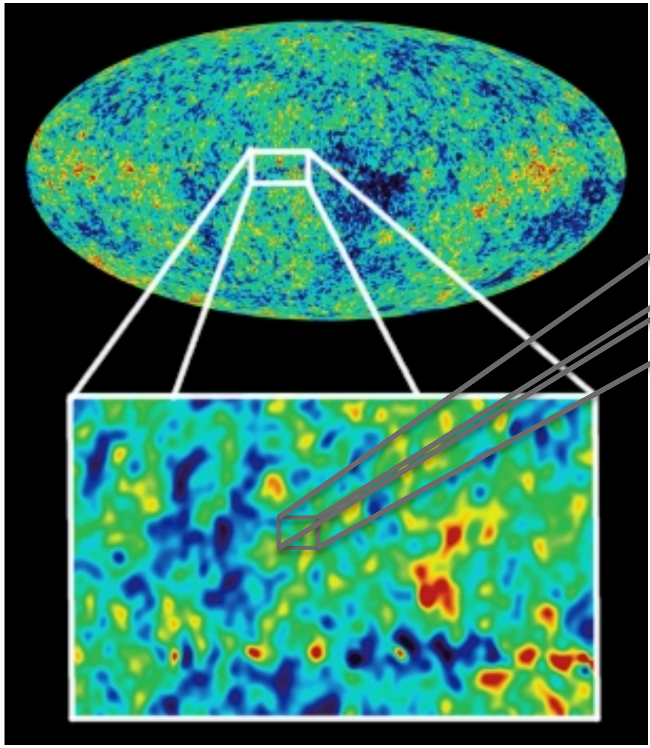
Dispersed information and non-linear inference in Bayesian cosmology

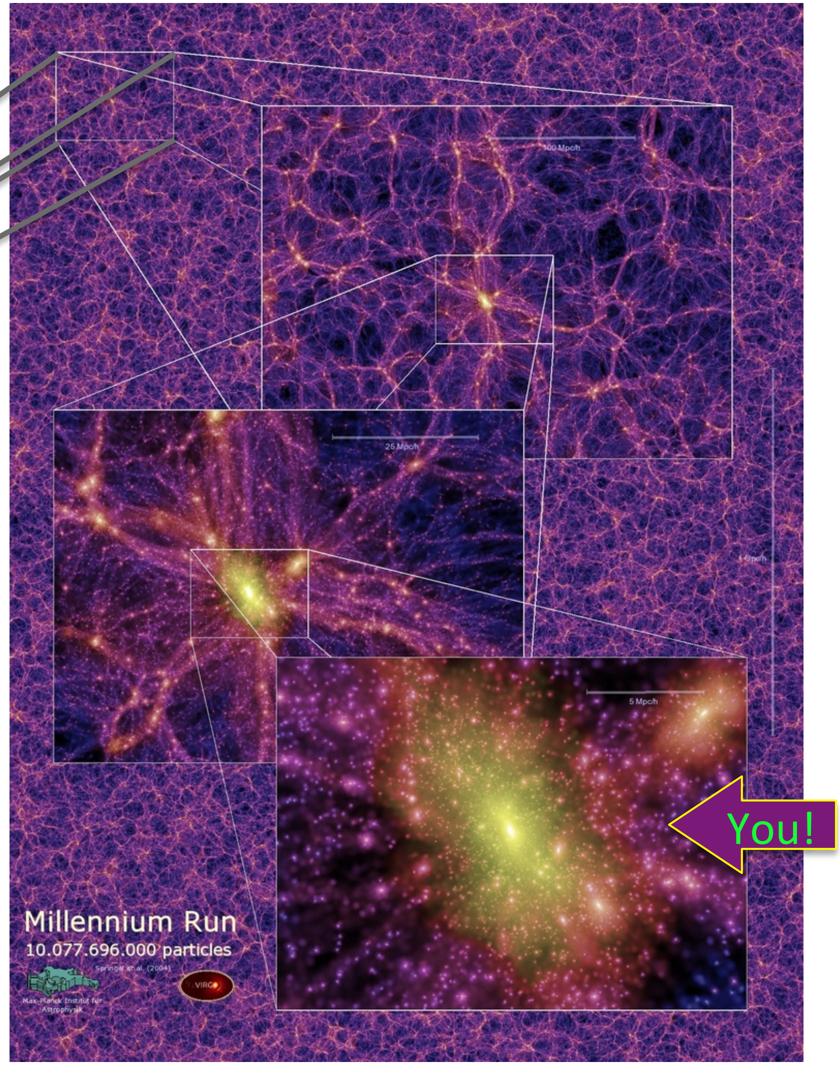
Benjamin Wandelt
Institut d'Astrophysique de Paris (IAP)
Université Pierre et Marie Curie (Paris 06), CNRS
Sorbonne Universités

wandelt@iap.fr

Cosmostatistics: the Big Picture



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



Dark matter distribution today (simulated)

Goals of cosmostatistics

- The cosmological agenda for the coming decade is to
  - Learn about the cosmic beginning
  - Understand the cosmic constituents, in particular Dark Matter and Dark Energy
  - Understand cosmological evolution from cosmic seeds to current observations
- Any given observable (e.g. CMB, galaxy survey) is informative (often weakly) about all of these goals in some way
- Decisive inference for each of these goals requires *joint inference* from several complementary data sets. This challenges the status quo.

Example

- Cosmic Constituents: Dark Energy
- Goal: infer D.E. Parameters
- Use
  - Two-point correlations P(k,z), BAO
  - Lensing
  - Halo abundances
  - Void properties
  - Primary CMB, ISW and secondaries
  - Supernovae
- **Joint analysis** provides powerful constraints and propagates the information between all parts of the analysis
- Joint analysis also avoids using data twice
  - P(k,z), BAO, halos, voids lensing all probe the same galaxies… traditional approaches do separate analyses and combine after the fact as if independent!

Issues in Cosmostatistics

- Awash in data, but fundamental limits to information:
  - On large scales: **causality**
  - On small scales: **non-linearity**

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

⇒Linear methods are OK on intermediate scales

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

Beyond the linear and the Gaussian

- Even for Gaussian fields non-Gaussian statistics arise for covariance estimation (power spectrum inference, parameter inference)
- 21st century precision cosmology deals with non-linear problems:
  - Gravitational non-linearity! Lensing, galaxy surveys
  - Primordial non-Gaussianity?
  - Non-linearities and non-Gaussianity can also arise in interesting ways when dealing with data imperfections and systematics

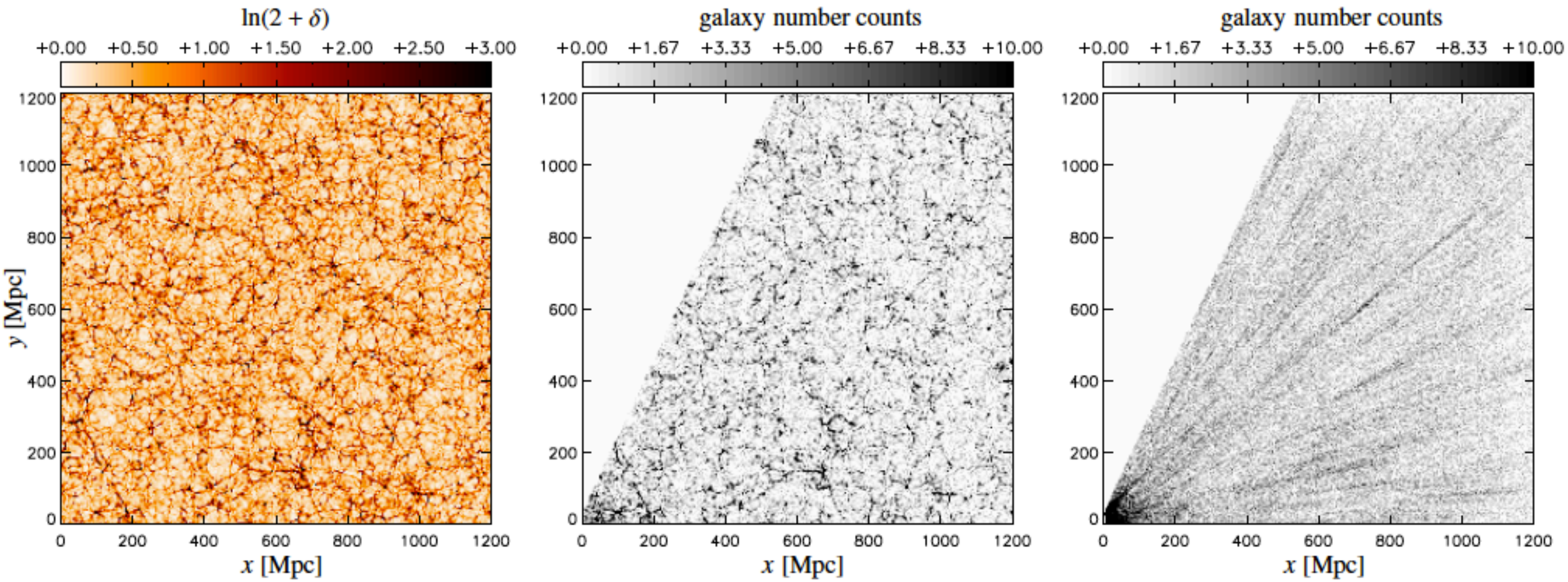
Examples of Bayesian solutions to non-linear problems in Cosmostatistics

- Global Bayesian inference from photometric redshift surveys
- Bayesian lens reconstruction from CMB data
- Precision cosmography with Cosmic Voids from spectroscopic surveys

The majority of ongoing and future surveys will resolve distance poorly (photometry vs spectroscopy)

Your massive survey here

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

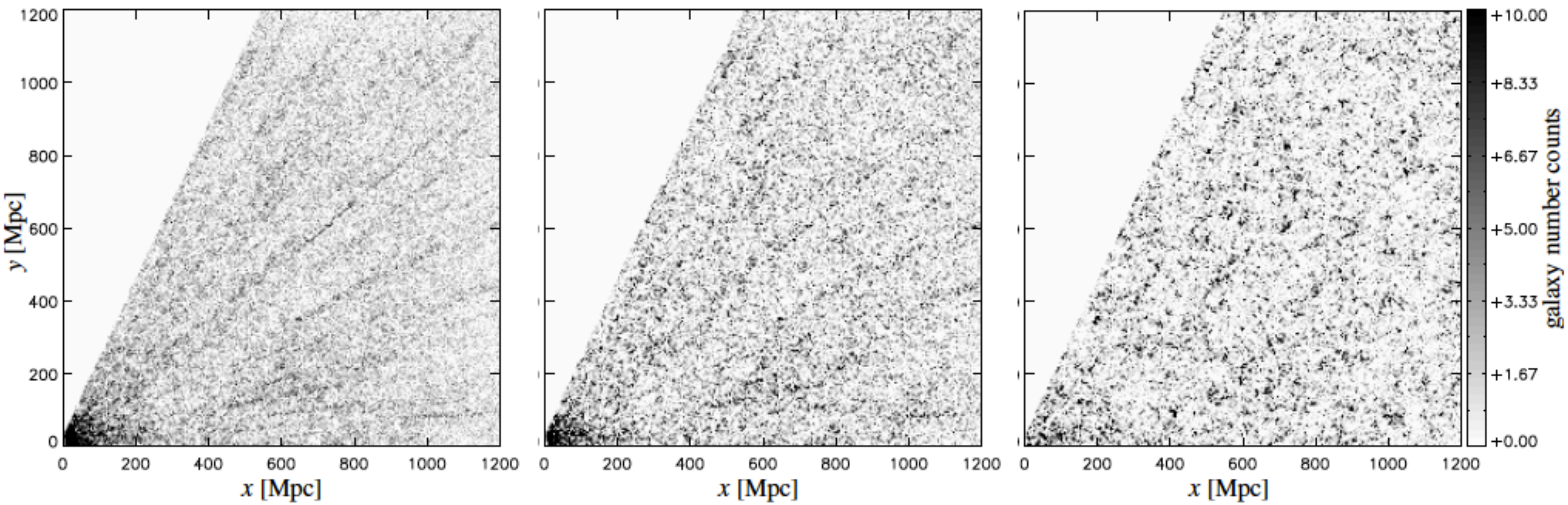
Assumptions:

- 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 (1.6x10^7 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

