Indicator Power Spectra: Surgical Excision of Non-linearities and Covariance Matrices for Counts in Cells

István Szapudi

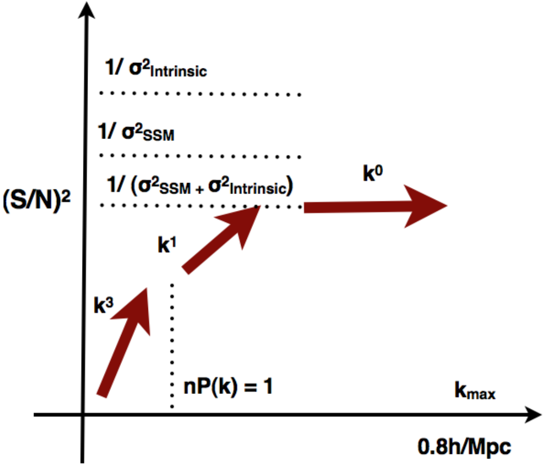
Institute for Astronomy
University of Hawaii

February 6, 2023

Cosmological Information in LSS Surveys

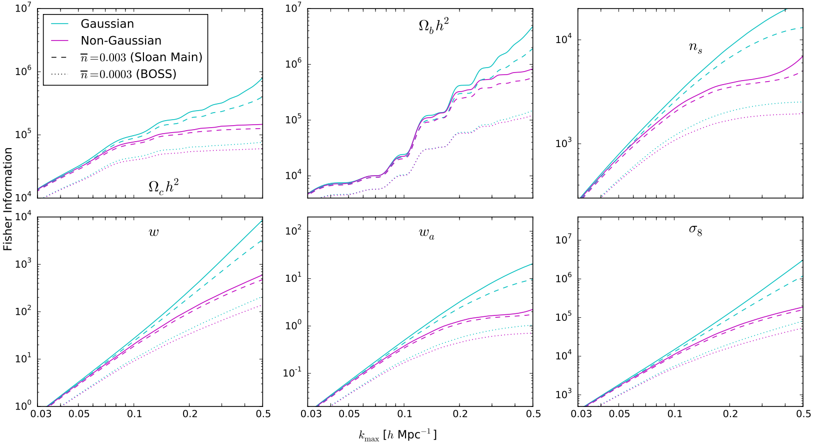
- Standard power spectrum extracts a small fraction of the available information from LSS surveys, unlike for CMB
- Good news: many ($\propto k^3$) high k modes are available even for smaller deep surveys
- Bad news: high k's in $P(k)$ don't contain information: information plateau due to the tri-spectrum, beat coupling (BC) and integral constraint (IC; super survey modes)

Summary of plateaux



Forcasting the plateaux

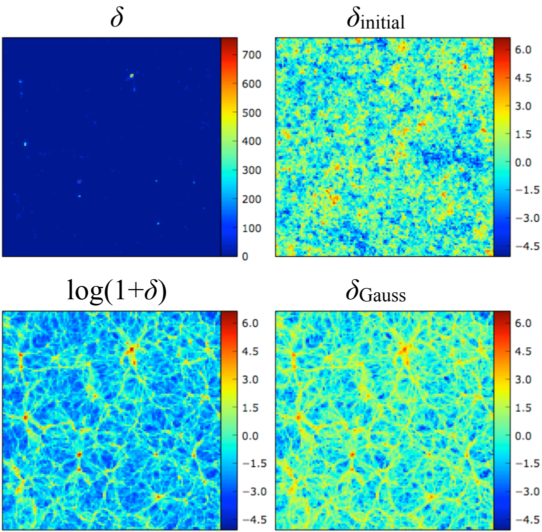
Repp Etal 2015



Cosmological Information in LSS Surveys

- Standard power spectrum extracts a small fraction of the available information from LSS surveys, unlike for CMB
- Good news: many ($\propto k^3$) high k modes are available even for smaller deep surveys
- Bad news: high k's in $P(k)$ don't contain information: information plateau due to the tri-spectrum, beat coupling (BC) and integral constraint (IC; super survey modes)
- Silver lining: BC and IC cancel to 10% accuracy
- Standard solution: add higher order statistics: single or multi-point moments
- Worst news: higher moments contain a small fraction of available information, usually at high computational cost
- Sufficient statistics, density pdf, and indicator functions

