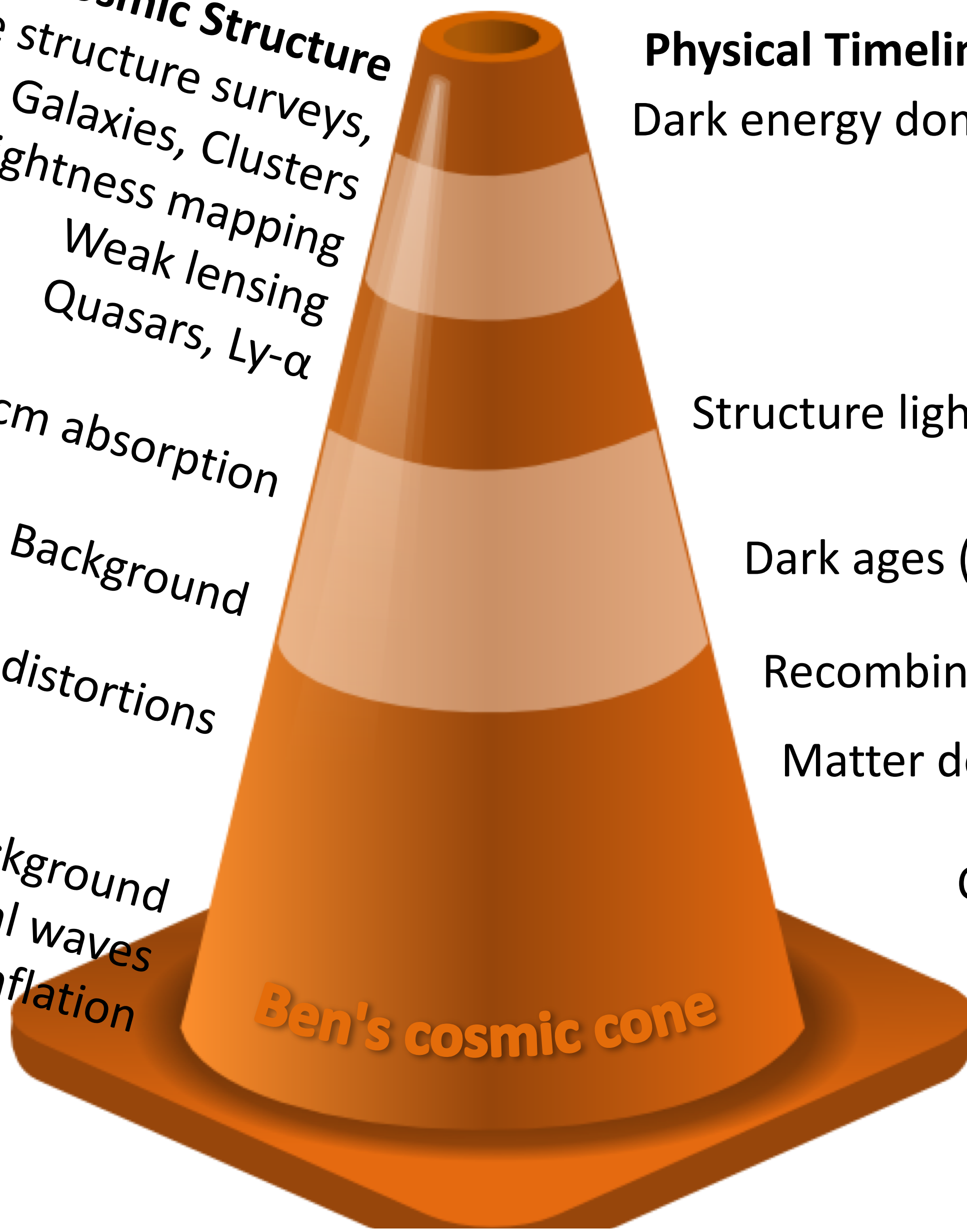
How to learn from cosmological data

Ben Wandelt
Flatiron Institute

Institut d'Astrophysique, Paris (IAP)
Lagrange Institute, Paris (ILP)
Sorbonne University
Princeton University

All of cosmology

You, Today



Tracers of Cosmic Structure

Large scale structure surveys,
Galaxies, Clusters
21-cm brightness mapping
Weak lensing
Quasars, Ly-α

21-cm absorption

Cosmic Microwave Background

CMB Spectral distortions

Neutrino background
Gravitational waves
from inflation

Physical Timeline

Dark energy dominates – acceleration

Structure lights up, reionization

Dark ages (absorption only)

Recombination (CMB is emitted)

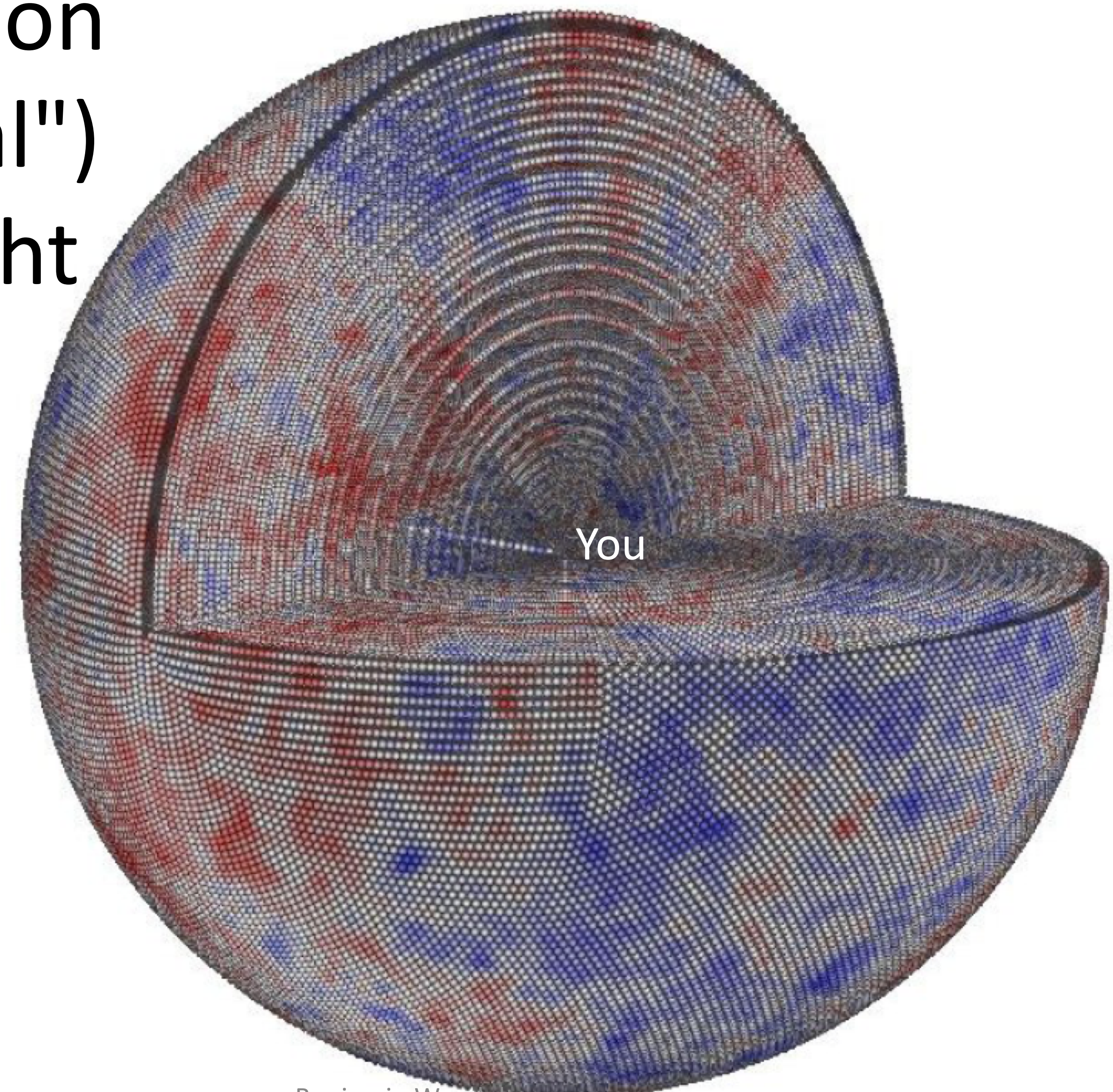
Matter dominates

Creation of the elements
Radiation dominates
End of inflation

What we *want* to learn

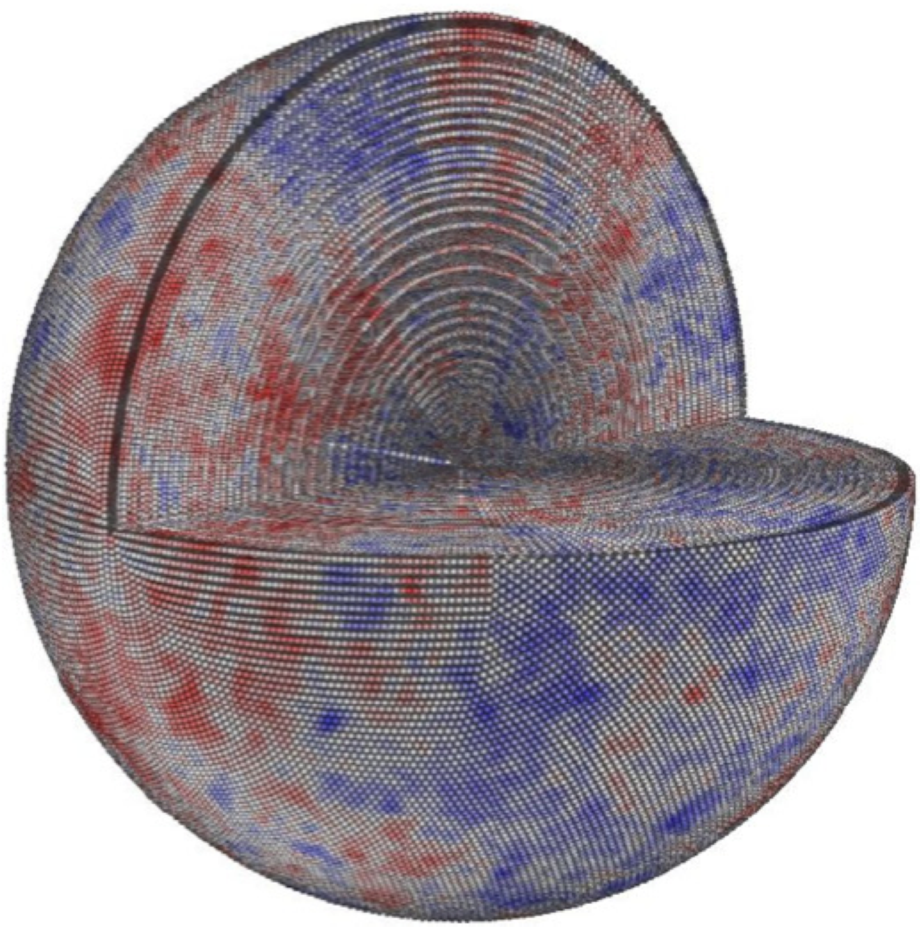
- *How did the Universe begin?*
- *How did structure appear in the Universe?*
- *What is the Universe made of?*
- *What are the properties of dark matter?*
- *What are the properties of dark energy?*

Curvature perturbation ("potential") on the light cone

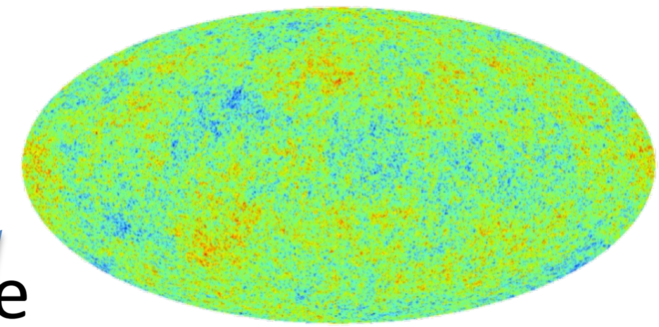


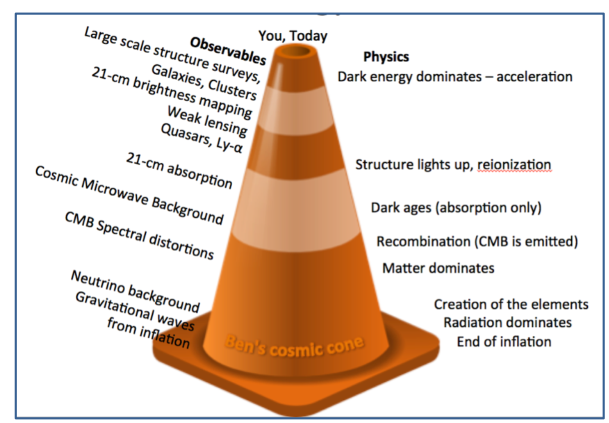
Benjamin Wandelt

Primordial perturbations give rise to all observations

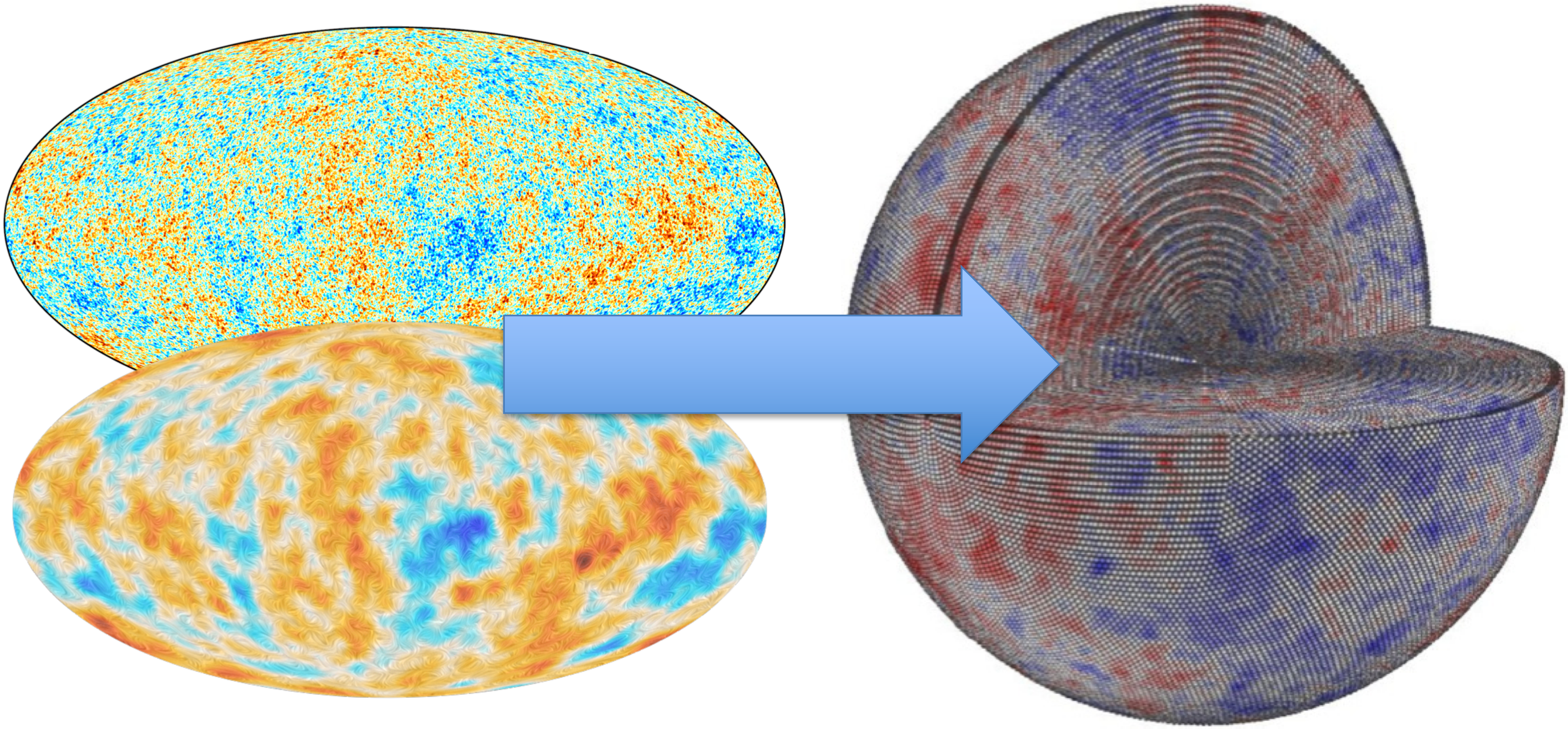


Radiative Transfer
+
Gravity
+
Astro-physics





The linear physics CMB time machine



Observed cosmic microwave background sky

Primordial quantum perturbations

The linear physics CMB time machine

Tracers of Cosmic Structure

Large scale structure surveys,
Galaxies, Clusters
21-cm brightness mapping
Weak lensing
Quasars, Ly-α

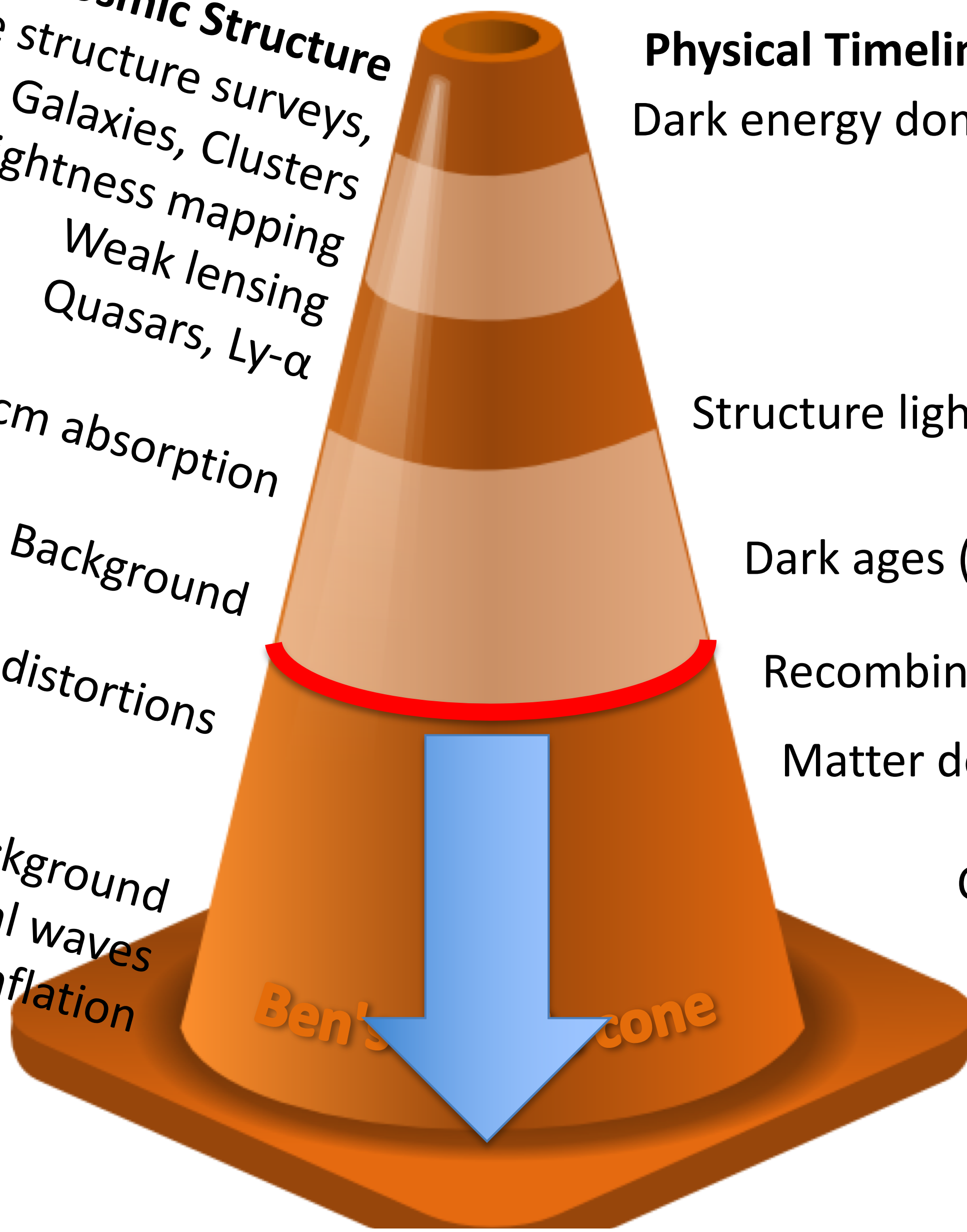
21-cm absorption

Cosmic Microwave Background

CMB Spectral distortions

Neutrino background
Gravitational waves
from inflation

You, Today



Physical Timeline

Dark energy dominates – acceleration

Structure lights up, reionization

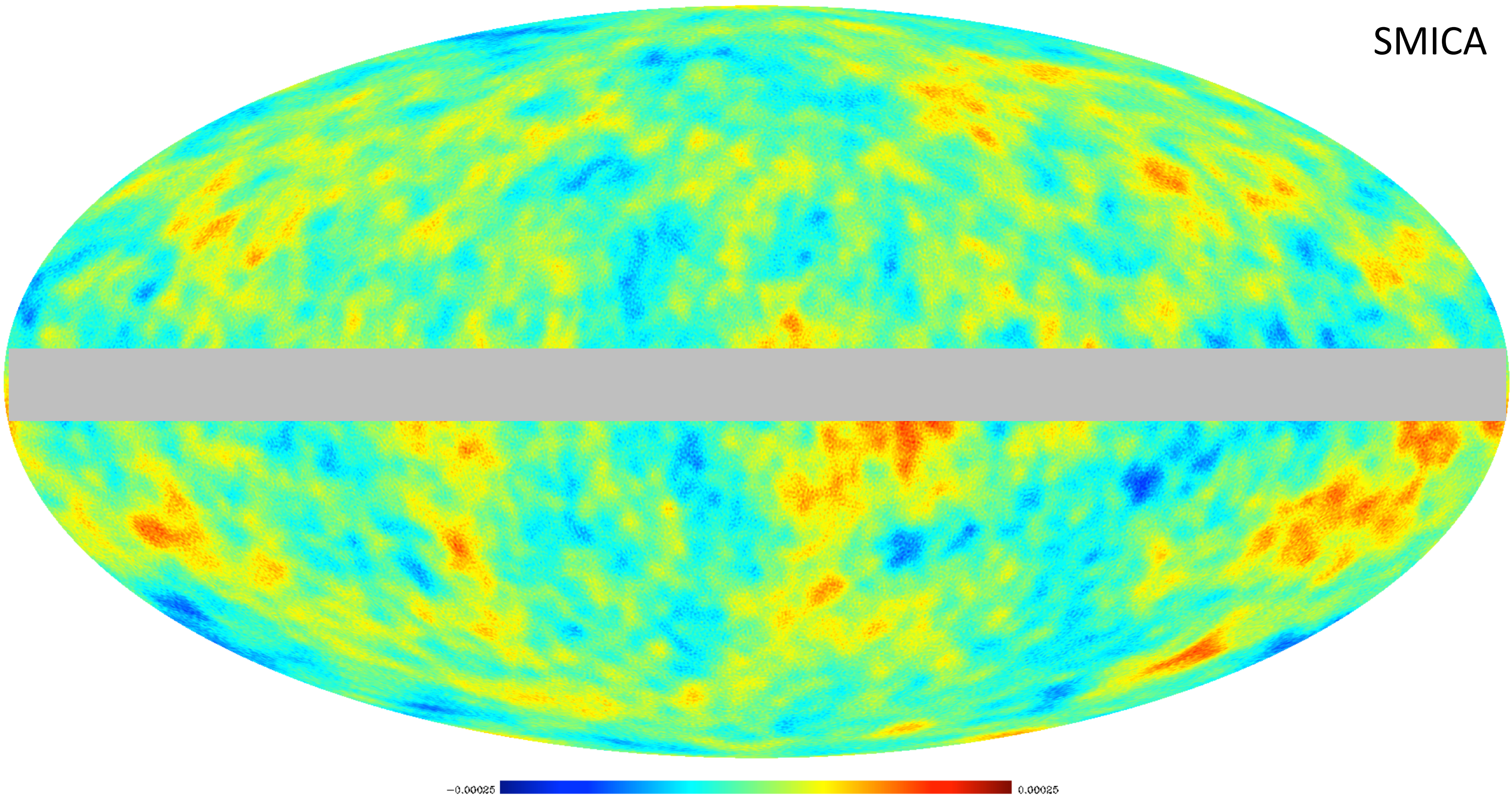
Dark ages (absorption only)

