How did structure appear in the Universe? A Bayesian approach

Florent Leclercq

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

August 27th, 2014

In collaboration with:

Jacopo Chevallard (U. São Paulo), Héctor Gil-Marín (U. Portsmouth), Nico Hamaus (IAP), Jens Jasche (IAP), Alice Pisani (IAP), Emilio Romano-Díaz (U. Bonn), Paul M. Sutter (IAP/Ohio State U.), Svetlin Tassev (U. Princeton), Benjamin Wandelt (IAP/U. Illinois), Matías Zaldarriaga (IAS Princeton)

How did structure appear in the Universe?

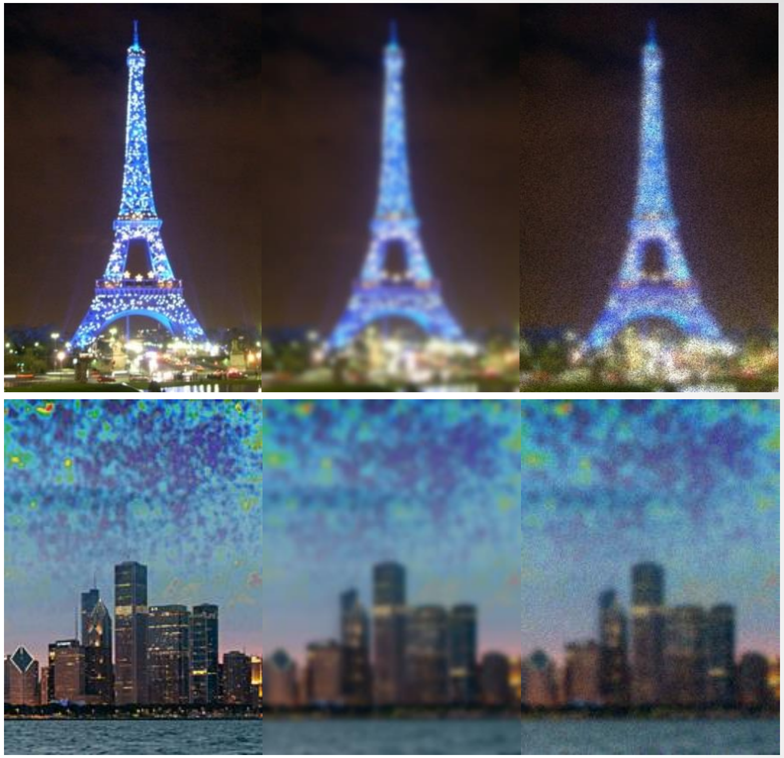
A joint problem!

- **How did the Universe begin?**
  - What are the statistical properties of the initial conditions?
- **How did the large-scale structure take shape?**
  - What is the physics of dark matter and dark energy?
- Usually these problems are addressed in isolation.
- This talk:
  - A case for physical inference of four-dimensional dynamic states
  - A description of methodology and progress towards enriching the standard for analysis of galaxy surveys
  - From theory to data, from data to theory

FL, Pisani & Wandelt 2014, arXiv:1403.1260

(Lectures Varenna 2013 and
Paris École Doctorale for Astronomy and Astrophysics)

Why Bayesian inference?



- Why do we need Bayesian inference?

Inference of signals = ill-posed problem

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

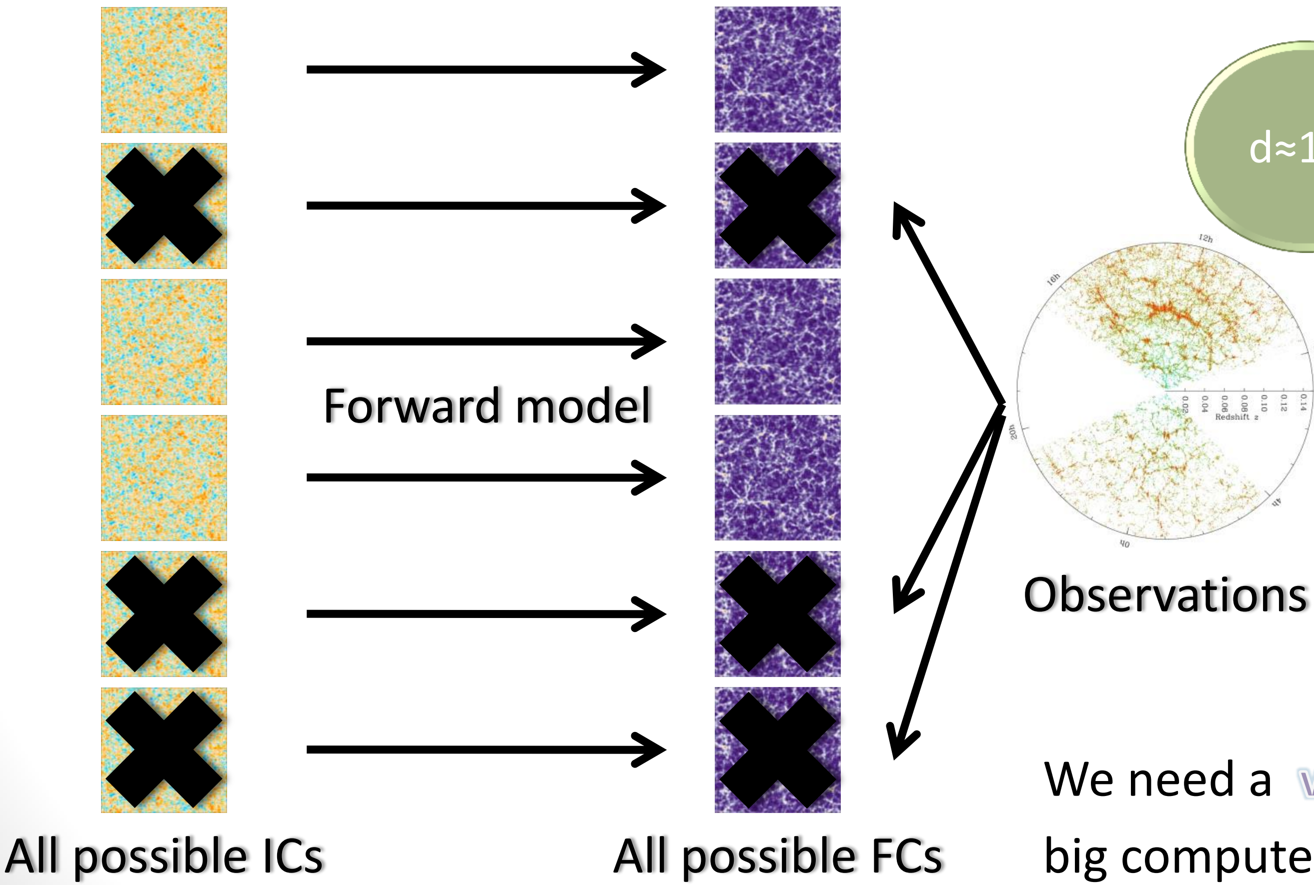
➡ **No unique recovery is possible!**

"What is the formation history of the Universe?" ➡ "What is the probability distribution of possible formation histories (signals) compatible with the observations?"

