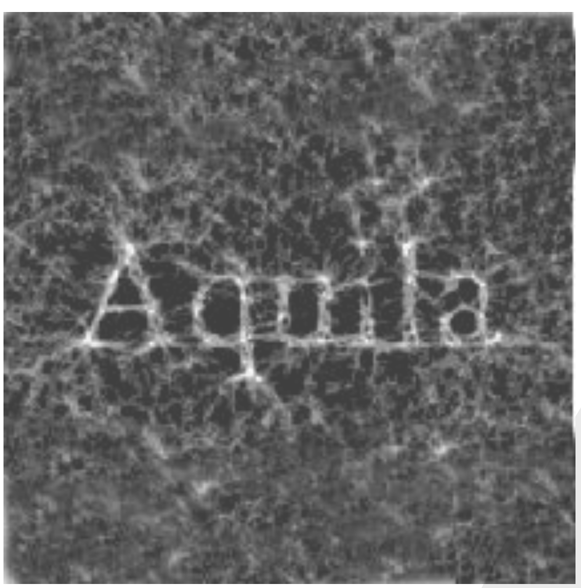
$$\mathbf{x} = \mathbf{x}_{\rm fl}(\tau)$$

Going forward: Bayesian Inference from Galaxy Clustering



aquila-consortium.org

Fabian Schmidt
MPA

European Research Council
Established by the European Commission

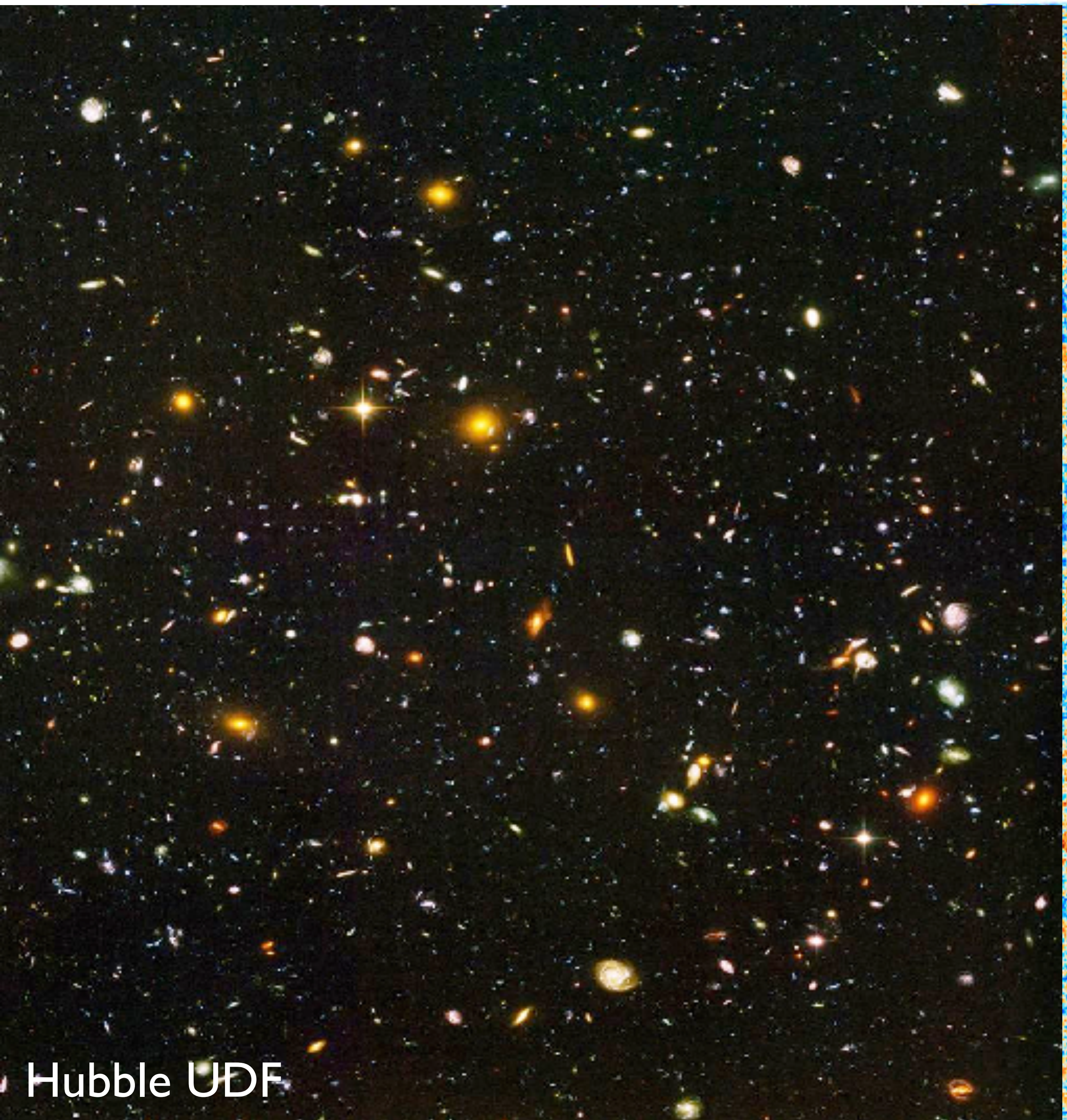
with

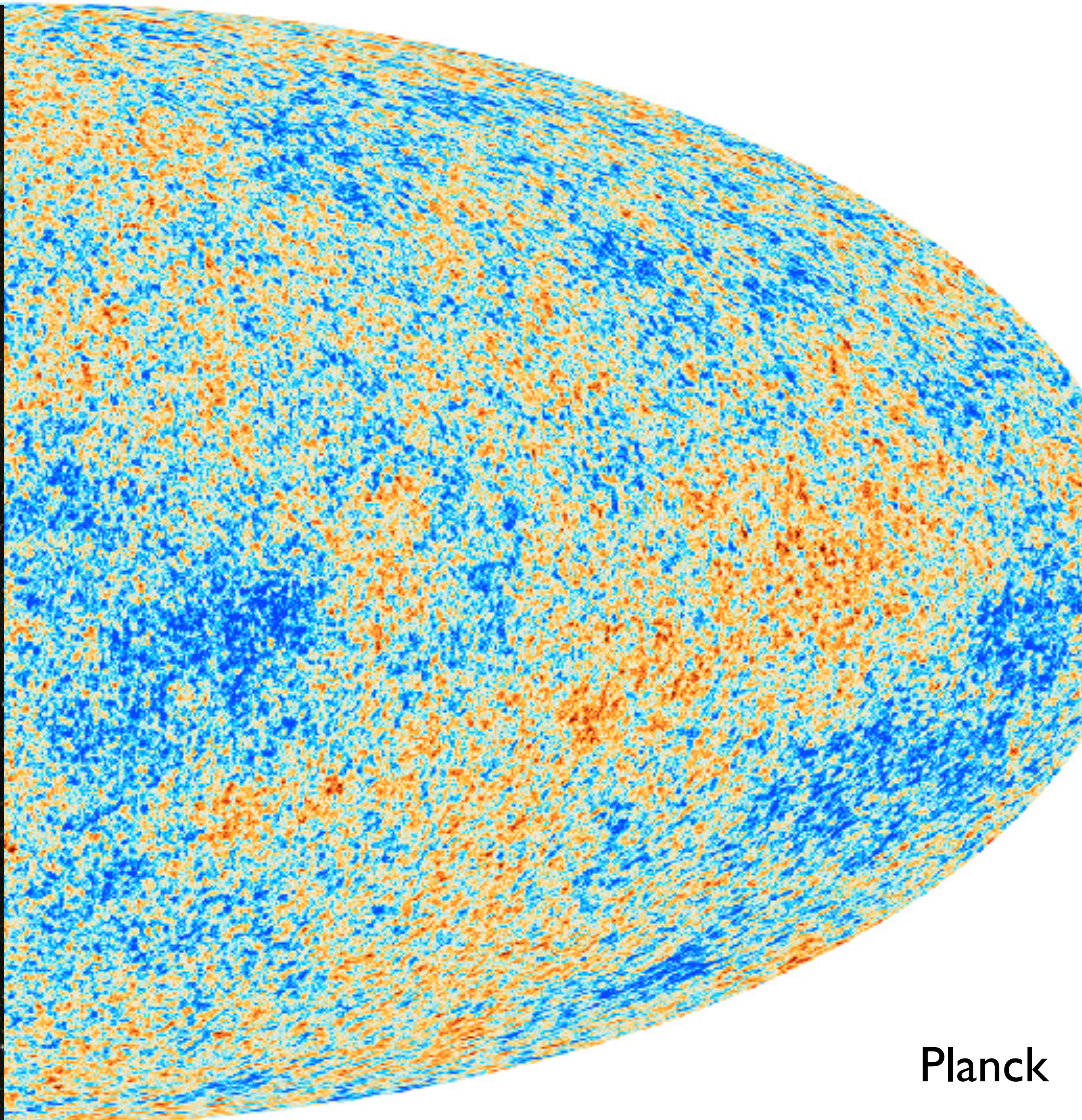
Giovanni Cabass, Franz Elsner, Jens Jasche,
Guilhem Lavaux, Minh Nguyen, Martin Reinecke

$$\mathbf{q} = \mathbf{x}_{\rm fl}(0)$$

Cambridge/LMU Workshop Jan 5, 2021

Galaxy clustering: unlike the CMB, every data point is nonlinear!





