Bayesian data interpretation with large scale cosmological models

Jens Jasche

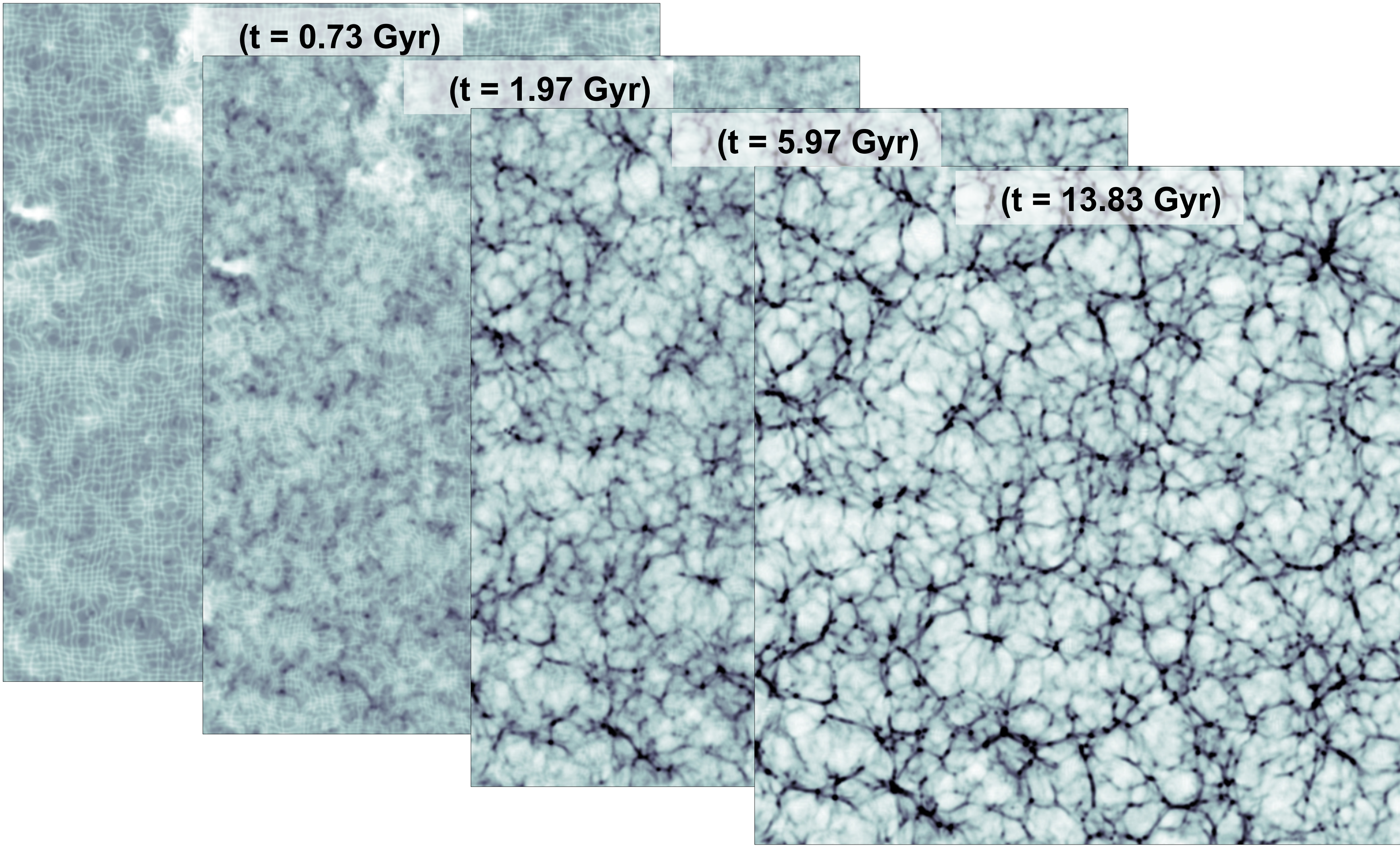
Guilhem Lavaux, Florent Leclercq, Doogesh Kodi Ramanah, Harry Desmond, Natalia Porqueres, Fabian Schmidt, Franz Elsner Benjamin Wandelt
and the
Aquila consortium

Methods for Statistical Inference

Paris, 25 October 2018

The cosmic large scale structure...

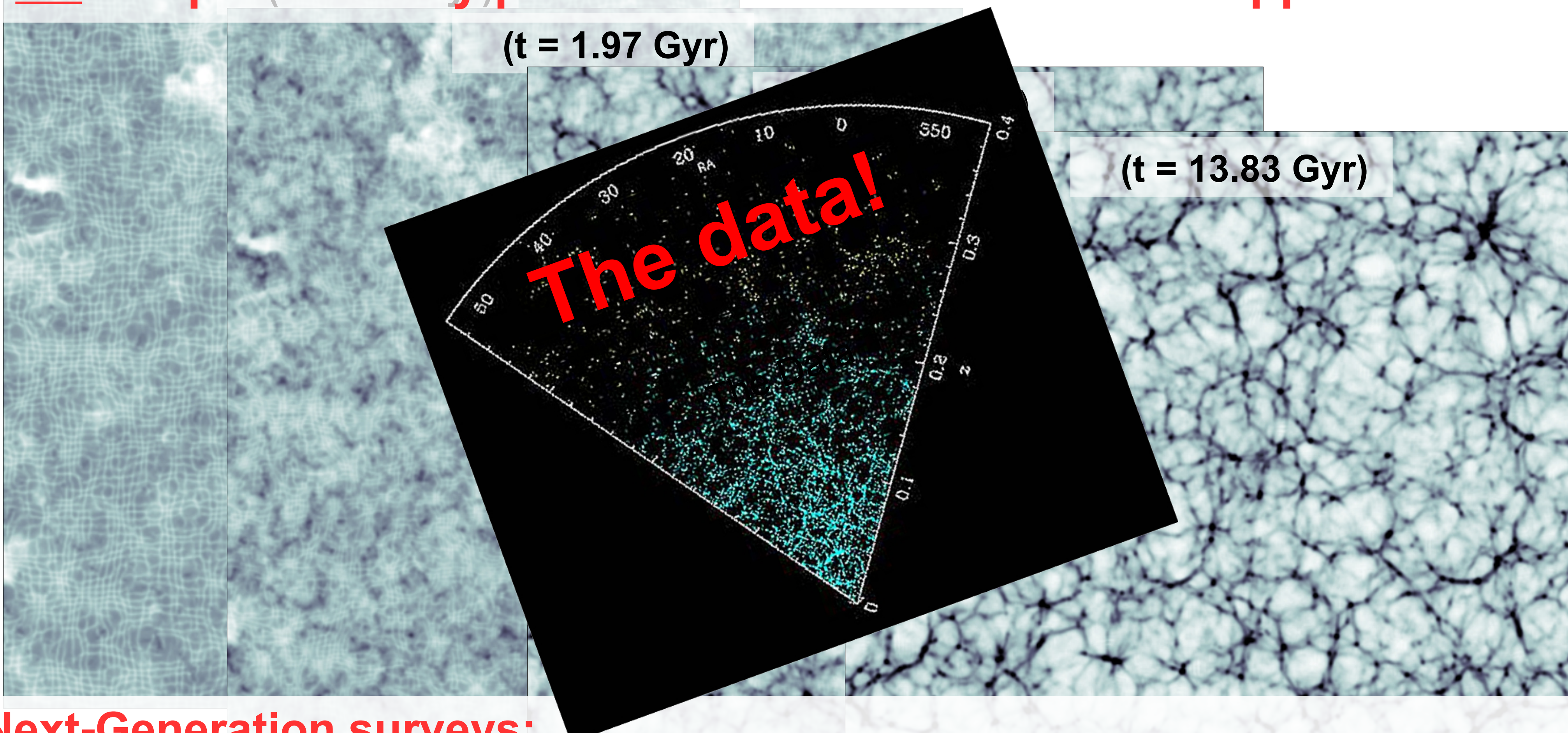
... A source of knowledge!



The cosmic large scale structure...

... A source of knowledge!

<u>No</u> unique recovery possible!!! We need statistical approaches!



Next-Generation surveys:

- Dominated by Systematics
- ~ 80 % of the total signal comes from non-linear structures

see e.g. LSST Science collaboration (2009)
Schäfer (2017) (arXiv:1701.04469)

Bayesian forward modeling

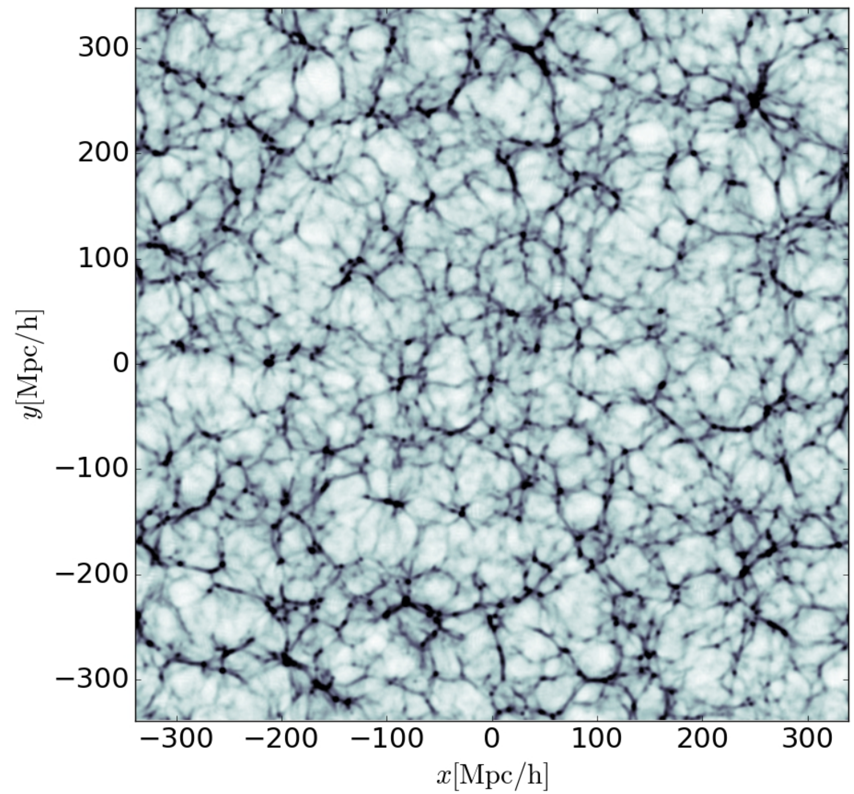
Bayesian inference

Jasche, Wandelt (2013)

- Need posterior distribution

$$\mathscr{P}(\boldsymbol{s}|\boldsymbol{d}) = \mathscr{P}(\boldsymbol{s}) \frac{\mathscr{P}(\boldsymbol{d}|\boldsymbol{s})}{\mathscr{P}(\boldsymbol{d})}$$

Final State



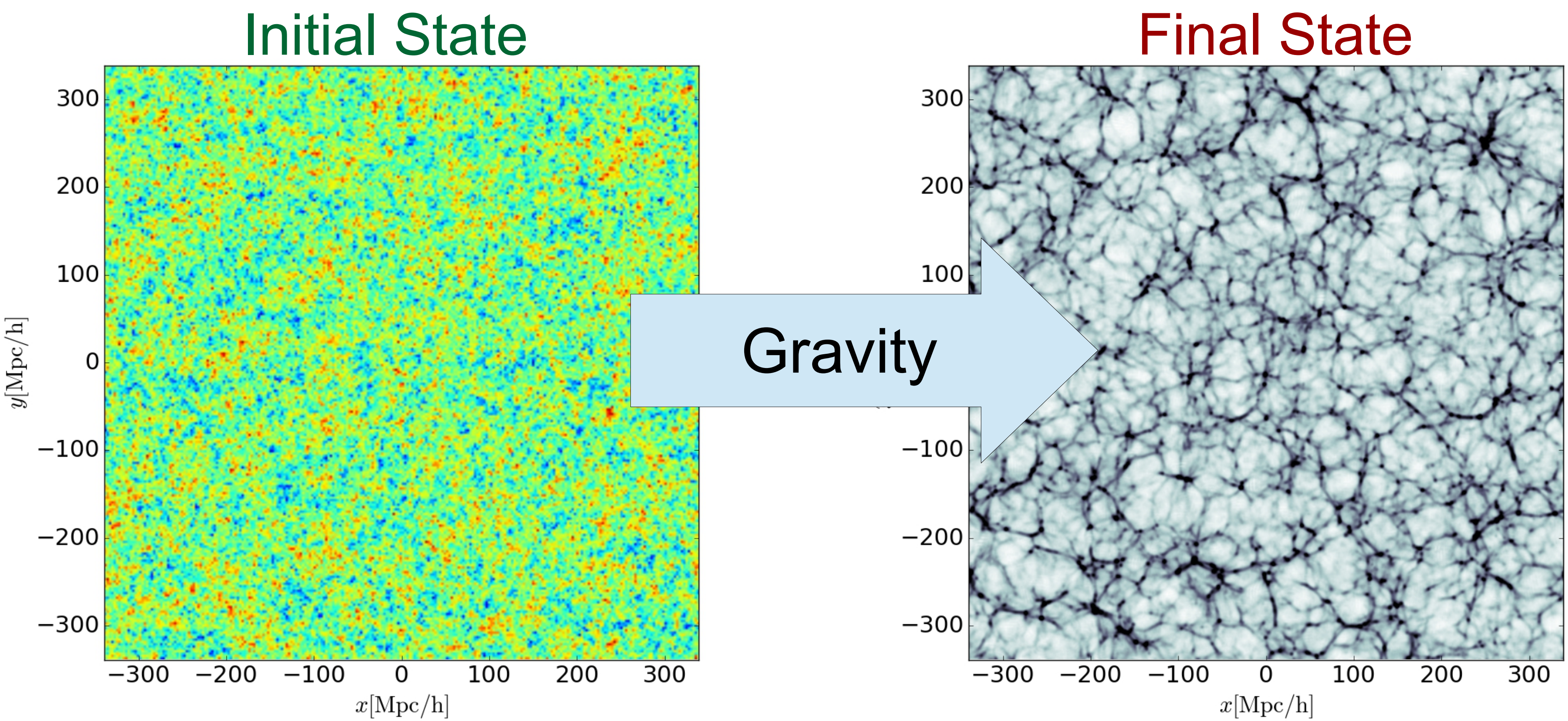
Bayesian forward modeling

Bayesian inference

Jasche, Wandelt (2013)

- Need posterior distribution

$$\mathscr{P}(\boldsymbol{s}|\boldsymbol{d}) = \mathscr{P}(\boldsymbol{s})\,\frac{\mathscr{P}(\boldsymbol{d}|\boldsymbol{s})}{\mathscr{P}(\boldsymbol{d})}$$



A large scale Bayesian inverse problem

Bayesian Forward modeling:

Jasche, Wandelt (2013)
Lavaux, Jasche (2016)

Prior model

$$\mathscr{P}(\boldsymbol{s}|\boldsymbol{S}) = \frac{\mathrm{e}^{-\frac{1}{2}\boldsymbol{s}^{\mathrm{T}}\boldsymbol{S}^{-1}\boldsymbol{s}}}{\sqrt{\det(2\,\pi\,\boldsymbol{S})}}$$

Structure formation model

$$\mathscr{P}(\boldsymbol{\delta}|\boldsymbol{s}) = \prod_i \delta^D\left(\delta_i - G_i(\boldsymbol{s})\right)$$

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}a} = \frac{\vec{p}}{\dot{a}a^2}$$

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}a} = -\frac{3}{2}H_0^2\Omega_m\frac{\nabla^2\Phi}{Ha^2}$$

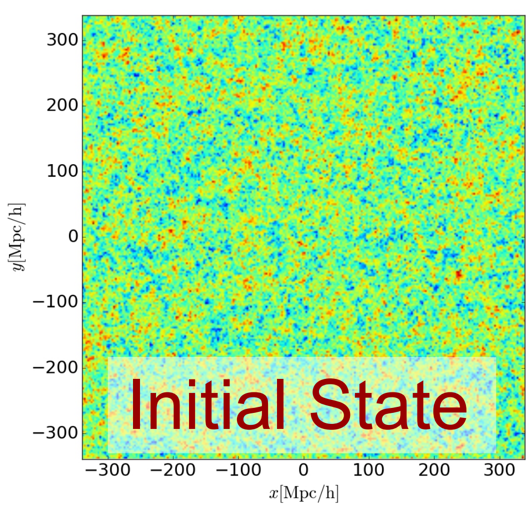
Data model

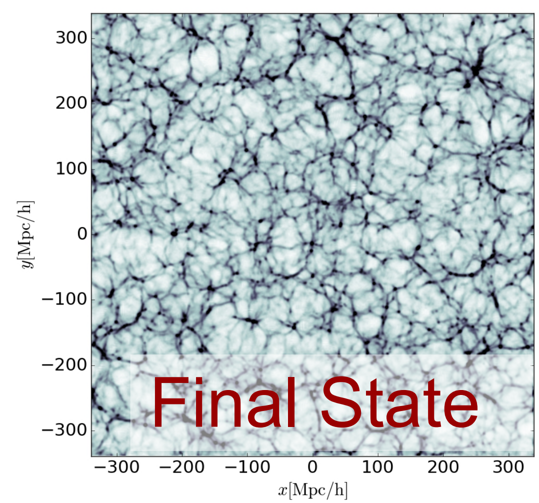
$$\mathscr{P}(\boldsymbol{N}|\boldsymbol{\lambda}(\boldsymbol{\delta})) = \prod_i \frac{\mathrm{e}^{-\lambda_i}\lambda_i^{N_i}}{N_i\,!}$$

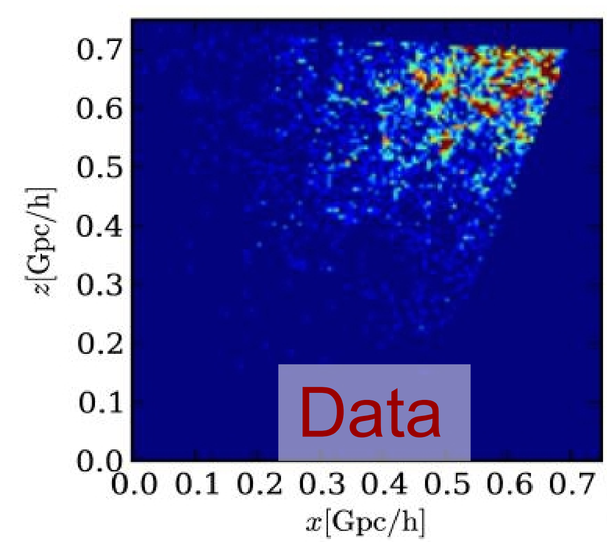
Galaxy bias model

$$\lambda_i = R_i\bar{N}\,(1+\delta)^{\beta}\,\mathrm{e}^{-\rho_g\,(1+\delta)^{-\epsilon_g}}$$

See e.g. Neyrinck et al. 2014
Ata et al. 2015
Lavaux & Jasche 2016







$$\dim(\boldsymbol{s}) \sim 10^7 \text{ parameters}$$

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln\left(\mathcal{P}(\mathbf{x})\right)$

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln(\mathcal{P}(\mathbf{x}))$
- The Hamiltonian : $H = \frac{1}{2}\,\mathbf{p}^T\,\mathbf{M}^{-1}\mathbf{p} + \psi(\mathbf{x})$

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln\left(\mathcal{P}(\mathbf{x})\right)$
- The Hamiltonian : $H = \frac{1}{2}\mathbf{p}^T\,\mathbf{M}^{-1}\mathbf{p} + \psi(\mathbf{x})$

Nuisance parameter!!!

