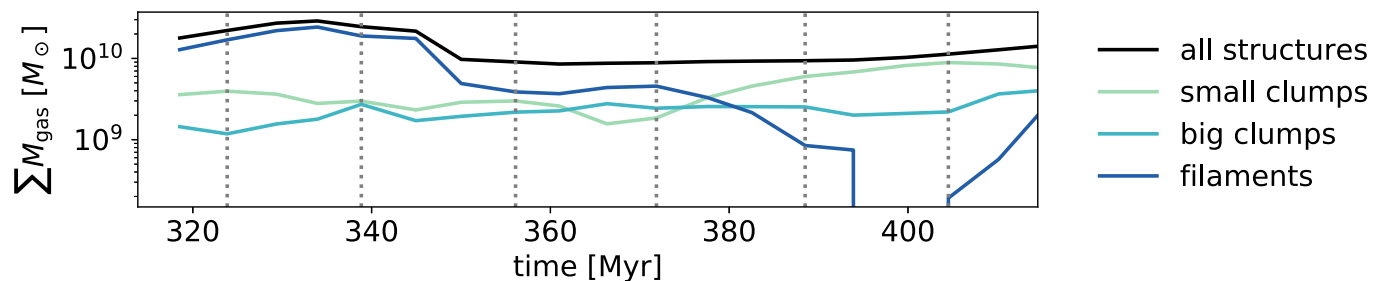
(b) Time evolution of dense gas mass contained in the three categories over the same period of time. Vertical grey lines mark the outputs shown in the top panel.

Fig. 9: Visual time evolution of one episode of AGN feedback that starts around $t = 350$ Myr (top plot) and time evolution of dense gas mass contained in the three categories over the same period of time (bottom plot).

locities. The outflow or inflow dominated nature of the flow (left or middle panels respectively) cannot easily be recovered from line-of-sight velocity maps. This difficulty in distinguishing between flow patterns in the frame of the cluster, and line-of-sight flow patterns, is even more obvious in Fig. 12, which shows the radial velocity probability distributions for the three snapshots in Fig. 11, as well as that for the three line-of-sight velocities (here aligned with the *x*-axis, *y*-axis and *z*-axis of the box respectively). In all three cases, the line-of-sight velocities fail to recover the radial velocity pattern, even if the flow is clearly inflow or outflow dominated, and predict a more gaussian-like pattern with a mean velocity close to zero. The Gaussian distribution of line-of-sight velocities is expected for infalling or outflowing gas distributed roughly spherically around the clus-

ter center. Therefore, the chaotic velocity patterns observed in nearby clusters are not necessarily evidence for the absence of coherent radial flow of the gas.

3.5. Condensation

As first proposed in McCourt et al. (2012), and then shown in idealised simulations by Sharma et al. (2012), dense gas can form out of the hot ICM via local thermal instability, even if the cluster is globally thermally stable. Thermal instabilities are thought to occur when the ratio of t_{cool}/t_{ff} falls below a given value, with a maximum value for the formation of molecular gas
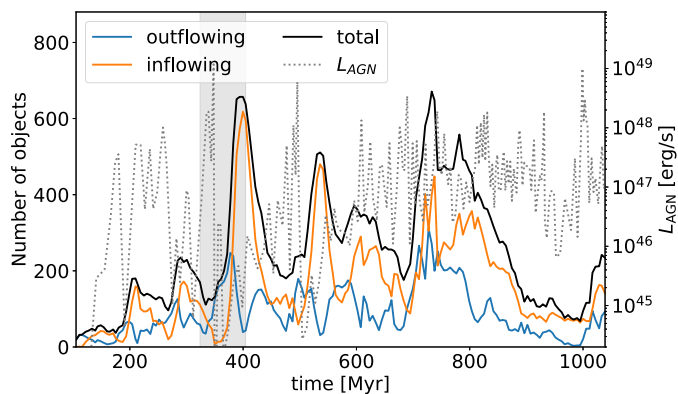
Fig. 10: The total number of inflowing and outflowing clumps The AGN luminosity is shown as a dotted line for comparison. The solid background highlights the event shown in Fig. 9.

in the range $10 - 30$ as seen in observations (Hogan et al. 2017; Pulido et al. 2018; Olivares et al. 2019).

In the simulation presented here, we can use the tracer particles to estimate the condensation rate of dense gas. As each tracer particle has a unique identification number and traces 2×10^4 M$_\odot$ of gas mass, the trajectories of tracer particles can be used to track gas flows throughout the simulation. The total mass of gas transferred from the hot, diffuse to dense phase between two simulation outputs can be estimated by counting the number of tracer particles that pass from the diffuse phase to the dense phase between two simulation outputs. The condensation rate $\dot{M}_{\rm condensed}$ is then found by dividing the newly condensed gas mass $M_{\rm condensed}$ by the time it took to assemble it.

As can be seen in the left hand panels of Fig. 13, our simulation confirms that condensation primarily occurs when $t_{\rm cool}/t_{\rm ff} < 20$. This is somewhat higher than prediction from idealised cooling simulations (Sharma et al. 2012; McCourt et al. 2012), most likely because the hot gas along the jet drives up the spherically averaged cooling time, but in line with observed values (Hogan et al. 2017; Olivares et al. 2019). Profiles of $t_{\rm cool}/t_{\rm ff}$ during the cooling dominated phases, which produce the bulk of the condensation, are generally ordered, with a clear minimum around 10 kpc. During heating dominated phases, by contrast, profiles show a much wider range of shapes as gas heated by the AGN rises to large radii in the form of hot bubbles, which significantly increase the cooling time both in the center and at larger radii. Some condensation continues during the heating dominated phases, and while the condensation remains confined to < 20 kpc from the cluster centers, the values of $t_{\rm cool}/t_{\rm ff}$ can be as high as 50 even for actively cooling clusters. We postulate that this continued condensation is due to the multiphase structure of the ICM and the directionality of AGN feedback. Both $t_{\rm cool}$ and $t_{\rm ff}$ are calculated for the hot ICM only, and it takes even strong AGN feedback bursts some time to reach large volume filling factors and shut off condensation completely.

This theory is confirmed by the condensation time-series in Fig. 14, which shows that condensation is highest towards the minimum of heating-dominated phases and falls to zero as the AGN feedback continues of impact the ICM. Fig. 14 also shows that at the end of cooling-dominated phases, condensation occurs preferentially onto filamentary structures, but by the end of heating-dominated phases and the beginning of the next cooling-dominated phases, condensation occurs preferentially onto small

and big clumps, in line with the uplifting - shattering - recondensation picture presented in Section 3.4.

As can be seen in Fig. 14, the total condensation rate of the cluster varies with time, ranging from a minimum of 3 M$_\odot$ yr^{-1} at the beginning of cooling dominated times to a maximum of up to 1.8×10^3 M$_\odot$ yr^{-1} towards the end of cooling dominated phases. While the bulk of condensation takes place onto filaments, smaller and big clumps dominate when condensation rates are low. As discussed in the context of the clusters SFR in Section 3.1, this condensation rate is high in comparison to the observed condensation rate for Perseus, which is in the range of $50 - 100$ M$_\odot$ (Fabian 2012). In future work, we will explore is this over-cooling occurs because of the omission of non-thermal energies from cosmic rays in the work presented here, which are expected to be able to offset as much as 60 % of the thermal cooling in a cluster environment (Pfrommer 2013; Jacob & Pfrommer 2017a,b; Ruszkowski et al. 2017).

While the areas of high condensation rate are confined to the minima of the $t_{\rm cool}/t_{\rm ff}$ profiles, dense gas can be found over a much wider range of radii (see righthand panels of Fig. 13), and significant amounts of dense gas can also be observed during heating-dominated times. This is due to the fact that existing dense gas free-falls onto the cluster center from its formation location around 10 kpc, and is uplifted to larger radii due to its interactions with AGN feedback. The location at which dense gas is observed is therefore not a perfect proxy for where it is formed, as the kinematics in active clusters are complex and subject to hysteresis.

This can be seen in more detail when comparing the radial and velocity distributions for stacked samples of newly condensed gas (left panel) and dense gas (right panel) in Fig. 15. While some amount of condensation occurs over the full parameter space of radii and velocities occupied by dense clumps, the distribution in both radius and velocity is different for newly condensed gas and dense gas in general. As shown in both the mass distribution in Fig. 15, and in the probability distributions in Fig. 16, dense gas is preferentially found at the cluster center, whereas condensation preferentially occurs at larger radii, with a peak of the distribution at 10 kpc. In velocity space, both existing dense gas and new condensation are preferentially infalling, but condensation has a broader distribution towards negative values, with a mean velocity at -155.6 km/s for condensation compared to -104.3 km/s for dense gas. Overall, only 75.9 % of gas is infalling, while 82.1 % of condensation occurs onto infalling clumps. This means that while the bulk of newly condensed gas is infalling, with an average velocity-weighted condensation rate onto inflowing gas of $= 3.09$ M$_\odot$/yr, there is also evidence for gas condensation onto outflowing clumps, which have an average velocity-weighted condensation rate of 0.65 M$_\odot$/yr.

We therefore conclude that condensation occurs preferentially onto infalling clumps within the radial range of $5 - 15$ kpc, but approximately a fifth of all condensation occurs onto outflowing gas.

4. Discussion

In this paper, we have studied the formation, evolution and destruction of dense gas in the center of a Perseus-like cluster, under the influence of a spin-driven AGN jet, with a particular focus on the role of uplifting and condensation on the kinematics and morphology of the dense gas.
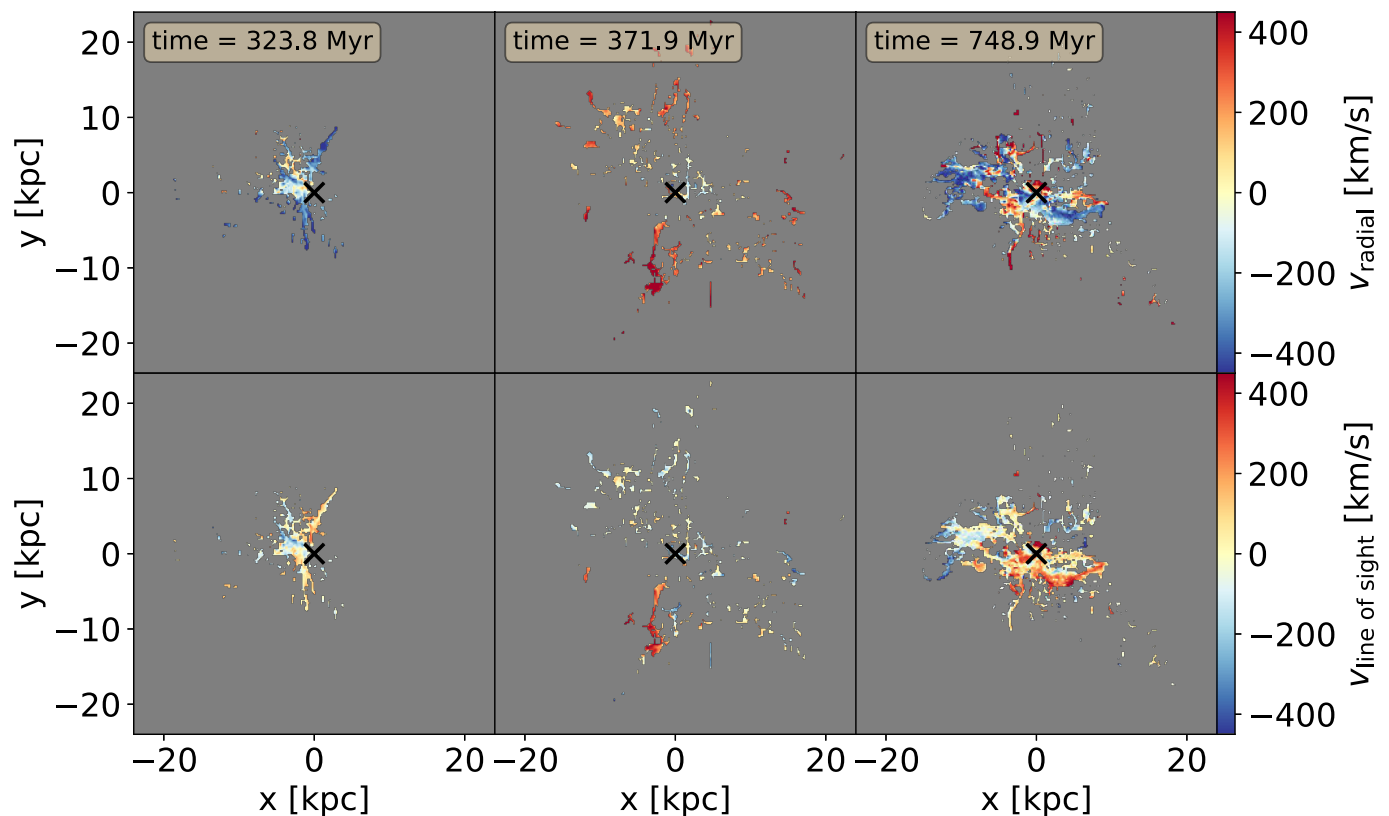
Fig. 11: Density weighted velocity projections of the dense gas at three different points in time. The top row shows the radial velocity for each snapshot, the bottom row the corresponding line of sight velocity (here chosen to be the z-axis of the simulation box).
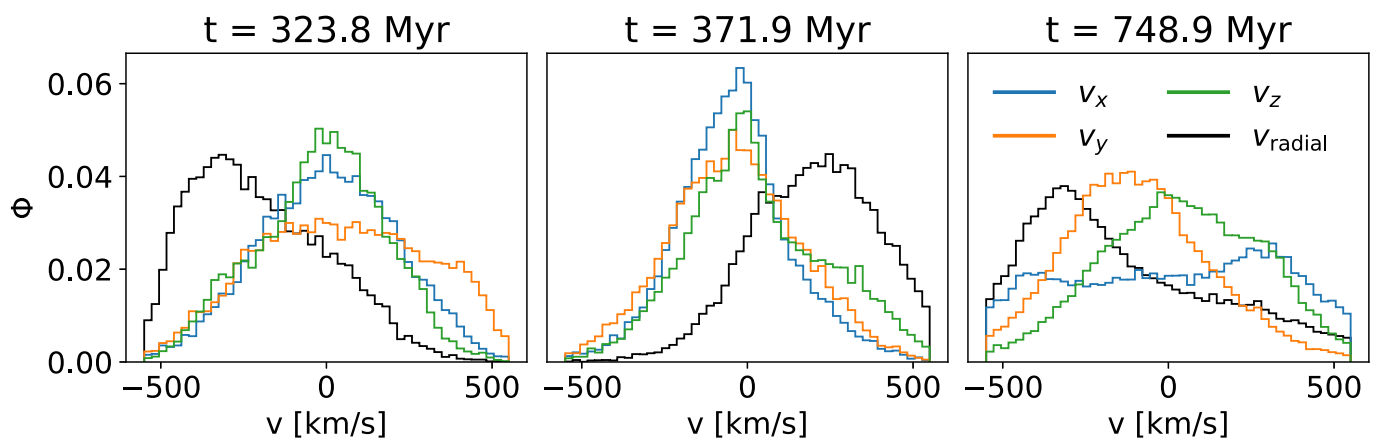

Fig. 12: Distribution of resolution elements in radial velocity, and line of sight velocity along the x-axis, y-axis and z-axis of the simulation box respectively, for the three snapshots in time shown in Fig. 11.

