Primordial non-Gaussianity with future large scale structure measurements

William Coulton
with Paco Villaescusa-Navarro, Ben Wandelt, Gabriel Jung, Drew Jamieson, Michele Liguori, Licia Verde, Dionysios Karagiannis, and Marco Baldi

On arXiv earlier this month: 2206.01624 & 2206.01619

An alternative method for Fisher forecasting with simulations

Ubiquity of Fisher forecasts (in cosmology?)

- Fisher forecasts are useful for quick estimates:
    - of the impact of design choices for upcoming surveys (e.g. Astro2020 and Snowmass)
    - of the detectability of my new model of dark matter, dark energy, tension-solving idea, etc…
    - of the information content for my new statistical probe (e.g. topological measures, N>>2-correlation functions etc)
    - (your use case here)

The anatomy of the Fisher forecast

- The Fisher information is defined as

$$\mathcal{I}_{ij} = \langle \frac{\partial \ln \mathcal{L}}{\partial \theta_i} \frac{\partial \ln \mathcal{L}}{\partial \theta_j} \rangle$$

θ - parameters

$\mathcal{L}$ - likelihood

- Useful as $\mathrm{Var}[\theta_\mathrm{i}] \geq \mathcal{I}_{ii}^{-1}$, i.e. bounds the error on unbiased parameter estimators

- For a Gaussian (with parameter indep. covariance)

$$\mathcal{I}_{ij} = \frac{\partial \mu_a}{\partial \theta_i} \mathbf{C}_{ab}^{-1} \frac{\partial \mu_b}{\partial \theta_j}$$

- Thus to perform the forecast we need the derivative of the mean, μ, with respect to the parameters and the covariance, $\mathbf{C}$

Monte Carlo Fisher estimates

- Often these components are analytically intractable thus

$$\widehat{\mathcal{F}}_{ij} = \widehat{\frac{\partial \mu_a}{\partial \theta_i}} \, \widehat{\mathbf{C}_{ab}^{-1}} \, \widehat{\frac{\partial \mu_b}{\partial \theta_j}}$$

- It is straight forward to see this is a biased estimator

$$\langle \widehat{\mathcal{F}}_{ij} \rangle = \frac{\partial \mu_a}{\partial \theta_i} \mathbf{C}_{ab}^{-1} \frac{\partial \mu_b}{\partial \theta_j} + \mathrm{Cov}\left[\widehat{\frac{\partial \mu_a}{\partial \theta_i}}, \widehat{\frac{\partial \mu_b}{\partial \theta_j}} \right] \mathbf{C}_{ab}^{-1}$$

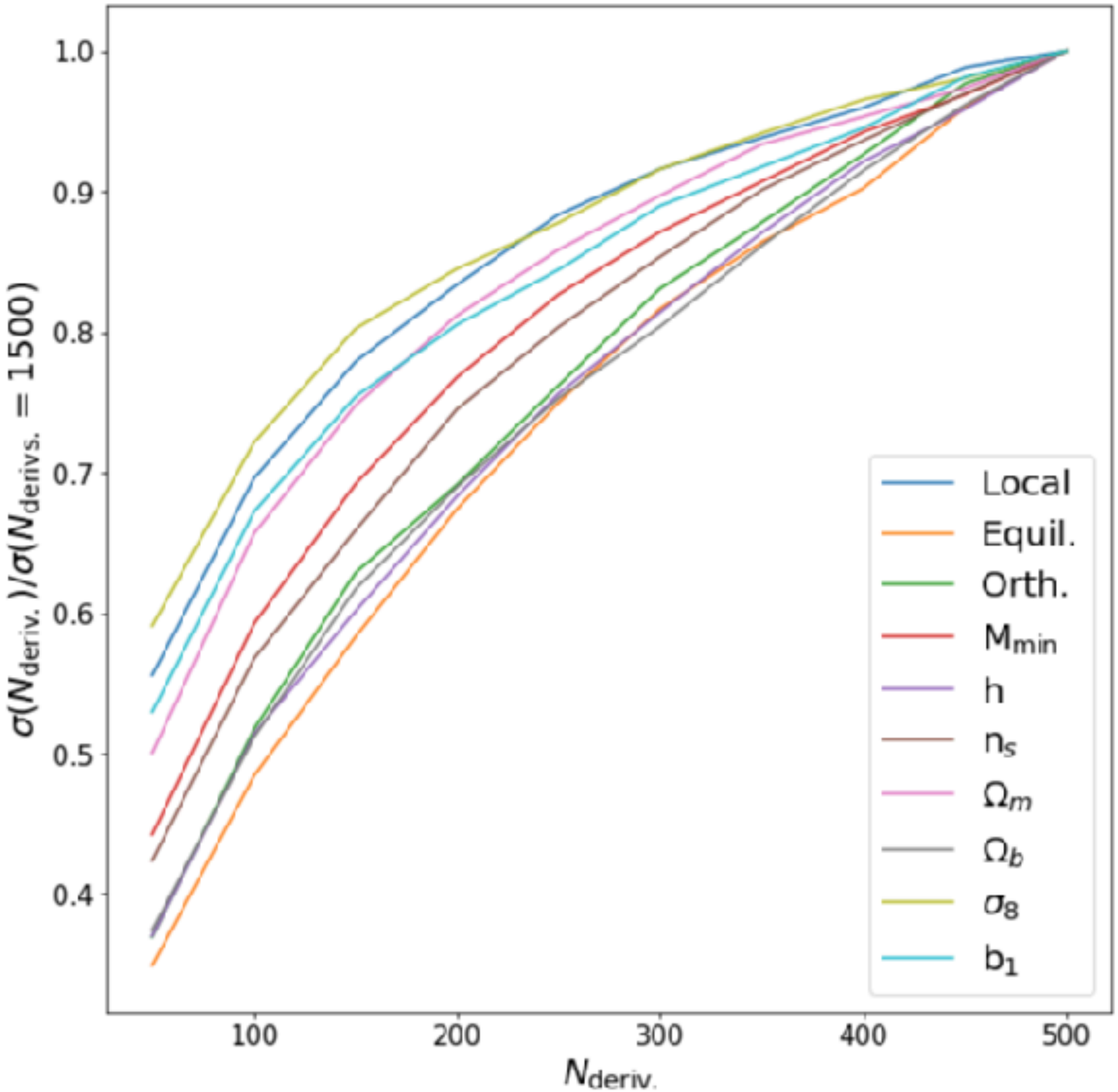
- This is biased high, and thus parameter constraints will be overly optmistic*

*note: immediately a few techniques jump to mind to obtain an unbiased estimate, however generally these do not leave the Fisher information invertible

How important are these biases?

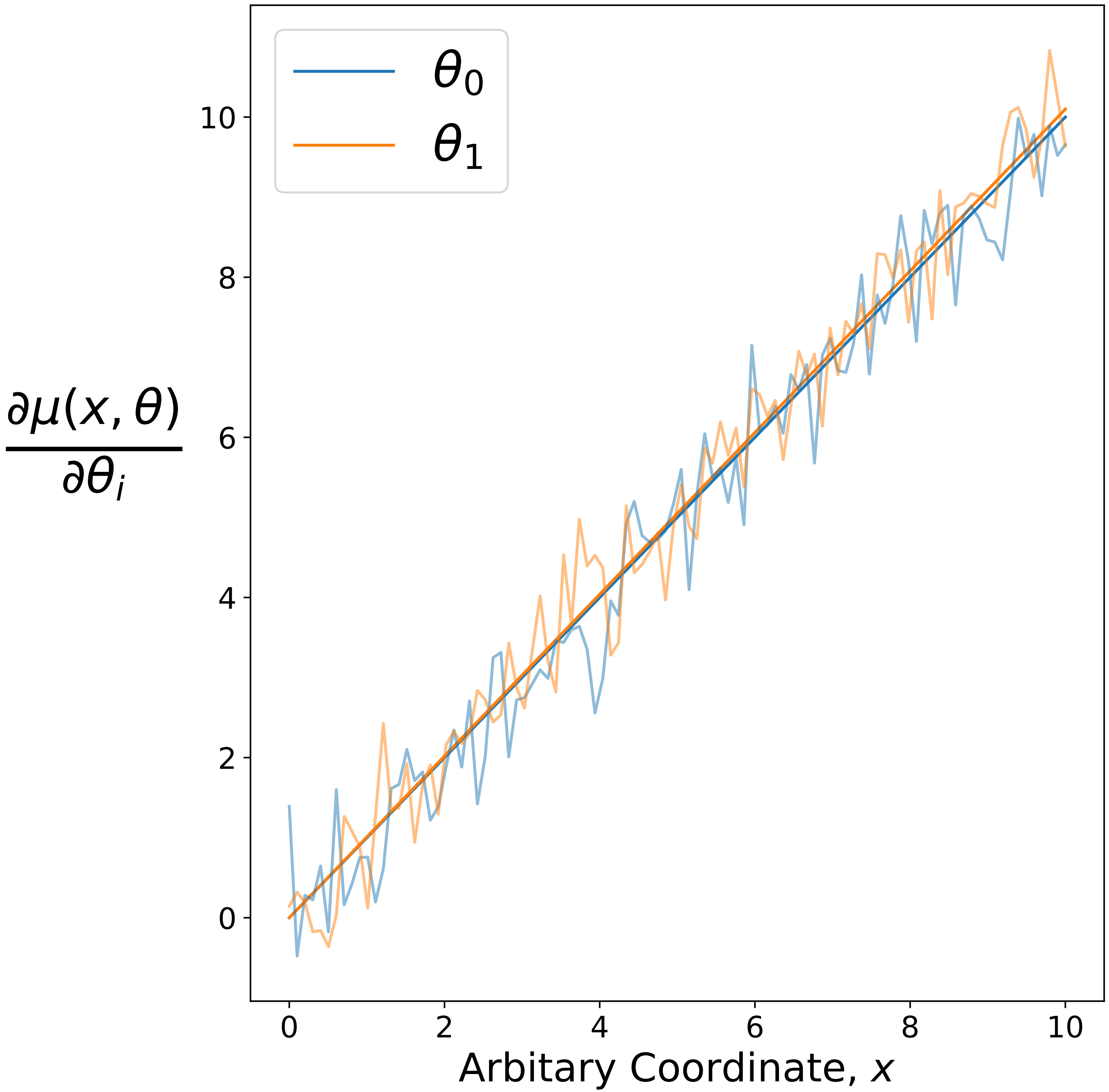
Stability of parameter constraints from a bispectrum analysis:

The constraints also scale as $\sim \sqrt{N_{\mathrm{deriv}}}$ **!**



Why does this arise?

An extreme example of how noise biases Fisher forecasts: derivatives generally look less similar!



Our approach: compressed Fisher

- Consider a set of compressed statistics **Alsing et al (2018a, 2108b)**

$$\mathbf{t} = \frac{\partial \ln \mathscr{L}(\mathbf{x})}{\partial \theta}\Big|_{\theta=\theta_*}$$

these have: $\mathcal{F}^{tt} = \mathcal{F}$!

- Perform the Monte Carlo estimate of the Fisher information in the compressed space!

- How does this help?

This scales ~ dimension of problem

$$\langle \widehat{\mathcal{F}}^{tt}_{ij} \rangle = \frac{\partial \mu^t_a}{\partial \theta_i} \mathbf{C}^{\mathbf{tt}\,-1}_{ab} \frac{\partial \mu^t_b}{\partial \theta_j} + \mathrm{Cov}\left[\widehat{\frac{\partial \mu^t_a}{\partial \theta_i}}, \widehat{\frac{\partial \mu^t_b}{\partial \theta_j}}\right] \mathbf{C}^{\mathbf{tt}\,-1}_{ab}$$

The cost: suboptimality

- As an explicit example for a Gaussian likelihood:

$$\mathbf{t} = \frac{\partial \mu_a}{\partial \theta_i} C^{-1}_{ab} \left(d_b - \mu_b \right)$$

- Now this compression requires the same ingredients as the Fisher forecast!

- Solution is to use some of the simulations to estimate the compression

$$\mathbf{t} = \widehat{\frac{\partial \mu_a}{\partial \theta_i}} \, \widehat{C}^{-1}_{ab} \left(d_b - \widehat{\mu_b} \right)$$

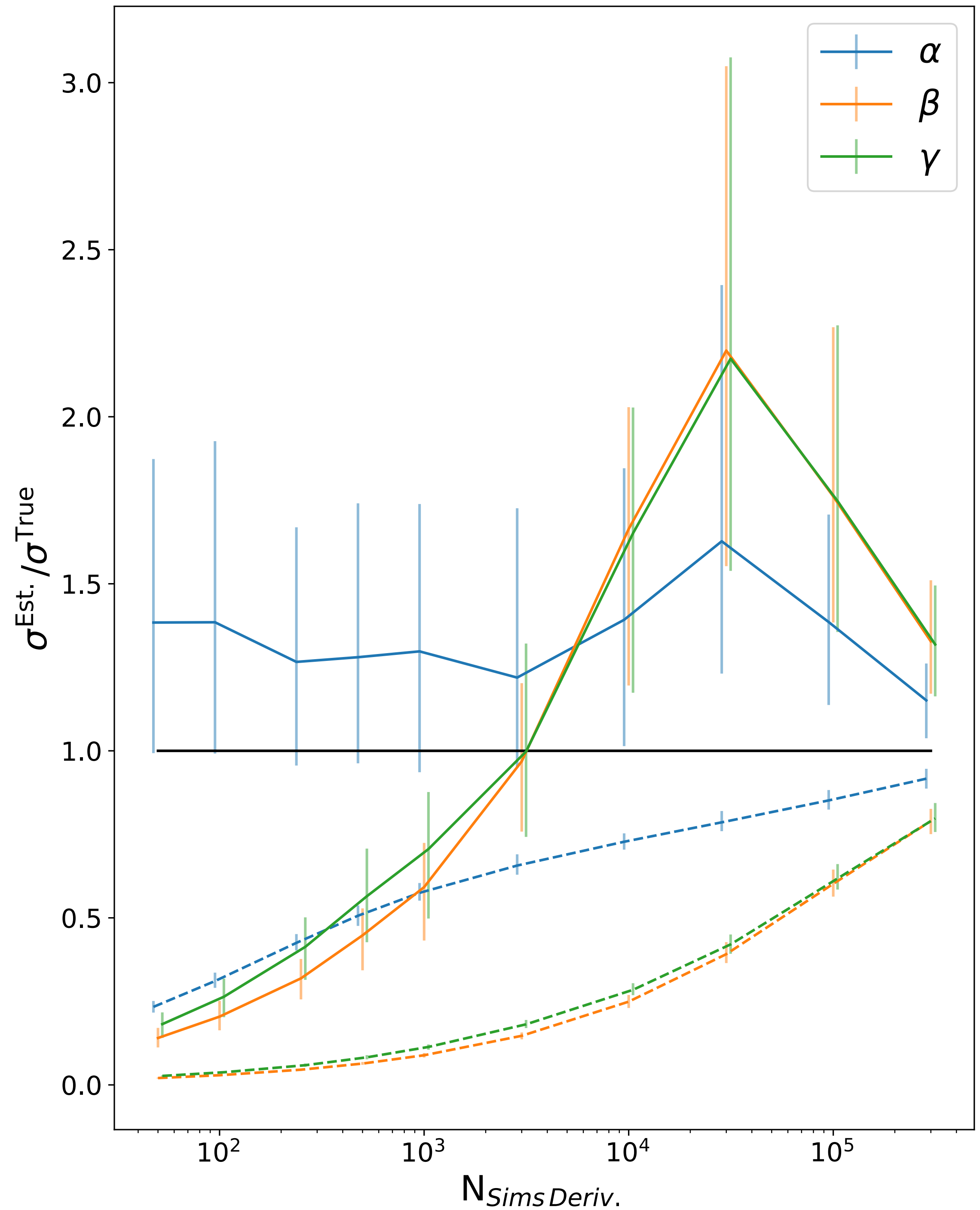
and the reminder for the compressed Fisher estimate.