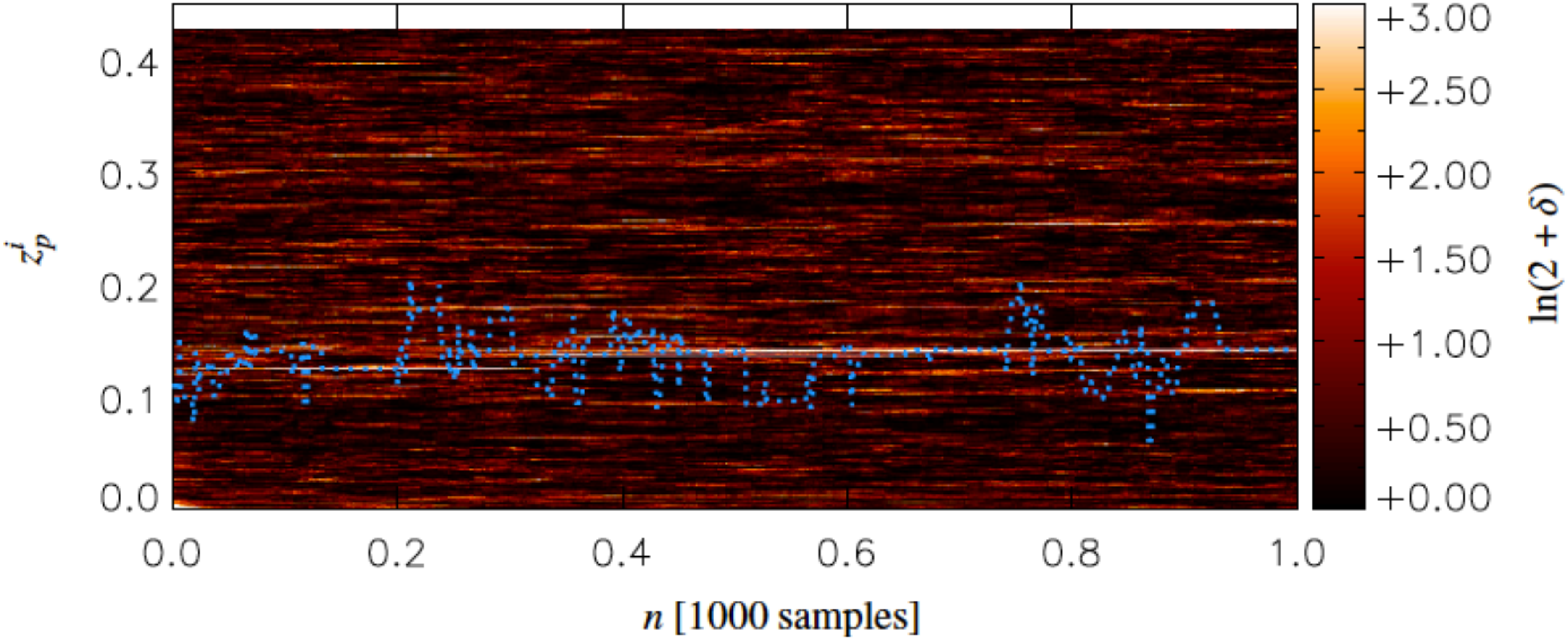
The reconstruction in action

Sample 1 Sample 5 Sample 100



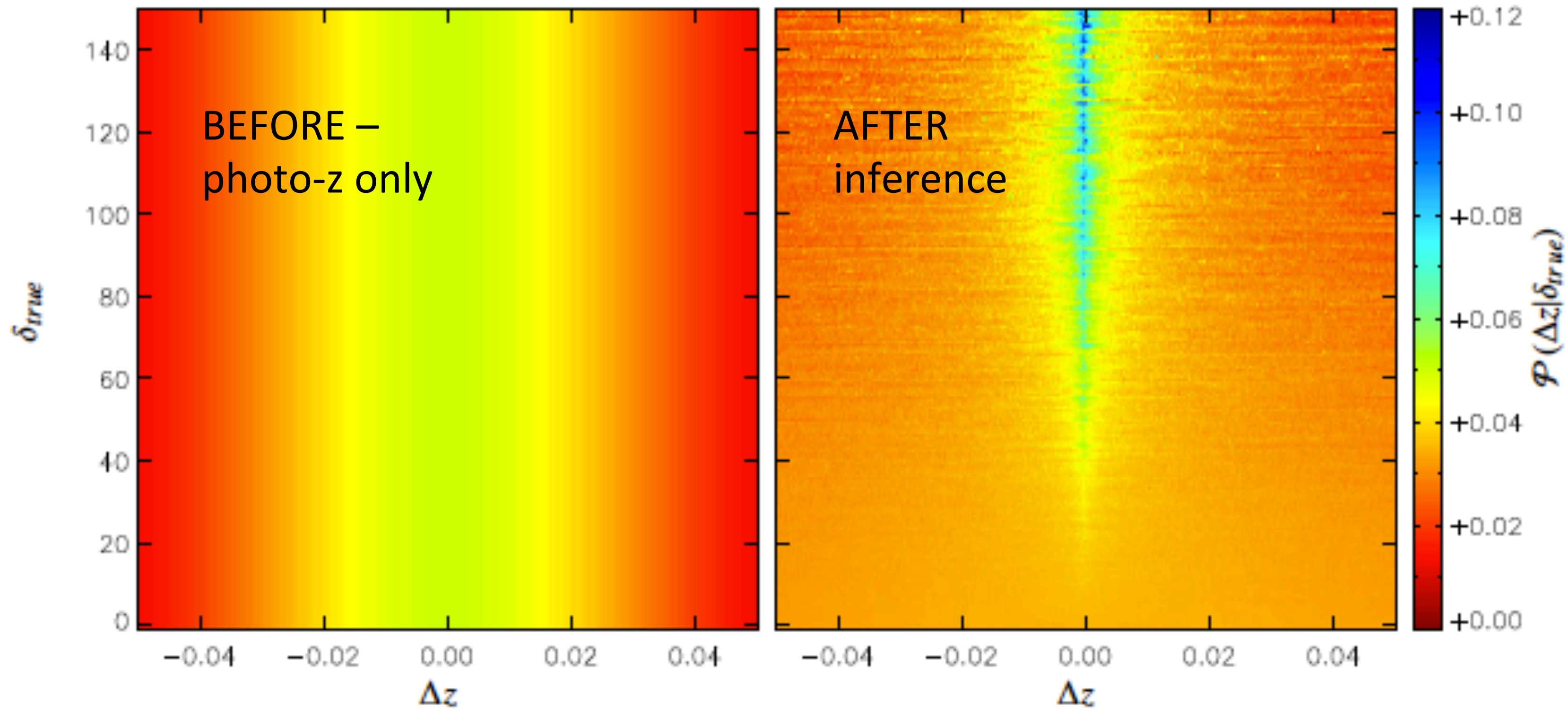
Jasche and Wandelt, arxiv:1106.2757

Galaxies random walk through the (dynamically updated) density reconstructions



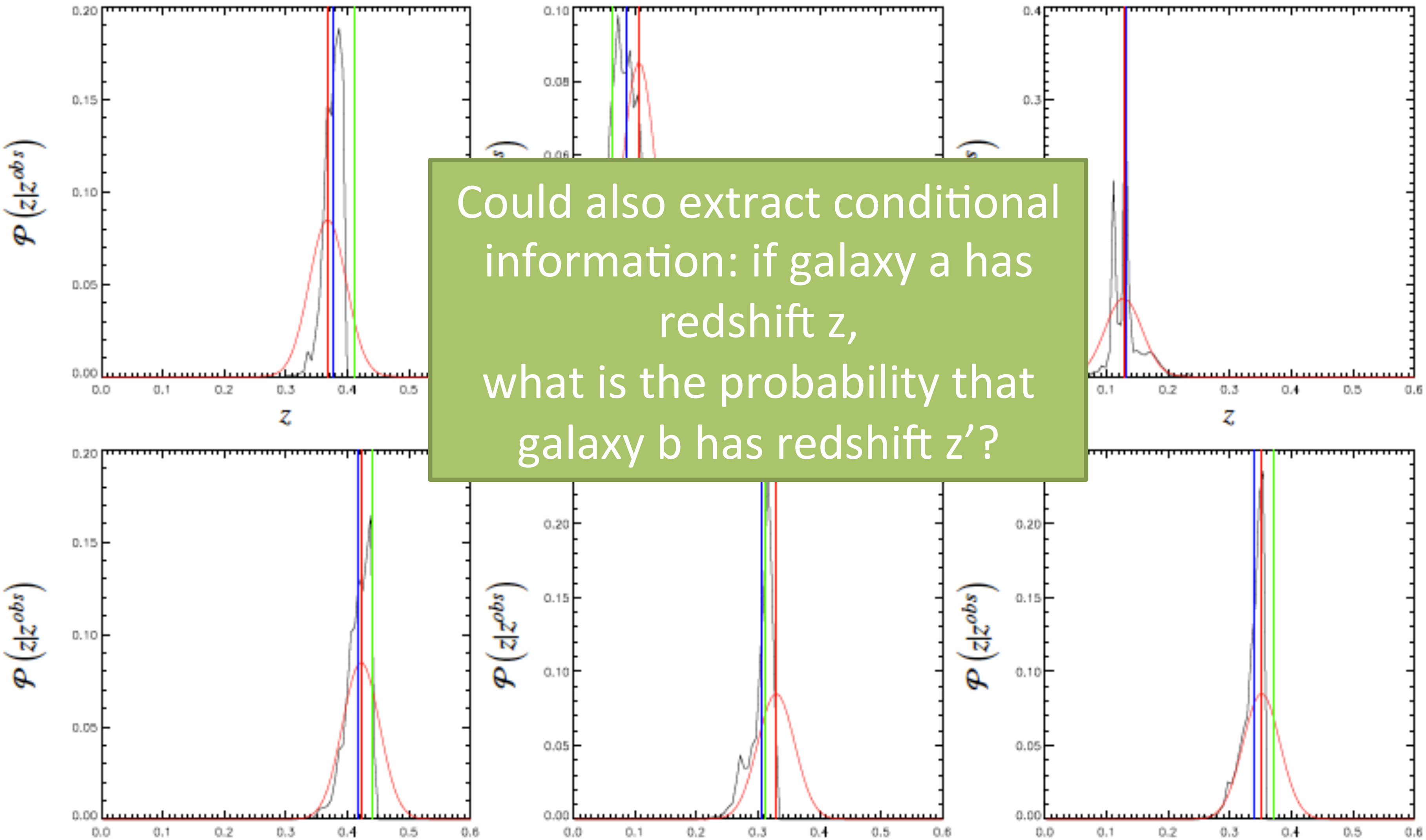
Jasche and Wandelt, arxiv:1106.2757

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

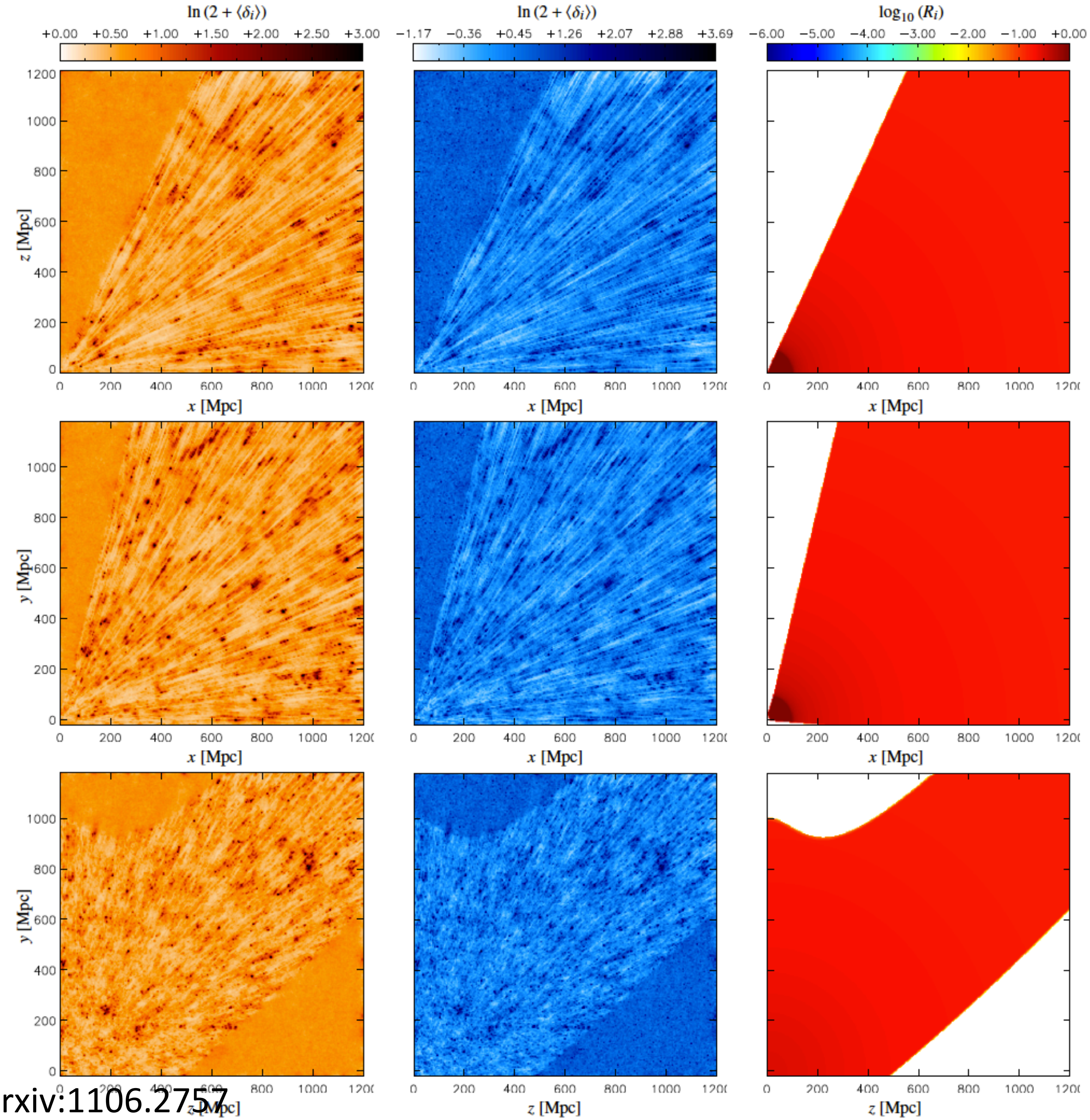


Could also extract conditional information: if galaxy a has redshift z,
what is the probability that galaxy b has redshift z'?

Jasche and Wandelt, arxiv:1106.2757

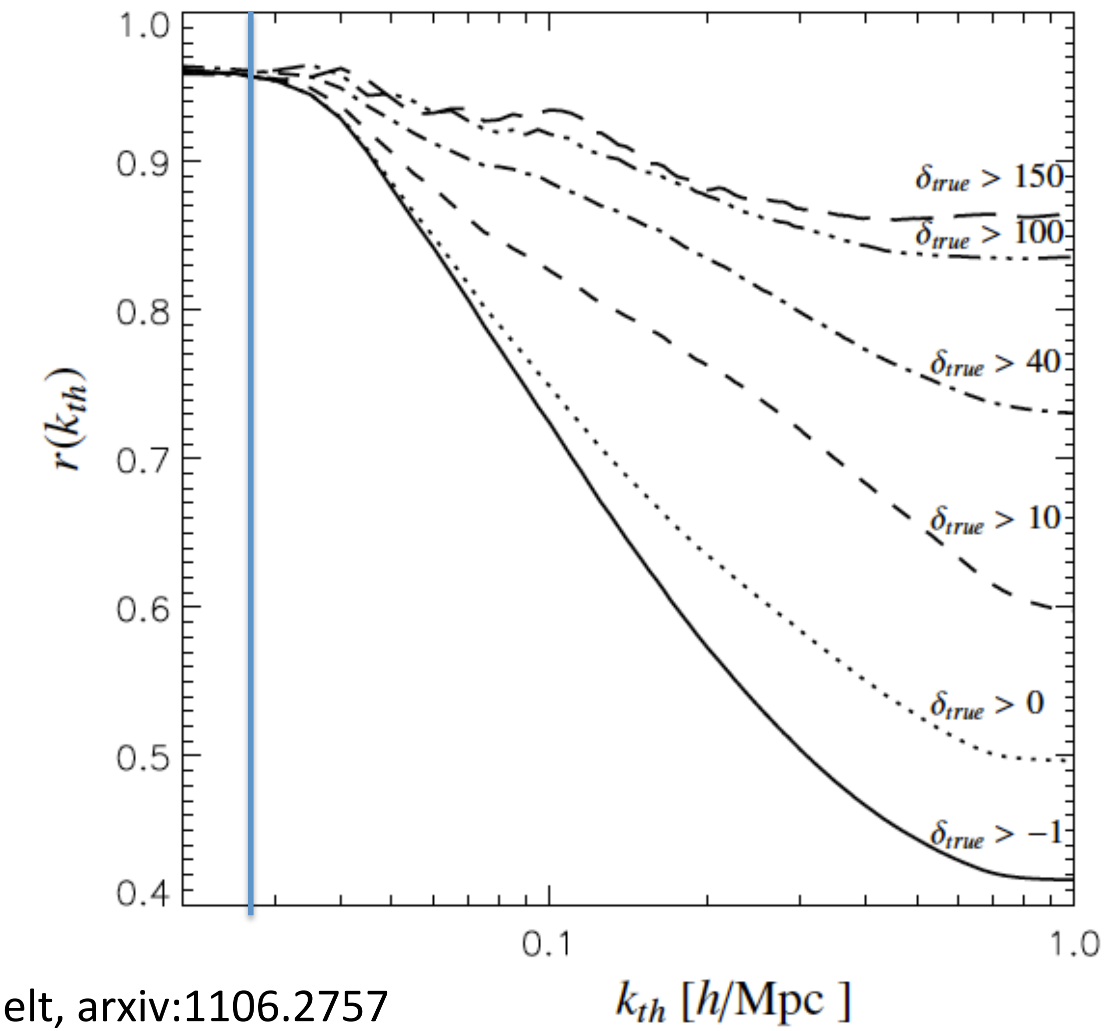
Full sampled representation of the posterior