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln\left(\mathcal{P}(\mathbf{x})\right)$
- The Hamiltonian : $H = \frac{1}{2}\mathbf{p}^T\,\mathbf{M}^{-1}\,\mathbf{p} + \psi(\mathbf{x})$

Nuisance parameter!!!

$$(\mathbf{x},\ \mathbf{p}) \Rightarrow \begin{aligned} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} &= \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p} \\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} &= -\frac{\partial H}{\partial \mathbf{x}} = -\frac{\partial \psi(\mathbf{x})}{\partial \mathbf{x}} \end{aligned} \Rightarrow (\mathbf{x}',\ \mathbf{p}')$$

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln\left(\mathcal{P}(\mathbf{x})\right)$
- The Hamiltonian : $H = \frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p} + \psi(\mathbf{x})$

Nuisance parameter!!!

$$(\mathbf{x},\ \mathbf{p}) \Rightarrow \begin{aligned} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} &= \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p} \\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} &= -\frac{\partial H}{\partial \mathbf{x}} = -\frac{\partial \psi(\mathbf{x})}{\partial \mathbf{x}} \end{aligned} \Rightarrow (\mathbf{x}',\ \mathbf{p}')$$

see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

MCMC in high dimensions

HMC: Use Classical mechanics to solve statistical problems!

- The potential : $\psi(\mathbf{x}) = -\ln\left(\mathcal{P}(\mathbf{x})\right)$
- The Hamiltonian : $H = \frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p} + \psi(\mathbf{x})$

Nuisance parameter!!!

$$(\mathbf{x},\ \mathbf{p}) \Rightarrow \begin{aligned} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} &= \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p} \\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} &= -\frac{\partial H}{\partial \mathbf{x}} = -\frac{\partial \psi(\mathbf{x})}{\partial \mathbf{x}} \end{aligned} \Rightarrow (\mathbf{x}',\ \mathbf{p}')$$

Randomize $\mathbf{p}$ and accept $\mathbf{x}'$: $\alpha = \min\left[1, \mathrm{e}^{-(H'-H)}\right] = 1$

HMC beats the "curse of dimensionality" by:

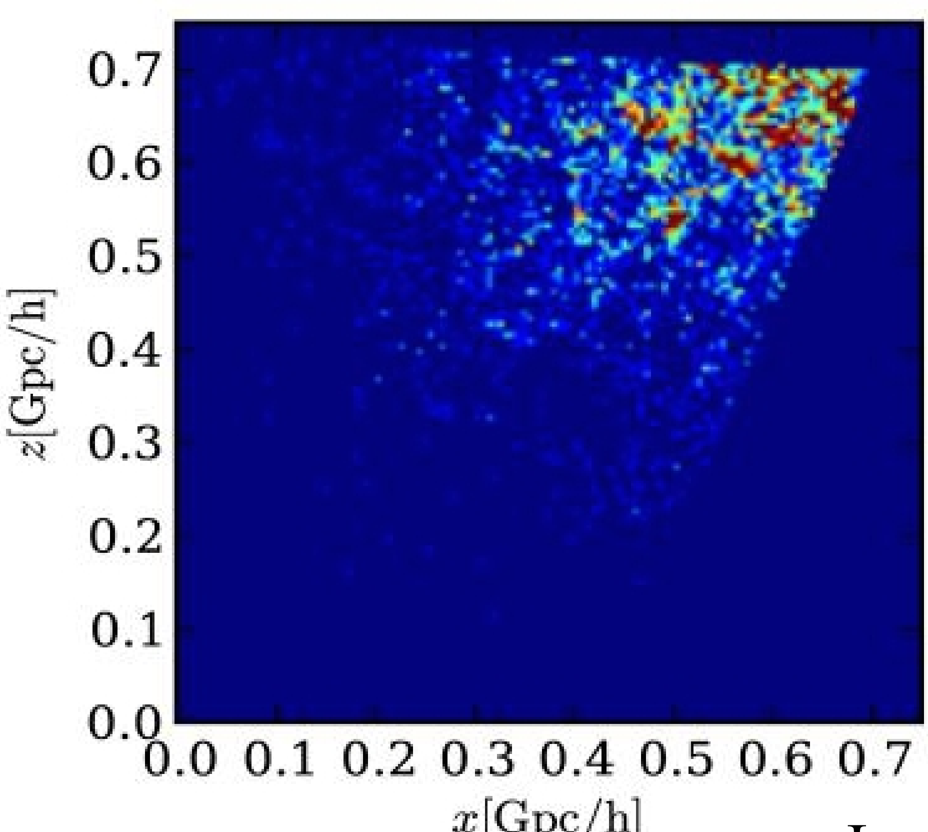
- Exploiting gradients
- Using conserved quantities

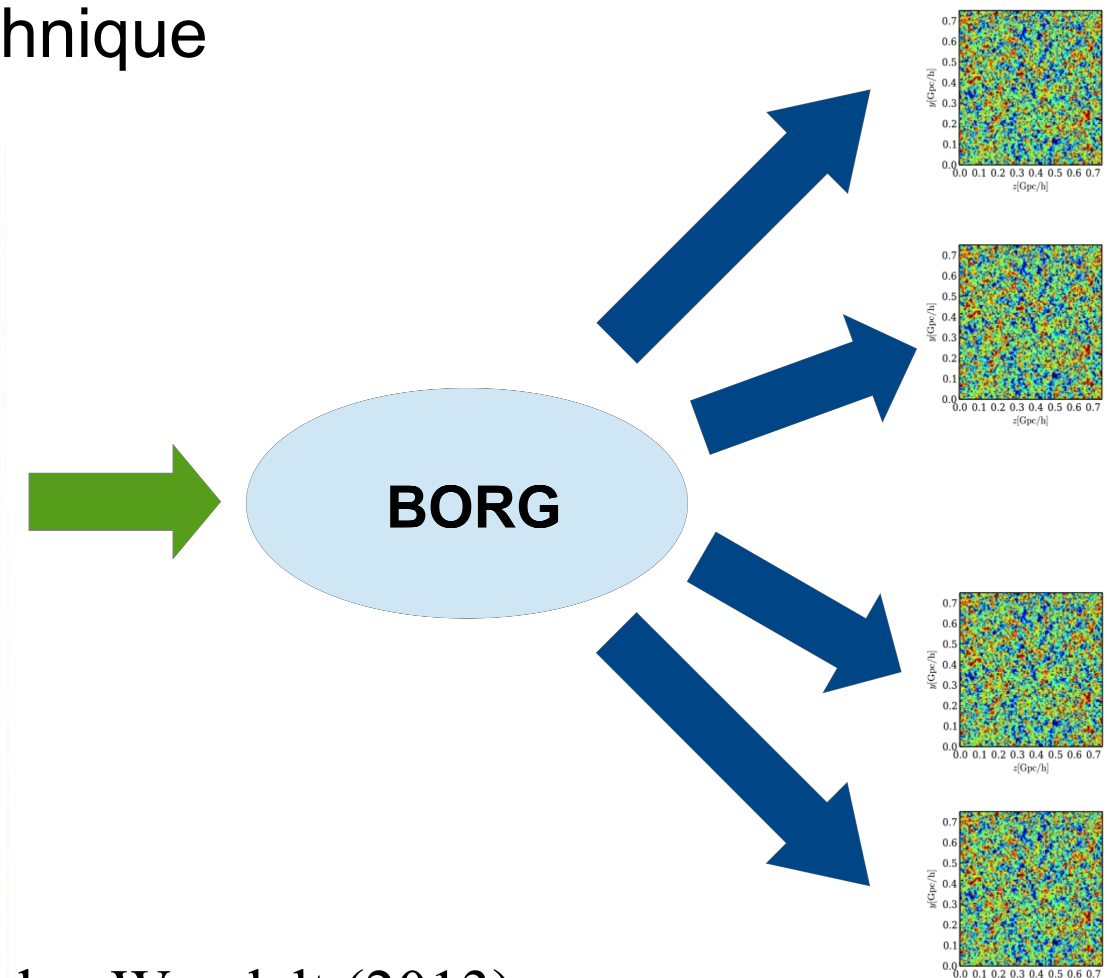
see e.g. Duane et al. (1987)
Neal (2012)
Betancourt (2017)

Bayesian Inference of initial conditions

BORG (**B**ayesian **O**rigin **R**econstruction from **G**alaxies)

- Uses dynamical LSS model (2LPT, PM) within Likelihood
- Solves a statistical initial conditions problem
- Exploits HMC sampling technique





Jasche, Wandelt (2013)

Full Bayesian analysis

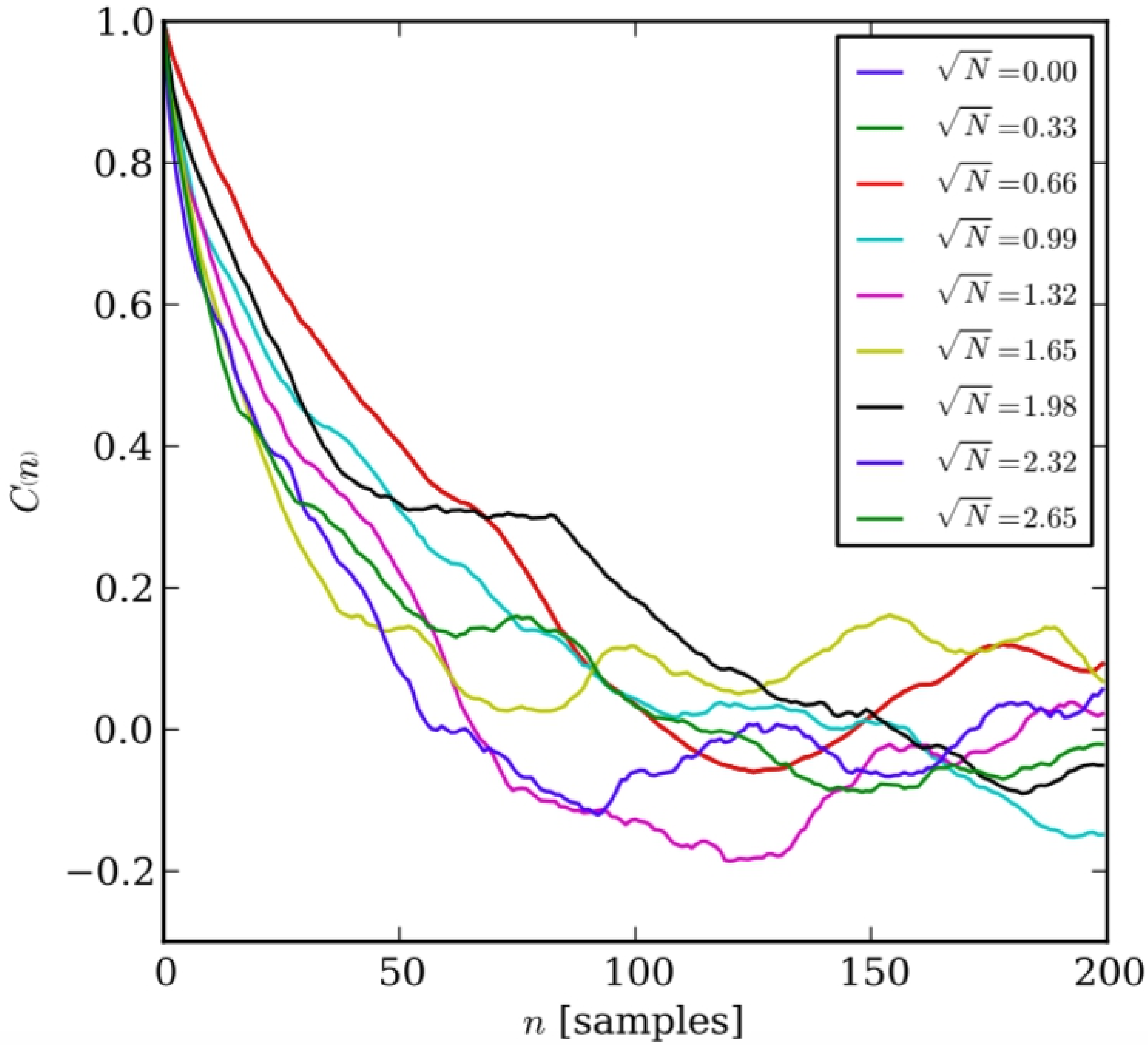
Bayesian Origin Reconstruction from Galaxies (BORG)

- Simultaneously constrained initial and final conditions
- Data application: SDSS DR7 main sample / 2LPT - Poisson

Jasche et al. 2015 (arXiv:1409.6308)

A comment on correlation length

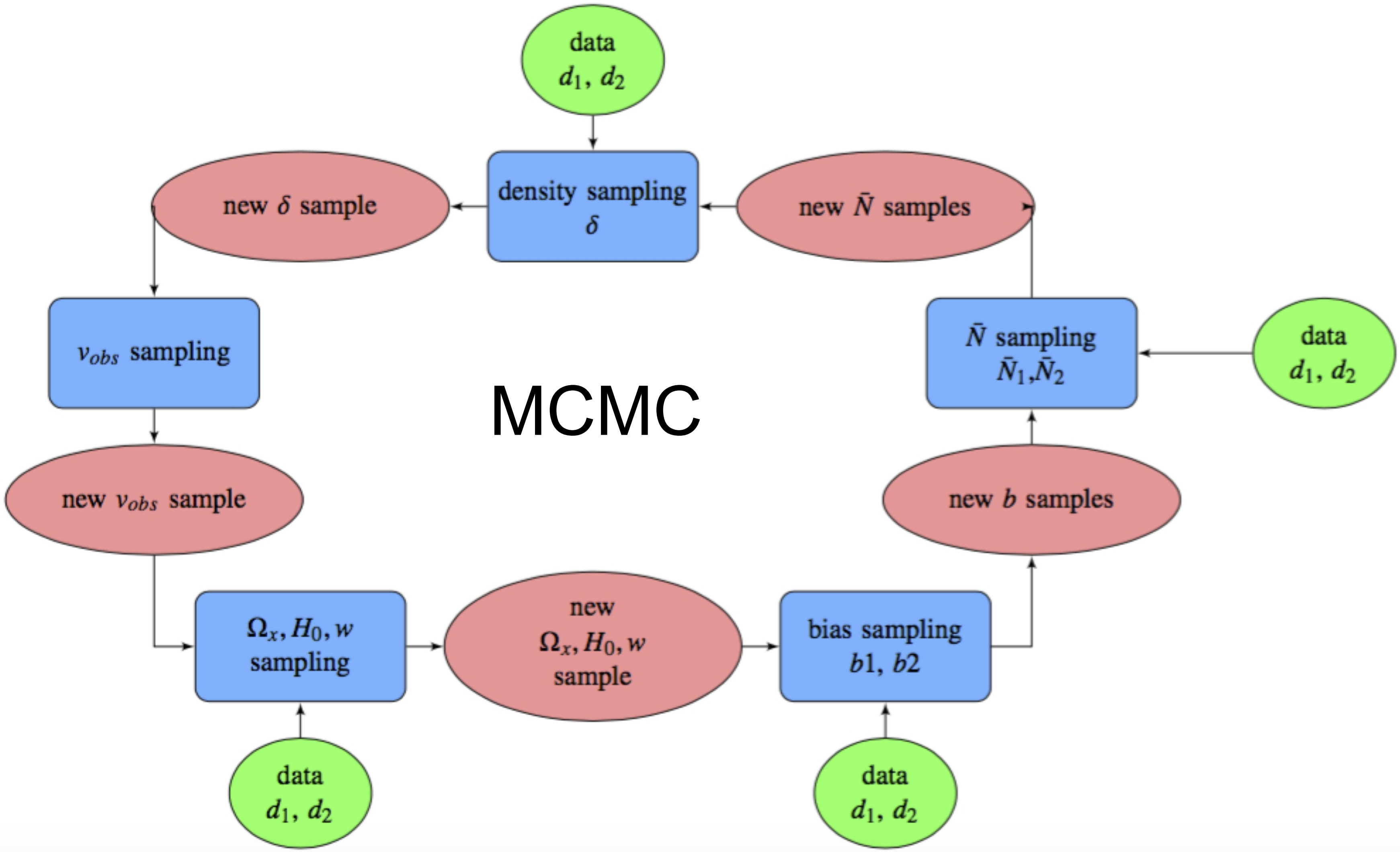
Correlation length ~ 100 samples



Jasche &Wandelt 2013 (arXiv:1203.3639)

BORG³: A Modular statistical programing engine

A MCMC framework to build flexible data models

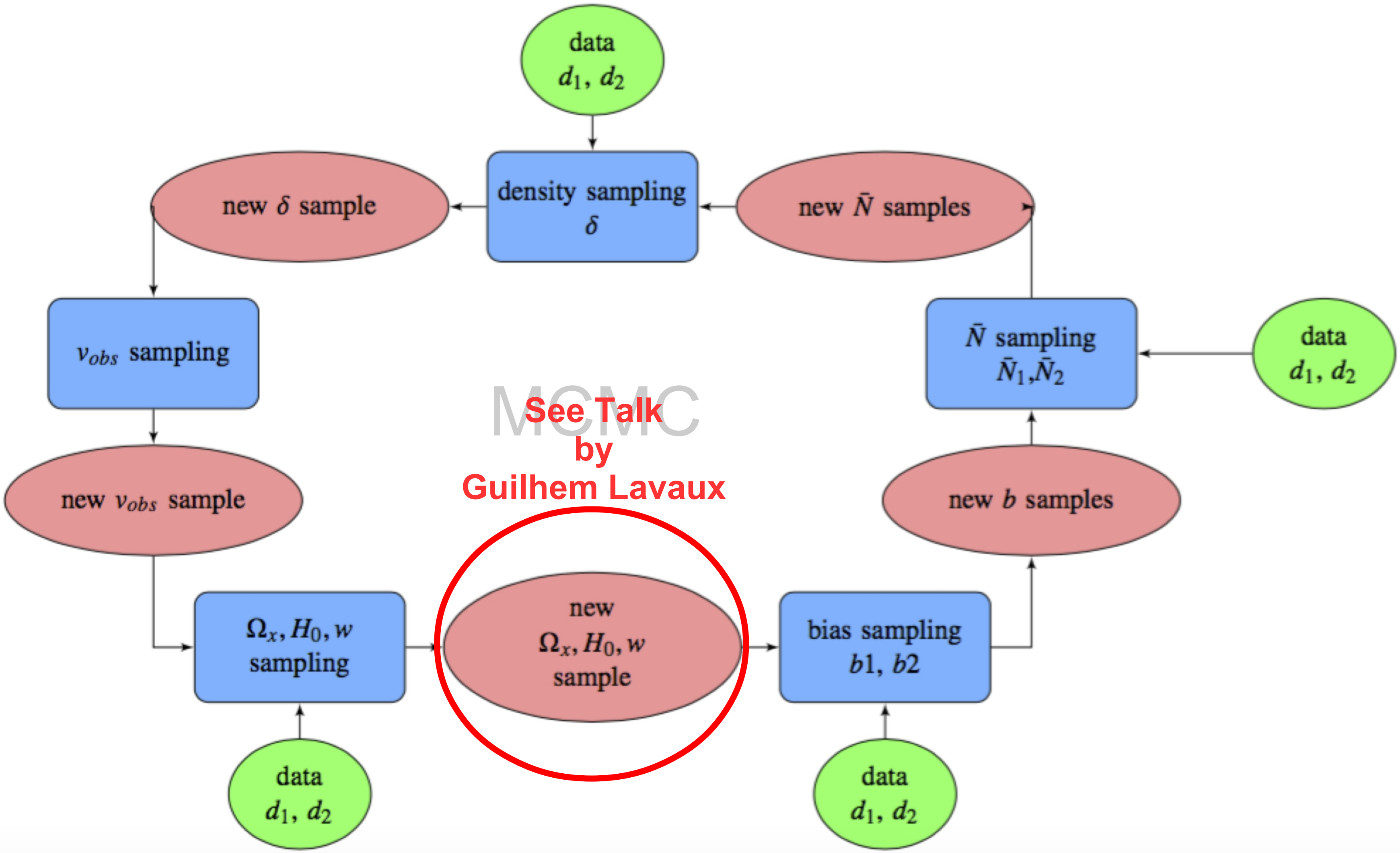


Straightforward combinations of HMC with other samplers:

- Use e.g. slice sampling for unit acceptance (Neal 2003)

BORG[3]: A Modular statistical programing engine

A MCMC framework to build flexible data models

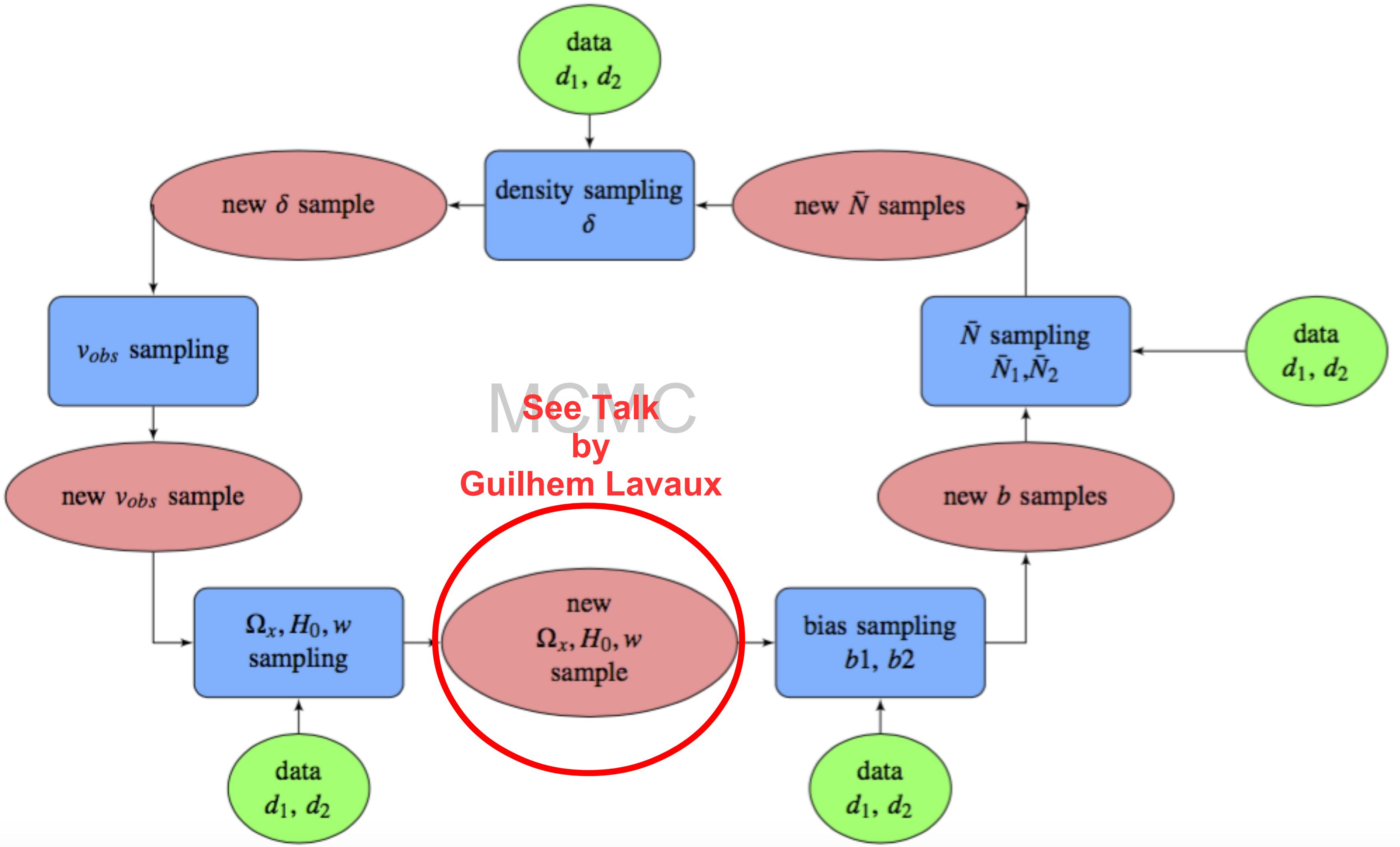


Straightforward combinations of HMC with other samplers:

- Use e.g. slice sampling for unit acceptance (Neal 2003)

BORG³: A Modular statistical programing engine

A MCMC framework to build flexible data models



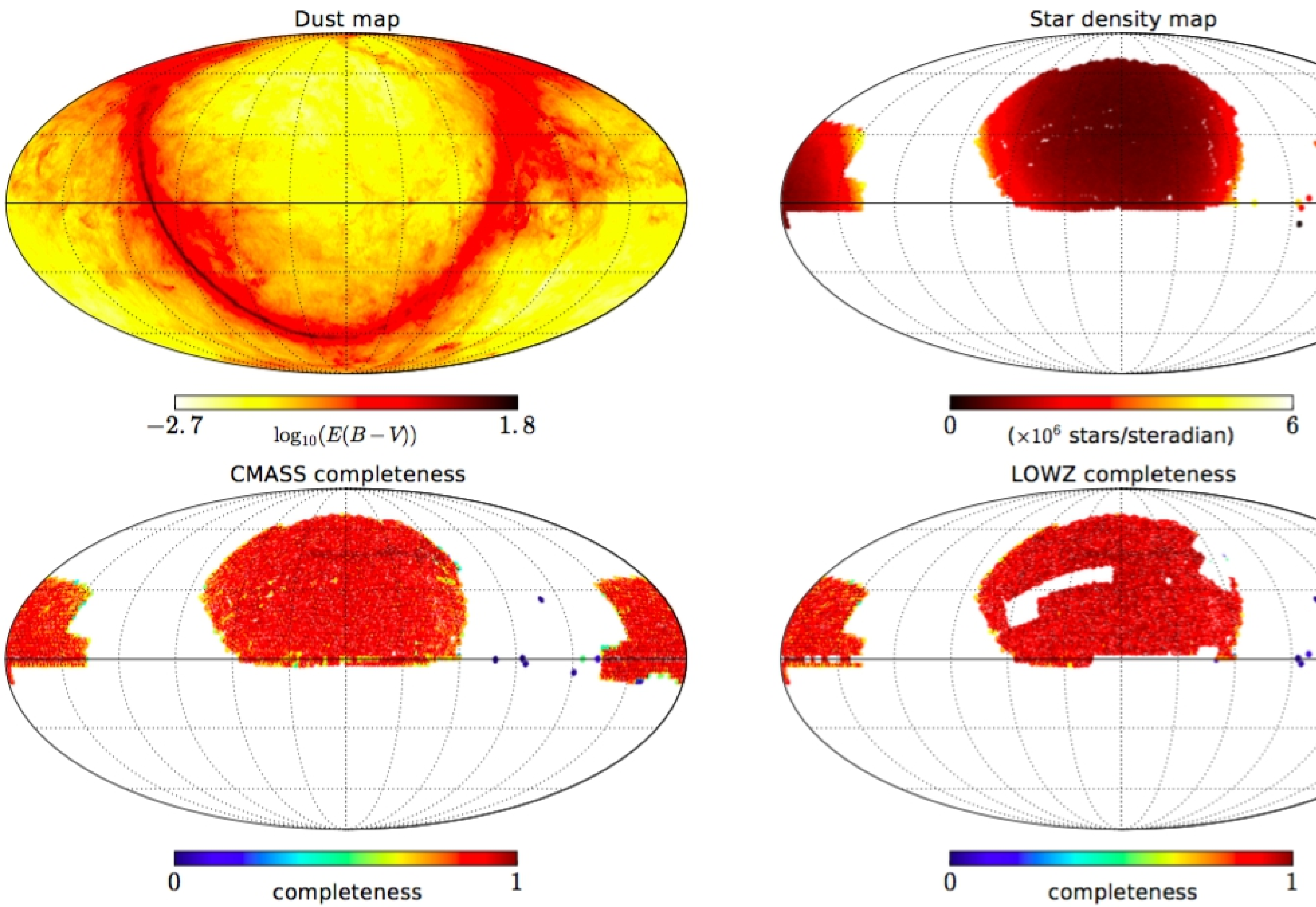
Marginalize out nuisance parameters.

Straightforward combinations of HMC with other samplers:

- Use e.g. slice sampling for unit acceptance (Neal 2003)

Systematics: Foreground contaminations

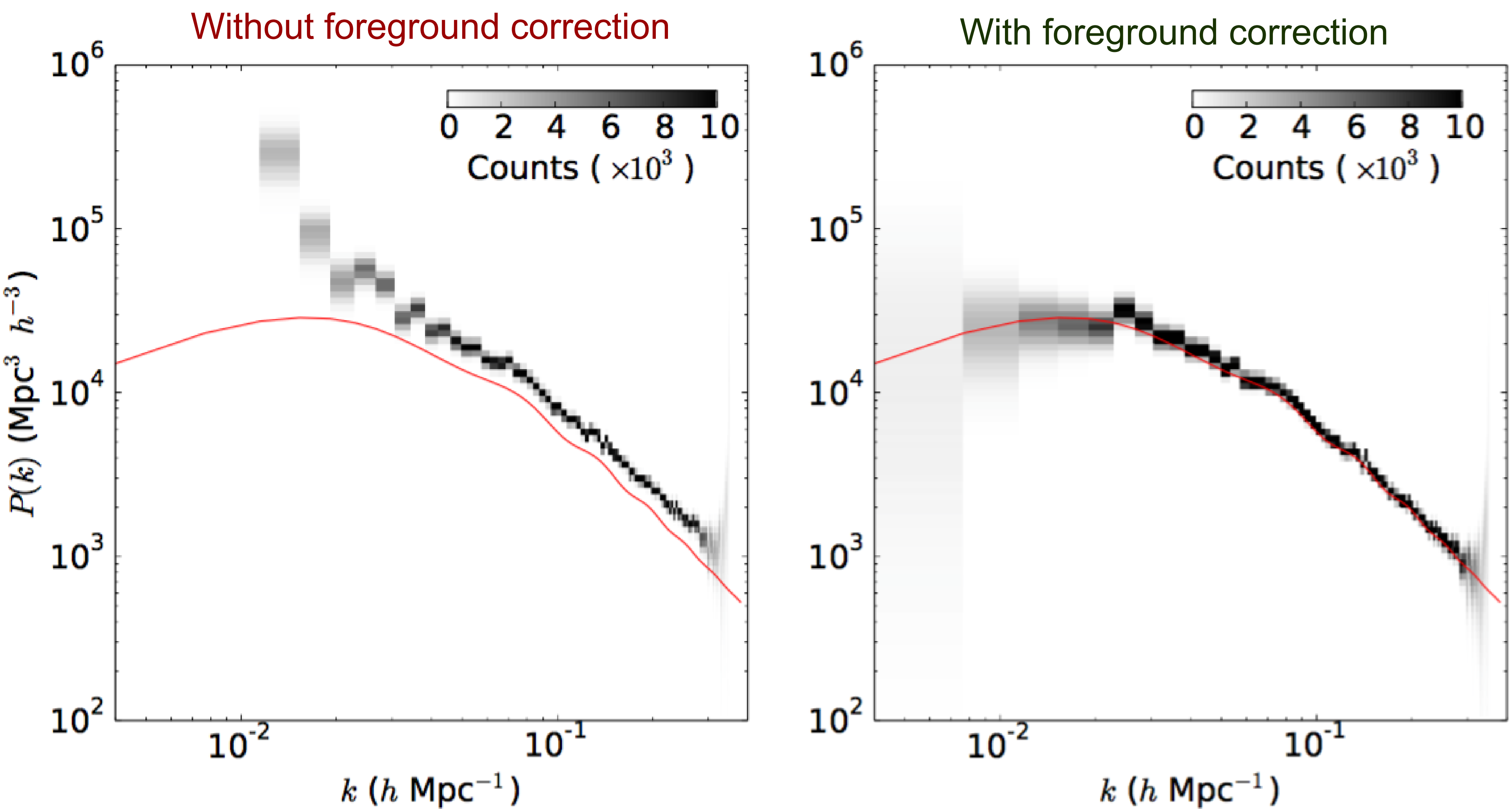
Foreground effect contaminate the inference (see e.g. Leistedt & Peiris (2014))



Jasche & Lavaux 2017 (arXiv:1706.08971)

Systematics: Foreground contaminations

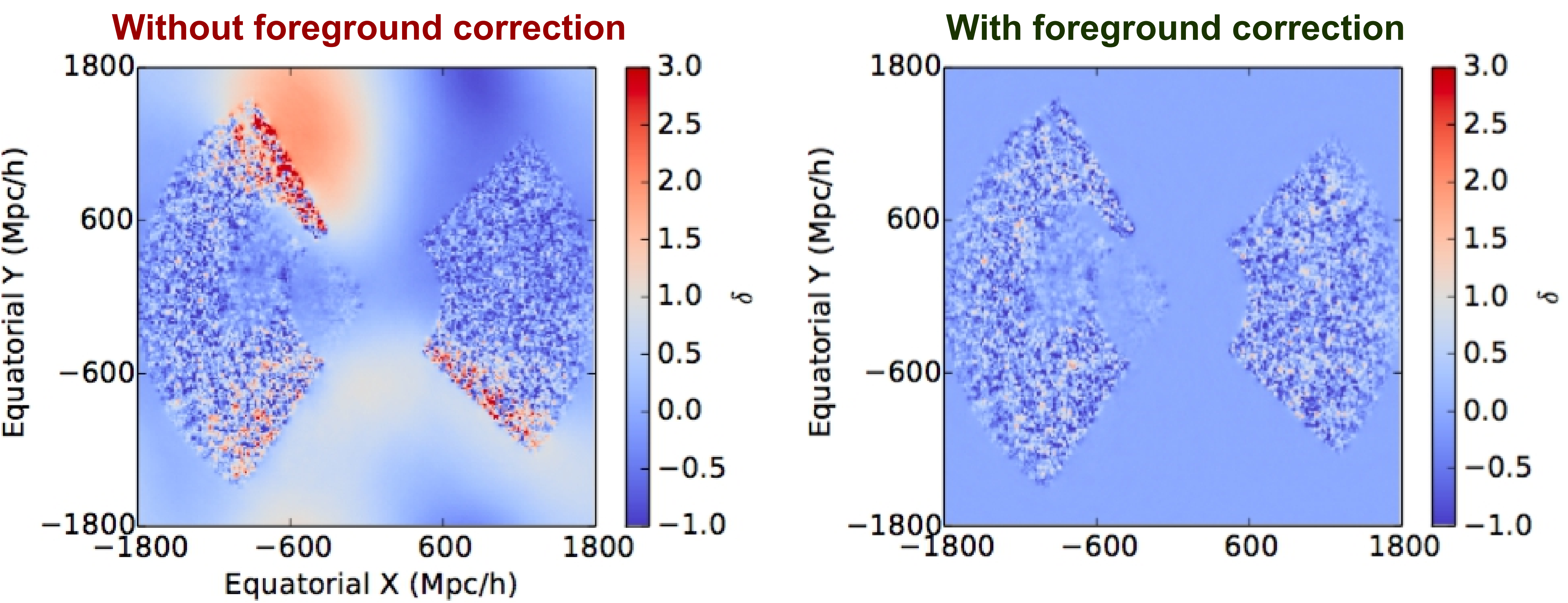
Mock data emulating LOWZ + CMASS



Jasche & Lavaux 2017 (arXiv:1706.08971)

Systematics: Foreground contaminations

Mock data emulating LOWZ + CMASS



Use inferred 3D density field as diagnostics.