- The cost of this is a suboptimal compression and therefore the Fisher forecast errors will be larger than the truth

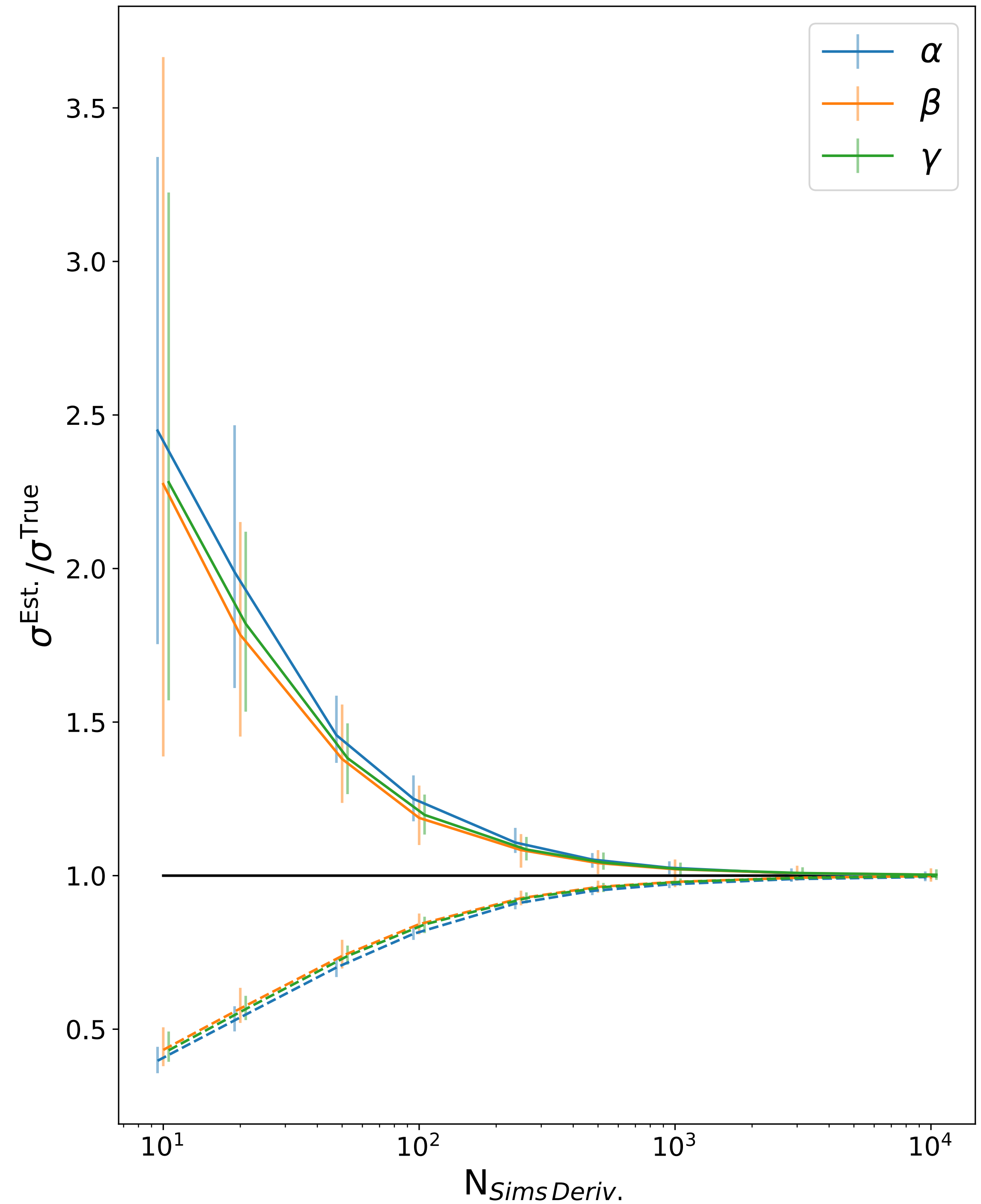
How well does it work?

Example forecast constraining power for two 3 parameter toy models

Poisson model



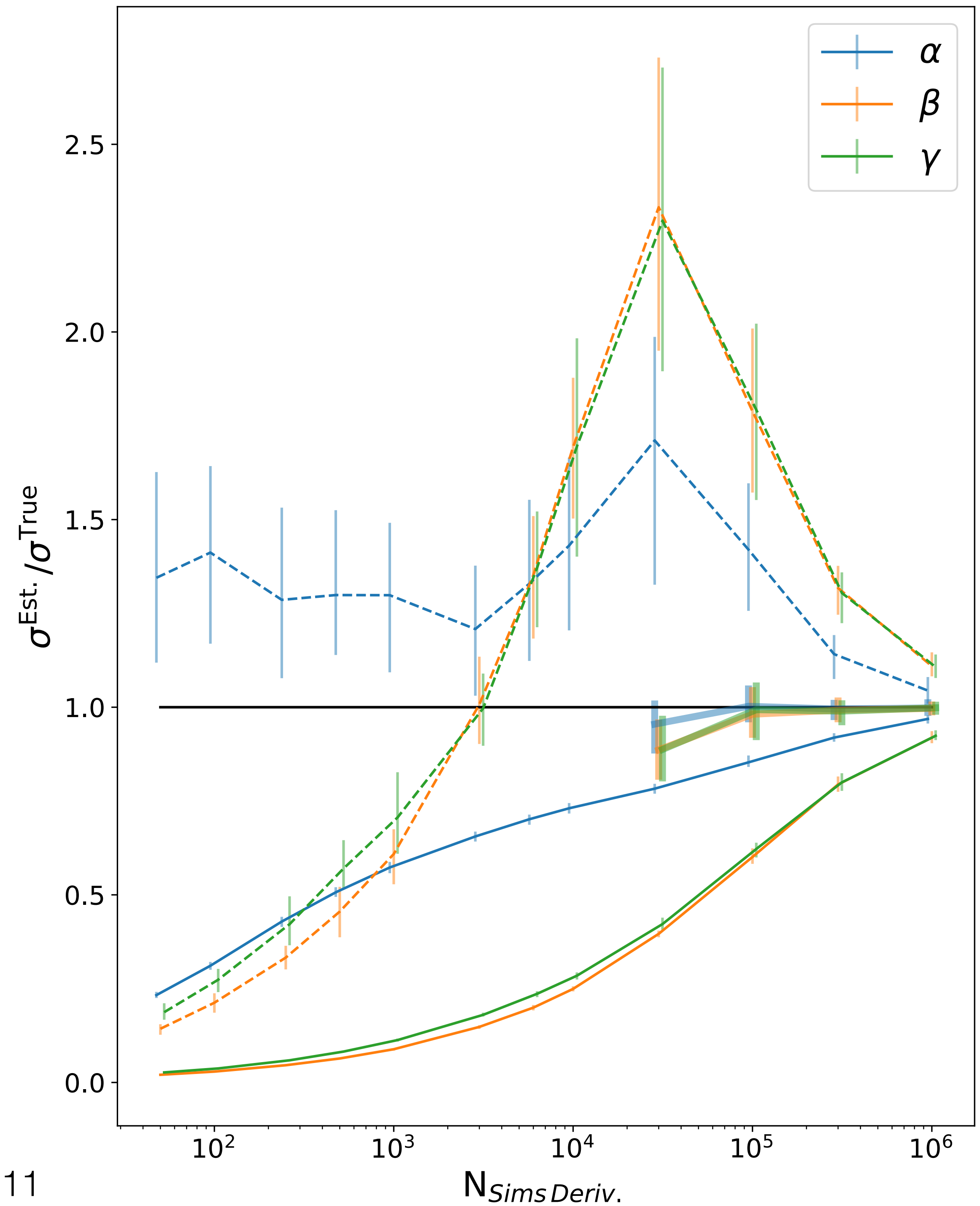
Gaussian model
(use "seed" matching to in the derivatives)



Can we do even better?

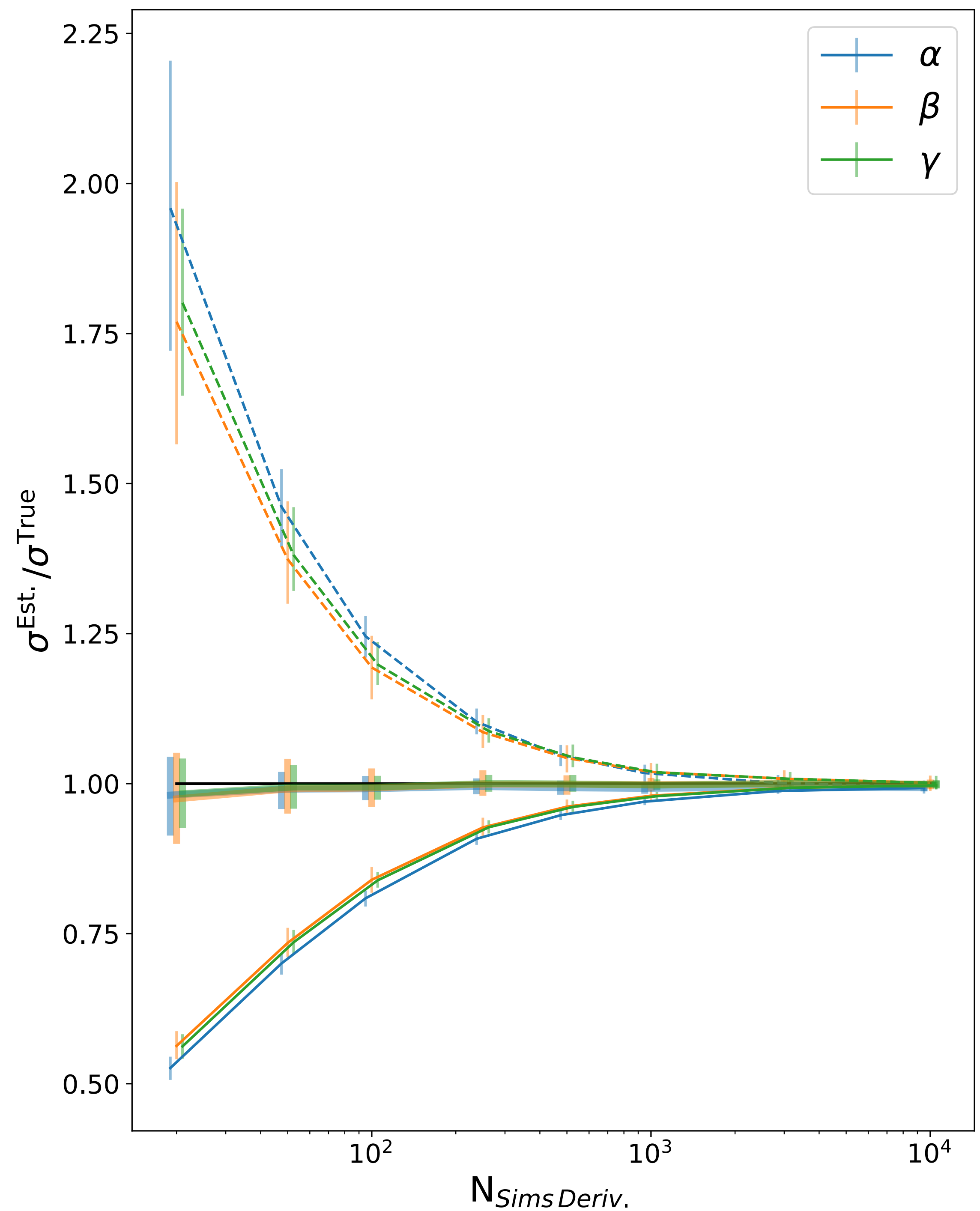
Example forecast constraining power for two 3 parameter toy models

Poisson model



Gaussian model

(use "seed" matching to in the derivatives)

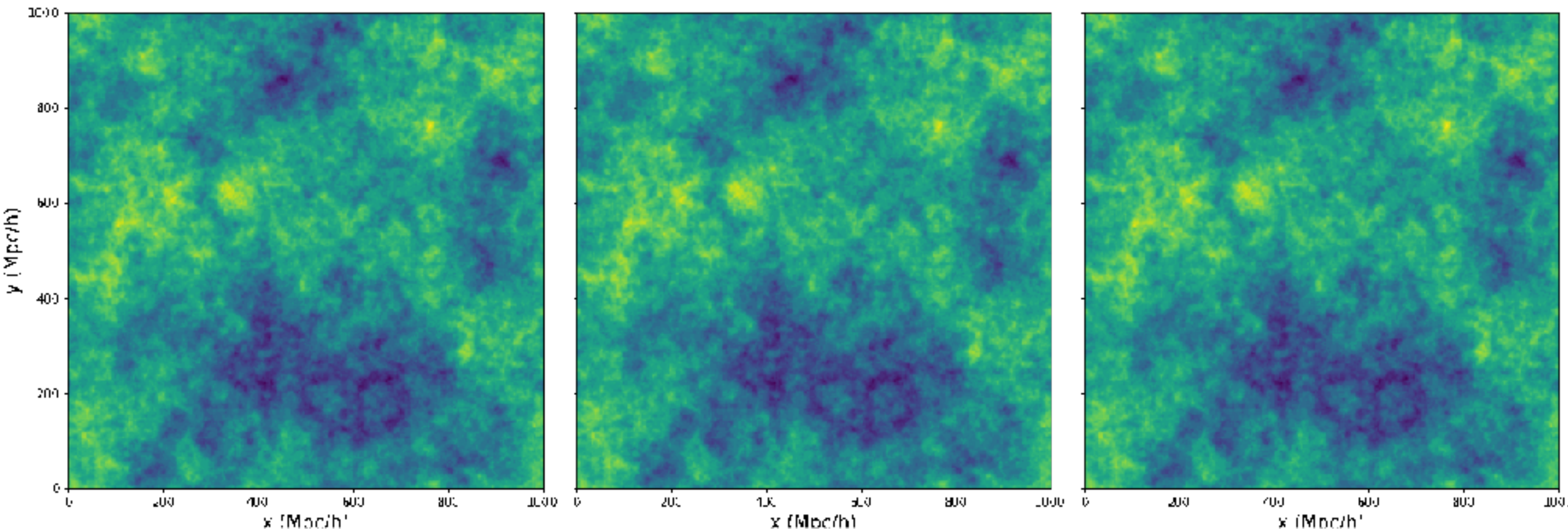


The search for primordial non-Gaussianity

The what and why of PNG

- Primordial non-Gaussianity is a catch-all for cases when the primordial anisotropies deviate from a Gaussian distribution!

Local-type **Equilateral-type** **Orthogonal-type**



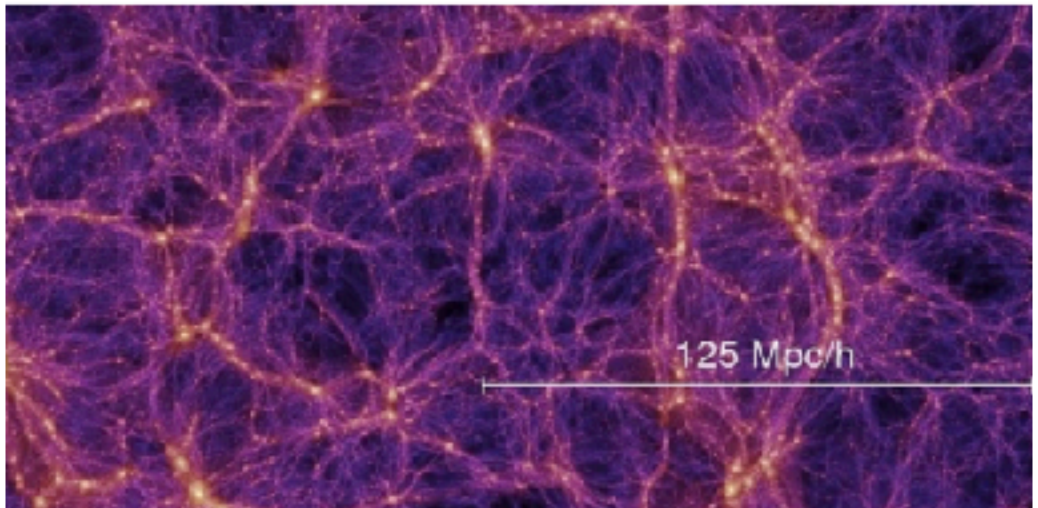
- Different types of PNG encode different, unique aspects of the physics of the primordial universe!

