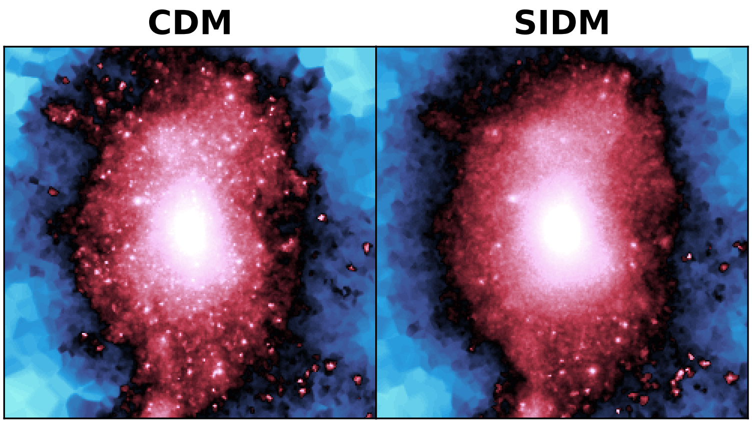
Topology of the cosmic web in cold and self-interacting dark matter models

LIU Report

Institut d'Astrophysique de Paris (IAP)



SZPILFIDEL Adrian

Master 2 International Research Track (IRT)
Observatoire de Paris, Paris Sciences et Lettres (PSL)

Supervisors: LAIGLE Clotilde (IAP), BOLDRINI Pierre (LIRA)

From September 8 to January 6, 2026

Acknowledgments

I would first like to thank Clotilde Laigle for dedicating her time and for giving me the opportunity to discover the physics of the cosmic web. I also thank Pierre Boldrini for his help on this project and for providing the simulation data. I am grateful to the Institut d'Astrophysique de Paris (IAP) for welcoming me and offering excellent conditions during my internship.

Abstract

Context. Dark matter (DM) was discovered almost a century ago and yet, it is still one of the biggest mysteries of modern astrophysics. The dominant cosmological model, namely ΛCDM , which, despite being astonishingly precise for describing the formation of structures on cosmological scales ($\gtrsim 10\,\mathrm{Mpc}$), faces many discrepancies between observations and predictions on galactic scales ($\lesssim 1\,\mathrm{Mpc}$). For instance, the predicted density distribution of dark matter in galaxies is higher than what has been measured, and also the number of satellites predicted by ΛCDM is not recovered by observations. In this context, alternative dark matter models are proposed to match the observations on galactic scales while recovering the result of ΛCDM on larger scales. One such alternative scenario describes dark matter as self-interacting. In the present work, we compare the topology of the large-scale structure of self-interacting Dark Matter (SIDM) with Cold Dark Matter (CDM). Additionally, the objective is also to evaluate whether these differences are expected to be observable.
Method. We use the genus of the isodensity surfaces to compare their topology in the two DM models on different scales. The analysis is carried out on snapshots of DM-only cosmological simulations, using boxes of size $48\,\mathrm{Mpc}/h$, filled with 576^3 particles. We also perform the analysis on finer scales centered around halos, using box sizes of the order of the halo sizes, for 24 different halos with masses ranging from 10^{12} to $10^{14}\,\mathrm{M}_\odot/h$ over redshifts $z = 0$ to $z = 2$.
Results. We present a first comparison of the topology of the large-scale structure in CDM and SIDM simulations using the genus statistics. We find that CDM is systematically clumpier than SIDM in regions centered on massive dark matter halos, especially for scales $\lesssim 0.05$ Mpc/h. On larger scales, the difference between the two models decreases, confirming that SIDM and CDM behave similarly on large scales. When computed on larger volumes ($> 10\,r_{\mathrm{vir}}$), the genus measured converges to similar values in the two models, and it becomes indistinguishable when computed on the entire $48\,\mathrm{Mpc}/h$ box. To quantify the differences, we use the clump abundance A_{C} and consider the ratio $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$, which reaches its largest amplitude at $z = 0$. As expected, this ratio depends on the SIDM self-interaction cross-section and approaches unity as the cross-section decreases. We also observe that the differences between the two models increase with time (decreasing redshift). Finally, we apply this analysis directly on the halo positions, to evaluate the feasibility of an observational application. The result shows that the topological difference between CDM and SIDM is also detectable using only the halo distribution. However, more work is needed to assess the robustness of this result in the presence of observational uncertainties.
Conclusions. Topological diagnostics (genus, clump abundance) provide a powerful probe of dark matter self-interactions. It allows one to identify precisely the scale at which most of the difference happens. The strength of this method, with respect to other probes of DM models (e.g. satellite counts), is that it does not require extracting halos. The genus is based on field analysis, since it relies on the full DM distribution and not on a discrete dataset. Further work will test to which extent this method can be applied on observational surveys, with forthcoming datasets such as Euclid.

Key words. cosmology: large-scale structure of Universe / dark matter / methods: numerical / statistical

Résumé (version française)

Contexte. La matière noire (DM) a été mise en évidence il y a près d'un siècle et demeure l'un des plus grands mystères de l'astrophysique moderne. Le modèle cosmologique dominant, ΛCDM , bien qu'extraordinairement précis pour décrire la formation des structures aux échelles cosmologiques ($\gtrsim$ 10 Mpc), présente des divergences avec les observations aux échelles galactiques ($\lesssim$ 1 Mpc). Par exemple, la distribution de densité de matière noire prédite dans les galaxies est plus élevée que ce qui est mesuré, et le nombre de satellites prévu par ΛCDM n'est pas retrouvé. Dans ce contexte, des modèles alternatifs de DM ont été proposés afin de s'accorder avec les observations à petite échelle tout en préservant les prédictions de ΛCDM à grande échelle. L'un de ces scénarios introduit des auto-interactions de la DM. Dans ce travail, nous comparons la topologie de la structure à grande échelle de la DM auto-interagissante (SIDM) et de la DM froide (CDM). Un autre objectif est d'évaluer la détectabilité potentielle de ces signatures dans les futures données d'Euclid.
Méthode. Nous utilisons le genus des surfaces d'isodensité pour comparer leur topologie entre les deux modèles à différentes échelles. L'analyse est effectuée sur des instantanés de simulations cosmologiques DM-only, réalisées dans des boîtes de taille 48 Mpc/h contenant 576^3 particules. Nous réalisons aussi l'étude à des échelles plus fines centrées sur des halos, avec des tailles de boîte de l'ordre de la taille des halos, pour 24 halos de masses entre 10^{12} et 10^{14} $\mathrm{M}_\odot/h$ sur les redshifts $z = 0$ à $z = 2$.
Résultats. Nous présentons une première comparaison de la topologie de la structure à grande échelle dans des simulations CDM et SIDM via la statistique du genus. Nous trouvons que la CDM est systématiquement plus grumeleuse que la SIDM dans des régions centrées sur des halos massifs, en particulier pour des échelles $\lesssim 0.05$ Mpc/h. À plus grande échelle, la différence entre les deux modèles diminue, confirmant que SIDM et CDM se comportent de façon similaire sur les grandes échelles. Lorsque le genus est calculé sur des volumes plus grands ($> 10\,r_{\mathrm{vir}}$), il converge vers des valeurs similaires dans les deux modèles, et devient indiscernable sur la boîte entière de 48 Mpc/h. Pour quantifier ces différences, nous utilisons l'abondance de grumeaux A_{C} et considérons le rapport $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$, qui atteint son amplitude maximale à $z = 0$. Comme attendu, ce rapport dépend de la section efficace d'auto-interaction et tend vers l'unité lorsque celle-ci diminue. Nous avons également observé que les différences entre les deux modèles s'accentuent avec le temps (diminution du redshift). Enfin, nous avons appliqué cette analyse directement sur les positions des halos, afin d'évaluer la faisabilité d'une application observationnelle. Le résultat montre que la différence de topologie entre CDM et SIDM est également détectable via la distribution des halos seuls. Cependant, des travaux supplémentaires sont nécessaires pour évaluer la robustesse de ce résultat en présence d'effets observationnels.
Conclusions. Les diagnostics topologiques (genus, abondance de grumeaux) constituent une sonde puissante des auto-interactions de la matière noire. Ils permettent d'identifier précisément l'échelle à laquelle se concentrent les différences. L'un des atouts de cette méthode, par rapport à d'autres approches (p. ex. le comptage de satellites), est qu'elle ne nécessite pas l'extraction de halos: le genus repose sur une analyse de champ, utilisant la distribution complète de DM plutôt qu'un jeu de données discret. Des travaux futurs testeront dans quelle mesure cette méthode est applicable aux relevés observationnels, avec des jeux de données à venir tels qu'Euclid.

Mots-clés. cosmologie: structure à grande échelle de l'Univers / matière noire / méthodes: numériques / statistiques

Contents

Introduction

ΛCDM has become the dominant cosmological model describing the evolution of mass and energy in the Universe, following the discovery of the accelerated cosmic expansion by Riess et al. 1998. The most precise measurements to date indicate that the Universe is spatially flat and composed of matter $\Omega_\mathrm{m} = 0.315 \pm 0.007$ and dark energy $\Omega_\Lambda = 0.685 \pm 0.007$ (Planck Collaboration et al. 2020). The matter component includes ordinary baryonic matter and an invisible form of matter that interacts with itself and the former only through gravity. Until now, this dark matter (DM) has been assumed cold and collisionless. Within the ΛCDM framework, the cold dark matter (CDM) has not been associated with any known elementary particle of the standard model, and its nature remains an open question. ΛCDM has been remarkably successful in reproducing observations at cosmological scales ($\gtrsim 10\,\mathrm{Mpc}$), where simulations and large-scale surveys show excellent agreement with predictions (Eisenstein et al. 2005; Beutler et al. 2017; Planck Collaboration et al. 2020; Alam et al. 2021). However, on smaller, galactic scales (1–100 kpc), tensions remain between theoretical predictions and observations. Dwarf galaxies often exhibit central cores rather than the steep cusps predicted by CDM: the cusp–core problem (Blok 2010; Boldrini 2021). CDM-based cosmological simulations predict a density profile that diverges as $\rho \propto r^{-1}$ toward the center of DM halos (Navarro et al. 1997). In contrast, observations of dwarf galaxies suggest a flat density profile in their central regions (Moore 1994; De Blok et al. 2001). CDM-based cosmological simulations also predict more satellite galaxies than are observed (the missing-satellites problem; Klypin et al. 1999), and Milky Way–like halos appear to host subhalos that are too dense compared to data (the too-big-to-fail problem; Boylan-Kolchin et al. 2011). For a comprehensive review of small-scale challenges to CDM, see Bullock et al. 2017.

