



LSS power spectrum analysis: progress report

Uroš Seljak
UC Berkeley
Stockholm, July 27 2017

Three LSS challenges

- Perturbative modeling of galaxy RSD (Hand etal)
- Removal of systematics (Pinol etal, Hand etal)
- Covariance matrix challenge (Mohammed etal, Li etal)

Collaborators: N. Hand, Z. Vlah, Y. Feng, Y. Li, Z. Slepian, I. Mohammed, F. Beutler, L. Pinol. B. Cahn, M. White

RSD: 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: Vlah et al and Okumura et al papers based on distribution function approach (Seljak & McDonald 2009)
- There is always more information on small scales, but some of it is hopelessly corrupted by nonlinear effects that cannot be modeled in PT: need to parametrize our ignorance of small scales
- Our ignorance must obey all symmetries (e.g. $k^2 P_L(k)$ 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+PT₃

PT modeling of BOSS/DESI redshift space clustering

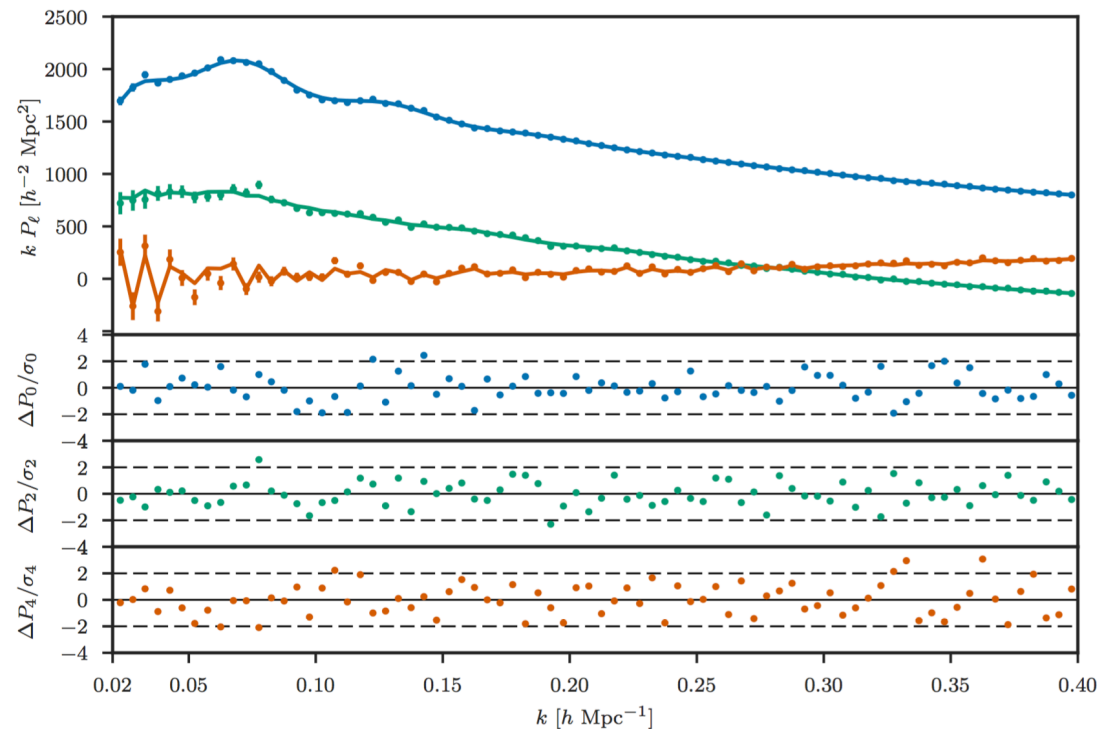
Hand, Seljak et al 2017

Usual PT terms at 1-loop level, plus a number of ignorance terms

Our model achieves 1-2% residual error to $k=0.4h/\text{Mpc}$ by introducing many (13) physical parameters in the PT model: central and satellite galaxy fractions, each with a bias and Fingers-of-God damping, 1-halo contribution from central-satellite pairs, halo exclusion...

A factor of 2 improvement in scale relative to previous work

Conclusion: one needs a lot of astrophysical parameters on small scales to supplement PT



Information gain with scale for BOSS

	$k_{\max} = 0.2 \text{ h/Mpc}$	$k_{\max} = 0.3 \text{ h/Mpc}$	$k_{\max} = 0.4 \text{ h/Mpc}$	truth
$f\sigma_8$	$0.455^{+0.015}_{-0.015}$	$0.467^{+0.012}_{-0.013}$	$0.469^{+0.010}_{-0.009}$	0.472
$f\sigma_8$ [fixed AP]	$0.457^{+0.010}_{-0.010}$	$0.466^{+0.009}_{-0.008}$	$0.468^{+0.007}_{-0.007}$	0.472
$\alpha_{\perp}$	$1.003^{+0.005}_{-0.006}$	$1.004^{+0.005}_{-0.005}$	$1.005^{+0.004}_{-0.004}$	1.000
$\alpha_{\parallel}$	$1.009^{+0.010}_{-0.009}$	$1.006^{+0.008}_{-0.007}$	$1.009^{+0.007}_{-0.008}$	1.000
$b_1\sigma_8$	$1.266^{+0.009}_{-0.009}$	$1.265^{+0.008}_{-0.008}$	$1.268^{+0.008}_{-0.008}$	1.272
f_s	$0.122^{+0.019}_{-0.018}$	$0.143^{+0.013}_{-0.013}$	$0.143^{+0.008}_{-0.008}$	0.104
f_{cB}	$0.104^{+0.033}_{-0.030}$	$0.124^{+0.022}_{-0.023}$	$0.122^{+0.013}_{-0.015}$	0.089
f_{sB}	$0.438^{+0.203}_{-0.197}$	$0.438^{+0.136}_{-0.123}$	$0.466^{+0.081}_{-0.079}$	0.399
σ_c	$1.134^{+0.214}_{-0.238}$	$0.906^{+0.088}_{-0.111}$	$0.930^{+0.062}_{-0.065}$	—
σ_{sA}	$4.239^{+0.476}_{-0.413}$	$3.737^{+0.372}_{-0.464}$	$3.443^{+0.278}_{-0.270}$	—
$\chi^2/\text{d.o.f.}$	$113/(108 - 13) = 1.19$	$159/(168 - 13) = 1.03$	$241/(228 - 13) = 1.12$	

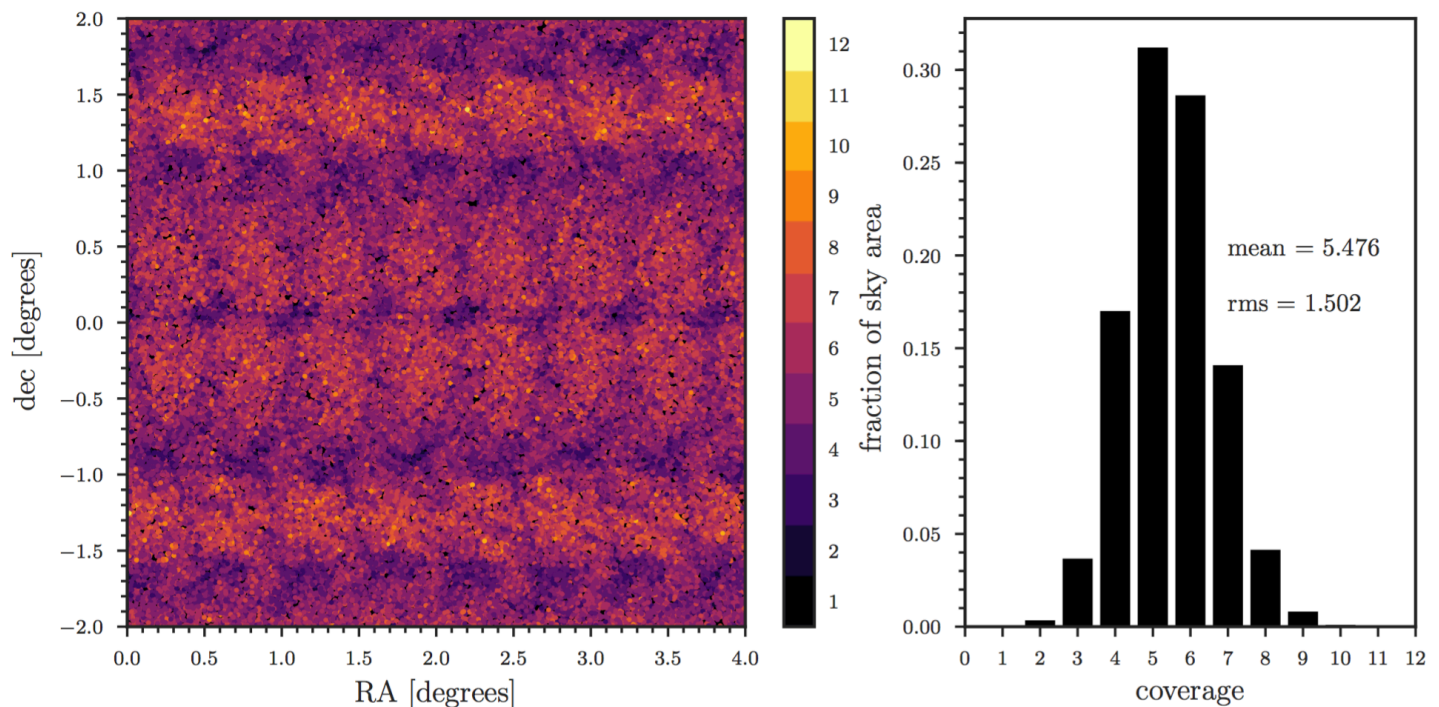
Hand, Seljak et al, 2017

- At the same $k_{\max}$ errors 10-20% larger than current analyses: price one pays for more parameters
- With $k_{\max}=0.4\text{h/Mpc}$ we expect 30% improvement relative to $k=0.2\text{h/Mpc}$ for BOSS
- Improvements could be better for ELGs (less shot noise)

