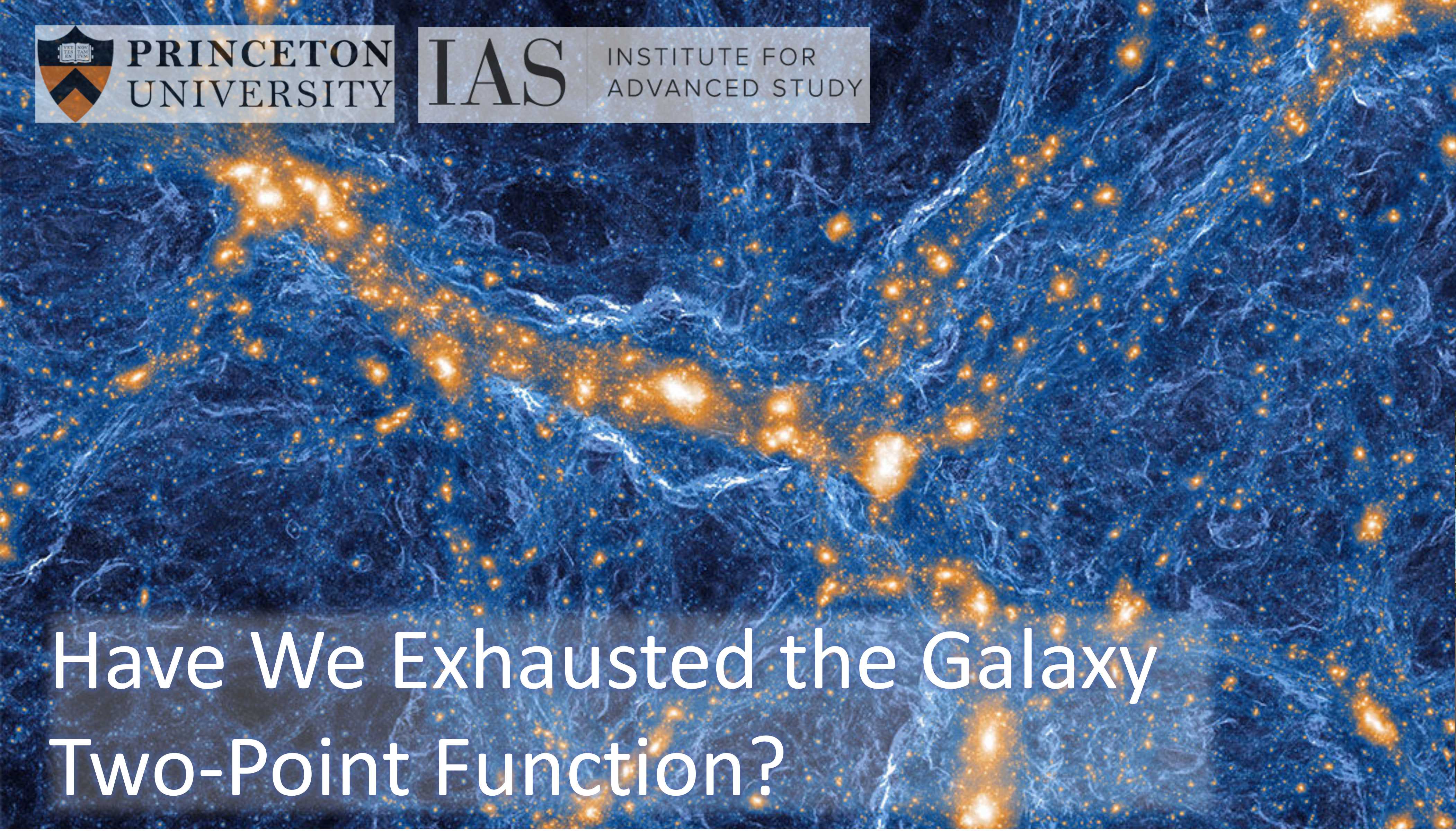
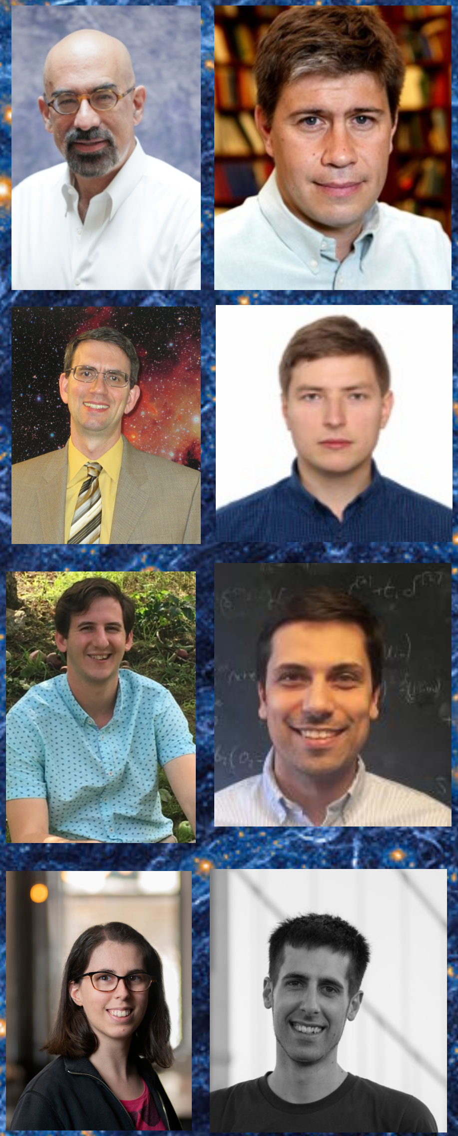


PRINCETON UNIVERSITY

IAS INSTITUTE FOR ADVANCED STUDY

Have We Exhausted the Galaxy Two-Point Function?



Oliver Philcox (Princeton / IAS)

Gravity Group, Princeton University, 4/2/21

Based on: 1912.01010, 2005.01739 2009.03311, 2006.10055, 2012.09389, 2102.08384

Extracting Information from Galaxy Surveys

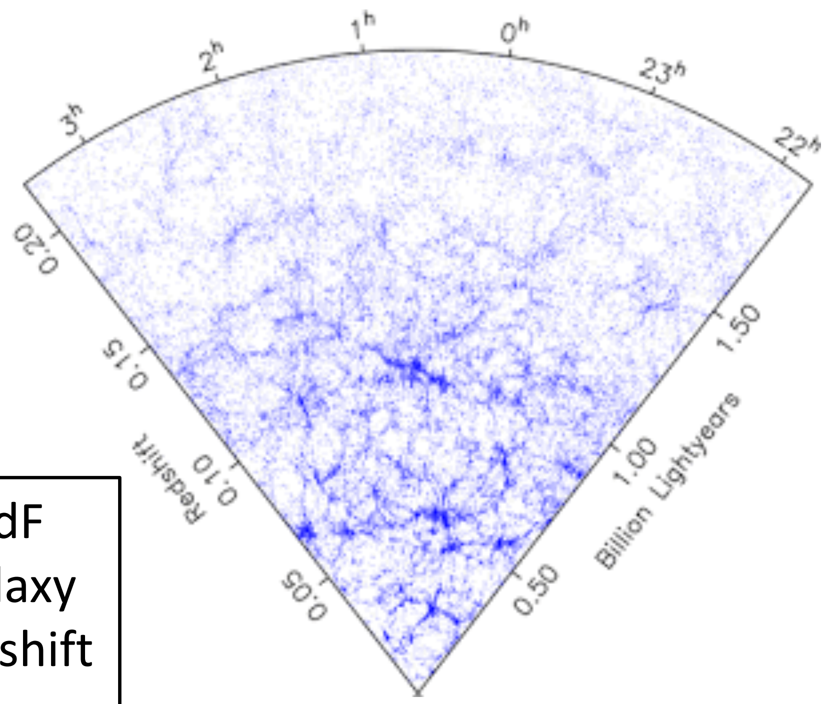
- Fundamental observable: the galaxy **overdensity** field

$$\delta(\mathbf{r}) = \frac{1}{\bar{n}} \left[n_g(\mathbf{r}) - n_r(\mathbf{r}) \right]$$

Random particle positions

Mean density

Galaxy positions



2dF Galaxy Redshift Survey

Extracting Information from Galaxy Surveys

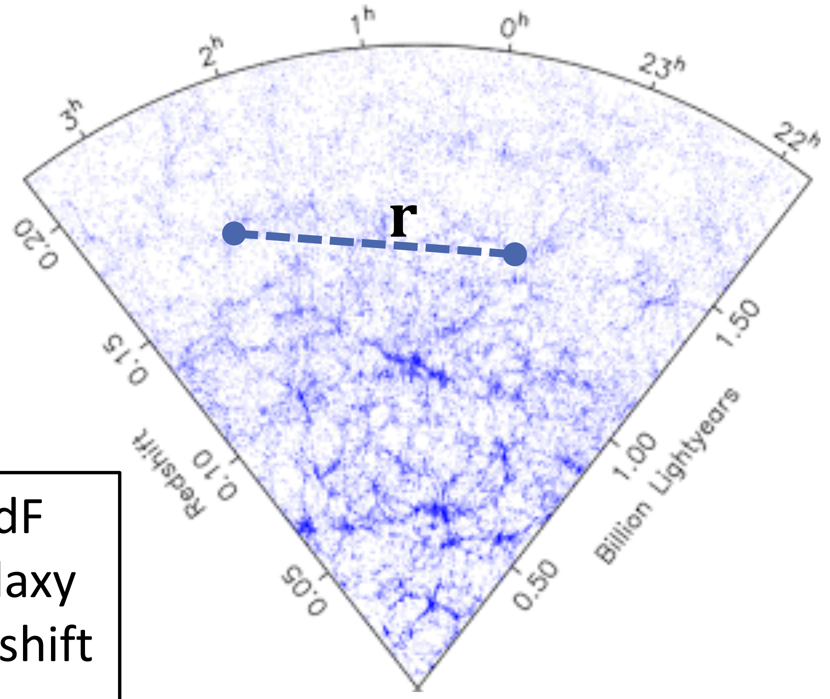
- Fundamental observable: the galaxy **overdensity** field

$$\delta(\mathbf{r}) = \frac{1}{\bar{n}}\left[n_g(\mathbf{r}) - n_r(\mathbf{r})\right]$$

Random particle positions

Mean density

Galaxy positions



2dF Galaxy Redshift Survey

- Analyze with **summary statistics:**
  - Two-point correlation function (2PCF), $\xi(\mathbf{r})$

$$\xi(r) = \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle$$

  - Power spectrum, $P(\mathbf{k})$

$$(2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')P(k) = \langle \delta(\mathbf{k})\delta(\mathbf{k}')\rangle$$

Extracting Information from Galaxy Surveys

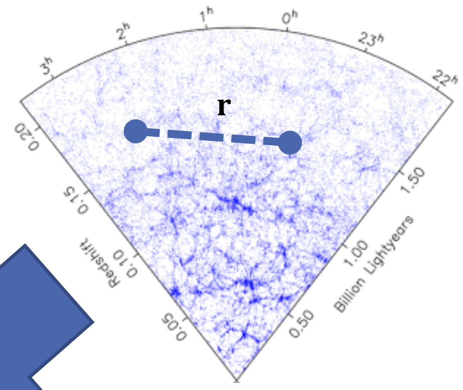
- Fundamental observable: the galaxy **overdensity** field

$$\delta(\mathbf{r}) = \frac{1}{\bar{n}}\left[n_g(\mathbf{r}) - n_r(\mathbf{r})\right]$$

Mean density

Galaxy positions

Random particle positions

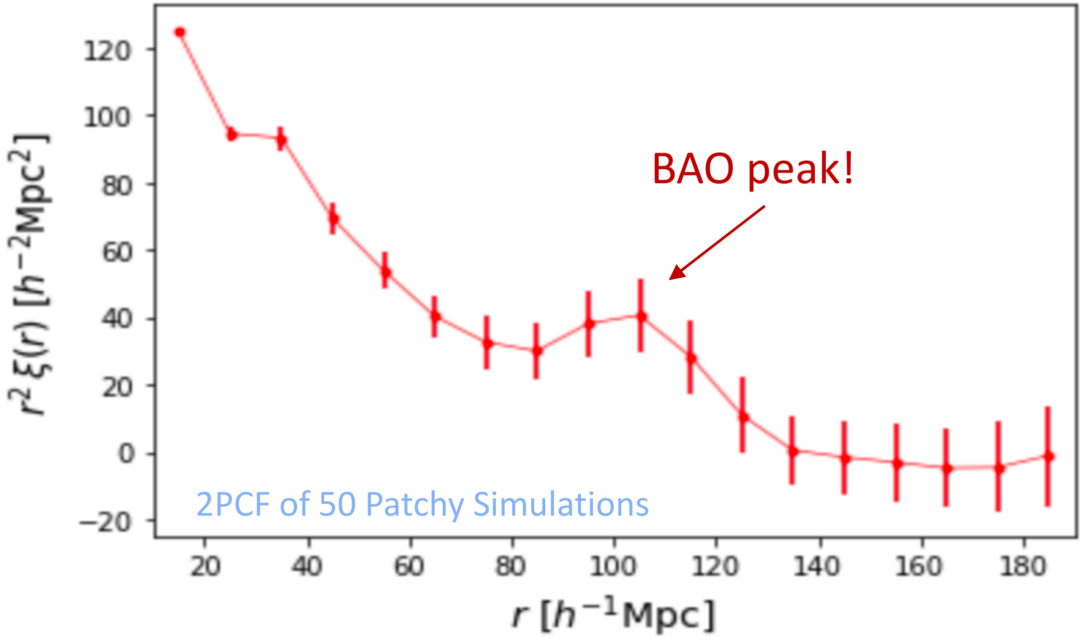


- Analyze with **summary statistics:**
  - Two-point correlation function (2PCF), $\xi(\mathbf{r})$

$$\xi(r) = \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle$$

  - Power spectrum, $P(\mathbf{k})$

$$(2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')P(k) = \langle \delta(\mathbf{k})\delta(\mathbf{k}')\rangle$$



Extracting Information from Galaxy Surveys

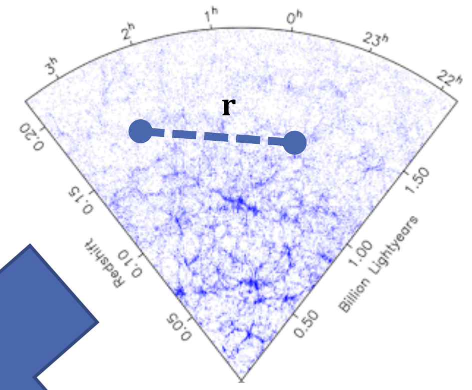
- Fundamental observable: the galaxy **overdensity** field

$$\delta(\mathbf{r}) = \frac{1}{\bar{n}}\left[n_g(\mathbf{r}) - n_r(\mathbf{r})\right]$$

Mean density

Galaxy positions

Random particle positions

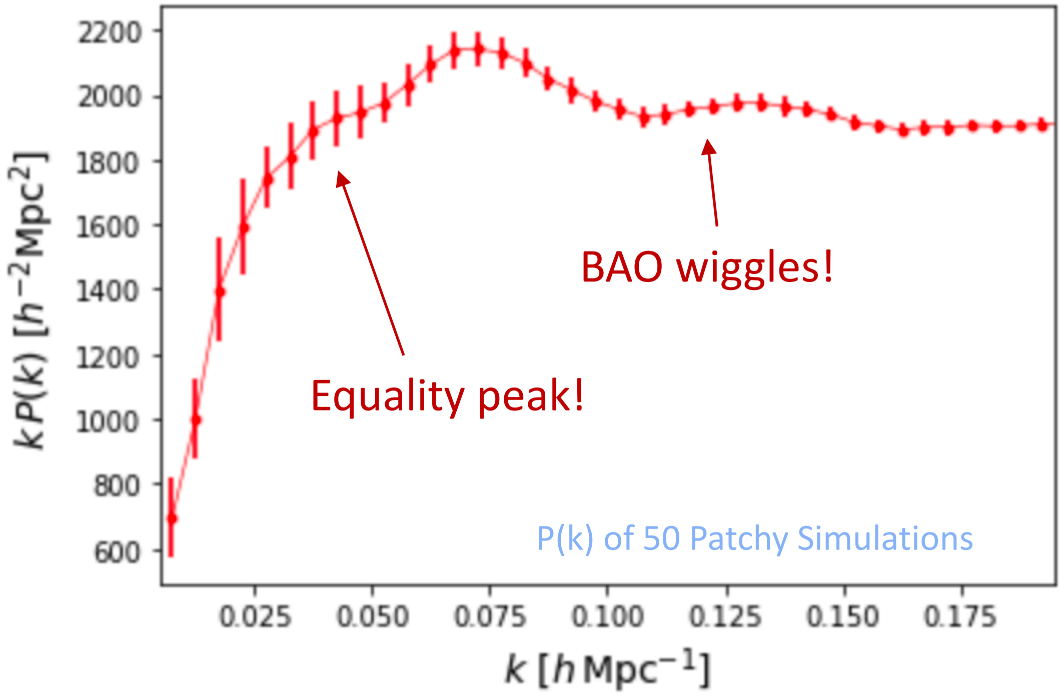


- Analyze with **summary statistics:**
  - Two-point correlation function (2PCF), $\xi(\mathbf{r})$

$$\xi(r) = \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle$$

  - Power spectrum, $P(\mathbf{k})$

$$(2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')P(k) = \langle \delta(\mathbf{k})\delta(\mathbf{k}')\rangle$$



Understanding Anisotropy

- **Redshift-space distortions** lead to **anisotropy**
  - Parametrize by **galaxy separation** and angle to **line-of-sight**, $\hat{\mathbf{n}}$

$$\xi(\mathbf{r}) = \sum_\ell \xi_\ell(r) L_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}}) \qquad P(\mathbf{k}) = \sum_\ell P_\ell(k) L_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})$$

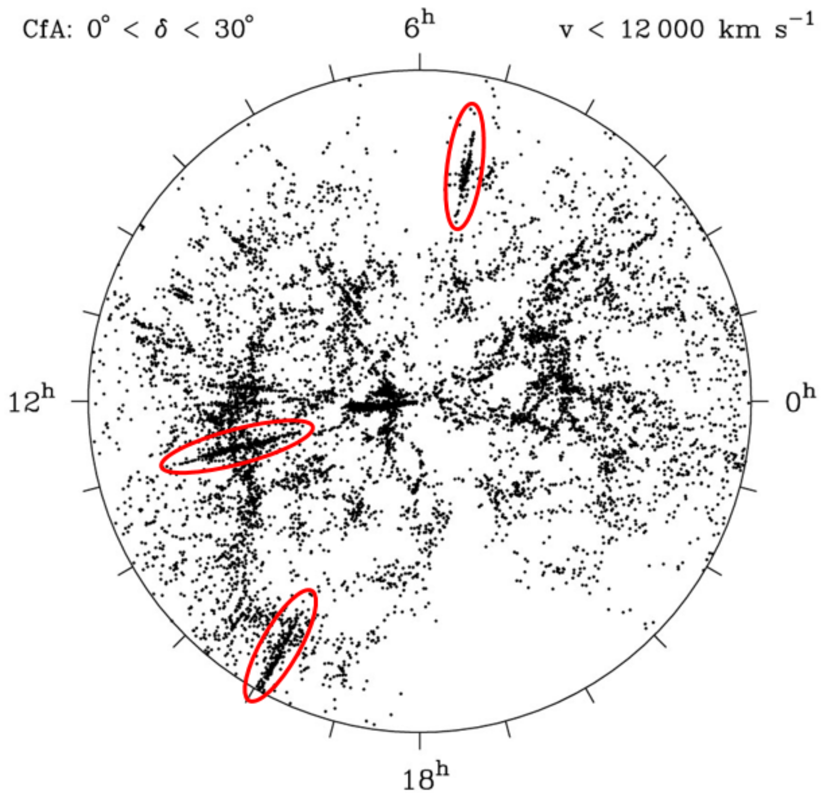
  - Define the **multipoles**:

$$\hat{\xi}_\ell(r) = (2\ell+1) \int \frac{d\Omega_r}{4\pi} \int d\mathbf{x}\, \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) L_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}})$$

$$\hat{P}_\ell(k) = \frac{(2\ell+1)}{V} \int \frac{d\Omega_k}{4\pi} \int d\mathbf{r}_1\, d\mathbf{r}_2\, e^{-i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)} \delta(\mathbf{r}_1)\delta(\mathbf{r}_2) L_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})$$

$= \delta(\mathbf{k})\delta^*(\mathbf{k})$

Legendre Polynomials



CfA Redshift Survey

Understanding Anisotropy

- **Redshift-space distortions** lead to **anisotropy**
  - Parametrize by **galaxy separation** and angle to **line-of-sight**, $\hat{\mathbf{n}}$

$$\xi(\mathbf{r}) = \sum_\ell \xi_\ell(r) L_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}}) \qquad P(\mathbf{k}) = \sum_\ell P_\ell(k) L_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})$$

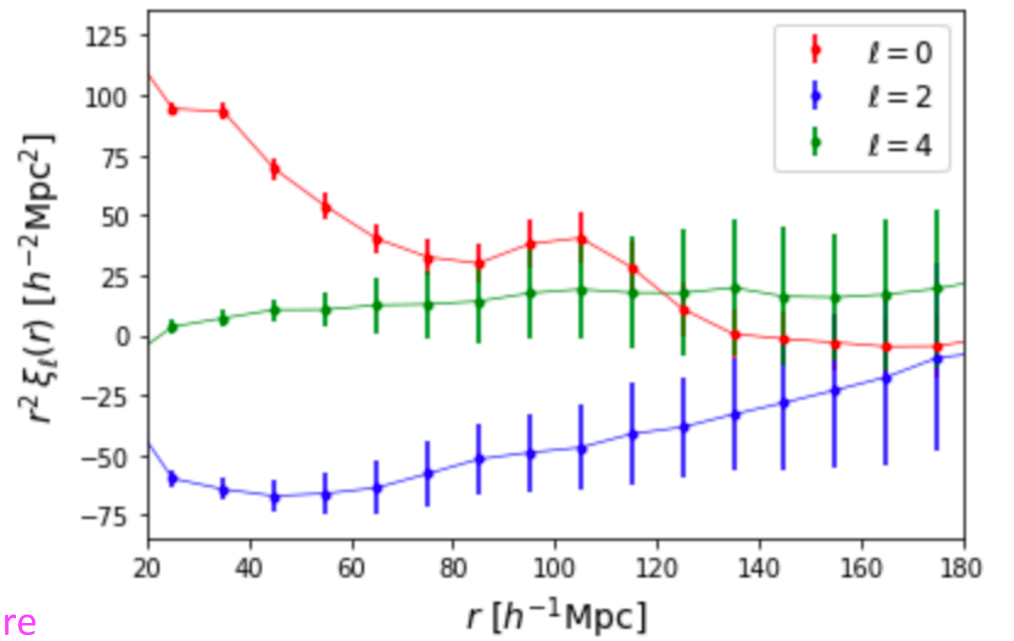
  - Define the **multipoles**:

$$\hat{\xi}_\ell(r) = (2\ell+1) \int \frac{d\Omega_r}{4\pi} \int d\mathbf{x}\, \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) L_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}})$$

$$\hat{P}_\ell(k) = \frac{(2\ell+1)}{V} \int \frac{d\Omega_k}{4\pi} \int d\mathbf{r}_1\, d\mathbf{r}_2\, e^{-i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)} \delta(\mathbf{r}_1)\delta(\mathbf{r}_2) L_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})$$

$= \delta(\mathbf{k})\delta^*(\mathbf{k})$

Legendre Polynomials



MultiDark-Patchy Mocks (Kitaura+16)

Understanding Anisotropy

- **Redshift-space distortions** lead to **anisotropy**
  - Parametrize by **galaxy separation** and angle to **line-of-sight**, $\hat{\mathbf{n}}$

$$\xi(\mathbf{r}) = \sum_\ell \xi_\ell(r) L_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}}) \qquad P(\mathbf{k}) = \sum_\ell P_\ell(k) L_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})$$

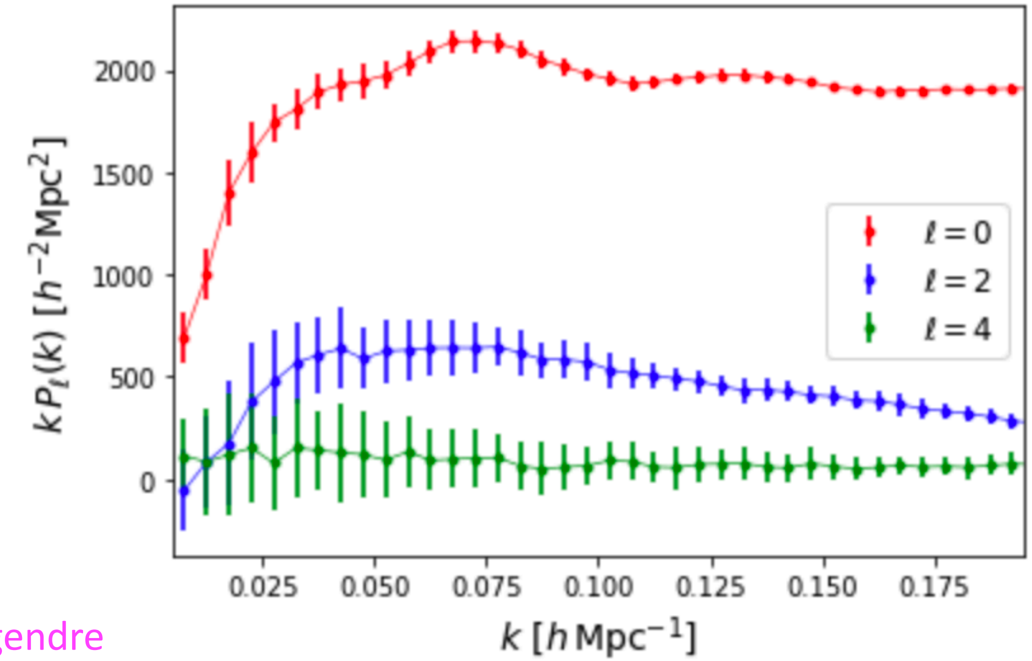
  - Define the **multipoles**:

$$\hat{\xi}_\ell(r) = (2\ell+1) \int \frac{d\Omega_r}{4\pi} \int d\mathbf{x}\, \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) L_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}})$$

$$\hat{P}_\ell(k) = \frac{(2\ell+1)}{V} \int \frac{d\Omega_k}{4\pi} \int d\mathbf{r}_1\, d\mathbf{r}_2\, e^{-i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)} \delta(\mathbf{r}_1)\delta(\mathbf{r}_2) L_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})$$

= $\delta(\mathbf{k})\delta^*(\mathbf{k})$

Legendre Polynomials

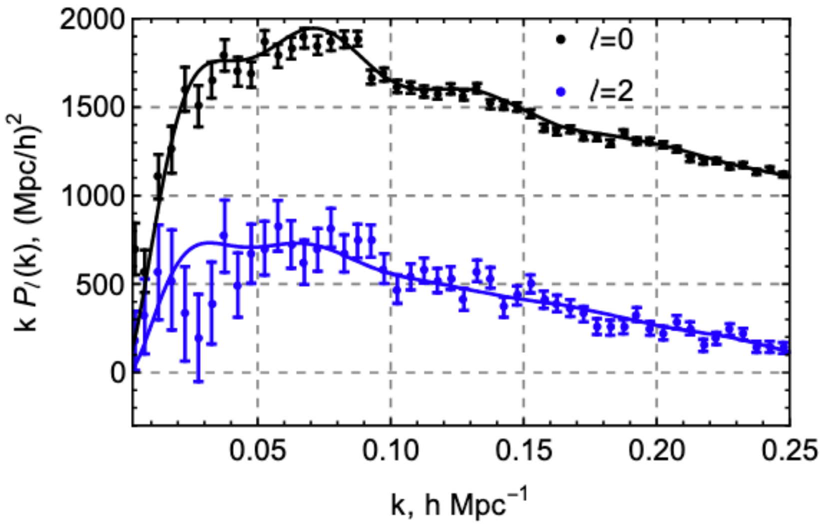


MultiDark-Patchy Mocks (Kitaura+16)

Parameter Inference

CMB-Strength Parameter Constraints, including 1.6% on H_0!

