The Abdus Salam
International Centre
for Theoretical Physics

2419-24

Workshop on Large Scale Structure

30 July - 2 August, 2012

Bayesian cosmological inference from Large Scale Structure data

B. Wandelt
IAP - U. Pierre et Marie Curie

Bayesian cosmological inference from Large Scale Structure data

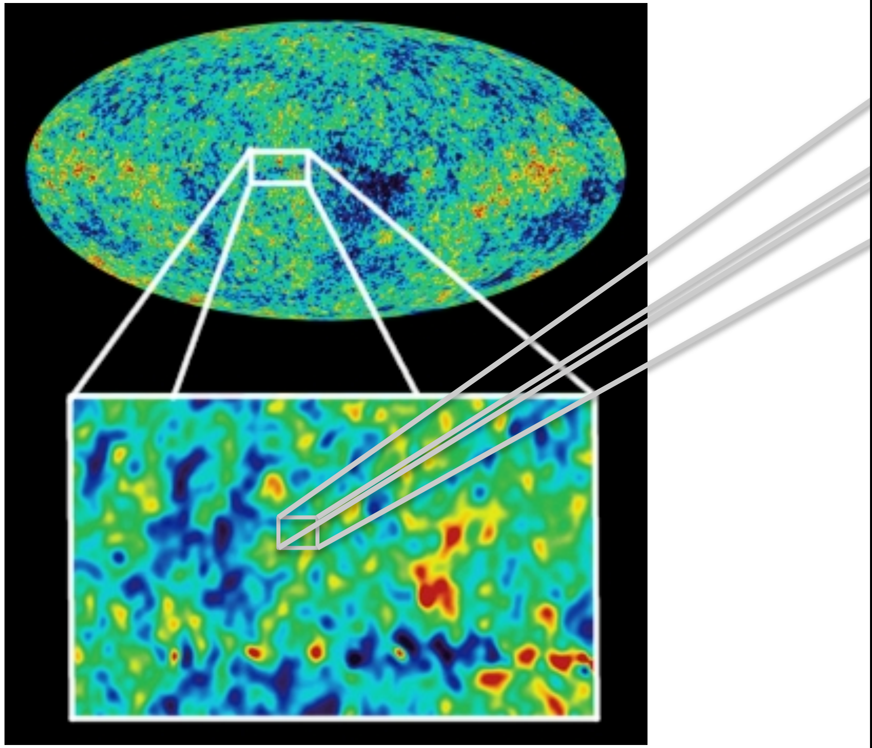
Ben Wandelt

Institut d'Astrophysique de Paris
Lagrange Institute, Paris
Sorbonne University

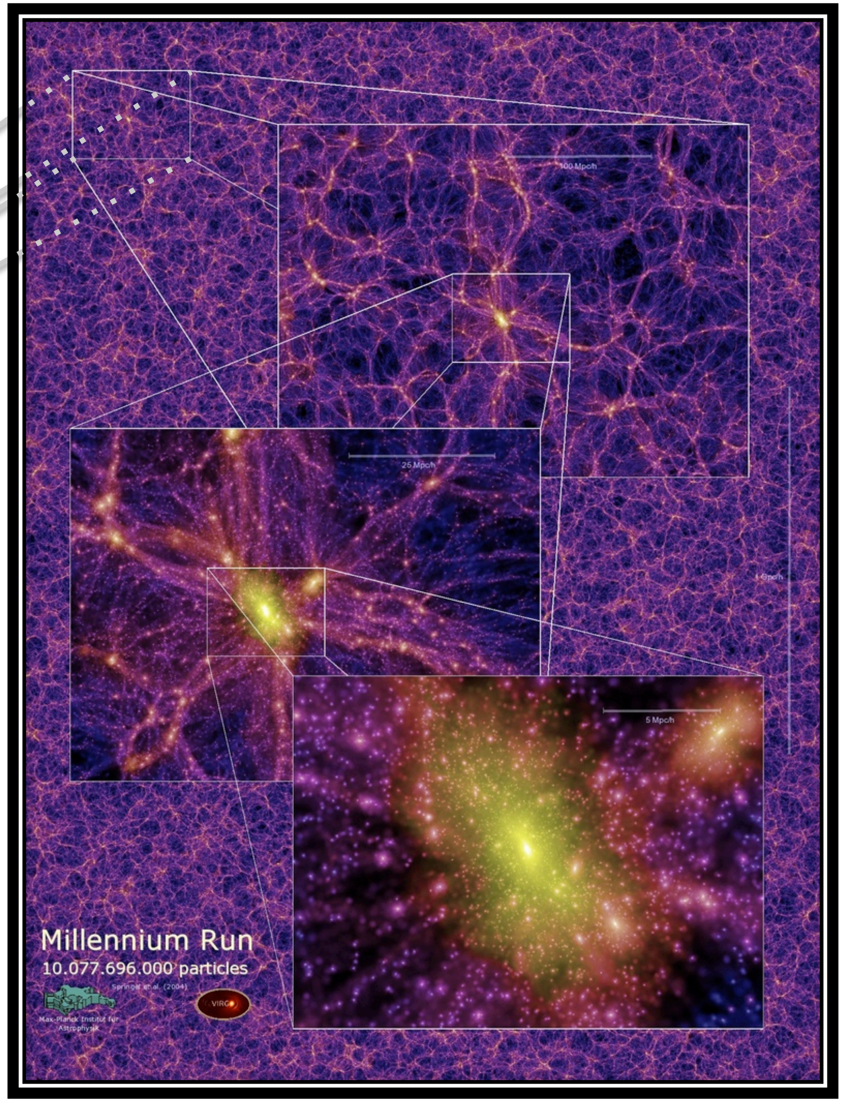
References

- Bayesian photo-z
  Jens Jasche, Wandelt arxiv:1106.2757
- Bayesian dynamical histories
  Jens Jasche, Wandelt arxiv:1203.3639
- Stacked voids
  Guilhem Lavaux, Wandelt arxiv:1110.0345
- Void catalog / cosmicvoids.net
  Paul Sutter, Lavaux, Wandelt, Weinberg arxiv:1207.2524

Inference from Large Scale Structure: the Big Picture



Primordial perturbations as seen in the Cosmic Microwave Background anisotropies (WMAP)



Dark matter distribution today (simulated)

Cosmological inference goals

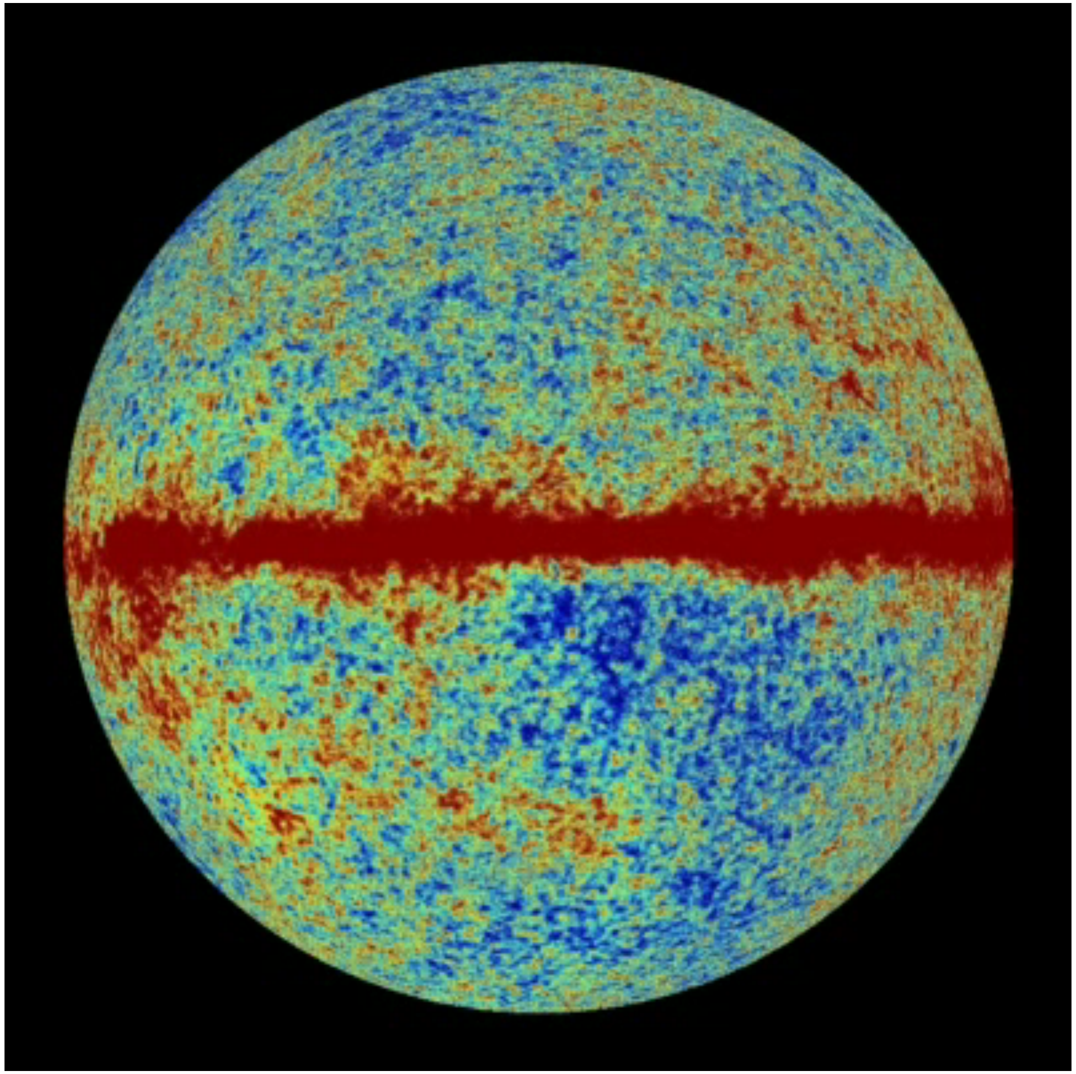
- How did the Universe begin?
- How did structure appear in the Universe?
- How did it evolve until today?
- What is the Universe made of?
- What are the properties of dark matter?
- What are the properties of dark energy?
- What is the geometry of the Universe?
- Is the Universe "symmetric?"

Learning from observations

Cosmic structure is *stochastic. Information is contained in modes.*

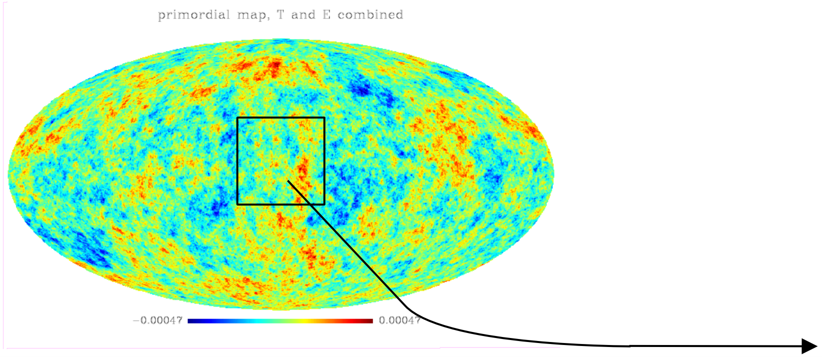
How can we access the largest number of modes?

Control systematics and noise?

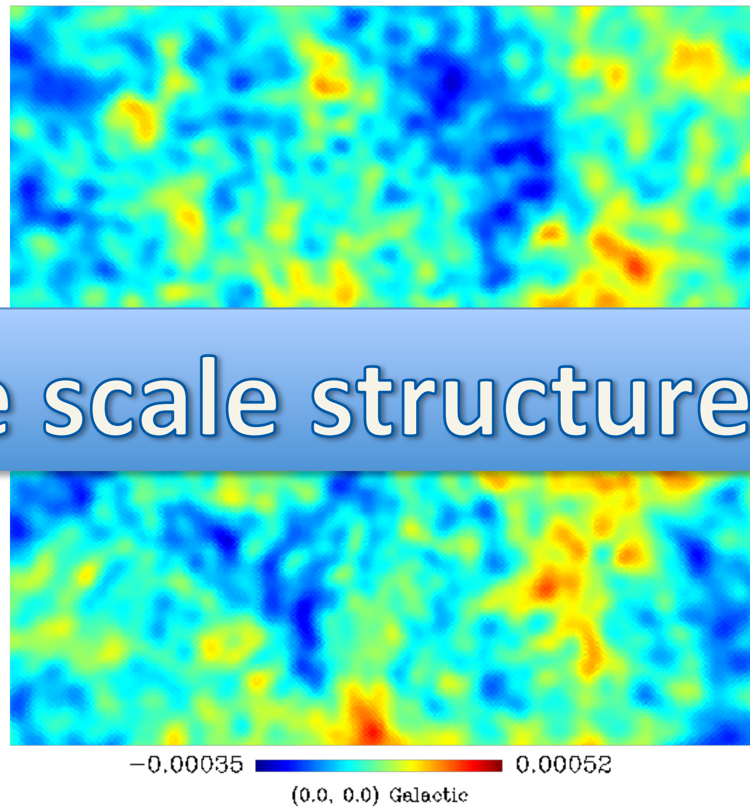


Microwave sky (WMAP)

Probing initial conditions with the CMB



Primordial curvature fluctuations

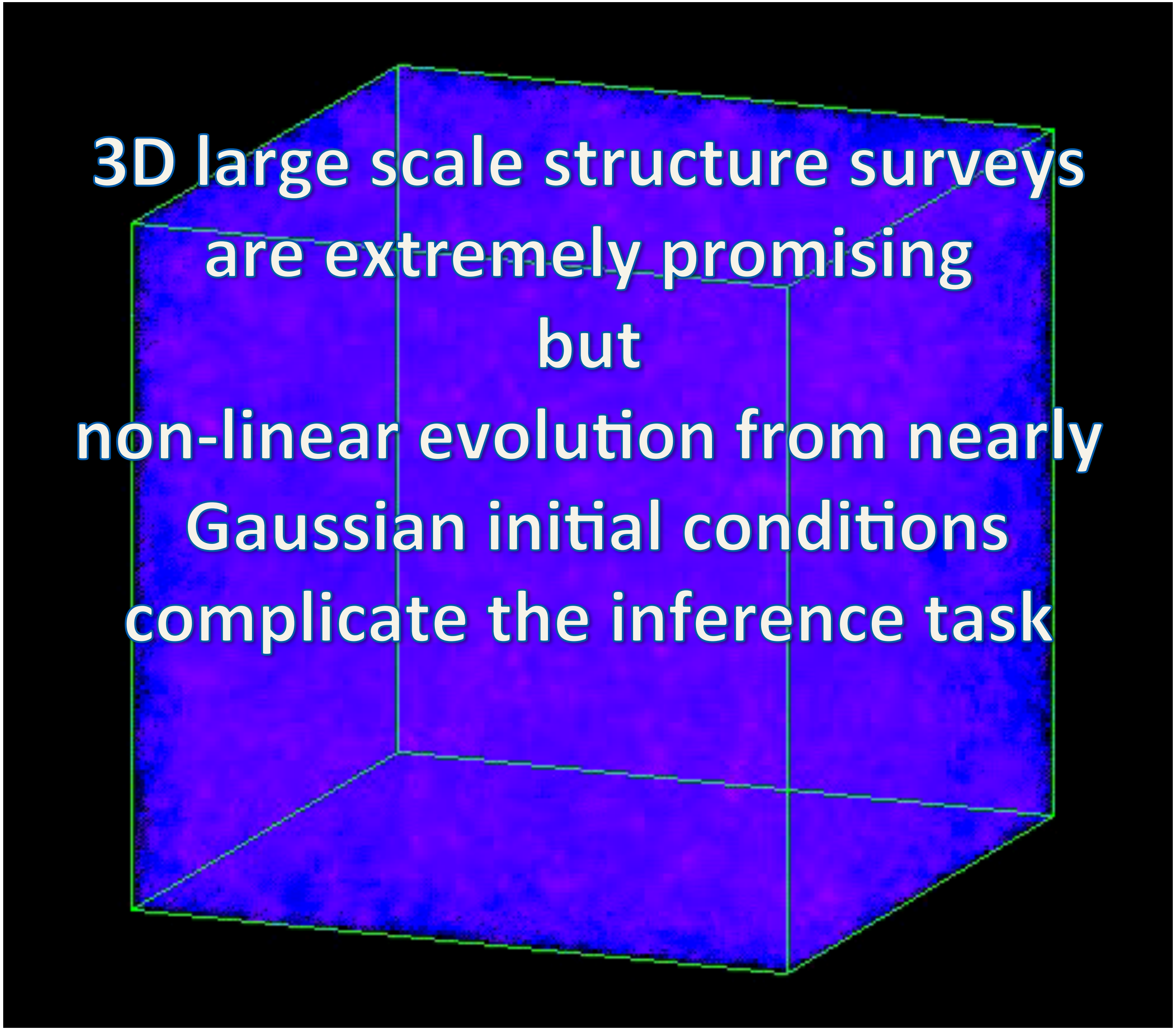


. What about large scale structure?