Recombination (CMB is emitted)

Matter dominates

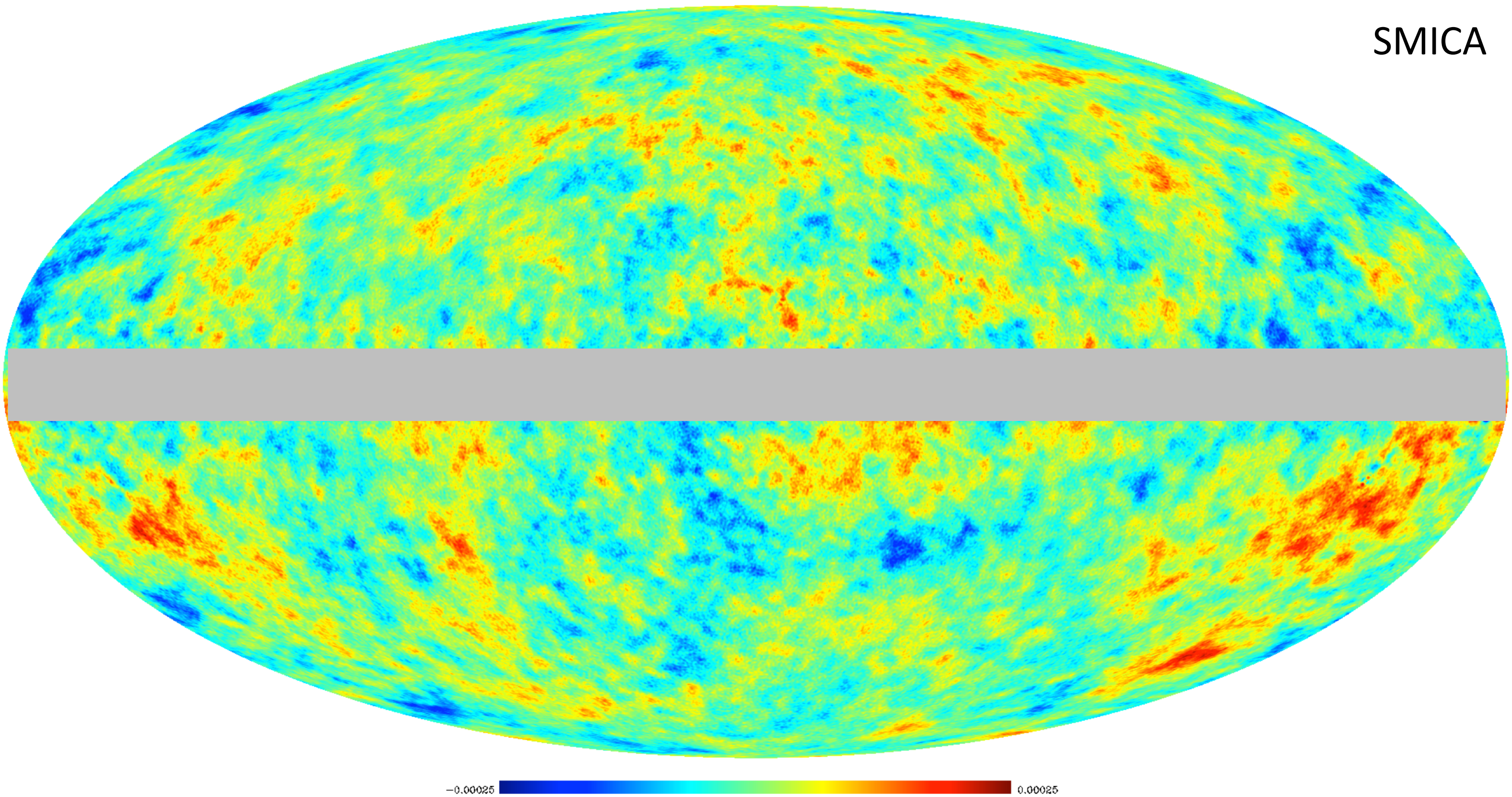
Creation of the elements
Radiation dominates
End of inflation

Primordial curvature perturbations at $r = \eta_{CMB}$ from Planck T data



Method: Komatsu, Spergel, Wandelt 2005

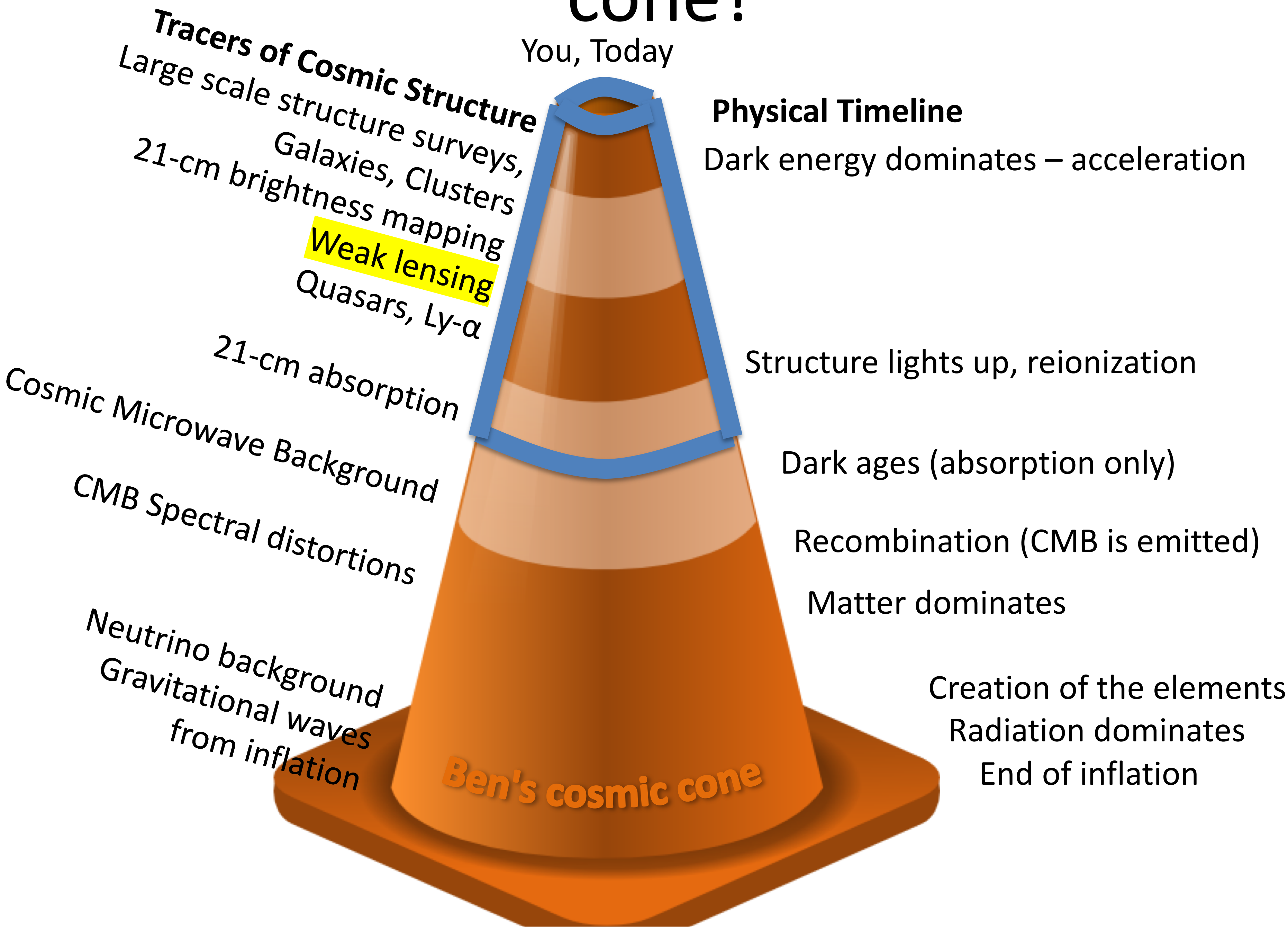
Primordial curvature perturbations at r= η_{CMB} from both Planck T and polarization data

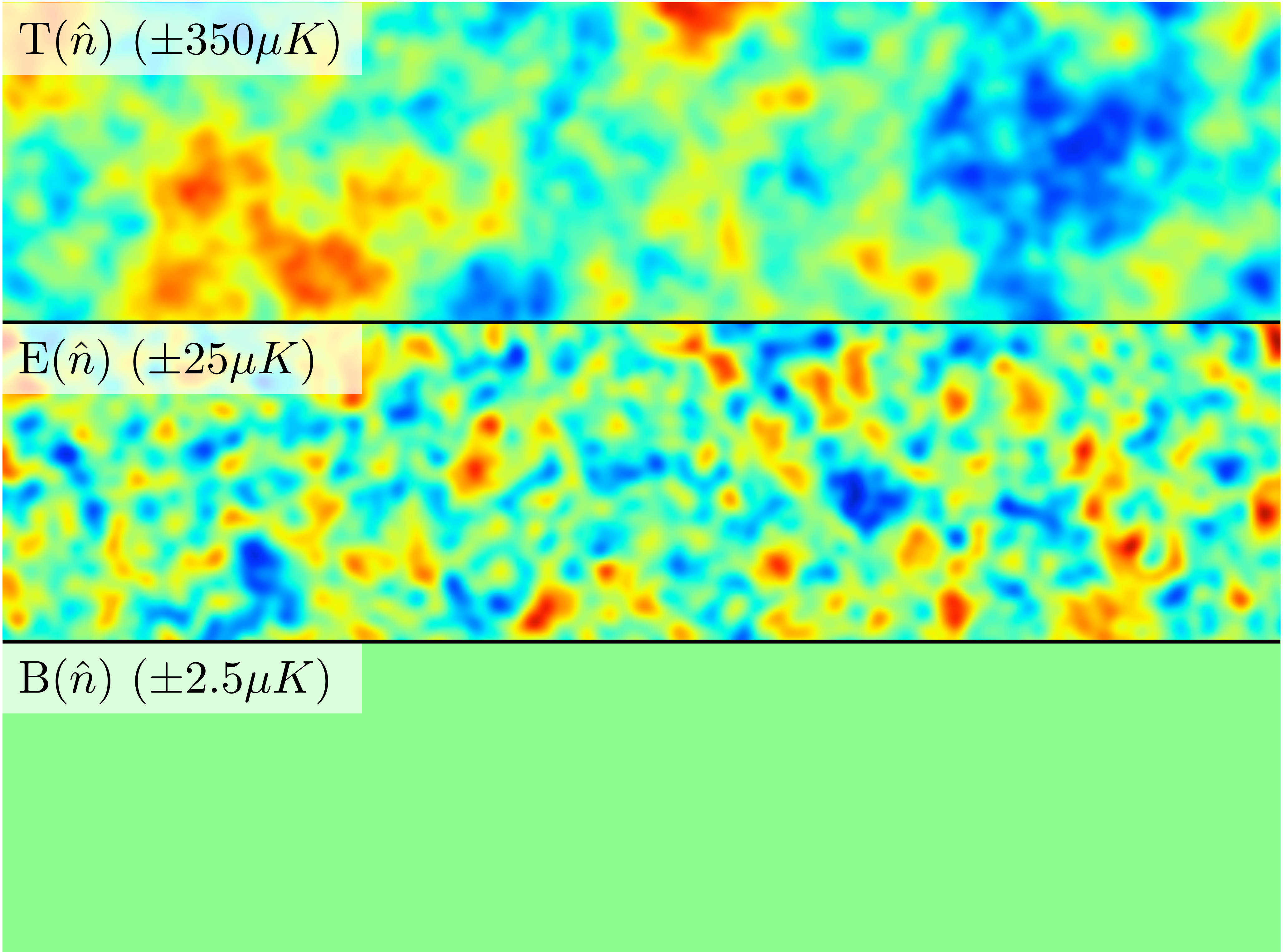


Method: Yadav & Wandelt 2005

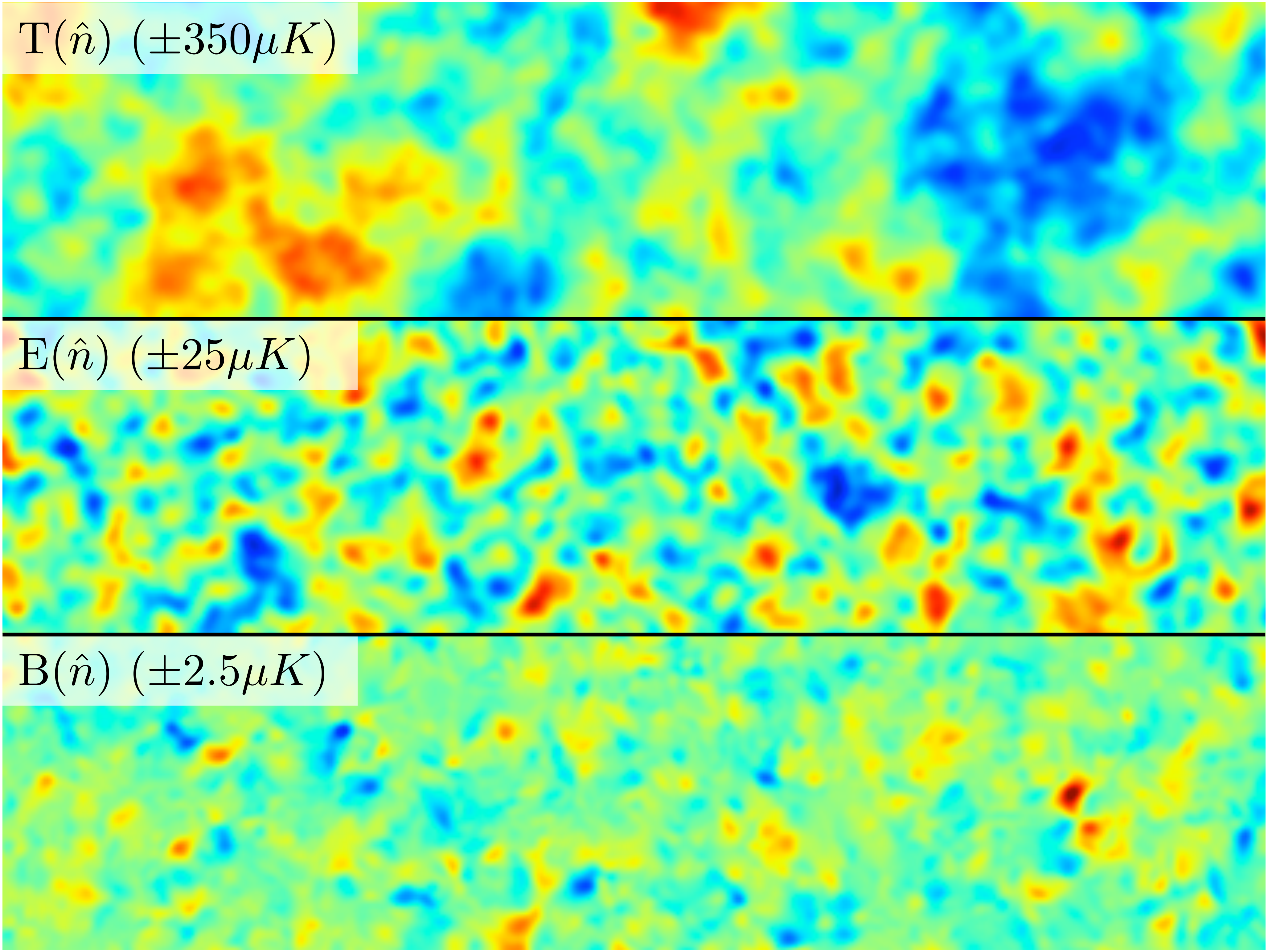
… leading to many results, e.g. Planck constraints on inflation (spectra, non-Gaussianity) and other early universe paradigms

So, what about the rest of the light cone?





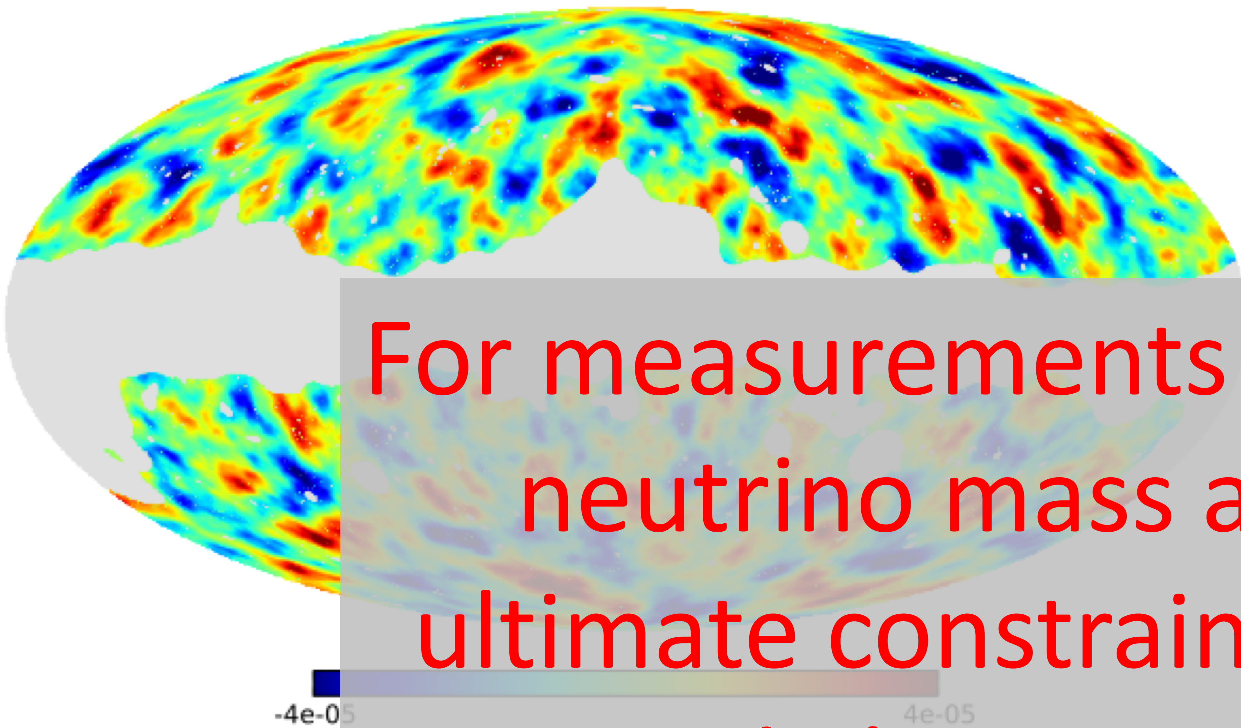
$T(\hat{n})$ ($\pm 350\mu K$)
$E(\hat{n})$ ($\pm 25\mu K$)
$B(\hat{n})$ ($\pm 2.5\mu K$)

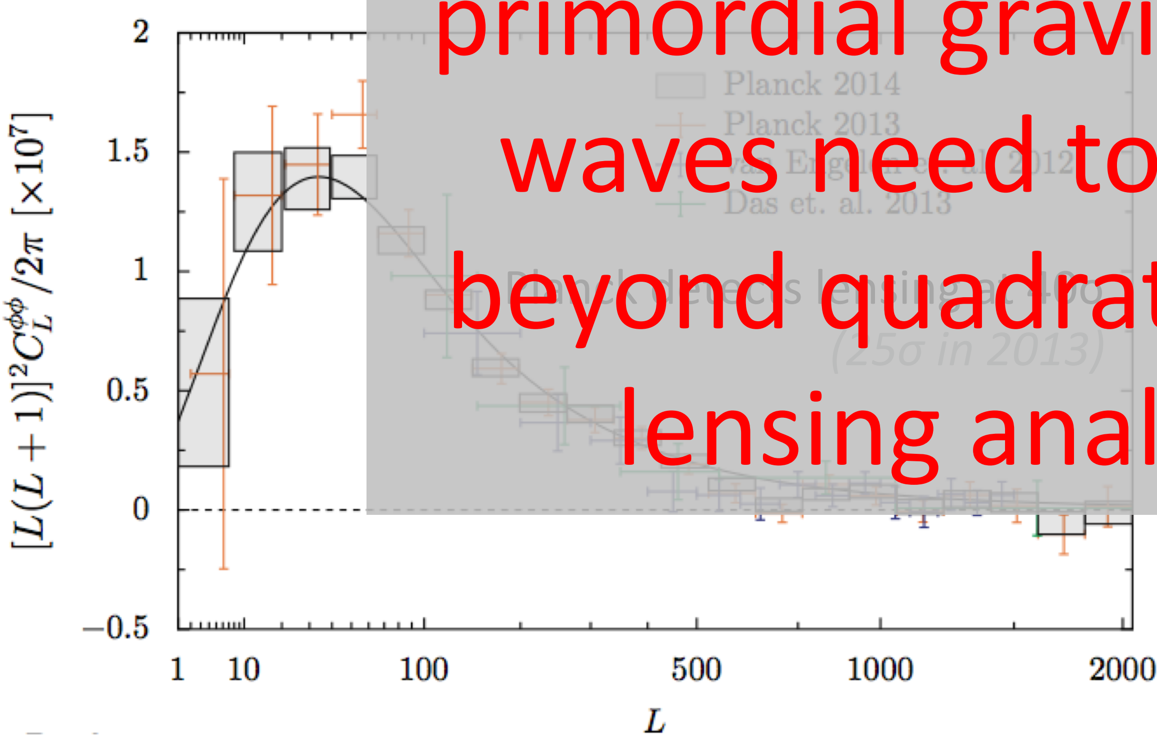


T($\hat{n}$) ($\pm 350\mu K$)
E($\hat{n}$) ($\pm 25\mu K$)
B($\hat{n}$) ($\pm 2.5\mu K$)

All clumping mass in the Universe, as it lenses the CMB

\- seen using the quadratic estimator





For measurements of the neutrino mass and ultimate constraints on primordial gravitational waves need to do go beyond quadratic CMB lensing analysis

Bayesian lensing potential reconstruction

- A fully Bayesian approach would be **optimal**. But studying the posterior pdf for the lensing potential given the data is a "doubly intractable" problem that has remained unsolved for ~20 years.
- All standard Bayesian inference approaches fail in practice
- The key problem is the *lensing determinant*

A new way to think about weak lensing

Expansion:

$$T(x + \nabla\phi(x)) = \left[\sum_{i=0}^{N} \frac{1}{i!} \left[\nabla\phi(x)\right]^i \nabla^i\right] T(x)$$

LensFlow:

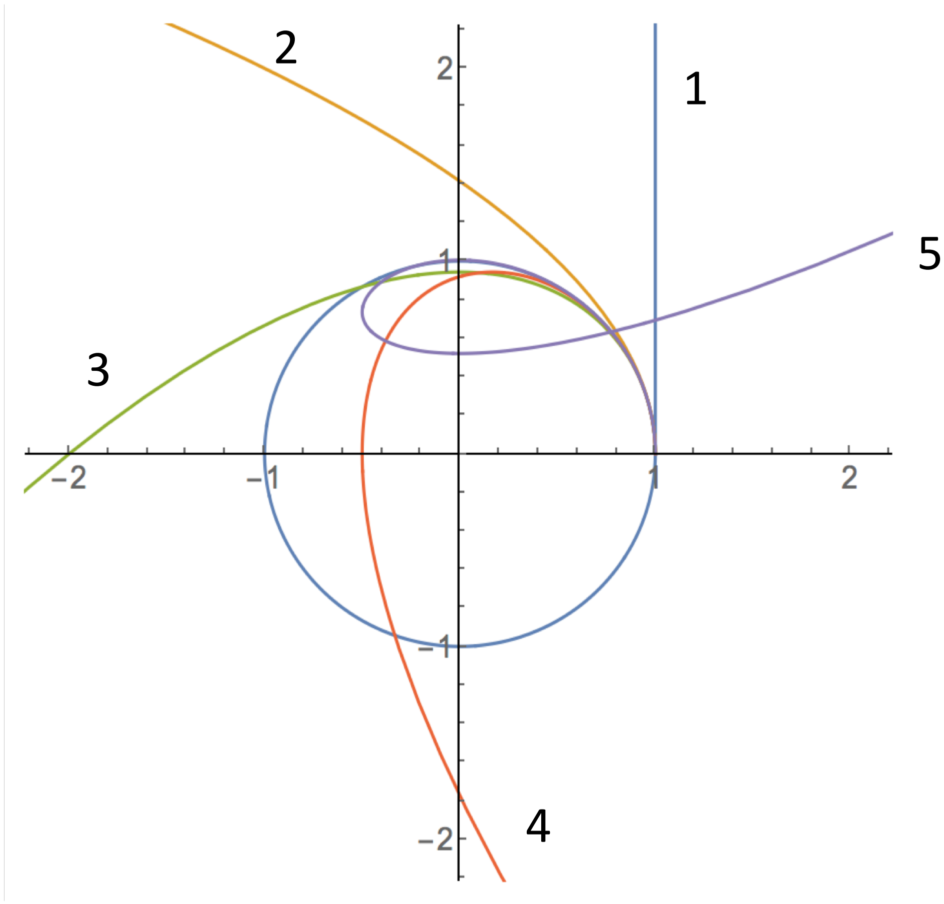
$$T(x + \nabla\phi(x)) = \left[\prod_{i=0}^{N} \left(1 + \frac{1}{N} p_i \nabla\ \right)\right] T(x)$$

