Analytic models of large-scale structure: modeling CMB lensing cross-correlations with perturbation theory

Martin White
with Chirag Modi & Zvonimir Vlah

UC Berkeley/LBNL

Paris, Dec 2017

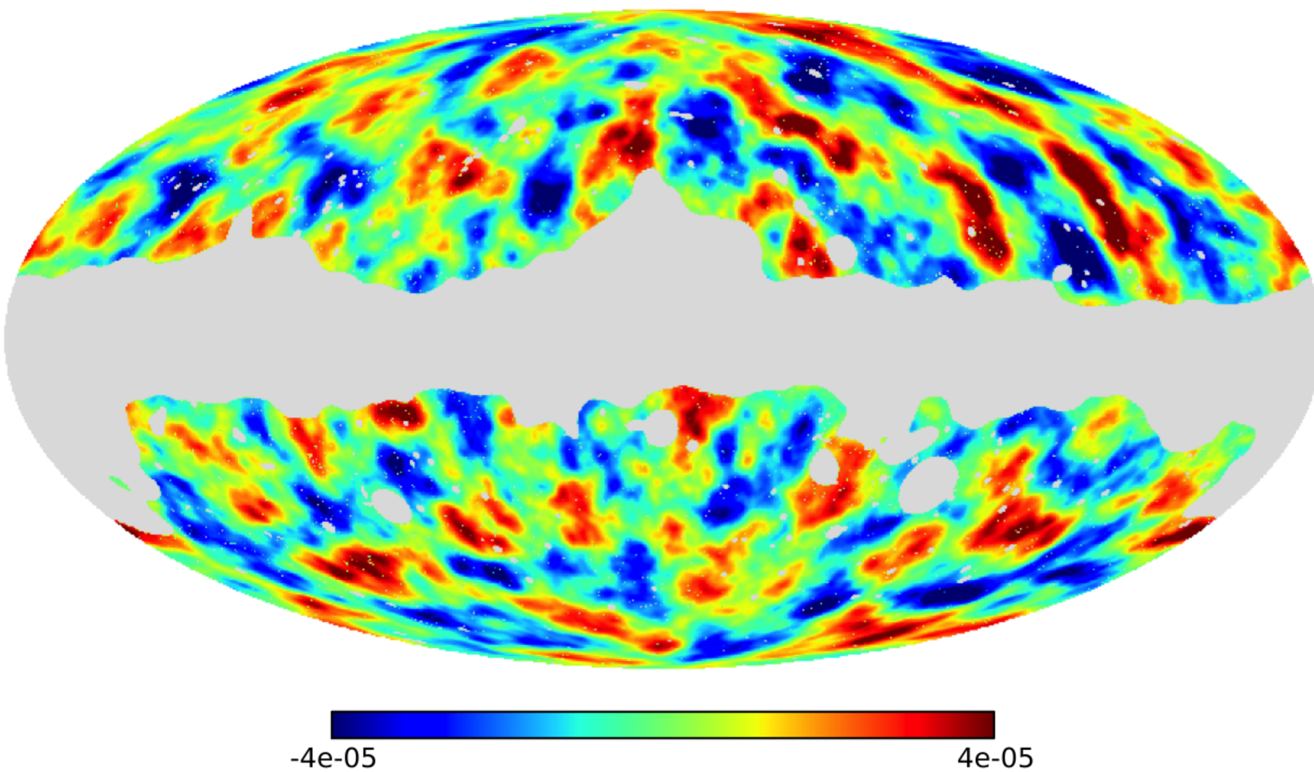
Perturbation theory

- Over the last several decades, cosmological perturbation theory has developed steadily.
- CMB anisotropies are "everyone's favorite", linear, cosmological perturbation theory calculation ...
- ... a field to which Joe Silk made numerous foundational contributions and in which he trained many of the leading practitioners.
- Arguably, CMB anisotropies form the gold standard for cosmological inference and cosmological knowledge.
- A well controlled, analytic calculation which can be compared straightforwardly to observations.

