Perturbative models for Euclid analysis

Martin White (UCB/LBNL)
w/ Stephen Chen (UCB), Zvonimir Vlah (CERN)

https://github.com/sfschen/velocileptors

Scope

- Today I want to talk about state-of-the-art, perturbative models for modeling the “broadband” power spectrum of biased tracers in redshift-space.
- I will talk about some implications for measurements.
- I will not talk about modeling the power spectrum post reconstruction.
- I will not talk about relative velocity effects.
- I will not talk about wide-angle corrections.
- I will not talk about modeling cross-correlations with lensing.
- **However** the same models and basic parameters can be applied to these things too!
  - ... and, yes, we have written papers on them ...

Perturbation theory I

New surveys are probing higher z and larger volumes:

- The Universe at "high" redshift is more linear, better correlated with the 'primordial' Universe.
- Start to get very high precision measurements of modes we can model well.
  - Larger volume $\Rightarrow$ smaller $P(k)$ errors at each k, and 'small' errors at low k.
  - Need to go beyond linear, even at $k \simeq 0.05\,h\,\mathrm{Mpc}^{-1}$.
  - This is a regime where PT approaches work very well (small corrections to almost-linear quantities).
- Large-scale structure at high-z offers many of the same 'advantages' of primary CMB anisotropy: a well controlled, analytic calculation which can be compared straightforwardly to observations.
- **PT approaches are more powerful for bigger surveys!**

Perturbation theory II

- We have models using both the Eulerian and Lagrangian formulation of PT – allows an estimate of robustness.
- Models work in both Fourier and configuration space (and pre- and post-reconstruction) ... focus today on pre-reconstruction Fourier space statistics.
- Full model has 8 free parameters: 3 bias, 3 counterterm and 2 stochastic (e.g. shot noise, FoG) though can drop some of these with little impact on the fits.
  - FoG model is more flexible than e.g. Gaussian or Lorentzian. Priors?
- Using a variety of tricks we can compute $P(\mathbf{k})$ quickly (seconds, comparable to CAMB/CLASS) for use in MCMC codes – code is public (Python).

N-body simulations

- To validate our model we compare to a series of public, large N-body simulations run by the ANL group using HACC.
- Four 4096^3 particle simulations, total volume $256\,h^{-3}\mathrm{Gpc}^3$ or > 25 full-sky surveys at $z \simeq 0.8 \pm 0.05$.
- Very difficult to find large volume, high resolution, percent-level accurate simulations.
  - Evidence of percent-level errors in the velocities due to time-stepping choices.

RSD: generating function

Going into redshift space involves a "shift" in configuration space (due to peculiar velocity, $\mathbf{u}$) and so a phase in Fourier space:

$$1 + \delta_s(\mathbf{s}, \tau) = \int d^3\mathbf{x} \, \left(1 + \delta_g(\mathbf{x}, \tau)\right) \, \delta_D(\mathbf{s} - \mathbf{x} - \mathbf{u})$$

$$(2\pi)^3 \delta_D(\mathbf{k}) + \delta_s(\mathbf{k}) = \int d^3\mathbf{x} \, \left(1 + \delta_g(\mathbf{x}, \tau)\right) \, e^{i\mathbf{k}\cdot(\mathbf{x}+\mathbf{u}(\mathbf{x}))},$$

This is most easily handled using a generating function

$$\tilde{M}(\mathbf{J}, \mathbf{k}) = \frac{k^3}{2\pi^2} \int d^3r \, e^{i\mathbf{k}\cdot\mathbf{r}} \, \left\langle (1 + \delta_g(\mathbf{x}_1))(1 + \delta_g(\mathbf{x}_2)) e^{i\mathbf{J}\cdot\Delta\mathbf{u}} \right\rangle_{\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{r}}$$

in terms of which

$$\frac{k^3}{2\pi^2} P_s(\mathbf{k}) = \frac{k^3}{2\pi^2} \int d^3r \, e^{i\mathbf{k}\cdot\mathbf{r}} \, \left\langle (1 + \delta_g(\mathbf{x}_1))(1 + \delta_g(\mathbf{x}_2)) e^{i\mathbf{k}\cdot\Delta\mathbf{u}} \right\rangle_{\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{r}}$$

$$= \tilde{M}(\mathbf{J} = \mathbf{k}, \mathbf{k})$$

Moment expansion

All models for RSD involve choices in evaluating $\tilde{M}(\mathbf{J} = \mathbf{k}, \mathbf{k})$.