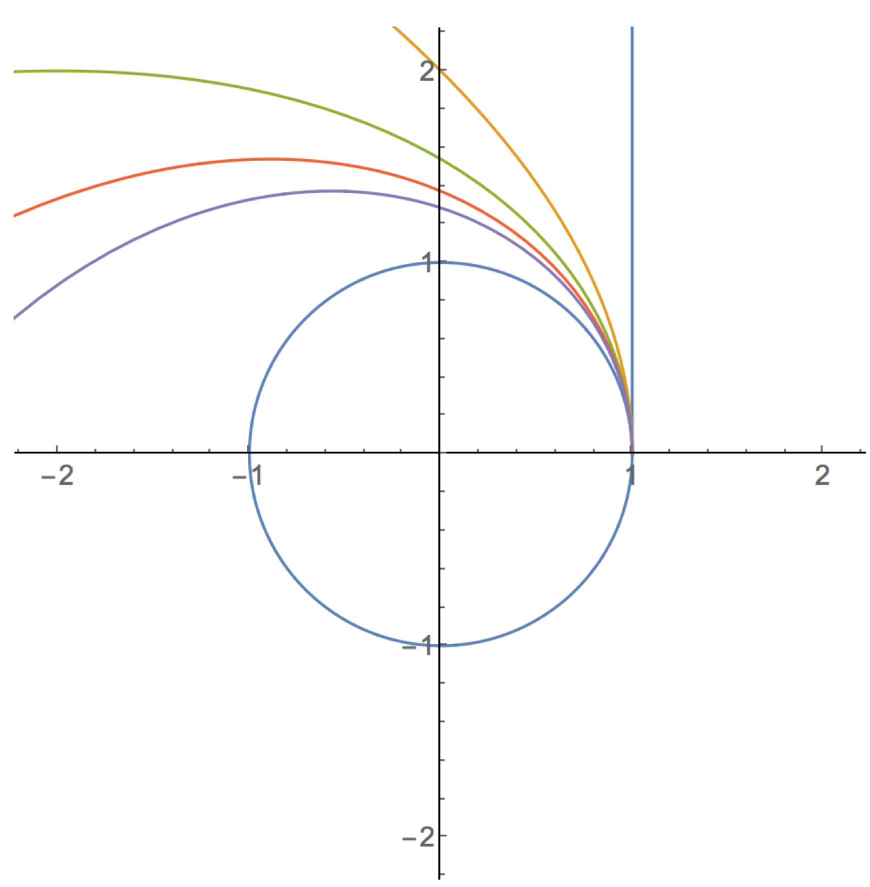
Conserves the lensing determinant!

$$p_t^i = (\nabla^j \phi)(M_t^{-1})^{ji}$$

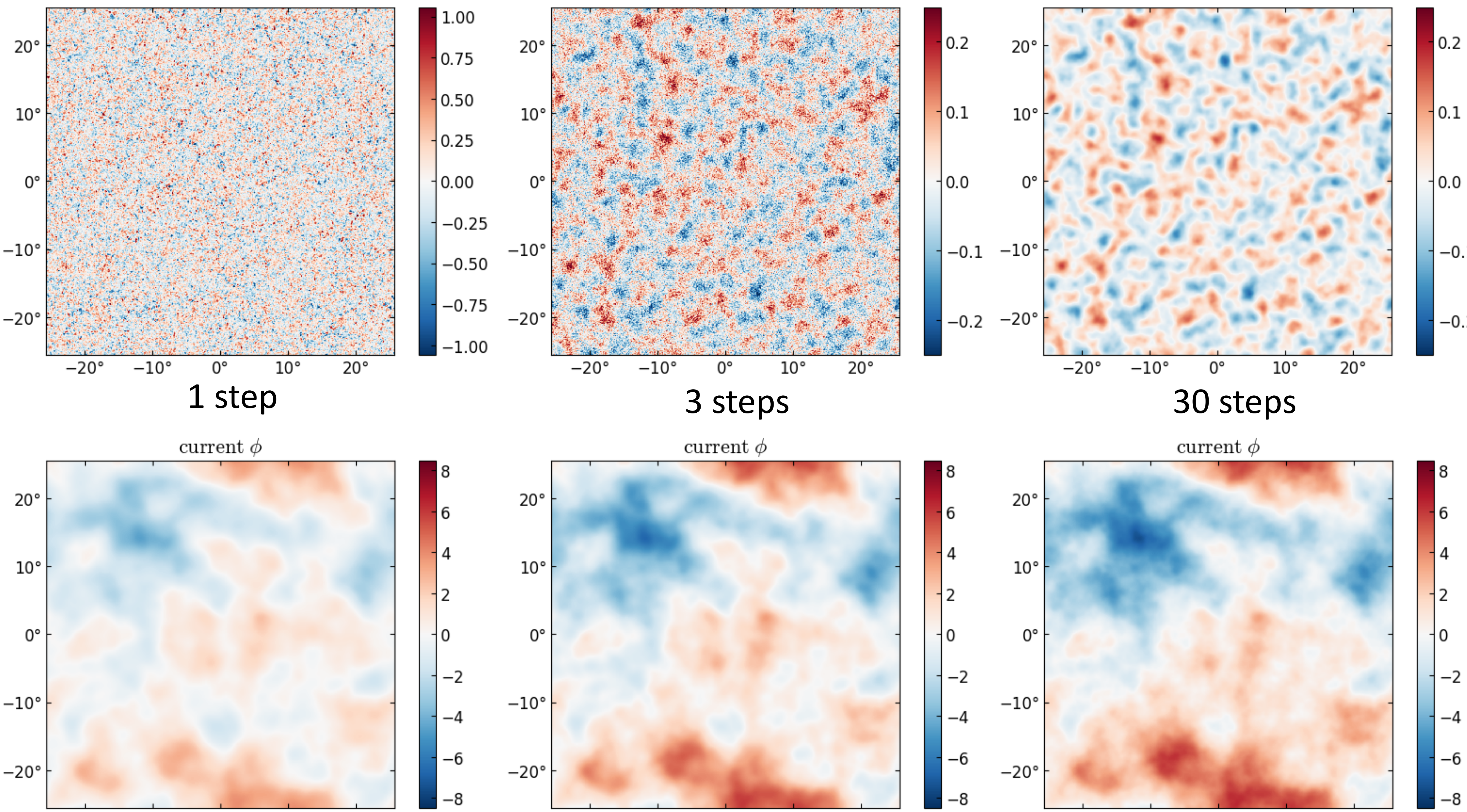
Millea, Anderes & Wandelt arXiv:1708.06753

Taylor and “Flow” approximations to rotations



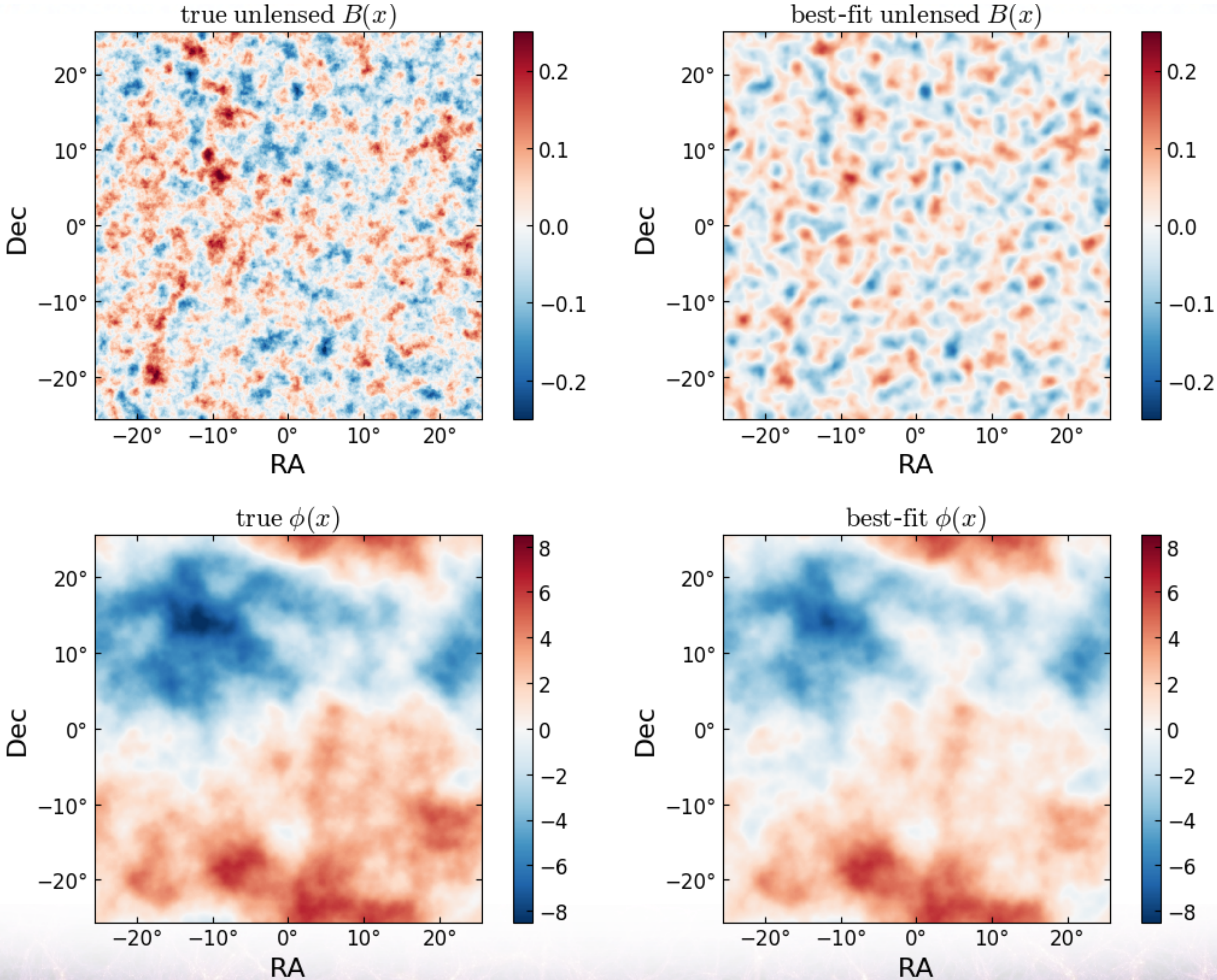


Example de-lensing

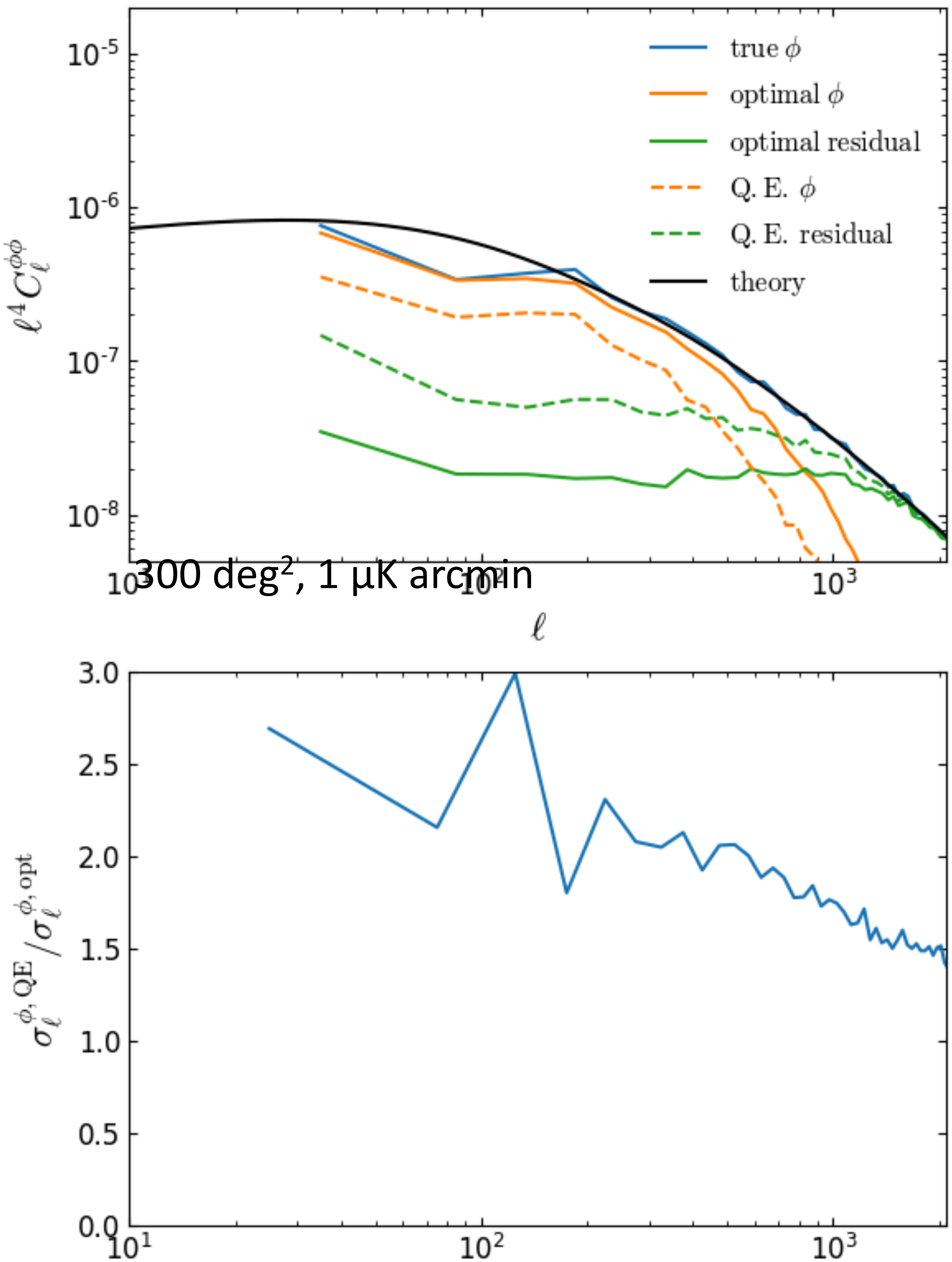


Millea, Anderes & Wandelt arXiv:1708.06753

Example de-lensing



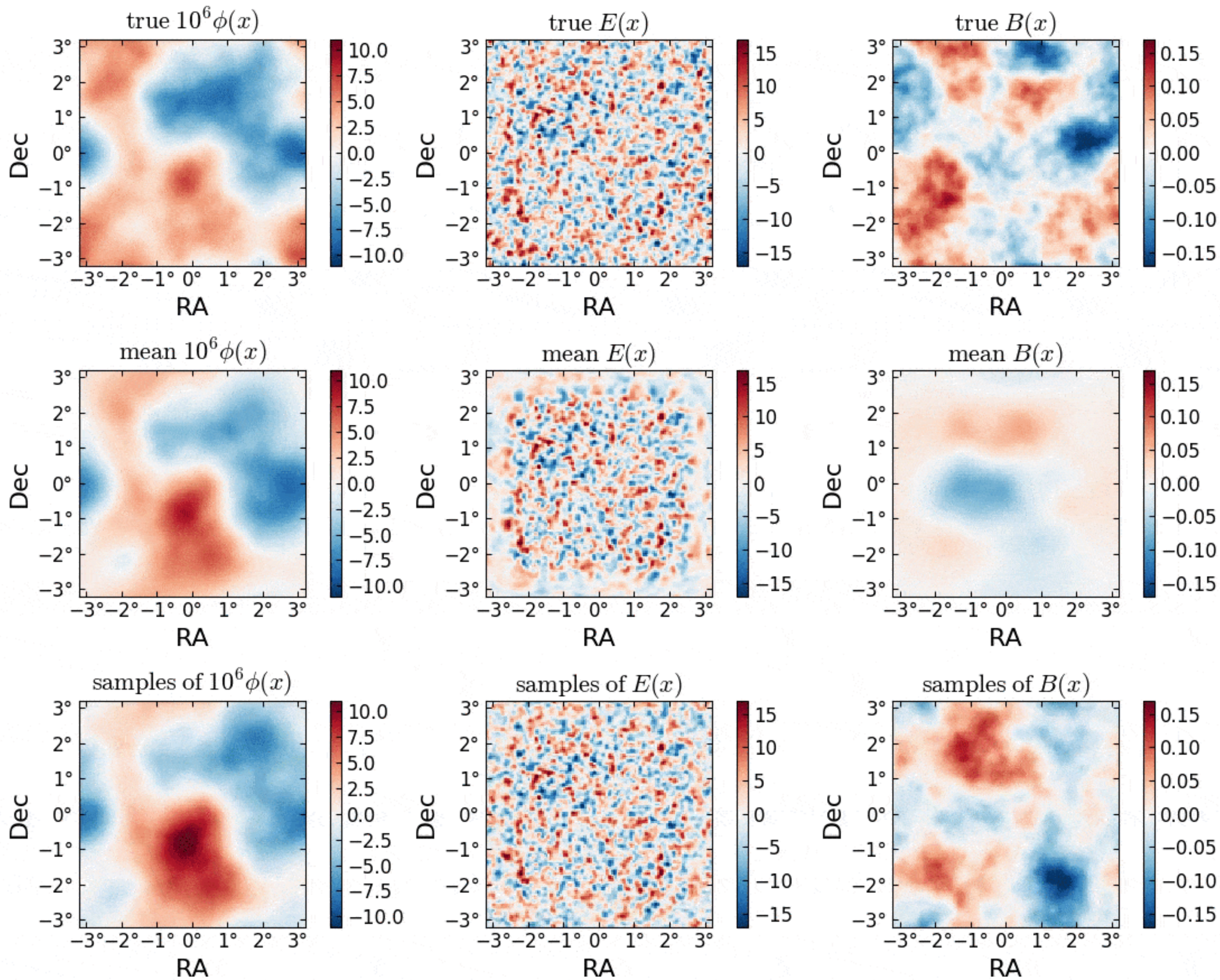
Millea, Anderes & Wandelt arXiv:1708.06753



Millea, Anderes & Wandelt, arXiv:1708.06753

Bayesian lensing inference gives 50-200% improvement over standard quadratic estimator

Full delensing B-mode reconstruction and r inference

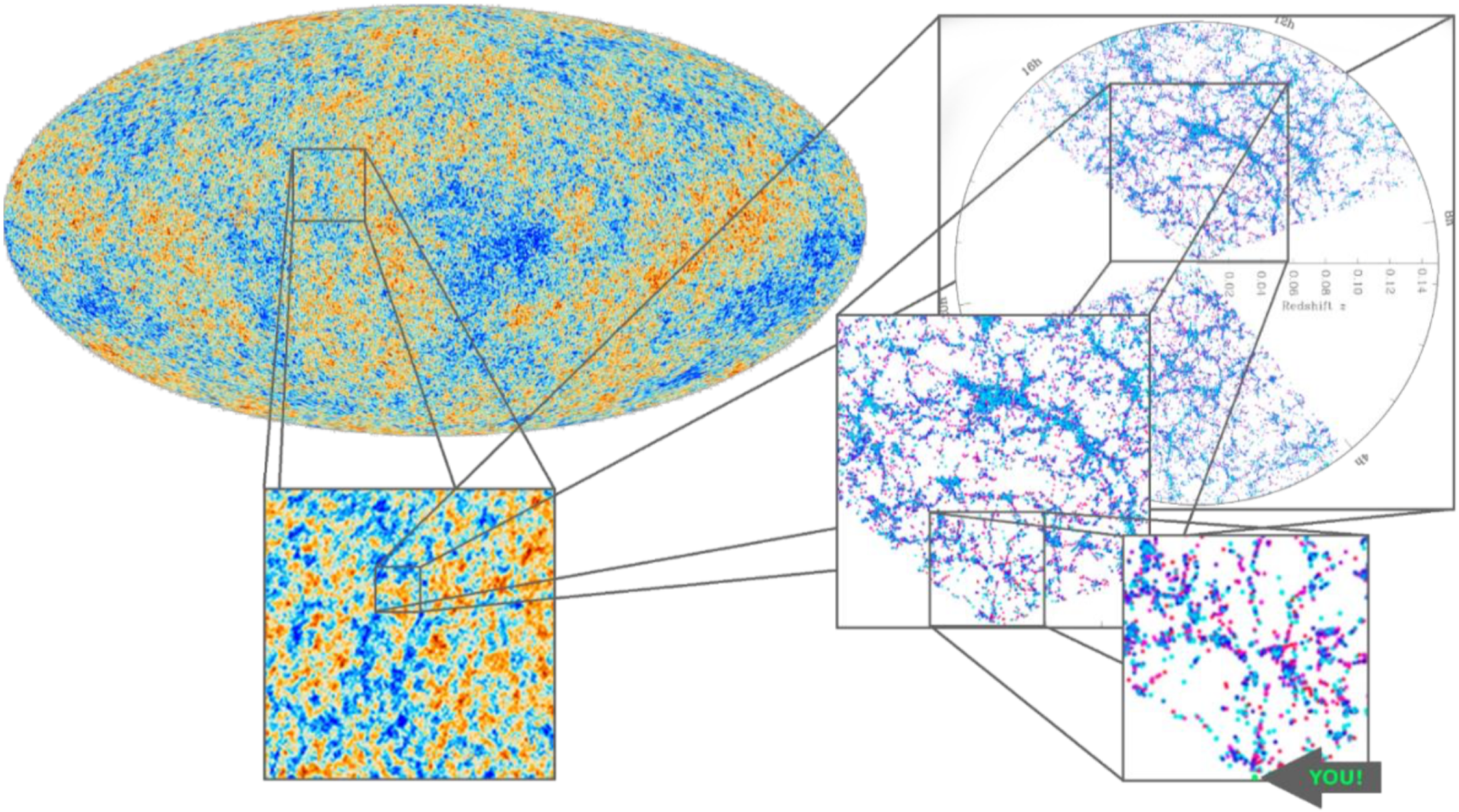


Millea, Anderes & Wandelt, in prep.