BOSS DR12 Spectra



$$P_{g,\ell}(k) = P^{\rm tree}_{g,\ell}(k) + P^{\rm 1-loop}_{g,\ell}(k) + P^{\rm noise}_{g,\ell}(k) + P^{\rm ctr}_{g,\ell}(k)$$

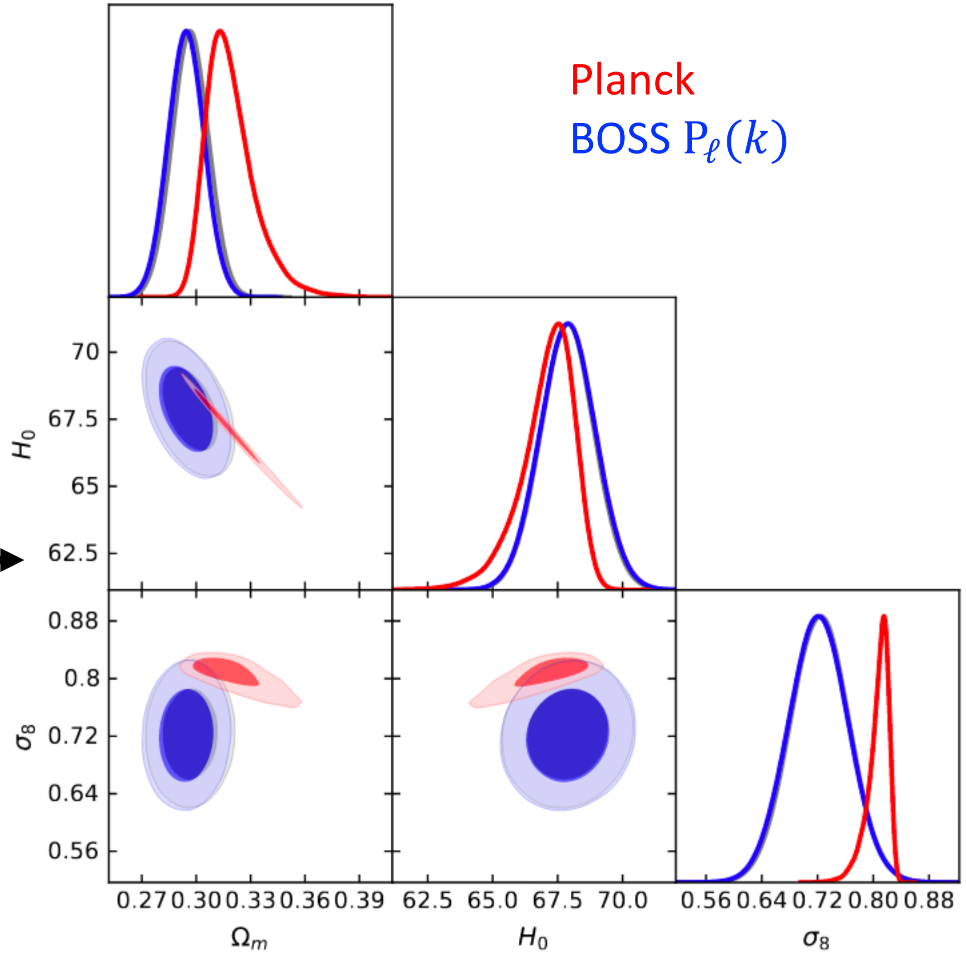
Linear Theory | 1-loop PT | Shot-noise | Counterterms

Theory Model

(Effective Field Theory)

MCMC

Mock Datasets



Planck

BOSS $P_\ell(k)$

Ivanov+19, Philcox+20a

Beyond 2-Point Statistics

The Universe is **non-Gaussian**

Information in **higher-point** functions, e.g.

- Bispectrum / 3PCF [Gil-Marín+16, Slepian+15, d'Amico+19]

$$(2\pi)^3\delta_D\left(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3\right)B(\mathbf{k}_1,\mathbf{k}_2) = \langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle$$

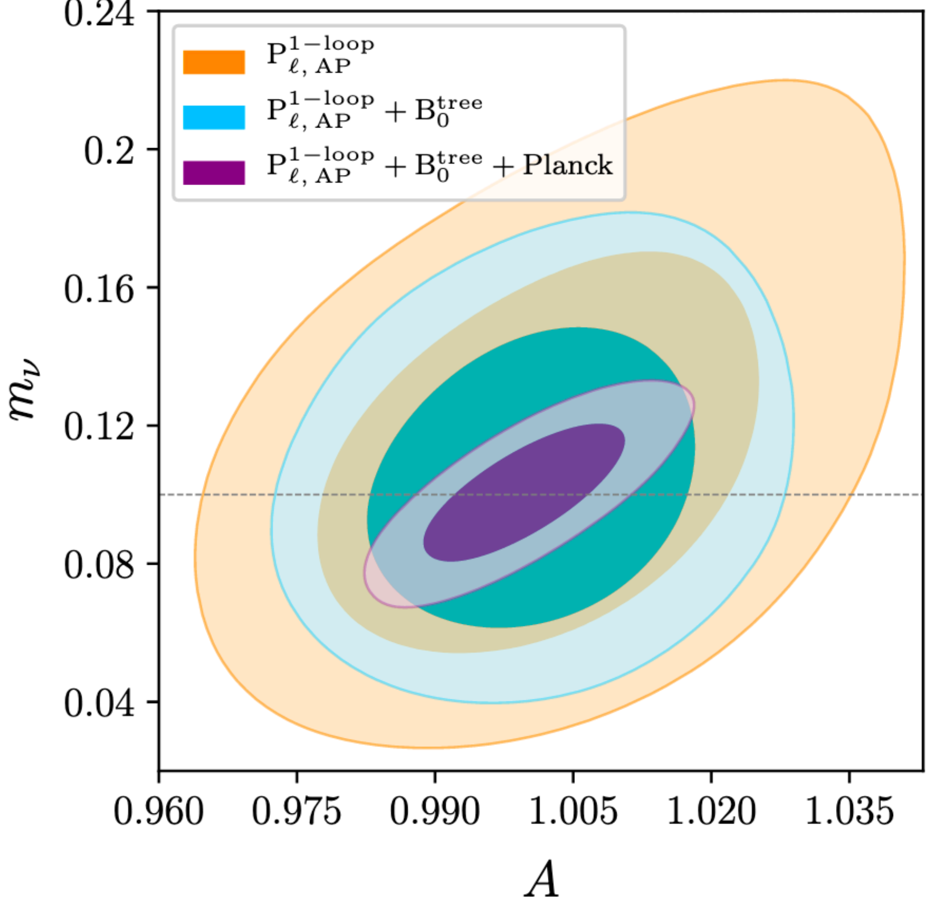
- Trispectrum / 4PCF [Gualdi+20]

$$(2\pi)^3\delta_D\left(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3+\mathbf{k}_4\right)T(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = \langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\delta(\mathbf{k}_4)\rangle$$

These get steadily larger and harder to measure.

- Not used in many cosmological analyses yet!

3-point Forecasts for Euclid



Chudaykin & Ivanov 19

Cosmology from $P_\ell(k)$: A Summary

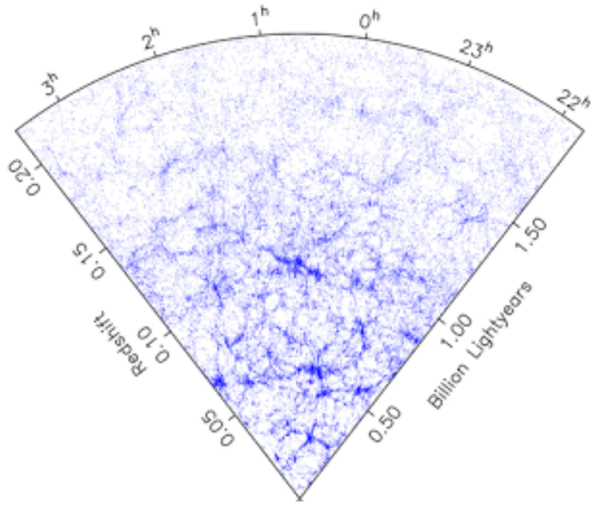
- Fundamental observable: the galaxy **overdensity** field

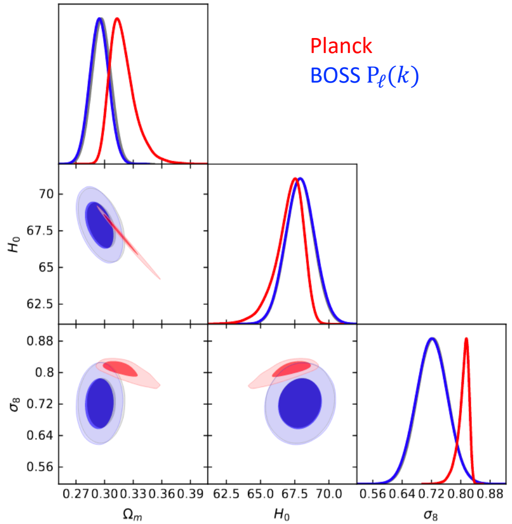
- $P_\ell(k)$ parametrized by pair separation and **line-of-sight angle**

- Power spectrum estimators measure $|\delta(\mathbf{k})|^2 L_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})$

- Computed using **Fast Fourier Transforms** (FFTs)

- Compare data and theory with MCMC





Cosmology from $P_\ell(k)$: A Summary

- Fundamental observable: the galaxy **overdensity** field — *Is this the best field to use?*
- $P_\ell(k)$ parametrized by pair separation and **line-of-sight angle** — *How do we define this angle?*
- Power spectrum estimators measure $|\delta(\mathbf{k})|^2 L_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{n}})$ — *Can we estimate it more optimally?*
- Computed using **Fast Fourier Transforms** (FFTs) — *Are FFTs always the most efficient?*
- Compare data and theory with MCMC — *Can data-compression help?*

1. Parametrizing Anisotropy

Philcox & Slepian 21

Choosing the Line-of-Sight

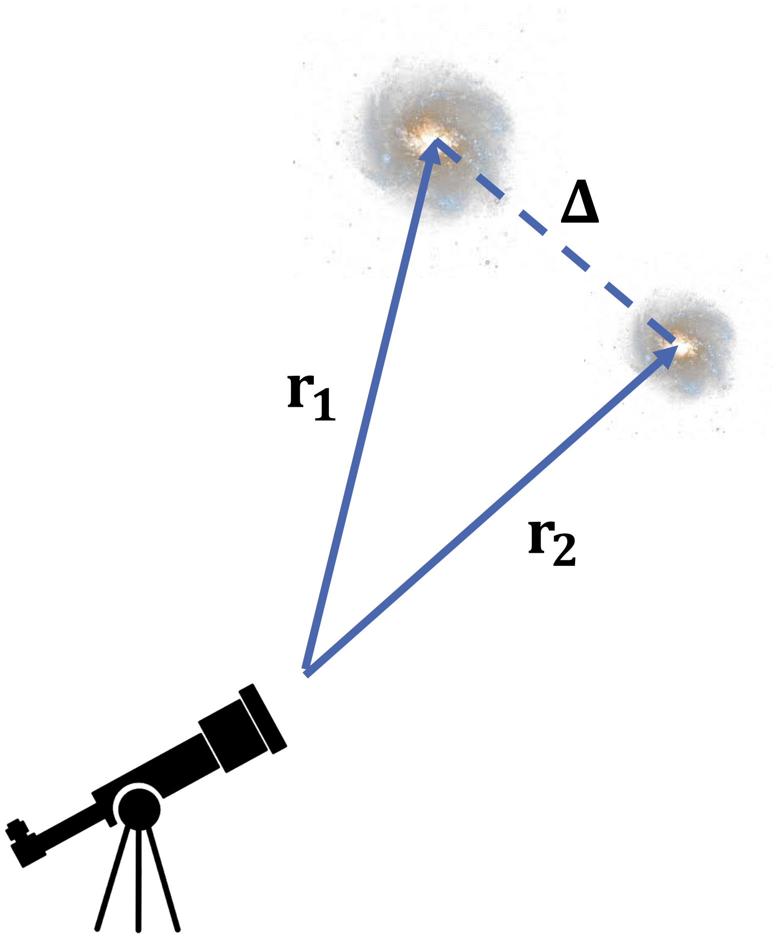
- Galaxy correlation function depends on the angle between the separation vector $\mathbf{\Delta}$ and the line-of-sight $\hat{\mathbf{n}}$:

Angular Dependence

$$\hat{\xi}_\ell(r) = \frac{2\ell+1}{V}\int d\mathbf{r}_1\, d\mathbf{r}_2\, \delta(\mathbf{r}_1)\delta(\mathbf{r}_2) L_\ell(\hat{\mathbf{\Delta}}\cdot\hat{\mathbf{n}})\left[\frac{\delta_D(r-\Delta)}{4\pi r^2}\right]$$

Density Fields

Binning



Slepian & Eisenstein 15,
Philcox & Slepian 21

Choosing the Line-of-Sight

- Galaxy correlation function depends on the angle between the separation vector $\mathbf{\Delta}$ and the line-of-sight $\hat{\mathbf{n}}$:

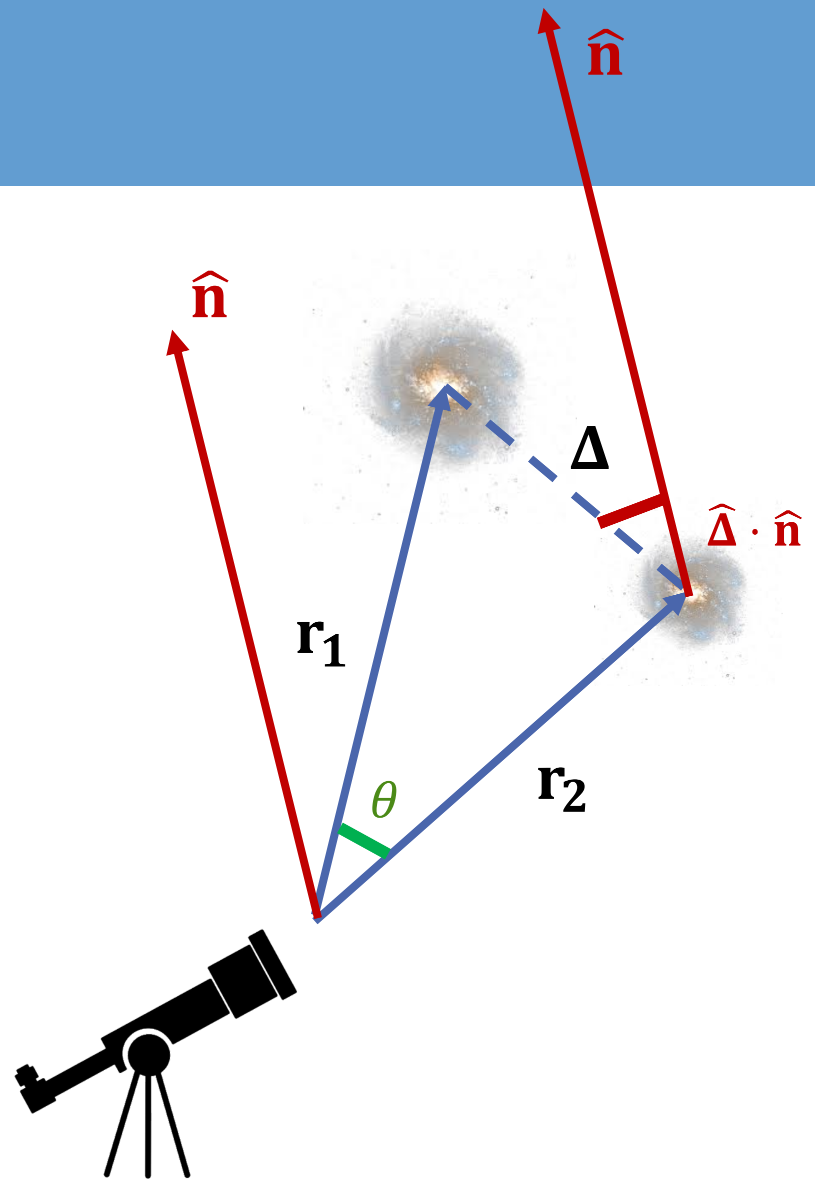
Angular Dependence

$$\hat{\xi}_\ell(r) = \frac{2\ell+1}{V} \int d\mathbf{r}_1\, d\mathbf{r}_2\, \delta(\mathbf{r}_1)\delta(\mathbf{r}_2) L_\ell(\hat{\mathbf{\Delta}} \cdot \hat{\mathbf{n}}) \left[\frac{\delta_D(r-\Delta)}{4\pi r^2}\right]$$

Density Fields

Binning

- Options:
  - **Fixed $\hat{\mathbf{n}}$** : $\mathcal{O}(\theta^0)$ error, for survey size θ



Slepian & Eisenstein 15,
Philcox & Slepian 21

Choosing the Line-of-Sight

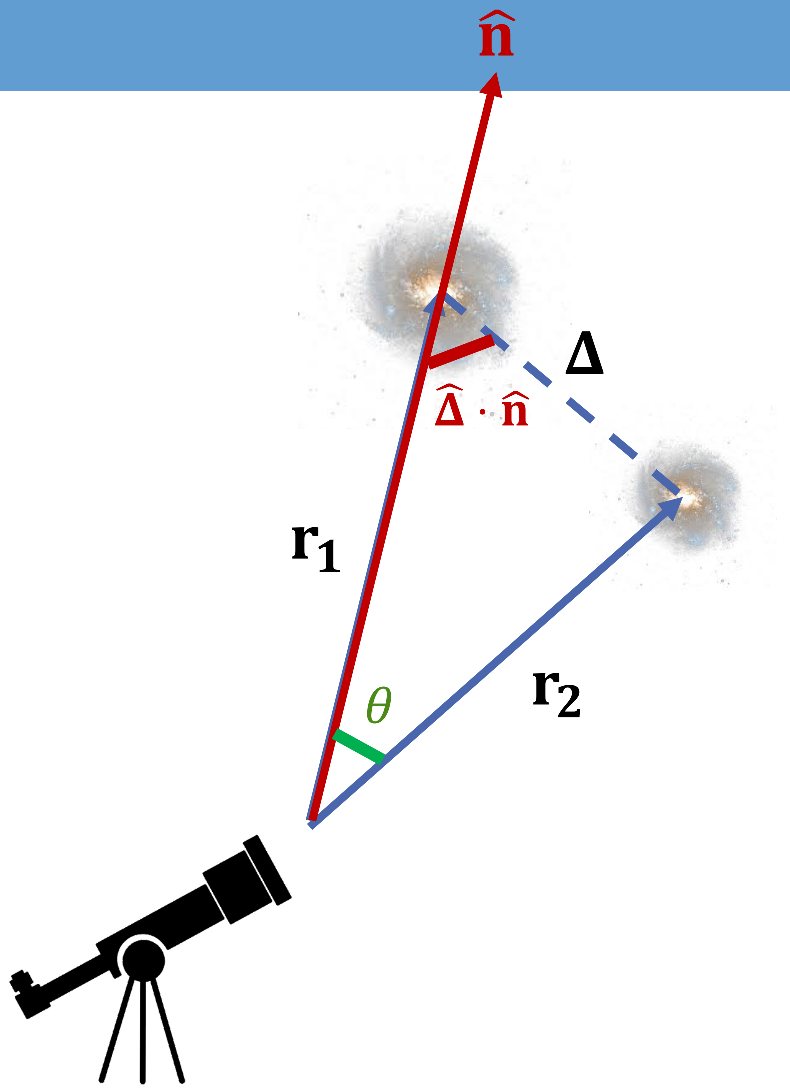
- Galaxy correlation function depends on the angle between the separation vector $\boldsymbol{\Delta}$ and the line-of-sight $\hat{\mathbf{n}}$:

Angular Dependence

$$\hat{\xi}_\ell(r) = \frac{2\ell+1}{V}\int d\mathbf{r}_1\, d\mathbf{r}_2\, \delta(\mathbf{r}_1)\delta(\mathbf{r}_2) L_\ell(\hat{\boldsymbol{\Delta}}\cdot\hat{\mathbf{n}}) \left[\frac{\delta_D(r-\Delta)}{4\pi r^2}\right]$$

Density Fields | Binning

- Options:
  - **Fixed $\hat{\mathbf{n}}$** : $\mathcal{O}(\theta^0)$ error, for survey size θ
  - **Yamamoto** approximation: $\hat{\mathbf{n}} = \hat{\mathbf{r}}_1$, $\mathcal{O}(\theta^2)$ error



Slepian & Eisenstein 15, Philcox & Slepian 21

Choosing the Line-of-Sight

- Galaxy correlation function depends on the angle between the separation vector $\mathbf{\Delta}$ and the line-of-sight $\hat{\mathbf{n}}$:

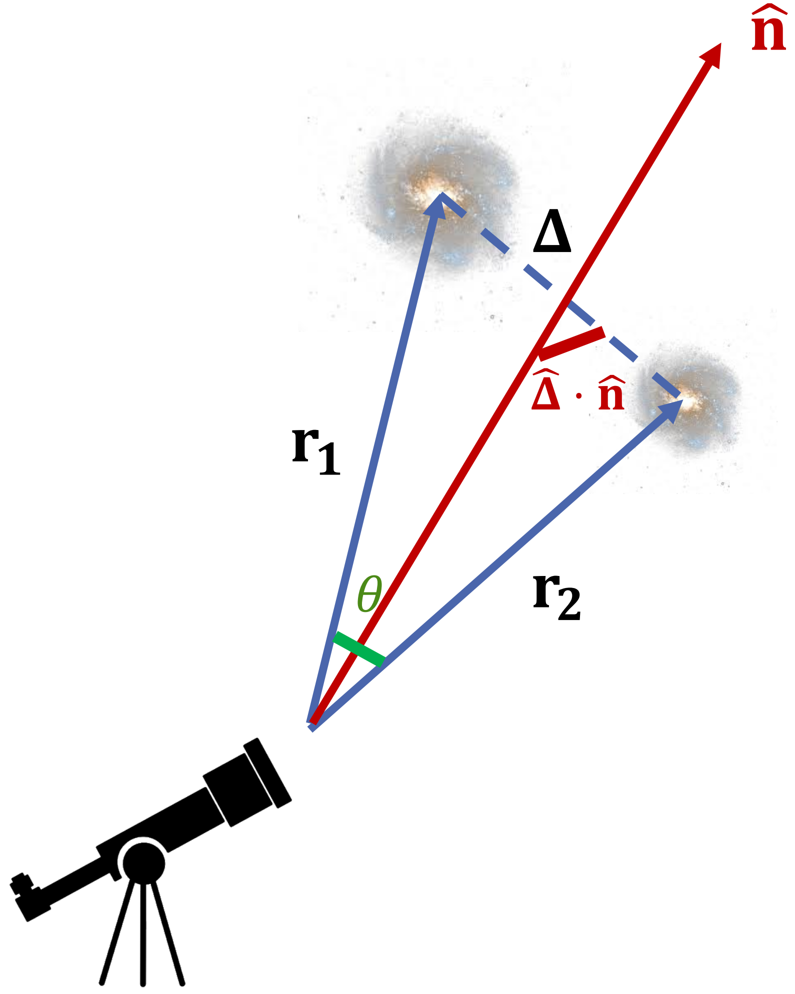
Angular Dependence

$$\hat{\xi}_\ell(r) = \frac{2\ell+1}{V}\int d\mathbf{r}_1\, d\mathbf{r}_2\, \underbrace{\delta(\mathbf{r}_1)\delta(\mathbf{r}_2)}_{\text{Density Fields}}\overbrace{L_\ell(\hat{\mathbf{\Delta}}\cdot\hat{\mathbf{n}})}^{\text{Angular Dependence}}\underbrace{\left[\frac{\delta_D(r-\Delta)}{4\pi r^2}\right]}_{\text{Binning}}$$

Density Fields

Binning

- Options:
  - **Fixed** $\hat{\mathbf{n}}$: $\mathcal{O}(\theta^0)$ error, for survey size θ
  - **Yamamoto** approximation: $\hat{\mathbf{n}} = \hat{\mathbf{r}}_1$, $\mathcal{O}(\theta^2)$ error
  - **Midpoint** method: $\hat{\mathbf{n}} = \widehat{\mathbf{r}_1 + \mathbf{r}_2}$, $\mathcal{O}(\theta^{4+})$ error



Slepian & Eisenstein 15,
Philcox & Slepian 21

Lines-of-Sight in the Power Spectrum

- Same for the **power spectrum**:

Fourier Transform

Angular Dependence

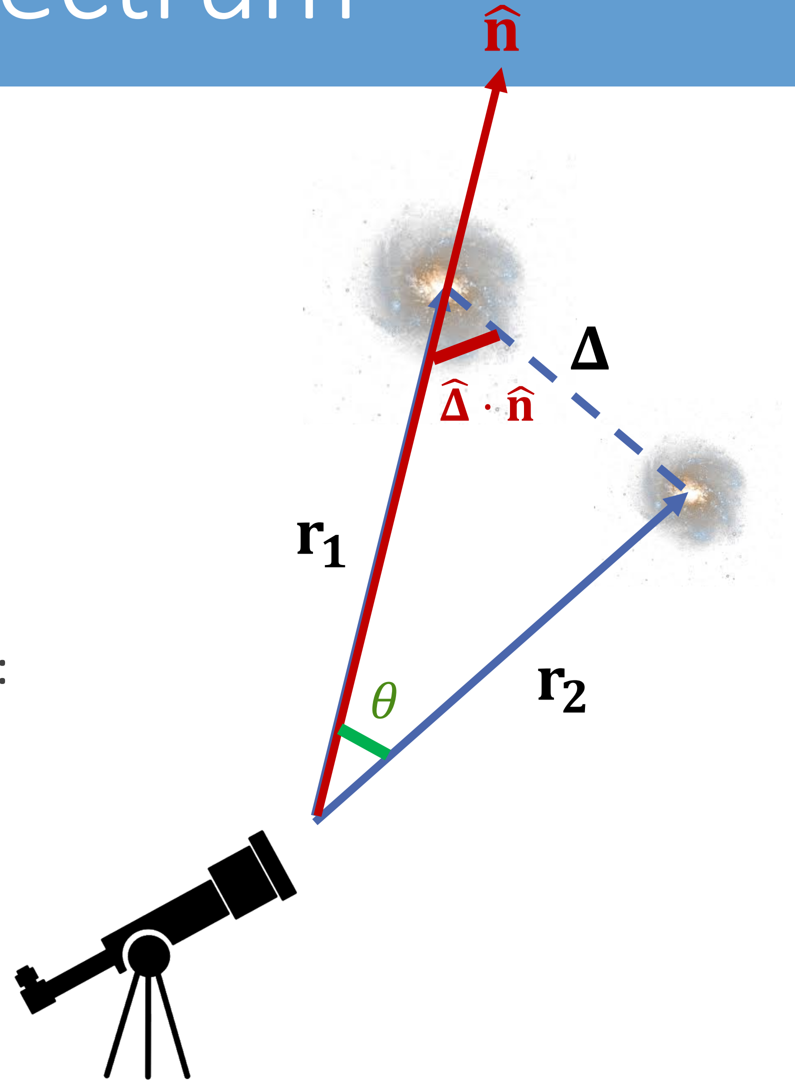
$$\hat{P}_\ell(k) = \frac{2\ell+1}{V}\int_{\Omega_k}\int d\mathbf{r}_1\, d\mathbf{r}_2\, e^{-i\mathbf{k}\cdot(\mathbf{r}_2-\mathbf{r}_1)}\delta(\mathbf{r}_1)\delta(\mathbf{r}_2)L_\ell(\hat{\mathbf{k}}\cdot\hat{\mathbf{n}})$$

Density Fields

- This is **easy to implement** for the Yamamoto approximation, $\hat{\mathbf{n}} = \hat{\mathbf{r}}_1$:

$$\hat{P}_\ell^{\rm Yama}(k) = \frac{4\pi}{V}\int_{\Omega_k}\left[\sum_{m=-\ell}^{\ell} Y_\ell^{m*}(\hat{\mathbf{k}})\mathcal{F}\left[Y_\ell^m\delta\right](\mathbf{k})\right]\delta^*(\mathbf{k})$$

Spherical Harmonics



Hand+17, Philcox & Slepian 21

Lines-of-Sight in the Power Spectrum

- Same for the **power spectrum**:

Fourier Transform — Angular Dependence

$$\hat{P}_\ell(k) = \frac{2\ell+1}{V}\int_{\Omega_k}\int d\mathbf{r}_1\, d\mathbf{r}_2\, e^{-i\mathbf{k}\cdot(\mathbf{r}_2-\mathbf{r}_1)}\delta(\mathbf{r}_1)\delta(\mathbf{r}_2)L_\ell(\hat{\mathbf{k}}\cdot\hat{\mathbf{n}})$$

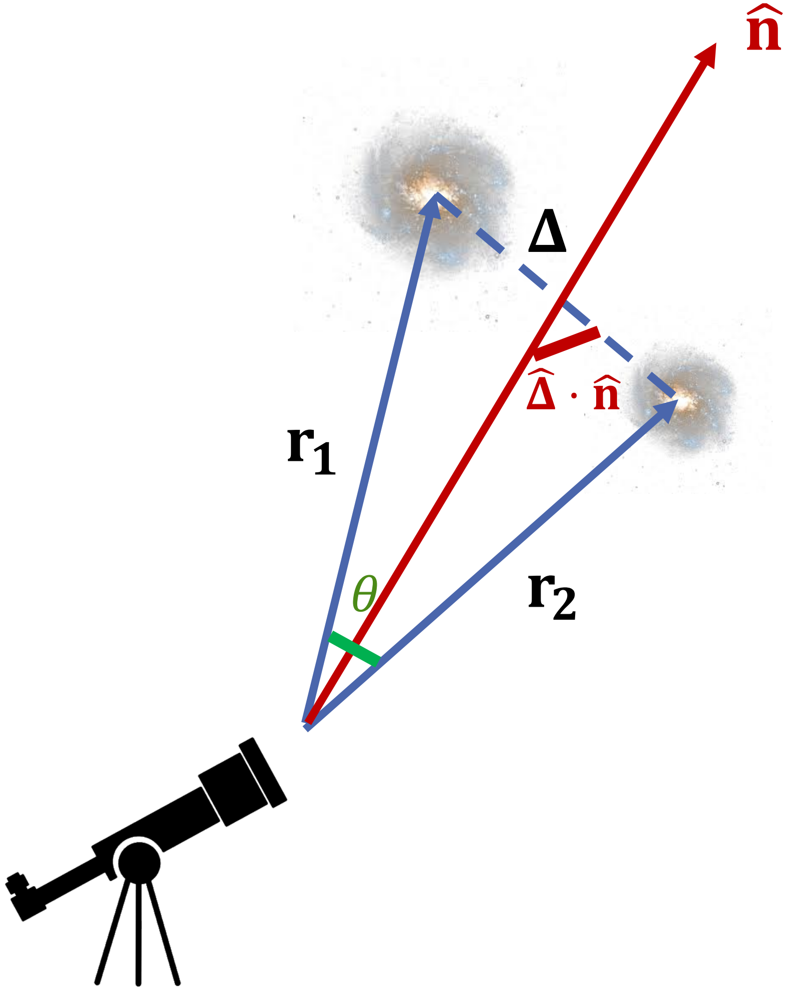
Density Fields

- This is **easy to implement** for the Yamamoto approximation, $\hat{\mathbf{n}} = \hat{\mathbf{r}}_1$:

$$\hat{P}_\ell^{\rm Yama}(k) = \frac{4\pi}{V}\int_{\Omega_k}\left[\sum_{m=-\ell}^{\ell} Y_\ell^{m*}(\hat{\mathbf{k}})\mathcal{F}\left[Y_\ell^m\delta\right](\mathbf{k})\right]\delta^*(\mathbf{k})$$

Spherical Harmonics

- But **not separable** for the midpoint method!



Hand+17,
Philcox & Slepian 21

Implementing the Midpoint Method

- Use a trick to make the integrals separable:

  - Expand in powers of $\theta \sim \Delta/r_1$:

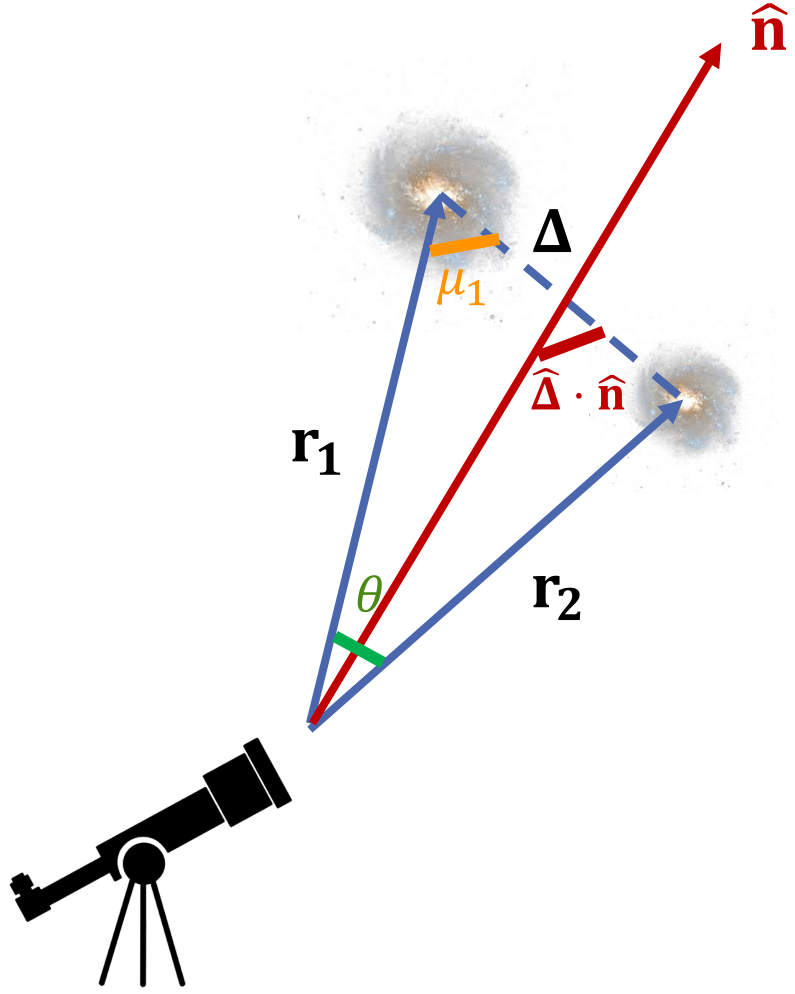
Survey Angle, $\ll 1$

$$L_\ell(\hat{\mathbf{\Delta}}\cdot\widehat{\mathbf{r}_1+\mathbf{r}_2}) = \sum_{\alpha=0}^{\infty}\sum_{J=0}^{\ell+\alpha} f_J^{\alpha,\ell}\left(\frac{\Delta}{2r_1}\right)^{\alpha} L_J(\hat{\mathbf{\Delta}}\cdot\mathbf{r}_1)$$

Coefficients

Yamamoto Piece

$$\begin{aligned}L_2(\hat{\mathbf{\Delta}}\cdot\widehat{\mathbf{r}_1+\mathbf{r}_2}) \;=\;& L_2(\mu_1)+\frac{6}{5}\left(\frac{\Delta}{2r_1}\right)[L_1(\mu_1)-L_3(\mu_1)]\\ &+\frac{1}{35}\left(\frac{\Delta}{2r_1}\right)^2[7L_0(\mu_1)-55L_2(\mu_1)+48L_4(\mu_1)]\\ &-\frac{4}{105}\left(\frac{\Delta}{2r_1}\right)^3[9L_1(\mu_1)-49L_3(\mu_1)+40L_5(\mu_1)]\\ &+\frac{1}{385}\left(\frac{\Delta}{2r_1}\right)^4[11L_0(\mu_1)+165L_2(\mu_1)-816L_4(\mu_1)+640L_6(\mu_1)]\\ &+\ldots\end{aligned}$$



Slepian & Eisenstein 15
Philcox & Slepian 21

Implementing the Midpoint Method

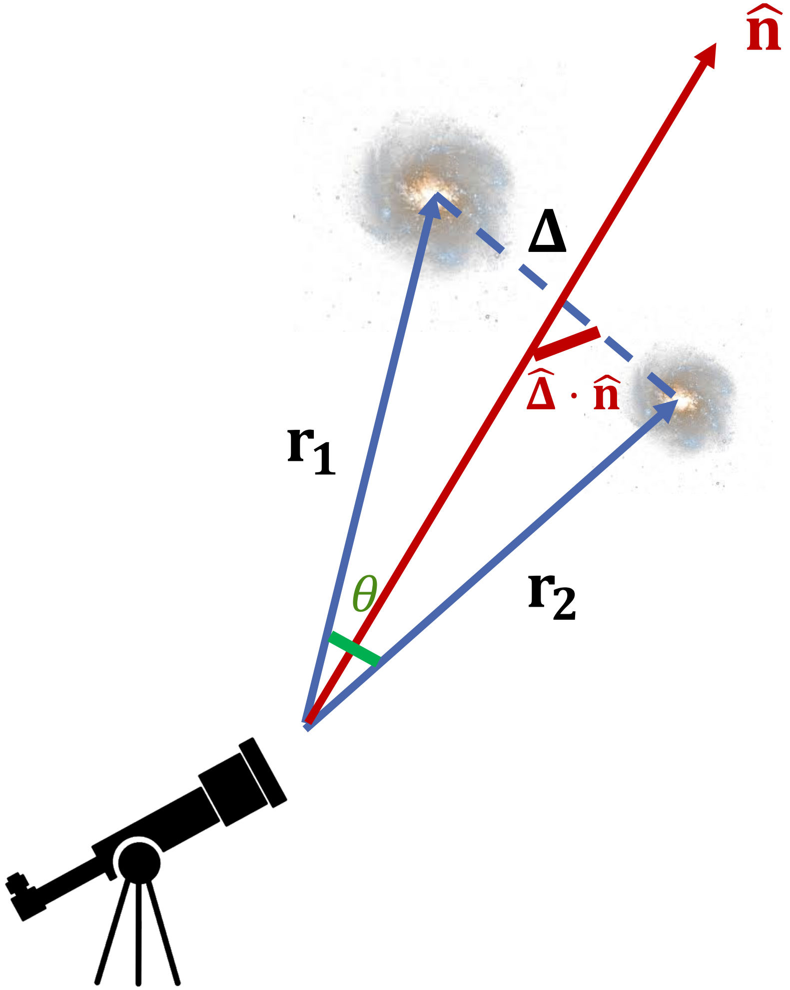
- Use a trick to make the integrals separable:
  - Expand in powers of $\theta \sim \Delta/r_1$:

Survey Angle, ≪ 1

$$L_\ell(\hat{\mathbf{\Delta}}\cdot\widehat{\mathbf{r}_1+\mathbf{r}_2}) = \sum_{\alpha=0}^{\infty}\sum_{J=0}^{\ell+\alpha} f_J^{\alpha,\ell}\left(\frac{\Delta}{2r_1}\right)^{\alpha} L_J(\hat{\mathbf{\Delta}}\cdot\mathbf{r}_1)$$

Coefficients — Yamamoto Piece

  - Can now compute the 2PCF using Fourier transforms!
  - Also applies to the **power spectrum**
  - Same computational complexity as Yamamoto approximation

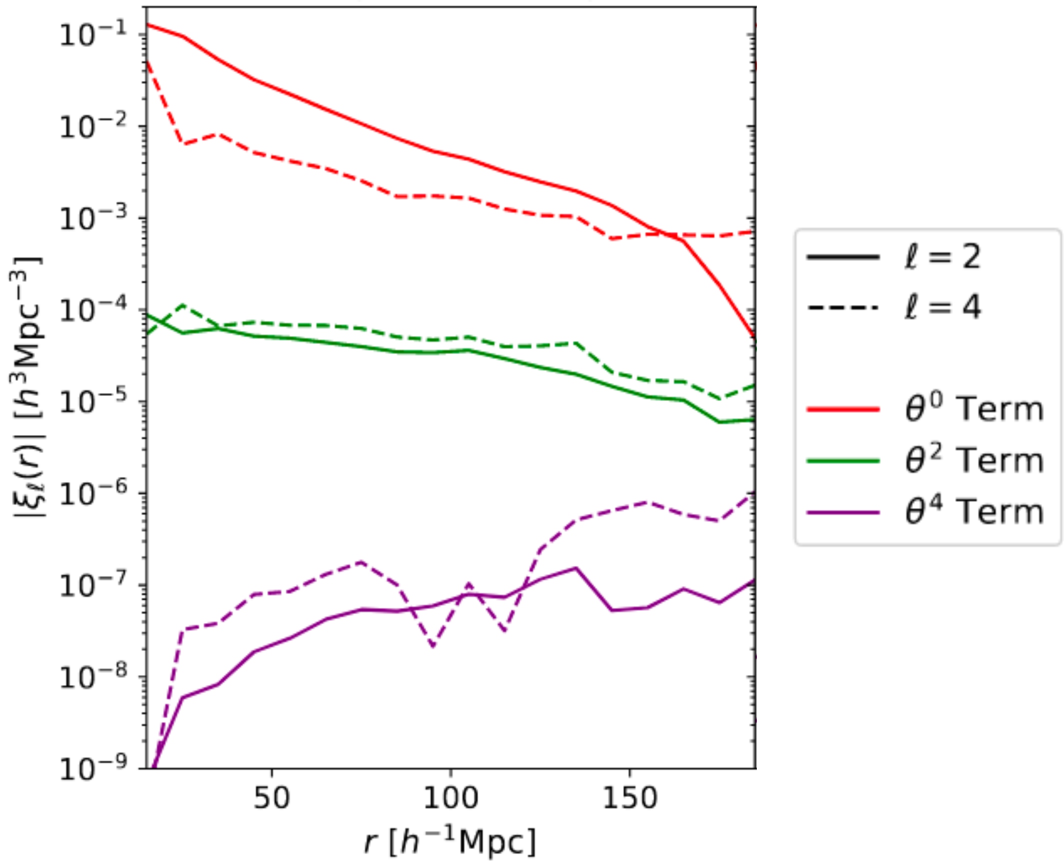


Castorina & White 18
Philcox & Slepian 21

The Midpoint Method in Practice

- BOSS correlation function:
  - $\theta \sim 0.1 - 0.2$ at the BAO scale
  - Larger $r \Rightarrow$ Larger corrections
  - Still $\ll 1\sigma$ for BOSS

2PCF Wide-Angle Corrections



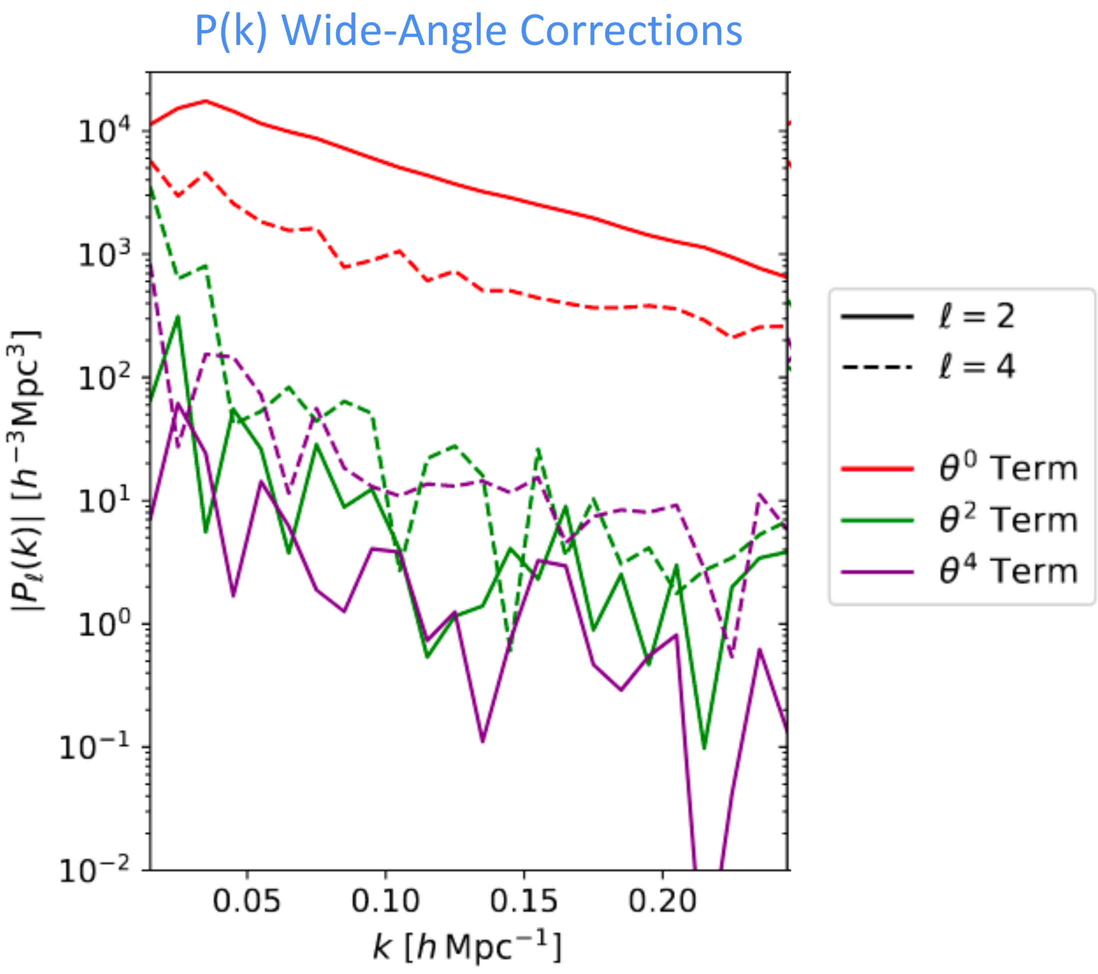
Philcox & Slepian 21

The Midpoint Method in Practice

- BOSS correlation function:
  - $\theta \sim 0.1 - 0.2$ at the BAO scale
  - Larger *r* ⇒ Larger corrections
  - Still $\ll 1\sigma$ for BOSS

- BOSS $P(k)$:
  - Spectrum is an integral over **all** *r* in survey
  - $\theta \sim 1$ for the largest-modes
  - Corrections **are** marginally important at **all** *k*

- Most important for **wide surveys** at **low redshifts**



Philcox & Slepian 21

2. Optimal Power Spectrum Estimation

“Throwing the window out the window…” – Z. Slepian

Philcox 20b

The FKP Estimator

Power spectrum **isn't** just $|\delta(\mathbf{k})|^2$.

- Neglects **inhomogeneous noise** and **survey window functions**

1. Define $\delta(\mathbf{r})$ as the difference between **galaxy** and **random** densities
2. Add an **FKP weight** to incorporate **Poisson noise** densities (and systematics)

This is the **optimal solution** on small-scales with Poisson noise

But:

- **Not** optimal on large scales
- Measures the **window-convolved** power spectrum

$$\hat{P}(k) = \int \frac{d\Omega_k}{4\pi} \int d\mathbf{r}_1\, d\mathbf{r}_2\, e^{-i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)}\delta(\mathbf{r}_1)\delta(\mathbf{r}_2)$$

Galaxies *Randoms*

$$\delta(\mathbf{r}) \to \frac{w(\mathbf{r})[n_g(\mathbf{r}) - \alpha_r n_r(\mathbf{r})]}{I^{1/2}}, \quad I \equiv \int d\mathbf{r}\, w^2(\mathbf{r})\bar{n}^2(\mathbf{r})$$

Systematics

$$w(\mathbf{r}) = \frac{w_{\rm sys}(\mathbf{r})}{1 + P_{\rm FKP}\, n(\mathbf{r})}$$

Poisson Noise Correction, $P_{\rm FKP} \sim 10^4$

Feldman+94
Tegmark+98

Optimal Estimators

Maximize the **likelihood** for data, **d**, with band-powers **p** and pixel covariance C(**p**)

$$-2\log L[\mathbf{d}](\mathbf{p}) = \mathbf{d}^T \mathsf{C}^{-1}(\mathbf{p})\mathbf{d} + \mathrm{Tr}\log \mathsf{C}(\mathbf{p}) + \text{const.}$$

← *Gaussian likelihood*

Gives a **maximum-likelihood** estimator for the **unwindowed** power spectrum**:**

Fisher matrix (from sims)

$$\hat{p}_\alpha^{\mathrm{ML}} = p_\alpha^{\mathrm{fid}} + \sum_\beta F_{\alpha\beta}^{-1}\left(\hat{q}_\beta - \bar{q}_\beta\right)$$

Known Fiducial Model

Quadratic estimator (from data)

Bias term (from sims)

Estimator is a **quadratic** function of the data, $\hat{q}_\beta$

Tegmark 97
Philcox 20b

Implementing the ML Estimator

$$\hat{p}_\alpha^{\mathrm{ML}} = p_\alpha^{\mathrm{fid}} + \sum_\beta F_{\alpha\beta}^{-1}\left(\hat{q}_\beta - \bar{q}_\beta\right)$$

- Need the **quadratic estimator** $\hat{q}_\beta$:

$$\hat{q}(k) = \int \frac{d\Omega_k}{4\pi} \int d\mathbf{r}\, d\mathbf{r}'\, e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \left[\mathsf{C}^{-1}\mathbf{d}\right](\mathbf{r}) \left[\mathsf{C}^{-1}\mathbf{d}\right](\mathbf{r}')$$

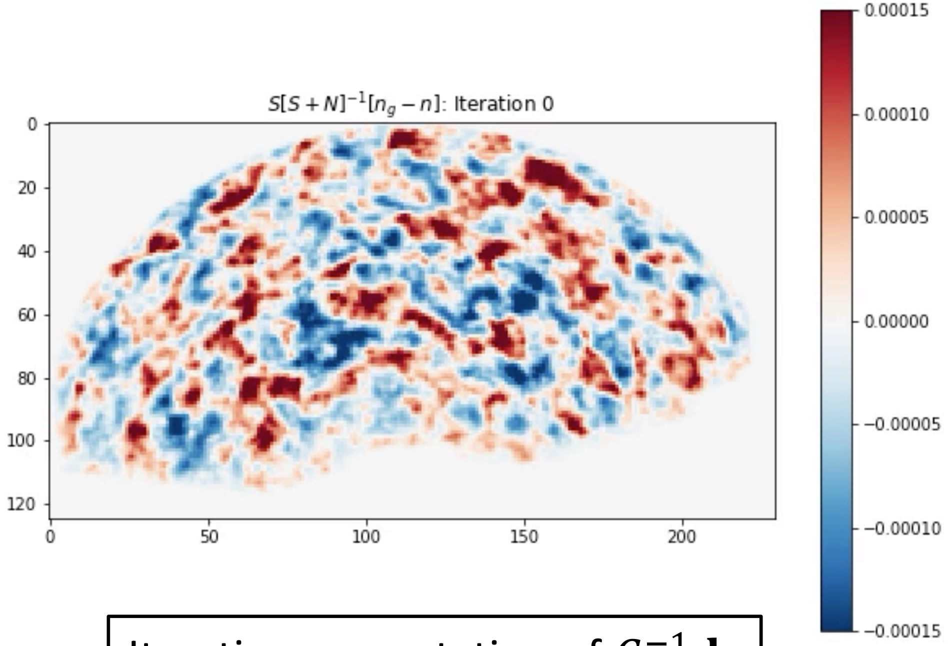
  - Just a power spectrum of the **inverse-covariance** weighted data

