Cosmostatistics: the initial conditions and the large-scale structure of the Universe

Florent Leclercq
Institut d'Astrophysique de Paris
École polytechnique ParisTech

July 31th, 2013

with: Jens Jasche (IAP), Héctor Gil-Marín (U. Portsmouth/U. Barcelona), Benjamin Wandelt (IAP/U. Illinois)

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 infered.

The ***initial conditions*** of the Universe are particular with respect to other physical phenomena.

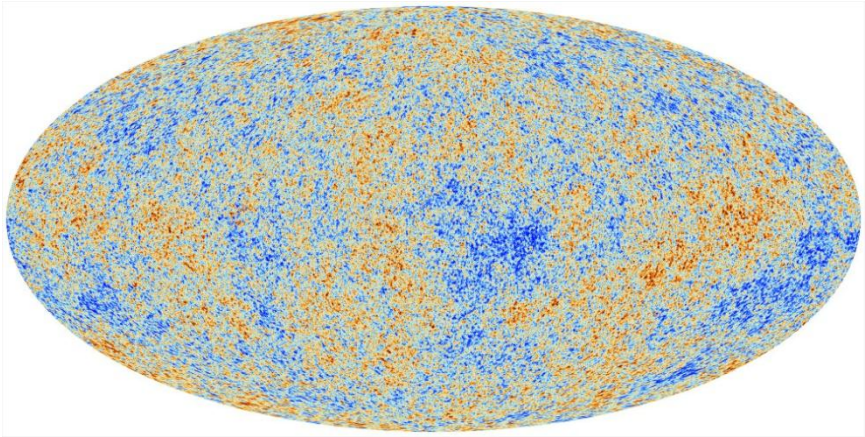
Cosmostatistics of the initial conditions

- ***"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 of **using the departures from homogeneity** observed in astronomical surveys to **distinguish between cosmological models**.

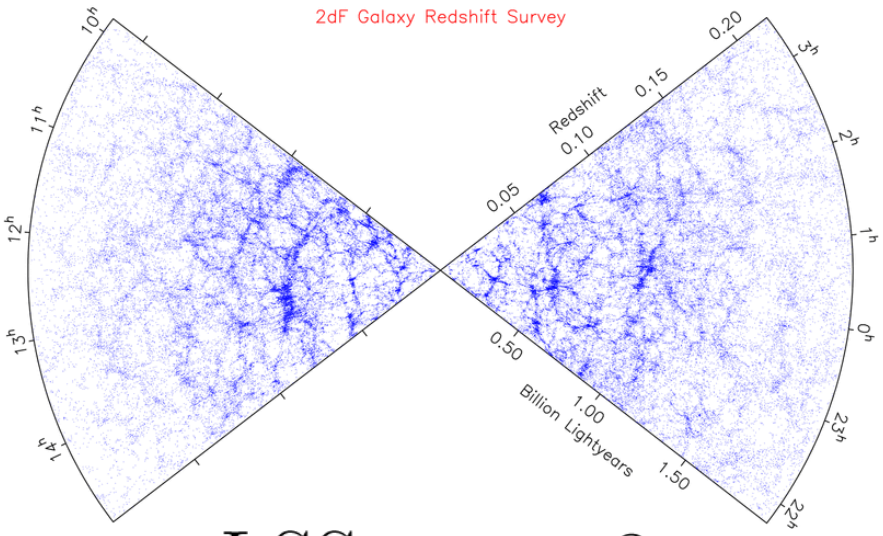
- Huge data sets, but fundamental limits to information:
  - on large scales: **causality**
  - on small scales: **non-linearity**

CMB



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

LSS



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

A time-machine (380,000 yrs ➡ 10^{-35} s):

linear perturbation theory

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

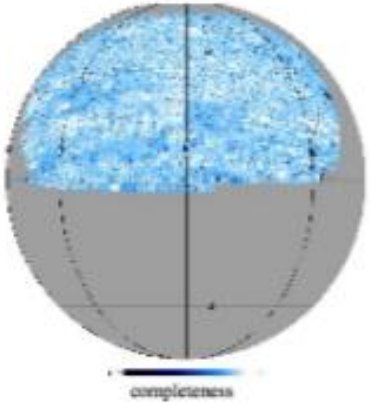
Yadav & Wandelt 2005, arXiv:astro-ph/0505386

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, biases, cosmic variance
  - Systematic uncertainties
  - Cosmic variance



No unique recovery is possible!