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



Jasche and Wandelt, arxiv:1106.2757

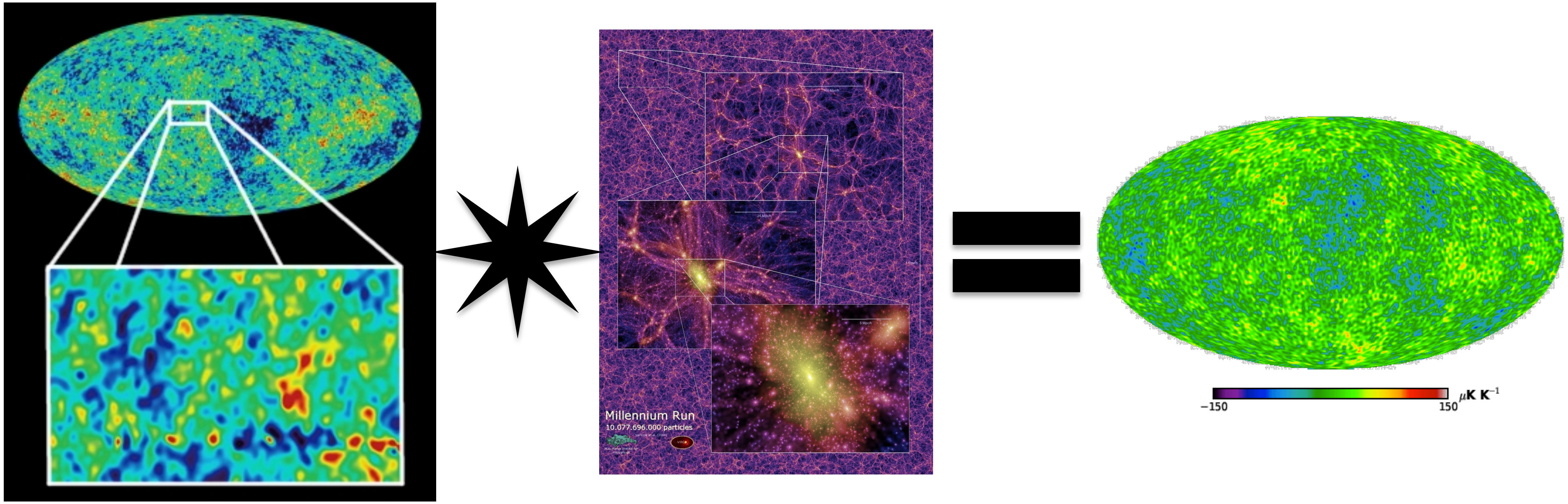
High density regions are super-resolved (cross-correlation between input and posterior mean density field for different density thresholds)



Jasche and Wandelt, arxiv:1106.2757

Comments on Bayesian large scale structure analysis

- Robust to prior misspecification: the entire simulation study used an full n-body sim while the model assumes a Poisson-Lognormal model
- Our approach is completely independent of and complementary to the means by which the photometric redshift is derived.
- Each tracer object can have its own photo-z PDF
- A decisive gain is achieved by combining a well-motivated physical prior with data where information is dispersed over millions of “weak” measurements



BAYESIAN LENS RECONSTRUCTION FROM CMB DATA

Lavaux&Wandelt, in prep.

Why Bayesian lens reconstruction?

- Trustworthy error propagation
- “Principled” rather than “ad hoc” analysis
- Optimal use of information-the standard quadratic estimator is known to be lossy
- Interesting by-product: posterior statistics of unlensed, primary CMB
- Spin-off: a fast and highly accuracy interpolation method on the sphere.

Basics and first attempt

- Lensing model $d = L(\phi)s + n$
- Resulting posterior

$$P(s, d|S, N, \phi) \propto \exp\left(-\frac{1}{2}\mathcal{L}(s, d, S, N, \phi)\right)$$

$$\mathcal{L}(s, d, S, N, \phi) = \alpha s^\dagger T S^{-1} T^{-1} s + (d - L(\phi)s)^\dagger N^{-1}(d - L(\phi)s)$$

- Could expand to linear term and Gibbs sample. Tried this (with Antony Lewis) about 8 years ago
- Data strongly correlates primary CMB and lensing potential. Gibbs sampling moves only infinitesimally.
- **Fail**.

Why not solve the exact problem?

- Exact, marginal posterior is

$$P(\phi|S,N,d,C^{\phi}) = \frac{1}{\mathcal{N}'(S,N,d,C^{\phi})} P(d|S,N,\phi)P(\phi|C^{\phi})$$

$$P(d|S,N,\phi) =$$

$$\left(\det 2\pi \tilde{S}(S,N,\phi)\right)^{-1/2} \exp\left(-\frac{1}{2} d^{\dagger} \tilde{S}^{-1}(S,N,\phi) d\right)$$

$$\tilde{S}(S,N,\phi) = N + \alpha L(\phi) T S T^{-1} L^{\dagger}(\phi)$$

- N_{pix}^3 operations to evaluate normalization
- Impossible to evaluate. No Metropolis-Hastings, no Hamiltonian Sampling, no importance sampling etc.
- **Fail.**

Exchange sampling (Murray 2007)

- … solves the problem of not being able to evaluate the normalization for a posterior pdf of the form

$$p(\theta)\,\frac{f(y;\theta)}{\mathcal{Z}(\theta)}$$

- Augment this problem with another parameter set ϑ’ and another (fake) data set d’. New joint posterior is

$$p(y,\theta,x,\theta') = p(\theta)\,\frac{f(y;\theta)}{\mathcal{Z}(\theta)}\,q(\theta'\leftarrow\theta;y)\,\frac{f(x;\theta')}{\mathcal{Z}(\theta')}$$

- Marginalization gives the old posterior
- Sample by alternating two transitions:
  - Sample (ϑ’,d’|ϑ,d) – this is a Gibbs sample.
  - Propose exchange move ϑ’ <–> ϑ. Accept with acceptance ratio

$$\frac{p(\theta')\,q(\theta\leftarrow\theta';y)\,f(y;\theta')\,f(x;\theta)}{p(\theta)\,q(\theta'\leftarrow\theta;y)\,f(y;\theta)\,f(x;\theta')}$$

Note cancellation of Z(ϑ)!

Details

- L(ϕ) implemented using FLINTS (Lavaux &Wandelt 2010) for fast lensing and transpose lensing
- Initialize using quadratic estimator

$$\phi'_{\ell,m} = \alpha_\ell \phi_{\ell,m} + \gamma_\ell(\delta\phi_{\ell,m} + \epsilon_{\ell,m}) \qquad \alpha_\ell^2 + r_\ell^2 \gamma_\ell^2 = 1$$

  – Drift-diffusion proposal based on re-weighted quadratic estimator

$$\langle|\phi_{\ell,m}|^2\rangle = C_\ell^\phi \quad \langle|\delta\phi_{\ell,m}|^2\rangle = r_\ell C_\ell^\phi \qquad r_\ell = \min\left(1, \frac{E_\ell}{C_\ell^\phi}\right)$$

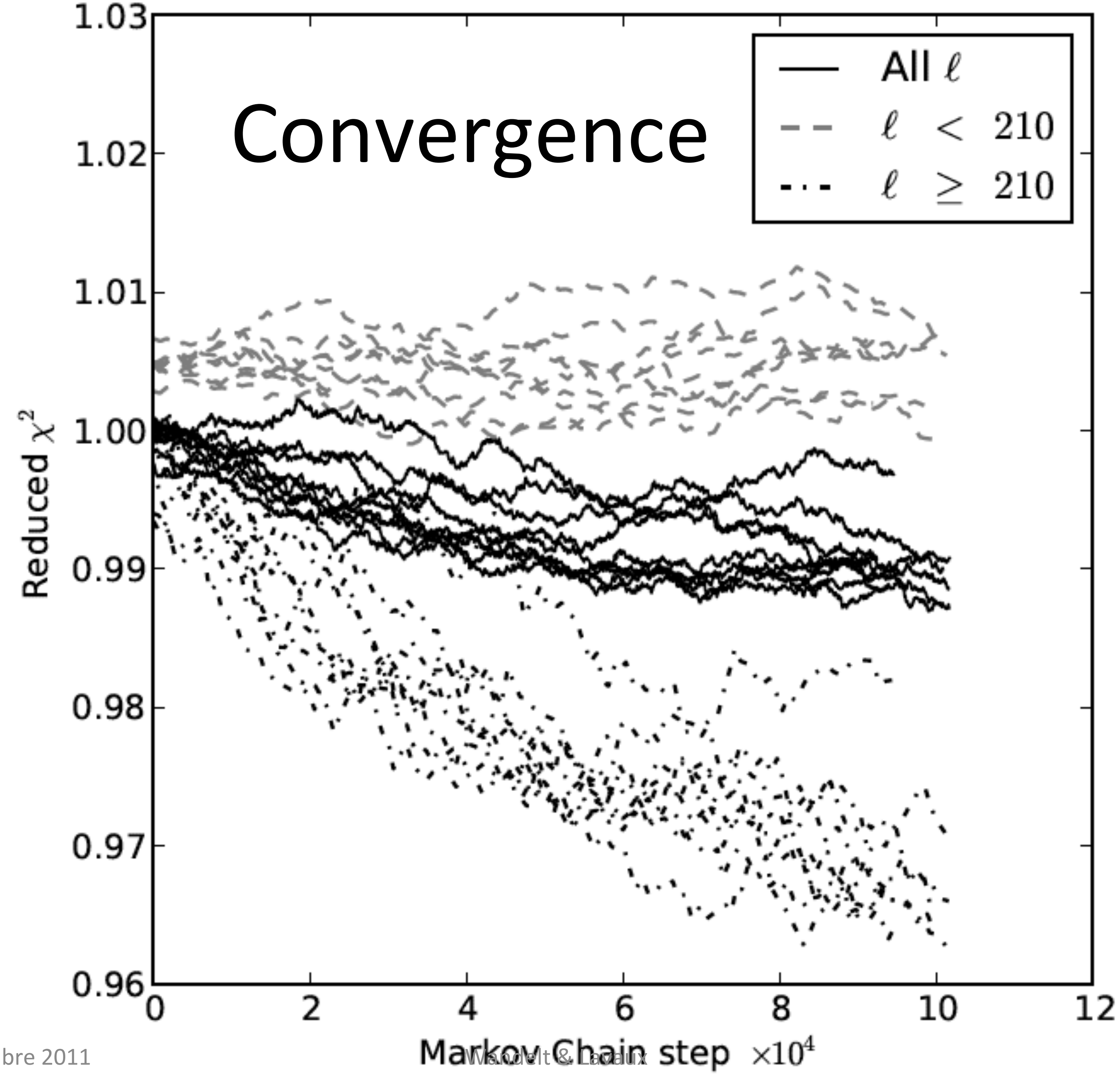
$$Q(\phi',\phi) = \frac{1}{2}\sum_{\ell=2}^{+\ell_{\max,\phi}} \frac{1}{\gamma_\ell^2 E_\ell} \sum_{m=-\ell}^{+\ell} \left|\phi'_{\ell,m} - \alpha_\ell \phi_{\ell,m}\right|^2$$

