N-body Simulations and Cosmological Statistics
Wigner 121, Budapest

István Szapudi

Institute for Astronomy
University of Hawaii

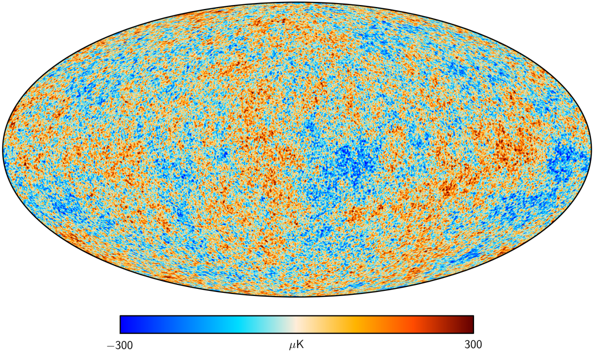
Collaborators: István Csabai, Gábor Rácz, Andrew Repp

09/19/2023

Table of contents

Anisotropies in the observed Universe

Cosmological principle: statistical homogeneity and anisotropy
CMB map of the early ($380ky$ after BB) density fluctuations [1]:

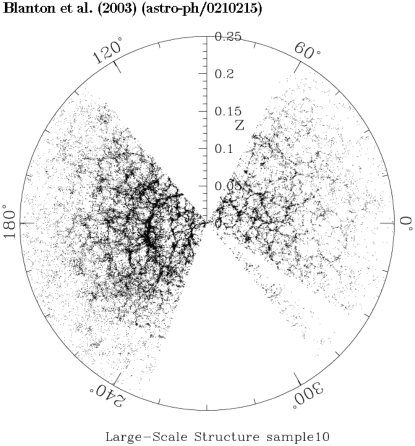


Typical overdensity fluctuations: $\Delta T/T \sim \rho/\overline{\rho} - 1 = \delta \sim 10^{-5}$

Inhomogeneities in the observed Universe

Late-time large-scale structures from the small initial fluctuations.

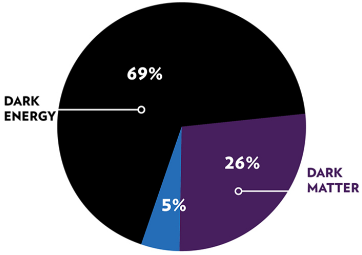
Large-scale galaxy survey map [2]:



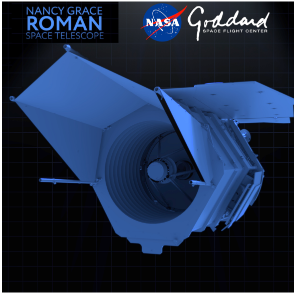
Typical overdensity fluctuations: $\Delta\rho/\bar{\rho} = \delta \sim 10^2$

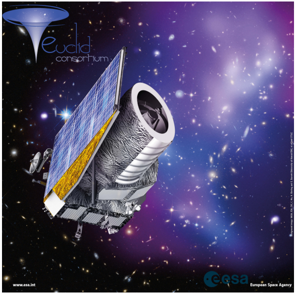
Standard Model of Cosmology

- Since structure formation depends on the expansion history and the law of gravitation, the large-scale structure statistics probe cosmological models.
- Growth and expansion history (and other probes such as nucleosynthesis) constrain parameters of the concordance ΛCDM model.



Upcoming Surveys





Upcoming Surveys

These surveys will provide new data from the large-scale structure by mapping the positions and distances of billions of galaxies.

Calculating the evolution of cosmic structures in different cosmological models will be crucial for interpretation.

Evolution of the cosmic structures

We expand fluctuations as a superposition of waves:

$$\delta(\vec{k}) = A(\vec{k}) \cdot e^{i\varphi(\vec{k})}$$

- $\vec{k}$: Wavenumber (inverse wavelength)[1/Mpc] [1]
- A: Amplitude of the fluctuation [dimensionless]
- φ: Phase of the fluctuation [dimensionless]

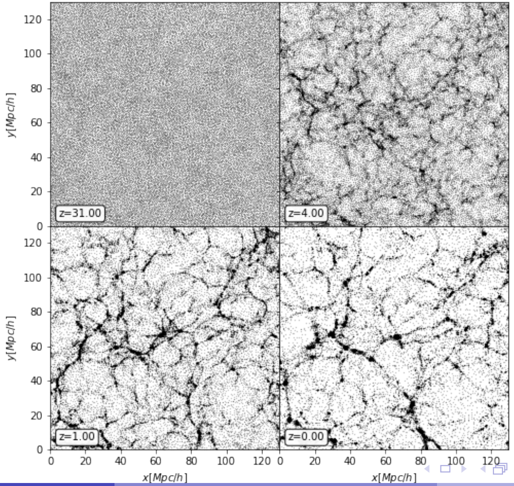
Linear theory: if $A(\vec{k})$ amplitudes are small, the modes grow independently:

$$\delta(\vec{k}, t) = \frac{D(t)}{D(t_0)} \delta(\vec{k}, t_0)$$

[1] 1 Mpc=3, 261, 563 ly

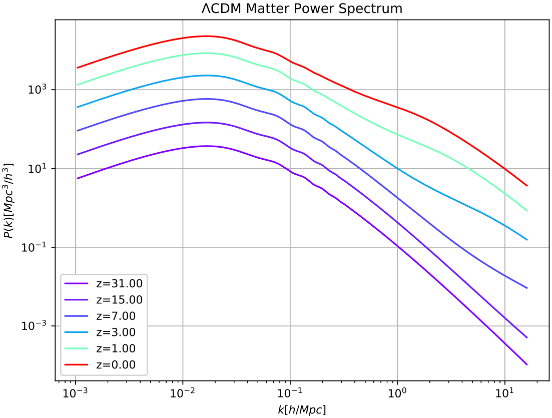
Evolution of the cosmic structures

Cosmological N-body simulations follow the orbits of gravitationally interacting particles for non-linear structure formation.