Fourier space: correlation functions

Higher-order correlation functions:

$$\langle \delta_{\vec{k}_1}\delta_{\vec{k}_2}\delta_{\vec{k}_3}\rangle \equiv \delta_D(\vec{k}_1+\vec{k}_2+\vec{k}_3)\, B(k_1,k_2,k_3)$$ **bispectrum**

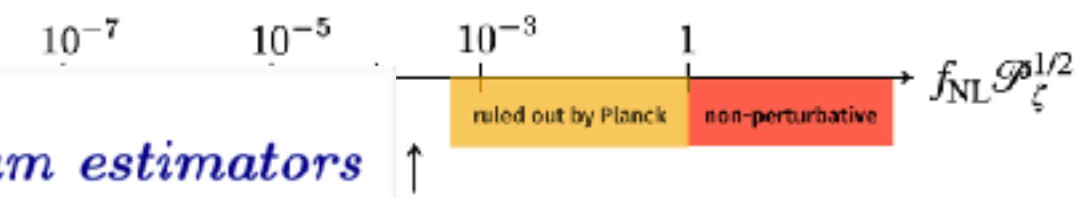
$$\langle \delta_{\vec{k}_1}\delta_{\vec{k}_2}\delta_{\vec{k}_3}\delta_{\vec{k}_4}\rangle \equiv \delta_D(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4)\, T(\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4)$$ **trispectrum**



Th
to

Primordial non-Gaussianity

Bispectrum $\frac{\langle\zeta\zeta\zeta\rangle}{\langle\zeta\zeta\rangle^2} \propto f_{\rm NL}$ Equilateral configuration (e.g. higher-derivative kinetic terms)



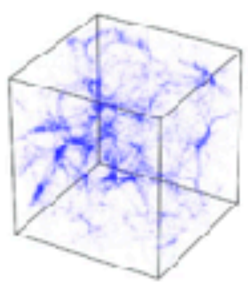
ure Cosmic Microwave Background experiments

surveys (SPHEREx, LSST, ...)

ing? mu-distortions?

$= -26 \pm 47$ Planck (2018)

The power spectrum and bispectrum estimators



$k_f = \frac{2\pi}{L} \to$ Fundamental mode
$k \to$ Center of the shell
$\Delta k \to$ Width of the bin

Power spectrum estimator

$$\hat{P}(k) = \frac{k_f^3}{N_k}\sum_{\boldsymbol{q}\in k}\delta(\boldsymbol{q})\delta(-\boldsymbol{q})$$

Number of modes in the shell

$$N_k \equiv \sum_{\boldsymbol{q}\in k} \simeq \frac{1}{k_f^3}\int_{k-\Delta k/2}^{k+\Delta k/2} dq\, q^2 d\Omega = 4\pi\frac{k^2\Delta k}{k_f^3} + \mathcal{O}(\Delta k^3)$$

Bispectrum estimator

$$\hat{B}(k_1,k_2,k_3) = \frac{k_f^3}{N_t(k_1,k_2,k_3)}\sum_{\boldsymbol{q}_1\in k_1}\sum_{\boldsymbol{q}_2\in k_2}\sum_{\boldsymbol{q}_3\in k_3}\delta_K(\boldsymbol{q}_{123})\,\delta(\boldsymbol{q}_1)\delta(\boldsymbol{q}_2)\delta(\boldsymbol{q}_3)$$

Number of fundamental triangles

$$N_t(k_1,k_2,k_3) \equiv \sum_{\boldsymbol{q}_1\in k_1}\sum_{\boldsymbol{q}_2\in k_2}\sum_{\boldsymbol{q}_3\in k_3}\delta_K(\boldsymbol{q}_{123})$$

R.Scoccimarro, M.Zaldarriaga, and L.Hui (1999)

E.Sefusatti, M.Crocce, Martin, S.Pueblas, R.Scoccimarro

What is the bispectra?

- Typical to work in Fourier or Harmonic space:

$$\Delta T(\mathbf{n}) = \sum Y_{\ell m}(\mathbf{n})\, a_{\ell m} \quad \text{or} \quad \delta(\mathbf{x}) = \int \frac{\mathrm{d}^3 k}{(2pi)^3} \delta(\mathbf{k}) e^{i\mathbf{x}\cdot\mathbf{k}}$$

- If the field is purely Gaussian then the fluctuations can be fully described by the power-spectrum

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = \delta_{mm'} \delta_{\ell\ell'} C_\ell \quad \text{or} \quad \langle \delta(\mathbf{k}) \delta^*(\mathbf{k}') \rangle = (2\pi)^3 \delta^{(3)}(\mathbf{k} + \mathbf{k}) P(k)$$

- The bispectrum is the harmonic equivalent of the three point function

- For a homogeneous and isotropic universe it has the form:

$$\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle \sim B(\ell_1, \ell_2, \ell_3) \quad \text{or} \quad \langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle \sim B(k_1, k_2, k_3)$$

- Vanishes for Gaussian fluctuations

Komatsu (2002)
Spergel and Goldberg (1999)

Why study the bispectra?

- More formally, if the anisotropies are Gaussian

$$P(a) = \frac{1}{(2\pi)^{N_{\rm harm}/2}|C|^{1/2}} \exp\left[-\frac{1}{2}\sum_{lm}\sum_{l'm'} a^*_{lm}(C^{-1})_{lm,l'm'}a_{l'm'}\right],$$

- Perform an Edgeworth expansion around the Gaussian and the bispectrum is the leading order correction:

$$P(a) = \frac{1}{(2\pi)^{N_{\rm harm}/2}|C|^{1/2}} \exp\left[-\frac{1}{2}\sum_{lm}\sum_{l'm'} a^*_{lm}(C^{-1})_{lm,l'm'}a_{l'm'}\right]$$
$$\times \left\{1 + \frac{1}{6}\sum_{\text{all } l_i m_j} \langle a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\rangle \left[(C^{-1}a)_{l_1m_1}(C^{-1}a)_{l_2m_2}(C^{-1}a)_{l_3m_3}\right.\right.$$
$$\left.\left.-3(C^{-1})_{l_1m_1,l_2m_2}(C^{-1}a)_{l_3m_3}\right]\right\}.$$

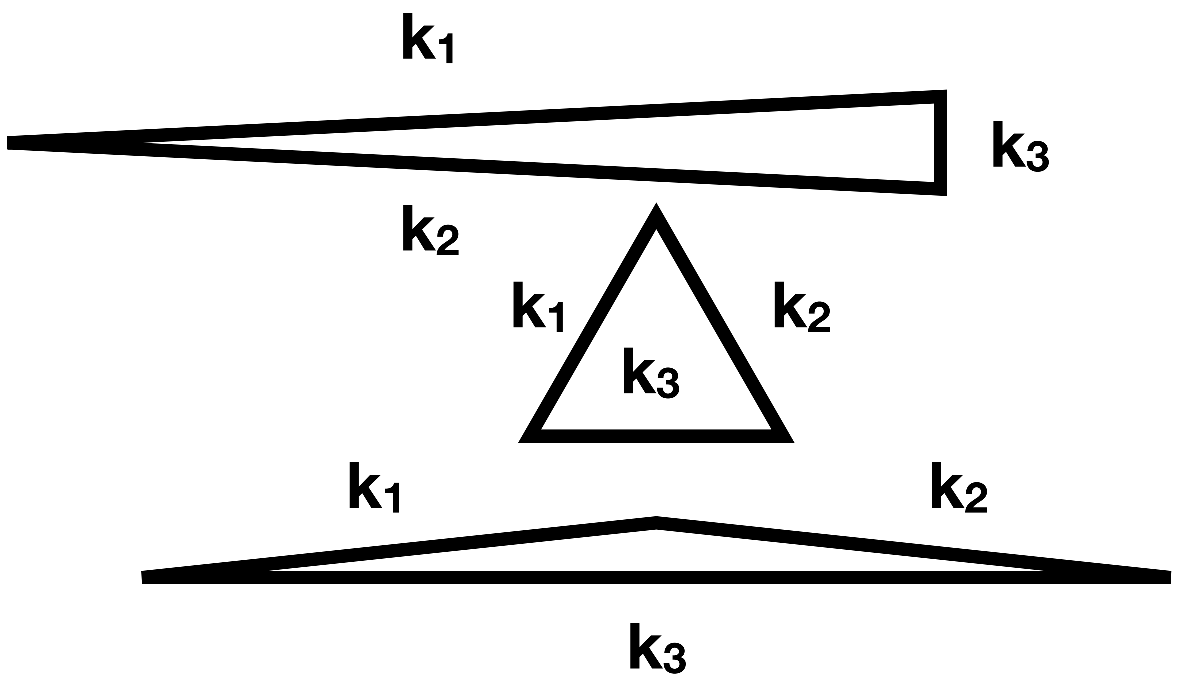
See e.g. Komatsu 2010 for a review

Why study the bispectra?

- Unique window into physics of early universe
    - Highly complementary to B mode searches
- Theoretical models of inflation give us predictions

$$\langle \zeta(k_1)\zeta(k_2)\zeta(k_3)\rangle \propto \ shape \times f_{NL}$$

- Three commonly studied shapes:
    - Local - Multi-field inflation?
    - Orthogonal / Equilateral - $c_s \neq 1$?
    - Folded - Non-bunch Davies initial conditions?



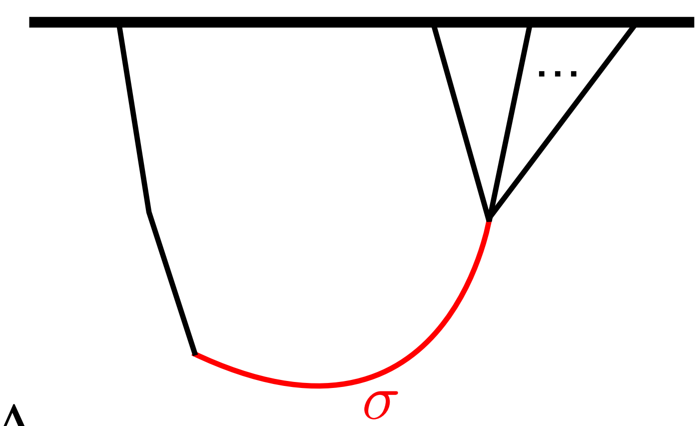
See e.g. Chen (2010) for a review

Interesting correlation regimes

- Correlations produced by inflation contact terms most important when (given de Sitter symmetries):

$$F(k_1 \sim k_2 \ldots k_N) \sim \frac{1}{k^{3(N-1)}}$$

- Correlations induced by mediators:

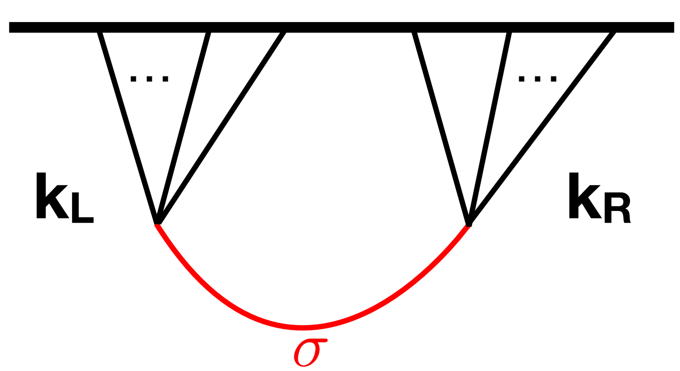


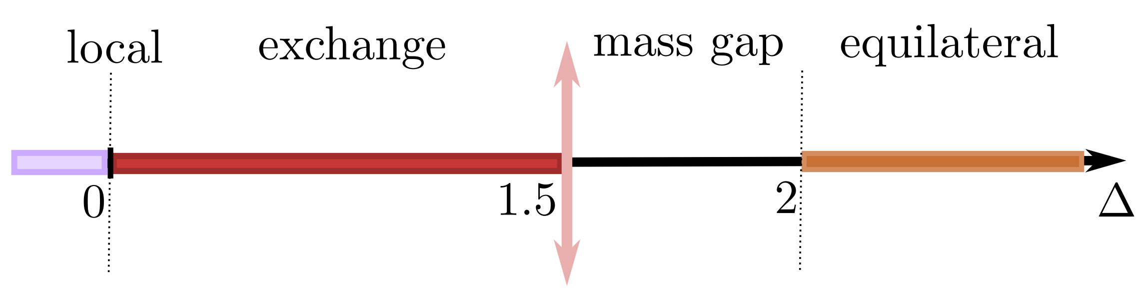
“Squeezed”

$$F(k_1 <<< k_2 \ldots k_N) \sim \frac{1}{k_1^3 k^{3(N-2)}} \left(\frac{k_1}{k}\right)^{\Delta_s}$$

“Collapsed”

$$F(k_I <<< k_L \; k_R) \sim \frac{1}{k_I^3 k_R^{3(M-2)} k_L^{3(N-M-1)}} \left(\frac{k_I^2}{k_R k_L}\right)^{\Delta_c}$$





- Δ encodes interesting new physics!

How to measure non-Gaussianity?

- To date measurements driven by CMB constraints as

$$a_{\ell m} \propto \zeta(\mathbf{k})$$

- Ideally measure every configuration

$$\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle \propto \langle \zeta(\mathbf{k}_1) \zeta(\mathbf{k}_2) \zeta(\mathbf{k}_3) \rangle$$

- Computational prohibitive and signal is weak

- Two broad solutions:

  - Compress the data: e.g. Binned bispectrum
    This averages over 'nearby' configurations

  - Compress the information: e.g. KSW or Modal estimators
    These compress the data to a set of template amplitudes.
    Crudely:

$$\hat{f}_{\mathrm{NL}} \propto \sum b^{\mathrm{theory}}_{\ell_1 \ell_2 \ell_3} a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3}$$

Komatsu, Spergel and Wandelt (2005)
Fergursson et al (2009)
Bucher et al (2013,2015)

Future Constraints

Shape ($\zeta\zeta\zeta$)	Current	SO constraint
Local	-0.9 ± 5.1	3
Equilateral	-26 ± 47	24
Orthogonal	-38 ± 23	13

For context, when $f_{NL} \sim 1$

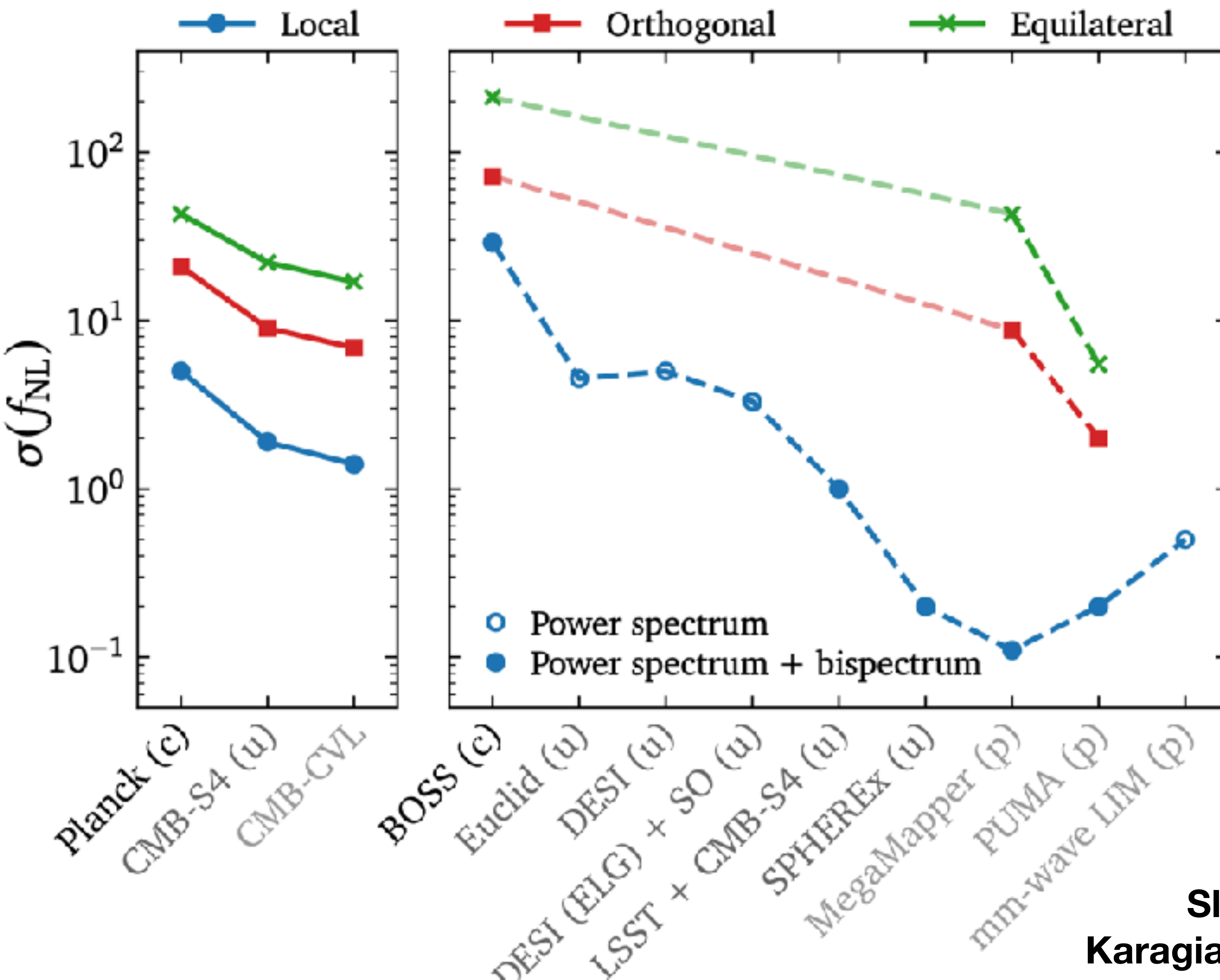
$$\frac{\langle \zeta(k_1)\zeta(k_2)\zeta(k_3) \rangle}{\sqrt{P_\zeta(k_1)P_\zeta(k_2)P_\zeta(k_3)}} \sim 10^{-5}$$

