New Approaches to Galaxy Clustering

Fabian Schmidt
MPA

with
Giovanni Cabass, Vincent Desjacques, Franz Elsner, Jens Jasche, Donghui Jeong, Guilhem Lavaux, Mehrdad Mirbabayi, Minh Nguyen, Matias Zaldarriaga

q = x_{fl}(0)

x = x_{fl}(\tau)
Motivation

• The clustering of galaxies (large-scale structure, LSS) is historically one of the key probes of cosmology

  Peebles; Efstathiou+ ’90 predicted a positive cosmological constant $\Lambda$ from LSS observations

• From ~1998 until recently, most spectacular results came from “cleaner” probes - Supernovae and the cosmic microwave background (CMB)

• Now, again, in a new golden age of LSS with plenty of experiments under way: BOSS, DES, DESI, PFS, SphereX, Euclid, WFIRST, ...
Motivation

- Using large-scale structure, we can learn about

- Inflation (or, which mechanism generated seeds of structure?)
- Dark Energy and Gravity (is General Relativity correct?)
- Dark Matter (does it cluster as expected?)
- the formation history of galaxies, clusters, and the IGM

Uniquely broad set of science opportunities!
Motivation

• **Inflation**: reconstruct the properties of the initial conditions, and look for gravitational waves

• **Dark Energy and Gravity**: the growth of structure depends sensitively on the expansion history of the Universe, and the nature of gravity

  Growth equation: \[ D'' + aH D' = 4\pi G \bar{\rho} D \]

• **Dark Matter**: how “cold” is cold dark matter? What is the sum of neutrino masses?
Challenge: unlike the CMB, every data point is nonlinear!
Cold Dark Matter cosmology in a nutshell

- Assume scale-invariant, adiabatic, approx. Gaussian initial conditions
- Large-scale fluctuations are small (still linear today)
- Structure forms hierarchically from small to large scales
- Perturbative expansion in fluctuations on large scales

Millennium simulation / MPA
Cold Dark Matter cosmology in a nutshell

- Assume scale-invariant, adiabatic, approx. Gaussian initial conditions
- Large-scale fluctuations are small (still linear today)
- Structure forms hierarchically from small to large scales
- Perturbative expansion in fluctuations on large scales
Theory of Large-Scale Structure

• Well-established tools:
  • linear Boltzmann
  • N-body* methods

* and hydrodynamics

\[ \frac{aH}{k} \]

\[ \delta(k, t) = \frac{\rho_m(k, t)}{\bar{\rho}_m} - 1 \]
Theory of Large-Scale Structure

• Foundation: separation between nonlinear scale and horizon

\[ k_{NL} \approx 0.1 h \text{ Mpc}^{-1} \gg aH \]

• Linear theory: Fourier modes evolve independently; solved problem

• However, bulk of information in LSS is on nonlinear scales \((N_{\text{modes}} \sim k_{\text{max}}^3)\)
How do we compare theory with data?

- Assume we observe the matter density field\(^*\) \(\rho(x) = \bar{\rho}[1 + \delta(x)]\)

- Given cosmological model, theory predicts
  1. Statistics of initial conditions \(\delta_{\text{in}}(x) = \lim_{t \to 0^+} \delta(x, t)\)
  2. How a given \(\delta_{\text{in}}(x)\) evolves into the final density field

\(^*\) Drop time argument throughout for clarity; assume fixed observation time
How do we compare theory with data?

- If matter density field was Gaussian,
  - PDF of $\delta(x)$ is multivariate Gaussian, with diagonal covariance in Fourier space

- Then all the information would be contained in the power spectrum

\[
\langle \delta(k)\delta^*(k') \rangle = (2\pi)^3 \delta_D(k - k') P(k)
\]
How do we compare theory with data?

- If matter density field was Gaussian,
  - PDF of $\delta$ is multivariate Gaussian, with diagonal covariance in Fourier space

- Then all the information would be contained in the power spectrum
  \[
  \langle \delta(k)\delta^*(k') \rangle = (2\pi)^3 \delta_D(k-k') P(k)
  \]

- However, final matter density is clearly non-Gaussian!
Inference beyond the power spectrum

• Assume we observe the matter density field\(^*\) \[ \rho(x) = \bar{\rho}[1 + \delta(x)] \]

• Given cosmological parameters \(\theta\), theory predicts
  1. Statistics of initial conditions
  2. How a given \(\delta_{\text{in}}(x)\) evolves into the final density field
Inference beyond the power spectrum

• Assume we observe the matter density field
  \[ \rho(x) = \bar{\rho}[1 + \delta(x)] \]

• Given cosmological parameters \( \theta \), theory predicts
  1. Statistics of initial conditions
  2. How a given \( \delta_{\text{in}}(x) \) evolves into the final density field