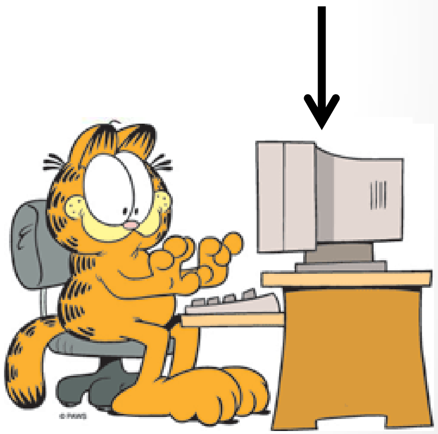
$$p(s|d)p(d) = p(d|s)p(s)$$

Bayesian forward modeling: the ideal scenario

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



$d \approx 10^7$

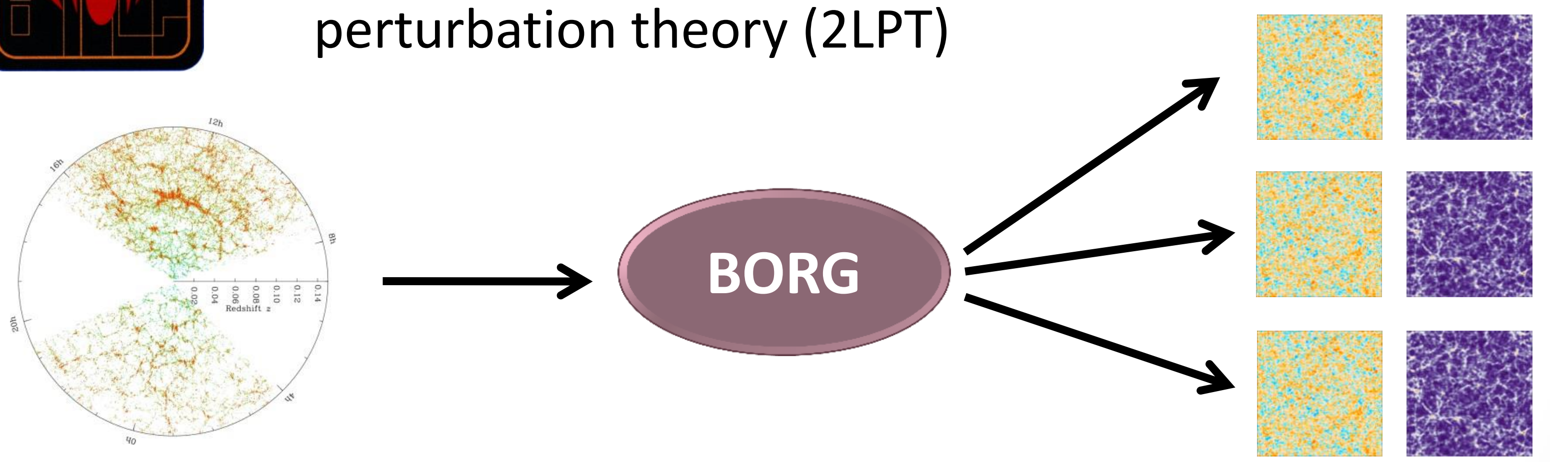


We need a very, very, very big computer!

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)



Observations

Samples of possible 4D states

Jasche & Wandelt 2013, arXiv:1203.3639

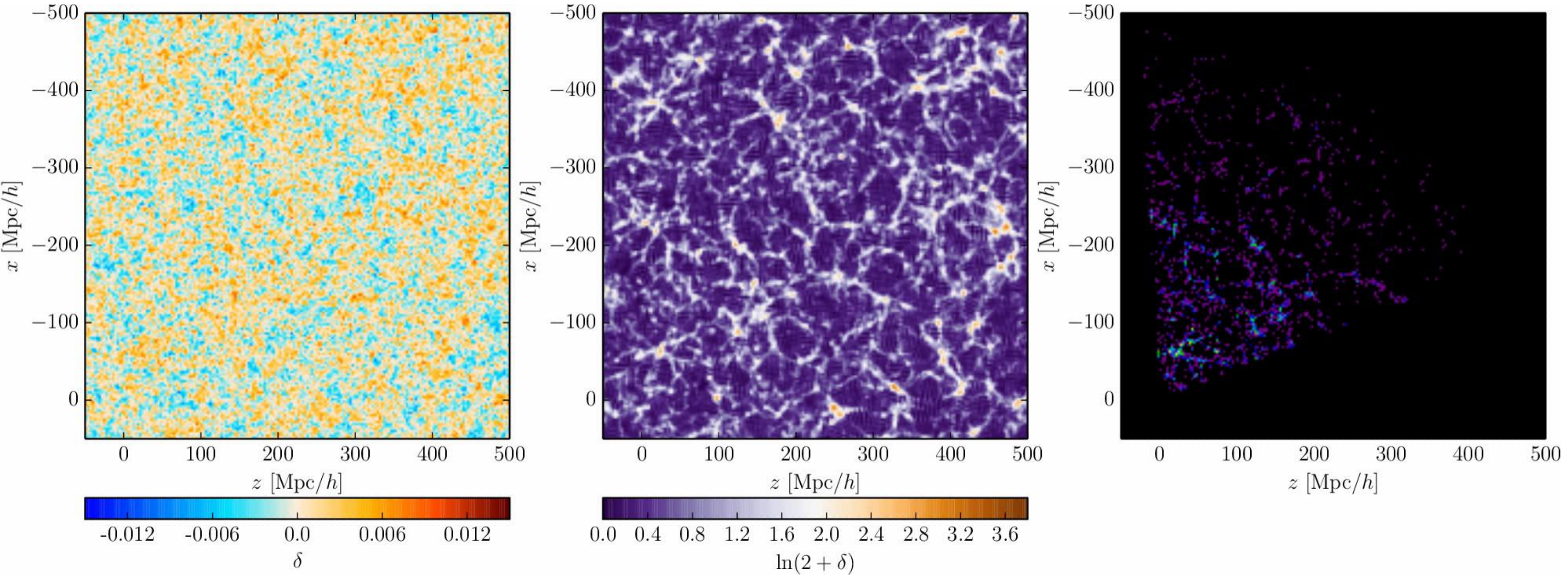
see also:
Kitaura 2013, arXiv:1203.4184
Wang, Mo, Yang & van den Bosch 2013, arXiv:1301.1348

The BORG SDSS run

- 463,230 galaxies from the NYU-VAGC based on SDSS DR7
- Comoving cubic box of side length 750 Mpc/h, with periodic boundary conditions
- 256^3 grid, resolution 3 Mpc/h ⇒ ≈ 17 millions parameters
- 12,000 samples, four-dimensional maps
- ≈ 3 TB disk space
- 10 months wallclock time on 16-32 cores

Jasche, FL & Wandelt, in prep.