Drift term εis quadratic estimator on lensing subtracted data

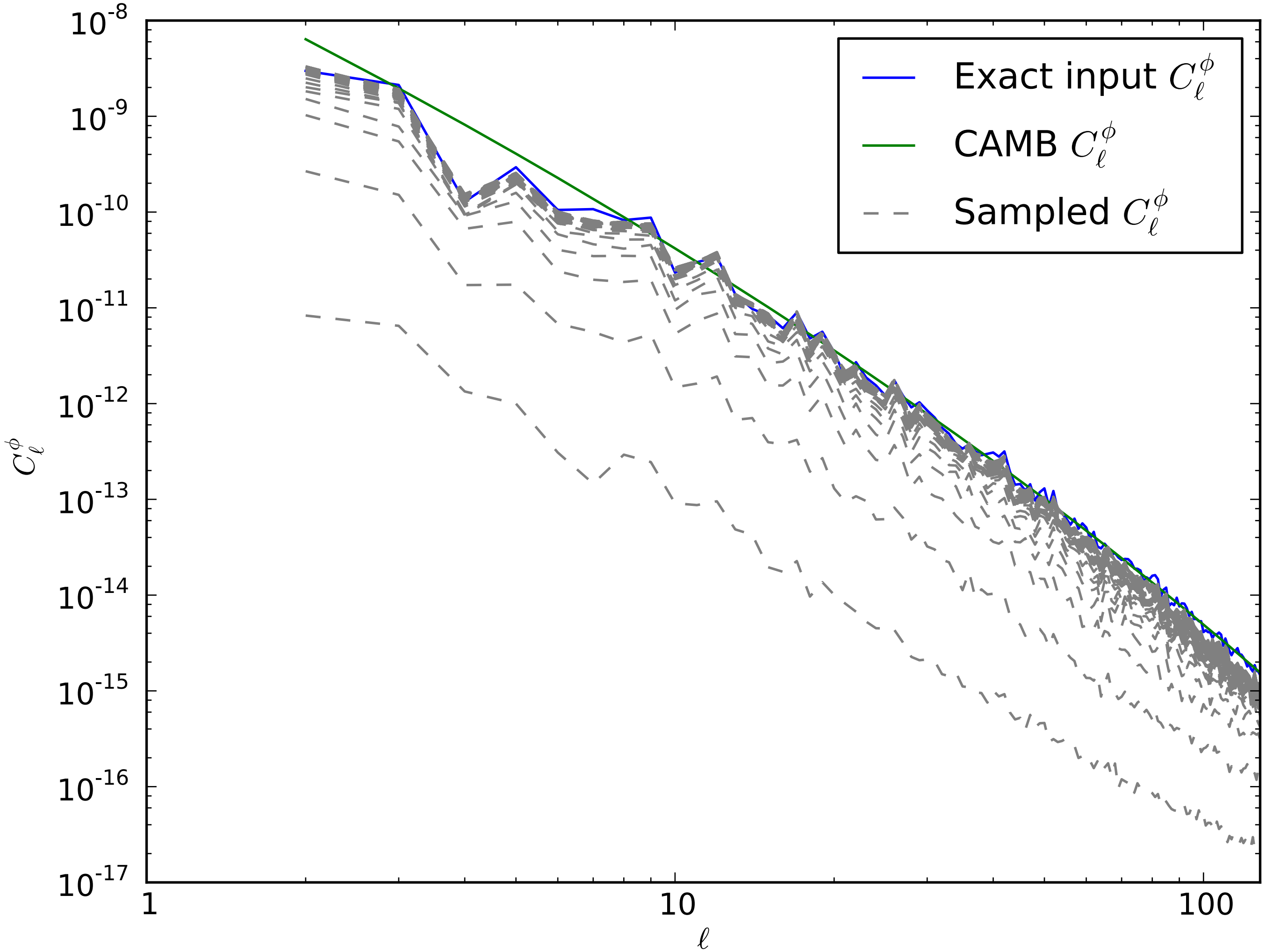
$$\tilde{d} = s_0 + (d - L(\phi)s_0)$$

nside=512, lmax=256

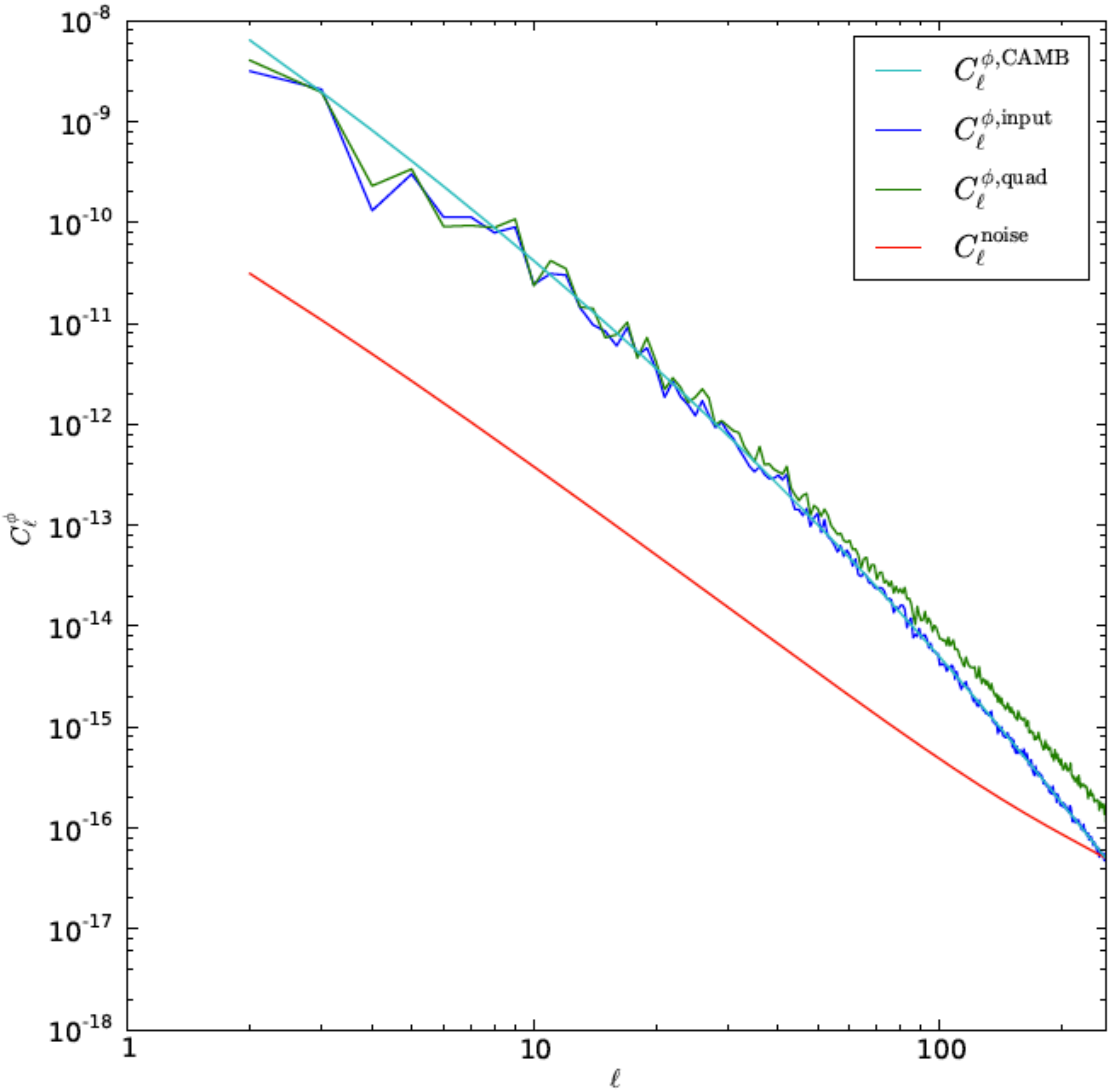
Convergence



Burn-in

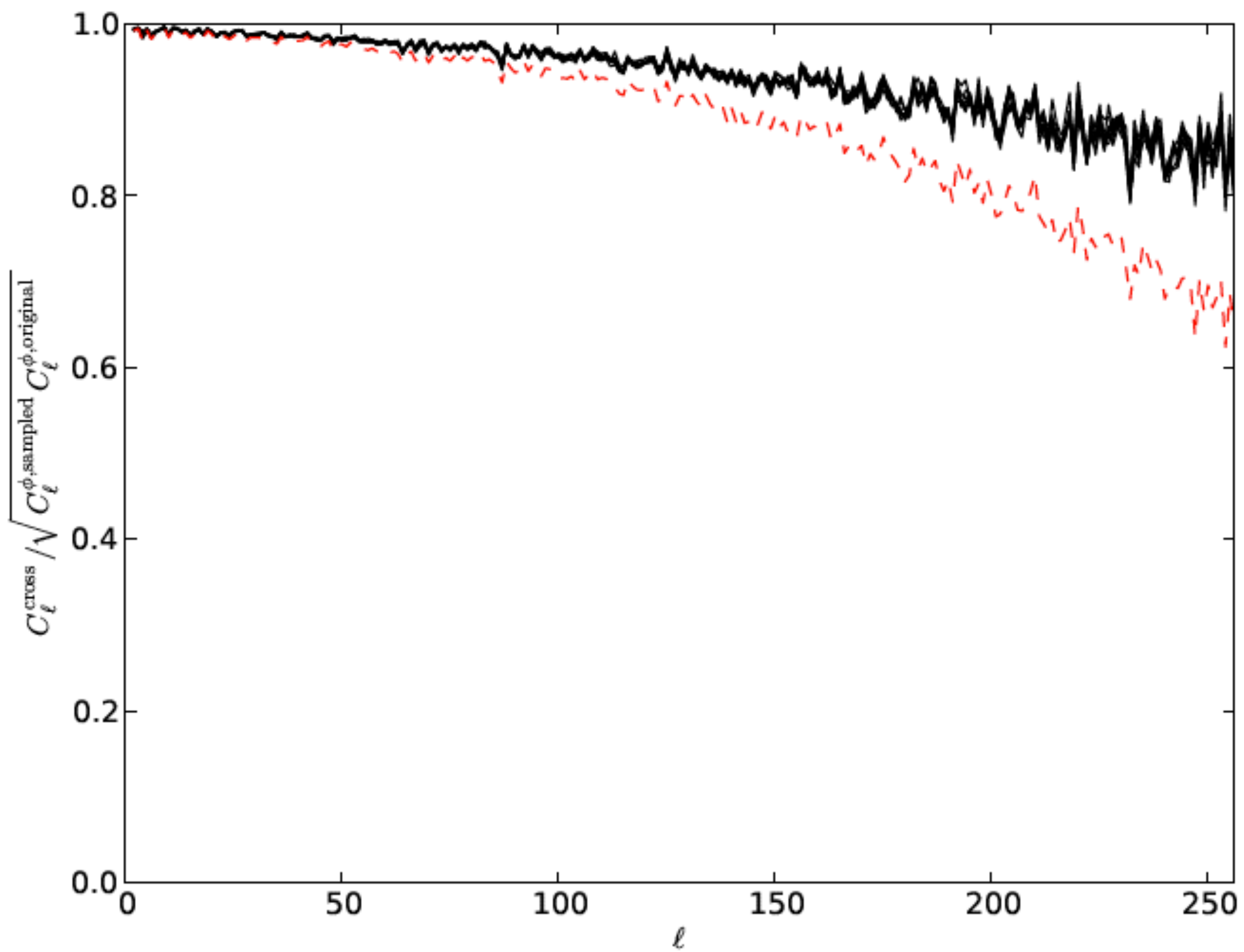


Exact input C_ℓ^ϕ
CAMB C_ℓ^ϕ
Sampled C_ℓ^ϕ
10^{-8}
10^{-9}
10^{-10}
10^{-11}
10^{-12}
10^{-13}
10^{-14}
10^{-15}
10^{-16}
10^{-17}
C_ℓ^ϕ
1
10
100
ℓ

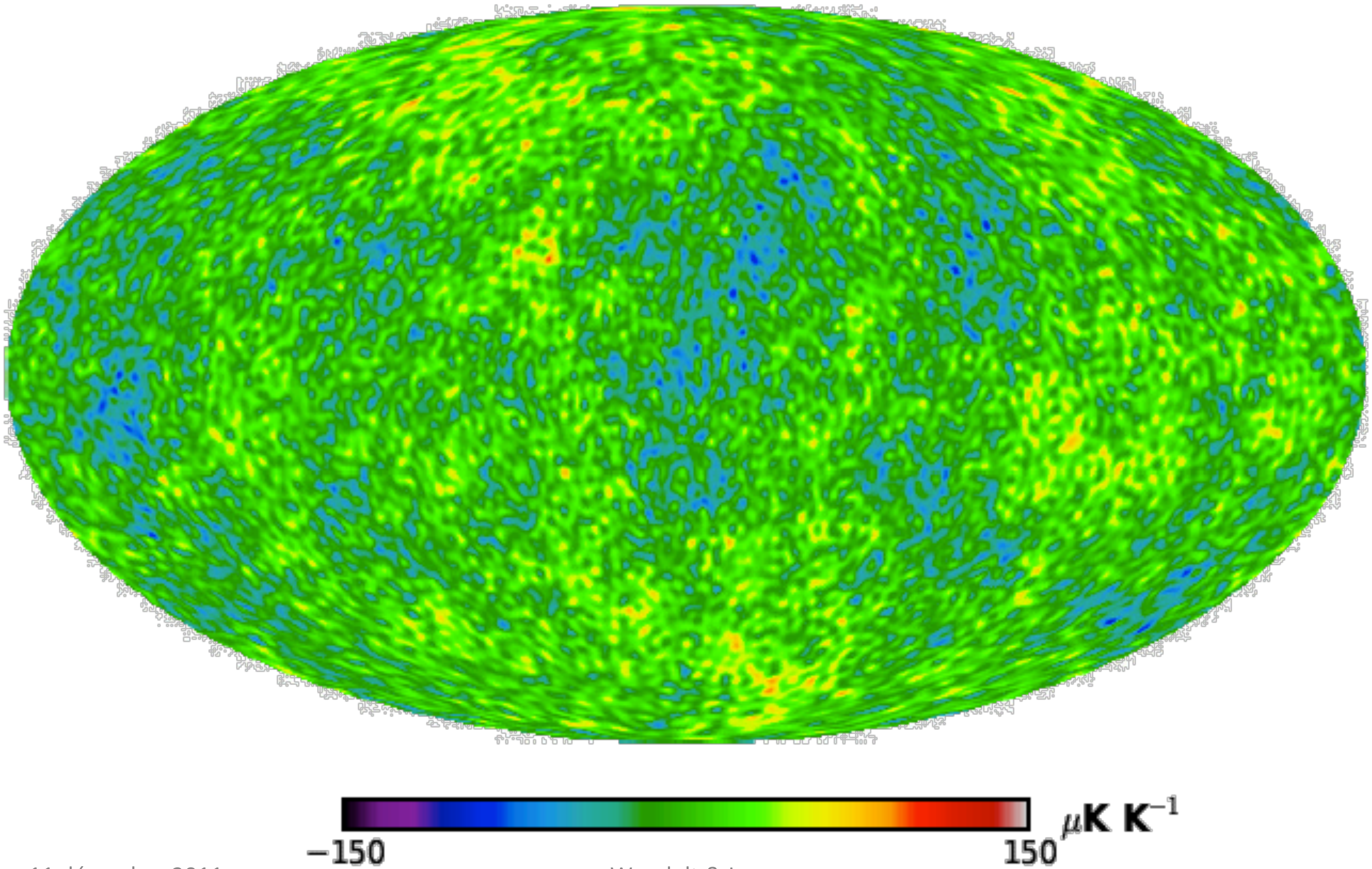


$C_\ell^{\phi,\mathrm{CAMB}}$
$C_\ell^{\phi,\mathrm{input}}$
$C_\ell^{\phi,\mathrm{quad}}$
C_ℓ^{noise}
10^{-8}
10^{-9}
10^{-10}
10^{-11}
10^{-12}
10^{-13}
10^{-14}
10^{-15}
10^{-16}
10^{-17}
10^{-18}
C_ℓ^ϕ
1
10
100
ℓ

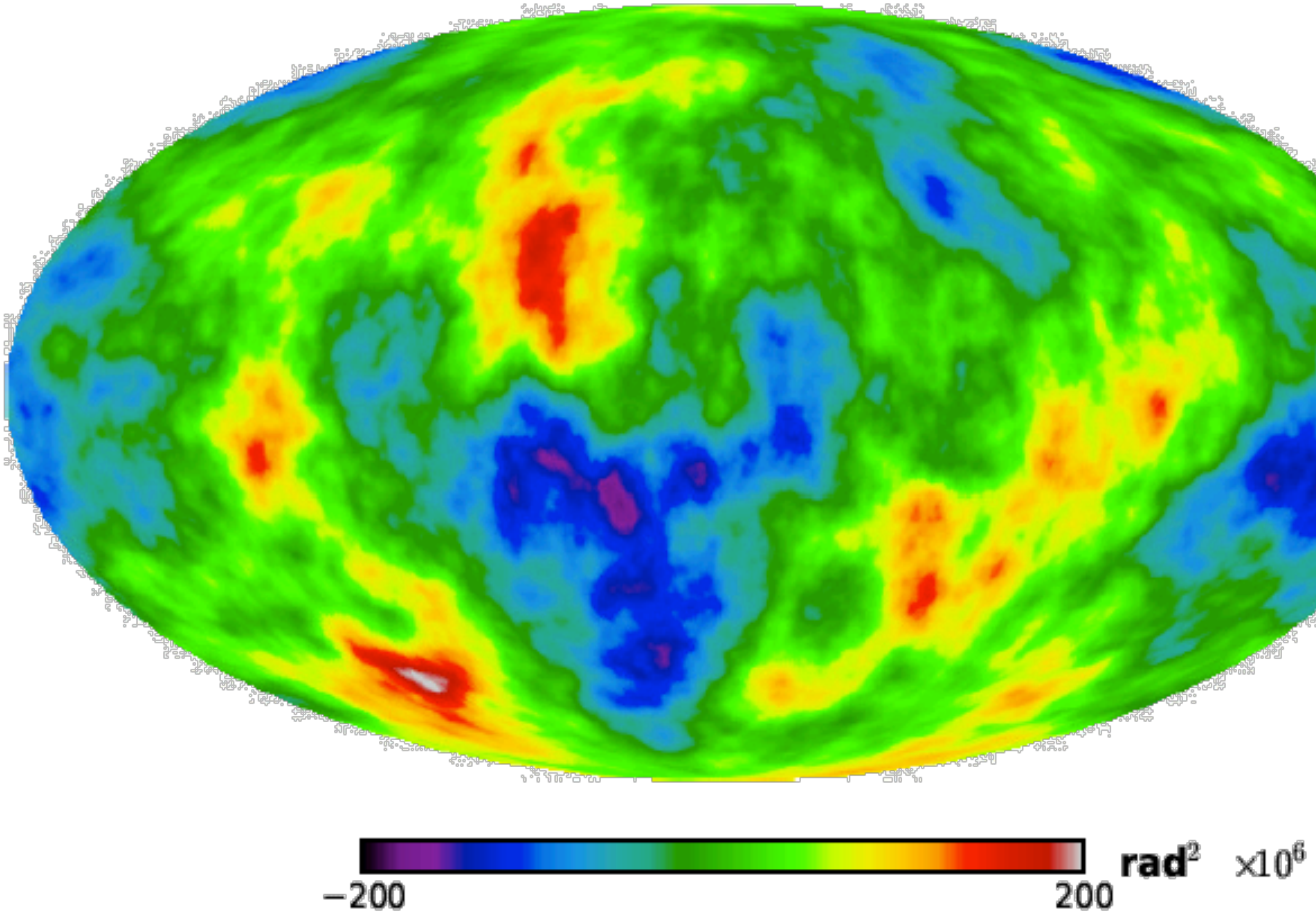
X-Correlation between input and reconstructed lensing potential