Planck Collaboration (2019)

The Simons Observatory Collaboration XI (2018)

PNG from LSS

Forecast constraints on PNG from LSS highlight challenges for non-local shapes



Slozar et al (2019)
Karagiannis et al (2019)
Coulton, Beringue and Meerburg (2020) ++

The challenge of measuring PNG

- In principle there are many more LSS modes

$$\sigma(f_{\mathrm{NL}}) \propto \frac{1}{\mathrm{Volume} \times k_{\mathrm{max}}^3}$$

- However for LSS measurements

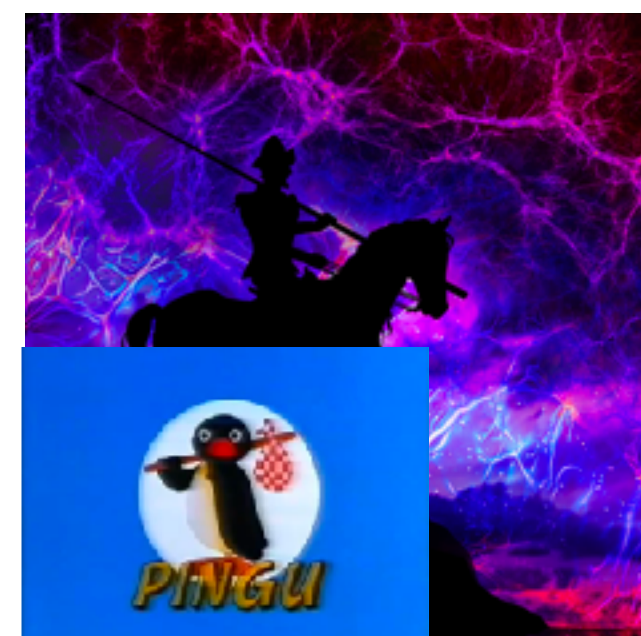
$$\delta_g \propto \zeta + \zeta^2 + \zeta^3 \ldots\ldots$$

- More problematically

$$\langle \delta_g(\mathbf{k}_1)\delta_g(\mathbf{k}_2)\delta_g(\mathbf{k}_3)\rangle_{\mathrm{primordial}} \sim \langle \delta_g(\mathbf{k}_1)\delta_g(\mathbf{k}_2)\delta_g(\mathbf{k}_3)\rangle_{\mathrm{evolution}}$$

Quijote -PrImordial-Non-GaUssianity

- Extension to the Quijote suite of simulations
  - Large suite of cosmological N-body sims.
  - 1 Gpc/h per side, 512^3 particles
  - Designed for use in ML applications and Fisher forecasting
  - Varies 7 cosmological parameters: $h, \Omega_m, \Omega_b, \sigma_8, w, \sum m_\nu, n_s$
- Added N-body simulations four types of primordial non-Gaussianity
  - $f_{\mathrm{NL}}^{\mathrm{Local}}, f_{\mathrm{NL}}^{\mathrm{Equil}}, f_{\mathrm{NL}}^{\mathrm{Orth-CMB}}, f_{\mathrm{NL}}^{\mathrm{Orth-LSS}}$
  - 1000 sims of each type



Villaescusa-Navarro et al (2020)

How to generate non-Gaussian ICs?

- Local-type PNG is so named as it corresponds to:

$$\phi(\mathbf{x}) = \phi^G(\mathbf{x}) + f_{\rm NL}(\phi^G(\mathbf{x})^2 - < \phi^G(\mathbf{x})^2 >)$$

- Most generally we can generate a bispectrum via

$$\phi(\mathbf{x}) = \phi^G(\mathbf{x}) + f_{\rm NL} \int \frac{d^3k_1 d_2^3}{(2\pi)^6} \delta^{(3)}(\mathbf{k_1} + \mathbf{k_2} + \mathbf{k}) B(\mathbf{k_1}, \mathbf{k_2}, \mathbf{k}) \frac{\phi^G(-\mathbf{k_1})}{P(k_2)} \frac{\phi^G(-\mathbf{k_2})}{P(k_3)}$$

- For separable bispectra :

$$\phi(\mathbf{x}) = \phi^G(\mathbf{x}) + f_{\rm NL} H[F[\phi^G(\mathbf{x})], G[\phi^G(\mathbf{x})]]$$

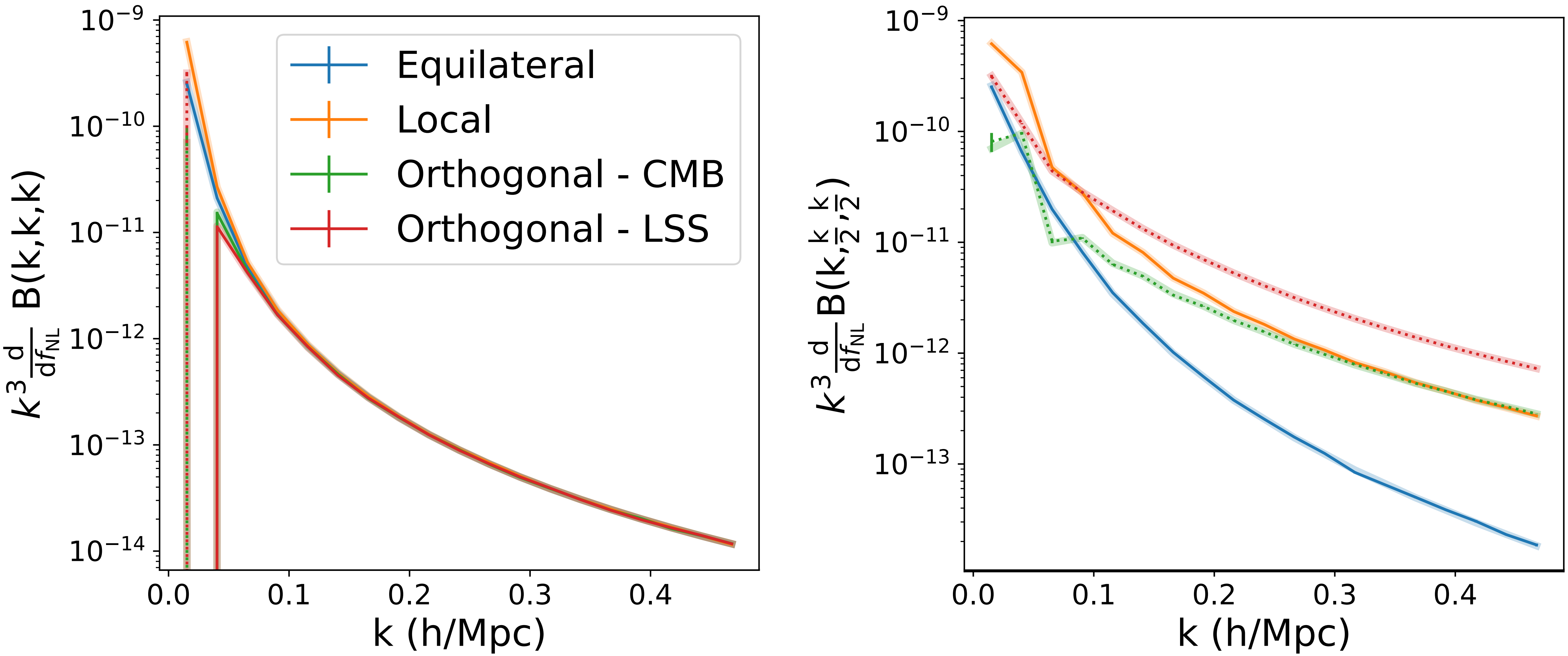
- Care must be taken that this does not lead to unphysical corrections to the power spectrum

- These primordial ICs are linearly evolved to redshift zero and then backscaled to z=100.

Ligouri et al (2003), Smith et al (2006), Scoccimarro et al (2012)

Tests of the ICs

Two slices of the primordial bispectrum in our simulations compared to theoretical predictions



Shaded bands are the theory and the solid (dotted) lines are the measured bispectrum

Our question:

- How much can we gain by using standard methods but going beyond the analytic regime ($k \approx 0.1\,\mathrm{h/Mpc}$ and z=0.0)?
- Consider both the matter density and the halo density fields
- Measure the power spectrum

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

  we use the halo monopole and quadrupole, the matter power spectrum and the matter-halo cross spectrum

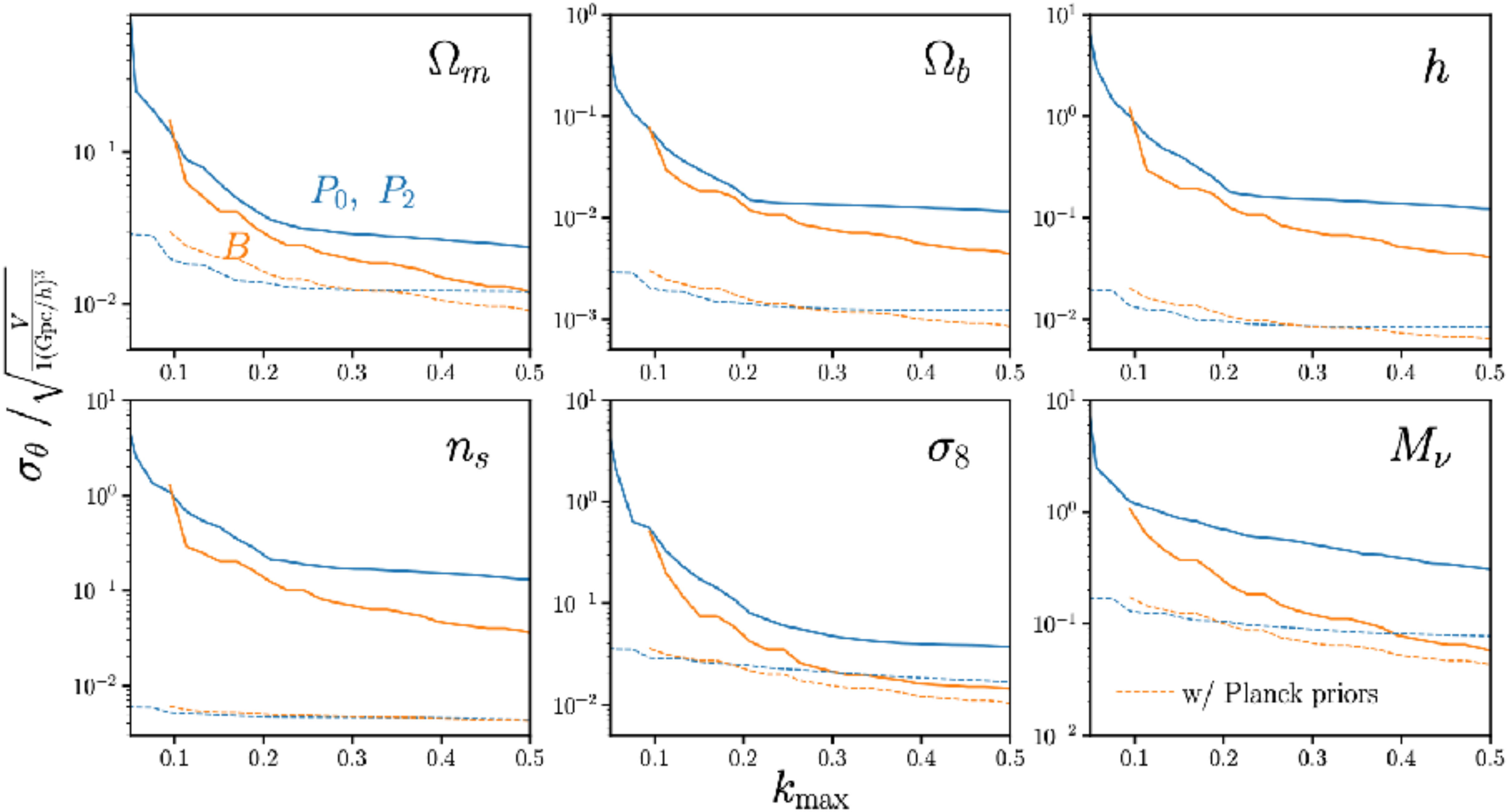
- Measure the bispectrum

$$\langle \delta(\mathbf{k}_1)\delta^(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle = (2\pi)^3\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)B(k_1,k_2,k_3)$$

- In the primordial space the bispectrum captures all the information on the types of PNG considered here

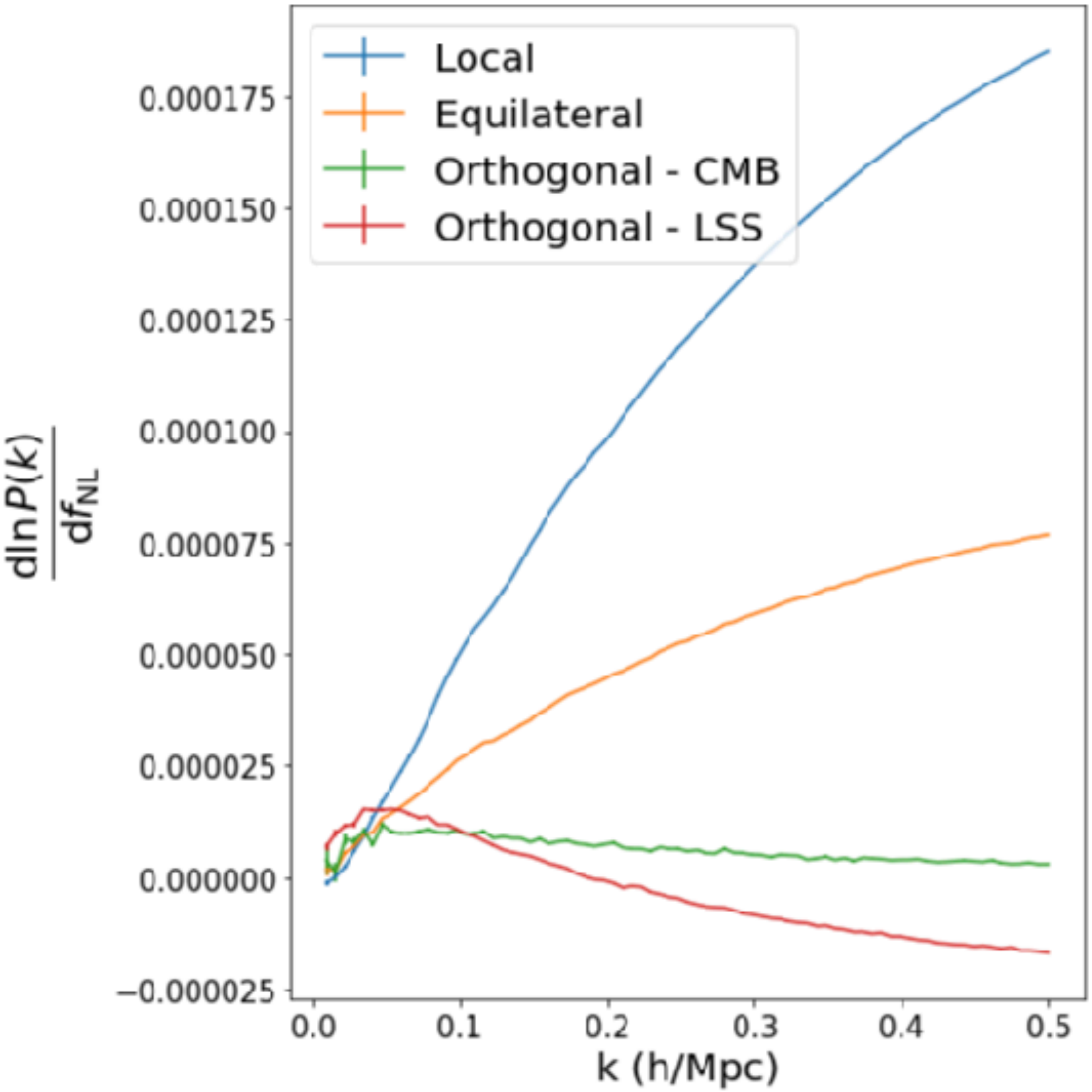
PNG from LSS?

Hahn et al (2020) showed there is lots of information beyond perturbative regime



How is the power spectrum impacted?