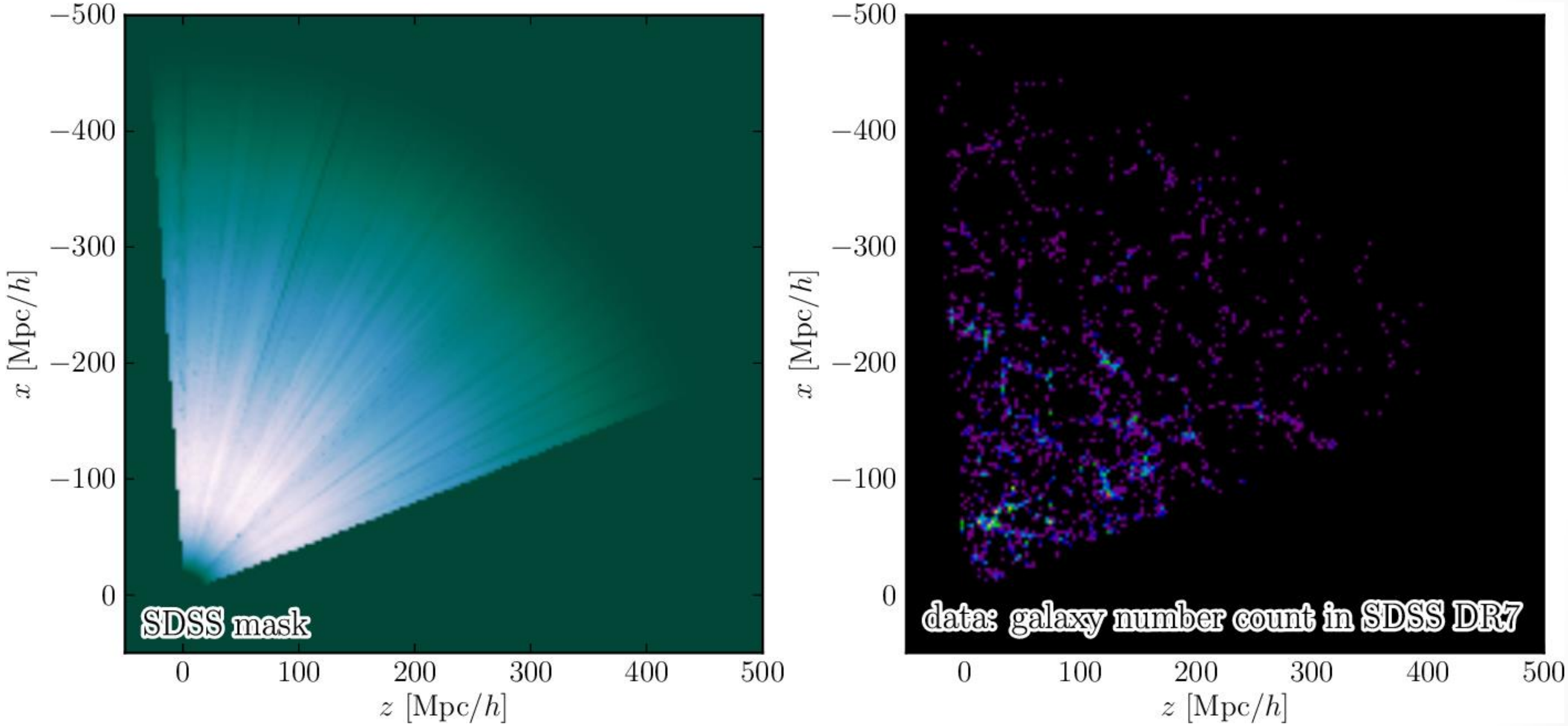
BORG at work – chronocosmography



Initial conditions | Final conditions | Observations

Jasche, FL & Wandelt, in prep.

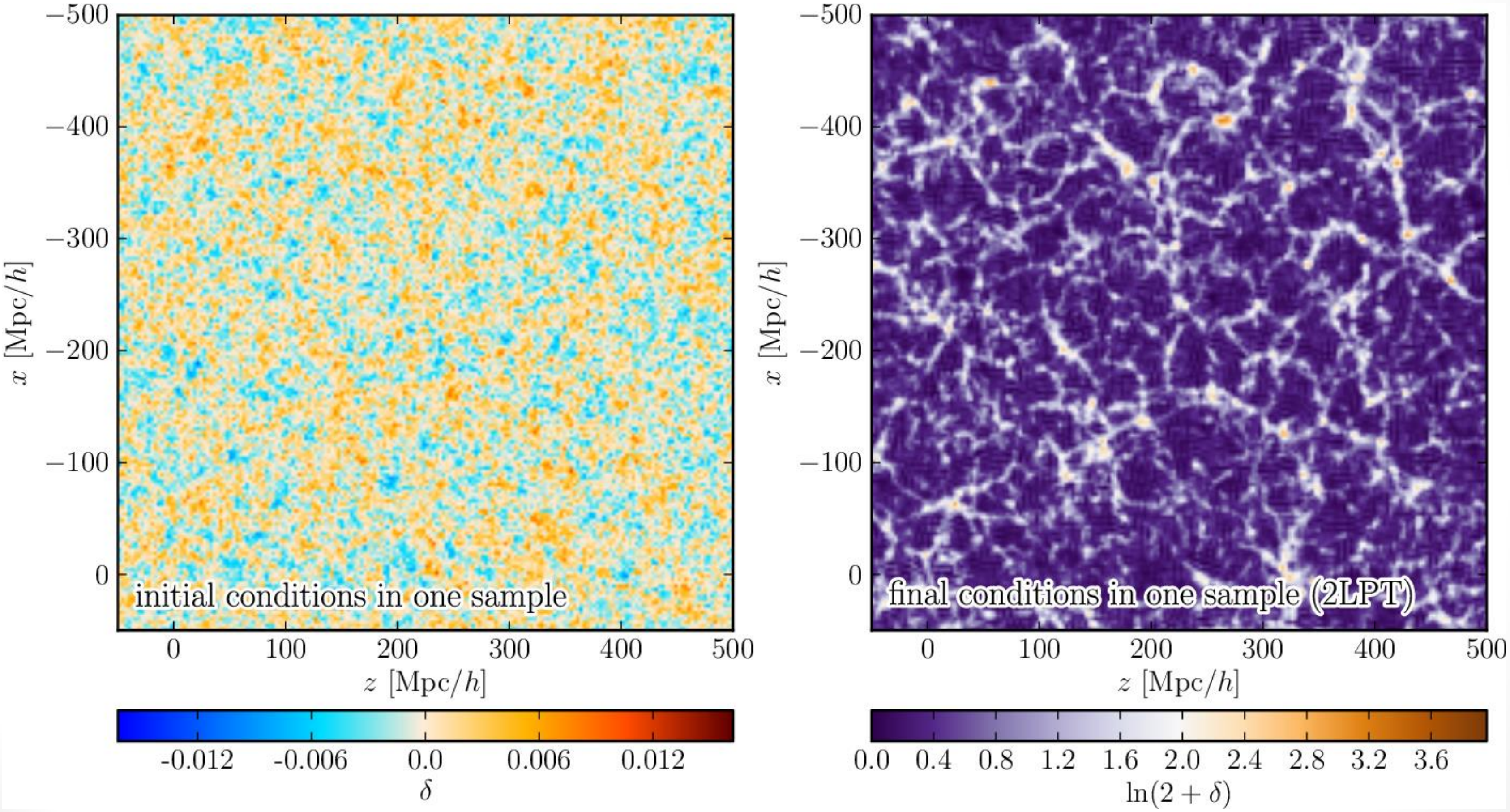
Bayesian chronocosmography from SDSS DR7