Systematics: DESI target selection

- DESI has on average 5 passes over each patrol area (5 arcmin^2), but sometimes more or less. Mean number of ELGs is 3.5 per patrol area: not every ELG will have redshift determined
- A transverse modulation, this one can be modeled with randoms

Pinol et al 2016

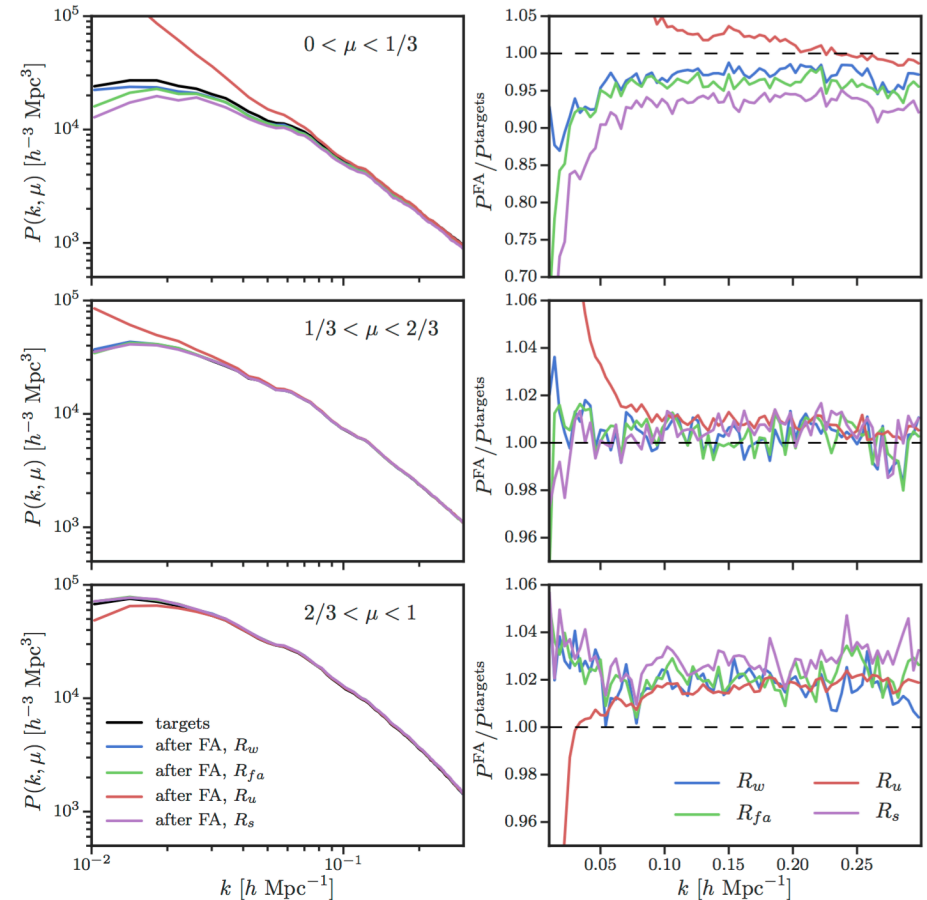


How to compute randoms?

- What happens if we have parts of a survey with varying depths? We need to compute randoms for each survey depth i , normalizing properly each separately: $w_i = N_{\text{gal},i} / N_{\text{ran},i}$
- This is the situation for DESI: we will have different coverage, from 0 to 11 passes: different depths
- Randoms have to be computed separately for each coverage (Pinol et al 2016): we need to know coverage
- There are other randoms one can define, for example passing randoms through fiber assignment procedure, or just use galaxy positions and randomized redshifts (Burden et al 2016)
- They are more ad hoc, but also reduce systematics

Results on wedges after 5 passes

- “Correct randoms” give best results (blue)
- Uniform randoms worst results (red)
- Over most range of k effects of order 1-3% for correct randoms
- No data should be thrown away because of this: easy to correct 1-3% effects with a simulation
- No increase in variance



How to handle transverse (angular) systematics: $\mu=0$?

Many systematics (PSF, extinction, star-galaxy separation) are transverse

Hand et al 2017

By removing $\mu=0$ bin we remove them

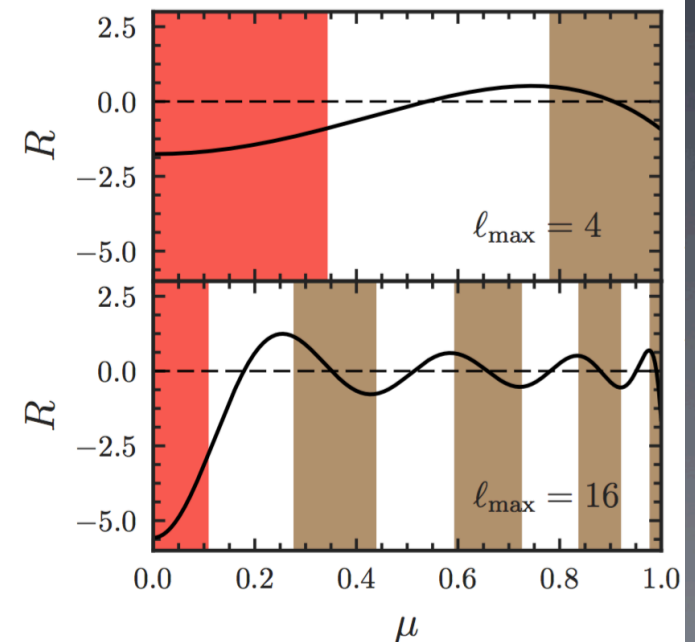
Problem: the analysis is done with multiple moments, $\mu=0$ cannot be completely isolated

Solution: we rotate multipoles to construct wedges with intervals such that $\mu=0$ contribution is explicitly cancelled except for 1st bin

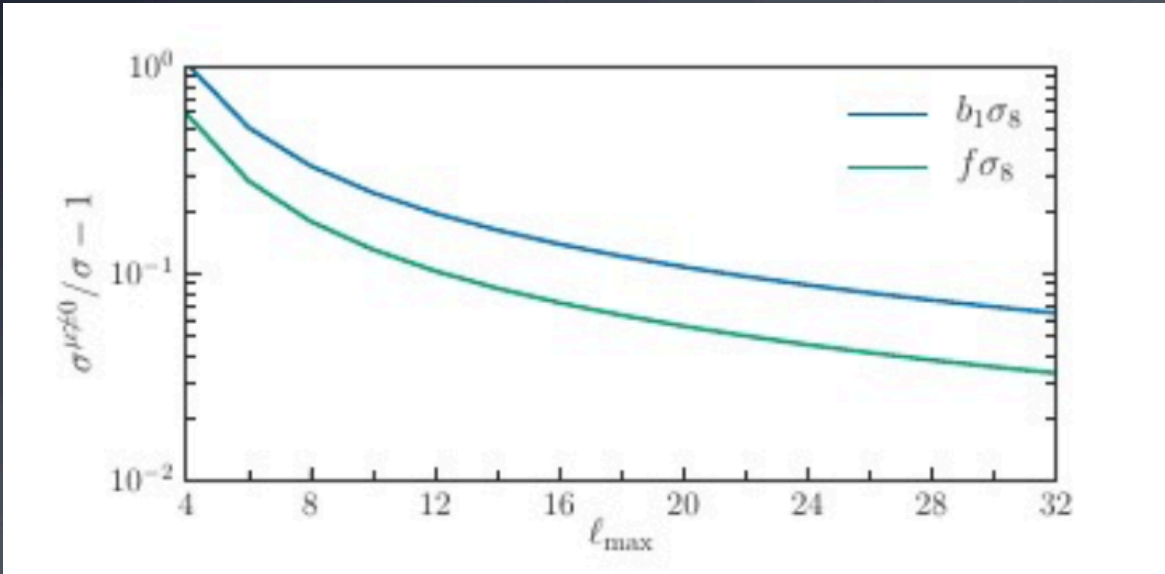
Using spherical harmonic decomposition instead of Cartesian is faster than previous implementations: $2l+1$ instead of $(l+1)(l+2)/2$

At $l_{\max}=16$ we need 153 FFTs instead of 525

$$\begin{aligned}\hat{P}(k, \mu) &= \hat{P}^{\text{obs}}(k, \mu) + P_c(k) \sum_{\ell=0}^{\ell_{\max}} \frac{2\ell+1}{2} \mathcal{L}_{\ell}(0) \mathcal{L}_{\ell}(\mu), \\ &= \sum_{\ell=0}^{\ell_{\max}} \hat{P}_{\ell}(k) \mathcal{L}_{\ell}(\mu) + P_c(k) \frac{\ell_{\max}+1}{2} \mathcal{L}_{\ell_{\max}+1}(0) \frac{\mathcal{L}_{\ell_{\max}+1}(\mu)}{\mu}\end{aligned}$$