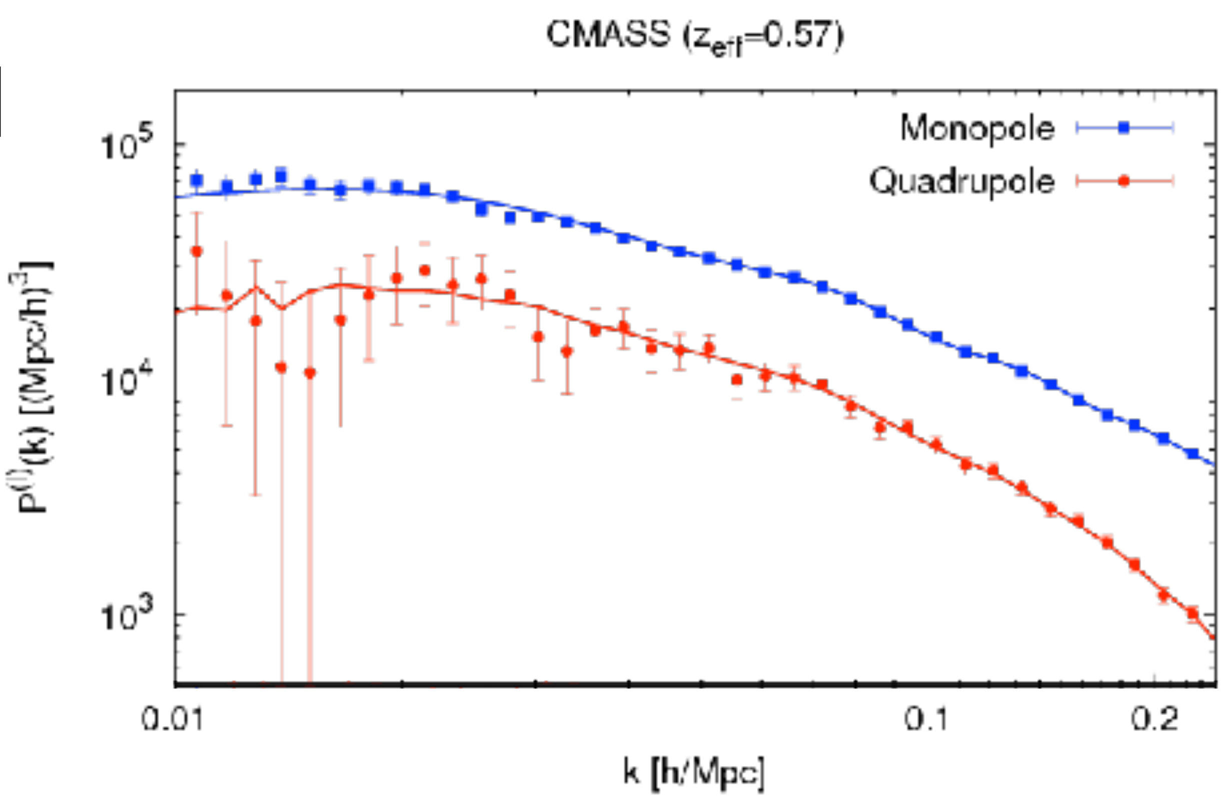
Outline

- Two goals of the talk:
  - Argue that there is *much more (trustable) information in galaxy clustering* than what we are using so far
  - Show that we can deal with complexities of galaxies *rigorously* on large scales

How do we compare theory with data?

- If galaxy density field was Gaussian,
  - i.e. PDF of δ(x) is multivariate Gaussian, with diagonal covariance in Fourier space
- Then all the information would be contained in the power spectrum

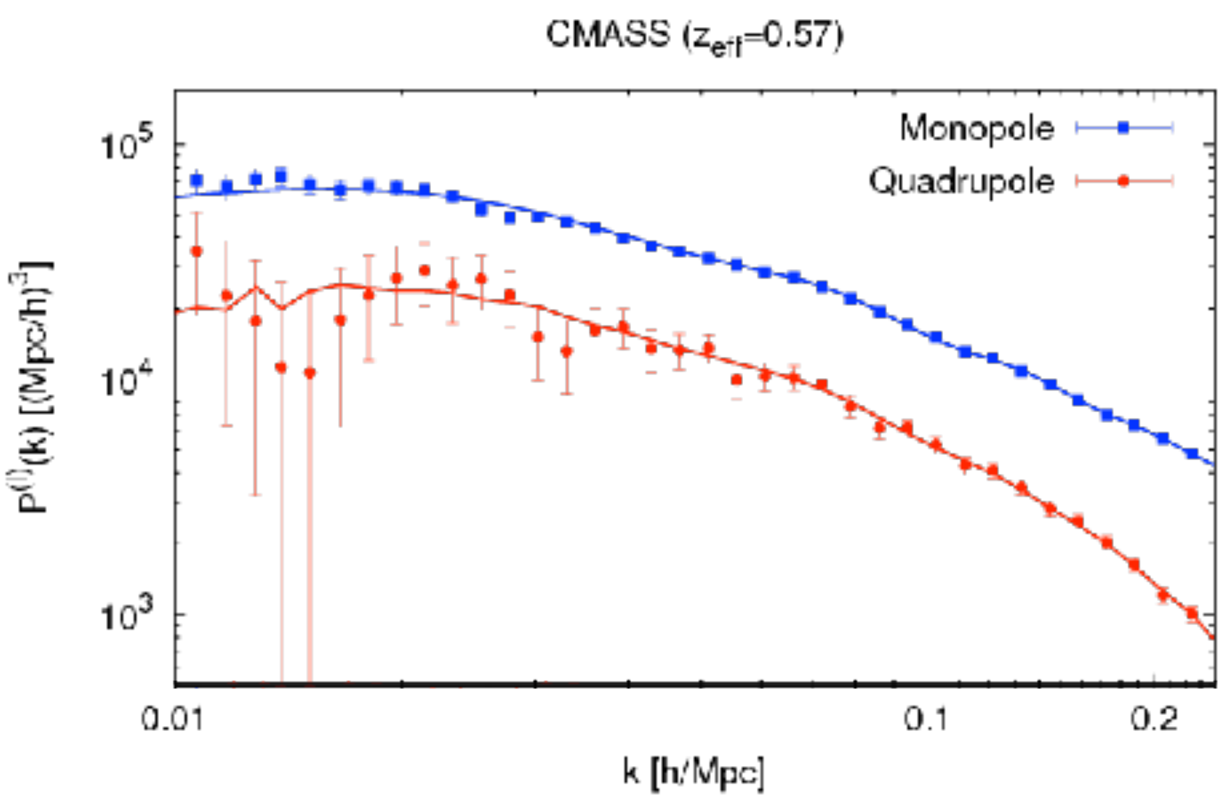
$$\langle \delta(\boldsymbol{k})\delta^*(\boldsymbol{k}')\rangle = (2\pi)^3\delta_D(\boldsymbol{k}-\boldsymbol{k}')\,P(k)$$



Gil-Marin et al, 2016

How do we compare theory with data?