- Need the the covariance for each pair of **pixels**:

Poisson Noise

$$\mathsf{C}(\mathbf{r},\mathbf{r}') = n(\mathbf{r})n(\mathbf{r}') \int_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \sum_\ell P_\ell(k) L_\ell(\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}') + (1+\alpha)n(\mathbf{r})\delta_D(\mathbf{r}-\mathbf{r}')$$

Signal



Iterative computation of $C^{-1}\mathbf{d}$ from an initial guess

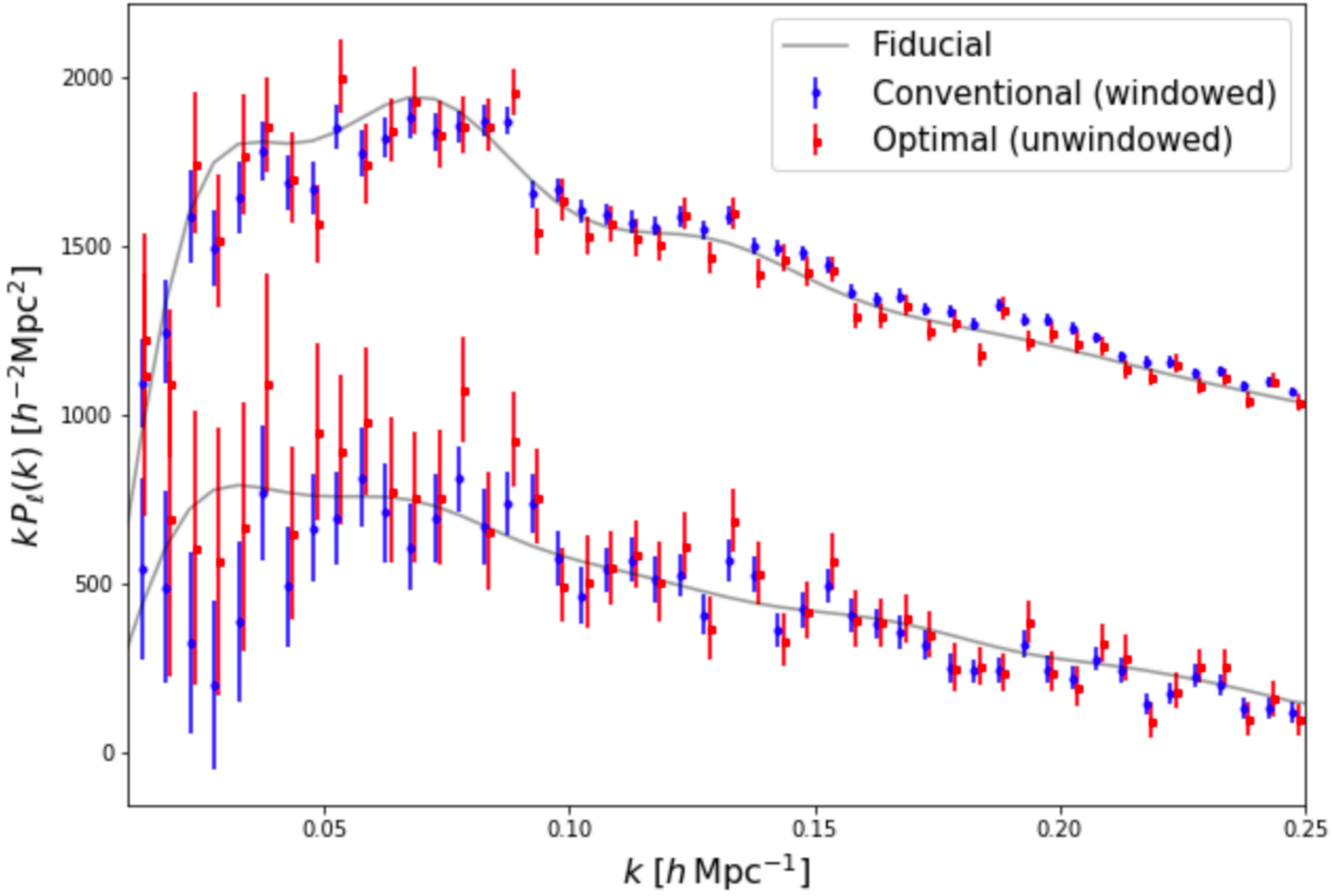
- This covariance is **gigantic** ($N_{\mathrm{pix}} \times N_{\mathrm{pix}}$)
  - Never store directly
  - Invert using **conjugate gradient descent** methods

Implementing the ML Estimator

$$\hat{p}^{\mathrm{ML}}_\alpha = p^{\mathrm{fid}}_\alpha + \sum_\beta F^{-1}_{\alpha\beta}\left(\hat{q}_\beta - \bar{q}_\beta\right)$$

Pipeline:

1. Choose a **fiducial cosmology**
2. Compute the **quadratic estimator** on the data, $\hat{q}_\beta$
3. Repeat on **simulations** to get bias, $\bar{q}_\beta$ and Fisher matrix, $F_{\alpha\beta}$
4. Combine to get the **power spectrum**
5. ***Optional:*** Repeat with new cosmology



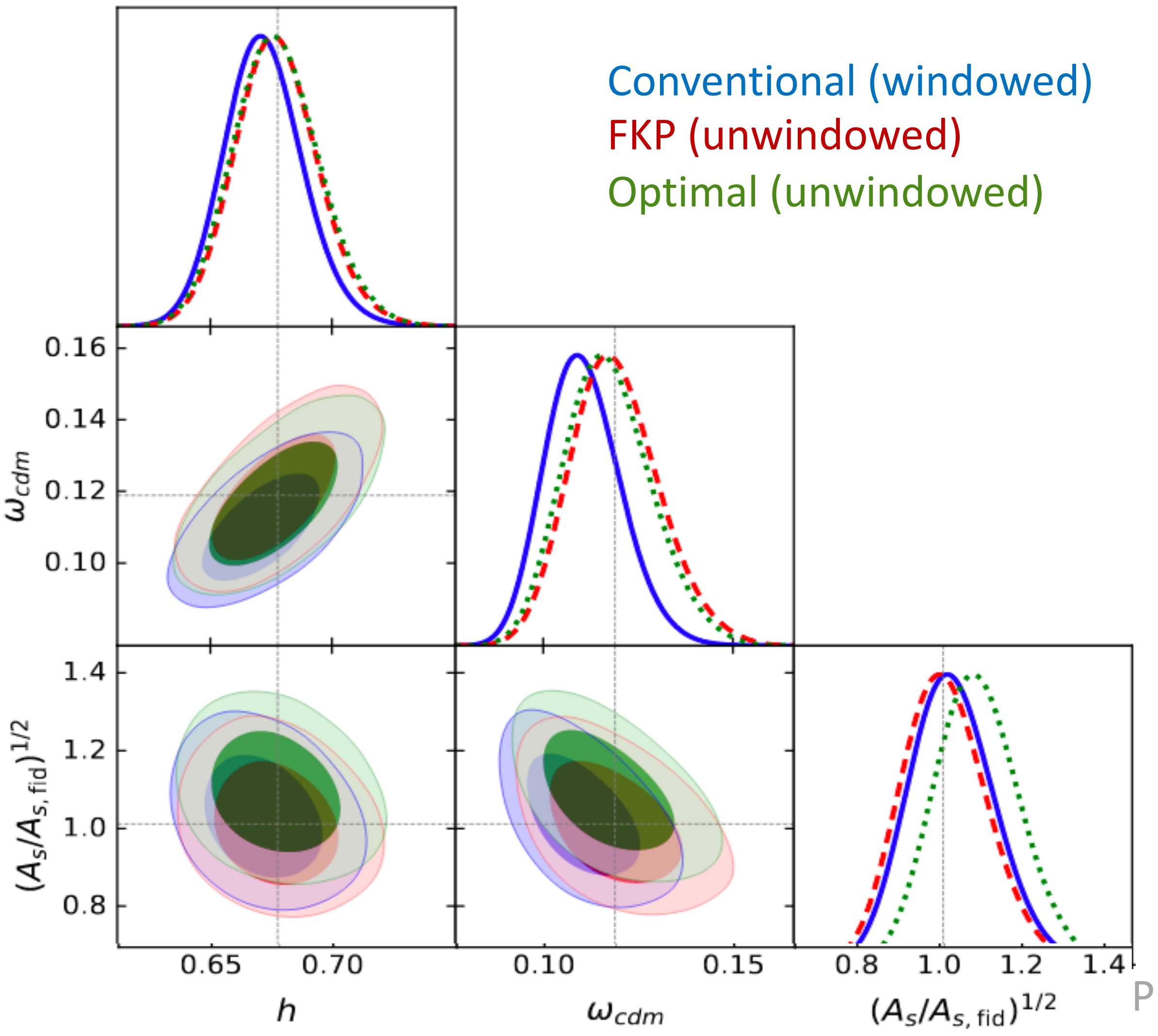
Philcox 20b

Implementing the ML Estimator

$$\hat{p}_\alpha^{\rm ML} = p_\alpha^{\rm fid} + \sum_\beta F_{\alpha\beta}^{-1}\left(\hat{q}_\beta - \bar{q}_\beta\right)$$

Pipeline:

1. Choose a **fiducial cosmology**
2. Compute the **quadratic estimator** on the data, $\hat{q}_\beta$
3. Repeat on **simulations** to get bias, $\bar{q}_\beta$ and Fisher matrix, $F_{\alpha\beta}$
4. Combine to get the **power spectrum**
5. ***Optional:*** Repeat with new cosmology



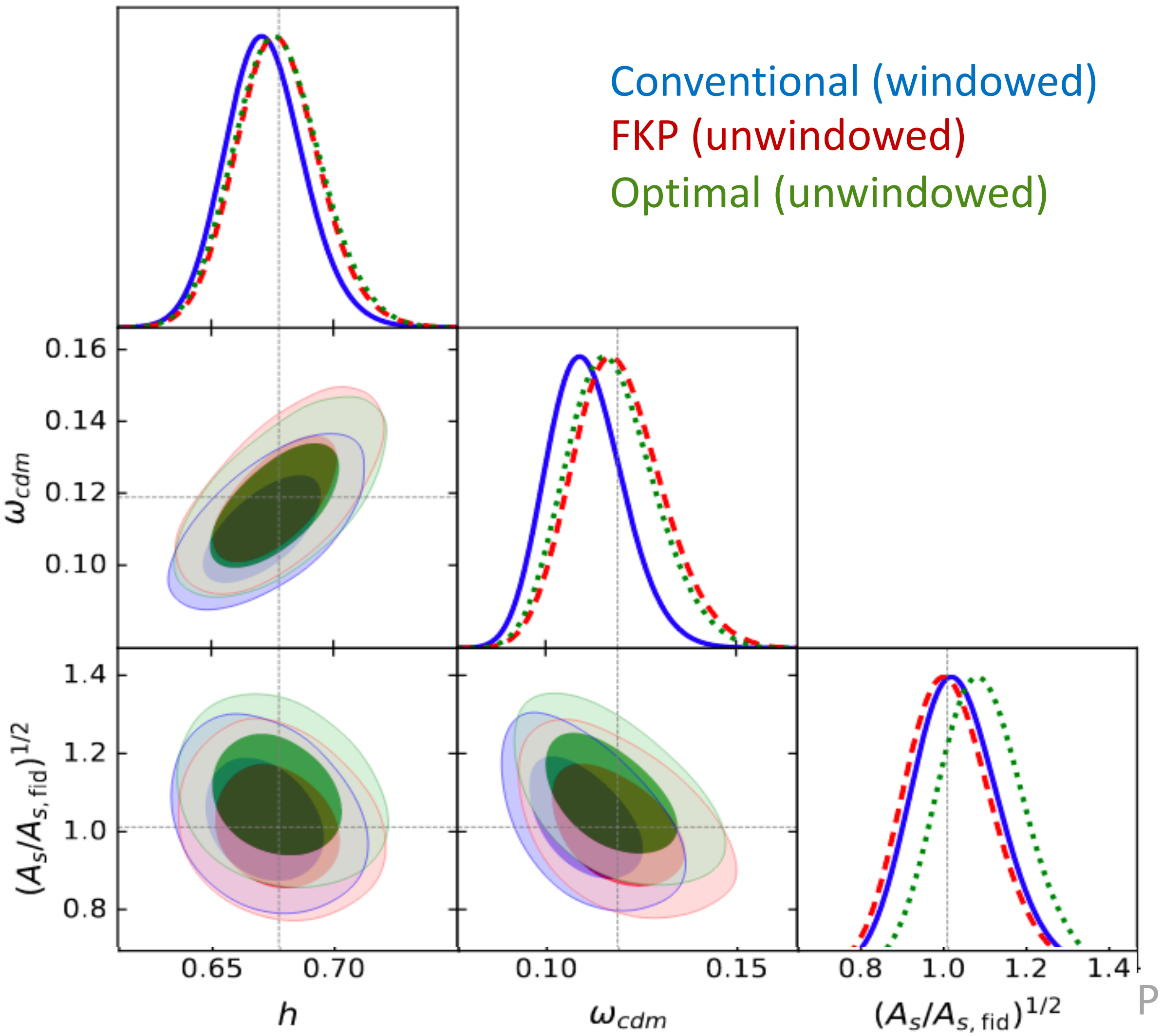
Philcox 20b

Is this useful?

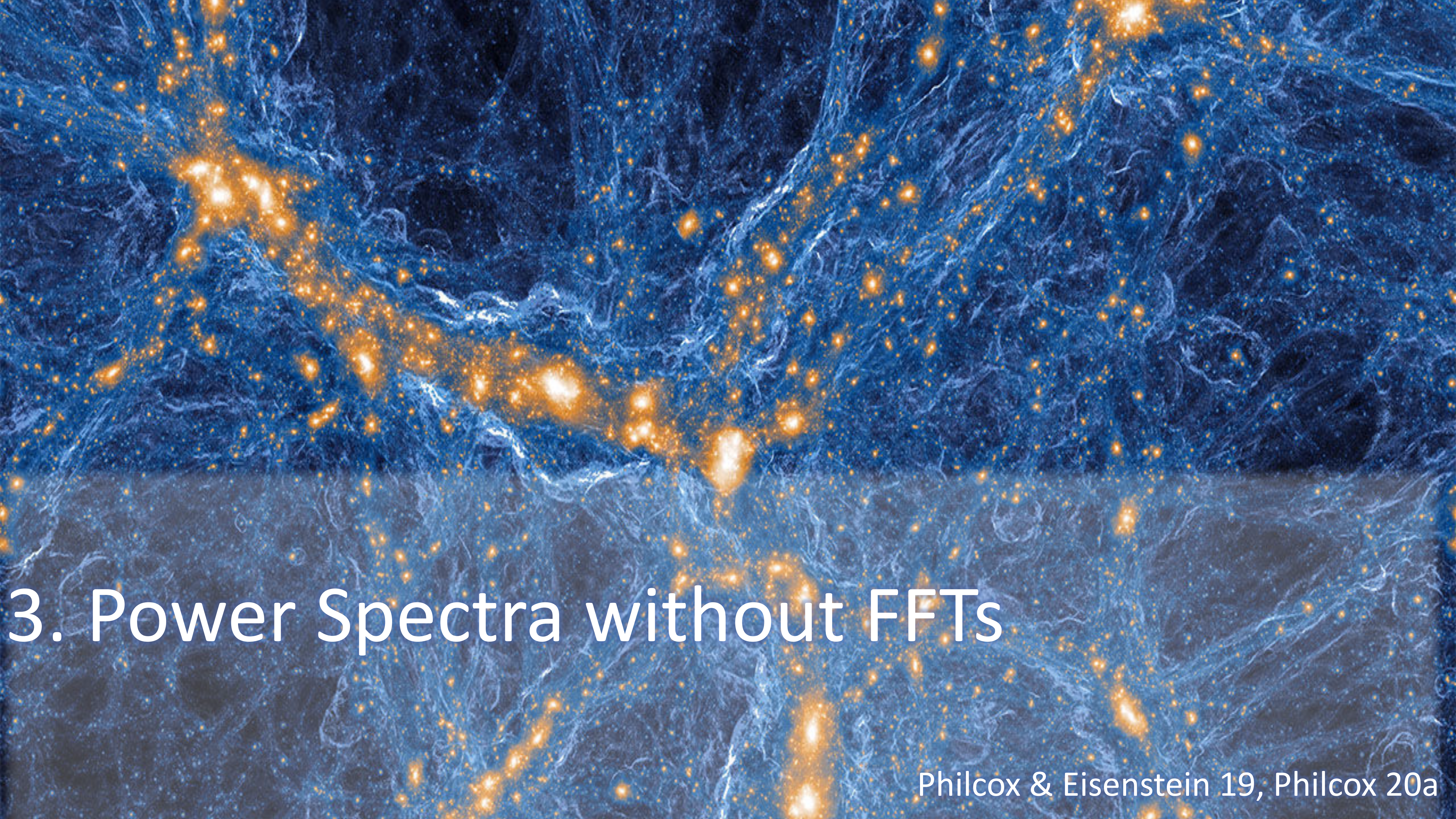
- **Benefits**:
  - No window-convolution
  - **Optimal** error-bars if Gaussian
  - Less gridding
  - Less shot-noise

  Best for **small**, **dense**, **anisotropic** surveys, and **large-scale** modes

- Especially useful for $f_{\rm NL}$ and the **bispectrum**



Philcox 20b



3. Power Spectra without FFTs

Philcox & Eisenstein 19, Philcox 20a

Configuration-Space $P(k)$ Estimators

- $P(k)$ usually estimated using Fast Fourier Transforms

$$\hat{P}(k) = \int \frac{d\Omega_k}{4\pi} \left|\mathrm{FFT}\left[\delta\right](\mathbf{k})\right|^2$$

  - **Complexity**: $\mathcal{O}(N_g \log N_g)$ for N_g grid points

- Small scales need **large** N_g $\Rightarrow$ **slow computation** and **high memory** usage!

$$\mathit{Time} \propto k_{\max} \log k_{\max}$$

- 2PCF estimated by **counting pairs of particles** with $\mathcal{O}(N^2)$ complexity

$$\xi^a = \int d\mathbf{r}_1\, d\mathbf{r}_2\, \delta(\mathbf{r}_1)\delta(\mathbf{r}_2)\Theta^a(|\mathbf{r}_1 - \mathbf{r}_2|) = \sum_{i\neq j} w_i w_j \Theta^a(|\mathbf{r}_i - \mathbf{r}_j|)$$

Weights

Binning function

Sum over galaxies

- This is **fast** on small scales!

$$\mathit{Time} \propto N n R_{\max}^3$$

Philcox & Eisenstein 19, Philcox 20

Configuration-Space $P(k)$ Estimators

- We can do the same for $P(k)$:

$$P(k) \propto \int \frac{d\Omega_k}{4\pi} \sum_{i\neq j} w_i w_j e^{-i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j)} = \sum_{i\neq j} w_i w_j j_0(k|\mathbf{r}_i-\mathbf{r}_j|)$$

Weights

0^{th} order Bessel function

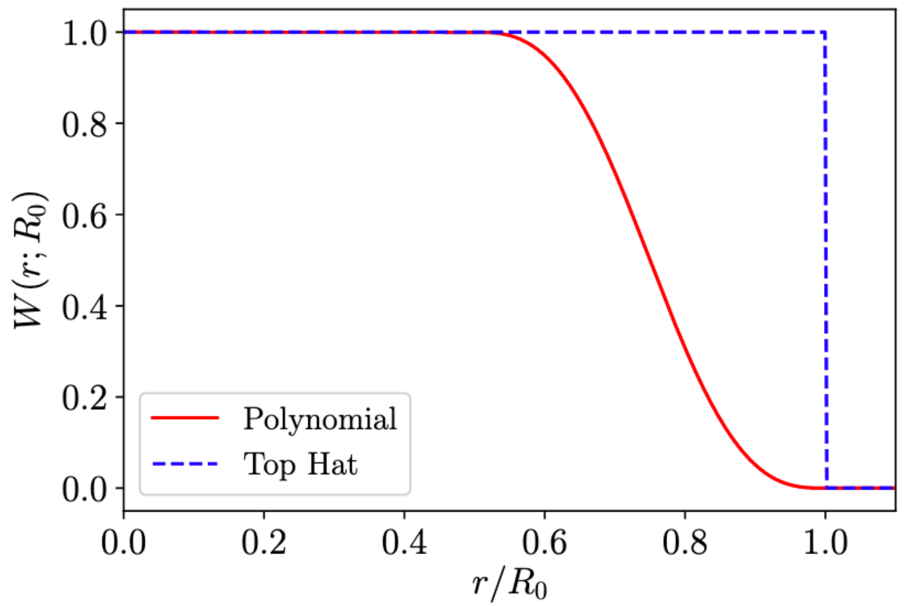
Sum over galaxies

- But we need to sum over all N^2 pairs of galaxies in the survey!

  - Only sum up to some maximum radius R_0, via a smooth function $W(r;R_0)$

$$P(k;R_0) \propto \sum_{i\neq j} w_i w_j j_0(k|\mathbf{r}_i-\mathbf{r}_j|) W(|\mathbf{r}_i-\mathbf{r}_j|;R_0)$$

$$Time \propto NnR_0^3 \propto k_{\min}^{-3}$$

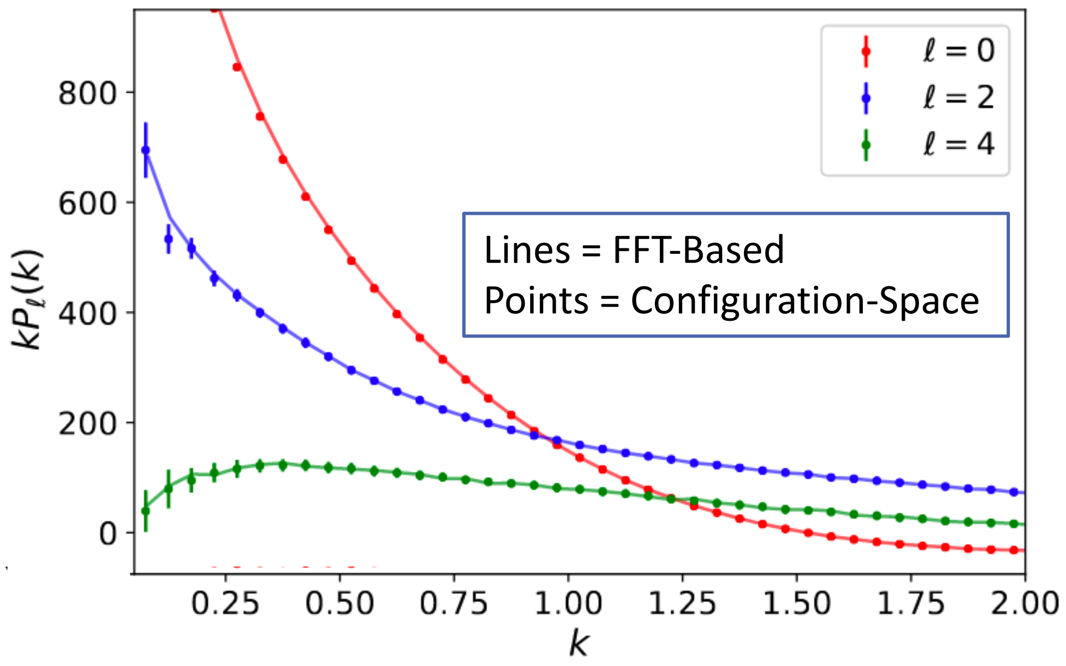


Philcox & Eisenstein 19, Philcox 20

Configuration-Space $P_\ell(k)$ Estimators

<u>Benefits</u>

- **Speed**
  - Time scales as k_{min}^{-3}
- **Memory**
  - No storage of large FFT grids
- **Aliasing**
  - No gridding!
- **Shot-noise**
  - Removes self-counts -> Poissonian shot-noise!
- **Window function**
  - Can remove survey window, just as for 2PCF



Philcox & Eisenstein 19, Philcox 20

Configuration-Space $P_\ell(k)$ Estimators

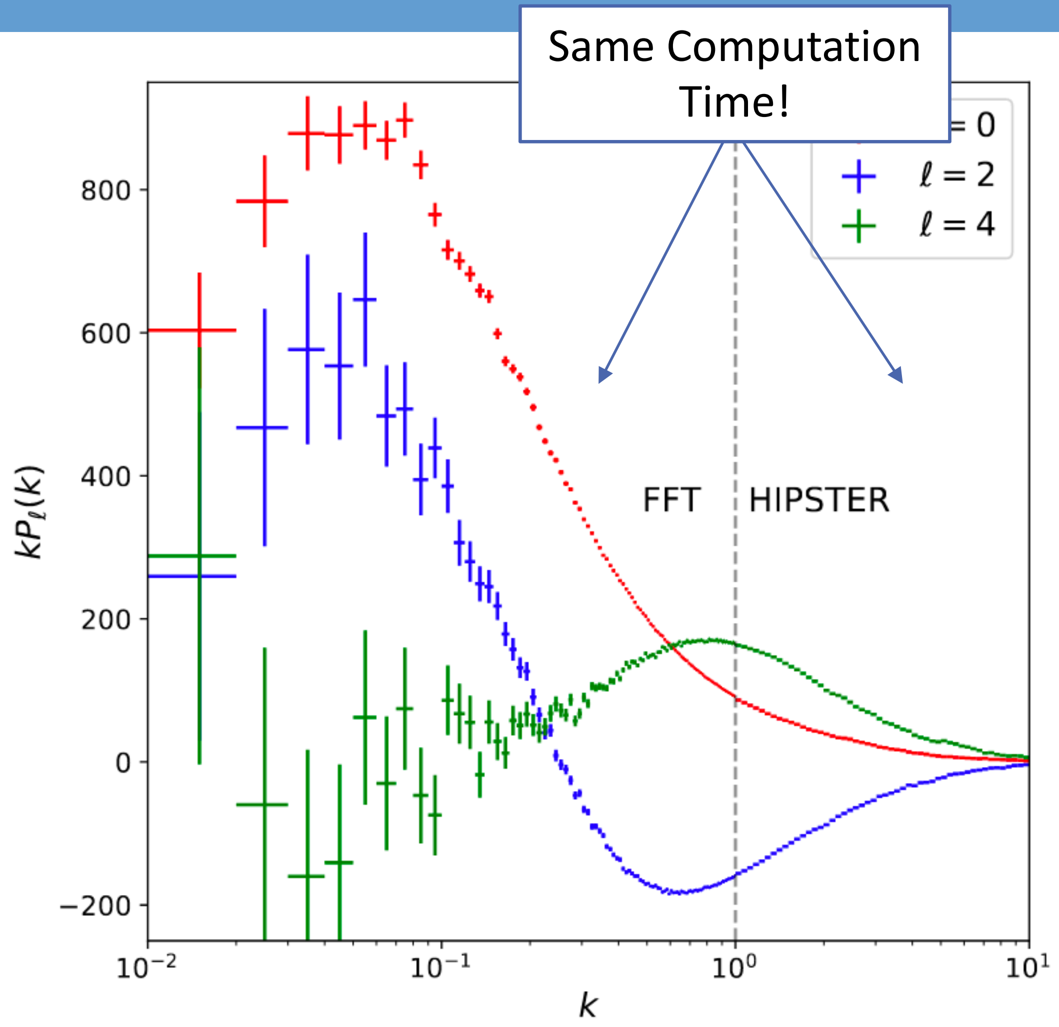
- Implemented in the **HIPSTER** code

- **Combine** with FFT-based treatments:

  - FFTs are fastest on **large** scales (time $\sim k_{\max} \log \mathrm{k}_{\max}$)
  - HIPSTER is fastest on **small** scales (time $\sim k_{\min}^{-3}$)

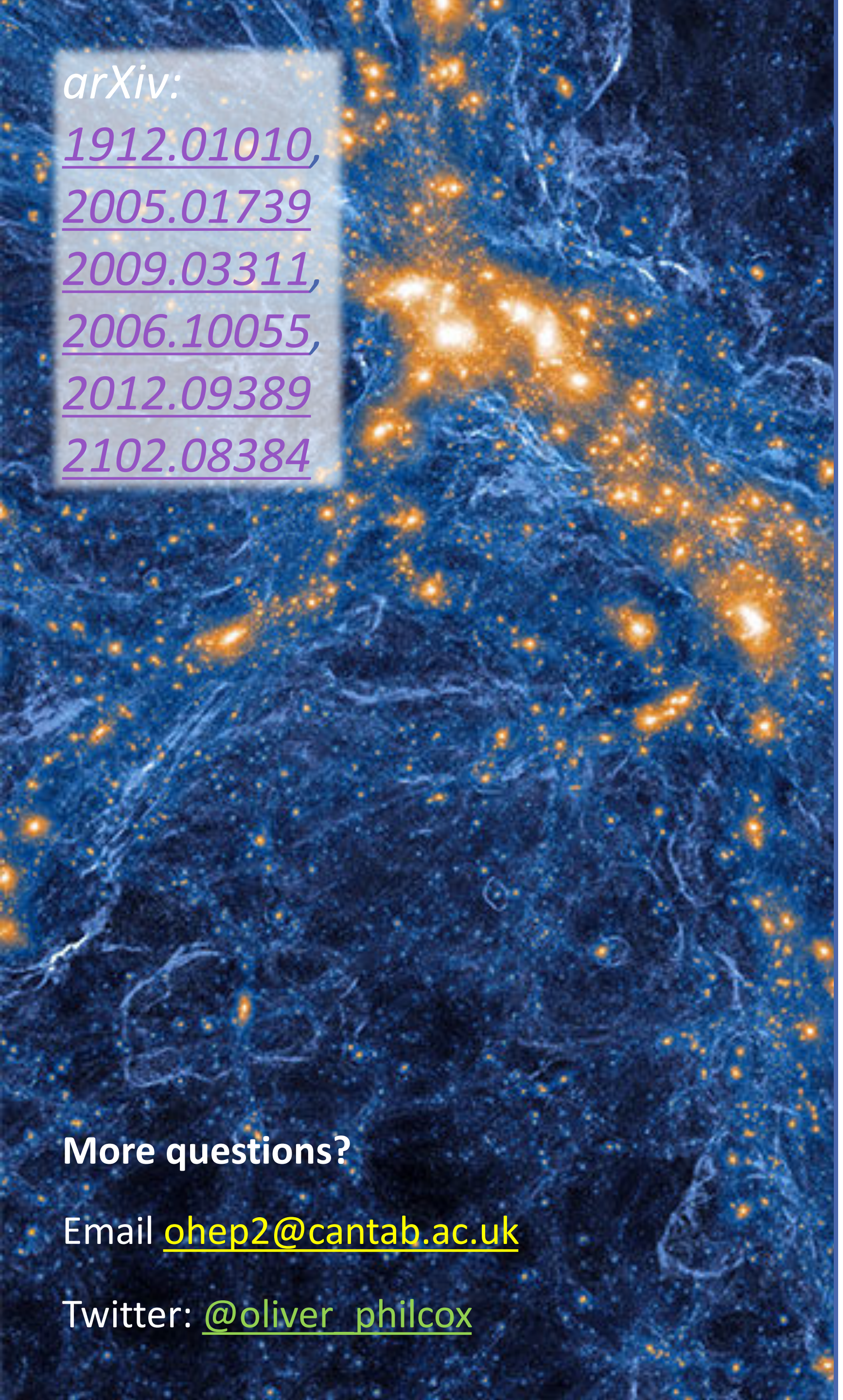
- Can be similarly applied to **bispectra**

  - Time $\propto N n^2 R_0^6 \propto k_{\min}^{-6}$
  - Same scaling with number density as for $P(k)$!



Philcox & Eisenstein 19, Philcox 20

HIPSTER.rtfd.io



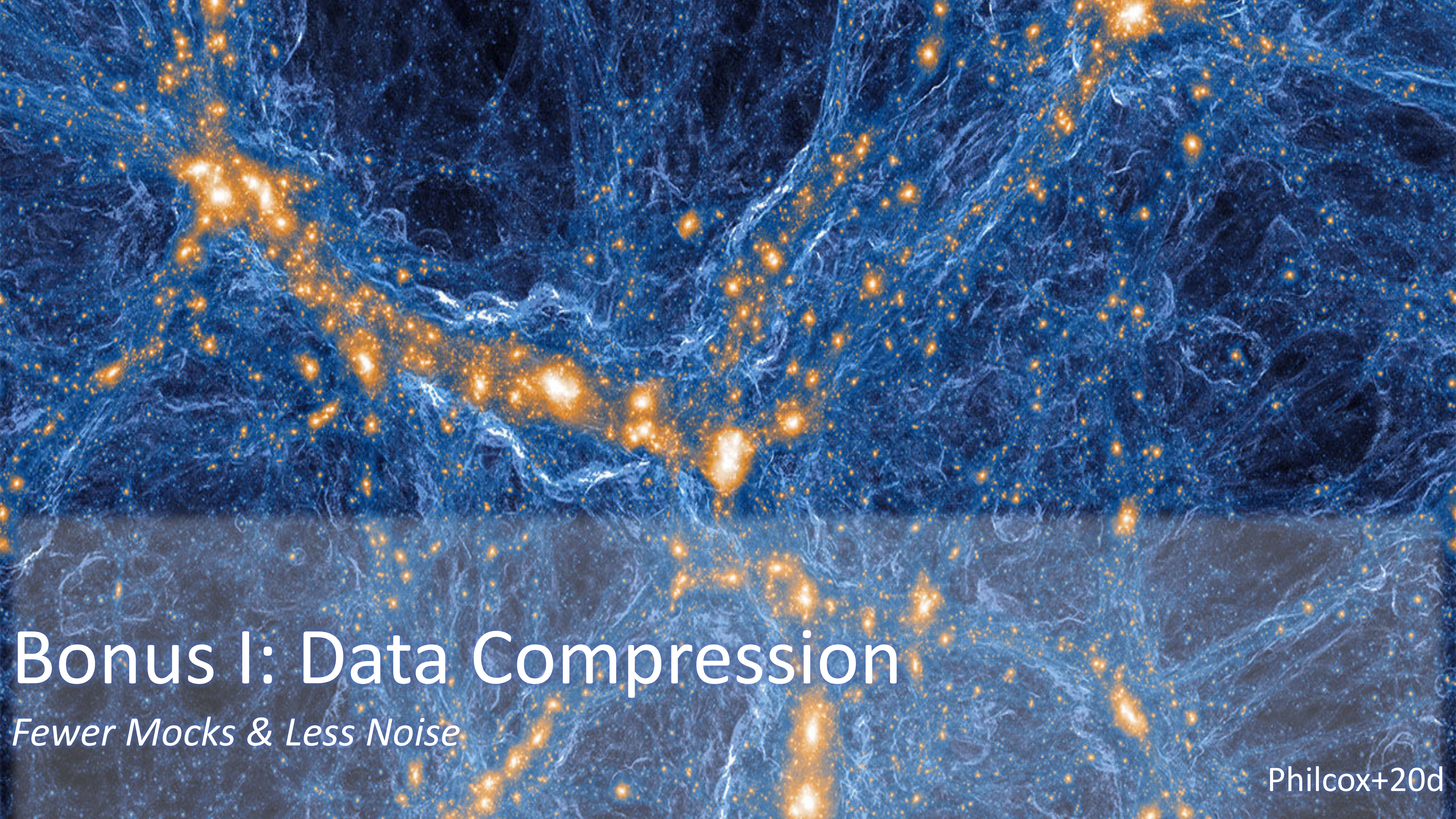
Conclusions

- We're not finished with the galaxy power spectrum yet!

- Recent updates include:
  - More accurate lines-of-sight
  - Closer to optimal large-scale $P_\ell(k)$ estimation
  - Faster small-scale computation *without* FFTs
  - *(Powerful analysis-specific data compression)*

Coming soon:

- Estimating the bispectrum and beyond!



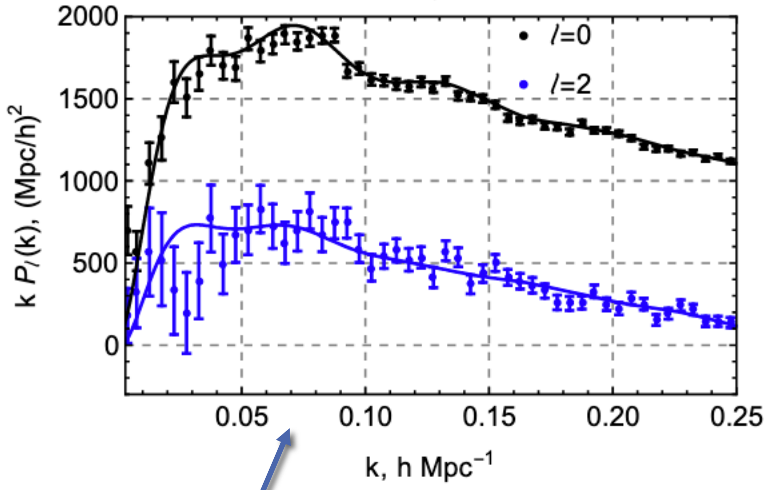
Bonus I: Data Compression

Fewer Mocks & Less Noise

Philcox+20d

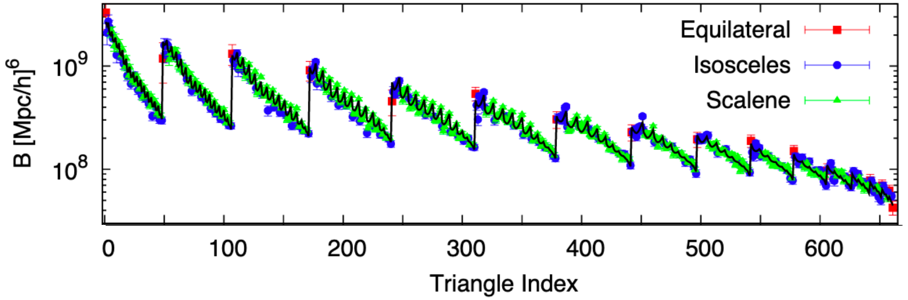
The Curse of Dimensionality

- $P_\ell(k)$ is **high-dimensional**, e.g.;
  - BOSS has ~ 100 bins
  - Only use these to measure ~ 10 parameters
- Conventional likelihoods use a **sample covariance**
  - Need $N_{mocks} > N_{bins}$ to invert
  - Too few mocks ⇒ **parameter shifts** or **error inflation**
- We should **compress** our data!



Power Spectrum

Bispectrum



Gil-Marin+16, Ivanov+19

Data Compression via PCA

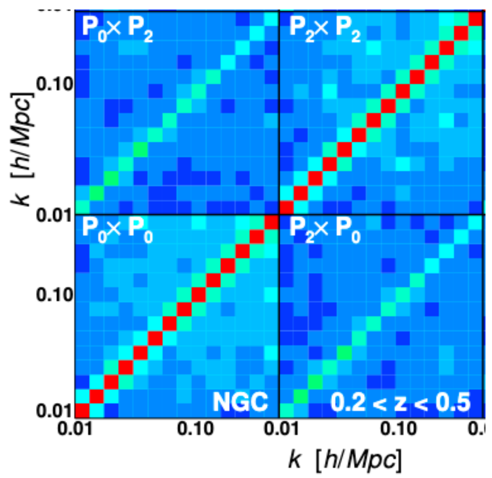
Beutler+16

- A canonical approach: [e.g. Scoccimarro 2000]
  - Compute the theoretical **covariance** matrix
  - Perform a **P**rincipal **C**omponent **A**nalaysis
  - Project the data onto the first few components

- This chooses the basis vectors that contribute most to the signal-to-noise

- Signal-to-noise isn't everything!

Power Spectrum Covariance



PCA

$$P(k) \approx \sum_i \mathrm{a}_i W_i(k)$$

Coefficients

Basis Vectors

See also MOPED: Heavens+00, Alsing+18 , KL: Tegmark+97

Data Compression via Subspace Projection

New* approach

- Draw sets of parameters from the **priors**
- Compute the **theory model** at each point
- Perform a **S**ingular **V**alue **D**ecomposition on the **noise-weighted** samples
- Use these **basis vectors** to perform the compression

Picks out directions contributing most to the **log-likelihood**

Parameters used in the analysis

$$\theta = \{\omega_{\rm cdm}, A_s/A_{s,{\rm fid}}, h, ...\} \times \{b_1, b_2, b_{G_2}, b_4, c_{s,0}, c_{s,2}, P_{\rm shot}\}$$

$$X_a(\theta) \equiv \sum_{ab} \mathsf{C}_{ab}^{-1/2}\left[P_b(\theta) - \overline{P}_b\right]$$

Covariance Estimate *Theory Model*

$$X_{ia} = \sum_\alpha U_{i\alpha} D_\alpha V_{\alpha a}$$

Basis Vectors

$$X_a^{(i)} \approx \sum_{\alpha=1}^{N_{\rm SV}} c_\alpha^{(i)} V_{\alpha a}$$

Subspace Coefficients

*somewhat inspired by gravitational wave analyses [e.g. Roulet+19]

Philcox+20d

Data Compression via Subspace Projection

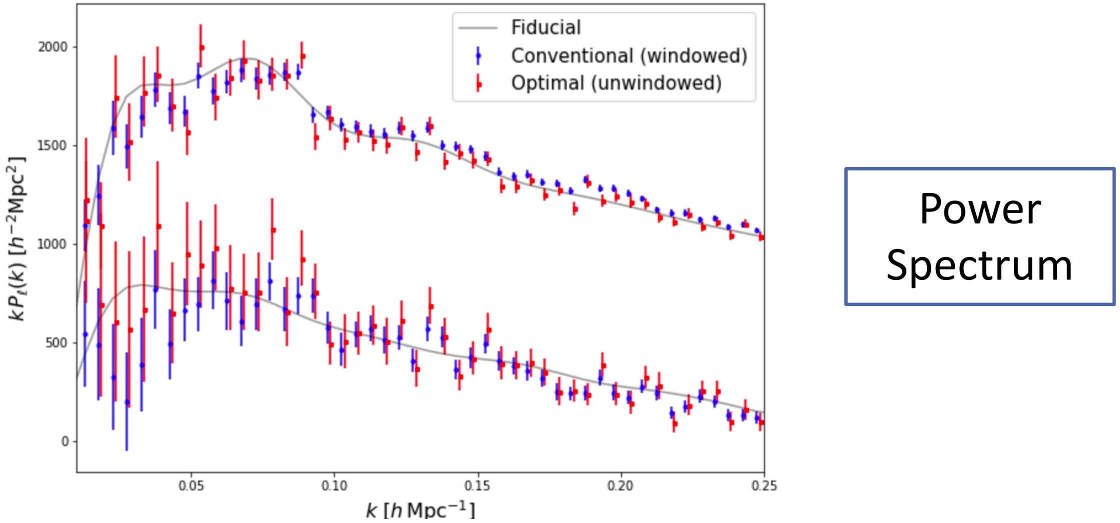
- This is the **best** linear compression for a **specific** analysis
- Set the **number** of basis vectors **robustly**
- Estimate coefficients **optimally**

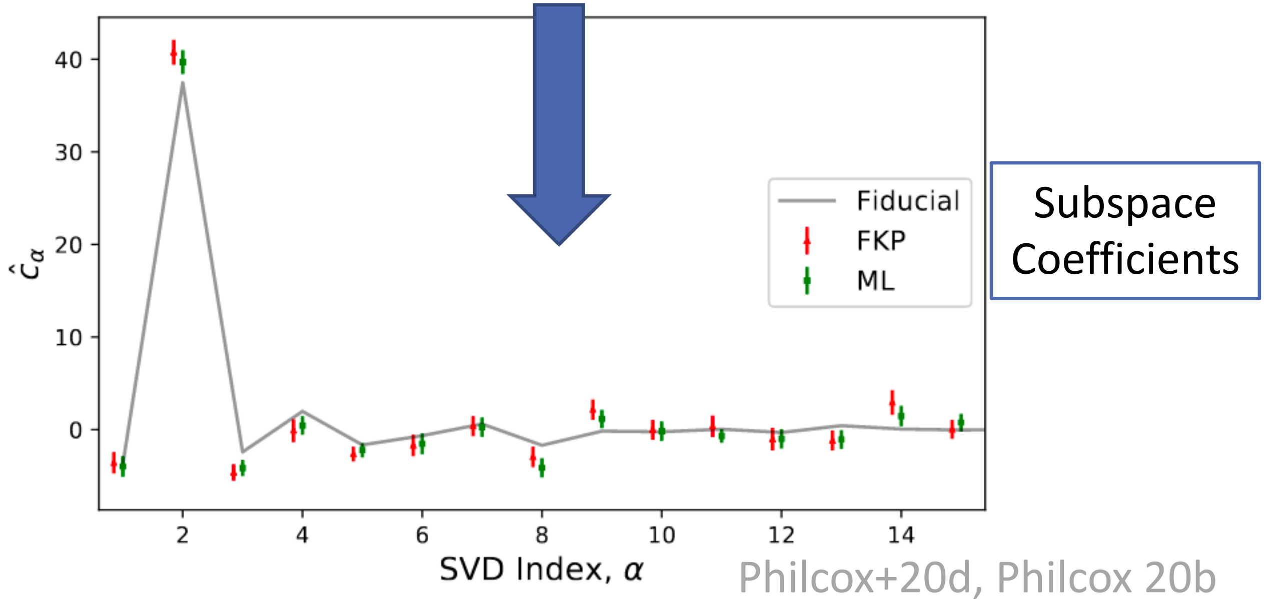
For BOSS 10-parameter analysis:

- **100**-bin P(k) -----> **12** subspace coefficients
- **2135**-bin $B(k_1,k_2)$ ----> **8** subspace coefficients

Applicable to **any** analysis given:

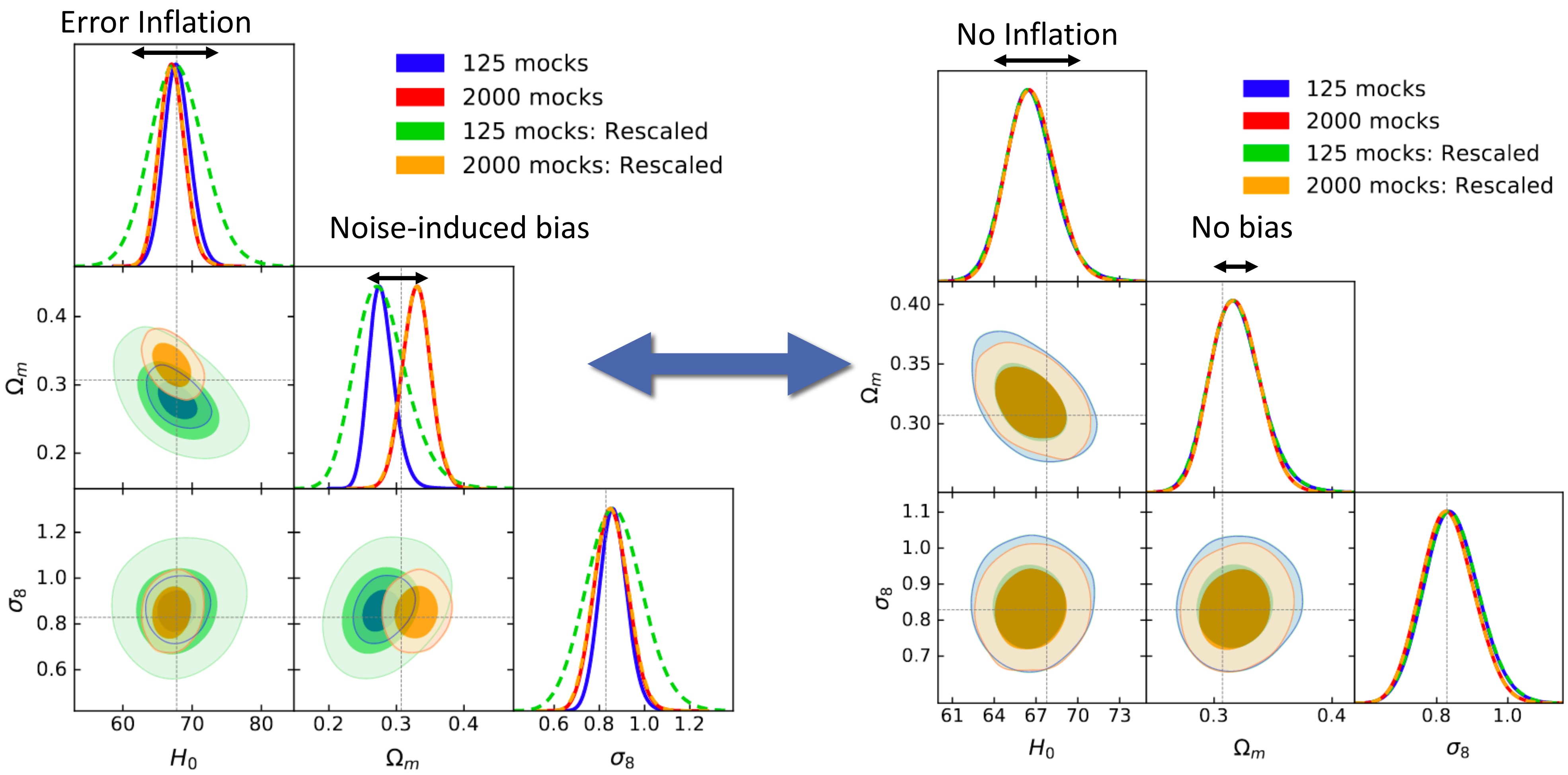
1. Theory Model
2. Parameter Priors
3. Rough Covariance Estimate





Philcox+20d, Philcox 20b

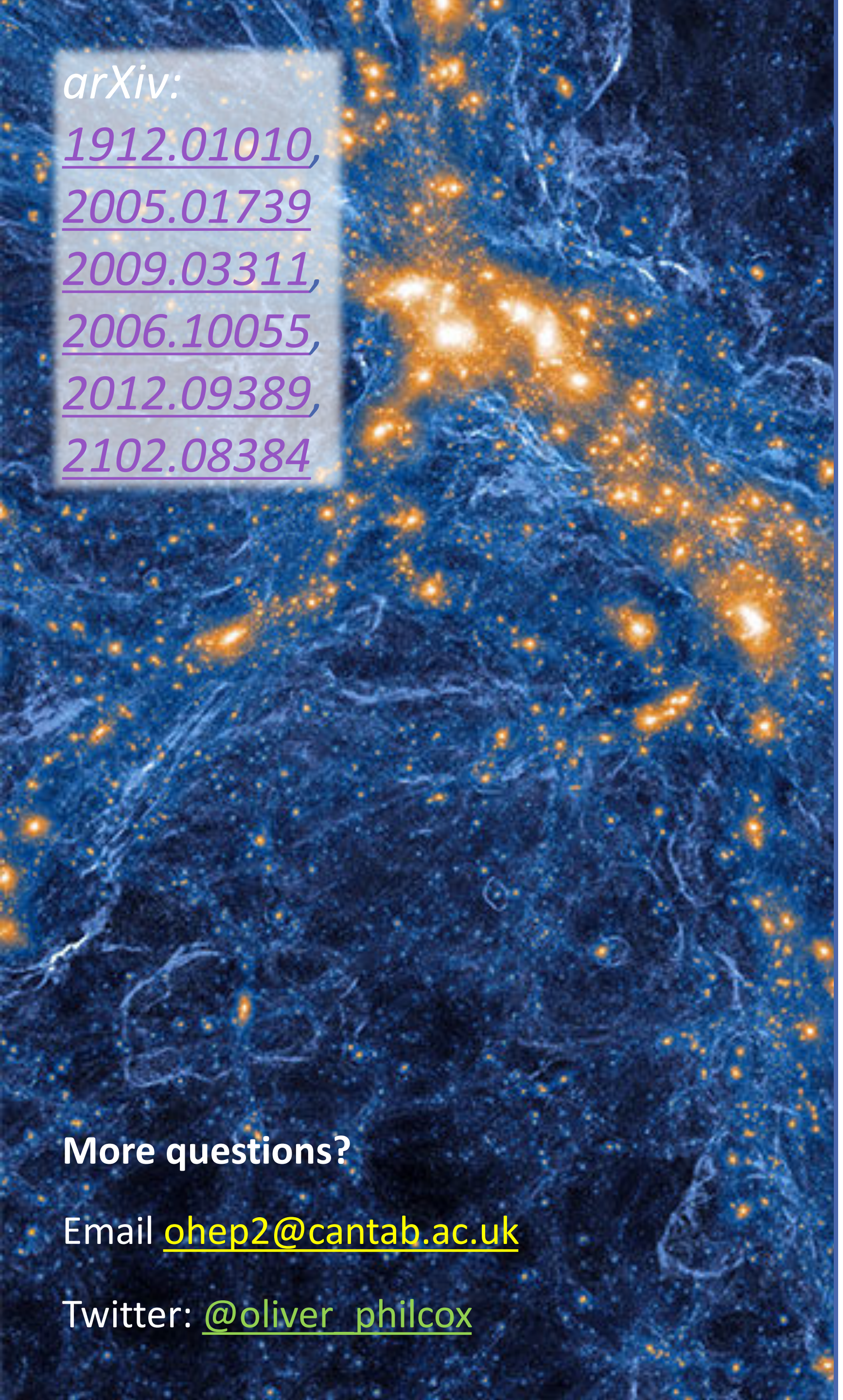
Too Few Mocks -> Parameter Biases



(a) 96-bin Power Spectrum

(c) 12 Subspace Coefficients

Philcox+20d



Conclusions

- We're not finished with the galaxy power spectrum yet!