- A good question: "What is the probability distribution of possible signals compatible with the observations? "

from J. Jasche

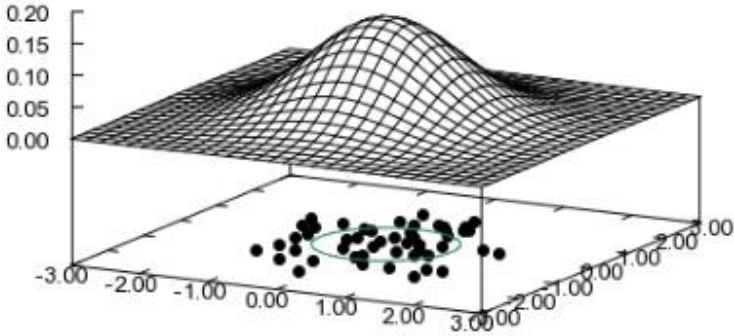
Bayesian inference of the ICs

- Problems:
  - High dimensional (10^7 parameters)
  - A large number of correlated parameters

  ➡ No reduction of the problem size is possible!
  - Complex posterior distribution
- Numerical approximation: for dim > 4: sampling the posterior distribution

$$\mathcal{P}(s|d) \rightarrow \mathcal{P}_N(s|d) = \frac{1}{N}\sum_{i=1}^{N}\delta^D(s - s_i)$$

- But how to "get the dots" ?

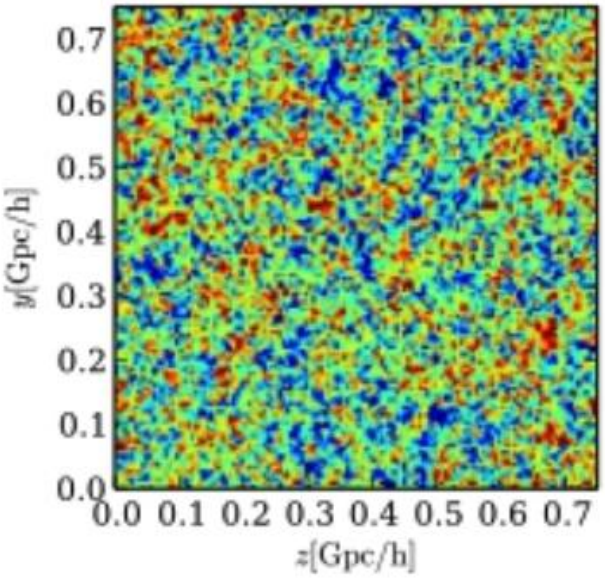


from J. Jasche

4D physical inference of the ICs

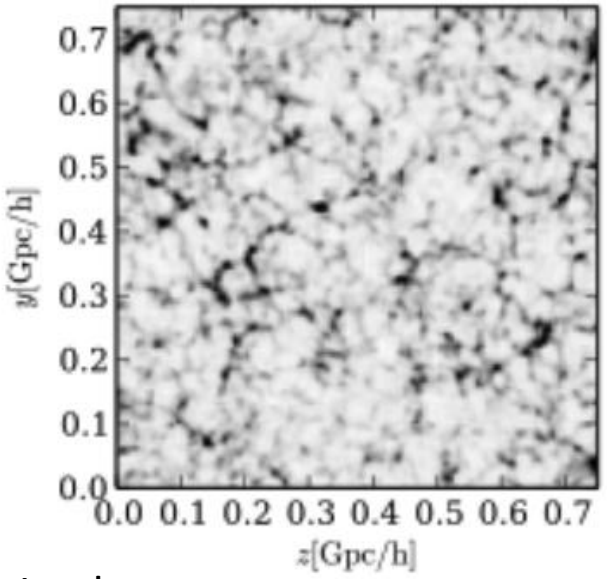
- Physical motivation:
  - Complex final state
  - Simple initial state
- A "direct only" problem