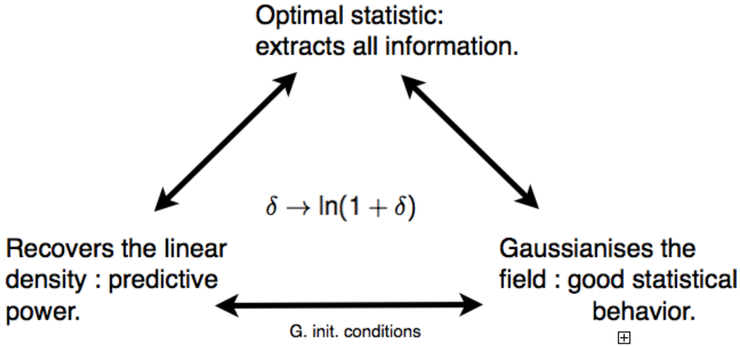
Logarithmic mapping



Sufficient Statistics: All Information on a Parameter

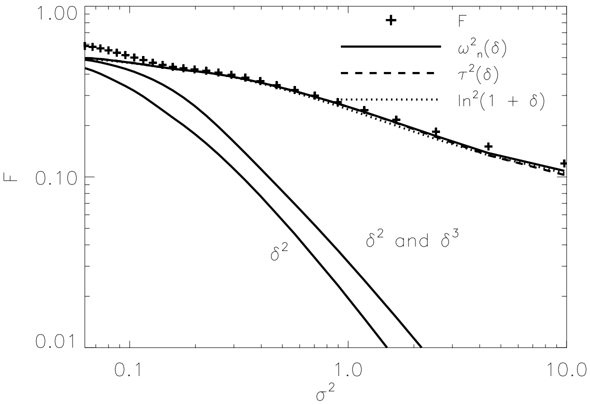
Carron & Szapudi (2013 MNRAS 434, 2961; 2014, MNRAS 439, L11)

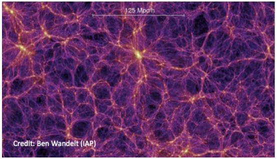
$$\partial_\alpha \ln p(\delta) \simeq \tau^2(\delta) \simeq \left(\frac{(1+\delta)^{(n+1)/3} - 1}{(n+1)/3}\right)^2 \simeq \ln(1+\delta)^2$$



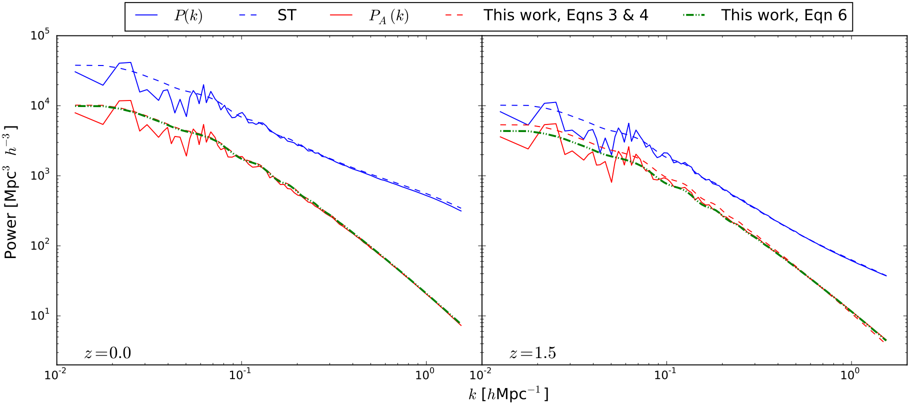
This follows from $n \simeq -1$, and Gaussian initial conditions.