These discrepancies have motivated the exploration of alternative models. A promising scenario for solving these problems introduces dark matter with self-interactions, known as Self-Interacting Dark Matter (SIDM), first developed by Carlson et al. 1992 and Spergel et al. 2000. SIDM has been later investigated through cosmological simulations (e.g. Fischer et al. 2022; Correa et al. 2022; Despali et al. 2025). This model introduces new free parameters for dark matter self-interactions, such as the cross-section σ_0. SIDM behaves like CDM in the limit of null cross-section. However, assuming non-zero cross-section leads to different dark matter distribution on galactic scales. Dave et al. 2001; Rocha et al. 2013; Elbert et al. 2015 showed that SIDM can produce cored density profiles in dwarf galaxies through gravitational heating, alleviating the cusp–core problem: the SIDM particles gain kinetic energy via self-interactions, leading to more extended orbits and, therefore, a reduction in central density. This process redistributes energy, preventing the formation of bound structures in dense regions, resulting in more diffuse structures. Zavala et al. 2013; Kaplinghat et al. 2019 demonstrated that the 'too-big-to-fail' problem might be solved by SIDM. Fischer et al. 2022 found that cosmological simulations using small-angle scattering for SIDM results in a reduced number of satellites, addressing the missing satellites problem. For a review of SIDM models and their implications on small-scale structures, see Tulin et al. 2018. These studies suggest that SIDM can provide a better fit to observations on galactic scales while remaining consistent with large-scale structure formation.

In practice, obtaining robust constraints on DM self interactions from galaxy cores is extremely challenging: complex baryonic physics in these regions, such as bars, gas inflows/outflows, and bursty stellar feedback, redistribute mass and energy and can mimic or erase SIDM signatures. A more effective strategy is to probe larger scales, which mitigate baryonic degeneracies while still lying in a regime where SIDM gravitational heating yields smoother and more diffuse structures. This is precisely where topological diagnostics can help distinguish SIDM from CDM. On larger scales than few Mpc, matter is distributed along a cosmic web (Bond et al. 1996; Pogosyan et al. 1998). This complex network consists of filaments and walls, intersecting at nodes, surrounding low-density regions (voids). The dark matter distribution, being highly anisotropic at these scales, induces gravitational potential wells that drive the infall of baryonic matter along filaments toward their intersections (nodes), where halos grow, shaping the large-scale structure of the Universe. Even though baryonic matter is not modeled in the simulation used for this work, it is useful to clarify the process of galaxy formation that is driven by DM. As gas accretes into a halo, it is shock-heated close to the virial temperature, then radiatively cools, loses pressure support, and fragments (Rees et al. 1977; White and Frenk 1991; White and Rees 1978; Mo et al. 2010). This sequence of shock heating, cooling, and fragmentation triggers star formation (Barkana et al. 2001), building the first galaxies in the densest knots of the cosmic web. Therefore, the observed galaxy distribution is a powerful tracer for probing the underlying dark matter distribution (Davis et al. 1982; De Lapparent et al. 1986; Geller et al. 1989). Measuring the geometry of the cosmic web provides insight into its growth and content. Among the various existing methods for studying the topology of the cosmic web, one approach is to use the genus of the density contours (Gott et al. 1987; Park, Choi, et al. 2005), which quantifies the connectivity of the density field. The gravitational heating induced by self-collisions in SIDM, limiting substructure formation, might lead to different topology than CDM. Thus, the genus is expected to be sensitive to these topological differences. This method has the advantage of being non-parametric, thereby avoiding biases in the analysis. This statistical tool has been used widely in cosmology, for instance to constrain cosmological parameters (Appleby et al. 2018) or on alternative dark matter models (Watts et al. 2017).

The objective of this work is to compare the topology of the cosmic web formed in CDM and SIDM cosmologies using DM-only cosmological simulations. These are N-body simulations that use initial conditions consistent with ΛCDM cosmology. We investigate the topological properties of DM density distribution in both CDM and SIDM frameworks comparing the results on two different parameter sets for SIDM. We also investigate how the differences between the two models evolve with redshift. This comparison aims to assess how dark matter self-interactions modify the connectivity of large-scale structures and how such effects could be tested with upcoming data from large-scale surveys such as the Euclid mission. Euclid survey aims to probe the nature of dark energy and dark matter, it will provide high-resolution optical imaging, as well as near-infrared imaging and spectroscopy, over about 14,000 deg^2 of extragalactic sky (Euclid et al. 2025). With its unprecedented volume and depth, Euclid will map the large-scale structure of the Universe with high precision, making it an ideal dataset for testing cosmological models, including SIDM.

This report is organized as follows. In section 1, I present the genus statistic and how it is applied to large-scale structure analysis of cosmological simulations. In section 2, I report the results of the comparison of the cosmic web topology in CDM and SIDM frameworks. Finally, I summarize and discuss possible perspectives on this work.

Introduction (french version)

ΛCDM est devenu le modèle cosmologique dominant décrivant l'évolution de la masse et de l'énergie dans l'Univers, à la suite de la découverte de l'expansion cosmique accélérée par Riess et al. 1998. Les mesures les plus précises à ce jour indiquent que l'Univers est spatialement plat et composé de matière $\Omega_{\rm m} = 0.315 \pm 0.007$ et d'énergie noire $\Omega_\Lambda = 0.685 \pm 0.007$ (Planck Collaboration et al. 2020). La composante de matière inclut la matière baryonique ordinaire et une forme invisible de matière interagissant principalement par la gravité. Jusqu'à présent, cette matière noire (DM) est supposée froide et sans collisions. Dans le cadre ΛCDM , la matière noire froide (CDM) n'est associée à aucune particule élémentaire connue du modèle standard, et sa nature demeure une question ouverte. ΛCDM reproduit remarquablement bien les observations aux échelles cosmologiques ($\gtrsim$ 10 Mpc), où simulations et grands relevés s'accordent avec les prédictions (Eisenstein et al. 2005; Beutler et al. 2017; Planck Collaboration et al. 2020; Alam et al. 2021). En revanche, à des échelles galactiques (1–100 kpc), des tensions persistent entre théorie et observations. Les galaxies naines montrent souvent des cœurs centraux plutôt que les cusps abrupts prédits par la CDM: le problème cusp-core (Blok 2010; Boldrini 2021). Les simulations CDM prédisent un profil de densité divergeant comme $\rho \propto r^{-1}$ vers le centre des halos (Navarro et al. 1997), alors que les observations suggèrent des profils plats (Moore 1994; De Blok et al. 2001). La CDM prédit aussi davantage de satellites qu'observé (problème des satellites manquants ; Klypin et al. 1999) et des sous-halos trop denses (problème too-big-to-fail ; Boylan-Kolchin et al. 2011). Pour une synthèse des défis à petite échelle, voir Bullock et al. 2017.

Ces tensions ont motivé l'exploration de modèles alternatifs. Un scénario prometteur introduit des auto-interactions de la DM, la Self-Interacting Dark Matter (SIDM), proposée par Carlson et al. 1992 et Spergel et al. 2000, puis étudiée en simulations cosmologiques (p. ex. Fischer et al. 2022; Correa et al. 2022; Despali et al. 2025). Ce modèle introduit des paramètres libres, notamment la section efficace σ_0. La SIDM se comporte comme la CDM pour $\sigma_0 \to 0$, mais une section non nulle modifie la distribution de DM aux échelles galactiques. Dave et al. 2001; Rocha et al. 2013; Elbert et al. 2015 ont montré que la SIDM peut produire des cœurs via un chauffage gravitationnel: les particules gagnent de l'énergie cinétique par auto-interactions, adoptent des orbites plus étendues et réduisent la densité centrale, rendant les structures plus diffuses. Zavala et al. 2013; Kaplinghat et al. 2019 suggèrent que le problème too-big-to-fail peut être résolu avec la SIDM. Fischer et al. 2022 trouvent qu'une diffusion à petit angle en SIDM réduit le nombre de satellites, répondant au problème des satellites manquants. Globalement, la SIDM semble mieux correspondre aux observations à petite échelle tout en restant compatible avec la formation de structure à grande échelle.

En pratique, contraindre les auto-interactions via les cœurs galactiques est difficile à cause de la physique baryonique complexe (barres, flux de gaz, rétroaction stellaire) pouvant imiter ou effacer les signatures SIDM. Une stratégie plus robuste consiste à sonder des échelles plus grandes, réduisant ces dégénérescences, tout en restant dans un régime où le chauffage gravitationnel SIDM engendre des structures plus lisses et diffuses. C'est là que des diagnostics topologiques peuvent distinguer SIDM de CDM. Au-delà de quelques Mpc, la matière se distribue en toile cosmique (Bond et al. 1996; Pogosyan et al. 1998), réseau de filaments et de parois se rejoignant aux nœuds et entourant des vides. La DM, fortement anisotrope, façonne des puits de potentiel qui guident l'accrétion baryonique vers les halos. La distribution observée des galaxies est donc un traceur puissant de la DM (Davis et al. 1982; De Lapparent et al. 1986; Geller et al. 1989). Mesurer la géométrie de la toile éclaire sa croissance et son contenu. Parmi les méthodes topologiques, le genus des contours de densité (Gott et al. 1987; Park, Choi, et al. 2005) quantifie la connectivité du champ. Le chauffage gravitationnel en SIDM, en limitant les sous-structures, pourrait mener à une topologie différente, rendant le genus sensible à ces écarts. La méthode est non paramétrique et largement utilisée, y compris pour contraindre des paramètres cosmologiques (Appleby et al. 2018) ou des modèles alternatifs de DM (Watts et al. 2017).

L'objectif de ce travail est de comparer la topologie de la toile cosmique formée en CDM et SIDM avec des simulations DM-only. Nous étudions les propriétés topologiques de la densité de DM dans les deux cadres, pour deux jeux de paramètres SIDM, et analysons l'évolution des différences avec le redshift. Cette comparaison vise à évaluer comment les auto-interactions modifient la connectivité des structures à grande échelle et comment ces effets peuvent être testés avec les futurs relevés tels que Euclid. Le télescope Euclid sondera la nature de l'énergie sombre et de la DM, fournira une imagerie optique haute résolution ainsi que de l'imagerie et de la spectroscopie en proche IR sur $\sim$ 14,000 deg^2 de ciel extragalactique (Euclid et al. 2025), offrant un jeu de données idéal pour tester des modèles, y compris la SIDM.

Ce rapport est organisé comme suit. En section 1, nous présentons la statistique du genus et son application à l'analyse de la structure à grande échelle dans les simulations. En section 2, nous rapportons les résultats de la comparaison de la topologie de la toile cosmique en CDM et SIDM. Enfin, nous résumons et discutons des perspectives de ce travail.

1 Methodology

This section describes the data and methods used to analyze the topology of the cosmic web in CDM and SIDM cosmologies. I first present the cosmological simulation data employed in this study, followed by a detailed explanation of the genus statistics and its application for characterizing the large-scale structure.

1.1 Cosmological simulation data

The analysis presented in this work has been carried out using data from cosmological simulations in CDM and SIDM[1] (Fischer et al. 2021) frameworks, based on the code OpenGadget-3 (Springel et al. 2001; Springel 2005; Groth et al. 2023). The simulation runs over a Hubble time (13.8 Gyr). I use snapshots of the simulations, corresponding to different times or redshifts. I have used 3D boxes of physical size 48 Mpc/h, filled with 576^3 particles for 4 different redshifts from $z = 3$ to $z = 0$. Here h is the reduced Hubble constant defined by:

$$h = \frac{H_0}{100\,\mathrm{km\,s^{-1}/\,Mpc}},$$

where H_0 is the Hubble constant[2]. The particle coordinates are given in comoving coordinates and do not include the expansion of the Universe. The mass resolution of dark matter in the simulation is $4.37 \times 10^7\,\mathrm{M}_\odot/h$.

