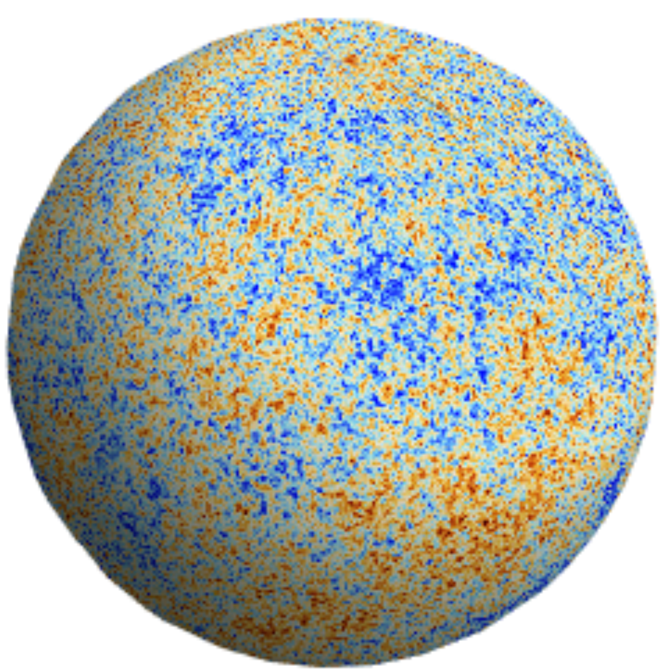
Cosmological information in perturbative forward modeling

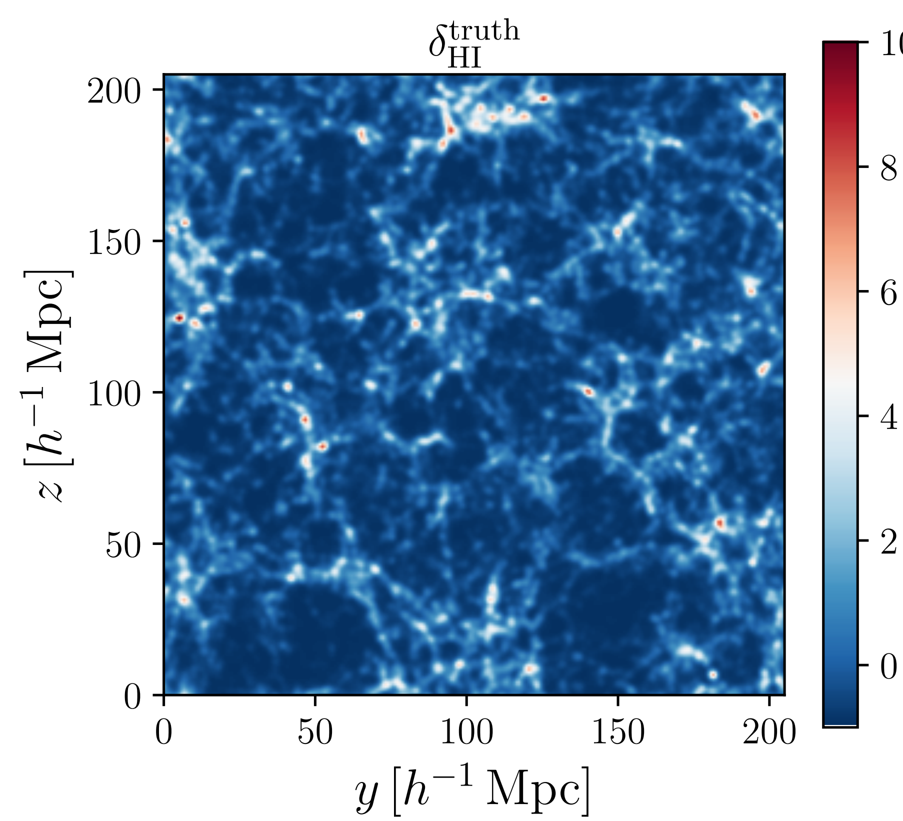
based on
Cabass, MS, Zaldarriaga (2023)