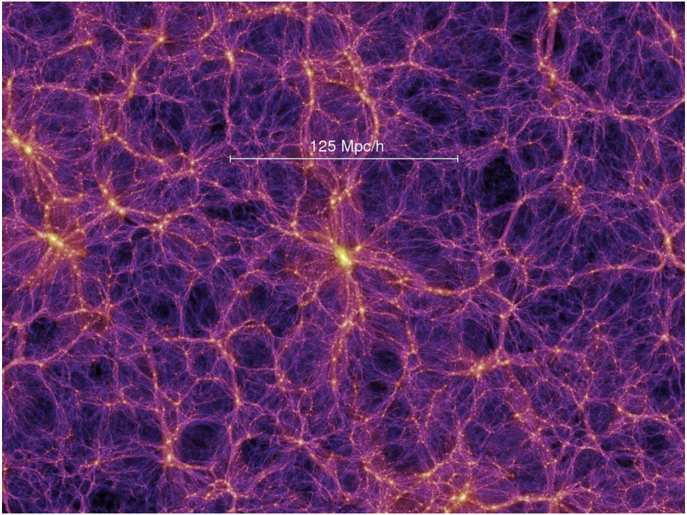
Power Spectrum

The most used statistical quantity to describe the density field is the power spectrum $P(k) = \frac{1}{(2\pi)^3} < \delta(\vec{k})\delta^*(\vec{k}) >_{k=|\vec{k}|}$.



Millennium Simulation $500h^{-1}\mathrm{Mpc}$, $N \simeq 10^{10}$

Gold standard/3600 refs



Sample Variance

Finite volume and resolution $\Longrightarrow$

- Sample variance: Each k wavenumber is resolved by a finite number of modes.

$$\Delta P_{SV}(k) = \sqrt{\frac{2}{\Delta N_m(k)}} P(k),$$

  where $\Delta N_m(k)$ is the number of modes per k-bin.

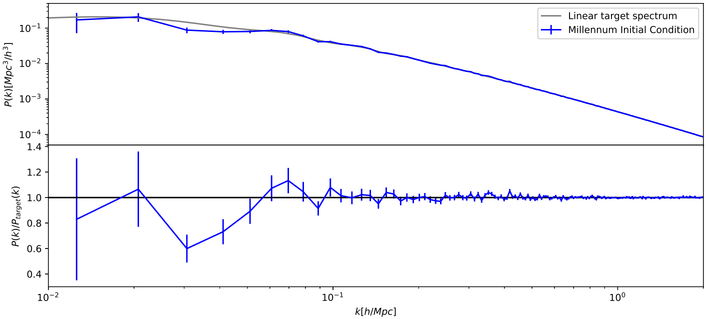
- Shot-noise: From discrete Poisson sampling.

$$\Delta P_{SN}(k) = V/N^3,$$

  where V is the volume, and N is the number of tracers of the density field (galaxies or particles)

- non-linearity/non-Gaussianity: off diagonal elements in the covariance matrix: beat coupling and super-survey covariance
- Millennium: rare fluctuation

Sample Variance in Simulations



- Number of modes are scaling with k^3, $\Longrightarrow$ the random fluctuations in the initial power spectrum will be stronger at the largest scales.
- For a given scale, this effect can be reduced by increasing the simulation volume.

Reducing Cosmic Variance in Simulations

N-body simulations are costly, therefore, it is imperative to develop methods to reduce cosmic variance.
Previously known methods:

- Ensembles of simulations (brute force)
- Paired: $\varphi \longrightarrow \varphi + \pi$ phase shift between the simulations [3]
- Paired-and-Fixed: Phase shifted pairs with fixed amplitudes (No variance) [4]

Complementary initial conditions

Rácz etal 2023

For a given initial condition, it is possible to generate a complementary pair, in which the initial amplitudes are chosen in such a way, that the average of the two initial $P(k)$ matches the $P_{target}(k)$ power spectrum.

$$P_{target}(k) = \frac{1}{2}\left(P_O(k) + P_C(k)\right)$$

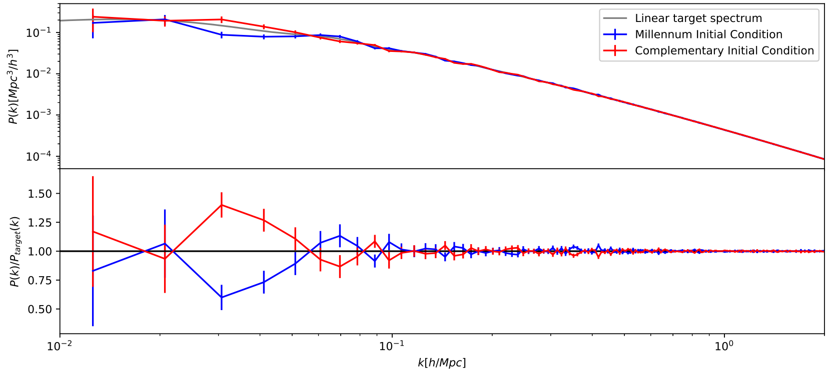
To further reduce the effects of phase-coupling, we also choose to use

$$\varphi(\vec{k}) \longrightarrow \varphi(\vec{k}) + \pi$$

inverted phases in the complementary pair.

