Title: Large Scale Structure Beyond the 2-Point Function

Speakers: Oliver Philcox

Series: Cosmology & Gravitation

Date: December 07, 2021 - 11:00 AM

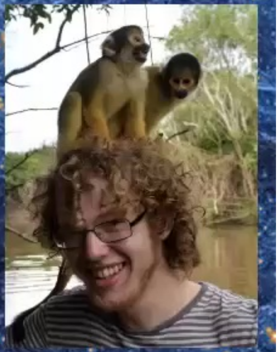
URL: https://pirsa.org/21120000

Abstract: Quantum fluctuations in inflation provide the seeds for the large scale distribution of matter today. According to the standard paradigm, these fluctuations induce density perturbations that are adiabatic and Gaussian distributed. In this limit, all the information is contained within the two-point correlation function, or equivalently, the power spectrum. Today, the distribution of matter is far from Gaussian, with structures forming across a vast range of scales. Despite this, almost all analyses of observational data are performed using two-point functions. This begs the question: what information lies in higher-point statistics?

In this seminar, I will present a pedagogical overview of the non-Gaussian correlation functions, and demonstrate how they can be used both to sharpen constraints on known physical parameters, and to provide stringent tests of new physics occurring in the early Universe. One of the major barriers to constraining cosmology from the higher-point functions is computational: measuring the statistics with conventional techniques is infeasible for current and future datasets. I will discuss new methods capable of reducing the computational cost by orders of magnitude, and show how this facilitates a number of exciting new tests of the cosmological model.

PRINCETON UNIVERSITY | IAS INSTITUTE FOR ADVANCED STUDY

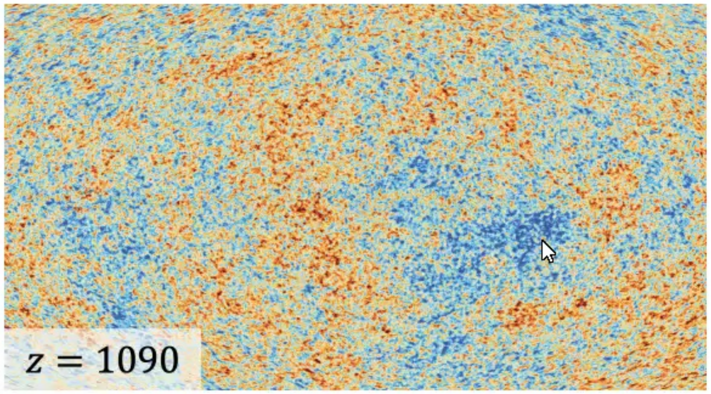
# Large Scale Structure Beyond the Two-Point Function



**Oliver Philcox (Princeton / IAS)**

Cosmology & Gravitation Seminar, Perimeter 12/06/21

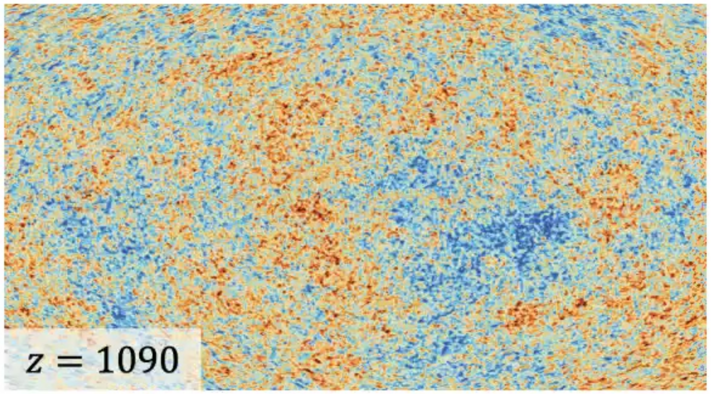
# THE EARLY UNIVERSE IS GAUSSIAN




2

Planck 2018, IllustrisTNG

# THE EARLY UNIVERSE IS GAUSSIAN




$$\delta(\mathbf{k}) \sim \mathcal{N}\left(0, P_L(\mathbf{k})\right)$$

- All information contained in the power spectrum
- **No** higher order statistics needed!



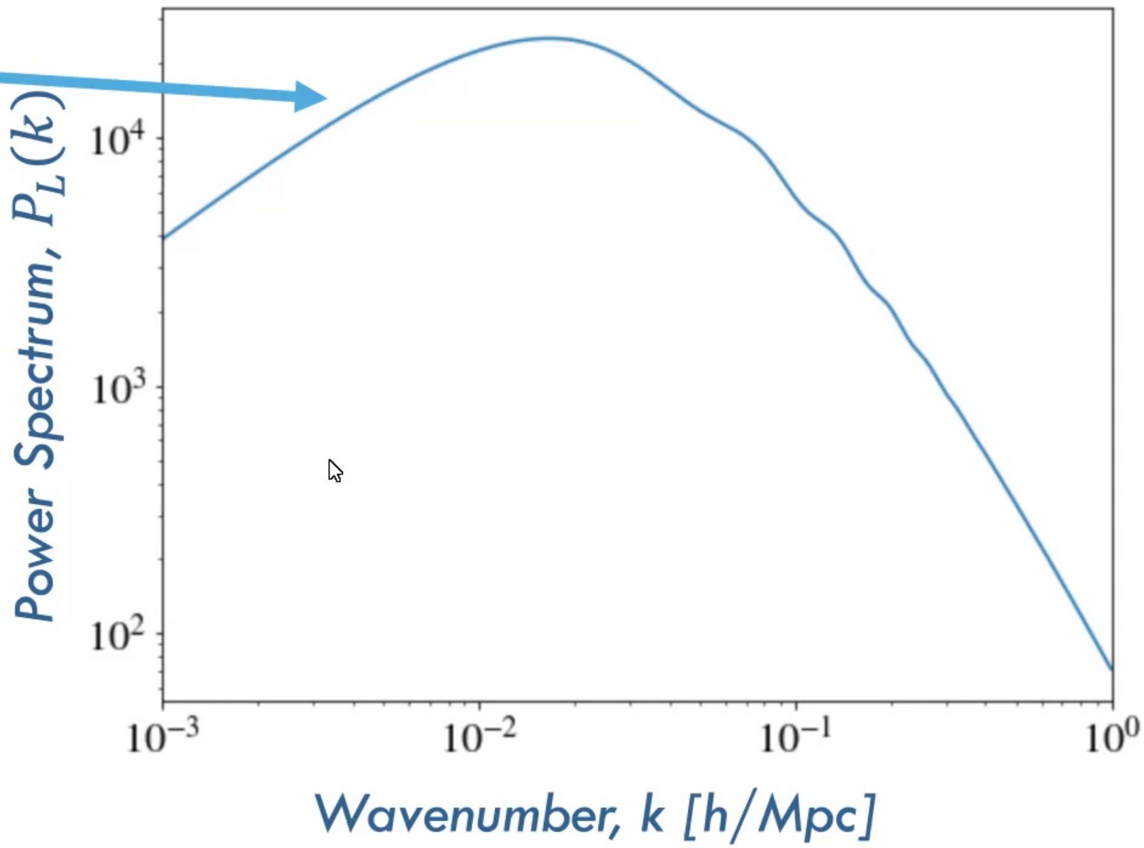
Planck 2018, IllustrisTNG

# LINEAR POWER SPECTRUM

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}')\rangle = (2\pi)^3 \delta_{\mathrm{D}}\left(\mathbf{k}+\mathbf{k}'\right) P_L(\mathbf{k})$$

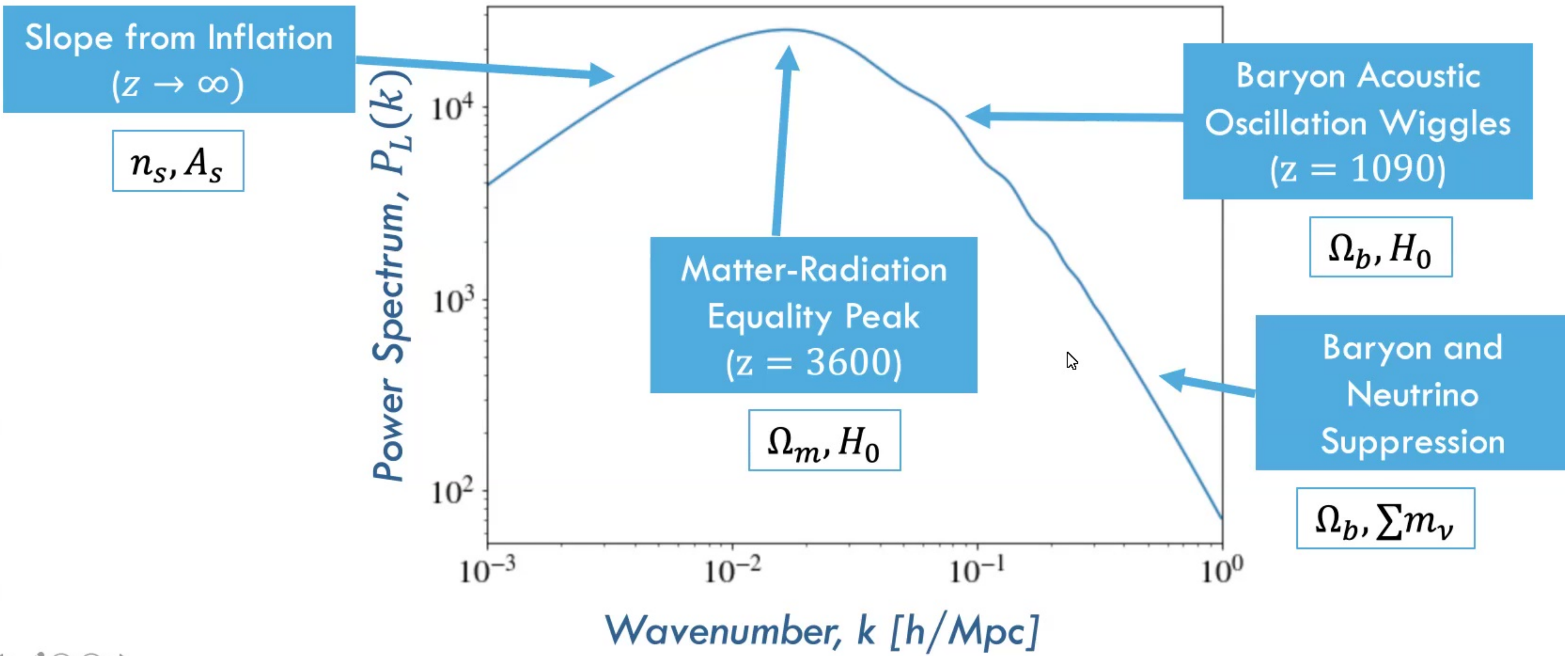
Slope from Inflation ($z \to \infty$)

$n_s, A_s$

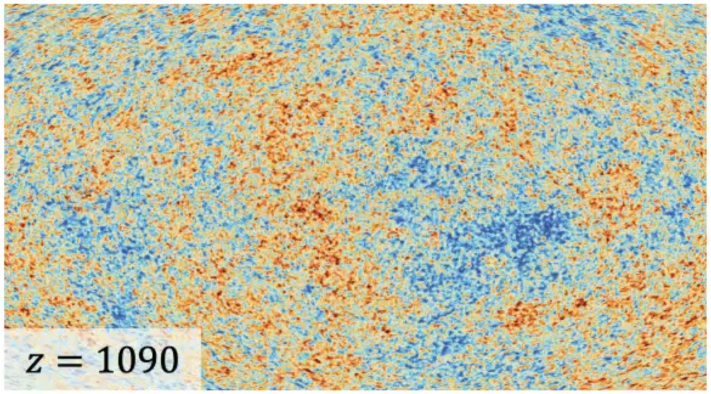


# LINEAR POWER SPECTRUM

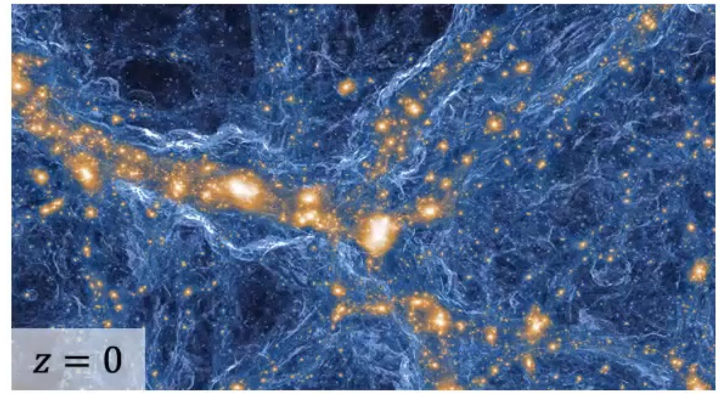
$$\langle \delta(\mathbf{k})\delta(\mathbf{k}')\rangle = (2\pi)^3 \delta_\mathrm{D}\left(\mathbf{k}+\mathbf{k}'\right) P_L(\mathbf{k})$$



# THE LATE UNIVERSE IS NOT GAUSSIAN




Gravitational Collapse




$$\delta(\mathbf{k}) \sim \mathcal{N}\left(0, P_L(\mathbf{k})\right)$$

- All information contained in the power spectrum
- **No** higher order statistics needed!

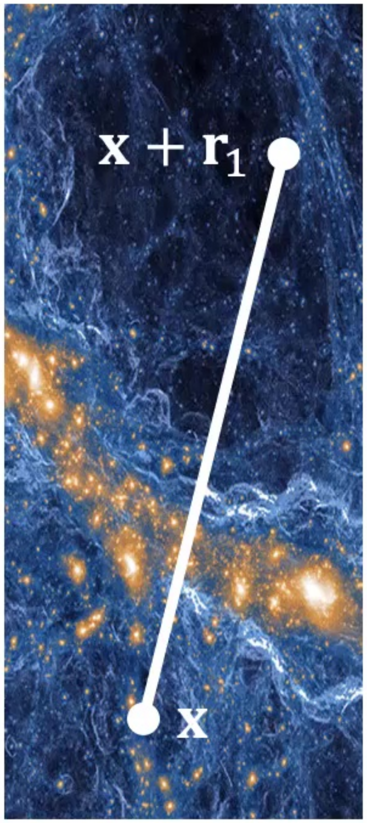
$$\delta(\mathbf{k}) \not\sim \mathcal{N}\left(0, P_L(\mathbf{k})\right)$$

- **Not** all information contained in the power spectrum
- Higher-order statistics needed!