4.1. Filament lifetimes

One notable result of our simulations is that extended gas structures form preferentially during comparatively AGN quiet times, and are readily destroyed in the interaction with AGN feedback. While this interaction between dense gas and AGN feedback is one of the requirements for effective self-regulation of cooling in the cluster, it also means that the lifetime of filaments is limited by the length of AGN duty cycles.

While as much as 25 % of the dense gas mass survives the interaction with the hot, AGN driven outflows, larger structures

are broken into smaller structures in the process. The result is a volume-filling distribution of small clumps, which are at first outflowing and then fall back onto the cluster center. Such a clumpy morphology of the dense gas is not supported by observations, which show more extended, filamentary structures (Conselice et al. 2001; Fabian et al. 2006). Two possible explanations come to mind.

One possibility is that the dense filaments are too readily destroyed in our simulations. If physical processes not modelled here, such as notably magnetic fields, could support the filaments against fragmentation, they might survive their interaction with
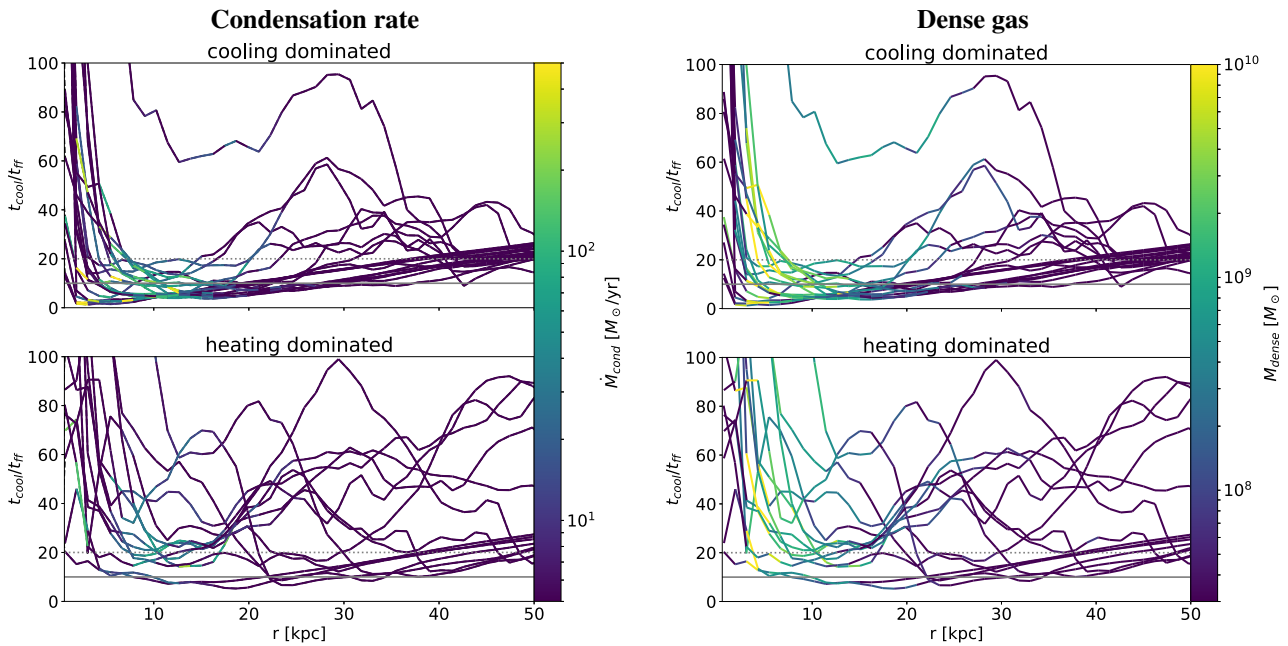
Fig. 13: Cluster profiles of the cooling time (t_{cool}) to free fall time (t_{ff}) ratio at different snapshots of the simulation. The cluster profiles are sampled each 25 Myr across the full time evolution of the simulation. t_{cool} is calculated for each cell in the simulation, using its instantaneous density, temperature and cooling function as computed by RAMSES. t_{ff} is calculated using all mass (DM, gas, stars and the SMBH) contained within a given radius. Profiles are colour-coded by condensation rate (left) or dense gas mass (right), based on the condensation rate and dense gas mass onto clumps at that radius. Snapshots during cooling dominated times (top panel) and heating dominated times (bottom panel) are plotted separately for clarity.
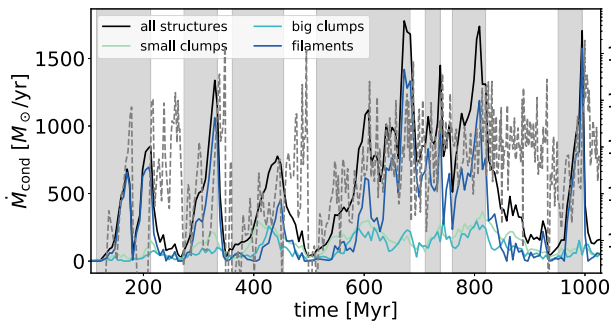
Fig. 14: Time evolution for the gas condensation rate onto the dense structures in the simulation. See text for how the condensation rate is calculated.

Fig. 15: Phase plot showing the total condensation rate (left) and total dense gas mass (right) over a range of radial positions and radial velocities of the clumps. Data shown here is stacked over all clumps at all snapshots of the simulation.

the AGN jet and retain their extended morphology for longer. This theory is supported by work on the survival rate of isolated clumps accelerated by hot, magnetised winds (Shin et al. 2008; McCourt et al. 2015; Xu & Lazarian 2018), which show that magnetised winds draw spherical clouds out into extended, filamentary structures instead of evaporating them or breaking them into smaller clumps. From this point of view, we over-estimate the fragmentation rate of dense filaments into the hot ICM.

The other possibility is that we underestimate the ability of AGN feedback to destroy dense clumps, for example by under-resolving the mixing layers at the outer clump surface (Gronke & Oh 2018), or simply due to lack of resolution to follow the fragmentation process to smaller scales. This theory is supported by our high-resolution companion simulation, which showed that the fraction of dense gas that survives this particular uplifting event falls from 25 % at a resolution of $\Delta x_{\text{min}} = 120$ pc to 19 %

at a resolution of $\Delta x_{\text{min}} = 30$ pc. The fact that the minimum clump size remains at the resolution limit shows that this process is by no means converged, and higher resolution would likely lead to even smaller clumps and even lower dense gas survival rates. This question has been investigated further by McCourt et al. (2018), who report that for individual clouds accelerated by a hot wind, even a sub-pc scale resolution is insufficient for fragmentation to converge. Based on work by Armillotta et al. (2017), the survival rates for small gas clumps in hot winds is very low, which suggests that we would expect the gas currently contained in our small, compact gas clumps to break into an even
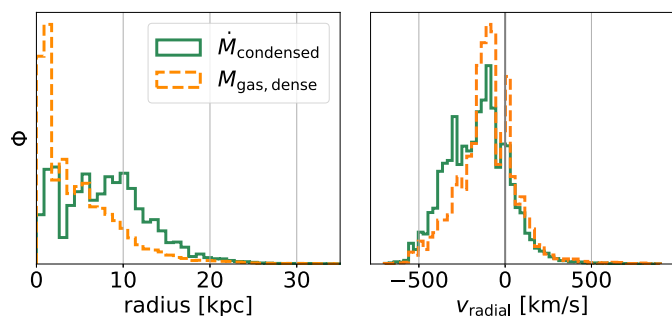
Fig. 16: Probability distribution of clump radius (left) and clump radial velocity (right) weighted by condensation rate and total dense mass respectively.

large number of even smaller clumps until it evaporates entirely and mixes back into the ICM. From this point of view, we are under-estimating the fragmentation rate of small clumps, as well as under-estimating the ability of the AGN to evaporate dense gas.

4.2. The width of filaments

Throughout this paper, we have shown that extended dense gas structures readily form in the cluster center. While our filamentary dense gas structures show maximal extents of 1-10 kpc, in agreement with observations (Conselice et al. 2001), many of our structures appear much wider than the observed 70 pc.

Resolution will play a role in determining the width of the filaments, particularly for very thin filaments which currently have a width close to the resolution limit, such as the long, thing structures seen in the left two panels of Fig. 6. A comparison simulation with higher resolution of $\Delta x_{min} = 30$ pc, run for only a span of 15 Myr, produced thinner filaments than the fiducial simulation at 120 pc. However, many filaments seen in fiducial simulation, such as for example the extended structures in the right hand panel of Fig. 6, are well resolved at the current resolution and therefore not influenced by improvements in resolution.

One process not modelled here, which is thought to play an important role in the morphology of filaments, are magnetic fields and anisotropic thermal conduction along magnetic field lines. In the presence of anisotropic thermal conduction, in combination with magnetic fields, the characteristic thermal collapse length scale (the field length) becomes much larger along field lines than perpendicular to it (Field 1965), as thermal energy is preferentially redistributed along field lines. Collapse therefore preferentially occurs perpendicular to magnetic field lines, smearing spherical collapse out along magnetic field lines. Isolated simulations have shown that in the presence of magnetic fields, local thermal instabilities do indeed produce more extended filamentary gas structures (McCourt et al. 2012; Ji et al. 2018; Xu & Lazarian 2018) compared to more clumpy dense gas for the purely hydrodynamical runs. While this process could help smear dense, round clumps into long, extended filaments, it is unlikely to make the existing filaments thinner. Understanding why the filaments reported here take their particular shapes, and how their morphology might change in the presence of magnetic fields and cosmic rays, will be the subject of future work.

Another limitation of our work is that with many structures shown here at the resolution limit of the simulation, it will be impossible to resolve the detailed internal structure observed for filaments, which consist of dense molecular clumps surrounded

by an H-α envelope (Salomé, P. et al. 2006; Salomé et al. 2011). With a more complex internal structure and gas dynamics, we would expect the energy balances of filaments to change, with as of yet poorly understood consequences for their morphology.

5. Conclusions

In this paper, we have investigated the formation and evolution of dense, dense gas in the center of a Perseus-like cluster under the influence of a spin driven AGN jet, using hydrodynamical simulations.

We showed that:

1. Under the influence of the AGN jet, the cluster undergoes repeated cycles of cooling dominated phases, when dense gas builds up in the cluster center, and heating dominated phases, when the total amount of dense gas decreases. Cycle lengths are on the order of 100 Myr, but show significant variation. (Section 3.1)
2. For low black hole spin values, the chaotic cold accretion onto the cluster center is able to continuously reorient the spin axis, with characteristic reorientation timescales of the order of 10 Myr, allowing the jet to sweep out the full parameter space in both polar and azimuthal angle. (Section 3.2)
3. The morphology of dense gas is highly variable throughout the simulation, with between 20 and 620 individual dense structures present at a given point in time. (Section 3.3)
4. Major axis lengths of individual clumps range from the resolution limit of the simulation up to more than 30 kpc. Larger clumps have more complex, filamentary morphologies than smaller objects, which tend to be rounder and compacter. (Section 3.3)
5. We find evidence for uplifting of existing dense gas by the AGN. During a strong feedback episode, larger, infalling structures fragment into smaller clumps under the influence of the hot outflows driven by the AGN. In the process, they lose up to 75 % of their gas mass and become entrained and ejected from the cluster center. (Section 3.4)
6. A second round of fragmentation into even smaller clumps occurs at the top of the ballistic orbit, before surviving dense clumps fall back onto the cluster center where they re-coalesce into larger objects. (Section 3.4)
7. Condensation takes place preferentially when $t_{cool}/t_{ff} < 20$, which occurs primarily during cooling dominated phases of the cluster, and in the radial range of $5-15$ kpc. Heating dominated phases see more disturbed profiles of t_{cool}/t_{ff} without a clear minimum as the ICM is heating by the AGN feedback. (Section 3.5)
8. Dense gas continues to be observable even during heating-dominated phases, and is preferentially found at smaller radii than condensation, i.e. at $r < 5$ kpc, but can be found as far out as 30 kpc due to uplifting. The presence of dense gas is therefore not a reliable tracer for condensation. (Section 3.5)
9. While 82.1 % of condensation of gas from the hot ICM onto dense clumps occurs on infall, there is also evidence for continued condensation for outflowing gas, with outflowing dense clumps having an average velocity-weighted condensation rate of 0.65 $M_\odot$/yr, compared to 3.09 $M_\odot$/yr for infalling clumps. (Section 3.5)
10. Both direct uplifting of dense gas and condensation of gas from the hot, diffuse to the dense phase in outflowing gas has been invoked to explain the unstructured velocity maps observed in nearby clusters. While we find evidence for both mechanisms, and confirm a general lack of rotation in the

dense gas, we also caution that the observed line of sight velocities fail to show coherent radial flow patterns even when they are present in the dense gas (Section 3.4).