Perturbation theory

- As we move to lower redshifts we need to start worrying about structure going non-linear and about the relation between the matter field and what we see (bias).
- As surveys get larger and more powerful more of the modes we measure <u>well</u> are "quasi-linear" $\Rightarrow$ analytic models.
- The last decade has seen an explosion of work on perturbative models of large-scale structure – at Berkeley we have been developing analytic models based on Lagrangian perturbation theory.
- Our original goal was baryon acoustic oscillations (BAO) and redshift-space distortions (RSD). But I will argue these tools (and others like them) are "perfect" for the coming world of survey cross-correlations...

Planck lensing map



Planck Collaboration (2015)

Coming of age

Planck was definitely not the first experiment to

- to measure lensing,
- ... by large scale structure,
- ... of the CMB

however it was the first experiment to measure CMB lensing by large scale structure over a significant fraction of the sky and with enough signal to noise that it provided a sharp test of the theory and could drive fits.

In some sense *Planck* was a "coming of age" for CMB lensing, and a taste of things to come – much of the science from future CMB surveys will come from lensing.

The landscape

A natural "by-product" of next generation CMB experiments to constrain primordial gravitational waves is high fidelity CMB lensing maps.

- CMB lensing is sensitive to the matter field and to the space-space metric perturbation, over a broad redshift range.
- CMB lensing has radically different systematics than cosmic shear (and measures[†] κ, not γ).
- CMB redshift is very well known (but can't change it)!
- CMB lensing surveys tend to have large $f_{\rm sky}$, but relatively poor resolution.
- The lensing kernel peaks at $z \sim 2-3$ and has power to $z \gg 1$, where galaxy lensing becomes increasingly difficult.
- The CMB is behind "everything" ... but projection is a big issue.

Optical surveys

We will also have major new imaging and spectroscopic facilities ...

- Dark Energy Survey (DES)
- DECam Legacy Survey (DECaLS)
- Dark Energy Spectroscopic Instrument (DESI)
- Subaru Hyper Suprime-Cam (HSC)
- Large Synoptic Survey Telescope (LSST)
- Euclid
- Wide-Field Infrared Survey Telescope (WFIRST)

These facilities can map large areas of sky to unprecedented depths!

The opportunity

A new generation of deep imaging surveys and CMB experiments offers the possibility of using cross-correlations to

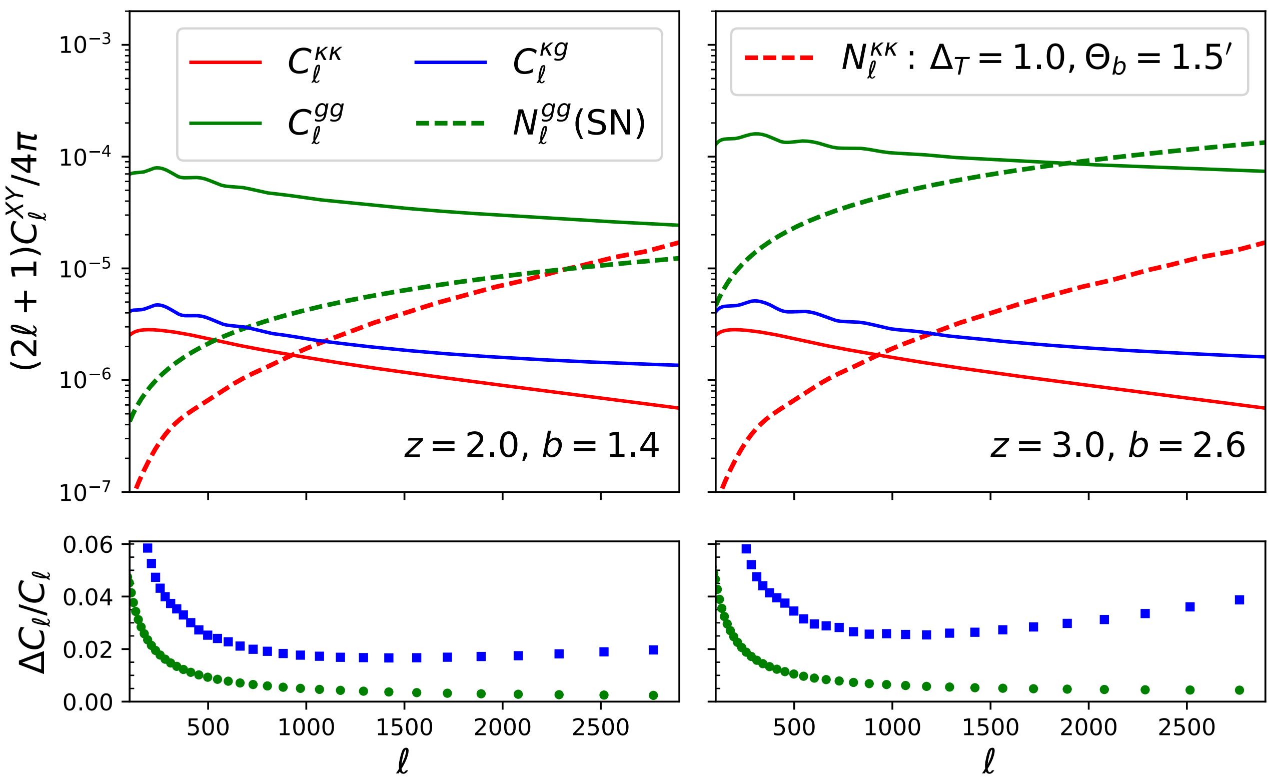
- constrain the early Universe
- test General Relativity
- probe the galaxy-halo connection
- measure the growth of large-scale structure