For example, simply expanding the exponential gives the moment expansion

$$P_s(\mathbf{k}) = \sum_{n=0}^{\infty} \frac{i^n}{n!} \, k_{i_1} \cdots k_{i_n} \, \tilde{\Xi}^{(n)}_{i_1 \cdots i_n}(\mathbf{k})$$

where

$$\Xi^{(n)}_{i_1 \cdots i_n} = \langle (1 + \delta_1)(1 + \delta_2) \Delta \mathbf{u}_{i_1} \cdots \Delta \mathbf{u}_{i_n} \rangle$$

The 1^{st} moment is the mean pairwise velocity, $\Xi^{(1)}_i = v_{12,i}(\mathbf{r})$, and the 2^{nd} is the pairwise velocity dispersion, $\Xi^{(2)}_{ij} = \sigma_{12,ij}(\mathbf{r})$.

Both are also of interest to peculiar velocity or kSZ surveys!

Moments

On the scales most relevant for cosmology, almost all of the expansions turn out to be very close to the moment expansion.

- Including more moments in the expansion increases the accuracy up to $k|\Delta\mathbf{u}| \approx 1$, where the expansion breaks down.
- Higher moments tend to be dominated by increasingly non-linear physics (e.g. finger of god).
- Up to $n = 3$ essentially saturates the expansion, and $n = 3$ is dominated by the disconnected piece:

$$\begin{aligned}\Xi^{(3)}_{ijk}(\mathbf{r}) &= \langle (1+\delta_1)(1+\delta_2)\Delta\mathbf{u}_i\Delta\mathbf{u}_j\Delta\mathbf{u}_k\rangle \\ &\approx \langle (1+\delta_1)(1+\delta_2)\Delta\mathbf{u}_i\rangle\,\langle \Delta\mathbf{u}_j\Delta\mathbf{u}_k\rangle \\ &\approx \sigma_v^2\ \delta_{jk}\ \Xi^{(1)}_i(\mathbf{r}).\end{aligned}$$

(but we've computed through $n = 4$).

Perturbation theory

- We use Eulerian or Lagrangian PT to model $P(k)$, $v_i(k)$ and $\sigma_{ij}(k)$ and disconnected $\Xi^{(3)}_{ijk}$.
- We include EFT contributions to handle UV-sensitive terms.
- We use a 3rd order bias expansion, e.g. in LPT

$$
\begin{aligned}
1+\delta_g(\mathbf{q},\tau_0) &= 1 + b_1\delta_0(\mathbf{q}) \\
&+ \frac{1}{2}b_2\left(\delta_0^2(\mathbf{q}) - \left\langle\delta_0^2\right\rangle\right) \\
&+ b_s\left(s_0^2(\mathbf{q}) - \left\langle s_0^2\right\rangle\right) \\
&+ b_3\ O_3(\mathbf{q}) + \cdots \\
&+ b_\nabla\nabla^2\delta_0(\mathbf{q}) + \epsilon(\mathbf{q}),
\end{aligned}
$$

- The EPT and LPT results can be mapped into one another through a "dictionary" relating the bias coefficients.

(The b_s and b_∇ terms can account for assembly bias.)

Example: LPT $P(k)$

Real-space power spectrum in Lagrangian PT:

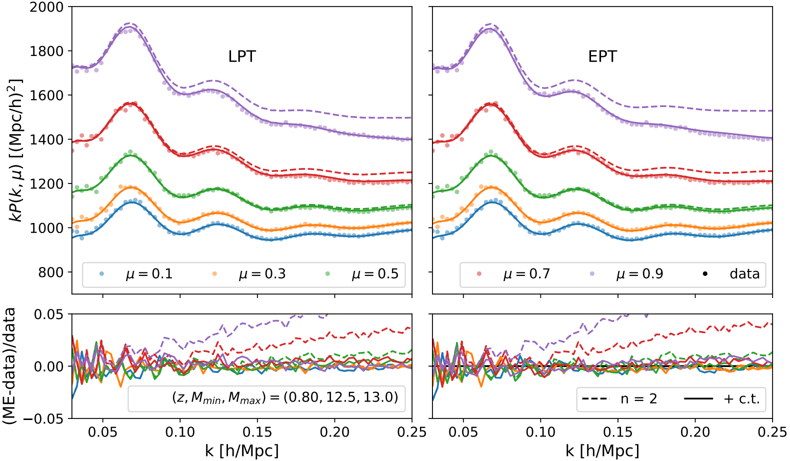
$$P(k) = \int d^3\mathbf{q}\ e^{i\mathbf{k}\cdot\mathbf{q}}\ e^{-\frac{1}{2}k_ik_jA_{ij}^{\rm lin}} \Big\{ 1 + 2ib_1k_iU_i + b_1^2\xi_L - \frac{1}{2}k_ik_jA_{ij}^{1-\rm loop} + \frac{i}{6}k_ik_jk_kW_{ijk} + \frac{1}{2}b_2^2\xi_L^2 + 2ib_1b_2\xi_Lk_iU_i - (b_2+b_1^2)k_ik_jU_iU_j + ib_2k_iU_i^{20} + ib_1^2k_iU_i^{11} - b_1k_ik_jA_{ij}^{10} + b_s(-k_ik_j\Upsilon_{ij} + 2ik_iV_i^{10}) + 2ik_ib_1b_sV_i^{12} + b_2b_s\chi + b_s^2\zeta + \underbrace{\alpha_P k^2}_{c.t.} + \cdots \Big\} + \underbrace{R_h^3}_{stoch.}$$

