Bayesian large-scale structure inference: initial conditions and cosmic voids

Florent Leclercq

Institut d'Astrophysique de Paris
Institut Lagrange de Paris
École polytechnique ParisTech

April 9th, 2014

In collaboration with:

Nico Hamaus (IAP/U. Illinois), Jens Jasche (IAP), Alice Pisani (IAP), Emilio Romano-Díaz (U. Bonn), Paul M. Sutter (IAP/Ohio State U.), Benjamin Wandelt (IAP/U. Illinois)

Why cosmology?

- Cosmology is the science of the Universe as a physical system, where “the Universe” means “everything that exists in the physical sense”.
  - Matter ✓
  - Ideas ✗
  - Laws ?
- Important ideas:
  - The Universe in its globality can be treated as a physical system
  - Science can deal with times and places we cannot experience (the observable Universe is a strict subset of the Universe)

Some specificities of cosmology

- **Unicity**. The experience is unique and irreproducible by physical experimentation. There is no exteriority nor anteriority. The properties of the Universe cannot be determined statistically on a set.
- **Energy**. The energy scales at stake in the Early Universe are orders of magnitude higher than anything we can reach on Earth.
- **Arrow of time**. Reasoning in cosmology is "bottom-up". The final state is known and the initial state has to be inferred.

⇒ The **initial conditions** of the Universe have a **particular status** with respect to other physical phenomena.

Cosmostatistics of the initial conditions

- **"Initial conditions"**: ICs for *gravitational evolution*…
  - AFTER inflation
  - AFTER Hot Big Bang phenomena

(primordial nucleosynthesis, decoupling, recombination, free-streaming of neutrinos, acoustic oscillations of the photon-baryon plasma, transition from radiation to matter dominated universe…)

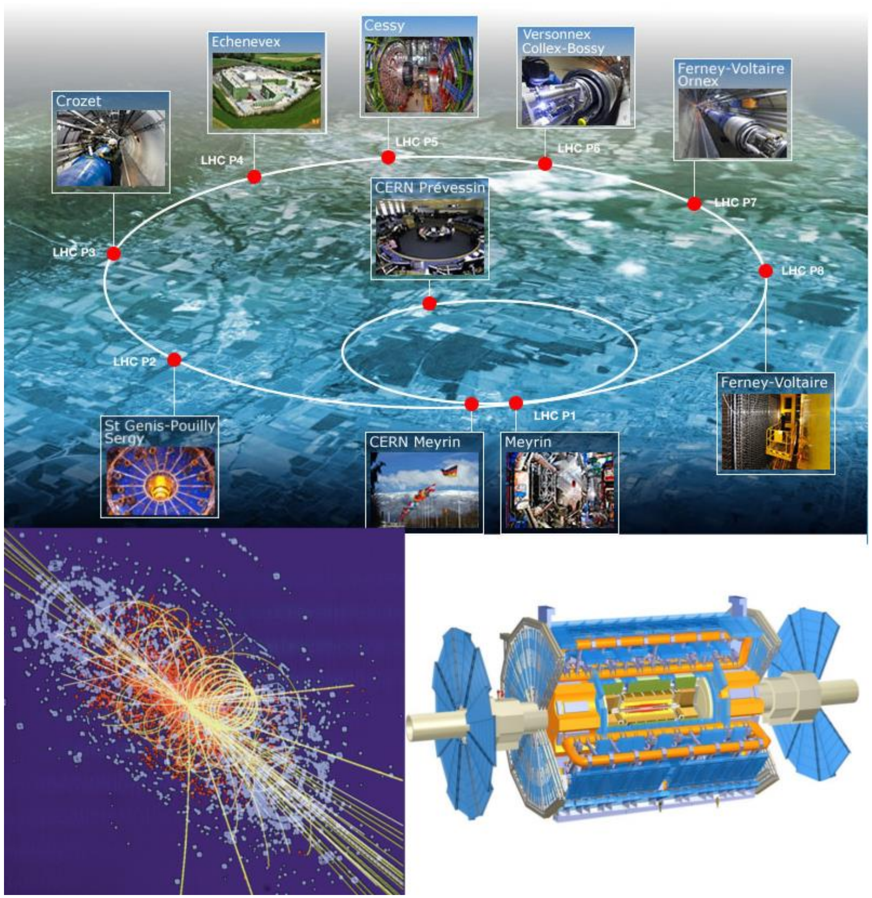
- **Cosmostatistics**: discipline dealing with stochastic quantities as seeds of structure in the Universe
  - prediction of cosmological observables from random inputs

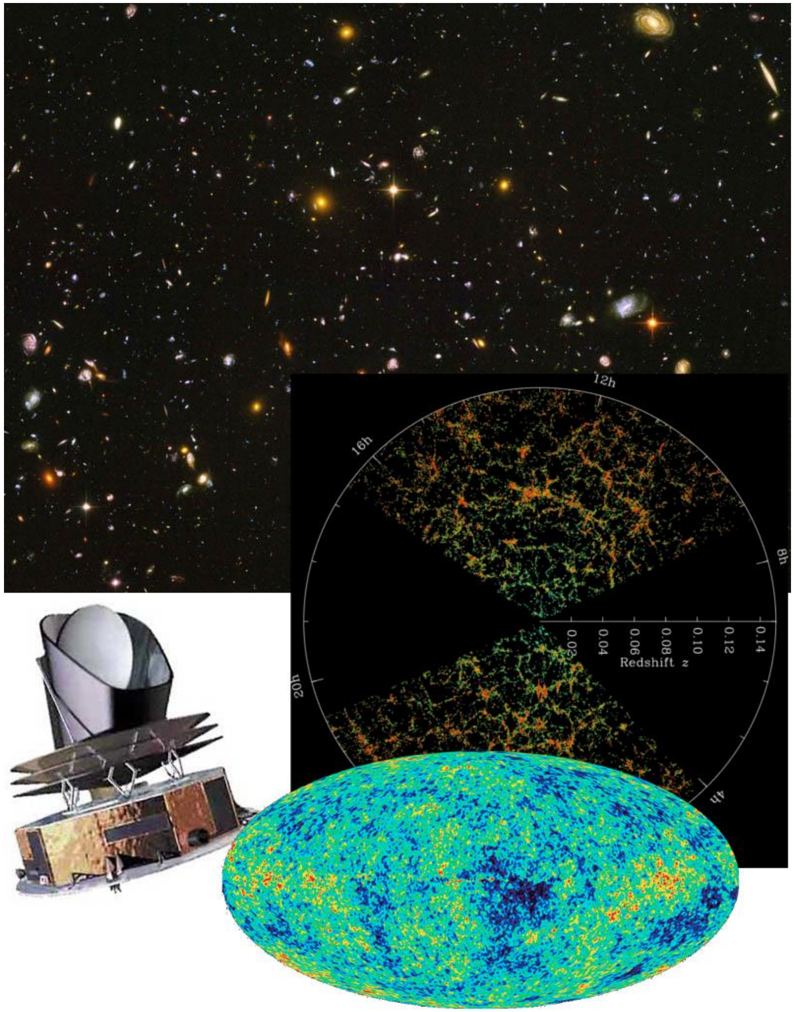
    (*from theory to data*)
  - use of the departures from homogeneity in astronomical surveys to distinguish between cosmological models

    (*from data to theory*)

see also FL, Pisani & Wandelt 2014, arXiv:1403.1260

High energy physics experiments



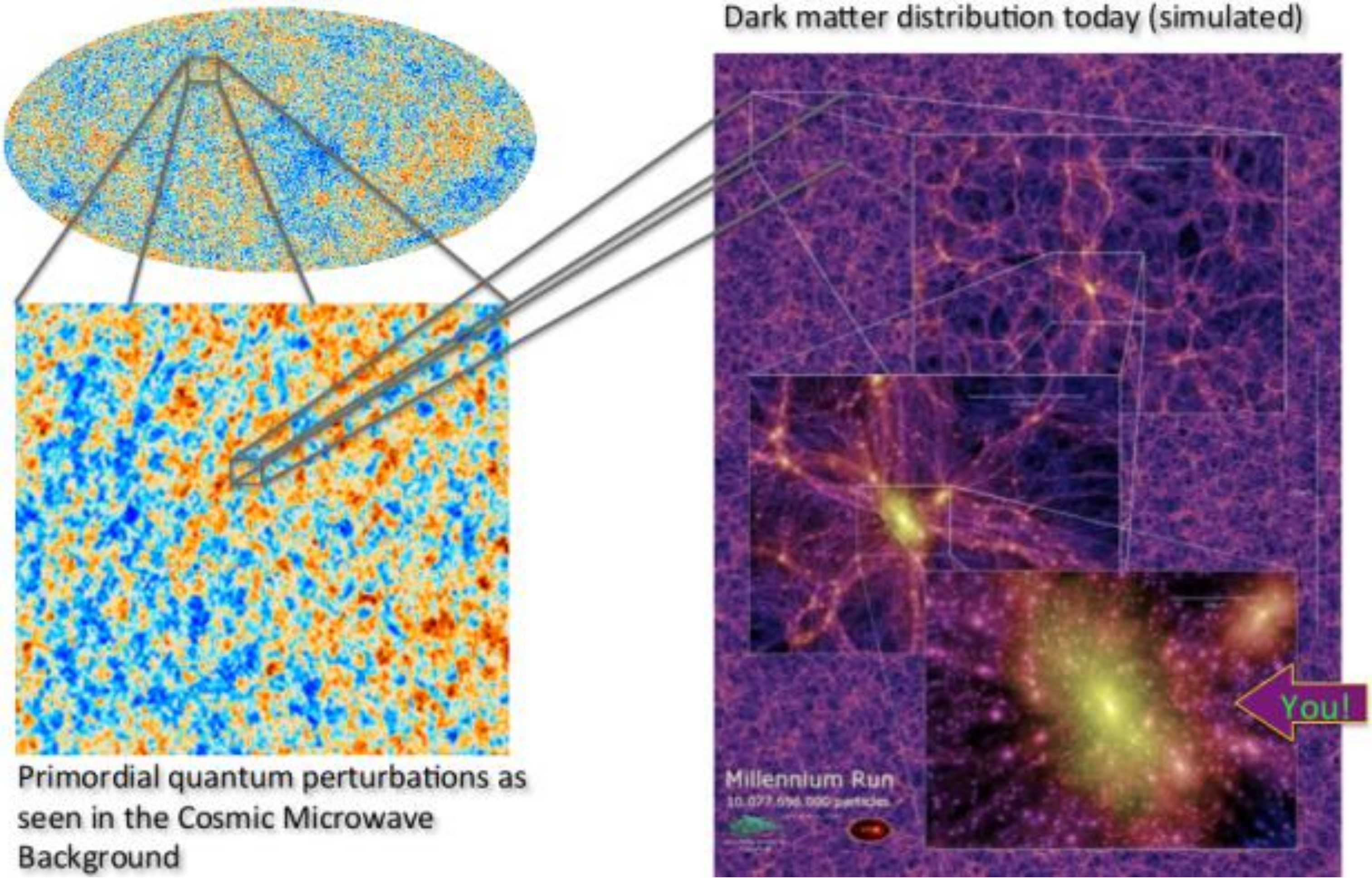


The inhomogeneous Universe: the big picture

You are here, make the best of it…

Planck

Dark matter distribution today (simulated)



Primordial quantum perturbations as seen in the Cosmic Microwave Background

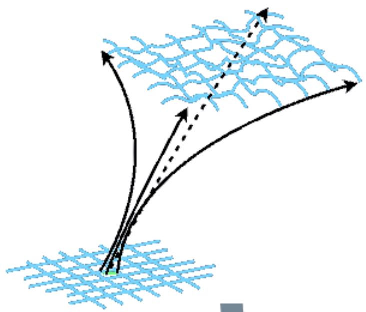
A call to modesty...