Planck 2018, IllustrisTNG

# NON-GAUSSIAN DENSITY ⇒ NON-GAUSSIAN STATISTICS




## Gaussian

1. **Power Spectrum:**

$$P(\mathbf{k}_1) = \langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\rangle'$$

2. **2-Point Correlation Function:**

$$\xi(\mathbf{r}_1) \equiv \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}_1)\rangle$$

## Non-Gaussian

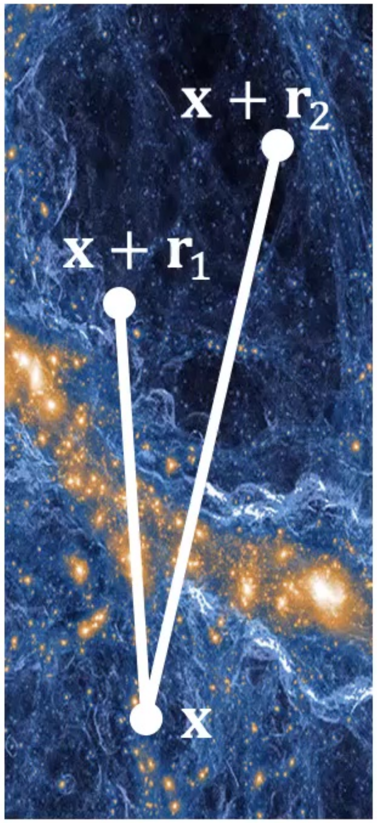
1. **Bispectrum:**

$$B(\mathbf{k}_1, \mathbf{k}_2) = \langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle'$$

2. **3-Point Correlation Function:**

$$\zeta_3(\mathbf{r}_1, \mathbf{r}_2) \equiv \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}_1)\delta(\mathbf{x}+\mathbf{r}_2)\rangle$$

*And beyond…*



# WHAT MAKES UP THE BISPECTRUM?

$$B_g(\mathbf{k}_1, \mathbf{k}_2) = \left[2b_1^3 F_2(\mathbf{k}_1, \mathbf{k}_2) + b_2 b_1^2 + 2b_{s^2} b_1^2 \left(\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2 - 1/3\right)\right] P_L(k_1) P_L(k_2) + 2 \text{ perms.}$$

The galaxy bispectrum depends on galaxy formation physics, gravity, and early-Universe cosmology.

▷ To obtain **all** the large-scale information in the initial conditions, we need:*

- Power Spectra / 2-Point Functions $\sim P_L(k)$
- Bispectra / 3-Point Functions $\sim P_L^2(k)$
- Trispectra / 4-Point Functions $\sim P_L^3(k)$
- *etc.*

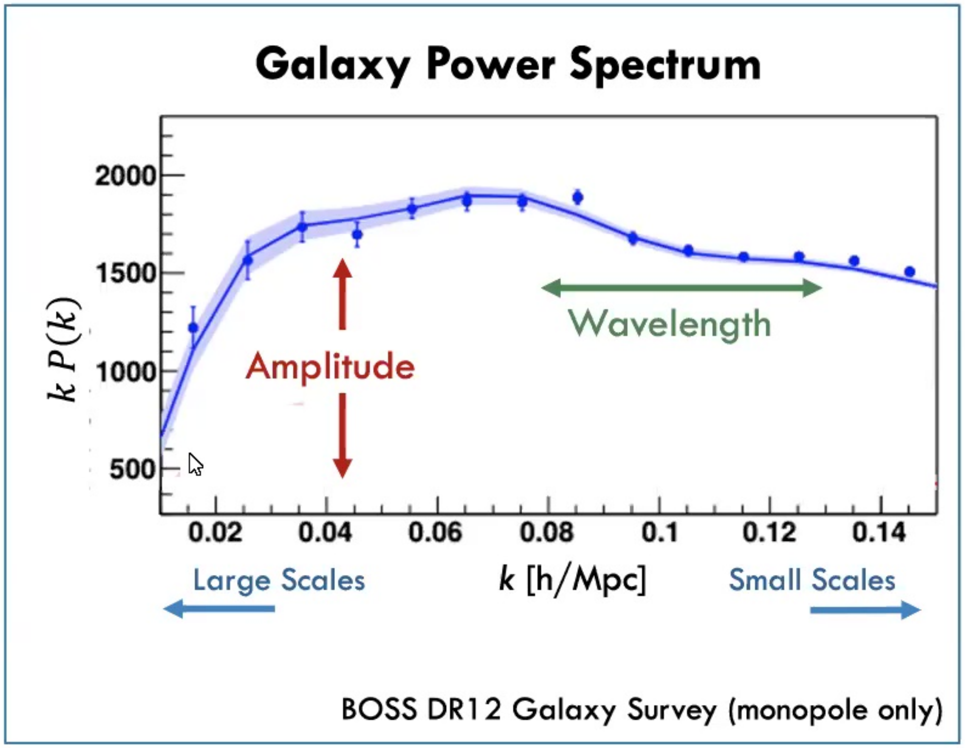
16

*ignoring higher-order perturbative effects, redshift-space distortions, renormalization, *etc.* e.g. Ivanov+21

# THE CURRENT STATE OF PLAY

- Analyze the galaxy **power spectrum** using a **scaling analysis**

- This measures:
  - Overall **amplitude** ( = primordial amplitude)
  - **Wiggle** positions ( = BAO feature)

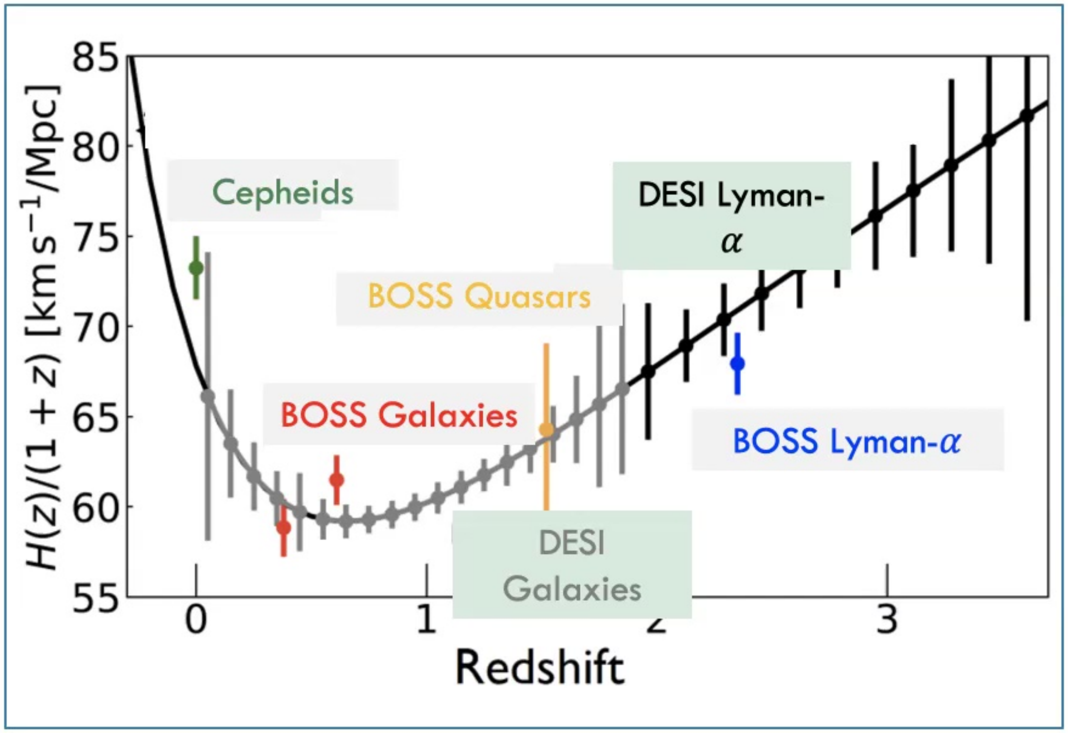





e.g. Beutler+17, Gil-Marin+15,17, Alam+20

# THE CURRENT STATE OF PLAY

- Analyze the galaxy **power spectrum** using a **scaling analysis**
- Measure **wiggle positions** ( = BAO feature) and **overall amplitude**
- Robust way to constrain **growth rate** and **expansion history** $H(z)$



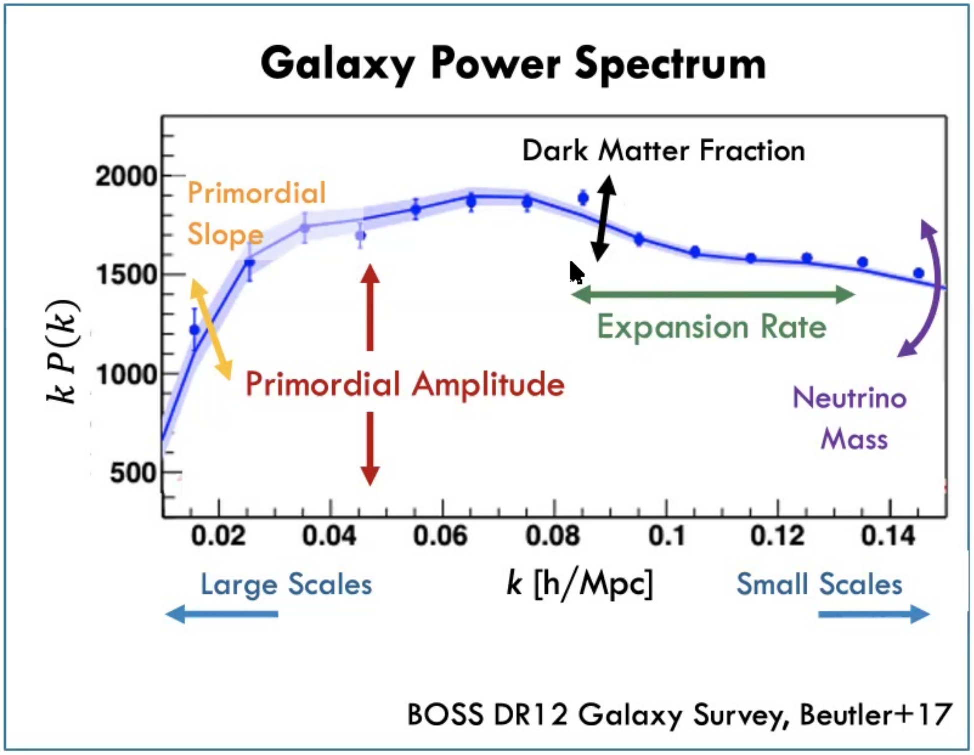



DESI / Gil-Marin

# THE CURRENT STATE OF PLAY

- This is *not* all the available information!

- Measure parameters **directly** from the **full shape** of the galaxy power spectrum

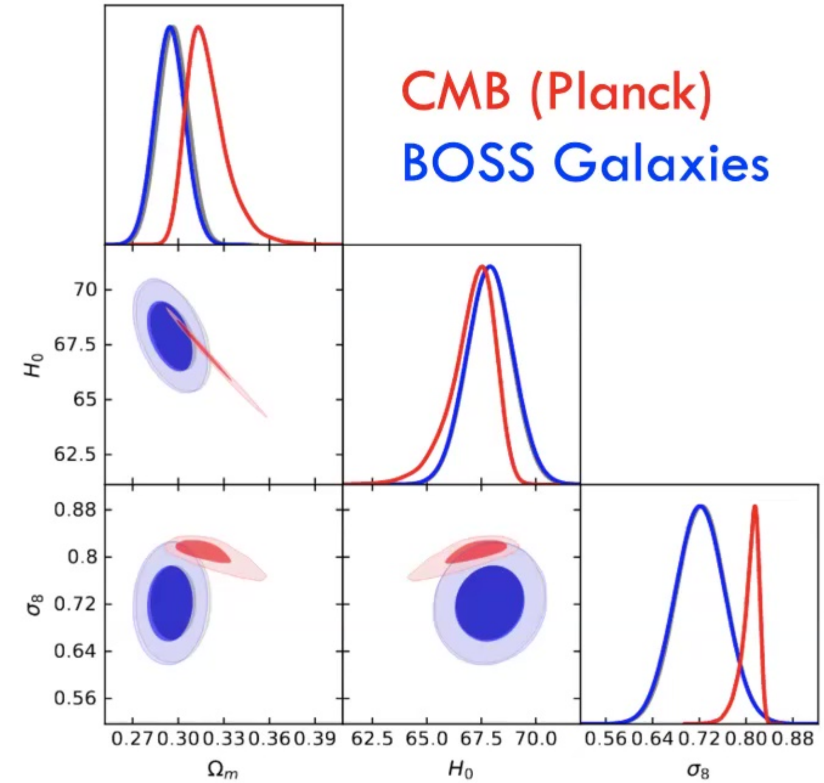





e.g. *Ivanov+19,20, d'Amico+19,20,* ***Philcox****+20ab, Chen+21, Kobayashi+21*

# THE CURRENT STATE OF PLAY

- This is *not* all the available information!
- Measure parameters **directly** from the **full shape** of the galaxy power spectrum
- Constrain parameters in **new** ways e.g. expansion rate from **equality** scale.
  [Farren, **Philcox** & Sherwin (in prep.)]