with e.g.

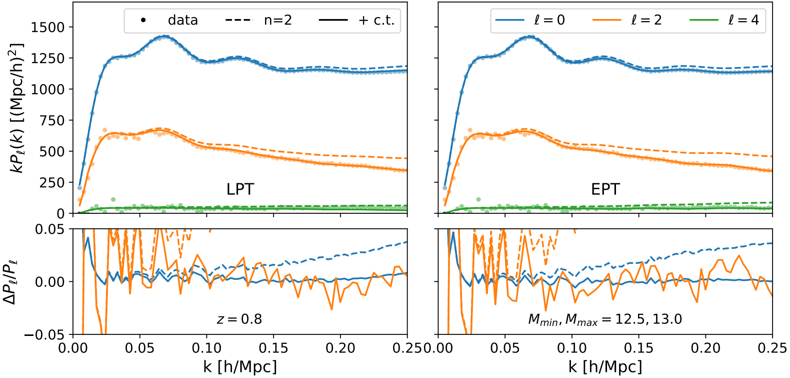
$$U_i(q) = \frac{\hat{q}_i}{2\pi^2}\int_0^\infty dk\ P_L(k)\left[-k\,j_1(kq)\right] + \mathcal{O}(P_L^2)$$

The integrals over k and q can be done efficiently using fast Fourier transforms – code is "intrinsically fast" and can also mesh easily with an emulator!

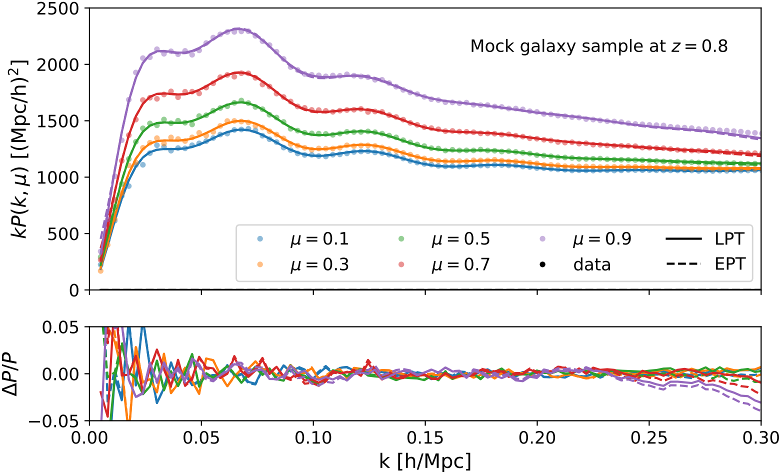
Power spectrum wedges: halos



Power spectrum multipoles: halos



PT power spectrum wedges: ‘galaxies’



Wedges vs. multipoles

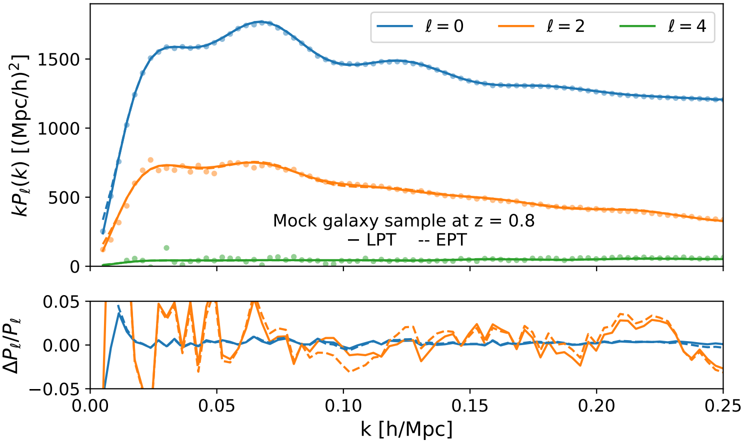
- Modern $P(\mathbf{k})$ measurement codes work natively in multipoles.
  - As do window function computations/corrections.
- There may be reasons to prefer wedges:
  - Redshift-space distortions are an expansion in $k_{\parallel} = k\mu$.
  - The $k_{\rm max}$ to which a model works is thus a function of μ
    - e.g. if FoG are "complex" this shows up at high $k_{\parallel}$ not k.
  - Easy to implement in a wedge analysis: $k_{\rm max}(\mu)$.
  - All multipoles have contributions from all μ!
- The theory can predict apodized wedges as easily as hard-edged wedges – allows lower $\ell_{\rm max}$ in $P(\mathbf{k})$ calculation.
  - Working out the optimal estimator would be an interesting exercise!
  - We have some understanding of error scaling with μ.