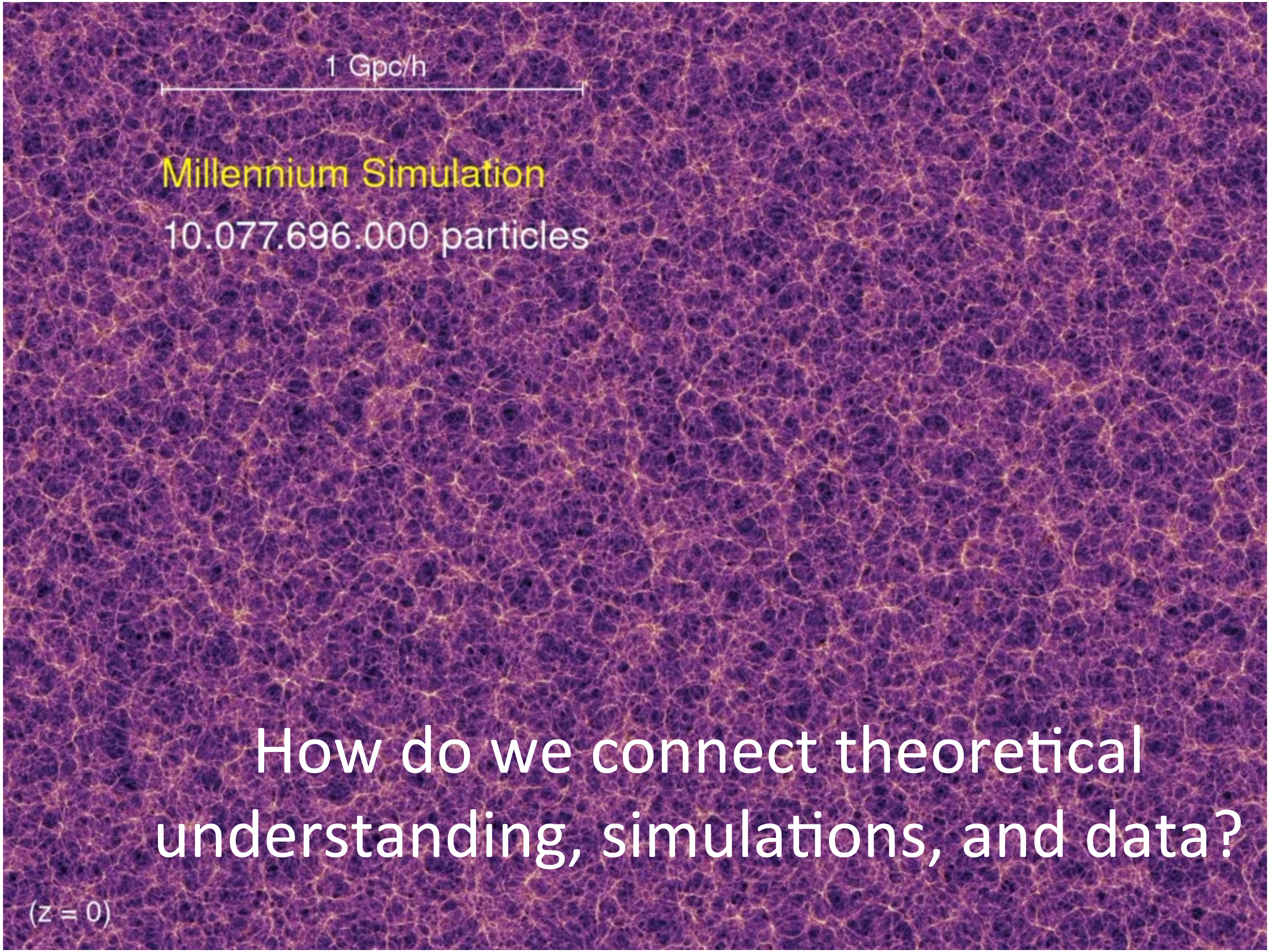
primordial curvature perturbation.

- Can“reverse” processing by linear physics and test model predictions beyond the power spectrum

Komatsu Spergel Wandelt (2005)
Yadav Wandelt (2005)



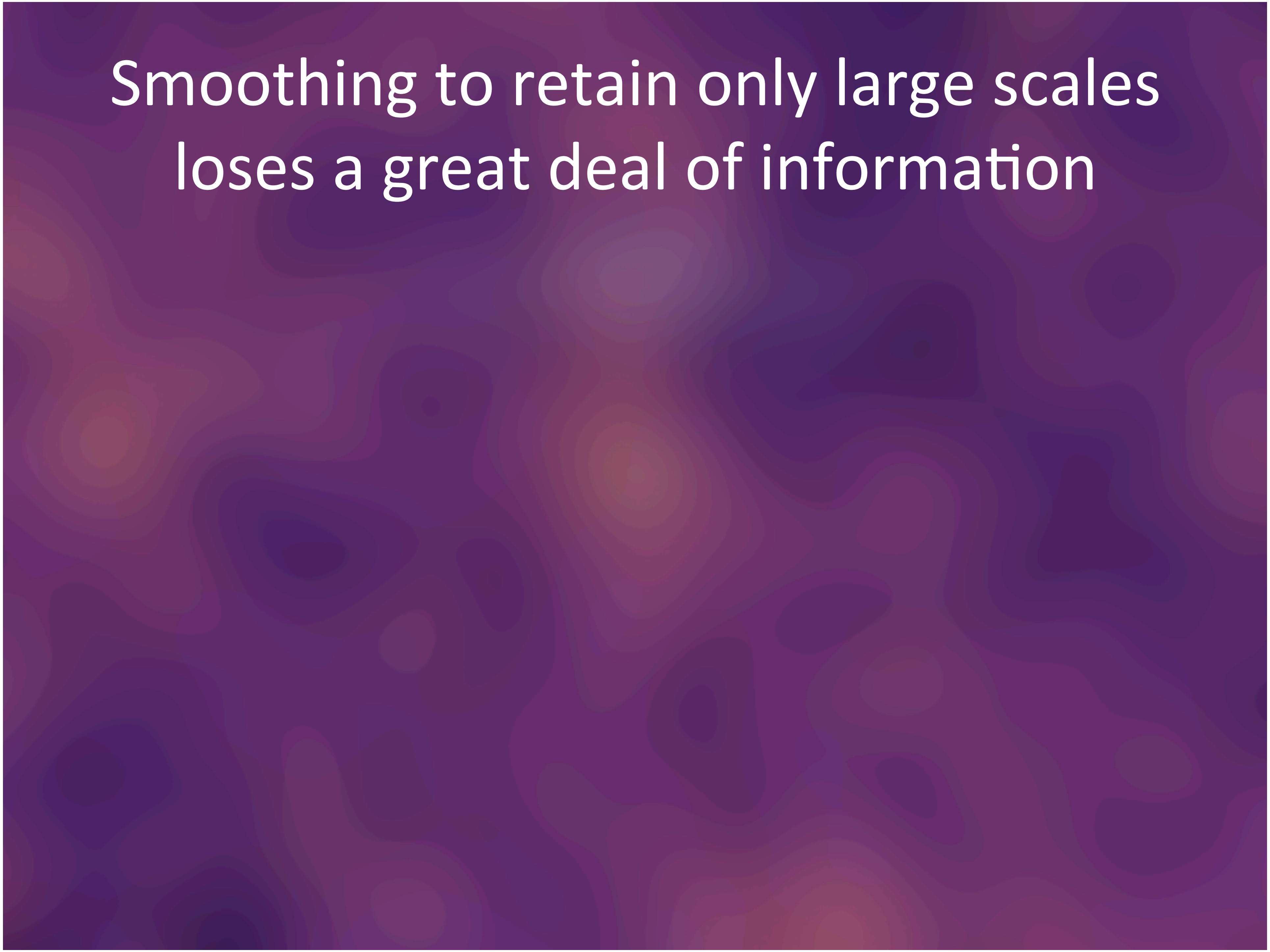
3D large scale structure surveys
are extremely promising
but
non-linear evolution from nearly
Gaussian initial conditions
complicate the inference task



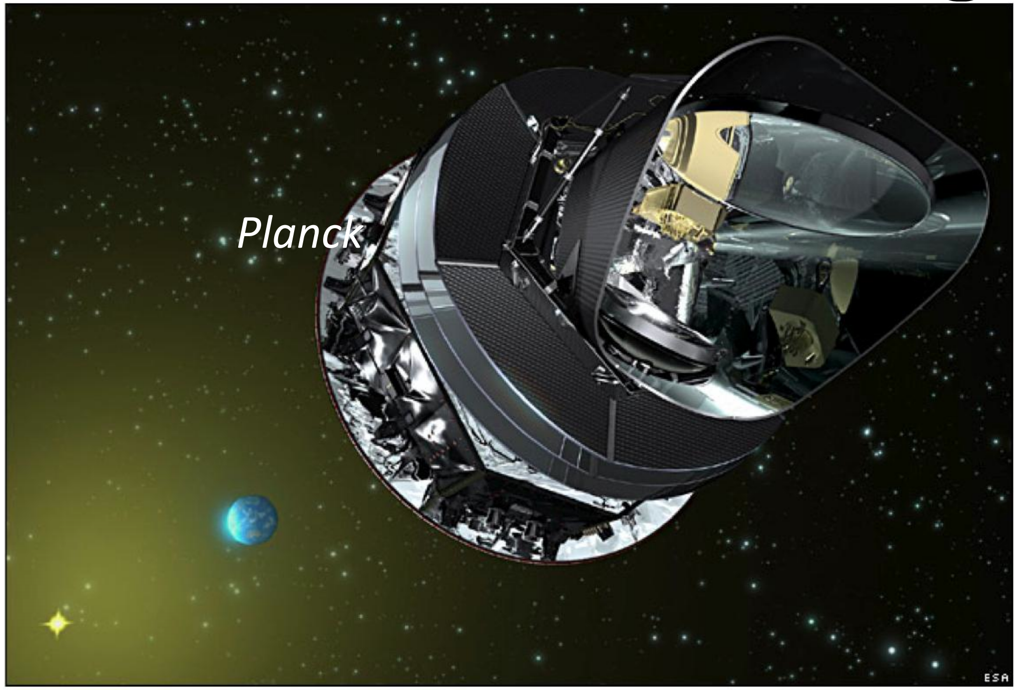
How do we connect theoretical understanding, simulations, and data?

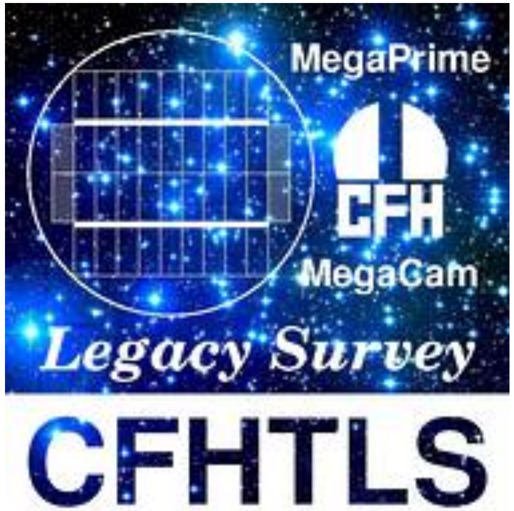
How to deal with non-linearity?

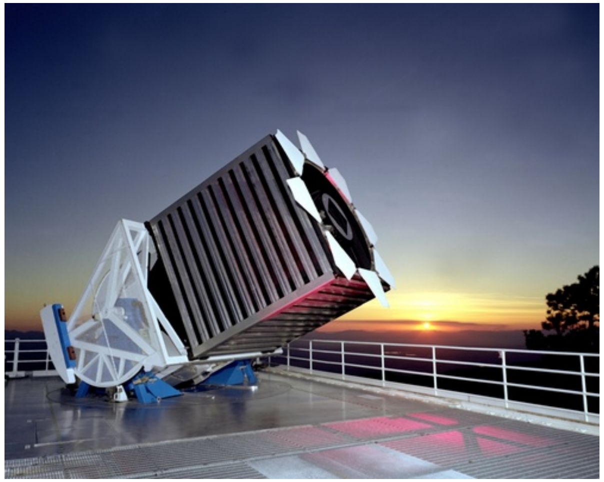
Smoothing to retain only large scales loses a great deal of information