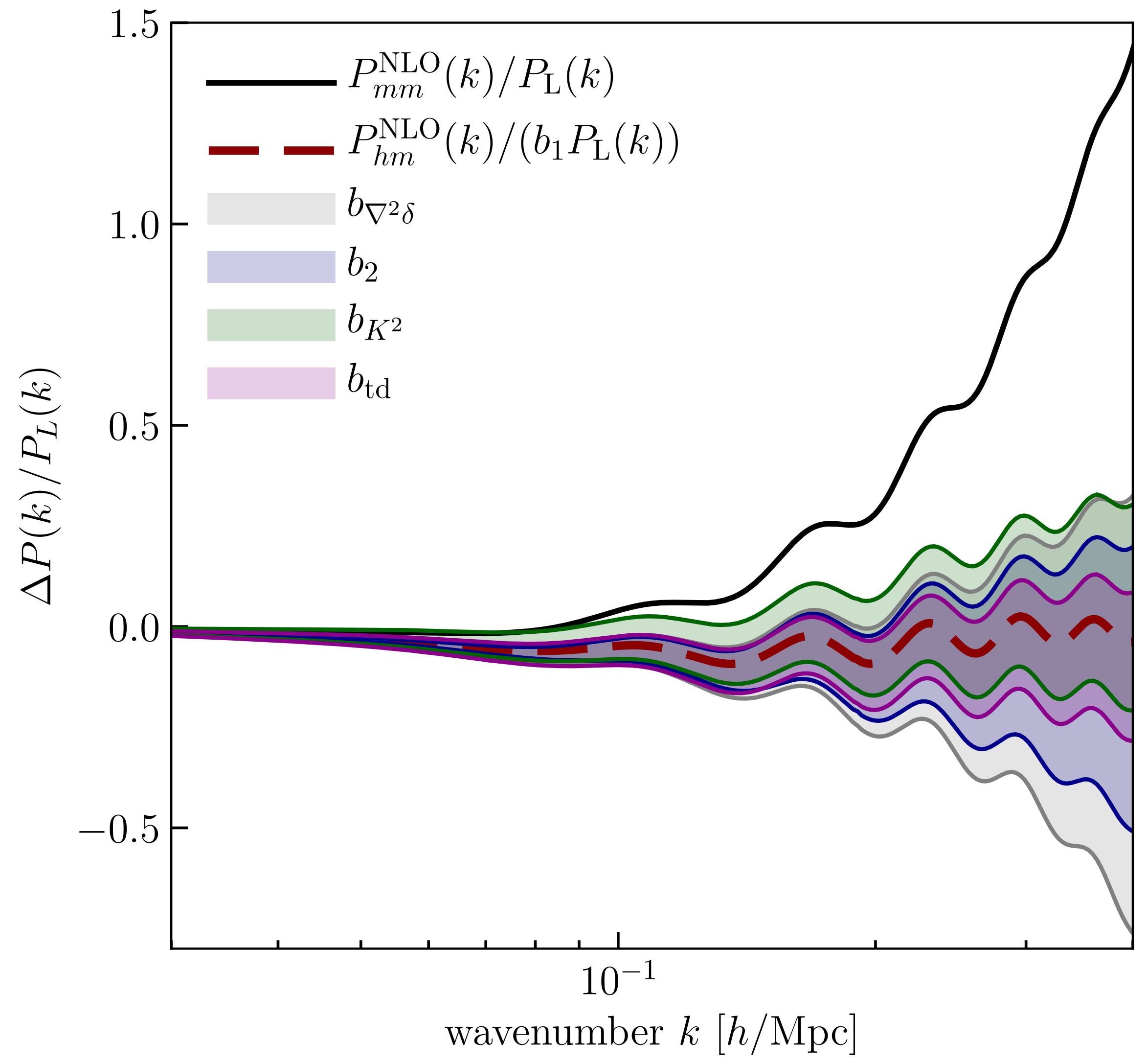
- We can of course still use the power spectrum on smaller scales
- However, need to add more nuisance parameters to have a reliable prediction



Gil-Marin et al, 2016

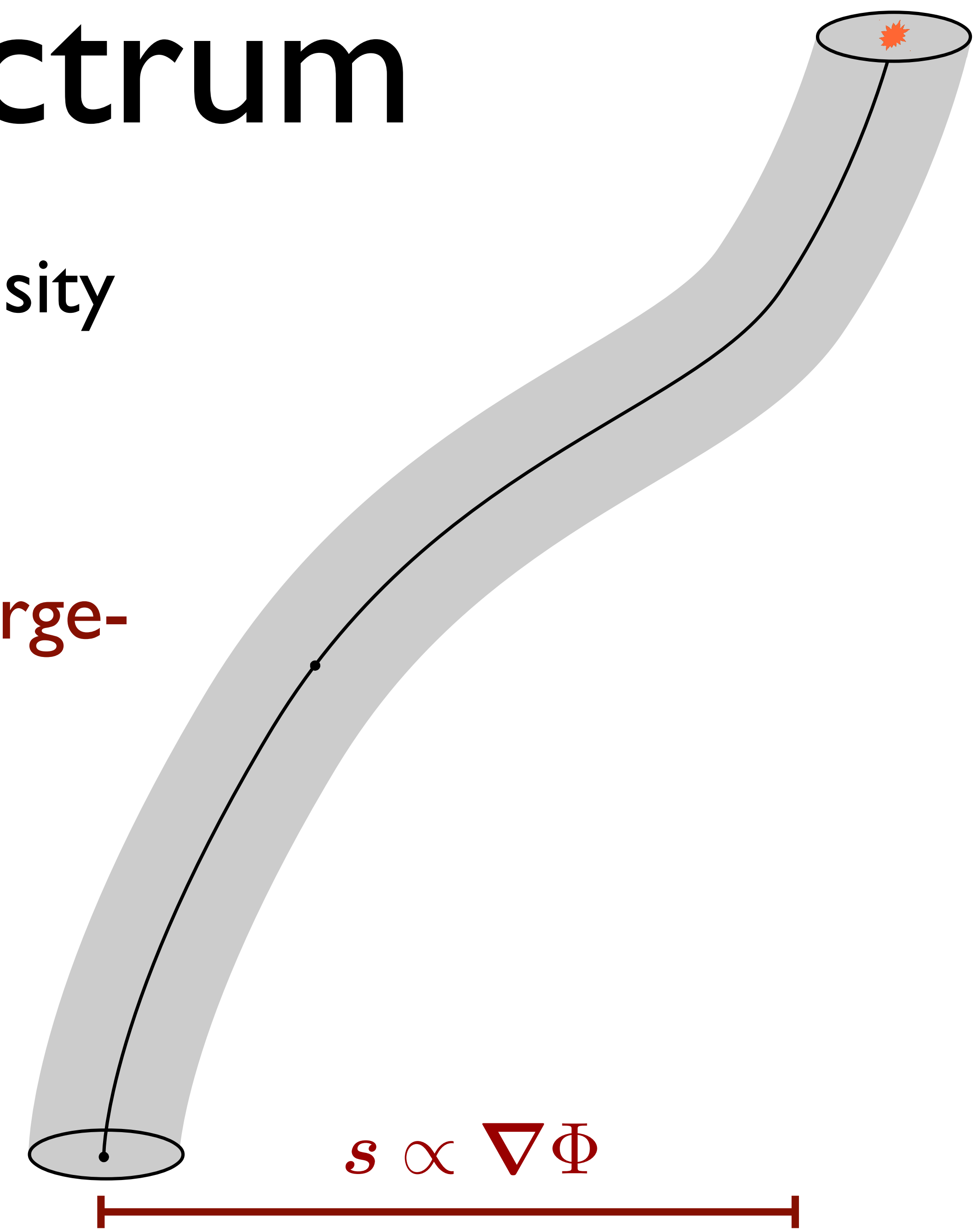
Why we *should* go beyond the power spectrum

- Beyond leading order: *5 additional parameters*
- Many contributions have very similar shape
- Free parameters *limit cosmological information that is available in power spectrum by itself*



Why we *should* go beyond the power spectrum

- At second and higher order, galaxy density contains displacement terms which are special:
  - Equivalence principle ensures that large-scale displacement is the same for galaxies and matter
  - Displacement term allows for disentangling bias and amplitude of fluctuations ($\mathcal{A}_s$ or σ_8)
- In the power spectrum, these are mixed in with other nonlinear bias contributions and impossible to disentangle



Inference beyond the power spectrum

- One approach: *higher-order statistics* such as the bispectrum
  - Issue: complicated data vector, even more complicated covariance…
  - Even if the bispectrum is eventually measured, going to the four-point function seems futuristic
- Another: *nonlinear transformation* of the data
  - BAO reconstruction, voids, density-split statistics, marked correlation functions, Minkowski functionals…

Inference beyond the power spectrum

- One approach: *higher-order statistics* such as the bispectrum
  - Issue: complicated data vector, even more complicated covariance…
  - Even if the bispectrum is eventually measured, going to the four-point function seems futuristic
- Another: *nonlinear transformation* of the data
  - BAO reconstruction, voids, density-split statistics, marked correlation functions, Minkowski functionals…
- But can we get *all the information from the galaxy density field at once?*

Inference beyond the power spectrum

- Given cosmological parameters θ, we can predict

1. Statistics of initial conditions — Prior $P_{\rm prior}\left(\vec{\delta}_{\rm in}, \theta\right)$
2. How a given $\delta_{\rm in}(\boldsymbol{x})$ evolves into the final density field — deterministic evolution $\vec{\delta}_{\rm fwd}[\vec{\delta}_{\rm in}, \theta]$

Inference beyond the power spectrum

- For the situation we are dealing with in cosmology, then, the *full posterior of cosmological parameters given the data* is then given by

$$P(\theta) = \int \mathcal{D}\vec{\delta}_{\rm in} P\left(\vec{\delta}_g \middle| \vec{\delta}_{\rm fwd}[\vec{\delta}_{\rm in}, \theta]\right) P_{\rm prior}\left(\vec{\delta}_{\rm in}, \theta\right)$$

$$= \int d\{b_O\} P\left(\vec{\delta}_g \middle| \vec{\delta}, \theta_i, \{b_O\}\right)$$

conditional probability: all physics of galaxy formation enters here

Inference beyond the power spectrum

- For the situation we are dealing with in cosmology, then, the *full posterior of cosmological parameters given the data* is then given by

$$P(\theta) = \int \mathcal{D}\vec{\delta}_{\rm in} P\left(\vec{\delta}_g \middle| \vec{\delta}_{\rm fwd}[\vec{\delta}_{\rm in}, \theta]\right) P_{\rm prior}\left(\vec{\delta}_{\rm in}, \theta\right)$$