Prior \( P_{\text{prior}} \left( \delta_{\text{in}}, \theta \right) \)
Inference beyond the power spectrum

- Assume we observe the matter density field* $\rho(x) = \bar{\rho}[1 + \delta(x)]$

- Given **cosmological parameters** $\theta$, theory predicts
  1. Statistics of initial conditions  
  2. How a given $\delta_{\text{in}}(x)$ evolves into the final density field

Prior
\[ P_{\text{prior}}(\delta_{\text{in}}, \theta) \]

Conditional probability, in absence of errors:
\[ P\left(\delta | \delta_{\text{in}}, \theta\right) = \delta_{D}^{\infty}\left(\delta - \delta_{\text{fwd}}[\delta_{\text{in}}, \theta]\right) \]
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

\[ P(\theta) = \int D\delta_{\text{in}} \, P\left( \delta_{\text{obs}} | \delta_{\text{in}}, \theta \right) P_{\text{prior}} \left( \delta_{\text{in}}, \theta \right) \]
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

\[
P(\theta) = \int \mathcal{D}\tilde{\delta}_\text{in} \ P \left( \tilde{\delta}_\text{obs} \bigg| \tilde{\delta}_\text{in}, \theta \right) P_{\text{prior}} \left( \tilde{\delta}_\text{in}, \theta \right)
\]

- Functional integral...
- Multivariate Gaussian
Inference beyond the power spectrum

\[ P(\theta) = \int D\delta_{\text{in}} P(\vec{\delta}_{\text{obs}} | \vec{\delta}_{\text{in}}, \theta) P_{\text{prior}}(\vec{\delta}_{\text{in}}, \theta) \]

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

\[ P(\theta) = \int \mathcal{D}\vec{\delta}_{\text{in}} \ P \left( \vec{\delta}_{\text{obs}} \left| \vec{\delta}_{\text{in}}, \theta \right) \right) \ P_{\text{prior}} \left( \vec{\delta}_{\text{in}}, \theta \right) \]

- How does this work in practice? Markov Chain Monte Carlo:
  - Discretize field on grid
  - Draw initial conditions from prior
  - Forward-evolve using gravity
  - 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), ...
Full Bayesian inference in practice

\[ \delta_{\text{in}}(x) \]

\[ \delta(x) \]
We don’t observe the matter distribution, however…
What we observe, is this...
Thus, we need

\[ P(\theta) = \int \mathcal{D}\delta_{in} \int \mathcal{D}\delta P(\vec{N}_g|\vec{\delta}) P(\vec{\delta}|\delta_{in}, \theta) P_{prior}(\delta_{in}, \theta) \]

\[ P \left( \vec{\mathcal{N}}_g \mid \vec{\delta} \right) \]
Theory of galaxy clustering

• We cannot yet simulate the formation of galaxies* fully realistically

• Need to abstract from the incomplete understanding on small scales

  • Only hope for rigorous results is on scales $k < k_{NL}$

• Goal: describe galaxy clustering up to a given scale and accuracy using a finite number of free bias parameters $b_0$, and stochastic amplitudes

* Everything in following will apply to any tracer of LSS.
Most of what I will talk about, and much more, can be found in:

Large-Scale Galaxy Bias

Vincent Desjacques\textsuperscript{a,b}, Donghui Jeong\textsuperscript{c}, Fabian Schmidt\textsuperscript{d}

\texttt{arXiv:1611.09787}
Galaxy formation

- Consider coarse-grained (large scale) view of region that forms a galaxy at conformal time $\tau$

- Formation happens over long time scale, but small spatial scale $R^*$
  
  - For halos, expect $R^* \lesssim R_L$

- Thus, a gradient expansion makes sense on large scales (small wavenumbers $k$)
• Effective field theory: write down all terms (in Lagrangian or equations of motion) that are consistent with symmetries

• Gravity: general covariance

• Galaxy density: 0-component of 4-vector (momentum density)
EFT approach in LSS

• Effective field theory: write down all terms (in Lagrangian or equations of motion) that are consistent with symmetries

• Gravity: general covariance

• Galaxy density: 0-component of 4-vector (momentum density)

• Order contributions by perturbative order, and number of spatial derivatives (gradient expansion)
EFT approach in LSS

• For large-scale structure (LSS), general covariance boils down to the statement that $\Phi$, $\nabla \Phi$ and $\nu$ cannot appear in bias expansion.