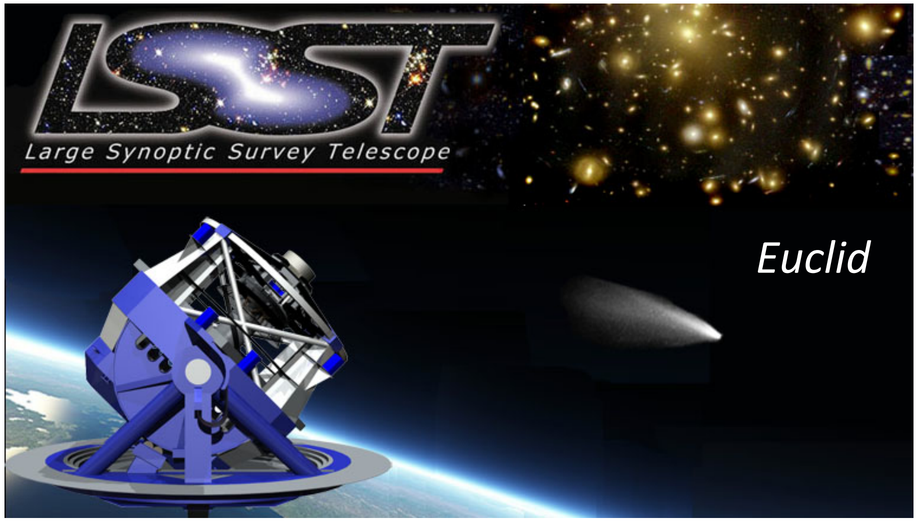


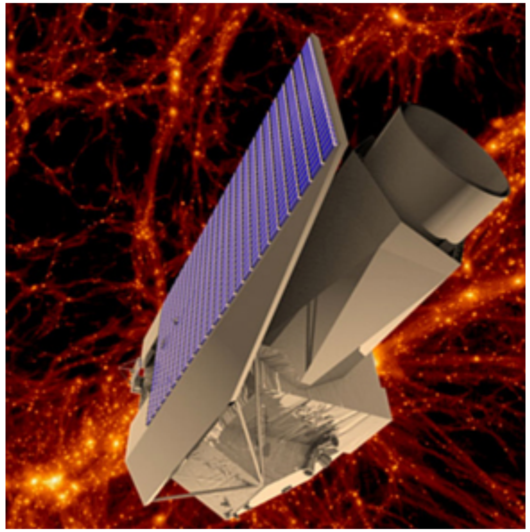
We live in the age of large surveys

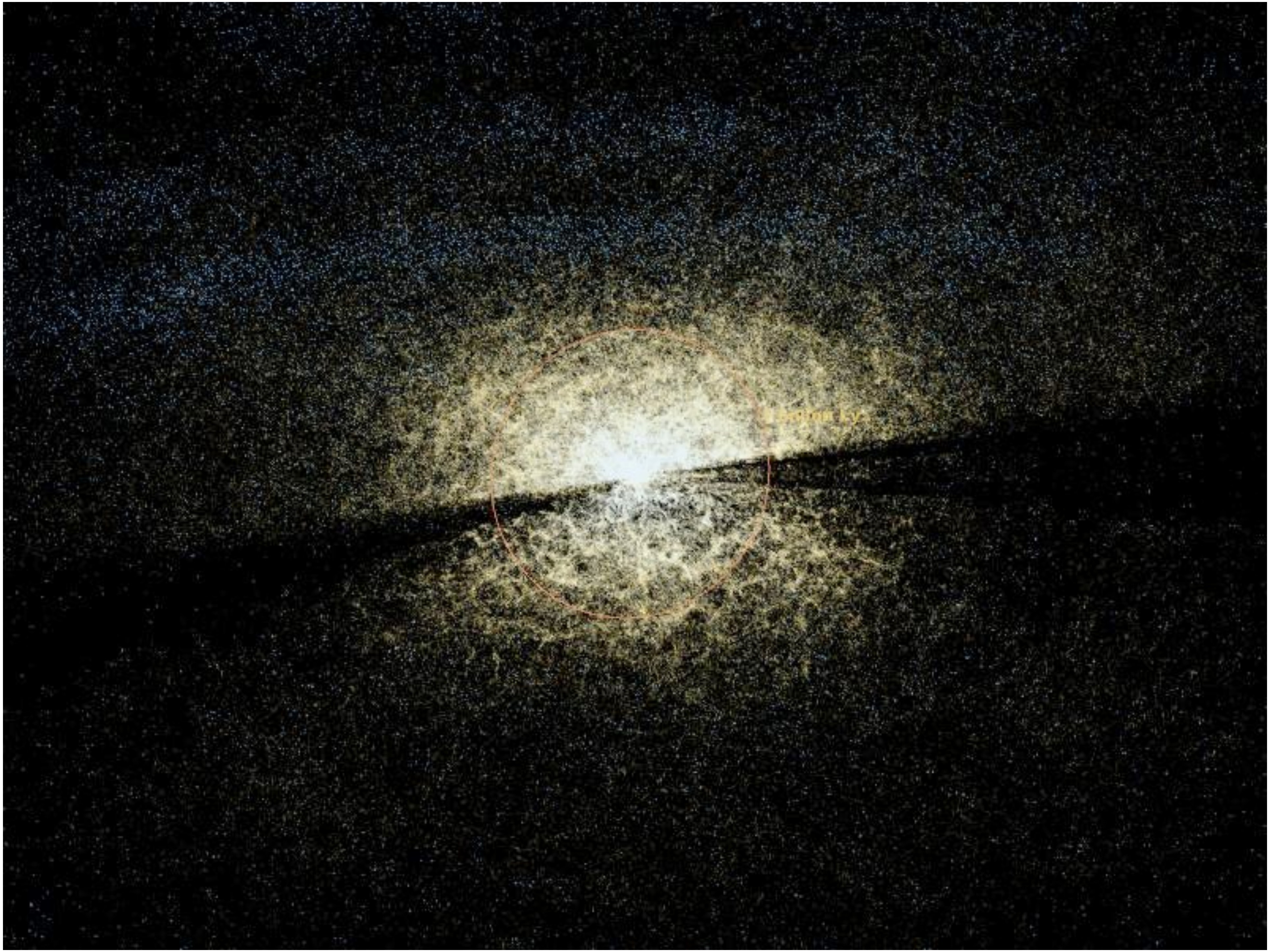




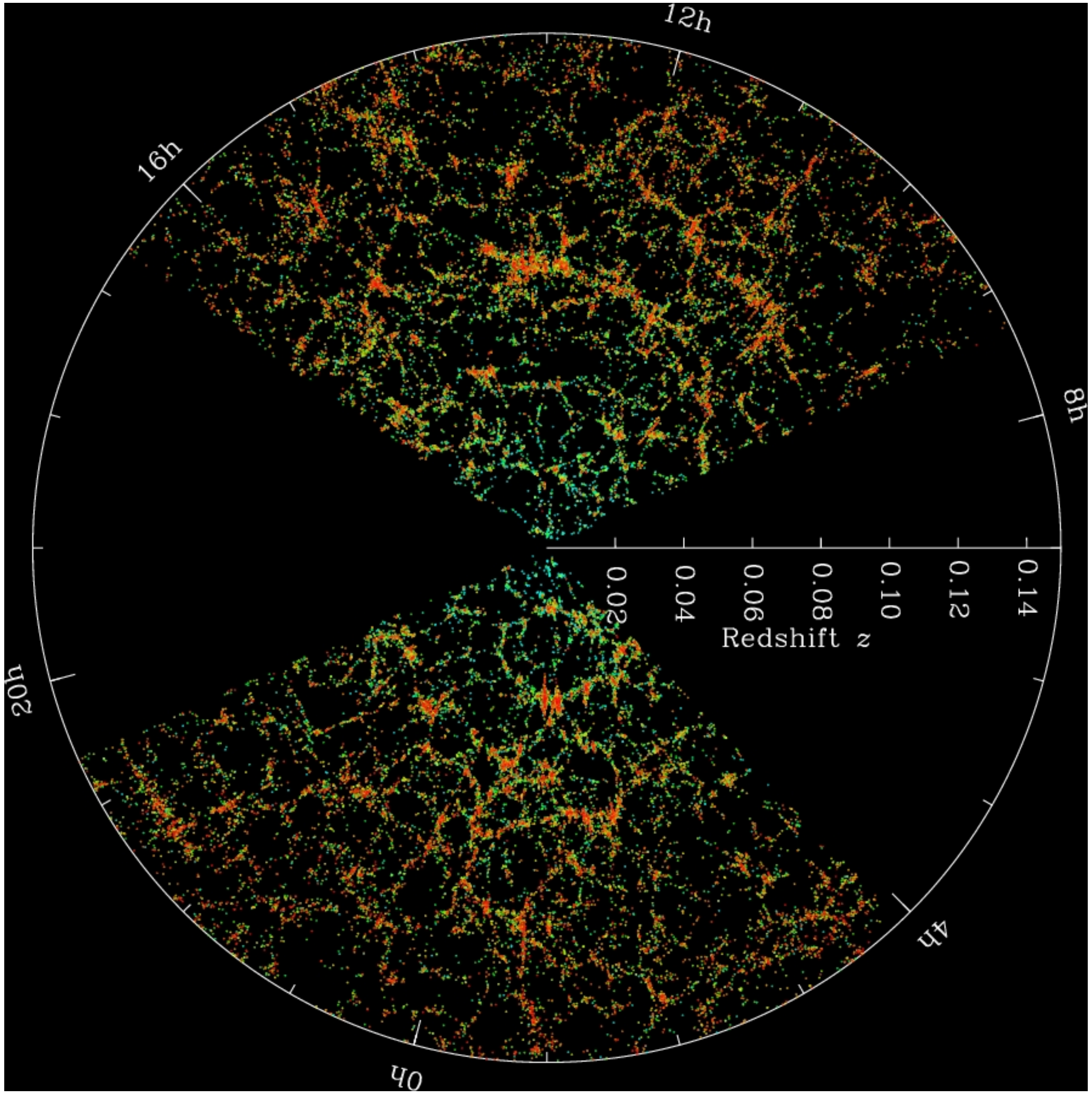








Data example: the SDSS survey



M. Blanton
and SDSS

Peculiarities of learning from LSS surveys

- Solid theoretical foundations – can formulate very good priors
- Awash in data
- But fundamental limits to information:
  - On large scales:
    causality: the observable universe finite
  - On small scales:
    non-linearity: erases primordial information

⇒ Large scales require careful statistical treatment to extract precious information from a relatively small number of modes

⇒ Linear methods are OK only on intermediate scales

⇒ Large potential gains in information when pushing to smaller, non-linear scales since number of modes grows as $1/(\text{length})^3$

⇒ Complex modeling – distances to galaxies not known precisely in very large surveys

⇒ I will present three new lines of approach to these problems developed in my group.

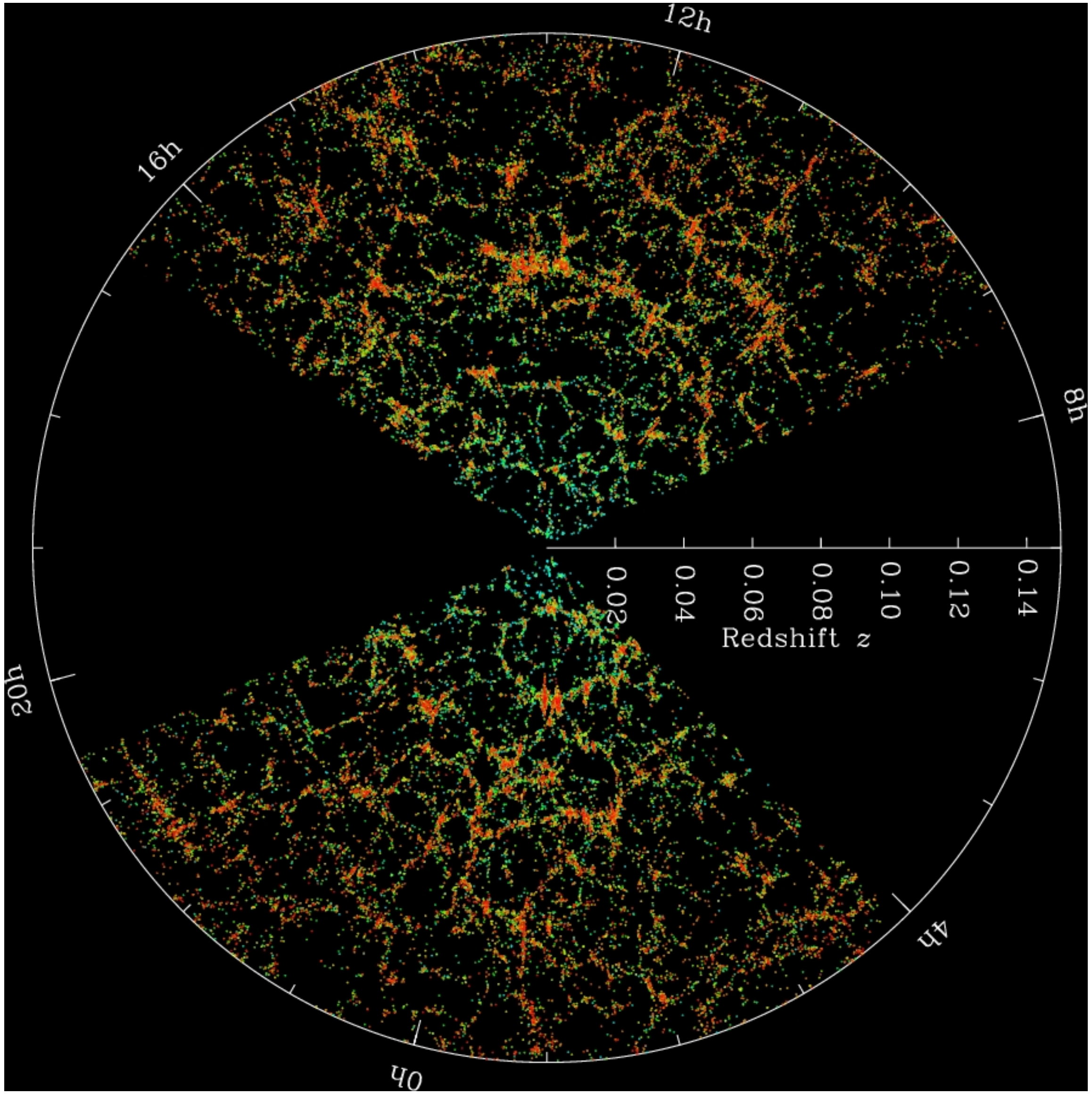
Three new lines of approach in two categories

- Filter/select features that are informative and where systematics can be controlled (frequentist)

- Model systematics and do a Bayesian analysis
  1) Instrumental systematics/missing information
     - photo-z errors
     - Galaxy bias
  2) Gravitational non-linearity

Instead of fighting non-linearity, can we embrace it?

Let's look at the data again

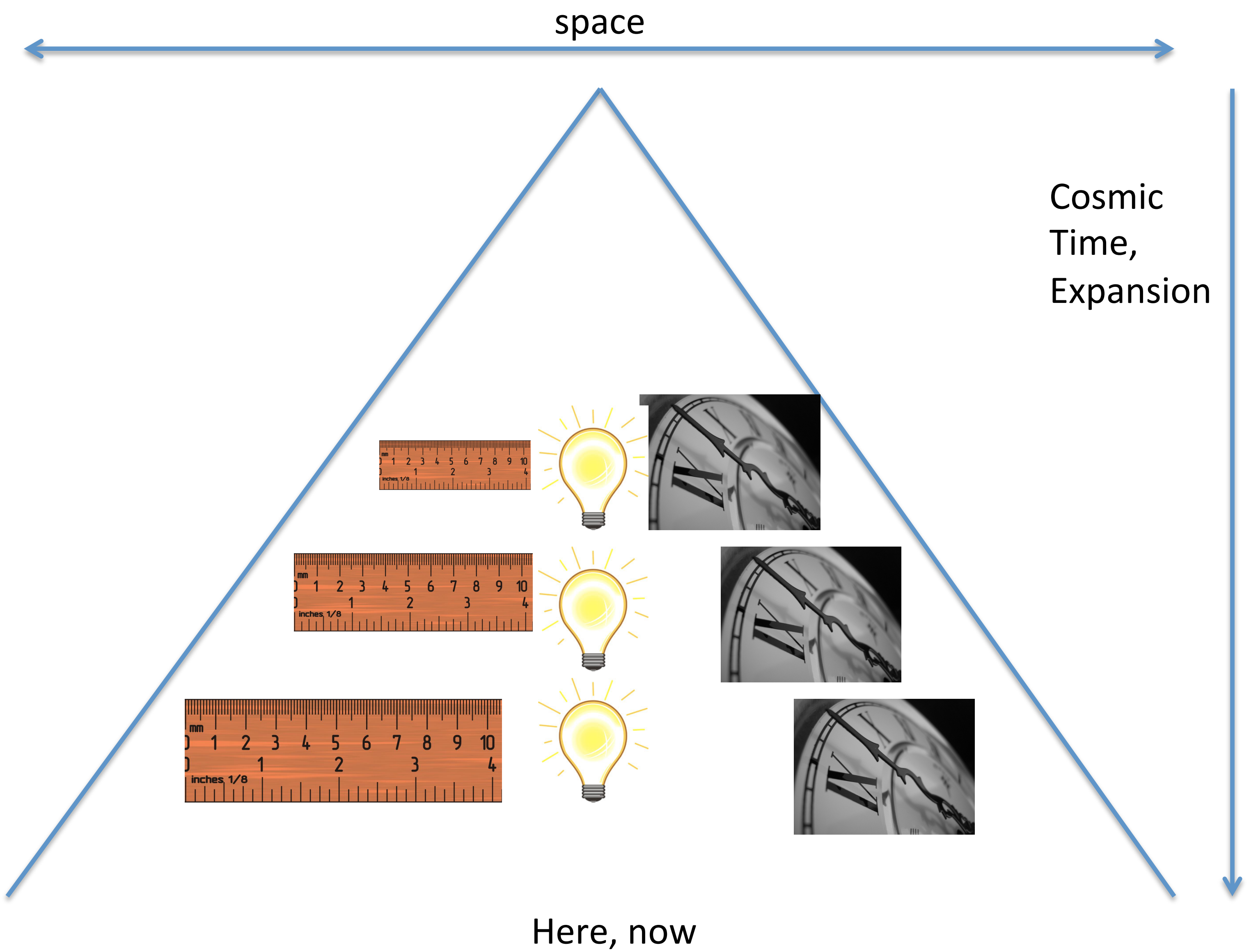


M. Blanton
and SDSS

Non-linearity creates foam like structure of voids

Can we use those non-linear voids to learn about dark energy?

Cosmography

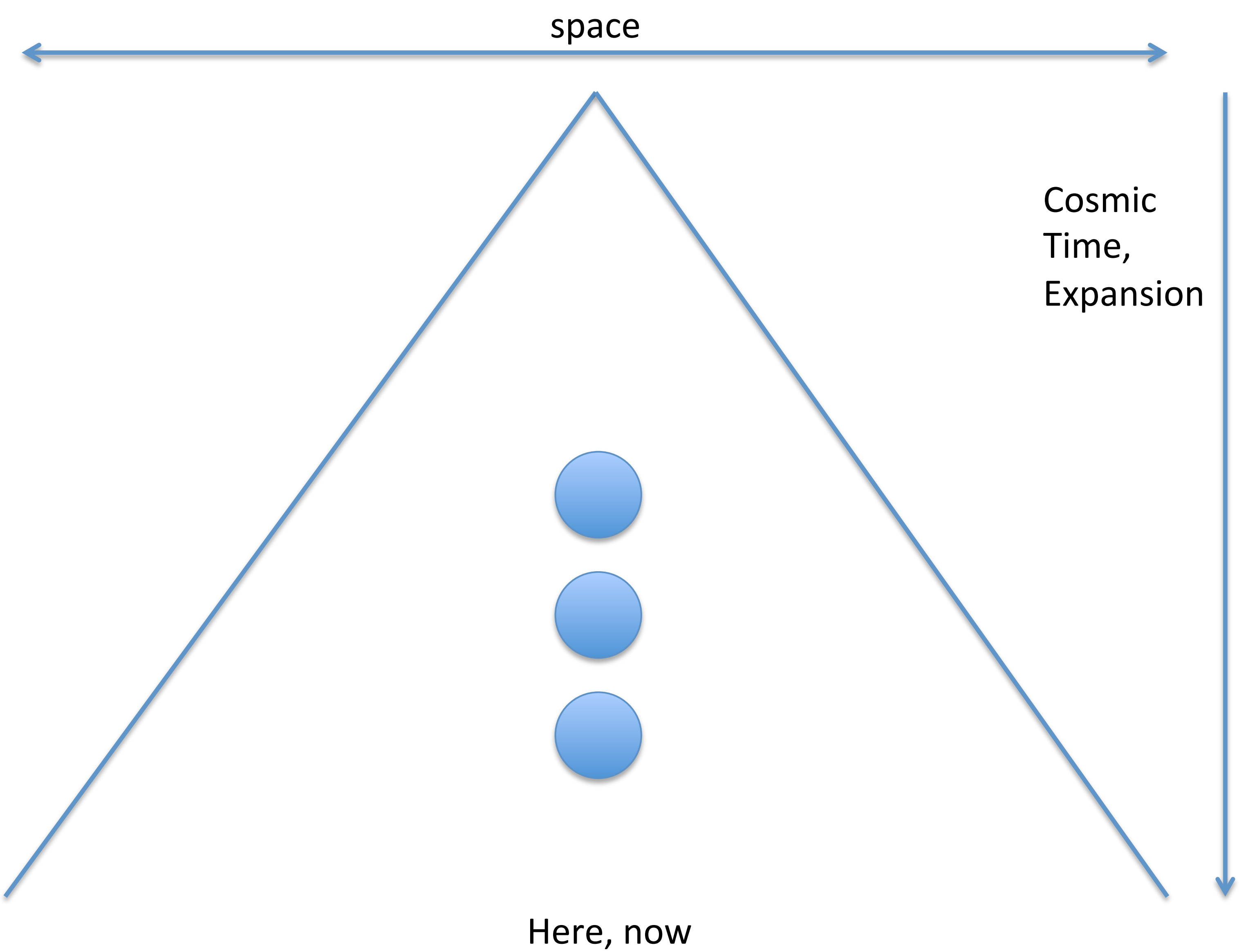


Cosmic stopwatches

- Good clocks (with long term stability) are hard to find
- But what if we could define cosmic stopwatches
- Standard spheres are like cosmic stopwatches
- This is the Alcock-Paczynsky technique
- First realized for voids by Ryden in 1995!