Multivariate Gaussian, diagonal covariance in Fourier space

Inference beyond the power spectrum

- For the situation we are dealing with in cosmology, then, the *full posterior of cosmological parameters given the data* is then given by

$$P(\theta) = \int \mathcal{D}\vec{\delta}_{\rm in} P\left(\vec{\delta}_g \middle| \vec{\delta}_{\rm fwd}[\vec{\delta}_{\rm in}, \theta]\right) P_{\rm prior}\left(\vec{\delta}_{\rm in}, \theta\right)$$

Very high-dimensional integral…

Inference beyond the power spectrum

$$P(\theta) = \int \mathcal{D}\vec{\delta}_{\rm in} P\left(\vec{\delta}_g \middle| \vec{\delta}_{\rm fwd}[\vec{\delta}_{\rm in}, \theta]\right) P_{\rm prior}\left(\vec{\delta}_{\rm in}, \theta\right)$$

- How does this work in practice? Markov Chain Monte Carlo:
  - Discretize fields on grid
  - Draw initial conditions from prior
  - Forward-evolve using gravity, compute likelihood
  - Compare with data and repeat
- Challenge: even with fairly coarse resolution, have to sample many millions of parameters
- Key: Hamiltonian Monte Carlo

Inference beyond the power spectrum

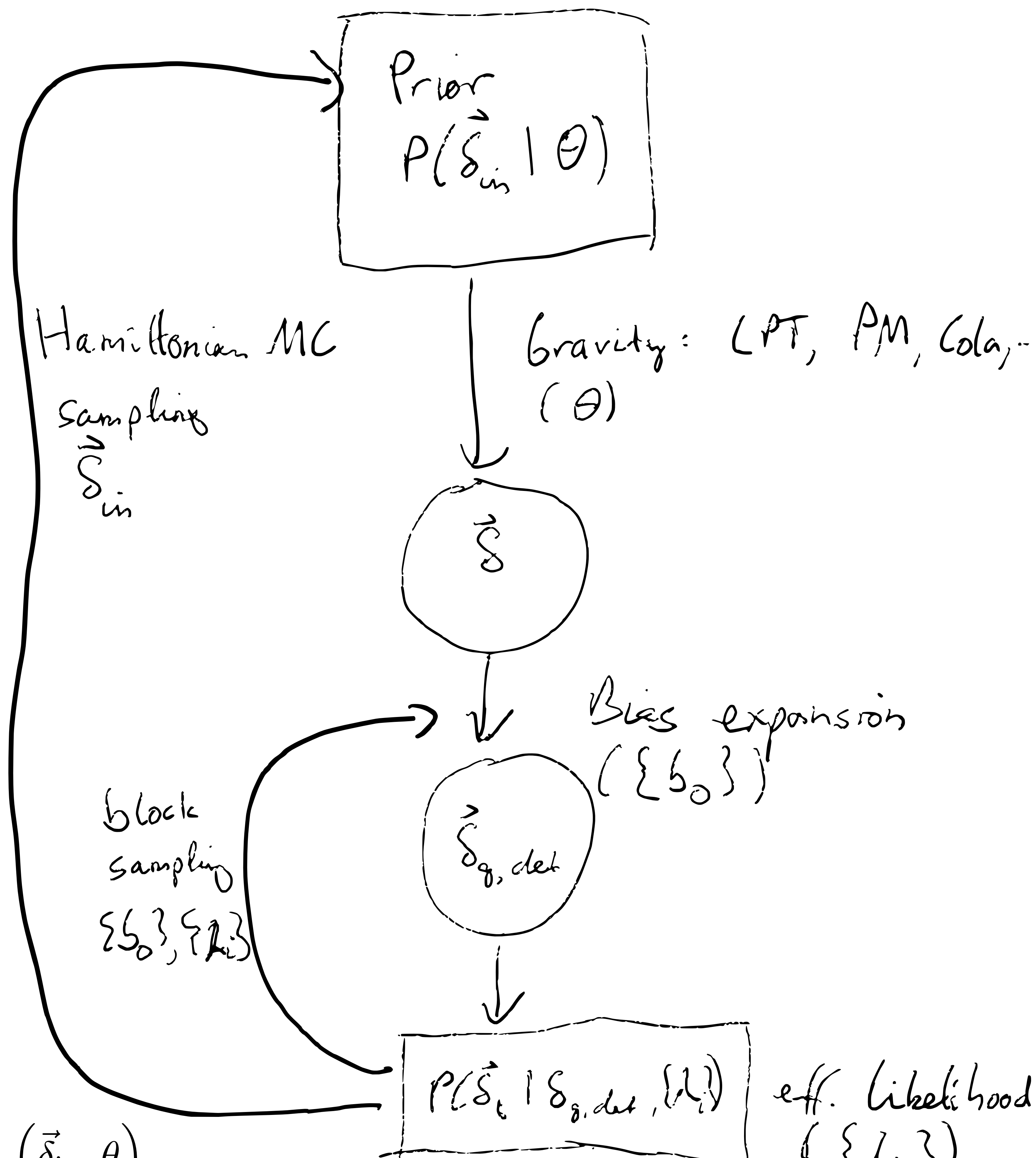
$$P(\theta) = \int \mathcal{D}\vec{\delta}_{\rm in} P\left(\vec{\delta}_g \middle| \vec{\delta}_{\rm fwd}[\vec{\delta}_{\rm in}, \theta]\right) P_{\rm prior}\left(\vec{\delta}_{\rm in}, \theta\right)$$

- How does this work in practice? Markov Chain Monte Carlo:
  - Discretize fields on grid
  - Draw initial conditions from prior
  - Forward-evolve using gravity, compute likelihood
  - Compare with data and repeat

- Lots of interest in this approach recently

Kitaura & Ensslin, Jasche & Wandelt, Wang, Mo et al, Seljak et al, Jasche & Lavaux (2017), …

Flowchart:



$$P(\theta) = \int \mathcal{D}\vec{\delta}_{\rm in} \int d\{b_O\}$$
$$P\left(\vec{\delta}_g \middle| \vec{\delta}_{\rm fwd}[\vec{\delta}_{\rm in}, \theta], \{b_O\}\right) P_{\rm prior}\left(\vec{\delta}_{\rm in}, \theta\right)$$

The likelihood for galaxies

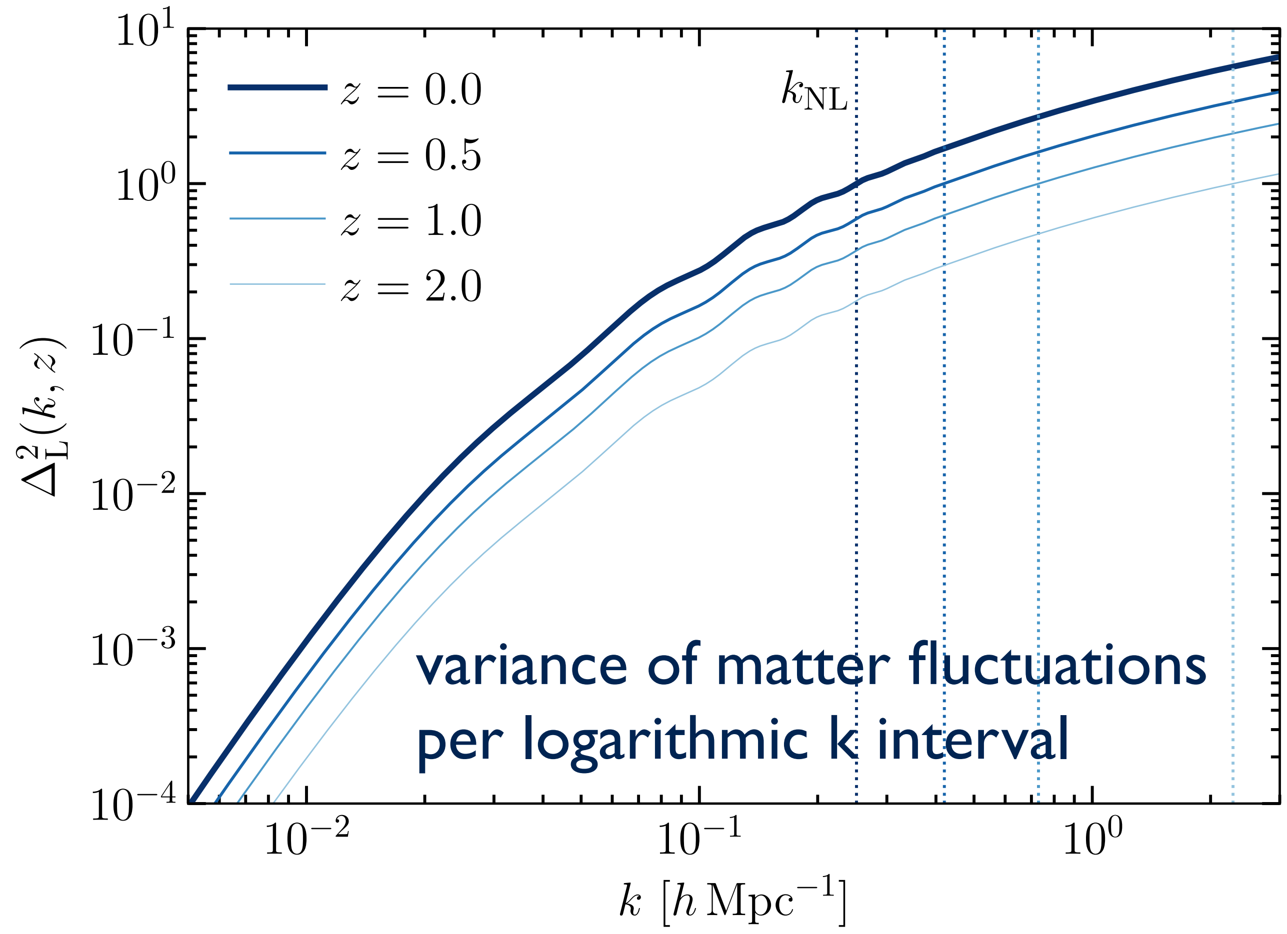
- The key physical ingredient needed in this approach is the conditional probability *(likelihood)* $P\left(\vec{\delta}_g \middle| \vec{\delta}, \theta_i, \{b_O\}\right)$
- A wrong likelihood leads to biased inference of initial conditions and cosmology
  - Although correlation coefficient between true and reconstructed IC is robust