Marko Simonović
University of Florence

The problem of optimal inference

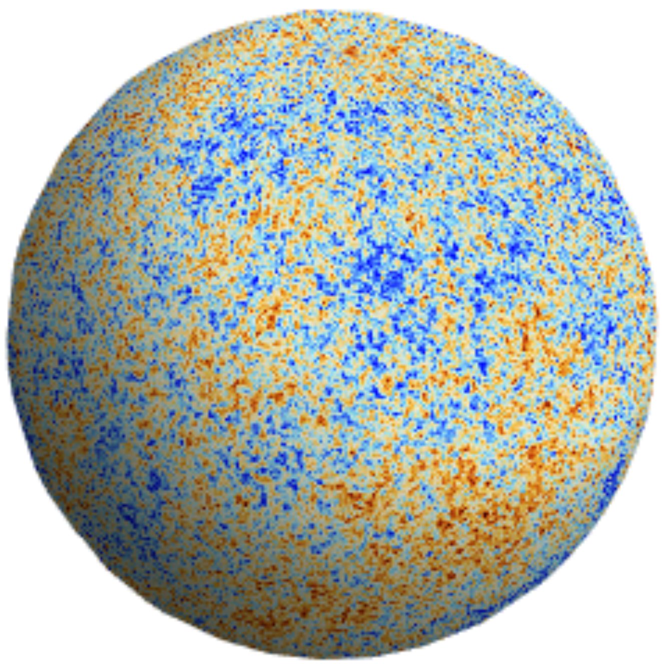
How do we optimally extract information from the maps?