Jasche, FL & Wandelt, in prep.

Data

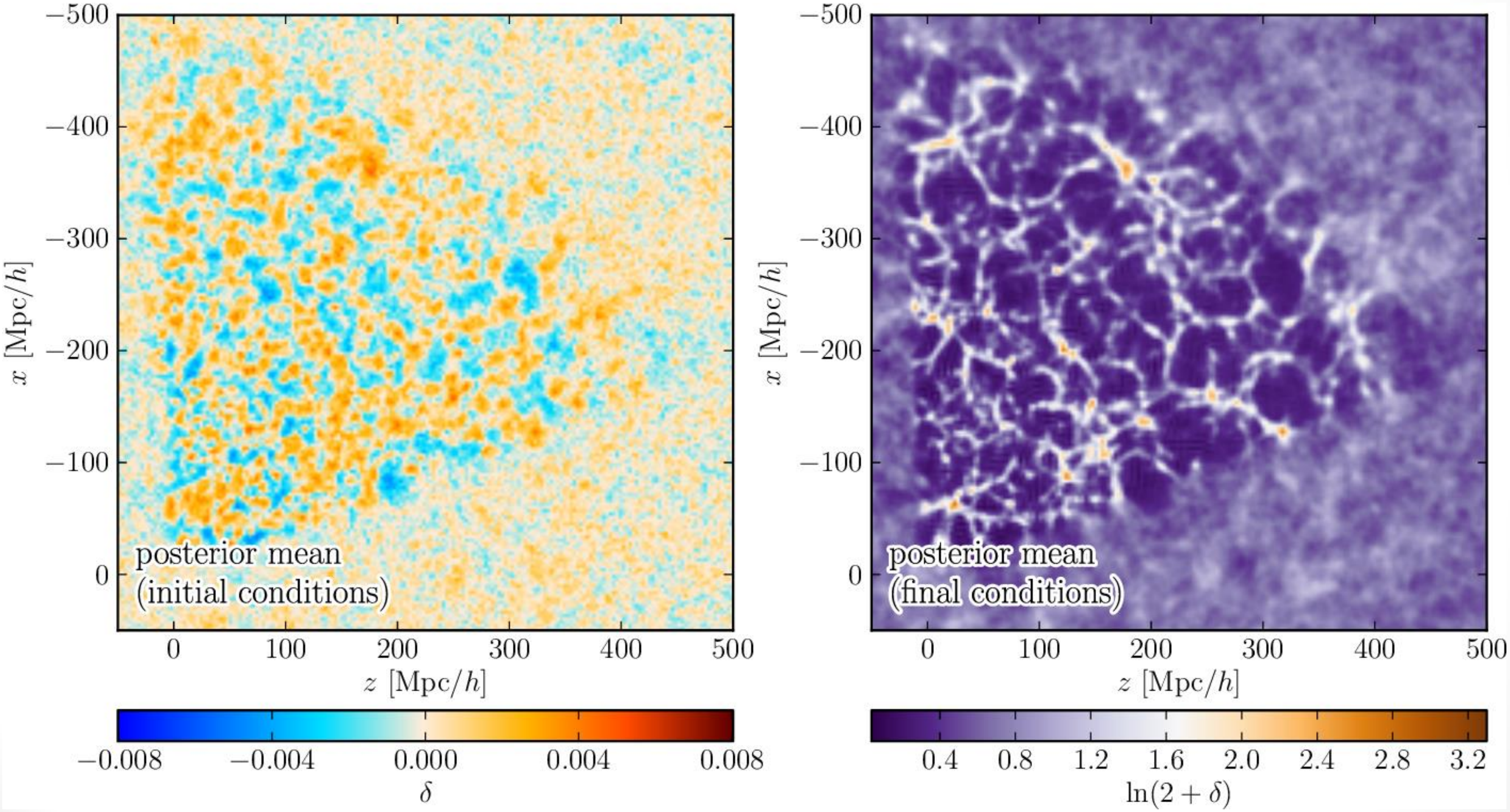
Bayesian chronocosmography from SDSS DR7



One sample

Jasche, FL & Wandelt, in prep.