Nguyen, FS et al; 2011.06587

- Is there a way to obtain a *likelihood that rigorously marginalizes over small-scale nonlinearities?*
  - Yes, using the effective field theory of LSS

Theory of galaxy clustering

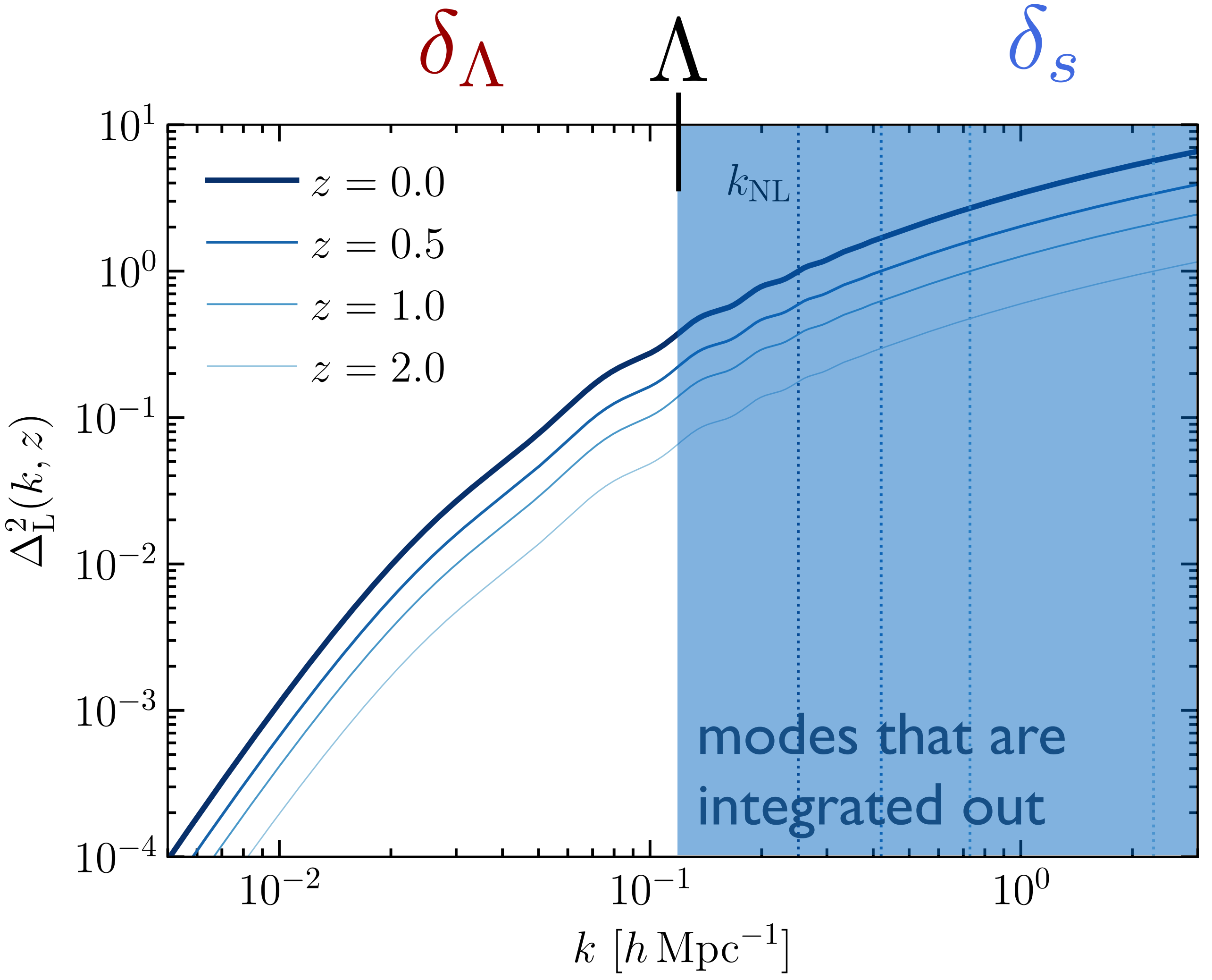
- Perturbations in our universe are small on large scales
  - Perturbation theory works on quasilinear scales $k < k_{NL}$
- Goal: describe galaxy clustering up to a given scale and accuracy using a finite number of free bias parameters and stochastic amplitudes



EFT approach

- Idea: trust our theory for $k < \Lambda$
- Split perturbations into large scale (< Λ) and small scale (>= Λ):

$$\delta(\boldsymbol{x}, \tau) \equiv \frac{\rho_m(\boldsymbol{x}, \tau)}{\bar{\rho}_,(\tau)} - 1 = \delta_\Lambda + \delta_s$$



- Then, we integrate out (marginalize over) perturbations with k > Λ

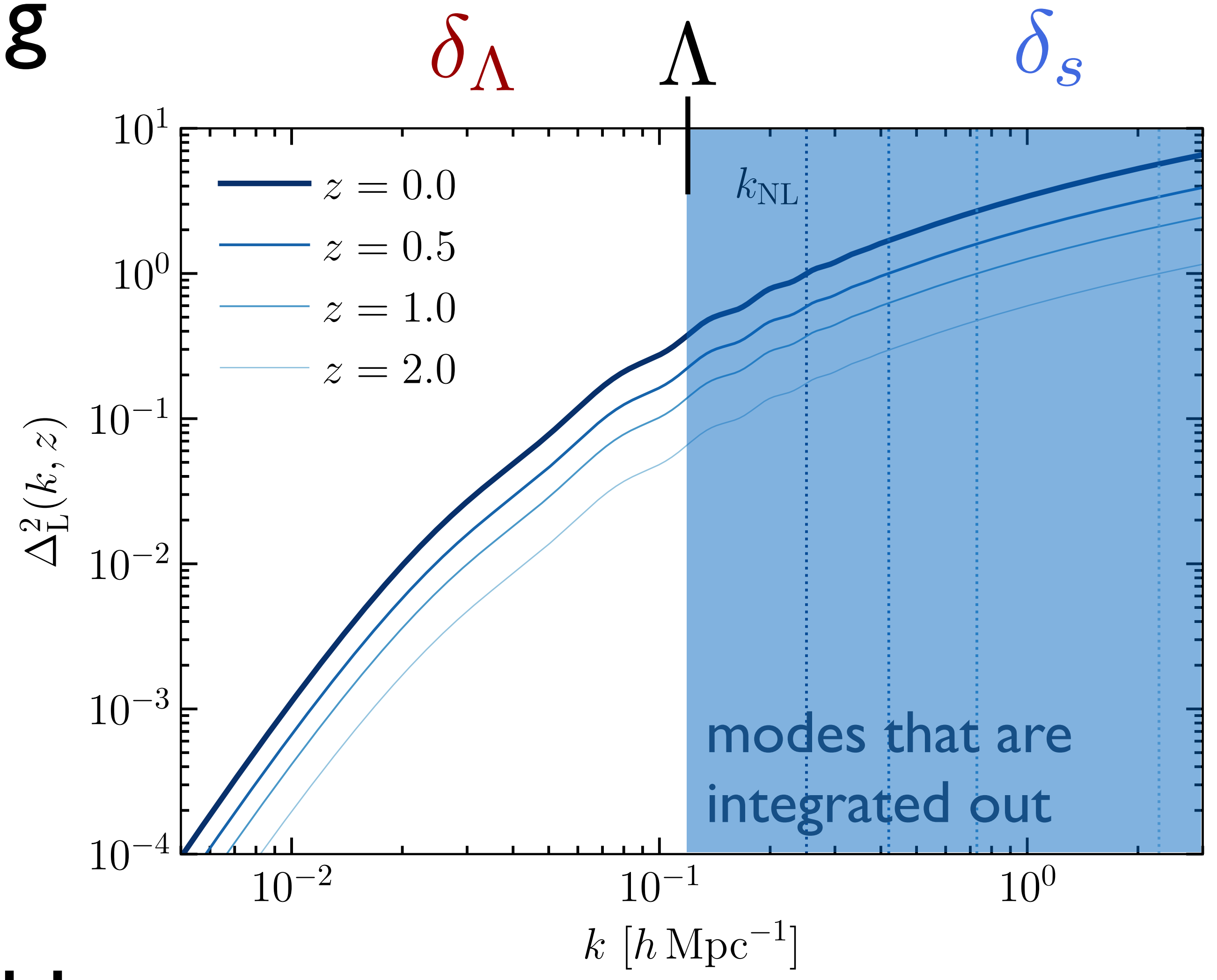
EFT approach

- Incorporate effect of large-scale perturbations explicitly using bias expansion, with free coefficients b_O

$$\delta_g(\boldsymbol{x}) = \sum_O b_O O(\boldsymbol{x}) + \varepsilon(\boldsymbol{x})$$