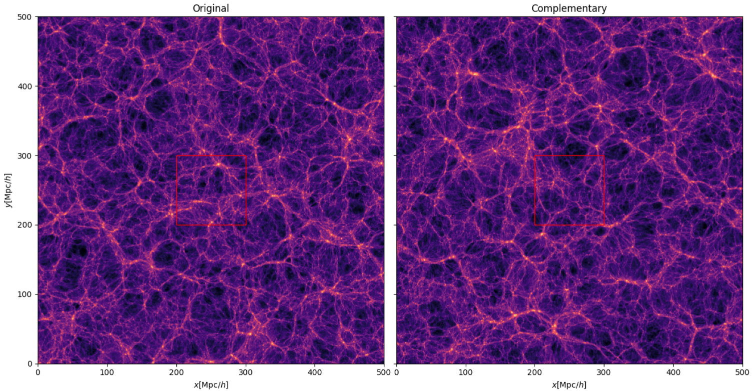
Millennium Complementary Initial Condition

Initial power spectrum:



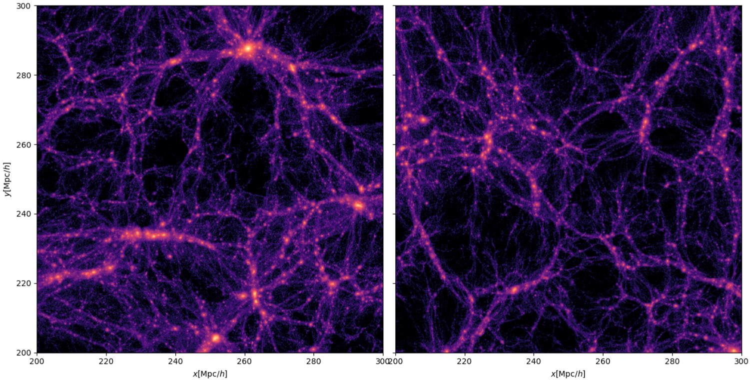
Millennium Complementary Pair

Dark Matter fields



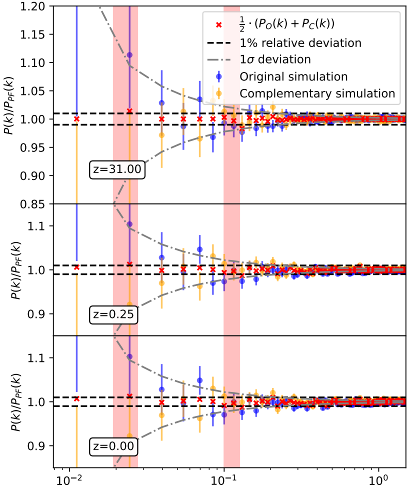
Millennium Complementary Pair

Dark Matter fields



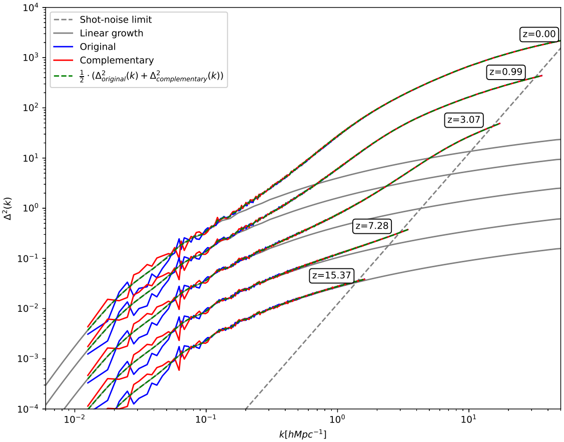
Complementary Simulations

We used the Paired-and-Fixed method as a reference to compare the different methods.



Millennium Complementary Pair

Dimensionless power spectrum $k^3 P(k)/2\pi^2$

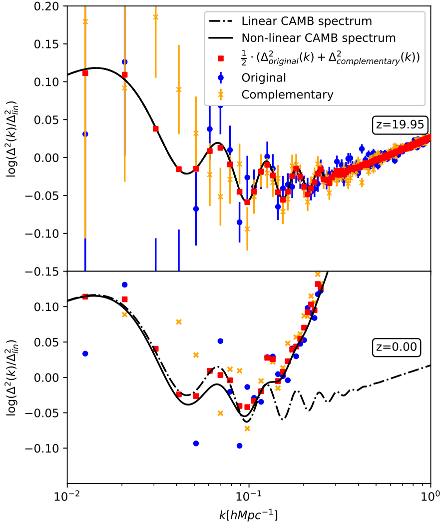


The original Millennium run in average, underestimated the power spectrum

- by 0.997% for the large scales
- by 0.0881% for small scales

at the $z = 0$ redshift.

Millennium Complementary Pair



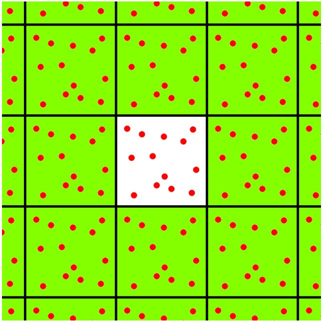
The fine wiggles in the $P(k)$ are completely resolved in the averaged spectrum.