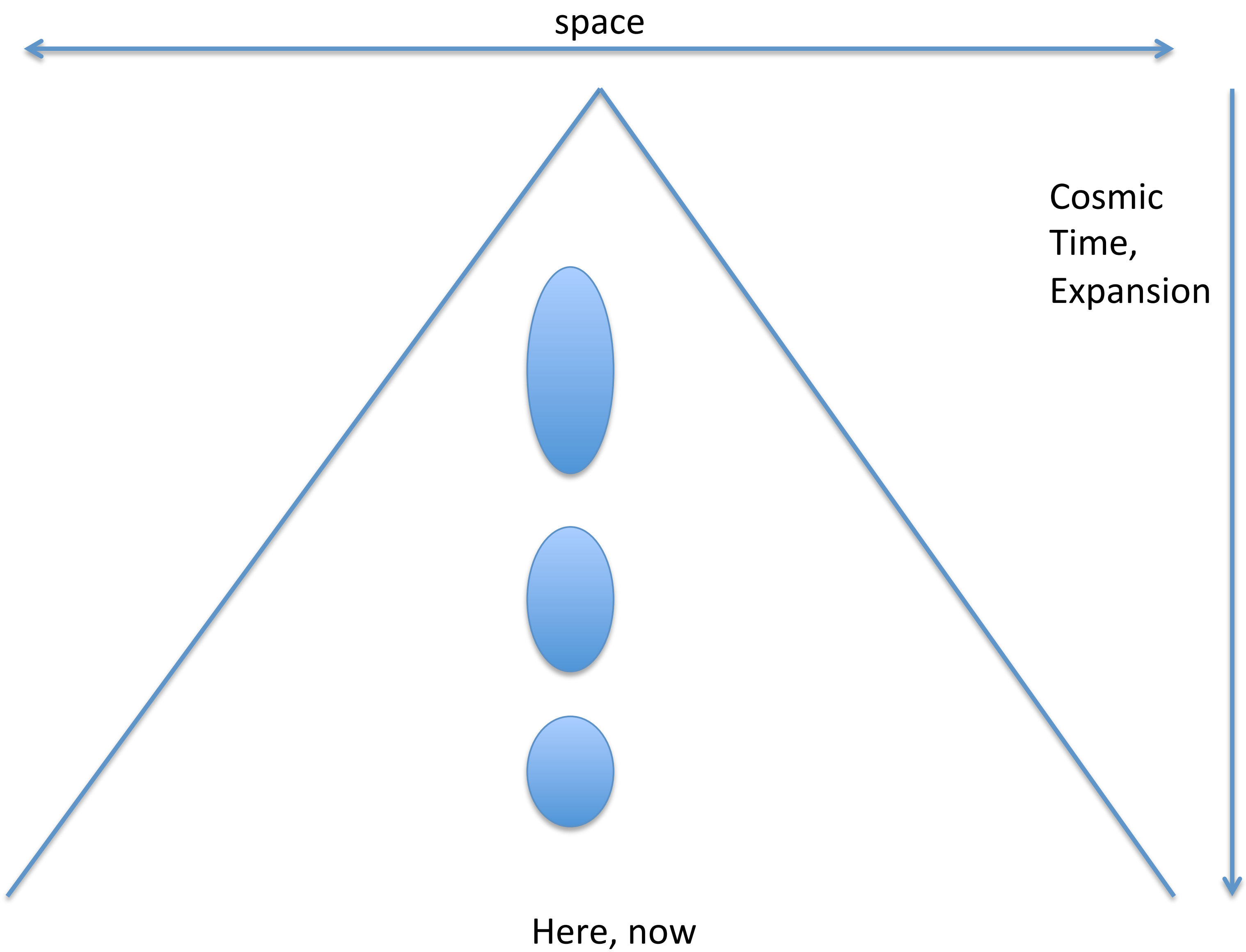
Lavaux & Wandelt, 2011

Cosmography with voids



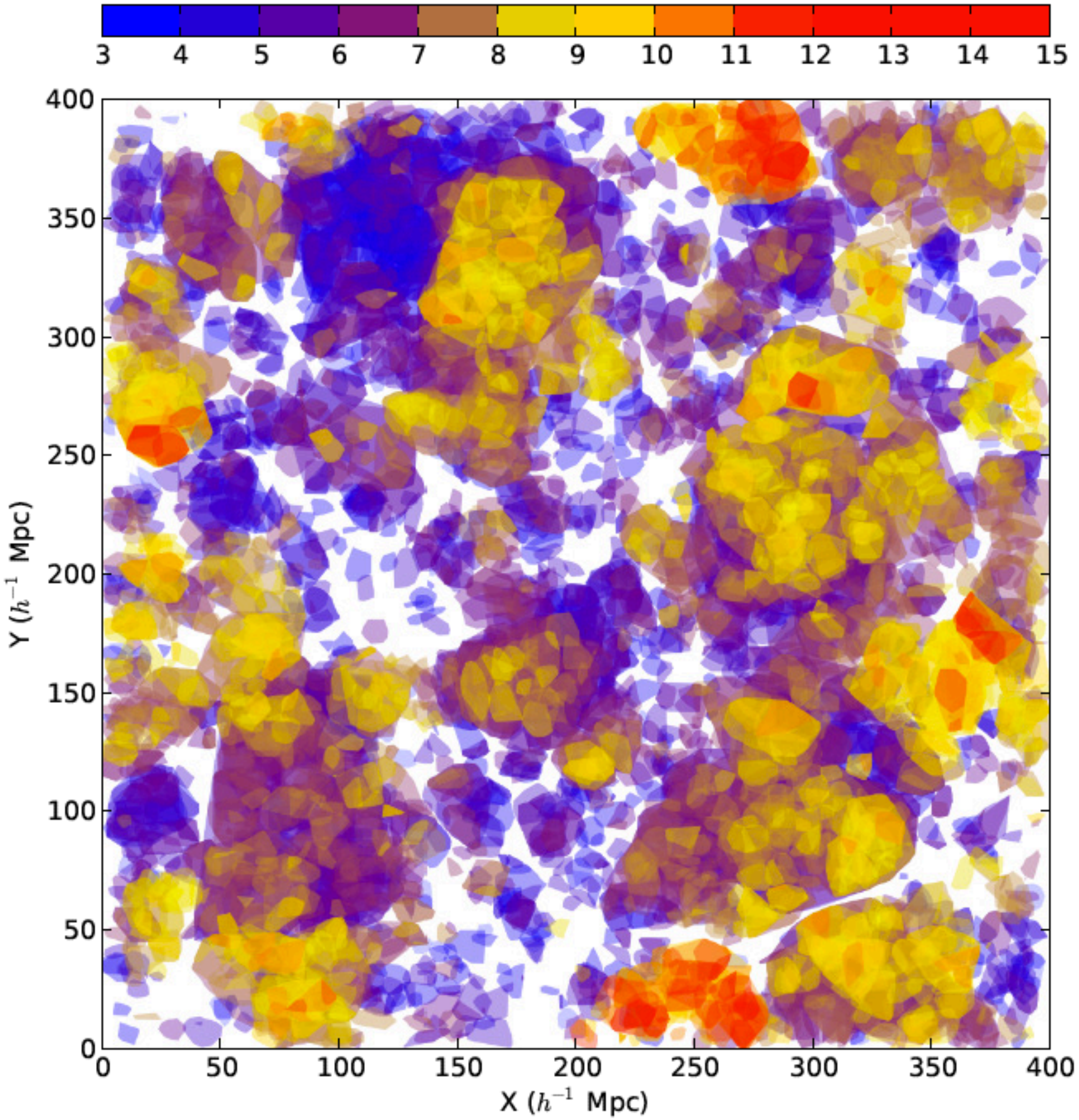
Lavaux & Wandelt, 2011

Cosmography with voids



Lavaux & Wandelt, 2011

Spatial signal processing: finding voids



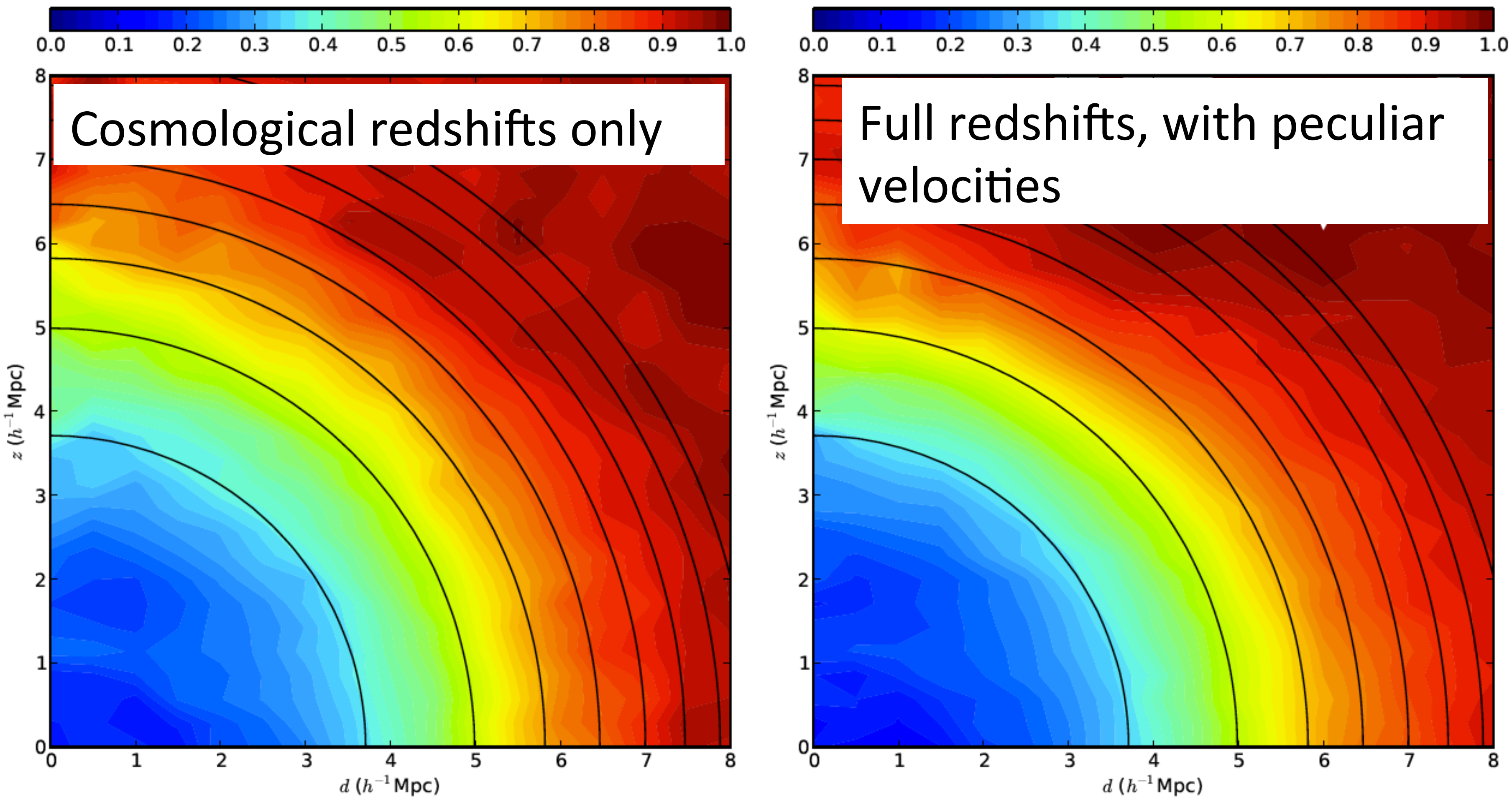
Lavaux & Wandelt, 2011

Challenges

- Voids are *not* spheres – they have complicated shapes
  - Solution: **void stacking**
    - Take particles in all detected void regions in a redshift shell, co-center and merge them
    - Voids are spherical *on average* (in physical coordinates).
- Tracers (galaxies) move – this distorts the voids systematically in redshift space
  - This is a much smaller effect for low density regions than high density regions

Lavaux & Wandelt, 2011

Void stacking



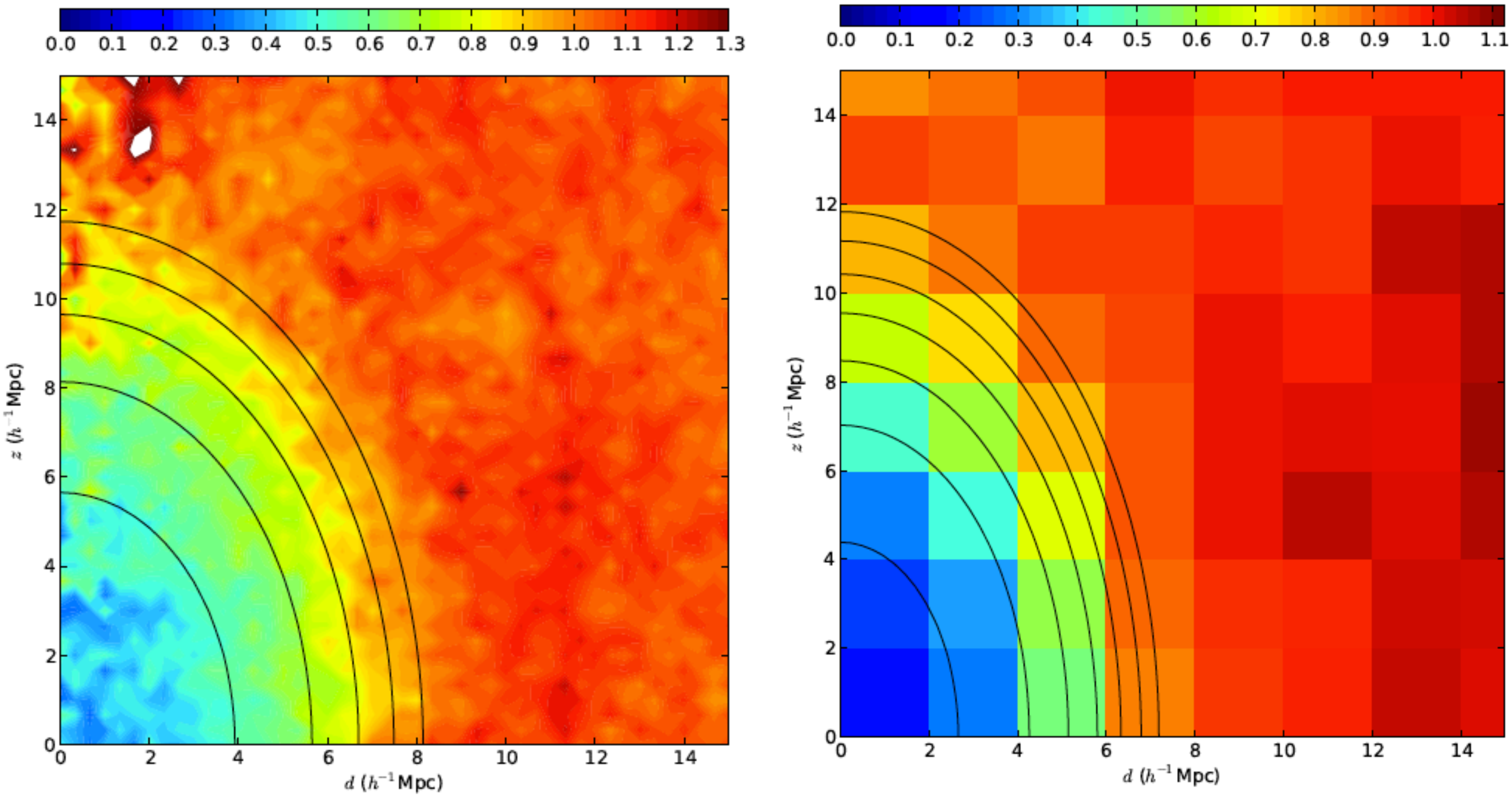
Lavaux & Wandelt, 2011

Challenges III

- Void walls have structure => shape noise
  - Some residual clumping remains after stacking
  - Solution: choose pixel size large enough (~2 h^{-1} Mpc) that clumps only contribute to one pixel.
  - This allows treating clumping noise as independent in pixel space.
  - Bayesian MCMC procedure for fitting an ellipsoidal cubic density profile, including a clumping noise parameter.