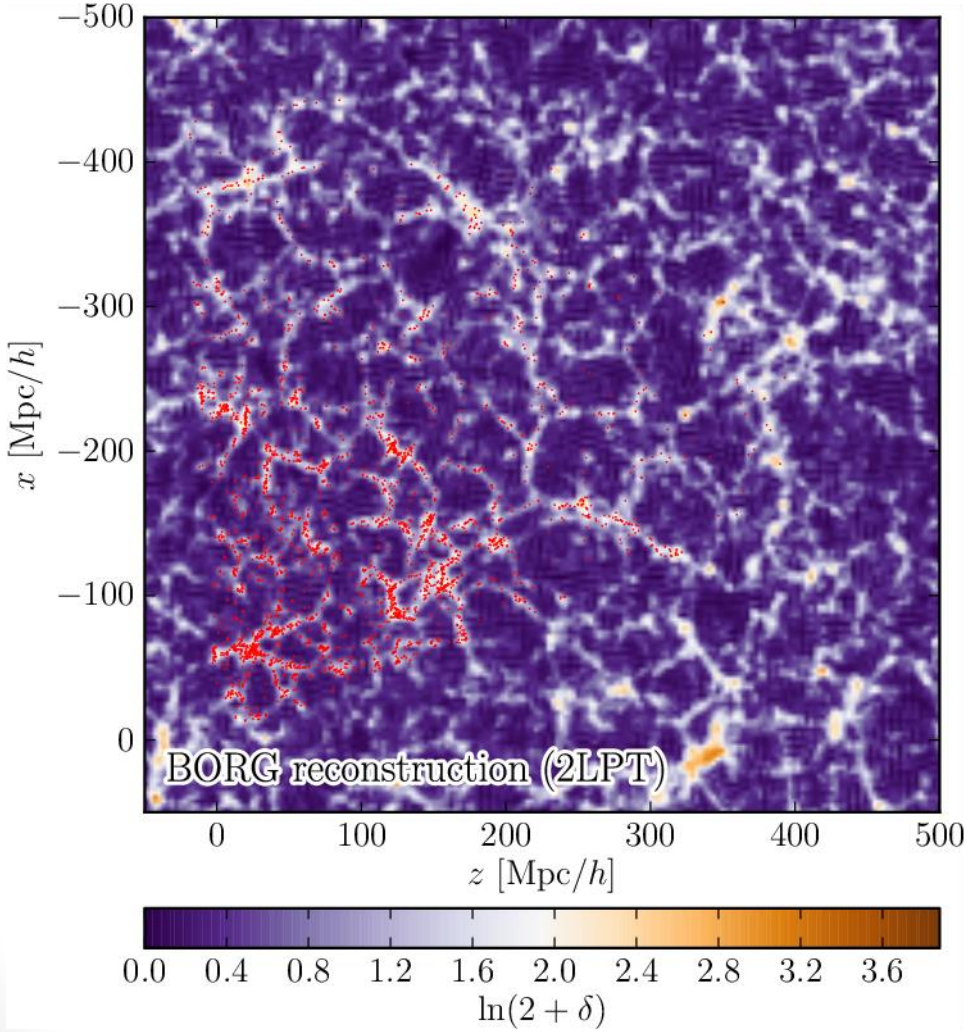
Bayesian chronocosmography from SDSS DR7



Posterior mean

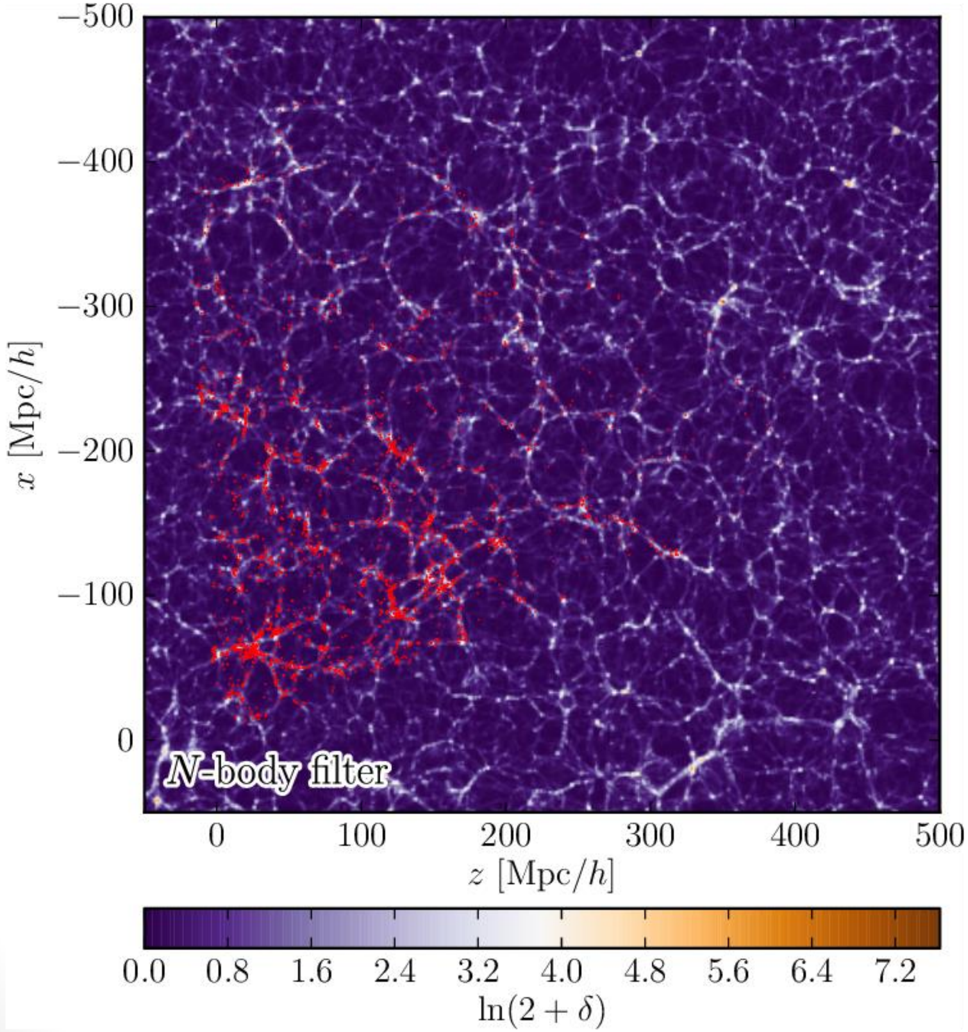
Jasche, FL & Wandelt, in prep.