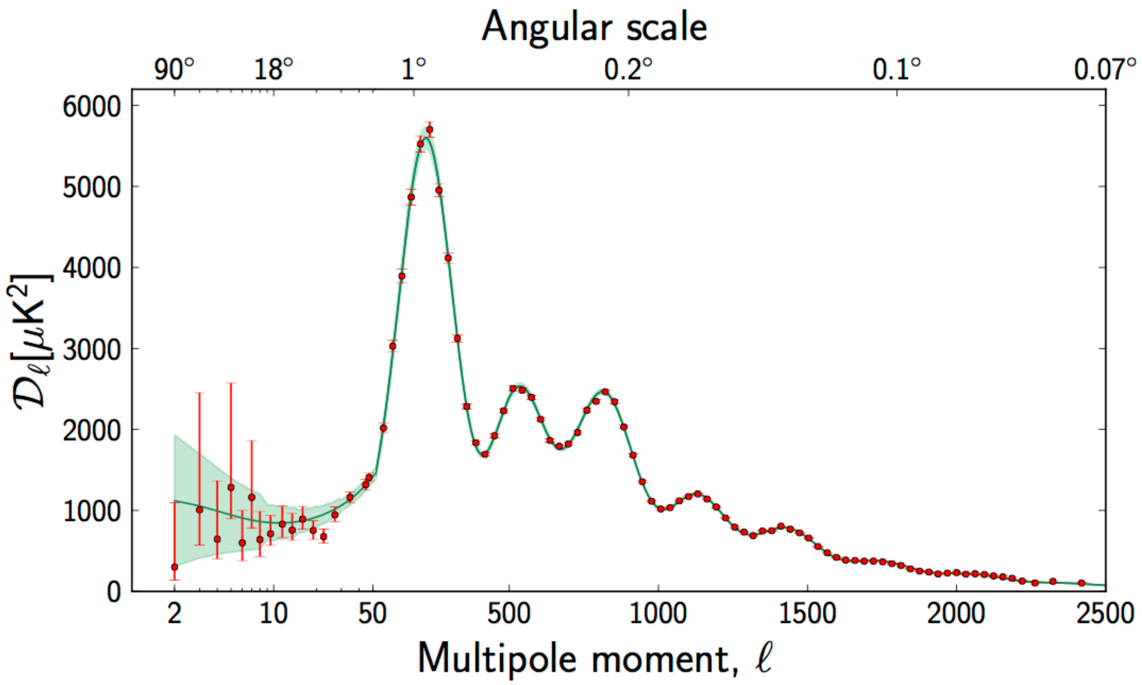




The problem of optimal inference

For a Gaussian field, the two point function contains all information

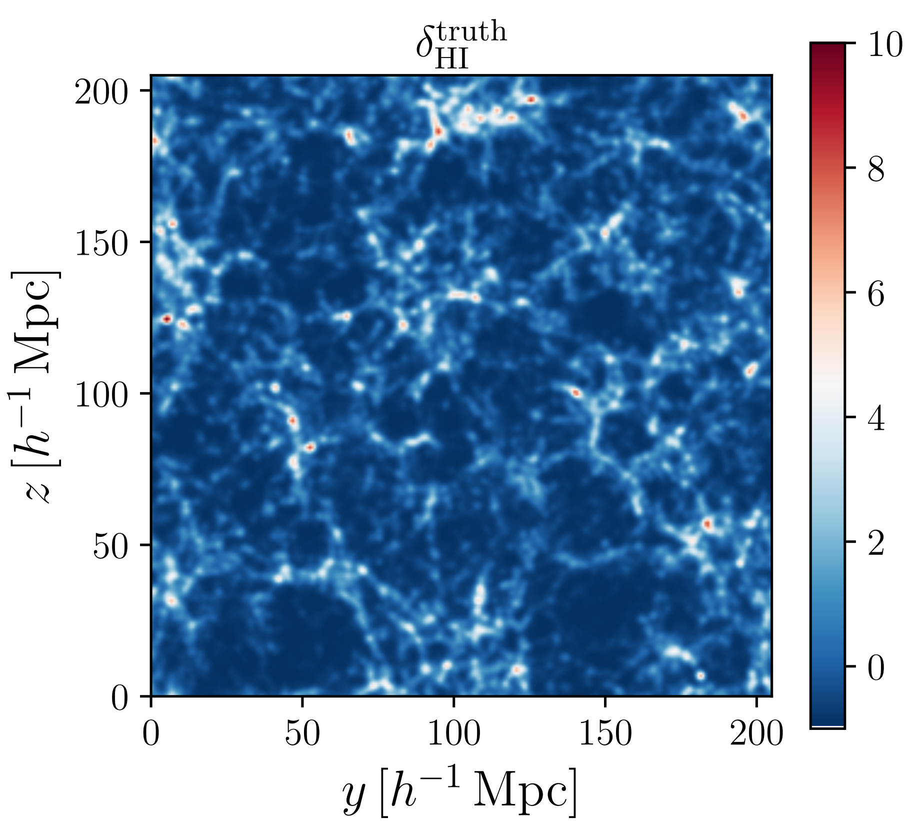


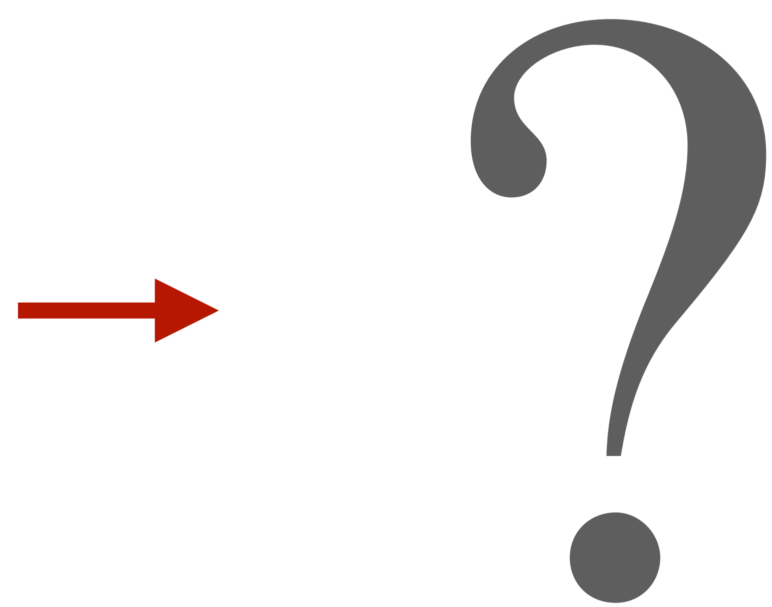


The problem of optimal inference

For a non-Gaussian field, the problem is much harder

1) What are the optimal observables?
2) What is the likelihood?





