Analytic approaches to large scale structure: perturbation theory and beyond

Uroš Seljak
UC Berkeley
Garching, July 22 2015
Why do theory

• Because it is fun
• To show off
• Because we are too lazy to run simulations
• Because we like to argue with fellow theorists
• Because it gives useful insights (maybe)
Perturbative approaches

- Perturbation theory: Eulerian (SPT), Lagrangian (LPT)
- Lowest order: 1 loop SPT, Zeldovich
- PT assumes $\delta<1$ and fails at the orbit crossing: halo formation
- 1d example (McQuinn & White): Zeldovich exact up to orbit crossing
LPT gets LSS right, but does not form high density halos.
In Zeldovich approximation particles just stream along straight line.
Beyond PT: parametrize PT ignorance

take the ratio. One can write the ratio of true $P(k)$ to $1\text{LPT} P(k)$ as a function $T^2(k)$ describing the PT ignorance, which starts as $k^2$:

effective field theory (EFT) approach

$$\frac{P_{\text{dm}}(k)}{P_{1\text{LPT}}(k)} = T_{1\text{LPT}}^2(k) = 1 + \alpha_{1\text{LPT}}(k)k^2 = \left(1 + \sum_{i=1}^{\infty} \alpha_{i,1\text{LPT}}k^{2i}\right)$$

EFT $\alpha$ is a parameter describing $1\text{LPT}$ error, of order (few Mpc)$^2$

---

Vlah & US 2015

Mcquinn & White 2015 [Mpc$^{-1}$]
PT challenges in 3-d

- In 1-d SPT expansion is convergent: e.g. 2 loop starts as $k^4P_L$, hence smaller than 1 loop $k^2P_L$ at low $k$

- In 3-d no longer so: higher order loops all start at $k^2P_L$ and receive large contributions from small scales which are spurious: in reality DM is trapped inside dark matter halos (shell crossings) on small scales. PT in 3-d is harder

- As a result higher loop terms can be completely wrong in their dominant terms

Vlah et al. 2014
3-d PT/EFT situation

- One can define many different PT approaches and many different ways to parametrize the ignorance: SPT 1 loop, SPT 2 loop, 1LPT, 2LPT, 3LPT... Higher order not obviously better. One can also define many different EFT schemes (EFT in SPT, LPT, IR resummation...)

- EFT needs to absorb a lot of PT problems: many EFT parameters, typical value of $\alpha$ of order $(1-3\text{Mpc}/h)^2$

- One can quantify the success by asking whether EFT parameter in front of $k^2 P_L$ is a constant (but beware of numerical issues at low $k$)

- EFT 1 loop SPT works to $k=0.1\text{h/Mpc}$, 2 loop to $0.2\text{h/Mpc}$ with 1 parameter (Baldauf etal, Foreman etal)
Primordial+BAO wiggles

- Zeldovich predicts wiggles well

PT+EFT predicts even better: tiny <1% effects, require special simulations to see them
EFT challenges for broadband

- One parameter EFT is only an approximation, one needs to add more free parameters at higher $k$.
- EFT has free parameters which are fit from simulations, so one can always choose a $k$ where the fit is perfect: danger of overfitting (e.g. Baldauf et al. 2015).
- For $k > 0.2h/Mpc$ stochastic terms become important, and EFT needs additional terms.
- Parametrizing ignorance with a single parameter works best at low $k$, but this is the regime where sampling variance errors dominate and linear theory works well.
- There is a limited dynamic range where EFT with a single parameter is applicable and relevant: $0.05h/Mpc < k < 0.2h/Mpc$ ($z=0$).
- For higher $k$ additive 1-halo term model is more suitable.
Halo model  
(with I. Mohammed, Z. Vlah)

- Approach 2: use PT (e.g. Zeldovich) approximation to describe large scales
- Use halo model to describe small scales and add
- Halo description: mass, profile: take moments
- Halo mass function $dn/\ln M = f(\nu) \, d\nu$
- $P_{1\text{halo}} = A_0 k^0 + A_2 k^2 + A_4 k^4 + \ldots$
  