Beyond the CMB

Benjamin Wandelt

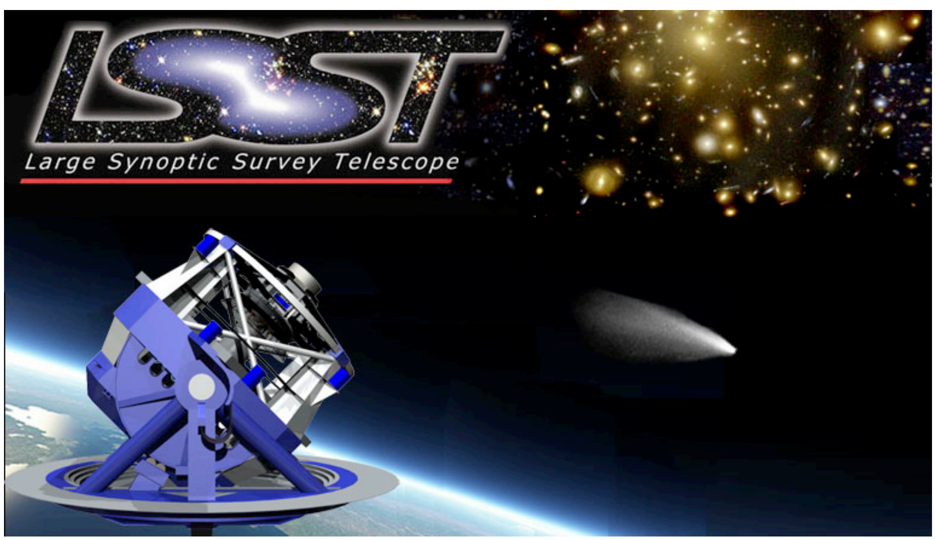
Large Scale Structure: many more modes…



… in the golden age of surveys

EUCLID
MAPPING THE GEOMETRY OF THE DARK UNIVERSE

MegaPrime
CFH
MegaCam
Legacy Survey
CFHTLS



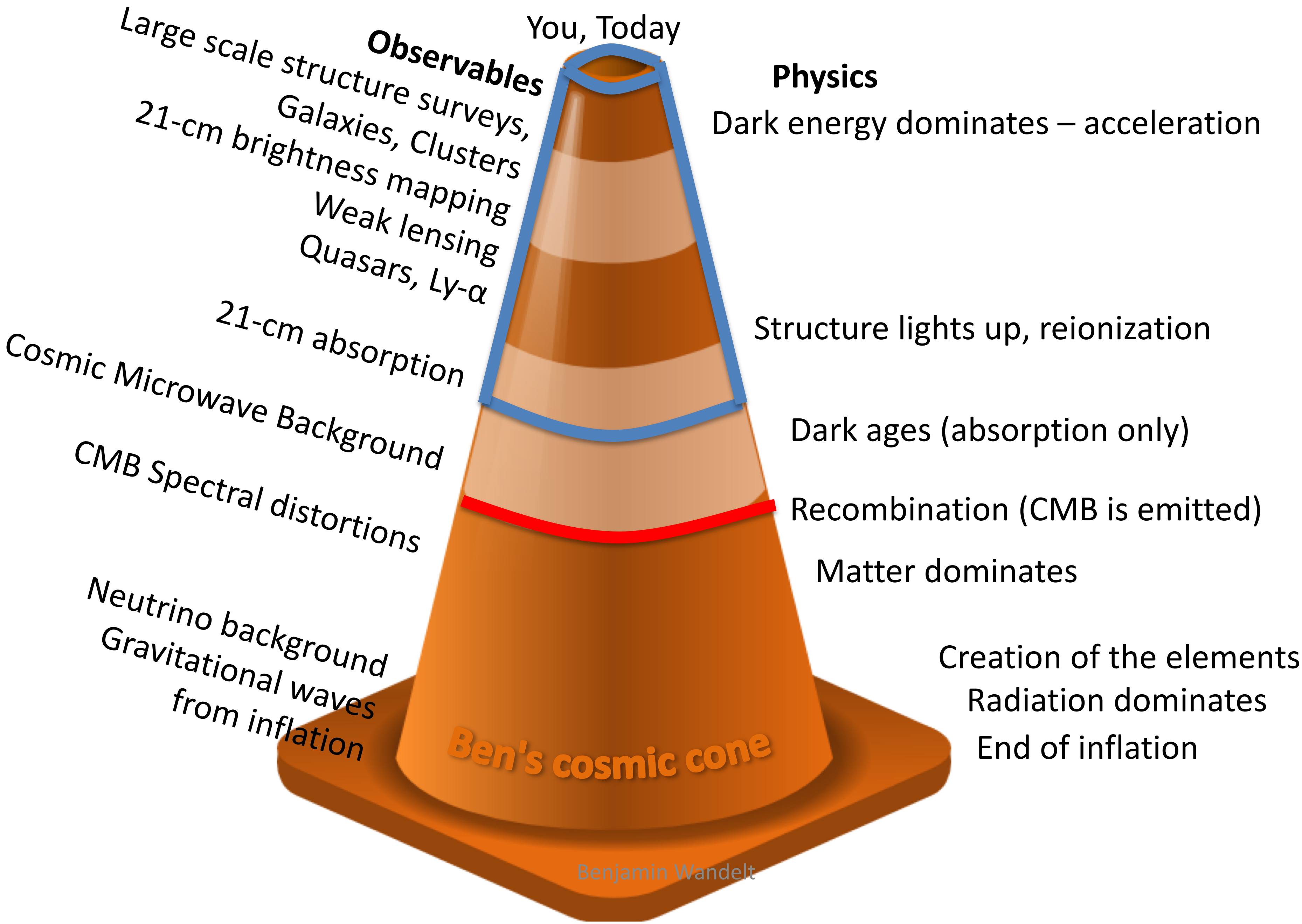
THE DARK ENERGY SURVEY

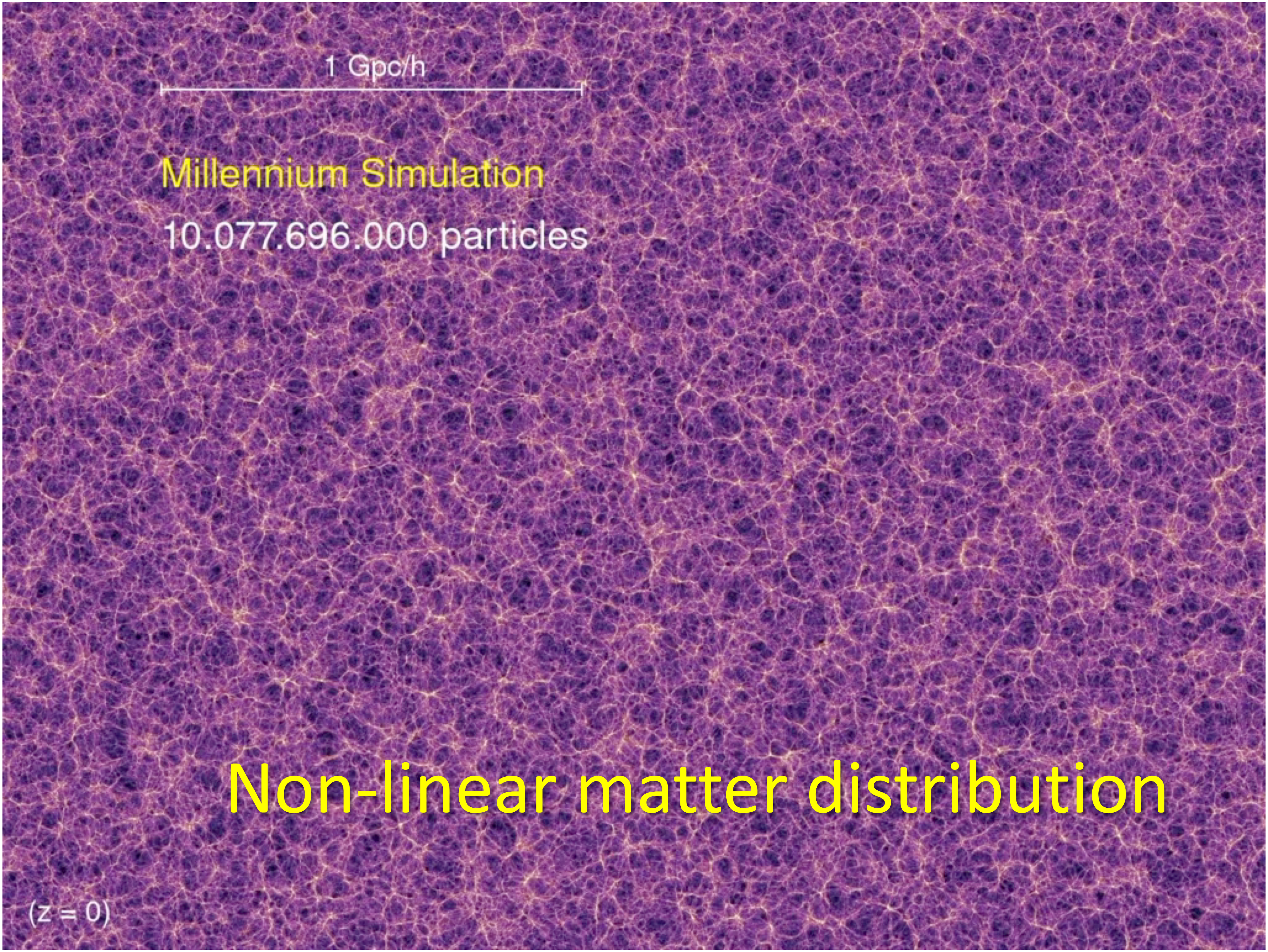
HYPER SUPRIME-CAM

WFIRST

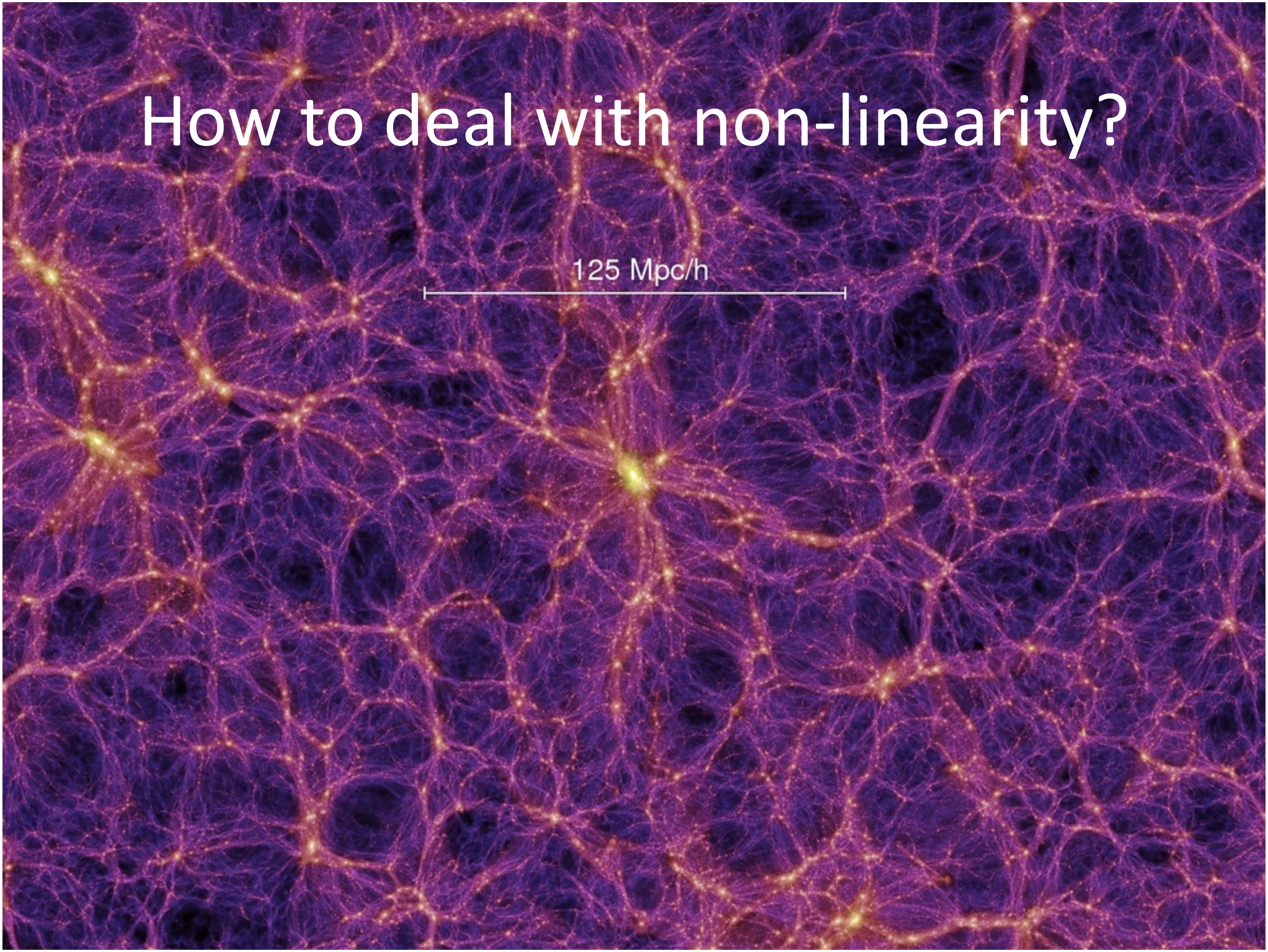
Your favorite survey here

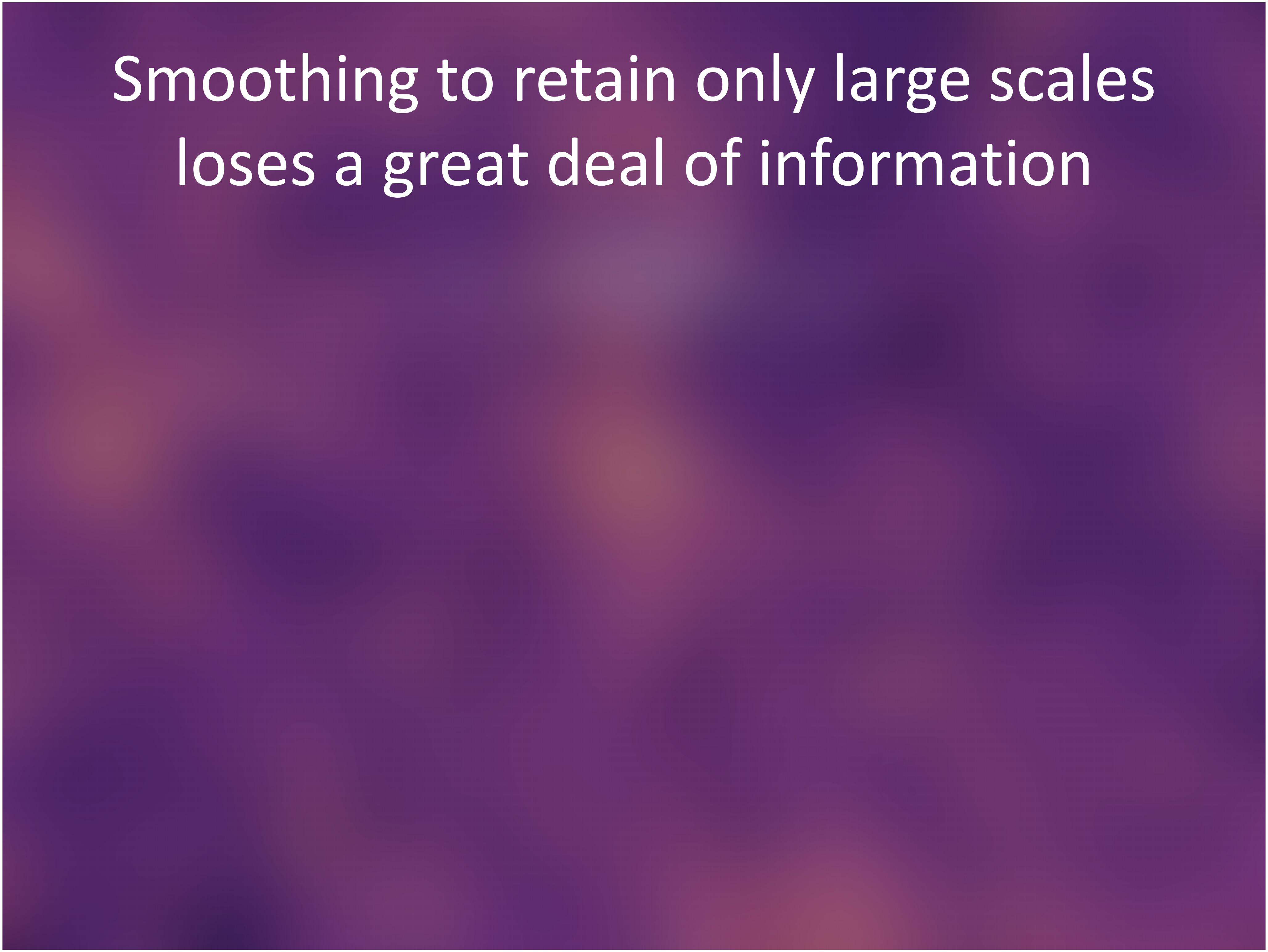
The promise of Large Scale Structure surveys




1 Gpc/h
Millennium Simulation
10.077.696.000 particles
Non-linear matter distribution
(z = 0)

How to deal with non-linearity?




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

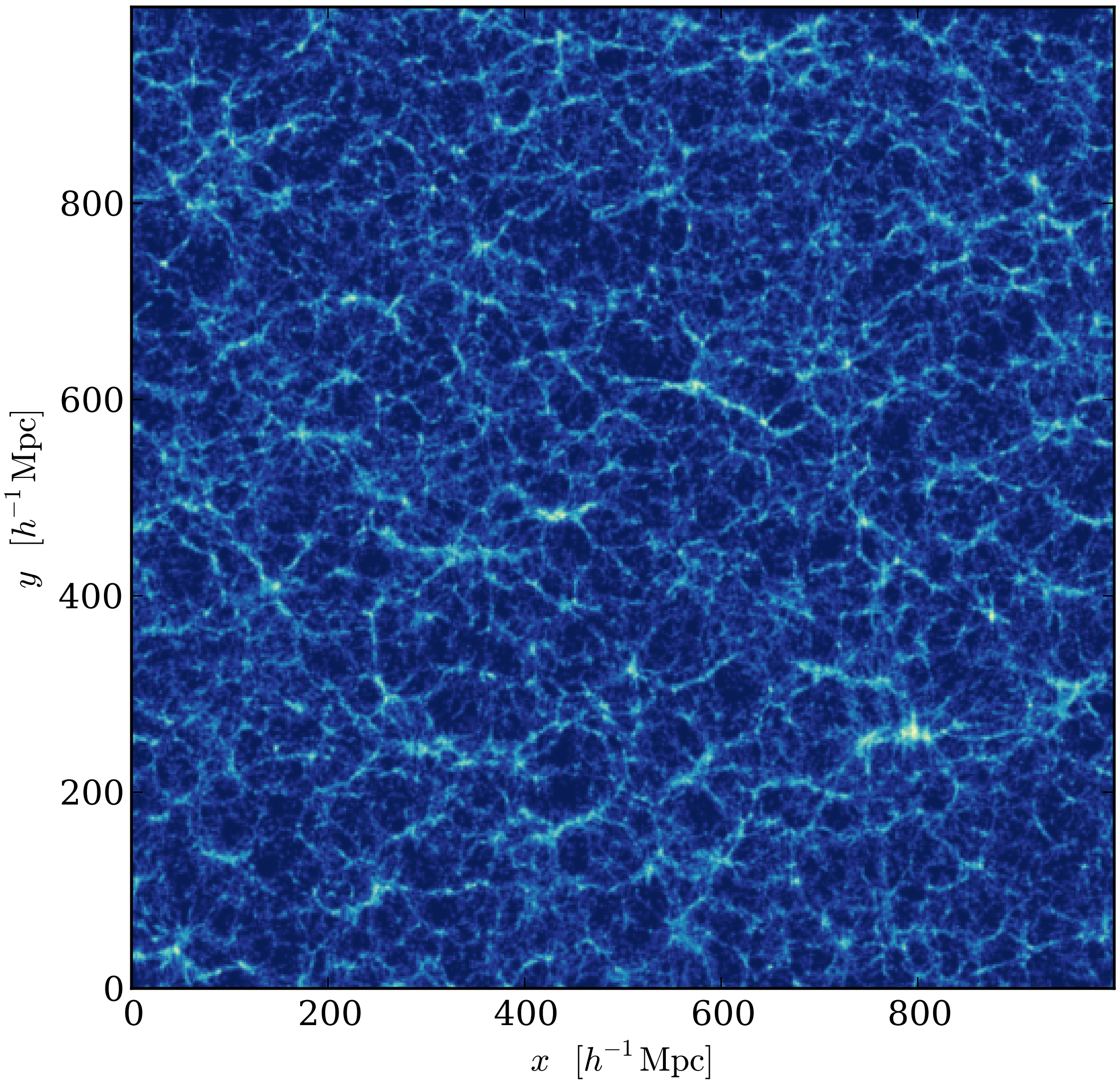
The large scale structure challenge

- ***Problem: to access the small scales need to deal with non-linearity and "bias"***

- Possible approaches:
  - **Avoid:** Observe at high redshift before density contrast became non-linear (CMB, 21cm, Ly-alpha?)
  - **Adapt:** Focus on less complicated parts of the Universe, e.g. those that retain more memory of the initial conditions: cosmic voids
  - **Attack:** Physical & statistical forward model of the survey, bias, etc. (perturbative or non-linear)

Adapt

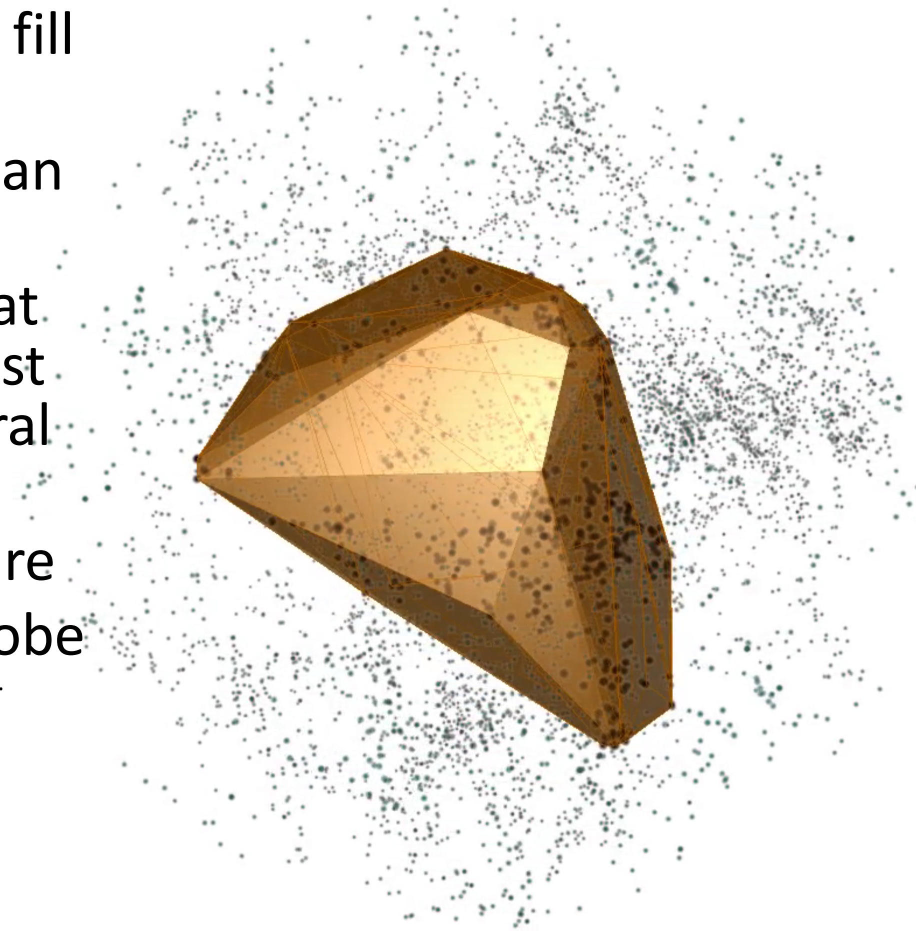
How do we summarize this?
In terms of the soap or the bubbles?