Fisher information analysis



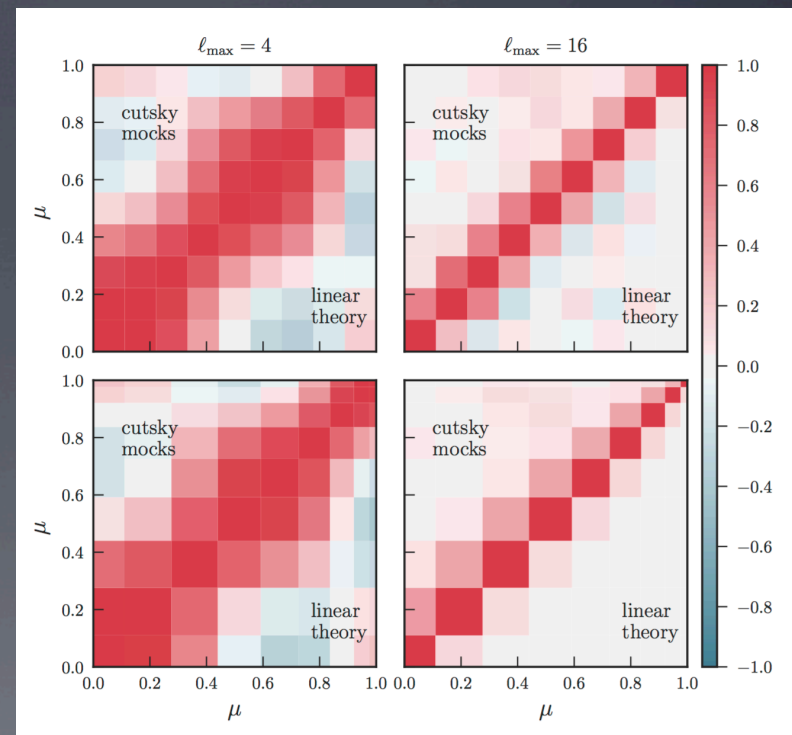
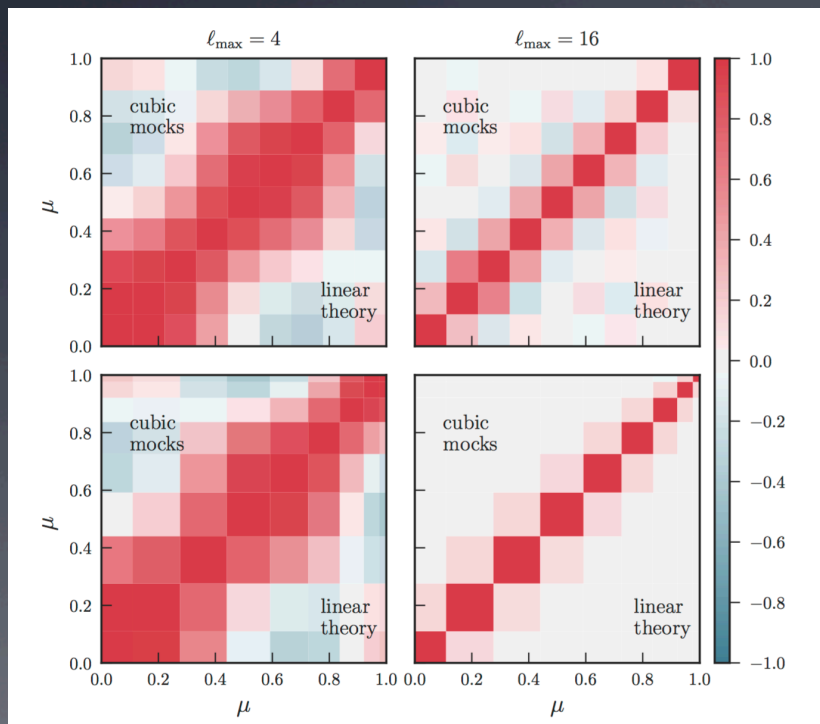
Hand et al 2017

With $\ell_{\max}=16$ we increase error on $f\sigma_8$ by 8% (vs 50% for $\ell_{\max}=4$)

By doing this analysis to $\ell_{\max}=16$ we have removed most of the systematics while preserving most of the information

Gaussian covariance matrix

- Non-uniform wedges result in a simpler analytic covariance matrix in the gaussian approximation
- periodic window



- Almost diagonal, $2P^2$, while for uniform wedges it is singular (high condition number) for $\ell_{\max} > 4$
- Remaining issue: window

Covariance matrix challenge

- 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: trispectrum (smooth response to long wavelength modes)

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

Mohammed & US 2014

PT approach to NG Covariance

- Modes from outside the survey (do not average to zero): tree level effects from survey window function very important (beat coupling or supersample covariance), 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
- Can be calibrated numerically with separate universe simulations
- Modes inside the survey (average to zero): use PT trispectrum

PT trispectrum: tree level

- Tree-level calculation (Scoccimarro et al 1999)

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

$$\bar{T}(k_i, k_j) = \int_{k_i} \frac{d^3 \mathbf{k}_1}{V_s(k_i)} \int_{k_j} \frac{d^3 \mathbf{k}_2}{V_s(k_j)} T(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_2)$$

$$\begin{aligned} \bar{T}(k_i, k_j) = & \int_{k_i} \frac{d^3 \mathbf{k}_1}{V_s(k_i)} \int_{k_j} \frac{d^3 \mathbf{k}_2}{V_s(k_j)} \left[12F_3(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_2) P_1^2 P_2 + 8F_2^2(\mathbf{k}_1 - \mathbf{k}_2, \mathbf{k}_2) P(|\mathbf{k}_1 - \mathbf{k}_2|) P_2^2 \right. \\ & \left. + 16F_2(\mathbf{k}_1 - \mathbf{k}_2, \mathbf{k}_2) F_2(\mathbf{k}_2 - \mathbf{k}_1, \mathbf{k}_1) P_1 P_2 P(|\mathbf{k}_1 - \mathbf{k}_2|) + (\mathbf{k}_1 \leftrightarrow \mathbf{k}_2) \right] \end{aligned}$$

- Dominates at very low k

1 loop functional derivative

- 1-loop power spectrum

$$\begin{aligned} P^{(1)}(k, z) &= P_{13}(k, z) + P_{22}(k, z) \\ &= 6 \int F_3^{(s)}(\mathbf{k}, \tilde{\mathbf{q}}, -\tilde{\mathbf{q}}) P_L(k, z) P_L(\tilde{q}, z) d^3 \tilde{q} + 2 \int \left(F_2^{(s)}(\mathbf{k} - \tilde{\mathbf{q}}, \tilde{\mathbf{q}}) \right)^2 P_L(|\mathbf{k} - \tilde{\mathbf{q}}|, z) P_L(\tilde{q}, z) d^3 \tilde{q} \end{aligned}$$

- Functional derivatives

$$\begin{aligned} N \frac{\delta P_{1-\text{loop}}(\mathbf{k})}{\delta P_L(\mathbf{q})} &= N \frac{\delta P_{22}(\mathbf{k})}{\delta P_L(\mathbf{q})} + 2N \frac{\delta P_{13}(\mathbf{k})}{\delta P_L(\mathbf{q})} \\ &= 4 \left(F_2^{(s)}(\mathbf{k} - \mathbf{q}, \mathbf{q}) \right)^2 P_L(|\mathbf{k} - \mathbf{q}|) + 6 F_3^{(s)}(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_L(k). \end{aligned}$$

$$\mathbf{V}(q, k) = \frac{P_L(q)}{\Delta^2(q)} \left\langle \frac{\delta P_{1-\text{loop}}(\mathbf{k})}{\delta P_L(\mathbf{q})} \right\rangle_{\Omega}$$