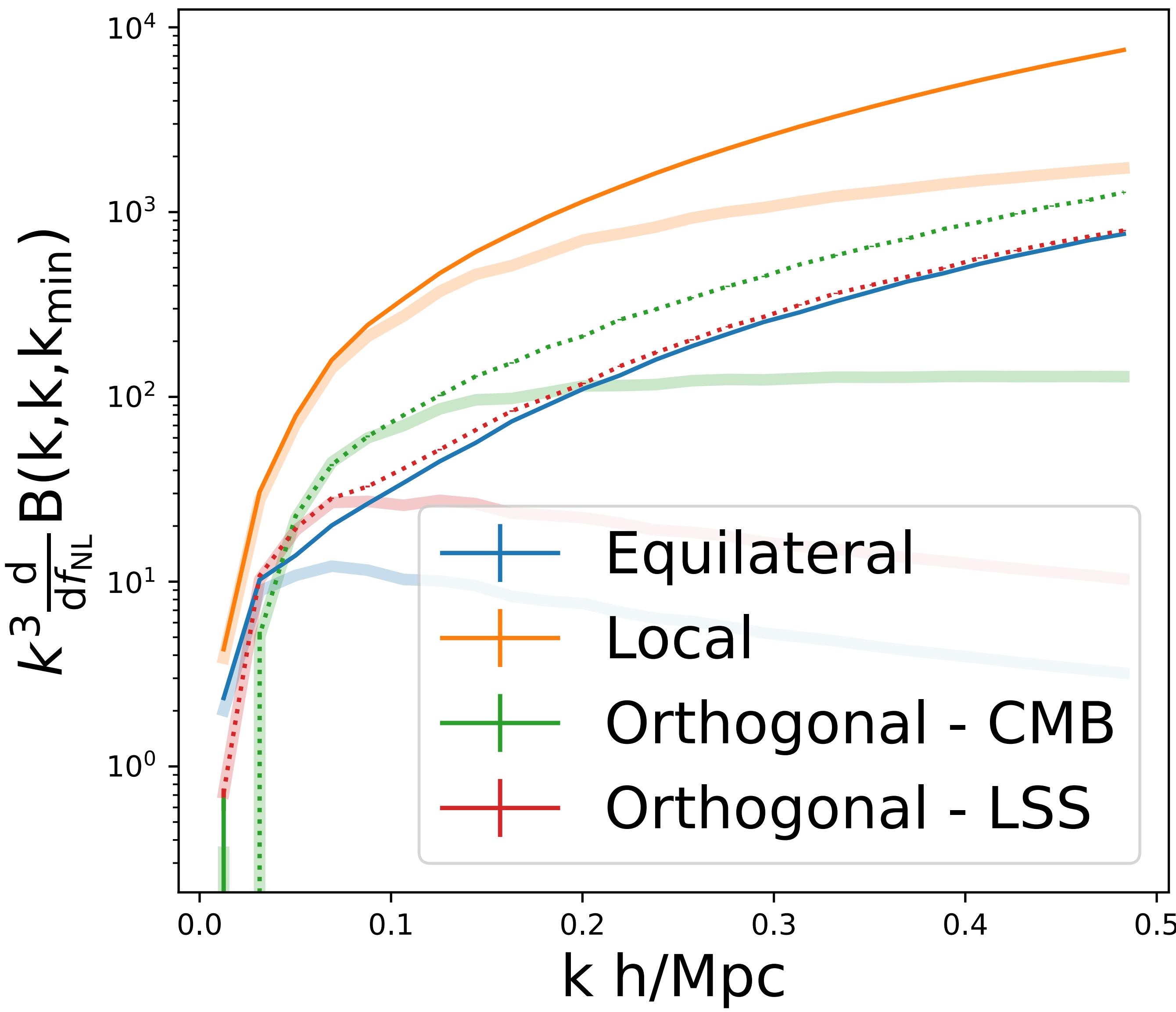
Matter power spectrum

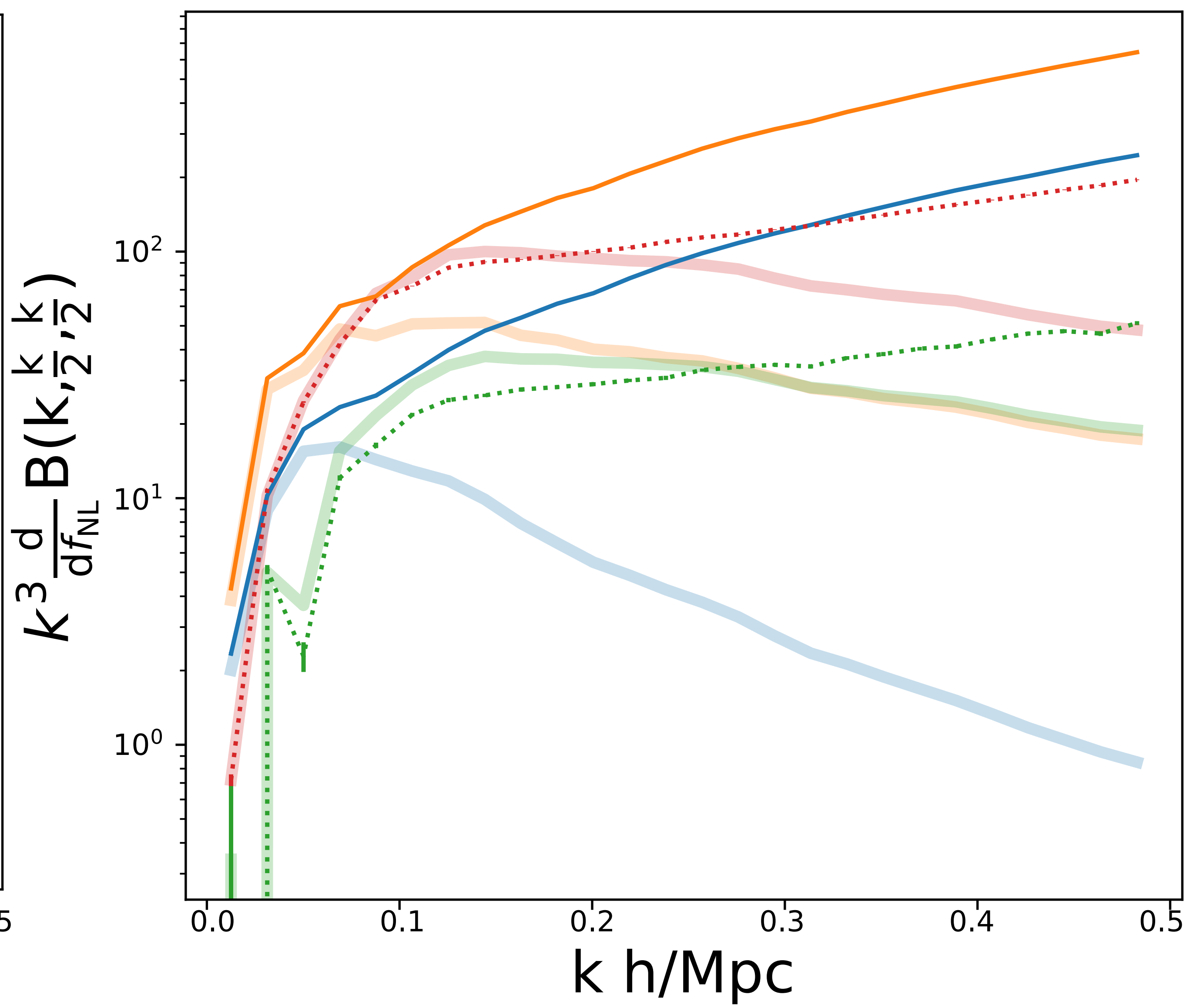


How is the matter bispectrum impacted?

Squeezed bispectrum of the matter field

Folded bispectrum of the matter field

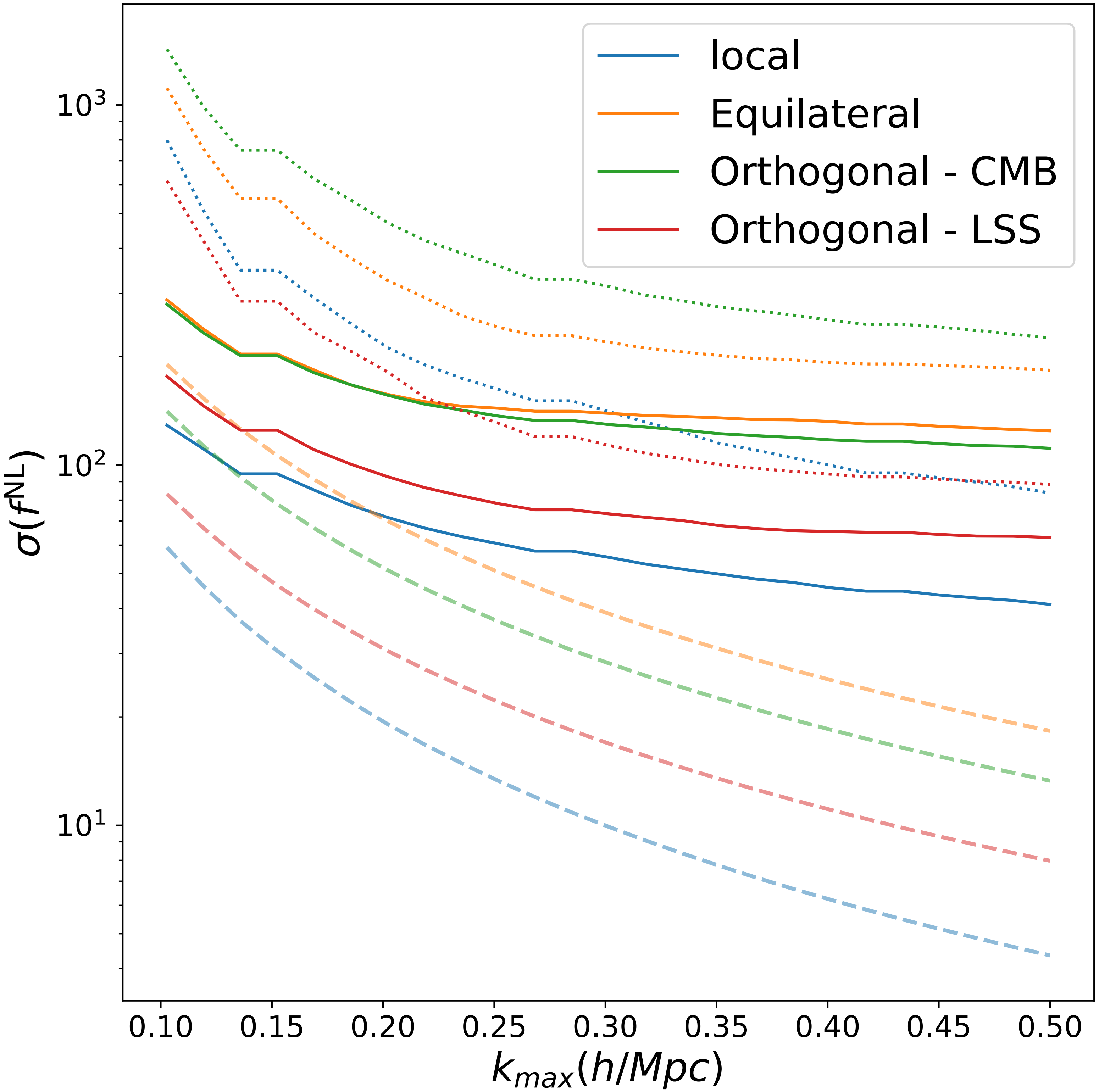




Shaded bands are the tree level theory

What is the information content in the matter field?

Bispectrum constraining power as a function of the maximum scale included in the analysis



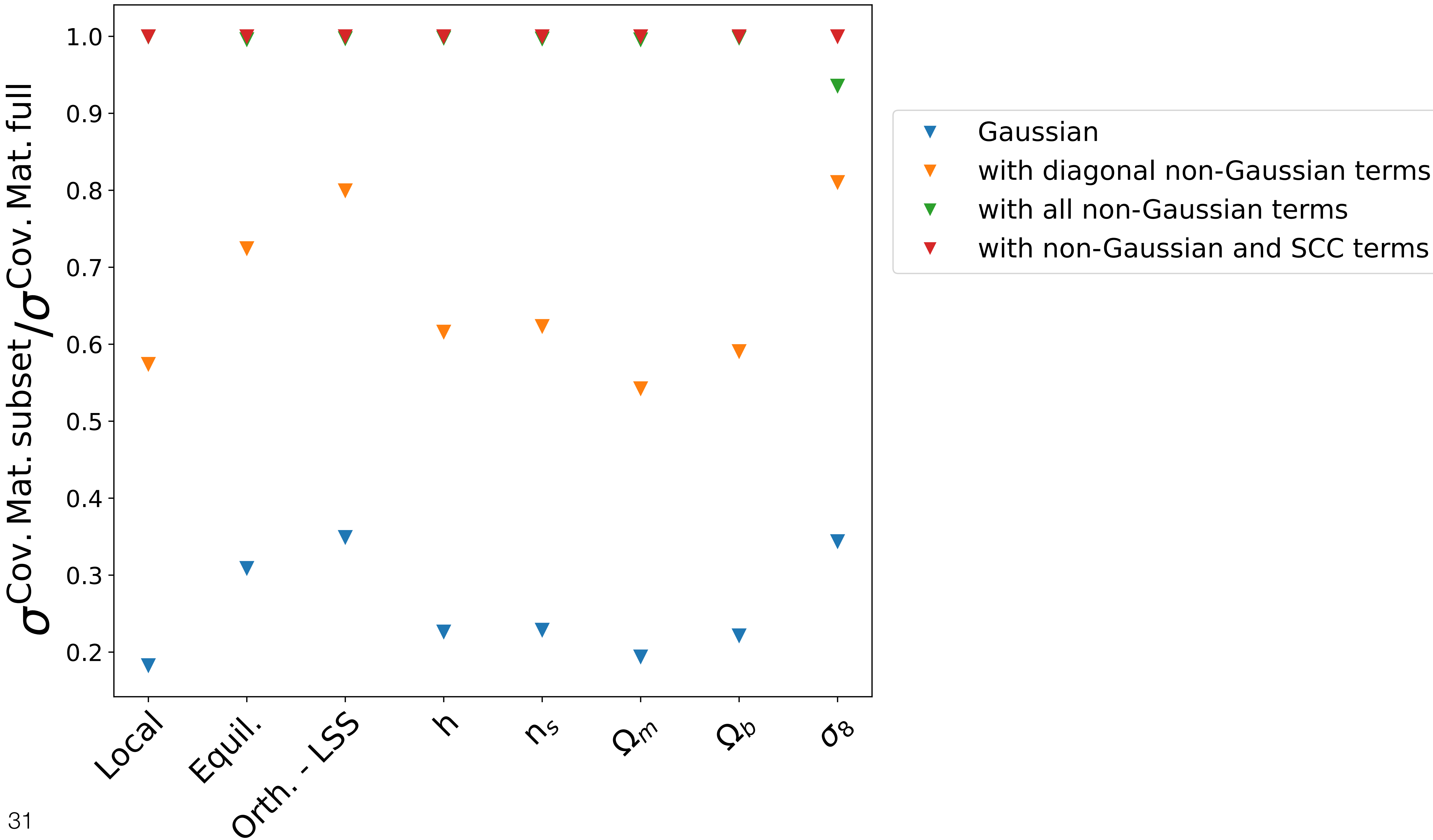
Dotted: including marginalization

Solid: without marginalizing cosmological params.