Jasche & Lavaux 2017 (arXiv:1706.08971)

Uncertainties: Photo-z

Photometric redshift surveys

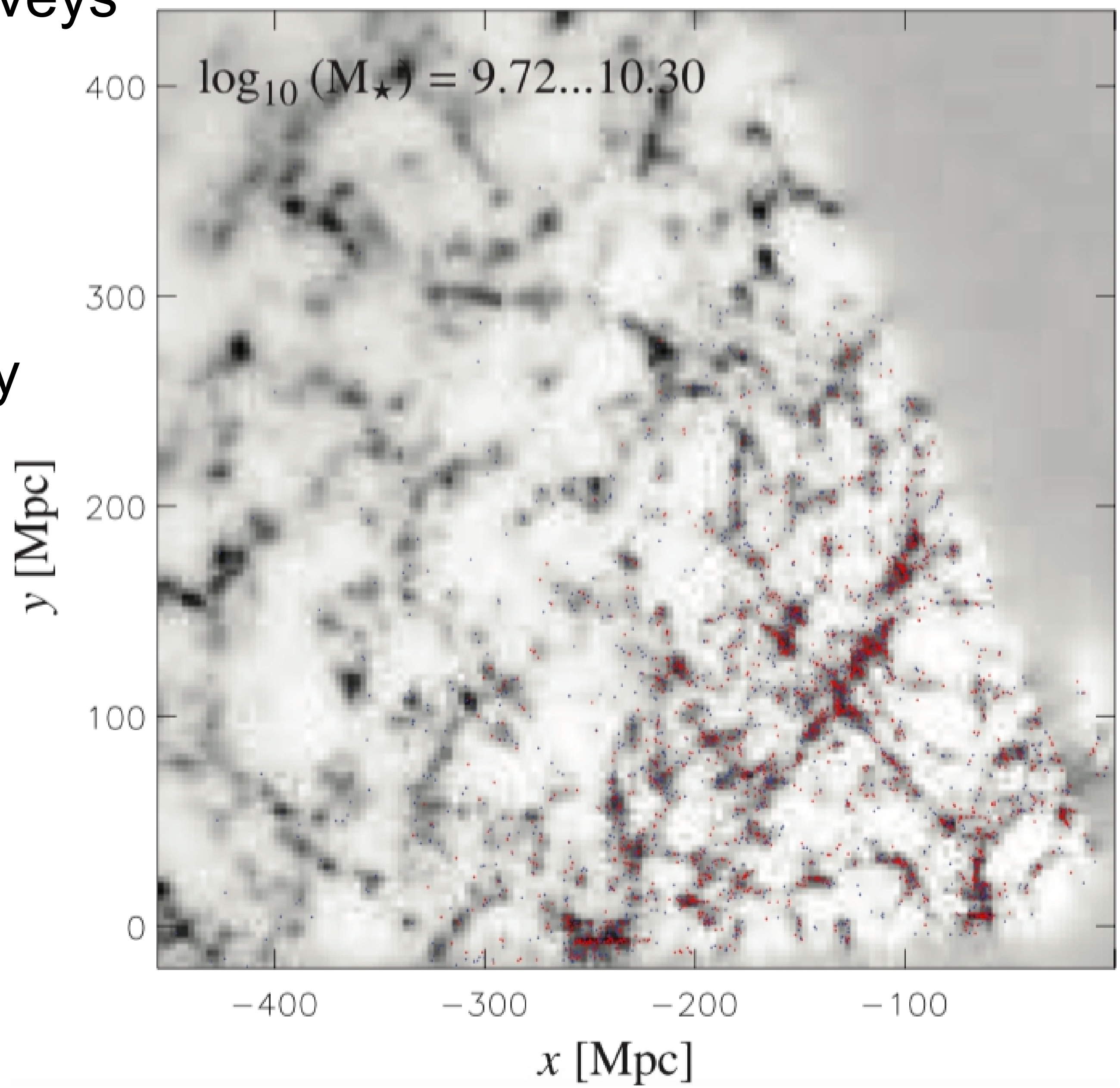
(e.g. LSST / Euclid)

- Deep volumes
- Millions of galaxies
- Low redshift accuracy

affects density fields and cosmological analyses

See e.g. Blake, Bridle 2005

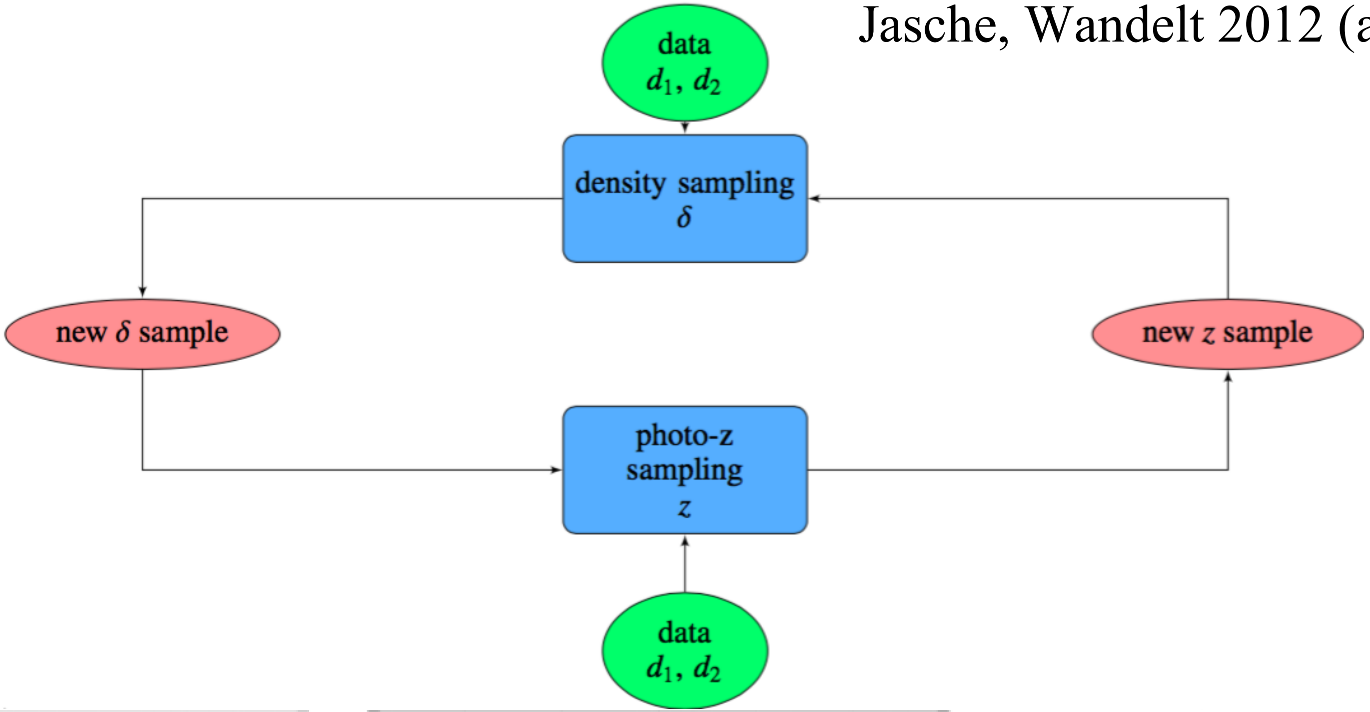
But: Galaxies trace the matter distribution!!

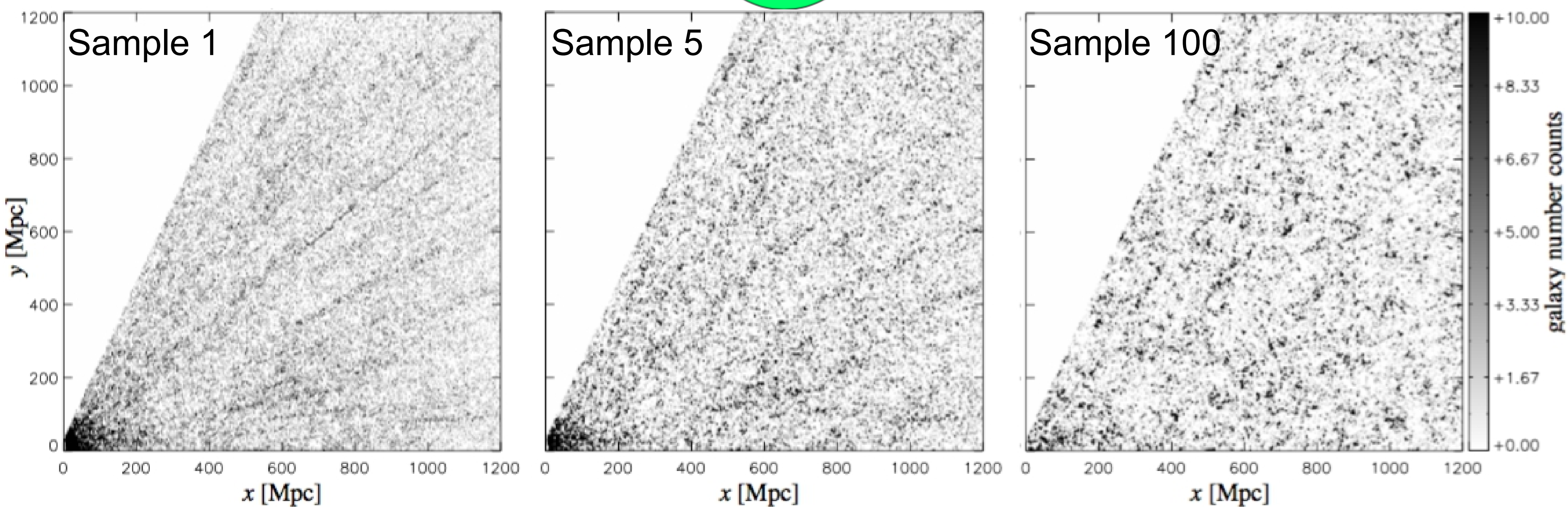


Jasche et al 2010 (arXiv:0911.2498)

Uncertainties: Photo-z

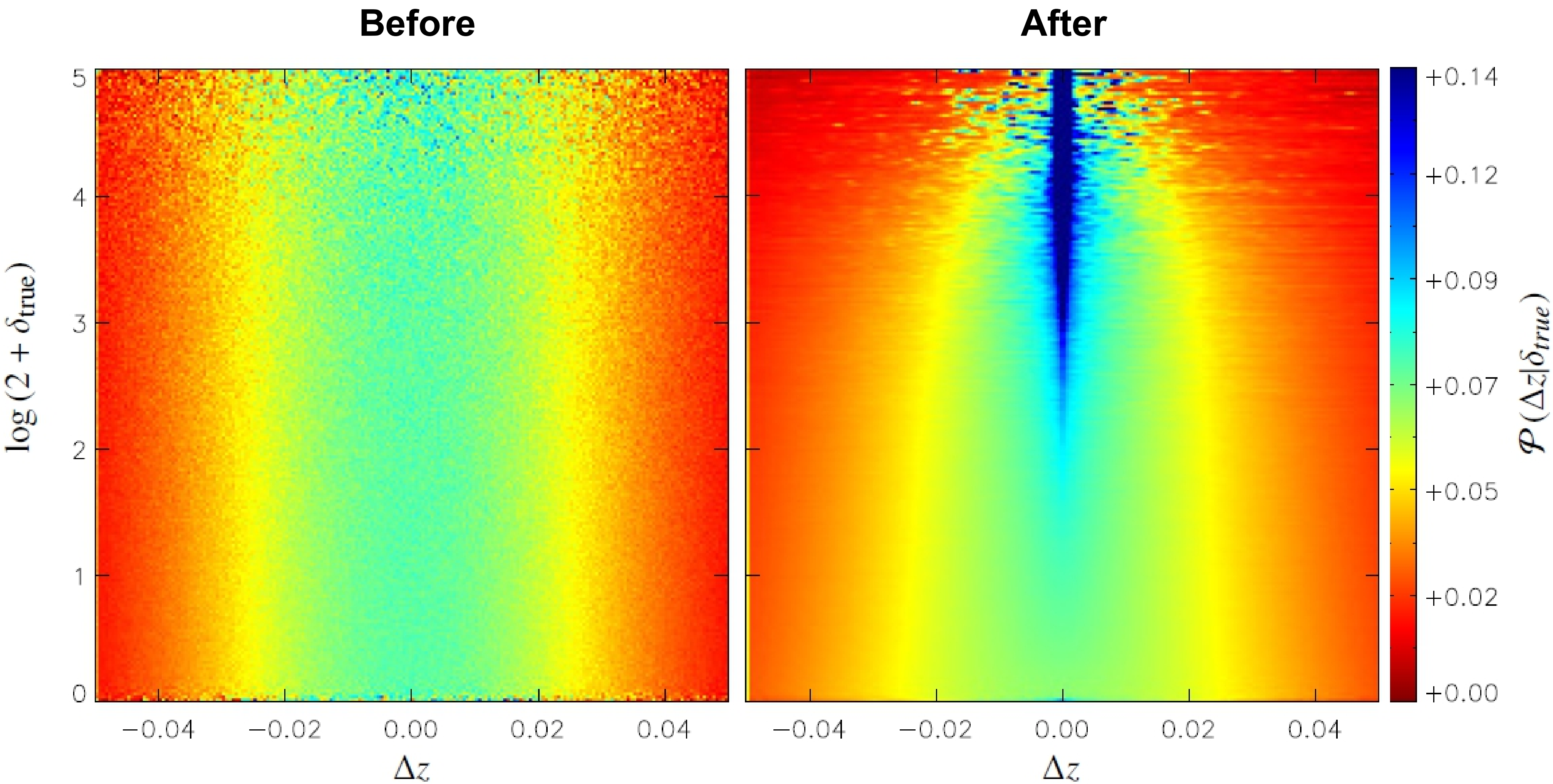
Jasche, Wandelt 2012 (arXiv:1106.2757)





Uncertainties: Photo-z

Proof of concept: Application to mock data



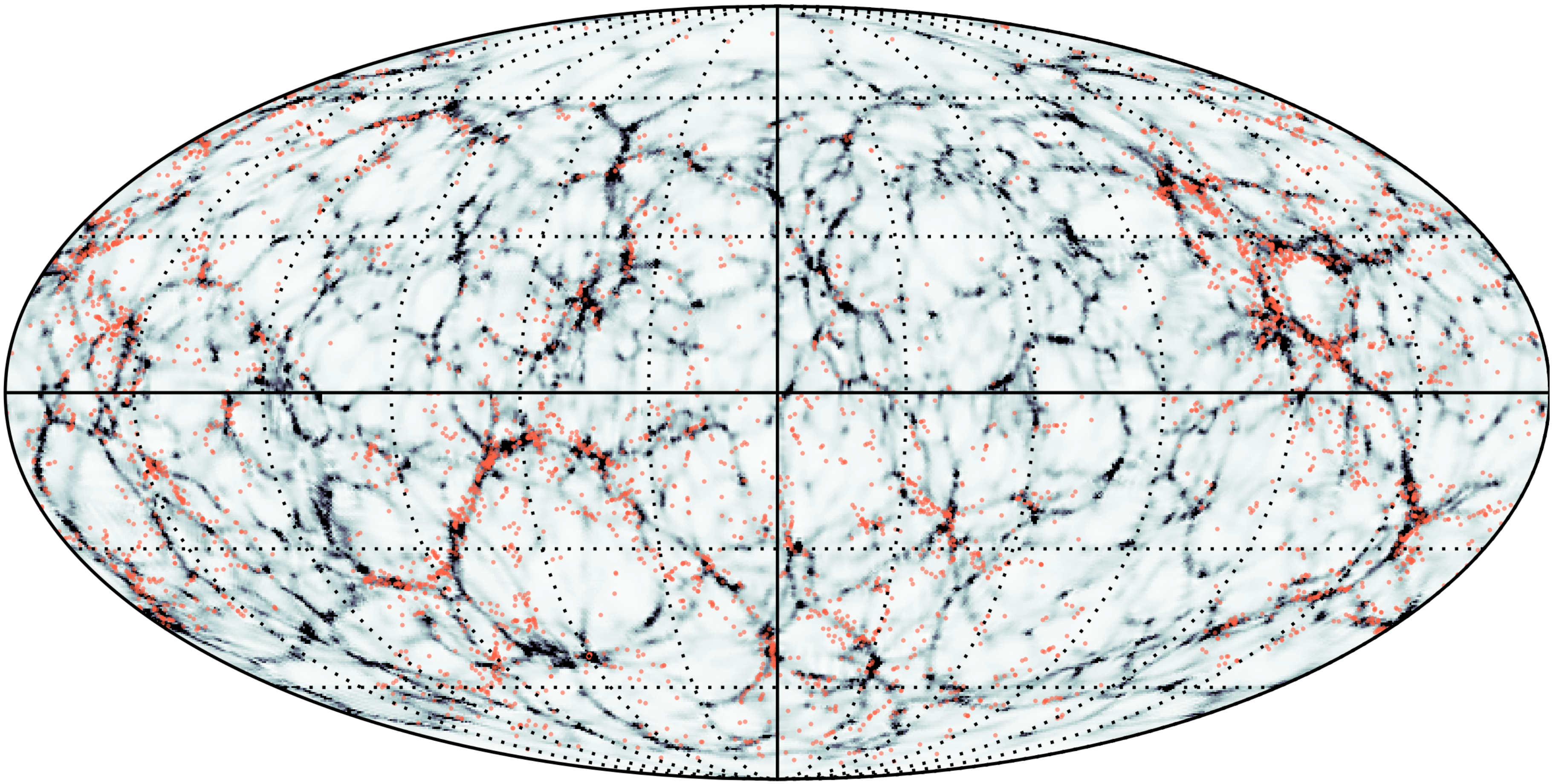
$$\delta z \sim 0.03$$

$$\delta z_f \sim 0.003$$

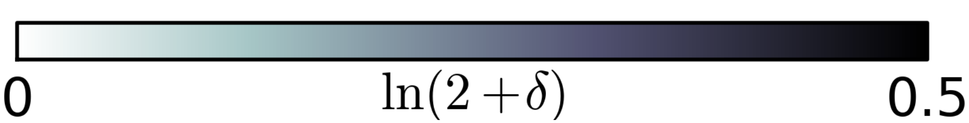
Jasche, Wandelt 2012 (arXiv:1106.2757)

The non-linear LSS of our Universe

Final conditions inferred from spectroscopic 2M++ data

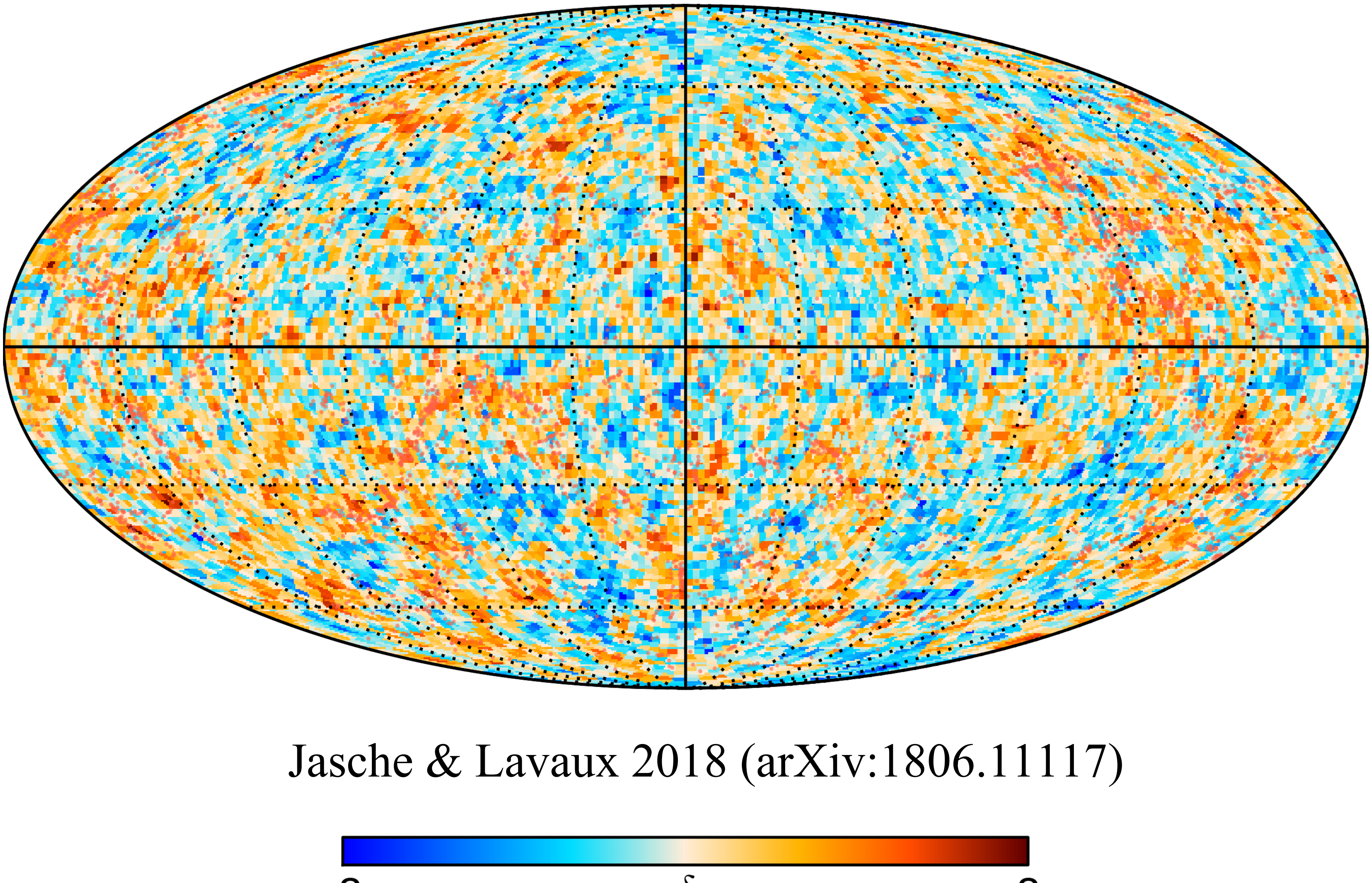


Jasche & Lavaux 2018 (arXiv:1806.11117)