For this kind of resolution at these scales, we would need over 2000 times larger volume for a single simulation approach. [5]

Periodic Boundary Conditions

All cosmological simulations since the 1950s

- Simulating the infinite universe in a finite memory
- Standard solution: repeat a finite volume infinite times
- Sharp cut in the power spectrum: $k < 2\pi/L_{box}$
- always smaller than the Newtonian $1/r^2$ force due to the periodic images
- Increasing particle number: too many high k, too few large scale mods



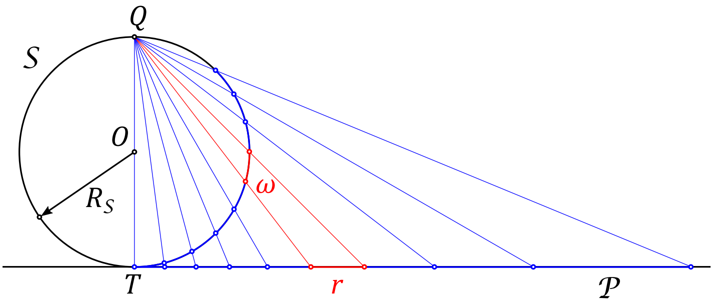
Compactifying the Universe

Map the infinite spatial extent of the Universe onto a finite volume by compactification:

$$1D : r \in (-\infty; +\infty) \longrightarrow \omega \in (-\pi; \pi)$$

Use constant resolution in the compact space. StePS uses inverse stereographic projection:

$$\omega = 2 \arctan\left(\frac{r}{2R_S}\right)$$



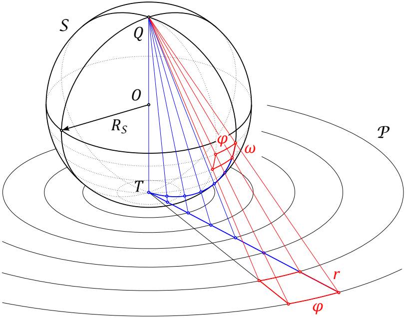
Compactifying the Universe

In 2D and 3D, if we use polar coordinates, only the radial coordinate is transformed:

$r \in [0; +\infty) \longrightarrow \omega[0; +\pi)$

$\vartheta \in [0; 2\pi) \longrightarrow \vartheta[0; 2\pi)$

$\varphi \in [0; \pi) \longrightarrow \varphi[0; \pi)$ (3D case)



Constant resolution in the compact sphere $\Longrightarrow$ decreasing resolution in radial direction, and constant angular resolution

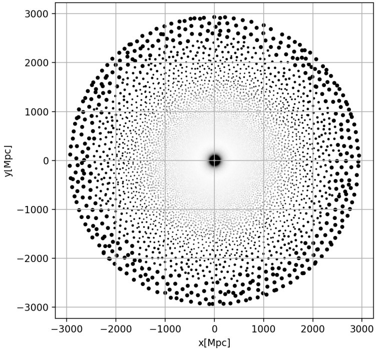
Compactifying the Universe

We use equal volume pixels in the compact space or equal space binning in ω

The North Pole pixel represents an infinite volume with infinite mass in the non-compact space. This point needs a special treatment.

Corresponds to an external (radial) force in comoving coordinates that prevents the collapse of the simulated volume

Homogeneous particle distribution:



Stereographically Projected Simulations

Solving the equations of motion:

- In compact space:
  - $\omega_i(t)$, $\varphi_i(t)$, $\vartheta_i(t)$
  - Too many floating point operation (hard to calculate physical distances)
  - No implementation yet
- In cartesian coordinates after decompactification:
  - $x_i(t)$, $y_i(t)$, $z_i(t)$
  - Fast
  - Existing N-body codes can be used

Initial conditions:

- Linear initial power spectrum
- Zel‘dovich and 2LPT approximation can be used
- The number of resolved modes is a function of ω (Zoom-in)
- In periodic geometry: $L_{box} > 4R_{sim}$ (Existing codes can be used)

Equations of motion

In comoving Cartesian coordinates:

$$m_i\ddot{\vec{x}}_i = \sum_j m_i m_j \frac{F(\vec{x}_i - \vec{x}_j, h_i + hj)}{a(t)^3} - 2m_i \frac{\dot{a}(t)}{a(t)}\dot{\vec{x}}_i + m_i \frac{4\pi G}{3}\overline{\rho}\vec{x}_i$$

- Smoothed gravitational force between the particles
- Hubble-drag
- Radial force: Force from the last pixel (gravitational pull)

In physical Cartesian coordinates:

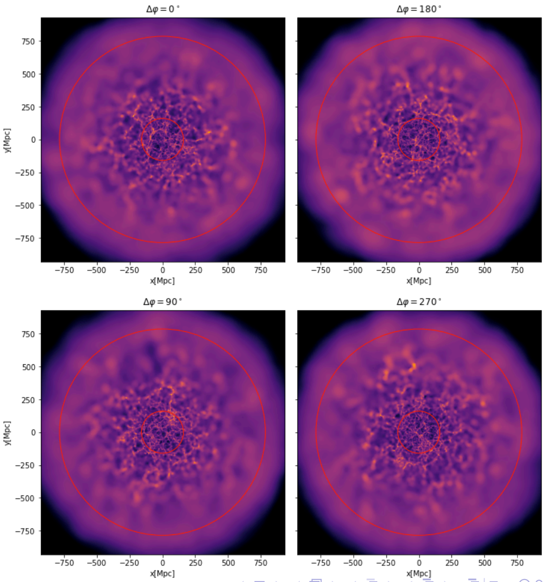
$$m_i\ddot{\vec{x}}_i = \sum_j m_i m_j F(\vec{x}_i - \vec{x}_j, h_i + hj) + m_i H_0^2 \Omega_\Lambda \vec{x}_i$$