In simulations, it is observed that dark matter forms halos and subhalos that are characterized by their mass and virial radius r_{vir}. The latter is defined as the radius within which the mean density is 200 times the critical density of the Universe; it can be interpreted as the halo size. I also use a catalog that reports the position of 3×10^5 halos and subhalos, sorted from the most massive to the least massive. This catalog is extracted using the SUBFIND algorithm implemented in the code OpenGadget-3 (Groth et al. 2023). This catalog is useful to choose dense substructures and zoom in on them, for a finer analysis as a function of the object mass and size. At redshift $z = 0$, the virial radius of such halos ranges from $0.01\,\mathrm{Mpc}/h$ to $1\,\mathrm{Mpc}/h$ while the mass ranges from dwarf galaxy masses ($10^8\,\mathrm{M}_\odot/h$) to galaxy cluster masses ($10^{14}\,\mathrm{M}_\odot/h$).

The model of SIDM introduces self-interactions between dark matter particles, characterized by a cross-section parameter σ_0. I am using the so-called frequently self-interacting dark matter (fSIDM) model, which corresponds to highly anisotropic scattering, with small-angle scattering (Fischer et al. 2022). It means that the DM particles are slightly deviated, with a small angle from their initial trajectory during an interaction. Consequently, the number of scatter events has to be high to have a significant effect on the DM distribution. This is why this model is called "frequent". This SIDM model is motivated by particle physics and it matches the observations better (Fischer et al. 2021). I use data for two different cross-sections: $\sigma_0 = 1\,\mathrm{cm}^2/\mathrm{g}$ and $\sigma_0 = 0.1\,\mathrm{cm}^2/\mathrm{g}$. The cross-section quantifies the probability of interaction between two dark matter particles. A higher cross-section means that dark matter particles are more likely to interact with each other, leading to more gravitational heating. SIDM particles gain kinetic energy through collisions. The speed increase causes particles to move to more extended orbits, resulting in a more diffuse dark matter distribution compared to the collisionless CDM predictions. Thus, a larger cross-section σ_0 leads to a smoother DM distribution and more pronounced differences with CDM on galactic scales. The larger cross-section ($1\,\mathrm{cm}^2/\mathrm{g}$) maximizes the effects of self-interactions, while the smaller cross-section ($0.1\,\mathrm{cm}^2/\mathrm{g}$) represents a more conservative scenario that is closer to CDM behavior. Nonetheless, on large scales ($\gtrsim 10$ Mpc/h), the effects of self-interactions are negligible, and SIDM behaves similarly to CDM. Using these two cross-sections allows us to explore the impact of self-interactions on the topology of the cosmic web and to assess how sensitive our results are to the strength of these interactions.

[1] https://www.darkium.org/

[2] The normalization of lengths and masses by h is commonly used by cosmologists to standardize the data using different values for H_0.

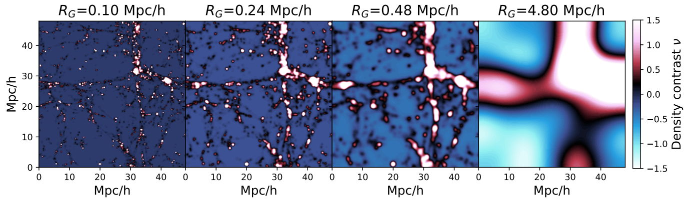


Figure 1: Density contrast field of a CDM snapshot at redshift $z = 0$, smoothed with Gaussian kernels of size from $0.1\,\mathrm{Mpc}/h$ to $4.8\,\mathrm{Mpc}/h$.

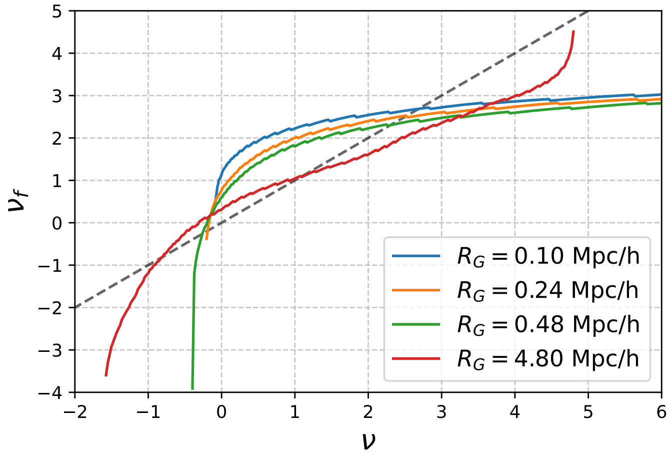


Figure 2: Volume-fraction threshold ν_{f} as a function of the density contrast ν. The dashed line represents the identity relation $\nu_{\mathrm{f}} = \nu$. The volume-fraction threshold is calculated on the full $48\,\mathrm{Mpc}/h$ CDM box at redshift $z = 0$, smoothed with 4 different Gaussian kernels.

1.2 Analysis of the cosmic web structure

To compare the topology of the cosmic web in CDM and SIDM cosmologies, I use the genus statistics, which quantifies the connectivity of the surface densities at different threshold values. It is one of the Minkowski functionals, and it has been introduced in large-scale structures studies by Gott et al. 1987; Mecke et al. 1993. To calculate it, I use a code developed by D. Pogosyan[3]. In the following, I describe how I start from the particle positions given by cosmological simulations and end with the genus as a function of the density. The main steps of the analysis are as follows:

- Interpolation of the DM density field from the discrete positions.
- Smoothing of the density field with a Gaussian kernel.
- Interpolation of a grid of the density contrast (ν) on a regularly spaced volume-fraction threshold grid.
- Calculation of the genus on 8 sub-boxes, as a function of the density contrast ν.

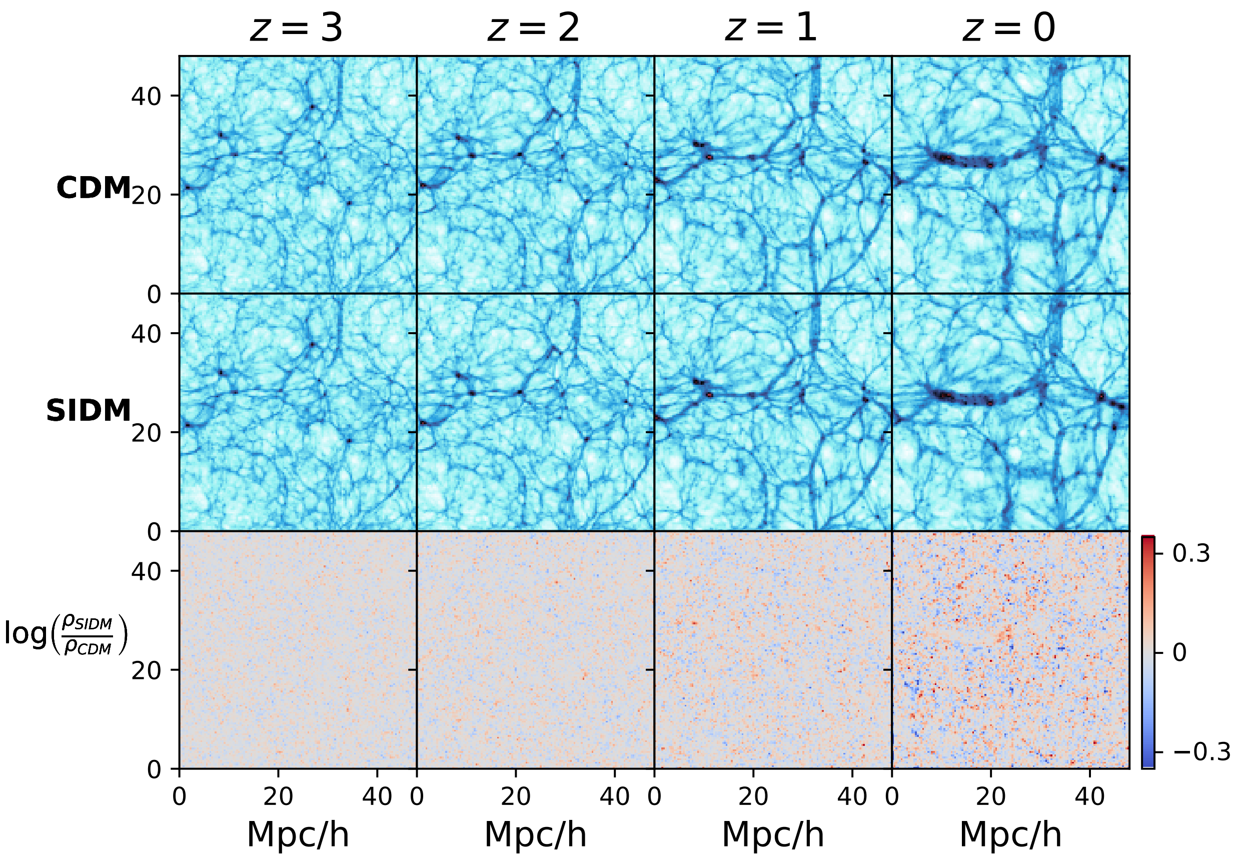


Figure 3: Interpolated density fields of dark matter. Each column corresponds to a redshift z from 3 to 0. The first row is the density in CDM simulation, the second row is SIDM with cross-section $\sigma_0 = 1\,\mathrm{cm}^2/\mathrm{g}$. The third row is the residuals of the two models. The quantities are shown in logarithmic scale and averaged on the z-axis over 20 pixel-slices.

1.2.1 Interpolation of the dark matter density field

In cosmological simulations which are N-body simulations, the dark matter distribution is discretized, as particles are modeled as point masses moving in a 3D space. We need an interpolated density field associating any point of the box to a value. Otherwise, low-density regions are fixed at 0 even though they may be close to a particle. The interpolation is based on a Delaunay tessellation, which connects nearby particles into a network of non-overlapping tetrahedra (Delaunay 1934). Inside each tetrahedron, physical quantities can be smoothly interpolated, allowing us to estimate the density at any point of the box. To calculate the tessellation in the box, I use the implementation of Sousbie 2011 in DisPerSE. This calculation being numerically demanding, I divide the simulation box into 8 sub-boxes when interpolating the entire box, and then interpolate the density field separately. Afterward, I reconstruct the entire box and check that there are no edge effects. The calculation for the full box takes about 8 CPU hours. The resulting density field is represented on a regular grid of 1000^3 pixels of size $0.048\,\mathrm{Mpc}/h$.

1.2.2 Smoothing of the density field

We now have a density value for each fraction of the box. However, we are interested in the structures at a particular scale. I smooth the density field by convolving the distribution with a 3D Gaussian kernel of width R_G. The effect of the smoothing is to filter out the density fluctuations on scales smaller than R_G. In the following, we call R_G the smoothing scale of the distribution. I represent the smoothed density field for different values of R_G in Figure 1. This illustrates how the smoothing process removes the density fluctuations that occur on smaller scales than R_G. The choice of the smoothing scale enables one to analyze the density field fluctuations at different length scales.