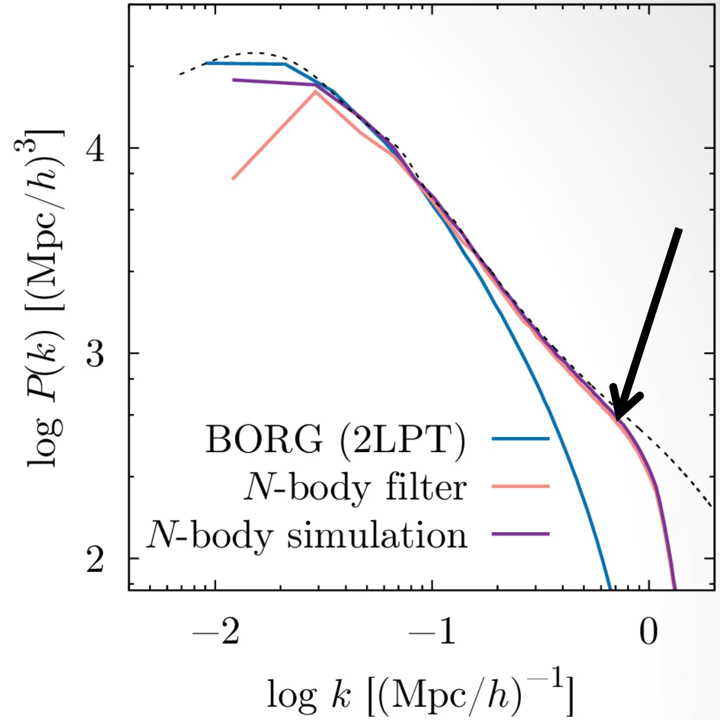
Non-linear filtering



FL, Jasche, Sutter, Hamaus & Wandelt, in prep. + Jasche, FL, Romano-Diaz & Wandelt, in prep.

Non-linear filtering





More on non-linear/non-Gaussian data models:

- Remapping LPT

FL, Jasche, Gil-Marín & Wandelt 2013, arXiv:1305.4642

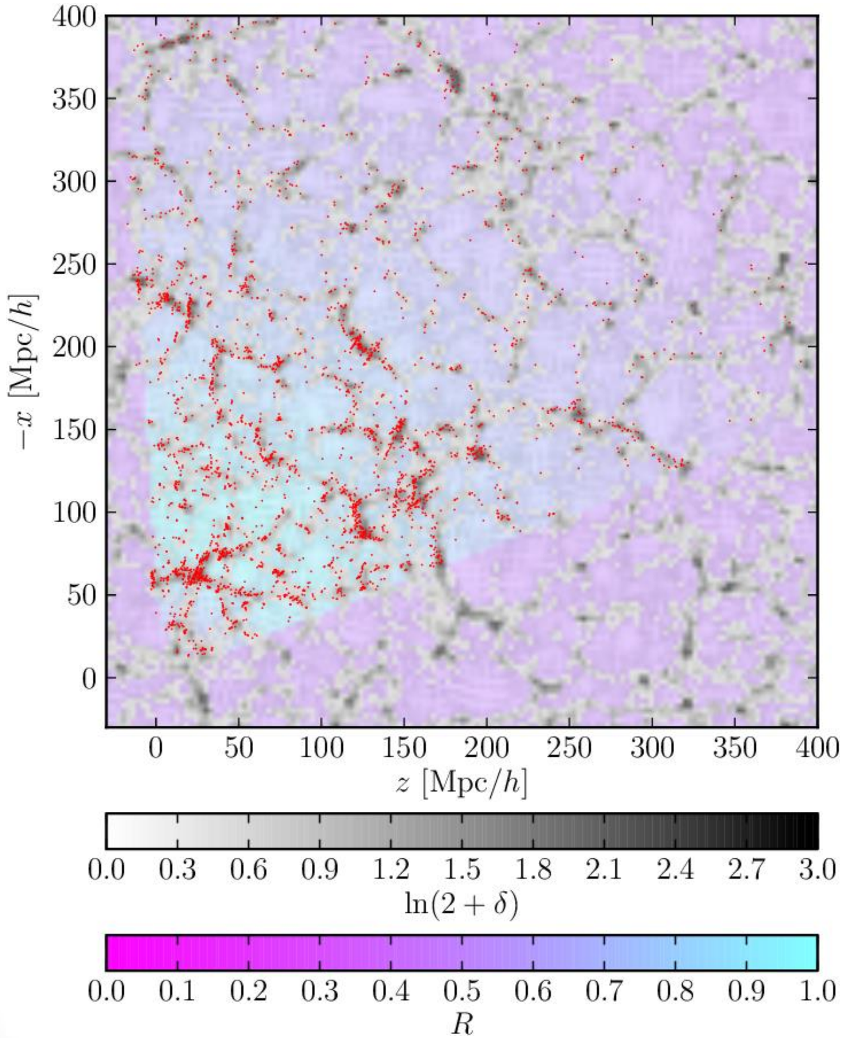
- COLA

Tassev, Zaldarriaga, Eisenstein 2013, arXiv:1301.0322

See also Metin Ata's poster

FL, Jasche, Sutter, Hamaus & Wandelt, in prep. + Jasche, FL, Romano-Diaz & Wandelt, in prep.

Dark matter voids in the SDSS



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

- How?

VIDE toolkit: Sutter *et al*. 2014, arXiv:1406.1191
www.cosmicvoids.net

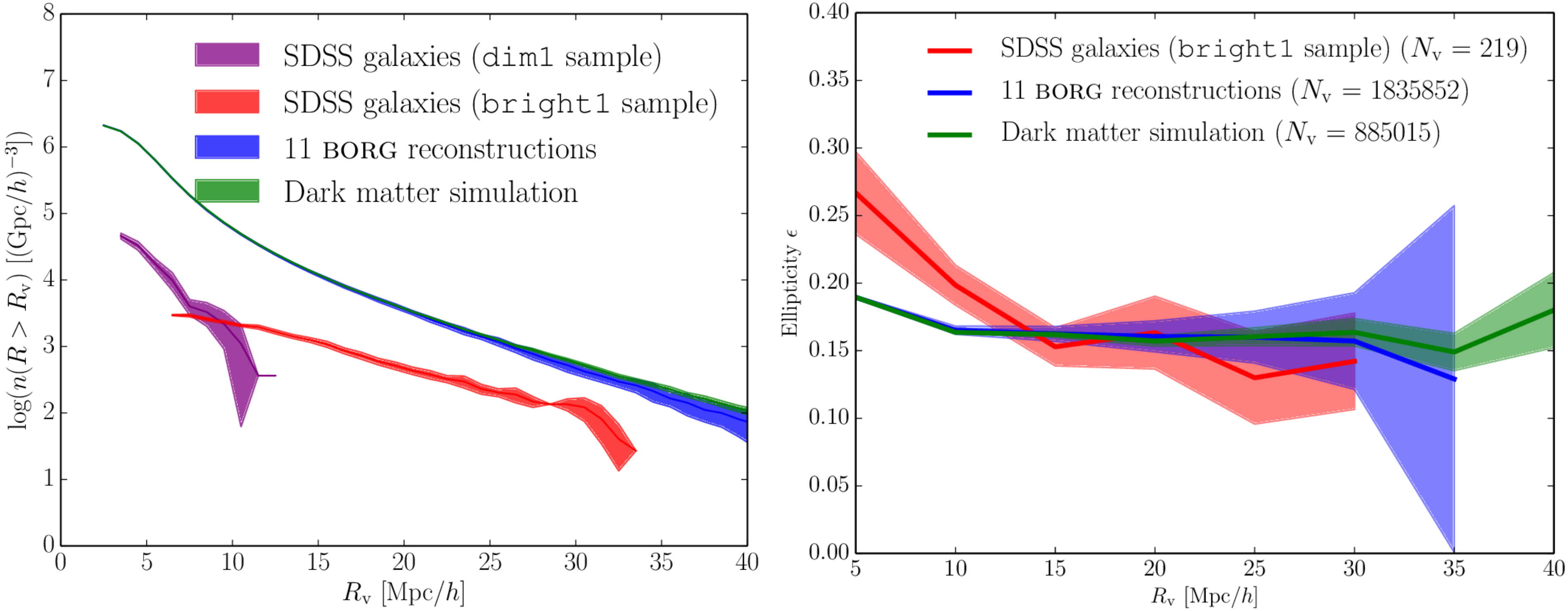
based on ZOBOV: Neyrinck 2007, arXiv:0712.3049