- Recent updates include:
  - More accurate lines-of-sight
  - Closer to optimal large-scale $P_\ell(k)$ estimation
  - Faster small-scale computation *without* FFTs
  - Powerful analysis-specific data compression

Coming soon:

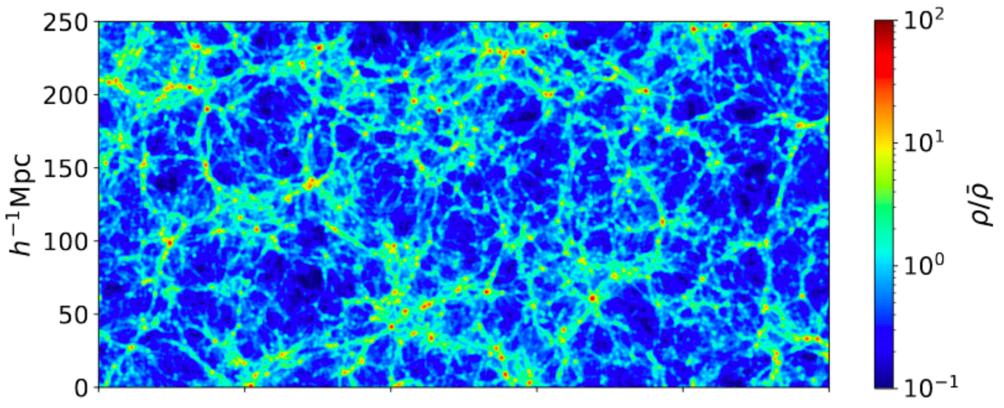
- Estimating the bispectrum and beyond!

Bonus II: Alternative 2-Point Statistics

Philcox+20c, Philcox+20e

Beyond the Density Field

- What should we compute the two-point function of?
  - For a **Gaussian** universe, the power-spectrum of galaxy overdensity contains **all** the information
- The Universe is **not** Gaussian:
  - Information **cascades** to the higher-point functions
  - **Low-density regions** carry a lot of cosmological information, and contribute little to δ [e.g. Pisani+19]
- Can use a **transformed** field, e.g.:
  - Reconstructed Density Fields [e.g. Eisenstein+07]
  - Log-normal Transforms [Neyrinck+09, Wang+11]
  - Gaussianized Density Fields [Weinberg 92, Neyrinck+17]
  - Marked Density Fields [Stoyan 84, White 16, Massara+20]

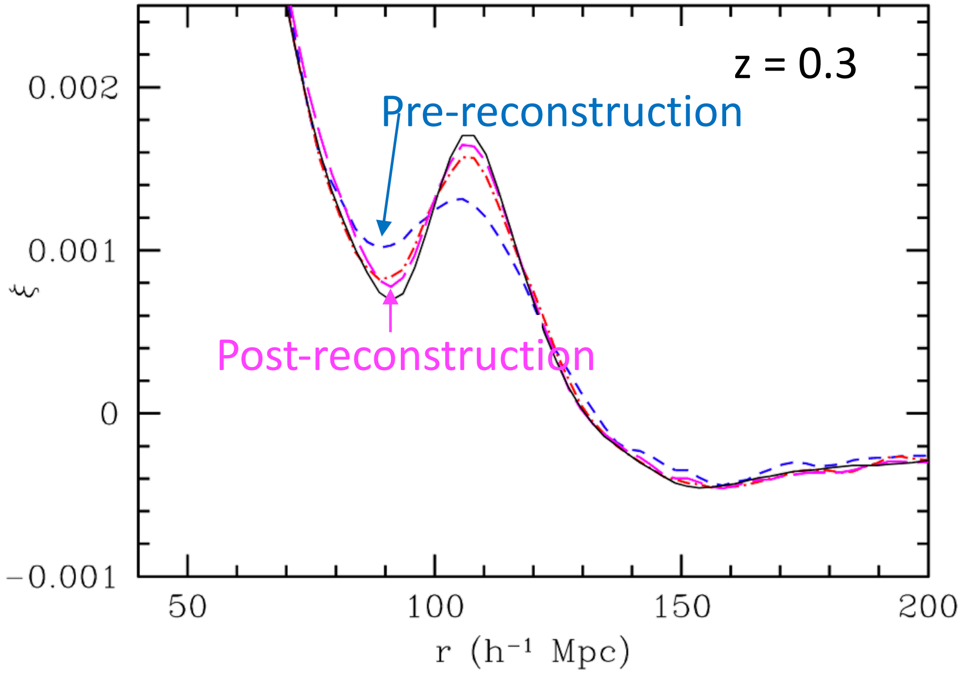


Overdensity Field

Massara+20

Beyond the Density Field

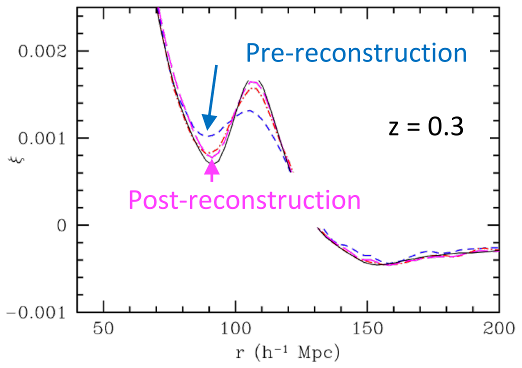
- What should we compute the two-point function of?
  - For a **Gaussian** universe, the power-spectrum of galaxy overdensity contains **all** the information
- The Universe is **not** Gaussian:
  - Information **cascades** to the higher-point functions
  - **Low-density regions** carry a lot of cosmological information, and contribute little to δ [e.g. Pisani+19]
- Can use a **transformed** field, e.g.:
  - Reconstructed Density Fields [e.g. Eisenstein+07]
  - Log-normal Transforms [Neyrinck+09, Wang+11]
  - Gaussianized Density Fields [Weinberg 92, Neyrinck+17]
  - Marked Density Fields [Stoyan 84, White 16, Massara+20]



Eisenstein+06, Padamanbhan+12

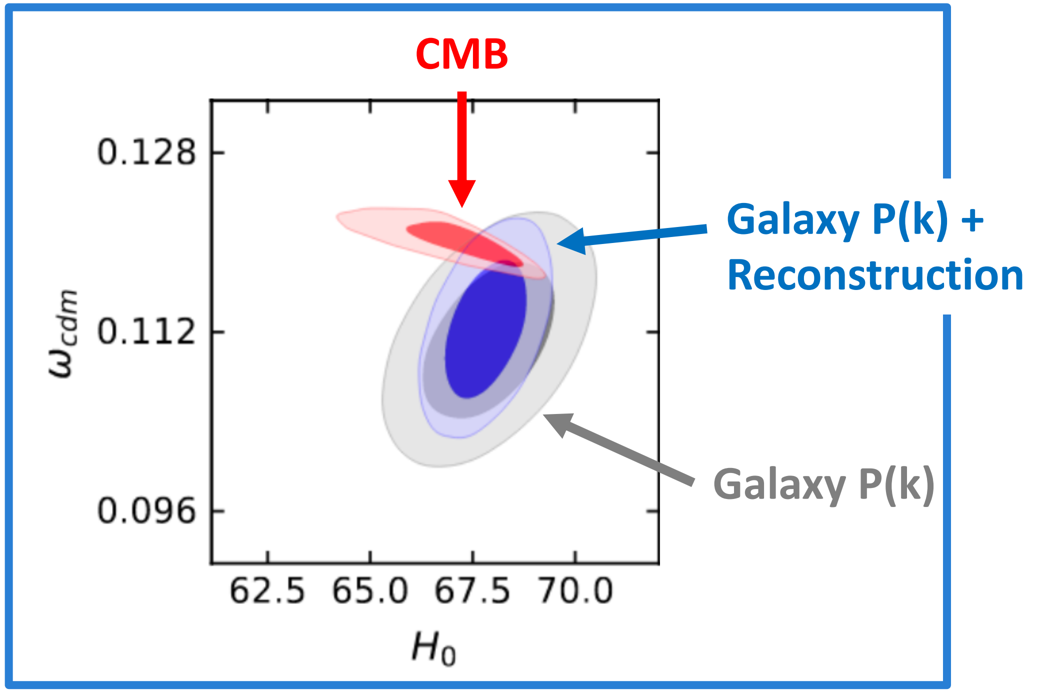
Beyond the Density Field

- What should we compute the two-point function of?
  - For a **Gaussian** universe, the power-spectrum of galaxy overdensity contains **all** the information



- The Universe is **not** Gaussian:
  - Information **cascades** to the higher-point functions
  - **Low-density regions** carry a lot of cosmological information, and contribute little to δ [e.g. Pisani+19]

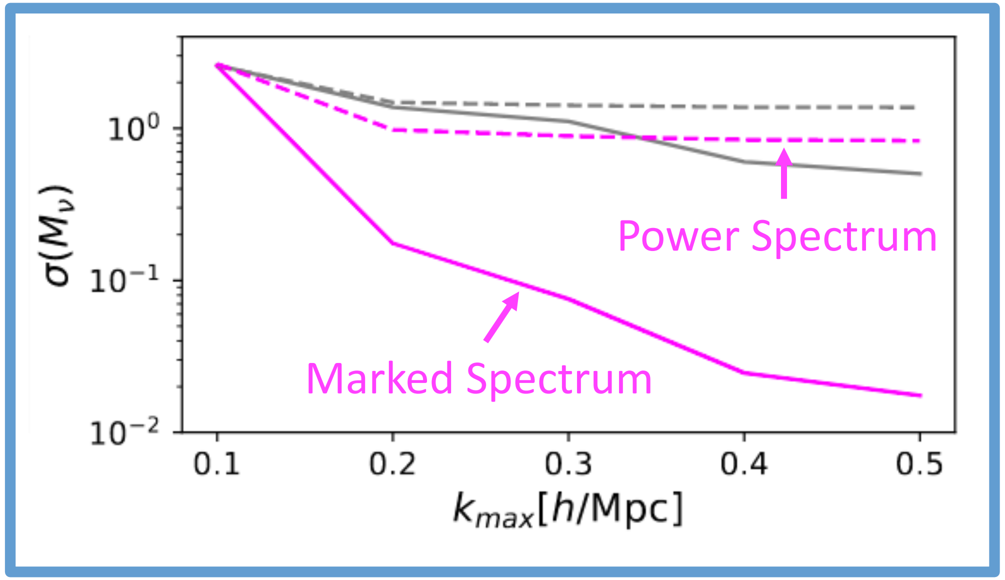
- Can use a **transformed** field, e.g.:
  - Reconstructed Density Fields [e.g. Eisenstein+07]
  - Log-normal Transforms [Neyrinck+09, Wang+11]
  - Gaussianized Density Fields [Weinberg 92, Neyrinck+17]
  - Marked Density Fields [Stoyan 84, White 16, Massara+20]



Philcox+20a

Beyond the Density Field

- What should we compute the two-point function of?
  - For a **Gaussian** universe, the power-spectrum of galaxy overdensity contains **all** the information
- The Universe is **not** Gaussian:
  - Information **cascades** to the higher-point functions
  - **Low-density regions** carry a lot of cosmological information, and contribute little to δ [e.g. Pisani+19]
- Can use a **transformed** field, e.g.:
  - Reconstructed Density Fields [e.g. Eisenstein+07]
  - Log-normal Transforms [Neyrinck+09, Wang+11]
  - Gaussianized Density Fields [Weinberg 92, Neyrinck+17]
  - Marked Density Fields [Stoyan 84, White 16, Massara+20]



Fisher Matrix Constraints on Neutrino Mass

Massara+20

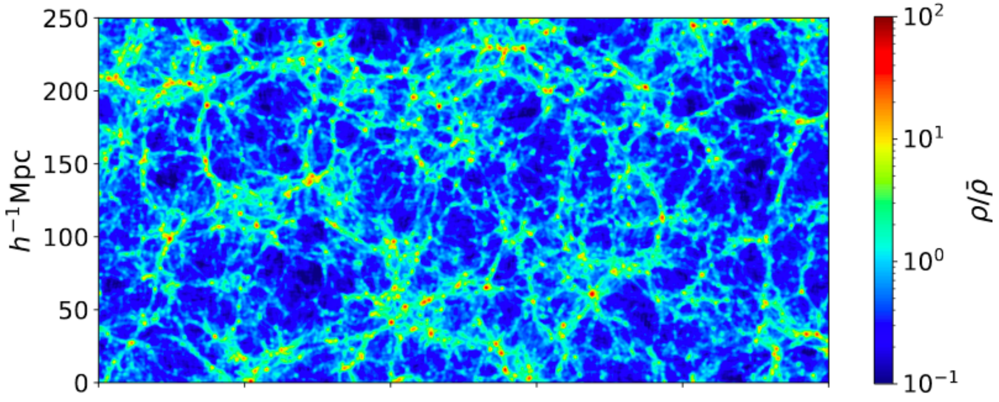
The Marked Density Field

- Define a new density field by weighting by the **mark**

$$m(\mathbf{x}) = \left(\frac{1+\delta_s}{1+\delta_s+\delta_R(\mathbf{x})}\right)^p$$

$$\rho_M(\mathbf{x}) = m(\mathbf{x})n(\mathbf{x}) = m(\mathbf{x})\bar{n}\left[1+\delta(\mathbf{x})\right]$$

depending on **smoothed** overdensity $\delta_R(\mathbf{x})$



Unweighted Density Field

Massara+20

The Marked Density Field

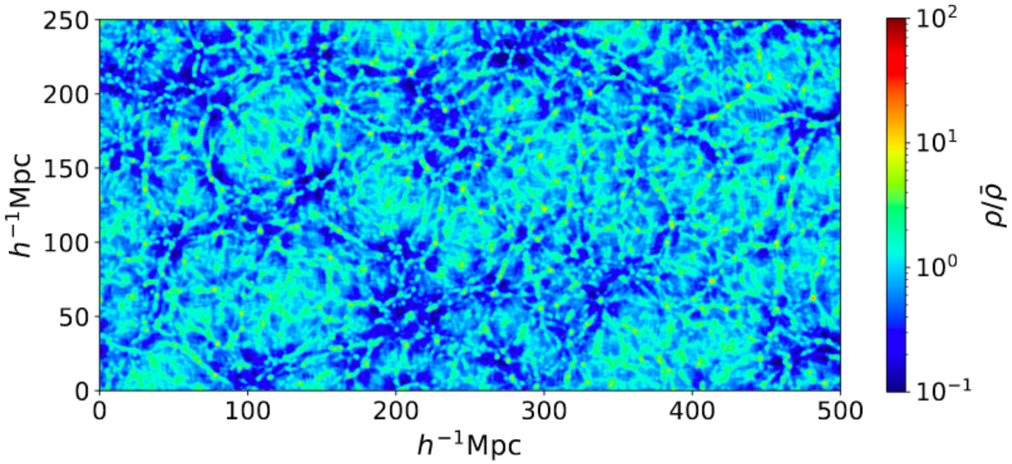
- Define a new density field by weighting by the **mark**

$$m(\mathbf{x}) = \left(\frac{1+\delta_s}{1+\delta_s+\delta_R(\mathbf{x})}\right)^p$$

$$\rho_M(\mathbf{x}) = m(\mathbf{x})n(\mathbf{x}) = m(\mathbf{x})\bar{n}\left[1+\delta(\mathbf{x})\right]$$

depending on **smoothed** overdensity $\delta_R(\mathbf{x})$

- Significantly enhances constraints on:
  - **Neutrino masses** [Massara+20]
  - **Modified gravity** [White 16]



Marked Density Field

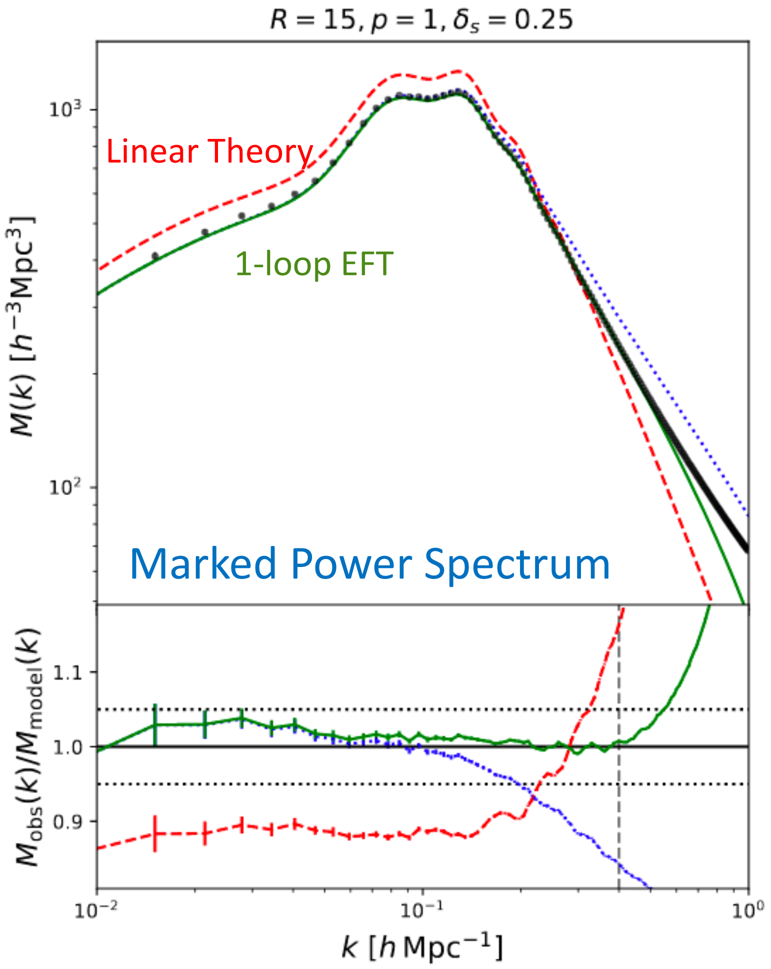
Massara+20

The Marked Density Field

- Can we **model** the marked spectrum?
  - Yes! Using **E**ffective **F**ield **T**heory

- Can we **understand** the impressive information content?
  - The mark couples **small-scale** non-Gaussianities to **large-scale** modes
  - So we find more neutrino information at low-k!

- But:
  - Modelling is **difficult** at low-z
  - Is it still useful for galaxies?

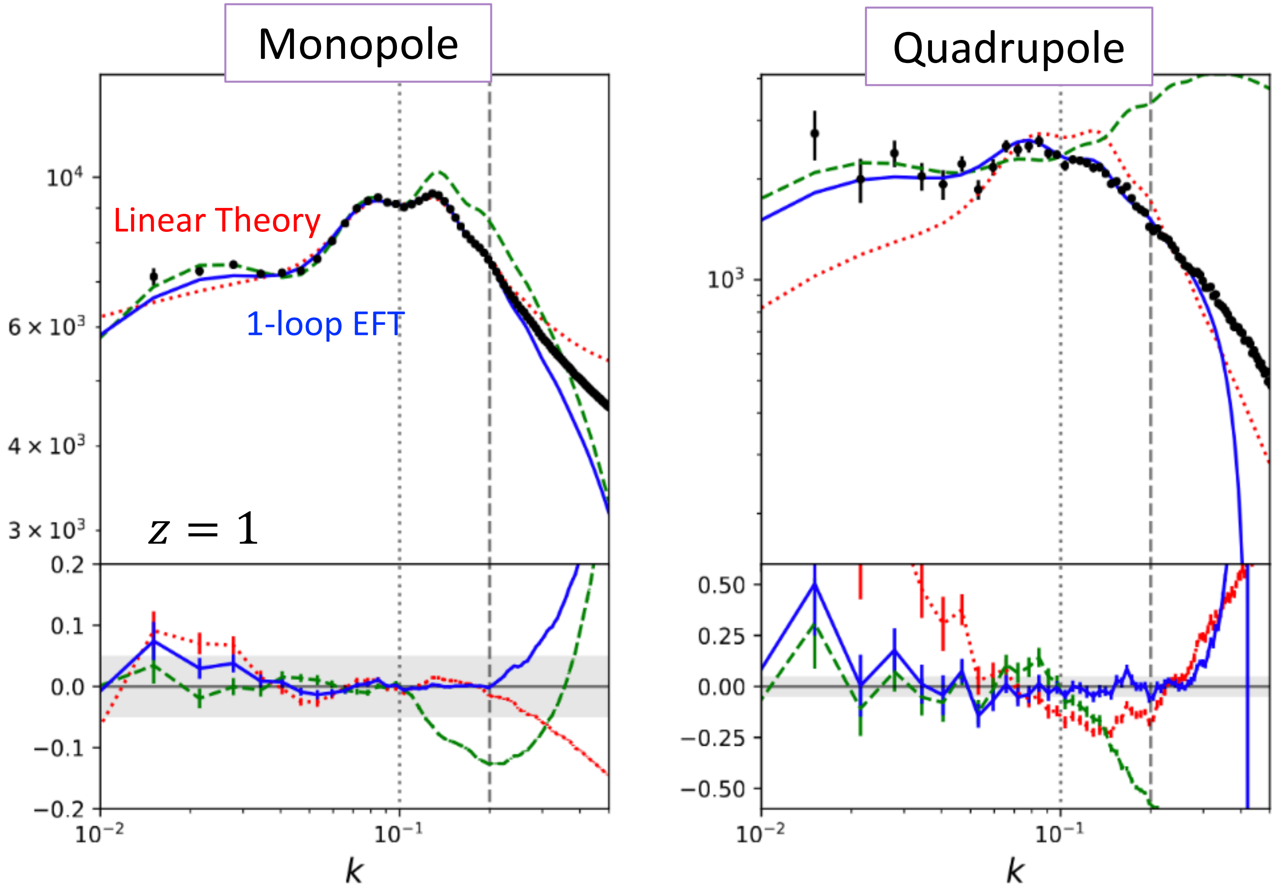


Matter at $z = 1$

Massara+20, Philcox+20ce

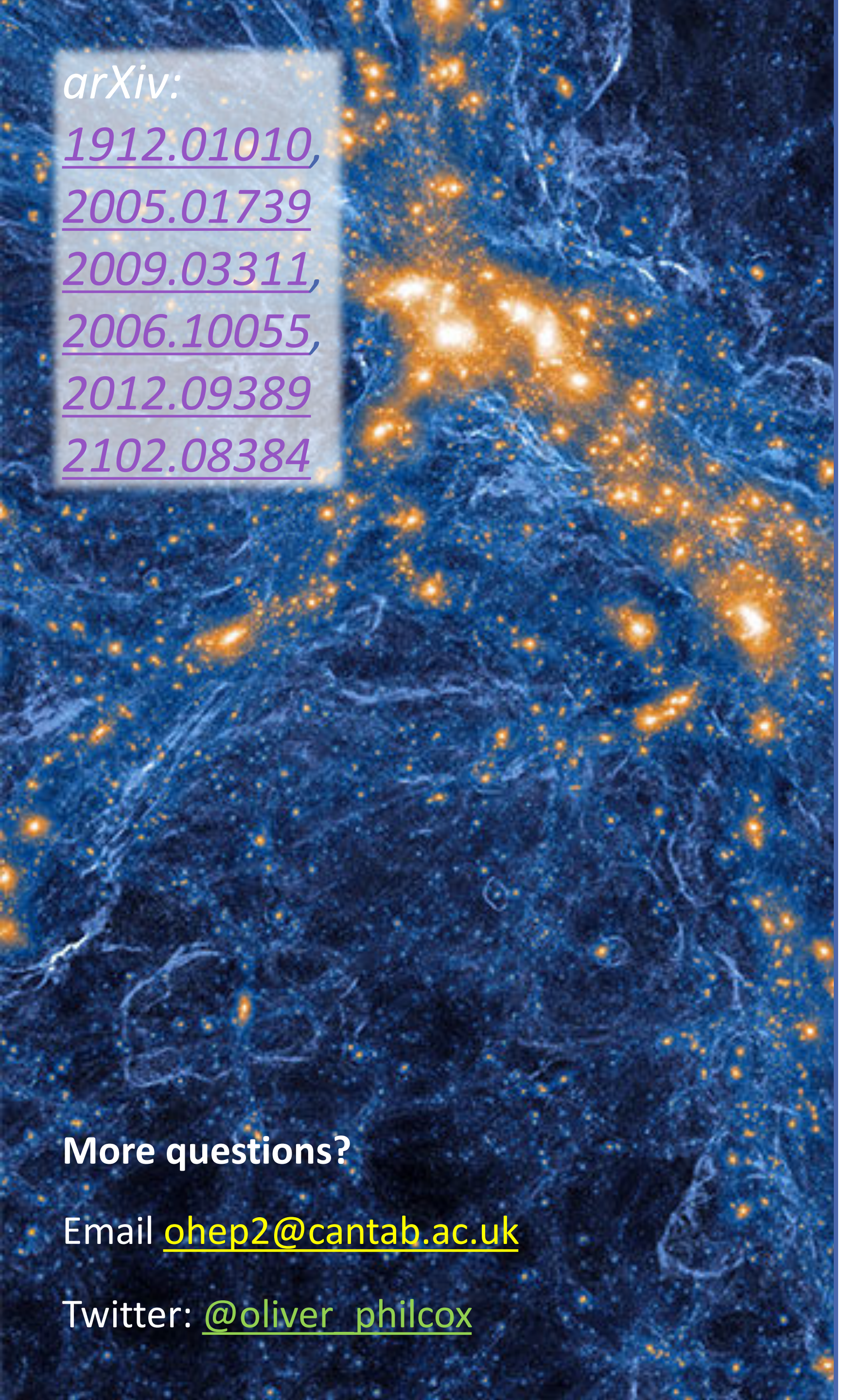
The Marked Density Field

- Can we **model** the marked spectrum?
  - Yes! Using **E**ffective **F**ield **T**heory
- Can we **understand** the impressive information content?
  - The mark couples **small-scale** non-Gaussianities to **large-scale** modes
  - So we find more neutrino information at low-k!
- But:
  - Modelling is **difficult** at low-z
  - Is it still useful for galaxies?