Benjamin Wandelt

The power of cosmic voids

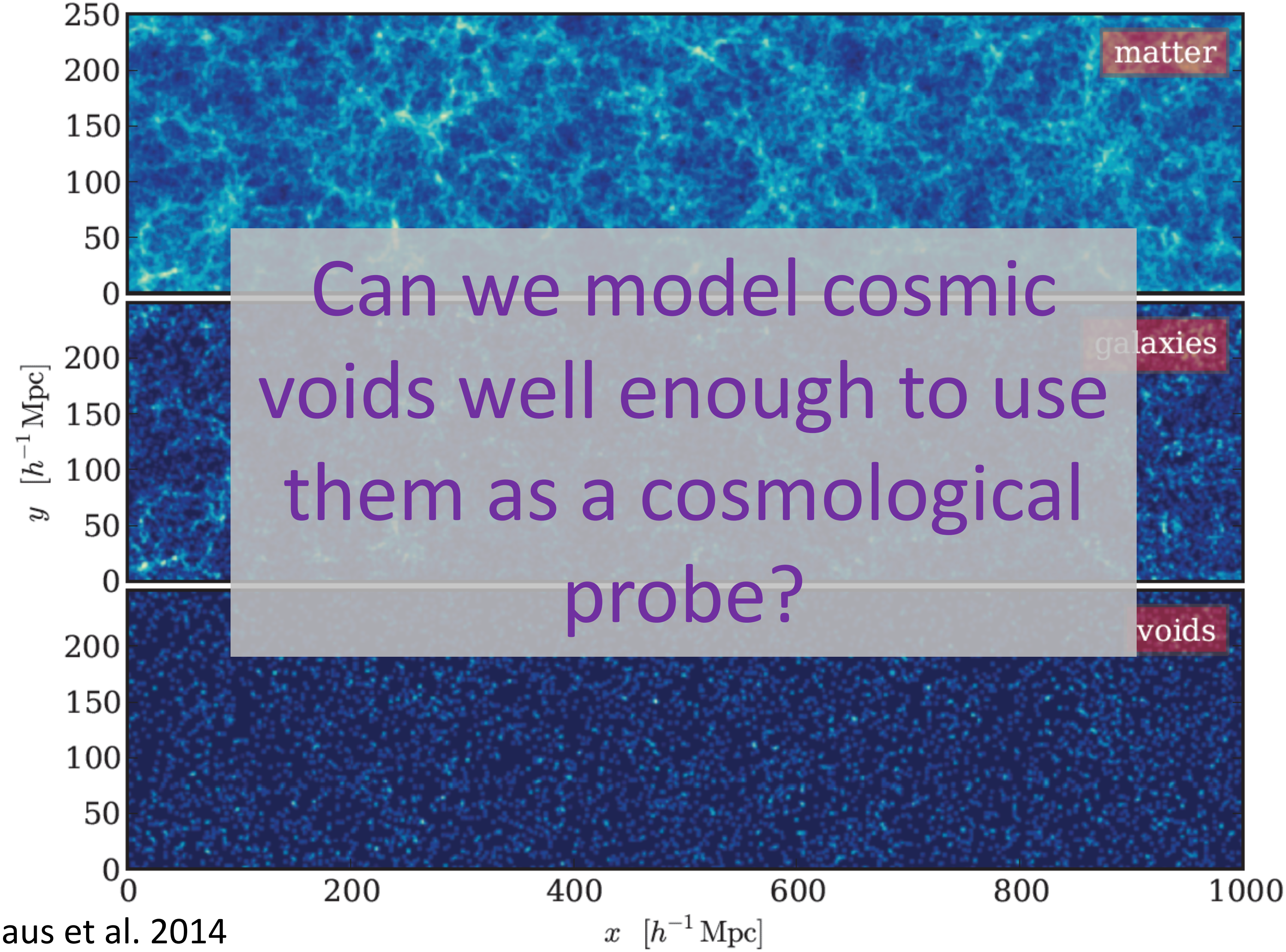
- Biggest "objects" in the Universe – fill most of the volume!
- Simpler dynamics and evolution than high density regions
- The first regions in the universe that are dominated by dark energy; most sensitive to modifications of General Relativity
- Sensitive to small scale DM structure
- A free, additional observational probe in current and future surveys: ~10^4 voids in Euclid!



Several active groups and a rapidly growing body of work

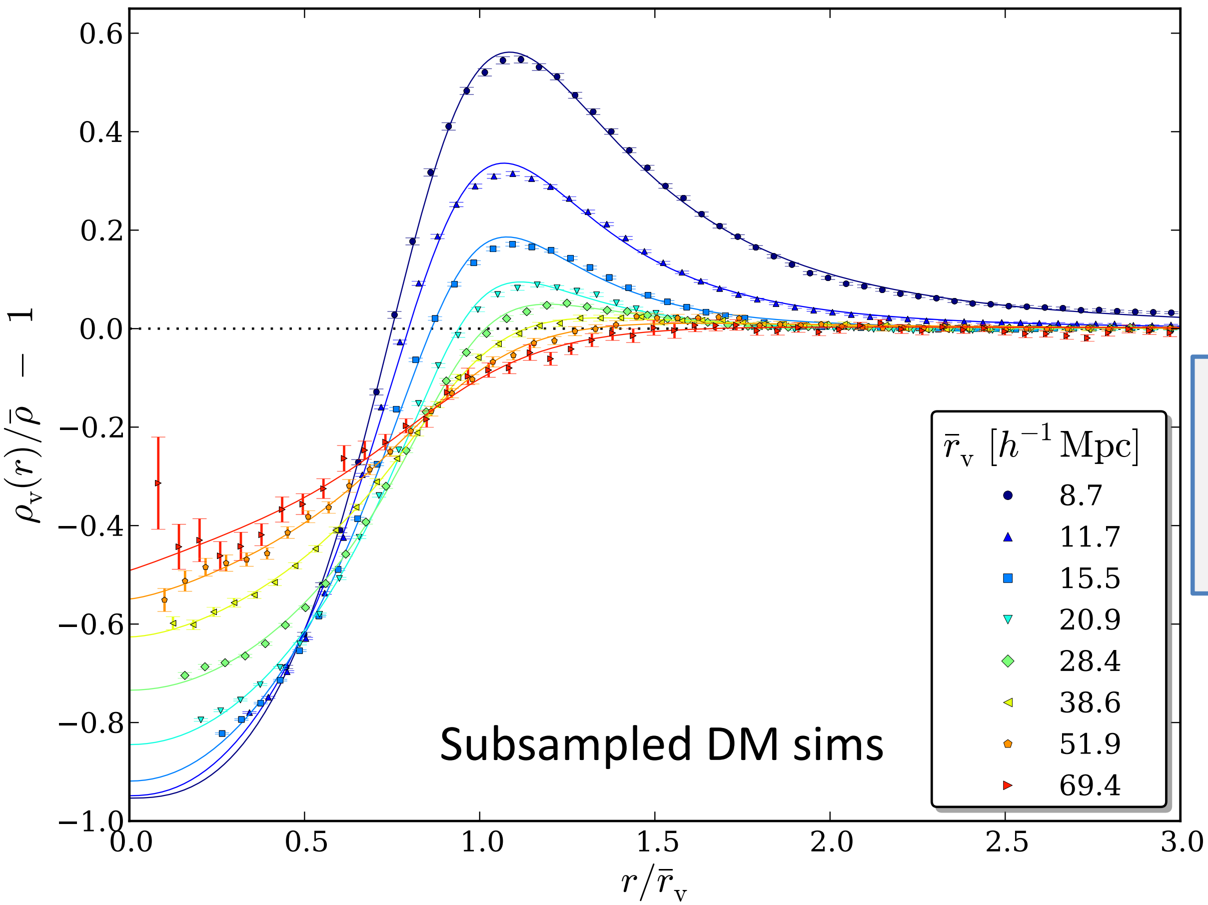
Google "VIDE bitbucket"

Matter, galaxies, voids in simulation



Hamaus et al. 2014

A universal profile for voids



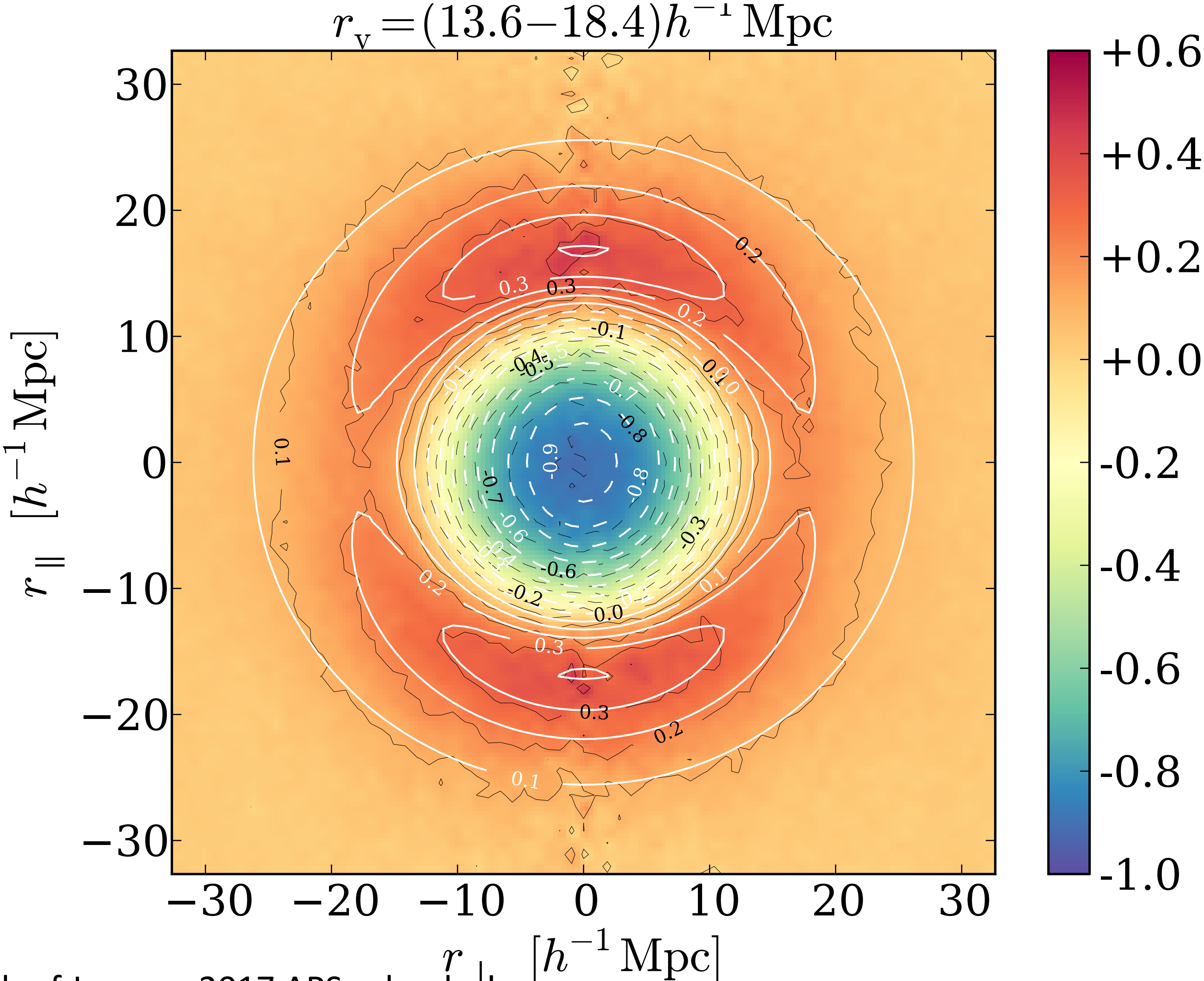
$$\frac{\rho_v(r)}{\bar{\rho}} - 1 = \delta_c \frac{1 - (r/r_s)^\alpha}{1 + (r/r_v)^\beta}$$

Scaling properties of voids allow reduction from 4 to 2 params

NL velocity profile matches this NL profile!

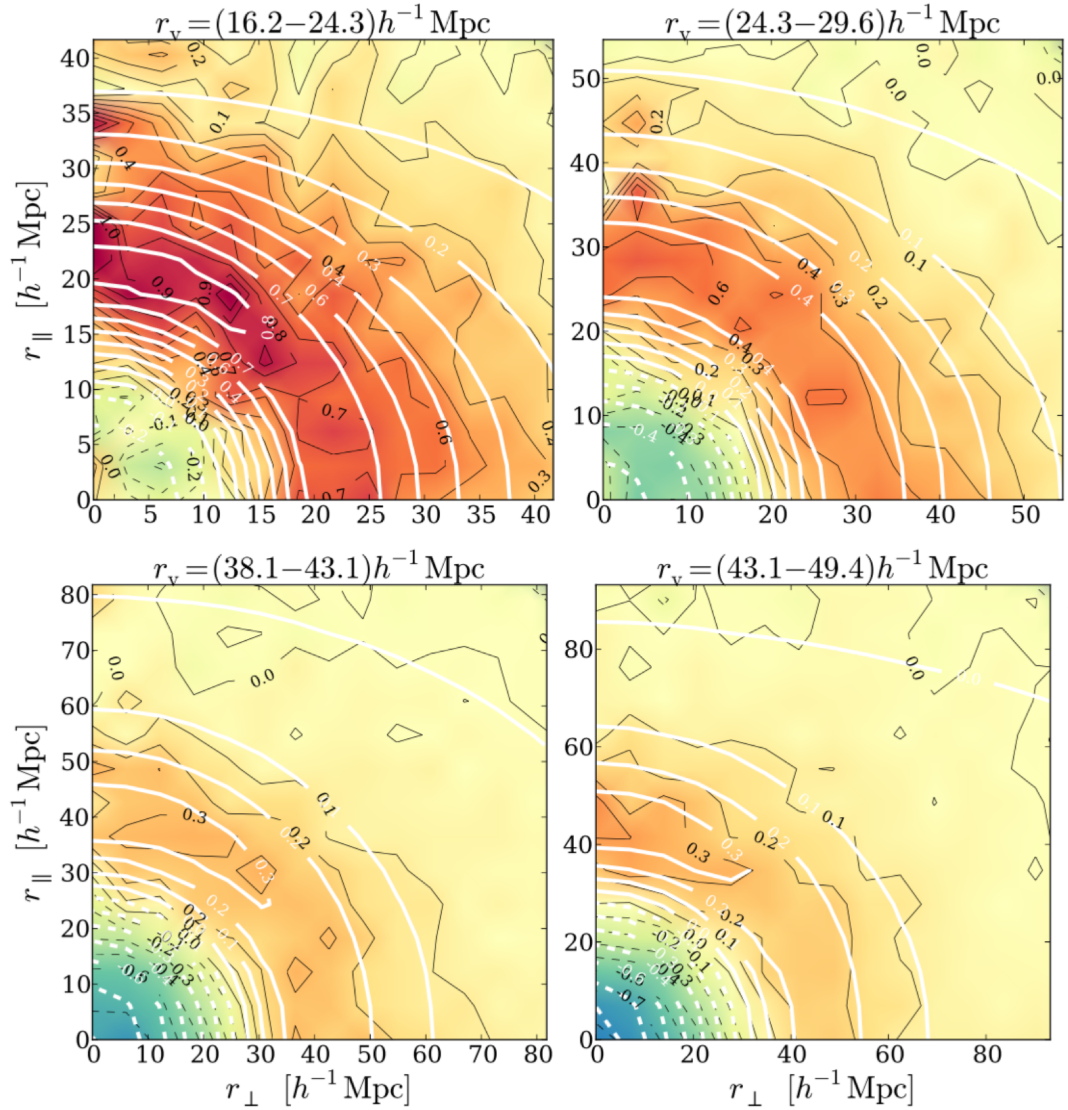
Hamaus, Sutter & Wandelt PRL 2014, arxiv:1403.5499

Void-galaxy correlation function in redshift space

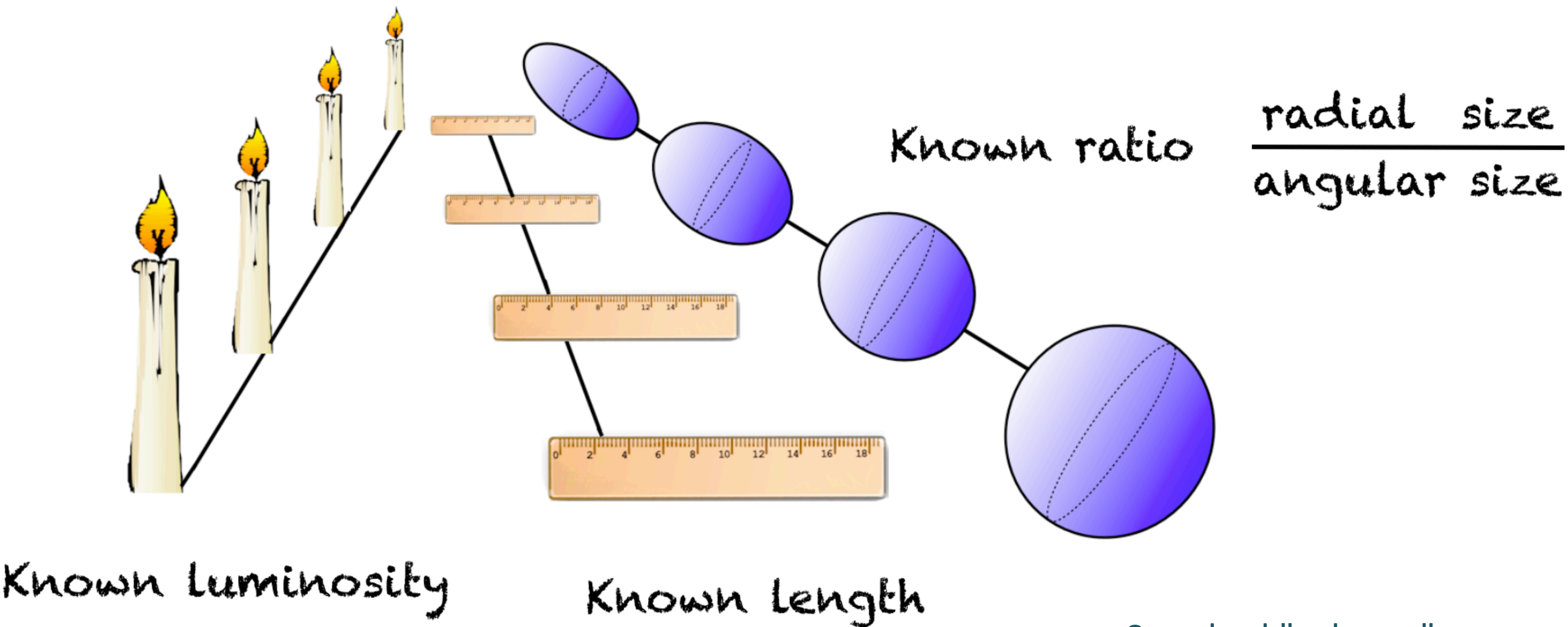


See month of June on 2017 APS calender!

Void-galaxy correlation function in SDSS DR11 CMASS



Cosmography with the Alcock-Pazcynski test



Standard "candles"

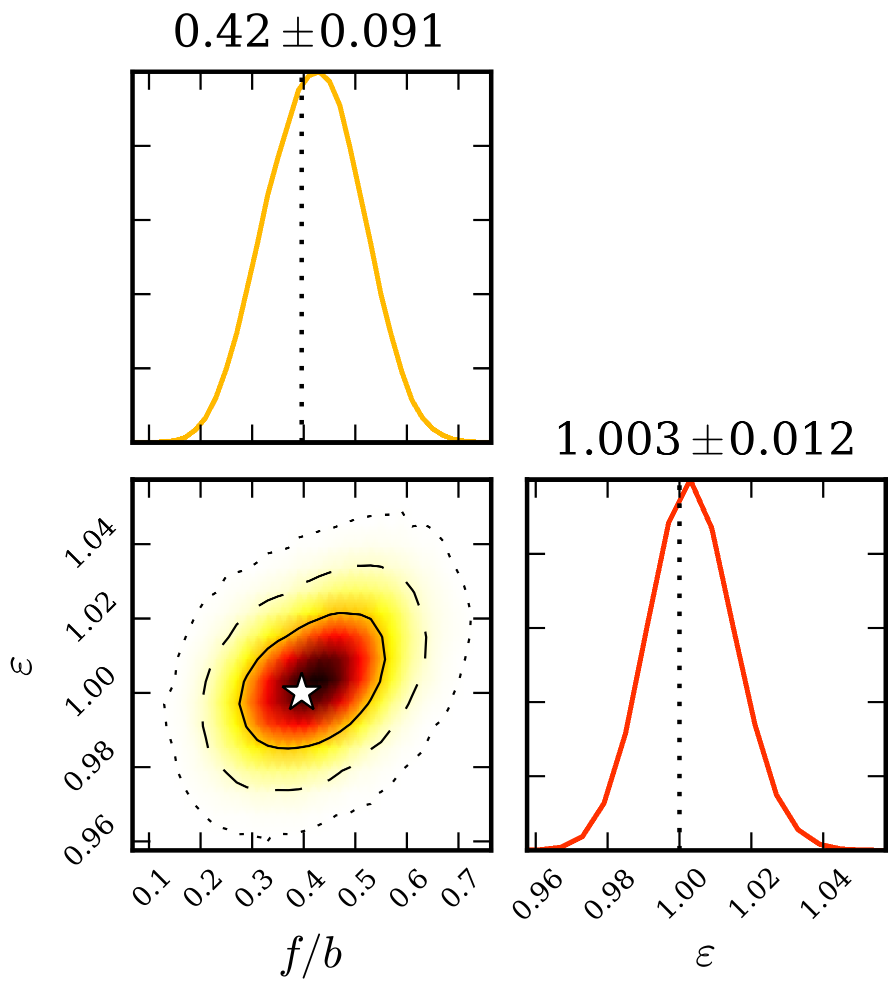
Standard "rulers"

Standard "spheres"

Joint measurement of growth of structure and of expansion geometry

Using BOSS data. AP measurement is 4 times tighter than RSD from SDSS DR12 galaxy clustering analysis!

(Gil-Marin *et al.* arXiv:1509.06386)

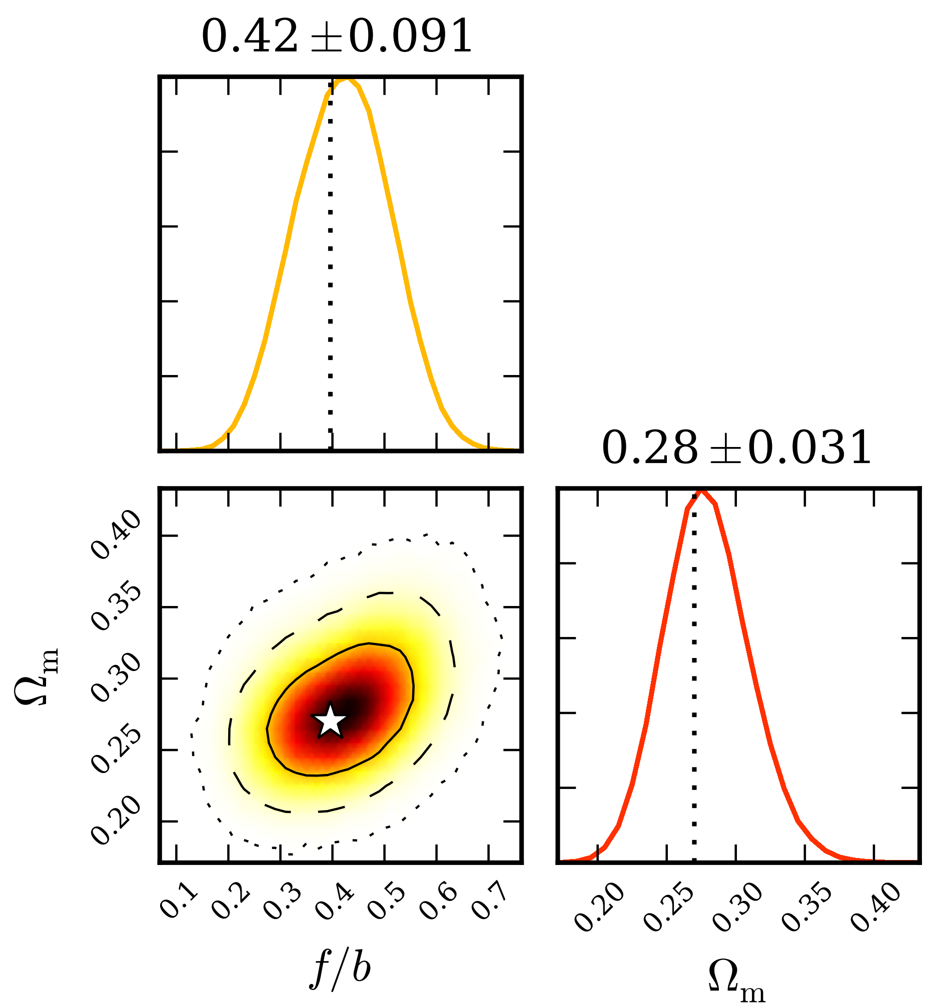


Preliminary Euclid forecasts => 30 times higher Figure of merit than standard BAO

Hamaus et al. arXiv:1602.01784

Joint constraint on growth f/b and on matter density

Assumes LCDM



Hamaus et al. arXiv:1602.01784

Bright future for cosmic voids

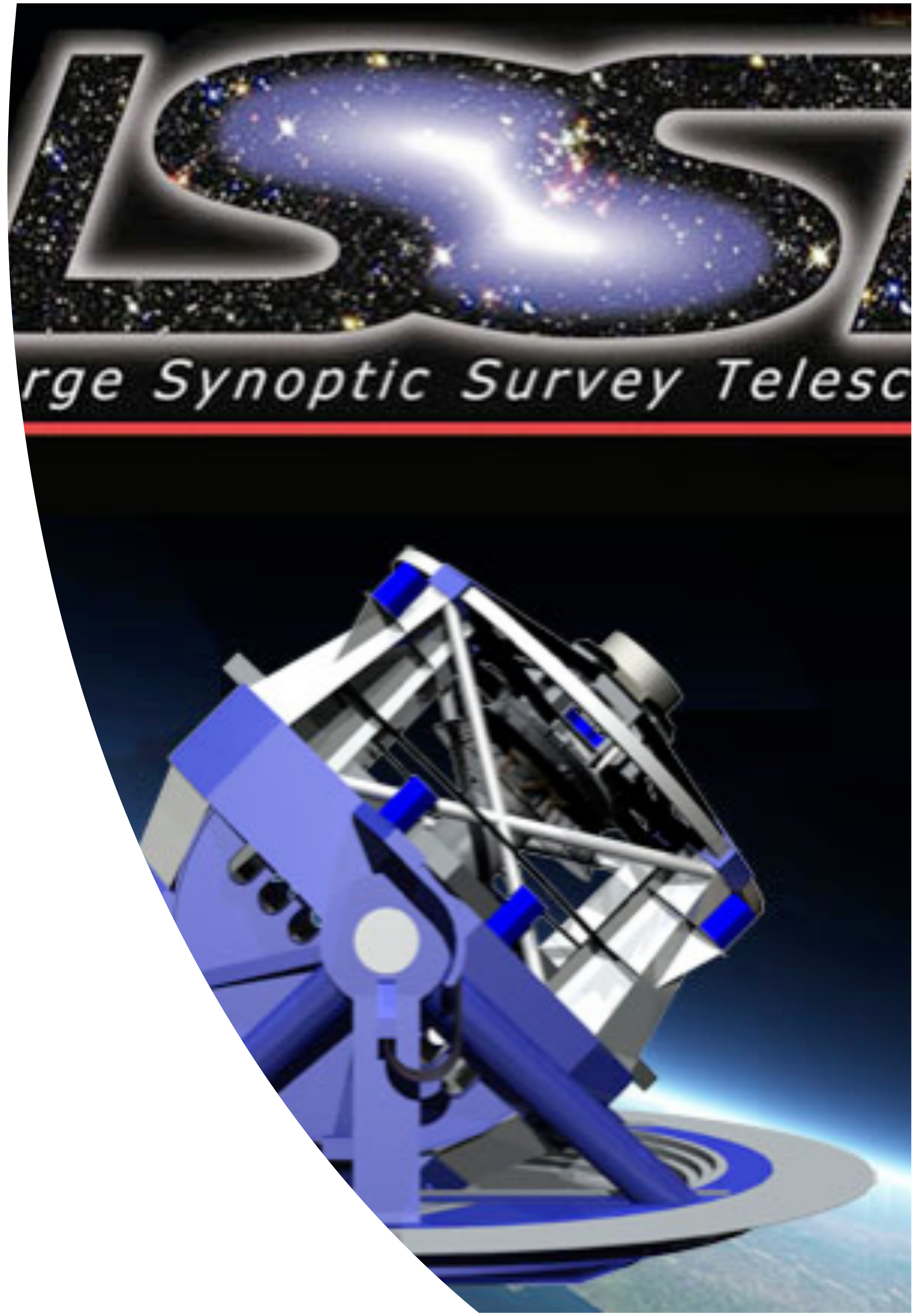
- Voids are not empty! Substructure in voids
  - Probes nature of dark matter
  - Contains information about small-scale primordial spectrum frozen in
  - Gives additional dynamical information
  - Contain very differently biased tracers

Next generation supernova cosmology

- Upcoming surveys will have tens of thousands of supernovae
- But there are too many supernovae to follow all of them up spectroscopically. Photometric information leads to type contamination and photo-z systematics.