Nishimichi et al 2015

low k limit

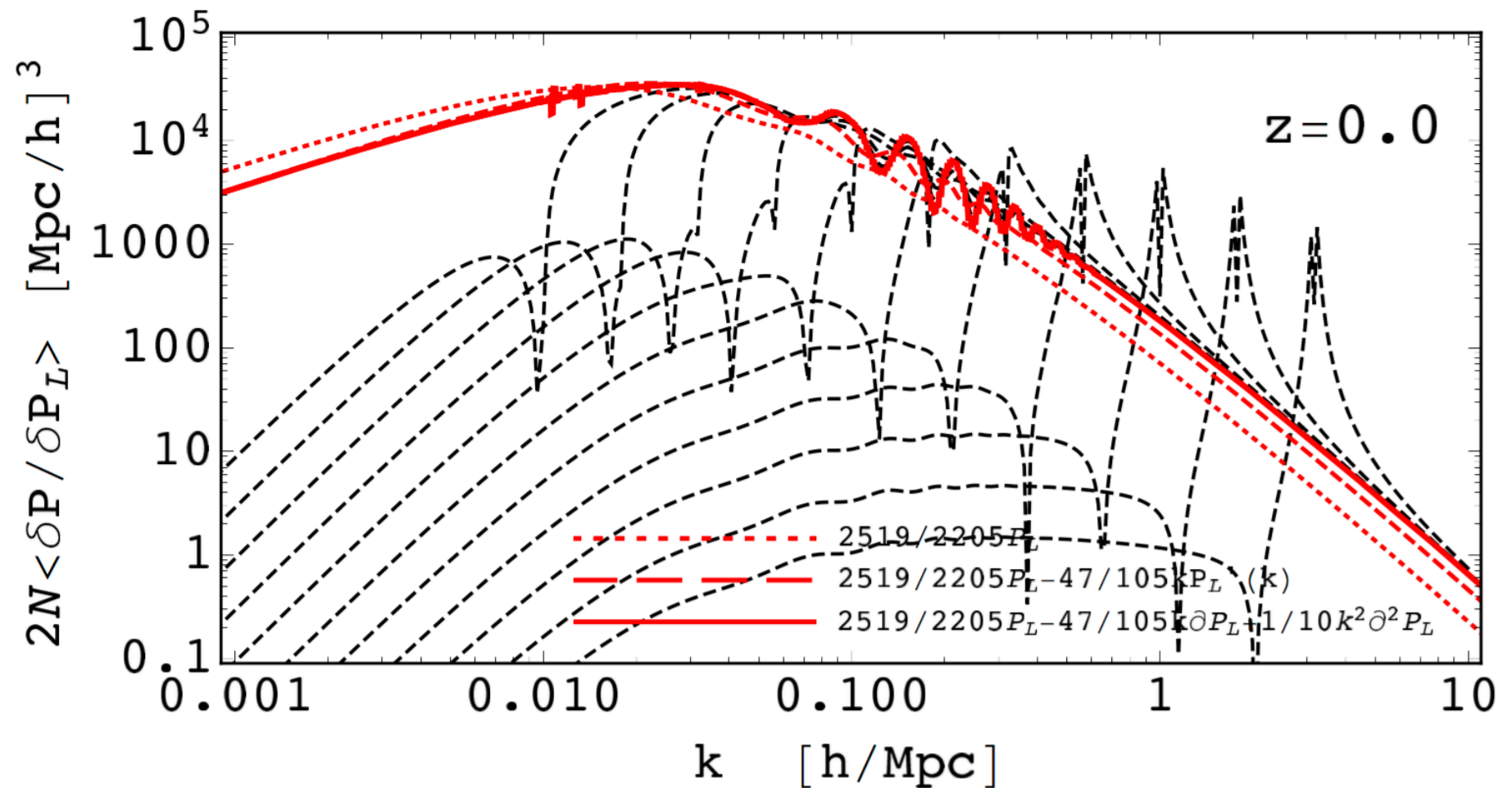
- 1-loop terms: sample variance of low k modes

$$\lim_{q/k \rightarrow 0} \mathbf{V}(q, k) = W(k)P(k)$$

$$P = P_{22} + P_{13} = \left\{ \frac{2519}{2205} P_{S0}(k) - \frac{47}{105} k P'_{S0}(k) + \frac{1}{10} k^2 P''_{S0}(k) \right\} \langle \delta_L^2 \rangle$$

$$\mathbf{W}_i = \frac{2519}{2205} E_2(k_i) - \frac{47}{105} \frac{d \ln P(k_i)}{d \ln k_i} + \frac{1}{10} \frac{d^2 \ln P(k_i)}{d \ln k_i^2}$$

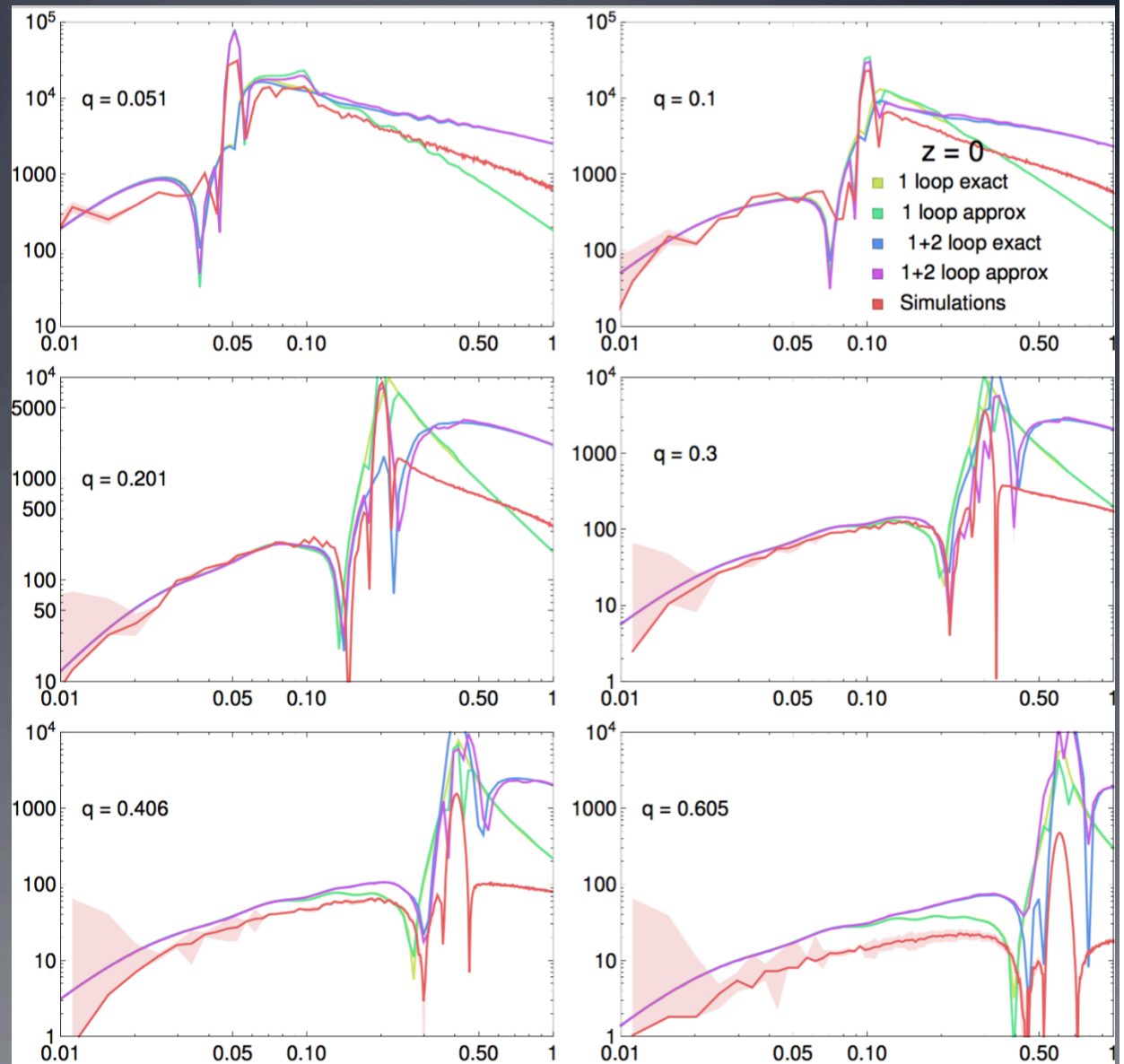
Functional derivatives



Simulations vs PT: functional derivatives

PT fails at high k
EFT improvements
modest (not shown)
PT fails at high q :
damping
(Nishimichi et al)

Kozarski, Vlah,
US, in prep



1 loop covariance

- 1-loop terms: sample variance of low k modes

$$\frac{\mathbf{Cov}_{ij}}{P(k_i)P(k_j)} = \left(\frac{1}{\pi^2} \int P_{\text{Lin}}^2(k) k^3 d \ln k \right) \mathbf{W}_i \mathbf{W}_j$$

$$S = \left(\frac{1}{V\pi^2} \int P_{\text{Lin}}^2(q) q^2 dq \right)$$

- In general

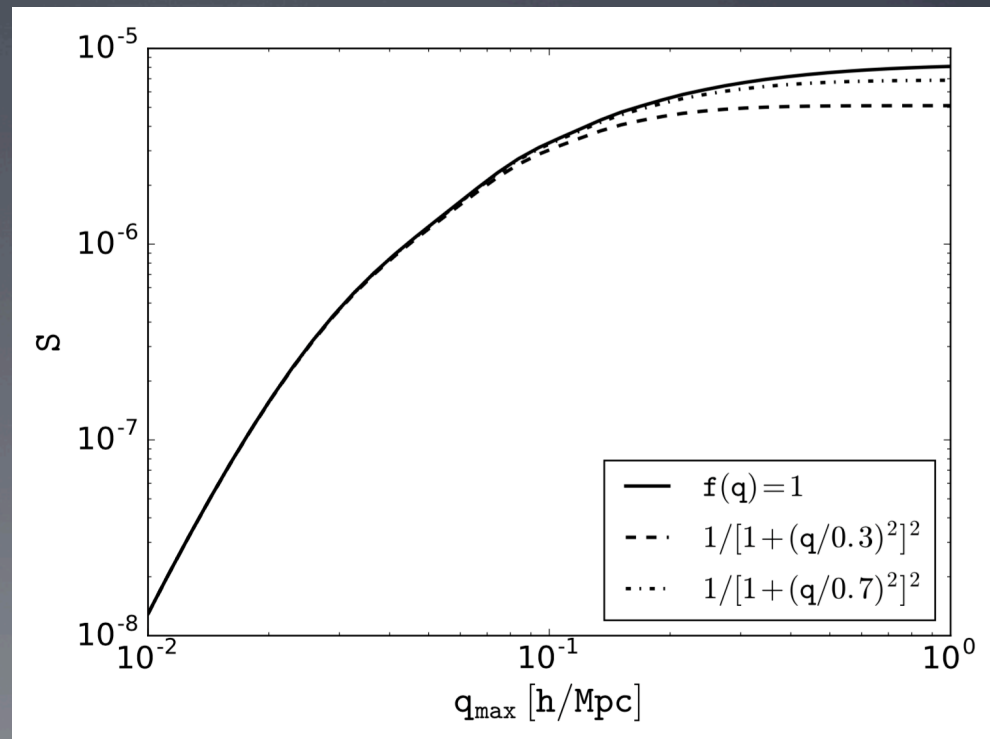
$$\mathbf{Cov}_{ij}^{1\text{-loop}} = \left(\frac{1}{V\pi^2} \int P_{\text{Lin}}^2(q) q^2 \mathbf{V}(q, k_i) \mathbf{V}(q, k_j) dq \right)$$