Galaxies at $z = 1$

Massara+20, Philcox+20ce



arXiv:
1912.01010, 2005.01739 2009.03311, 2006.10055, 2012.09389 2102.08384

Conclusions

- We're not finished with the galaxy power spectrum yet!

- Recent updates include:
  - More accurate lines-of-sight
  - Closer to optimal large-scale $P_\ell(k)$ estimation
  - Faster small-scale computation *without* FFTs
  - Powerful analysis-specific data compression
  - Statistics **beyond** the density field

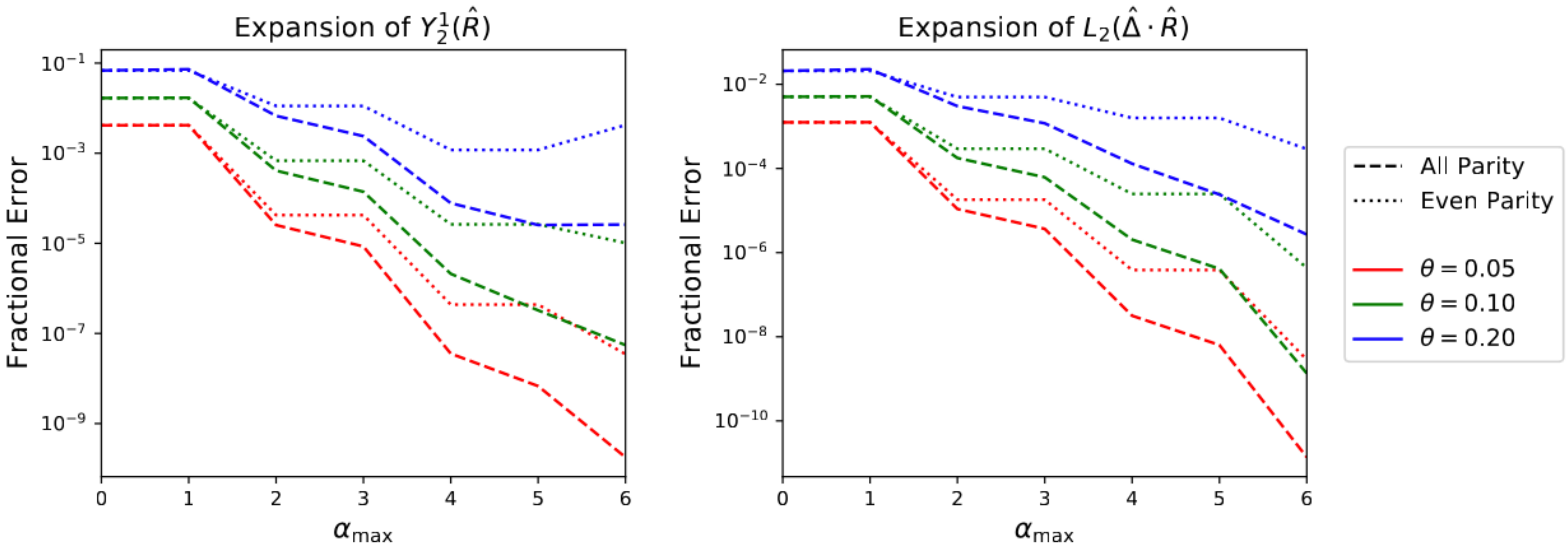
Coming soon:
  - Estimating the bispectrum and beyond!

More questions?