The non-linear LSS of our Universe

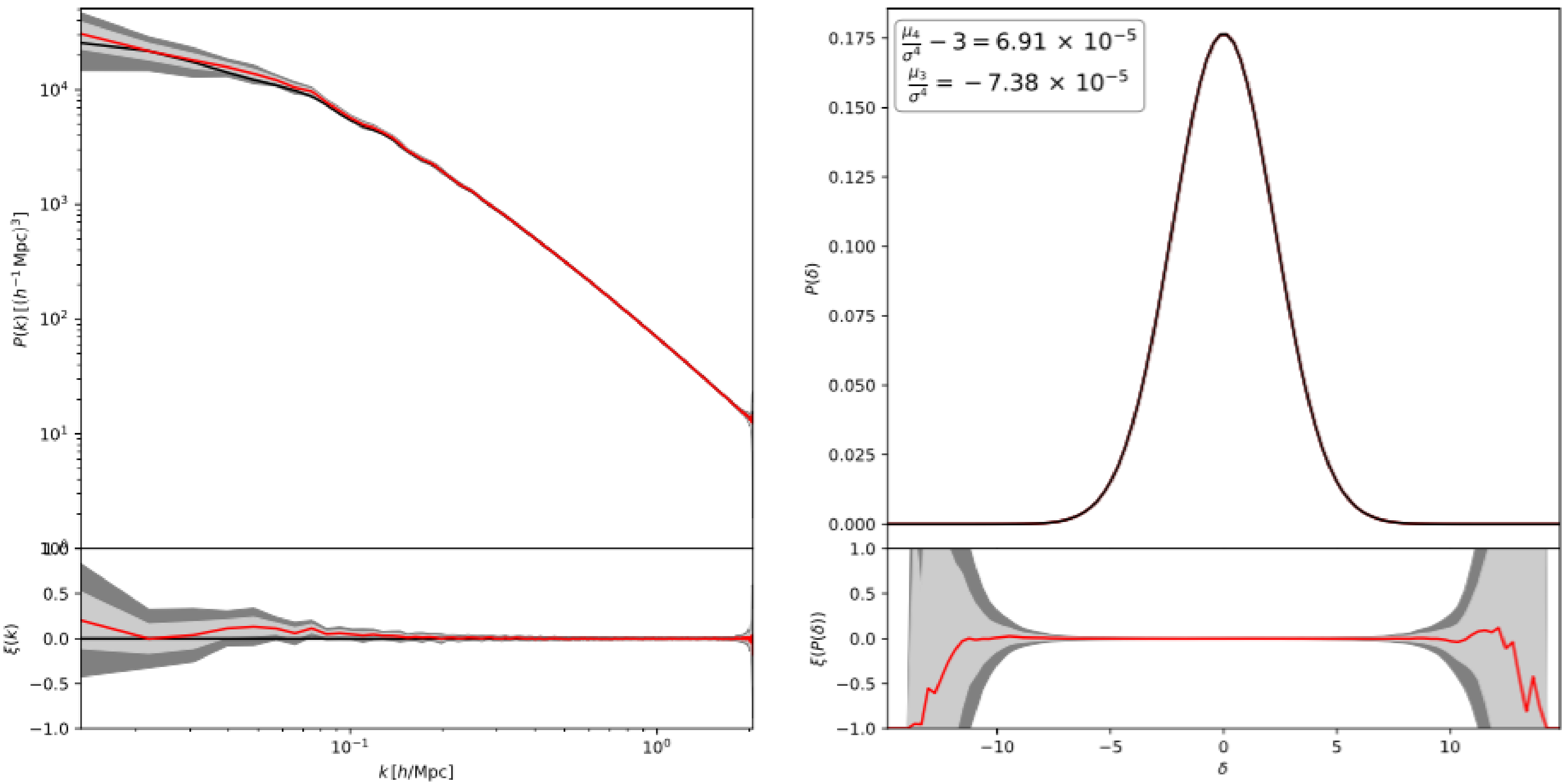
Initial conditions inferred from spectroscopic 2M++ data



Jasche & Lavaux 2018 (arXiv:1806.11117)

Posterior Power-Spectra

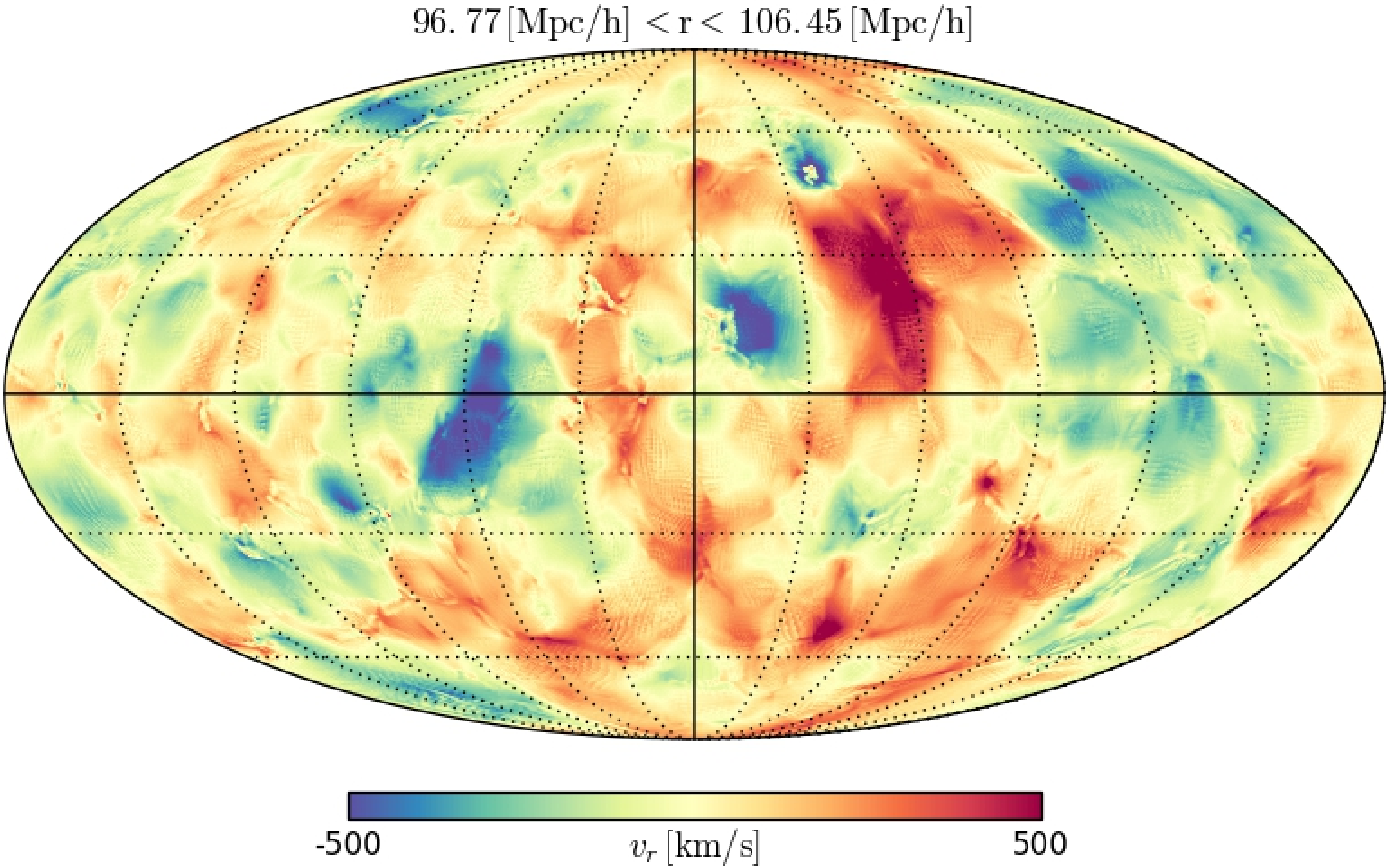
A posteriori tests of 2-pt and 1-pt functions



Unbiased inference throughout entire Fourier domain.

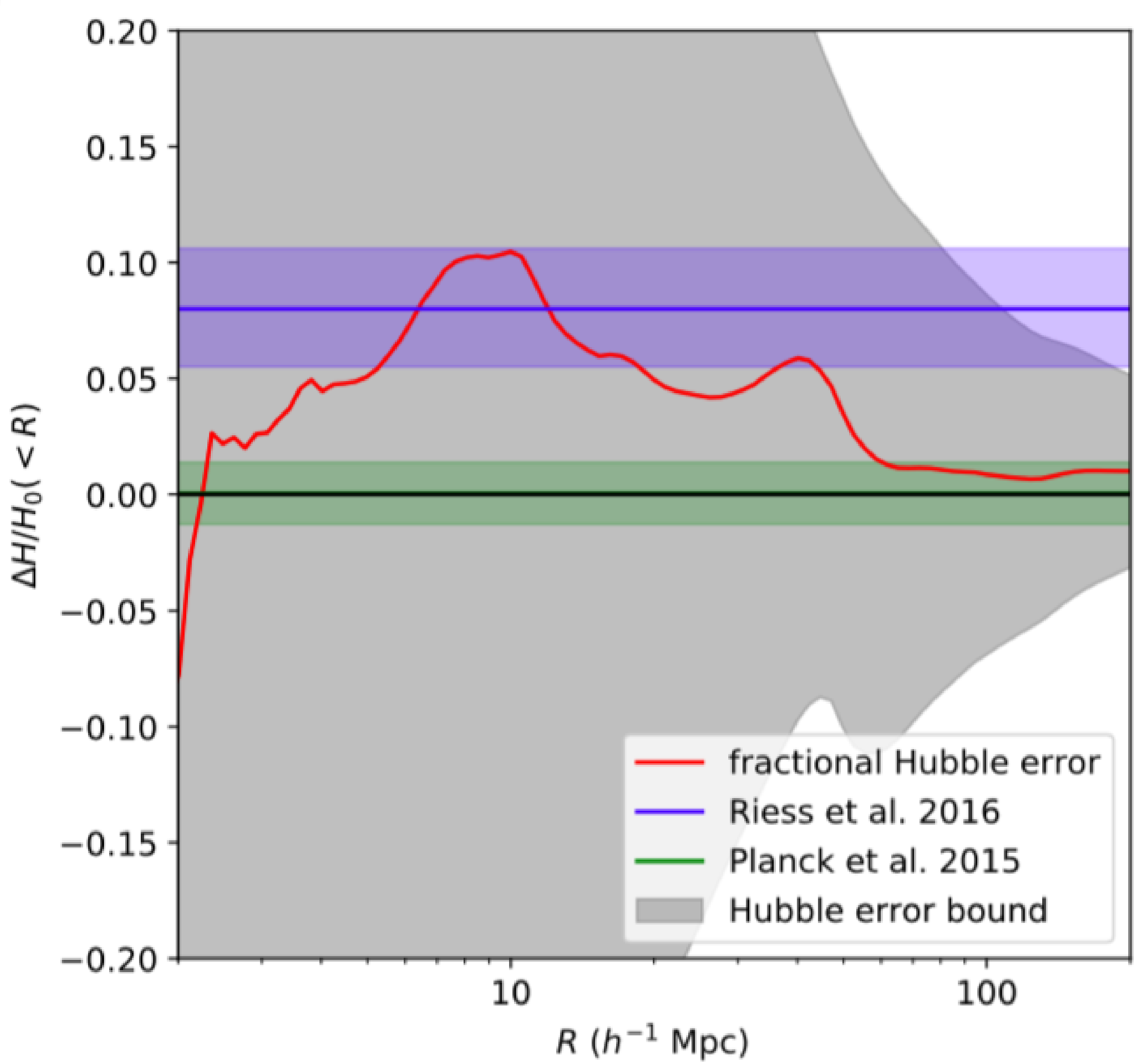
Jasche & Lavaux 2018 (arXiv:1806.11117)

Peculiar velocities and the Hubble flow



Jasche & Lavaux 2018 (arXiv:1806.11117)

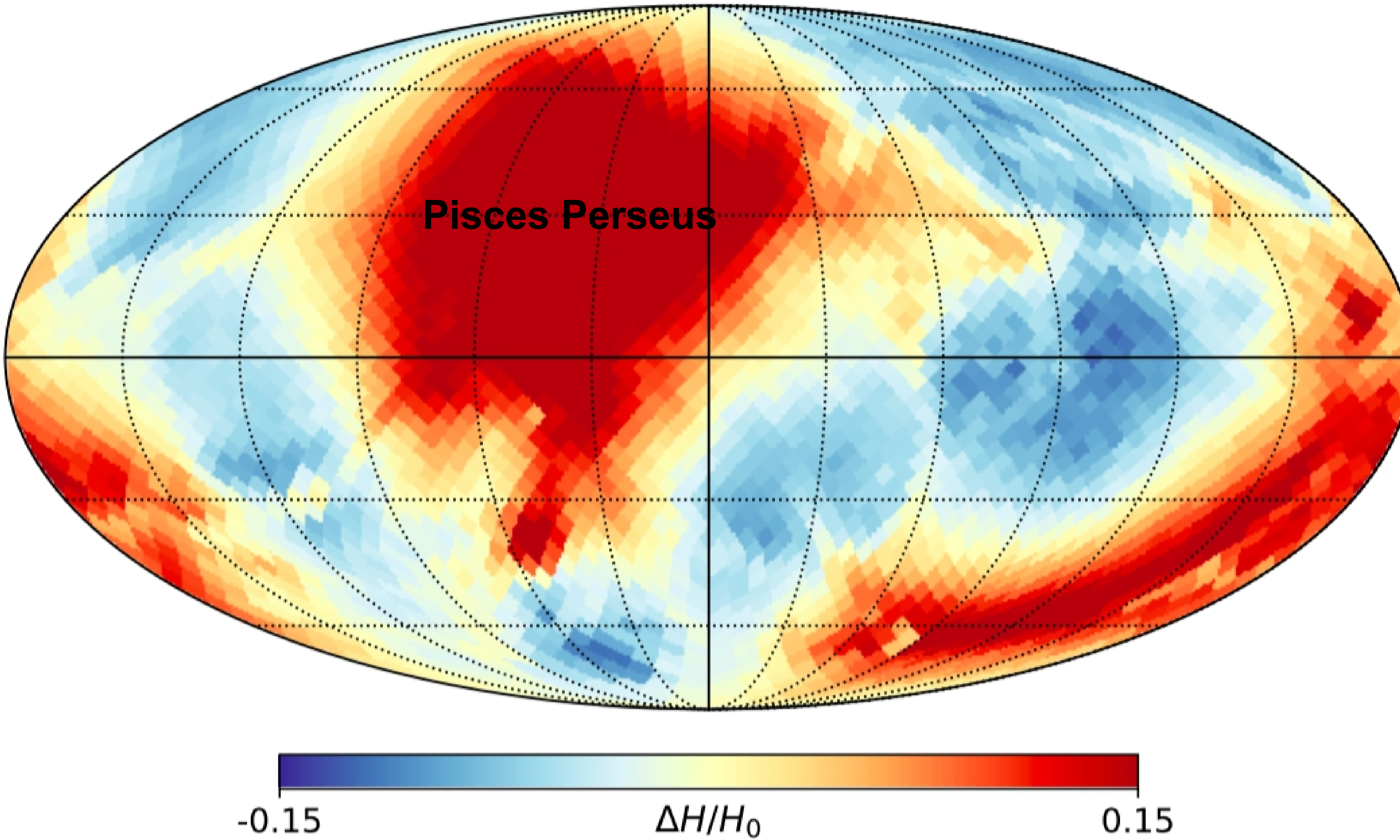
Fractional Hubble Uncertainties



Jasche & Lavaux 2018 (arXiv:1806.11117)