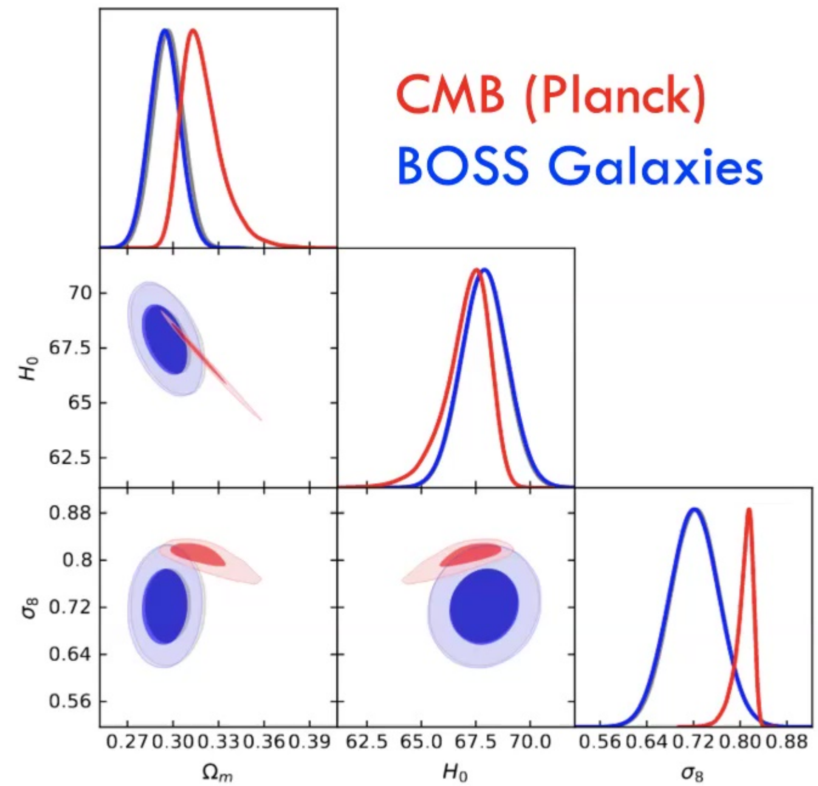

25

e.g. *Ivanov+19,20, d'Amico+19,20,* ***Philcox****+20ab, Chen+21, Kobayashi+21*

# THE CURRENT STATE OF PLAY

- This is *not* all the available information!
- Measure parameters **directly** from the **full shape** of the galaxy power spectrum
- Constrain parameters in **new** ways e.g. expansion rate from **equality** scale.
  [Farren, **Philcox** & Sherwin (in prep.)]

*Can we go **beyond** the power spectrum?*




25

e.g. *Ivanov+19,20, d'Amico+19,20,* ***Philcox****+20ab, Chen+21, Kobayashi+21*

# WHY USE HIGHER-ORDER STATISTICS?

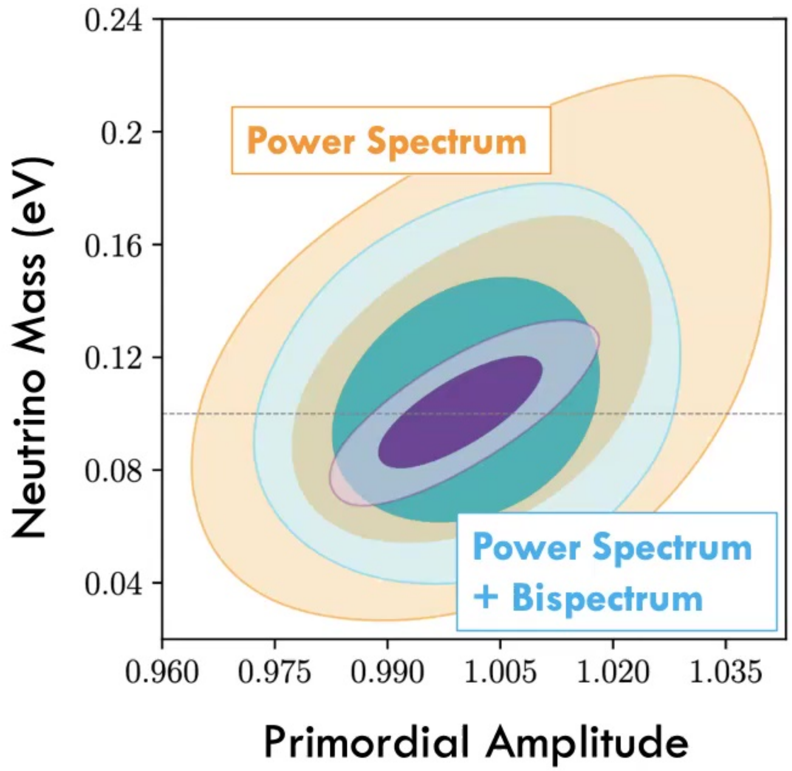
- **Sharpen** parameter constraints!

- **Break** parameter **degeneracies**!

  [e.g. $P_g \sim b_1^2\sigma_8^2,\ B_g \sim b_1^3\sigma_8^4$]

<u>Euclid Forecast</u>

- Bispectrum improves constraints by $\approx 40\%$
- $1\sigma$ constraint of $\sigma_{M_\nu} = 0.013$ eV [including *Planck*]




*Euclid Forecast*, Chudaykin+19

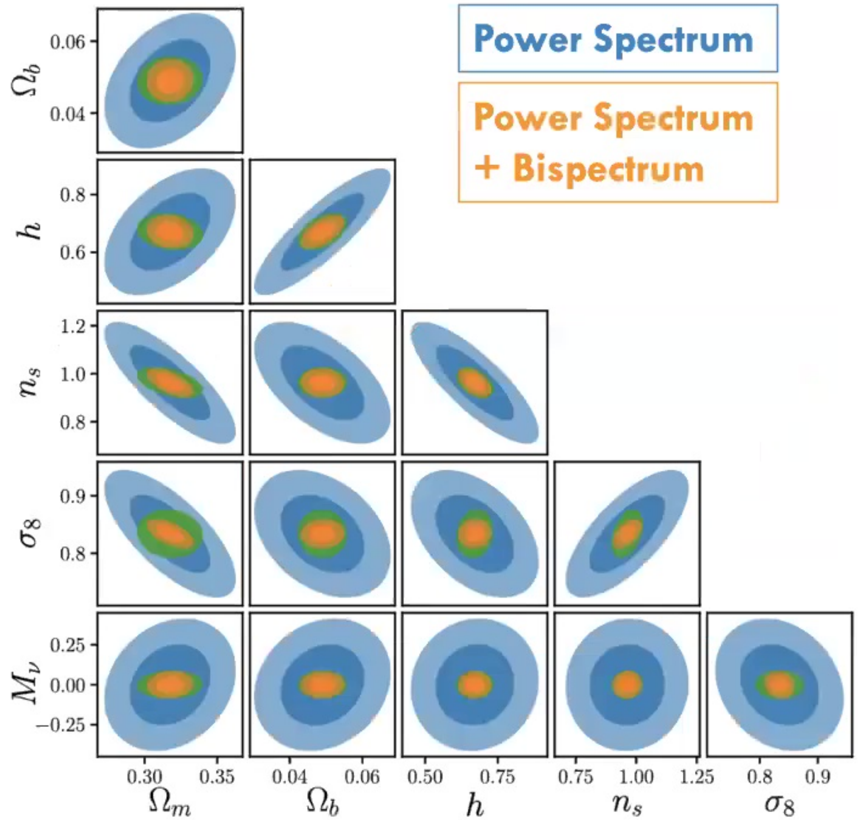
# NON-GAUSSIAN INFLATION

*Are the primordial perturbations **Gaussian** and **adiabatic**?*

# WHY USE HIGHER-ORDER STATISTICS?

- **Sharpen** parameter constraints!
- **Break** parameter **degeneracies**!

[e.g. $P_g \sim b_1^2\sigma_8^2,\ B_g \sim b_1^3\sigma_8^4$]




$1\ (\mathrm{Gpc}/h)^3$ *volume*, Hahn+20

<u>Simulation-Based Forecast</u>

- Galaxy Bispectrum improves constraints by $> 2\times$
- Neutrino constraint improves by $5\times$

27

# NON-GAUSSIAN INFLATION

*Are the primordial perturbations* ***Gaussian*** *and* ***adiabatic****?*

**Standard Model of Inflation:**

- Scalar field $\phi$ rolling down a potential $V(\phi)$

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}R - \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi)\right]$$

Gravity — Kinetic Energy — Potential

- Action, $S$, encodes **statistics** of the **primordial** curvature perturbations, $\zeta$

**Second Order ⇒ Power Spectrum**

$$S^{(2)} \Rightarrow P_\zeta(k) \approx A_s k^{n_s-4}$$

**Third Order ⇒ Bispectrum**

$$S^{(3)} \Rightarrow B_\zeta(\mathbf{k}_1, \mathbf{k}_2) \approx \frac{6}{5} f_{\rm NL} P_\zeta(k_1) P_\zeta(k_2) + 2 \text{ perms.}$$

Generates **non-Gaussianity** proportional to $f_{\rm NL}$

# NON-GAUSSIAN INFLATION

$$B_\zeta(\mathbf{k}_1, \mathbf{k}_2) \approx \frac{6}{5} f_{\rm NL} P_\zeta(k_1) P_\zeta(k_2) + 2 \text{ perms.}$$

The **consistency condition** states that

$$\lim_{k_1 \to 0} B_\zeta(\mathbf{k}_1, \mathbf{k}_2) = (1 - n_s) P_\zeta(k_1) P_\zeta(k_2)$$

$$f_{\rm NL} \sim (1 - n_s) \ll 1$$

*Non-Gaussianity is too small to be detected!*

**Non-standard** inflation can beat this, e.g.

- ▷ Multifield Inflation [Local Bispectrum]
- ▷ New Kinetic Terms [Equilateral Bispectrum]

$$S = \int d^4x \sqrt{-g} \left[ \frac{1}{2} R - \frac{1}{2} \partial^\mu \phi \partial_\mu \phi - V(\phi) \right]$$

31

Maldacena 03, Creminelli+04

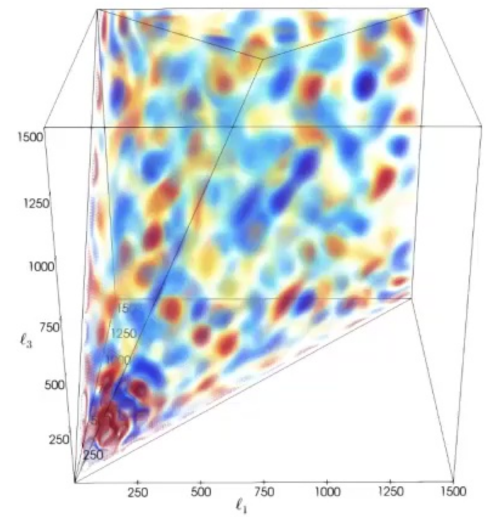
# NON-GAUSSIAN INFLATION

$$B_\zeta(\mathbf{k}_1, \mathbf{k}_2) \approx \frac{6}{5} f_{\rm NL} P_\zeta(k_1) P_\zeta(k_2) + 2 \text{ perms.}$$

**How do we measure this?**

1. **CMB Bispectrum**

*Planck* TTT Bispectrum




≈ 2× better with CMB-S4!

$f_{\rm NL}$ Constraints

| | |
|---|---|
| Local . . . . . . . . . . . . | 6.7 ± 5.6 |
| Equilateral . . . . . . . . . | 6 ± 66 |
| Orthogonal . . . . . . . . | −38 ± 36 |

Planck 2018 IX

# NON-GAUSSIAN INFLATION

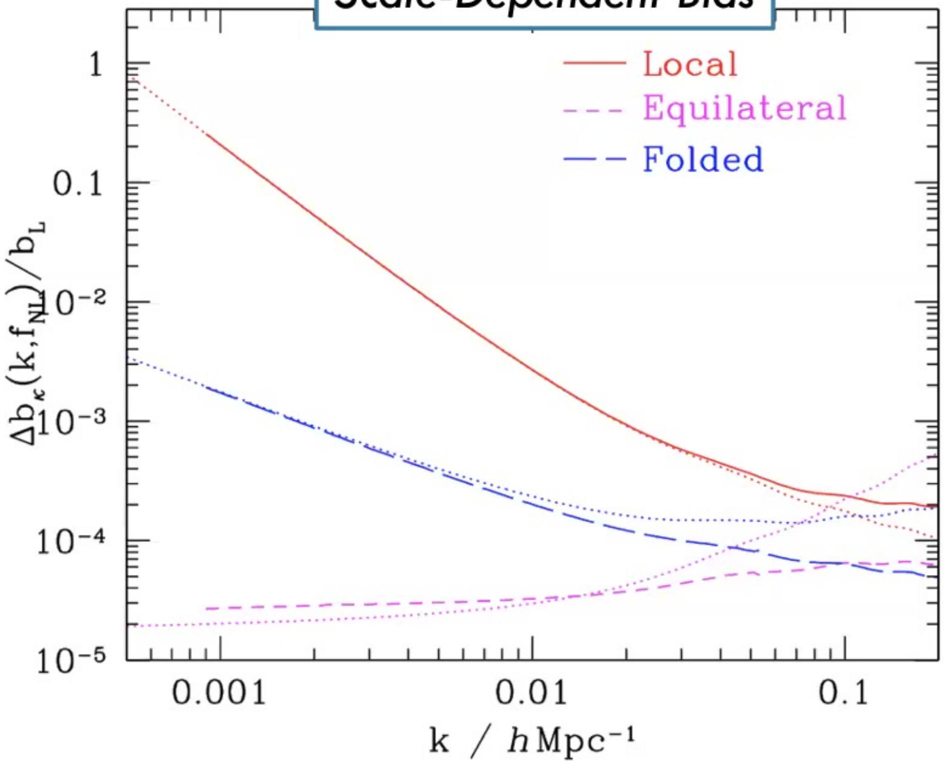
$$B_\zeta(\mathbf{k}_1, \mathbf{k}_2) \approx \frac{6}{5} f_{\rm NL} P_\zeta(k_1) P_\zeta(k_2) + 2 \text{ perms.}$$