Info. in the Millenium simulation density field



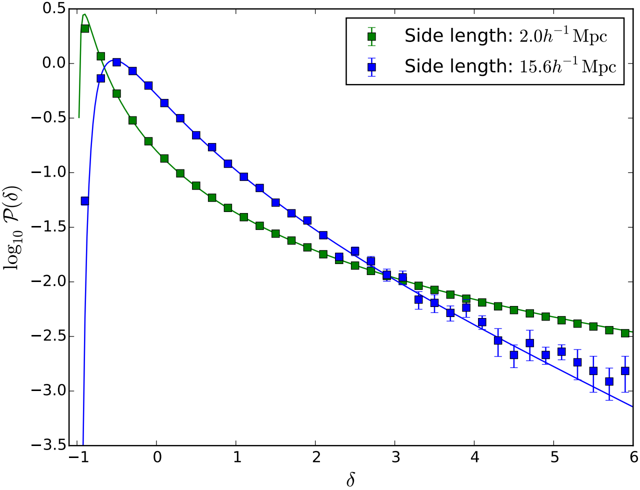


Precision Prediction for the Log Power Spectrum



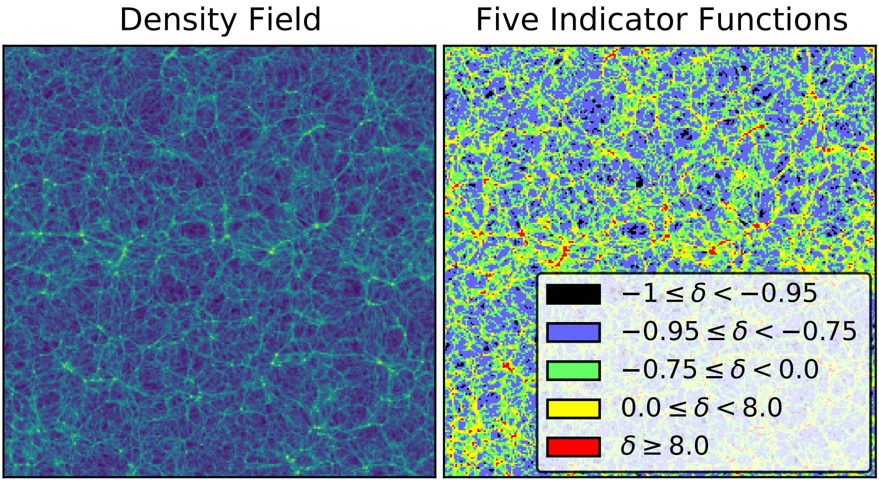
PDF

The density pdf breaks the degeneracy between σ_8 and bias (Repp & Szapudi 2018, 2020)



Repp & Szapudi (2022)

$$\mathcal{I}_B(x) = \begin{cases} 1 & x \in B \\ 0 & \text{otherwise} \end{cases}$$



Lognormal theory

$$\langle \mathcal{I}_1 \mathcal{I}_2 \rangle_r = \int d\nu_1\, d\nu_2\, W_{B_1}(\nu_1) W_{B_2}(\nu_2) \mathcal{P}(\nu_1, \nu_2)$$

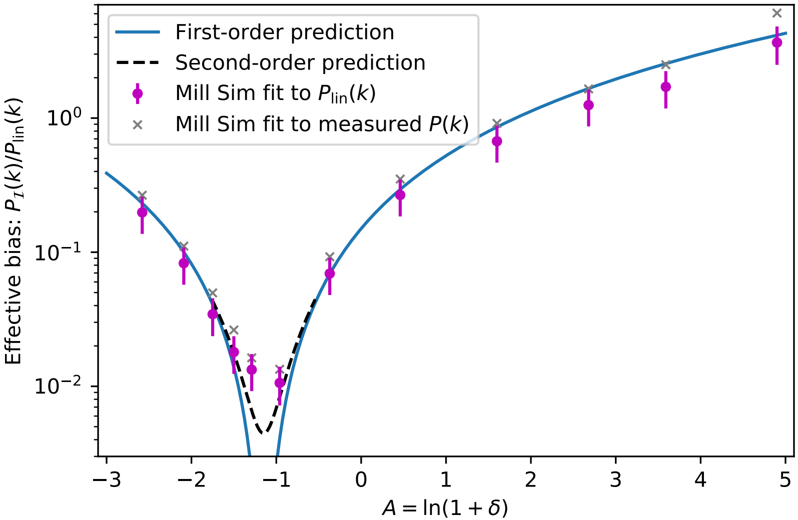
$$\mathcal{P}(\nu_1, \nu_2) = \frac{1 + \gamma \nu_1 \nu_2}{2\pi} \exp\left(-\frac{\nu_1^2 + \nu_2^2}{2} \right)$$
$$= (1 + \gamma \nu_1 \nu_2) \mathcal{P}(\nu_1) \mathcal{P}(\nu_2)$$

$$\xi_{12}(r) = \xi_A(r) \frac{(\langle A \rangle_{B_1} - \overline{A})(\langle A \rangle_{B_2} - \overline{A})}{\sigma_A^4} = \xi_A(r) \frac{\langle \nu \rangle_{B_1} \langle \nu \rangle_{B_2}}{\sigma_A^2}$$