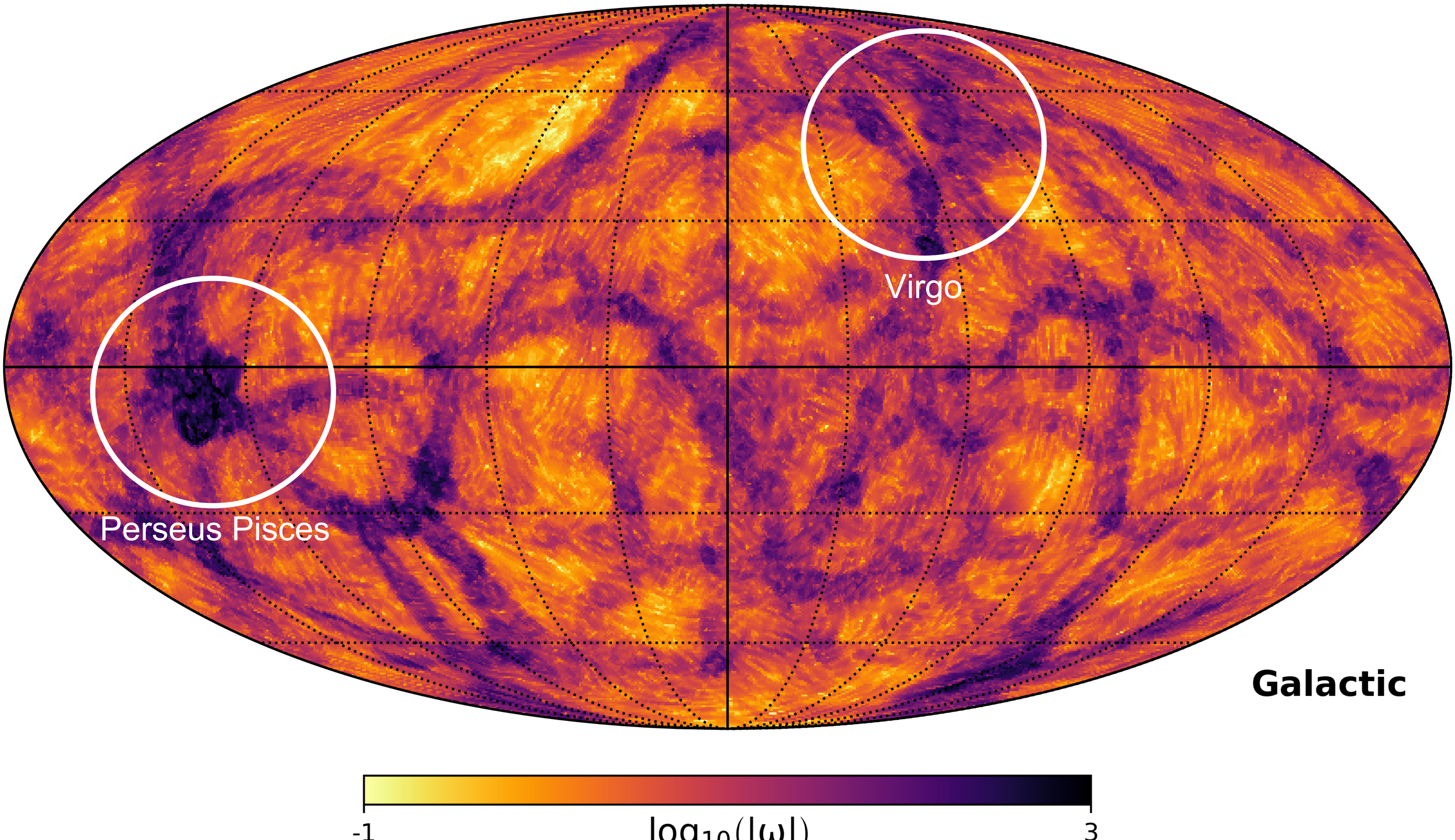
Fractional Hubble Uncertainties

Fractional Hubble Uncertainties R < 60 Mpc/h



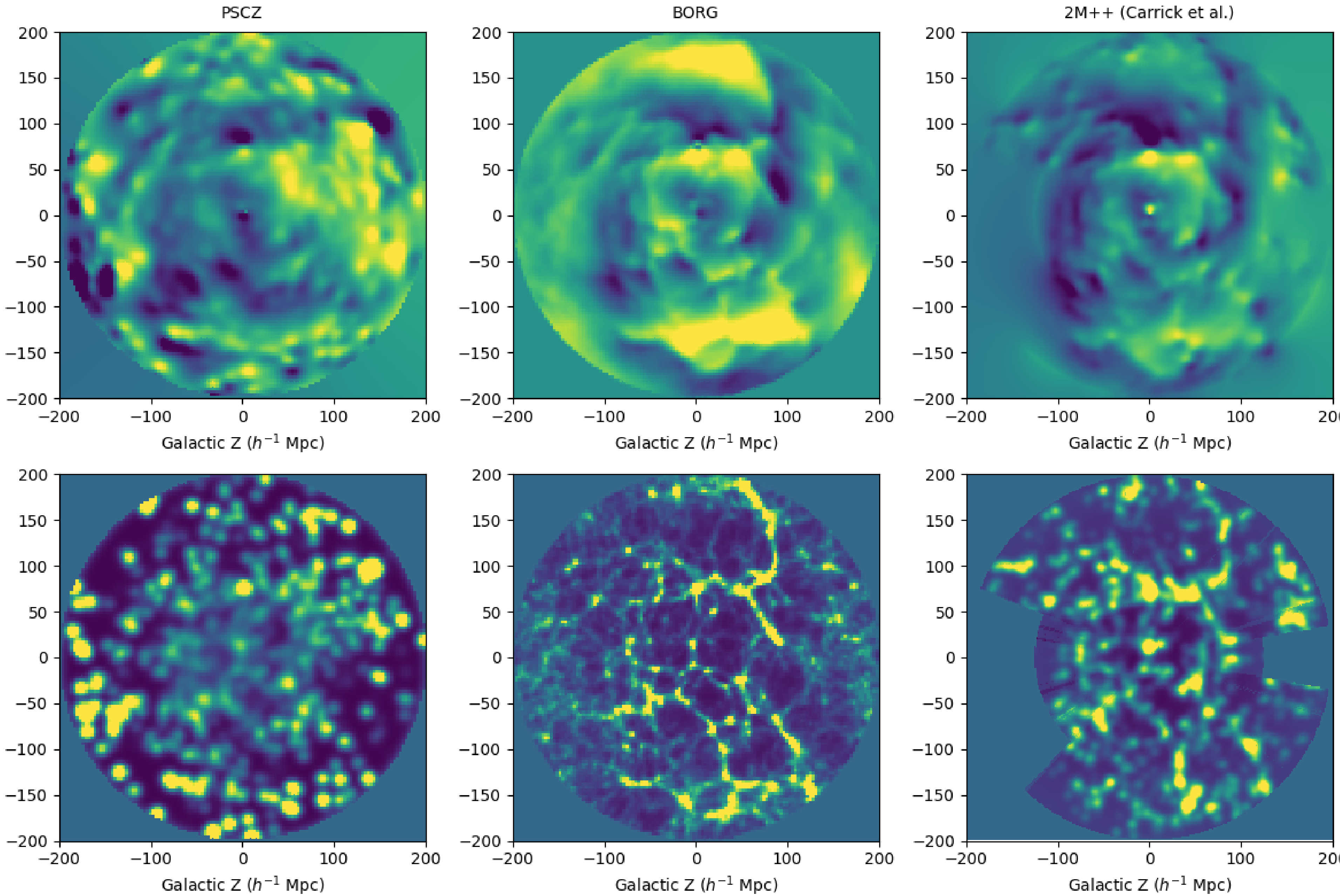
Jasche & Lavaux 2018 (arXiv:1806.11117)

Vorticity of the velocity field



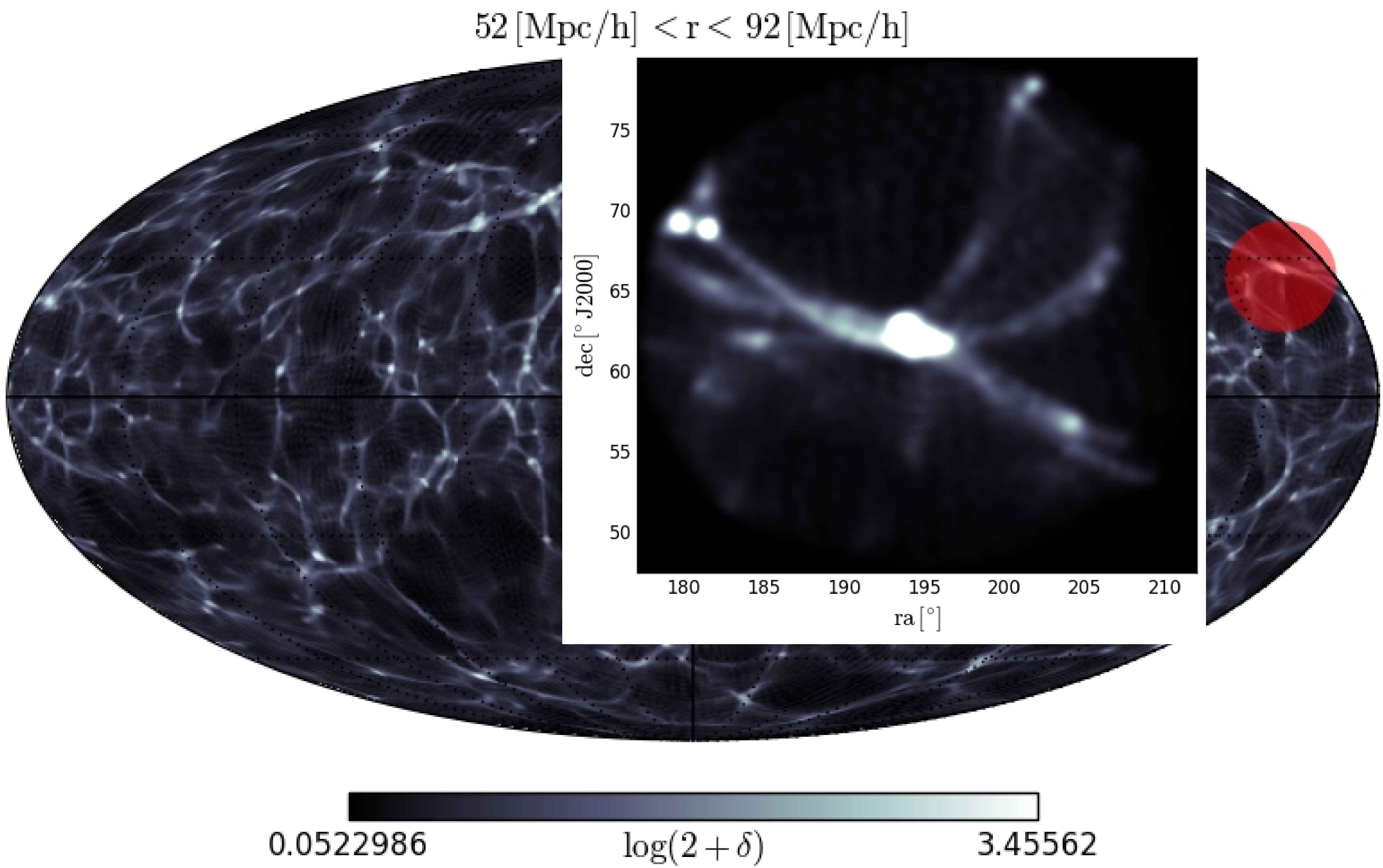
Jasche & Lavaux 2018 (arXiv:1806.11117)

Comparing local velocity fields



The Coma Cluster

The Coma Cluster



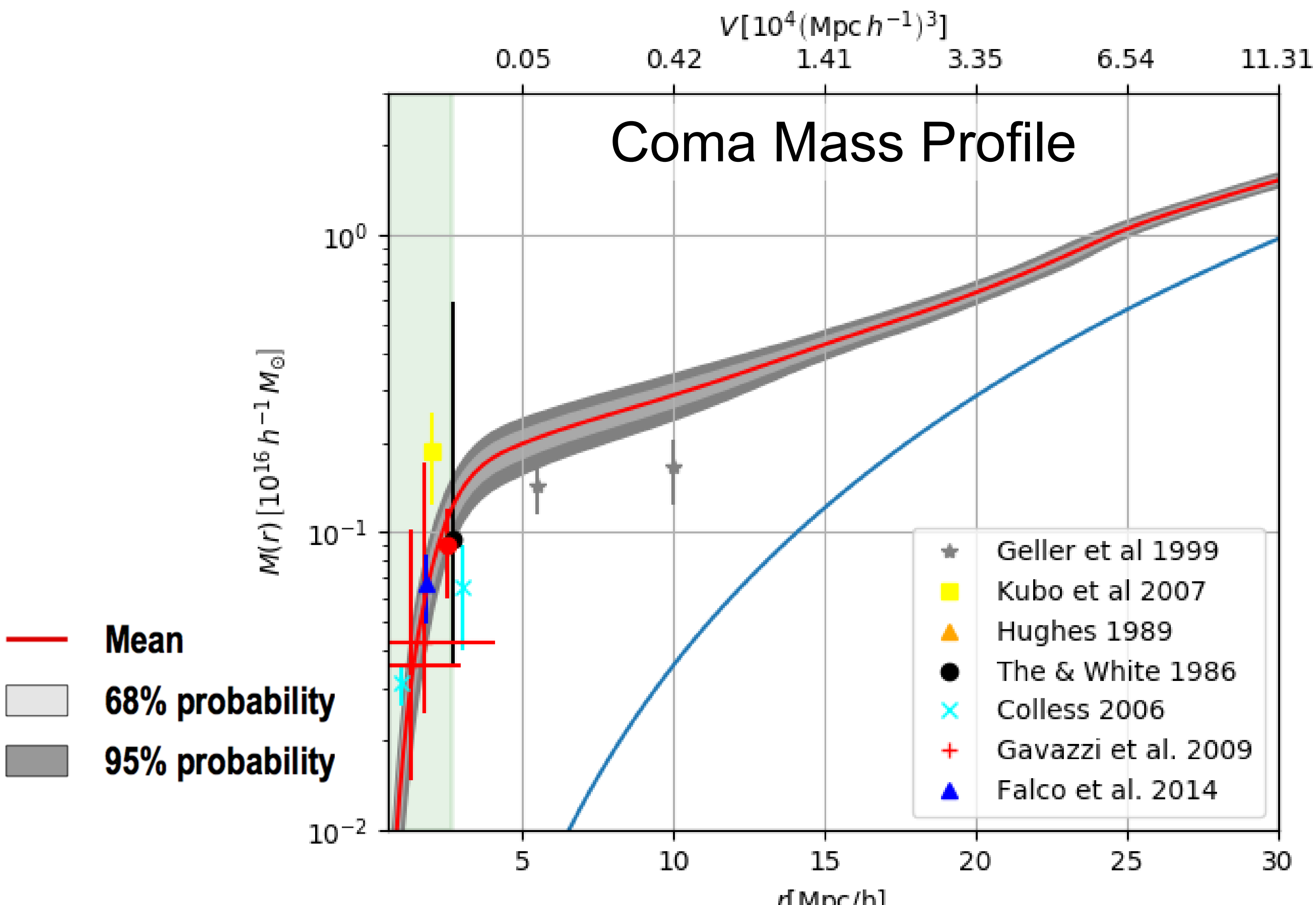
Chronography of the coma cluster

Preliminary work!

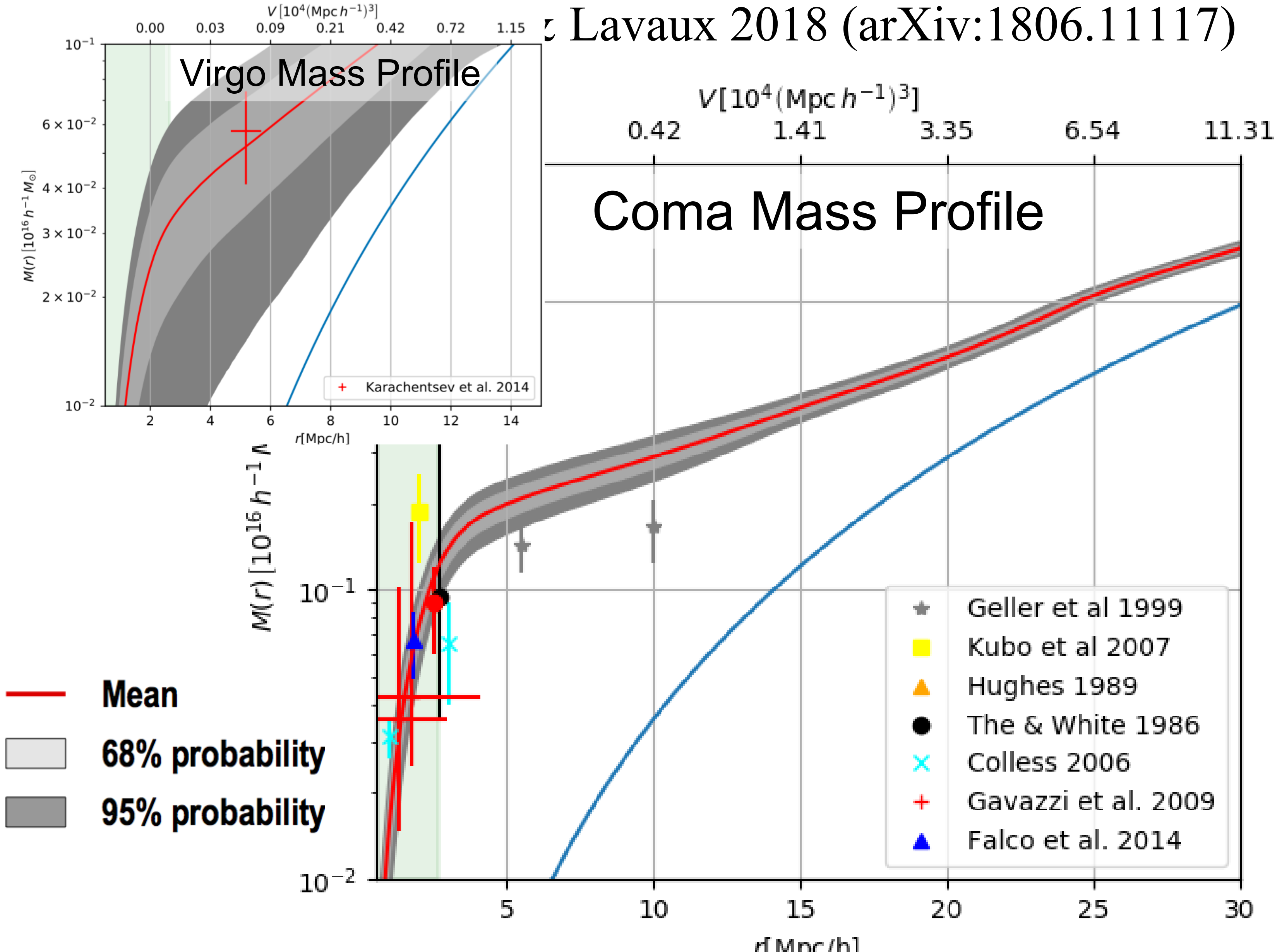
Jasche & Lavaux (in prep)

Dynamic mass estimates

Jasche & Lavaux 2018 (arXiv:1806.11117)

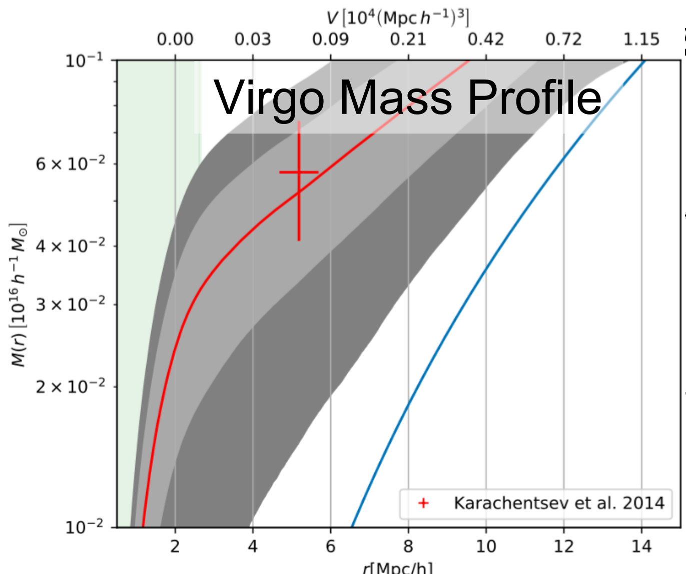


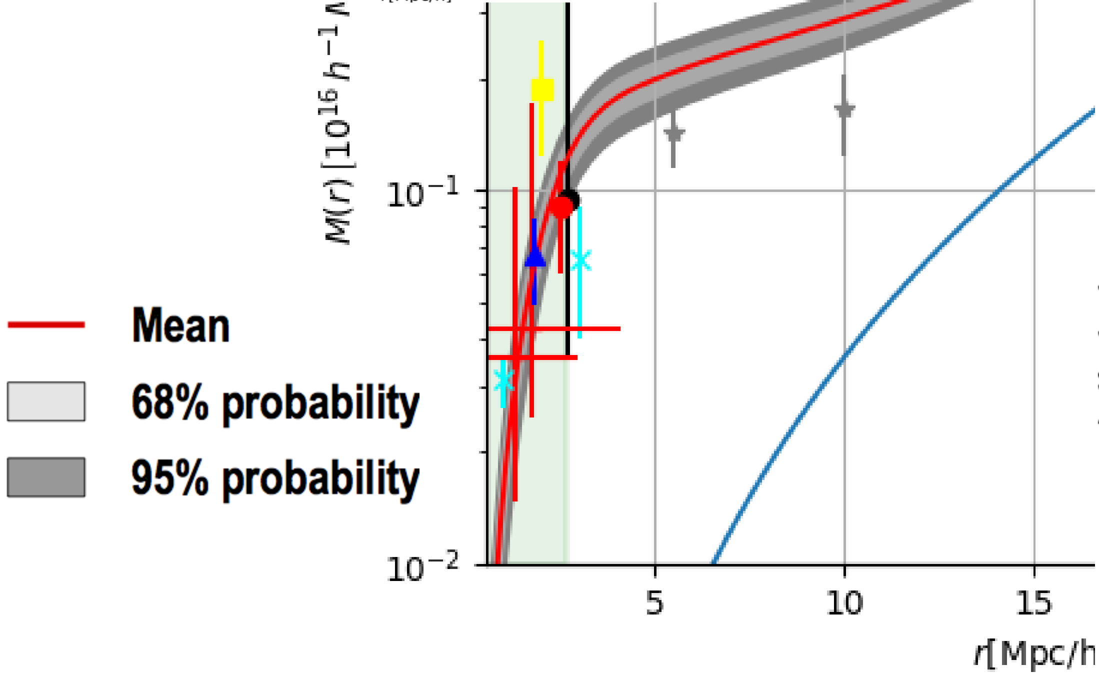
Dynamic mass estimates

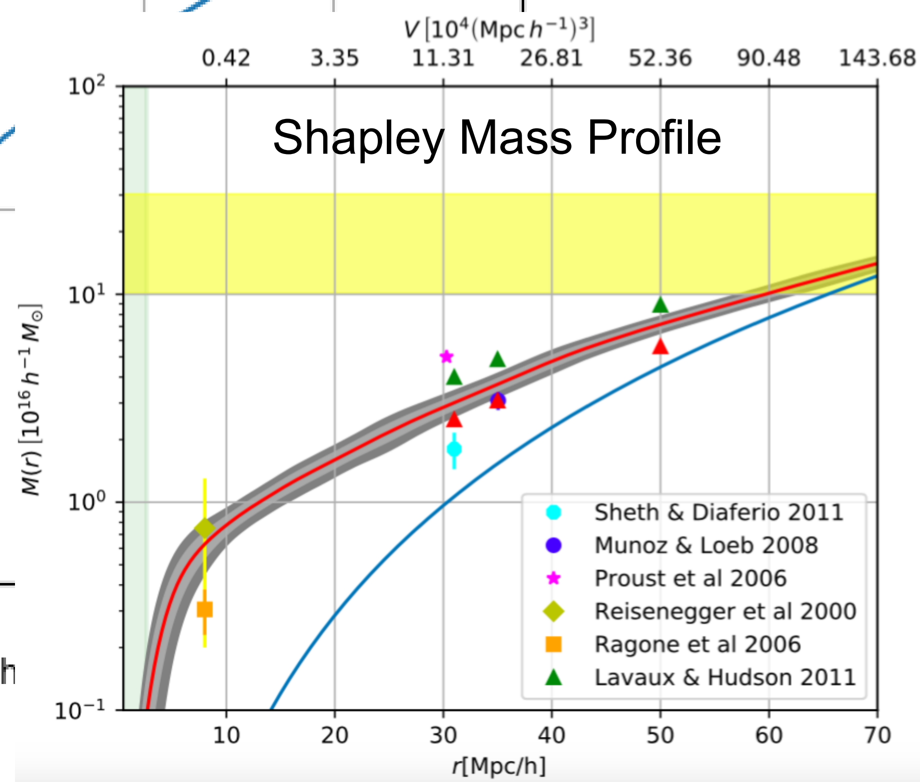


Dynamic mass estimates

z Lavaux 2018 (arXiv:1806.11117)