"*Hominem te esse*"

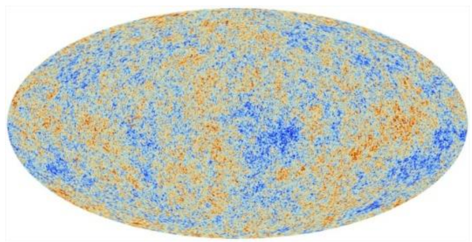
"*Memento mori*"

Inflation as the origin of structure

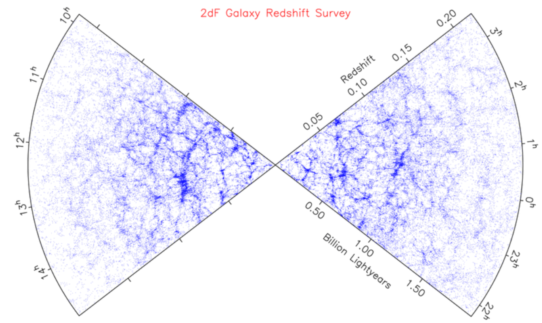
quantum fluctuations



matter and energy perturbations



clusters, halos, galaxies



inflation

gravitational collapse

Probe of this era

Statistics of these

$\delta\varphi$

inflaton field

Φ

curvature potential

δT
$\delta \rho$
δn

CMB

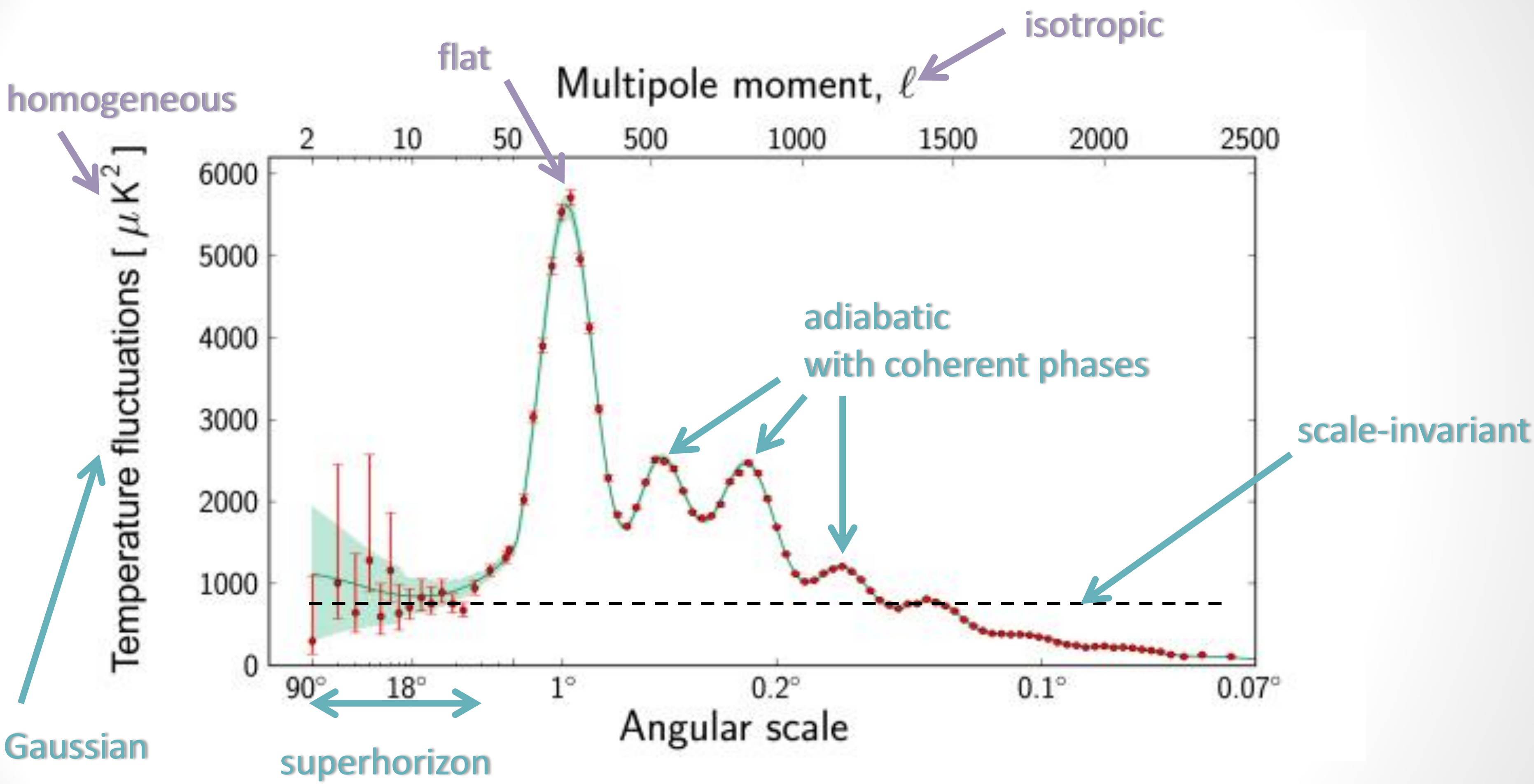
matter

galaxies

Powerful probes of inflation

if we can accurately model some relationships

The most boring Universe?



Planck collaboration 2013, arXiv:1303.5062

- Phenomenologically, inflation is a great success…

The case for physical reconstruction of the ICs

- … but what is the microphysics of inflation?
- Some challenges:
  - The eta problem: scale-invariant, superhorizon fluctuations require

    $$\eta = M_{\rm Pl}^2 \frac{V''}{V} = \frac{m_\phi^2}{3H^2} \ll 1$$

    How to achieve and stabilize this mass hierarchy?

  - Large-field inflation: observational gravitational waves mean

    $$r \approx 0.2 \quad \longleftrightarrow \quad \Delta\phi \gg M_{\rm Pl}$$

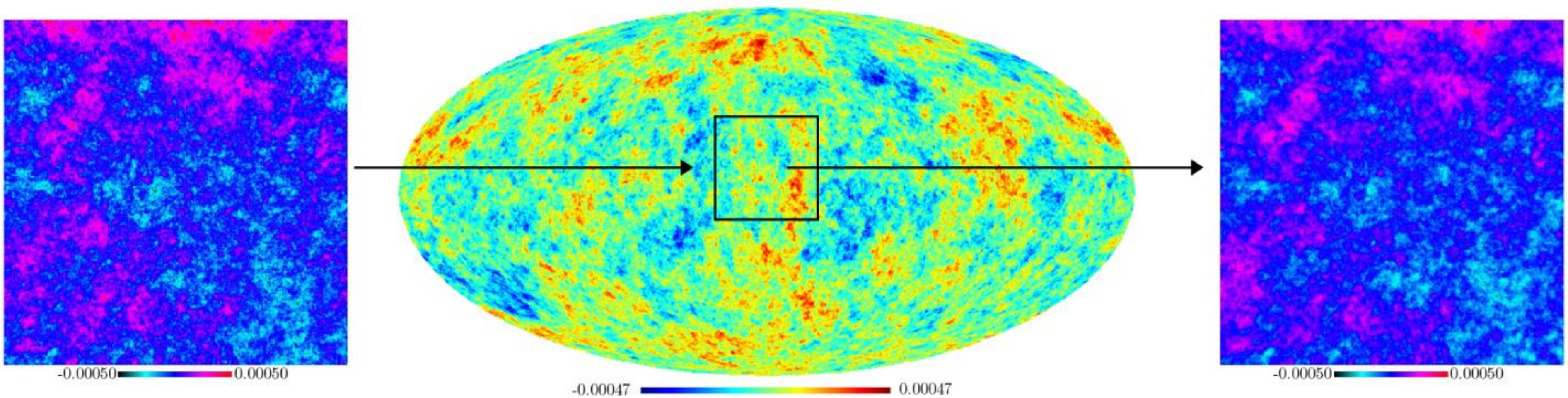
    Astrophysics — BICEP2 collaboration 2014, arXiv:1403.3985

    Quantum gravity — *Lyth bound,* Lyth 1997, arXiv:hep-ph/9606387

- Some open questions: *multi-field inflation? non-standard kinetic term? periods of fast-roll? non-trivial pre-inflationary state? non-Bunch-Davies vacuum?*

The CMB time-machine

- A time-machine (380,000 yrs ⇨ 10^{-35} s): linear perturbation theory



Input gravitational potential

adapted from Elsner & Wandelt 2009, arXiv:0909.0009

Reconstruction of the gravitational potential

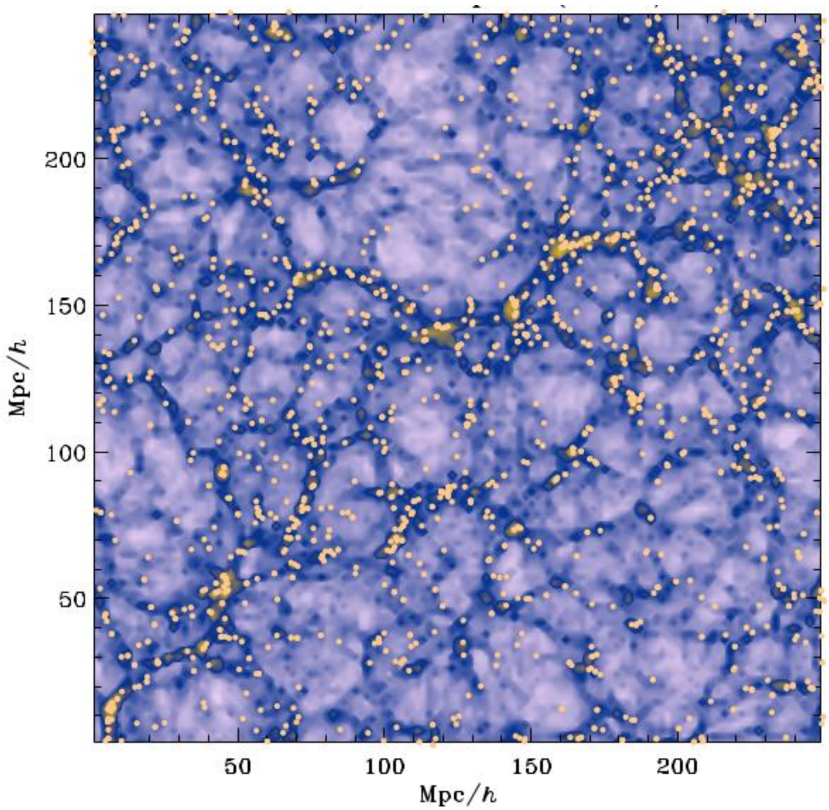
Komatsu, Spergel & Wandelt 2005, arXiv:astro-ph/0305189

Yadav & Wandelt 2005, arXiv:astro-ph/0505386

- Relies on:
  - Gaussian random fields
  - Linear transfer
  - Optimal inference of a GRF from a GRF: Wiener filtering

see also FL, Pisani & Wandelt 2014, arXiv:1403.1260

A large-scale structure in the Universe



Blue: matter distribution
Orange: dark matter halos / galaxies

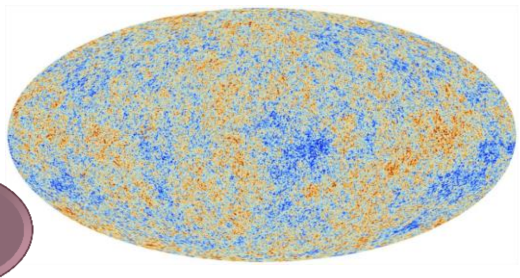
- Halos trace mass distribution (of *dark matter*).
- Halos are NOT randomly distributed: there exists a Large Scale Structure of the Universe
- How do we analyze this structure quantitatively?