- **How to do accurate cosmology with large number of supernovae?**

S. Mukherjee and B. Wandelt, in prep.



Accurate cosmology with supernova-galaxy cross-correlations

Idea: both supernovae and galaxies are biased tracers of the density field.

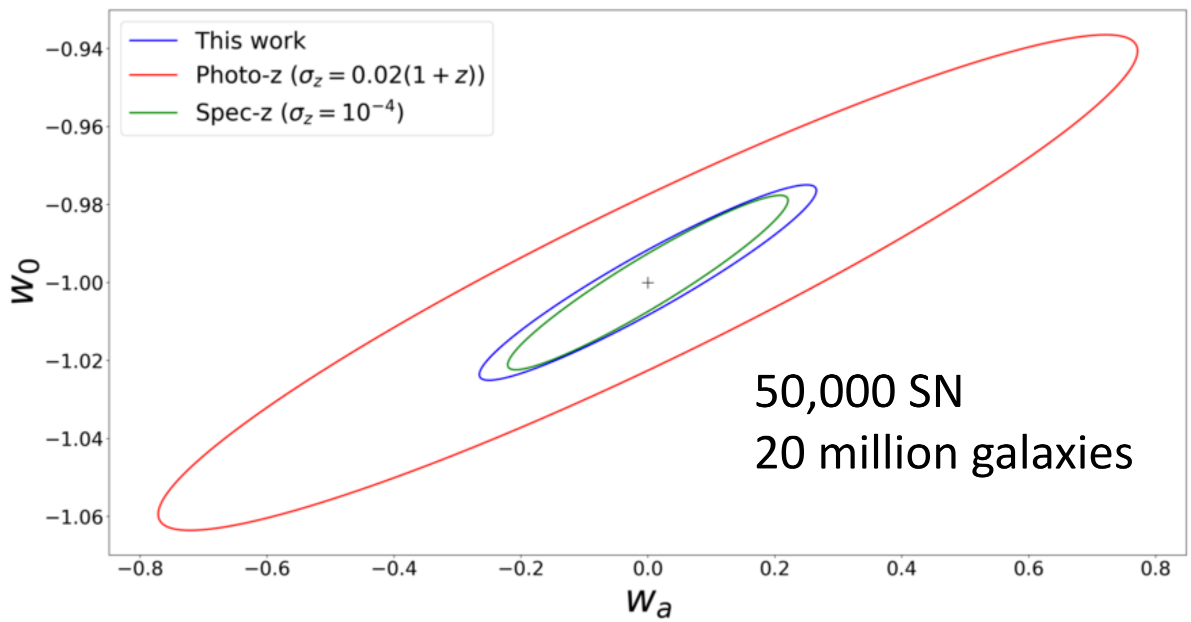
Their correlation is a function of the cosmological parameters.

Estimate cosmological parameters by maximizing the correlation!

S. Mukherjee and B. Wandelt, in prep.

Constraining the luminosity distance redshift relation using SN-galaxy cross correlations

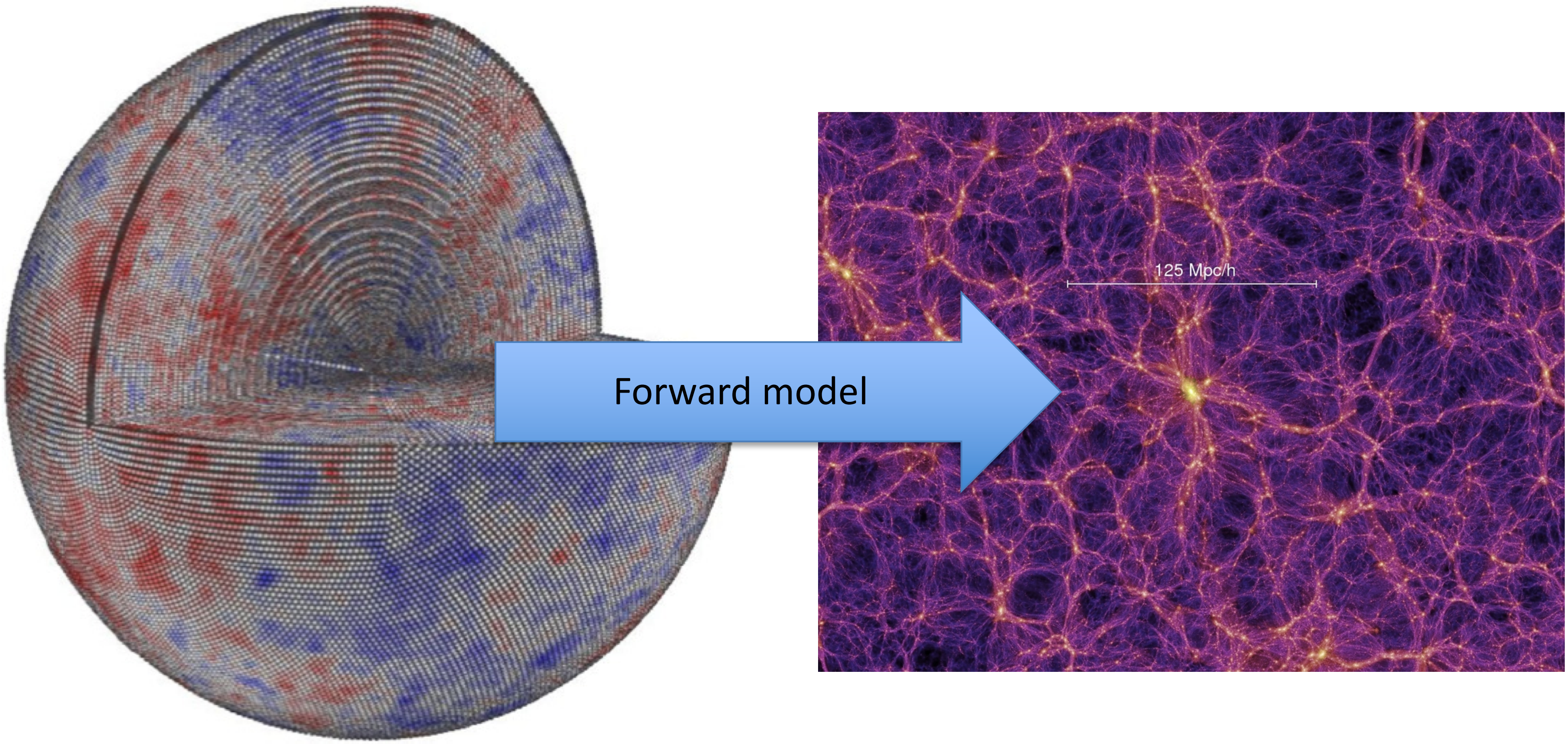
- Robust to type contamination
- Completely Insensitive to photo-z systematics



S. Mukherjee and B. Wandelt, in prep.

Attack

Going from "Adapt" to "Attack" mode: can we fit a full forward model of Lambda CDM to galaxy surveys?



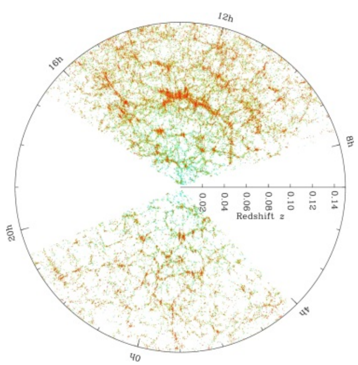
A fully *probabilistic* model of galaxy surveys

BORG: *Bayesian Origin Reconstruction from Galaxies*

- Gaussian prior + ***Gravity*** + likelihood for galaxies

  (includes survey model, bias model, automatic noise level calibration, selection function, mask, …)

- Hamiltonian Markov Chain Monte Carlo method



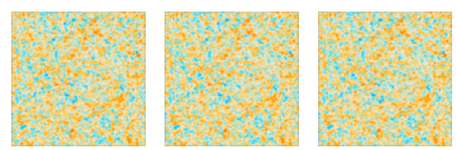
Observations

(galaxy catalog + meta-data: selection functions, completeness…)

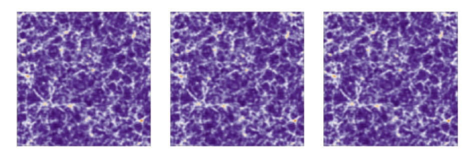
→ BORG →

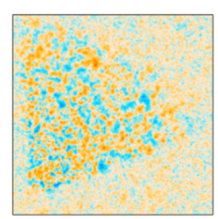
E.g. inferred dark matter densities

z=100

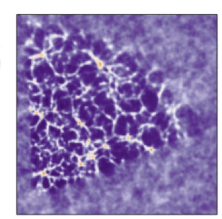


z=0





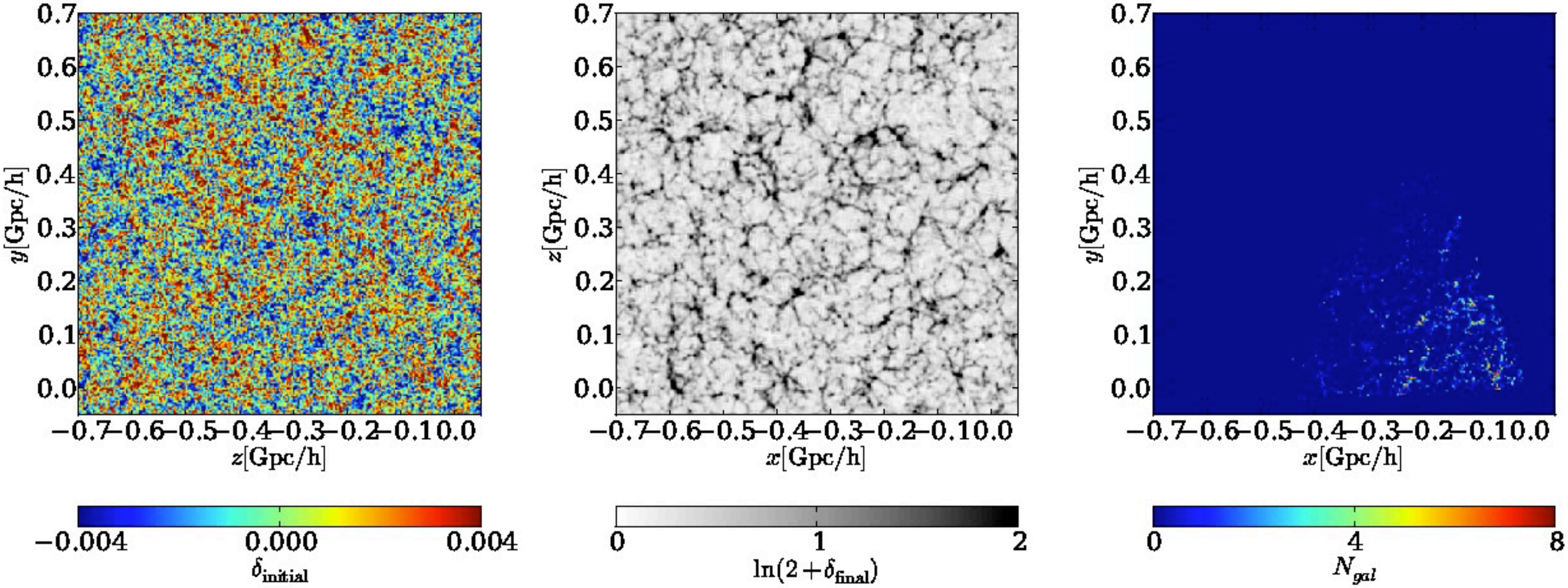
Summaries with quantified uncertainties



Jasche & Wandelt 2013, arXiv:1203.3639

Jasche, Leclercq & Wandelt 2015, arXiv:1409.6308

Bayesian LSS sampling - movie



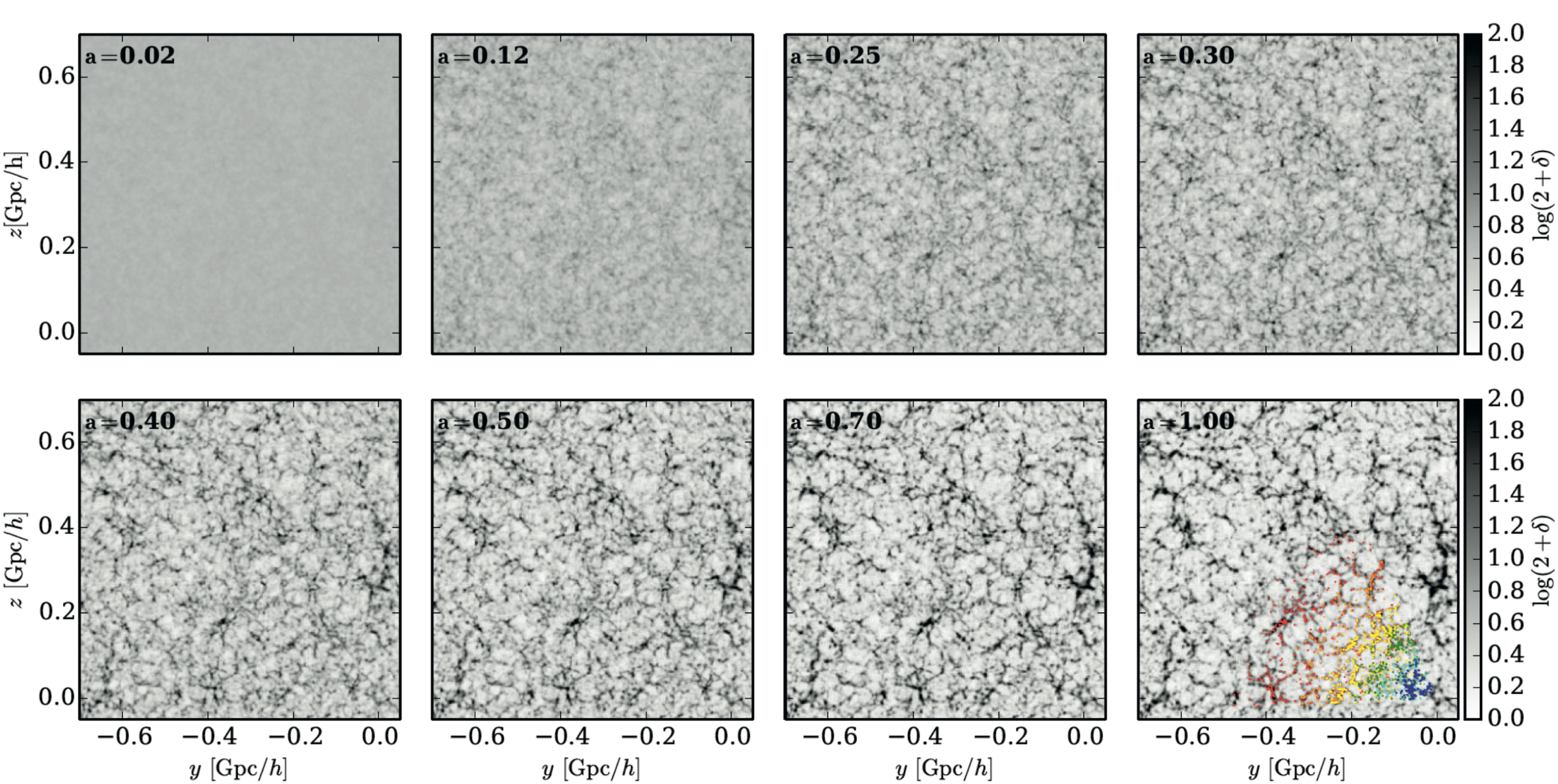
Initial conditions

Final conditions

Observations

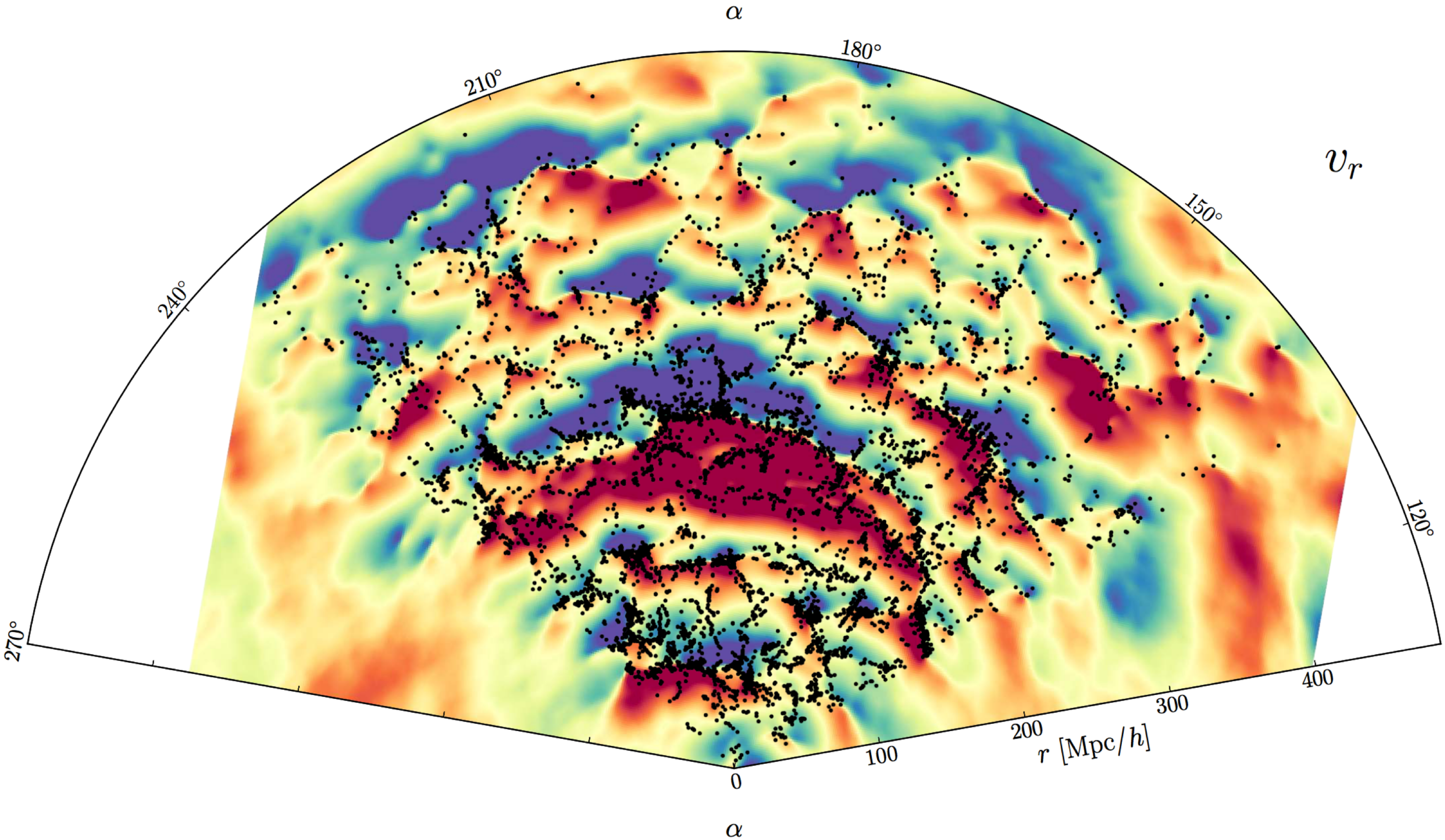
Jasche, Leclercq & Wandelt 2014, arXiv:1409.6308

A posterior sample of the formation history of our Universe



Jasche, Leclercq & BDW 2014, arXiv:1409.6308

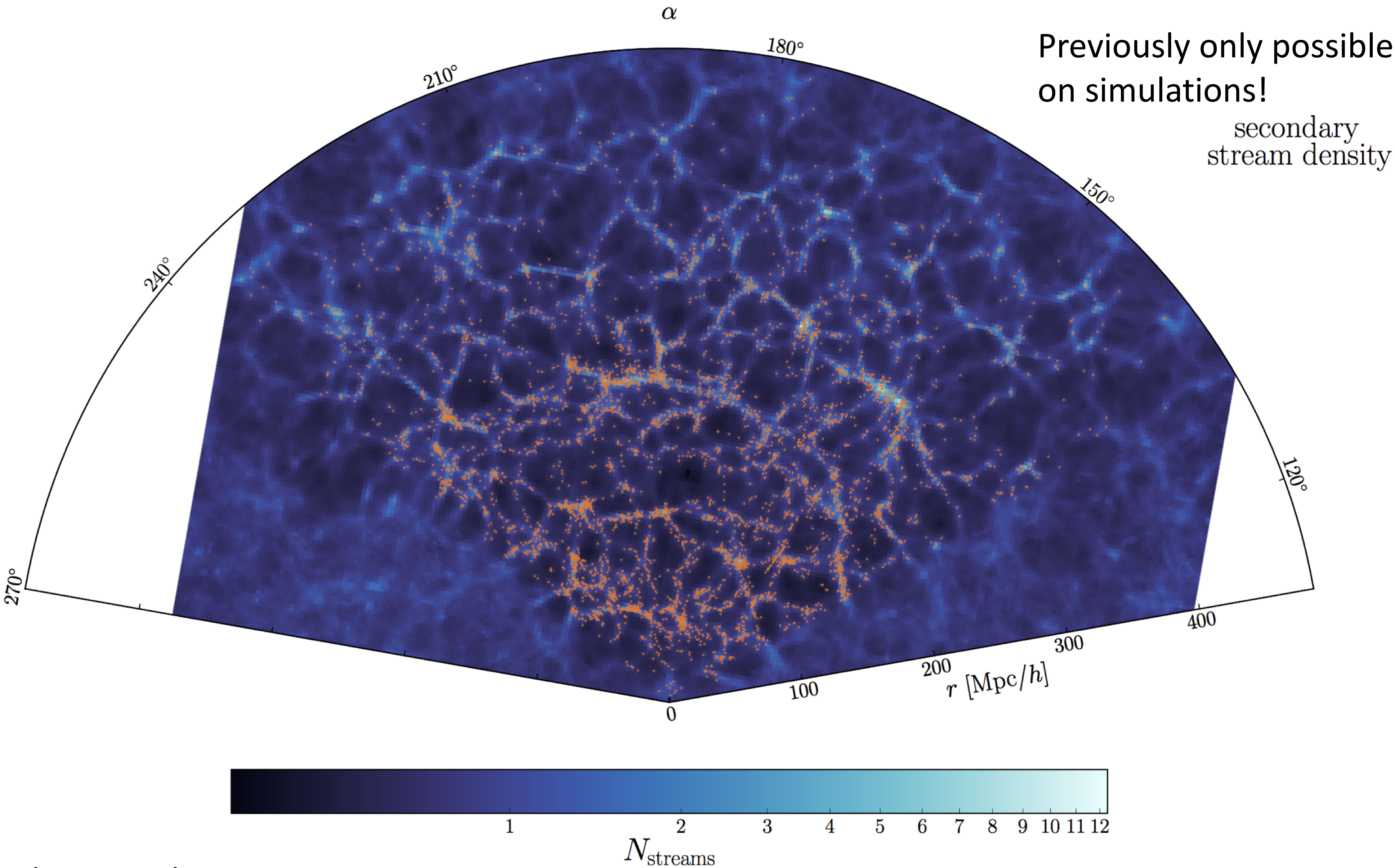
Bayesian LCDM predictions: Dynamical velocities