- Smoothed gravitational force between the particles
- Radial force: Cosmological constant (Λ)
- This option is not available in periodic topology

Power spectrum estimation from StePS simulations

Using initially phase-shifted of pairs or quartets of simulations will:

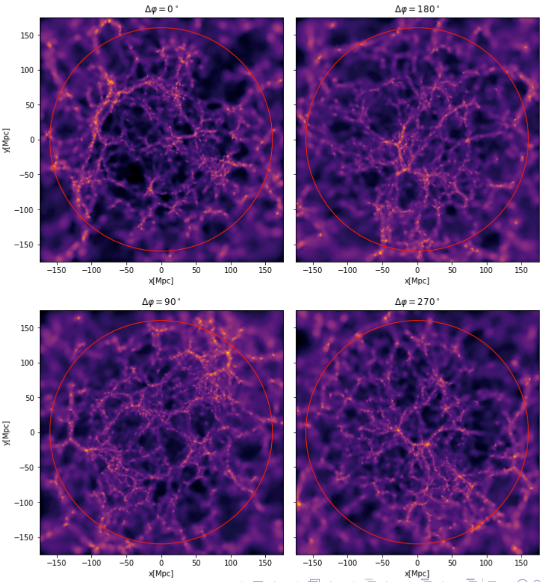
- reduce the effects of the cosmic variance (dominant effect: tidal fields)
- increase the volume of the sample



Power spectrum estimation from StePS simulations

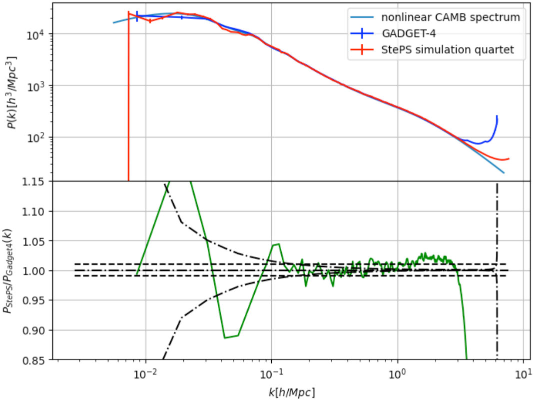
Using initially phase-shifted of pairs or quartets of simulations will:

- reduce the effects of the cosmic variance (dominant effect: tidal fields)
- increase the volume of the sample



Power spectrum estimation from StePS simulations

In this example, a quartet of StePS simulations with $1.8 \cdot 10^6$ particles was able to match a periodic Gadget-4 simulation with 10^9 particles.
$\Longrightarrow$ 3 orders of magnitude reduction in particle number is possible with StePS.

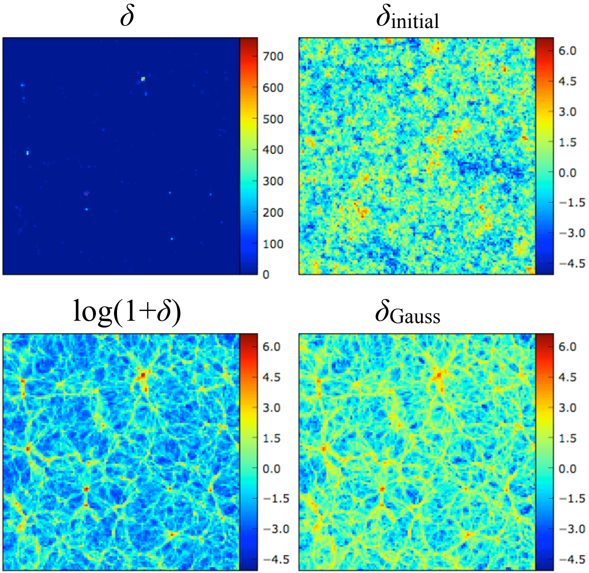


Gadget-4:36096 CPU hrs (Wall Clock: 48h; Xeon 768 Cores)
StePS:60(=4x15) GPU hrs (Wall clock: 4x7.5h; 2xNvidia V100)