Correlation functions and Fourier analysis

Reconstruction of the initial conditions…

- … a solved problem!
- And…

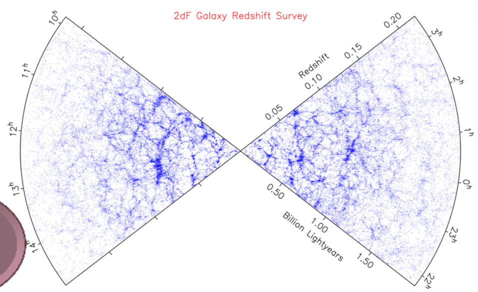
CMB:



2D

$$N_{\mathrm{mode}}^{\mathrm{CMB}} \propto l_{\mathrm{max}}^{2}$$

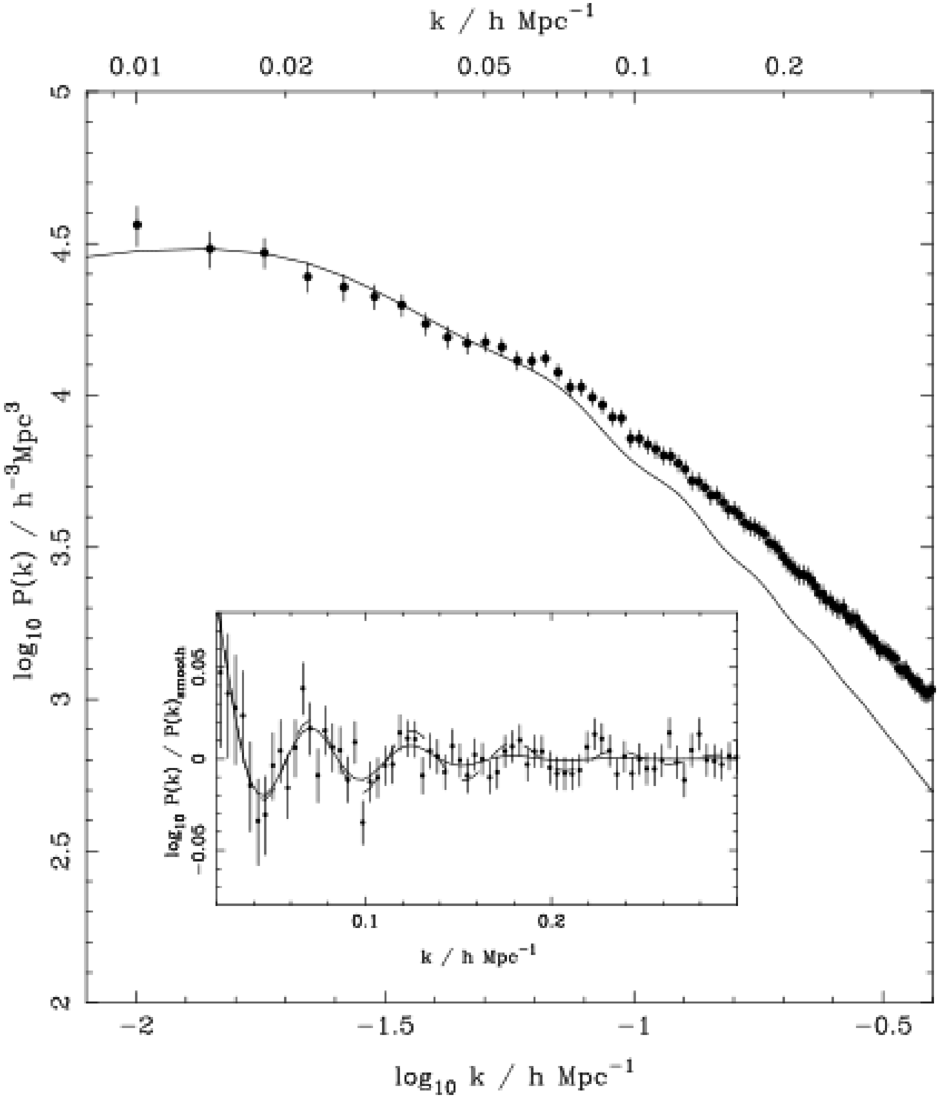
LSS:



3D

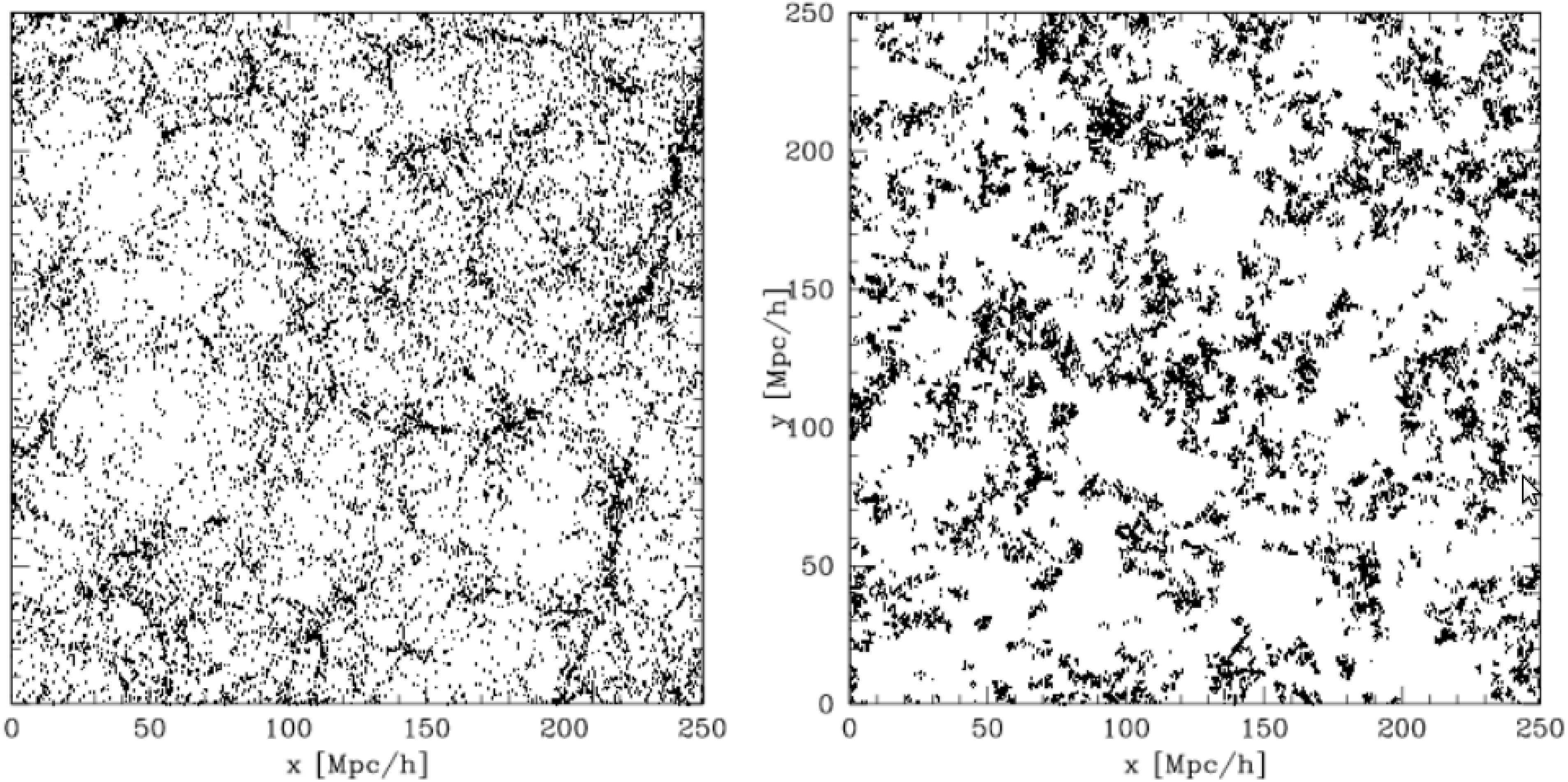
$$N_{\mathrm{mode}}^{\mathrm{LSS}} \propto k_{\mathrm{max}}^{3}$$

Where the Universe becomes non-Gaussian…



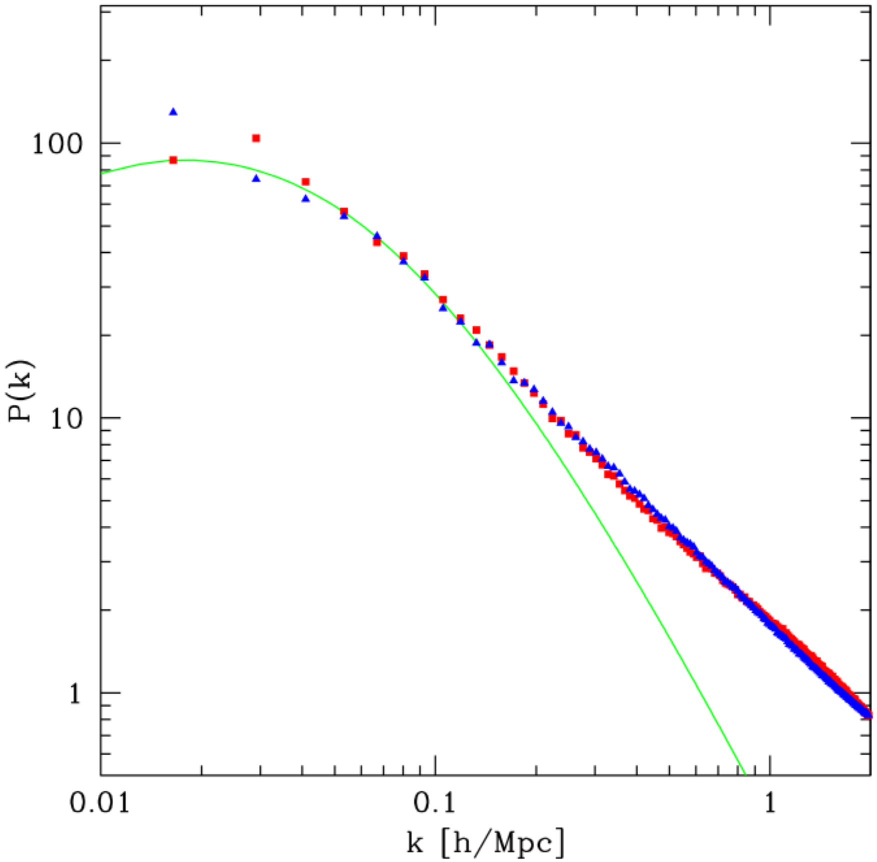
Percival *et al.* 2006, arXiv:astro-ph/0608636