Dashed: primordial information content

The importance of accurate covariance modelling

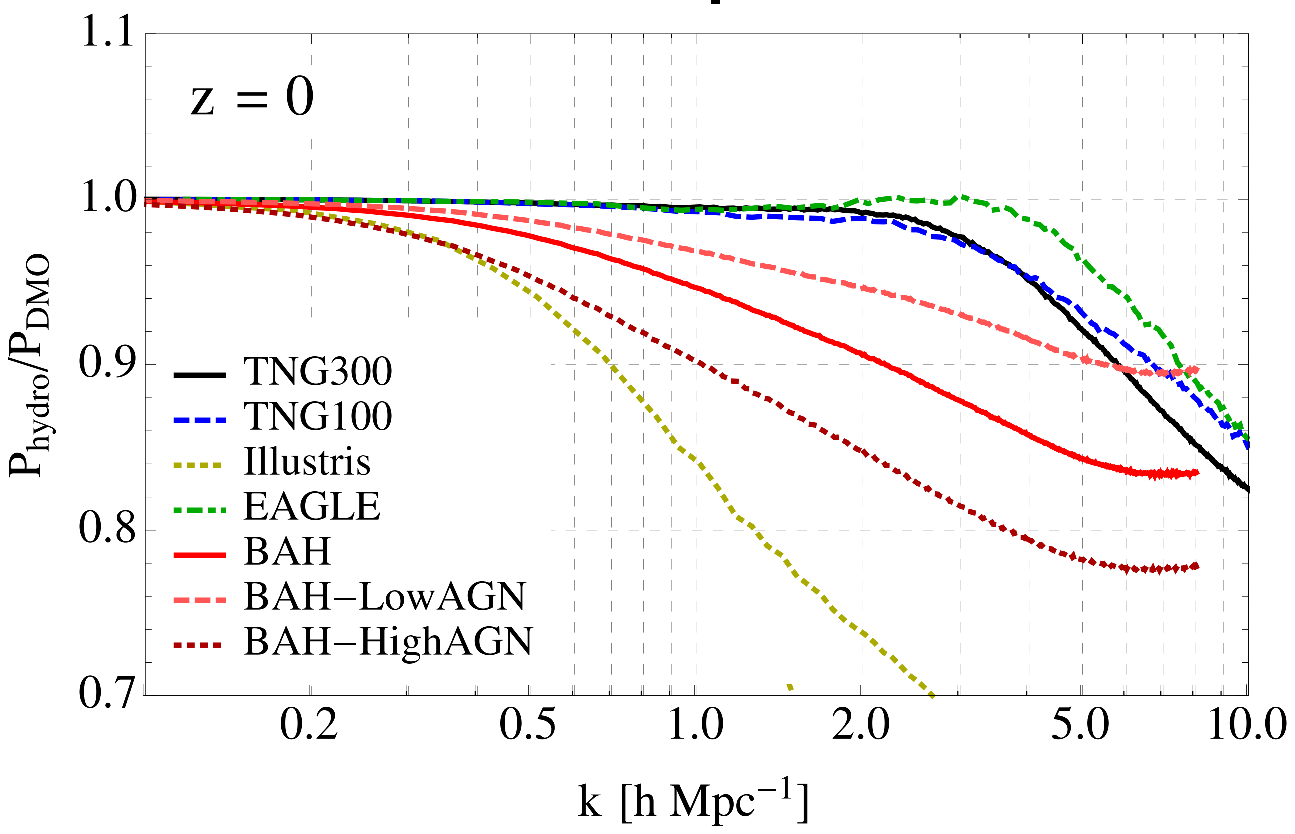
Forecasts constraints with the matter bispectrum when neglecting parts of the covariance matrix



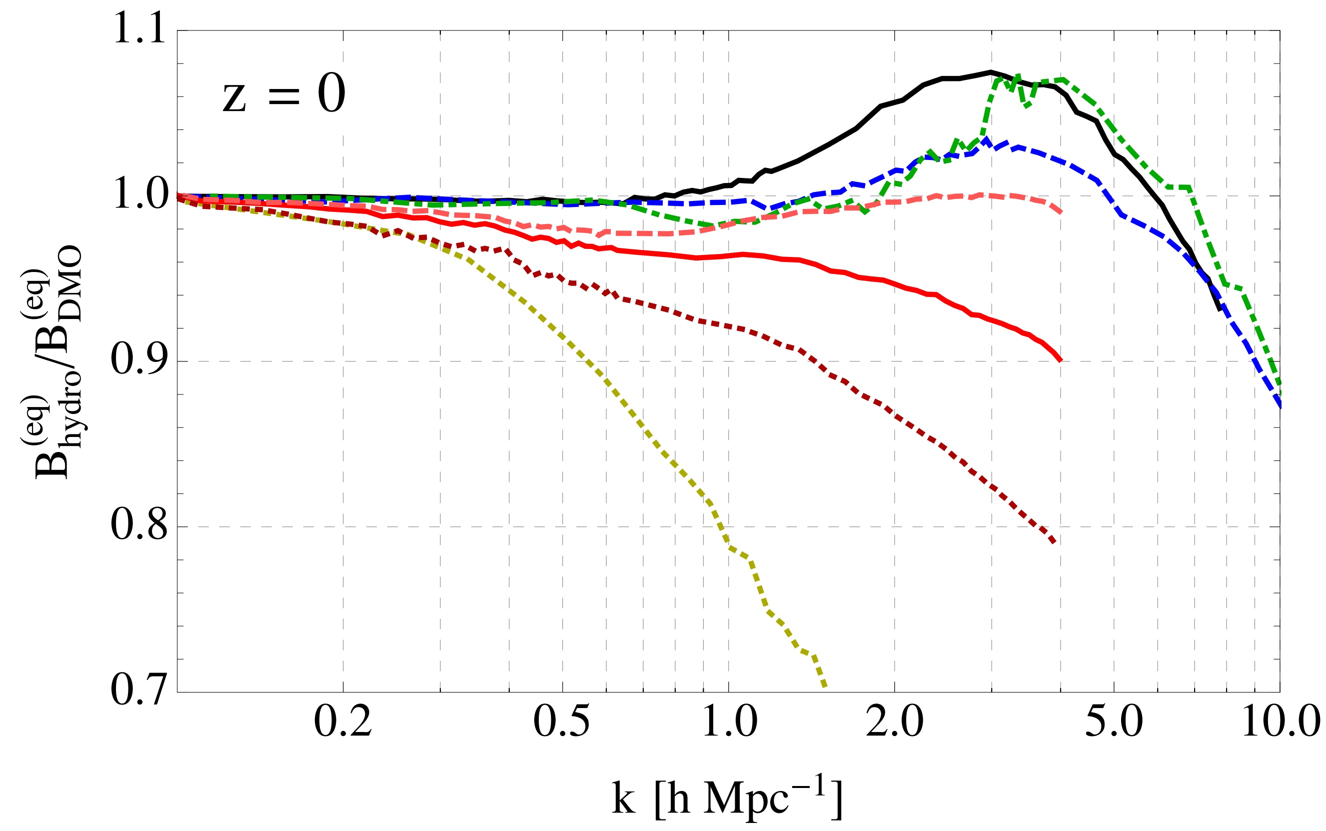
The impact of baryons

The ratio of hydrodynamical measurements to dark matter only for a range of state of the art models

Power spectrum



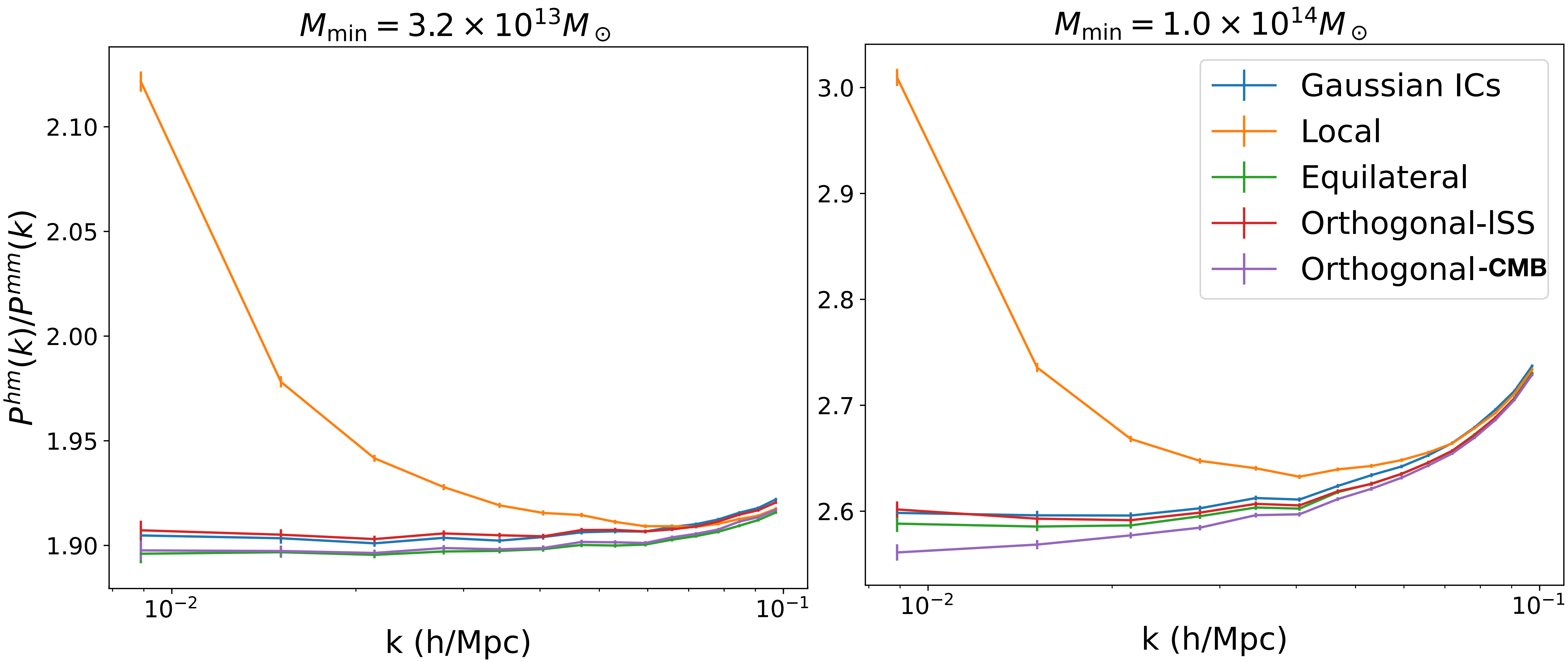
Bispectrum



Foreman, Coulton et al (2019)

Scale Dependent Bias

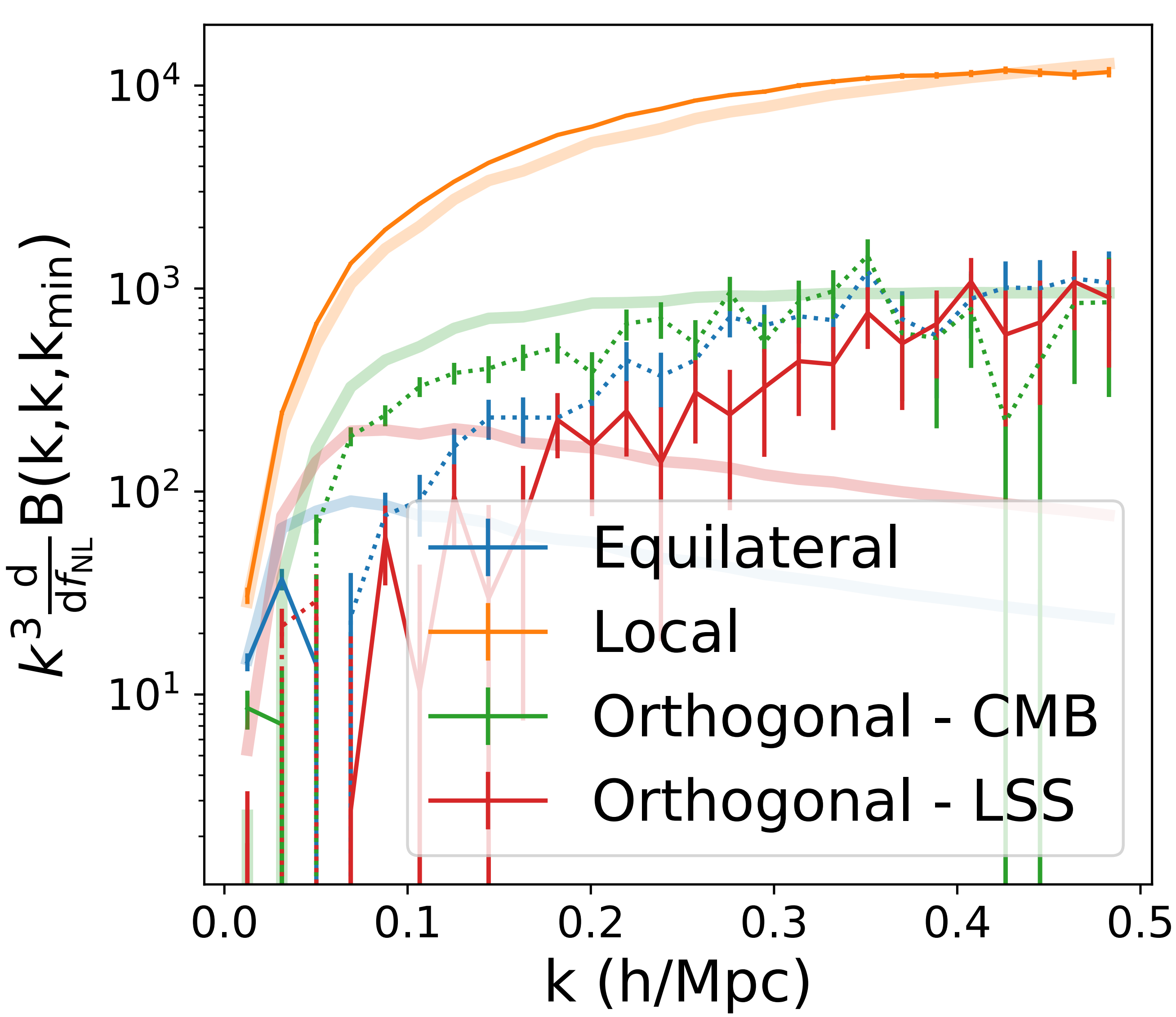
Halo power spectrum monopole for two difference tracers



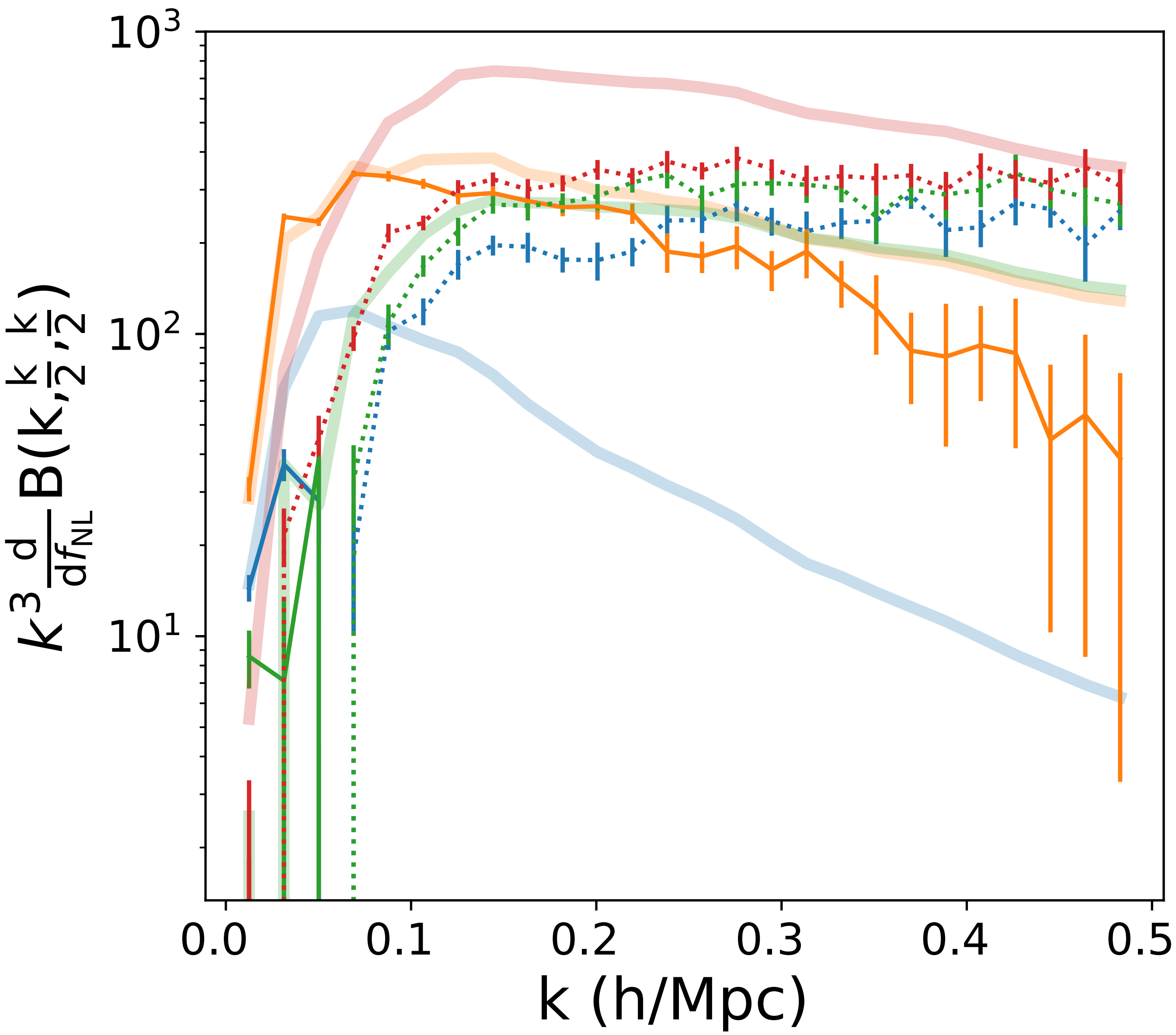
Schmidt et al (2010)

How is the halo bispectrum impacted?