The combination can be more than the sum of its parts!
In particular we can use the optical survey to isolate the κ contribution from narrow z slices, increase S/N and downweight systematics.

Improvements in data require concurrent improvements in the theoretical modeling in order to reap the promised science.

What is the right framework for analyzing such data?

The future is bright



Example: Measuring $P_{mm}(k, z)$

- A proper accounting of the growth of large scale structure through time is one of the main goals of observational cosmology – key quantity is $P_{mm}(k, z)$.
- Schematically we can measure $P_{mm}(k, z)$ by picking galaxies at z and

$$P_{mm}(k) \sim \frac{[bP_{mm}(k)]^2}{b^2 P_{mm}(k)} \sim \frac{[P_{mh}(k)]^2}{P_{hh}(k)} \sim \frac{\left[C^{\kappa g}_{\ell=k\chi}\right]^2}{C^{gg}_{\ell=k\chi}}$$

- Operationally we perform a joint fit to the combined data set.
  - With only the auto-spectrum there is a strong degeneracy between the amplitude (σ_8) and the bias parameters (b).
  - However the matter-halo cross-spectrum has a different dependence on these parameters and this allows us to break the degeneracy and measure σ_8 (and b).
- Need a model for the auto- and cross-spectra of biased tracers.

Need a model

Thus we need a model which can predict the auto- and cross-spectra of biased tracers at large and intermediate scales.

- Even though we are at high z and "large" scales it turns out that linear perturbation theory isn't good enough.
- Need to include non-linear corrections – and as soon as you do that you need to worry about scale-dependent bias, stochasticity and a whole host of other evils.

“Standard” model

- The most widely used model to date is based on the HaloFit fitting function for $P_{mm}(k)$ (auto-magically computed by CAMB and CLASS).
- Most analyses assume scale-independent bias (but this is barely sufficient even “now”).
- One extension, motivated by peaks theory, is to use $b(k) = b^E_{10} + b^E_{11} k^2$.
- We will find we need to augment this with a phenomenological k term

$$P_{mh}(k) = \left[b^E_{10} + b^E_{1\frac{1}{2}} k + b^E_{11} k^2 \right] P_{HF}(k)$$

$$P_{hh}(k) = \left[b^E_{10} + b^E_{1\frac{1}{2}} k + b^E_{11} k^2 \right]^2 P_{HF}(k)$$

Note the (necessary) assumption that $b_{hh} = b_{mh}$!

CLEFT model

(Large scales, high z, it sounds like a job for ...)