Gaussian vs non-Gaussian information

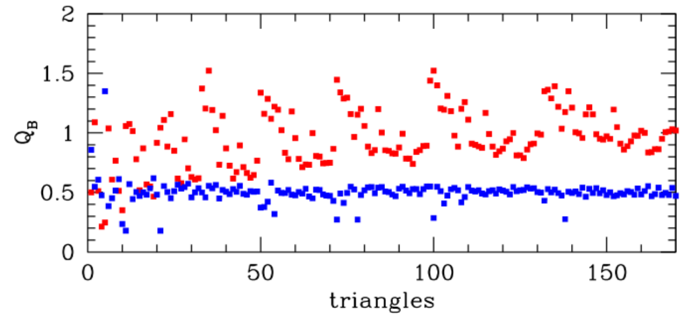


from R. Scoccimarro

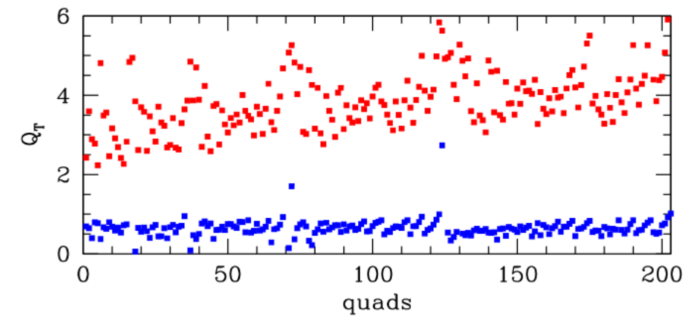
Gaussian vs non-Gaussian information



The two distributions have about the same power spectrum!



bispectrum



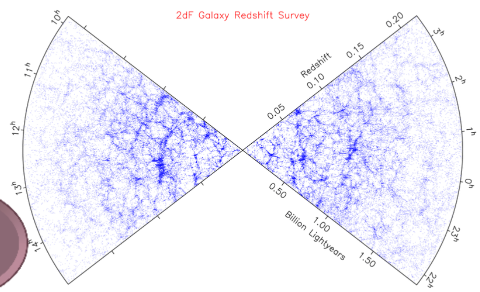
trispectrum

from R. Scoccimarro

Reconstruction of the initial conditions…

LSS:

3D



$$N_{\rm mode}^{\rm LSS} \propto k_{\rm max}^3$$

- The challenges : non-linearity and non-Gaussianity
  - Non-linear transfer functions in the Hot Big Bang phenomena
  - Gravitational evolution
  - Primordial non-Gaussianity (…?)
  - Data imperfection and systematics…

Can we go from the linear to the non-linear problem?

Bayesian inference of the ICs

- Why do we need Bayesian inference?

Inference of signals = ill-posed problem

- Noise
- Incomplete observations: survey geometry, selection effects
- Systematic uncertainties, biases
- Cosmic variance

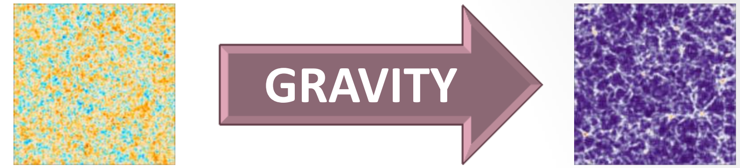
⇨ **No unique recovery is possible!**

"What are the initial conditions of the Universe?" ⇨ "What is the probability distribution of possible initial conditions (signals) compatible with the observations?"

$$p(s|d)p(d) = p(d|s)p(s)$$

Bayesian inference of the ICs

- Physical motivation:
  - Complex final state, simple initial state
  - A "**forward only**" problem
    (we have a generative model for the final state)

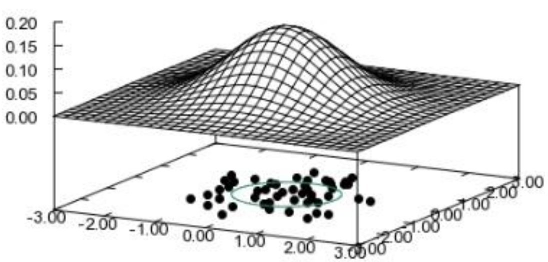


Initial state Final state

- Problems:
  - Highly dimensional inference (10^7 parameters)
  - A large number of **correlated** parameters

⇨ **No reduction of the problem size is possible!**

  - Potentially complex posterior distribution
- Numerical approximation: **sampling** the posterior



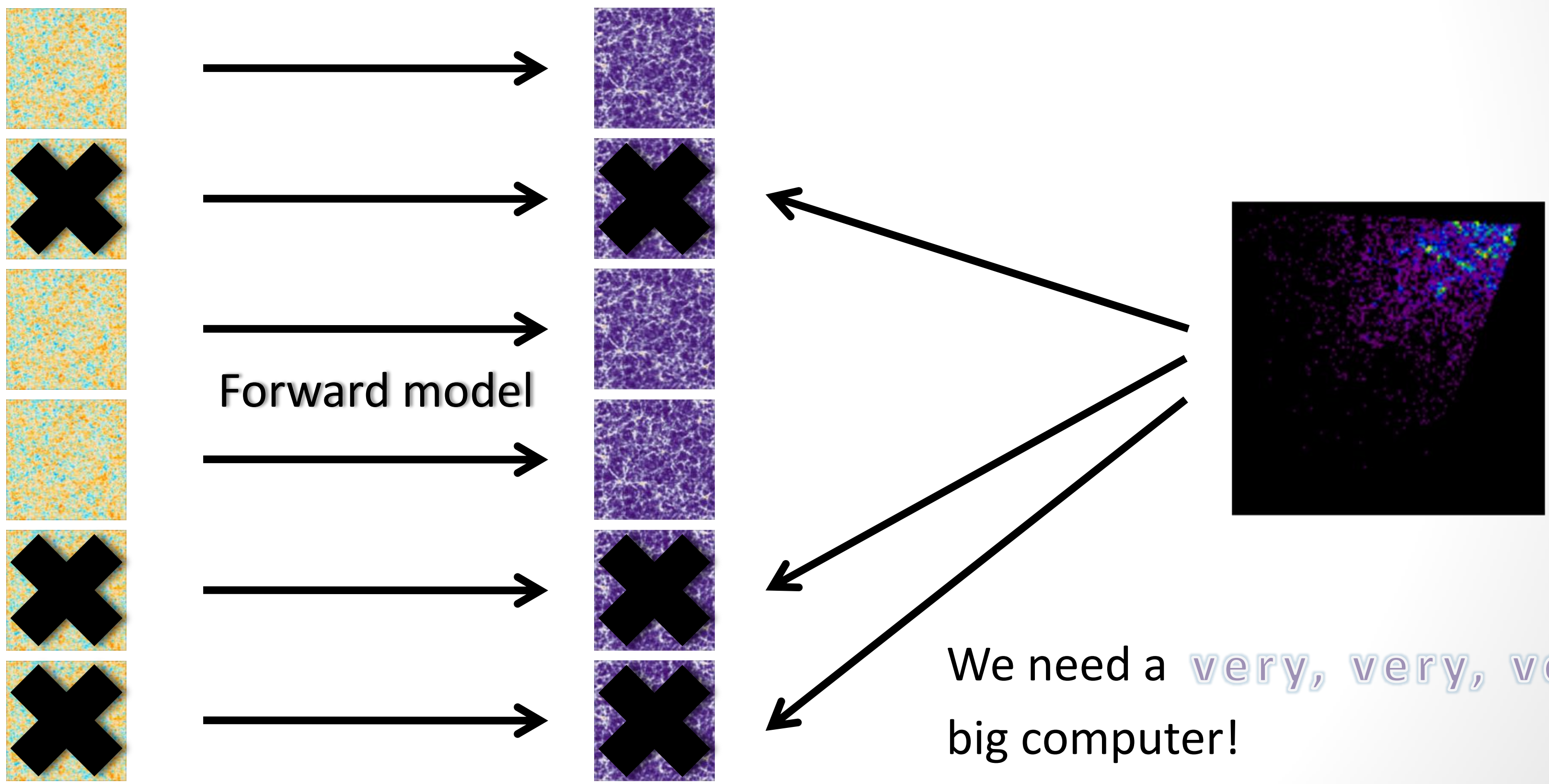
$$p(s|d) \to p_N(s|d) = \frac{1}{N}\sum_{i=1}^{N}\delta_{\mathrm{D}}(s-s_i)$$

- But how to "get the dots" ?

4D physical inference of the ICs

- The ideal scenario:

Forward model = N-body simulation + Halo occupation + Galaxy formation + Feedback + …



We need a very, very, very big computer!