High q damping

- At high q response is suppressed (Nishimichi et al)

$$\mathbf{Cov}_{ij}^{\text{NL}} = \left(\frac{1}{V\pi^2} \int \frac{P_{\text{Lin}}^2(q)}{(1 + (q/q_{\text{nl}})^2)^2} q^3 \mathbf{V}(q, k_i) \mathbf{V}(q, k_j) d\ln q \right)$$

- Suppresses high q modes
- Small effect at high z

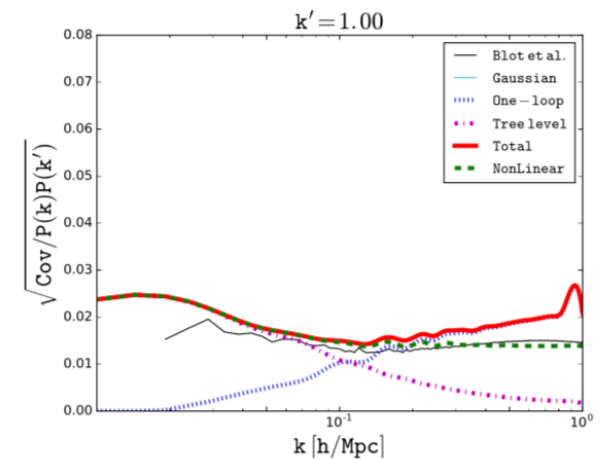
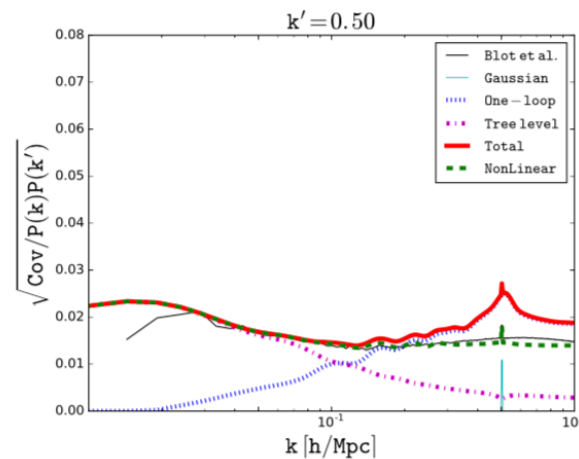
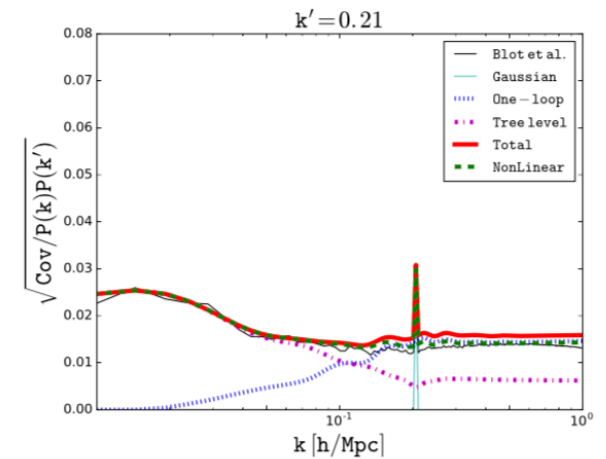
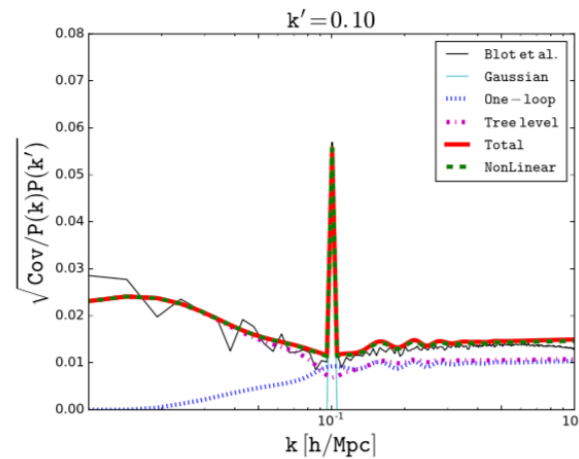
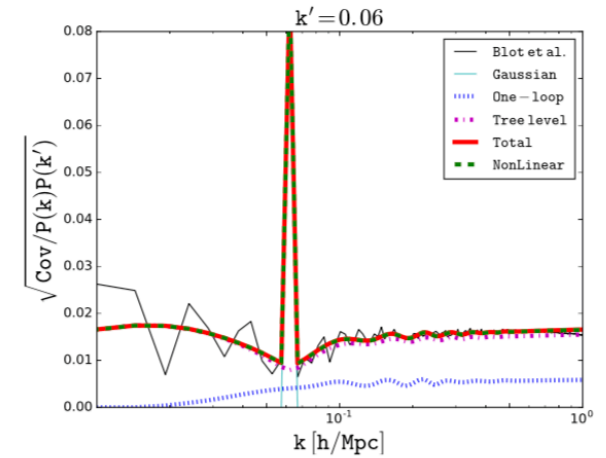
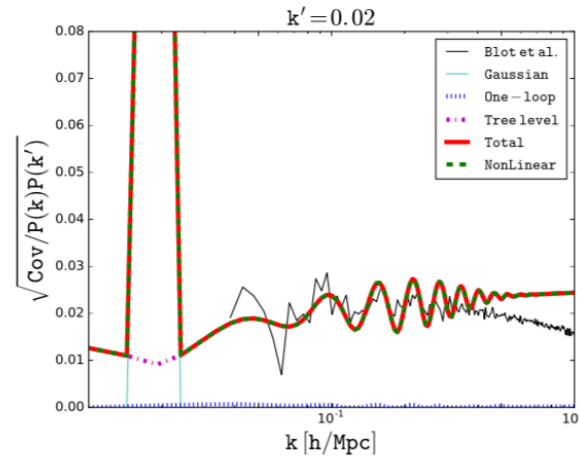