  Halo model not accurate enough, use simulations to determine $A_0, A_2, \text{etc}$

At low $k$ we need to impose mass and momentum conservation, forcing 1 halo term to vanish
Total power spectrum
PT model for galaxies

• Need to develop nonlinear models that are sufficiently general to allow for any reasonable nonlinear effects present in the data, while preserving as much of cosmological information as possible

• Some can be modeled by perturbation theory (PT)+biasing

• There is always more information on small scales, but most of it is hopelessly corrupted by nonlinear effects that cannot be modeled in PT

• One needs to model our ignorance obeys all symmetries (e.g. $k^2P_L(k)$ at low k) and all physics (the biasing parameters are physical, e.g. FoG is determined by halo mass...)

• In recent years a workhorse has been the halo model+biasing+PT
PT+halo model for halos

- Use PT to model halos, account for all bias terms
- Biasing: local, non-local, $k^2$, stochastic...
- General principle: everything that is allowed by symmetry is also present in reality (many biasing terms)
- Example: 1 loop SPT modeling of RSD for halos (Vlah et al. 2013)

RSD is never linear for $k>0.1\, h/\text{Mpc}$
Beyond PT: effect of satellites in real space

- Satellite-central pairs and satellite-satellite pairs inside halos create additional 1 halo term
- Leading term at low k: additional Poisson shot noise amplitude $\Sigma^2 = V/N_{cs}$ given by number of central-satellite pairs
- Leading correction due to radial distribution $\Sigma^2(1-k^2R^2_{vir})$, same also for 2-halo term
- Must vanish at high k to give 0 or 1/n

Okumura et al 2015
Beyond PT: redshift space distortions

- Supplemented by satellite velocities inside the halos (Fingers of God), inducing 2 halo term and 1 halo term
- FoG term is large: virial theorem $\sigma_{\text{vir}}^2 = 50R_{\text{vir}}^2$, must use resummed version, e.g. $\exp(-k^2\mu^2\sigma_{\text{vir}}^2)$
- Leading term: 2 halo correlation between central galaxies and satellites in different halos
- FoG also applies to 1 halo term

Okumura et al. 2015
Example: modeling of CMASS in simulations

Linear Kaiser never a good model

With SPT, 1 shot noise term and 1 FoG term one can go to $k=0.2\,h/\text{Mpc}$ (current state of the art: e.g. Beutler et al, based on models by Saito, Taruya, Scoccimarro)

To go beyond one needs more PT biasing parameters. New NL model achieves 1% to $k=0.4\,h/\text{Mpc}$ by introducing many physical parameters in the PT model: central and satellite galaxies, each with a bias and FoG, 1-halo contribution from central-satellite pairs, halo exclusion…”
Covariance matrix

- Simulations have a hard time converging on covariance matrix, its inverse is “hard”: e.g. 12,000 simulations in Blot et al. 2014
-Disconnected part: “gaussian” is easy: we should compute it analytically using window functions (note: this is not done currently)
- Connected part: smooth response to long wavelength modes

\[
\text{Cov}(P(k_i), P(k_j)) = P(k_i)P(k_j)V^{-1}\left(\frac{4\pi^2}{k_i^2\Delta k} \delta_{ij} + \delta_{A0}^2\right)
\]
PT approach to Covariance

- Modes from outside the survey (do not average to zero): tree level effects from survey window function very important (supersample variance), easy to calculate, depend on whether the mean density is computed from within the survey or not (Li, Takada, Hu 2014)

\[
\delta \ln P(k) = \left( \frac{47}{21} - \frac{1}{3} \frac{d \ln P}{d \ln k} \right) \delta_b = \left( \frac{68}{21} - \frac{1}{3} \frac{d \ln (k^3 P)}{d \ln k} \right) \delta_b
\]