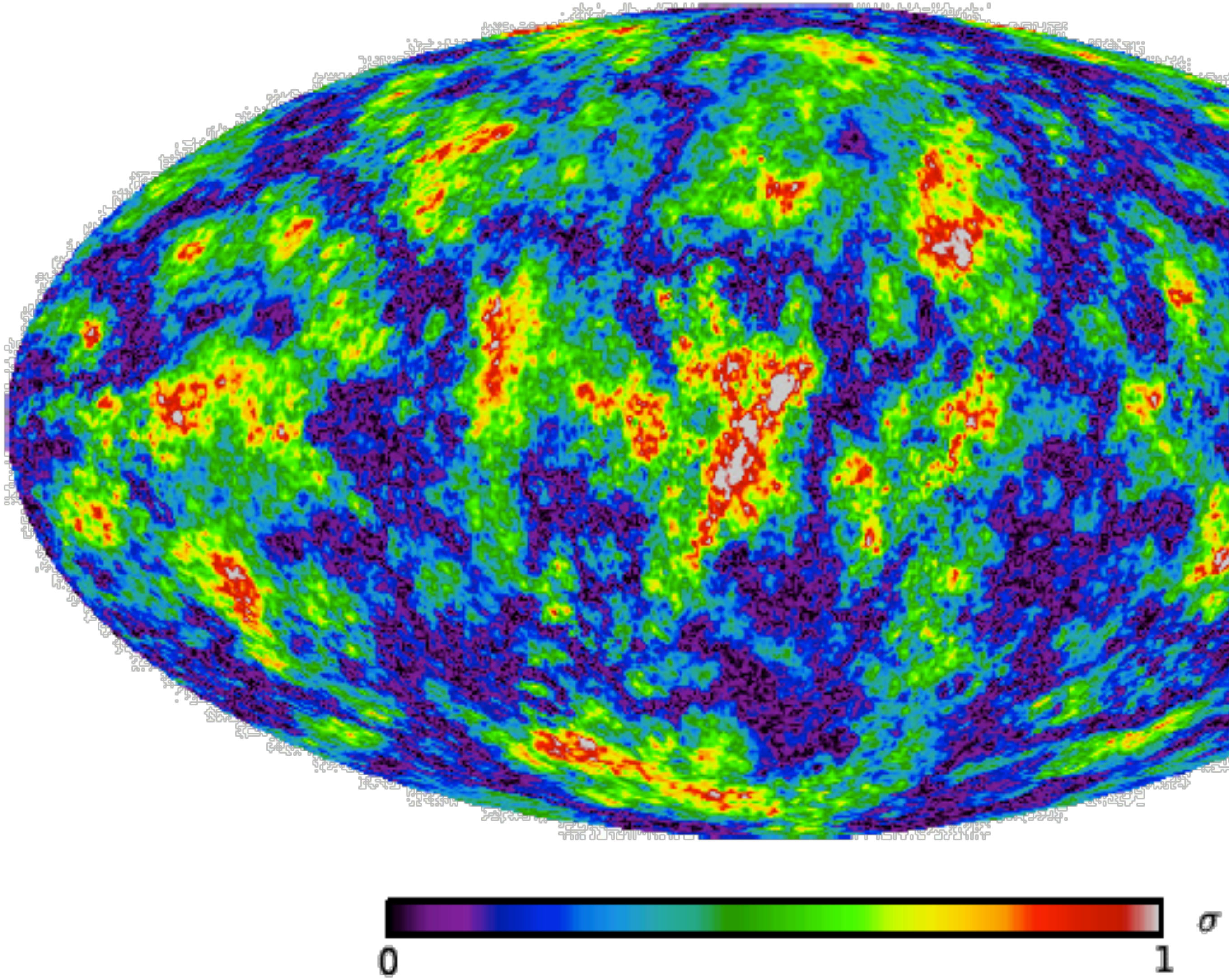
Lensed CMB input signal



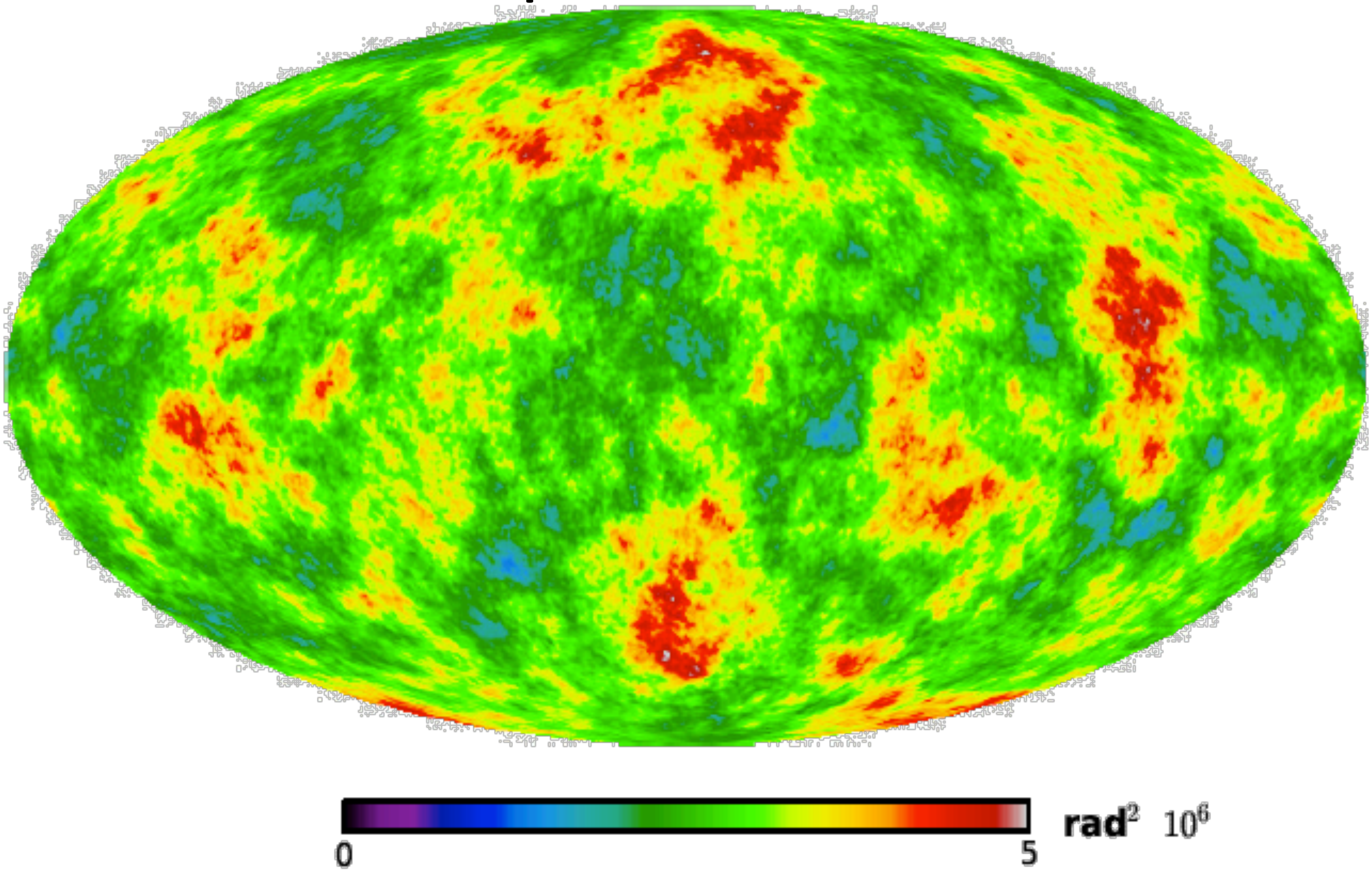
Input lensing potential



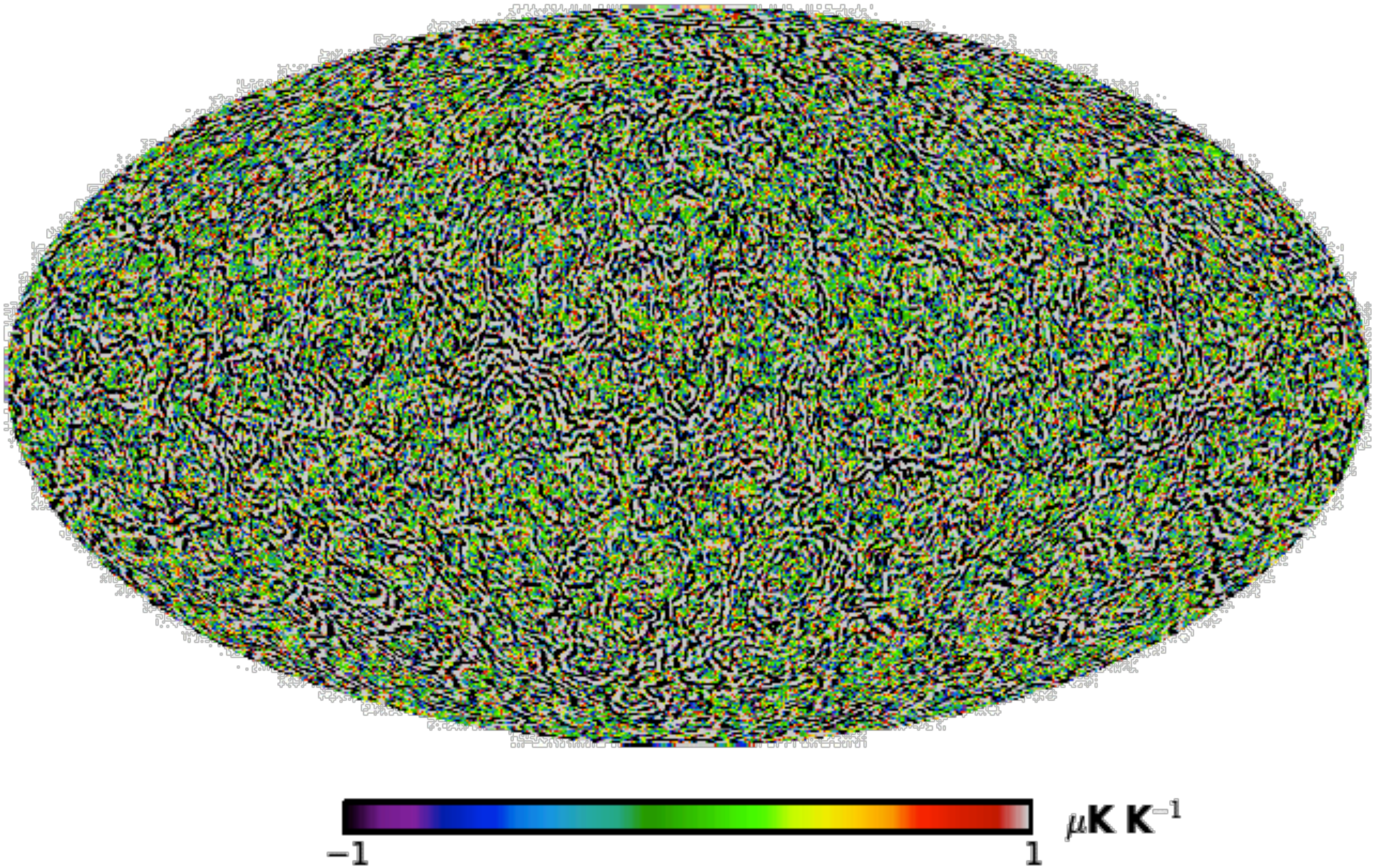
Difference between input and posterior mean lensing potential