Initial state



GRAVITY

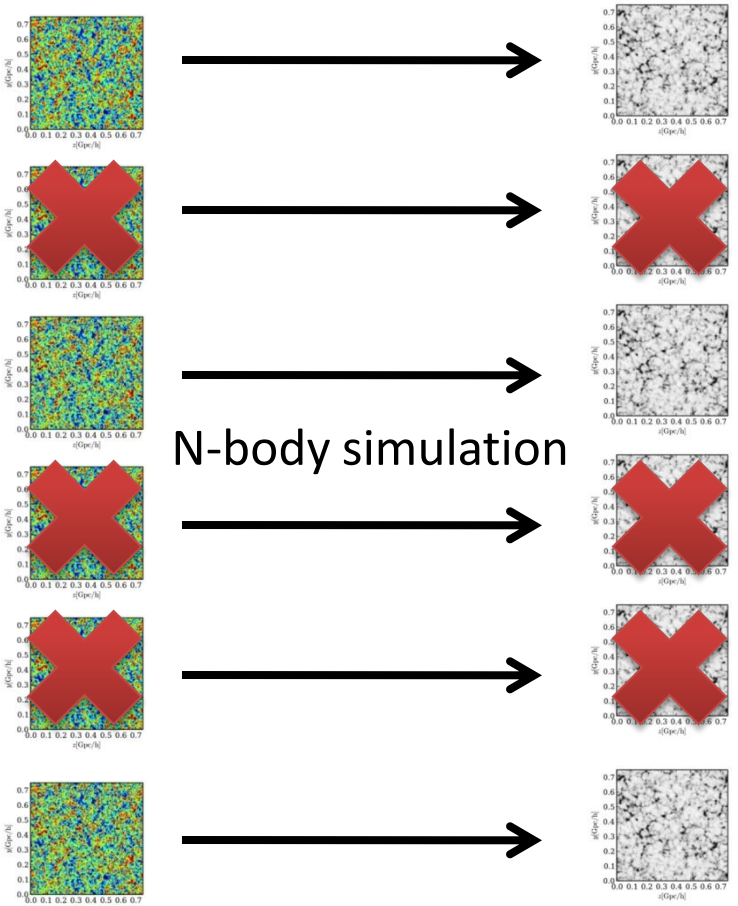
Final state



from J. Jasche

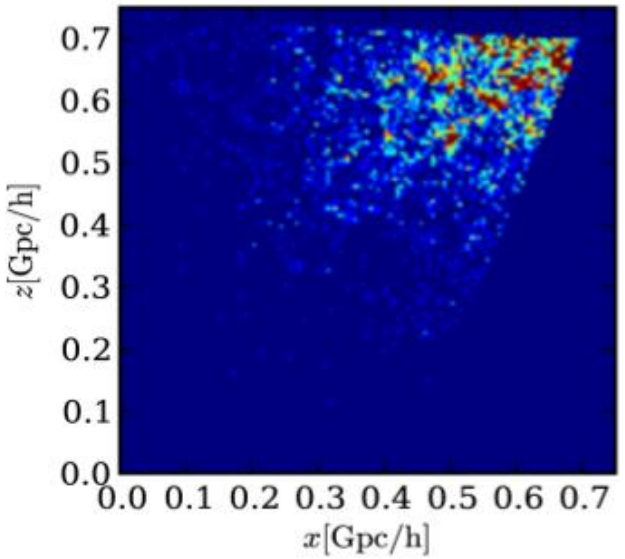
4D physical inference of the ICs

- The ideal scenario:



N-body simulation

observations



We need a very, very, very big computer!