References

Agertz, O., Kravtsov, A. V., Leitner, S. N., & Gnedin, N. Y. 2013, ApJ, 770, 25
Armillotta, L., Fraternali, F., Werk, J. K., Prochaska, J. X., & Marinacci, F. 2017, MNRAS, 470, 114
Beckmann, R. S., Devriendt, J., & Slyz, A. 2019, MNRAS, 483, 3488
Bridges, T. J. & Irwin, J. A. 1998, MNRAS, 300, 967
Brown, S., Emerick, A., Rudnick, L., & Brunetti, G. 2011, ApJ, 740, L28
Cadiou, C., Dubois, Y., & Pichon, C. 2019, A&A, 621, A96
Canning, R. E. A., Fabian, A. C., Johnstone, R. M., et al. 2010, MNRAS, 405, 115
Canning, R. E. A., Ryon, J. E., Gallagher, J. S., et al. 2014, MNRAS, 444, 336
Cattaneo, A. & Teyssier, R. 2007, MNRAS, 376, 1547
Choudhury, P. P. & Sharma, P. 2016, MNRAS, 457, 2554
Cielo, S., Babul, A., Antonuccio-Delogu, V., Silk, J., & Volonteri, M. 2018, A&A, 617, A58
Conselice, C. J., Gallagher, John S., I., & Wyse, R. F. G. 2001, AJ, 122, 2281
Crawford, C. S. & Fabian, A. C. 1992, MNRAS, 259, 265
Dubois, Y., Devriendt, J., Slyz, A., & Teyssier, R. 2010, MNRAS, 409, 985
Dubois, Y., Devriendt, J., Teyssier, R., & Slyz, A. 2011, MNRAS, 417, 1853
Dubois, Y., Volonteri, M., Silk, J., Devriendt, J., & Slyz, A. 2014, MNRAS, 440, 2333
Ebeling, H., Voges, W., Bohringer, H., et al. 1996, MNRAS, 281, 799
Fabian, A. C. 1994, ARA&A, 32, 277
Fabian, A. C. 2012, ARA&A, 50, 455
Fabian, A. C., Johnstone, R. M., Sanders, J. S., et al. 2008, Nature, 454, 968
Fabian, A. C., Sanders, J. S., Crawford, C. S., et al. 2003, MNRAS, 344, L48
Fabian, A. C., Sanders, J. S., Taylor, G. B., et al. 2006, MNRAS, 366, 417
Fabian, A. C., Walker, S. A., Russell, H. R., et al. 2016, MNRAS, 461, 922
Field, G. B. 1965, ApJ, 142, 531
Gaspari, M., Melioli, C., Brighenti, F., & D'Ercole, A. 2011, MNRAS, 411, 349
Gaspari, M., Ruszkowski, M., & Oh, S. P. 2013, MNRAS, 432, 3401
Gaspari, M., Ruszkowski, M., & Sharma, P. 2012, ApJ, 746, 94
Gendron-Marsolais, M., Hlavacek-Larrondo, J., Martin, T. B., et al. 2018, MNRAS, 479, L28
Gnedin, N. Y. & Hollon, N. 2012, ApJS, 202, 13
Gronke, M. & Oh, S. P. 2018, MNRAS, 480, L111
Hamer, S. L., Edge, A. C., Swinbank, A. M., et al. 2014, MNRAS, 437, 862
Hamer, S. L., Edge, A. C., Swinbank, A. M., et al. 2016, MNRAS, 460, 1758
Hatch, N. A., Crawford, C. S., & Fabian, A. C. 2007, MNRAS, 380, 33
Hatch, N. A., Crawford, C. S., Johnstone, R. M., & Fabian, A. C. 2006, MNRAS, 367, 433
Heckman, T. M., Baum, S. A., van Breugel, W. J. M., & McCarthy, P. 1989, ApJ, 338, 48
Hitomi Collaboration. 2016, Nature, 535, 117
Hogan, M. T., McNamara, B. R., Pulido, F. A., et al. 2017, ApJ, 851, 66
Jacob, S. & Pfrommer, C. 2017a, MNRAS, 467, 1449
Jacob, S. & Pfrommer, C. 2017b, MNRAS, 467, 1478
Ji, S., Oh, S. P., & McCourt, M. 2018, MNRAS, 476, 852
Kannan, R., Vogelsberger, M., Pfrommer, C., et al. 2017, ApJ, 837, L18
Kimm, T., Cen, R., Devriendt, J., Dubois, Y., & Slyz, A. 2015, MNRAS, 451, 2900
Klein, R. I., McKee, C. F., & Colella, P. 1994, ApJ, 420, 213
Li, Y. & Bryan, G. L. 2014a, ApJ, 789, 54
Li, Y. & Bryan, G. L. 2014b, ApJ, 789, 153
Li, Y., Ruszkowski, M., & Bryan, G. L. 2017, ApJ, 847, 106
Lim, J., Ohyama, Y., Chi-Hung, Y., Dinh-V-Trung, & Shiang-Yu, W. 2012, ApJ, 744, 112
Lynds, R. 1970, ApJ, 159
Martizzi, D., Quataert, E., Faucher-Giguère, C.-A., & Fielding, D. 2019, MNRAS, 483, 2465
McCourt, M., Oh, S. P., O'Leary, R., & Madigan, A.-M. 2018, MNRAS, 473, 5407
McCourt, M., O'Leary, R. M., Madigan, A.-M., & Quataert, E. 2015, MNRAS, 449, 2
McCourt, M., Sharma, P., Quataert, E., & Parrish, I. J. 2012, MNRAS, 419, 3319

McDonald, M., Gaspari, M., McNamara, B. R., & Tremblay, G. R. 2018, ApJ, 858, 45
McDonald, M., Veilleux, S., & Rupke, D. S. N. 2012, ApJ, 746, 153
McDonald, M., Veilleux, S., Rupke, D. S. N., & Mushotzky, R. 2010, ApJ, 721, 1262
McKinley, B., Tingay, S. J., Carretti, E., et al. 2018, MNRAS, 474, 4056
McKinney, J. C., Tchekhovskoy, A., & Blandford, R. D. 2012, MNRAS, 423, 3083
McNamara, B. & Nulsen, P. 2007, Annual Review of Astronomy and Astrophysics, 45, 117
Mittal, R., Whelan, J. T., & Combes, F. 2015, MNRAS, 450, 2564
Nagai, H., Onishi, K., Kawakatu, N., et al. 2019, arXiv e-prints, arXiv:1905.06017
Narayan, R. & Medvedev, M. V. 2001, ApJ, 562, L129
O'Dea, C. P., Baum, S. A., Privon, G., et al. 2008, ApJ, 681, 1035
Ogiya, G., Biernacki, P., Hahn, O., & Teyssier, R. 2018, arXiv e-prints, arXiv:1802.02177
Olivares, V., Salomé, P., Combes, F., et al. 2019, arXiv e-prints, arXiv:1902.09164
Perret, V., Renaud, F., Epinat, B., et al. 2014, A&A, 562, A1
Peterson, J. R. & Fabian, A. C. 2006, Phys. Rep., 427, 1
Pfister, H., Volonteri, M., Dubois, Y., Dotti, M., & Colpi, M. 2019, MNRAS, 486, 101
Pfrommer, C. 2013, ApJ, 779, 10
Prasad, D., Sharma, P., & Babul, A. 2015, ApJ, 811, 108
Prasad, D., Sharma, P., & Babul, A. 2018, ApJ, 863, 62
Pudritz, R. E., Hardcastle, M. J., & Gabuzda, D. C. 2012, Space Sci. Rev., 169, 27
Pulido, F. A., McNamara, B. R., Edge, A. C., et al. 2018, ApJ, 853, 177
Rafferty, D. A., McNamara, B. R., Nulsen, P. E. J., & Wise, M. W. 2006, ApJ, 652, 216
Reimer, A., Reimer, O., Schlickeiser, R., & Iyudin, A. 2004, A&A, 424, 773
Revaz, Y., Combes, F., & Salomé, P. 2008, A&A, 477, L33
Rosen, A. & Bregman, J. N. 1995, ApJ, 440, 634
Russell, H. R., McDonald, M., McNamara, B. R., et al. 2017, ApJ, 836, 130
Ruszkowski, M. & Oh, S. P. 2010, ApJ, 713, 1332
Ruszkowski, M., Yang, H.-Y. K., & Reynolds, C. S. 2017, ApJ, 844, 13
Salomé, P., Combes, F., Revaz, Y., et al. 2011, A&A, 531, A85
Salomé, Q., Salomé, P., Combes, F., Hamer, S., & Heywood, I. 2016, A&A, 586, A45
Salomé, P., Combes, F., Edge, A. C., et al. 2006, A&A, 454, 437
Salomé, P., Combes, F., Revaz, Y., et al. 2008, A&A, 484, 317
Schmidt, R. W., Fabian, A. C., & Sanders, J. S. 2002, MNRAS, 337, 71
Segers, M. C., Oppenheimer, B. D., Schaye, J., & Richings, A. J. 2017, MNRAS, 471, 1026
Sharma, P., McCourt, M., Quataert, E., & Parrish, I. J. 2012, MNRAS, 420, 3174
Shi, Y., Helou, G., Yan, L., et al. 2011, ApJ, 733, 87
Shin, M.-S., Stone, J. M., & Snyder, G. F. 2008, ApJ, 680, 336
Sparre, M., Pfrommer, C., & Vogelsberger, M. 2019, MNRAS, 482, 5401
Sutherland, R. S. & Dopita, M. A. 1993, ApJS, 88, 253
Teyssier, R. 2002, A&A, 385, 337
Tremblay, G. R., Combes, F., Oonk, J. B. R., et al. 2018, ApJ, 865, 13
Turk, M. J., Smith, B. D., Oishi, J. S., et al. 2011, The Astrophysical Journal Supplement Series, 192, 9
Vantyghem, A. N., McNamara, B. R., Edge, A. C., et al. 2017, ApJ, 848, 101
Vantyghem, A. N., McNamara, B. R., Russell, H. R., et al. 2018, ApJ, 863, 193
Voit, G. M. & Donahue, M. 2015, ApJ, 799, L1
Voit, G. M., Meece, G., Li, Y., et al. 2017, ApJ, 845, 80
Wang, C., Li, Y., & Ruszkowski, M. 2019, MNRAS, 482, 3576
Werner, N., Urban, O., Simionescu, A., & Allen, S. W. 2013, Nature, 502, 656
Xu, S. & Lazarian, A. 2018, arXiv e-prints, arXiv:1802.00987
Yang, H.-Y. K. & Reynolds, C. S. 2016a, ApJ, 829, 90
Yang, H. Y. K. & Reynolds, C. S. 2016b, ApJ, 818, 181
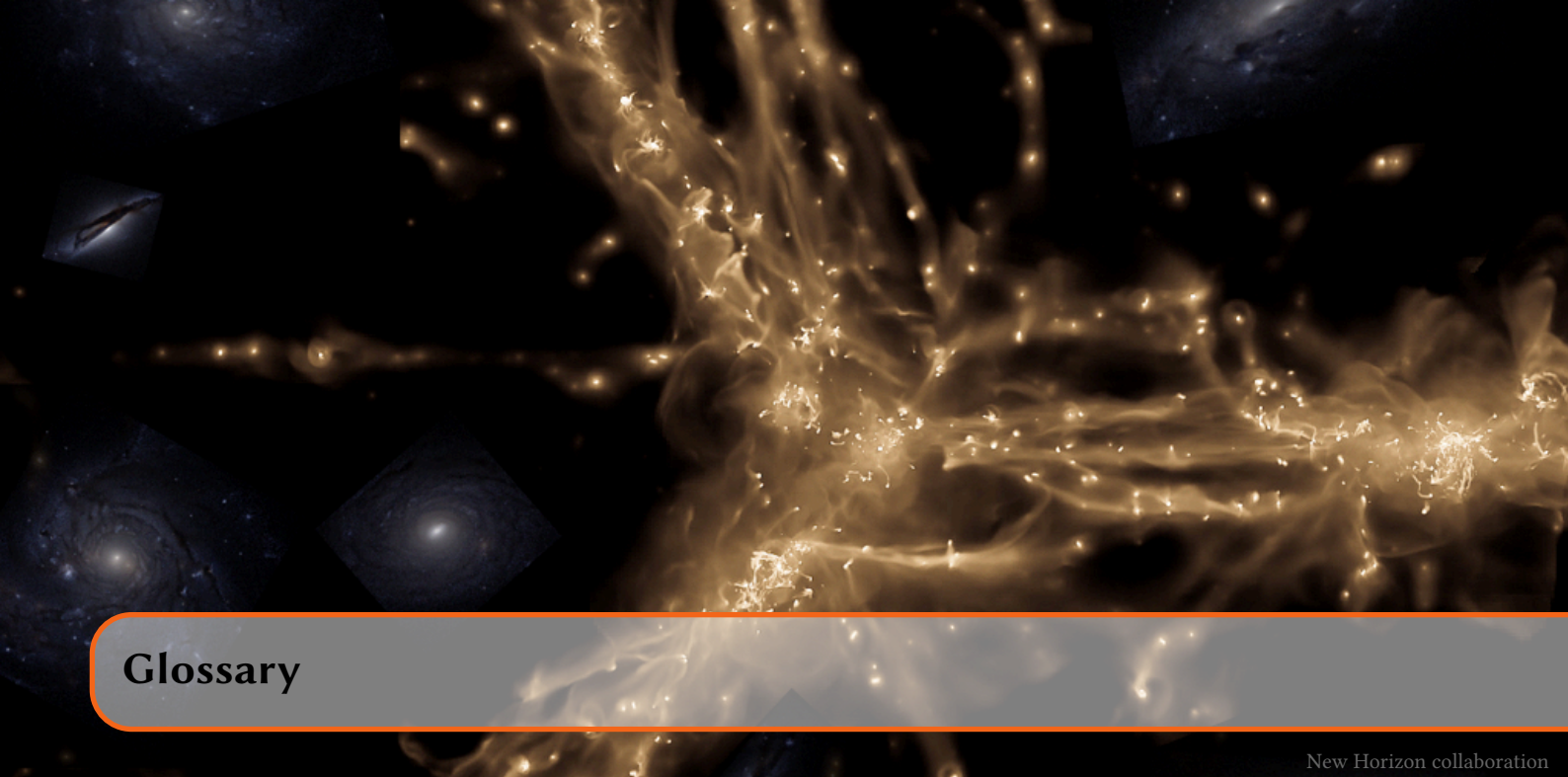
Glossary

Glossary

AGN Active Galactic Nuclei. 40, 46, 51, 176, 182, 185, 245

AM angular momentum. 152, 161, 162, 164, 166, 167, 169, 172, 173, 175–178, 180

AMR Adaptive Mesh Refinement. 40, 42, 133, 134, 152–154, 157, 162

CMB Cosmic Microwave Background. 3, 4, 9–11, 34

DM Dark Matter. 3, 10, 16, 32, 33, 40, 45, 47, 49, 54, 152, 157, 160, 172–175, 177, 190, 216

EdS Einstein de-Sitter. 17–19, 29

EoS Equation of State. 40, 43

FFT Fast Fourier Transform. 14

GAMA Galaxy And Mass Assembly. 189

IMF Initial Mass Function. 37

ΛCDM Λ Cold Dark Matter. 3, 4, 10, 14, 16, 17, 19, 23–25, 31–33, 49, 92, 102, 107, 108, 110, 160

LSS Last Scattering Surface. 10

PDF Probability Distribution Function. 11, 89–91, 95, 101, 102, 105–108, 111, 115, 117, 118, 120, 122, 123, 130, 131

PM Particle Mesh. 45

sAM specific angular momentum. 152, 153, 162, 163, 166, 167, 169, 170, 172, 174, 177, 178

SMBH Supermassive Black Hole. 32, 40, 46, 134, 135, 153, 157, 185, 186

SN Supernova. 42, 46

SPH Smooth Particle Hydrodynamics. 40, 133, 135, 161, 162

TTT Tidal Torque Theory. 5, 6, 28, 29, 54, 161, 167, 173

Bibliography

Abazajian, K. et al. (2003). "The First Data Release of the Sloan Digital Sky Survey". In: *Astron. J.* 126.4, p. 2081. DOI: 10/bj64qs (cit. on pp. 4–5).

Abbott, T. M. C. et al. (2019). "First Cosmology Results Using Type Ia Supernovae from the Dark Energy Survey: Constraints on Cosmological Parameters". In: *Astrophys. J.* 872.2, p. L30. DOI: 10/gf42qm (cit. on p. 4).