Finally, with $P_A(k) = b_A^2 P_{\mathrm{lin}}(k)$ (up to a constant)

$$P_{\mathcal{I}}(k) = b_{\mathcal{I}}^2 b_A^2 P_{\mathrm{lin}}(k) = \langle \nu \rangle_B^2 \frac{P_{\mathrm{lin}}(k)}{\sigma_{\mathrm{lin}}^2}.$$

$$P_{\mathcal{I}}(k) = \langle \nu \rangle^2 \frac{P_{\rm lin}(k)}{\sigma_{\rm lin}^2} + \frac{(1 - \langle \nu^2 \rangle)^2}{2} \frac{P_{\rm lin}^{(*2)}(k)}{\sigma_{\rm lin}^4}, \quad (1)$$



Models for the spectra

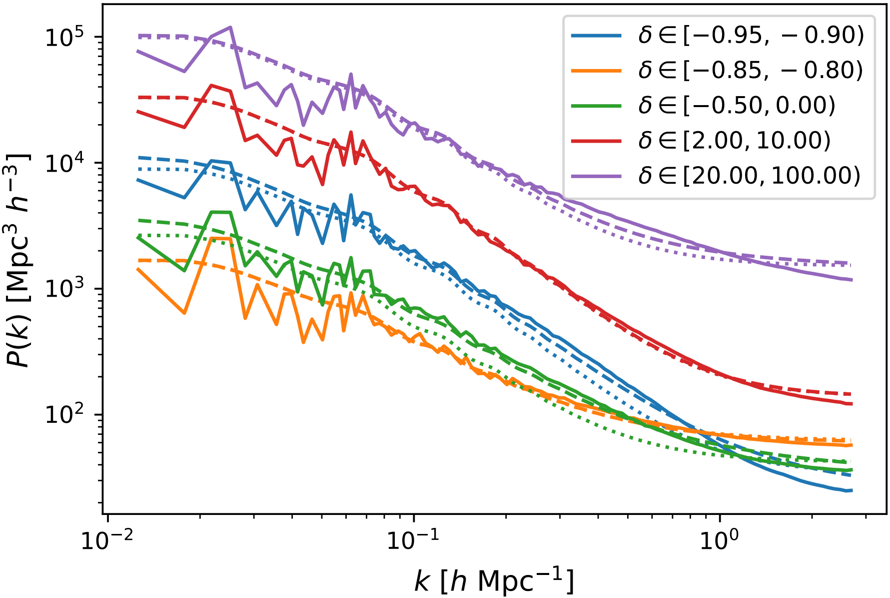
Linear model:

$$P_{\mathcal{I}}(k) = \langle \nu \rangle_B^2 \frac{P_{\rm lin}(k)}{\sigma_{\rm lin}^2} + C$$

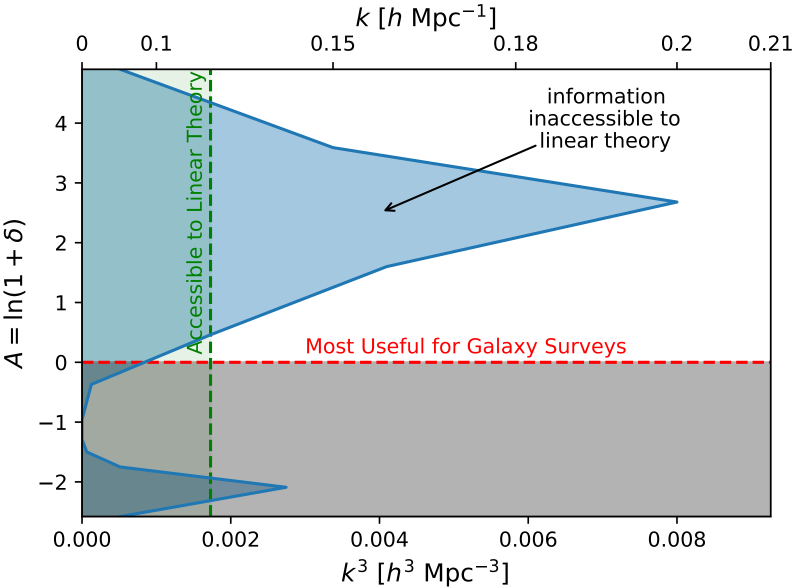
$$C = \delta V \left(\frac{1}{\mathcal{P}(B)} - 1 - \langle \nu \rangle_B^2 \right) .$$

C: the variance is constrained.
Extended model:

$$P_{\mathcal{I}}(k) = \langle \nu \rangle_B^2 \frac{P_{\rm lin}(k)}{\sigma_{\rm lin}^2} + Dk^n + C$$



$\delta \in [-0.95, -0.90)$
$\delta \in [-0.85, -0.80)$
$\delta \in [-0.50, 0.00)$
$\delta \in [2.00, 10.00)$
$\delta \in [20.00, 100.00)$
$P(k)$ [Mpc^3 h^{-3}]
10^5
10^4
10^3
10^2
10^{-2}
10^{-1}
10^0
k [h Mpc^{-1}]