1.2.3 Density contrast and volume-fraction threshold

The density field can be described by the density contrast defined by:

$$\nu = \frac{\rho - \bar{\rho}}{\sigma}\,, \tag{1}$$

where ρ is the smoothed density, $\bar{\rho}$ the mean density and σ the standard deviation of the density distribution. For a given threshold value of ν, an isodensity surface can be defined in the 3D space, which separates the region with higher density from the region with lower density than the threshold. In the following, I use the volume-fraction threshold ν_f which is, for a given value of ν, the density threshold value for which the isodensity surface encloses the same volume fraction than in a Gaussian field. To understand the dependence of the volume-fraction threshold $\nu_\mathrm{f}(\nu)$ on the density contrast, we refer to the Figure 2. As we smooth more (i.e R_G increases), $\nu_\mathrm{f}(\nu)$ becomes closer to the identity, which is expected for a Gaussian field. At lower smoothing scale, we see that an identical volume-fraction is reached for higher and higher density contrast, because at low scales, the field contains very highly-contrasted regions (dense halos), more than it contains very underdense regions (voids). Conversely, for low density contrast values, the volume-fraction

[3] Private communication.

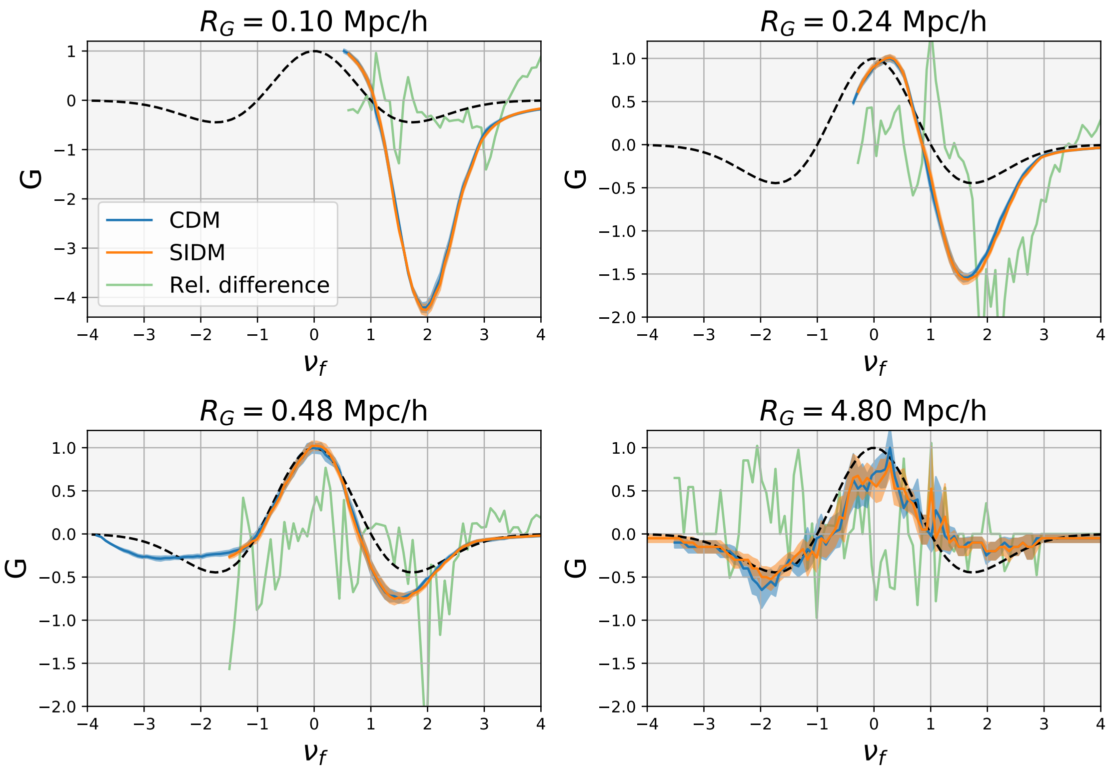


Figure 4: Comparison of SIDM and CDM genus curves. The panels represent the genus as a function of the volume-fraction threshold computed on the full 48 Mpc/h cosmological box at $z = 0$, using four different smoothing scales. The black dashed line represents the analytical genus for a Gaussian field. The solid curves for CDM and SIDM represents the mean genus calculated separately on 8 sub-boxes, normalized by the maximum value. The color filled areas shows the deviation from the mean of the sub-boxes calculation. The green solid line represents the difference between CDM and SIDM genus curves, normalized by the standard mean error of the sub-boxes calculation: $\delta G = \frac{G_{\rm SIDM} - G_{\rm CDM}}{\sqrt{\sigma^2_{\rm SIDM} + \sigma^2_{\rm CDM}}}$.

becomes very small, thus the volume-fraction threshold $\nu_{\rm f}$ goes to very negative values. This shows that the density field does not contain underdense regions as much as in a Gaussian field. In other words, the field is more asymmetric on small scales (low smoothing scale $R_{\rm G}$) and tends to a Gaussian distribution on large scales (high smoothing scale $R_{\rm G}$). The volume-fraction threshold $\nu_{\rm f}$ is widely used in the literature because it allows separating variations in the topology from variation in the density distribution between different simulations. For a given density contrast ν, the genus in CDM and SIDM could differ even though the topology is identical. This effect is corrected when using the volume-fraction threshold $\nu_{\rm f}$. Because we would like a regularly spaced grid of $\nu_{\rm f}$, I first calculate $\nu_{\rm f}(\nu)$, using a fine grid for ν.

However, by design, the code that calculates the genus needs the value of ν, so I interpolate a new grid of ν, over a regularly spaced $\nu_{\rm f}$ grid that I use for presenting the results.

1.2.4 Genus statistics

The physical properties of the large-scale structures can be studied through the topology of such surface density (Gott et al. 1987, Park, Choi, et al. 2005, Watts et al. 2017). Minkowski functionals are mathematical tools (Minkowski 1903), based on integral geometry that have been used in cosmology to characterize the large-scale structures through its connectivity, shape and content (see Mecke et al. 1993; Liu et al. 2025; Armijo et al. 2025).

The genus is defined by:

$$G = \text{nb. of holes} - \text{nb. of isolated regions}. \qquad (2)$$

Alternatively, one can figure the genus as the number of cuts through a surface required to separate it into two pieces. In topology, the meaning of a hole is similar to that of a torus. A high number of holes means that the topology of the surface density is "spongelike", that is, the connectivity is high, therefore the genus G is positive. Isolated regions are disconnected and closed surface density. When the number of such surfaces is high, the genus becomes negative, and we call the distribution "meatball like". It is now clear that the genus depends on the density threshold adopted. At low density threshold ν, the surface density draws many closed surfaces, enclosing low density regions: voids. On the other hand, at high density threshold, the surface density encloses dense regions: clusters. In both cases, the corresponding genus is negative. Then, at intermediate density threshold, close to the mean density value, the field is highly connected (spongelike), consequently, the surface has many holes, leading to a positive genus.

The Gauss-Bonnet theorem links the genus of a surface S to the integrated local Gaussian curvature κ over the surface:

$$G = \frac{1}{4\pi} \int_S \kappa \, \mathrm{d}a \qquad (3)$$

The genus as a function of the density contrast $g(\nu)$ for a Gaussian random field is known (Doroshkevich 1973; Adler

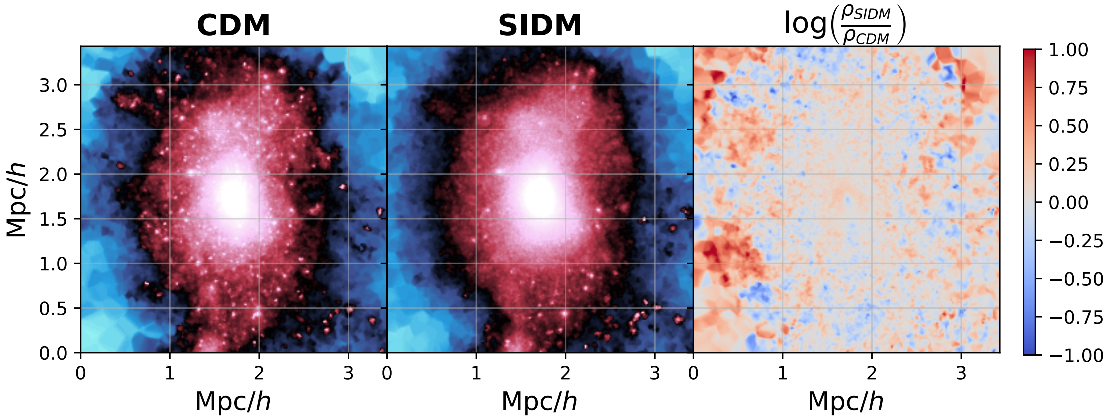


Figure 5: Density field of a sub-box of edge length $3.4\,\mathrm{Mpc}/h$, centered around the most massive halo in the box ($M = 1.3 \times 10^{14}\,\mathrm{M}_\odot/h$, $r_{\rm vir} = 1.1\,\mathrm{Mpc}/h$), for CDM (left), SIDM (middle) and the ratio of the two (right). The densities are represented in logarithmic scale and averaged over 20 pixel-slices of the z-axis.

1981; Hamilton et al. 1986):

$$g(\nu) \propto (1 - \nu^2)\,\mathrm{e}^{-\nu^2/2} \tag{4}$$

This function exhibits a symmetry in the density as $g(\nu) = g(-\nu)$, which is expected for a Gaussian field, i.e the level of connectivity should be equal in high density and low density regions. At infinitely high or low density, the genus goes to zero, as there are no isodensity surfaces. When computing the genus on the box, I divide the box in eight sub-boxes and compute the genus individually, then we can use the mean value and the dispersion from it of the eight boxes to interpret the results.

Statistical metrics have been developed from the genus, and we use some of them to better quantify the differences found in the genus of CDM and SIDM. In particular, I use the clump abundance $A_{\rm C}$ defined by Park, Gott Iii, et al. 2001; Park, Choi, et al. 2005; Park, Kim, et al. 2005 as:

$$A_{\rm C} = \int G\,\mathrm{d}\nu \tag{5}$$

The edges of the integral are $1.2 < \nu_{\rm f} < 2.2$. The interval is centered around the position of the minima of the Gaussian genus, and stays away from biases induced by the shift of the genus curve (Park, Choi, et al. 2005). This represents the area under the genus curve, for high density contrast, in the meatball-like region. The area under the curve for these density contrast values corresponds to abundance of isolated regions (or clumps) of dark matter. To compare the result in both models, we refer to the ratio of clump abundances $A_{C,\rm CDM}/A_{C,\rm SIDM}$. The value of this ratio is interpreted as follows: if $A_{C,\rm CDM}/A_{C,\rm SIDM} > 1$, it means that the genus in CDM is more negative than in SIDM in the overdensity region, meaning that CDM has more isolated high-density regions than SIDM, i.e. CDM is more clumpy than SIDM. Conversely, if $A_{C,\rm CDM}/A_{C,\rm SIDM} < 1$, it means that SIDM is more clumpy than CDM. Therefore, this ratio quantifies how much CDM and SIDM differ in terms of clumpiness. The higher the ratio, the more discrepancies there are between the two models. If the ratio is equal to one, it means that both models have the same clump abundance, i.e. neither CDM nor SIDM have a clumpier distribution than the other. We expect that $A_{C,\rm CDM}/A_{C,\rm SIDM} \geq 1$, indicating that CDM is able to form more clumps than SIDM due to the absence of self-interactions. The self-interactions in SIDM redistribute energy among particles. Therefore, it is more difficult for SIDM to form dense clumps, leading to a smoother distribution compared to CDM.