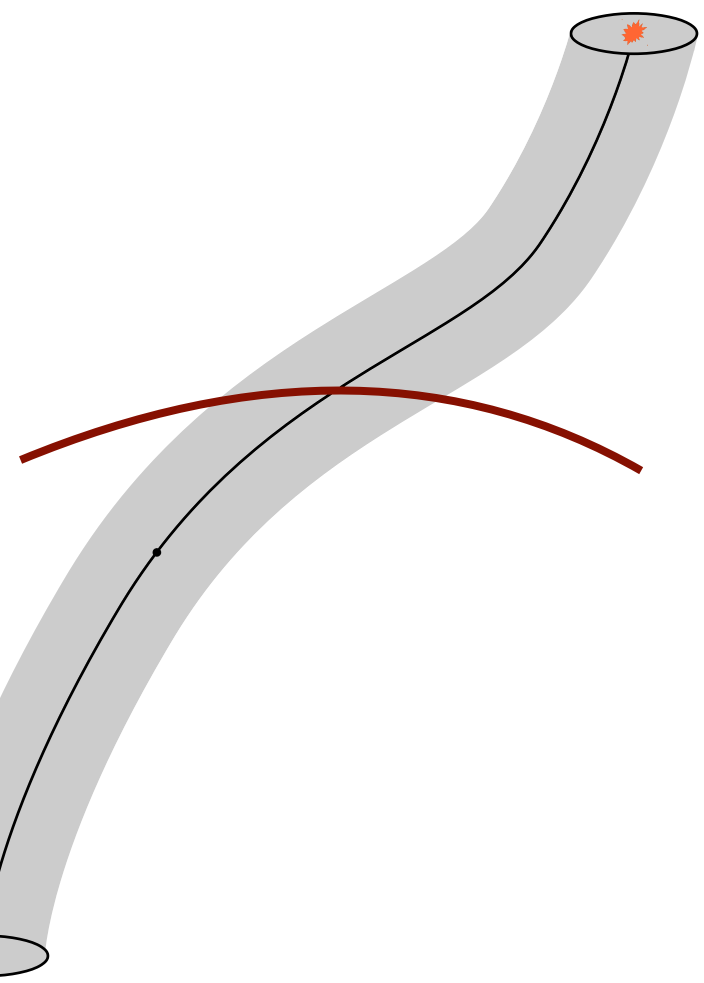
  - Fields O are constructed from δ_Λ

- Small-scale perturbations add noise ε



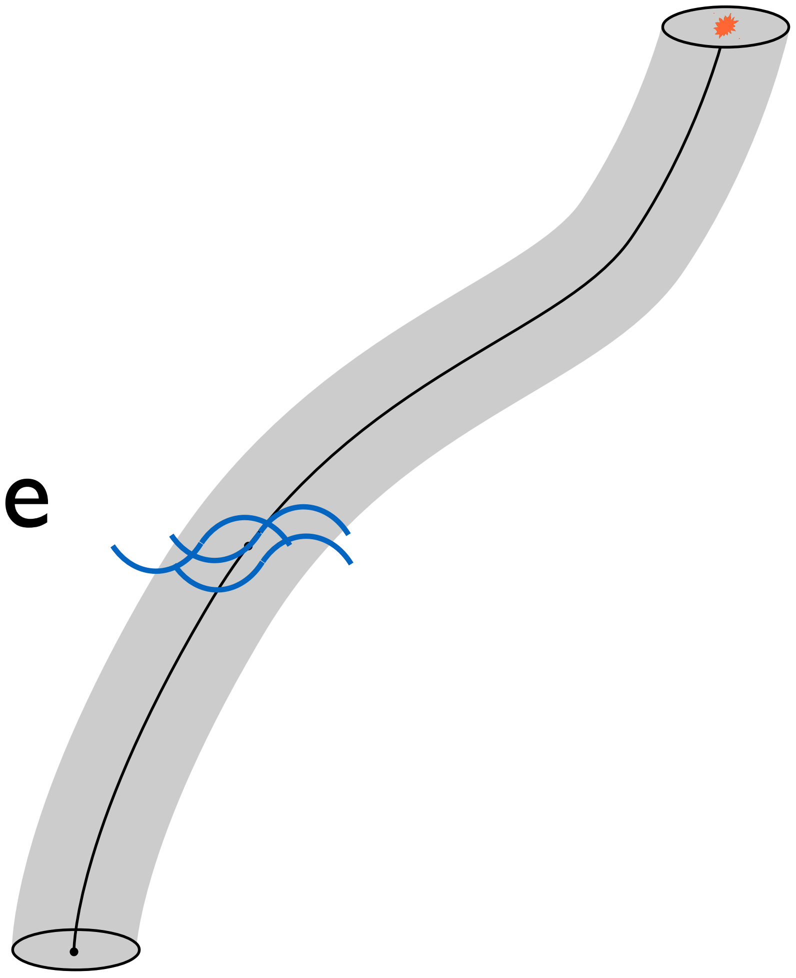
Bias

- Which bias terms $O(x)$ we need to include:
  - Well understood by now
  - Include dependence on full history of structure formation
  - Includes "local bias" (powers of matter density) as well as tidal fields, time and space derivatives thereof
- Displacement terms protected by equivalence principle have *fixed* coefficients!



Stochasticity

- ε arises from local (in real space) superposition of many small-scale perturbations
- Central limit theorem: $\varepsilon(k)$ is approximately Gaussian distributed (the lower k, the more Gaussian it is)
- Local in real space: power spectrum is white noise at low k, with corrections ~k^2:



$$\langle \varepsilon(\boldsymbol{k}) \varepsilon^*(\boldsymbol{k}') \rangle = (2\pi)^3 \delta_D(\boldsymbol{k} - \boldsymbol{k}') \left[P_\varepsilon + k^2 P_\varepsilon^{\{2\}} + \cdots \right]$$

Cabass, FS arXiv:1909.04022

EFT likelihood

- Given its Gaussianity, can analytically integrate out the noise to obtain the desired likelihood of the galaxy density field:

$$P\left(\vec{\delta}_g\middle|\vec{\delta}\right) \propto \left(\prod_{\boldsymbol{k}\neq 0}^{\Lambda} \sigma_0^2\right)^{-1/2} \exp\left[-\frac{1}{2}\sum_{\boldsymbol{k}\neq 0}^{\Lambda} \frac{1}{\sigma_0^2}\left|\delta_g(\boldsymbol{k}) - \delta_{g,\mathrm{det}}(\boldsymbol{k})\right|^2\right]$$

with $\delta_{g,\mathrm{det}}(\boldsymbol{k}) = \sum_O b_O O(\boldsymbol{k})$

(at leading order)
FS, Elsner, et al; 1808:02002

- Equivalent formulation exists in real space — Cabass, FS; 2004.00617
- Easy to go to higher orders in bias expansion
- In fact, can *analytically marginalize* over bias parameters
- Clear relative ordering of bias and stochastic terms

Cabass, FS; 1909.04022

Fixed phase test for halos

- To test this EFT likelihood, let's begin with a toy setup:
  - Take halos in full N-body simulation as our galaxy sample
  - Can we recover unbiased cosmology from a halo catalog of unknown selection, given initial conditions with an arbitrary normalization?
  - Cosmology: restrict to σ_8 (or equivalently $\mathcal{A}_s$)