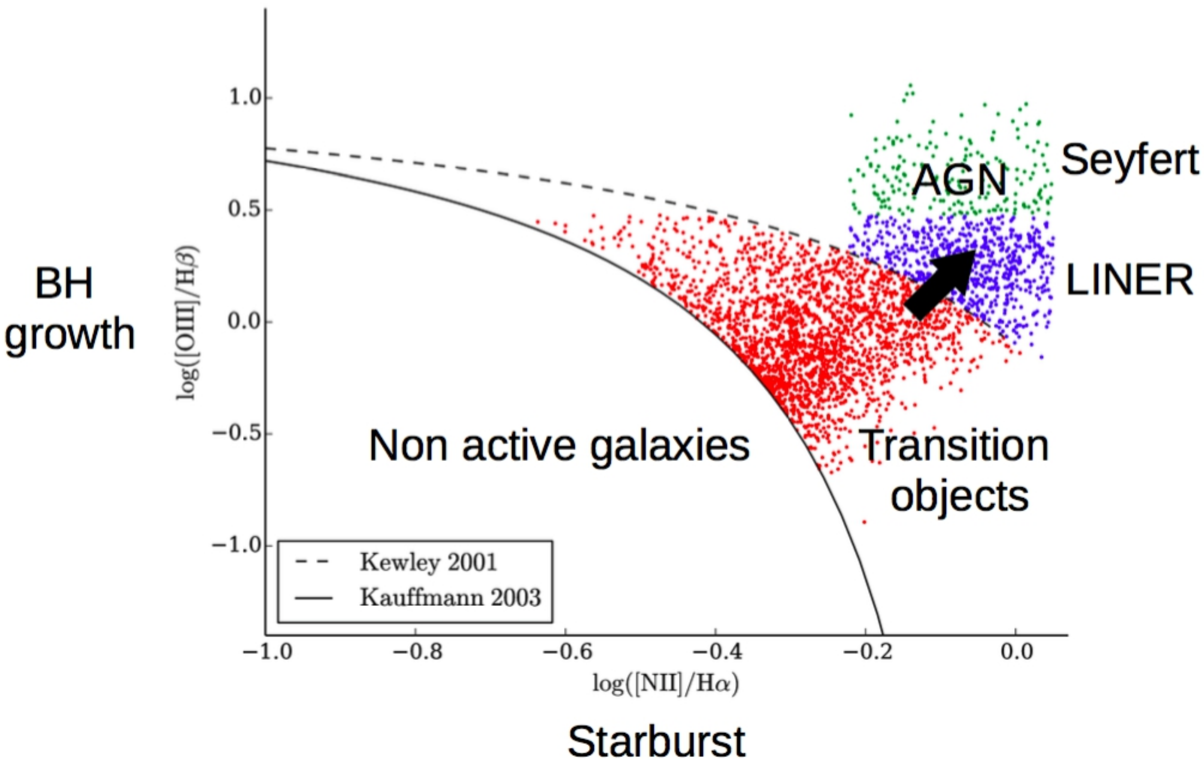




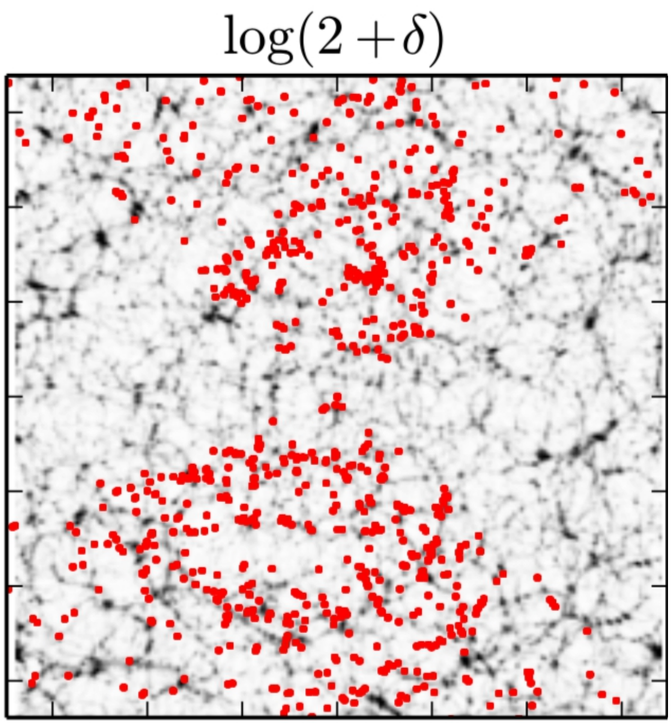


Imprints of the LSS on AGN evolution

Project by Natalia Porqueres (MPA)



Natalia Porqueres Rosa
(MPA, Garching)



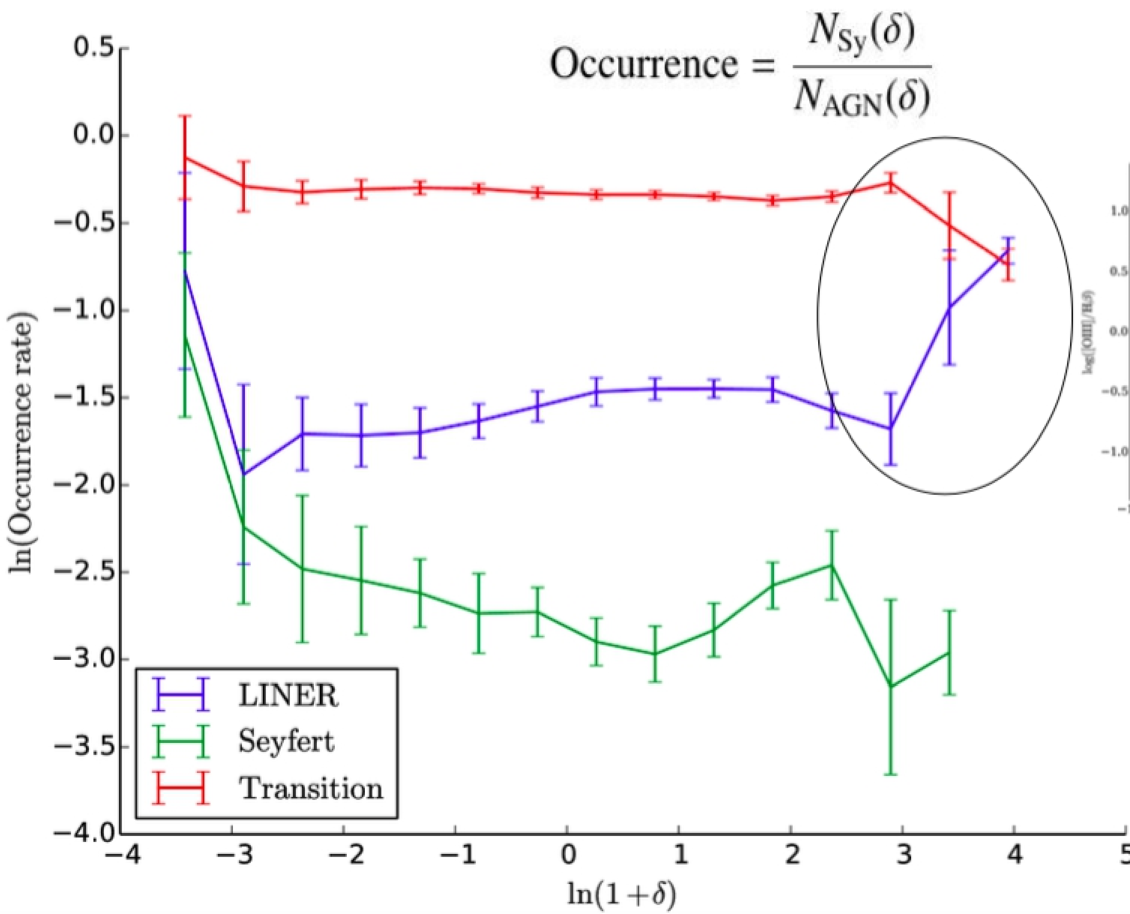
Transition Objects → **LINERs ???**

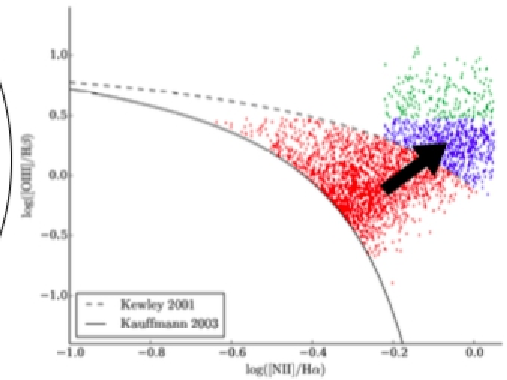
e.g. Constantin et al. (2008)

Porqueres et al. 2018 (arXiv:1710.07641)

Imprints of the LSS on AGN evolution

Project by Natalia Porqueres (MPA)





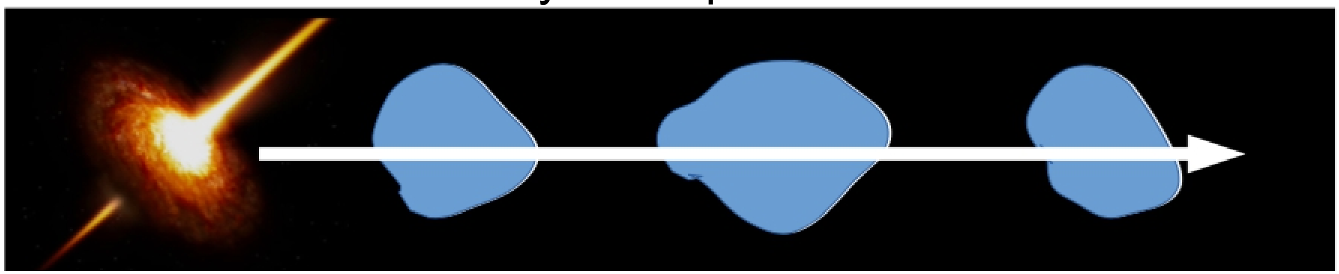
Transition objects convert to LINERs as suggested by Constantin et al. (2008)

Porqueres et al. 2018 (arXiv:1710.07641)

Bayesian inference of the high-z LSS

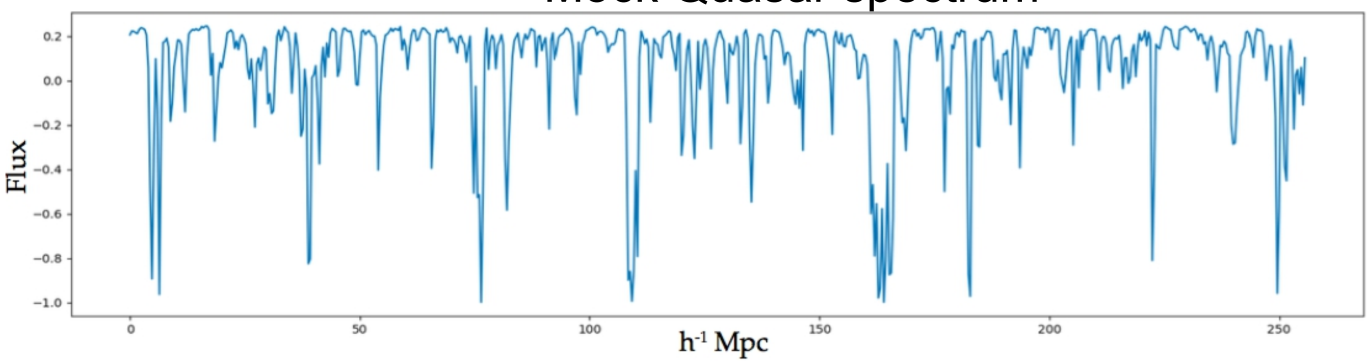
Project by Natalia Porqueres (MPA)

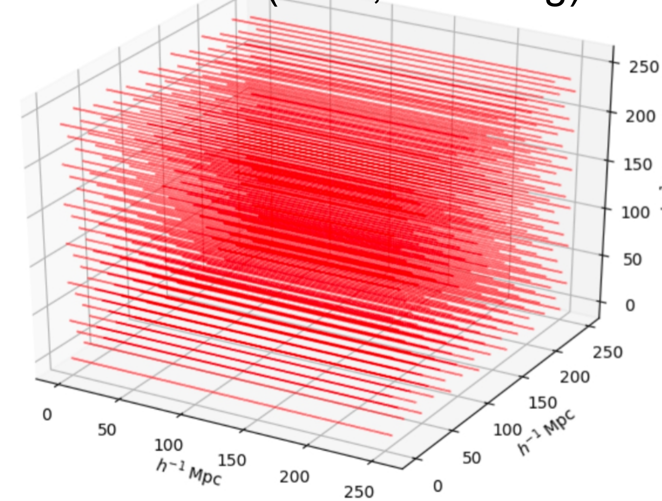
The Lyman alpha forest



Natalia Porqueres Rosa
(MPA, Garching)

Mock Quasar spectrum





Fluctuating Gunn-Peterson Approximation (FGPA)

$$F = e^{-A(1+\delta)^{\beta}}$$

$$A \propto (1+z)^6 T_0^{-0.7} \Gamma^{-1}$$

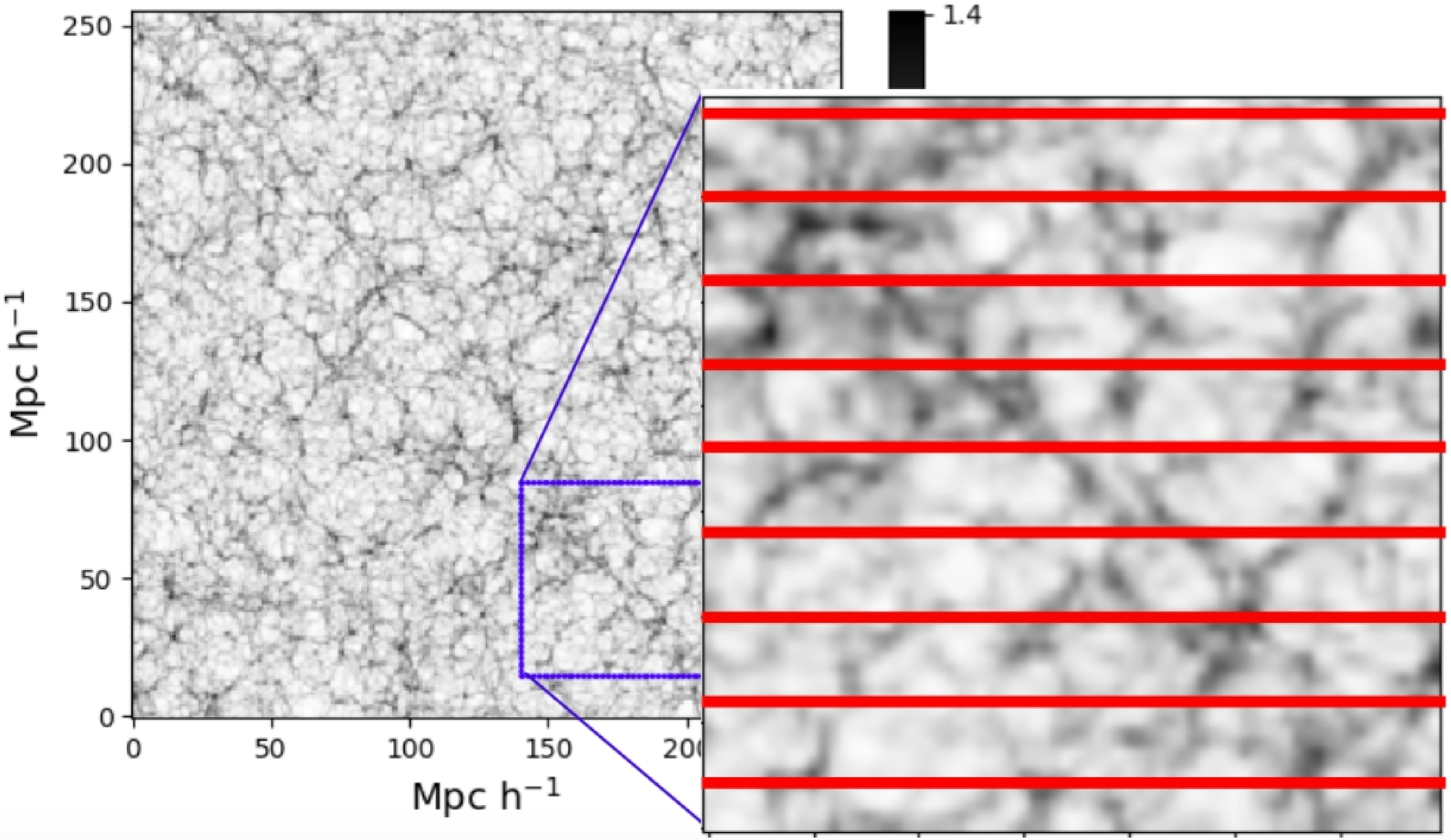
$$\beta = 2 - 0.7(\gamma - 1)$$

Porqueres et al. 2018 (in prep)

Bayesian inference of the high-z LSS

Preliminary results!