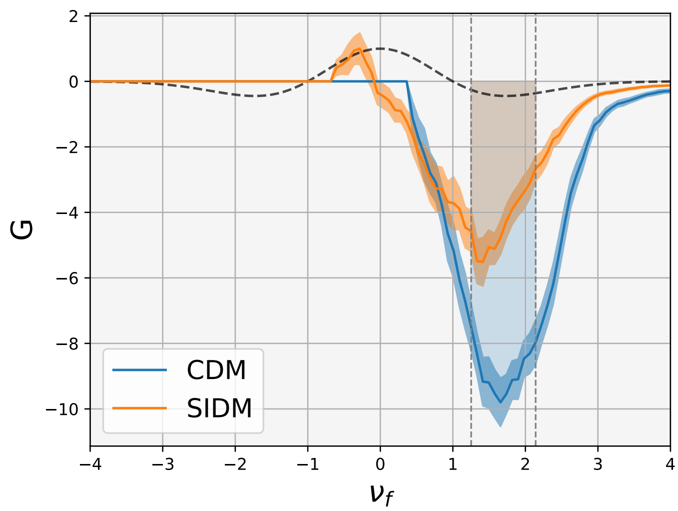


Figure 6: Genus curves of the density field smoothed with a kernel size $R_{\rm G} = 0.01\,\mathrm{Mpc}/h$ in a sub-box centered around the most massive halo in the box ($M = 1.3 \times 10^{14}\,\mathrm{M}_\odot/h$, $r_{\rm vir} = 1.1\,\mathrm{Mpc}/h$), for CDM (blue) and SIDM (orange). The solid curves represent the mean genus calculated separately on 8 sub-boxes, normalized by the maximum value. The color filled areas shows the deviation from the mean of the sub-boxes calculation. The vertical shaded blue and orange areas represent the area used to compute the clump abundance $A_{\rm C}$ in CDM and SIDM respectively.

One of the difficulties of this work was to connect all these steps properly, since each part is numerically costly. I have constructed a pipeline that computes all the steps described above, starting from the positions of particles in the box to the genus of the density field. The total computation time of all these steps on a complete box of 576^3 particles is about 70 CPU hours (interpolation: 12 CPU hours; smoothing: 0.75 CPU hours; volume-fraction $\nu_{\rm f}$: 1.5 CPU hours; genus: 56 CPU hours).

2 Results

2.1 Distribution and topology of the density field on large scales

Looking at the large-scale distribution of dark matter in CDM and SIDM is the first step of the comparison. It provides a general view of potential differences between the two models. A visualization of density fields of SIDM with cross-section $\sigma_0 = 1.0\,\mathrm{cm}^2/\mathrm{g}$ and CDM at different redshifts is shown in Figure 3. In both cases, we see that the matter distribution is rather homogeneous at high redshift and becomes progressively clumpier at later times, illustrating how dark matter collapses under gravity to form large filaments, halos and clusters of halos. We already see that, qualitatively, the distributions on large scales do not differ much between the two dark matter models. This is expected, as SIDM has been designed to solve small-scale problems of CDM, while remaining similar to CDM on large scales, which is in agreement with observations. To quantify the differences, I calculate the ratio of densities between CDM and SIDM at different redshifts, shown in the third row of Figure 3. The ratio of densities shows that CDM and SIDM are similar at redshift $z = 3$. Differences appear progressively, although the dispersion of the ratio remains limited, despite the high value of the cross-section parameter chosen for SIDM. At redshift $z = 0$, for the box interpolated with 1000^3 pixels, the mean value of the residuals is 1, which indicates that CDM and SIDM have equal density on average, and the standard deviation of the residuals is about 2.3, meaning most ratios lie within a factor of ~ 2. This conclusion is consistent with previous work showing SIDM agrees with CDM on large scales (Fischer et al. 2022). The maximal fluctuation amplitude of the residuals is found at $z = 0$, which is a first indication that more significant genus differences may occur at this redshift.

To further test the difference at large scales, I calculate the genus of the distribution. The calculation of the genus in the complete 48 Mpc/h box is represented in Figure 4. Each panel illustrates the genus calculated in a snapshot of the boxes at redshift $z = 0$ for CDM and SIDM using different smoothing scales R_G. Comparing the obtained genus curves with the black dashed curve, representing the genus in a Gaussian field (defined in (2)), we immediately note that both CDM and SIDM density field are non-Gaussian. Instead of a symmetric genus, as the Gaussian genus, we obtain much more isolated surface density, as illustrated by the negative values reached by the genus for high volume-fraction threshold. The non-Gaussianities increase when the smoothing scale is lower. This trend was expected, as we have already seen in Figure 2. This result shows that at small scales, the density field is highly asymmetric as there are many more high-density regions (halos) than low-density regions (voids). Indeed, the resulting genus curves indicate that at small scales ($R_\mathrm{G} = 0.1$ Mpc/h) and for high density contrast threshold, the topology of dark matter surface density is meatball-like, meaning that there are many clumps. This clumpy distribution at scales 0.1 Mpc/h is related to the dark matter halo in which galaxies are embedded. The number of isolated surface density decreases with the smoothing scales and their topology is more similar to a Gaussian field. For large smoothing scales, the number of structures is low, therefore the genus is noisier. The genus is cut off on the low density side when using small smoothing scales. This is because for low density contrast ν, the volume-fraction threshold ν_f needed for the density field to occupy the same volume as in a Gaussian field becomes infinitely negative, as we see in Figure 2. Consequently, we cannot calculate the genus for these values. We qualitatively see that the genus curves of CDM and SIDM are very similar for all smoothing scales. To quantify the differences, I plot the difference between CDM and SIDM genus curves, normalized by the standard mean error of the sub-boxes calculation (green solid line). For all smoothing scales, the difference remains within 3σ, indicating no significant difference between CDM and SIDM genus curves. Thus, the occasional peaks of the difference curve do not represent significant deviations. This conclusion holds for all redshifts from $z = 3$ to $z = 0$ (not shown here).

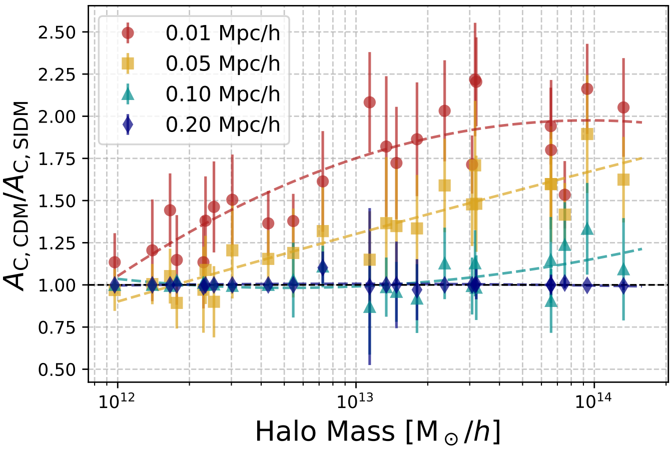


Figure 7: Clump abundance ratio as a function of the CDM halo mass. Each point corresponds to a halo-centered box of size $L_\mathrm{box} = 3r_\mathrm{vir}$. The color code indicates the smoothing scale used for the genus calculation. The error bars represent the uncertainty on the ratio, propagated from the dispersion of the genus calculated in the 8 sub-boxes. The color dashed lines represent polynomial fits of degree two.

In summary, at large scales, both the density field and the genus curves do not reveal significant differences between CDM and SIDM, even for a high cross-section value. This suggests that self-interactions in SIDM do not strongly affect the large-scale topology of the cosmic web compared to CDM. This result is consistent with previous findings that SIDM behaves similarly to CDM on large scales, as self-interactions primarily impact small-scale structures. In the next section, I focus on smaller scales by zooming in on halos to investigate potential differences in the density distribution and topology around these dense structures.

2.2 Distribution and topology of dark matter around halos

Neither the density field nor the genus computed on the full box reveal significant differences between CDM and SIDM at large scales. I therefore probe smaller scales by analyzing only fractions of the box, which allows us to use lower smoothing scales and to perform a finer characterization of the small-scale topology of the density field. I select regions centered on halos. I use the halos 3D positions reported by the halo catalog. I apply the same pipeline on the sub-box as in the previous subsection. I limit the sub-box size to $L_\mathrm{box} = 3\,r_\mathrm{vir}$ to include the halo and its immediate environ-

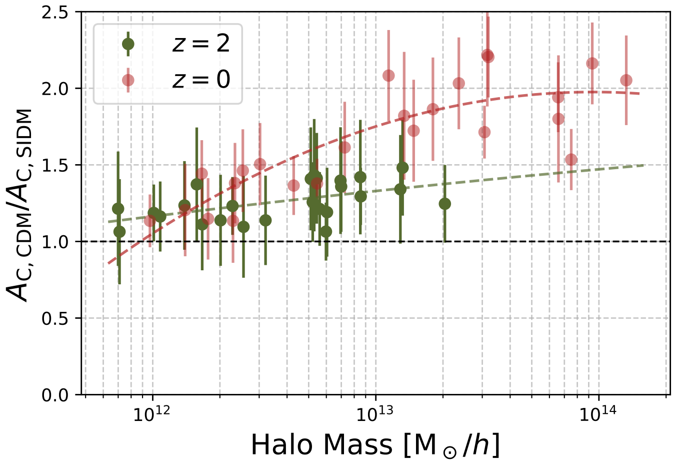


Figure 8: Evolution of the clump abundance ratio across redshift. Each point corresponds to the clump abundance ratio computed from the genus of smoothed density field in halo-centered box of size $L_{\mathrm{box}} = 3r_{\mathrm{vir}}$, using a smoothing scale of $R_{\mathrm{G}} = 0.01\,\mathrm{Mpc}/h$. The color dashed lines represent polynomial fits of degree two.