The problem of optimal inference

Using full hydro simulations we can in principle solve the problem

Simulation-based inference

1) run a huge number of simulations
2) “learn” the likelihood for the density field
3) compare numerical forward model to data, varying all ICs

This is not computationally feasible at the moment

Perturbative forward modeling

IC → nonlinear density field using PT

$$\delta_g = \delta_g[\delta, \boldsymbol{\theta}] + \epsilon_g$$

1) The map from ICs to the nonlinear field is simple

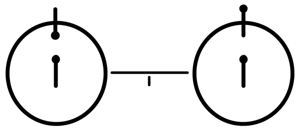
2) The likelihood is Gaussian on large scales

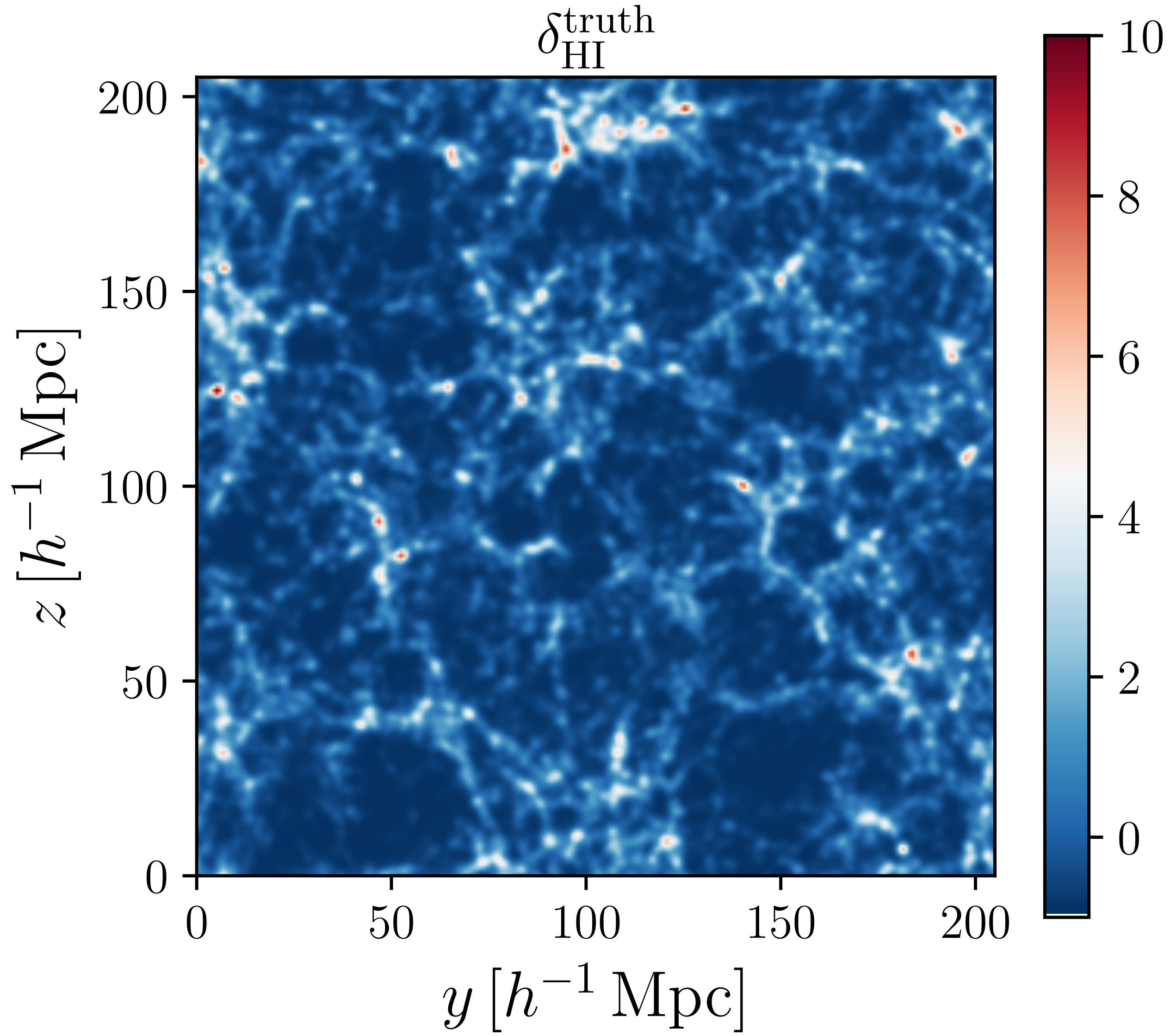
$$\mathcal{L}[\hat{\delta}_g|\delta, \boldsymbol{\theta}] = \text{normalization} \times \exp\left(-\frac{1}{2}\int_{\boldsymbol{k}} \frac{|\hat{\delta}_g(\boldsymbol{k}) - \delta_g[\delta, \boldsymbol{\theta}](\boldsymbol{k})|^2}{P_\epsilon}\right)$$