Adler, R. J. and J. E. Taylor (2007). *Random Fields and Geometry*. Vol. 115. DOI: 10.1007/978-0-387-48116-6 (cit. on p. 11).

Alam, S. et al. (2017). "The Clustering of Galaxies in the Completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological Analysis of the DR12 Galaxy Sample". In: *Monthly Notices of the Royal Astronomical Society* 470, pp. 2617–2652. DOI: 10/gbrxk2 (cit. on p. 4).

Alpaslan, M. et al. (2016). "Galaxy And Mass Assembly (GAMA): stellar mass growth of spiral galaxies in the cosmic web". In: *MNRAS* 457, pp. 2287–2300. DOI: 10.1093/mnras/stw134 (cit. on p. 6).

Alpaslan, M. et al. (2015). "Galaxy And Mass Assembly (GAMA): Trends in Galaxy Colours, Morphology, and Stellar Populations with Large-Scale Structure, Group, and Pair Environments". In: *Mon. Not. R. Astron. Soc.* 451.3, p. 3249. DOI: 10/f7q72s (cit. on p. 6).

Andrae, R. and K. Jahnke (2011). "Only Marginal Alignment of Disc Galaxies". In: *Mon. Not. R. Astron. Soc.* 418.3, p. 2014. DOI: 10/bj7n7w (cit. on p. 6).

Ayliffe, B. A., G. Laibe, D. J. Price, and M. R. Bate (2012). "On the Accumulation of Planetesimals near Disc Gaps Created by Protoplanets". In: *Mon. Not. R. Astron. Soc.* 423.2, p. 1450. DOI: 10/gf49mz (cit. on p. 183).

Bardeen, J. M., J. R. Bond, N. Kaiser, and A. S. Szalay (1986). "The Statistics of Peaks of Gaussian Random Fields". In: *Astrophys. J.* 304, pp. 15–15. DOI: 10/cxd2pr (cit. on pp. 26, 30, 130).

Beckmann, R. S. et al. (submitted). "Dense Gas Formation and Destruction in a Simulated Perseus-like Galaxy Cluster with Spin-Driven Black Hole Feedback". In: (cit. on pp. 8, 183, 189).

Bennett, C. L. et al. (2013). "Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results". In: *Astrophys. J. Suppl. Ser.* 208.2, p. 20. DOI: 10/gf42qr (cit. on p. 4).

Benson, A. J. (2010). "Galaxy Formation Theory". In: *Phys. Rep.* 495.2-3, pp. 33–86. DOI: `10/cdsh6x` (cit. on pp. 33, 184).

Berlok, T. and C. Pfrommer (2019). "The Impact of Magnetic Fields on Cold Streams Feeding Galaxies". In: *ArXiv E-Prints*, arXiv:1904.02167 (cit. on p. 176).

Bernardeau, F., S. Colombi, E. Gaztañaga, and R. Scoccimarro (2002). "Large-Scale Structure of the Universe and Cosmological Perturbation Theory". In: *Phys. Rep.* 367.1-3, p. 1. DOI: `10/c8ffxs` (cit. on pp. 102, 115).

Bertschinger, E. (1985). "Self-Similar Secondary Infall and Accretion in an Einstein-de Sitter Universe". In: *Astrophys. J. Suppl. Ser.* 58, p. 39. DOI: `10/d39c75` (cit. on p. 33).

Beygu, B. et al. (2016). "The Void Galaxy Survey: Star Formation Properties". In: *Mon. Not. R. Astron. Soc.* 458.1, p. 394. DOI: `10/f8nxbd` (cit. on pp. 5, 54).

Binney, J. (1977). "The Physics of Dissipational Galaxy Formation". In: *Astrophys. J.* 215, pp. 483–483. DOI: `10/cqsjwf` (cit. on p. 33).

Birnboim, Y. and A. Dekel (2003). "Virial Shocks in Galactic Haloes?" In: *Mon. Not. R. Astron. Soc.* 345.1, pp. 349–364. DOI: `10/bc52vf` (cit. on pp. 6, 33, 161, 174).

Bond, J. R., S. Cole, G. Efstathiou, and N. Kaiser (1991). "Excursion Set Mass Functions for Hierarchical Gaussian Fluctuations". In: *Astrophys. J.* 379, pp. 440–440. DOI: `10/dvk8nx` (cit. on pp. 21, 50, 181).

Bond, J. R. and S. T. Myers (1996). "The Peak-Patch Picture of Cosmic Catalogs. I. Algorithms". In: *Astrophys. J. Suppl. Ser.* 103, p. 1. DOI: `10/dmskgh` (cit. on pp. 22, 26, 53–54, 86, 90, 183).

Bond, J. R., L. Kofman, and D. Pogosyan (1996). "How Filaments of Galaxies Are Woven into the Cosmic Web". In: *Nature* 380.6575, p. 603. DOI: `10/fq7bzc` (cit. on pp. 50, 53).

Borzyszkowski, M., C. Porciani, E. Romano-Díaz, and E. Garaldi (2017). "ZOMG – I. How the Cosmic Web Inhibits Halo Growth and Generates Assembly Bias". In: *Mon. Not. R. Astron. Soc.* 469.1, pp. 594–611. DOI: `10/gbjbcb` (cit. on pp. 54, 106, 112, 183).

Bryan, G. L. et al. (2014). "ENZO: An Adaptive Mesh Refinement Code for Astrophysics". In: *The Astrophysical Journal Supplement Series* 211, p. 19. DOI: `10/gfx7v2` (cit. on pp. 42, 133).

Bullock, J. S. et al. (2016). "High Angular Momentum Halo Gas: A Feedback and Code–Independent Prediction of Lcdm ~". In: (cit. on p. 51).

Cadiou, C., Y. Dubois, and C. Pichon (2019). "Accurate Tracer Particles of Baryon Dynamics in the Adaptive Mesh Refinement Code Ramses". In: *A&A* 621, A96. DOI: `10/gf42rj` (cit. on pp. 8, 134–135, 162, 164, 182, 245).

Cai, Y.-C., N. Padilla, and B. Li (2015). "Testing gravity using cosmic voids". In: *MNRAS* 451, pp. 1036–1055. DOI: `10.1093/mnras/stv777` (cit. on p. 119).

Carroll, S. M., W. H. Press, and E. L. Turner (1992). "The Cosmological Constant." In: *Annu. Rev. Astron. Astrophys.* 30, p. 499. DOI: `10/fsf4wb` (cit. on pp. 17–18).

Castorina, E., A. Paranjape, O. Hahn, and R. K. Sheth (2016). "Excursion Set Peaks: The Role of Shear". In: (cit. on pp. 54, 183).

Catelan, P. and T. Theuns (1996). "Evolution of the Angular Momentum of Protogalaxies from Tidal Torques: Zel'dovich Approximation". In: *Mon. Not. R. Astron. Soc.* 282.2, p. 436. DOI: `10/gf42sf` (cit. on pp. 28, 167).

Cecil, G., J. Bland-Hawthorn, S. Veilleux, and A. V. Filippenko (2001). "Jet- and Wind-Driven Ionized Outflows in the Superbubble and Star-Forming Disk of NGC 3079". In: *ApJ* 555.1, pp. 338–355. DOI: `10/cdr9vw` (cit. on p. 37).

Cervantes-Sodi, B., X. Hernandez, and C. Park (2010). "Clues on the Origin of Galactic Angular Momentum from Looking at Galaxy Pairs". In: *Mon. Not. R. Astron. Soc.* 402.3, p. 1807. DOI: `10/fk7stc` (cit. on p. 6).

Chabrier, G. (2003). "Galactic Stellar and Substellar Initial Mass Function". In: *Publ. Astron. Soc. Pac.* 115.809, p. 763. DOI: `10/dmz3xv` (cit. on p. 37).

Chisari, N. E. et al. (2017). "Galaxy-Halo Alignments in the Horizon-AGN Cosmological Hydro-dynamical Simulation". In: *Mon. Not. R. Astron. Soc.* 472.1, p. 1163. DOI: 10/gcjfxw (cit. on pp. 6, 54, 161).

Codis, S. et al. (2012). "Connecting the Cosmic Web to the Spin of Dark Haloes: Implications for Galaxy Formation". In: *Mon. Not. R. Astron. Soc.* 427.4, p. 3320. DOI: 10/f4mzb4 (cit. on pp. 29, 54, 173).

Codis, S., C. Pichon, D. Pogosyan, F. Bernardeau, and T. Matsubara (2013). "Non-Gaussian Minkowski functionals & extrema counts in redshift space". In: *arXiv.org* 1, pp. 531–564 (cit. on pp. 102, 119).

Codis, S., C. Pichon, and D. Pogosyan (2015). "Spin Alignments within the Cosmic Web: A Theory of Constrained Tidal Torques near Filaments". In: *Mon. Not. R. Astron. Soc.* 452.4, pp. 3369–3393. DOI: 10/f7q92x (cit. on pp. 6, 50, 54, 114, 161).

Codis, S., D. Pogosyan, and C. Pichon (2018). "On the Connectivity of the Cosmic Web: Theory and Implications for Cosmology and Galaxy Formation". In: *Mon. Not. R. Astron. Soc.* 479.1, p. 973. DOI: 10/gf42tw (cit. on pp. 50, 107, 111, 123–124).

Colella, P. (1985). "A Direct Eulerian MUSCL Scheme for Gas Dynamics". In: *SIAM J. Sci. and Stat. Comput.* 6.1, pp. 104–117. DOI: 10/brwr97 (cit. on p. 45).

Colless, M., G. Dalton, S. Maddox, and et al. (2001). "The 2dF Galaxy Redshift Survey: spectra and redshifts". In: *MNRAS* 328, pp. 1039–1063. DOI: 10.1046/j.1365-8711.2001.04902.x (cit. on p. 5).

Cooray, A. and R. Sheth (2002). "Halo Models of Large Scale Structure". In: *Phys. Rep.* 372.1, p. 1. DOI: 10/fnqvfc (cit. on p. 6).

Corasaniti, P. S. and I. Achitouv (2011). "Excursion Set Halo Mass Function and Bias in a Stochastic Barrier Model of Ellipsoidal Collapse". In: *Phys. Rev. D* 84.2, p. 023009. DOI: 10/bhw379 (cit. on p. 23).

Cornuault, N., M. D. Lehnert, F. Boulanger, and P. Guillard (2018). "Are Cosmological Gas Accretion Streams Multiphase and Turbulent?" In: *Astron. Astrophys.* 610, A75. DOI: 10/gf42qp (cit. on pp. 161, 176).

Crittenden, R. G., P. Natarajan, U.-L. Pen, and T. Theuns (2001). "Spin-induced Galaxy Alignments and Their Implications for Weak-Lensing Measurements". In: *ApJ* 559, pp. 552–571. DOI: 10.1086/322370 (cit. on p. 28).

Dalal, N., M. White, J. R. Bond, and A. Shirokov (2008). "Halo Assembly Bias in Hierarchical Structure Formation". In: *Astrophys. J.* 687.1, pp. 12–21. DOI: 10/bzr6m2 (cit. on pp. 54, 181).

Dalla Vecchia, C. and J. Schaye (2012). "Simulating Galactic Outflows with Thermal Supernova Feedback". In: *Mon. Not. R. Astron. Soc.* 426.1, p. 140. DOI: 10/gf42q2 (cit. on p. 38).

Danovich, M., A. Dekel, O. Hahn, and R. Teyssier (2012). "Coplanar Streams, Pancakes and Angular-Momentum Exchange in High-z Disc Galaxies". In: *Mon. Not. R. Astron. Soc.* 422.2, pp. 1732–1749. DOI: 10/f3zpdk (cit. on pp. 6, 166, 175).

Danovich, M., A. Dekel, O. Hahn, D. Ceverino, and J. Primack (2015). "Four Phases of Angular-Momentum Buildup in High-z Galaxies: From Cosmic-Web Streams through an Extended Ring to Disc and Bulge". In: *Mon. Not. R. Astron. Soc.* 449.2, pp. 2087–2111. DOI: 10/gf42tp (cit. on pp. 6, 51, 152, 161–162, 166–167, 174).

de Lapparent, V., M. J. Geller, and J. P. Huchra (1986). "A slice of the universe". In: *ApJ* 302, pp. L1–L5. DOI: 10.1086/184625 (cit. on p. 5).

Dekel, A. et al. (2009). "Cold Streams in Early Massive Hot Haloes as the Main Mode of Galaxy Formation". In: *Nature* 457.7228, pp. 451–454. DOI: 10/bh6f6c (cit. on p. 6).

Dekel, A. and Y. Birnboim (2006). "Galaxy Bimodality Due to Cold Flows and Shock Heating". In: *Mon. Not. R. Astron. Soc.* 368.1, pp. 2–20. DOI: 10/c38j76 (cit. on pp. 161, 174).

Del Popolo, A., E. N. Ercan, and M. Gambera (2001). "The Effects of Shear and Rotation Anisotropy upon the Process of Gravitational Instability". In: *Baltic Astronomy* 10, pp. 629–649 (cit. on p. 54).

Di Matteo, T. et al. (2012). "Cold Flows and the First Quasars". In: *Astrophys. J.* 745.2, p. L29. DOI: 10/fzhdj5 (cit. on p. 161).

Diemer, B. (2018). "COLOSSUS: A Python Toolkit for Cosmology, Large-Scale Structure, and Dark Matter Halos". In: *Astrophys. J. Suppl. Ser.* 239.2, p. 35. DOI: 10/gf4n4x (cit. on p. 110).

Doroshkevich, A. G. (1973). "Spatial Structure of Perturbations and Origin of Galactic Rotation in Fluctuation Theory". In: *Astrophysics* 6.4, pp. 320–330. DOI: 10/ctbczf (cit. on p. 5).

— (1970). "The Space Structure of Perturbations and the Origin of Rotation of Galaxies in the Theory of Fluctuation." In: *Astrofizika* 6, pp. 581–600 (cit. on p. 28).

Driver, S. P. et al. (2011). "Galaxy and Mass Assembly (GAMA): Survey Diagnostics and Core Data Release". In: *Mon. Not. R. Astron. Soc.* 413.2, p. 971. DOI: 10/c8cqst (cit. on pp. 6, 86).

Dubois, Y. and R. Teyssier (2008). "On the Onset of Galactic Winds in Quiescent Star Forming Galaxies". In: *Astron. Astrophys.* 477.1, p. 79. DOI: 10/fpbtqw (cit. on p. 38).

Dubois, Y. et al. (2012). "Feeding Compact Bulges and Supermassive Black Holes with Low Angular Momentum Cosmic Gas at High Redshift". In: *Mon. Not. R. Astron. Soc.* 423.4, pp. 3616–3630. DOI: 10/f34xzg (cit. on pp. 134, 161, 164).

Dubois, Y. et al. (2013). "Blowing Cold Flows Away: The Impact of Early AGN Activity on the Formation of a Brightest Cluster Galaxy Progenitor". In: *Mon. Not. R. Astron. Soc.* 428.4, p. 2885. DOI: 10/f4wftr (cit. on pp. 161–162, 176, 186).