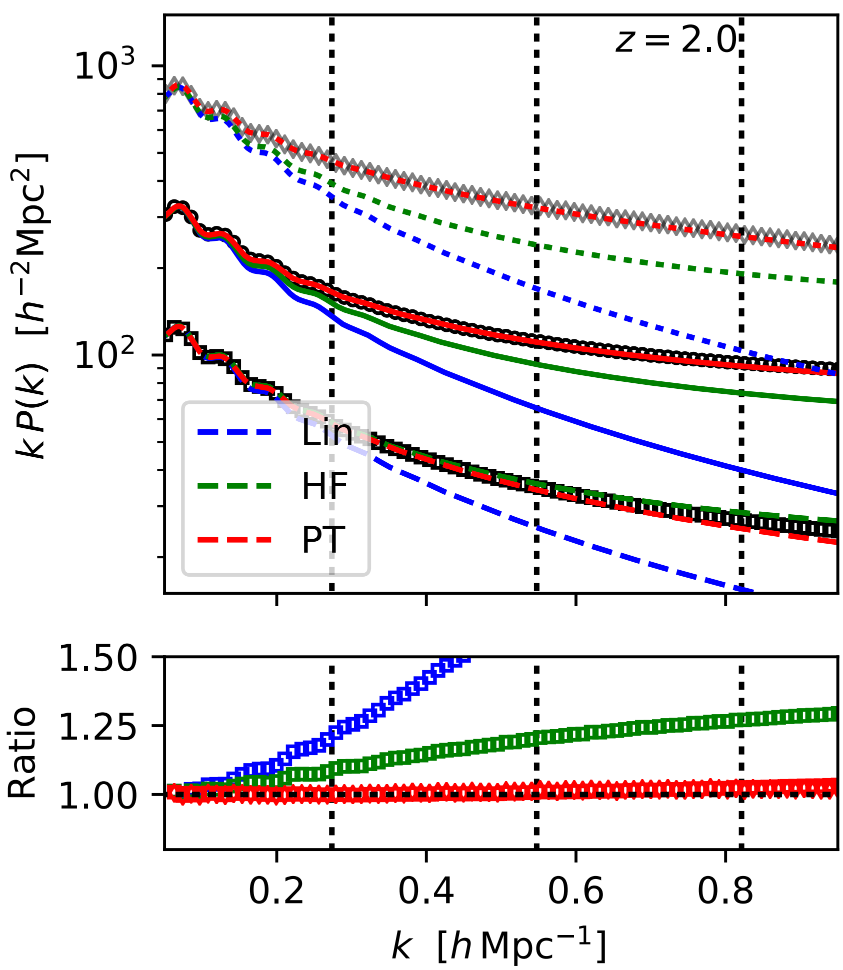
The Lagrangian PT framework we have been developing for many years naturally handles auto- and cross-correlations in real and redshift space for Fourier or configuration space statistics. For example:

$$P_{mg}(k) = \left(1 - \frac{\alpha\, k^2}{2}\right) P_Z + P_{1-\mathrm{loop}} + \frac{b_1}{2} P_{b_1} + \frac{b_2}{2} P_{b_2} + \cdots$$

where P_Z and $P_{1-\mathrm{loop}}$ are the Zeldovich and 1-loop matter terms, the b_i are Lagrangian bias parameters for the biased tracer, and α is a free parameter which accounts for k^2 bias and small-scale physics not modeled by PT.

Extend the highly successful linear perturbation theory analysis of primary CMB anisotropies which has proven so impactful!