**How do we measure this?**

1. CMB Bispectrum

2. **Galaxy Power Spectrum**

Scale-Dependent Bias




33

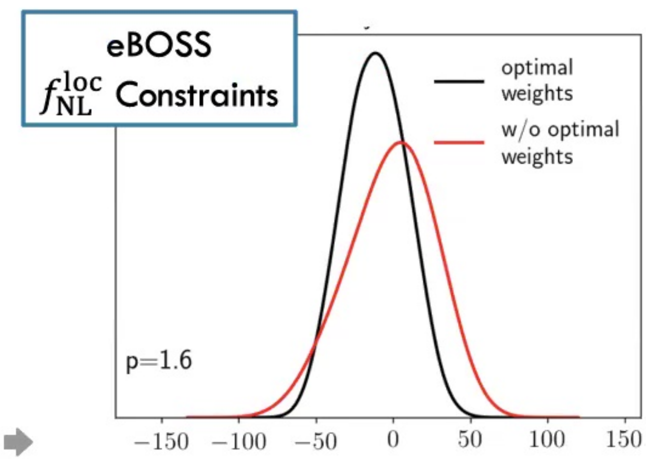
Desjacques & Seljak 10, eBOSS 21

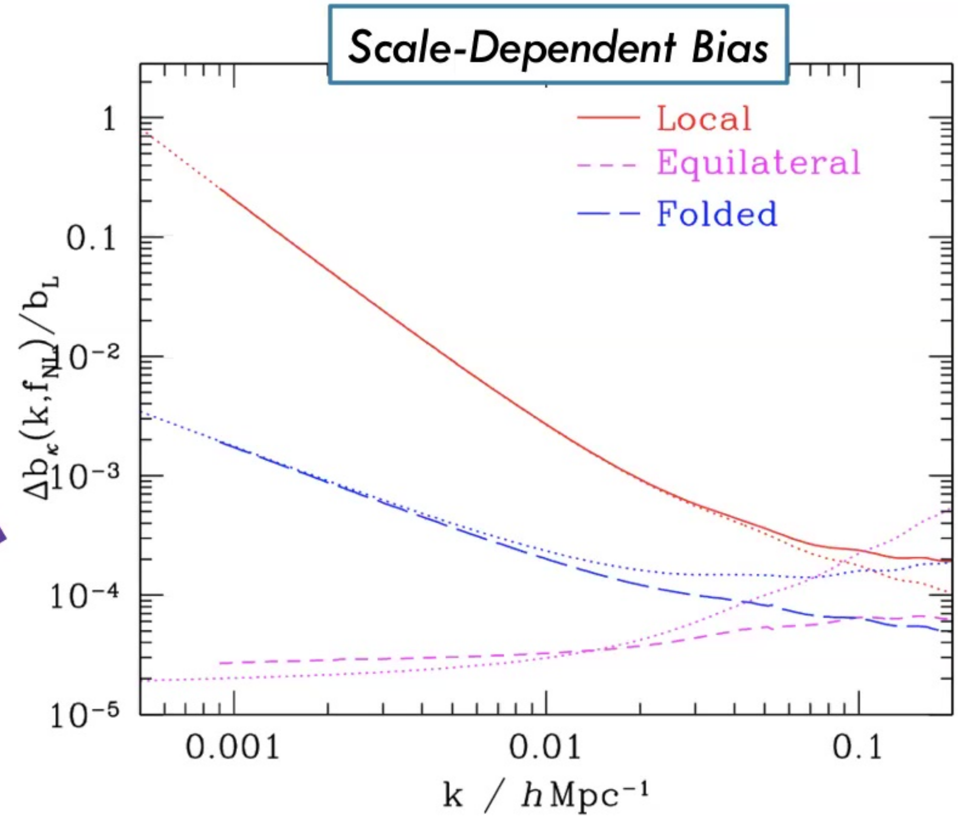
# NON-GAUSSIAN INFLATION

$$B_\zeta(\mathbf{k}_1, \mathbf{k}_2) \approx \frac{6}{5} f_{\rm NL} P_\zeta(k_1) P_\zeta(k_2) + 2 \text{ perms.}$$

**How do we measure this?**

1. CMB Bispectrum
2. **Galaxy Power Spectrum**




33

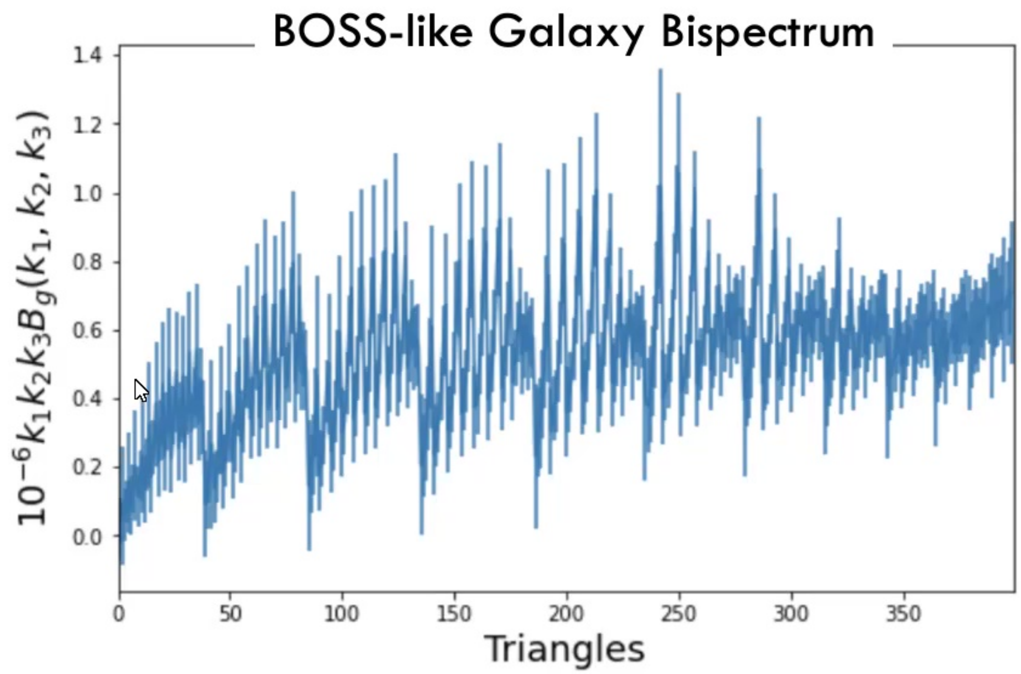
Desjacques & Seljak 10, eBOSS 21

# NON-GAUSSIAN INFLATION

$$B_\zeta(\mathbf{k}_1, \mathbf{k}_2) \approx \frac{6}{5} f_{\rm NL} P_\zeta(k_1) P_\zeta(k_2) + 2 \text{ perms.}$$

**How do we measure this?**

1. CMB Bispectrum
2. Galaxy Power Spectrum
3. **Galaxy Bispectrum**






Ivanov, **Philcox**+21, **Philcox** & Ivanov (in prep.)

# CHERN-SIMONS INTERACTIONS VIOLATE PARITY

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}R - \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi) - \frac{1}{4}f(\phi)F_{\mu\nu}F^{\mu\nu} + \frac{\gamma}{4}f(\phi)F_{\mu\nu}\tilde{F}^{\mu\nu}\right]$$

- Add a **gauge field** $A_\mu$ to the inflationary action, via $F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu$
- This can include a **Chern-Simons coupling** to the (pseudo-)scalar $\phi$ [motivated by baryogenesis]
- $f(\phi)F_{\mu\nu}\tilde{F}^{\mu\nu}$ violates **parity symmetry** ⇒ parity-violating correlators!

*Where should we look for these signatures?*

35

e.g. Lue+98, Jeong+12, Alexander+11, Shiraishi 16

# THE 2PCF + 3PCF ARE NOT SENSITIVE TO PARITY

**2-Point Correlation Function (2PCF):**

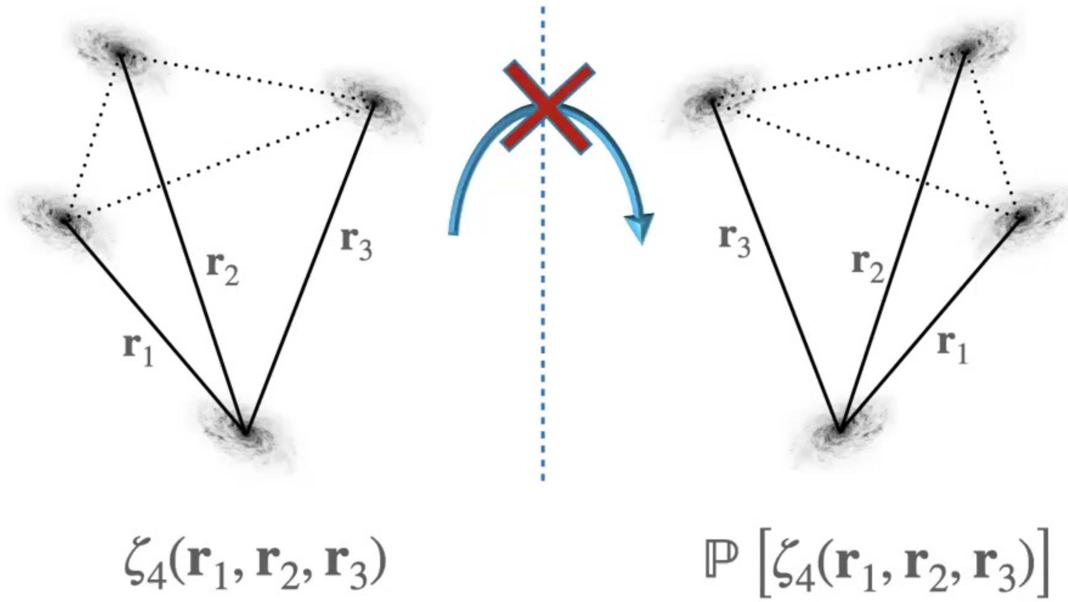
Parity Inversion = Rotation

**3-Point Correlation Function (3PCF):**

Parity Inversion = Rotation

**4-Point Correlation Function (4PCF):**

Parity Inversion ≠ Rotation




38

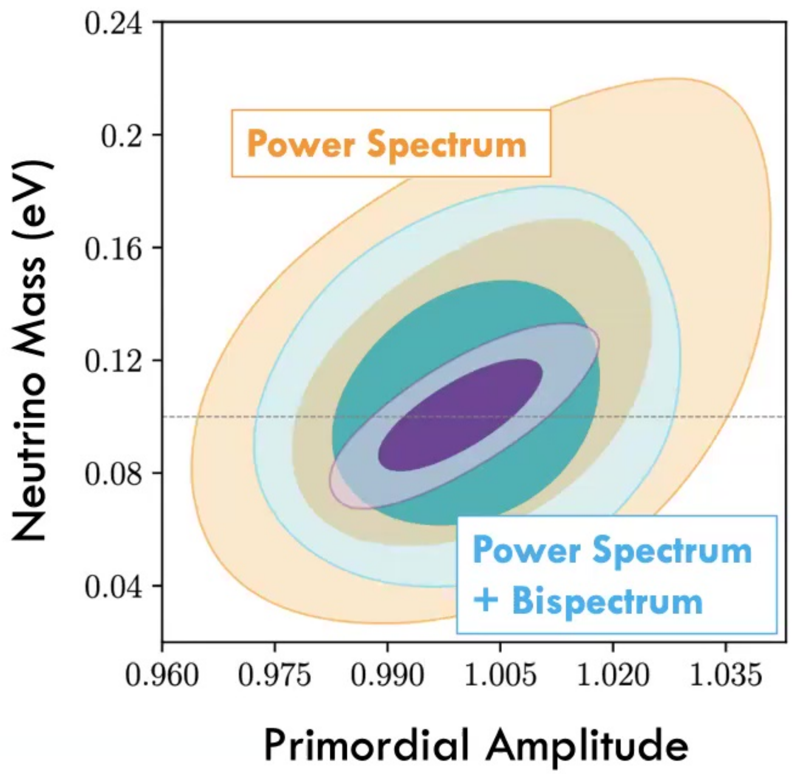
Lue+98, Jeong+12, Shiraishi 16, Cahn+21, **Philcox** (in prep.)

# WHY USE HIGHER-ORDER STATISTICS?

- ▷ **Sharpen** parameter constraints!
- ▷ **Break** parameter **degeneracies**!
- ▷ Test **non-standard** physics models!

**Why Use Large Scale Structure?**

- Signal-to-Noise is **cubic** in number of modes unlike CMB
- New physics constraints **don't** dilute with redshift




Chudaykin+19

# HOW TO MEASURE A BISPECTRUM

$$\hat{B}_g(k_1, k_2, k_3) = \int_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3 \in \text{bins}} \delta_g(\mathbf{k}_1)\delta_g(\mathbf{k}_2)\delta_g(\mathbf{k}_3)(2\pi)^3\delta_{\rm D}\left(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3\right)$$

**Problem:** We don't measure the density field directly.

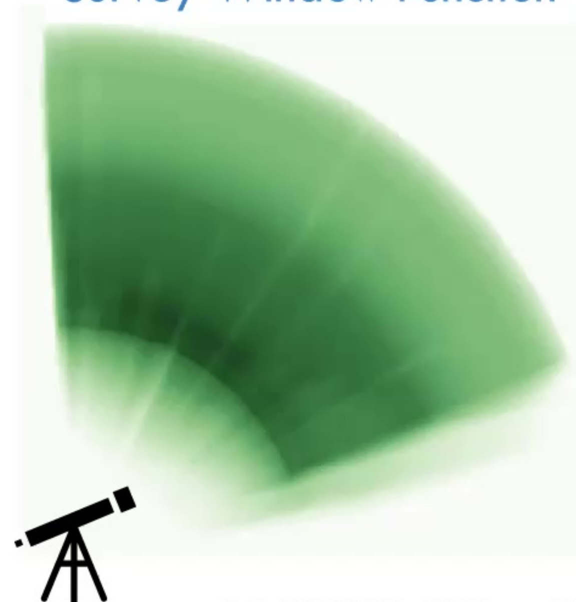
$$\delta_g(\mathbf{r}) \to W(\mathbf{r})\delta_g(\mathbf{r}) \qquad \delta_g(\mathbf{k}) \to \int \frac{d\mathbf{p}}{(2\pi)^3} W(\mathbf{k}-\mathbf{p})\delta_g(\mathbf{p})$$

*Window Function*

The measured bispectrum is a triple **convolution**

$$B_g(\mathbf{k}_1, \mathbf{k}_2) \to \int_{\mathbf{p}_1\mathbf{p}_2} W(\mathbf{k}_1-\mathbf{p}_1)W(\mathbf{k}_2-\mathbf{p}_2)W(\mathbf{p}_1+\mathbf{p}_2-\mathbf{k}_1-\mathbf{k}_2)B_g(\mathbf{p}_1,\mathbf{p}_2)$$

Survey Window Function



BOSS DR12, **Philcox 21**

41

# CONVOLUTION IS EXPENSIVE

$$B_g^{\mathrm{win}}(\mathbf{k}_1,\mathbf{k}_2) = \int_{\mathbf{p}_1\mathbf{p}_2} W(\mathbf{k}_1-\mathbf{p}_1)W(\mathbf{k}_2-\mathbf{p}_2)W(\mathbf{p}_1+\mathbf{p}_2-\mathbf{k}_1-\mathbf{k}_2)B_g(\mathbf{p}_1,\mathbf{p}_2)$$

- Window convolution is too costly to do repeatedly!