Dubois, Y. et al. (2014). "Dancing in the Dark: Galactic Properties Trace Spin Swings along the Cosmic Web". In: *Mon. Not. R. Astron. Soc.* 444.2, pp. 1453–1468. DOI: 10/f6kn7x (cit. on pp. 6, 37, 54, 161, 164).

Dubois, Y. et al. (2016). "The HORIZON-AGN Simulation: Morphological Diversity of Galaxies Promoted by AGN Feedback". In: *Monthly Notices of the Royal Astronomical Society* 463, pp. 3948–3964. DOI: 10/f9qzzx (cit. on pp. 6, 37, 48, 50).

Eardley, E. et al. (2015). "Galaxy and Mass Assembly (GAMA): The Galaxy Luminosity Function within the Cosmic Web". In: *Mon. Not. R. Astron. Soc.* 448.4, pp. 3665–3678. DOI: 10/f662n5 (cit. on p. 6).

Efstathiou, G., C. S. Frenk, S. D. M. White, and M. Davis (1988). "Gravitational clustering from scale-free initial conditions". In: *MNRAS* 235, pp. 715–748. DOI: 10.1093/mnras/235.3.715 (cit. on p. 53).

Eisenstein, D. J. and W. Hu (1999). "Power Spectra for Cold Dark Matter and Its Variants". In: *Astrophys. J.* 511, pp. 5–5. DOI: 10/b8d8fq (cit. on p. 106).

Eisenstein, D. J. et al. (2005). "Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies". In: *Astrophys. J.* 633.2, p. 560. DOI: 10/bk82m2 (cit. on p. 4).

Federrath, C. and R. S. Klessen (2012). "The Star Formation Rate of Turbulent Magnetized Clouds: Comparing Theory, Simulations, and Observations". In: *Astrophys. J.* 761.2, p. 156. DOI: 10/gf4pv2 (cit. on p. 37).

Gao, L., V. Springel, and S. D. M. White (2005). "The age dependence of halo clustering". In: *MNRAS* 363, pp. L66–L70. DOI: 10.1111/j.1745-3933.2005.00084.x (cit. on pp. 5, 53, 181).

Gao, L. and S. D. M. White (2007). "Assembly Bias in the Clustering of Dark Matter Haloes". In: *Mon. Not. R. Astron. Soc.* 377.1, p. L5. DOI: 10/fkgvs6 (cit. on pp. 5, 53).

Garaldi, E., E. Romano-Díaz, M. Borzyszkowski, and C. Porciani (2018). "ZOMG - III. The Effect of Halo Assembly on the Satellite Population". In: *Mon. Not. R. Astron. Soc.* 473.2, p. 2234. DOI: 10/gcxzqw (cit. on p. 183).

Gay, C., C. Pichon, and D. Pogosyan (2012). "Non-Gaussian statistics of critical sets in 2D and 3D: Peaks, voids, saddles, genus, and skeleton". In: *Phys. Rev. D* 85.2, 023011, p. 023011. DOI: 10.1103/PhysRevD.85.023011 (cit. on pp. 102, 119).

Gay, C. (2011). "Le Squelette Des Grandes Structures de l'Univers" (cit. on p. 107).

Geen, S., J. Rosdahl, J. Blaizot, J. Devriendt, and A. Slyz (2015). "A Detailed Study of Feedback from a Massive Star". In: *Mon. Not. R. Astron. Soc.* 448.4, p. 3248. DOI: 10/f66vkt (cit. on p. 164).

Geller, M. J. and J. P. Huchra (1989). "Mapping the universe". In: *Science* 246, pp. 897–903. DOI: 10.1126/science.246.4932.897 (cit. on p. 5).

Genel, S., M. Vogelsberger, D. Nelson, D. Sijacki, V. Springel, and L. Hernquist (2013). "Following the Flow: Tracer Particles in Astrophysical Fluid Simulations". In: *Mon. Not. R. Astron. Soc.* 435.2, pp. 1426–1442. DOI: 10/f5cxts (cit. on pp. 134, 162).

Gerritsen, J. P. E. (1997). "Star Formation and the Interstellar Medium in Galaxy Simulations" (cit. on p. 38).

Gnedin, N. Y., S. G. O. Glover, R. S. Klessen, and V. Springel (2015). "Star Formation in Galaxy Evolution: Connecting Numerical Models to Reality". In: *Star Formation in Galaxy Evolution: Connecting Numerical Models to Reality. Edited by N.Y. Gnedin et al. Springer, ISBN 978-3-662-47889-9* (cit. on p. 36).

Godunov, S. K. (1959). "A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics." In: *Mat. Sb., Nov. Ser.* 47, pp. 271–306 (cit. on p. 44).

Goh, T. et al. (2019). "Dark Matter Halo Properties versus Local Density and Cosmic Web Location". In: *Mon. Not. R. Astron. Soc.* 483.2, p. 2101. DOI: 10/gf5dt3 (cit. on p. 6).

Grassi, T. et al. (2014). "KROME - a Package to Embed Chemistry in Astrophysical Simulations". In: *Mon. Not. R. Astron. Soc.* 439.3, p. 2386. DOI: 10/f52mn8 (cit. on p. 35).

Guillet, T., D. Chapon, and M. Labadens (2013). "PyMSES: Python Modules for RAMSES". In: *Astrophysics Source Code Library*, ascl:1310.002 (cit. on p. 153).

Guo, Q., E. Tempel, and N. I. Libeskind (2015). "Galaxies in Filaments Have More Satellites: The Influence of the Cosmic Web on the Satellite Luminosity Function in the SDSS". In: *Astrophys. J.* 800.2, p. 112. DOI: 10/gf5dtz (cit. on p. 6).

Guzzo, L. et al. (2014). "The VIMOS Public Extragalactic Redshift Survey (VIPERS). An Unprecedented View of Galaxies and Large-Scale Structure at 0.5 < z < 1.2". In: *Astron. Astrophys.* 566, A108. DOI: 10/gf42qj (cit. on p. 6).

Haardt, F. and P. Madau (1996). "Radiative Transfer in a Clumpy Universe. II. The Ultraviolet Extragalactic Background". In: *Astrophys. J.* 461, p. 20. DOI: 10/crwj6k (cit. on pp. 40, 164).

Hahn, O., C. Porciani, C. M. Carollo, and A. Dekel (2007). "Properties of Dark Matter Haloes in Clusters, Filaments, Sheets and Voids". In: *Mon Not R Astron Soc* 375.2, pp. 489–499. DOI: 10/dxbdjj (cit. on pp. 5, 53).

Hahn, O., C. Porciani, A. Dekel, and C. M. Carollo (2009). "Tidal Effects and the Environment Dependence of Halo Assembly". In: *Mon. Not. R. Astron. Soc.* 398.4, pp. 1742–1756. DOI: 10/ckvjv9 (cit. on pp. 49, 54, 90, 106, 183).

Hahn, O. and T. Abel (2011). "Multi-Scale Initial Conditions for Cosmological Simulations". In: *Mon R Astron Soc* 415, pp. 2101–2121. DOI: 10/fhtgjf (cit. on p. 14).

Hahn, O. and A. Paranjape (2014). "The Locations of Halo Formation and the Peaks Formalism". In: *Mon. Not. R. Astron. Soc.* 438.1, pp. 878–899. DOI: 10/f5qxfv (cit. on p. 183).

Hanami, H. (2001). "Statistics of Merging Peaks of Random Gaussian Fluctuations: Skeleton Tree Formalism". In: *MNRAS* 327, pp. 721–738. DOI: 10/fwfn5t (cit. on pp. 50, 90, 92, 94, 98, 130).

Harten, A., P. Lax, and B. Leer (1983). "On Upstream Differencing and Godunov-Type Schemes for Hyperbolic Conservation Laws". In: *SIAM Rev.* 25.1, pp. 35–61. DOI: 10/cbvbks (cit. on p. 44).

Hirata, C. M. and U. Seljak (2004). "Intrinsic Alignment-Lensing Interference as a Contaminant of Cosmic Shear". In: *Phys. Rev. D* 70.6, p. 063526. DOI: 10/fkjzhr (cit. on p. 6).

Hopkins, P. F. et al. (2014). "Galaxies on FIRE (Feedback In Realistic Environments): Stellar Feedback Explains Cosmologically Inefficient Star Formation". In: *Mon. Not. R. Astron. Soc.* 445.1, p. 581. DOI: 10/f6nhn5 (cit. on pp. 37, 48).

Hopkins, P. F. (2015). "A New Class of Accurate, Mesh-Free Hydrodynamic Simulation Methods". In: *Mon. Not. R. Astron. Soc.* 450.1, p. 53. DOI: 10/gf42rb (cit. on p. 40).

Hoyle, F. (1949). *Problems of Cosmical Aerodynamics, Central Air Documents, Office, Dayton, OH.* Dayton, OH: Central Air Documents Office, p. 195 (cit. on pp. 28, 167).

Huang, K.-W., Y. Feng, and T. Di Matteo (2019). "The Early Growth of Supermassive Black Holes in Cosmological Hydrodynamic Simulations with Constrained Gaussian Realizations". In: *ArXiv E-Prints*, arXiv:1906.00242 (cit. on p. 185).

Hubble, E. (1929). "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae". In: *Proc. Natl. Acad. Sci.* 15.3, p. 168. DOI: 10/fwtf28 (cit. on pp. 2–3).

Huchra, J. P. et al. (2012). "The 2MASS Redshift Survey—Description and Data Release". In: *Astrophys. J. Suppl. Ser.* 199.2, p. 26. DOI: 10/gf42qk (cit. on p. 4).

Hwang, H. S. et al. (2016). "HectoMAP and Horizon Run 4: Dense Structures and Voids in the Real and Simulated Universe". In: *Astrophys. J.* 818.2, p. 173. DOI: 10/gf4qzw (cit. on p. 4).

Jennings, E., Y. Li, and W. Hu (2013). "The Abundance of Voids and the Excursion Set Formalism". In: *Mon. Not. R. Astron. Soc.* 434.3, p. 2167. DOI: 10/f48w3f (cit. on pp. 92, 110).

Jõeveer, M., J. Einasto, and E. Tago (1978). "Spatial Distribution of Galaxies and of Clusters of Galaxies in the Southern Galactic Hemisphere". In: *Mon. Not. R. Astron. Soc.* 185, p. 357. DOI: 10/gf5crk (cit. on p. 53).

Jones, B. J. T., R. van de Weygaert, and M. A. Aragón-Calvo (2010). "Fossil Evidence for Spin Alignment of Sloan Digital Sky Survey Galaxies in Filaments". In: *Mon. Not. R. Astron. Soc.* 408.2, p. 897. DOI: 10.1111/j.1365-2966.2010.17202.x (cit. on p. 6).

Jones, E., T. Oliphant, and P. Peterson (2001). *SciPy: Open Source Scientific Tools for Python* (cit. on p. 125).

Kac, M. (1943). "On the Average Number of Real Roots of a Random Algebraic Equation". In: *Bull. Amer. Math. Soc.* 49.4, pp. 314–320. DOI: 10/dxrg59 (cit. on p. 26).

Kaiser, N. (1984a). "On the Spatial Correlations of Abell Clusters". In: *Astrophys. J.* 284, pp. L9–L9. DOI: 10/fmnhhv (cit. on p. 53).

— (1984b). "On the spatial correlations of Abell clusters". In: *ApJ* 284, pp. L9–L12. DOI: 10.1086/184341 (cit. on p. 105).

Katz, N. (1992). "Dissipational Galaxy Formation. II. Effects of Star Formation". In: *Astrophys. J.* 391, p. 502. DOI: 10/c85cgw (cit. on p. 37).

Kauffmann, G., S. D. M. White, and B. Guiderdoni (1993). "The Formation and Evolution of Galaxies within Merging Dark Matter Haloes." In: *Mon. Not. R. Astron. Soc.* 264, p. 201. DOI: 10/gf42qh (cit. on p. 5).

Kennicutt, R. C. (1998). "The Global Schmidt Law in Star-Forming Galaxies". In: *Astrophys. J.* 498.2, p. 541. DOI: 10/dcm2qp (cit. on p. 37).

Kereš, D., N. Katz, D. H. Weinberg, and R. Dave (2005). "How Do Galaxies Get Their Gas?" In: *Mon. Not. R. Astron. Soc.* 363.1, pp. 2–28. DOI: 10/d3685q (cit. on pp. 33, 135, 161, 164).

Kereš, D., N. Katz, M. Fardal, R. Davé, and D. H. Weinberg (2009). "Galaxies in a Simulated ΛCDM Universe - I. Cold Mode and Hot Cores". In: *Mon. Not. R. Astron. Soc.* 395.1, p. 160. DOI: 10/ffh89x (cit. on pp. 6, 33).

Kerscher, M., I. Szapudi, and A. S. Szalay (2000). "A Comparison of Estimators for the Two-Point Correlation Function". In: *Astrophys. J.* 535.1, p. L13. DOI: 10/bmxrhk (cit. on p. 128).

Khandai, N. et al. (2015). "The MassiveBlack-II Simulation: The Evolution of Halos and Galaxies to Z~0". In: *Mon. Not. R. Astron. Soc.* 450.2, pp. 1349–1374. DOI: 10/f7gs9h (cit. on pp. 6, 48).

Kim, J.-h. et al. (2013). "The AGORA High-Resolution Galaxy Simulations Comparison Project". In: *ApJS* 210.1, p. 14. DOI: 10.1088/0067-0049/210/1/14 (cit. on p. 48).

Kim, J.-h. et al. (2016). "The AGORA High-Resolution Galaxy Simulations Comparison Project. II. Isolated Disk Test". In: *ApJ* 833.2, p. 202. DOI: 10/f9r5cx (cit. on pp. 48, 50).

Kimm, T., J. Devriendt, A. Slyz, C. Pichon, S. A. Kassin, and Y. Dubois (2011). "The Angular Momentum of Baryons and Dark Matter Halos Revisited". In: *ArXiv E-Prints*, arXiv:1106.0538 (cit. on pp. 51, 161, 174).

Kimm, T. and R. Cen (2014). "Escape Fraction of Ionizing Photons during Reionization: Effects Due to Supernova Feedback and Runaway OB Stars". In: *Astrophys. J.* 788.2, p. 121. DOI: 10/gfsvmc (cit. on p. 38).

Kimm, T., R. Cen, J. Devriendt, Y. Dubois, and A. Slyz (2015). "Towards Simulating Star Formation in Turbulent High-z Galaxies with Mechanical Supernova Feedback". In: *Mon. Not. R. Astron. Soc.* 451.3, pp. 2900–2921. DOI: 10/f7qwz8 (cit. on pp. 38, 164).