• No surprise: leading gravitational observable is tidal field $\partial_i \partial_j \Phi$. 
EFT bias expansion

- Can we write the EFT as local in time and space?
  - Only makes sense if spatial and time derivatives are suppressed
  - True for spatial derivatives, but not for time derivatives! Galaxies form over many Hubble times (as does matter field)
- Fundamental theory is nonlocal in time
Non-locality in time

- We can similarly deal with non-locality in time at higher order, since expansion continues to factorize:

\[ \delta(x, \tau) = D(\tau)\delta^{(1)}(x) + D^2(\tau)\delta^{(2)}(x) + \cdots \]

- Allows us to obtain a complete expansion of galaxy density field:

\[ n_g(x, \tau) = \bar{n}_g(\tau) \left[ 1 + \sum_O b_O(\tau)O(x, \tau) \right] \]

up to given desired order in perturbations
Complete bias expansion

\[
n_g(x, \tau) = \bar{n}_g(\tau) \left[ 1 + \sum_{O} b_O(\tau)O(x, \tau) + \epsilon(x, \tau) + \epsilon_{\delta}(x, \tau)\delta(x, \tau) \cdots \right]
\]

• The picture is not complete yet, since this relation can only hold in a “mean-field” sense

• Small-scale perturbations introduce stochasticity \( \epsilon \) (and higher-order terms)

• Cannot predict \( \epsilon \) as field, but know the form of statistics:

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

• In the end, stochasticity reduces to fixed number of additional free parameters
Complete bias expansion

- Some virtues of this expansion:
  - Complete in the EFT sense: closed under renormalization
  - Equivalent expansions in Eulerian and Lagrangian space
  - Bias parameters can be mapped from one frame to another unambiguously
  - Allows us to obtain likelihood as well -> later
Spatial nonlocality and scale-dependent bias

- Beyond large-scale limit: need to expand spatial nonlocality of galaxy formation

- Higher derivative biases are suppressed with scale $R^*$

- E.g., $R^2 \nabla^2 \delta \to \delta_g(k, \tau) = (b_1 + b_{\nabla^2 \delta} k^2 R^2) \delta(k, \tau)$

- This also allows for baryonic physics, which has to come with additional derivatives

  - Example: pressure perturbations $\delta p = c_s^2 \delta \rho$

  - Pressure force: $F = \nabla \delta p \propto \nabla \delta$

- At higher order in derivatives, time evolution no longer determined by gravity alone
Velocity bias

- Galaxy velocities are important probe of cosmology - but how are they related to matter velocity?

- Recall that the relative velocity between matter and galaxies is an observable, and thus cannot involve $\Phi$, $\nabla \Phi$, or $\nu$

- Leading contribution:

$$\nu_g - \nu_m = \beta \nabla \delta$$

- Two more derivatives -> suppressed by $k^2$
Velocity bias

- $v_g - v_m = \beta \nabla \delta$

- This is what we expect from pressure forces
  
  $F = \nabla \delta p \propto \nabla \delta$

- Also small-scale stochastic velocities, with power spectrum $\sim k^4$, which captures virial motions

- **Summary:** Galaxy velocities are unbiased on large scales

- Can then be used to constrain gravity and dark energy
Application: galaxy power spectrum

- Assume we can measure rest-frame galaxy density
  - That is, neglect redshift-space distortions and other projection effects

- Leading-order galaxy power spectrum at fixed time:
  \[ P_{gg}(k) = b_1^2 P_L(k) + P_{\varepsilon}^{(0)} \]

- Valid on very large scales
- 2 free parameters
Application: galaxy power spectrum

- Next-to-leading order (NLO): involve 2 additional quadratic, 1 cubic, and 2 higher-derivative parameters

- Quadratic and cubic terms scale like
  \[ \epsilon_{\text{loop}} \equiv \left( \frac{k}{k_{\text{NL}}} \right)^{3+n} \approx \left( \frac{k}{0.25 \, h \, \text{Mpc}^{-1}} \right)^{1.3} \]

- Controlled by shape of \( P(k) \) and nonlinear scale

- Higher-derivative contributions scale as \( \epsilon_{\text{deriv.}} \equiv k^2 R_*^2 \)

- Obviously, NLO corrections become important toward smaller scales (higher \( k \))

- Importantly: Two independent expansion parameters!
Many contributions have very similar shape

If only interested in power spectrum, can significantly reduce number of free parameters

But then we cannot make use of constraints imposed by the EFT / equivalence principle
Why we should go beyond the power spectrum

• Consider galaxy and matter over density at second order:

\[ \delta_g^{(2)} = b_1 \delta^{(2)} + \frac{1}{2} b_2 \delta^2 + b_K^2 (K_{ij})^2 \]

\[ \delta^{(2)} = \frac{17}{21} \delta^2 + \frac{2}{7} (K_{ij})^2 - s^k \partial_k \delta \]

• Equivalence principle ensures that large-scale displacement \( s \propto \nabla \Phi \) is the same for galaxies and matter (cf. velocities)

• Displacement term allows for disentangling bias and amplitude of fluctuations (\( \mathcal{A}_s \) or \( \sigma_8 \))