- Why?

Sparsity & **Bias**

Sutter et al. 2013, arXiv:1309.5087
Sutter et al. 2013, arXiv:1311.3301

Dark matter void properties



See Nico Hamaus's talk

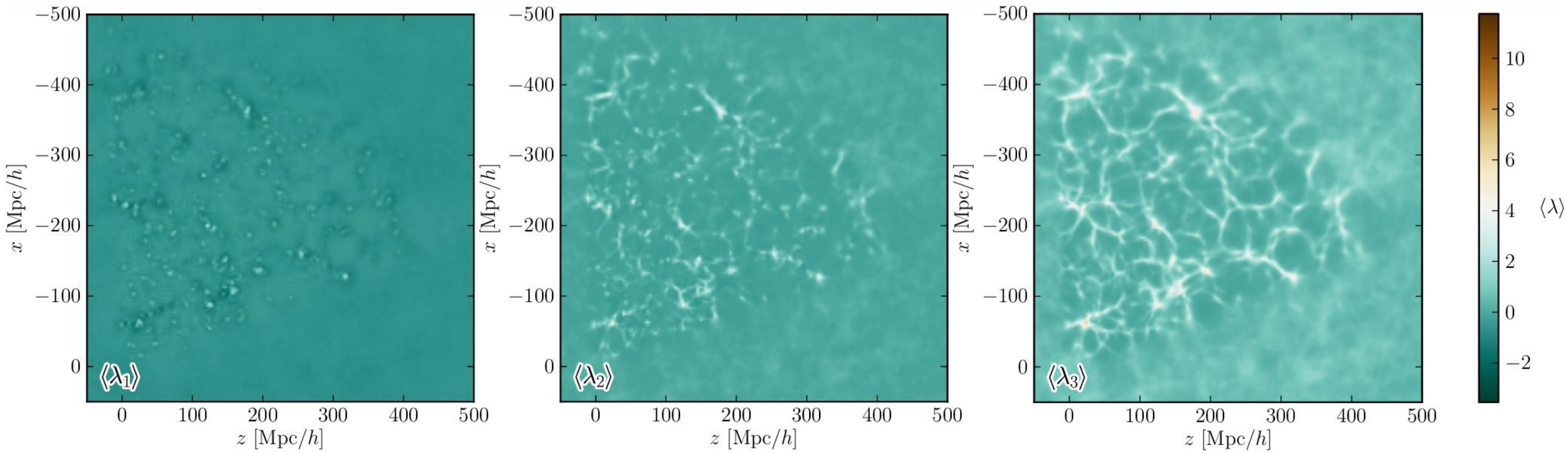
All catalogs will be made publicly available at
www.cosmicvoids.net

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

Tidal shear analysis

- $\lambda_1, \lambda_2, \lambda_3$: eigenvalues of the tidal field tensor, the Hessian of the gravitational potential: $T_{ij} = \partial_i \partial_j \Phi$
  - Voids: $\lambda_1, \lambda_2, \lambda_3 < 0$
  - Sheets: $\lambda_1 > 0$ and $\lambda_2, \lambda_3 < 0$
  - Filaments: $\lambda_1, \lambda_2 > 0$ and $\lambda_3 < 0$
  - Clusters: $\lambda_1, \lambda_2, \lambda_3 > 0$

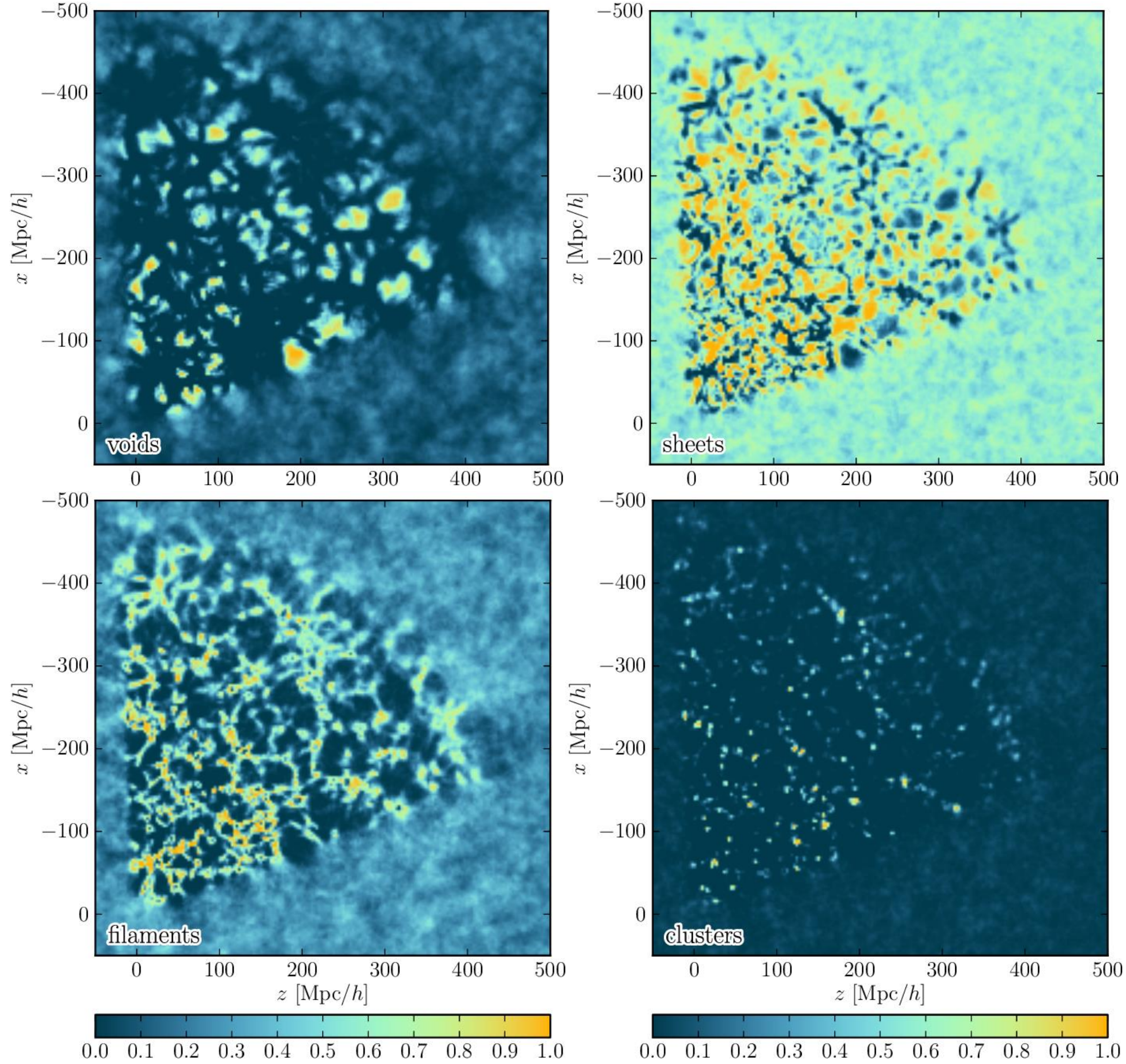
Hahn *et al*. 2006, arXiv:astro-ph/0610280
see also:
Forero-Romero *et al*. 2008, arXiv:0809.4135
Hoffman *et al*. 2012, arXiv:1201.3367