Perturbative forward modeling

Obuljen, MS, Schneider, Feldmann (2022)

Differences wrt the truth compatible with the shot noise



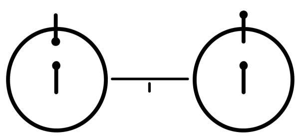


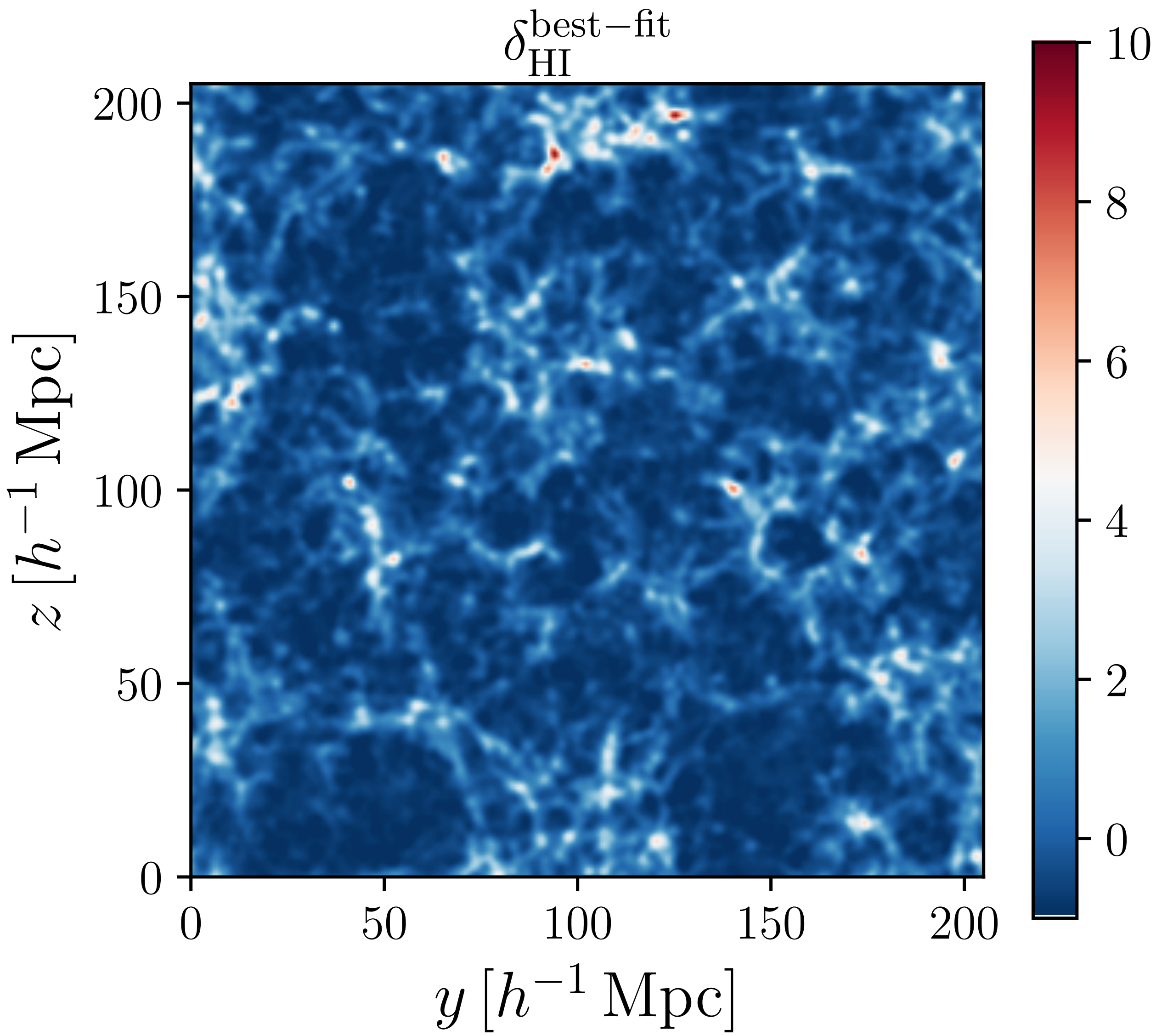
Hydro code

Perturbative forward modeling

Obuljen, MS, Schneider, Feldmann (2022)

Differences wrt the truth compatible with the shot noise





PT map

Perturbative forward modeling

$$\mathcal{P}[\boldsymbol{\theta}|\hat{\delta}_g] = \mathcal{N} \times \int \mathcal{D}\delta \, \exp\left(-\frac{1}{2}\int_{\boldsymbol{k}} \frac{|\delta(\boldsymbol{k})|^2}{P(k)} - \frac{1}{2}\int_{\boldsymbol{k}} \frac{|\hat{\delta}_g(\boldsymbol{k}) - \delta_g[\delta,\boldsymbol{\theta}](\boldsymbol{k})|^2}{P_\epsilon}\right) \times p(\boldsymbol{\theta})$$

This is usually solved numerically, but this is still very hard

Numerically it appears that the field level is more constraining than the analysis based on a few leading n-point functions

How can we understand this result? Can we trust it?

Perturbative forward modeling

1) Assume small noise

2) Assume perturbative model with a single expansion parameter