- Use $26/21$ instead of $68/21$ for local mean density

- Modes inside the survey (average to zero): use PT trispectrum
PT trispectrum

- Tree-level calculation (Scoccimarro et al. 1999)

\[
C_{ij} = \langle \hat{P}(k_i) \hat{P}(k_j) \rangle - \langle \hat{P}(k_i) \rangle \langle \hat{P}(k_j) \rangle = V_f \left[ \frac{2P_i^2}{V_s(k_i)} \delta_{ij} + \bar{T}(k_i, k_j) \right]
\]

\[
\bar{T}(k_i, k_j) = \int_{k_i} \int_{k_j} \frac{d^3k_1}{V_s(k_i)} \frac{d^3k_2}{V_s(k_j)} T(k_1, -k_1, k_2, -k_2)
\]

\[
\bar{T}(k_i, k_j) = \int_{k_i} \int_{k_j} \frac{d^3k_1}{V_s(k_i)} \frac{d^3k_2}{V_s(k_j)} \left[ 12F_3(k_1, -k_1, k_2) P_1^2 P_2 + 8F_2^2(k_1 - k_2, 2k_2) P(|k_1 - k_2|) P_2^2 + 16F_2(k_1 - k_2, k_2) F_2(k_2 - k_1, k_1) P_1 P_2 P(|k_1 - k_2|) + (k_1 \leftrightarrow k_2) \right]
\]

- 1-loop terms: sample variance of low k modes (Mohammed & US)

\[
P = P_{22} + P_{13} = \left\{ \frac{2519}{2205} P_{S_0}(k) - \frac{47}{105} k P'_{S_0}(k) + \frac{1}{10} k^2 P''_{S_0}(k) \right\} \langle \delta^2_L \rangle
\]

\[
\frac{\text{Cov}_{ij}}{P(k_i)P(k_j)} = \left( \frac{1}{\pi^2} \int P^2_{\text{Lin}}(k) k^3 d \ln k \right) W_i W_j
\]

\[
W_i = \frac{2519}{2205} E_2(k_i) - \frac{47}{105} \frac{d \ln P(k_i)}{d \ln k_i} + \frac{1}{10} \frac{d^2 \ln P(k_i)}{d \ln k_i^2}
\]

Large contribution from low k, hence large volume needed
PT vs simulations

$k' = 0.10$

- Li et al.
- Gaussian
- One-loop
- SSC (local)
- Tree level
- Gaussian + One-loop + Tree level + SSC (local)

$k \, [\text{h/Mpc}]$
Lessons: connected part of covariance is very correlated. For $k>0.2h/\text{Mpc}$ all the modes are strongly correlated. Correlations generated by long wavelength modes: beware of jackknife/bootstrap methods.
Covariance matrix as an external parameter

- Most of the connected covariance comes from a small scale response to long wavelength modes.

- The connected part can be written as a single eigenmode

\[ C_{ij} = A d_i d_j, \]

where \( i \) represents \( k_i \) amplitude and \( d_i \) is a response at that \( k_i \).
Covariance as a fit for additional external nuisance parameter

Treat $d_i$ as an external nuisance parameter:

$$P(k_i) = P_{\text{fid}}(k_i) + \alpha d_i$$

Marginalize over $\alpha$, possibly with a theoretical prior

Non-gaussian covariance matrix reduced by a factor of 100 ($k>0.1$)

Degeneracy with amplitude at low $k$ (prior useful), lifted at higher $k$ (prior not needed)

Preliminary: same concept also works on galaxies
Conclusions

- PT is alive and well: it may be able to describe clustering up to orbit crossing (1 halo term formation), but not beyond: in 3-d PT fails even before shell crossing
- PT alone cannot give accurate results, needs to be supplemented by non-PT modeling of halos: EFT, halo model...
- PT+EFT useful at low k (k<0.2h/Mpc), halo model also at higher k
- Galaxy biasing: all terms that can exist by symmetry exist, and at the halo level may be predictable (i.e. function of halo mass)
- 1-halo terms and RSD FoG terms non-perturbative and needed
- PT+halo model successful in RSD models up to k=0.4h/Mpc
- PT 1 loop SPT successful for modeling DM covariance matrix, and suggests covariance can be modeled as external nuisance parameter
- Full theory (beyond PT) still not available (except simulations)