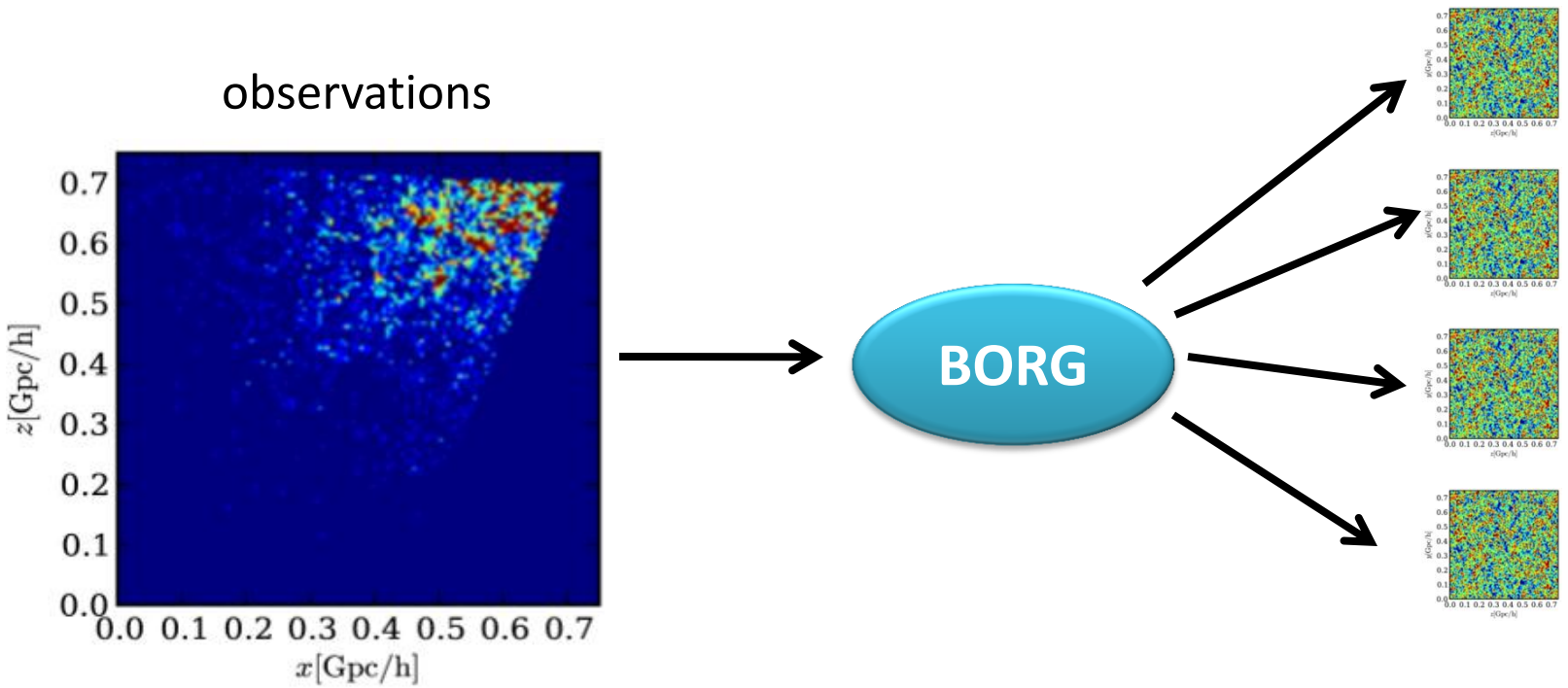
every possible ICs

every possible FCs

from J. Jasche

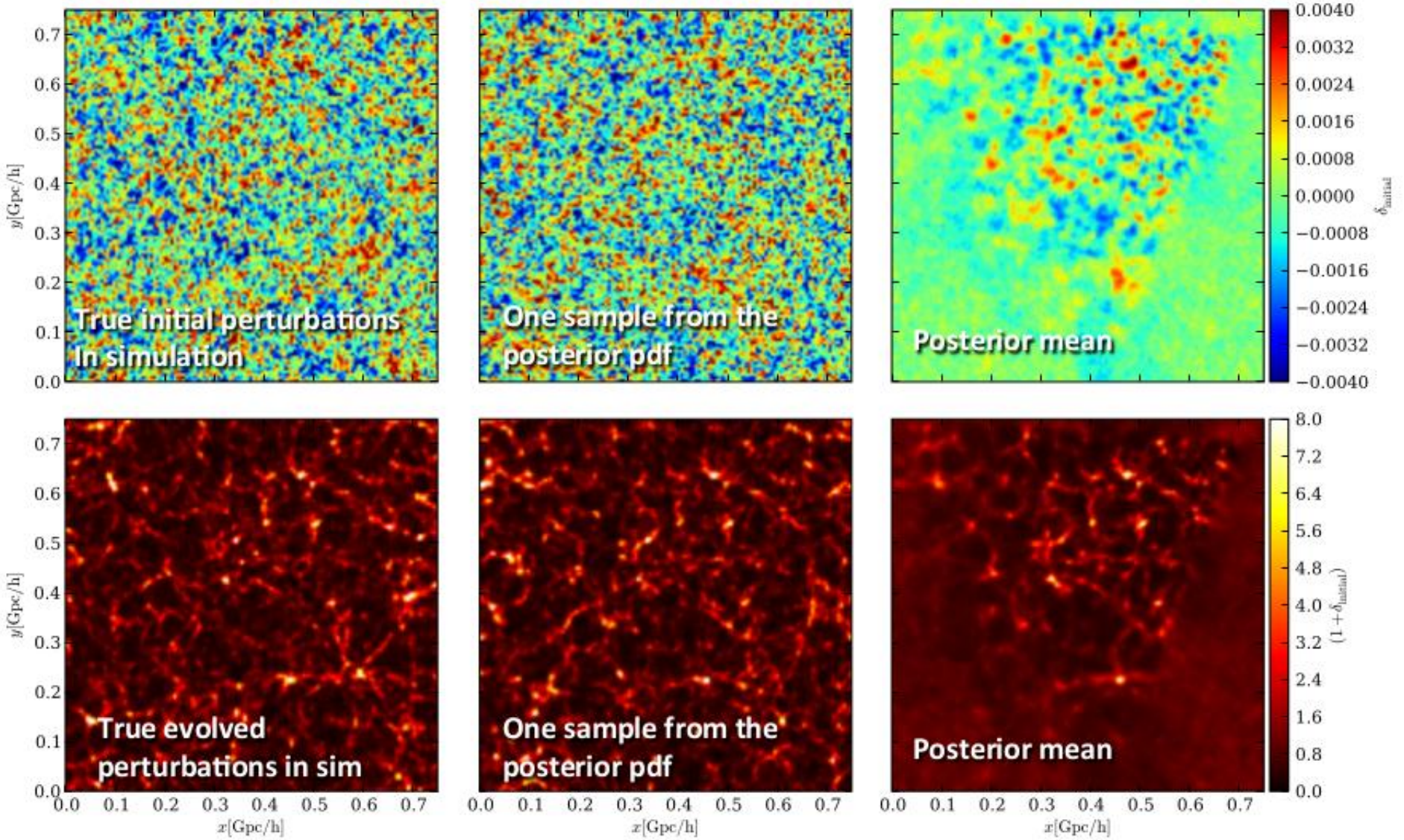
BORG: Bayesian Origin Reconstruction from Galaxies