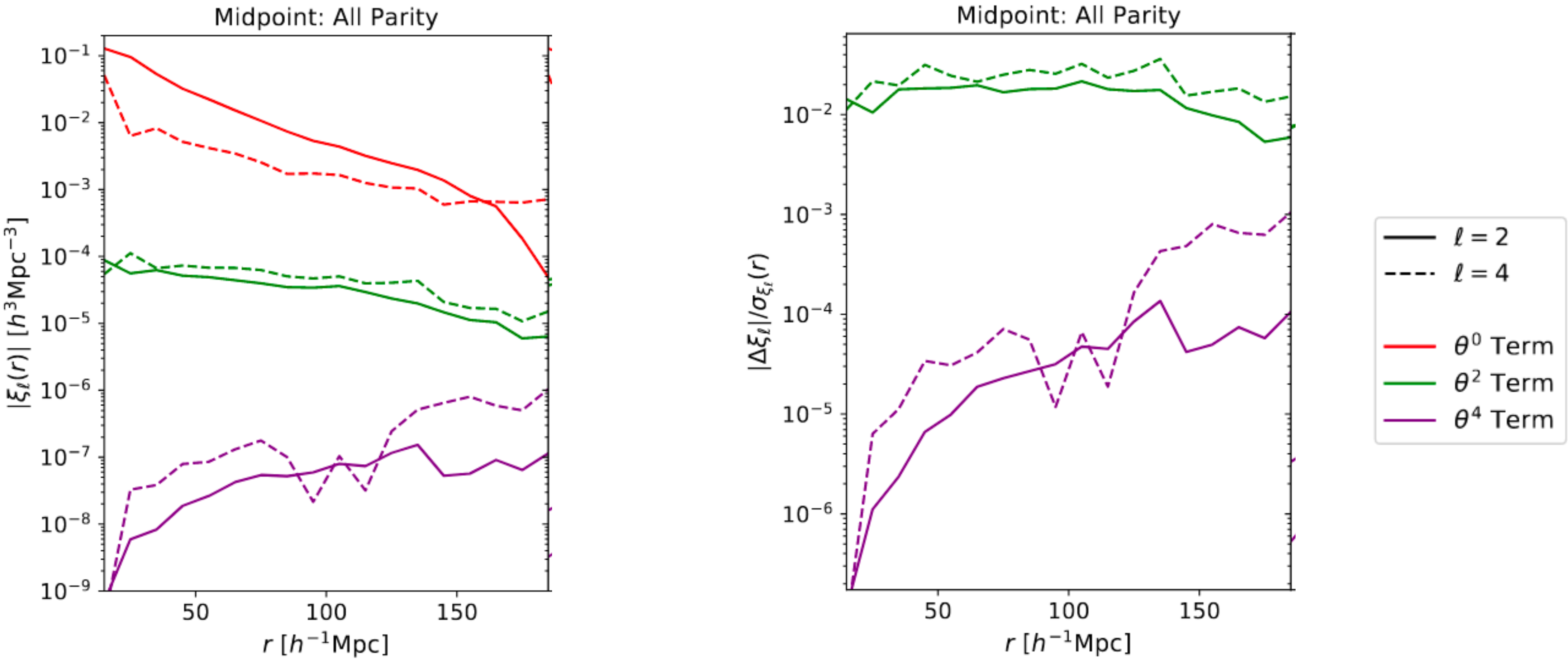
Email ohep2@cantab.ac.uk

Twitter: @oliver_philcox

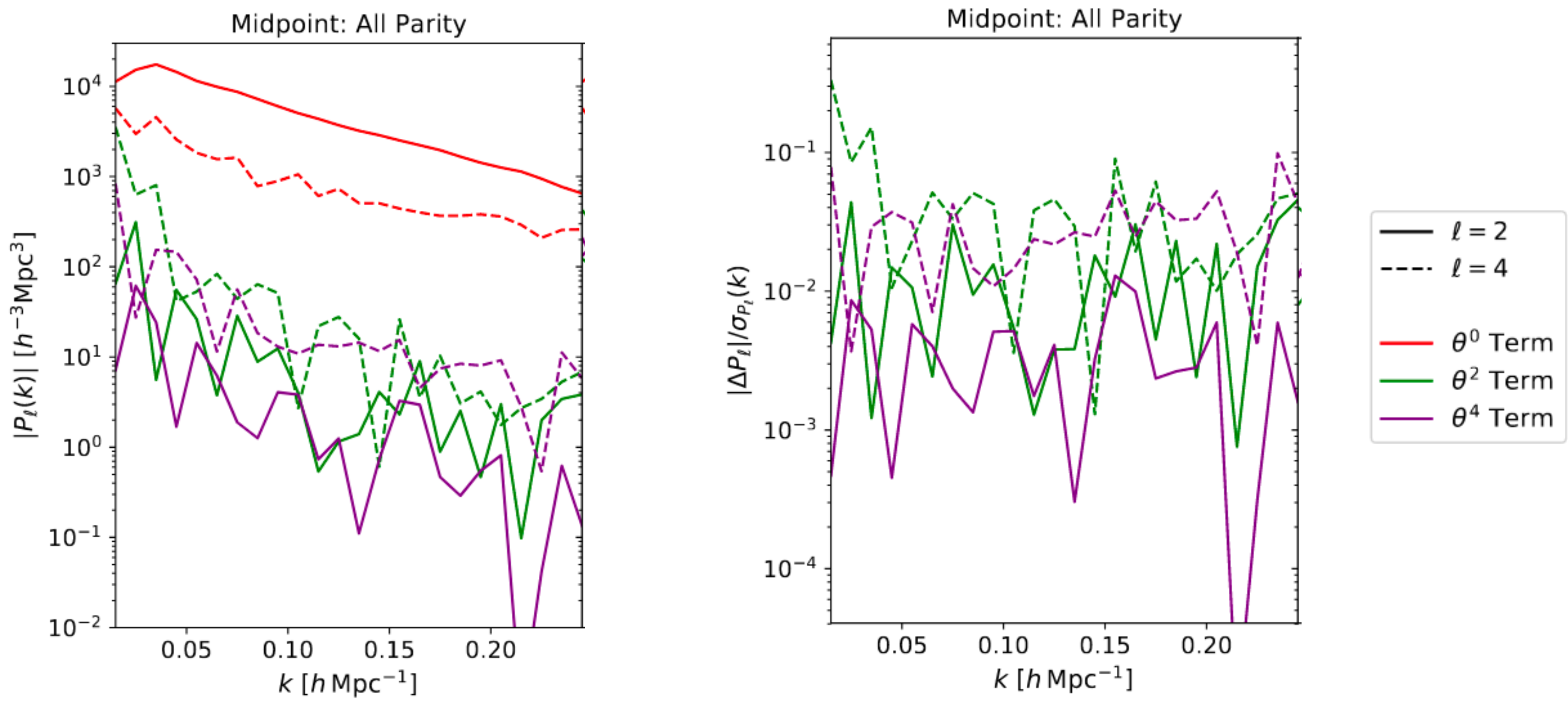
Shift Theorem Convergence



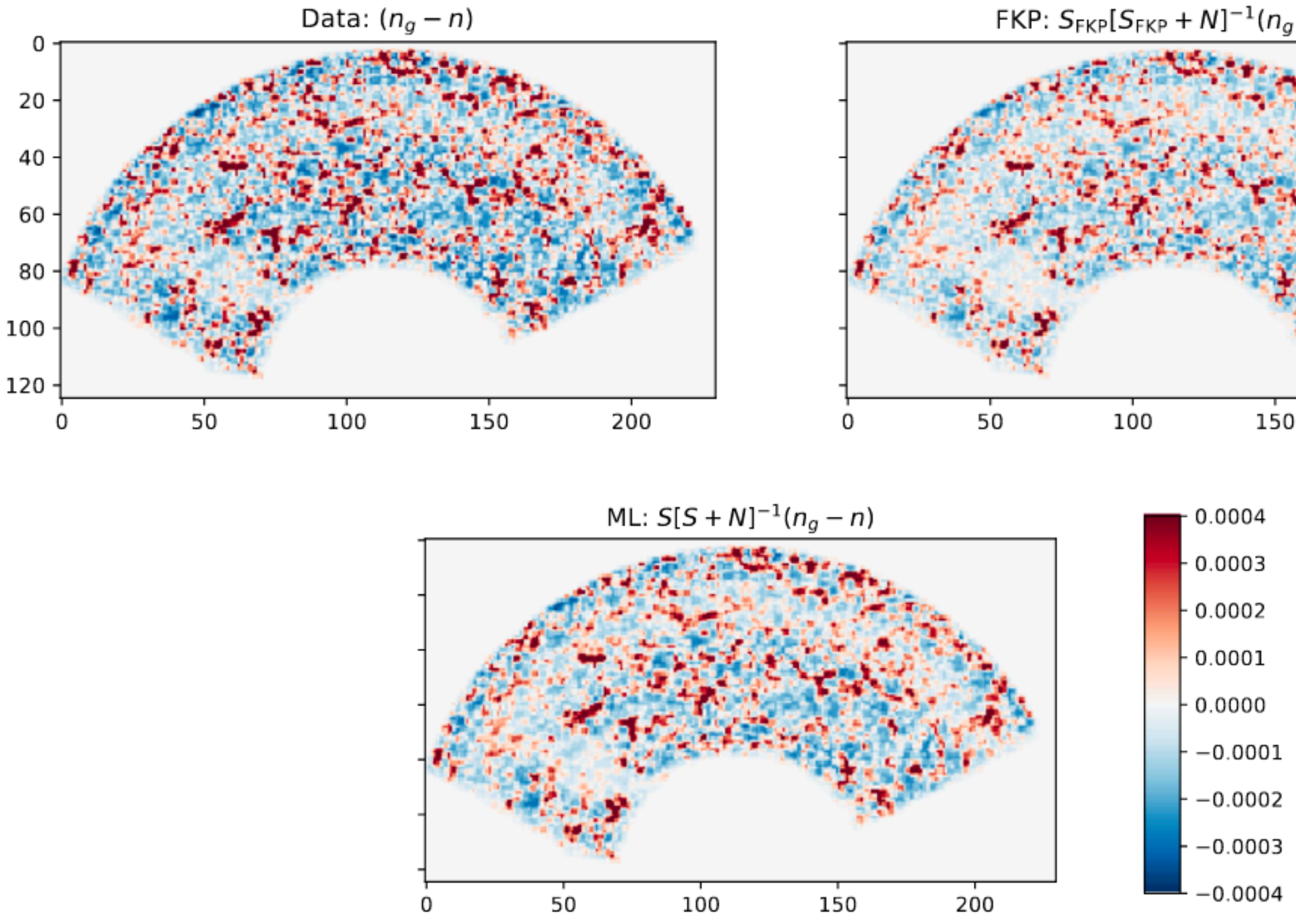
2PCF Wide-Angle Effects



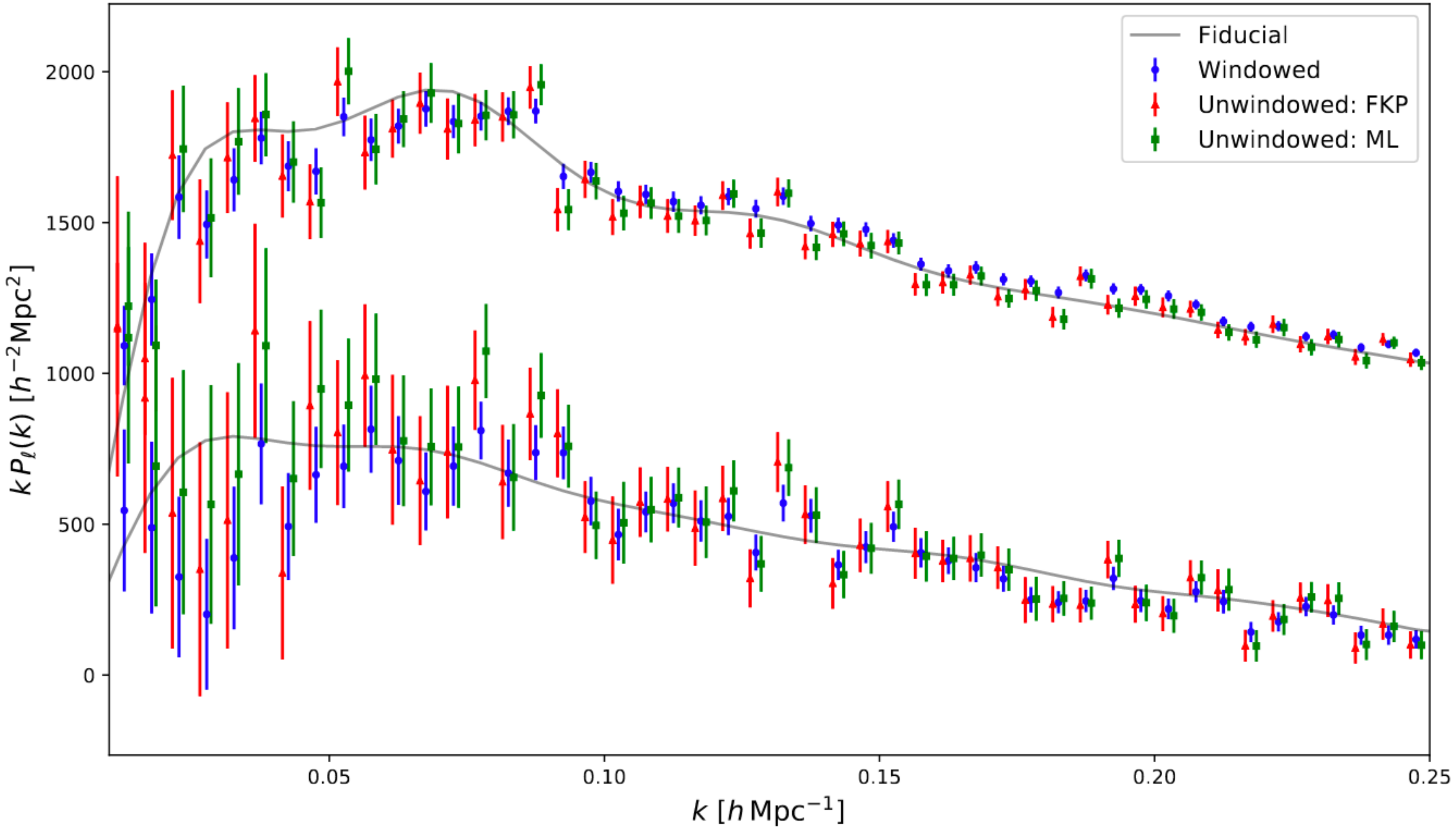
P(k) Wide-Angle Effects



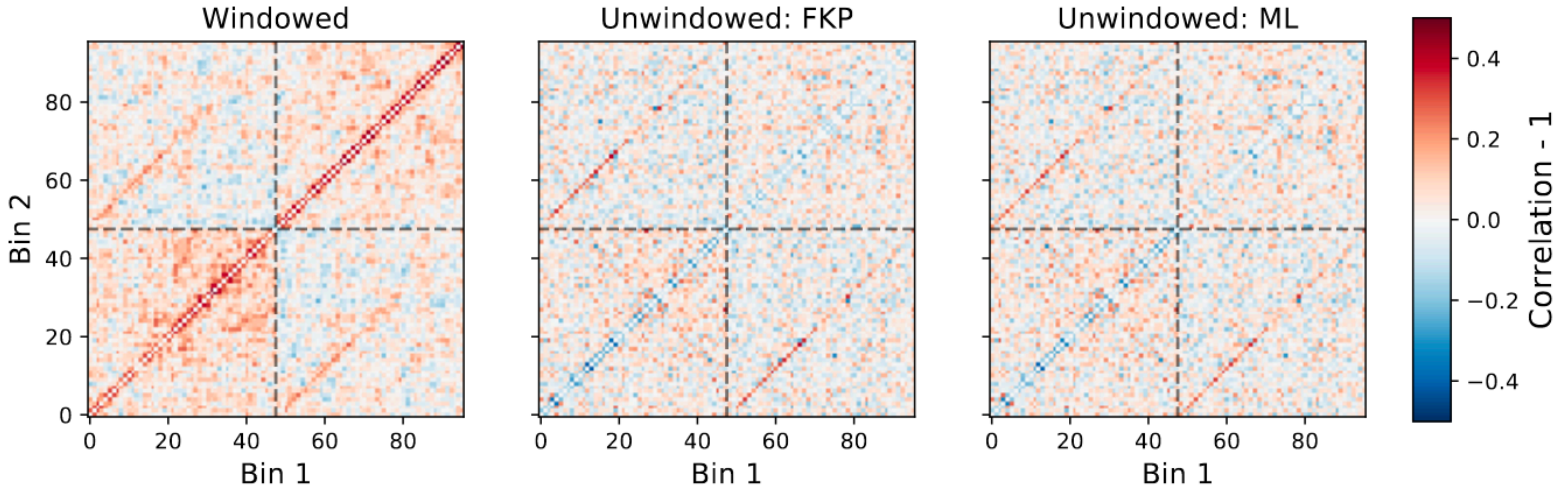
Optimal Estimators: Filtering



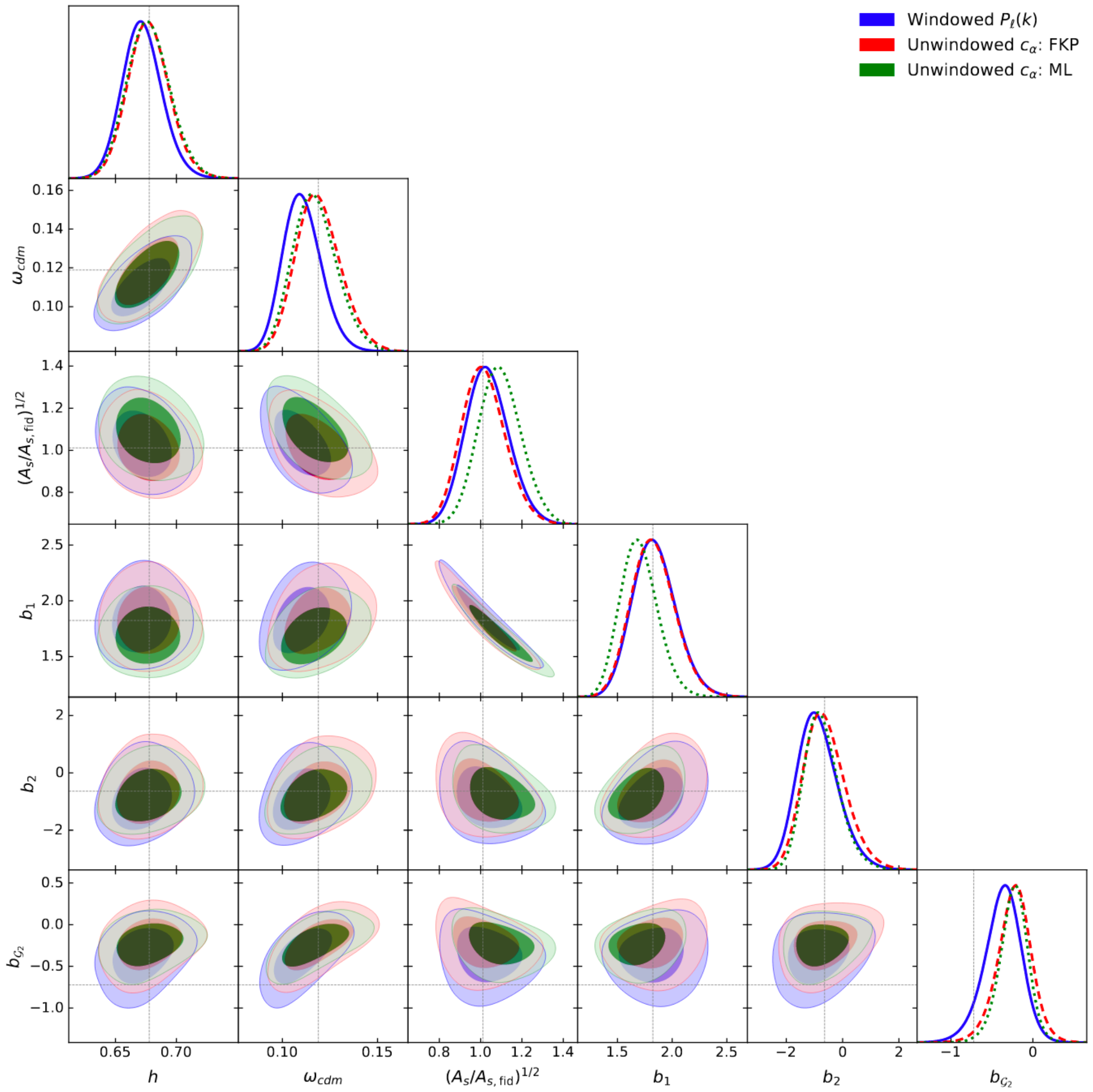
Optimal Estimators: Spectra



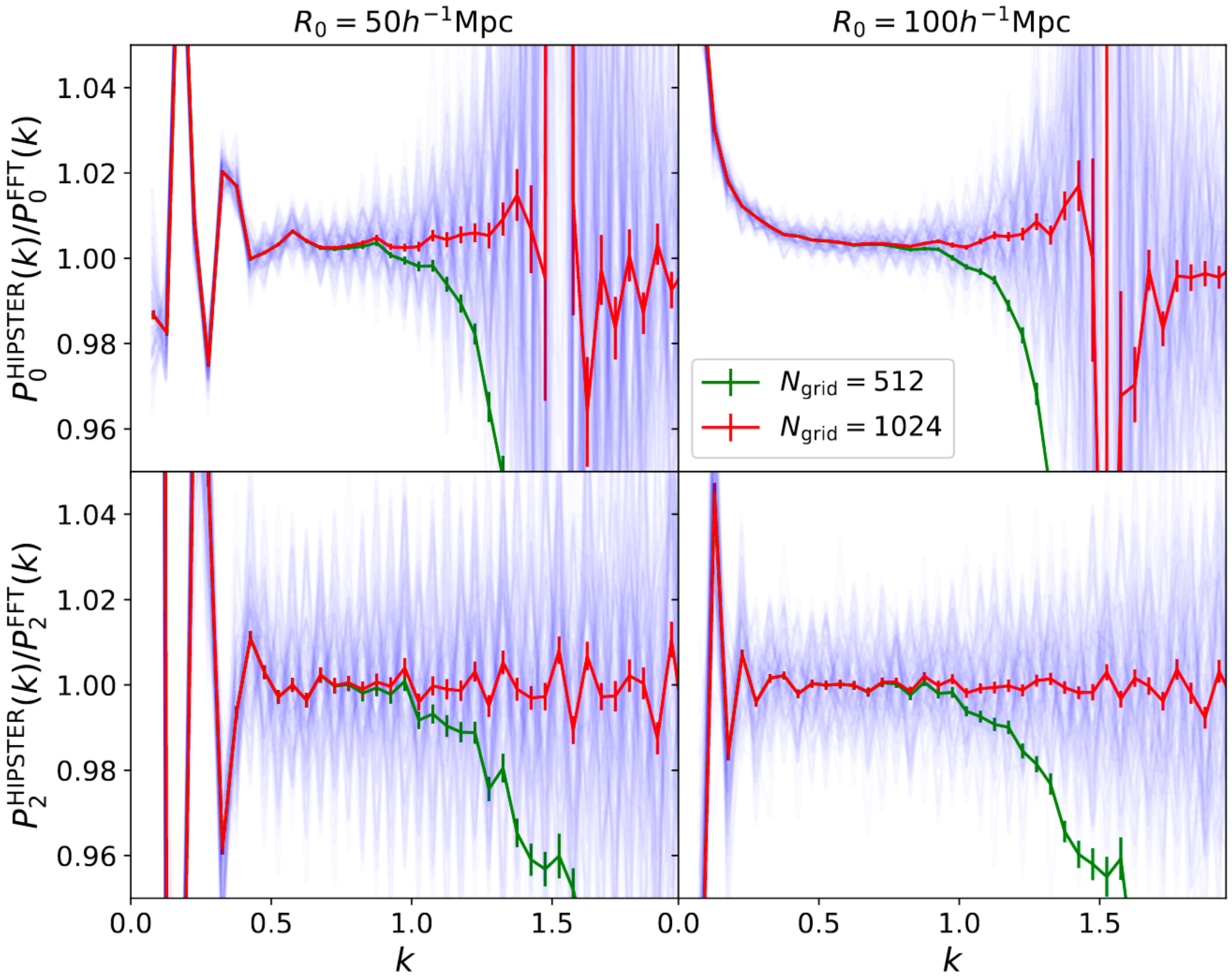
Optimal Estimators: Covariance



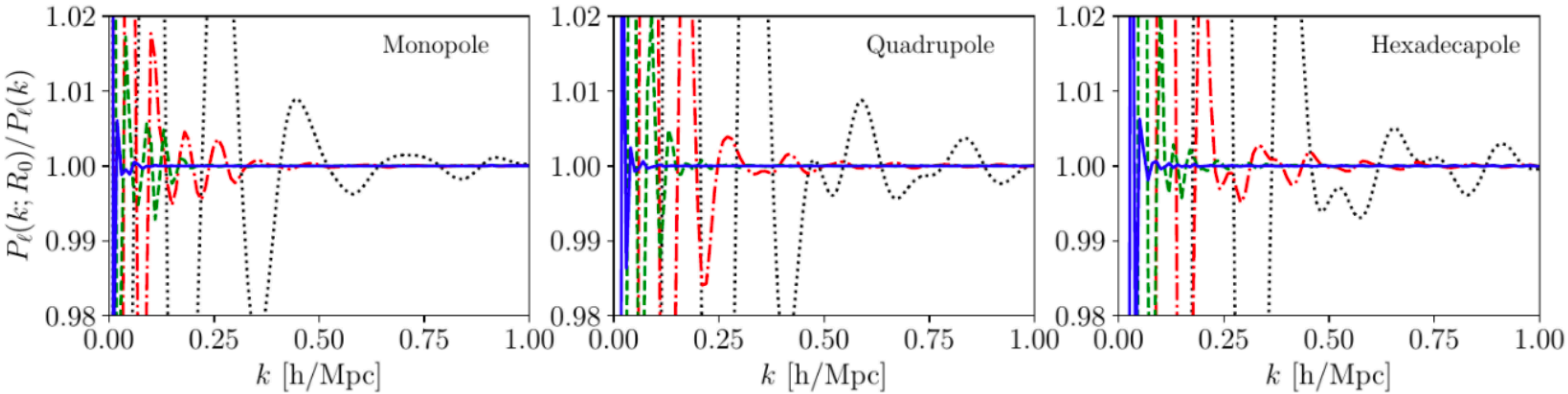
Optimal Estimators: Results



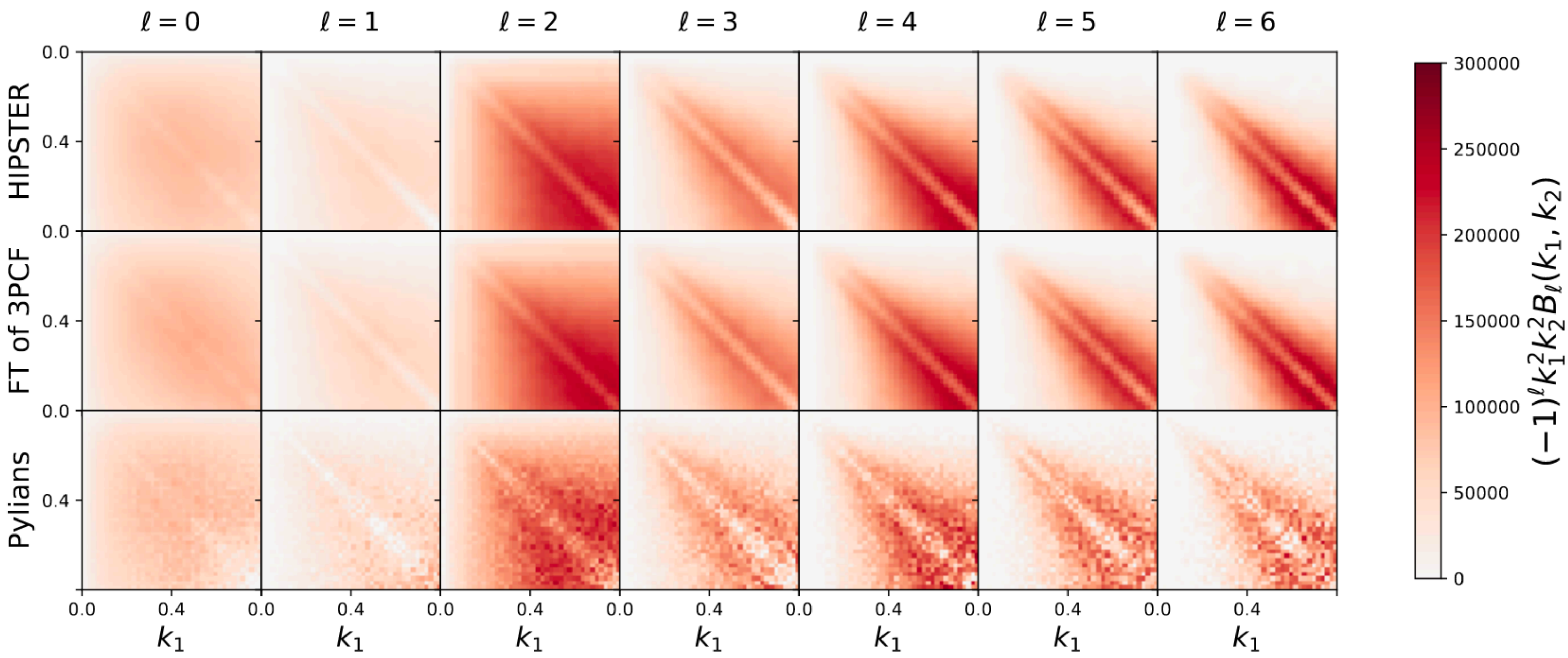
HIPSTER: Accuracy



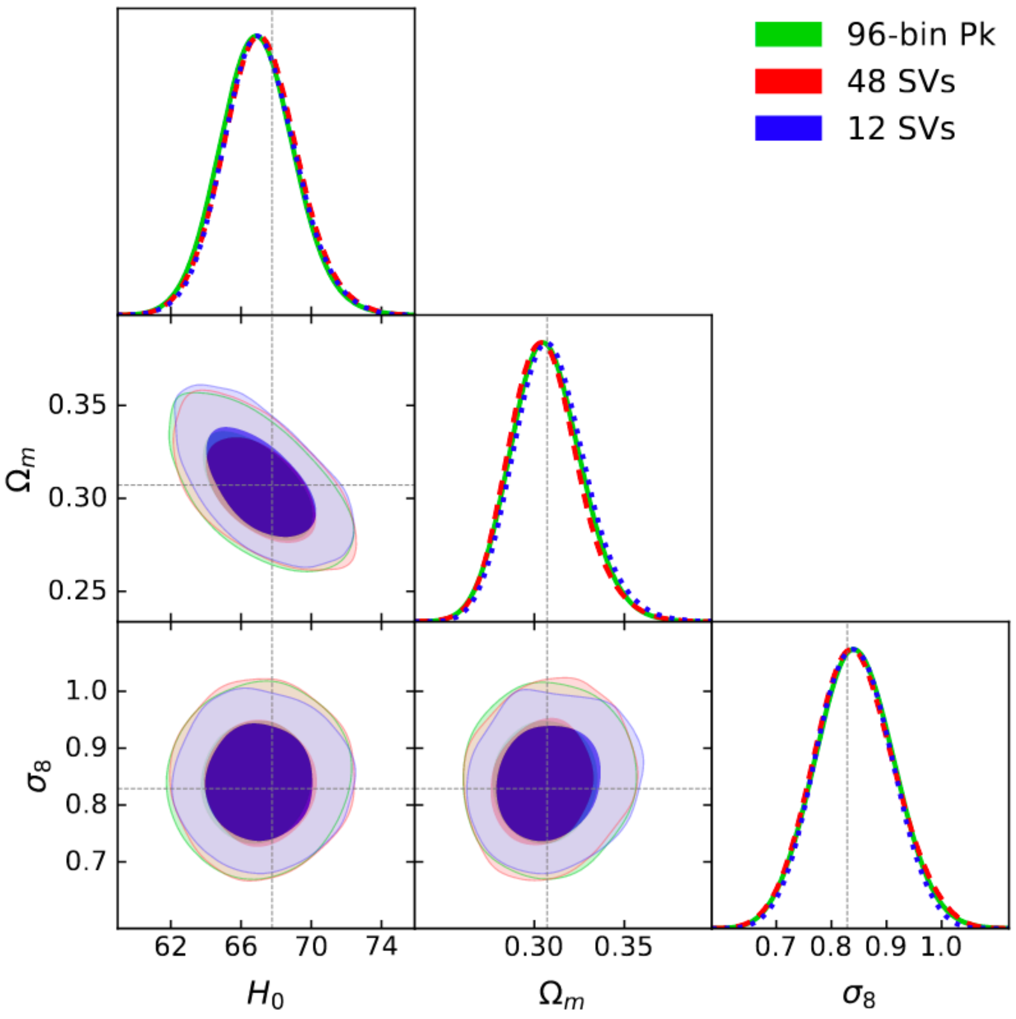
HIPSTER: Effects of Windowing

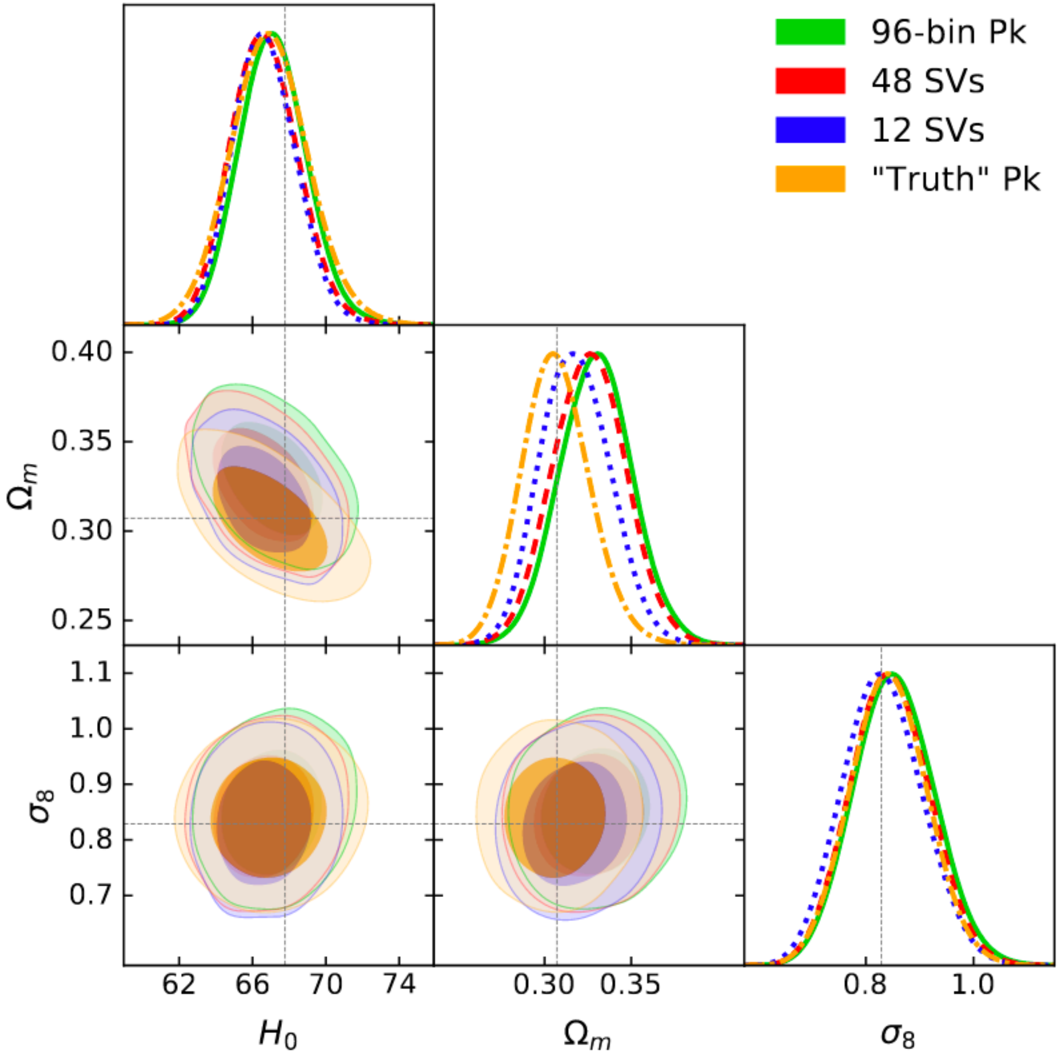


HIPSTER: Bispectra

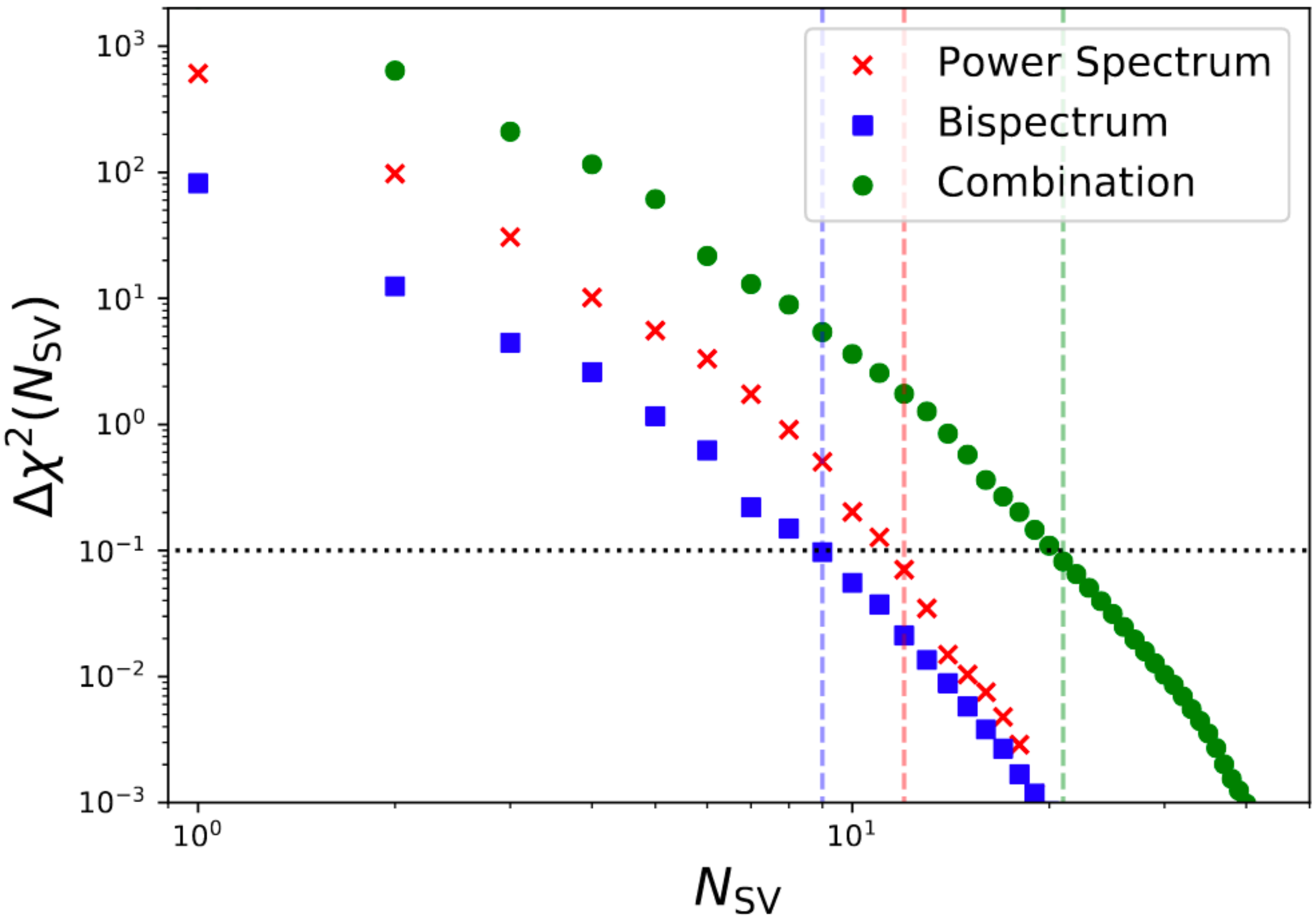


Compression: Mean of Mocks & Single Mock

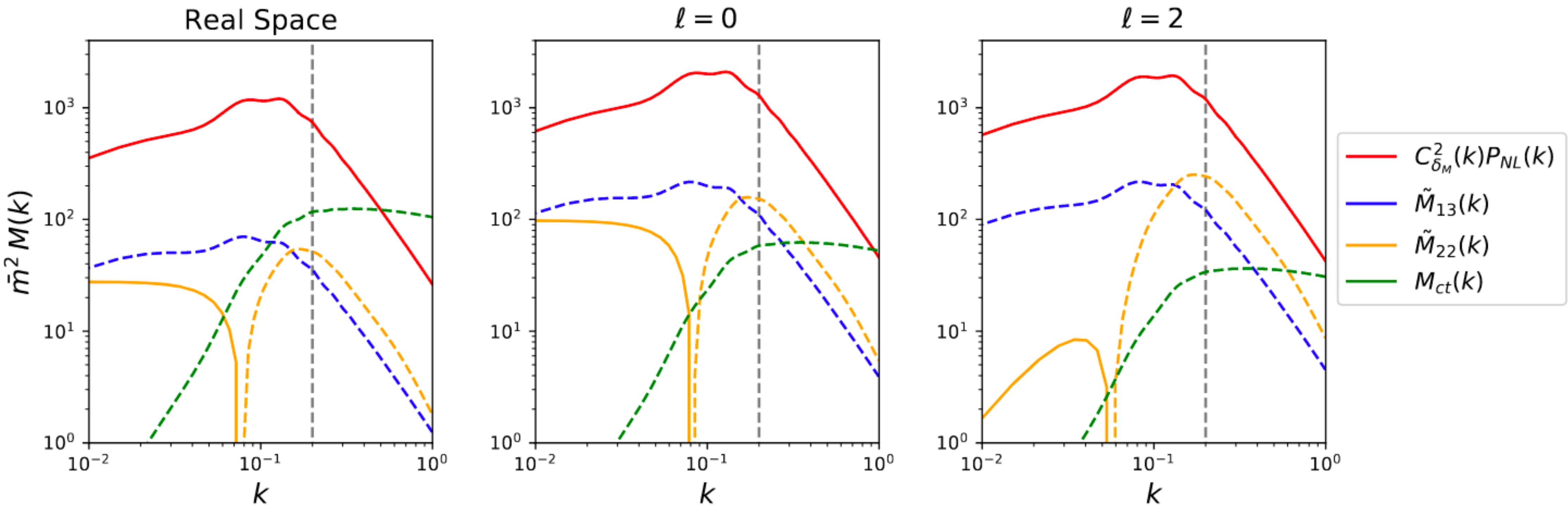




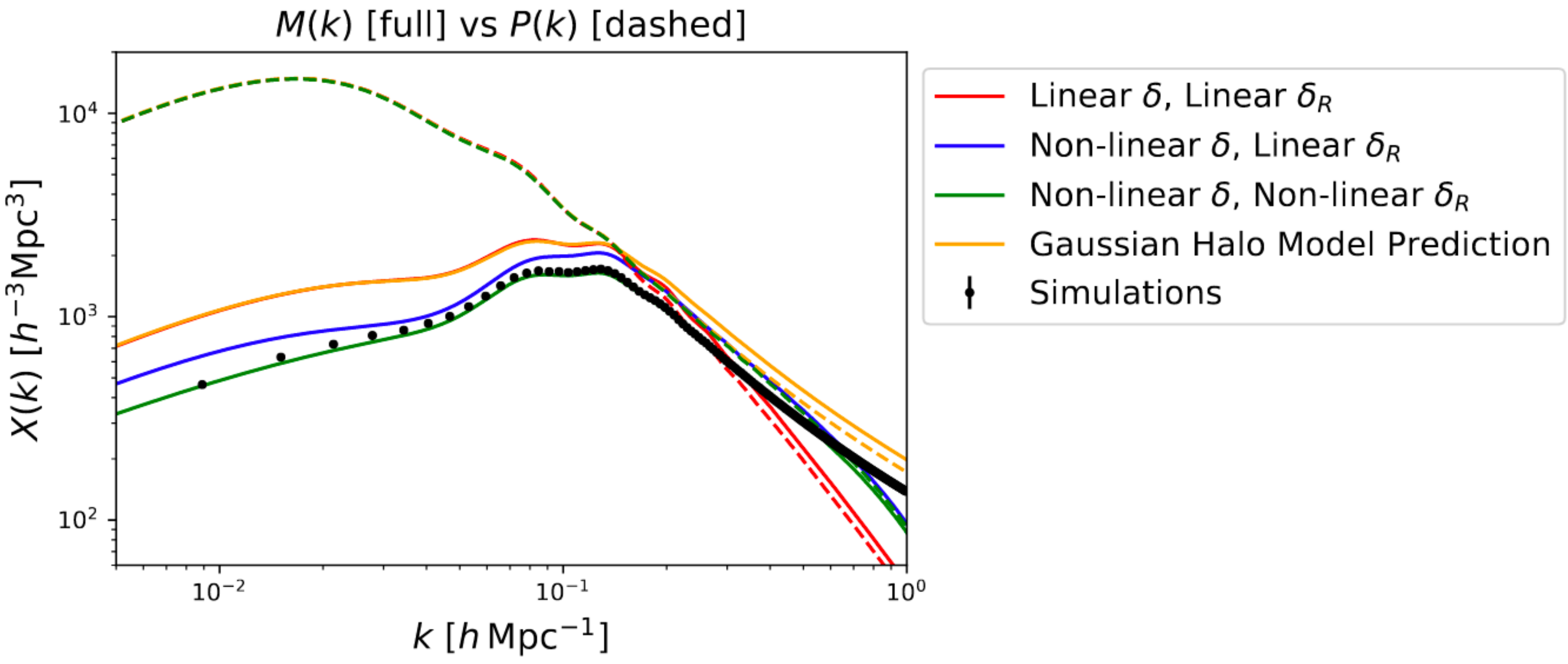
Compression: Number of Basis Vectors

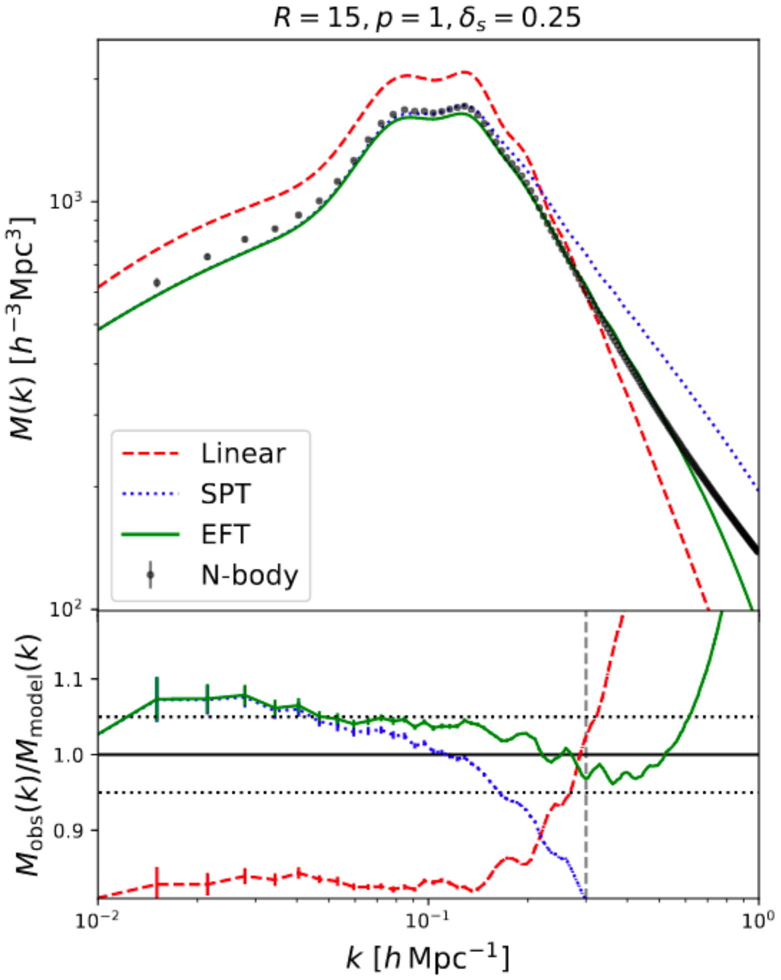


Marked Spectra: Matter Contributions



Marked Spectra: Information Content & Low-z





(a) $z = 0.5$