Kimm, T., H. Katz, M. Haehnelt, J. Rosdahl, J. Devriendt, and A. Slyz (2017). "Feedback-Regulated Star Formation and Escape of LyC Photons from Mini-Haloes during Reionization". In: *Mon. Not. R. Astron. Soc.* 466.4, p. 4826. DOI: 10/gf42r3 (cit. on pp. 37–38, 164).

Kleiner, D., K. A. Pimbblet, D. H. Jones, B. S. Koribalski, and P. Serra (2017). "Evidence for H I Replenishment in Massive Galaxies through Gas Accretion from the Cosmic Web". In: *Mon. Not. R. Astron. Soc.* 466.4, p. 4692. DOI: 10/gf5dtk (cit. on pp. 5, 53).

Kofman, L., D. Pogosyan, S. F. Shandarin, and A. L. Melott (1992). "Coherent Structures in the Universe and the Adhesion Model". In: *Astrophys. J.* 393, p. 437. DOI: 10/d5w4qj (cit. on p. 21).

Kraljic, K. et al. (2018). "Galaxy Evolution in the Metric of the Cosmic Web". In: *Mon. Not. R. Astron. Soc.* 474.1, p. 547. DOI: 10/gc3ccm (cit. on pp. 5–6, 8, 50, 53–54, 86, 161, 181, 183, 189, 216).

Kraljic, K. et al. (2019). "Galaxies Flowing in the Oriented Saddle Frame of the Cosmic Web". In: *Mon. Not. R. Astron. Soc.* 483.3, p. 3227. DOI: 10/gf4kjj (cit. on pp. 7–8, 50, 53–54, 86, 112, 161, 181, 183, 189).

Kravtsov, A. V., A. A. Klypin, and A. M. Khokhlov (1997). "Adaptive Refinement Tree: A New High-Resolution N-Body Code for Cosmological Simulations". In: *The Astrophysical Journal Supplement Series* 111, pp. 73–94. DOI: 10/fvf5nz (cit. on p. 42).

Kroupa, P. (2001). "On the Variation of the Initial Mass Function". In: *Mon. Not. R. Astron. Soc.* 322.2, p. 231. DOI: 10/d7c6df (cit. on pp. 37, 164).

Lacey, C. and S. Cole (1993). "Merger Rates in Hierarchical Models of Galaxy Formation". In: *Mon. Not. R. Astron. Soc.* 262.3, p. 627. DOI: 10/gf42q8 (cit. on pp. 18, 50, 181).

Lahav, O., P. B. Lilje, J. R. Primack, and M. J. Rees (1991). "Dynamical Effects of the Cosmological Constant." In: *Mon. Not. R. Astron. Soc.* 251, p. 128. DOI: 10/gf42q7 (cit. on p. 18).

Laigle, C. et al. (2015). "Swirling around Filaments: Are Large-Scale Structure Vortices Spinning up Dark Haloes?" In: *Mon. Not. R. Astron. Soc.* 446.3, p. 2744. DOI: 10/gf42sn (cit. on pp. 114, 160).

Laigle, C. et al. (2018). "COSMOS2015 Photometric Redshifts Probe the Impact of Filaments on Galaxy Properties". In: *Mon. Not. R. Astron. Soc.* 474.4, p. 5437. DOI: 10/gc4dzn (cit. on pp. 6–7, 54, 161).

Landy, S. D. and A. S. Szalay (1993). "Bias and Variance of Angular Correlation Functions". In: *The Astrophysical Journal* 412, pp. 64–71. DOI: 10/bt8j4g (cit. on p. 128).

Lavaux, G. and B. D. Wandelt (2012). "Precision Cosmography with Stacked Voids". In: *ApJ* 754, 109, p. 109. DOI: 10.1088/0004-637X/754/2/109 (cit. on p. 119).

Lebreuilly, U., B. Commerçon, and G. Laibe (2019). "Small Dust Grain Dynamics on Adaptive Mesh Refinement Grids. I. Methods". In: *Astron. Astrophys.* 626, A96. DOI: 10/gf49mx (cit. on p. 183).

Lee, J. and U.-L. Pen (2001). "Galaxy Spin Statistics and Spin-Density Correlation". In: *Astrophys. J.* 555.1, p. 106. DOI: 10/ct2g2h (cit. on p. 6).

Lee, J. and P. Erdogdu (2007). "The Alignments of the Galaxy Spins with the Real-Space Tidal Field Reconstructed from the 2MASS Redshift Survey". In: *Astrophys. J.* 671.2, p. 1248. DOI: 10/b3n7sq (cit. on p. 6).

Libeskind, N. I. et al. (2018). "Tracing the Cosmic Web". In: *Mon. Not. R. Astron. Soc.* 473.1, p. 1195. DOI: 10/gf43rh (cit. on p. 50).

Lindner, U. et al. (1996). "The distribution of galaxies in voids." In: *A&A* 314, pp. 1–12 (cit. on p. 119).

Ludlow, A. D., M. Borzyszkowski, and C. Porciani (2014). "The Formation of CDM Haloes – I. Collapse Thresholds and the Ellipsoidal Collapse Model". In: *Mon. Not. R. Astron. Soc.* 445.4, pp. 4110–4123. DOI: 10/gf42sj (cit. on pp. 49, 54, 183).

Maggiore, M. and A. Riotto (2010). "The Halo Mass Function from Excursion Set Theory. I. Gaussian Fluctuations with Non-Markovian Dependence on the Smoothing Scale". In: *ApJ* 711, pp. 907–927. DOI: 10.1088/0004-637X/711/2/907 (cit. on p. 23).

Malavasi, N. et al. (2017). "The VIMOS Public Extragalactic Redshift Survey (VIPERS): Galaxy Segregation inside Filaments at z = 0.7". In: *Mon. Not. R. Astron. Soc.* 465.4, p. 3817. DOI: 10/f9w3wf (cit. on pp. 5, 7, 53, 161).

Man, Z.-y., Y.-j. Peng, X. Kong, K.-x. Guo, C.-p. Zhang, and J. Dou (2019). "The Dependence of AGN Activity on Environment in SDSS". In: *Mon. Not. R. Astron. Soc.* 488.1, p. 89. DOI: 10/gf488z (cit. on p. 185).

Mandelker, N. et al. (2016). "Instability of Supersonic Cold Streams Feeding Galaxies - I. Linear Kelvin-Helmholtz Instability with Body Modes". In: *Mon. Not. R. Astron. Soc.* 463.4, p. 3921. DOI: 10/f9qw3s (cit. on pp. 161, 176).

Mandelker, N., D. Nagai, H. Aung, A. Dekel, D. Padnos, and Y. Birnboim (2019). "Instability of Supersonic Cold Streams Feeding Galaxies - III. Kelvin-Helmholtz Instability in Three Dimensions". In: *Mon. Not. R. Astron. Soc.* 484.1, p. 1100. DOI: 10/gf48wh (cit. on pp. 161, 176).

Manrique, A. and E. Salvador-Sole (1995). "The Confluent System Formalism. I. The Mass Function of Objects in the Peak Model". In: *Astrophys. J.* 453, p. 6. DOI: 10/cz7w49 (cit. on p. 90).

Martel, H. and P. R. Shapiro (1998). "A Convenient Set of Comoving Cosmological Variables and Their Application". In: *Mon. Not. R. Astron. Soc.* 297.2, p. 467. DOI: 10/dshghr (cit. on p. 46).

Martínez, H. J., H. Muriel, and V. Coenda (2016). "Galaxies Infalling into Groups: Filaments versus Isotropic Infall". In: *Mon. Not. R. Astron. Soc.* 455.1, p. 127. DOI: 10/f77ff9 (cit. on pp. 5, 54).

Martizzi, D. et al. (2019). "Baryons in the CosmicWeb of IllustrisTNG – II: The Connection among Galaxies, Halos, Their Formation Time and Their Location in the Cosmic Web". In: *ArXiv E-Prints*, arXiv:1907.04333 (cit. on pp. 6, 53).

McKinney, J. C., A. Tchekhovskoy, and R. D. Blandford (2012). "General Relativistic Magnetohydrodynamic Simulations of Magnetically Choked Accretion Flows around Black Holes". In: *Mon. Not. R. Astron. Soc.* 423.4, p. 3083. DOI: 10/gf42sd (cit. on p. 164).

Meisner, A. M. and D. P. Finkbeiner (2014). "A Full-Sky, High-Resolution Atlas of Galactic 12 Mm Dust Emission with WISE". In: *Astrophys. J.* 781.1, p. 5. DOI: 10/gf42qd (cit. on p. 184).

Mo, H. J. and S. D. M. White (1996). "An Analytic Model for the Spatial Clustering of Dark Matter Haloes". In: *Mon. Not. R. Astron. Soc.* 282.2, p. 347. DOI: 10/gf439t (cit. on pp. 50, 181).

Moore, A. W. et al. (2001). "Fast Algorithms and Efficient Statistics: N-Point Correlation Functions". In: *Min. Sky*, p. 71. DOI: 10/dpmqwz (cit. on p. 109).

Moresco, M. et al. (2016). "A 6% Measurement of the Hubble Parameter at Z~0.45: Direct Evidence of the Epoch of Cosmic Re-Acceleration". In: *Journal of Cosmology and Astro-Particle Physics* 05, p. 014. DOI: 10/gf42qq (cit. on p. 4).

Musso, M. and R. K. Sheth (2012). "One step beyond: the excursion set approach with correlated steps". In: *MNRAS* 423, pp. L102–L106. DOI: 10.1111/j.1745-3933.2012.01266.x (cit. on p. 24).

Musso, M., C. Cadiou, C. Pichon, S. Codis, K. Kraljic, and Y. Dubois (2018). "How Does the Cosmic Web Impact Assembly Bias?" In: *Mon. Not. R. Astron. Soc.* 476.4, pp. 4877–4906. DOI: 10/gdp56t (cit. on pp. 8, 50, 54, 86, 106, 110, 112, 115, 181, 183, 216).

Musso, M. and R. K. Sheth (2014a). "On the Markovian Assumption in the Excursion Set Approach: The Approximation of Markovian Velocities". In: *Mon. Not. R. Astron. Soc.* 443.2, p. 1601. DOI: 10/gf42qn (cit. on p. 25).

— (2014b). "The Importance of Stepping up in the Excursion Set Approach". In: *Mon. Not. R. Astron. Soc.* 438.3, p. 2683. DOI: 10/f5tfp2 (cit. on p. 25).

Navarro, J. F. and S. D. M. White (1993). "Simulations of Dissipative Galaxy Formation in Hierarchically Clustering Universes - Part One - Tests of the Code". In: *Mon. Not. R. Astron. Soc.* 265, p. 271. DOI: 10/gf42q3 (cit. on p. 38).

Nelson, D. et al. (2013). "Moving Mesh Cosmology: Tracing Cosmological Gas Accretion". In: *Mon. Not. R. Astron. Soc.* 429.4, p. 3353. DOI: 10/f4wr5r (cit. on pp. 33, 161–162, 165, 174).

Nelson, D. et al. (2015). "The Impact of Feedback on Cosmological Gas Accretion". In: *Mon. Not. R. Astron. Soc.* 448.1, p. 59. DOI: 10/gf42rr (cit. on p. 176).

Nelson, D. et al. (2019). "First Results from the TNG50 Simulation: Galactic Outflows Driven by Supernovae and Black Hole Feedback". In: *ArXiv E-Prints*, arXiv:1902.05554 (cit. on p. 38).

Neyrinck, M. C. (2014). "An Origami Approximation to the Cosmic Web". In: *Proc. Int. Astron. Union* 11.S308, pp. 97–102. DOI: 10/gf42sk (cit. on p. 21).

Obuljen, A., N. Dalal, and W. J. Percival (2019). "Anisotropic Halo Assembly Bias and Redshift-Space Distortions". In: *ArXiv E-Prints*, arXiv:1906.11823 (cit. on p. 184).

Ocvirk, P., C. Pichon, and R. Teyssier (2008). "Bimodal Gas Accretion in the Horizon-MareNostrum Galaxy Formation Simulation". In: *Mon. Not. R. Astron. Soc.* 390.4, pp. 1326–1338. DOI: 10/d2djdt (cit. on pp. 6, 33, 161–162).

Padnos, D., N. Mandelker, Y. Birnboim, A. Dekel, M. R. Krumholz, and E. Steinberg (2018). "Instability of Supersonic Cold Streams Feeding Galaxies-II. Non-Linear Evolution of Surface and Body Modes of Kelvin-Helmholtz Instability". In: *Mon. Not. R. Astron. Soc.* 477.3, p. 3293. DOI: 10/gdq3n9 (cit. on p. 176).

Pahwa, I. et al. (2016). "The Alignment of Galaxy Spin with the Shear Field in Observations". In: *Mon. Not. R. Astron. Soc.* 457.1, p. 695. DOI: 10/f8fwpq (cit. on p. 6).

Pan, D. C., M. S. Vogeley, F. Hoyle, Y.-Y. Choi, and C. Park (2012). "Cosmic Voids in Sloan Digital Sky Survey Data Release 7". In: *Mon. Not. R. Astron. Soc.* 421.2, p. 926. DOI: 10/gf42qf (cit. on p. 4).

Paranjape, A., O. Hahn, and R. K. Sheth (2018). "Halo Assembly Bias and the Tidal Anisotropy of the Local Halo Environment". In: *Mon. Not. R. Astron. Soc.* 476.3, pp. 3631–3647. DOI: 10/gdh285 (cit. on pp. 50, 54).

Park, M.-J. et al. (2019). "New Horizon: On the Origin of the Stellar Disk and Spheroid of Field Galaxies". In: *ArXiv E-Prints*, arXiv:1905.02216 (cit. on p. 48).

Paxton, B., L. Bildsten, A. Dotter, F. Herwig, P. Lesaffre, and F. Timmes (2011). "Modules for Experiments in Stellar Astrophysics (MESA)". In: *\apjs* 192, pp. 3–3. DOI: 10/b42mx4 (cit. on p. 37).

Paz, D. J., F. Stasyszyn, and N. D. Padilla (2008). In: *MNRAS* 389, 1127P (cit. on p. 6).

Peebles, P. J. E. (1969). "Origin of the Angular Momentum of Galaxies". In: *Astrophys. J.* 155, p. 393. DOI: `10/djfsxc` (cit. on pp. 5, 28, 54, 161, 167).

Pen, U.-L. (1997). "Generating Cosmological Gaussian Random Fields". In: *Astrophys. J.* 490.2, p. L127. DOI: `10/cmx3cc` (cit. on p. 14).

Penzias, A. A. and R. W. Wilson (1965). "A Measurement of Excess Antenna Temperature at 4080 Mc/s." In: *The Astrophysical Journal* 142, pp. 419–421. DOI: `10/dnwzfr` (cit. on p. 3).

Perlmutter, S. et al. (1999). "Measurements of Ω and Λ from 42 High-Redshift Supernovae". In: *Astrophys. J.* 517.2, p. 565. DOI: `10/dgjnwm` (cit. on p. 4).

Pichon, C. and F. Bernardeau (1999). "Vorticity generation in large-scale structure caustics". In: *A&A* 343, pp. 663–681 (cit. on p. 160).