Squeezed bispectrum of the halo field



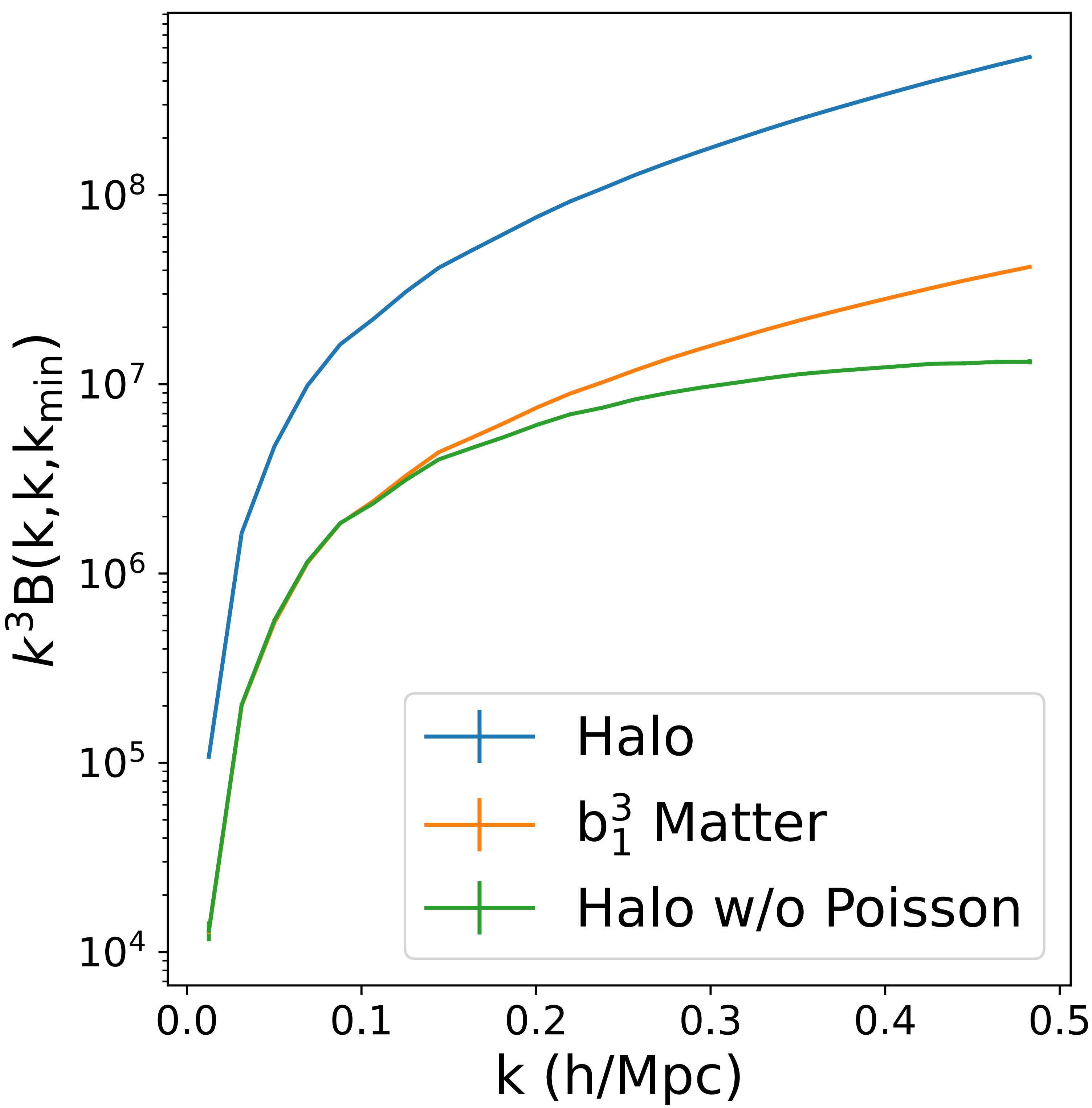
Folded bispectrum of the halo field



Shaded bands are the matter bispectrum, scaled by b_1^3

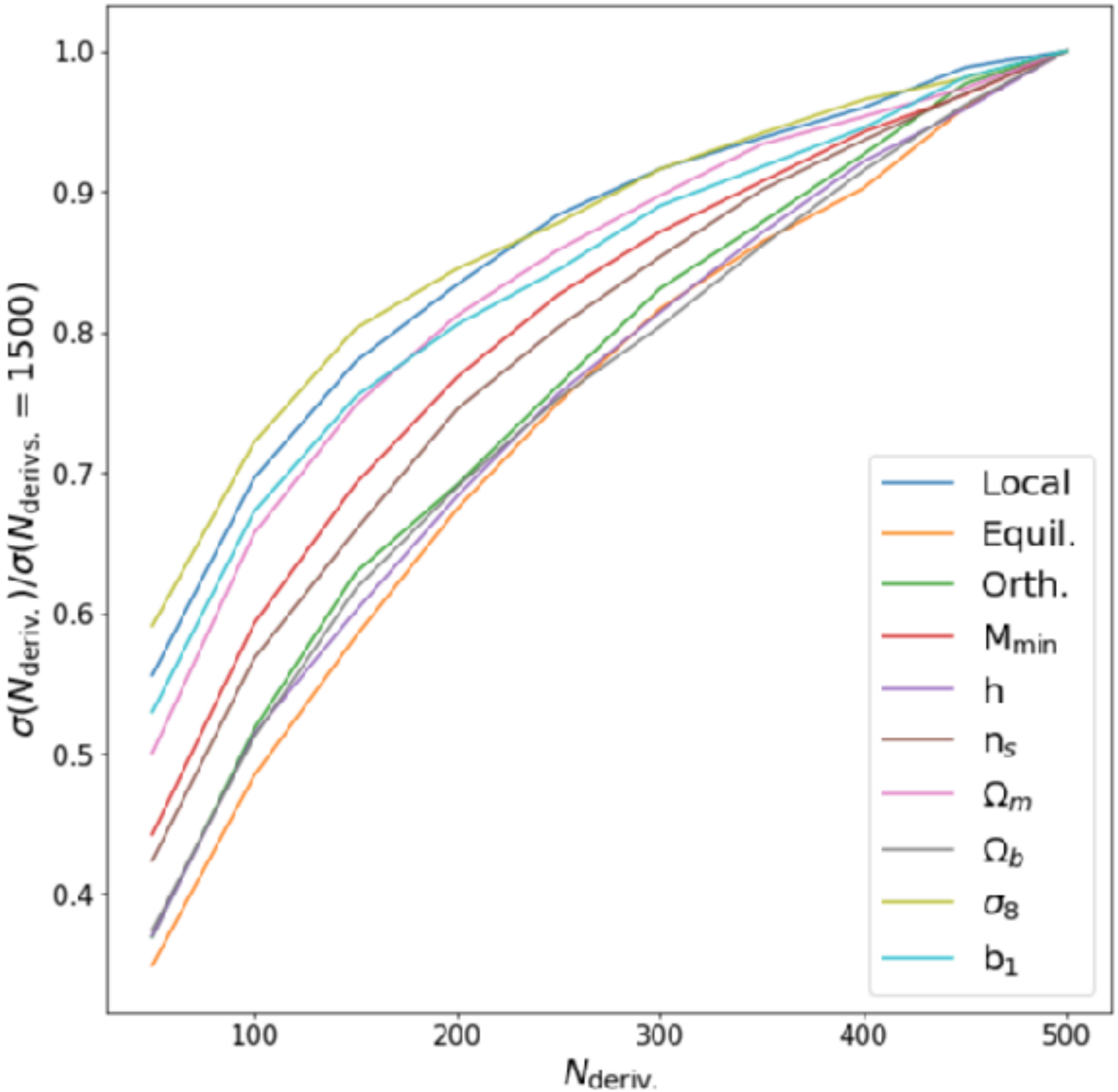
Why is the halo field so much noisier?

The squeezed slice of the bispectrum at the fiducial cosmology



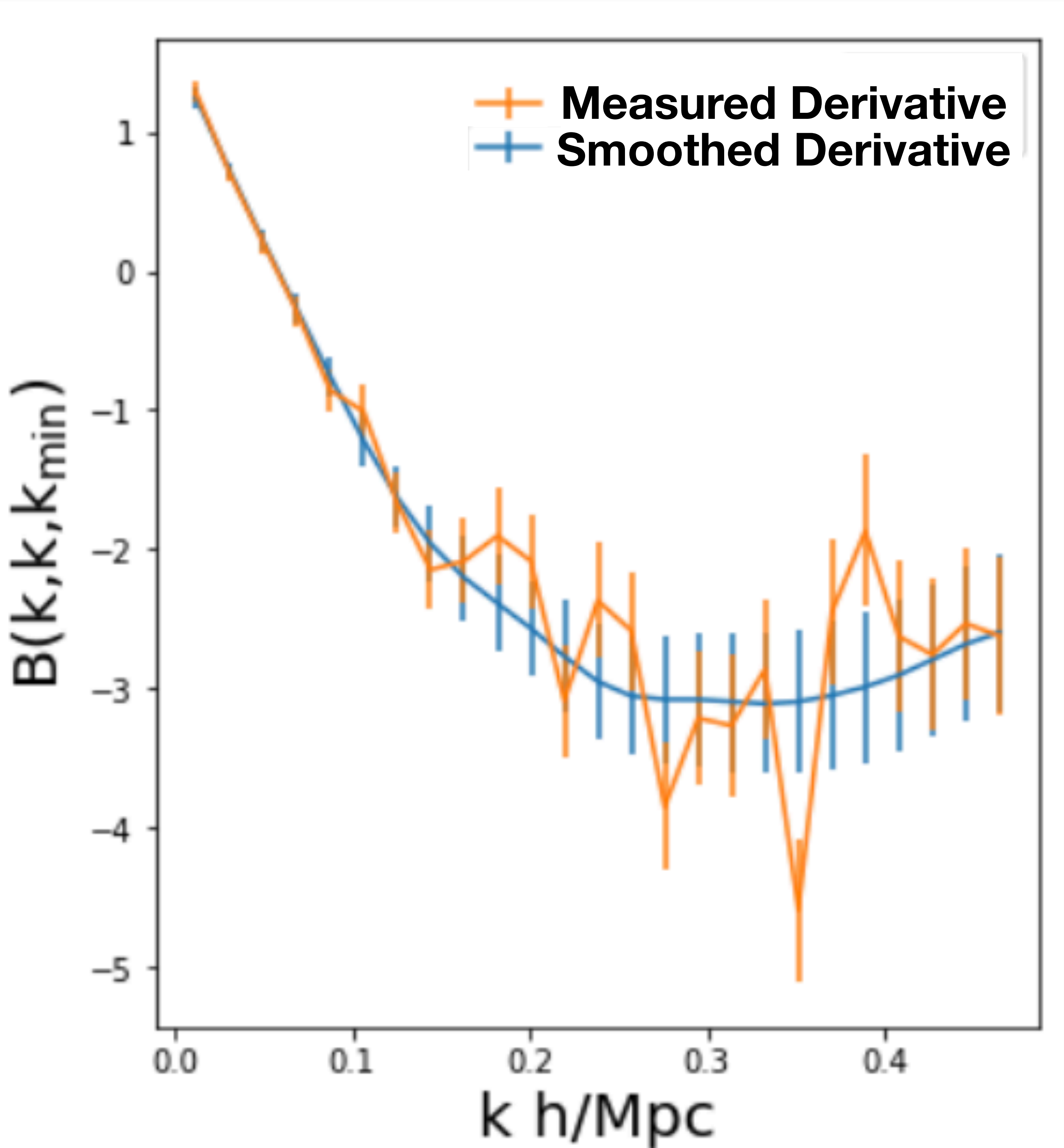
A complication?