- MCMC with Hamiltonian sampling
- Second-order Lagrangian perturbation theory



Jasche & Wandelt 2012, arXiv:1203.3639 from J. Jasche

Bayesian non-linear inference



Jasche & Wandelt 2012, arXiv:1203.3639

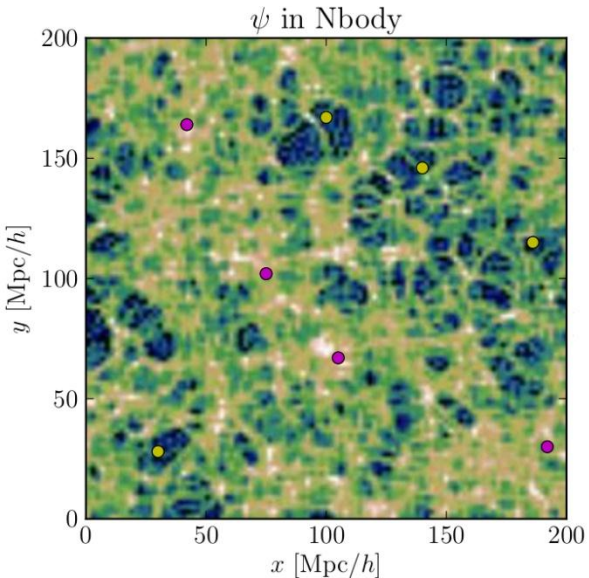
Samples of the posterior density

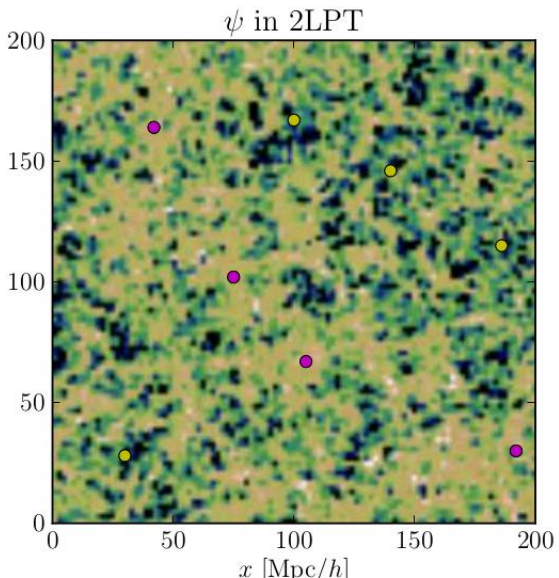
- Each sample: a **possible version of the truth**
- The **variation between samples** quantifies the uncertainty that results from having, e.g.
  - only one Universe (a more precise version of "cosmic variance")
  - imperfect data (mask, finite volume, finite number of galaxies, photometric redshifts…)