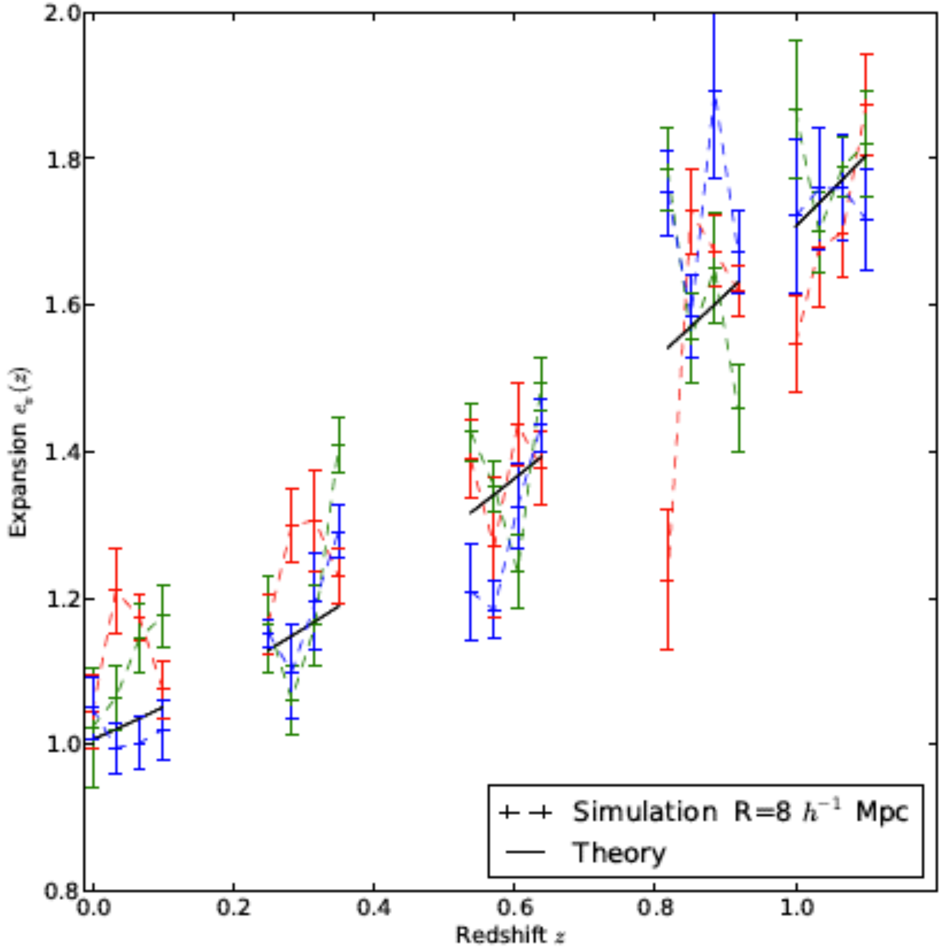
Lavaux & Wandelt, 2011

Coarser pixelization de-correlates clumping error



Lavaux & Wandelt, 2011

AP-Hubble diagram

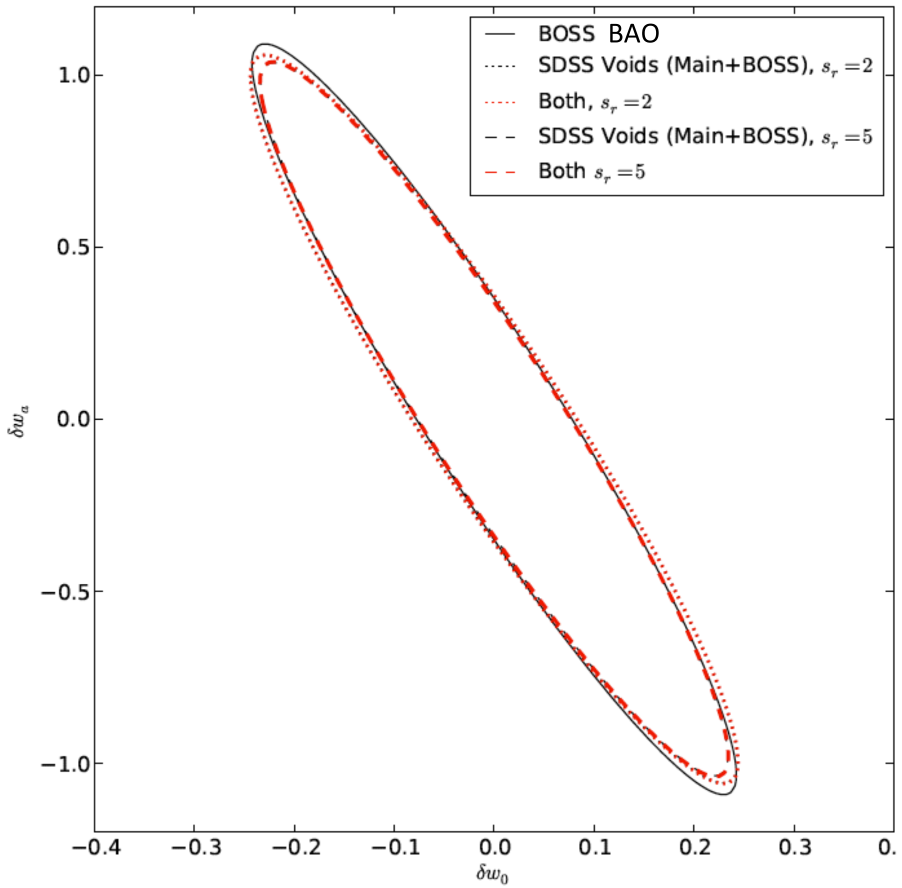


Lavaux & Wandelt, 2011

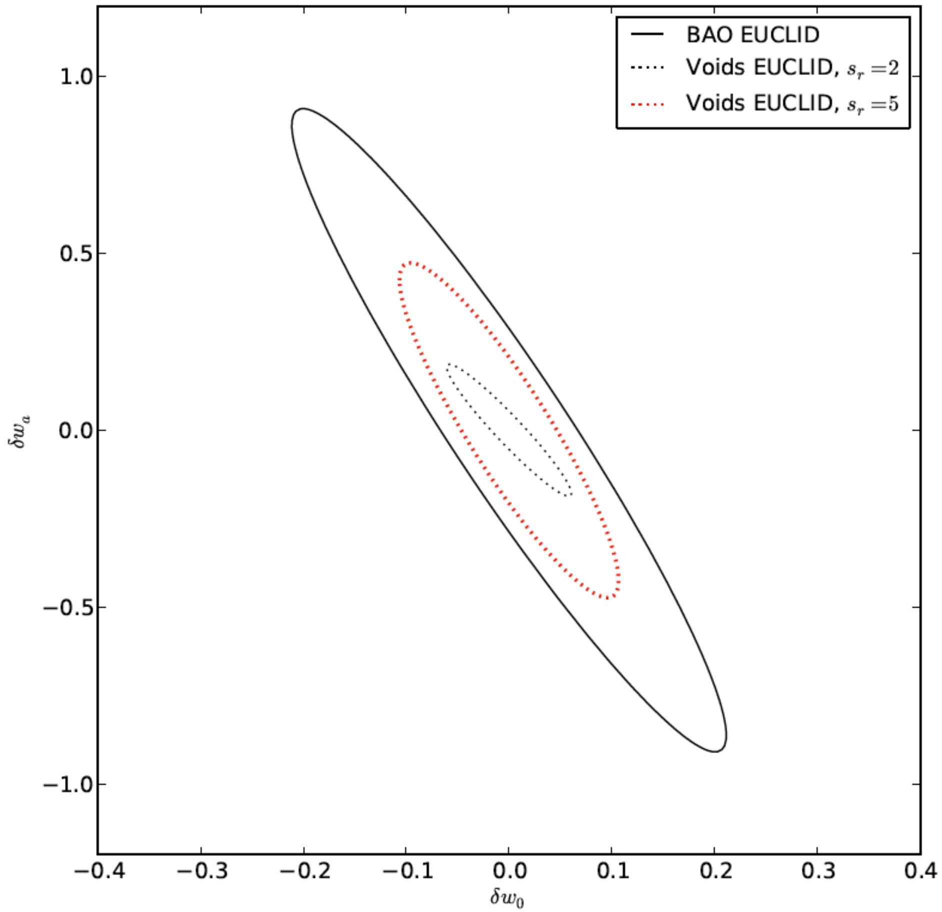
Sutter, Lavaux, Wandelt, Weinberg
arxiv:1208.1058

Dark energy constraint forecast

Lavaux & Wandelt, 2011

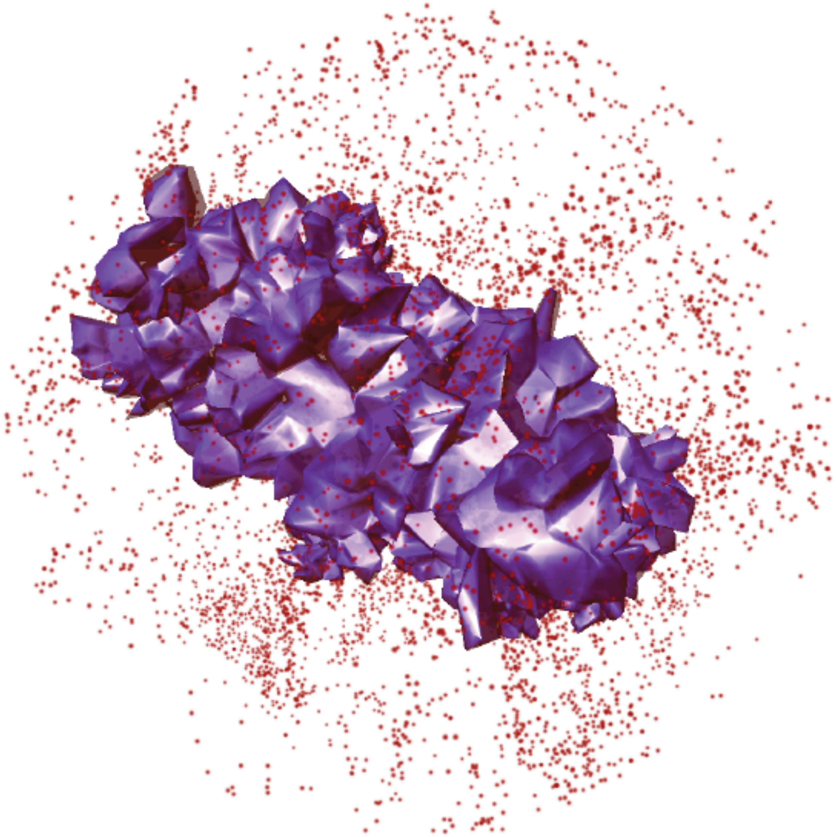


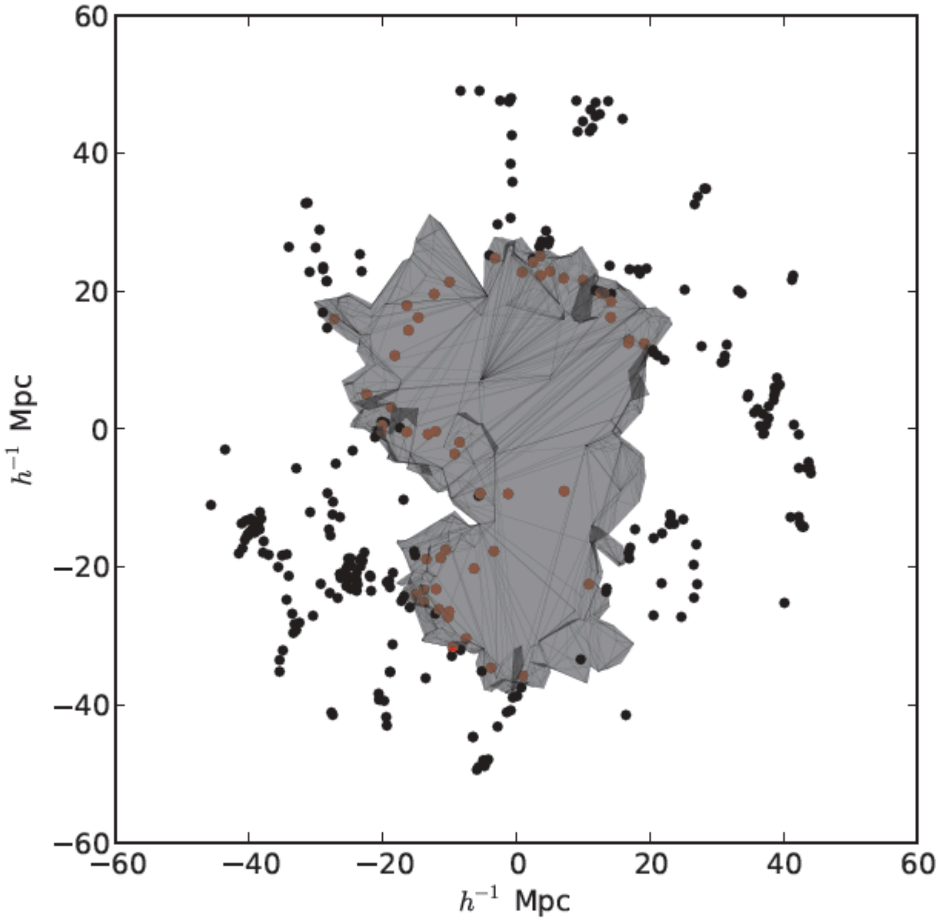
Comparable to DETF Baryon Acoustic Oscillation constraints for upcoming data



Outperforms BAO by a factor of O(10) for future data, such as EUCLID; voids alone yield double the combined FoM.