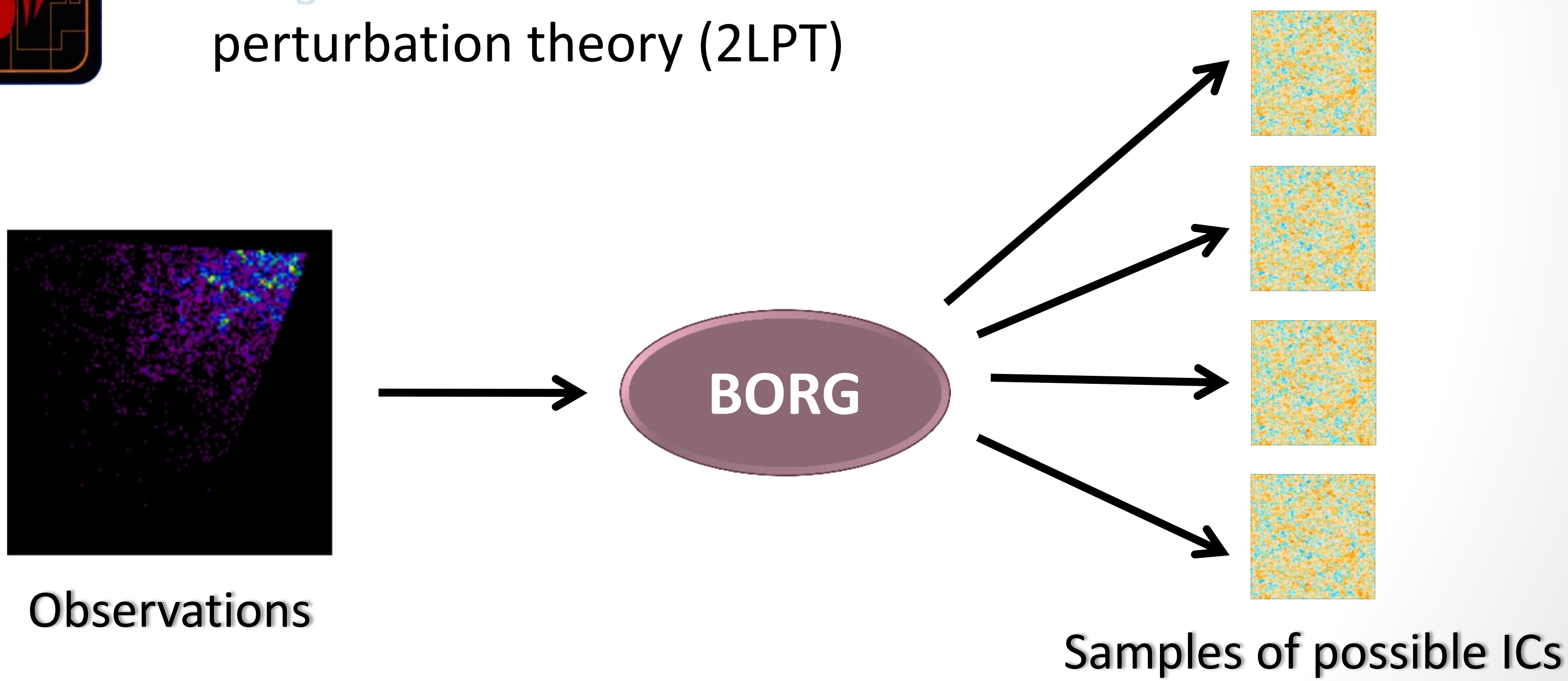
All possible ICs

All possible FCs

BORG: *Bayesian Origin Reconstruction from Galaxies*

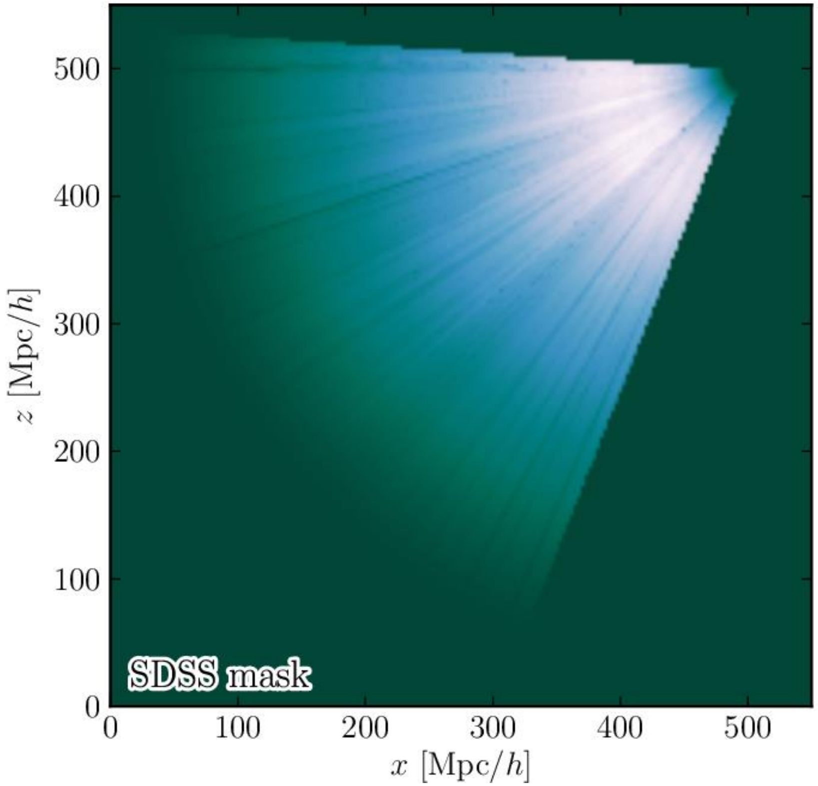
What makes the problem tractable:

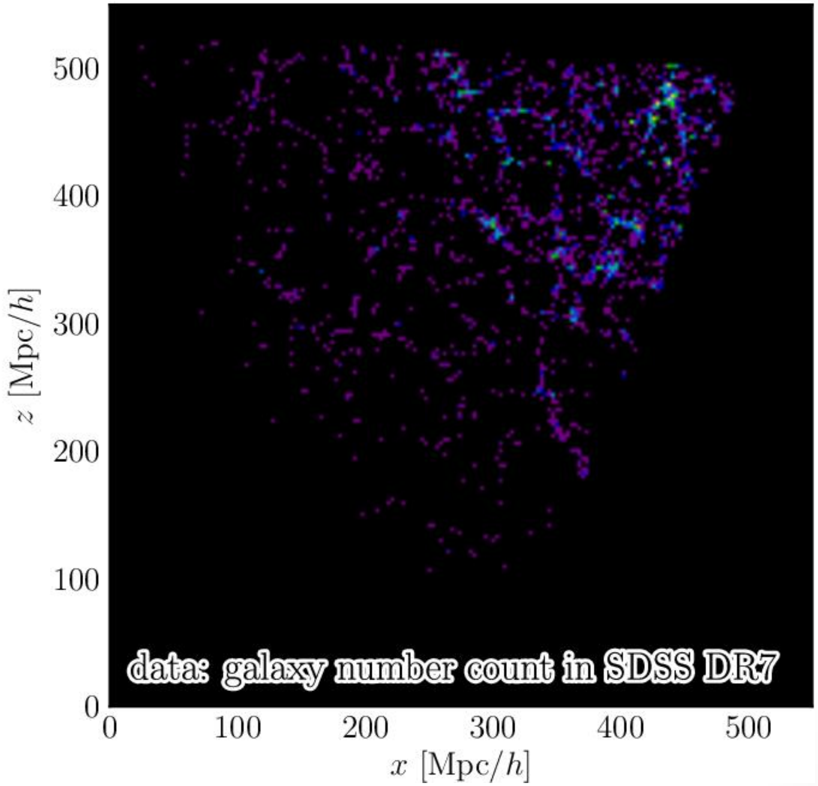
- **Sampler**: Hamiltonian Markov Chain Monte Carlo method
- **Physical model**: Second-order Lagrangian perturbation theory (2LPT)



Jasche & Wandelt 2012, arXiv:1203.3639

BORG: reconstructions from SDSS DR7

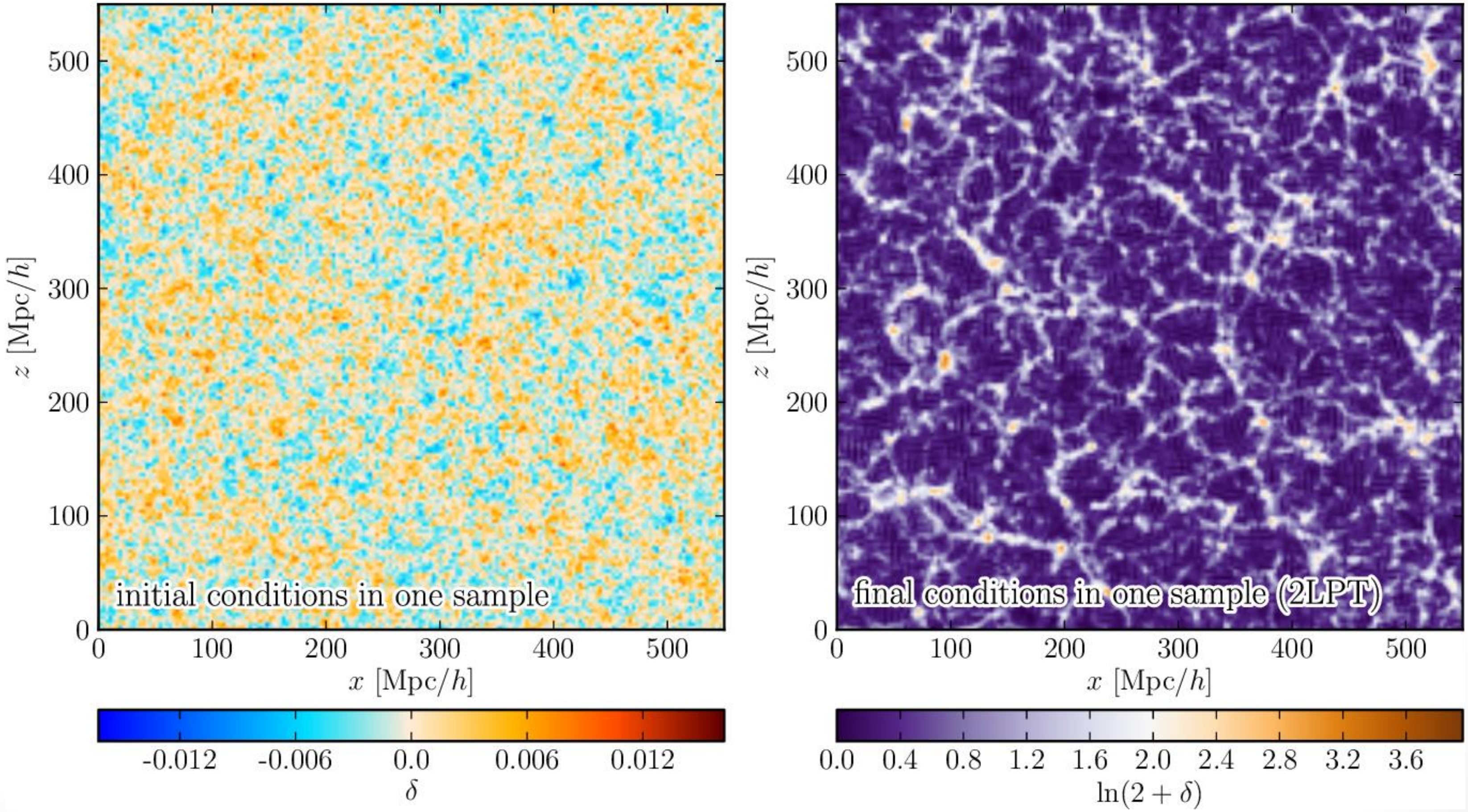




Data

Jasche, FL & Wandelt, in prep.