Marginalized constraints as function of number of derivative simulations

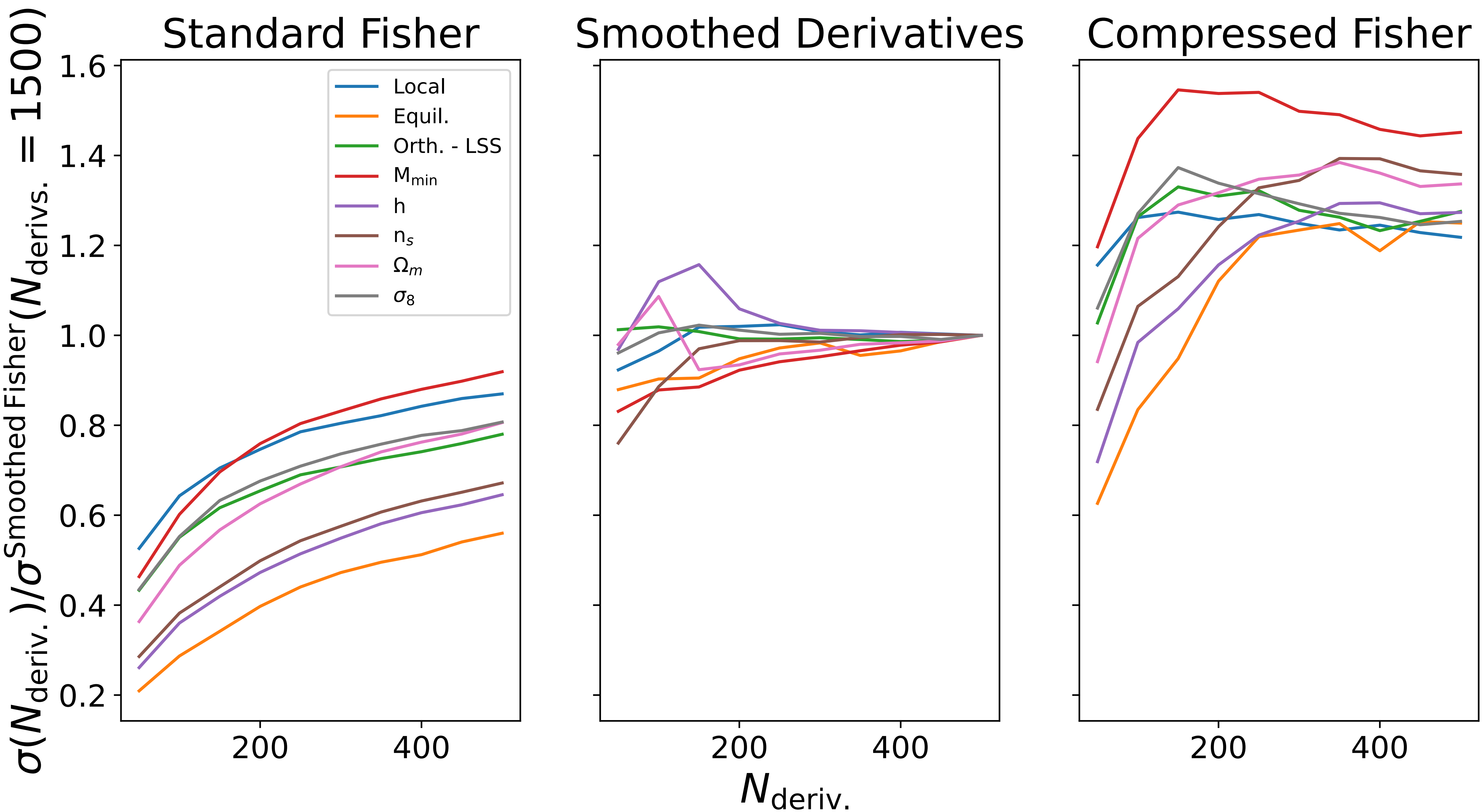


Smooth derivatives with a Gaussian Process

The response of the squeezed halo bispectrum to local primordial non-Gaussainity

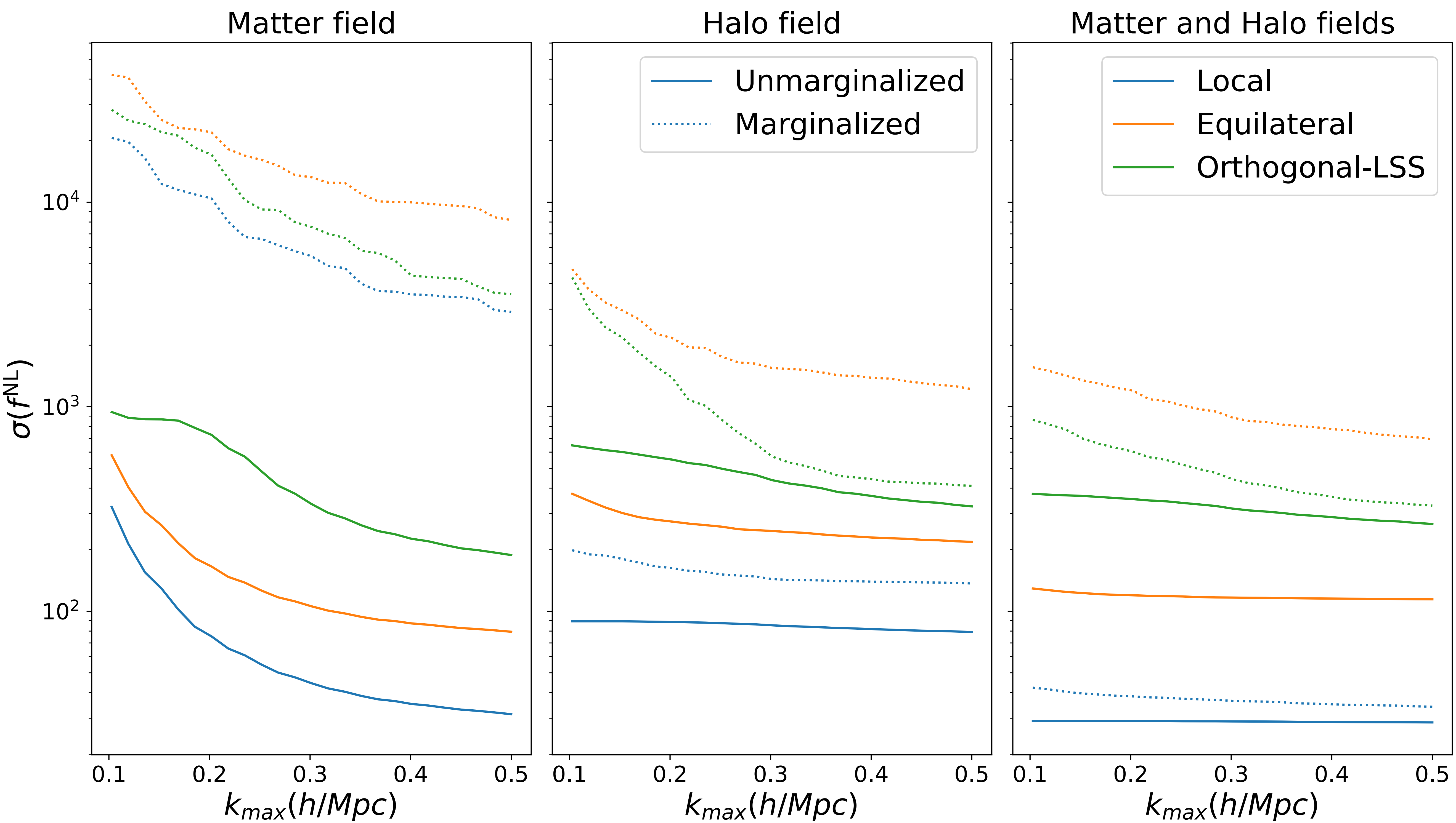


Convergence with mitigation methods

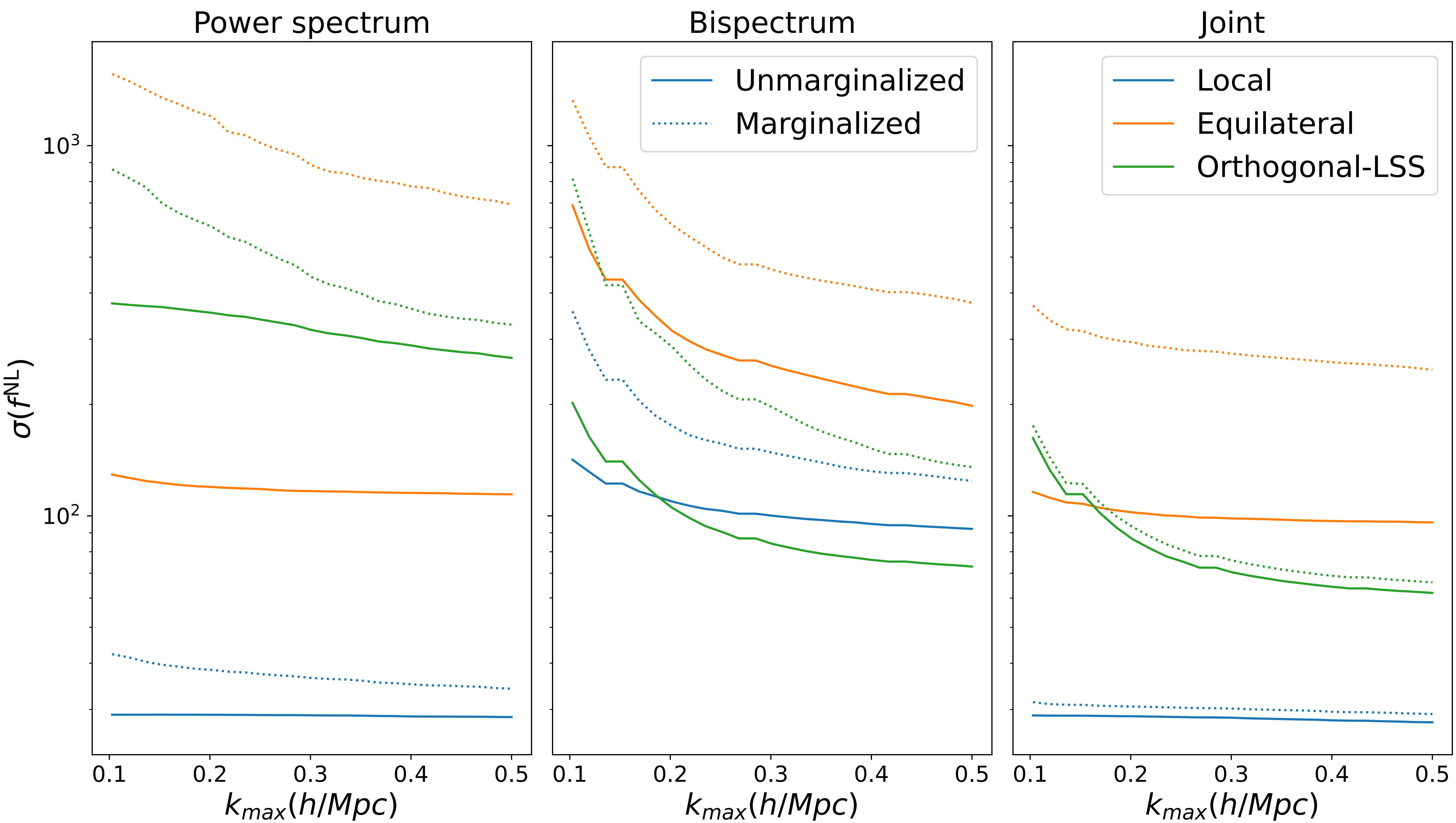


Power spectrum PNG constraints

Power spectrum constraints - the benefits of sample variance cancellation



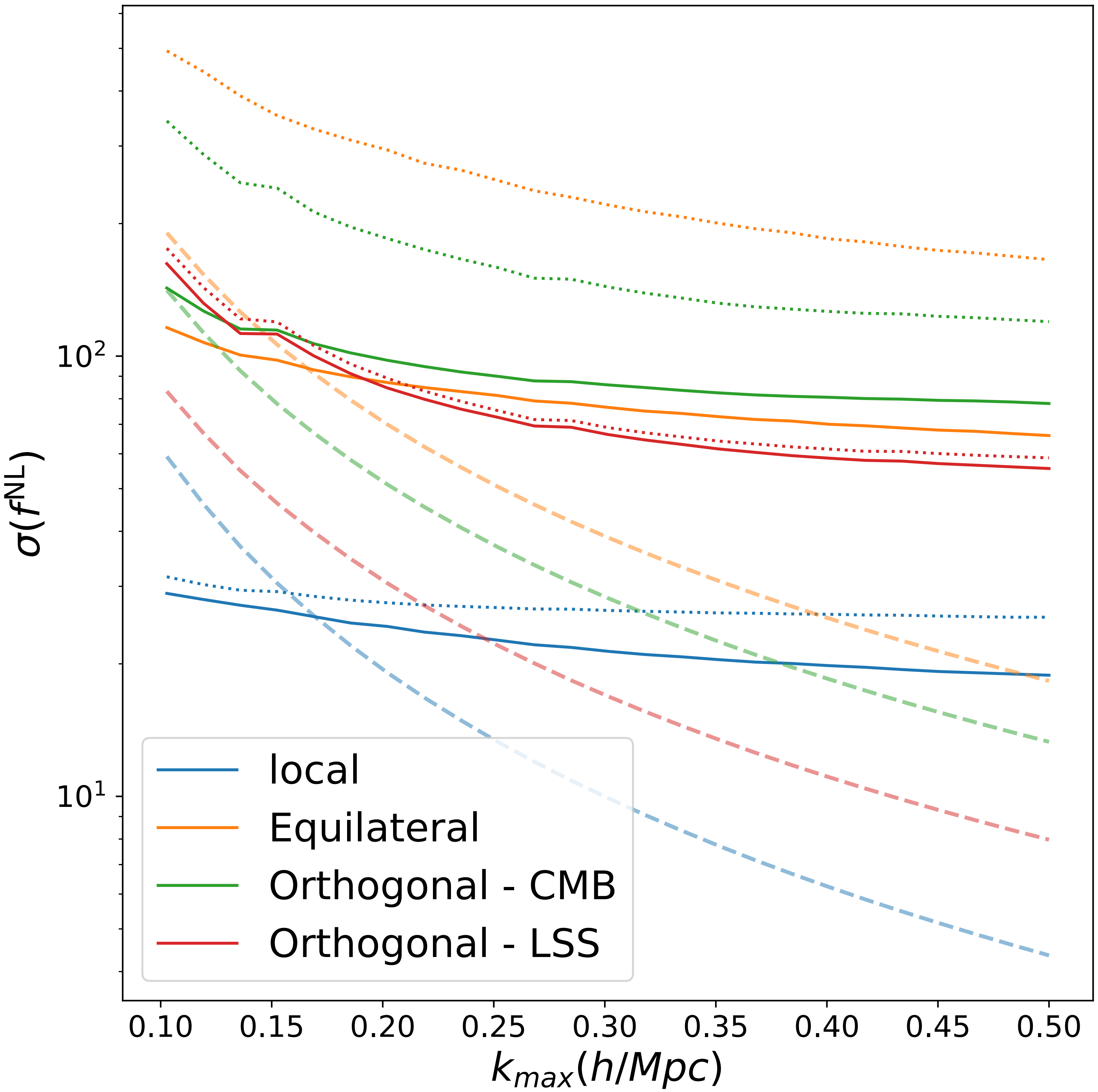
pnG constraints



Power spectrum = halo- monopole and quadrupole, matter power spectrum and halo-matter cross spectrum

pnG constraints

Joint power spectrum and bispectrum constraining power as a function of the maximum scale included in the analysis



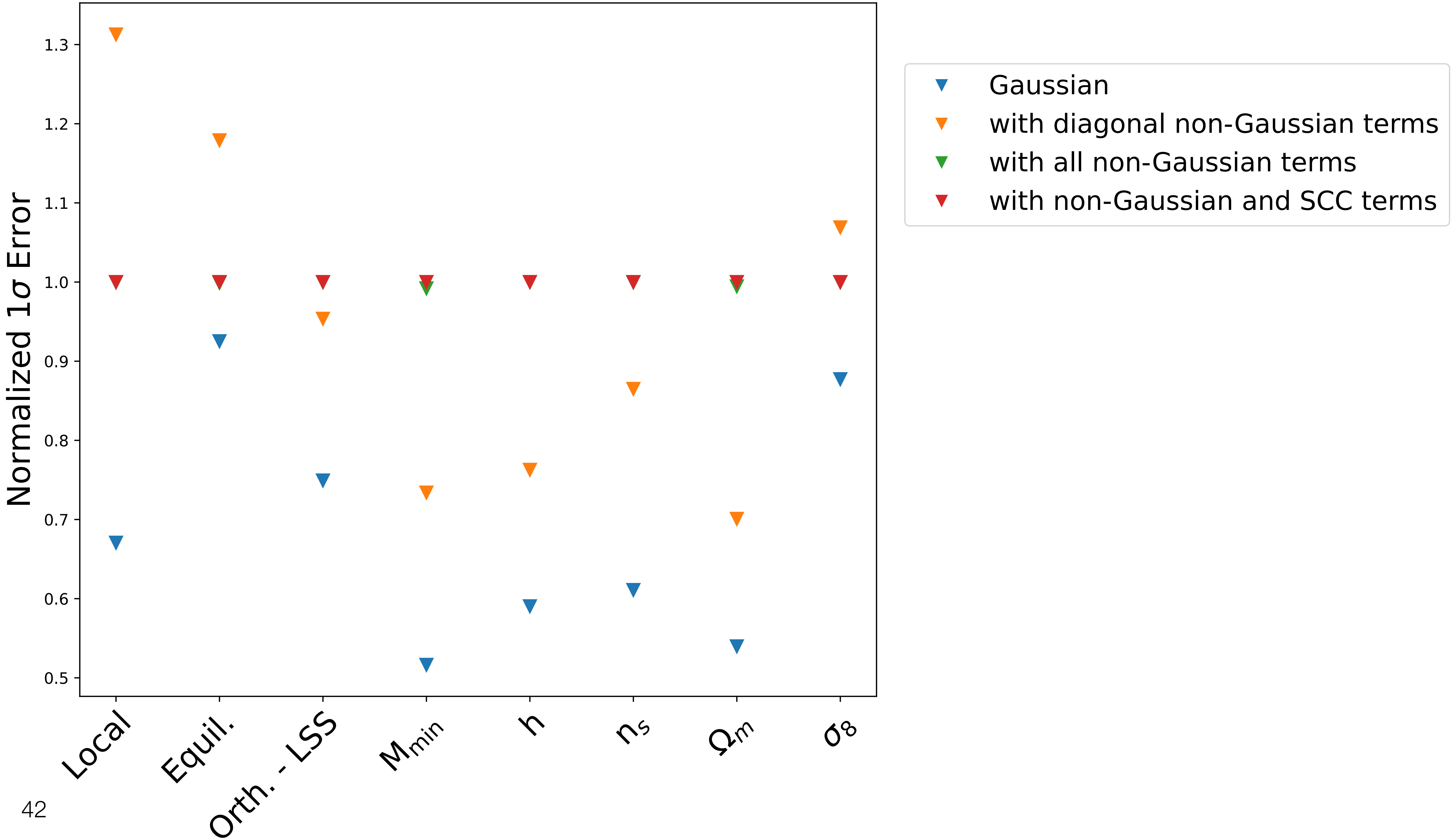
Dotted: including marginalization

Solid: without marginalizing cosmological params.

Dashed: primordial information content

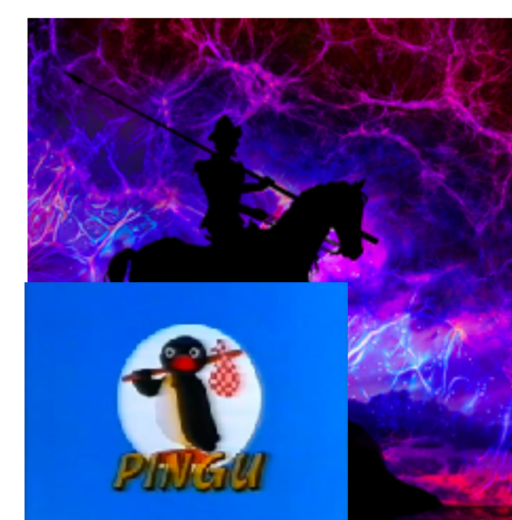
The importance of accurate covariance modelling

Forecasts constraints with the joint halo power spectrum and bispectrum when neglecting parts of the covariance matrix



Conclusions

- Monte Carlo Fisher forecasts over estimate the information content!
- Compressed Fisher forecasts provide an alternative estimator that is biased low
- Quijote-PNG simulations are available now: https://quijote-simulations.readthedocs.io/en/latest/png.html
- Sandbox for improving our understanding and ability to measure PNG



Thanks!!!