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

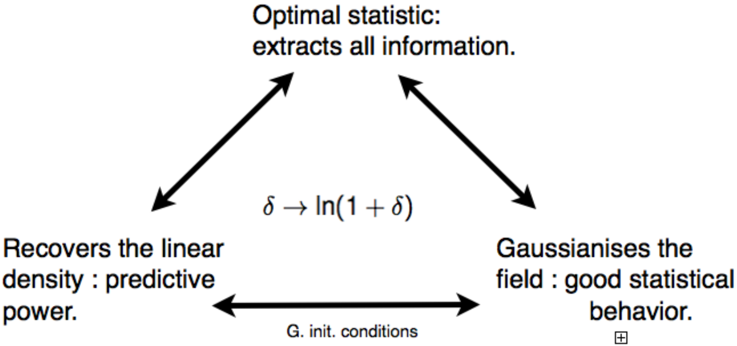
Logarithmic mapping



Sufficient Statistics: All Information on a Parameter

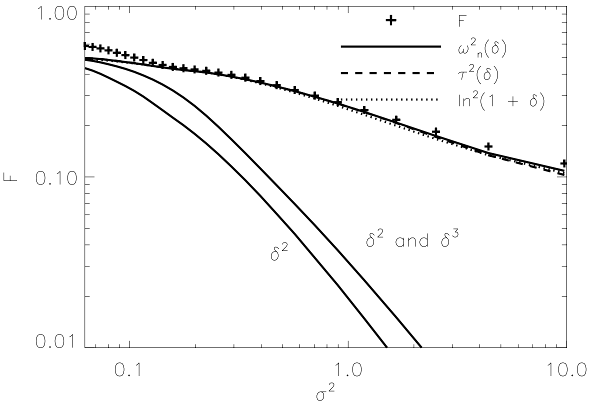
Carron & Szapudi (2013 MNRAS 434, 2961; 2014, MNRAS 439, L11)

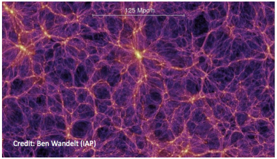
$$\partial_\alpha \ln p(\delta) \simeq \tau^2(\delta) \simeq \left(\frac{(1+\delta)^{(n+1)/3} - 1}{(n+1)/3} \right)^2 \simeq \ln(1+\delta)^2 \qquad (1)$$



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

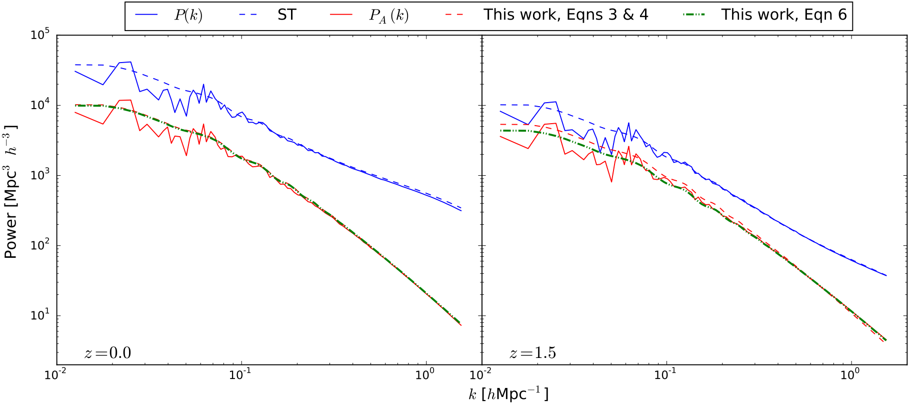
Info. in the Millenium simulation density field





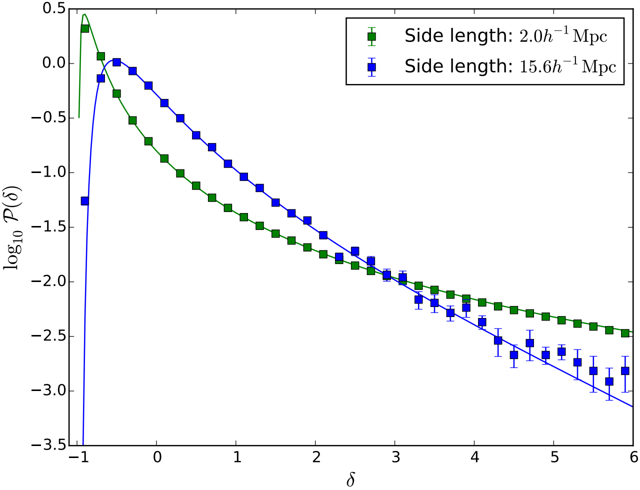
Precision Prediction for the Log Power Spectrum

Essential for cosmological parameters estimation



PDF

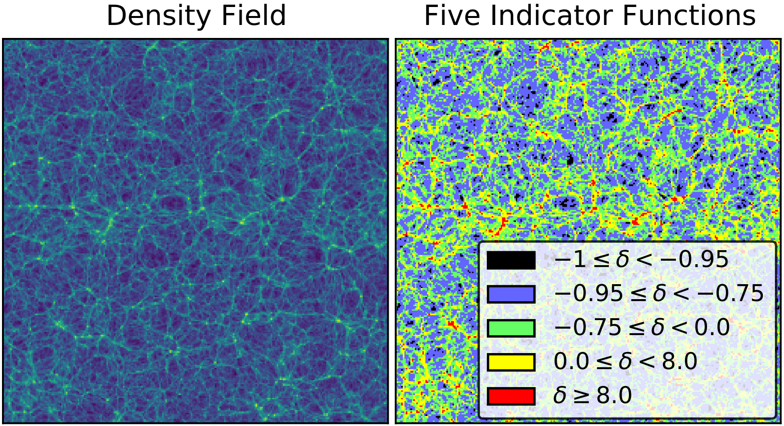
The density pdf breaks the degeneracy between σ_8 and bias (Repp & Szapudi 2018, 2020)



Indicator Functions: GUT of Statistics

Repp & Szapudi (2022)