FL, Jasche, Chevallard & Wandelt, in prep.

Dynamic structures inferred by BORG

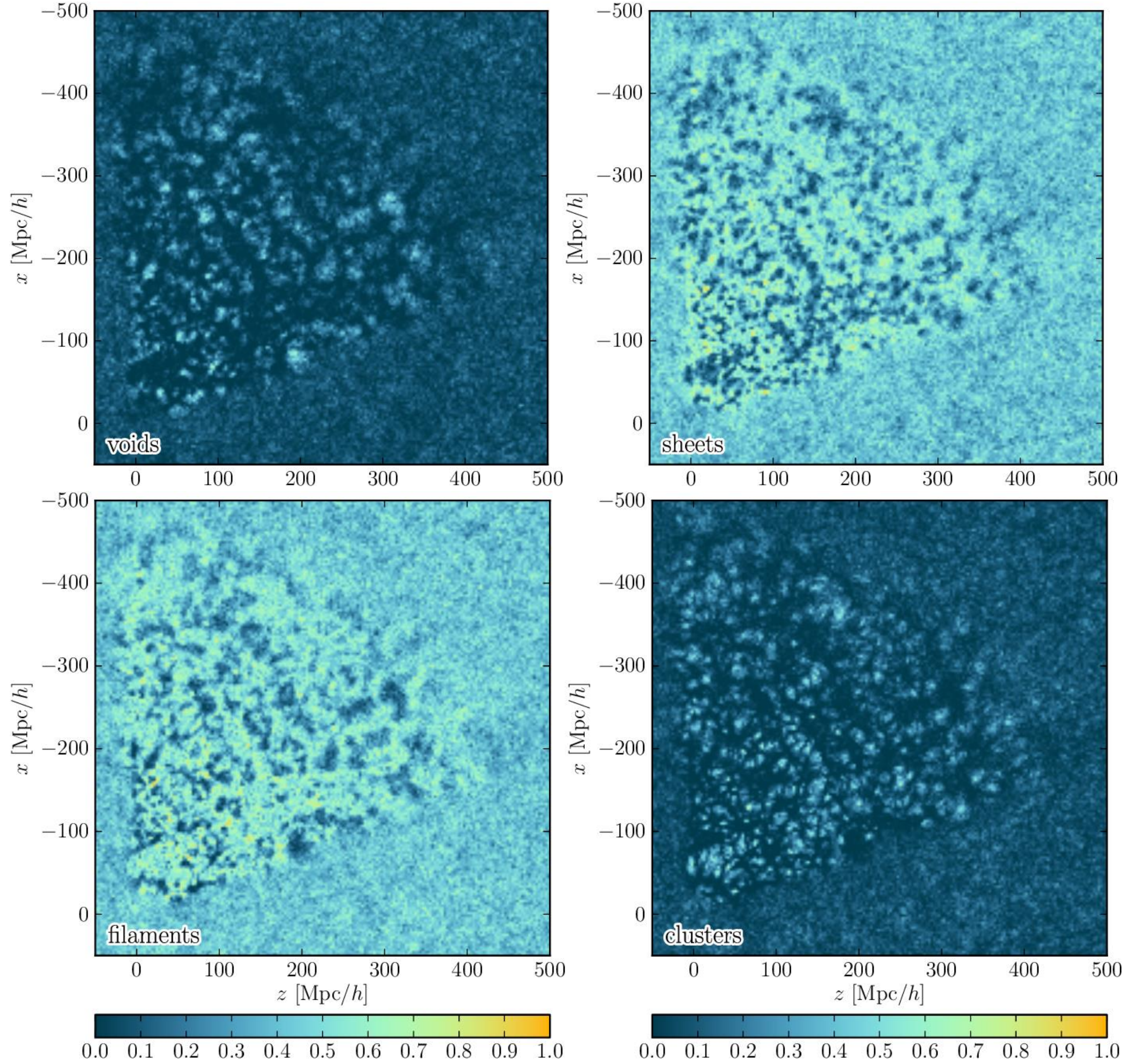
Final conditions



FL, Jasche, Chevallard & Wandelt, in prep.

Dynamic structures inferred by BORG

Initial conditions



FL, Jasche, Chevallard & Wandelt, in prep.

Summary & Conclusions

- Bayesian large-scale structure inference in 10 millions dimensions is possible!
  - Non-linear and non-Gaussian inference
  - Uncertainty quantification (noise, survey geometry, selection effects and biases)
- Application to data: four-dimensional chronocosmography
  - Simultaneous analysis of the morphology and formation history of the large-scale structure
  - Physical reconstruction of the initial conditions
  - Inference of cosmic voids at the level of the dark matter distribution
  - Characterization of the dynamic cosmic web underlying galaxies