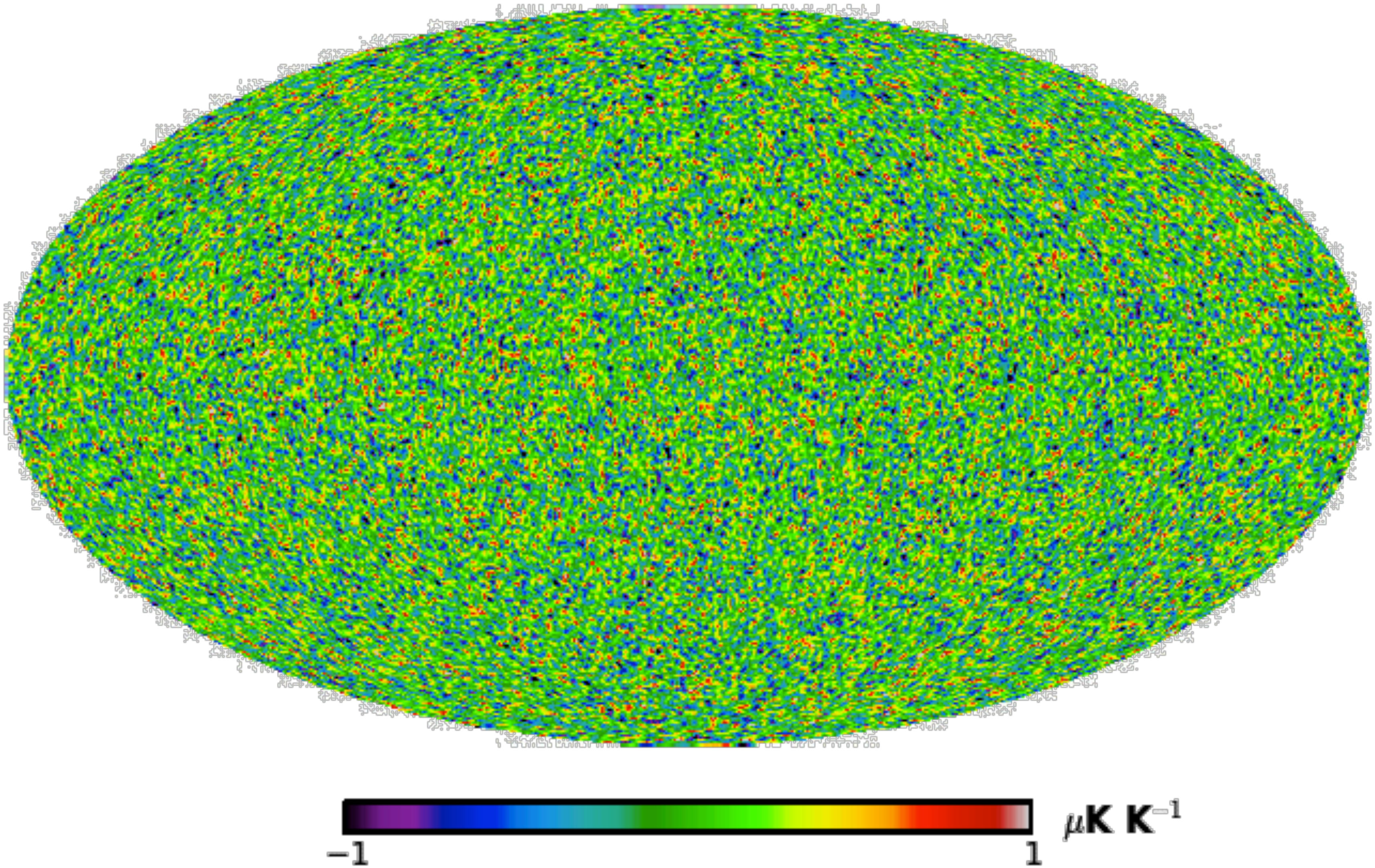
Posterior std dev. of lensing potential

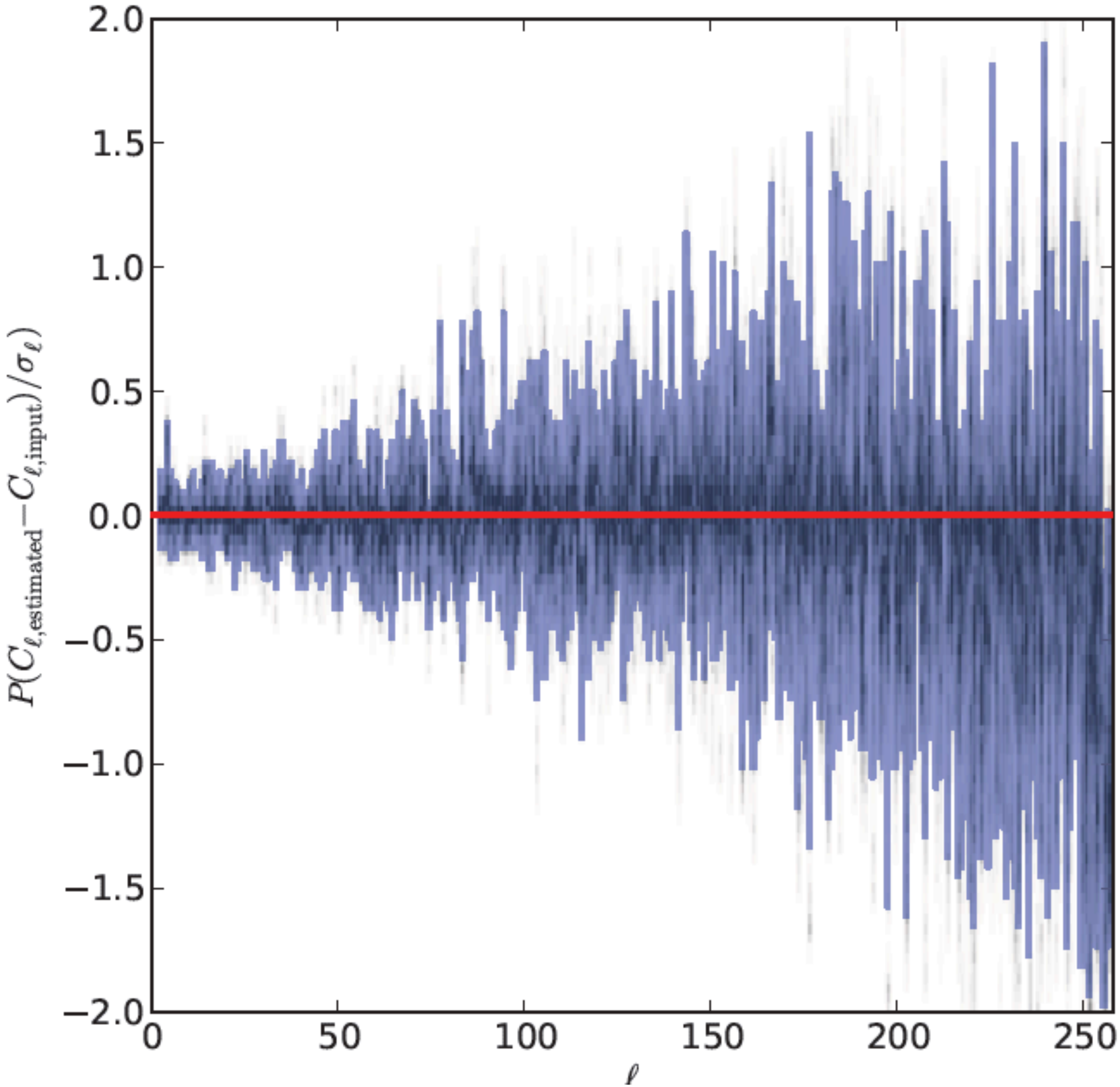


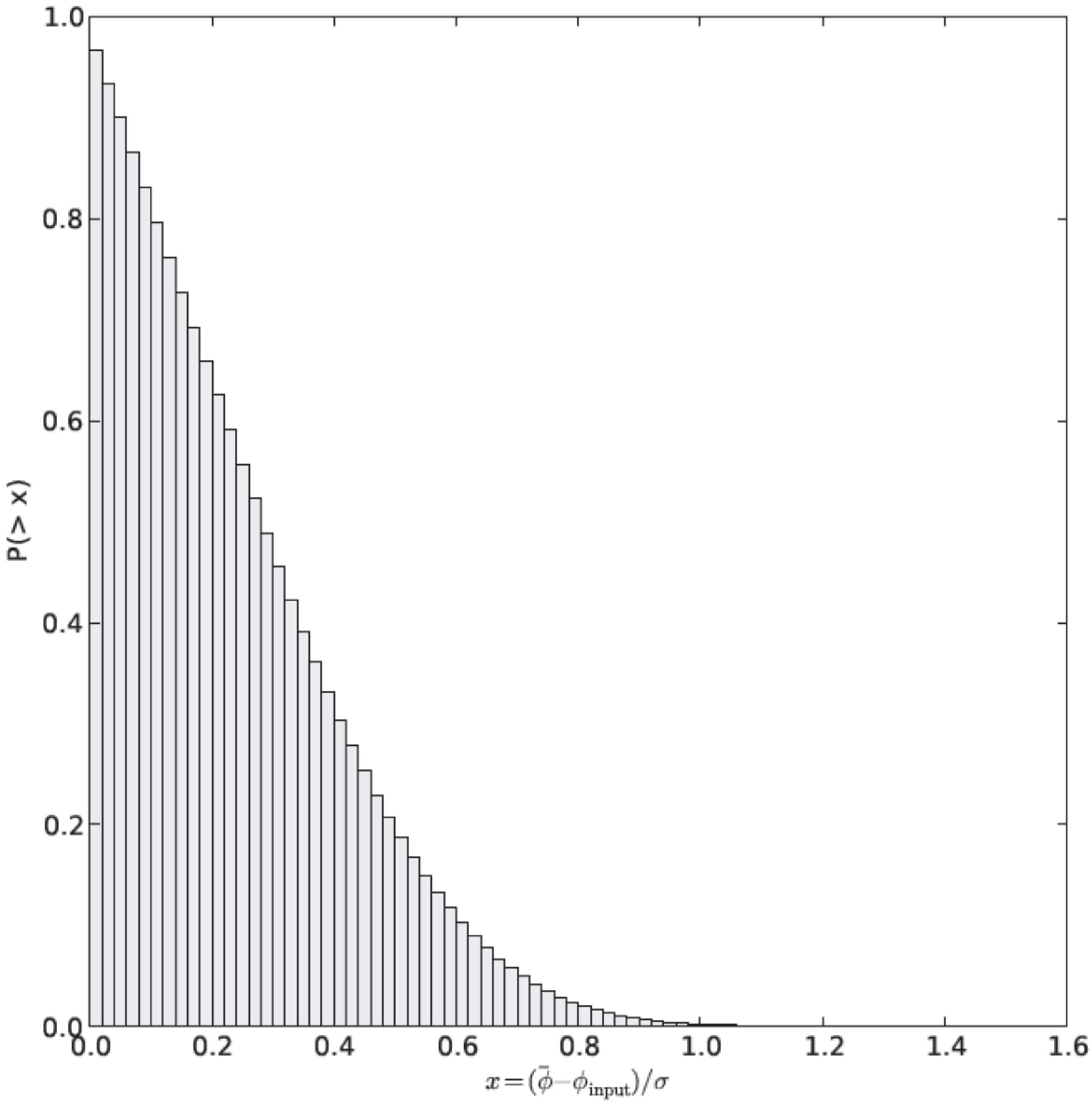
Input lensing signal



Difference between 1 reconstructed unlensed CMB and the true CMB




2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
0
50
100
150
200
250
$P(C_{\ell,\mathrm{estimated}} - C_{\ell,\mathrm{input}})/\sigma_\ell)$
ℓ


1.0
0.8
0.6
0.4
0.2
0.0
P(> x)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
$x = (\bar{\phi} - \phi_{\mathrm{input}})/\sigma$

Conclusions

- First implementation of fully Bayesian lensing analysis applied to idealized simulation
- Doubly hard sampling problem
- Interesting application of the exchange algorithm

PRECISION COSMOGRAPHY WITH STACKED COSMIC VOIDS

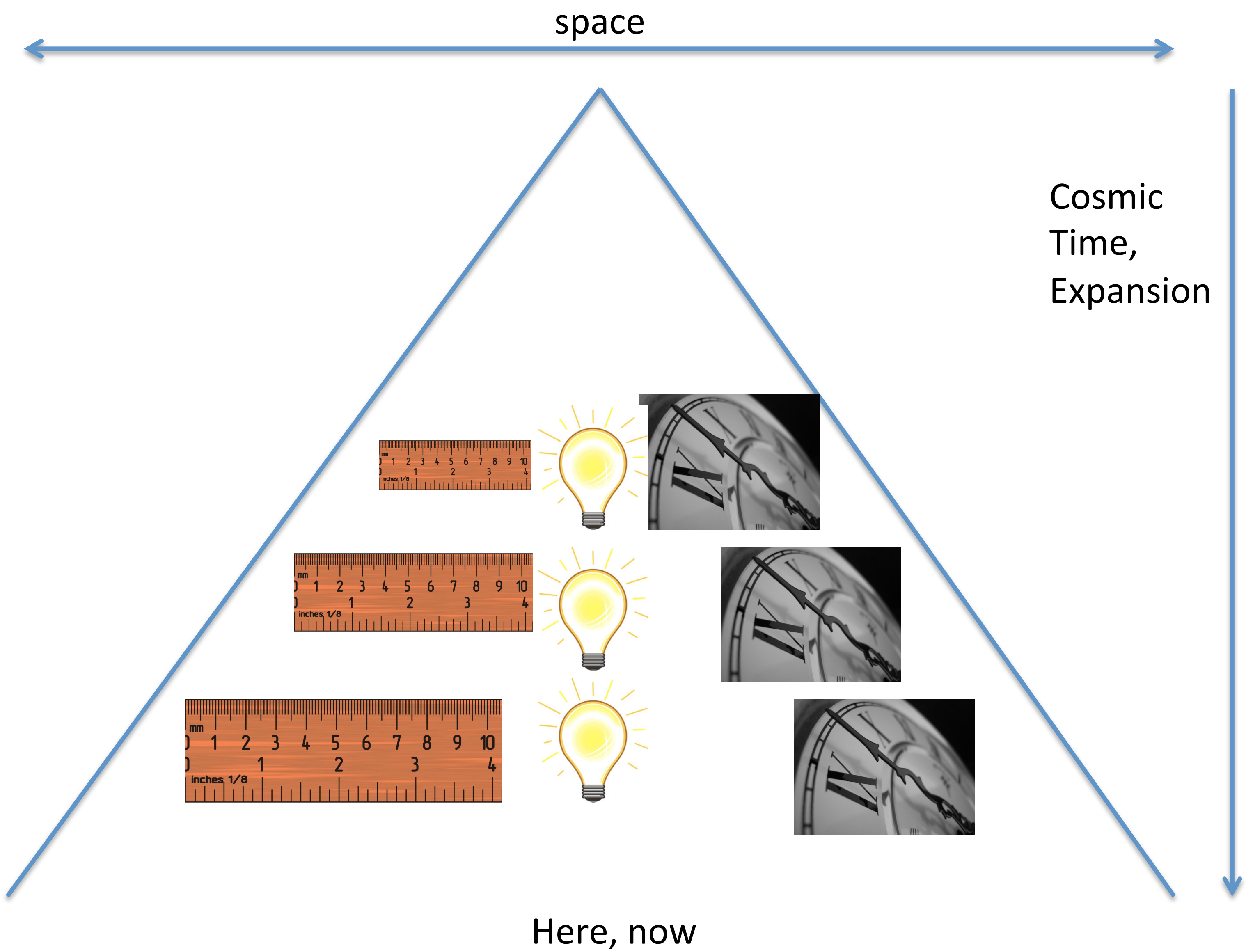
… or “voids as cosmic stopwatches,”

… or “cosmic voids as standard spheres”

… or “dark energy phenomenology from nothing”

Lavaux & Wandelt, arxiv:1110.0345)

Cosmography



Dark energy phenomenology

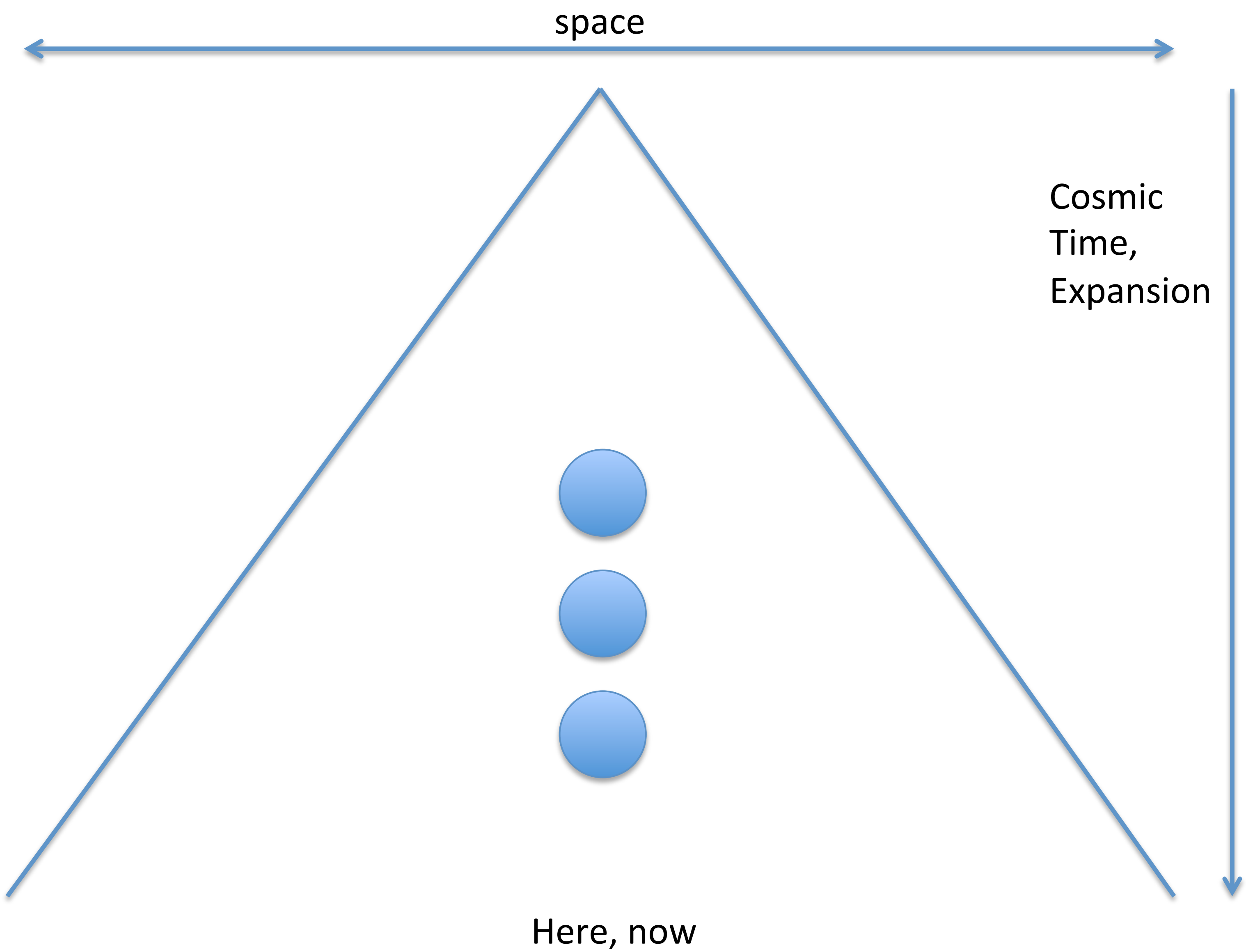
- Through GR, dark energy has a direct impact on the expansion history of the Universe:
  - Supernovae measure the way **standard candles** dim with expansion (luminosity distance, indirect)
  - The imprint primordial gas leaves on the dark matter gives **a standard ruler** whose length we can measure in the distant Universe (large scale, few modes).
  - If we had **cosmic clocks** we could directly find the relationship between expansion and cosmic time (Jimenez and Loeb 2002) (but need to understand galaxy spectra)

Cosmic stopwatches

- Good clocks (with long term stability) are hard to find
- But if we had standard spheres we could directly work out the cosmic expansion history by dividing their angular extent by their temporal extent
- This is a differential measurement
- Standard spheres are like cosmic stopwatches
- This is the Alcock-Paczynsky technique
- First realized for voids by Ryden in 1995!