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



Lognormal theory

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

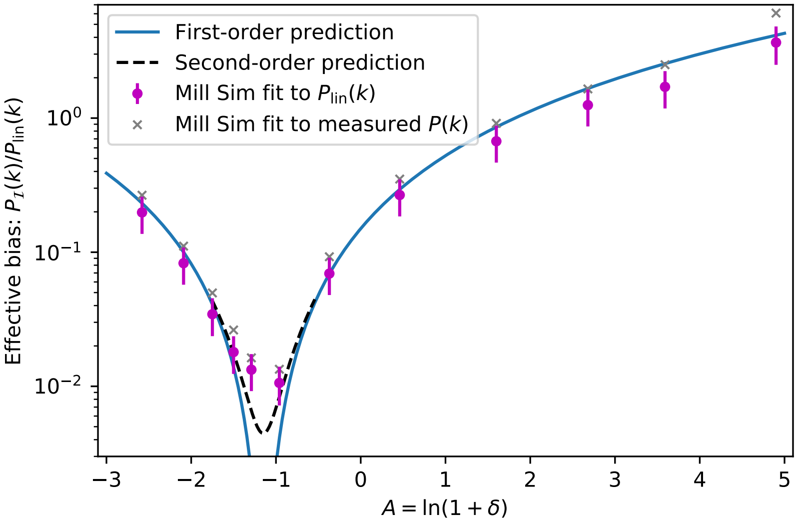
$$\mathcal{P}(\nu_1, \nu_2) = \frac{1 + \gamma \nu_1 \nu_2}{2\pi} \exp\left(-\frac{\nu_1^2 + \nu_2^2}{2} \right)$$
$$= (1 + \gamma \nu_1 \nu_2) \mathcal{P}(\nu_1) \mathcal{P}(\nu_2)$$

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

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

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

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



Models for the spectra

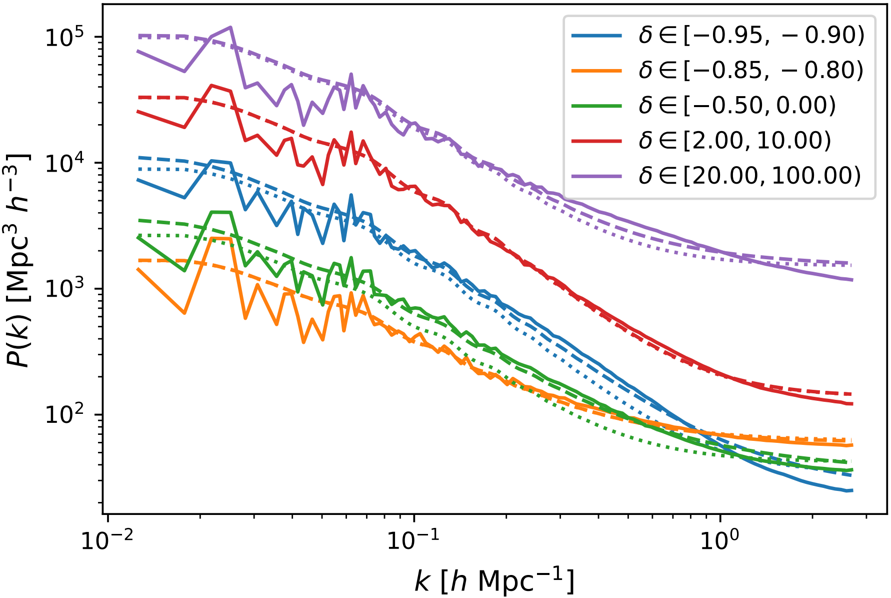
Linear model:

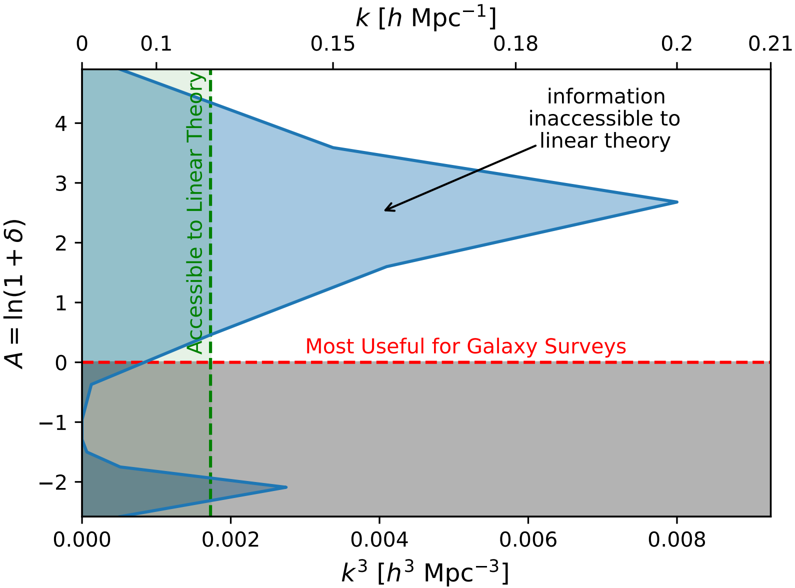
$$P_{\mathcal{I}}(k) = \langle \nu \rangle_B^2 \frac{P_{\rm lin}(k)}{\sigma_{\rm lin}^2} + C$$

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

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}]



$k\ [h\ \mathrm{Mpc}^{-1}]$
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
$k^3\ [h^3\ \mathrm{Mpc}^{-3}]$