$$\Delta^2(k) = \int \frac{d^3q}{(2\pi)^3} P(q)$$

Compute the Fisher matrix for the filed level analysis perturbatively

$$\mathcal{P}[\boldsymbol{\theta}|\hat{\delta}_g] = \mathcal{N} \times \int \mathcal{D}\delta \, \exp\left(-\frac{1}{2}\int_{\boldsymbol{k}} \frac{|\delta(\boldsymbol{k})|^2}{P(k)}\right) \delta_{\rm D}^{(\infty)}\left(\hat{\delta}_g - \delta_g[\delta, \boldsymbol{\theta}]\right)$$

$$\Downarrow$$

$$\mathcal{P}[\boldsymbol{\theta}|\hat{\delta}_g] = \int \mathcal{D}\delta_g \left|\frac{\partial \delta}{\partial \delta_g}\right| \exp\left(-\frac{1}{2}\int_{\boldsymbol{k}} \frac{|\delta[\delta_g, \boldsymbol{\theta}](\boldsymbol{k})|^2}{P(k)}\right) \delta_{\rm D}^{(\infty)}(\hat{\delta}_g - \delta_g)$$

Perturbative forward modeling

$$\delta_g(\boldsymbol{k}) = \sum_{n=1}^{+\infty} \int_{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_n} (2\pi)^3 \delta_{\rm D}^{(3)}(\boldsymbol{k} - \boldsymbol{p}_{1\cdots n})\, X_n(\boldsymbol{\theta}; \boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)\, \delta(\boldsymbol{p}_1) \cdots \delta(\boldsymbol{p}_n)$$