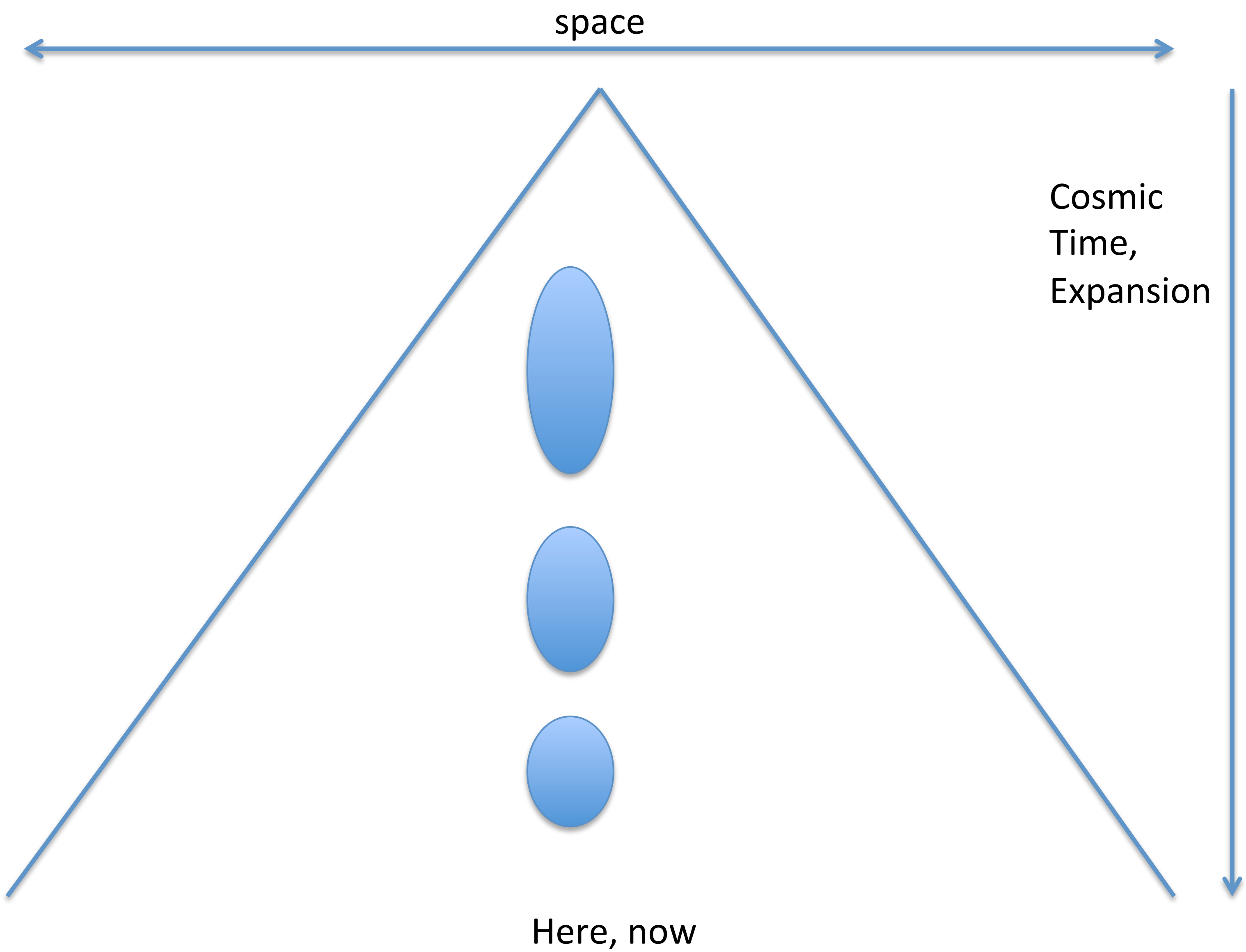
(Lavaux & Wandelt, arxiv:1110.0345)

Cosmography with voids



(Lavaux & Wandelt, arxiv:1110.0345)

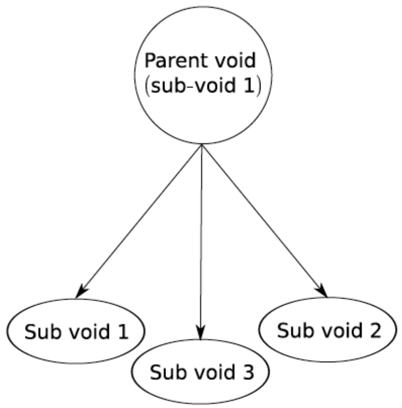
Cosmography with voids

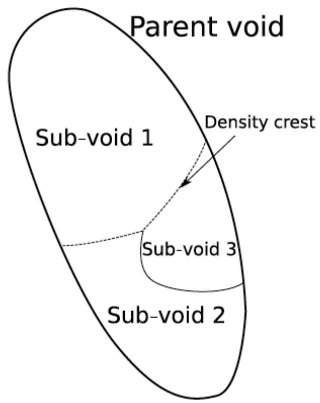


(Lavaux & Wandelt, arxiv:1110.0345)

Void definition

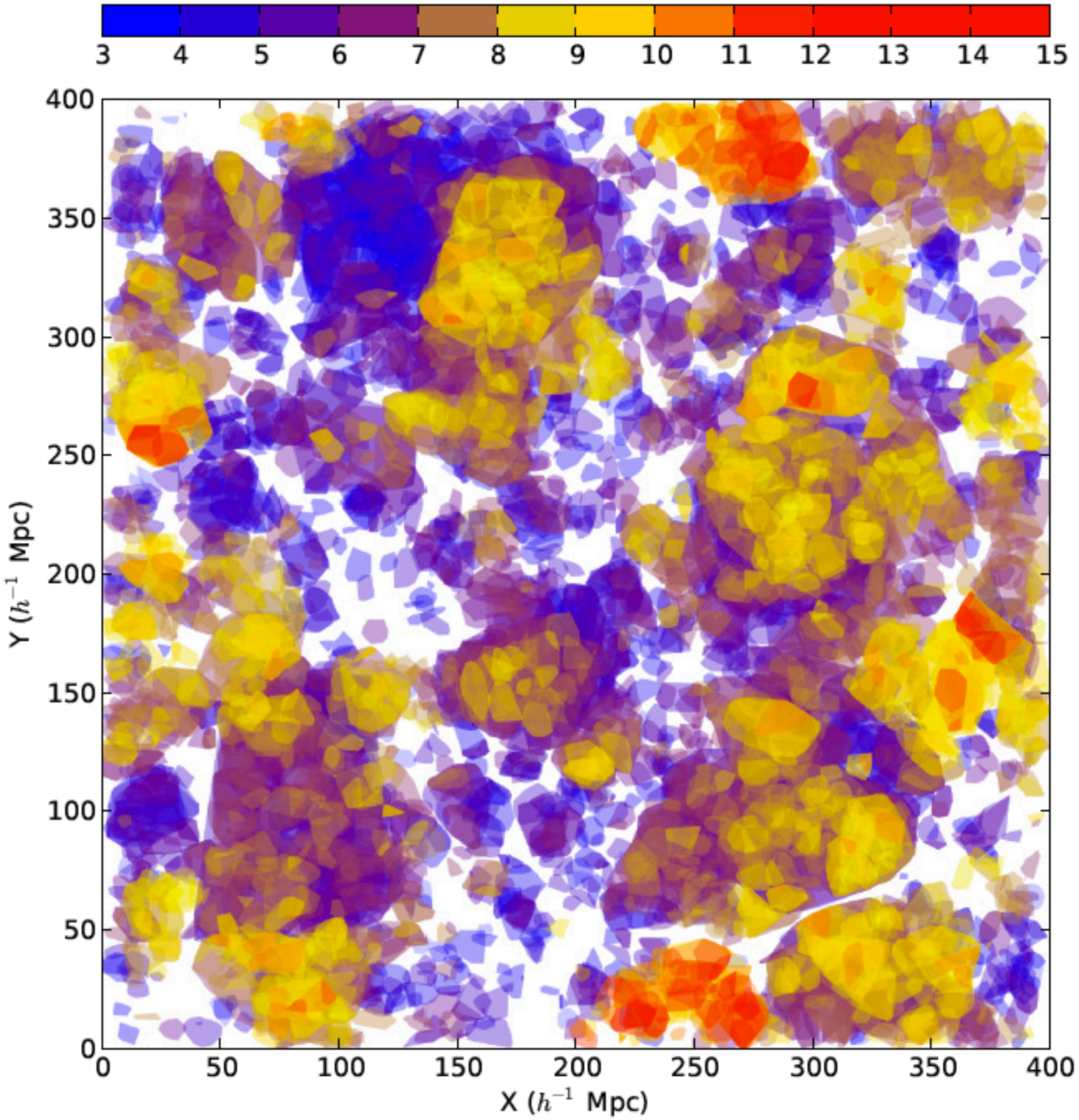
- Need a good definition of voids – spatial statistics
  - Modified ZOBOV scheme (Neyrinck 2008)
    - Uses Voronoi tesselation to find catchment basins associated with density minima
    - Does not assume a hard-coded void shape
    - Organizes voids in a tree-structure





(Lavaux & Wandelt, arxiv:1110.0345)

Void tree



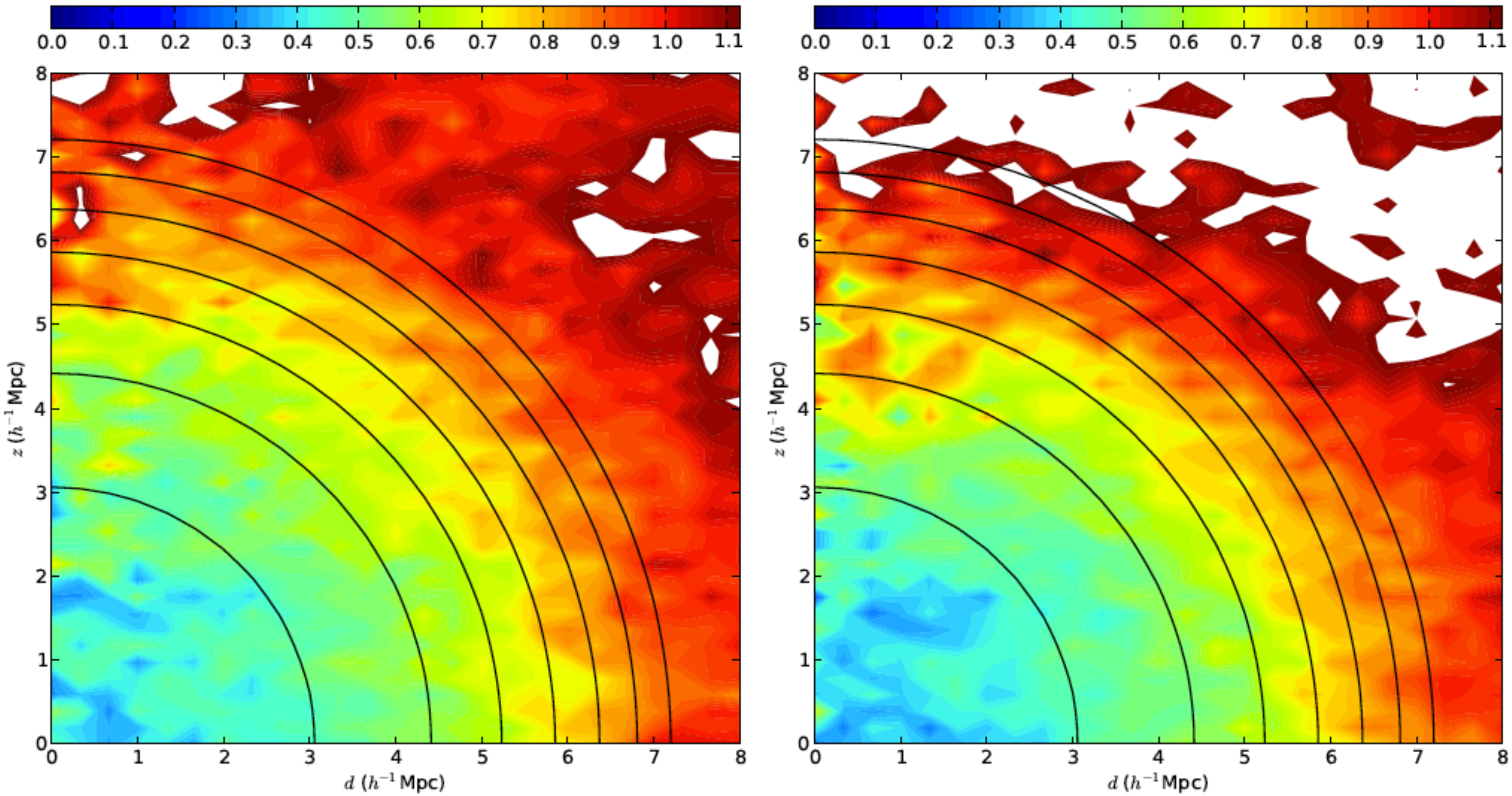
(Lavaux & Wandelt, arxiv:1110.0345)

Practical issues

- 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
  - Solution: model coherent motions in void.
    - In fact, the effect largely produces a constant bias that can be corrected.

(Lavaux & Wandelt, arxiv:1110.0345)

Void stacking



Cosmological redshifts only

Full redshift, with peculiar velocities

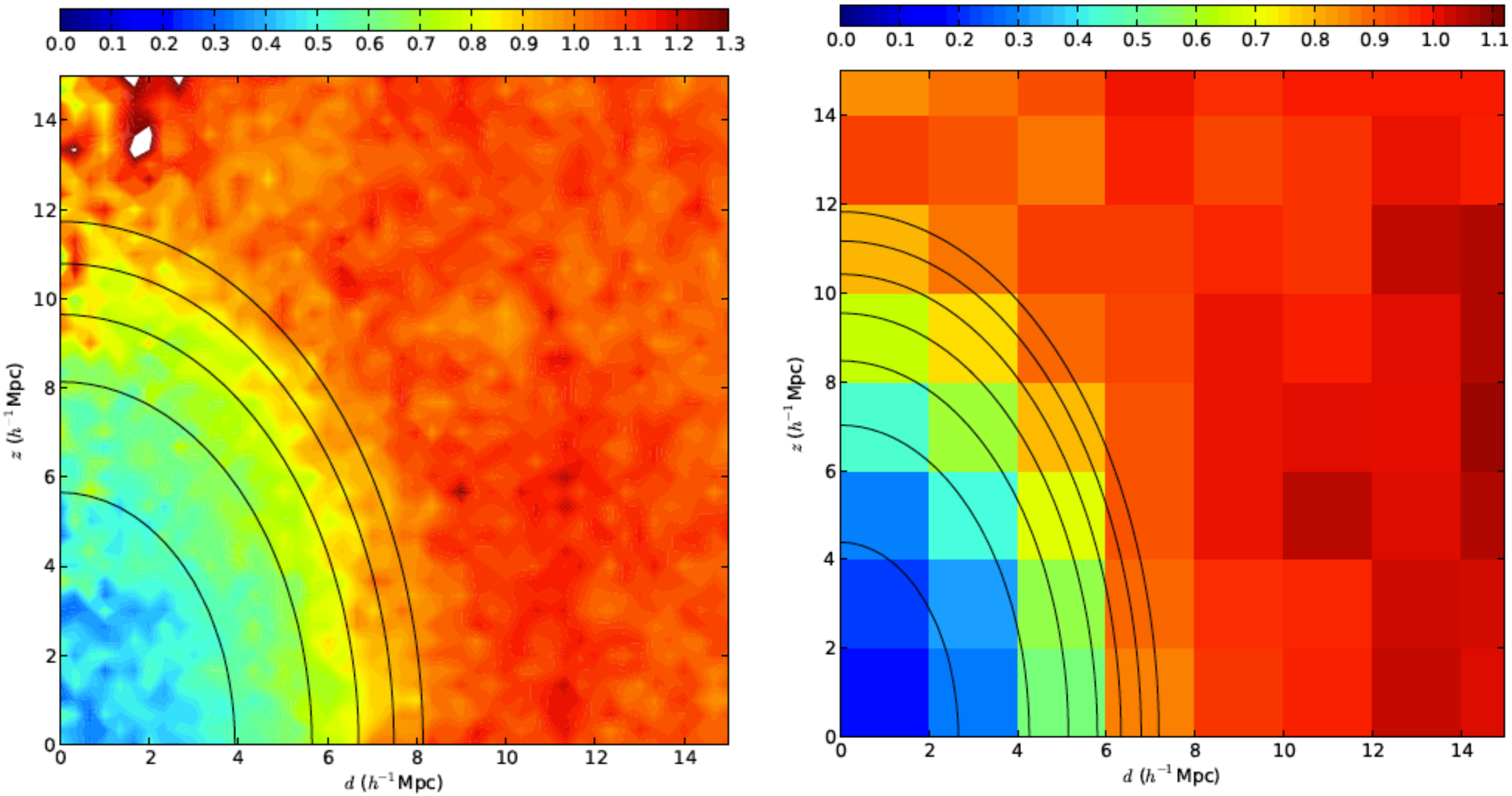
(Lavaux & Wandelt, arxiv:1110.0345)