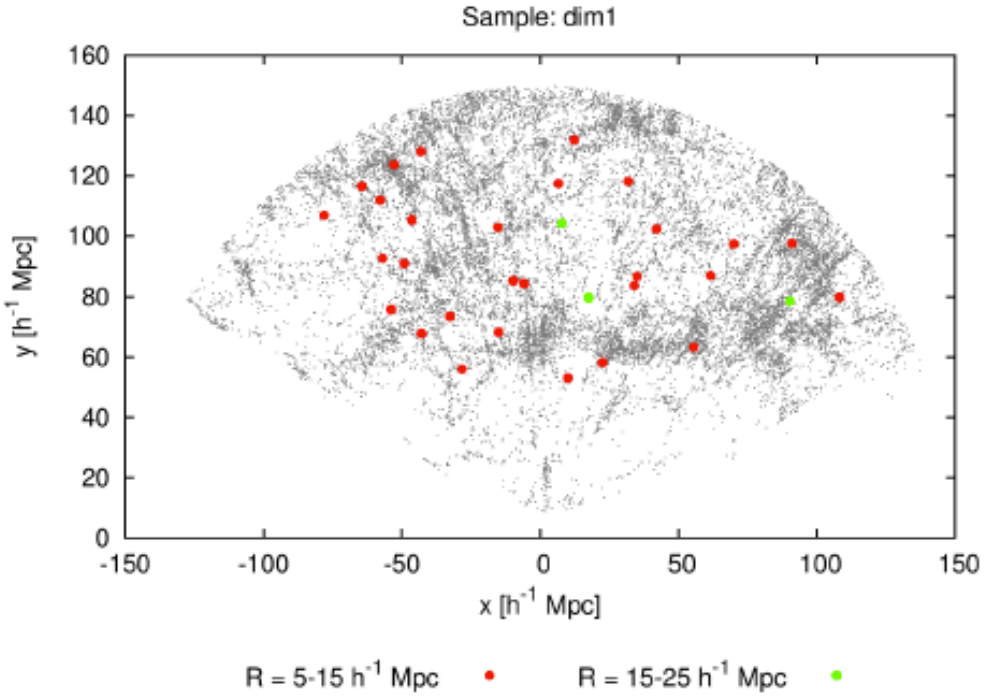
A real void

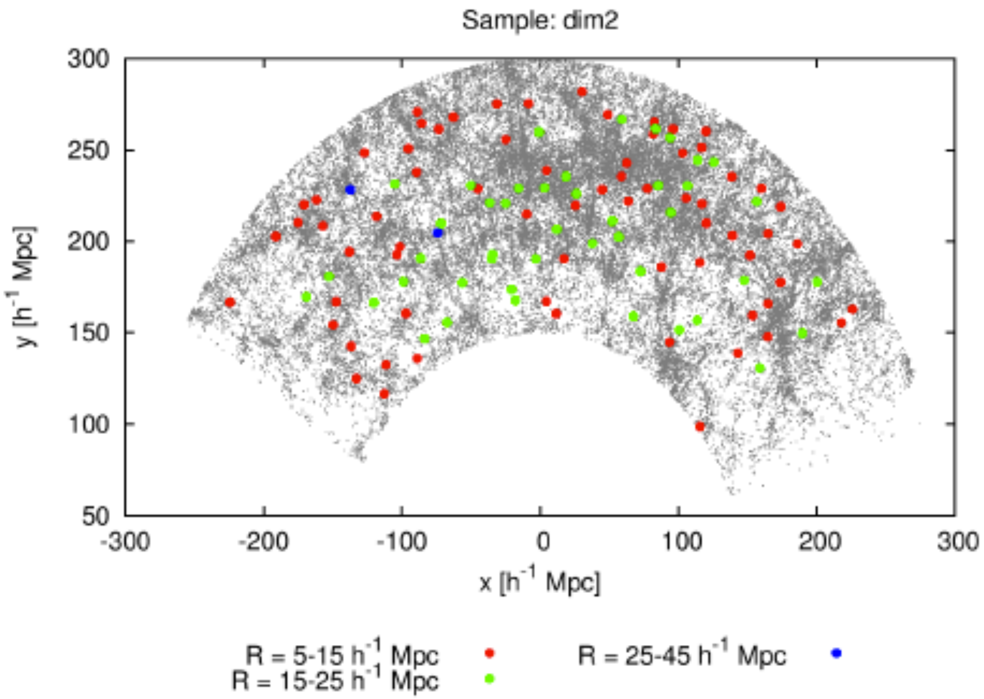




Sutter, Lavaux, Wandelt, Weinberg, arxiv:1207.2524

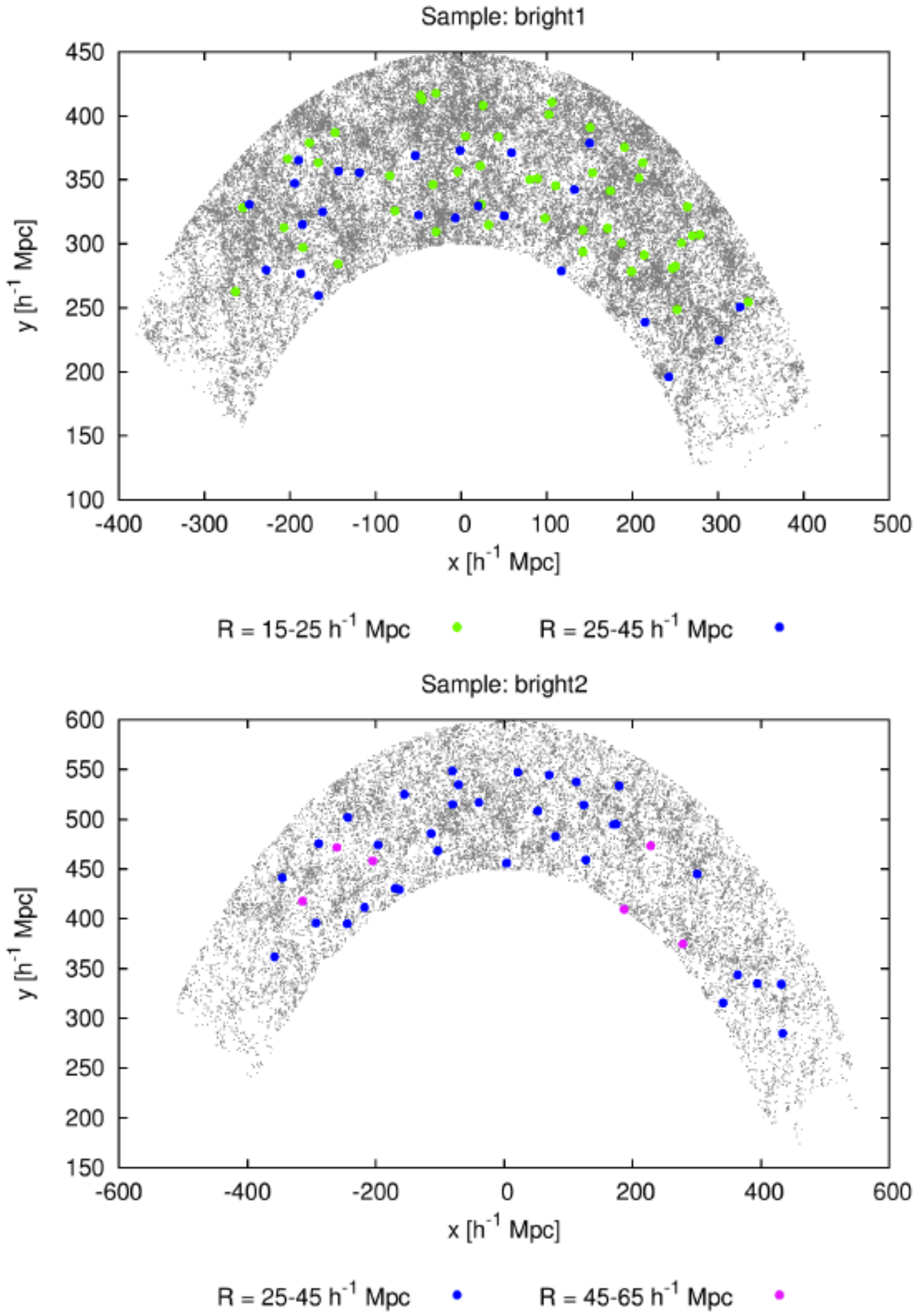
Voids in SDSS main sample I





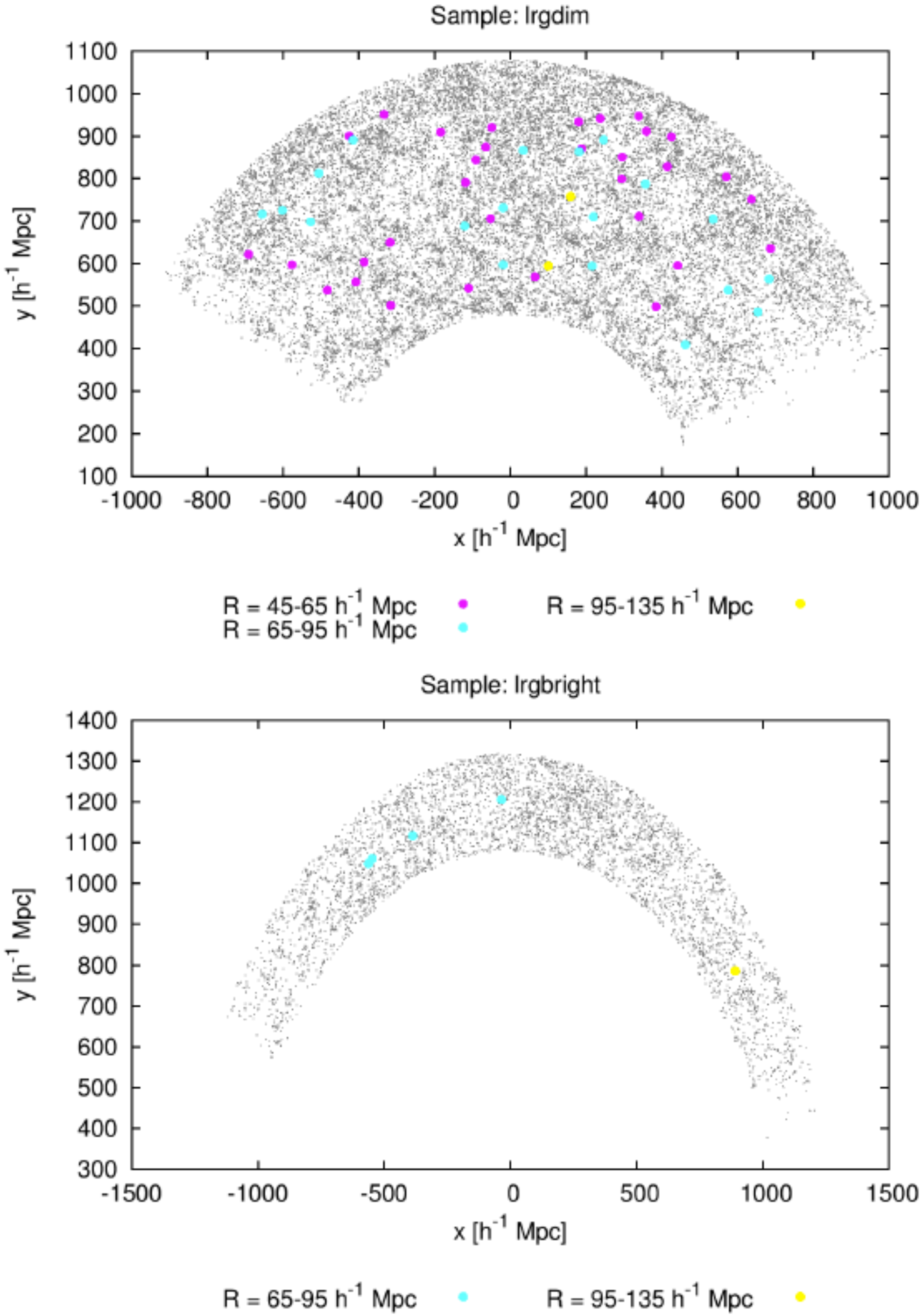
Sutter, Lavaux, Wandelt, Weinberg 2012

Voids in SDSS main sample II



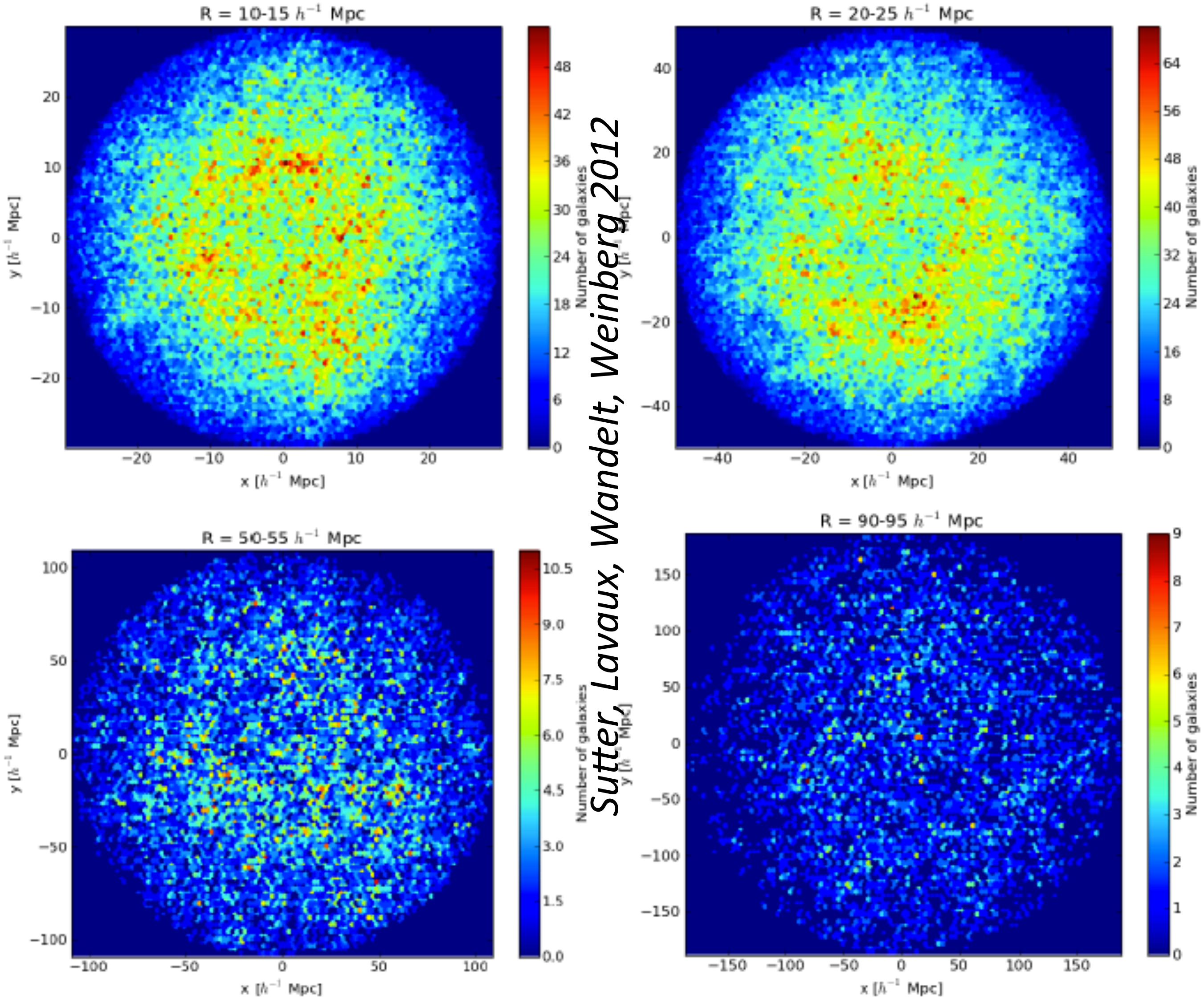
Sutter, Lavaux, Wandelt, Weinberg 2012

Voids in the SDSS LRG



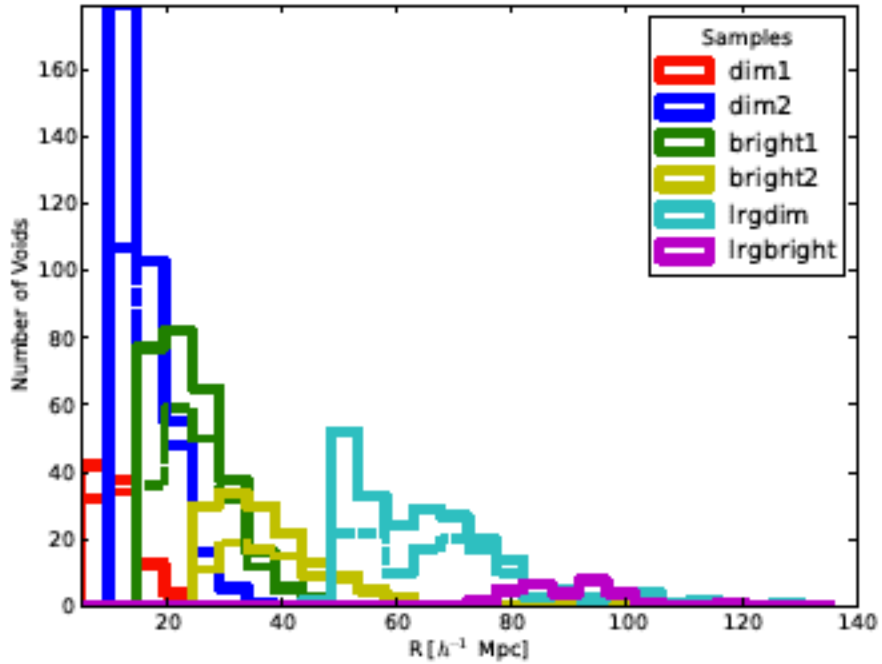
Sutter, Lavaux, Wandelt, Weinberg 2012

SDSS Void stacks projected on the sky



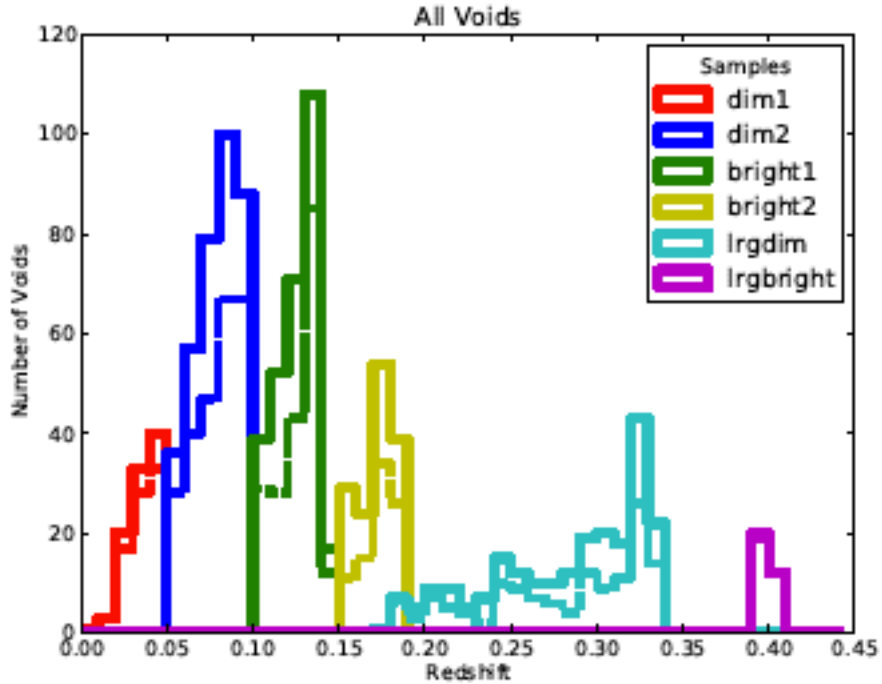
Sutter, Lavaux, Wandelt, Weinberg 2012

Void sizes



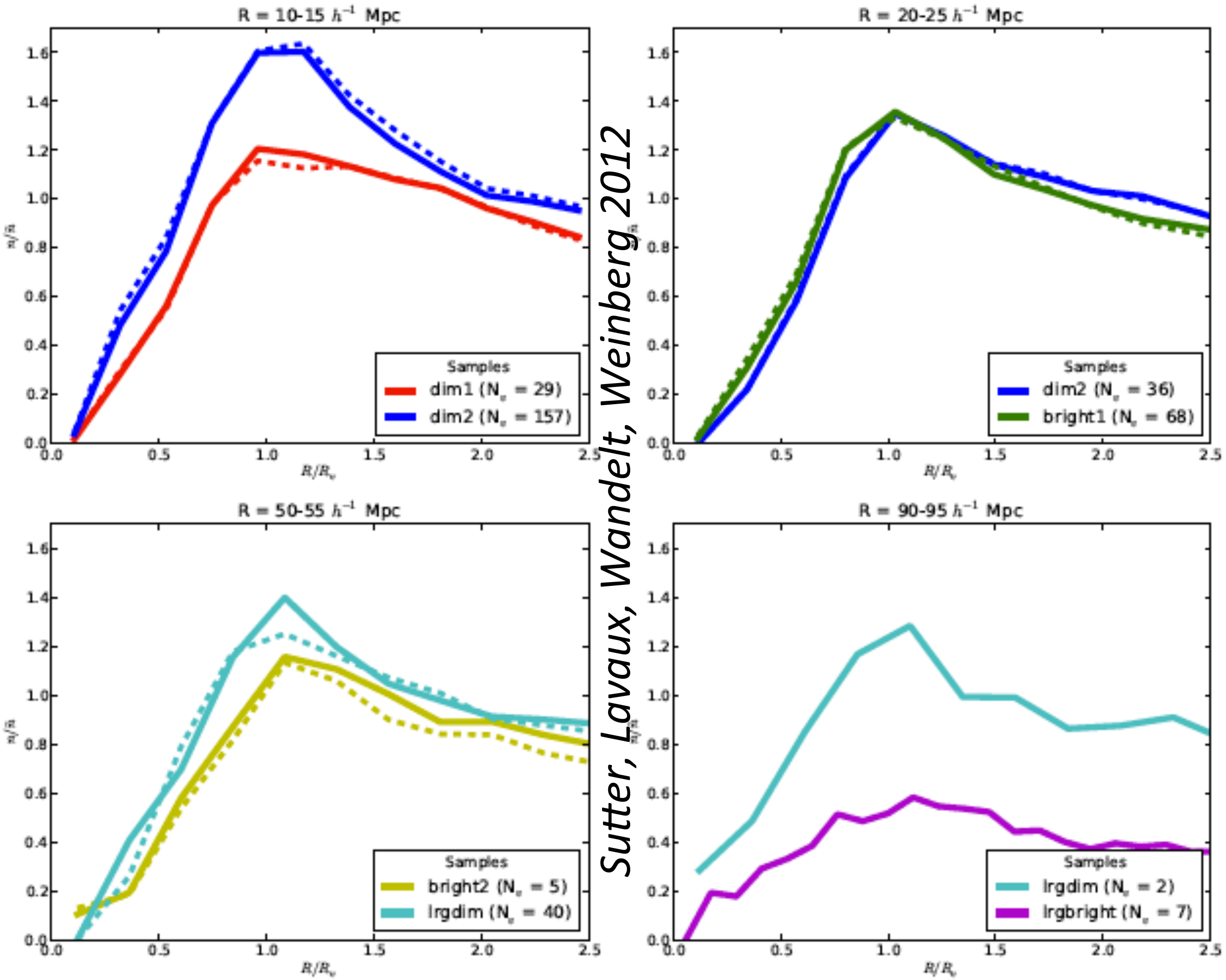
Sutter, Lavaux, Wandelt, Weinberg 2012

Void number



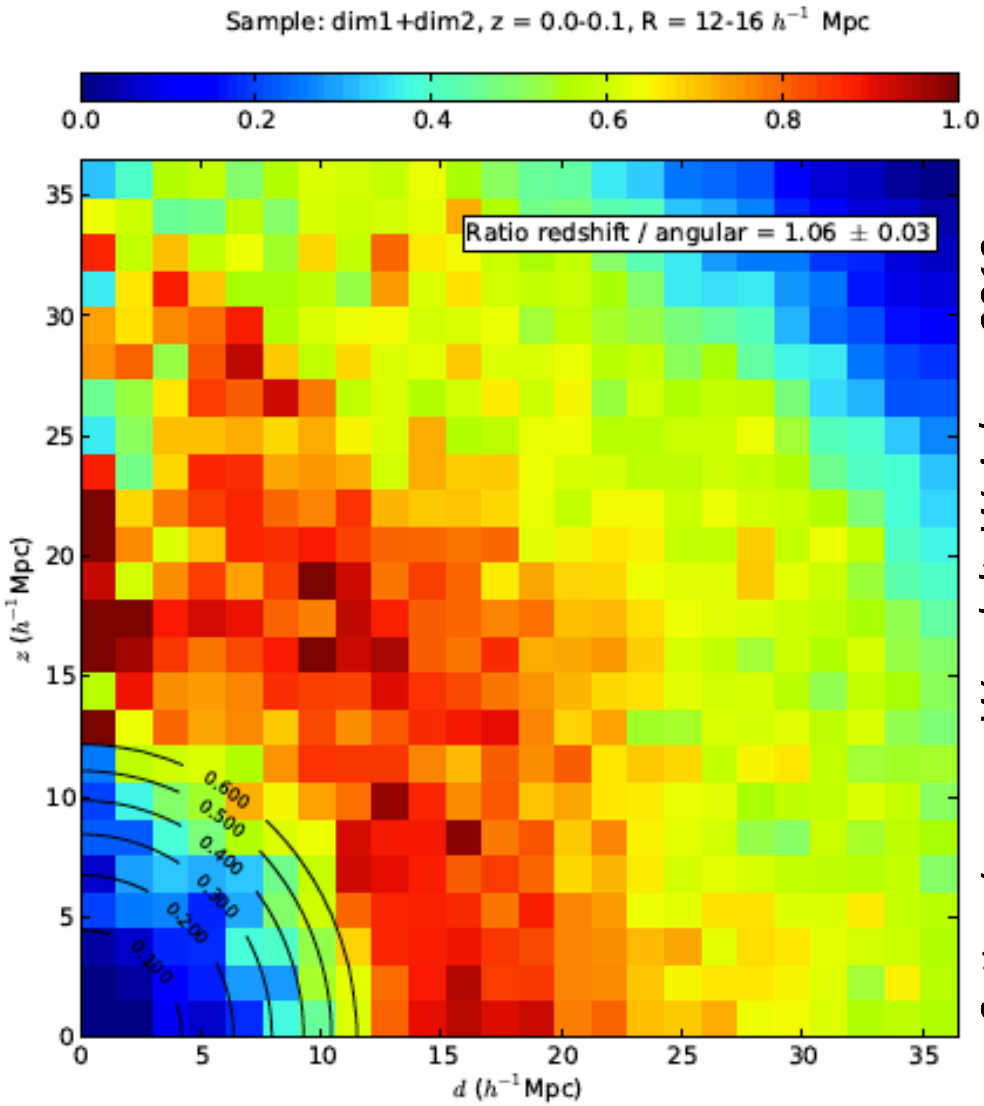
Sutter, Lavaux, Wandelt, Weinberg 2012

Void profiles



Sutter, Lavaux, Wandelt, Weinberg 2012

A void stack from Sutter et al. 2012



Sutter, Lavaux, Wandelt, Weinberg 2012

A publicly available catalog cosmicvoids.net

Cosmic Voids

Search this site

Home | Void Identification Algorithm | Public Catalogs | News and Updates | Contact

Welcome to the Public Cosmic Void Catalog

This is the repository for the public releases of a comprehensive cosmic void catalog from galaxy redshift surveys. This catalog is the product of a collaboration of P.M. Sutter (Illinois/IAP/OSU), Benjamin Wandelt (IAP/UPMC/Illinois), Guilhem Lavaux (Perimeter), and David Weinberg (OSU). Our void finder algorithm is based on ZOBOV, which used Voronoi tessellations and the watershed transform to identify voids. See here for the journal article describing our method used for defining and cataloging voids.