Pichon, C., D. Pogosyan, T. Kimm, A. Slyz, J. Devriendt, and Y. Dubois (2011). "Rigging Dark Haloes: Why Is Hierarchical Galaxy Formation Consistent with the inside-out Build-up of Thin Discs?" In: *Mon. Not. R. Astron. Soc.* 418.4, pp. 2493–2507. DOI: `10/dmvcx6` (cit. on pp. 6, 161, 174).

Planck Collaboration (2015). "Planck 2015 Results. XIII. Cosmological Parameters". In: *Astron. Astrophys.* 594, A13–A13. DOI: `10/f9scmm` (cit. on p. 163).

— (2018a). "Planck 2018 Results. VI. Cosmological Parameters". In: *ArXiv E-Prints*, arXiv:1807.06209 (cit. on pp. 4, 10–11, 13, 18, 106).

— (2018b). "Planck 2018 Results. XI. Polarized Dust Foregrounds". In: *ArXiv E-Prints*, arXiv:1801.04945 (cit. on p. 184).

Pogosyan, D., J. R. Bond, L. Kofman, and J. Wadsley (1996). "The Cosmic Web and Filaments in Cluster Patches". In: *Am. Astron. Soc. Meet. Abstr.* 189, p. 13.03 (cit. on p. 53).

— (1998). "Cosmic Web: Origin and Observables". In: *Wide Field Surveys in Cosmology*. Ed. by S. Colombi, Y. Mellier, & B. Raban, p. 61 (cit. on p. 98).

Pogosyan, D. et al. (2009a). "The Local Theory of the Cosmic Skeleton". In: *Mon. Not. R. Astron. Soc.* 396.2, p. 635. DOI: `10/btbmph` (cit. on p. 54).

— (2009b). "The local theory of the cosmic skeleton". In: *MNRAS* 396, pp. 635–667. DOI: `10.1111/j.1365-2966.2009.14753.x` (cit. on pp. 30, 100, 111, 120–121, 130).

Pontzen, A., R. Roškar, G. Stinson, and R. Woods (2013). "Pynbody: N-Body/SPH Analysis for Python". In: *Astrophysics Source Code Library*, ascl:1305.002 (cit. on p. 153).

Porciani, C., A. Dekel, and Y. Hoffman (2002). "Testing Tidal-Torque Theory - II. Alignment of Inertia and Shear and the Characteristics of Protohaloes". In: *Mon. Not. R. Astron. Soc.* 332.2, pp. 339–351. DOI: `10/ftg6pn` (cit. on pp. 21, 29).

Porqueres, N., J. Jasche, T. A. Enßlin, and G. Lavaux (2018). "Imprints of the Large-Scale Structure on AGN Formation and Evolution". In: *Astron. Astrophys.* 612, A31. DOI: `10/gf488x` (cit. on p. 185).

Porter, S. C., S. Raychaudhury, K. A. Pimbblet, and M. J. Drinkwater (2008). "Star Formation in Galaxies Falling into Clusters along Supercluster-Scale Filaments". In: *Mon. Not. R. Astron. Soc.* 388.3, p. 1152. DOI: `10/d6f3rg` (cit. on pp. 5, 53).

Press, W. H. and P. Schechter (1974). "Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation". In: *Astrophys. J.* 187, pp. 425–425. DOI: `10/fn2xj3` (cit. on pp. 21, 25, 111).

Price, D. J. et al. (2017). "Phantom: A Smoothed Particle Hydrodynamics and Magnetohydrodynamics Code for Astrophysics". In: *ArXiv E-Prints*, arXiv:1702.03930 (cit. on p. 133).

Prieto, J., A. Escala, M. Volonteri, and Y. Dubois (2017). "How AGN and SN Feedback Affect Mass Transport and Black Hole Growth in High-Redshift Galaxies". In: *Astrophys. J.* 836.2, p. 216. DOI: `10/gf2mzc` (cit. on pp. 152, 161, 174).

Prunet, S., C. Pichon, D. Aubert, D. Pogosyan, R. Teyssier, and S. Gottloeber (2008). "Initial Conditions For Large Cosmological Simulations". In: *ApJS* 178, pp. 179–188. DOI: 10.1086/590370 (cit. on pp. 106, 128).

Ramakrishnan, S., A. Paranjape, O. Hahn, and R. K. Sheth (2019). "Cosmic Web Anisotropy Is the Primary Indicator of Halo Assembly Bias". In: *ArXiv E-Prints*, arXiv:1903.02007 (cit. on pp. 106, 183).

Rasera, Y. and R. Teyssier (2006). "The History of the Baryon Budget. Cosmic Logistics in a Hierarchical Universe". In: *Astron. Astrophys.* 445.1, p. 1. DOI: 10/b358cr (cit. on p. 164).

Redner, S. (2001). *A Guide to First-Passage Processes.* Cambridge University Press. DOI: 10.1017/CBO9780511606014 (cit. on p. 25).

Rees, M. J. and J. P. Ostriker (1977). "Cooling, Dynamics and Fragmentation of Massive Gas Clouds: Clues to the Masses and Radii of Galaxies and Clusters." In: *Mon. Not. R. Astron. Soc.* 179, p. 541. DOI: 10/gf42qg (cit. on p. 5).

Rey, M. P. and A. Pontzen (2017). "Quadratic Genetic Modifications: A Streamlined Route to Cosmological Simulations with Controlled Merger History". In: 9 (June), pp. 1–9 (cit. on p. 185).

Rice, S. O. (1945). "Mathematical Analysis of Random Noise-Conclusion". In: *Bell Syst. Tech J* 24, p. 46. DOI: 10/gft5xr (cit. on p. 26).

Riess, A. G. et al. (1998). "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". In: *Astron. J.* 116.3, p. 1009. DOI: 10/ftbf8j (cit. on p. 4).

Robotham, A. S. G. et al. (2013). "Galaxy And Mass Assembly (GAMA): The Life and Times of L* Galaxies". In: *Mon. Not. R. Astron. Soc.* 431.1, p. 167. DOI: 10/f4wzxs (cit. on p. 6).

Rojas, R. R., M. S. Vogeley, F. Hoyle, and J. Brinkmann (2004). "Photometric Properties of Void Galaxies in the Sloan Digital Sky Survey". In: *Astrophys. J.* 617.1, p. 50. DOI: 10/dthcd2 (cit. on pp. 5, 54).

Romano-Díaz, E., E. Garaldi, M. Borzyszkowski, and C. Porciani (2017). "ZOMG – II. Does the Halo Assembly History Influence Central Galaxies and Gas Accretion?" In: *Mon. Not. R. Astron. Soc.* 469.2, pp. 1809–1823. DOI: 10/gbrz6j (cit. on p. 183).

Rosdahl, J. and J. Blaizot (2012). "Extended Lyα Emission from Cold Accretion Streams". In: *Mon. Not. R. Astron. Soc.* 423.1, p. 344. DOI: 10/f32fms (cit. on pp. 169, 176).

Rosdahl, J., J. Schaye, Y. Dubois, T. Kimm, and R. Teyssier (2017). "Snap, Crackle, Pop: Sub-Grid Supernova Feedback in AMR Simulations of Disc Galaxies". In: *Mon. Not. R. Astron. Soc.* 466.1, p. 11. DOI: 10/f93w65 (cit. on pp. 37–38).

Rosen, A. and J. N. Bregman (1995). "Global Models of the Interstellar Medium in Disk Galaxies". In: *Astrophys. J.* 440, p. 634. DOI: 10/c67hxg (cit. on p. 164).

Roth, N., A. Pontzen, and H. V. Peiris (2016). "Genetically Modified Haloes: Towards Controlled Experiments in LCDM Galaxy Formation". In: *Mon. Not. R. Astron. Soc.* 455.1, p. 974. DOI: 10/gf5bvg (cit. on p. 185).

Rubin, V. C. and W. K. Ford (1970). "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions". In: *Astrophys. J.* 159, p. 379. DOI: 10/cqkqht (cit. on pp. 2–3).

Salpeter, E. E. (1955). "The Luminosity Function and Stellar Evolution." In: *Astrophys. J.* 121, p. 161. DOI: 10/dqrfgf (cit. on p. 37).

Scannapieco, C. et al. (2012). "The Aquila Comparison Project: The Effects of Feedback and Numerical Methods on Simulations of Galaxy Formation". In: *Mon. Not. R. Astron. Soc.* 423.2, pp. 1726–1749. DOI: 10/f32tbv (cit. on pp. 48, 50).

Schaefer, B. M. (2009). "Galactic Angular Momenta and Angular Momentum Correlations in the Cosmological Large-Scale Structure". In: *International Journal of Modern Physics D* 18, pp. 173–222. DOI: 10.1142/S0218271809014388 (cit. on p. 161).

Schäfer, B. M. (2009). "Galactic Angular Momenta and Angular Momentum Correlations in the Cosmological Large-Scale Structure". In: *Int. J. Mod. Phys. D* 18.2, p. 173. DOI: 10/bvqwjg (cit. on pp. 28, 54).

Schaye, J. et al. (2015). "The EAGLE Project: Simulating the Evolution and Assembly of Galaxies and Their Environments". In: *Monthly Notices of the Royal Astronomical Society* 446, pp. 521–554. DOI: 10/gfx7t8 (cit. on pp. 6, 37, 50).

Schmidt, M. (1959). "The Rate of Star Formation." In: *Astrophys. J.* 129, p. 243. DOI: 10/fkm5r4 (cit. on p. 37).

Shapley, H. and H. D. Curtis (1921). "The Scale of the Universe". In: *Bulletin of the National Research Council, Vol. 2, Part 3, No. 11, p. 171-217* 2, pp. 171–217 (cit. on p. 1).

Sheth, R. K., H. J. Mo, and G. Tormen (2001). "Ellipsoidal Collapse and an Improved Model for the Number and Spatial Distribution of Dark Matter Haloes". In: *Mon. Not. R. Astron. Soc.* 323.1, pp. 1–12. DOI: 10/br49c6 (cit. on pp. 22, 54, 90).

Sheth, R. K. and G. Tormen (2004). "On the environmental dependence of halo formation". In: *MNRAS* 350, pp. 1385–1390. DOI: 10.1111/j.1365-2966.2004.07733.x (cit. on p. 181).

Sheth, R. K., K. C. Chan, and R. Scoccimarro (2013). "Nonlocal Lagrangian bias". In: *Phys. Rev. D* 87.8, 083002, p. 083002. DOI: 10.1103/PhysRevD.87.083002 (cit. on p. 54).

Sheth, R. K. and R. van de Weygaert (2004). "A Hierarchy of Voids: Much Ado about Nothing". In: *Mon. Not. R. Astron. Soc.* 350.2, p. 517. DOI: 10/drkqbz (cit. on p. 92).

Silk, J. (1977). "On the Fragmentation of Cosmic Gas Clouds. I. The Formation of Galaxies and the First Generation of Stars." In: *Astrophys. J.* 211, p. 638. DOI: 10/fvr8f4 (cit. on p. 5).

Smith, B. D. et al. (2017). "Grackle: A Chemistry and Cooling Library for Astrophysics". In: *Mon. Not. R. Astron. Soc.* 466.2, pp. 2217–2234. DOI: 10/f92fb6 (cit. on p. 35).

Somerville, R. S. and R. Davé (2015). "Physical Models of Galaxy Formation in a Cosmological Framework". In: *Annu. Rev. Astron. Astrophys.* 53, p. 51. DOI: 10/gfx7x6 (cit. on p. 37).

Sousbie, T., C. Pichon, S. Colombi, and D. Pogosyan (2008). "The 3D skeleton: tracing the filamentary structure of the Universe". In: *MNRAS* 383.4, pp. 1655–1670 (cit. on p. 50).

Sousbie, T., C. Pichon, and H. Kawahara (2011). "The Persistent Cosmic Web and Its Filamentary Structure - II. Illustrations". In: *Mon. Not. R. Astron. Soc.* 414.1, p. 384. DOI: 10/fgtd65 (cit. on pp. 91, 113).

Springel, V. and L. Hernquist (2003). "A Multi-Phase Model for Simulations of Galaxy Formation". In: *Astrophys. Supercomput. Using Part. Simul.* 208, p. 273 (cit. on p. 38).

Springel, V., N. Yoshida, and S. D. M. White (2001). "GADGET: A Code for Collisionless and Gasdynamical Cosmological Simulations". In: *New Astronomy* 6, pp. 79–117. DOI: 10/dfh7mw (cit. on pp. 115, 133).

Springel, V. (2005). "The Cosmological Simulation Code GADGET-2". In: *Mon. Not. R. Astron. Soc.* 364.4, p. 1105. DOI: 10/cx6gpd (cit. on p. 40).

Springel, V., C. S. Frenk, and S. D. M. White (2006). "The Large-Scale Structure of the Universe". In: *Nature* 440.7088, p. 1137. DOI: 10/c2qhbn (cit. on pp. 4, 50, 160).

Springel, V. et al. (2018). "First Results from the IllustrisTNG Simulations: Matter and Galaxy Clustering". In: *Monthly Notices of the Royal Astronomical Society* 475, pp. 676–698. DOI: 10/gdb2qb (cit. on p. 48).

Stewart, K. R. et al. (2013). "Angular Momentum Acquisition in Galaxy Halos". In: *Astrophys. J.* 769.1, p. 74. DOI: 10/gf48kr (cit. on pp. 51, 161, 174).

Stewart, K. R. et al. (2017). "High Angular Momentum Halo Gas: A Feedback and Code-Independent Prediction of LCDM". In: *Astrophys. J.* 843.1, p. 47. DOI: 10/gf48kq (cit. on pp. 33, 50, 161).

Stinson, G. S., C. Brook, A. V. Macciò, J. Wadsley, T. R. Quinn, and H. M. P. Couchman (2013). "Making Galaxies In a Cosmological Context: The Need for Early Stellar Feedback". In: *Mon. Not. R. Astron. Soc.* 428.1, p. 129. DOI: 10/f4v5t6 (cit. on p. 38).

Sutherland, R. S. and M. A. Dopita (1993). "Cooling Functions for Low-Density Astrophysical Plasmas". In: *Astrophys. J. Suppl. Ser.* 88, pp. 253–327. DOI: 10/cwm2cc (cit. on p. 164).

Szapudi, I. and A. S. Szalay (1999). "The Variance of a New Class of N-Point Correlation Estimators in Poisson and Binomial Point Processes". In: (cit. on p. 128).

Tegmark, M., M. R. Blanton, M. A. Strauss, and et al. (2004). "The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey". In: *ApJ* 606, pp. 702–740. DOI: 10.1086/382125 (cit. on p. 5).

Tempel, E., R. S. Stoica, and E. Saar (2013). "Evidence for Spin Alignment of Spiral and Elliptical/S0 Galaxies in Filaments". In: *Mon. Not. R. Astron. Soc.* 428.2, p. 1827. DOI: 10/f42fg8 (cit. on pp. 6, 54).