PT vs simulations

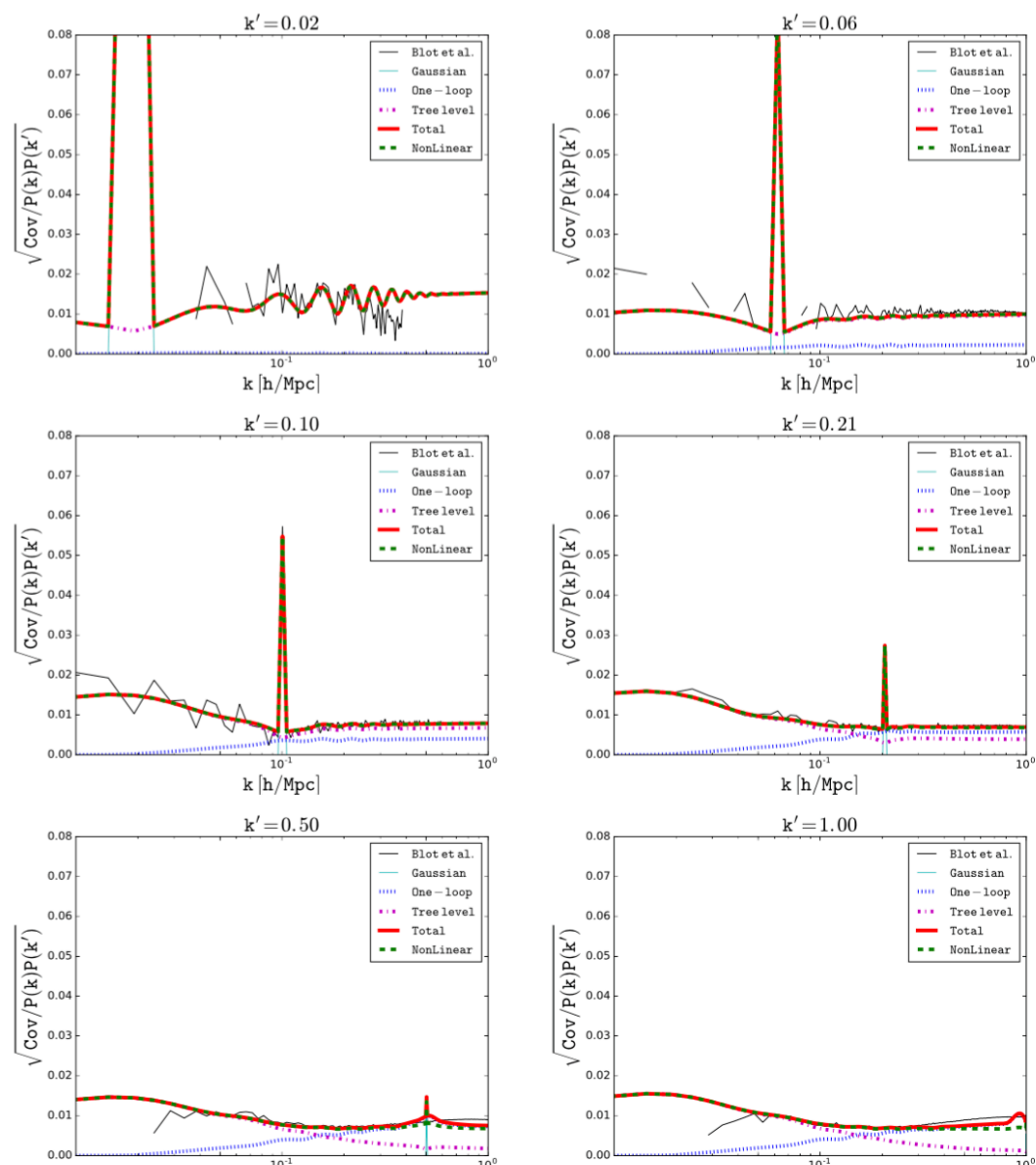
$Z=0$

Mohammed, US, Vlah
2016

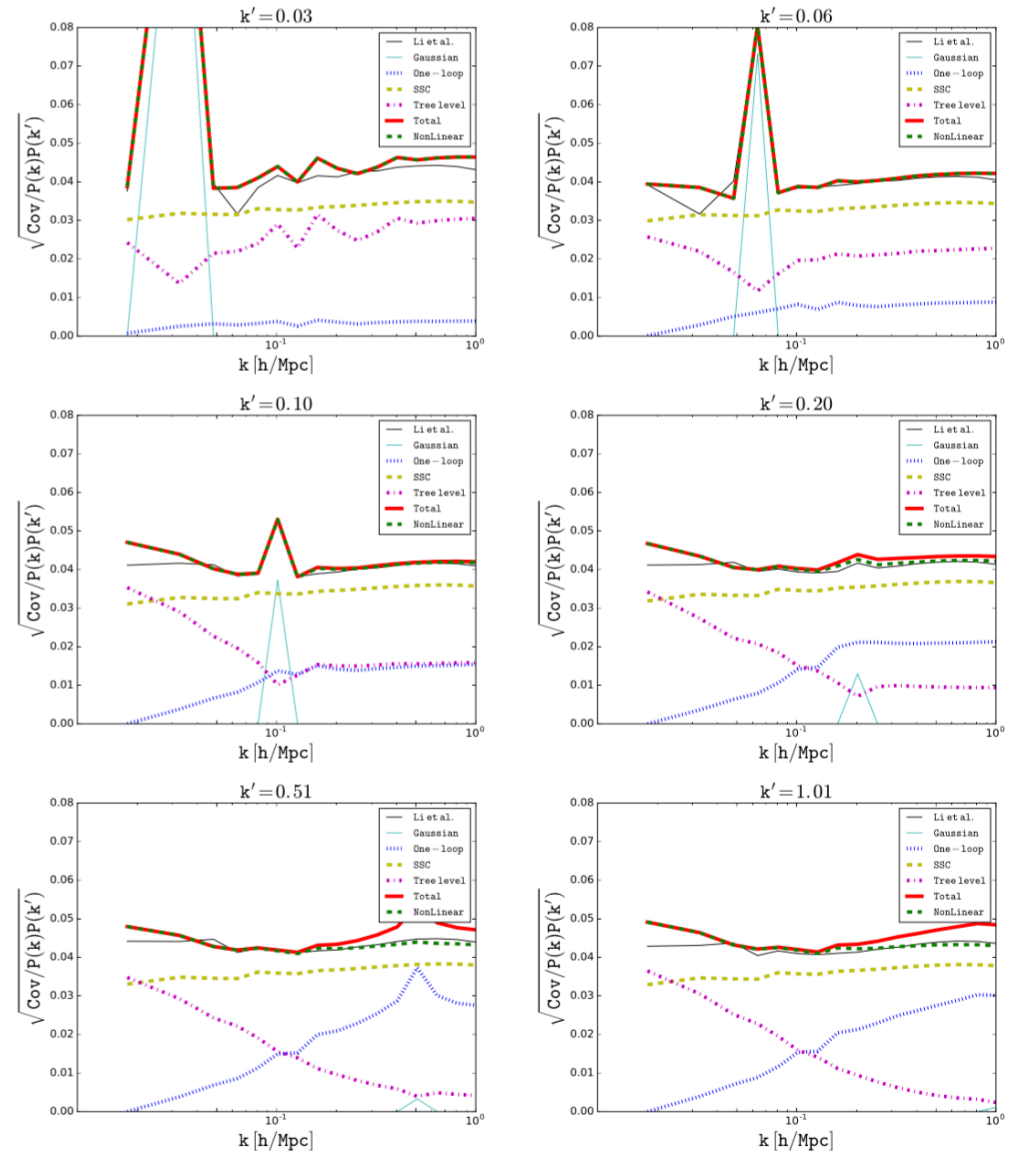
Excellent agreement
Much better than it
should be given that PT
fails at high k



PT vs sims z=1



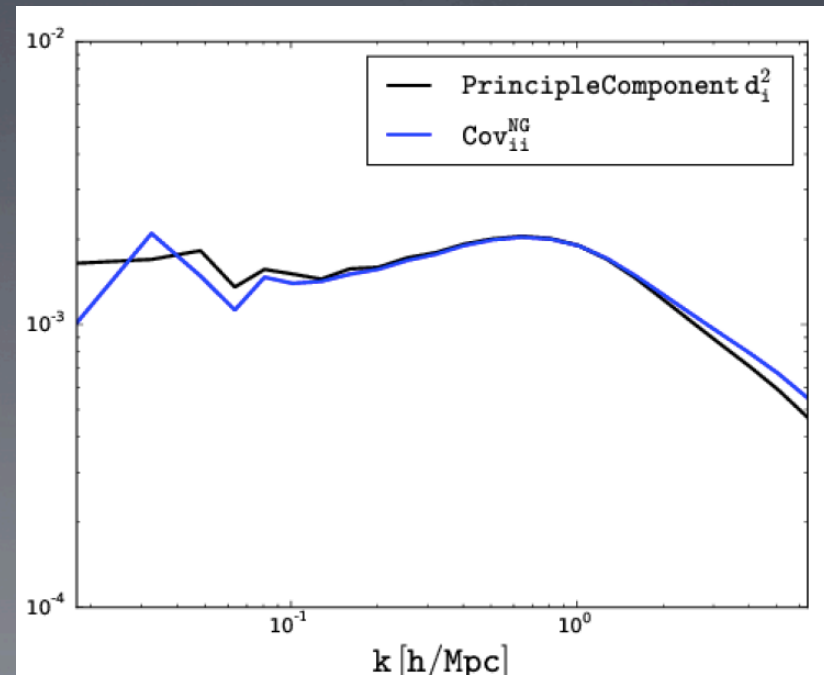
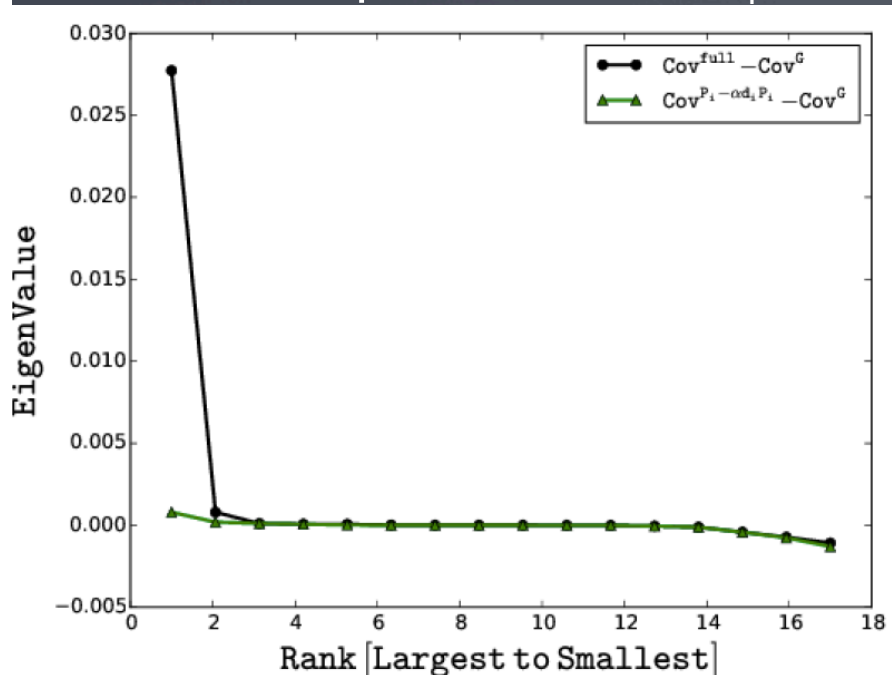
Total covariance (with beat coupling or SSC)