Catalog at a Glance:

The catalog contains all the information required to reproduce the journal article. This means that the catalog contains the raw ZOBOV-generated catalog and all derived data products, such as:

- Void barycenters, redshifts, effective radii, and redshifts
- Redshifts and sky positions of member galaxies
- One-dimensional radial profiles of stacked voids
- Two-dimensional projections of stacked voids
- Redshift-dependent void number counts
- Void size distributions

Catalog Objectives

The majority of very large ongoing and future surveys will be photometric rather than spectroscopic

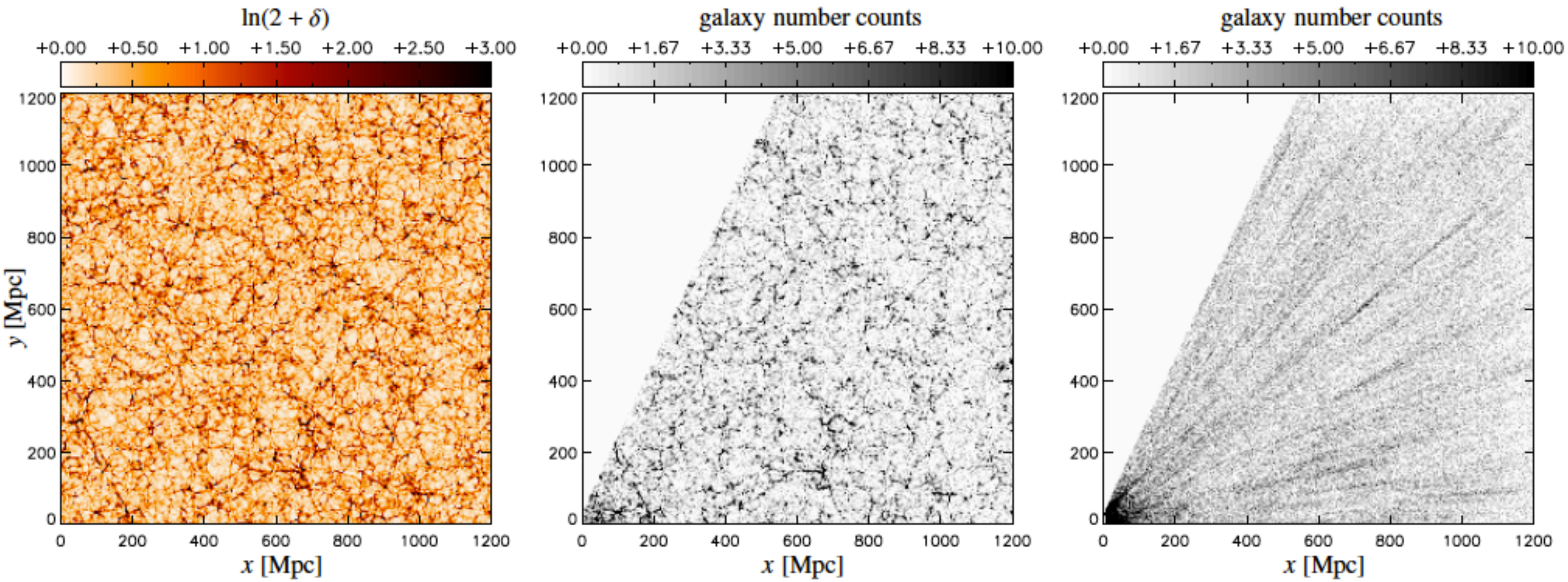
DARK ENERGY SURVEY

Pan-STARRS

MegaPrime
CFH
MegaCam
Legacy Survey
CFHTLS

LSST
Large Synoptic Survey Telescope

Redshifts based on photometry wipe out 3D structure on ~100Mpc/h scales



If we believe that the universe is homogenous and isotropic we can add this as prior information

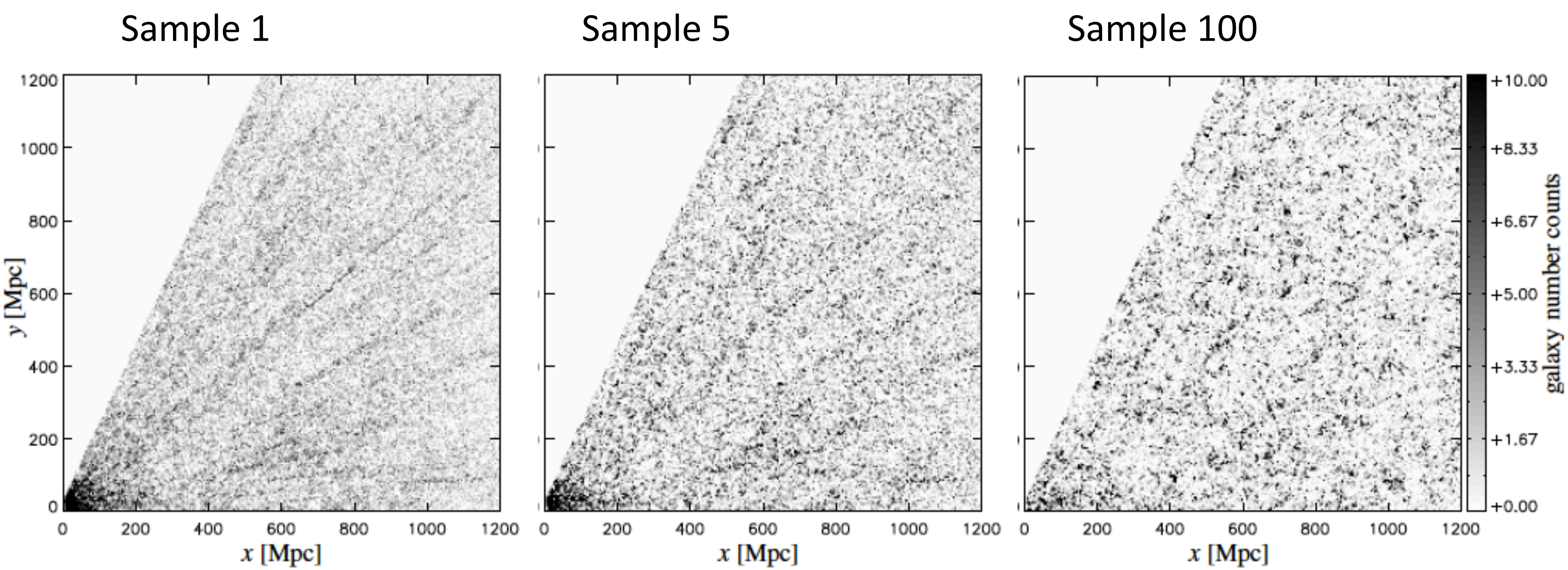
Bayesian joint/global reconstruction of cosmic density field and galaxy positions from a photo-z survey

Assumptions:

- Correlated, isotropic, log-normal model for the density field
- Galaxies are modeled as a Poisson sample from the density
- Inputs:
  - >20 million photo-z pdfs
  - P(k) for a cosmological model (can also be jointly inferred)
- Technique:
  - Block Metropolis-Gibbs sampling
  - Hamiltonian sampler for density field (1.6×10^7 parameters)
- Outputs:
  - samples from the density field
  - photo-z posterior pdfs
- Note that our simulations are from cosmological density fields – they violate the log-normal prior

Jasche and Wandelt arxiv: 1106.2757

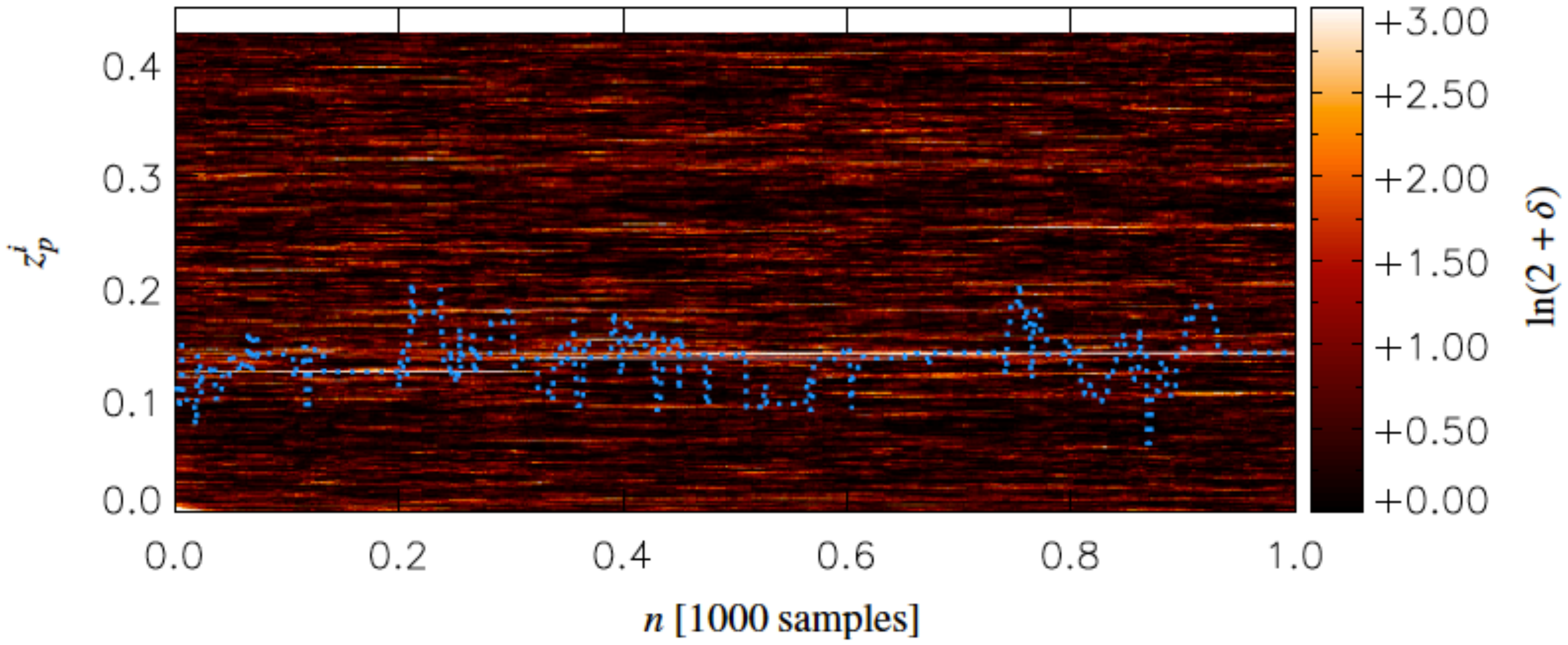
The reconstruction in action



Movie

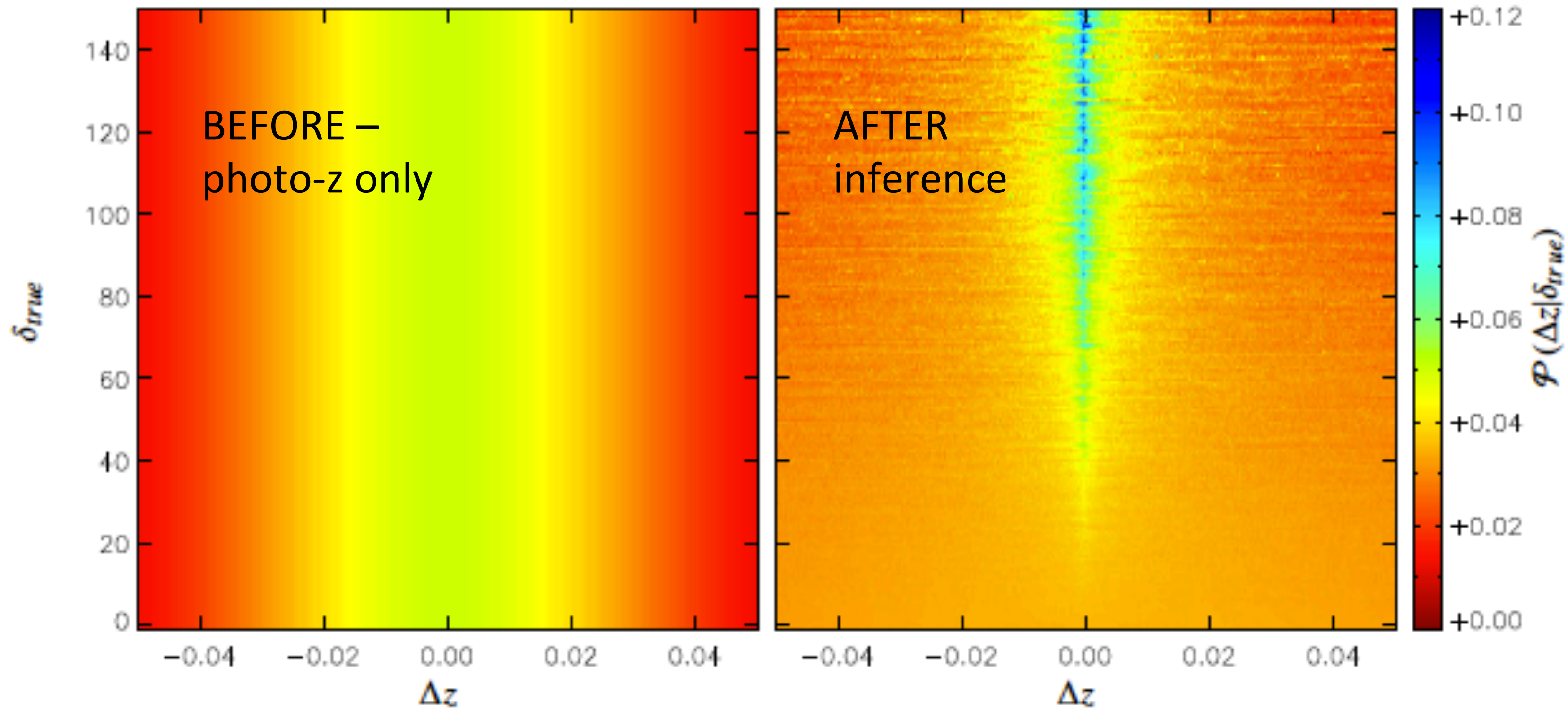
Jasche and Wandelt, arxiv:1106.2757

Galaxies random walk through the (dynamically updated) density reconstructions



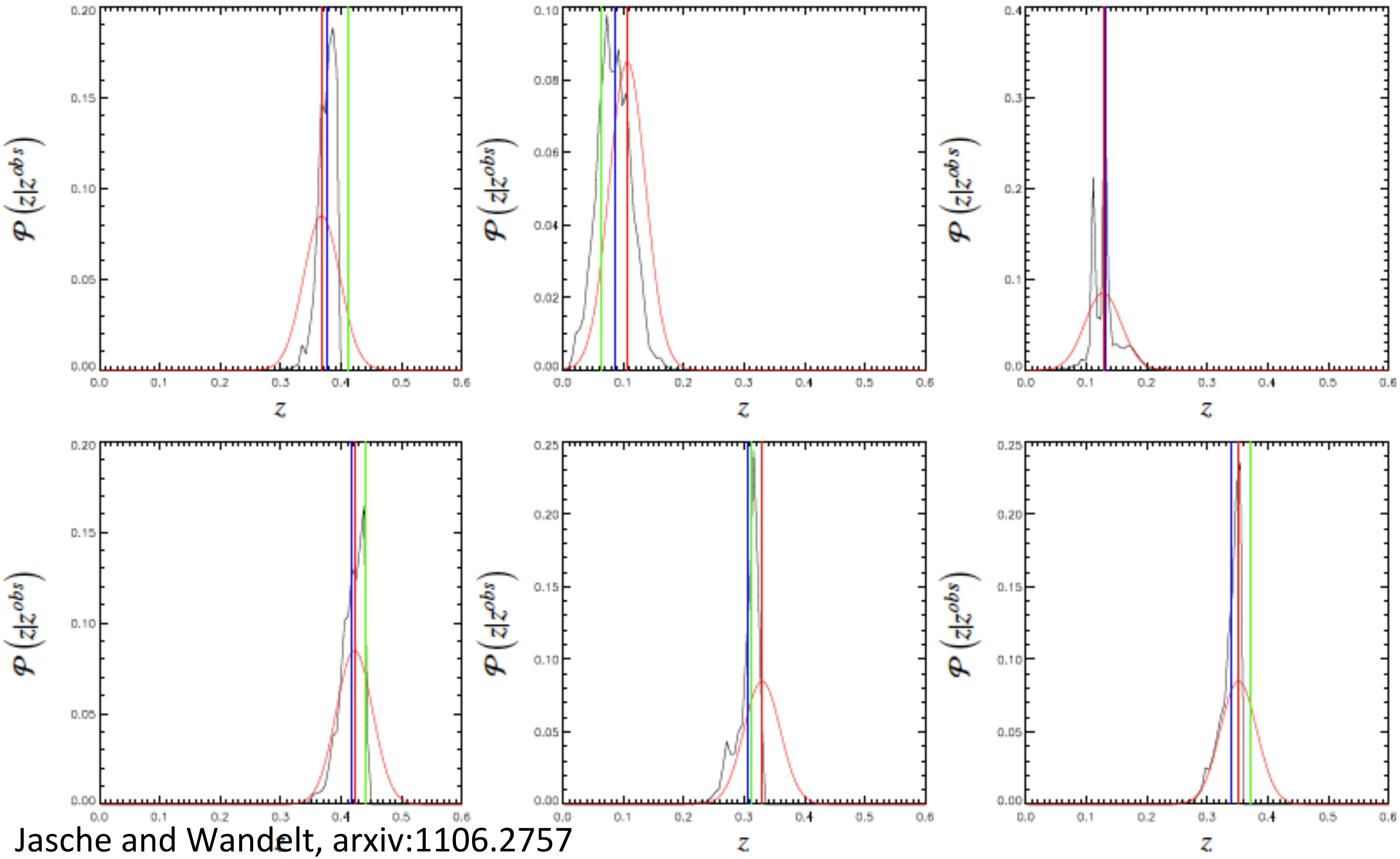
Jasche and Wandelt, arxiv:1106.2757

Inferred redshift locations are much better than photo-zs in high density regions