Tempel, E. and N. I. Libeskind (2013). "Galaxy Spin Alignment in Filaments and Sheets: Observational Evidence". In: *The Astrophysical Journal Letters* 775, p. L42. DOI: 10/gf42r5 (cit. on pp. 6, 160).

Tenneti, A., S. Singh, R. Mandelbaum, T. D. Matteo, Y. Feng, and N. Khandai (2015). "Intrinsic alignments of galaxies in the MassiveBlack-II simulation: analysis of two-point statistics". In: *MNRAS* 448, pp. 3522–3544. DOI: 10.1093/mnras/stv272 (cit. on p. 6).

Teyssier, R. (2002). "Cosmological Hydrodynamics with Adaptive Mesh Refinement. A New High Resolution Code Called RAMSES". In: *Astron. Astrophys.* 385.1, pp. 337–364. DOI: 10/fcr8x4 (cit. on pp. 42, 133).

Teyssier, R., A. Pontzen, Y. Dubois, and J. I. Read (2013). "Cusp-Core Transformations in Dwarf Galaxies: Observational Predictions". In: *Mon. Not. R. Astron. Soc.* 429.4, p. 3068. DOI: 10/f4wn9h (cit. on p. 38).

Tillson, H., J. Devriendt, A. Slyz, L. Miller, and C. Pichon (2015). "Angular Momentum Transfer to a Milky Way Disc at High Redshift". In: *Mon. Not. R. Astron. Soc.* 449.4, p. 4363. DOI: 10/f7dz4m (cit. on pp. 51, 162, 174).

Toro, E. F., M. Spruce, and W. Speares (1994). "Restoration of the Contact Surface in the HLL-Riemann Solver". In: *Shock Waves* 4.1, p. 25. DOI: 10/cf8sv9 (cit. on p. 44).

Toro, E. F. (2009). *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction.* 3. ed. Berlin: Springer. 724 pp. (cit. on p. 44).

Tozzi, P. and C. Norman (2001). "The Evolution of X-Ray Clusters and the Entropy of the Intracluster Medium". In: *Astrophys. J.* 546.1, p. 63. DOI: 10/dxmnj7 (cit. on p. 33).

Trebitsch, M., J. Blaizot, J. Rosdahl, J. Devriendt, and A. Slyz (2017). "Fluctuating Feedback-Regulated Escape Fraction of Ionizing Radiation in Low-Mass, High-Redshift Galaxies". In: *Monthly Notices of the Royal Astronomical Society* 470, pp. 224–239. DOI: 10/gf42r2 (cit. on pp. 37, 164).

Trujillo, I., C. Carretero, and S. G. Patiri (2006). "Detection of the Effect of Cosmological Large-Scale Structure on the Orientation of Galaxies". In: *ApJ* 640, pp. L111–L114. DOI: 10.1086/503548 (cit. on p. 6).

Turk, M. J. et al. (2011). "Yt: A Multi-Code Analysis Toolkit for Astrophysical Simulation Data". In: *Astrophys. J. Suppl. Ser.* 192.1, p. 9. DOI: 10/ft6md2 (cit. on pp. 153–154).

Tweed, D., J. Devriendt, J. Blaizot, S. Colombi, and A. Slyz (2009). "Building Merger Trees from Cosmological N -Body Simulations". In: *Astron. Astrophys.* 506.2, pp. 647–660. DOI: 10/fwx3tf (cit. on p. 162).

Van den Bosch, F. C., T. Abel, R. A. C. Croft, L. Hernquist, and S. D. M. White (2002). "The Angular Momentum of Gas in Protogalaxies. I. Implications for the Formation of Disk Galaxies". In: *Astrophys. J.* 576.1, p. 21. DOI: 10/c4x48f (cit. on p. 5).

Van Leer, B. (1984). "On the Relation Between the Upwind-Differencing Schemes of Godunov, Engquist–Osher and Roe". In: *SIAM J. Sci. and Stat. Comput.* 5.1, pp. 1–20. DOI: 10/bx6zfd (cit. on pp. 44–45).

Vogelsberger, M., S. Genel, D. Sijacki, P. Torrey, V. Springel, and L. Hernquist (2013). "A Model for Cosmological Simulations of Galaxy Formation Physics". In: *Mon. Not. R. Astron. Soc.* 436.4, p. 3031. DOI: 10/f5kh8c (cit. on p. 37).

Vogelsberger, M. et al. (2014). "Introducing the Illustris Project: Simulating the Coevolution of Dark and Visible Matter in the Universe". In: *Mon. Not. R. Astron. Soc.* 444.2, p. 1518. DOI: 10/f6kp3r (cit. on pp. 6, 48, 50).

Wadsley, J. W., J. Stadel, and T. Quinn (2004). "Gasoline: A Flexible, Parallel Implementation of TreeSPH". In: *New Astronomy* 9, pp. 137–158. DOI: 10/fn6fwg (cit. on pp. 40, 133).

Wechsler, R. H., A. R. Zentner, J. S. Bullock, A. V. Kravtsov, and B. Allgood (2006). "The Dependence of Halo Clustering on Halo Formation History, Concentration, and Occupation". In: *ApJ* 652, pp. 71–84. DOI: 10.1086/507120 (cit. on pp. 5, 53, 181).

Weinberg, D. H., M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo (2013). "Observational Probes of Cosmic Acceleration". In: *Phys. Rep.* 530.2, p. 87. DOI: 10/gdz64b (cit. on pp. 4, 10).

Welker, C., J. Devriendt, Y. Dubois, C. Pichon, and S. Peirani (2014). "Mergers Drive Spin Swings along the Cosmic Web." In: *Mon. Not. R. Astron. Soc.* 445, p. L46. DOI: 10/gf42tk (cit. on pp. 6, 53–54).

White, M. (2014). "The Zel'dovich Approximation". In: *Mon. Not. R. Astron. Soc.* 439.4, p. 3630. DOI: 10/f52mqm (cit. on p. 21).

White, S. D. M. (1984). "Angular Momentum Growth in Protogalaxies". In: *Astrophys. J.* 286, pp. 38–41. DOI: 10/dsjrcs (cit. on pp. 5, 28, 167).

White, S. D. M. and C. S. Frenk (1991). "Galaxy Formation through Hierarchical Clustering". In: *Astrophys. J.* 379, p. 52. DOI: 10/bkstjz (cit. on p. 33).

Wiersma, R. P. C., J. Schaye, and B. D. Smith (2009). "The Effect of Photoionization on the Cooling Rates of Enriched, Astrophysical Plasmas". In: *Mon. Not. R. Astron. Soc.* 393.1, p. 99. DOI: 10/d68bjp (cit. on p. 36).

Wolfire, M. G., C. F. McKee, D. Hollenbach, and A. G. G. M. Tielens (2003). "Neutral Atomic Phases of the Interstellar Medium in the Galaxy". In: *Astrophys. J.* 587.1, p. 278. DOI: 10/dvcsx4 (cit. on p. 35).

Yang, Y. et al. (2017). "Interstellar-Medium Mapping in M82 through Light Echoes around Supernova 2014J". In: *Astrophys. J.* 834.1, p. 60. DOI: 10/gf42q4 (cit. on p. 36).

Zeldovich, Y. B. (1970). "Gravitational Instability: An Approximate Theory for Large Density Perturbations". In: *Astron. Astrophys.* 5, pp. 84–89. DOI: 10.1007/s13398-014-0173-7.2 (cit. on p. 4).

Zentner, A. R. (2007). "The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth". In: *International Journal of Modern Physics D* 16, pp. 763–815. DOI: 10.1142/S0218271807010511 (cit. on p. 54).

Sujet : L'impact des grandes structures de l'Univers sur la formation des halos de matière noire et leur galaxie

Résumé : À grande échelle, il est frappant de voir que la distribution anisotrope de la matière forme un large réseau de vides délimités par des murs qui, avec les filaments présents à leurs intersections, tissent la toile cosmique. La matière qui doit former plus tard les halos de matière noire et leurs galaxies afflue vers les nœuds compacts se situant à l'intersection des filaments et garde dans ce processus une empreinte/le souvenir de la toile cosmique.

Dans cette thèse, je développe une extension contrainte de la théorie de l'excursion dans son approximation dite "du franchissement vers le haut" pour prédire la masse, le taux d'accrétion et le temps de formation des halos de matière noire au voisinage des proto-filaments (qui sont identifiés comme des points-selles du potentiel). Les points-selles sont utilisés comme un référentiel local dans lequel l'évolution des propriétés physiques et morphologiques des galaxies est quantifiée aux grandes échelles / à grande échelle. À masse fixe, le modèle prédit que le taux d'accrétion et le temps de formation varient avec l'orientation et la distance au point-selle, confirmant que le biais d'assemblage est sensible aux forces de marées de la toile cosmique. Les halos peu massifs, s'étant formés tôt et "affamés" sont regroupés le long de l'axe principal des filaments, tandis que les halos plus massifs, plus jeunes sont répartis autour des nœuds. Les différents gradients observés pour différentes quantités, tels que la masse typique et le taux d'accrétion, ont pour origine l'anisotropie du point-selle et leur dépendance distincte aux moyennes et aux variances du champ. Pour les faibles décalages vers le rouge, ce modèle prédit qu'à masse fixe il y a un excès de galaxies rouges dans des directions préférentielles, comme l'ont montré des relevés spectroscopiques (GAMA) et photométriques (COSMOS), mais aussi les simulations hydrodynamiques (Horizon-AGN).

J'ai également calculé les taux de fusions par analyse multi-échelle des conditions initiales pour prédire l'assemblage anisotrope des halos et comprendre son impact sur la formation des galaxie. Outre les fusions de halos, j'ai aussi pris en compte les fusions de murs et de filaments qui ont un effet sur l'accrétion galactique et j'ai calculé leur propriétés à un et deux points en fonction du temps cosmique. J'ai établi le lien entre les taux de fusion et la connectivité. J'ai ensuite exploité ce lien pour estimer l'effet des structures à grande échelle sur le biais d'assemblage. Cette théorie décrit l'anisotropie de la toile cosmique, qui est un élément important pour décrire conjointement l'évolution de la physique et de la dynamique des galaxies dans leur environnement, en particulier les alignements intrinsèques ou les diversités morphologiques.

Afin d'étudier l'accrétion cosmique à de plus faibles échelles, j'ai implémenté une nouvelle méthode de particules traceuses dans le code à raffinement de grille adaptatif Ramses. Cette méthode est basée sur un échantillonnage de Monte-Carlo et est capable de reconstruire la trajectoire lagrangienne du gaz et son retraitement. Je démontre que la distribution spatiale des particules traceuses reproduit précisément celle du gaz, et je propose une extension capable de suivre tout le cycle des baryons dans leurs échanges avec les étoiles et les trous noirs. Cette approche est particulièrement adaptée aux problèmes astrophysiques qui requièrent simultanément d'avoir une résolution efficace des chocs avec un solveur de Godounov et de suivre l'histoire lagrangienne des baryons. Je l'utilise ensuite dans plusieurs simulations zoomées pour étudier l'acquisition du moment angulaire par les galaxies via leur accrétion bi-modale pour les hauts décalages vers le rouge. J'y observe que l'amplitude et l'orientation du moment angulaire du gaz froid sont conservées jusque dans le halo interne où le moment angulaire contribue à l'augmentation de la rotation des galaxies. Les couples de pressions sont plus importants en amplitude mais, de par leur turbulence, ils sont incohérents et sont dominés globalement par les couples gravitationnels. Les couples de la matière noire dominent dans le halo externe, ceux des étoiles dominent dans le disque.

Mots clés : galaxies – matière noire – toile cosmique – cosmologie – accrétion anisotrope

Subject : How do large scale structures impact dark matter halo and galaxy formation

Abstract: The strikingly anisotropic large-scale distribution of matter is made of an extended network of voids delimited by sheets, with filaments at their intersection which together form the cosmic web. Matter that will later form dark matter halos and their galaxies flows towards compact nodes at filaments' intersections and in the process, retains the imprint of the cosmic web.

In this thesis, I predict the mass, accretion rate, and formation time of dark matter halos near proto-filaments (identified as saddle points of the potential) using a conditional version of the excursion set theory in its so-called up-crossing approximation. The (filament-type) saddle points provides a local frame in which to quantify the induced physical and morphological evolution of statistical properties of galaxies on large scales. The model predicts that at fixed mass, mass accretion rate and formation time vary with orientation and distance from the saddle, demonstrating that assembly bias is indeed susceptible to the tides imposed by the cosmic web. Starved, early-forming halos of smaller mass lie preferentially along the main axis of filaments, while more massive and younger halos are found closer to the nodes. Distinct gradients for distinct quantities such as typical mass and accretion rate occur because the saddle condition is anisotropic, and because the statistics of these observables depend on both the means and their covariances. The signature of this model corresponds at low redshift to an excess of reddened galaxies at fixed mass along preferred directions, as recently reported in spectroscopic (GAMA) and photometric (COSMOS) surveys and in hydrodynamical simulations (Horizon-AGN).

I also compute the rate of merger events in the multi-scale initial conditions to forecast special events driving the anisotropic assembly of dark matter halos and understand their impact on galaxy formation. Beyond halo mergers, I consider all sets of mergers, including wall and filament mergers, as they impact the geometry of galactic infall. Their one- and two-points statistics are computed as a function of cosmic time. I establish the relation between merger rates and connectivity, which is then used to assess the impact the large scale structures on assembly bias. The anisotropy of the cosmic web, as encoded in this theory, is a significant ingredient to describe jointly the physics and dynamics of galaxies in their environment, e.g. in the context of intrinsic alignments or morphological diversity.

In order to explore the impact of cosmic infall on smaller scales I implemented a novel tracer particles algorithm in the Eulerian adaptive mesh refinement code Ramses. The tracer particles are based on a Monte Carlo approach and keep tracks of where fluid elements originate, so as to follow their Lagrangian trajectories and re-processing history. I show that they reproduce the gas distribution very accurately and I extend them to also trace the stars and black holes through the full cycle of baryons. These tracer particles are ideal to study complex astrophysical systems where both efficiency of shock-capturing Godunov schemes and a Lagrangian follow-up of the fluid are required simultaneously, in particular in cold flows. Thanks to this accurate tracer particle algorithm, the acquisition and loss of angular momentum of both cold and hot accretion flows onto galaxies at high redshift can be studied reliably. I find that the amplitude and orientation of the specific angular momentum of the cold gas is preserved down to the inner halo where the angular momentum contributes to the spin-up of galaxies, while for the hot gas it is lost at larger radii. Pressure torques, stronger in magnitude than gravitational torques are, however, spatially incoherent, which leads them to have no significant impact on the redistribution of angular momentum of the accretion flows. Gravitational torques, which dominate globally, are the main driver of the loss of angular momentum of the accretion flows in those halos, with dark matter gravitational torques dominating in the outer halo and stellar gravitational torques dominating in the disk.