Elsner, FS, et al; 1906.07143

Cosmology from halos

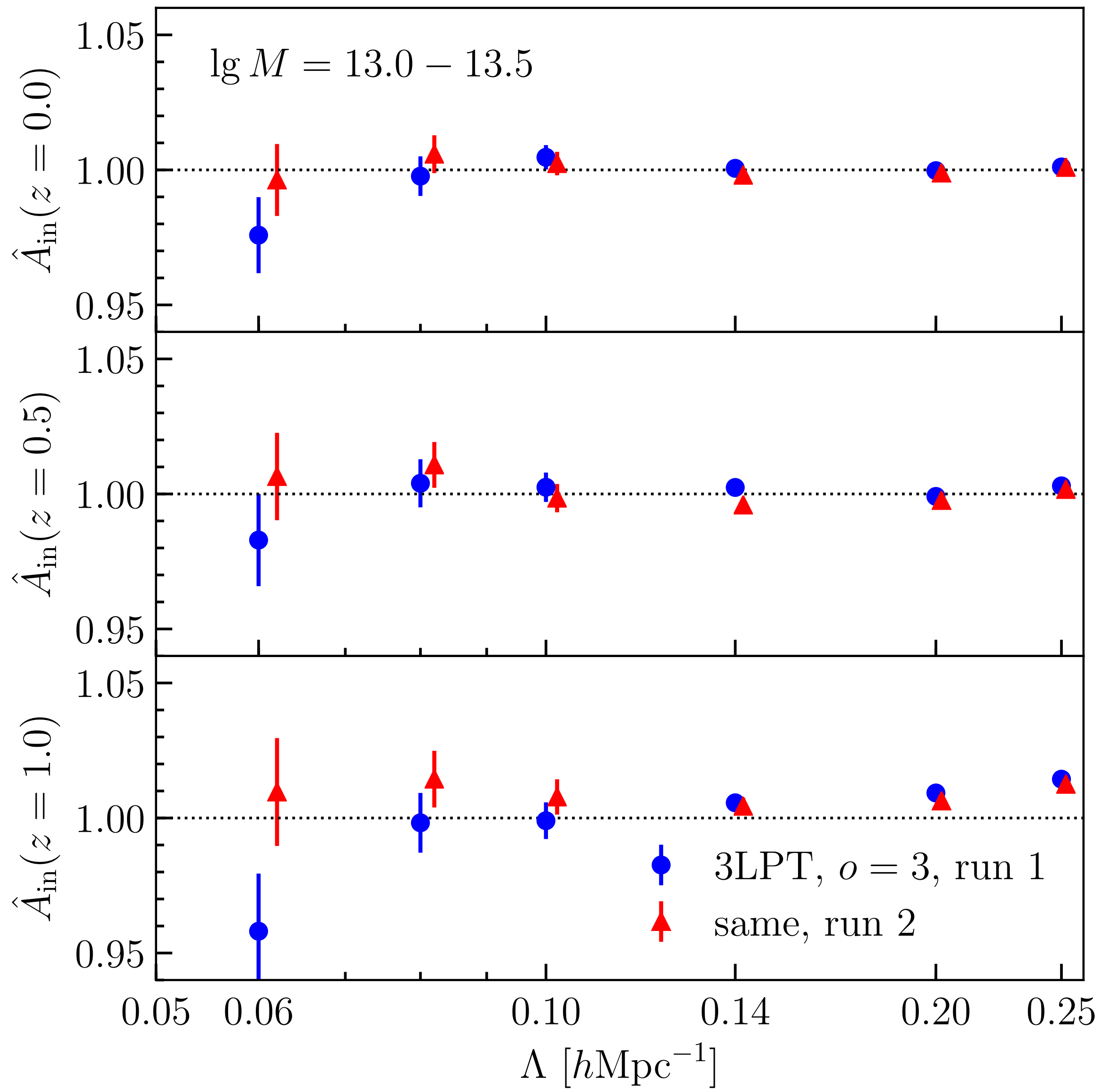
- Can we recover unbiased σ_8 from a halo catalog of unknown selection?
- Note: perfect degeneracy between b_1 and σ_8 at linear order; nonlinear information essential

Maximum-likelihood value of σ_8:

$$A_{\rm in} \equiv \frac{\sigma_8}{\sigma_8^{\rm fid}}$$

as a function of cutoff Λ
(maximum wavenumber used)

3LPT, third-order bias expansion



FS, et al; arXiv:2004.06707; FS, arXiv:2009.14176

Cosmology from halos

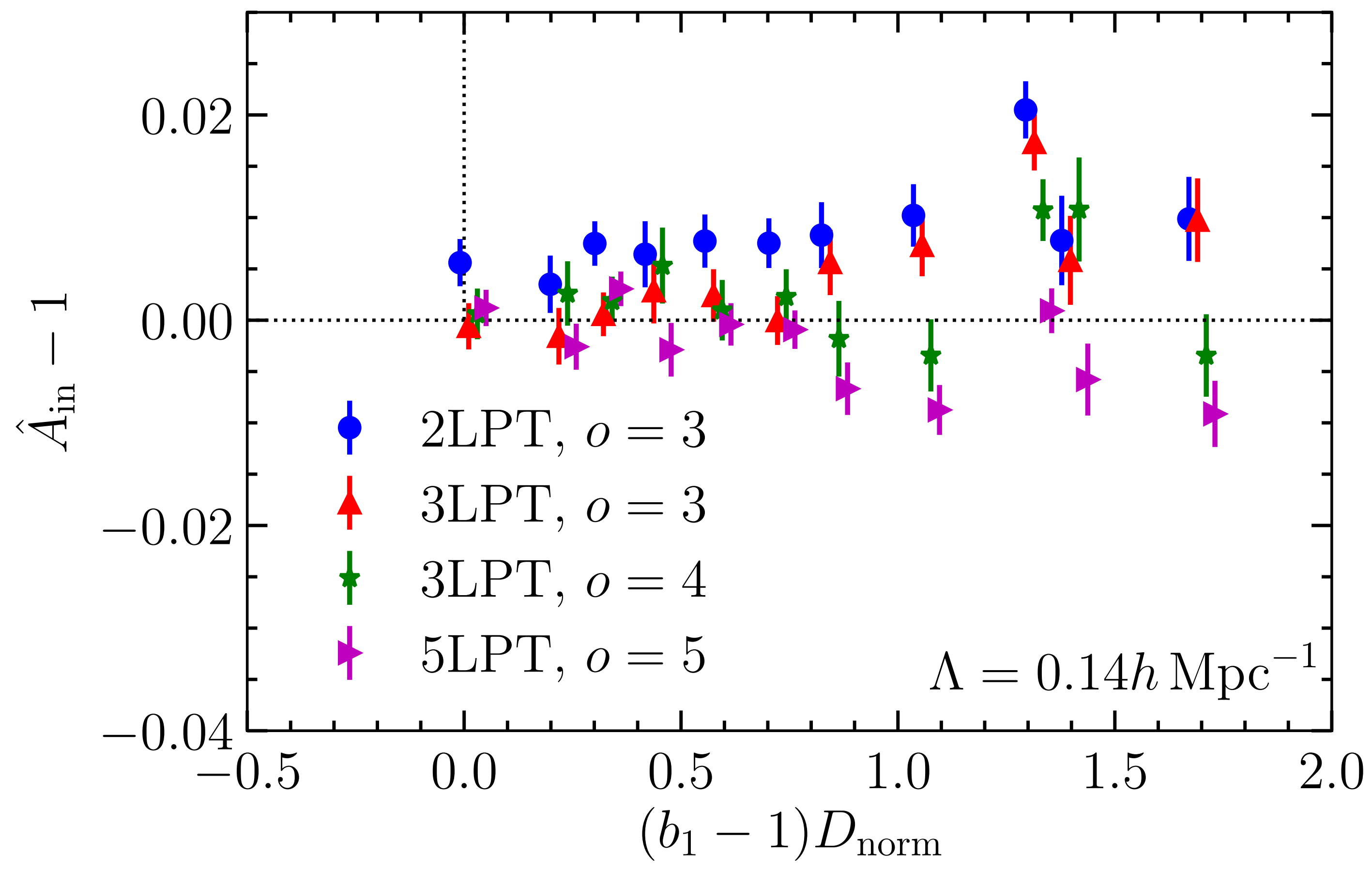
- Can we recover unbiased σ_8 from a halo catalog of unknown selection?
- Note: perfect degeneracy between b_1 and σ_8 at linear order; nonlinear information essential