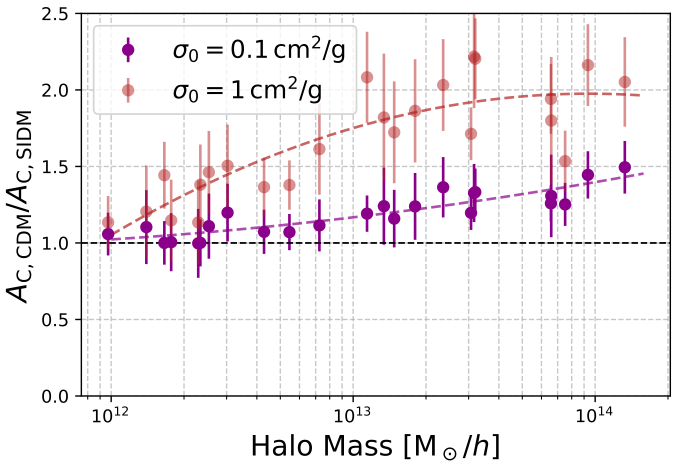


Figure 9: Self-interaction cross-section effect on the clump abundance ratio. The clump abundance ratio is represented as a function of the CDM halo mass for two value of the cross-section for SIDM. Purple points corresponds to low cross-section, associated to more collisionless SIDM. Red points are computed using the maximal value for the cross-section, maximizing the self-interactions in SIDM. The color dashed lines represent polynomial fits of degree two.

ment. The physical size of the sub-box therefore depends on the halo size. It ranges from $0.03\,\mathrm{Mpc}/h$ for small halos to $3.3\,\mathrm{Mpc}/h$ for the most massive halo in the box. This choice of sub-box size ensures that we capture the halo and its surroundings, where self-interactions in SIDM are expected to have the most significant impact on the density distribution and topology. In Figure 5, I show the density field around the most massive halo of the simulation. The comparison reveals a clear difference in substructure abundance: CDM appears very clumpy, with many subhalos, whereas the SIDM distribution is noticeably smoother. This contrast is also visible in the residuals (right panel) as numerous blue speckles, while the red regions around the halo indicate enhanced density in SIDM outskirts, suggesting that matter which has collapsed to the center in CDM remains in the neighborhood in SIDM. The corresponding genus curves for this halo (Figure 6) show the largest CDM–SIDM discrepancies at the smallest smoothing scales (e.g., $R_{\mathrm{G}} = 0.01\,\mathrm{Mpc}/h$) and at high density thresholds (between $\nu_{\mathrm{f}} = 1$ and $\nu_{\mathrm{f}} = 4$). For $R_{\mathrm{G}} = 0.01\,\mathrm{Mpc}/h$, CDM exhibits a markedly lower genus than SIDM at high ν_{f}, with a maximum difference of 8σ, indicating more isolated high-density clumps in CDM. We also quantify this with the clump abundance A_{C} which represents the area below the genus curve in the high-density region, as defined in (5). The shaded areas in Figure 6 illustrate the integration range used to compute A_{C} for both models. For this halo and at $R_{\mathrm{G}} = 0.01\,\mathrm{Mpc}/h$, I find a clump abundance ratio $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}} \simeq 2$, indicating that CDM is roughly twice clumpier than SIDM around this massive halo.

I generalize this analysis to 23 other halos, spanning a mass range from $10^{12}\,\mathrm{M}_\odot/h$ to $10^{14}\,\mathrm{M}_\odot/h$. For each halo, I compute the genus in a sub-box of size $L_{\mathrm{box}} = 3r_{\mathrm{vir}}$ centered on the halo, using various smoothing scales from $R_{\mathrm{G}} = 0.01\,\mathrm{Mpc}/h$ to $R_{\mathrm{G}} = 0.2\,\mathrm{Mpc}/h$. I calculate the clump abundance ratio $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$ for each halo and smoothing scale. The results are summarized in Figure 7, where I plot the ratio $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$ versus halo mass. At $R_{\mathrm{G}} = 0.01\,\mathrm{Mpc}/h$, CDM is significantly more clumpy than SIDM, with ratios up to 2.25 for the most massive halos, consistent with the expectation that SIDM self-interactions suppress small-scale substructures in dense regions. As the smoothing scale increases, the ratio approaches unity, indicating that differences diminish on larger scales. Figure 7 also shows that massive halos are the most discriminating environments: at $R_{\mathrm{G}} = 0.01\,\mathrm{Mpc}/h$, CDM remains more clumpy down to $M \simeq (1\text{–}2) \times 10^{12}\,\mathrm{M}_\odot/h$, and exceeds SIDM by a factor of $\gtrsim 2$ for $M > 10^{13}\,\mathrm{M}_\odot/h$. The minimum halo mass for which CDM is measurably clumpier increases with R_{G}; for $R_{\mathrm{G}} > 0.1\,\mathrm{Mpc}/h$, no significant difference is detected and the genus becomes similar in both models.

In order to test the analyzed region size threshold, beyond which the difference between CDM and SIDM are not noticeable anymore, I have done the same analysis with larger box sizes. I find that beyond a size $L_{\mathrm{box}} = 10r_{\mathrm{vir}}$, for smoothing scales larger than or equal $R_{\mathrm{G}} = 0.05\,\mathrm{Mpc}/h$, the genus is no longer able to differentiate the topology of both dark matter models. More details about this test are given in Appendix Appendix A.

The large-scale structure topology evolves with time, as gravitational collapse proceeds. I have repeated the previous analysis at redshift $z = 2$ to investigate how the clump abundance ratio $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$ changes with redshift. The result is presented in Figure 8. I find that at $z = 2$, the

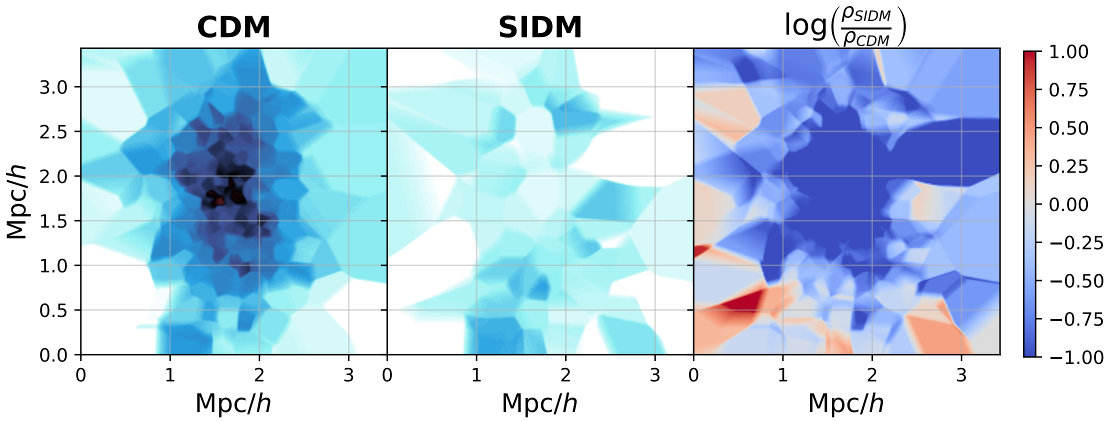


Figure 10: Density field interpolated on the DM halos positions, in the same sub-box than Figure 5 for CDM (left), SIDM (middle) and the ratio of the two (right). The densities are represented in logarithmic scale and averaged over 20 pixel-slices of the z-axis.

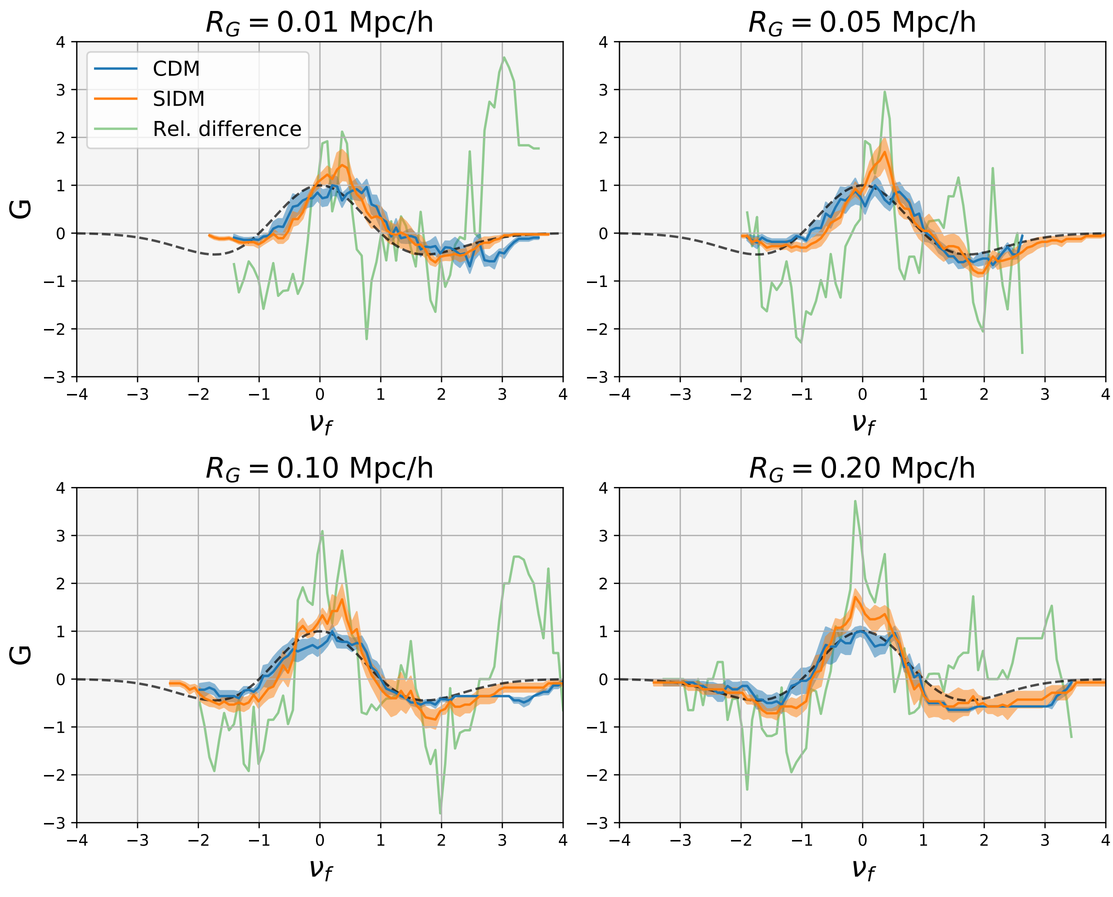


Figure 11: Genus curves of the density field interpolated on the halo positions. The density field is smoothed with 4 different kernel size R_G in entire box, for CDM (blue) and SIDM (orange). The solid curves represent the mean genus calculated separately on 8 sub-boxes, normalized by the maximum value. The color filled areas shows the deviation from the mean of the sub-boxes calculation. The solid green curve displays the difference between SIDM minus CDM, normalized by the total deviation from the mean. The green solid line represents the difference between CDM and SIDM genus curves, normalized by the standard mean error of the sub-boxes calculation.

clump abundance ratio is closer to unity than at $z = 0$, indicating that differences between CDM and SIDM are less pronounced at earlier times. This trend is expected, as self-interactions in SIDM have had less time to affect the small-scale structure formation at higher redshifts. We could already predict this trend, based on the density field presented earlier, in Figure 3, where we have noticed that the largest residuals amplitude are found at low redshifts. The genus computed at different redshifts presented here confirms this statement. The dependence of the clump abundance ratio on the halo mass at $z = 2$ is qualitatively similar to that at $z = 0$, with more massive halos showing larger differences between CDM and SIDM. However, the overall amplitude of the ratio is reduced at $z = 2$. This suggests that observations targeting lower redshift structures may be more sensitive to distinguishing between CDM and SIDM based on their topological properties. However, it is still relevant to study higher redshift to distinguish a physical effect from a systematic effect.