Jasche and Wandelt, arxiv:1106.2757

Radial location (redshift) posteriors compared to input from photo-z estimator



Jasche and Wandelt, arxiv:1106.2757

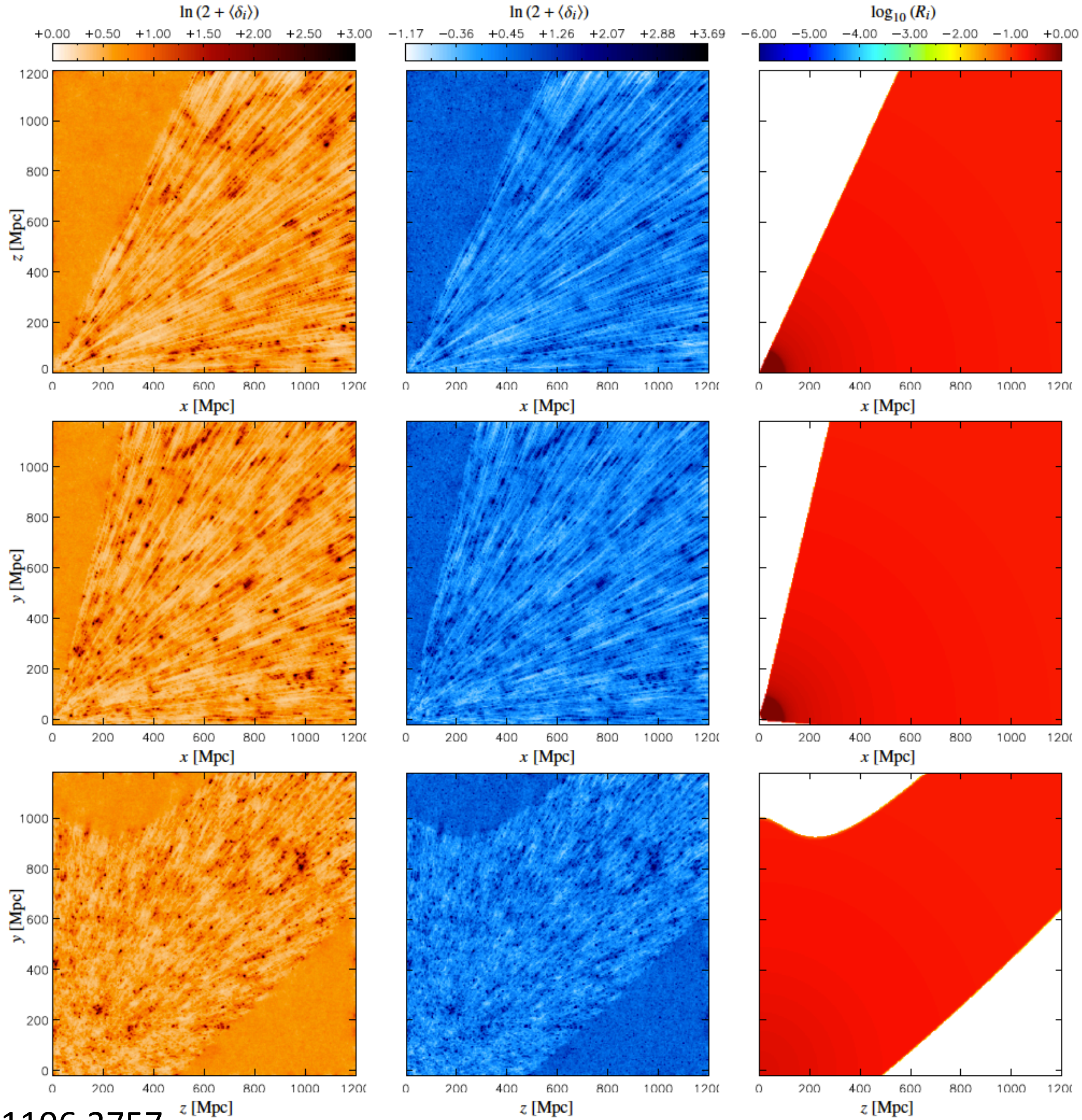
Radial location (redshift) posteriors compared to input from photo-z estimator

Could also extract conditional information: if galaxy a has redshift z,
what is the probability that galaxy b has redshift z'?

Jasche and Wandelt, arxiv:1106.2757

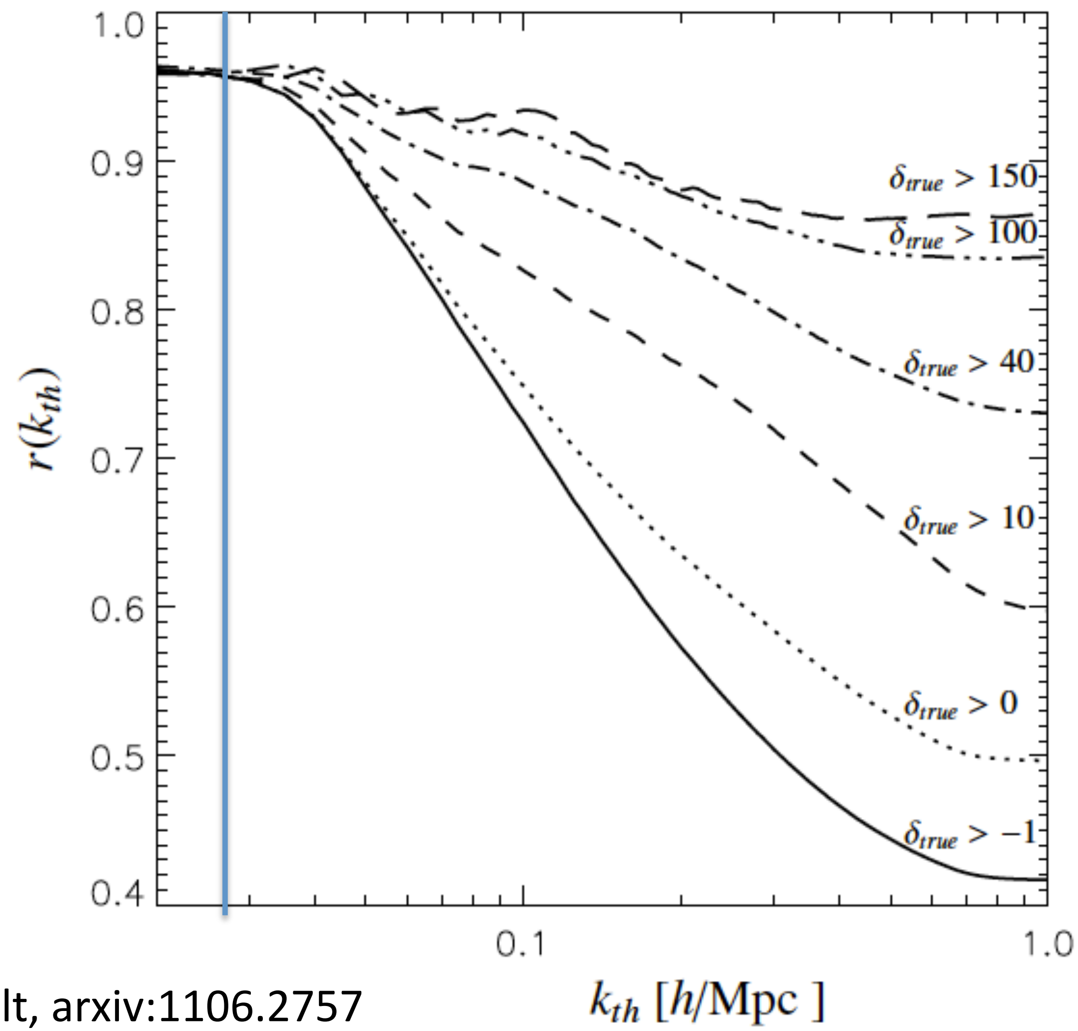
Full, sampled representation of the posterior

Posterior mean (column 1) and variance (column 2) of reconstructed density field + mask (column 3)



Jasche and Wandelt, arxiv:1106.2757

High density regions are super-resolved (cross-correlation between input and posterior mean density field for different density thresholds)



Jasche and Wandelt, arxiv:1106.2757

Better redshifts will give better reconstructions

- Our approach is completely independent of and complementary to the means by which the photometric redshift is derived.
- Information can be separately specified in terms of a different pdf for each galaxy

  $\Rightarrow$ can merge photometric and spectroscopic samples!
- Interesting case where a decisive gain is achieved by combining millions of "weak" measurements with physical prior information.

Jasche and Wandelt, arxiv:1106.2757

The ultimate dream for LSS data analysis:

"What if we could just solve for the non-linear evolution?"

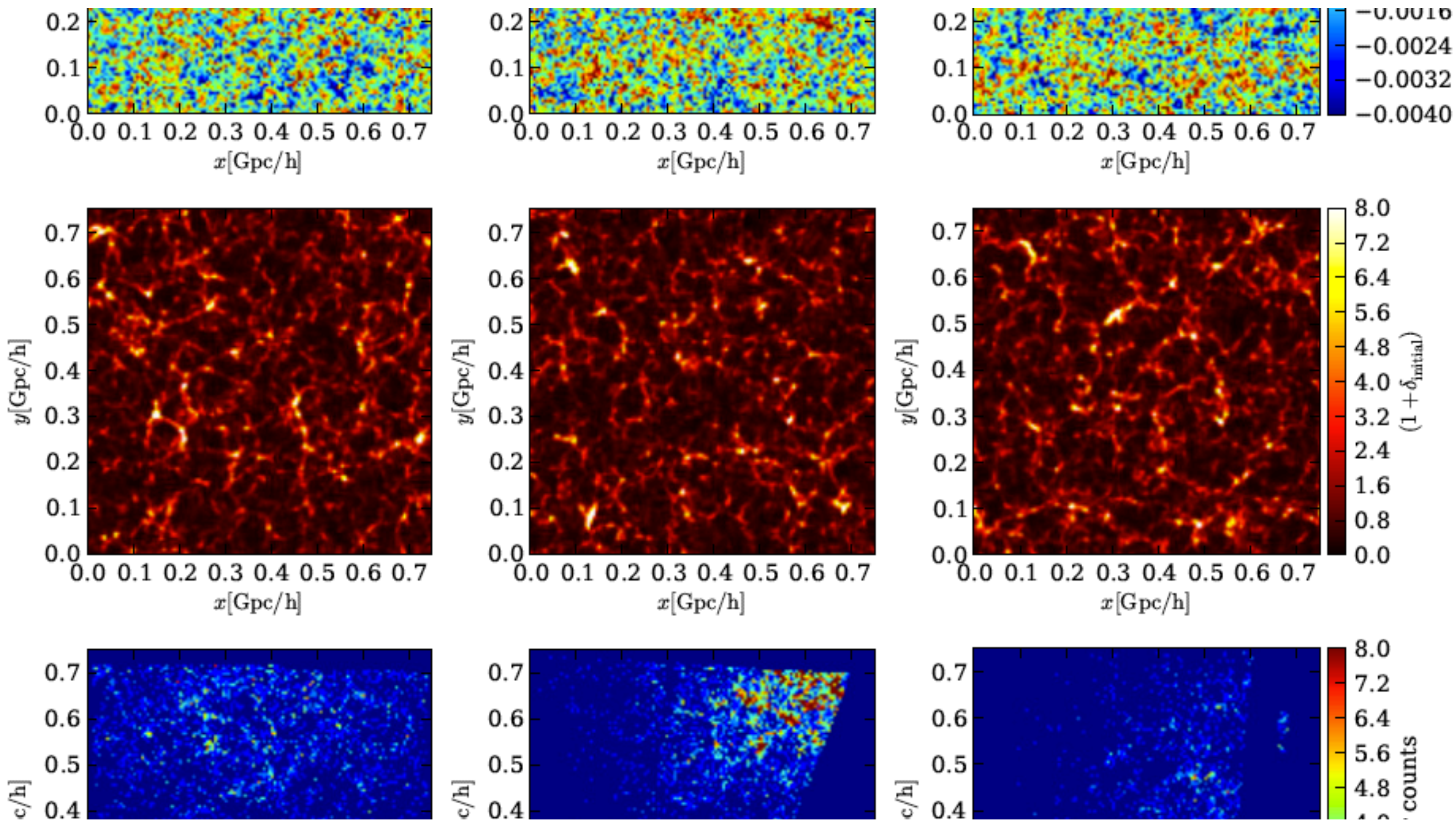
The ultimate Large Scale Structure inference

- The initial conditions are well-described as a (nearly) Gaussian random field
- What if we could evolve *all possible initial conditions* to the redshift where we observe galaxies?
- Then we could build a pdf over the space of initial conditions consistent with the observations

Bayesian large scale structure inference using physical dynamics

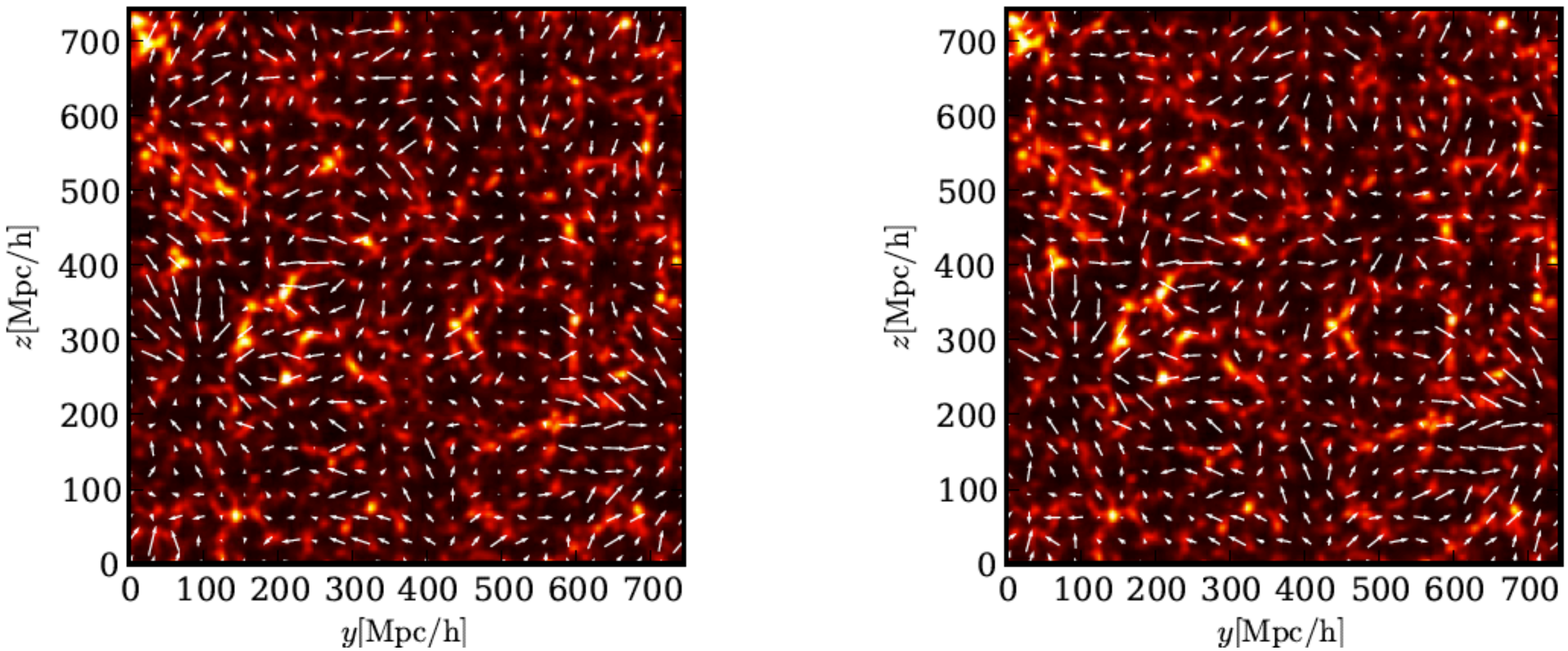
- We have taken the first baby steps towards this goal
- Jasche and Wandelt (arxiv:1203.3639) shows the first implementation of fitting non-linear evolution histories to a (simulated) galaxy survey.
- Gravity is implemented as 2^{nd} order Lagrangian perturbation theory

Initial condition reconstruction



Jasche & Wandelt arxiv:1203.3639

Dynamical reconstruction-velocity fields



Conclusions

- “Stacked Voids” are a new, purely geometrical dark energy observable. First estimates suggest tremendous additional potential for constraints on Dark Energy
- cosmicvoids.net
- Global analysis of survey data can add tremendous value to photo-z surveys
- New approach to connect the initial conditions to observations via dynamics
- Non-linear, principled cosmological inference with ~10^7 parameters is becoming feasible.

APPENDICES

Survey forecasts

Survey	Fraction of sky	Luminosity function	Limiting magnitude	$z_{\max}$	Number of galaxies
SDSS-DR7	24%	$\phi_* = 1.46\,10^{-2}\ h^3\mathrm{Mpc}^{-3}$ $M_* = -20.83$ $\alpha = -1.20$ (SDSS Collaboration & Blanton 2000)	$r = 17.77$	0.3	$1.7\,10^6$
SDSS-DR7 (LRG)	24%	$\phi_* = 2.63\,10^{-5}\ h^3\mathrm{Mpc}^{-3}$ $M_* = -19.42$ $\alpha = 3.90$ (Cool et al. 2008)	$r = 19.5$	0.45	10^5
BOSS	24%	same as the SDSS	$r = 20$	0.7	$1.5\ 10^6$
EUCLID	36%	$\phi_* = 1.16\,10^{-2}\ h^3\mathrm{Mpc}^{-3}$ $M_* = -23.39$ $\alpha = -1.09$ (Kochanek et al. 2001; Jones et al. 2006)	$H = 24$	1.5	$\sim 1.6\,10^8$

Method	Data	FoM
BAO	BOSS	86
Voids ($s_r = 2$)	SDSS+BOSS LRG	70
Voids ($s_r = 5$)		69
BAO+Voids ($s_r = 2$ or 5)	SDSS+BOSS	96
Voids ($s_r = 2$)	EUCLID	~11 530
Voids ($s_r = 5$)		~650
BAO		185
BAO+Voids ($s_r = 2$)		~11 600
BAO+Voids ($s_r = 5$)		~950

Lavaux & Wandelt 2011