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

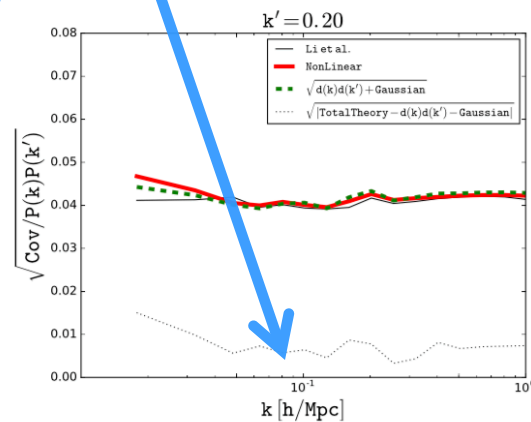
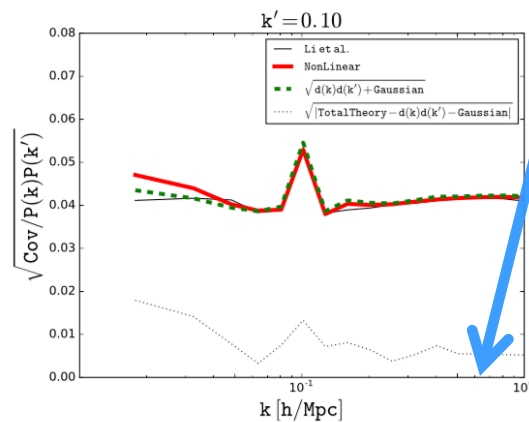
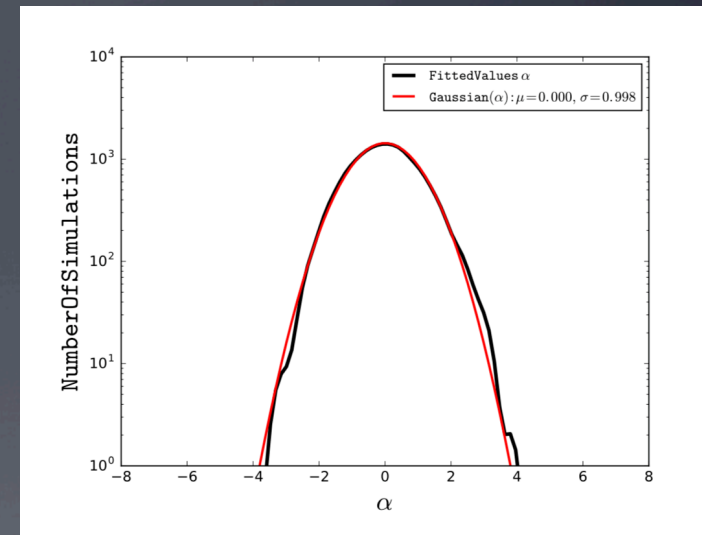
$C_{ij} = \langle \alpha^2 \rangle d_i d_j$, where i represents k_i amplitude and d_i is a response at that k_i



PDF of largest eigenmode

$$P'(k_i) = P(k_i) - \alpha d_1(k_i) \langle P(k_i) \rangle$$

- We can determine α from each realization
- PDF is gaussian
- dominated by SSC
- Residual connected term is small

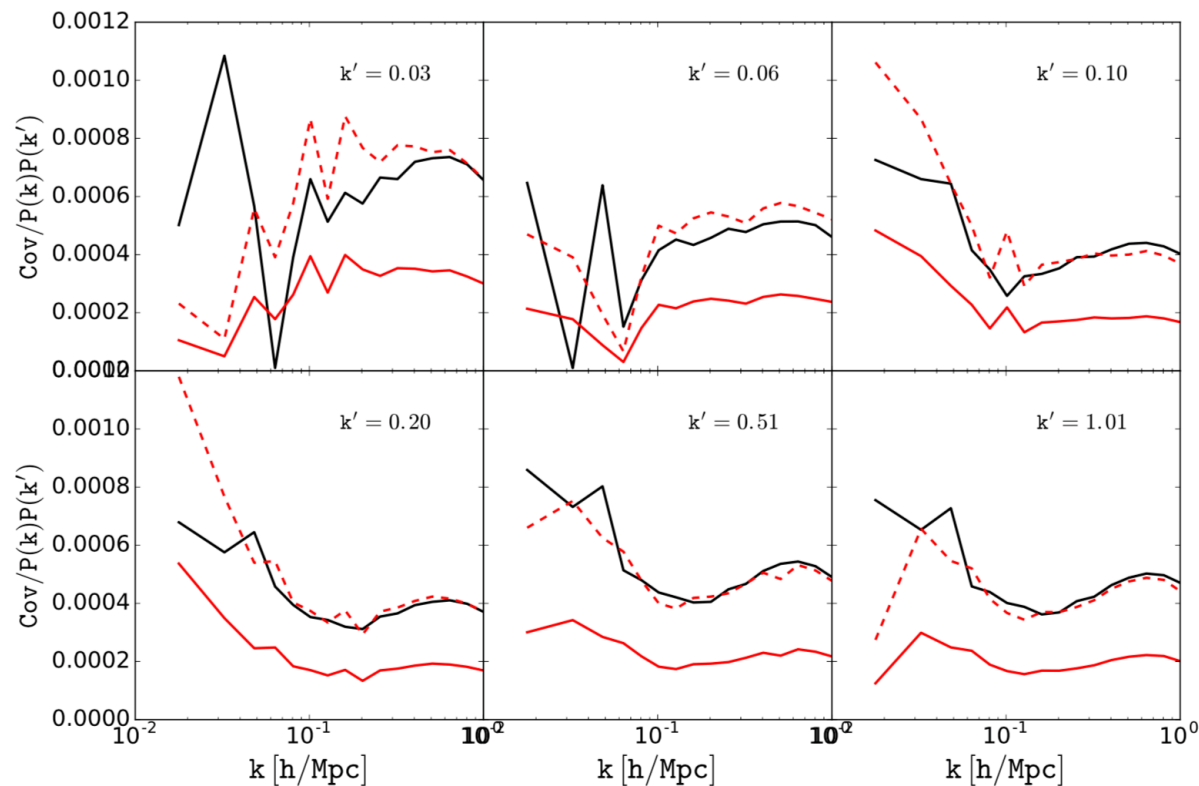


How to get covariance matrix from simulations

- Observed survey power spectrum is a true power spectrum convolved with the survey window
- True power spectrum unrelated to the survey volume and can be determined from small simulations
- Survey covariance matrix is a true trispectrum convolved with the survey window
- Trispectrum unrelated to the survey volume and can be determined from a small volume
- Special care needed to take our beat coupling effect (SSC)

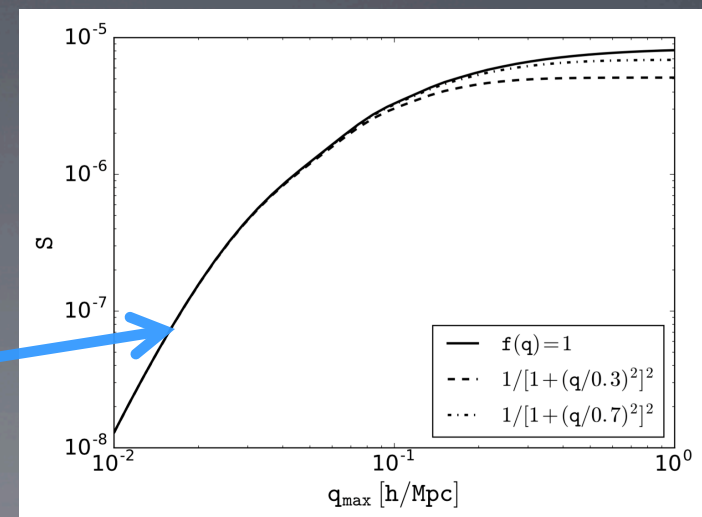
Small vs large volume simulations

- Blot et al $(650\text{Mpc}/h)^3$ versus Li et al $(500\text{Mpc}/h)^3$
- Rescaling by volume works very well
- The two simulations basically agree



Implications and future directions

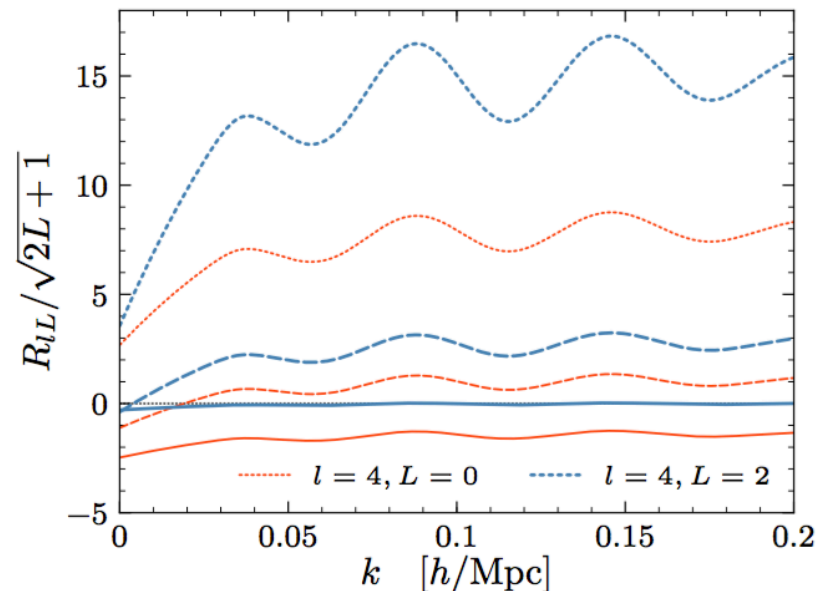
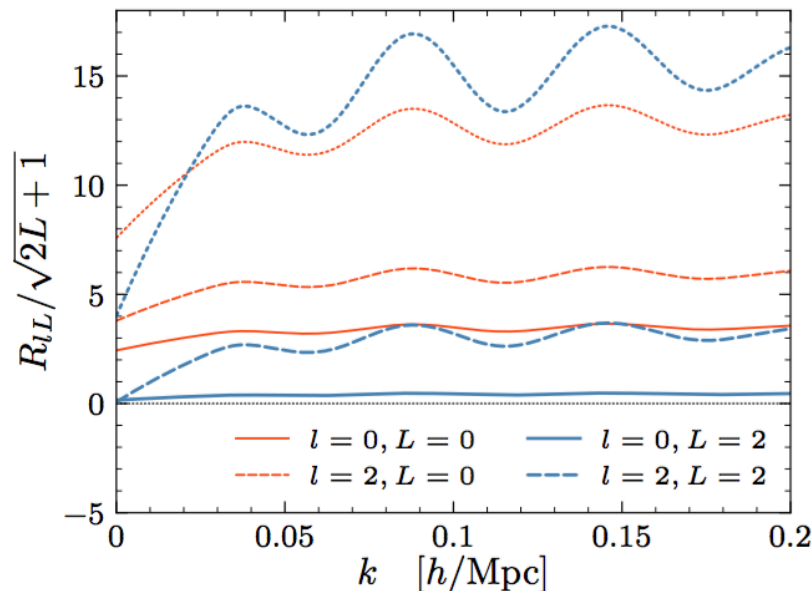
- PT approaches quite successful for dark matter
- For high k one probably needs simulations
- For galaxies: PT or simulations?
- Covariance matrix has a disconnected term ($2P^2$): analytic
- Connected term: trispectrum
- No need to relate trispectrum to the survey
- Need all the modes: minimum box size 300Mpc/h