Statistical model for shape inference

- Void walls have structure
  - 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, arxiv:1110.0345)

Coarser pixelization de-correlates clumping error



(Lavaux & Wandelt, arxiv:1110.0345)

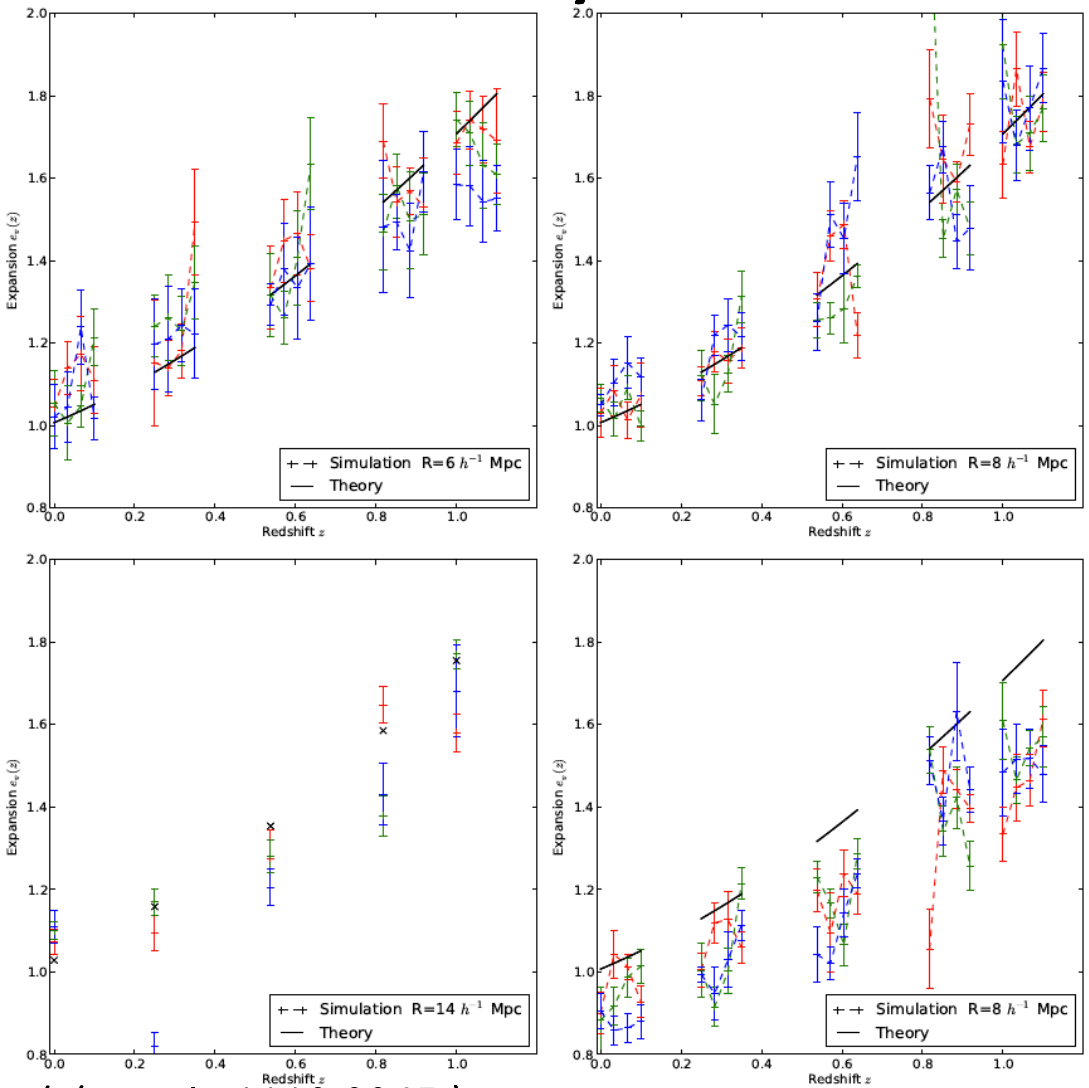
Hubble parameter H(z) recovered from N-body simulation

Void radius: 4 Mpc/h

Void radius: 8 Mpc/h

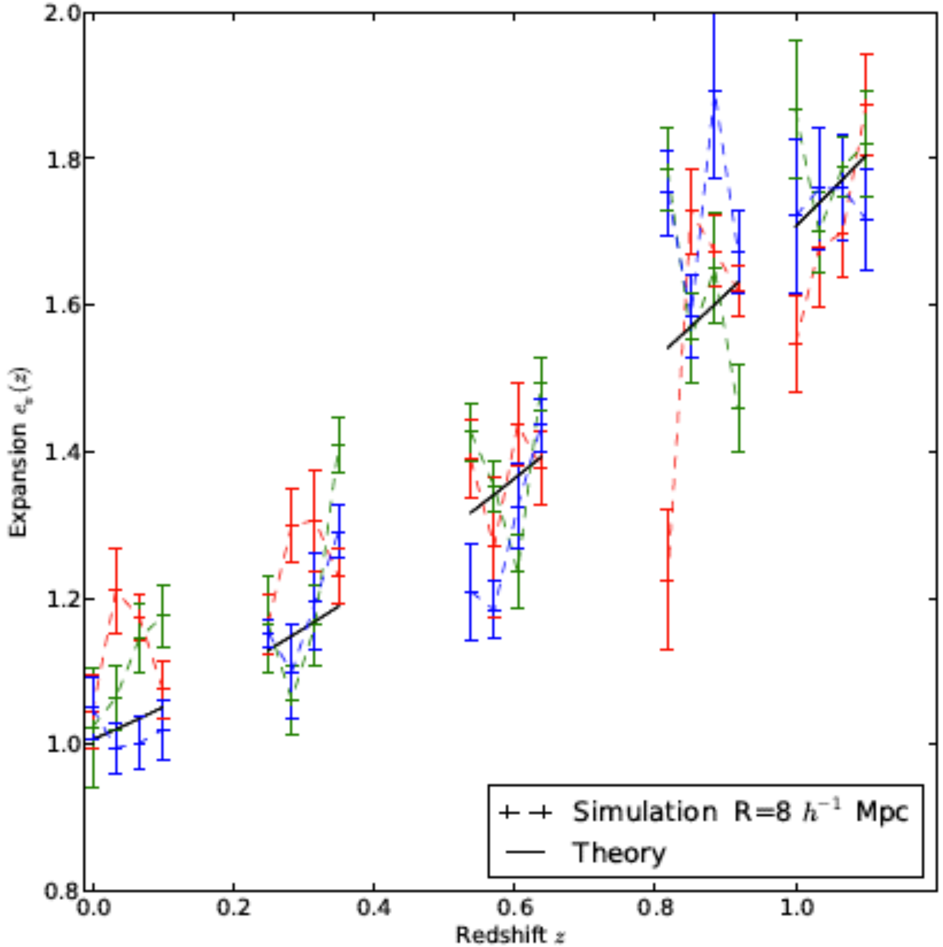
Void radius: 14 Mpc/h

Void radius: 8 Mpc/h with peculiar velocities



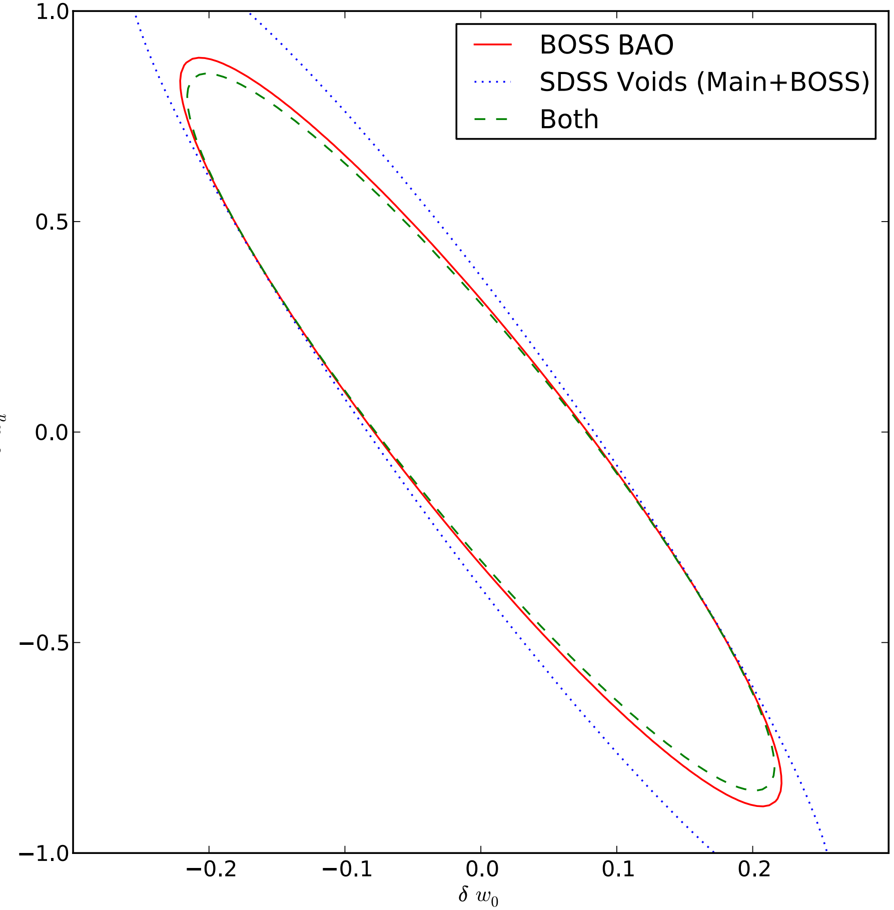
(Lavaux & Wandelt, arxiv:1110.0345)

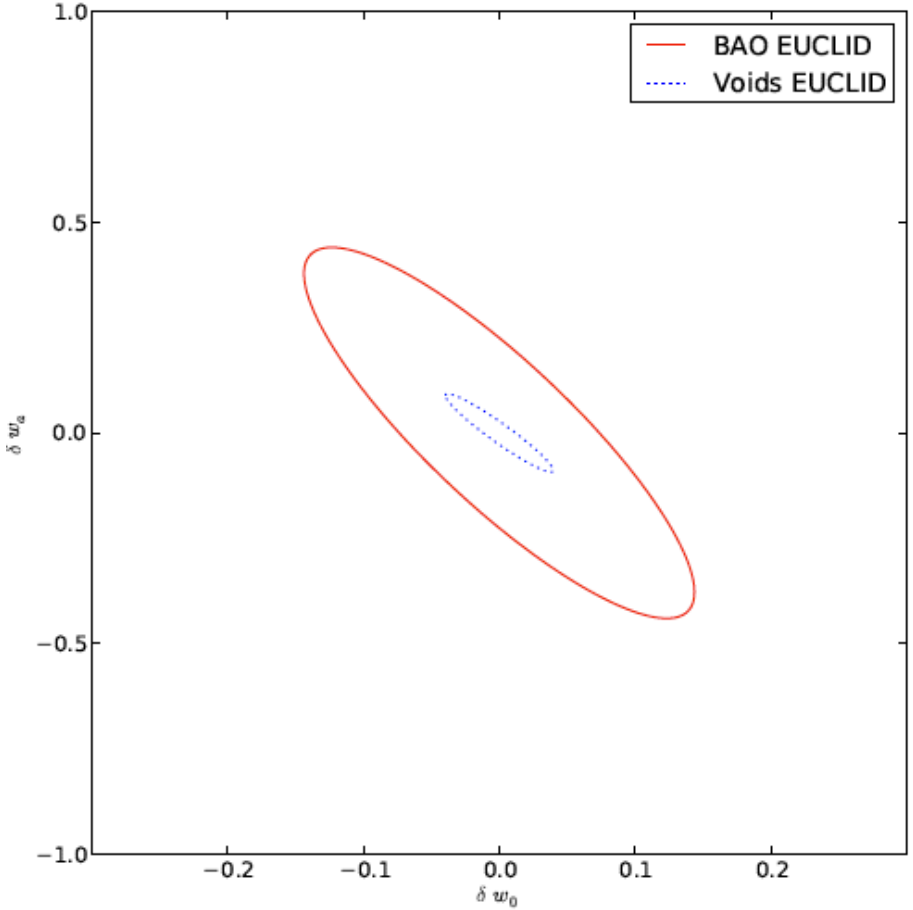
Simple de-biasing gives good H(z) estimate even with peculiar velocities



(Lavaux & Wandelt, arxiv:1110.0345)

Dark energy constraint forecast





(Lavaux & Wandelt, arxiv:1110.0345)

Comparable to Baryon Acoustic Oscillation constraints for current data

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

Next steps

- Each step (void identification, stacking/averaging, shape inference) involves choices and parameters. We have only scratched the surface.
- More detailed peculiar velocity modeling will yield additional signal to noise
- Need more realistic simulations including more real-world systematics:
  - Biased tracers of the density
  - Complicated masks
  - But since this is a purely geometrical measurement this signal is very robust
- Further, complementary constraints could come from additional inferences about the growth of cosmic structure from the intrinsic shapes of voids – which this method ignores. (e.g., Park & Lee 2007; Biswas et al. 2010; Lavaux & Wandelt (2010))

Conclusions

- Non-linear, principled cosmological inference with ~10^7 parameters is becoming feasible
- Global analysis of survey data can add tremendous value to photo-z surveys
- New sampling techniques allow progress for doubly hard inference problems, like lensing
- "Stacked Voids" are a new, purely geometrical dark energy observable. First estimates suggest tremendous additional potential for constraints on Dark Energy
- Opens up many interesting statistics questions
  - Spatial statistics
  - Robustness
  - Optimal survey design for the AP test with stacked voids

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)	$r = 18$	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)	$r = 20$	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	SDSS+BOSS LRG	63
BAO+Voids	SDSS+BOSS	91
Voids	EUCLID	~5 500
BAO	EUCLID	185
BAO+Voids	EUCLID	~5 500

(Lavaux & Wandelt, arxiv:1110.0345)