Conclusions

- Euclid will take large-scale structure to "the next level".
- There are a wealth of modes beyond $z \simeq 0.5$.
- We can measure "low and intermediate" k modes with very small uncertainties.
  - **PT approaches are more powerful for bigger surveys!**
- These will allow precisions tests of SM and GR, and improve constraints on parameters by substantial factors (or find something new!).
  - Already percent-ish level constraints at lower z are turning up much-discussed "tensions".
- This presents an interesting, and very 'principled', theoretical challenge.
- We have been developing PT-based tools for modeling next-generation spectroscopic (and imaging) surveys – which we expect will grow in power as the surveys become ever larger.

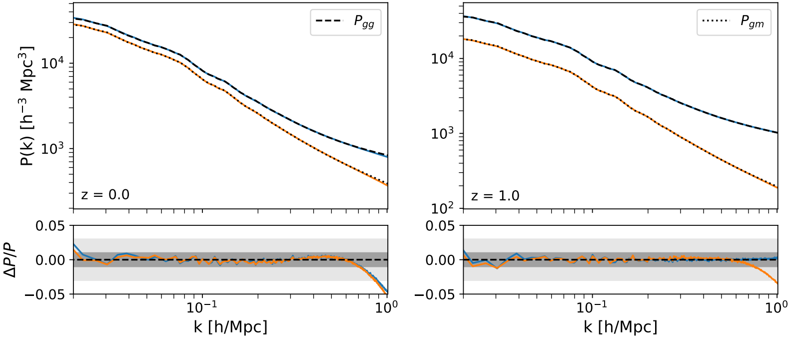
The End!

PT power spectrum multipoles: 'galaxies'



Hybrid emulator

Combine N-body simulations with symmetries-based bias expansion to build an efficient emulator for $P(k)$ ...



arXiv:1910.07097

Lowest order I

$$
\begin{aligned}
P_{\rm tree} \quad = \quad & 4\pi \int q^2\, dq\ e^{-(1/2)k^2(X_L+Y_L)} \Bigg\{ \\
& \left[1 + b_1^2\left(\xi_L - k^2 U_L^2\right) - b_2\left(k^2 U_L^2\right) + \frac{b_2^2}{2}\xi_L^2\right] j_0(kq) \\
+ \quad & \sum_{n=1}^{\infty}\left[1 - 2b_1 \frac{q\,U_L}{Y_L} + b_1^2\left(\xi_L + \left[\frac{2n}{Y_L} - k^2\right] U_L^2\right)\right. \\
& + b_2\left(\frac{2n}{Y_L} - k^2\right) U_L^2 \\
& \left. -2b_1 b_2 \frac{q\,U_L\,\xi_L}{Y_L} + \frac{b_2^2}{2}\xi_L^2\right]\left(\frac{k\,Y_L}{q}\right)^n j_n(kq)\Bigg\}
\end{aligned}
$$

For cross-correlations: $b_1 \to \frac{1}{2}\left(b_1^A + b_1^B\right)$, $b_1^2 \to b_1^A b_1^B$, etc.

Lowest order II

Where

$$
\begin{aligned}
\xi_L(q) &= \frac{1}{2\pi^2}\int_0^\infty dk\; P_L(k)\left[k^2\, j_0(kq)\right] \\
X_L(q) &= \frac{1}{2\pi^2}\int_0^\infty dk\; P_L(k)\left[\frac{2}{3} - 2\frac{j_1(kq)}{kq}\right] \\
Y_L(q) &= \frac{1}{2\pi^2}\int_0^\infty dk\; P_L(k)\left[-2j_0(kq) + 6\frac{j_1(kq)}{kq}\right] \\
U_L(q) &= \frac{1}{2\pi^2}\int_0^\infty dk\; P_L(k)\left[-k\, j_1(kq)\right]
\end{aligned}
$$

The integrals over q can be done efficiently using fast Fourier transforms or other methods.

The full expressions contain "1-loop" terms which are integrals of P_L^2.