Maximum-likelihood value of σ_8:

$$A_{\rm in} \equiv \frac{\sigma_8}{\sigma_8^{\rm fid}}$$

as a function of $(b_1\text{-}1)\,D(z)$
(Λ fixed)

varying LPT and bias orders



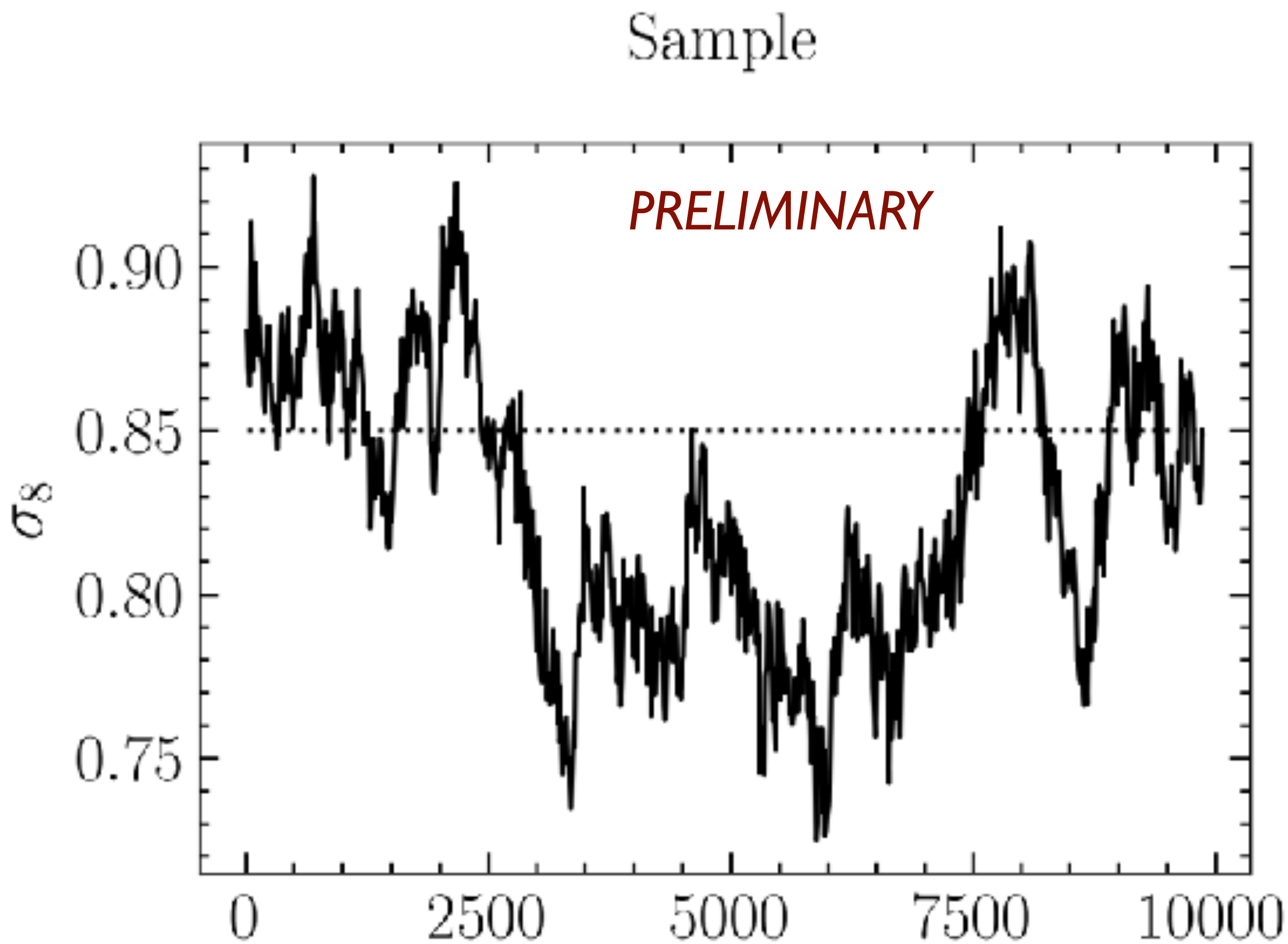
FS, et al; arXiv:2004.06707; FS, arXiv:2009.14176

How much information is there actually?

- We get statistical tiny error bars on cosmology if we fix the phases: *very conservatively* $\Delta\sigma_8$ <~ 0.8% for (2000 Mpc/h)^3 volume for halos
    - About one order of magnitude smaller than expected error from power spectrum/bispectrum analysis!
- But in reality, we don't know the initial conditions (phases) of course
- What can we ultimately expect once we also sample those?

How much information is there actually?

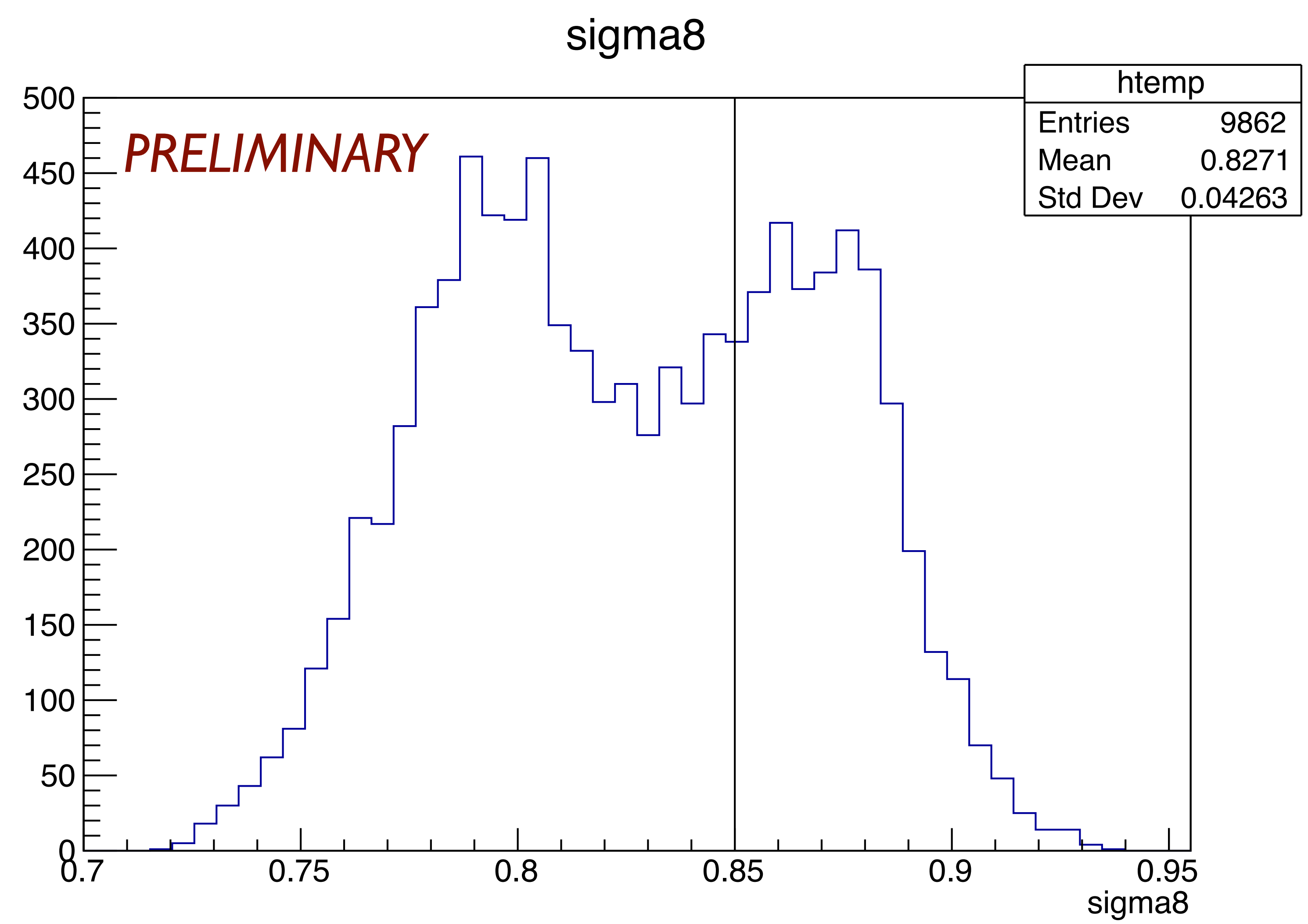
- Work in progress… EFT likelihood implemented in BORG code (Elsner/Reinicke/Jasche/Lavaux)
- Test on "mock sample" generated from likelihood itself; conservative cutoff
- ~5% constraint on σ_8; comparable to linear RSD constraints, but completely independent - based on nonlinear information



$\Lambda = 0.1h\,\mathrm{Mpc}^{-1}$; $V = 8h^{-3}\mathrm{Gpc}^3$; $\bar{n} \simeq 2\times 10^{-4}h^3\mathrm{Mpc}^{-3}$

How much information is there actually?

- Work in progress… EFT likelihood implemented in BORG code (Elsner/Reinicke/Jasche/Lavaux)
- Test on "mock sample" generated from likelihood itself; conservative cutoff
- ~5% constraint on σ_8; comparable to linear RSD constraints, but completely independent - based on nonlinear information



$\Lambda = 0.1h\,\mathrm{Mpc}^{-1}$; $V = 8h^{-3}\mathrm{Gpc}^3$; $\bar{n} \simeq 2 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$

Summary

- Full Bayesian inference of density field is expected to yield substantially improved cosmology constraints, at least on σ_8 and very probably also f_{NL}
- Much simpler in any forward approach to go to higher orders (see 5th order results) and to incorporate nontrivial observational effects

Discussion points

- Where does the additional information come from?
- ~~What about covariance?~~ *(None)*
- What about mask? Systematics?
- RSD? Cabass, arXiv:2007.14988