The inverse model:

$$\delta(\boldsymbol{k}) = \sum_{n=1}^{+\infty} \int_{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_n} (2\pi)^3 \delta_{\rm D}^{(3)}(\boldsymbol{k} - \boldsymbol{p}_{1\cdots n}) Y_n(\boldsymbol{\theta}; \boldsymbol{p}_1, \ldots, \boldsymbol{p}_n) \delta_g(\boldsymbol{p}_1) \cdots \delta_g(\boldsymbol{p}_n)$$

$$\begin{aligned}
Y_1(\boldsymbol{\theta}) &= X_1^{-1}(\boldsymbol{\theta}) \ , \\
Y_2(\boldsymbol{\theta}; \boldsymbol{p}_1, \boldsymbol{p}_2) &= -X_1^{-3}(\boldsymbol{\theta}) X_2(\boldsymbol{\theta}; \boldsymbol{p}_1, \boldsymbol{p}_2) \ , \\
Y_3(\boldsymbol{\theta}; \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) &= \frac{2}{3} X_1^{-5}(\boldsymbol{\theta}) \Big[X_2(\boldsymbol{\theta}; \boldsymbol{p}_1, \boldsymbol{p}_2 + \boldsymbol{p}_3) X_2(\boldsymbol{\theta}; \boldsymbol{p}_2, \boldsymbol{p}_3) + X_2(\boldsymbol{\theta}; \boldsymbol{p}_2, \boldsymbol{p}_1 + \boldsymbol{p}_3) X_2(\boldsymbol{\theta}; \boldsymbol{p}_1, \boldsymbol{p}_3) \\
&\quad + X_2(\boldsymbol{\theta}; \boldsymbol{p}_3, \boldsymbol{p}_1 + \boldsymbol{p}_2) X_2(\boldsymbol{\theta}; \boldsymbol{p}_1, \boldsymbol{p}_2) - \frac{3}{2} X_1(\boldsymbol{\theta}) X_3(\boldsymbol{\theta}; \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) \Big] \ .
\end{aligned}$$

Perturbative forward modeling

A typical example: amplitude of the linear field and linear bias

perturbative posterior → Fisher matrix

$$\frac{1}{\sigma_A^2} = 2V \int_{\boldsymbol{k},\boldsymbol{p}} \left[X_2^2(\boldsymbol{p}, \boldsymbol{k}-\boldsymbol{p}) \frac{P(p)P(|\boldsymbol{k}-\boldsymbol{p}|)}{P(k)} + 2X_2(\boldsymbol{p}, \boldsymbol{k}-\boldsymbol{p})X_2(-\boldsymbol{p}, \boldsymbol{k})P(p) \right]$$

The exact same error as for P+B analysis in the same model!

Perturbative forward modeling

More generally, one can show that the two approaches match order by order in perturbation theory

In the limit of small $\Delta^2(k)$ the P+B+… analysis is optimal

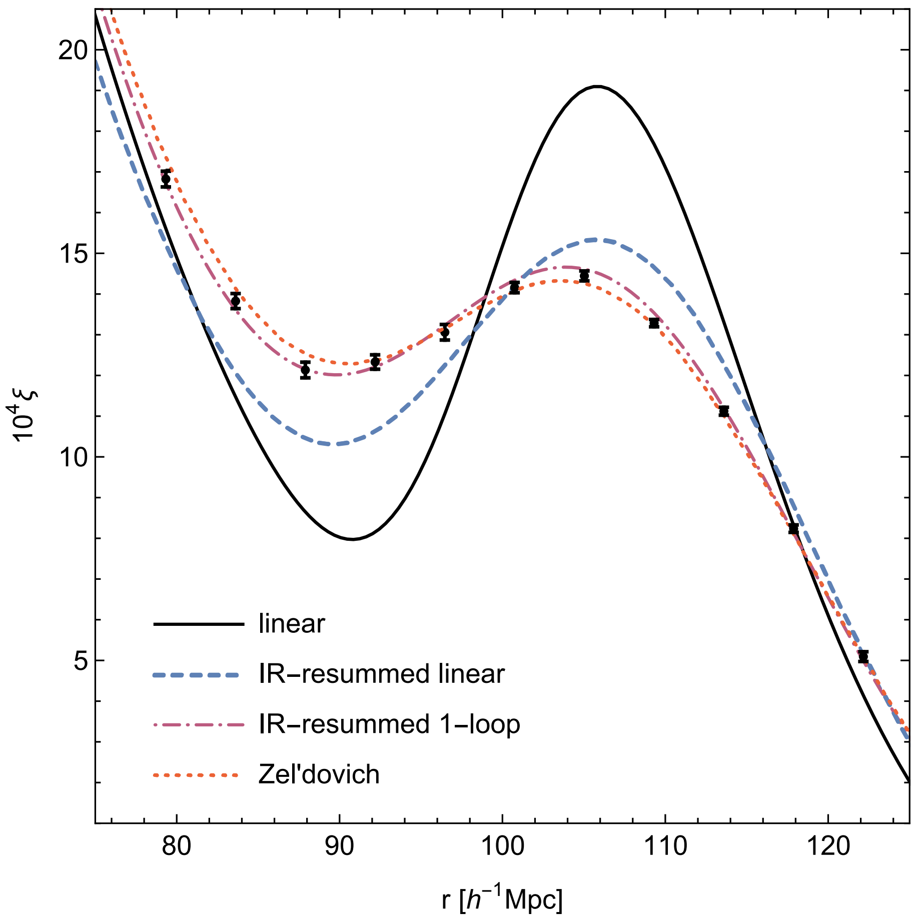
This result implies that on large scales we cannot gain much using alternative summary statistics, such as voids