SIDM introduces cross-section as a free parameter for dark matter self-interactions. We have investigated how the clump abundance ratio varies with the SIDM cross-section by analyzing simulations with different cross-section values. The result is shown in Figure 9. I find that as the cross-section increases, the clump abundance ratio $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$ also increases, indicating that stronger self-interactions lead to a more pronounced suppression of small-scale structure in SIDM compared to CDM. This trend is consistent across different halo masses, with more massive halos exhibiting larger ratios for higher cross-sections. This result is consistent with the idea that SIDM should behave as CDM when the cross-section vanishes. It also suggests that measurements of the clump abundance ratio in observed structures could provide constraints on the SIDM cross-section, helping to distinguish between different dark matter models.

2.3 Forecast for observations: genus of the halo density field

The results of the topological analysis presented above demonstrate that the genus statistic can effectively distinguish between CDM and SIDM cosmologies at small scales around halos. However, the analysis has been performed on the DM density field, which is not directly observable, as DM does not emit light and is not directly detectable in observations. In order to connect these findings to observations, we need to consider observable tracers of the underlying DM distribution. DM is the dominant component in matter in the Universe and it drives the shape of the large-scale structures through gravitational attraction. Therefore, observing the distribution of galaxies and galaxy clusters constitutes a powerful probe of the underlying DM distribution. Moreover, the genus has already been measured in galaxy surveys such as the Sloan Digital Sky Survey (Park, Choi, et al. 2005; Watts et al. 2017), making it a priori a relevant statistic to compare with simulations. Future surveys such as the Euclid survey (Euclid et al. 2025) will deliver a galaxy catalog complete to $M < 10^8\,\mathrm{M}_\odot$ at low redshift in the deep fields at the end of the survey and therefore offer the opportunity to draw the density distribution very precisely.

In the DM-only cosmological simulations, we have access to the halo/subhalo catalog which constitutes a more observable feature because DM halos host galaxies, and galaxies are directly detectable by telescopes. We want to know if the genus is sensitive to the difference between CDM and SIDM when computed on the halo distribution. Here, I present the topological analysis applied directly on the distribution of DM halos and subhalos, and not on individual DM particles. The method used is the same as described in section 1. The resulting halos number-density field is represented in Figure 10, for the same region shown Figure 5 with the DM distribution. It shows that the halo density field in CDM is denser than in SIDM, meaning that more halos are found in CDM than in SIDM around massive halos. This result is expected, as self-interactions in SIDM limit the formation of small-scale structures, leading to fewer halos compared to CDM, consistent with the previous findings. The residuals map confirms this trend, with blue regions indicating a lower halo density in SIDM compared to CDM. The resulting genus curves computed on this region are presented in Figure 11. The genus curves computed on the halo density field exhibit similar qualitative features as those computed on the full DM density field, though noisier. For the lowest smoothing scale ($R_{\mathrm{G}} = 0.01\,\mathrm{Mpc}/h$), CDM shows a lower genus than SIDM at high density thresholds, around $\nu_{\mathrm{f}} = 3$ and the relative error illustrated by the solid green line indicates a 1.5σ difference. This shows that more isolated high-density clumps are found in CDM, consistent with previous findings. However, the differences are less pronounced and the error bars are larger due to the coarse sampling of the density fields when using halos instead of DM particles. For all smoothing scales, the genus curves show that SIDM is more connected (higher genus) than CDM at null density threshold, indicating fewer isolated clumps in SIDM and more connected isodensity surfaces. This trend is consistent with the expectation that self-interactions in SIDM suppress small-scale structures. However, we did not detect such a signal when using DM particles to reconstruct the density fields (subsection 2.2). The explanation of this shift is not straightforward to interpret. A possible reason could be that the interpolation of the density field from halo positions does not take into account the mass of each halo, treating all halos equally regardless of their mass. In this approach, dense substructures are just points, whereas in the full DM density field, they contribute significantly to the density. This could lead to a shift in the density contrast where differences between CDM and SIDM are most pronounced. In the future, we will weight the halos by their mass when interpolating the density field, which will provide a more accurate representation of the underlying DM distribution. Overall, the genus topological analysis carried out on the halo density field is still sensitive to DM self-interactions, which is promising to carry out this analysis on galaxy surveys in the future. However, other observational uncertainties or systematics will have to be accounted for (genus in 3D versus in projection on the sky in observations, etc). That said, the drawback of this method is that the number of halos is much lower than the number of DM particles, leading to a much coarser interpolation of the density field. Consequently, the genus curves are much noisier than those obtained from the full DM density field.

Conclusion and discussion

In this work, we have investigated the topological differences in the cosmic web in CDM and SIDM cosmologies using cosmological DM-only simulations. We have employed the genus statistics to quantify the connectivity of isodensity surfaces in the dark matter density field, focusing on regions around halos where self-interactions are expected to have the most significant impact. The genus provides a powerful probe of dark matter self-interactions, through a topological analysis of the connectivity of smoothed density fields. In particular, I show that CDM is significantly clumpier than SIDM at scales lower than $0.1\,\mathrm{Mpc}/h$, as it forms a systematically higher number of isolated high-density regions than found for SIDM. I use the clump abundance ratio $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$ to quantify the differences between the two models. The structure formation is mitigated in SIDM due to the self-interactions that give more kinetic energy to the dark matter particles, resulting in a lower clump abundance. This ratio appears to be significantly larger than unity at smoothing scales lower than $0.05\,\mathrm{Mpc}/h$, when limiting the analysis in halo-centered regions of size a few virial radii. We observe that the discrepancies between the two models become more pronounced with increasing halo mass. The topological properties of the cosmic web are evolving across redshift. The amplitude of the residuals is found to be maximal at low redshift, as SIDM and CDM have had more time to diverge. The genus is a potential probe to constrain the cross-section free parameter of SIDM, as the clump abundance ratio $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$ is sensitive to its value. To connect these results to observations, I have also computed the genus of the density field interpolated directly from the halo and subhalos positions. Since galaxies are expected to form in the center of dark matter halos, the distribution of halos traces the underlying dark matter distribution. The resulting genus curves still exhibit differences between CDM and SIDM, although the signal is noisier due to the lower number of halos compared to DM particles. The genus calculated on the halo density field could be directly compared to observations, as the genus has already been measured in galaxy surveys. Future improvements of this analysis will be include weighting the halos by their mass when interpolating the density field, which will provide a more accurate representation of the underlying DM distribution. The investigation of observational systematics will also be necessary to apply this method to real data. These improvements will be addressed in future work, and will be presented in a forthcoming paper (in preparation). Alternatively, the present topological analysis could be applied to compare other DM theories where galactic cores are formed (such as warm or fuzzy DM) to CDM. One would expect from the gravitational heating induced by such alternative theories to produce similar effects on the genus of the isodensity surfaces. However, distinguishing alternative DM theories between each other may be more challenging. Overall, this work shows that topological analyses in galaxy surveys offer promising perspectives to constrain the nature and self-interactions of dark matter, leveraging DM-dominated gravitational dynamics while reducing reliance on complex baryonic modeling.

As a personal remark, I found this project very enriching as it allowed me to explore a novel approach to studying dark matter properties through topological analysis. The interdisciplinary nature of the work, combining cosmological simulations, topology, and dark matter physics, provided a unique learning experience. I appreciated the opportunity to develop computational skills and gain insights into the large-scale structure of the Universe. Overall, this project has deepened my understanding of both cosmology and data analysis techniques, which will be valuable for my future research endeavors.

Conclusion (french version)

Dans ce travail, nous avons étudié les différences topologiques de la toile cosmique dans des cosmologies CDM et SIDM à l'aide de simulations cosmologiques DM-only. Nous avons employé la statistique du genus pour quantifier la connectivité des isosurfaces de densité du champ de matière noire, en nous concentrant sur des régions autour des halos où les auto-interactions sont supposées avoir l'effet le plus marqué. Le genus constitue une sonde puissante des auto-interactions de la DM via une analyse topologique de la connectivité de champs de densité lissés. En particulier, nous montrons que la CDM est significativement plus grumeleuse que la SIDM à des échelles inférieures à 0.1 Mpc$/h$, formant un nombre systématiquement plus élevé de régions de haute densité isolées que la SIDM. Nous utilisons le rapport d'abondance de grumeaux $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$ pour quantifier ces différences. La formation de structures est atténuée en SIDM du fait des auto-interactions qui apportent davantage d'énergie cinétique aux particules de DM, entraînant une abondance de grumeaux plus faible. Ce ratio est significativement supérieur à l'unité pour des échelles de lissage inférieures à 0.05 Mpc$/h$, lorsque l'analyse est limitée à des régions centrées sur les halos de quelques rayons viriels. Nous observons que les écarts entre les deux modèles s'accentuent avec la masse du halo. Les propriétés topologiques de la toile cosmique évoluent avec le redshift: l'amplitude des résidus est maximale à bas redshift, les modèles SIDM et CDM ayant eu plus de temps pour diverger dans leur comportement. Le genus et l'abondance de grumeaux constituent une sonde potentielle pour contraindre la section efficace d'auto-interaction de la SIDM, le ratio $A_{\mathrm{C,CDM}}/A_{\mathrm{C,SIDM}}$ étant sensible à sa valeur.

Afin de relier ces résultats aux observations, nous avons également calculé le genus du champ de densité interpolé directement à partir des positions des halos et sous-halos. Puisque les galaxies se forment au centre des halos de DM, la distribution des halos trace la distribution sous-jacente de DM. Les courbes de genus obtenues présentent encore des différences entre CDM et SIDM, bien que le signal soit plus bruité en raison du nombre de halos plus faible que celui des particules de DM. Le genus calculé sur le champ de densité des halos peut être comparé directement aux observations, cette statistique ayant déjà été mesurée dans des relevés de galaxies. Des améliorations futures consisteront à pondérer les halos par leur masse lors de l'interpolation du champ de densité, afin de mieux représenter la distribution sous-jacente de DM, et à étudier les systématiques observationnelles pour appliquer la méthode aux données réelles. Ces améliorations seront traitées dans des travaux futurs (en préparation). Alternativement, l'analyse topologique présentée ici pourrait être appliquée pour comparer d'autres théories de DM produisant des cœurs galactiques (DM tiède ou floue) à la CDM; toutefois, distinguer ces théories entre elles pourrait s'avérer plus délicat.

Globalement, ce travail montre que l'analyse topologique de la structure à grande échelle dans les relevés de galaxies offre une voie prometteuse pour contraindre la nature et les auto-interactions de la matière noire, en s'appuyant sur une dynamique dominée par la gravité de la DM et en limitant la dépendance à une modélisation baryonique complexe.

References

Adler, R. J. *The Geometry of Random Fields.* Jan. 1, 1981.

Alam, S. et al. "Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory". In: *Physical Review D* 103.8 (Apr. 28, 2021).

Appleby, S. et al. "Cosmological Parameter Estimation Using the Genus Amplitude—Application to Mock Galaxy Catalogs". In: *The Astrophysical Journal* 853.1 (Jan. 20, 2018).

Armijo, J. et al. "Cosmological constraints using Minkowski functionals from the first year data of the Hyper Suprime-Cam". In: *Monthly Notices of the Royal Astronomical Society* 537.4 (Feb. 17, 2025).