- Common approximation: apply the window **only** to the power spectrum

$$B_g(\mathbf{k}_1,\mathbf{k}_2) \supset P_L(k_1)P_L(k_2)$$

**But:**

- This gives **systematic errors** on large scales
- Spectra cannot be used to search for new physics!



Gil-Marin 15,17, Sugiyama+18, **Philcox** 21

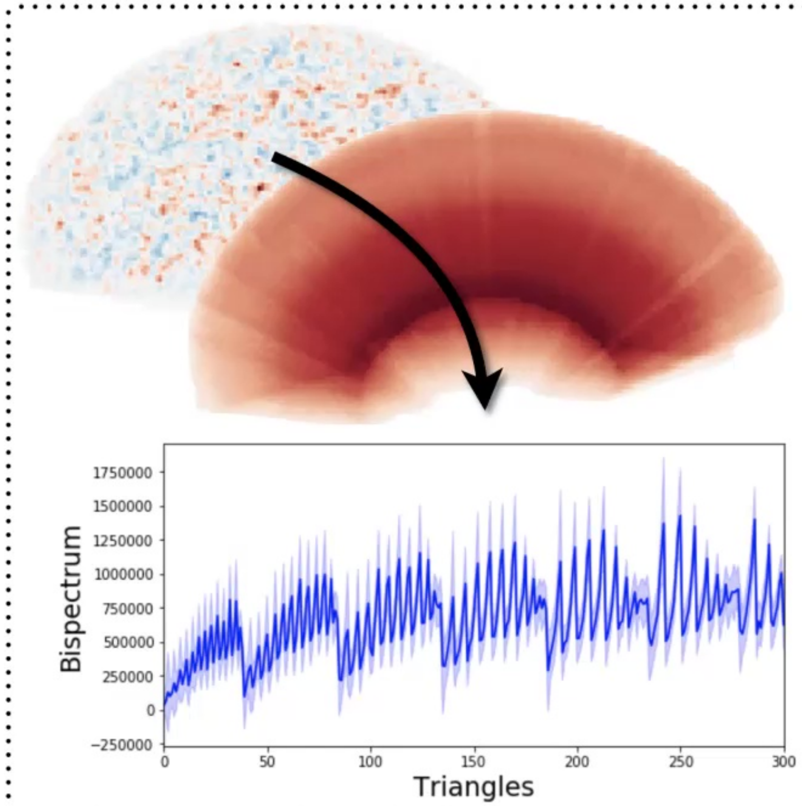
# BISPECTRA WITHOUT WINDOWS

**Alternatively:** estimate the **unwindowed** bispectrum directly

$$B_g^{\rm win}(\mathbf{k}_1,\mathbf{k}_2) = \int_{\mathbf{p}_1\mathbf{p}_2} W(\mathbf{k}_1-\mathbf{p}_1)W(\mathbf{k}_2-\mathbf{p}_2)W(\mathbf{p}_1+\mathbf{p}_2-\mathbf{k}_1-\mathbf{k}_2)B_g(\mathbf{p}_1,\mathbf{p}_2)$$

- Derive a **maximum-likelihood** estimator for the **true** bispectrum

$$\nabla_{B_g} L[\mathrm{data}|B_g] = 0 \quad \Rightarrow \quad \widehat{B}_g = \cdots$$




See GitHub.com/oliverphilcox/BOSS-Without-Windows



**Philcox** 20, 21

# BISPECTRA WITHOUT WINDOWS

**New Approach**

- Start from the **likelihood** for data **d,** using an Edgeworth expansion

$$L[\mathbf{d}](\mathbf{b}) = L_G[\mathbf{d}](\mathbf{b}) \left[1 + \frac{1}{3!}\mathsf{B}^{ijk}\left\{[\mathsf{C}^{-1}\mathbf{d}]_i\,[\mathsf{C}^{-1}\mathbf{d}]_j\,[\mathsf{C}^{-1}\mathbf{d}]_k - (\mathsf{C}^{-1}_{ij}d_k + 2\text{ perms.})\right\} + \cdots\right]$$

Gaussian Piece   Three-Point Function, $\mathsf{B}^{ijk} \equiv \langle d^i d^j d^k \rangle$   Covariance, $\mathsf{C}^{ij} = \langle d^i d^j \rangle$

- This depends on **survey geometry** through $\mathsf{C}^{ij}$ and **bispectrum** through $\mathsf{B}^{ijk}$

$$\nabla_{\mathbf{b}} \log L[\mathbf{d}](\mathbf{b}) = \mathbf{0}$$

- Optimize for true bispectrum, **b**:

$$\hat{b}^{\rm ML}_\alpha = \sum_\beta F^{-1,{\rm ML}}_{\alpha\beta}\hat{q}^{\rm ML}_\beta,$$

$$\hat{q}^{\rm ML}_\alpha = \frac{1}{6}\mathsf{B}^{ijk}_{,\alpha}\,[\mathsf{C}^{-1}\mathbf{d}]_i\left([\mathsf{C}^{-1}\mathbf{d}]_j\,[\mathsf{C}^{-1}\mathbf{d}]_k - 3\mathsf{C}^{-1}_{jk}\right)$$

*Cubic Estimator*

$$F^{\rm ML}_{\alpha\beta} = \frac{1}{6}\mathsf{B}^{ijk}_{,\alpha}\mathsf{B}^{lmn}_{,\beta}\mathsf{C}^{-1}_{il}\mathsf{C}^{-1}_{jm}\mathsf{C}^{-1}_{kn},$$

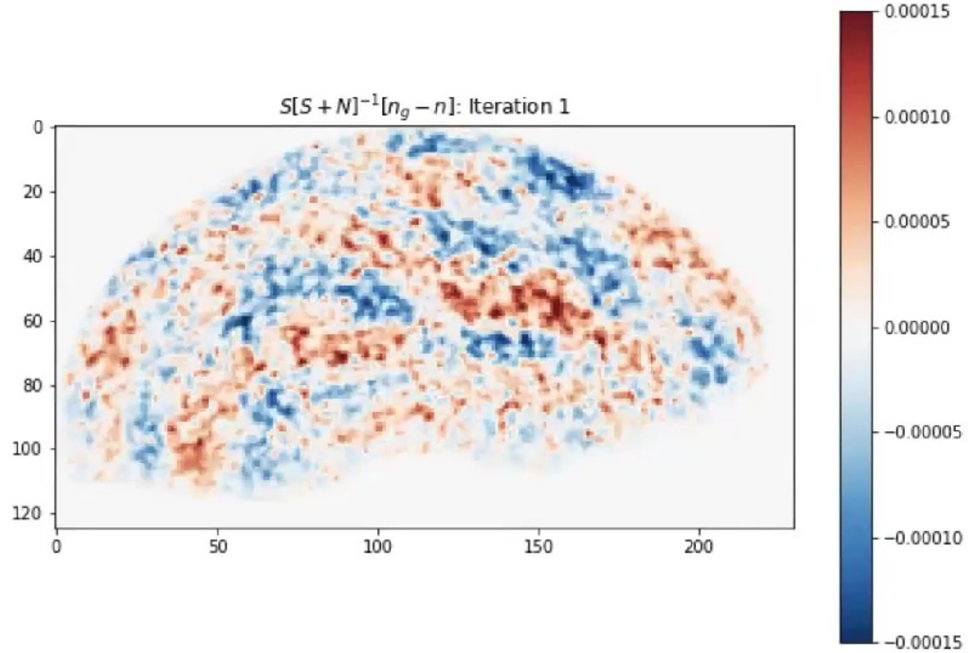
*Fisher Matrix*



**Philcox** 20, 21

# INVERSE VARIANCE WEIGHTING

Compute $\mathrm{C}^{-1}\boldsymbol{d}$ iteratively via
**conjugate gradient descent**