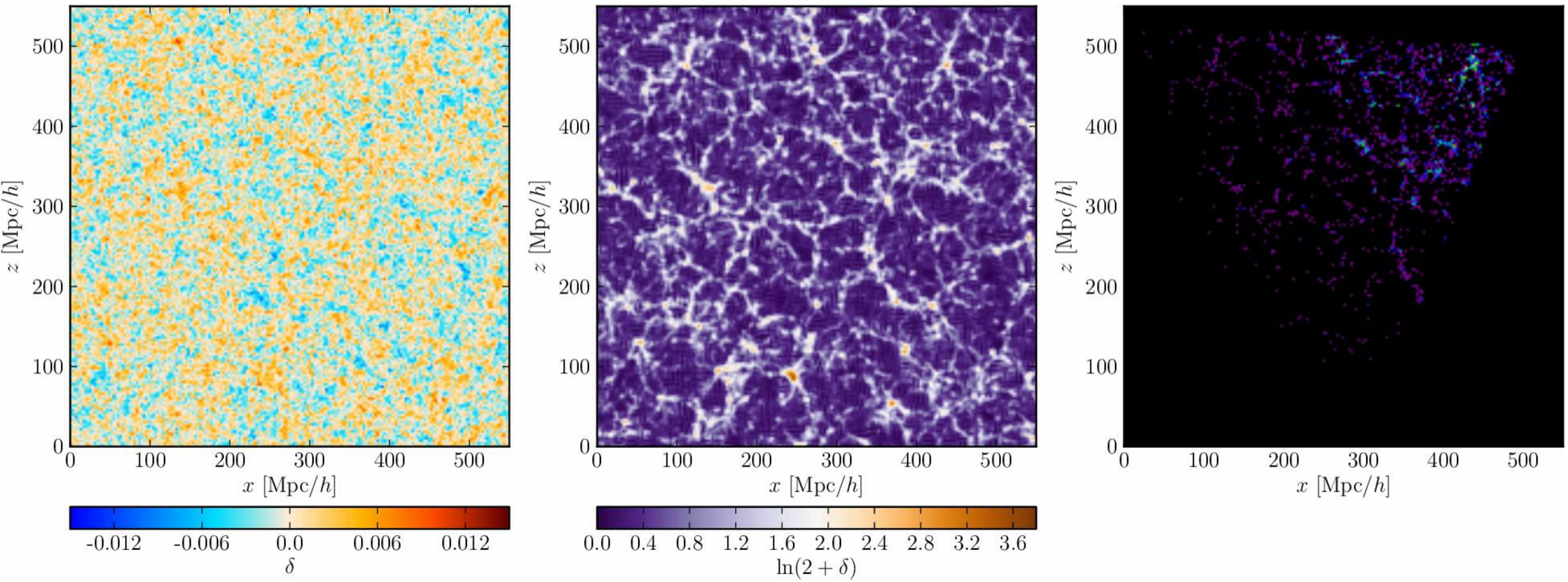
BORG: reconstructions from SDSS DR7



One sample

Jasche, FL & Wandelt, in prep.

BORG at work



Initial conditions

Final conditions

Observations

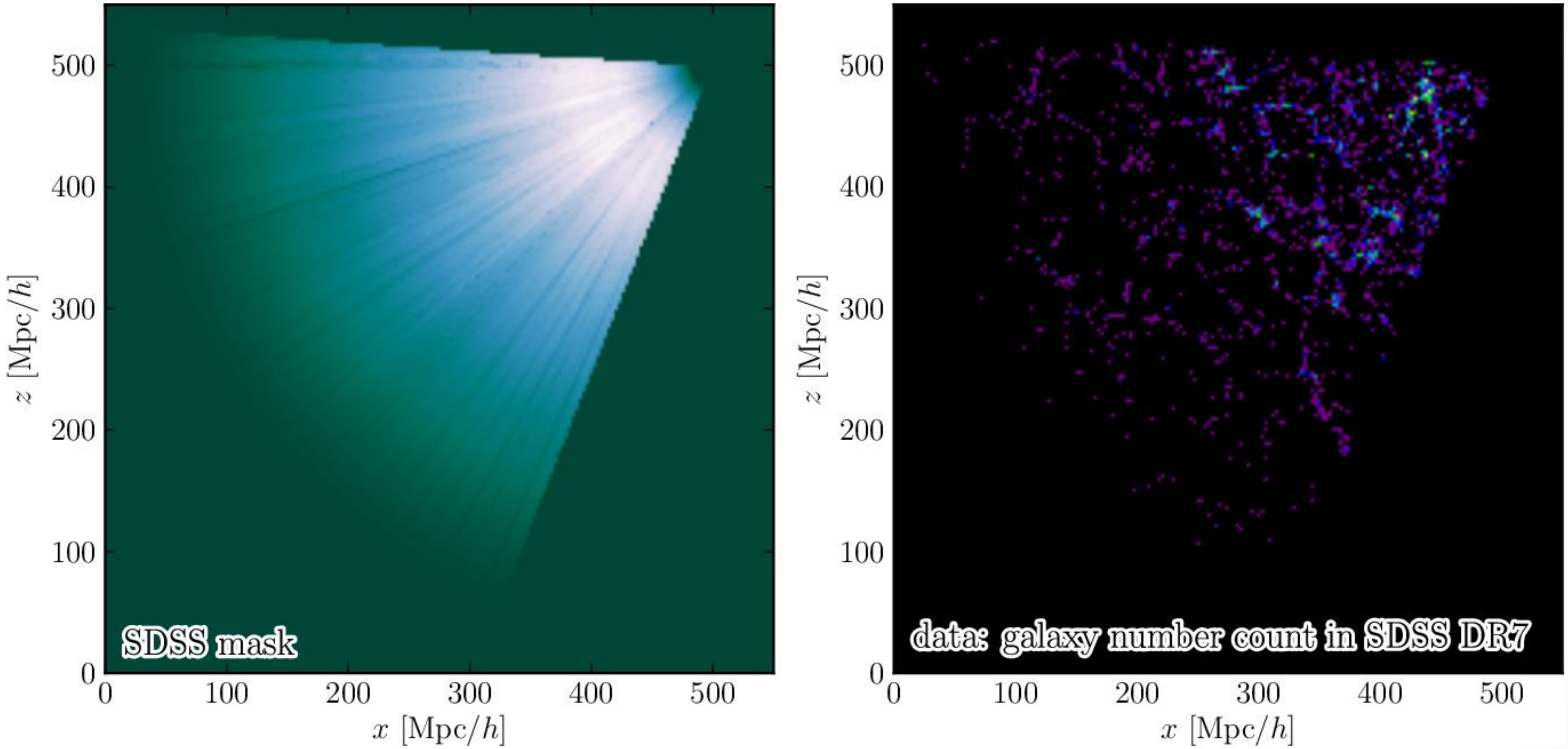
Samples of the posterior density

- Each sample: a **possible version of the truth**

- The **variation** between samples **quantifies the uncertainty** that results from having
  - only one Universe (a more precise version of "cosmic variance")
  - incomplete observations (mask, finite volume and number of galaxies, selection effects)
  - imperfect data (noise, biases, photometric redshifts…)

see also FL, Pisani & Wandelt 2014, arXiv:1403.1260

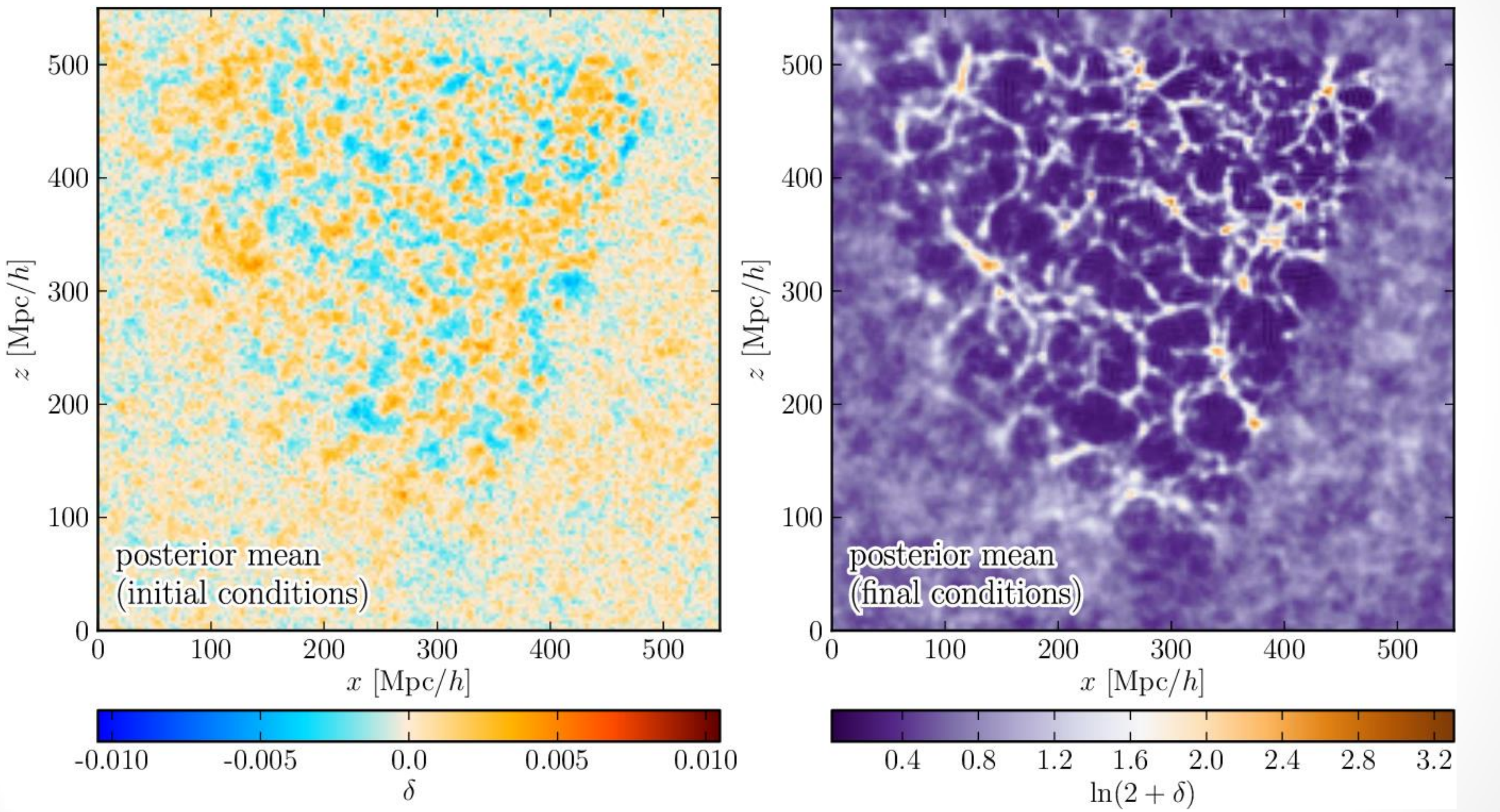
BORG: reconstructions from SDSS DR7



Data

Jasche, FL & Wandelt, in prep.