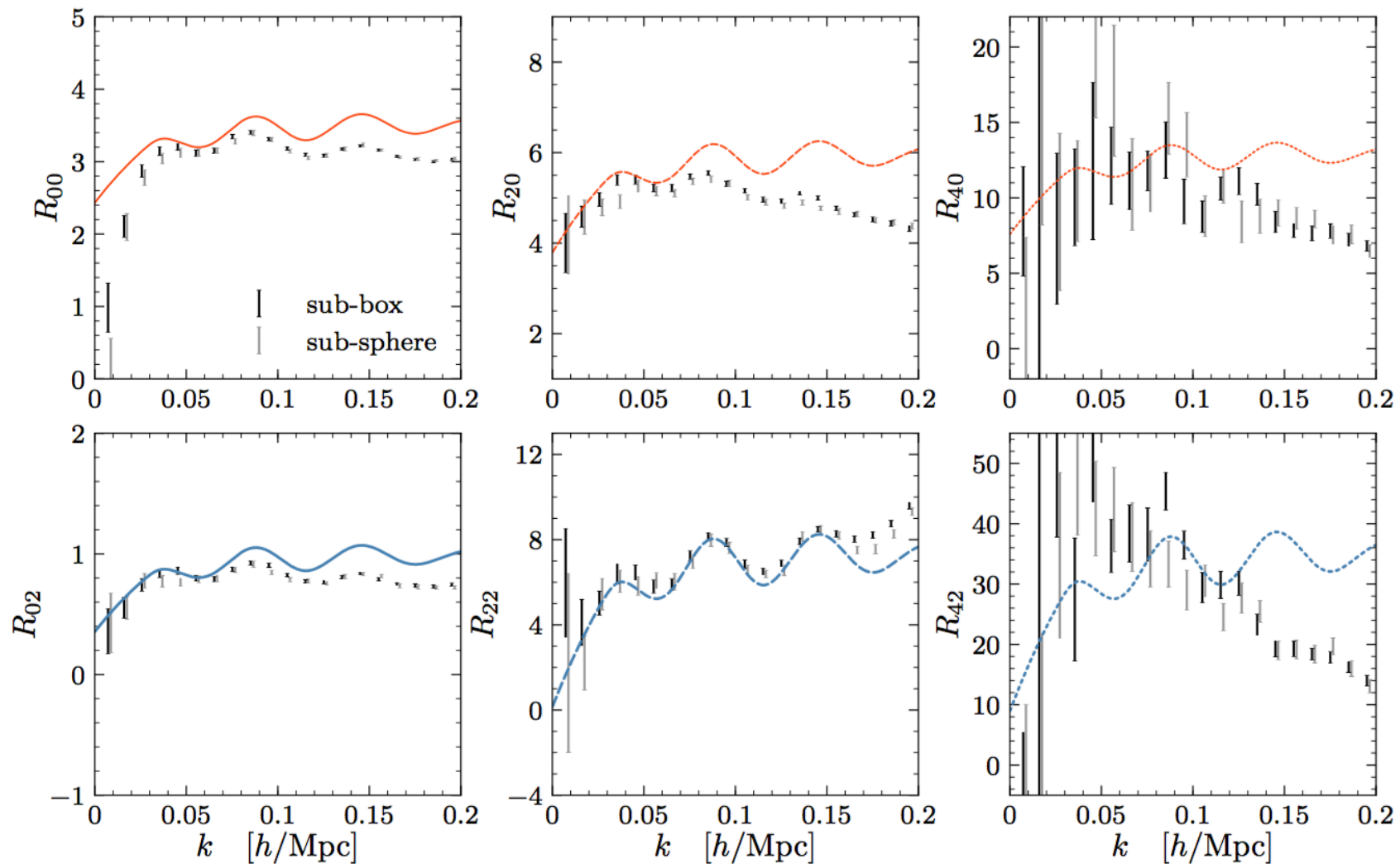
What about galaxies/RSD?

- SSC and trispectrum both important, SSC analytic
- We also need to model tidal responses etc (Li et al, in prep)
- Local SSC terms can be negative or zero due to b and f

$$R_{lL}(k) = G_{lL} + D_{lL} \frac{d \ln k^3 P}{d \ln k} - (2b_1 + \frac{2}{3}f)\delta_{L0} - \frac{4}{3}f\delta_{L2}$$

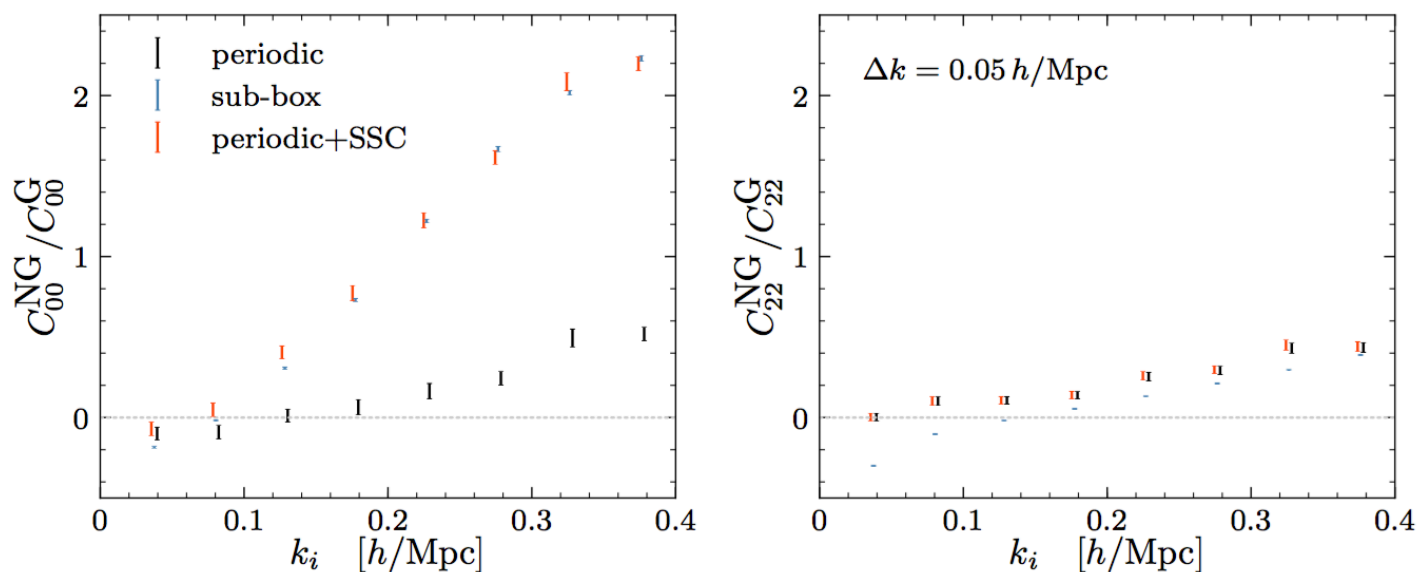


Comparison to simulations



Does NG Cov matter?

- monopole $l=0$ dominated by SSC, which can be absorbed by a change in bias
- quadrupole NG Cov dominated by bispectrum, small
- In a BOSS volume the effect of order 0.5% on f , much less for DESI
- Overall effect on f much smaller than the errors



Roadmap to future analyses

- We can compute multipoles efficiently using Yamamoto estimator and using Hand et al FFT based methods (no need for spherical harmonics Bessel basis?)
- In the presence of transverse systematics we can rotate these into non-uniform pseudo-wedges and throw away the first one
- Gaussian covariance matrix is near diagonal
- Nongaussian covariance matrix can be ignored if $k < 0.4 h/\text{Mpc}$
- Modeling of pseudo-wedges or multipoles: PT based approach has reached its limit with $k = 0.4 h/\text{Mpc}$. Current models have no residuals. Beyond that emulators may provide additional improvements