Inferring 3d field from 1D los:

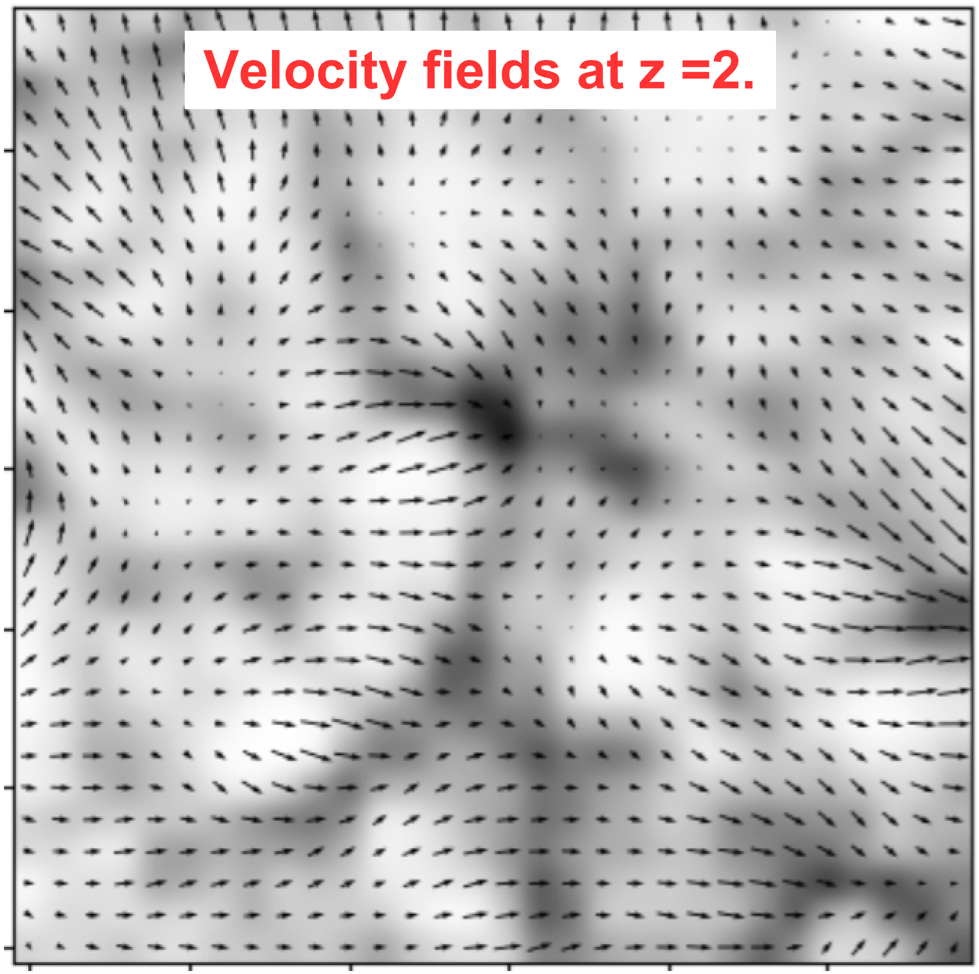


Porqueres et al. 2018 (in prep)

Bayesian inference of the high-z LSS

Some sample applications:

Preliminary results!

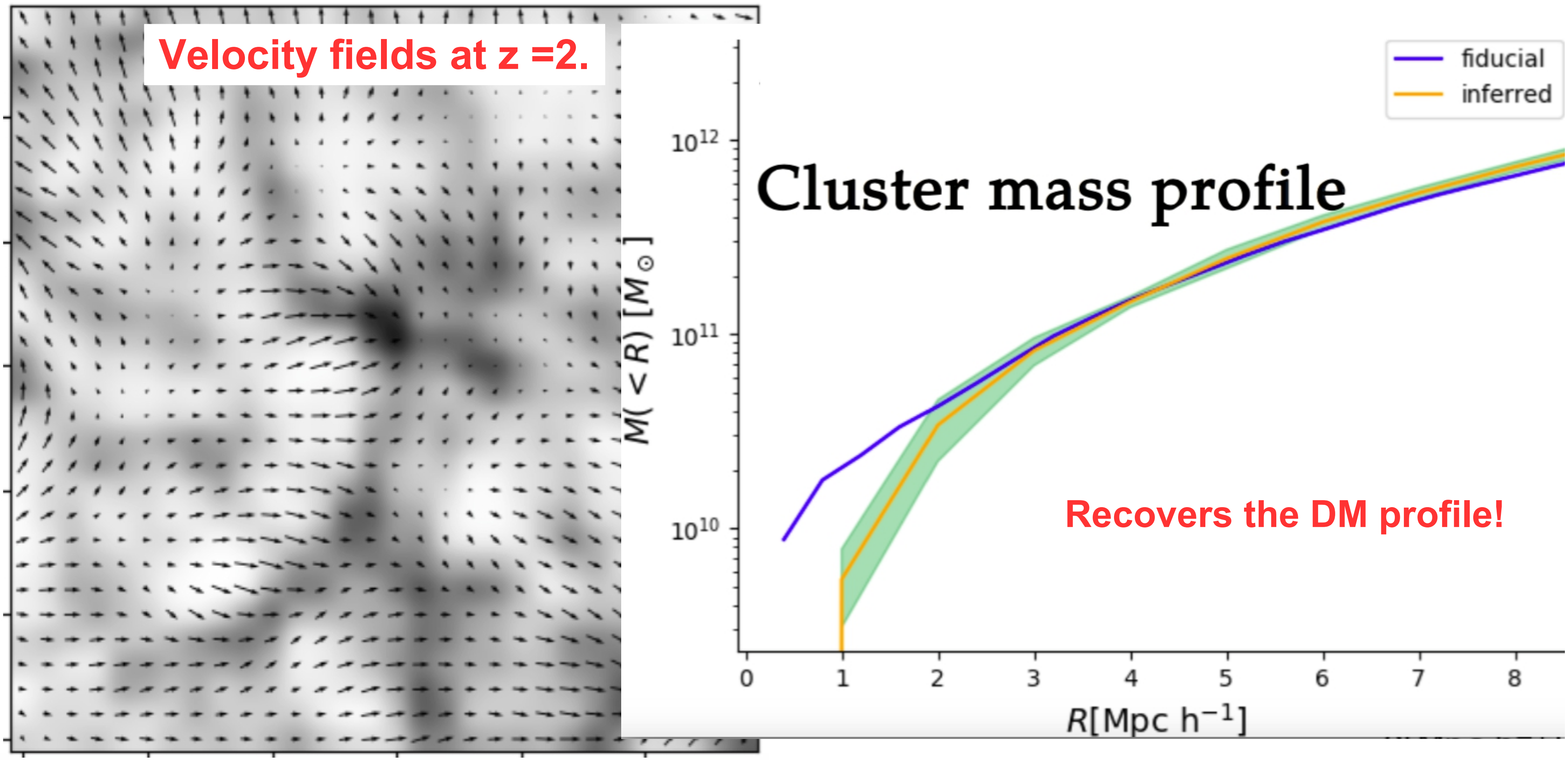


Porqueres et al. 2018 (in prep)

Bayesian inference of the high-z LSS

Some sample applications:

Preliminary results!

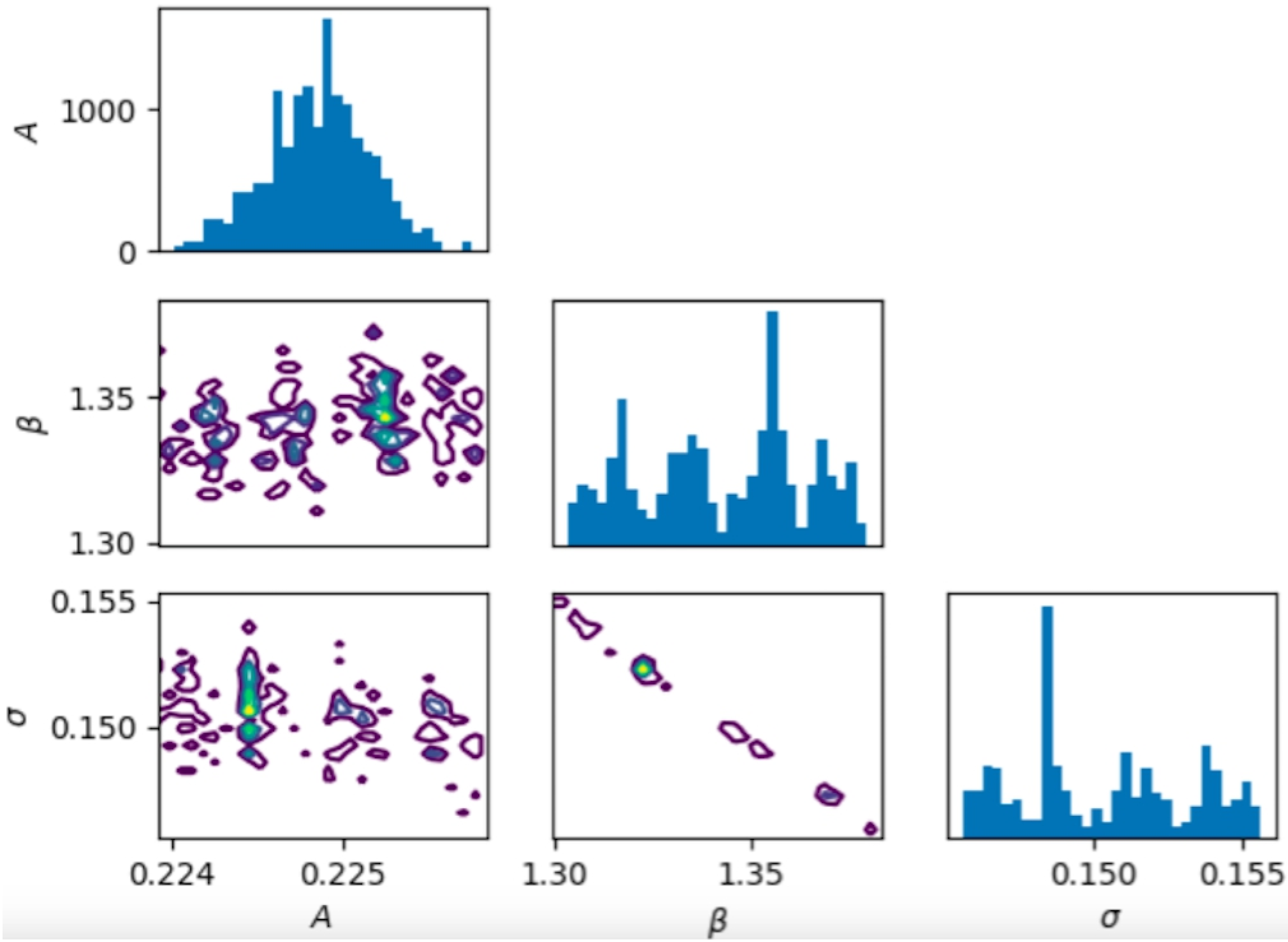


Porqueres et al. 2018 (in prep)

Bayesian inference of the high-z LSS

Detailed MCMC treatment of Baryonic meta parameters

Preliminary results!



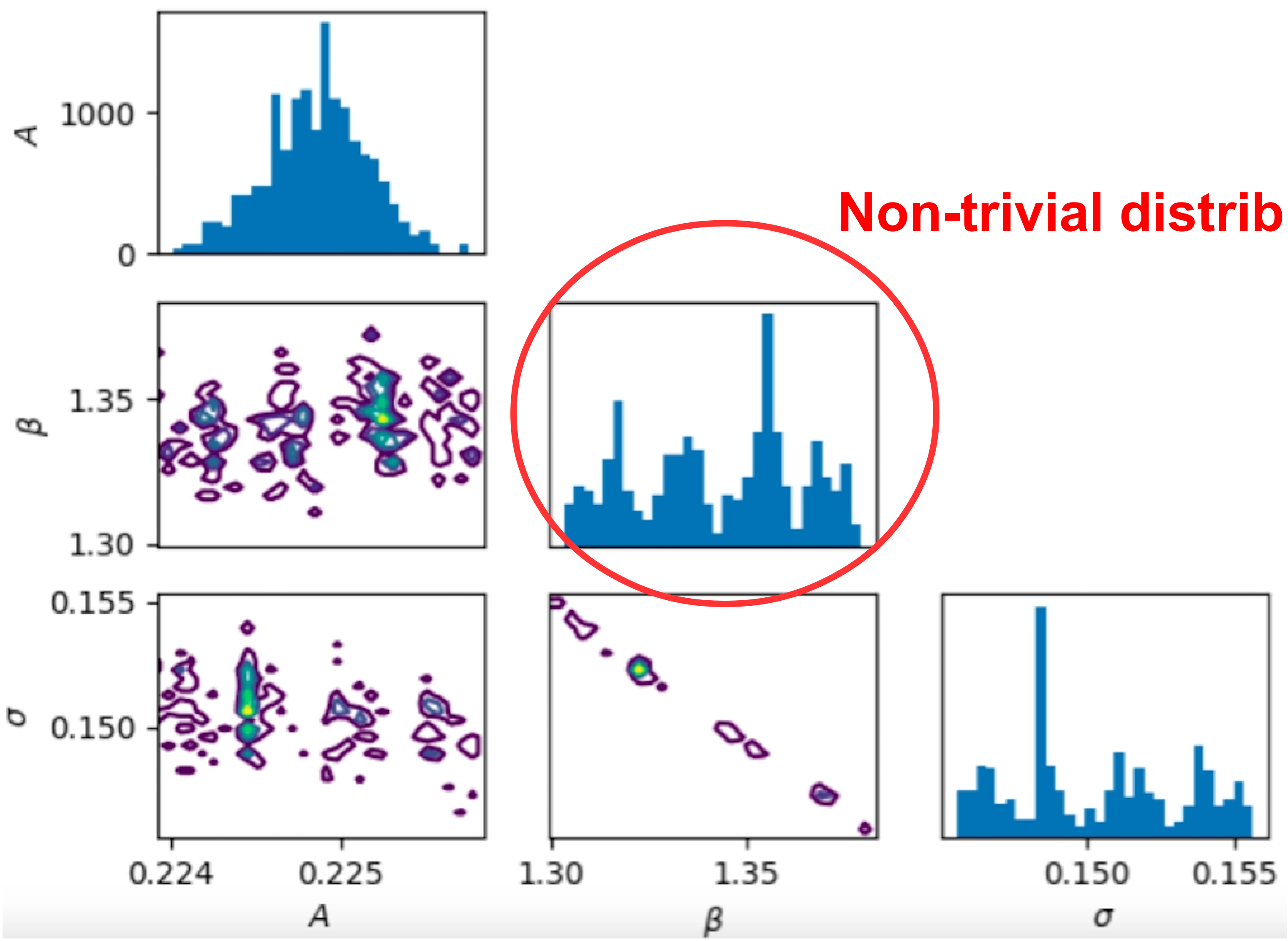
Porqueres et al. 2018 (in prep)

Bayesian inference of the high-z LSS

Detailed MCMC treatment of Baryonic meta parameters

Preliminary results!

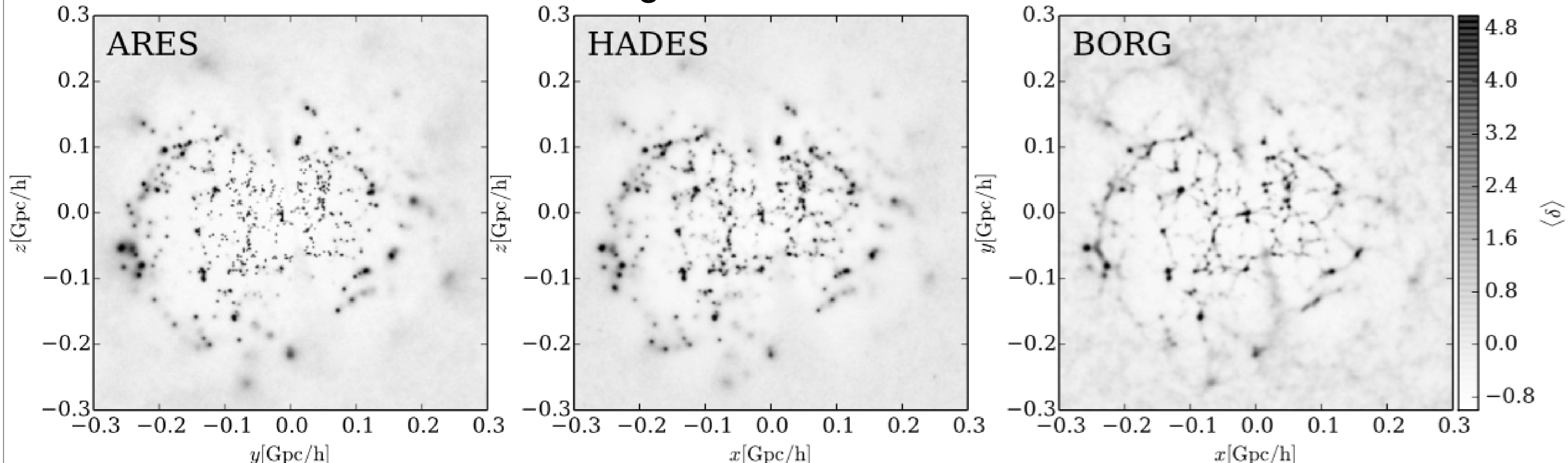
Non-trivial distributions!!!



Porqueres et al. 2018 (in prep)

Comparing inference schemes

Gaussian | Log-normal-Poisson | 2LPT-Poisson



a.k.a: Wiener-filtering
Zaroubi et al. 1994
Erdogdu et al. 2004
Kitaura & Ensslin 2008
Grannet et al. 2015

log-normal-filtering
Kitaura 2010
Jasche & Kitaura 2010

Jasche&Wandelt 2012

Which scheme performs best?

Ask the data!

$$A_{ij} = \ln\left(\mathcal{P}(d|\delta_i)\right) - \ln\left(\mathcal{P}(d|\delta_j)\right)$$

	ARES	**HADES**	**BORG**
ARES	0	-219580.31	-383482.25
HADES	219580.31	0	-163901.94
BORG	383482.25	163901.94	0.

Summary & Conclusion

BORG combines physical modeling with data science:

- Dynamical modeling accounts for non-Gaussian statistics
- Flexible data modeling via HMC and block sampling
- Detailed treatment of systematics and uncertainties
- Solves complex data models in high-d

Scientific results:

- Characterization of initial conditions
- Accurate & Detailed reconstructions of the DM field
- Dynamical reconstructions (velocity + vorticity)
- **We arrive at a consistent dynamical picture of our Universe**


Thank You!
Visit us at:
aquila-consortium.org