Exceptions

The real universe is more complicated and has other expansion parameters

Large displacements and the BAO peak

$$\Sigma^2 \approx \frac{1}{6\pi^2} \int_{\ell_{\rm BAO}^{-1}}^{k_{\rm NL}} {\rm d}q \, P(q)$$



Affects only features

Field level is optimal, but similar to the BAO reconstruction

Exceptions

Large covariance for simple estimators

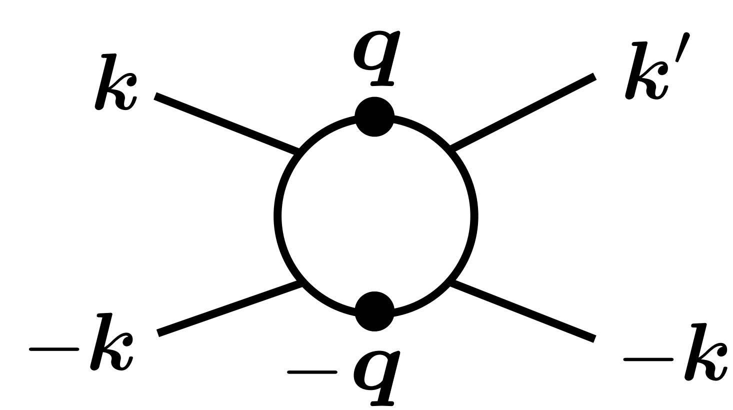
$$\tilde{\mathcal{E}} = \frac{1}{N_{\rm pix}} \int_{\boldsymbol{k}} \frac{|\hat{\delta}_g(\boldsymbol{k})|^2}{P(k)}$$

$$\mathrm{var}(\tilde{\mathcal{E}}) = \frac{2V}{N_{\rm pix}^2} \int_{\boldsymbol{k}} \frac{P_g^2(k)}{P^2(k)} + \frac{V}{N_{\rm pix}^2} \int_{\boldsymbol{k},\boldsymbol{k}'} \frac{T_g(\boldsymbol{k}, -\boldsymbol{k}, \boldsymbol{k}', -\boldsymbol{k}')}{P(k)P(k')}$$

$$T_g(\boldsymbol{k}, -\boldsymbol{k}, \boldsymbol{k}', -\boldsymbol{k}') \supset P(k)P(k') \int_{\boldsymbol{q}} P^2(q)$$

var(δ) can be small

var(δ^2) can in principle be very large



Conclusions

Field level analysis based on PT is equivalent to P+B+…

This is based on several assumptions, which are mildly violated even in LCDM

Are there more exceptions?

What does this imply for alternative summary statistics?

Are these arguments still valid for $\Delta^2(k)$ not too small?

Are there new large parameters related to new physics?