• In terms of summary statistics: need to measure three-point function (bispectrum)

• But there are more such robust terms at higher orders
Full Bayesian inference in the EFT approach

• Recall that we need an expression for the conditional probability $P \left( \tilde{N}_g \big| \delta \right)$

• Again, we want to abstract from unknown small-scale details of galaxy formation - likelihood needs to absorb these details

• Describe galaxy counts in term of discretized fractional overdensity: $\tilde{N}_g = \bar{N}[1 + \tilde{\delta}_g]$
An EFT approach to the likelihood

• Recall that the noise in the galaxy field is approximately Gaussian with analytic power spectrum:
  \[ \delta_g(k) - \delta_{g,\text{det}}(k) = \varepsilon(k) \]
  \[ \delta_{g,\text{det}}(k) = \sum_O b_O O(k) \]

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

• Direct consequence of perturbative expansion
An EFT approach to the likelihood

- Hence, write the conditional probability in Fourier space:

\[
P(\theta) = \int \mathcal{D} \tilde{\delta}_{\text{in}} \int d\{b_O\} \, P(\tilde{\delta}_g | \tilde{\delta}_{\text{fwd}}[\tilde{\delta}_{\text{in}}, \theta], \{b_O\}) \, P_{\text{prior}}(\tilde{\delta}_{\text{in}}, \theta)
\]

with

\[
P(\tilde{\delta}_g | \delta) \propto \left( \prod_{k \neq 0} \sigma^2(k) \right)^{-1/2} \exp \left[ -\frac{1}{2} \sum_{k \neq 0} \frac{1}{\sigma^2(k)} |\delta_g(k) - \delta_{g,\text{det}}(k)|^2 \right]
\]

and

\[
\delta_{g,\text{det}}(k) = \sum_{O} b_O O(k)
\]

\[
\sigma^2(k) = \sigma_0^2 + \sigma_2^2 k^2 + \sigma_4^2 k^4
\]
Flowchart:

\[ P(\theta) = \int D\tilde{\delta}_{\text{in}} \int d\{b_O\} \]

\[ P(\tilde{\delta}_g|\tilde{\delta}_{\text{fwd}}[\tilde{\delta}_{\text{in}}, \theta], \{b_O\}) \] \( P_{\text{prior}} \left( \tilde{\delta}_{\text{in}}, \theta \right) \)

Hamiltonian MC sampling \( \tilde{\delta}_{\text{in}} \)

Gravity: \( \text{CPT, PM, Cola,} \cdots \)

Bias expansion \( \{b_O\} \)

Block sampling \( \{b_O\} \)

\[ P(\tilde{\delta}_c | \tilde{\delta}_g, \text{det}, \{\lambda_i\}) \] eff. likelihood \( \{\lambda_i\} \)
EFT likelihood implementation

- Concrete implementation: 2LPT forward evolution with

\[ O \in \{ \delta, \delta^2 - \langle \delta^2 \rangle, (K_{ij}^2) - \langle (K_{ij})^2 \rangle, \nabla^2 \delta \} \]

with coefficients \( \left\{ b_1, \frac{b_2}{2}, b_K^2, c_{\nabla^2 \delta} \right\} \).

(Actually, renormalized versions of these operators, constructed from density sharp-k filtered at \( 2k_{\text{max}} \))

- 7 free parameters, plus \( k_{\text{max}} \) as fixed metaparameter
EFT likelihood implementation

- Includes complete bias expansion up to second order and leading higher-derivative term
- Can show that this Fourier-space EFT likelihood captures information equivalent to
  - Galaxy power spectrum at 1-loop order
  - Galaxy bispectrum at leading order
  - Full 2LPT BAO reconstruction

*Combined!
First results

\[ \alpha \equiv \sigma_8 / \sigma_8^{\text{true}} \]

\[ \alpha[k \leq 0.05 \, h \, \text{Mpc}^{-1}] \in [0.99, 1.04] \]

- From halos in N-body simulations
- Right now, phases are fixed to true values
Summary

• LSS contains a wealth of information on dark energy, growth of structure, and the early Universe

• To use this, we need to understand nonlinear (and nonlocal) relation between initial conditions and observed galaxies

• We now have a complete framework for galaxy biasing (on perturbative scales)
  • Also for galaxy velocities

• Leads to well-defined prediction for all $n$-point functions of galaxies
  • “Just compute”

• Lots of free parameters, however
Summary

Next challenges:

- How much information in nonlinear galaxy clustering, given these many free parameters?
- How best to extract it?

Right approach in principle: full Bayesian inference ("forward modeling") with explicit marginalization over phases.

We have made recent progress in understanding how the EFT approach provide us with a robust likelihood for this purpose.