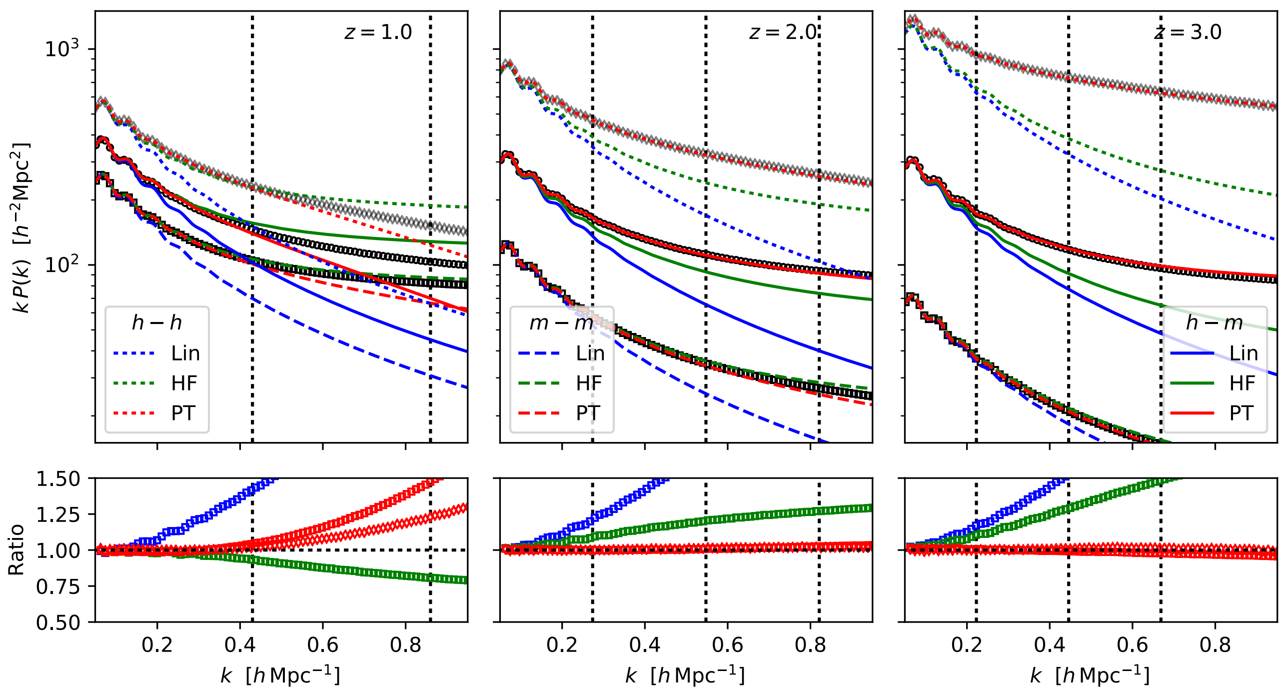
Comparison with N-body



Let's look at the ingredients going into the prediction of C_ℓ^{XY}, for three cases:

- Linear theory, constant bias.
- HaloFit, constant bias (for now!).
- PT, $b_1 - b_2$.