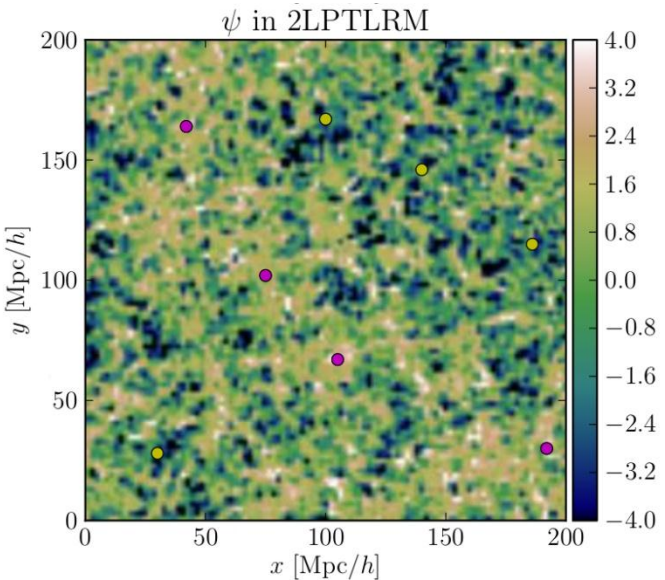
BORG at work

Beyond 2LPT?

- Recall the number of usable modes goes like k^3
- We need numerically efficient and flexible tools to model cosmic structure formation in the NL regime
- A proposal: remapping of 2LPT in the mildly non-linear regime FL, Jasche, Gil-Marín & Wandelt 2013, arXiv:1305.4642

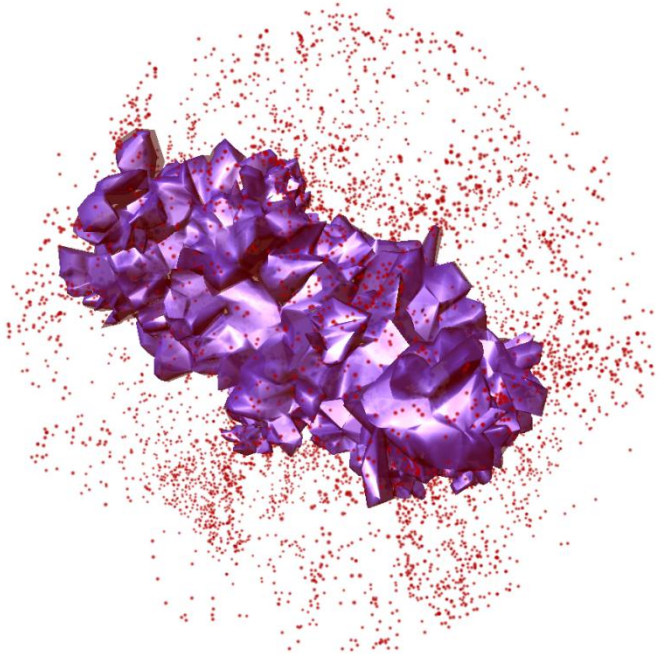






Aside: cosmology with voids

- A public void catalog from the Sloan Digital Sky Survey DR7:



Sutter, Lavaux,
Wandelt & Weinberg 2012, arXiv: 1207.2524
http://www.cosmicvoids.net

- **Number count**: void **size** determination
- **Dynamics**: **linear or weakly non-linear** regime
- First steps toward a systematic study of void statistics:
  - One-point function Sutter , Lavaux, Alizadeh, Biswas, FL & Wandelt, in prep.
  - Two-point function Hamaus *et al* 2013, arXiv:1307.2571; FL & Wandelt, in prep.

Concluding thoughts

- BORG: A **non-linear time machine** using Bayesian posterior exploration to infer primordial quantities from late-time observations
- Need for **efficient tools to model cosmic structure formation** in the non-linear regime

- Cosmological physical reconstruction of the initial conditions of the Universe becomes feasible.
  - BAO, clusters, voids
  - Non-Gaussianity
  - Isocurvature perturbations
  - Gravitational waves in LSS…

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