Covariance matrix theory

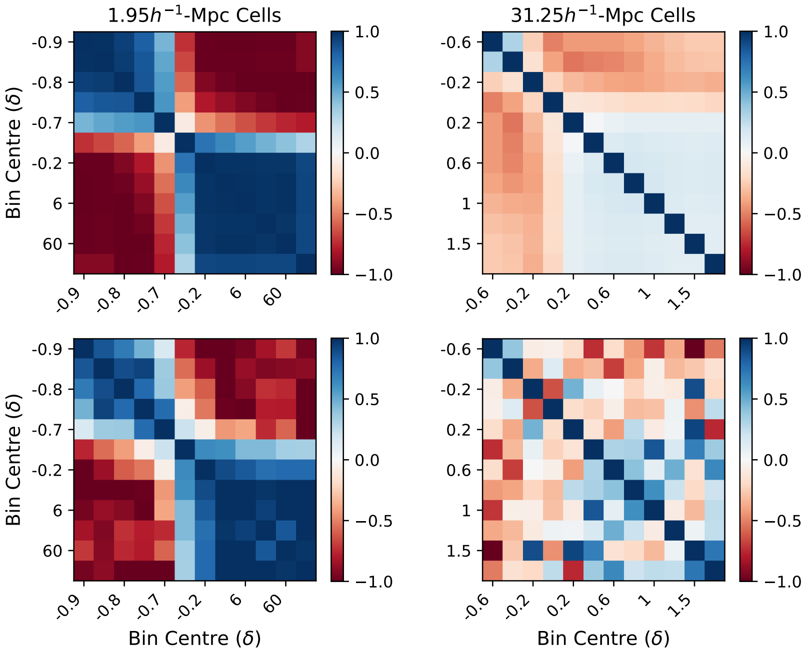
$$\sigma^2_{\mathcal{P}} = \frac{\mathcal{P}\,(1-\mathcal{P})}{N_c} + \frac{(N_c-1)\bar{\xi}^{\neq}_{\mathcal{I}}}{N_c}\mathcal{P}^2$$

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

$$N_C^2\,\bar{\xi}_{\mathcal{I}} = (N_c^2 - N_c)\bar{\xi}^{\neq}_{\mathcal{I}} + N_c\,\xi_{\mathcal{I}}(0)$$

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

$$= \frac{1}{N_c-1}\left(N_c\,\bar{\xi}_{\mathcal{I}} - \frac{1}{\mathcal{P}} + 1\right).$$



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 (GUTS)
- extract more information 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
- Recently: analytical expression of covariance matrix between levels: optimal estimator can be constructed
- The separation of density levels will help in understanding bias and redshift distortions.
- Future: MCMC parameter estimation

Summary

The StePS metod:

- zoom-in simulation with continuously and smoothly changing resolution.
- Hubble-volume simulations with just $\sim 10^7$ particles
- Comoving or static coordinates. Dark energy as a force proportional to the distance from the origin.
- Pairs or quartets can be used for sub-percent predictions
- rotating simulations

The Complementary-Original pairs:

- From any given initial condition
- The pair is opposite in terms of phases and amplitude
- cosmic variance estimation
- Average evolution with with reduced computational costs

Publications

Complementary Cosmological Simulations:

- Preprint: **Complementary Cosmological Simulations**, 2022, G. Rácz, A. Kiessling, I. Csabai, I. Szapudi *Astronomy & Astrophysics* 672, A59 (2023) https://arxiv.org/abs/2210.15077

StePS:

- **StePS: A Multi-GPU Cosmological N-body Code for Compactified Simulations**, 2019, G. Rácz, I. Szapudi, L. Dobos, I. Csabai, A. S. Szalay *Astronomy and Computing 28, [100303].* https://arxiv.org/abs/1811.05903
- **Compactified Cosmological Simulations of an Infinite Universe**, 2018, G. Rácz, I. Szapudi, I. Csabai, L. Dobos *Monthly Notices of the Royal Astronomical Society* 477, 1949-1957 https://arxiv.org/abs/1711.04959

Publications

Indicator Functions:

- **Indicator power spectra: surgical excision of non-linearities and covariance matrices for counts in cells**, 2022, A. Repp, I. Szapudi *Monthly Notices of the Royal Astronomical Society* 509, 586-594
  `https://arxiv.org/abs/2108.01673`

References I

Planck Collaboration.
Planck 2015 results. I. Overview of products and scientific results.
, 594:A1, September 2016.

Michael R. Blanton, David W. Hogg, Neta A. Bahcall, J. Brinkmann, Malcolm Britton, Andrew J. Connolly, István Csabai, Masataka Fukugita, Jon Loveday, Avery Meiksin, Jeffrey A. Munn, R. C. Nichol, Sadanori Okamura, Thomas Quinn, Donald P. Schneider, Kazuhiro Shimasaku, Michael A. Strauss, Max Tegmark, Michael S. Vogeley, and David H. Weinberg.
The Galaxy Luminosity Function and Luminosity Density at Redshift z = 0.1.
, 592(2):819–838, August 2003.

References II

Andrew Pontzen, Anže Slosar, Nina Roth, and Hiranya V. Peiris.
Inverted initial conditions: Exploring the growth of cosmic structure and voids.
, 93(10):103519, May 2016.

Raul E. Angulo and Andrew Pontzen.
Cosmological N-body simulations with suppressed variance.
, 462(1):L1–L5, October 2016.

Juhan Kim, Changbom Park, III Gott, J. Richard, and John Dubinski.
The Horizon Run N-Body Simulation: Baryon Acoustic Oscillations and Topology of Large-scale Structure of the Universe.
, 701(2):1547–1559, August 2009.