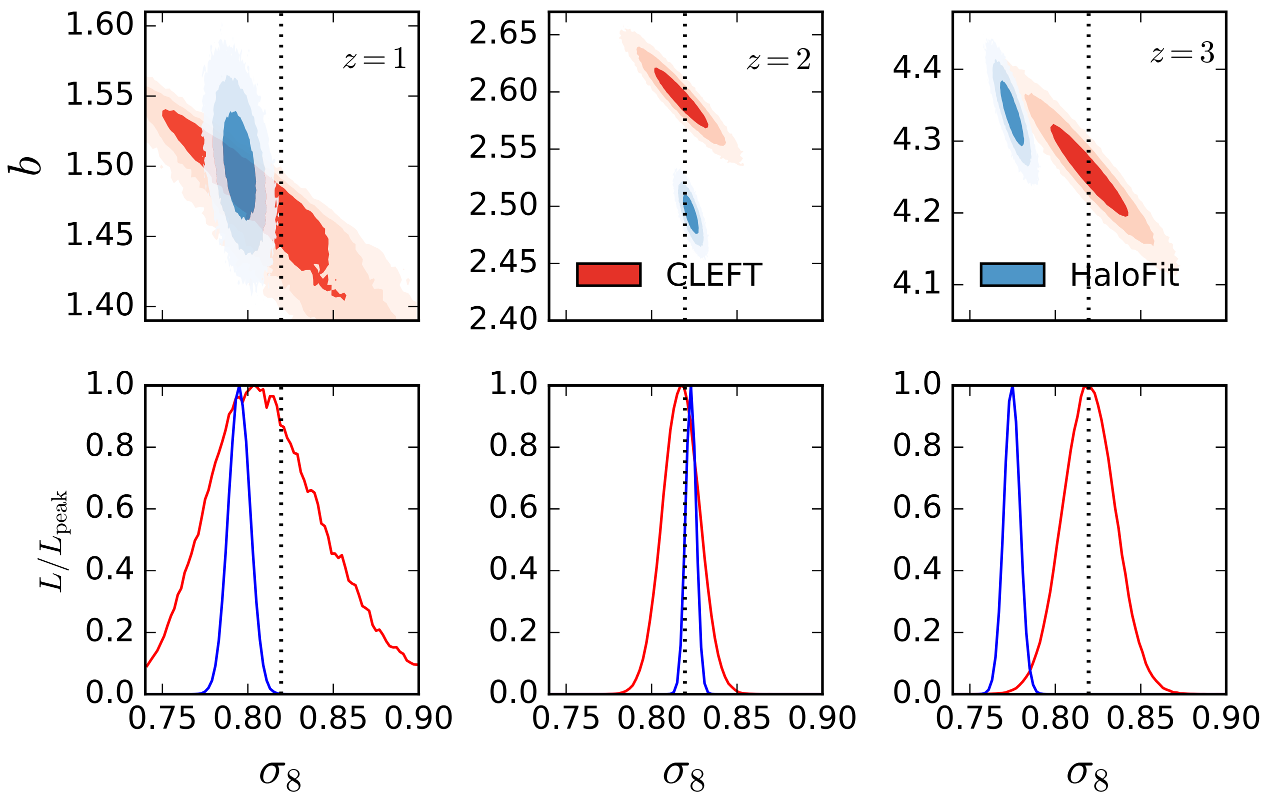
Comparison with N-body



Model fit

- Consider a future experiment, motivated by LSST and CMB-S4 but it could be a number of things.
- Imagine cross-correlating the CMB lensing map with the (gold sample) galaxies in a slice $\Delta z = 0.5$ at $z = 1$, 2 and 3.
  - $i_{\rm lim} = 25.3$.
  - $\theta_b = 1.5'$, $\Delta_T = 1\,\mu$K-arcmin.
- Compare two 'models':
  - HALOFIT with $b(k) = b_{10}^E + b_{1\frac{1}{2}}^E\,k + b_{11}^E\,k^2$.
  - Perturbation theory with b_1, b_2 (and α_i).
- Concentrate on just measuring an amplitude of matter clustering, σ_8.
- Jointly fit $C_\ell^{\kappa g}$ and C_ℓ^{gg} ...

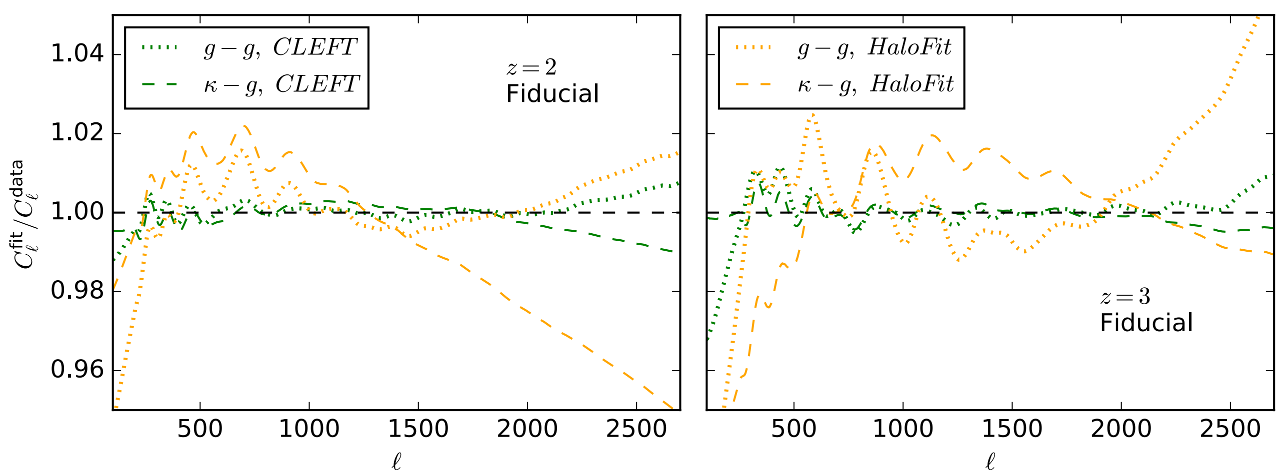
Model fit



(b means something different in each theory)

Model fit

The likelihoods hide a lot of information about how the fit is performing. If we look at the best fit models:

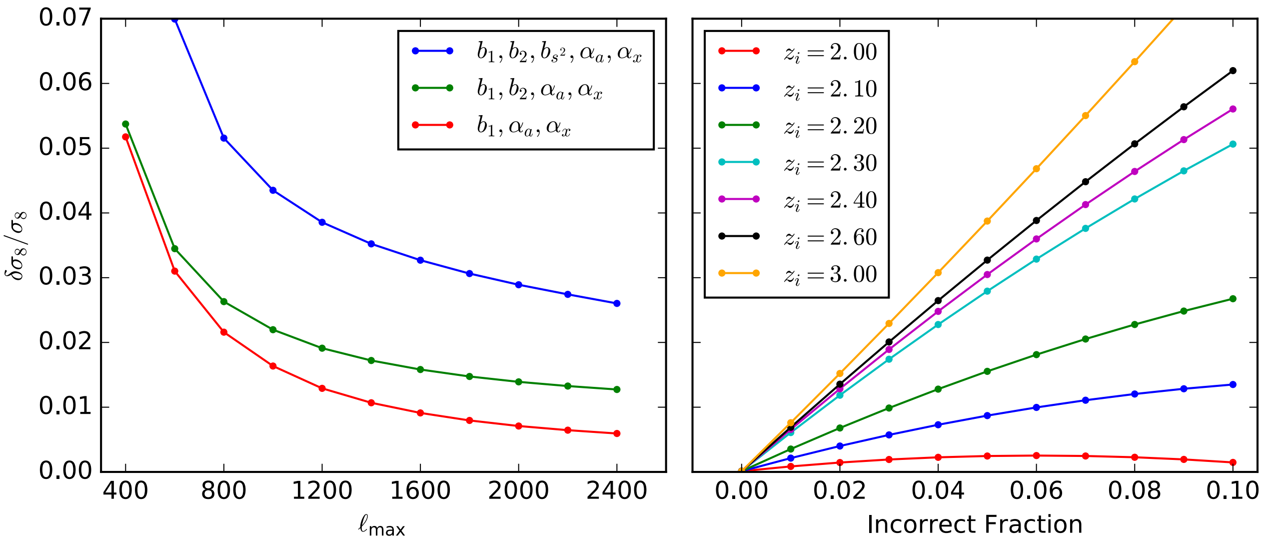


Model fit

- Part of the issue with HaloFit is with the fit to P_{mm}, much of it is with the $b(k)$ assumption.
- At high z, modeling bias is *at least* as important as modeling non-linear structure formation.
- In the EFT language: k_{NL} shifts to higher k at higher z, but the scale associated with halo formation (the Lagrangian radius) remains constant for fixed halo mass.
- In general there is a "sweet spot", where b is not *too* scale dependent but non-linearity is not *too* pronounced.
- How $b_{ij}(k)$ depends upon complex tracer selection is unknown.

Knowing dN/dz

We can use the Fisher forecasting formalism to investigate where the signal is coming from, degeneracies, and biases.



Can work at relatively low ℓ, but need to know dN/dz well.

Future directions

- There are good reasons to work in configuration space, not Fourier space ... (with compensated filters?)
- Go to 2-loop, so we can work to lower z and higher ℓ.
- Add $m_\nu > 0$ or MG, v_{bc}, ...
- More explicit modeling of lensing.
- Inclusion of baryonic effects using EFT techniques.
- Look at non-Gaussianity from inflation (low ℓ).
- Combining 3D surveys with 2D surveys. More modes to a fixed ℓ, but more difficult to model.
- Clean low z. Can model $C_\ell^{\kappa\kappa}(> z_{\rm min})$ and the decorrelations using PT.
- Simultaneously fitting dN/dz and σ_8 using clustering redshifts.
- Multi-tracer techniques (Schmitfull & Seljak 2017).

Conclusions

- We are on the cusp of a dramatic increase in the quality and quantity of both CMB and imaging data.
- The combination of CMB and galaxy data can be more than the sum of its parts.
- As always, better data requires "better" modeling.
  - With primary anisotropies, linear theory is 99% of the story.
  - At lower redshift this is no longer the case.
- We need to model both non-linear matter clustering *and* bias.
- Fitting functions for P_{mm} are good to $\mathcal{O}(5-15\%)$, but the error bars will be smaller than this.
- Once b is not a constant, $b_{hh} \neq b_{mh}$.
- The combination of high redshift and "large" scales makes this an attractive problem for analytic/perturbative attack.
- Generalizes to other high-z probes, in real- and redshift-space (e.g. LIM).

Thank you Joe

... and ...

Happy Birthday!