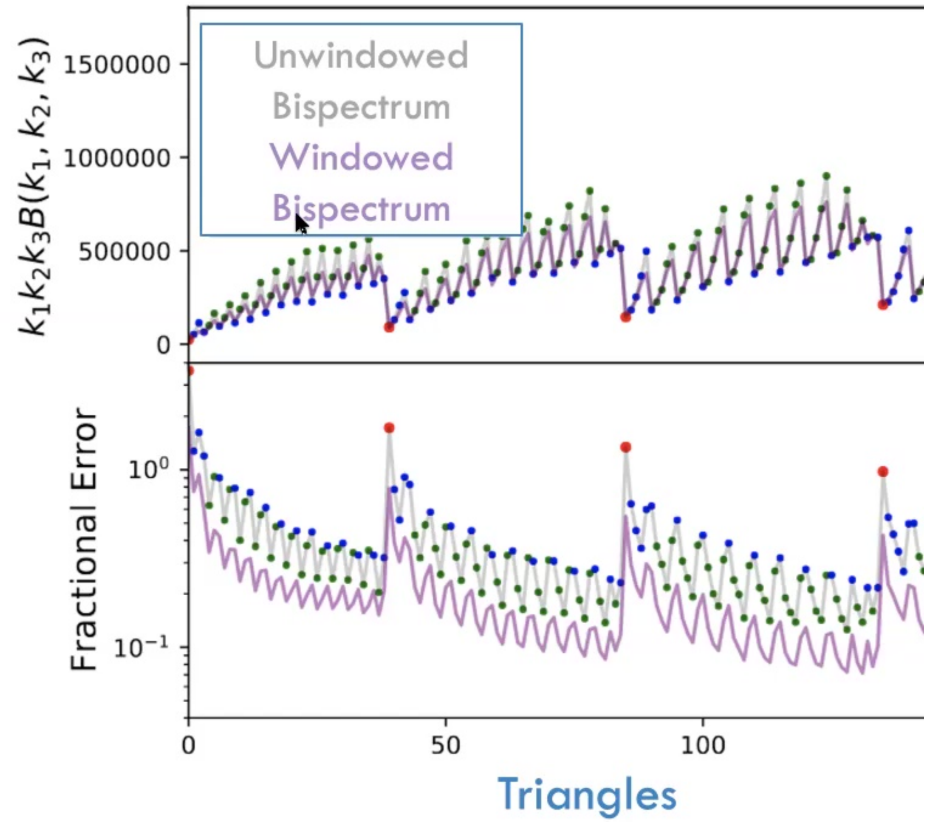

45

**Philcox** 20, 21

# BISPECTRA WITHOUT WINDOWS

Properties of the **cubic estimator:**

1. Unbiased
2. Minimum variance [as $B(k_1, k_2, k_3) \to 0$]
3. Window-free [effectively a deconvolution]

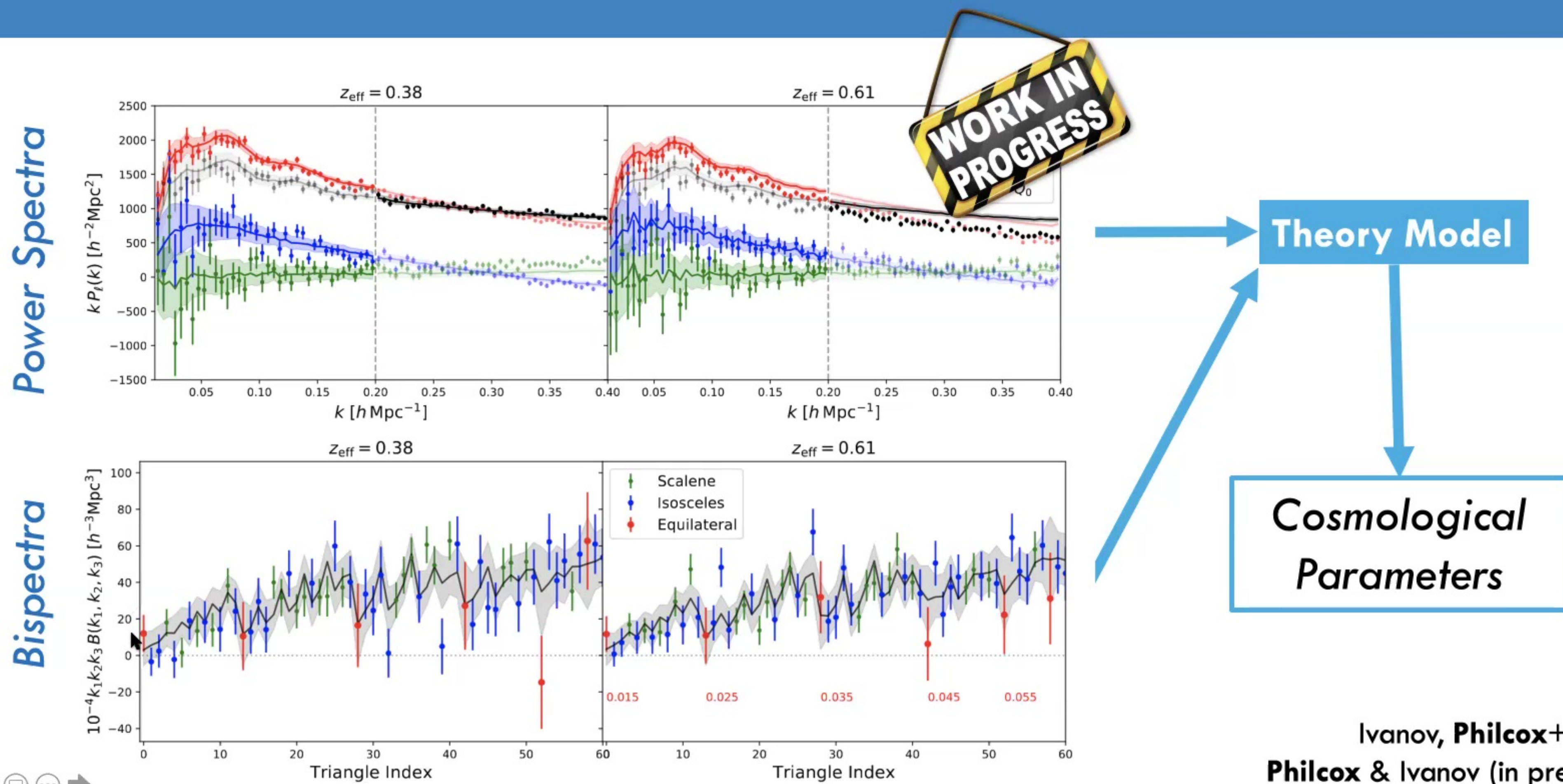



See GitHub.com/oliverphilcox/BOSS-Without-Windows

46

**Philcox** 20, 21

# BOSS WITHOUT WINDOWS

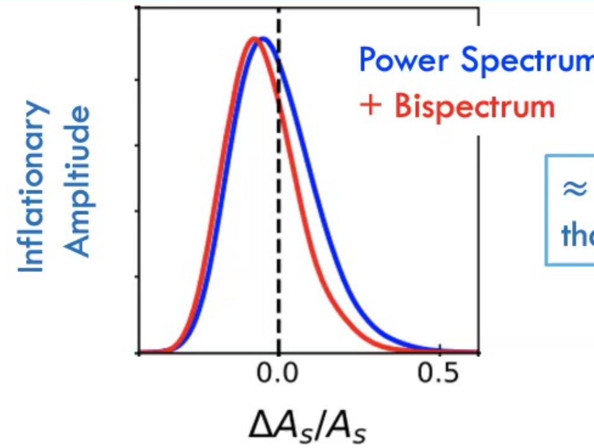



Ivanov, **Philcox**+21

**Philcox** & Ivanov (in prep.)

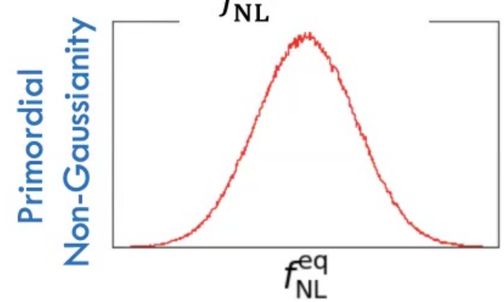
# WHAT WILL WE MEASURE?

- Tighter constraints on **cosmological** and **galaxy formation** parameters
  - $\sigma_8$ improves by 10%
  - Tidal bias improves by 50%
- Bounds on **all** flavors of **Primordial Non-Gaussianity**
  - First equilateral-type measurement from LSS




≈ 8× weaker than *Planck*

$\sigma_{f_{NL}^{eq}} \approx 500$



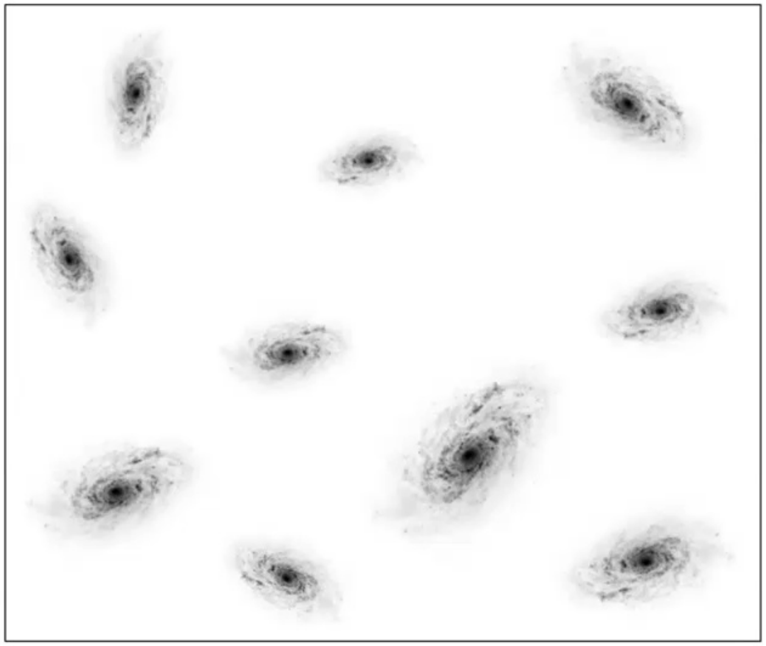

≈ 7× weaker than *Planck*

Ivanov, **Philcox**+21
**Philcox** & Ivanov (in prep.)

# HOW TO MEASURE A CORRELATION FUNCTION

$$\hat{\zeta}_{g,4}(r_1, r_2, r_3, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3, \hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) = \int \frac{d\mathbf{x}}{V} \int_{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3 \in \text{bins}} \delta_g(\mathbf{x}) \delta_g(\mathbf{x} + \mathbf{r}_1) \delta_g(\mathbf{x} + \mathbf{r}_2) \delta_g(\mathbf{x} + \mathbf{r}_3)$$

Measuring the 4PCF involves counting **quadruplets** of galaxies



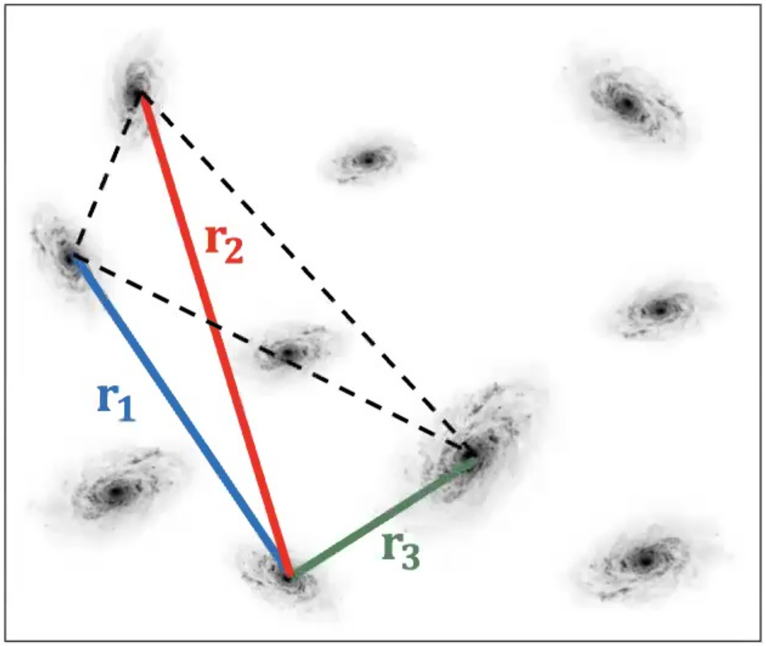
52

**Philcox**+21

# HOW TO MEASURE A CORRELATION FUNCTION

$$\hat{\zeta}_{g,4}(r_1, r_2, r_3, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3, \hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) = \int \frac{d\mathbf{x}}{V} \int_{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3 \in \mathrm{bins}} \delta_g(\mathbf{x})\delta_g(\mathbf{x}+\mathbf{r}_1)\delta_g(\mathbf{x}+\mathbf{r}_2)\delta_g(\mathbf{x}+\mathbf{r}_3)$$

Measuring the 4PCF involves counting **quadruplets** of galaxies



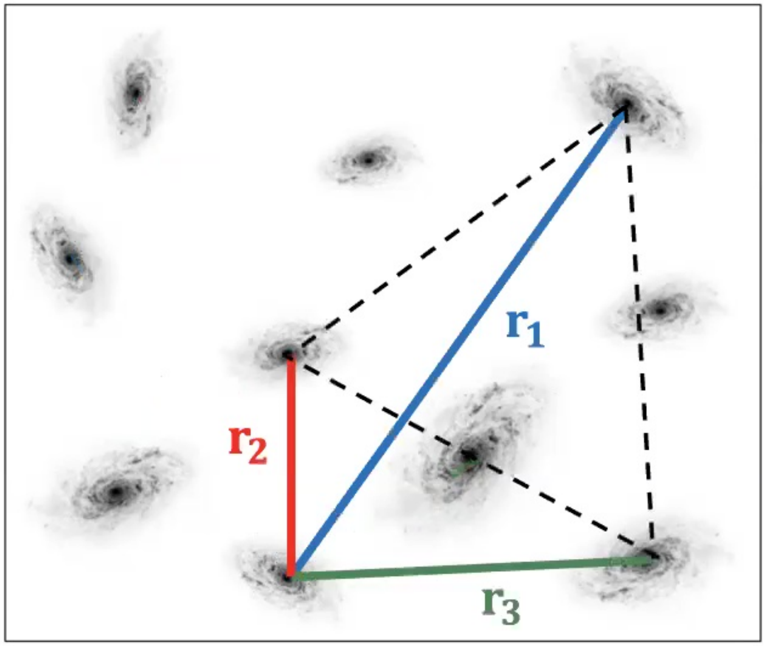

53

**Philcox**+21

# HOW TO MEASURE A CORRELATION FUNCTION

$$\hat{\zeta}_{g,4}(r_1, r_2, r_3, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3, \hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) = \int \frac{d\mathbf{x}}{V} \int_{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3 \in \text{bins}} \delta_g(\mathbf{x})\delta_g(\mathbf{x}+\mathbf{r}_1)\delta_g(\mathbf{x}+\mathbf{r}_2)\delta_g(\mathbf{x}+\mathbf{r}_3)$$

Measuring the 4PCF involves counting **quadruplets** of galaxies



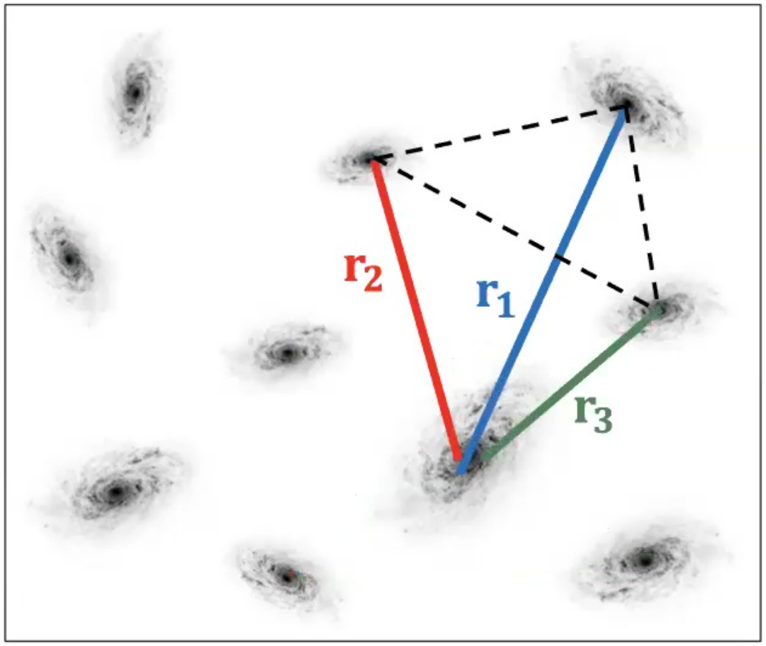

54

**Philcox**+21

# HOW TO MEASURE A CORRELATION FUNCTION

$$\hat{\zeta}_{g,4}(r_1, r_2, r_3, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3, \hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) = \int \frac{d\mathbf{x}}{V} \int_{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3 \in \text{bins}} \delta_g(\mathbf{x})\delta_g(\mathbf{x}+\mathbf{r}_1)\delta_g(\mathbf{x}+\mathbf{r}_2)\delta_g(\mathbf{x}+\mathbf{r}_3)$$

Measuring the 4PCF involves counting **quadruplets** of galaxies




55

**Philcox**+21

# HOW TO MEASURE A CORRELATION FUNCTION

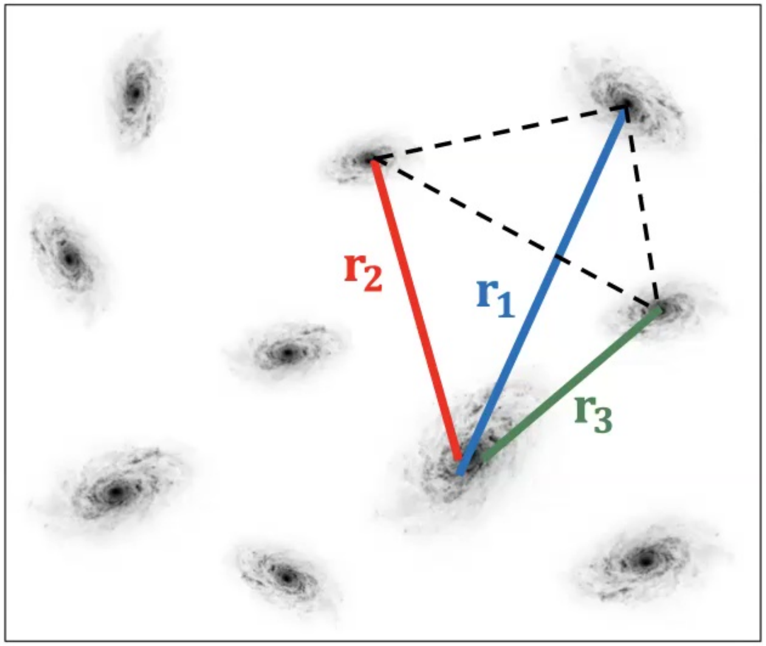
$$\hat{\zeta}_{g,4}(r_1, r_2, r_3, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3, \hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3) = \int \frac{d\mathbf{x}}{V} \int_{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3 \in \text{bins}} \delta_g(\mathbf{x})\delta_g(\mathbf{x}+\mathbf{r}_1)\delta_g(\mathbf{x}+\mathbf{r}_2)\delta_g(\mathbf{x}+\mathbf{r}_3)$$

Measuring the 4PCF involves counting **quadruplets** of galaxies

Total number of quadruplets:

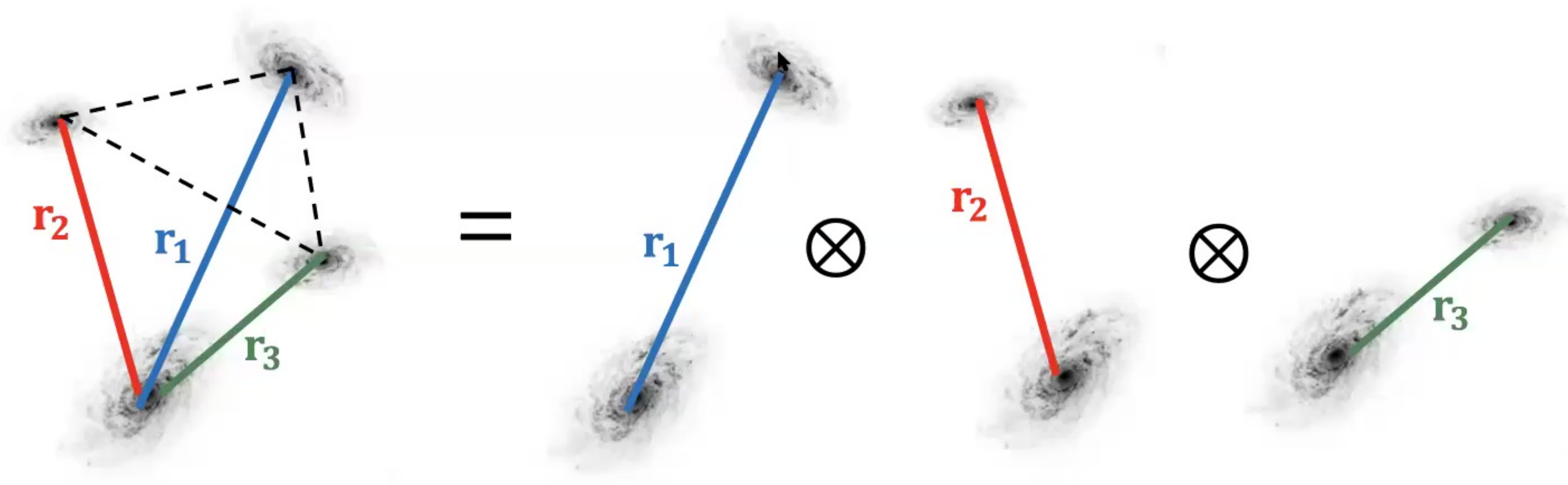
$$\mathcal{O}(N_{\text{gal}}^4)$$

**This is too many to count…**




**Philcox**+21

# ONE TETRAHEDRON = THREE VECTORS




*3 lengths + 3 angles*

$(r_1, r_2, r_3, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_3, \hat{\mathbf{r}}_2 \cdot \hat{\mathbf{r}}_3)$

*1 length + 1 direction*

$(r_1, \hat{\mathbf{r}}_1)$

*1 length + 1 direction*

$(r_2, \hat{\mathbf{r}}_2)$

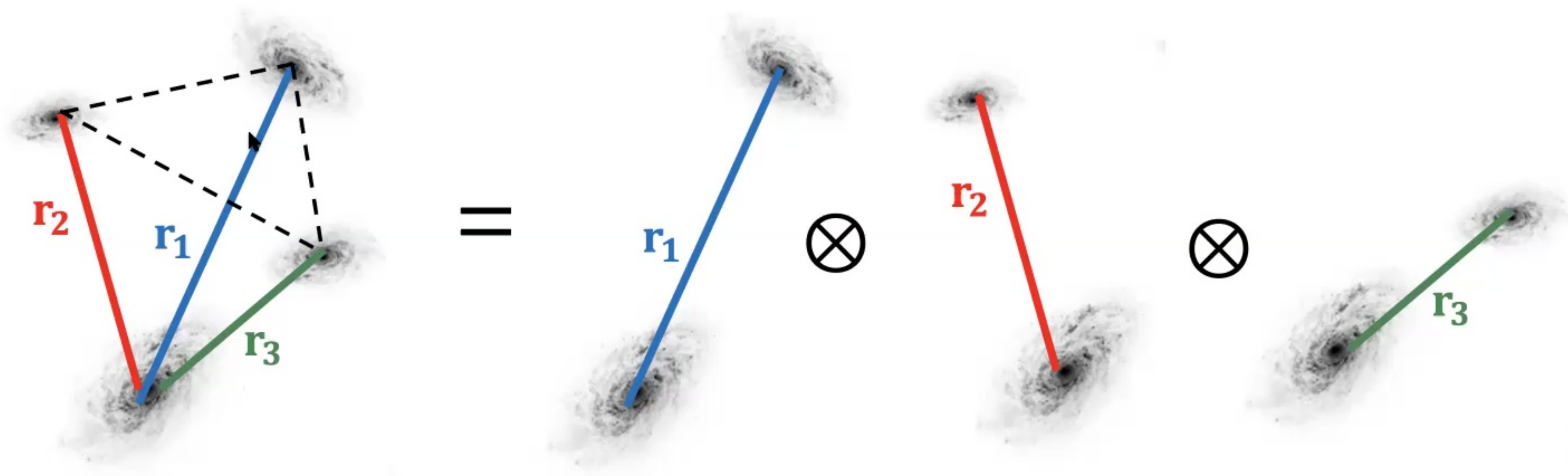
*1 length + 1 direction*

$(r_3, \hat{\mathbf{r}}_3)$

57

**Philcox**+21

# ONE TETRAHEDRON = THREE VECTORS




| 3 lengths + 3 multipoles | 1 length + 2 multipoles | 1 length + 2 multipoles | 1 length + 2 multipoles |
|---|---|---|---|
| $(r_1, r_2, r_3, \ell_1, \ell_2, \ell_3)$ | $(r_1, \ell_1, m_1)$ | $(r_2, \ell_2, m_2)$ | $(r_3, \ell_3, m_3)$ |

58

Philcox+21

# ANGULAR MOMENTUM BASIS

Expand 4PCF in basis of **isotropic functions**

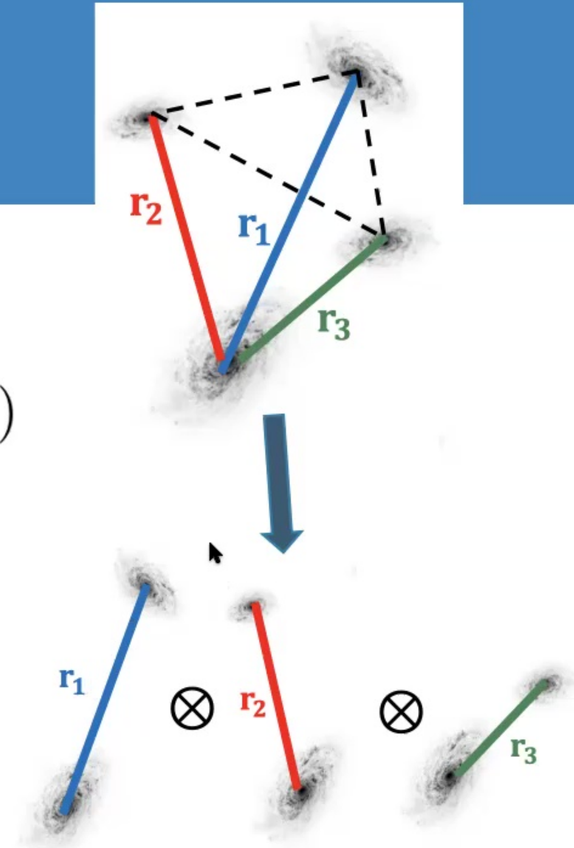
$$\zeta_4(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = \sum_{\ell_1\ell_2\ell_3} \zeta_{\ell_1\ell_2\ell_3}(r_1,r_2,r_3)\mathcal{P}_{\ell_1\ell_2\ell_3}(\hat{\mathbf{r}}_1,\hat{\mathbf{r}}_2,\hat{\mathbf{r}}_3)$$

Coefficients    Basis Functions

Basis formed from **angular momentum addition** in 3D

$$\mathcal{P}_{\ell_1\ell_2\ell_3}(\hat{\mathbf{r}}_1,\hat{\mathbf{r}}_2,\hat{\mathbf{r}}_3) = \sum_{m_1m_2m_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} Y^*_{\ell_1m_1}(\hat{\mathbf{r}}_1)Y^*_{\ell_2m_2}(\hat{\mathbf{r}}_2)Y^*_{\ell_3m_3}(\hat{\mathbf{r}}_3)$$

This is **separable** in $\hat{\mathbf{r}}_1$, $\hat{\mathbf{r}}_2$, $\hat{\mathbf{r}}_3$






Cahn+20, **Philcox**+21

# A SEPARABLE BASIS ⇒ A QUADRATIC ESTIMATOR

$$\hat{\zeta}_{\ell_1\ell_2\ell_3}(r_1,r_2,r_3) = \sum_{m_1m_2m_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \int d\mathbf{x}\, \delta_g(\mathbf{x}) \left[\int_{\mathbf{r}_1} \delta_g(\mathbf{x}+\mathbf{r}_1) Y_{\ell_1 m_1}(\hat{\mathbf{r}}_1)\right] \left[\int_{\mathbf{r}_2} \delta_g(\mathbf{x}+\mathbf{r}_2) Y_{\ell_2 m_2}(\hat{\mathbf{r}}_2)\right] \left[\int_{\mathbf{r}_3} \delta_g(\mathbf{x}+\mathbf{r}_3) Y_{\ell_3 m_3}(\hat{\mathbf{r}}_3)\right]$$

The estimator **factorizes** into **independent** pieces



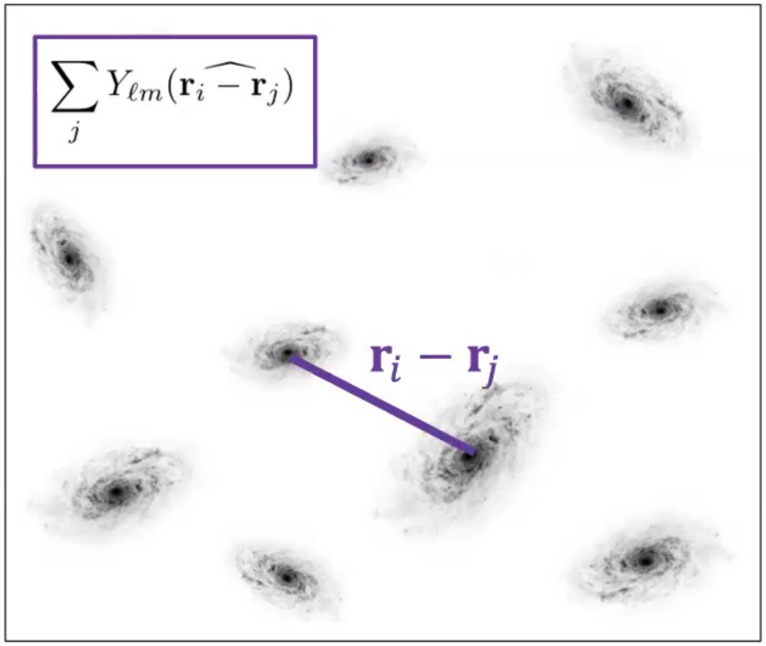
Philcox+21

# A SEPARABLE BASIS ⇒ A QUADRATIC ESTIMATOR

$$\hat{\zeta}_{\ell_1\ell_2\ell_3}(r_1,r_2,r_3) = \sum_{m_1m_2m_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \int d\mathbf{x}\, \delta_g(\mathbf{x}) \left[\int_{\mathbf{r}_1} \delta_g(\mathbf{x}+\mathbf{r}_1) Y_{\ell_1 m_1}(\hat{\mathbf{r}}_1)\right] \left[\int_{\mathbf{r}_2} \delta_g(\mathbf{x}+\mathbf{r}_2) Y_{\ell_2 m_2}(\hat{\mathbf{r}}_2)\right] \left[\int_{\mathbf{r}_3} \delta_g(\mathbf{x}+\mathbf{r}_3) Y_{\ell_3 m_3}(\hat{\mathbf{r}}_3)\right]$$

The estimator **factorizes** into **independent** pieces

To compute the 4PCF: count *pairs* of galaxies

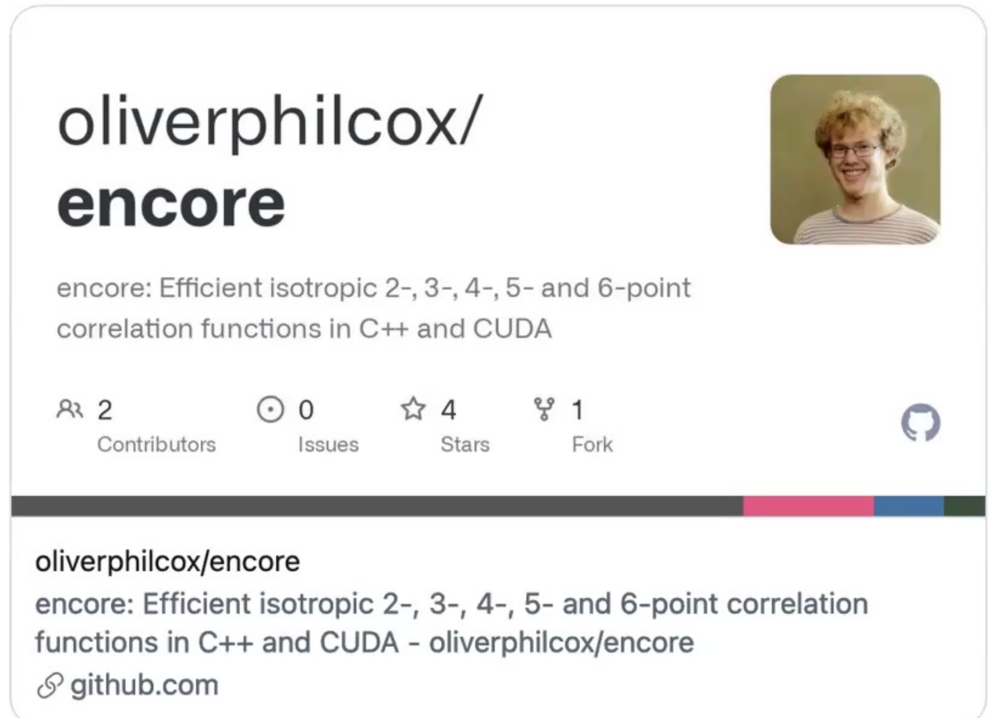





Philcox+21

# ENCORE: ULTRA-FAST N-POINT FUNCTIONS

- Public C++/CUDA code
- Computes isotropic 2-, 3-, 4-, 5- and 6-point correlation functions
- Corrects for **survey geometry**
- Requires ~ 10 CPU-hours to compute 4PCF of current data