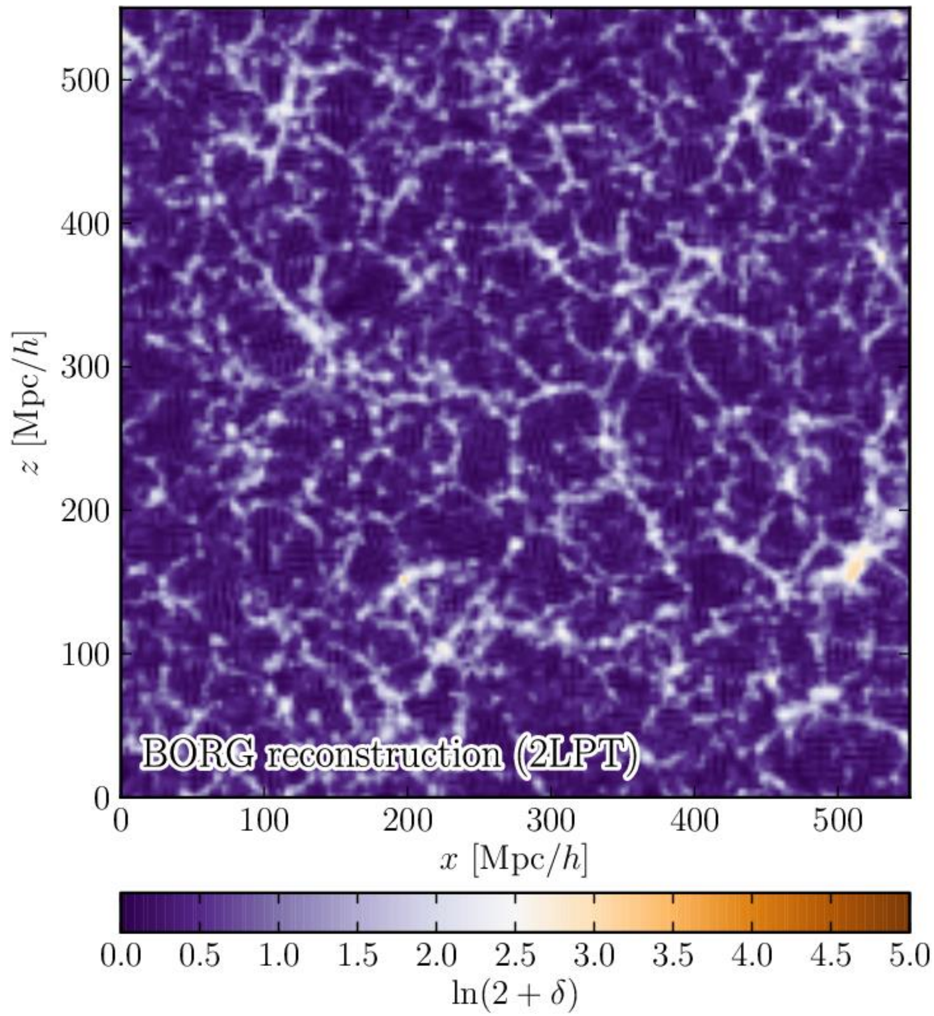
BORG: reconstructions from SDSS DR7



Posterior mean

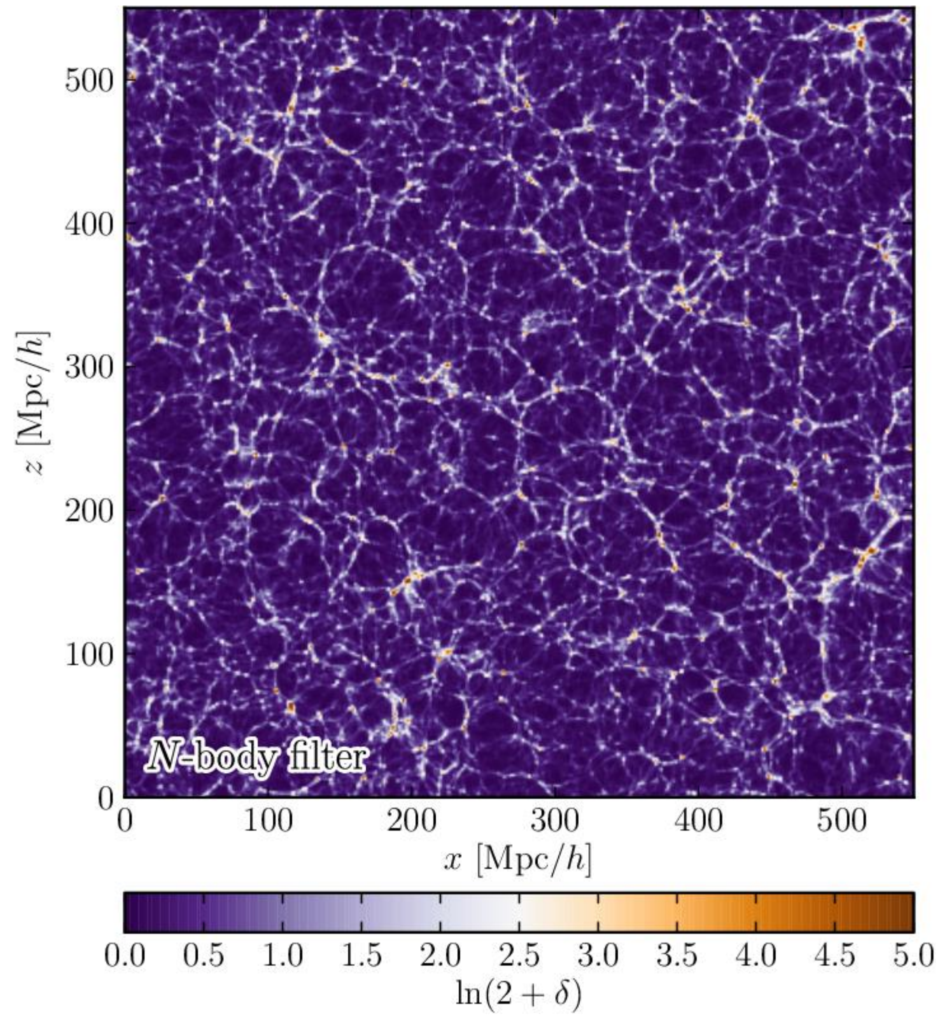
Jasche, FL & Wandelt, in prep.

Data-constrained non-linear realizations



Jasche, FL, Romano-Diaz & Wandelt, in prep.

Data-constrained non-linear realizations

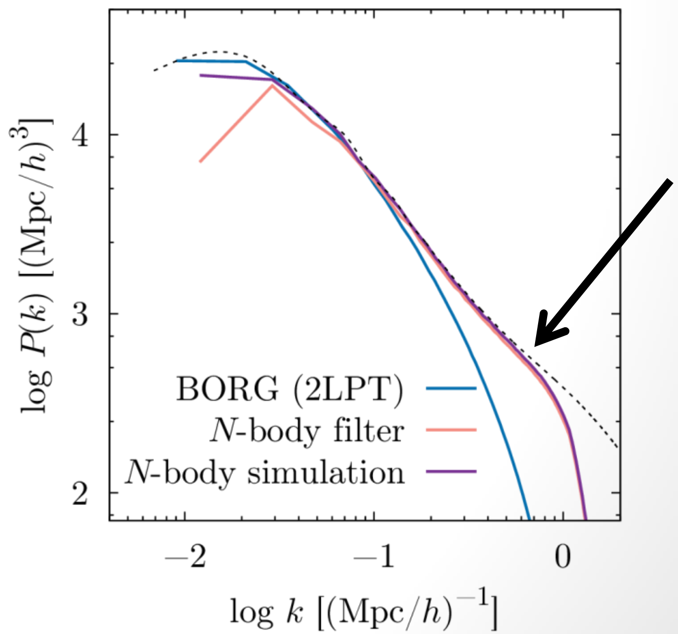


Jasche, FL, Romano-Diaz & Wandelt, in prep.

Data-constrained non-linear realizations

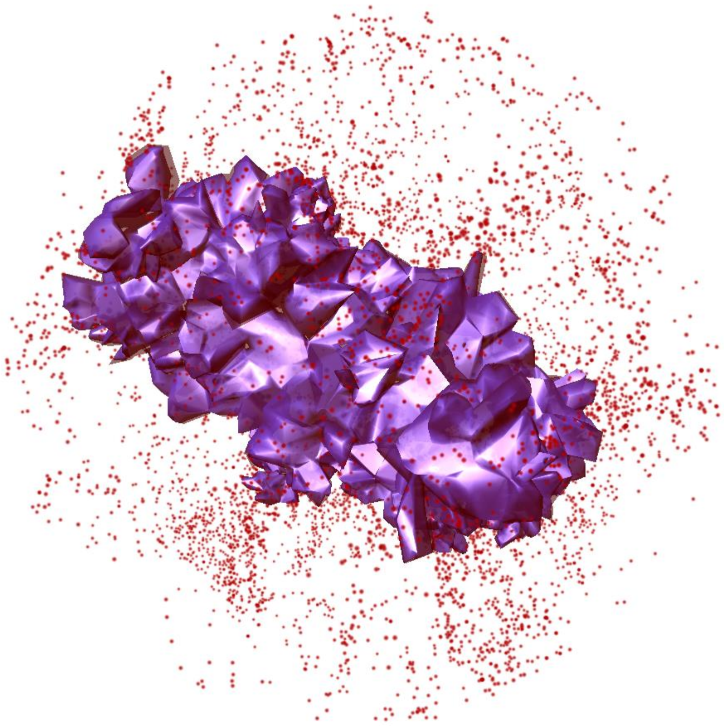
- A dynamic physical model naturally introduces some correlations between the constrained and unconstrained parts
- Constrained resimulations act as hypothesis generating machines, whose predictions can be tested with complementary observations in the actual sky.

- With a full N-body simulation, we address the non-linear regime of structure formation!



Jasche, FL, Romano-Diaz & Wandelt, in prep.