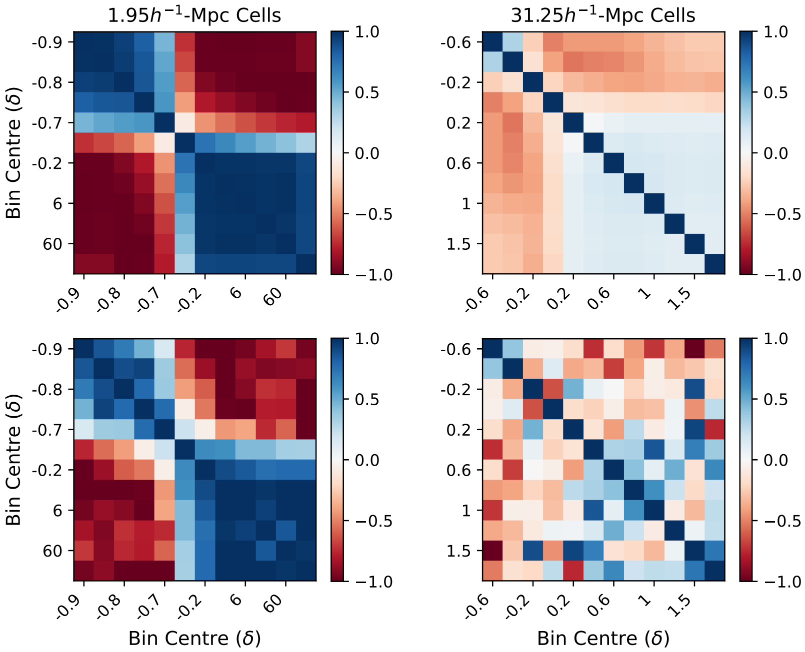
k [h Mpc^{-1}]
0
0.1
0.15
0.18
0.2
0.21
information
inaccessible to
linear theory
Accessible to Linear Theory
Most Useful for Galaxy Surveys
$A = \ln(1+\delta)$
4
3
2
1
0
−1
−2
0.000
0.002
0.004
0.006
0.008
k^3 [h^3 Mpc^{-3}]

Covariance matrix theory

$$\sigma^2_{\mathcal{P}} = \frac{\mathcal{P}\,(1-\mathcal{P})}{N_c} + \frac{(N_c-1)\overline{\xi}^{\neq}_{\mathcal{I}}}{N_c}\mathcal{P}^2$$

$$\sigma_{\mathcal{P}_1\mathcal{P}_2} = \frac{-\mathcal{P}_1\mathcal{P}_2}{N_c} + \frac{N_c-1}{N_c}\overline{\xi}_{\mathcal{I}_1\mathcal{I}_2}\mathcal{P}_1\mathcal{P}_2$$

$$\begin{aligned} N_C^2\,\overline{\xi}_{\mathcal{I}} &= (N_c^2 - N_c)\overline{\xi}^{\neq}_{\mathcal{I}} + N_c\,\xi_{\mathcal{I}}(0) \\ \overline{\xi}^{\neq}_{\mathcal{I}} &= \frac{1}{N_c-1}(N_c\,\overline{\xi}_{\mathcal{I}} - \sigma^2_{\mathcal{I}}) \\ &= \frac{1}{N_c-1}\left(N_c\,\overline{\xi}_{\mathcal{I}} - \frac{1}{\mathcal{P}} + 1\right). \end{aligned}$$



1.95h^{-1}-Mpc Cells
31.25h^{-1}-Mpc Cells
Bin Centre (δ)
-0.9
-0.8
-0.7
-0.2
6
60
-0.6
-0.2
0.2
0.6
1
1.5
1.0
0.5
0.0
−0.5
−1.0

Summary: Indicator Spectra

- Indicator functions repackage and slice information in an intuitive fashion
- Traditional power spectra, sufficient statistics (log-mapping) and PDFs can be reconstructed from IS
- IS extract more information than $P(k)$ with the same code and modest resources
- We have an precise theory to predict the full set of IS
- Linear theory accuracy is k-dependent
- For "optimal levels" valid up to $k \simeq 0.3$ from $\simeq 0.1$
- We can predict covariance matrices for density PDFs
- Covariance matrices of power spectra are more diagonal
- The seperation of density levels will help understanding bias and redshift distortions.
- Future: MCMC parameter estimation