See GitHub.com/oliverphilcox/encore

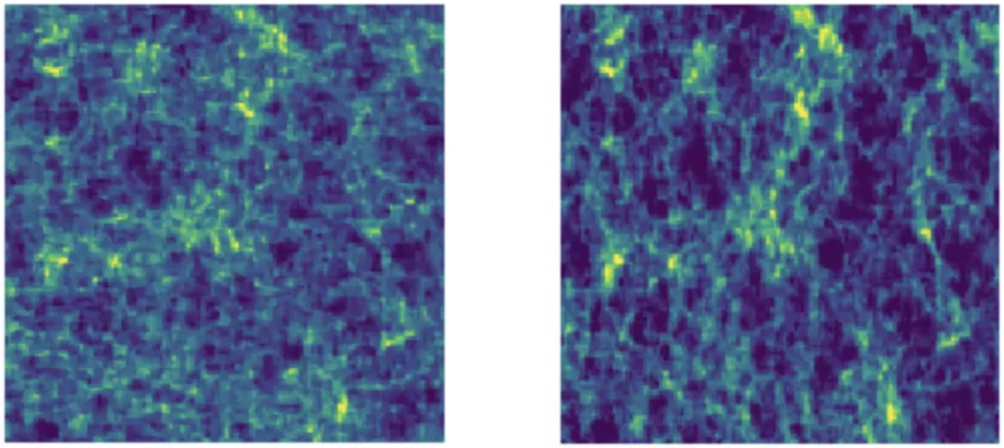
63

**Philcox**+21

# BEYOND THE 4-POINT FUNCTION

This generalizes **beyond** the 4PCF

- 5PCF, 6PCF, …
- **Anisotropic** correlation functions



*Real Space* *Redshift Space*

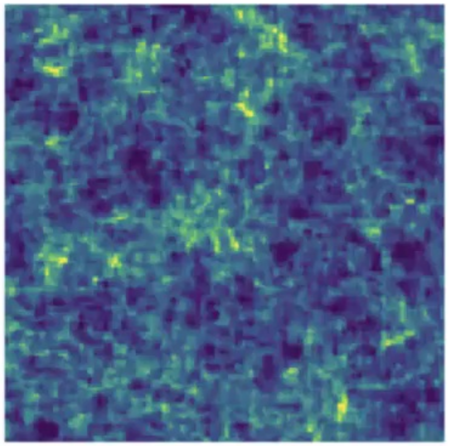


**Philcox**+21

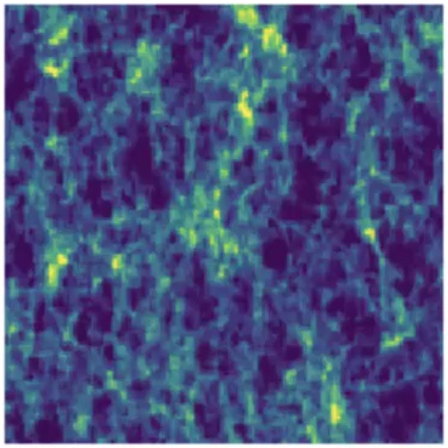
# BEYOND THE 4-POINT FUNCTION

This generalizes **beyond** the 4PCF

- 5PCF, 6PCF, …
- **Anisotropic** correlation functions
- Non-Flat Universes
- Two, Three, Four, … Dimensions



*Real Space*



*Redshift Space*

Requires the addition of N angular momenta in D dimensions [*i.e.* $\mathfrak{so}(D)$ *Lie algebra*]

See GitHub.com/oliverphilcox/NPCFs.jl



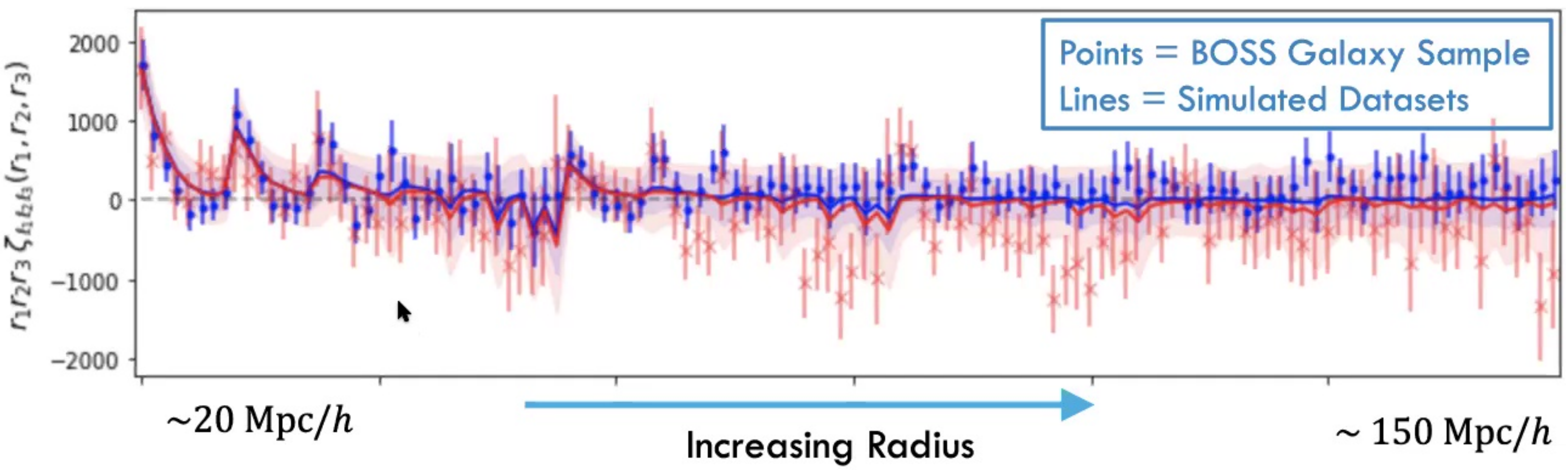
**Philcox**+21

# MEASURING THE 4-POINT FUNCTION

Compute the 4PCF from $\sim 10^6$ **BOSS galaxies**

**Do we detect a signal?**

4PCF (one of 42 components)

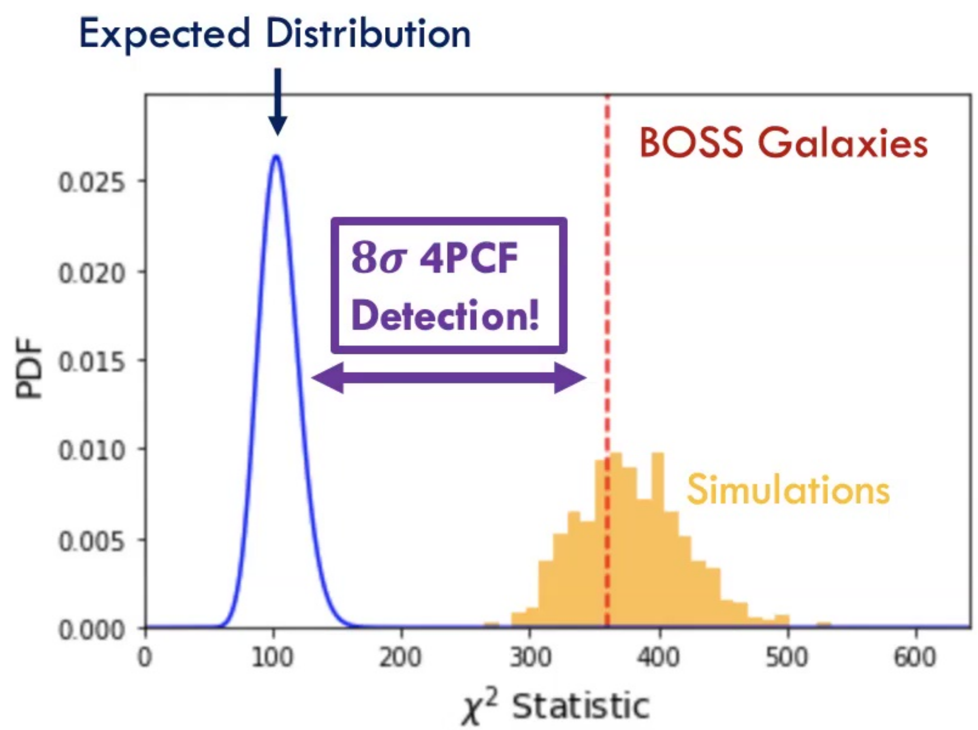



67

**Philcox+21**

# CAN WE DETECT THE GRAVITATIONAL 4PCF?

- Perform a $\chi^2$-test to search for a **gravitational** 4PCF
- Null Hypothesis: **4PCF = 0.**
- **Strong** detection of non-Gaussianity!




71

**Philcox**+21

# WHAT'S NEXT FOR THE 4-POINT FUNCTION?

- Create a **theory** model and **quantify** information content:
  - Allows **$\Lambda$CDM information** to be extracted

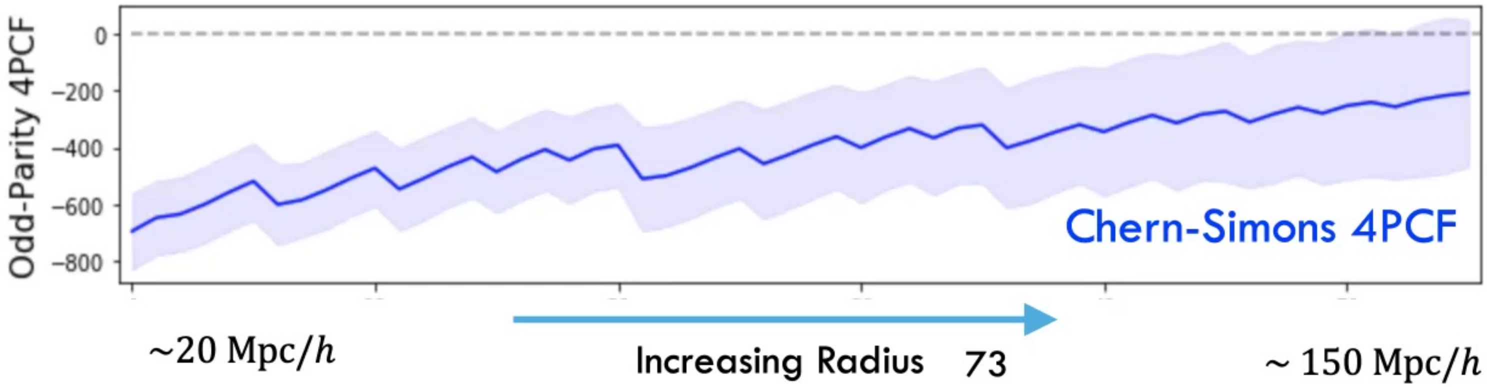


Cahn+21, **Philcox** (in prep.)

# WHAT'S NEXT FOR THE 4-POINT FUNCTION?

- Create a **theory** model and **quantify** information content:
  - Allows **ΛCDM information** to be extracted
- Search for **parity-violating** physics in the BOSS 4PCF

$$\zeta_4(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) - \mathbb{P}[\zeta_4(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)]$$




73

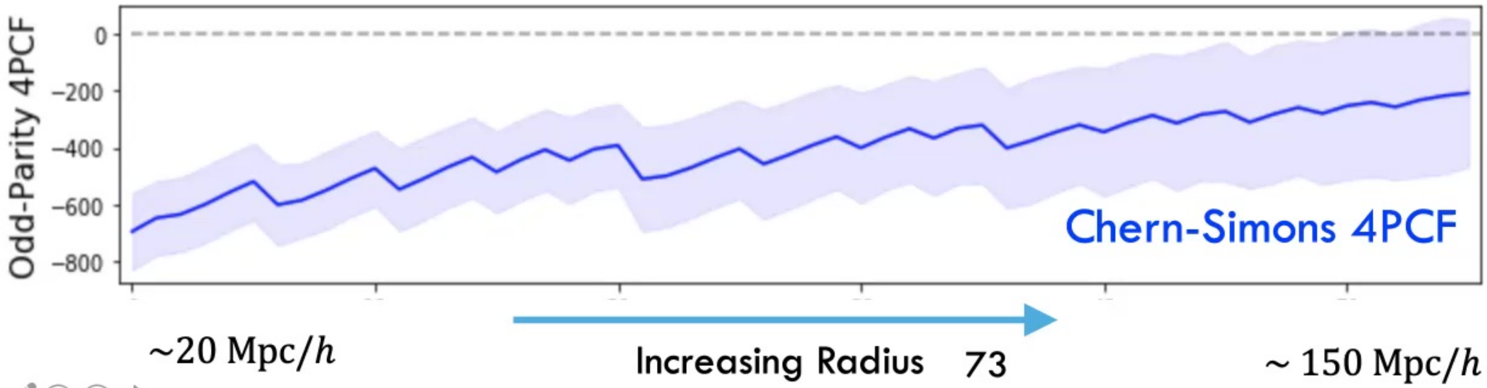
Cahn+21, **Philcox** (in prep.)

# WHAT'S NEXT FOR THE 4-POINT FUNCTION?

- Create a **theory** model and **quantify** information content:
  - Allows **ΛCDM information** to be extracted
- Search for **parity-violating** physics in the BOSS 4PCF

$$\zeta_4(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) - \mathbb{P}[\zeta_4(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)]$$

- Apply to **DESI** data [2× higher precision] and combine with the **CMB**




73

Cahn+21, **Philcox** (in prep.)


# CONCLUSIONS

- Non-Gaussian statistics:
  1. **Sharpen** cosmological constraints
  2. Probe **non-standard** physics in the early Universe
- **Fast** and **accurate** estimators now available
- Extract **more** information from LSS surveys **without** additional cost