Leclercq et al. 2017

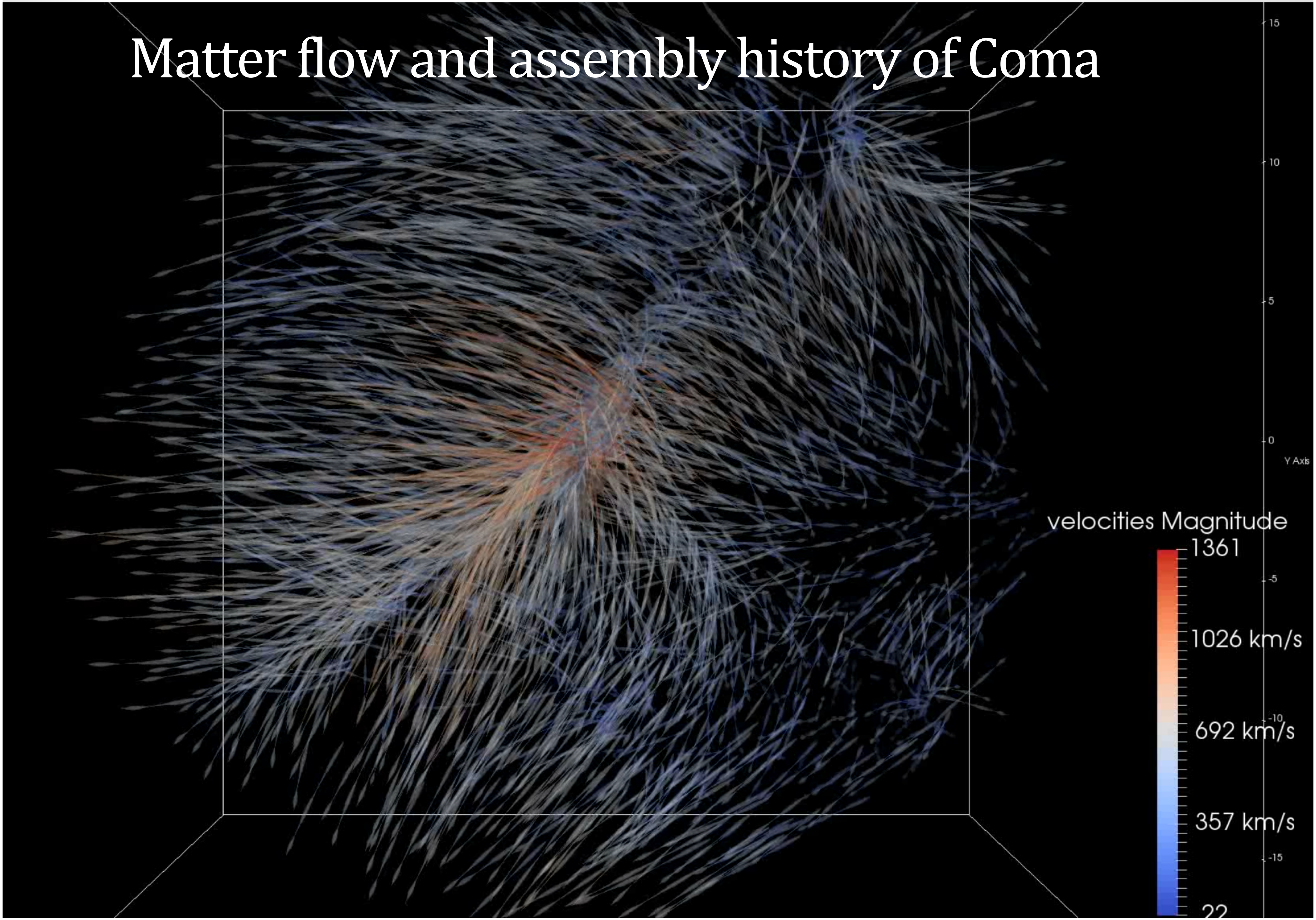
Posterior mean of Lagrangian stream density

Previously only possible on simulations!



Leclercq et al. 2017

Matter flow and assembly history of Coma



A next generation approach to “Attack mode:” fully simulation-based inference

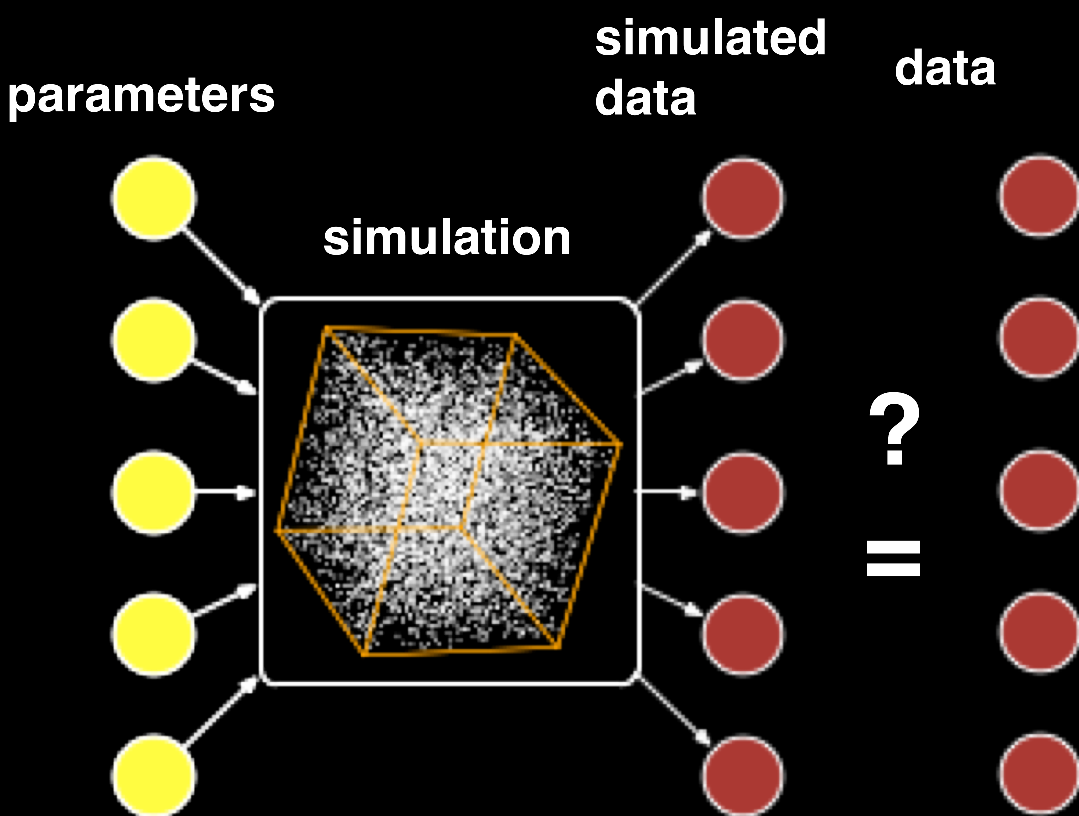
With Justin Alsing, Tom Charnock, Guilhem Lavaux, Francisco Villaescusa, Stephen Feeney

What if we can only do simulations?

$$P(\boldsymbol{\theta}|\mathbf{d}) = \frac{P(\mathbf{d}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\mathbf{d})}$$

$$\mathbf{d}^* \leftarrow \mathrm{simulation}(\mathbf{d}^*|\boldsymbol{\theta})$$

Likelihood-free inference 101



Draw from prior:

$$\boldsymbol{\theta} \leftarrow P(\boldsymbol{\theta})$$

Simulate data:

$$\mathbf{d}^* \leftarrow P(\mathbf{d}^*|\boldsymbol{\theta})$$

If $\rho(\mathbf{d}^*, \mathbf{d}) < \epsilon$

accept;

else:

reject;

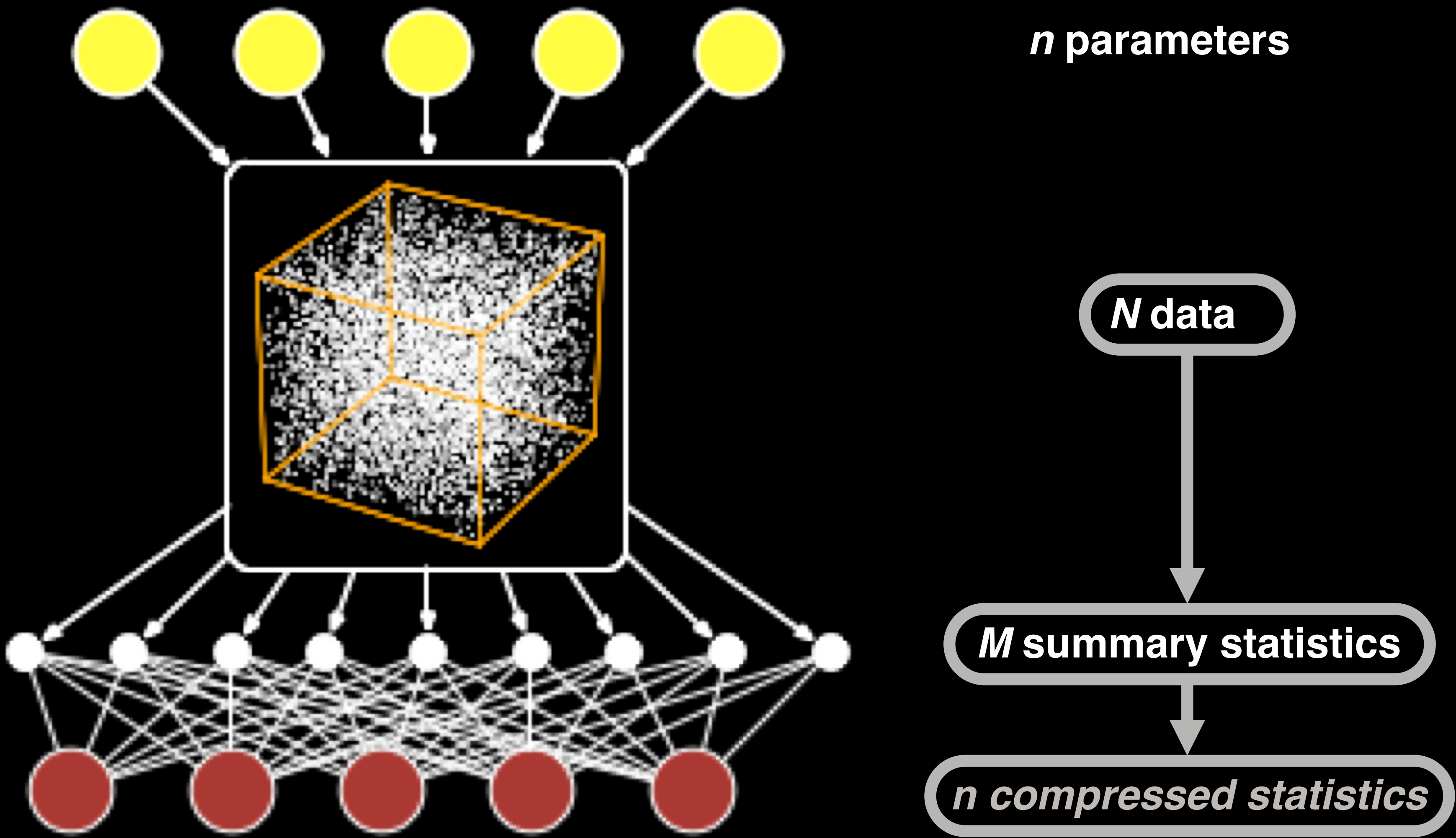
In the limit $\epsilon \to 0, \ \{\boldsymbol{\theta}\} \leftarrow P(\boldsymbol{\theta}|\mathbf{d})$

Likelihood-free inference 101

How to reduce data-space?

How to explore parameter-space?

Massive data compression



Massive data compression

Fisher information

$$\mathbf{F} \equiv -\mathbb{E}_{\boldsymbol{\theta}}(\nabla\nabla^T\mathcal{L})$$

Information inequality

$$\mathbb{V}_{\boldsymbol{\theta}}(t_\alpha) \geq [\nabla\mathbb{E}_{\boldsymbol{\theta}}(\mathbf{t})^T\mathbf{F}^{-1}\nabla\mathbb{E}_{\boldsymbol{\theta}}(\mathbf{t})]_{\alpha\alpha}$$

Can derive *n* compressed quantities that contain all the Fisher information!

Alsing & Wandelt arXiv:1712.00012

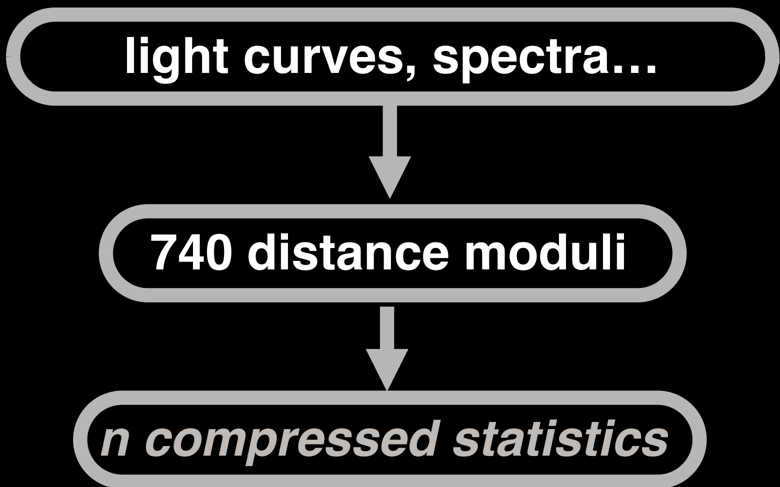
Exploration of parameter space

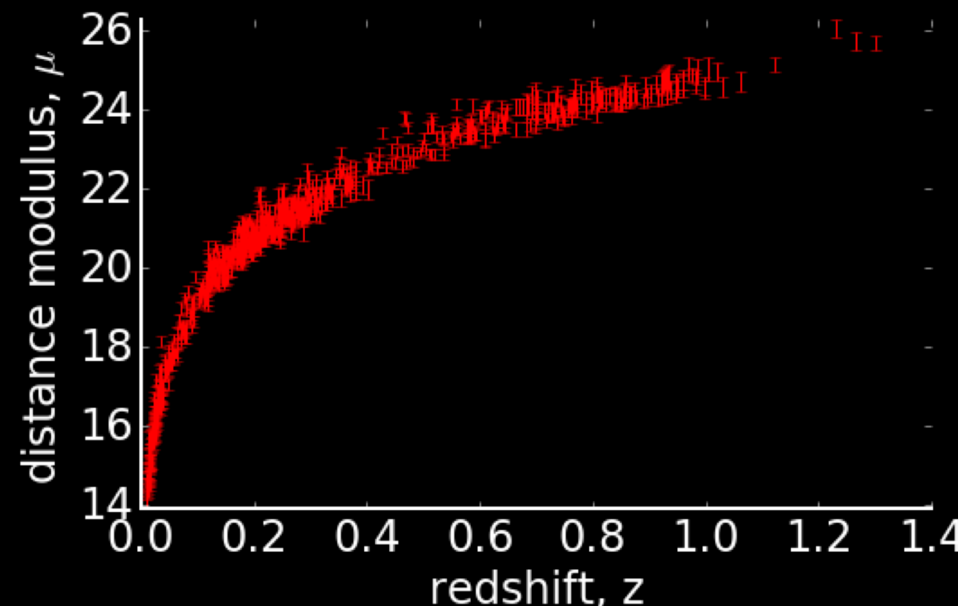
Density estimation Likelihood free inference (DELFI)

Learn *joint* probability density of parameters and compressed data using a Gaussian Mixture Model

Alsing, Feeney &Wandelt arXiv: 1801.01497

Case study: JLA SNe





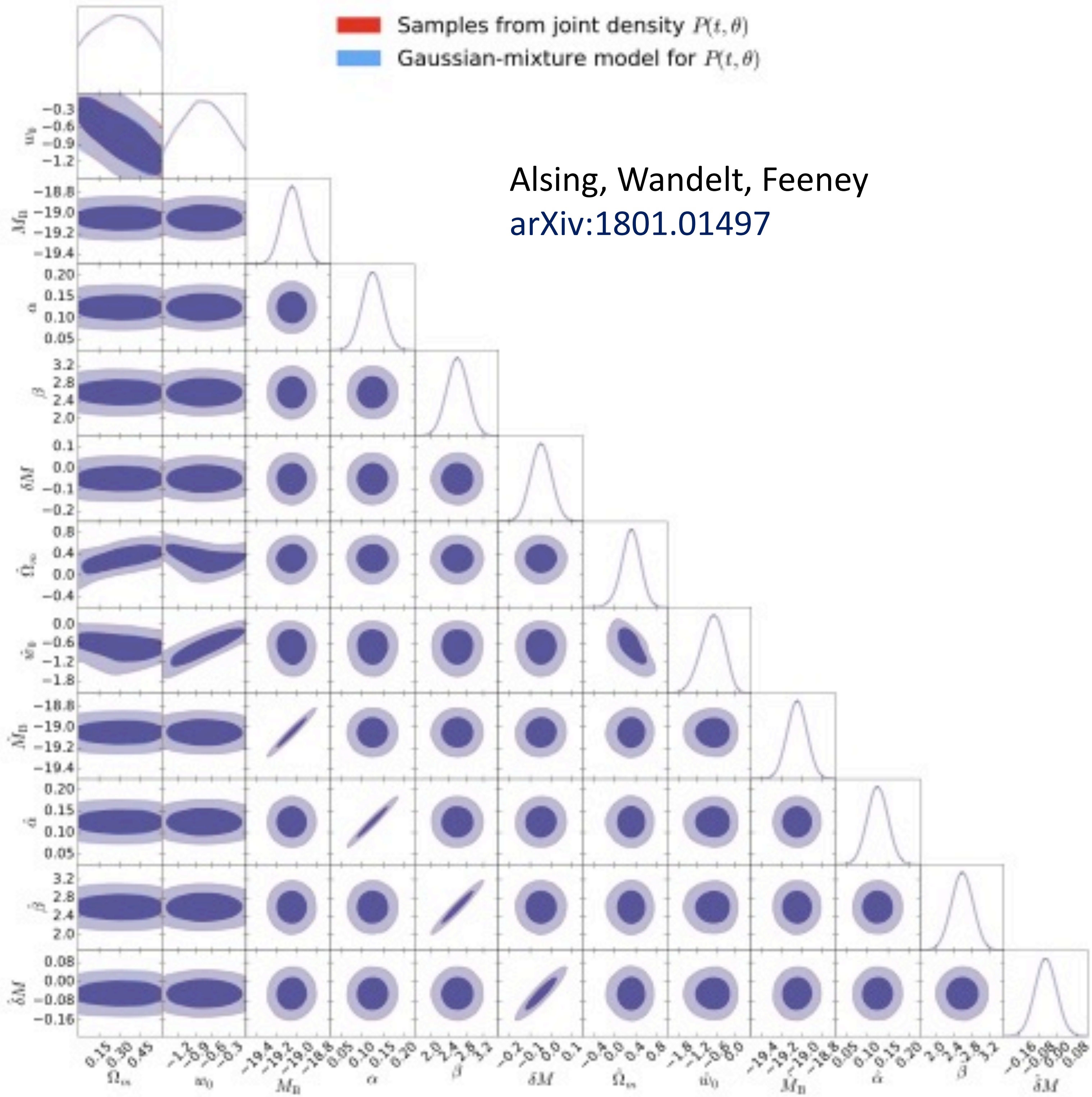
$$\mathcal{L} = -\frac{1}{2}(\mathbf{d} - \boldsymbol{\mu})^T \mathbf{C}^{-1} (\mathbf{d} - \boldsymbol{\mu}) - \frac{1}{2} \ln|\mathbf{C}|$$

$$\boldsymbol{\mu} = \boldsymbol{\mu}(\Omega_{\mathrm{m}}, w_0, M, \alpha, \beta, \delta m)$$

$$\mathbf{C} = \mathbf{C}(\alpha, \beta)$$

Betoule et al 2014

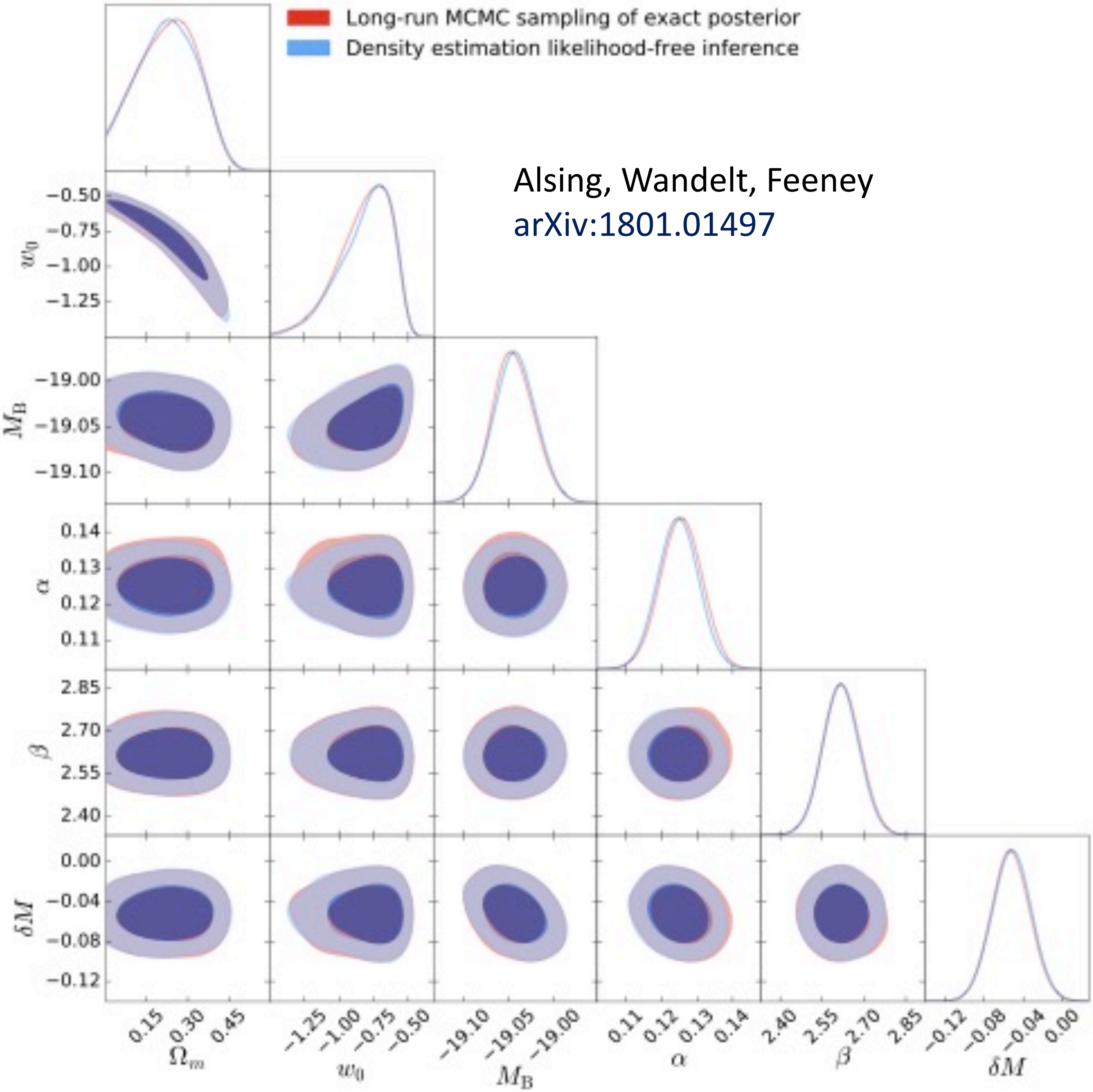
Fit to joint density $P(\boldsymbol{\theta}, \mathrm{d}^*)$