Cosmology with voids



Sutter, Lavaux, Wandelt & Weinberg 2012, arXiv: 1207.2524

- Public void catalog from the SDSS DR7 and DR9

Sutter, Lavaux, Wandelt & Weinberg 2012, arXiv: 1207.2524

Sutter, Lavaux, Wandelt, Weinberg & Warren 2013, arXiv:1310.7155

http://www.cosmicvoids.net

- Science results
  - Alcock-Paczynski test

Sutter, Lavaux, Wandelt & Weinberg 2012, arXiv: 1208.1058

  - Void-galaxy correlations

Hamaus, Wandelt , Sutter, Lavaux & Warren 2013, arXiv: 1307.2571

  - Density profile

Pisani, Lavaux , Sutter & Wandelt 2013, arXiv: 1307.2571

Hamaus, Sutter & Wandelt 2014, arXiv:1403.5499

  - Integrated Sachs-Wolfe effect

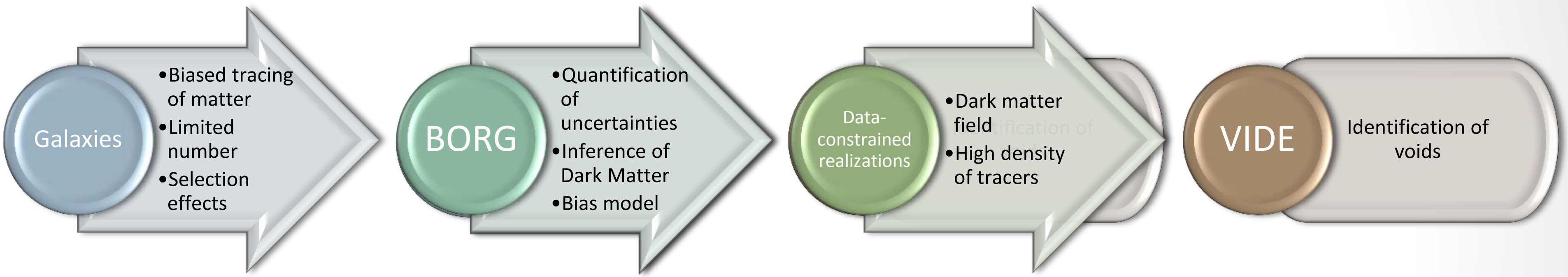
Planck collaboration 2013, arXiv:1303.5079

  - Gravitational lensing by voids

Melchior, Sutter, Sheldon, Krause & Wandelt 2013, arXiv:1309.2045

Dark matter voids in the SDSS galaxies

- How?

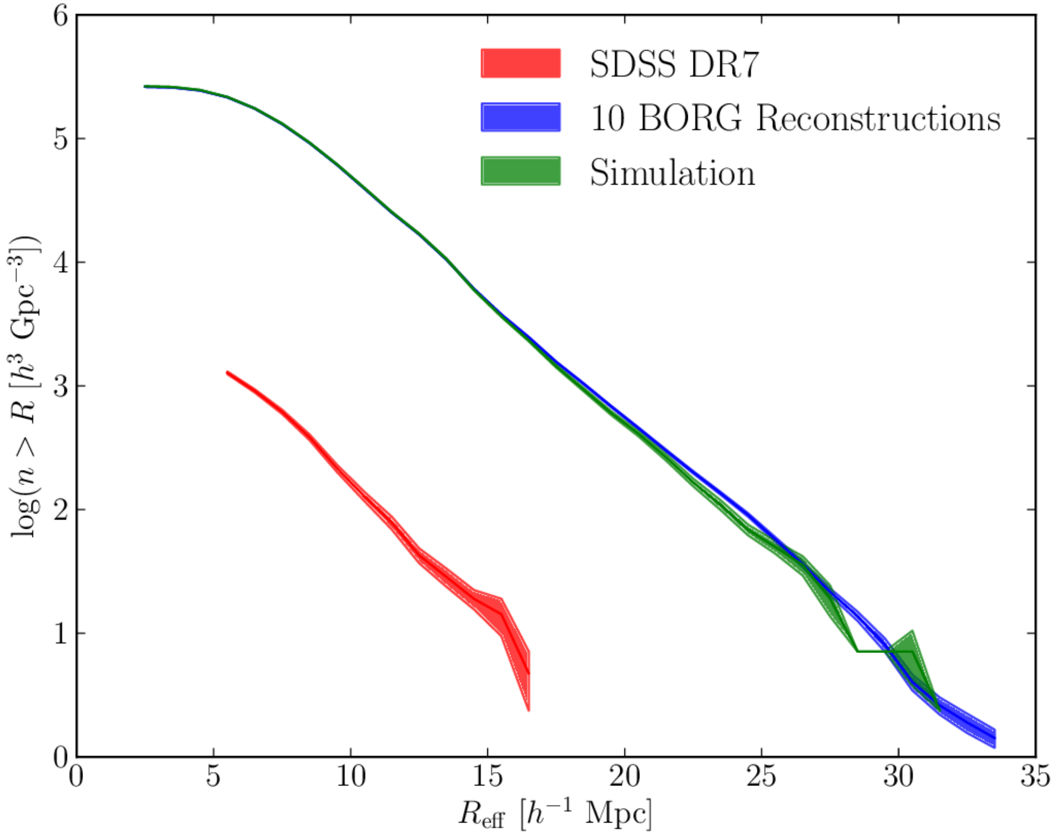


- Why? What is made possible by our technology:
  - Bias. Voids are defined in the dark matter distribution, not in galaxies.
  - Shot noise. Galaxies sparsely sample the dark matter distribution. We get 10x more dark matter voids than galaxy voids.

FL, Jasche, Sutter, Hamaus & Wandelt, in prep.

Properties of dark matter voids

- For usual void statistics, results are consistent with *N*-body simulations.

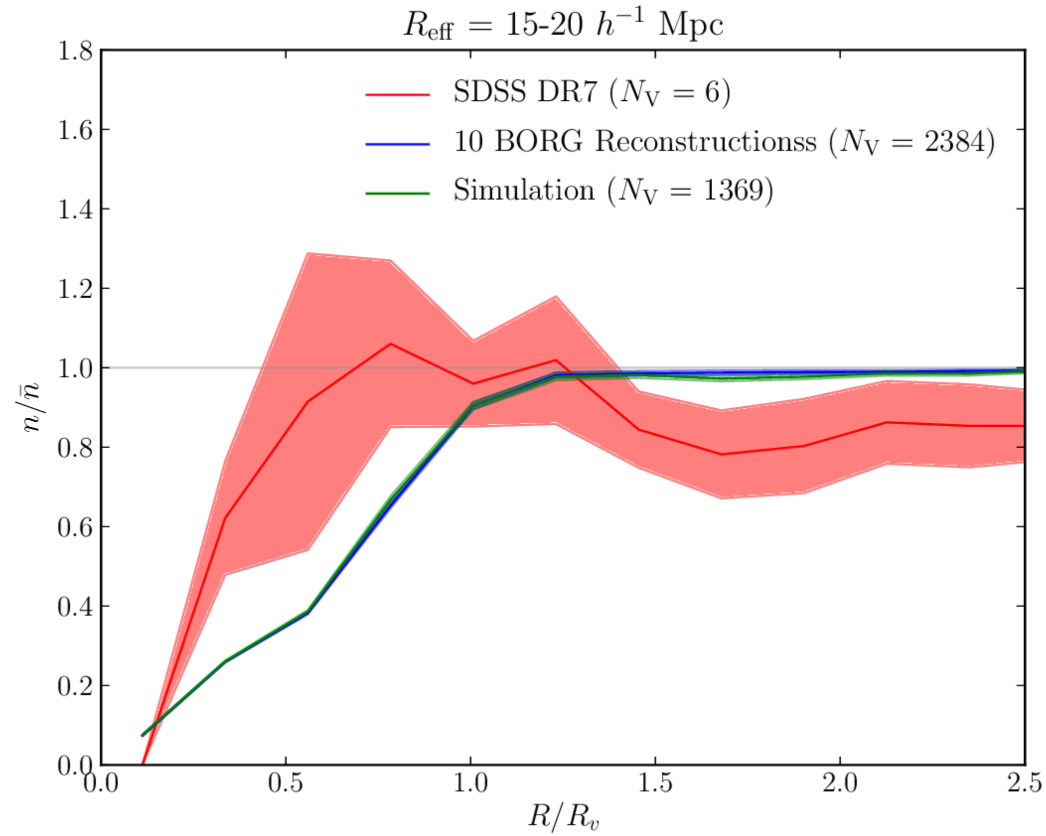


void number count

FL, Jasche, Sutter, Hamaus & Wandelt, in prep.

Properties of dark matter voids

- For usual void statistics, results are consistent with *N*-body simulations.

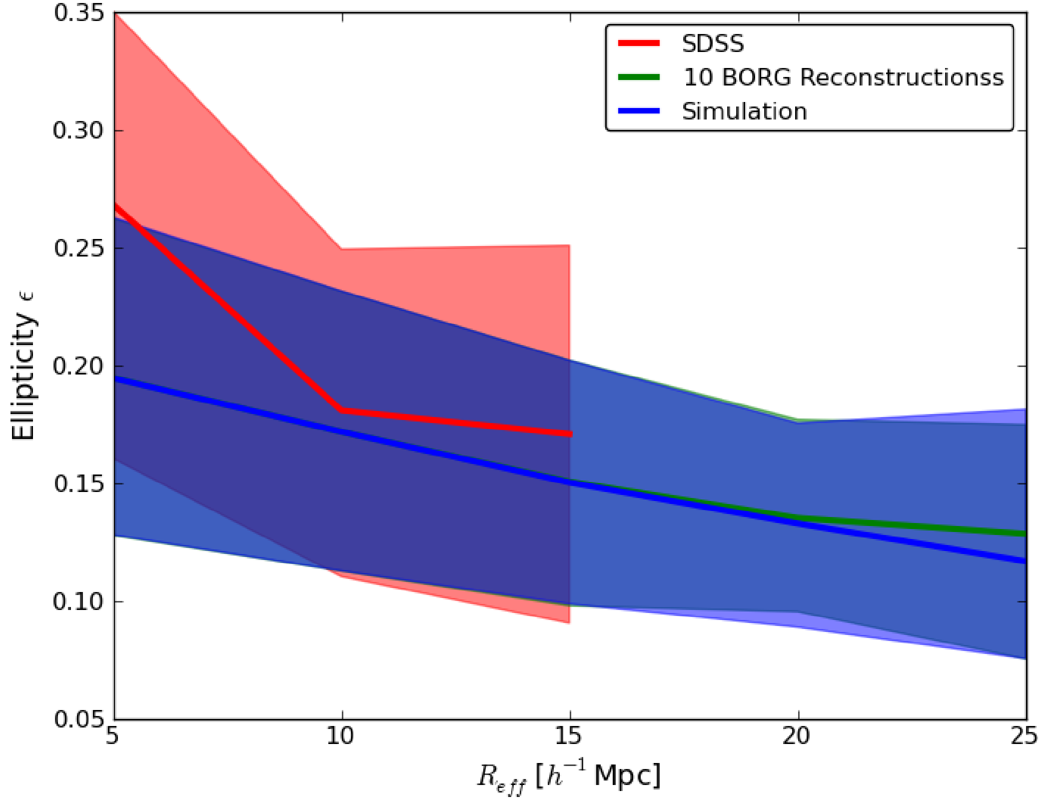


void density profile

FL, Jasche, Sutter, Hamaus & Wandelt, in prep.

Properties of dark matter voids

- For usual void statistics, results are consistent with *N*-body simulations.



void ellipticities

FL, Jasche, Sutter, Hamaus & Wandelt, in prep.

Concluding thoughts

- Cosmological physical reconstruction of the initial conditions of the Universe is becoming feasible.

BORG: A non-linear time machine using Bayesian posterior exploration to infer primordial quantities from late-time observations.

- An new, enhanced dark matter voids catalog.
- Additional great science is waiting behind the door.
  - Baryon acoustic oscillations, clusters, galaxies
  - Non-Gaussianity
  - Isocurvature perturbations
  - Gravitational waves in the large-scale structure…

Don't fight non-linearity to get cosmological information – embrace it!