Barkana, R. et al. "In the beginning: the first sources of light and the reionization of the universe". In: *Physics Reports* 349.2 (July 2001).

Beutler, F. et al. "The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: anisotropic galaxy clustering in Fourier space". In: *Monthly Notices of the Royal Astronomical Society* 466.2 (Apr. 11, 2017).

Blok, W. J. G. d. "The Core-Cusp Problem". In: *Advances in Astronomy* 2010.1 (Jan. 2010).

Boldrini, P. "The cusp-core problem in gas-poor dwarf spheroidal galaxies". In: *Galaxies* 10.1 (Dec. 30, 2021).

Bond, J. R. et al. "How filaments of galaxies are woven into the cosmic web". In: *Nature* 380.6575 (Apr. 1996).

Boylan-Kolchin, M. et al. "Too big to fail? The puzzling darkness of massive Milky Way subhaloes". In: *Monthly Notices of the Royal Astronomical Society: Letters* 415.1 (July 1, 2011).

Bullock, J. S. et al. "Small-Scale Challenges to the \$Λ\$CDM Paradigm". In: *Annual Review of Astronomy and Astrophysics* 55.1 (Aug. 18, 2017).

Carlson, E. D. et al. "Self-interacting dark matter". In: *The Astrophysical Journal* 398 (Oct. 1992).

Correa, C. A. et al. "TangoSIDM: tantalizing models of self-interacting dark matter". In: *Monthly Notices of the Royal Astronomical Society* 517.2 (Oct. 21, 2022).

Dave, R. et al. "Halo Properties in Cosmological Simulations of Self-interacting Cold Dark Matter". In: *The Astrophysical Journal* 547.2 (Feb. 2001).

Davis, M. et al. "A survey of galaxy redshifts. II - The large scale space distribution". In: *The Astrophysical Journal* 253 (Feb. 1982).

De Blok, W. J. G. et al. "High-Resolution Rotation Curves of Low Surface Brightness Galaxies. II. Mass Models". In: *The Astronomical Journal* 122.5 (Nov. 2001).

De Lapparent, V. et al. "A slice of the universe". In: *The Astrophysical Journal* 302 (Mar. 1986).

Delaunay, B. "Sur la sphère vide". In: *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na* 1934 (1934).

Despali, G. et al. "Introducing the AIDA-TNG project: Galaxy formation in alternative dark matter models". In: *Astronomy & Astrophysics* 697 (May 2025).

Doroshkevich, A. G. "Spatial structure of perturbations and origin of galactic rotation in fluctuation theory". In: *Astrophysics* 6.4 (1973).

Eisenstein, D. J. et al. "Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies". In: *The Astrophysical Journal* 633.2 (Nov. 10, 2005).

Elbert, O. D. et al. "Core formation in dwarf haloes with self-interacting dark matter: no fine-tuning necessary". In: *Monthly Notices of the Royal Astronomical Society* 453.1 (Oct. 11, 2015).

Euclid et al. "Euclid. I. Overview of the Euclid mission". In: *Astronomy & Astrophysics* 697 (May 2025).

Fischer, M. S. et al. "N -body simulations of dark matter with frequent self-interactions". In: *Monthly Notices of the Royal Astronomical Society* 505.1 (May 29, 2021).

– "Cosmological simulations with rare and frequent dark matter self-interactions". In: *Monthly Notices of the Royal Astronomical Society* 516.2 (Sept. 7, 2022).

Geller, M. J. et al. "Mapping the Universe". In: *Science* 246.4932 (Nov. 17, 1989).

Gott III, J. R. et al. "A Quantitative Approach to the Topology of Large-Scale Structure". In: *The Astrophysical Journal* 319 (Aug. 1, 1987).

Groth, F. et al. "The cosmological simulation code OpenGadget3 – implementation of meshless finite mass". In: *Monthly Notices of the Royal Astronomical Society* 526.1 (Sept. 21, 2023).

Hamilton, A. J. S. et al. "The topology of the large-scale structure of the universe". In: *The Astrophysical Journal* 309 (Oct. 1986).

Kaplinghat, M. et al. "Too big to fail in light of Gaia". In: *Monthly Notices of the Royal Astronomical Society* 490.1 (Nov. 21, 2019).

Klypin, A. et al. "Where Are the Missing Galactic Satellites?" In: *The Astrophysical Journal* 522.1 (Sept. 1999).

Liu, W. et al. "Cosmological constraints from the Minkowski functionals of the BOSS CMASS galaxy sample". In: *Journal of Cosmology and Astroparticle Physics* 2025.5 (May 1, 2025).

Mecke, K. R. et al. *Robust Morphological Measures for Large-Scale Structure in the Universe.* Dec. 14, 1993.

Minkowski, H. "Volumen und Oberflche". In: *Mathematische Annalen* 57.4 (Dec. 1903).

Mo, H. et al. *Galaxy Formation and Evolution.* 1st ed. Cambridge University Press, May 20, 2010.

Moore, B. "Evidence against dissipation-less dark matter from observations of galaxy haloes". In: *Nature* 370.6491 (Aug. 1994).

Navarro, J. F. et al. "A Universal Density Profile from Hierarchical Clustering". In: *The Astrophysical Journal* 490.2 (Dec. 1997).

Park, C., Y.-Y. Choi, et al. "Topology Analysis of the Sloan Digital Sky Survey: I. Scale and Luminosity Dependences". In: *The Astrophysical Journal* 633.1 (Nov. 2005).

Park, C., J. R. Gott Iii, et al. "Topology of the Galaxy Distribution in the Hubble Deep Fields". In: *The Astrophysical Journal* 553.1 (May 20, 2001).

Park, C., J. Kim, et al. "Effects of Gravitational Evolution, Biasing, and Redshift Space Distortion on Topology". In: *The Astrophysical Journal* 633.1 (Nov. 2005).

Planck Collaboration et al. "*Planck* 2018 results: VI. Cosmological parameters". In: *Astronomy & Astrophysics* 641 (Sept. 2020).

Pogosyan, D. et al. *Cosmic Web: Origin and Observables.* Oct. 5, 1998.

Rees, M. J. et al. "Cooling, dynamics and fragmentation of massive gas clouds: clues to the masses and radii of galaxies and clusters". In: *Monthly Notices of the Royal Astronomical Society* 179.4 (Aug. 1, 1977).

Riess, A. G. et al. "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". In: *The Astronomical Journal* 116.3 (Sept. 1998).

Rocha, M. et al. "Cosmological simulations with self-interacting dark matter – I. Constant-density cores and substructure". In: *Monthly Notices of the Royal Astronomical Society* 430.1 (Mar. 21, 2013).

Sousbie, T. "The persistent cosmic web and its filamentary structure - I. Theory and implementation: Persistent cosmic web - I: Theory and implementation". In: *Monthly Notices of the Royal Astronomical Society* 414.1 (June 11, 2011).

Spergel, D. N. et al. "Observational evidence for self-interacting cold dark matter". In: *Physical Review Letters* 84.17 (Apr. 24, 2000).

Springel, V. "The cosmological simulation code gadget-2". In: *Monthly Notices of the Royal Astronomical Society* 364.4 (Dec. 2005).

Springel, V. et al. "GADGET: a code for collisionless and gasdynamical cosmological simulations". In: *New Astronomy* 6.2 (Apr. 2001).

Tulin, S. et al. "Dark Matter Self-interactions and Small Scale Structure". In: *Physics Reports* 730 (Feb. 2018).

Watts, A. L. et al. "Large-scale structure topology in non-standard cosmologies: impact of dark sector physics". In: *Monthly Notices of the Royal Astronomical Society* 468.1 (June 11, 2017).

White, S. D. M. and M. J. Rees. "Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering". In: *Monthly Notices of the Royal Astronomical Society* 183.3 (July 1, 1978).

White, S. D. M. and C. S. Frenk. "Galaxy formation through hierarchical clustering". In: *The Astrophysical Journal* 379 (Sept. 1991).

Zavala, J. et al. "Constraining self-interacting dark matter with the Milky Way's dwarf spheroidals". In: *Monthly Notices of the Royal Astronomical Society: Letters* 431.1 (May 1, 2013).

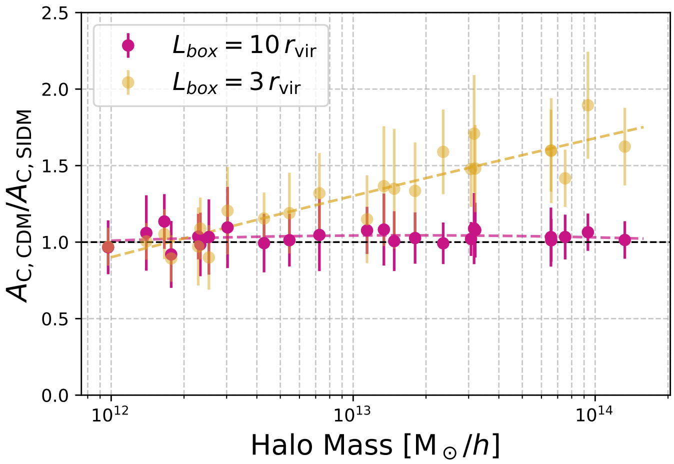


Figure 12: Effect of the region size on clump abundance ratio. Each point corresponds to the ratio of clump abundance computed from the genus of smoothed density field in halo-centered box of size $L_{\mathrm{box}} = 3\,r_{\mathrm{vir}}$ (yellow, same points as in Figure 7) and $L_{\mathrm{box}} = 10\,r_{\mathrm{vir}}$ (magenta). The smoothing scale used for the genus calculation is $R_{\mathrm{G}} = 0.05\,\mathrm{Mpc}/h$. The error bars represent the uncertainty on the ratio, propagated from the dispersion of the genus calculated in the 8 sub-boxes. The color dashed lines represent polynomial fits of degree two.

A Impact of the box size on the clump abundance

The genus computed on the halo-centered sub-boxes of size $L_{\mathrm{box}} = 3\,r_{\mathrm{vir}}$ reveals significant differences between CDM and SIDM at small smoothing scales. However, we have also seen that these differences vanish when considering the full $48\,\mathrm{Mpc}/h$ box. To test the robustness of this result, I have repeated the analysis using larger sub-boxes of size $L_{\mathrm{box}} = 10\,r_{\mathrm{vir}}$. The clump abundance ratio obtained with these larger boxes is shown in Figure 12. I find that the clump abundance ratio approaches unity for all halo masses when using the larger boxes, indicating that the differences between CDM and SIDM are no longer detectable. This suggests that the choice of box size is crucial for resolving the topological differences induced by dark matter self-interactions. Larger boxes include more low-density regions, such as filaments and voids, which dilute the impact of self-interactions on small-scale structure that happens in dense regions. To check the contribution of all these components, I have done the calculation once more, but using a spherical null mask on the halo. Separating the genus calculation this way avoids mixing effects that could be different in clusters and filaments. The result shows similar results as obtained without the mask. This confirms that the inability to differentiate CDM and SIDM at large box size is due to the inclusion of less dense regions, where the topology is more similar. It also confirms that the differences between CDM and SIDM are expected to arise in the densest regions. Therefore, to effectively probe SIDM effects using genus statistics, it is essential to focus on smaller regions centered around massive halos.