Alsing, Wandelt, Feeney
arXiv:1801.01497

(10000 simulations)

Posterior inference



Alsing, Wandelt, Feeney
arXiv:1801.01497

FLATIRON INSTITUTE
Division of Simons Foundation

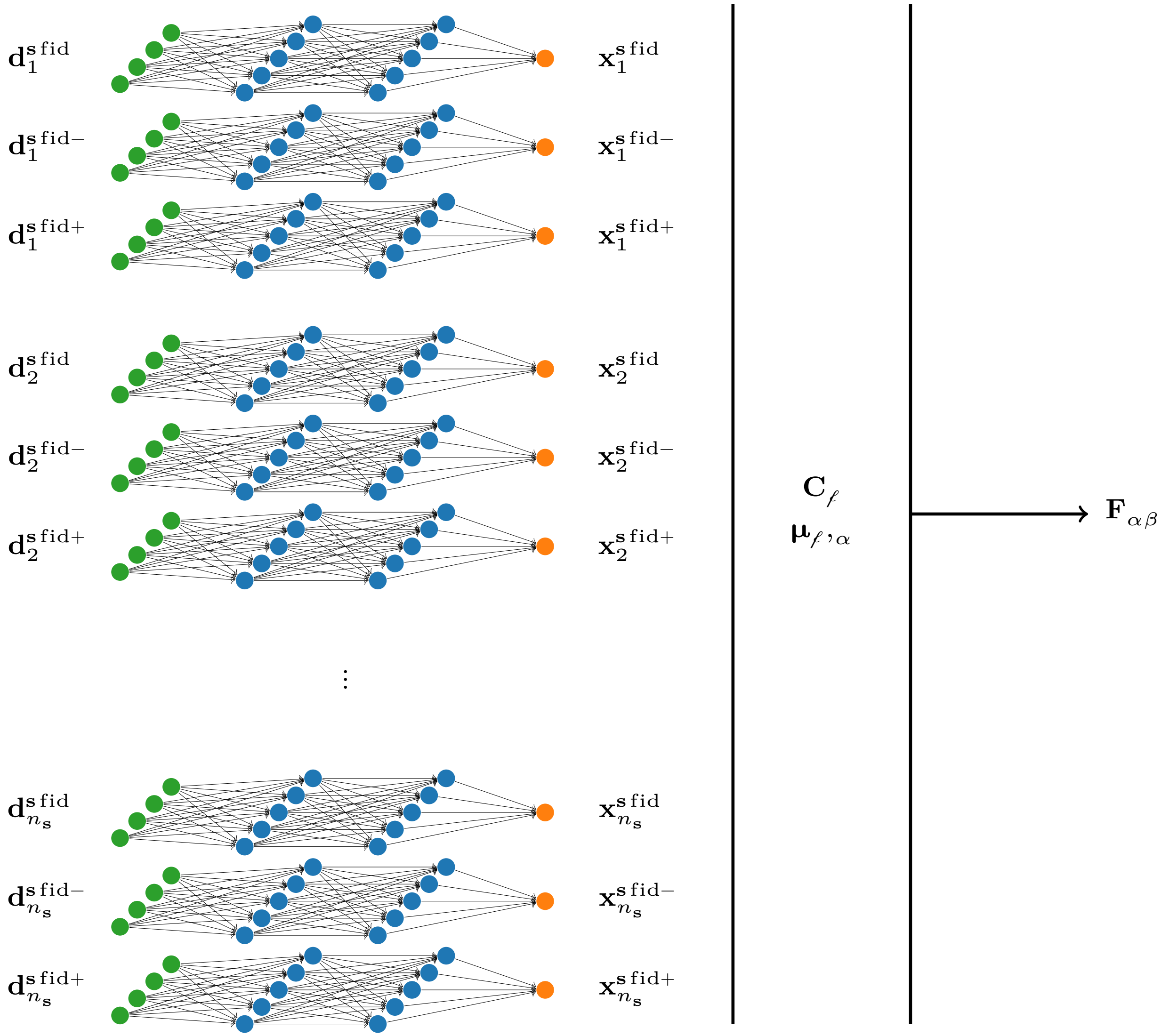
But what if you don’t know how to compute informative summaries of your data?

Automatic Physical Inference

Charnock, Lavaux, Wandelt (arXiv:1802:03537)

- Obviates the need to “guess” heuristic, informative summaries of the data
- An neural network is trained through reinforcement learning to compute functions of the data (*summaries)* that maximize the information about the parameters of the model.
- The network generates non-linear informative summaries that constrain model parameters

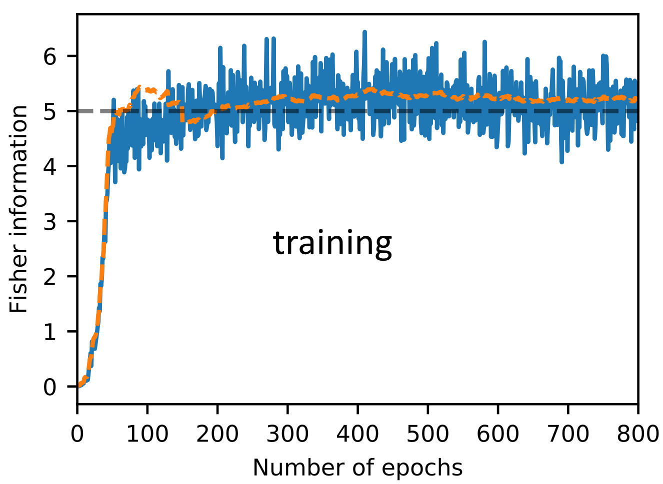
Information maximizing neural network

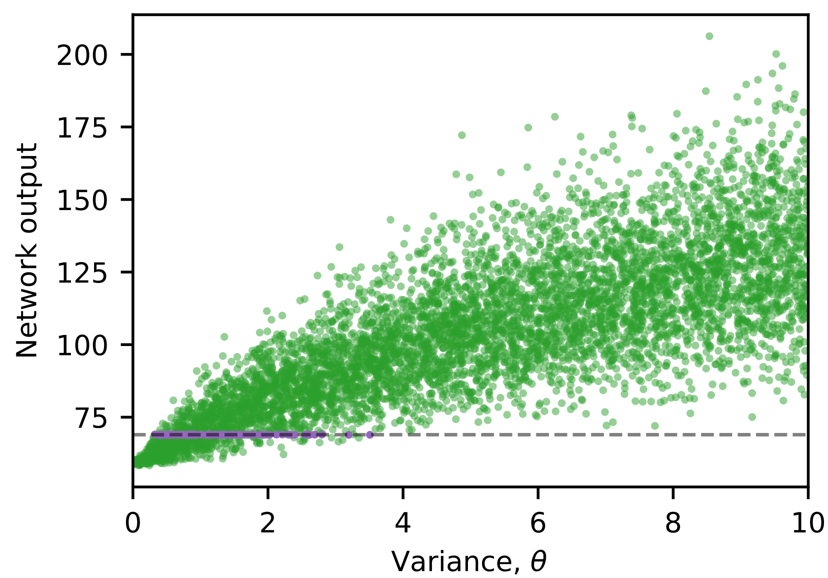


Charnock, Lavaux, Wandelt (arXiv:1802:03537)

Example 1: variance inference

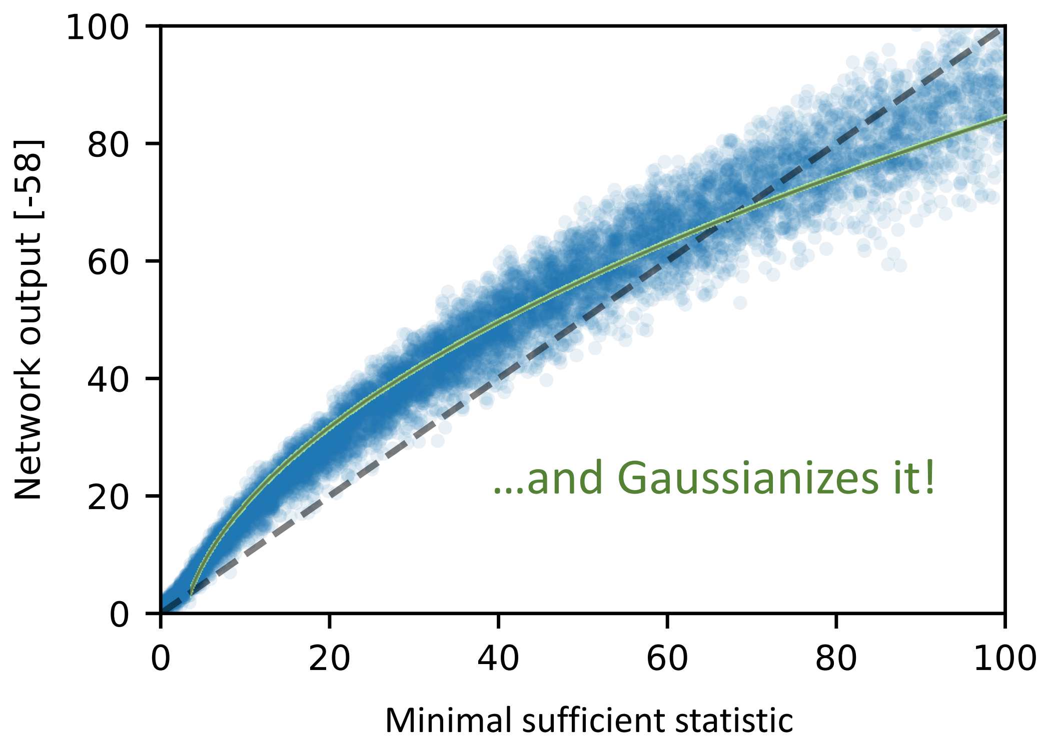
- Perfect information $F = 5$ in this problem
- Linear summaries give $\mathrm{F} = 0.5$





Charnock, Lavaux, Wandelt (arXiv:1802:03537)

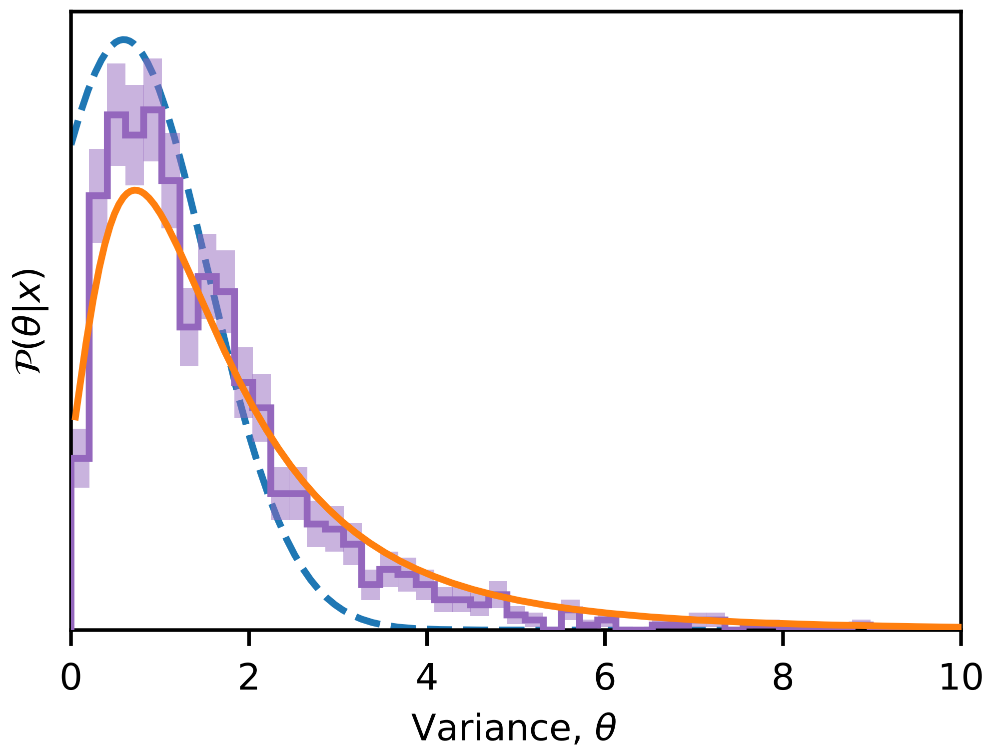
The IMNN finds a minimal sufficient statistic for this inference problem



Charnock, Lavaux, Wandelt (arXiv:1802:03537)

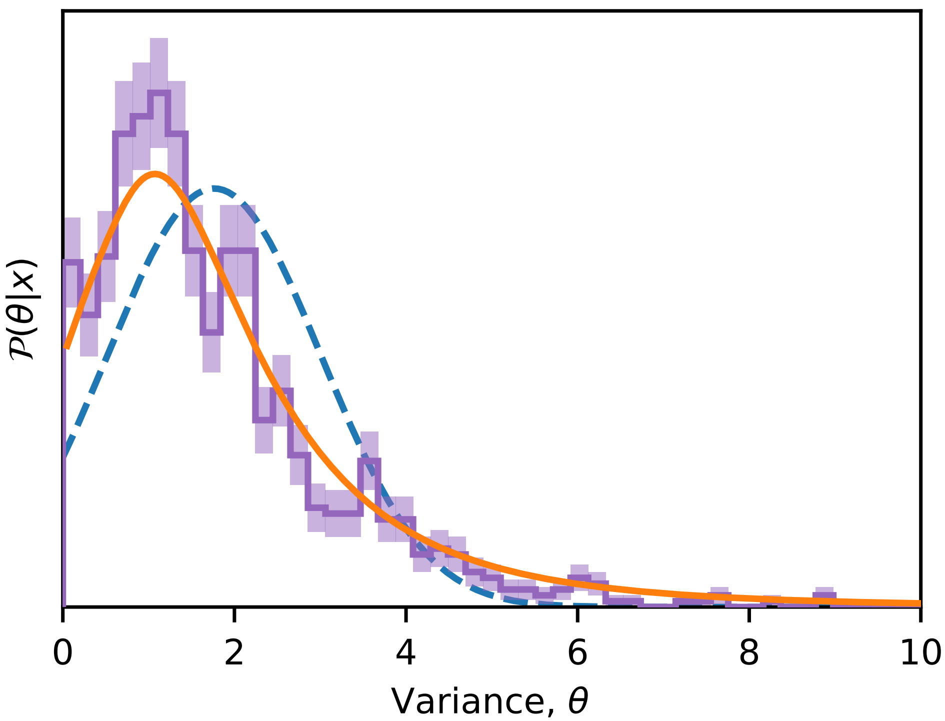
FLATIRON INSTITUTE
Division of Simons Foundation

Example 2: Automatic physical inference with noisy data



Charnock, Lavaux, Wandelt (arXiv:1802:03537)

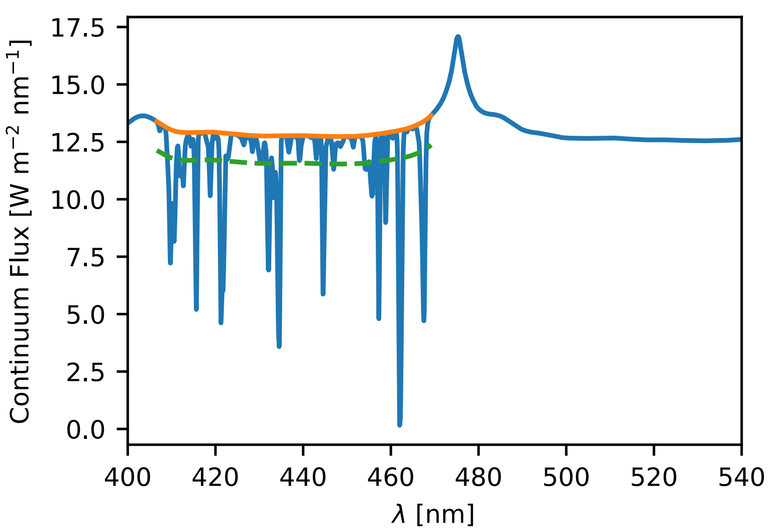
Example 3: Automatic physical inference with unknown noise

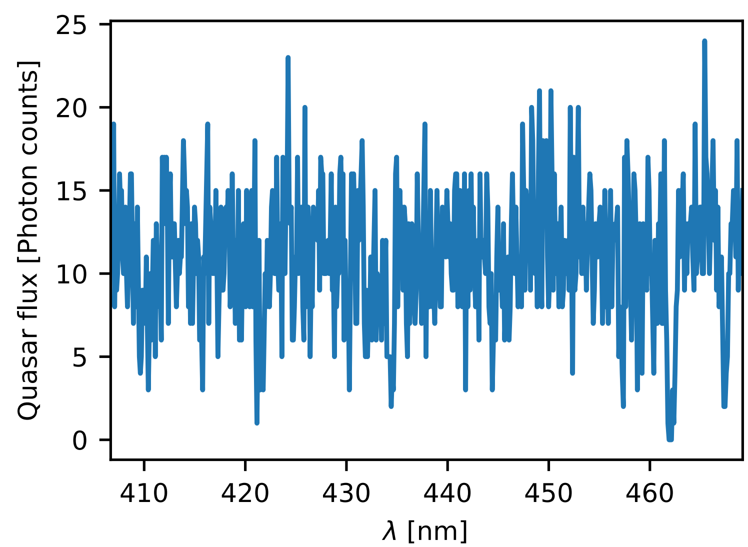


Charnock, Lavaux, Wandelt (arXiv:1802:03537)

Example 4: Lyman-α forest inference

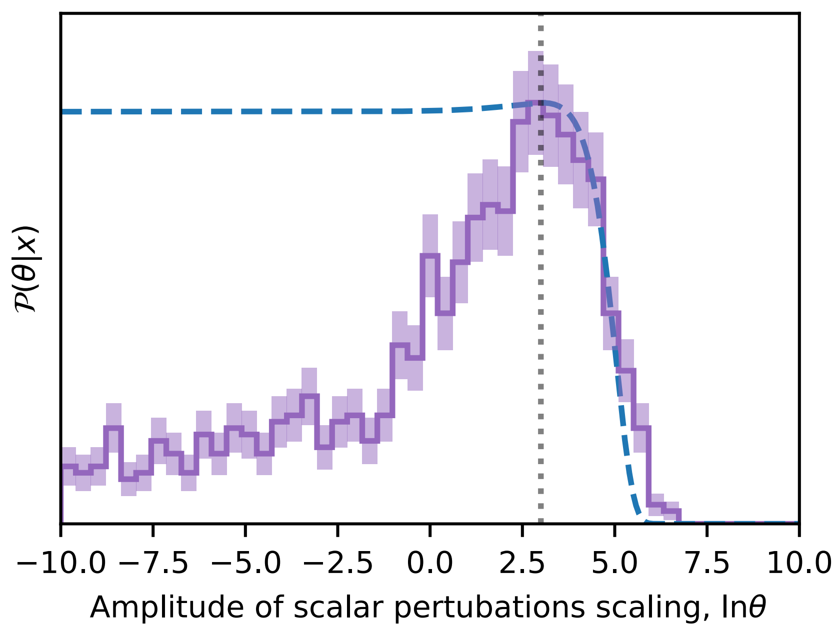
- The idea is to infer the variance of the underlying density field from a non-linearly transformed, photon-noise dominated Lyman-α forest spectrum





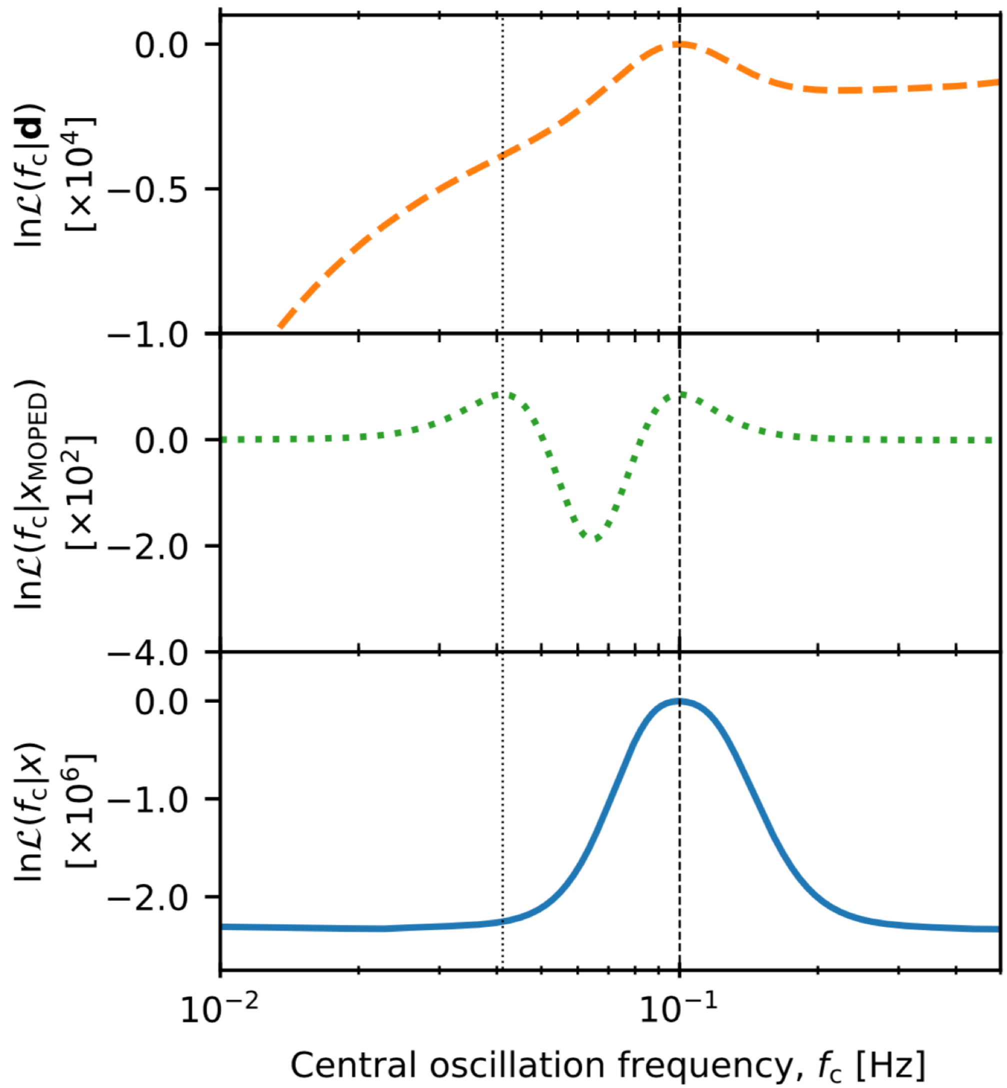
Charnock, Lavaux, Wandelt (arXiv:1802:03537)

Example 4: Lyman-α forest inference



Charnock, Lavaux, Wandelt (arXiv:1802:03537)

Example 5: LISA gravitational wave chirp



Full likelihood (LH)

Gaussian likelihood based on linear compression

Gaussian LH based on IMNN compression

The Information Maximizing Network summary gives the correct unique likelihood peak.

Charnock, Lavaux, Wandelt (arXiv:1802:03537)

Summary

- Cosmic Structure is a rich source of information on (and off) the cosmic light cone.
- **Avoid**: focus on linear regime
- **Adapt**: pick smart summaries (e.g. voids)/use X-corr probes
- **Attack**: computational inference with full physics model
- With non-linear inference techniques we can now reconstruct our cosmic history and unlock a much larger range of scales to probe dark matter and dark energy
- Automatic physical inference using information-maximizing neural networks and tractable likelihood-free inference is a new approach to extracting scientific, information from complex data and physical simulations.

The code for the IMNN paper is is published on zenodo

https://doi.org/10.5281/zenodo.1175196