Cosmology with the Nearby Universe through full statistical reconstruction of wide galaxy surveys

Institut d'Astrophysique de Paris (CNRS)

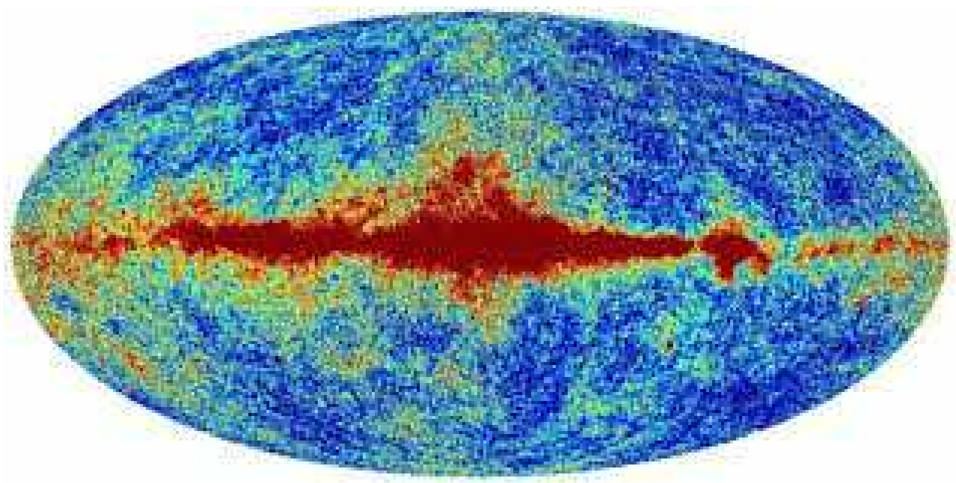
Guilhem Lavaux

Collaboration with

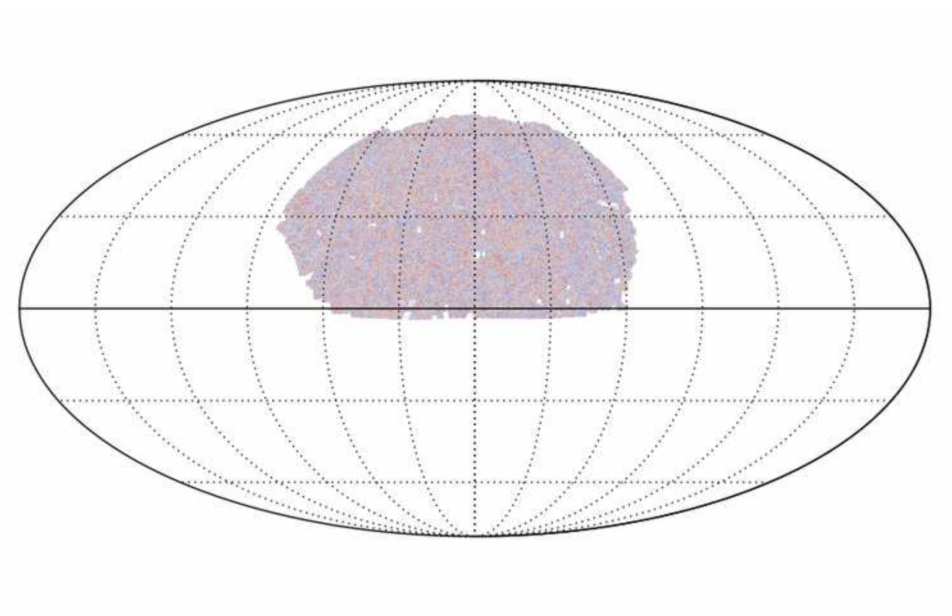
- Jens Jasche (TUM/ExC)
- Doogesh Kodi Ramanah (IAP/ILP)
- Benjamin Wandelt (IAP/ILP/UPMC)

Cosmology with large scale structures

Dirty CMB data

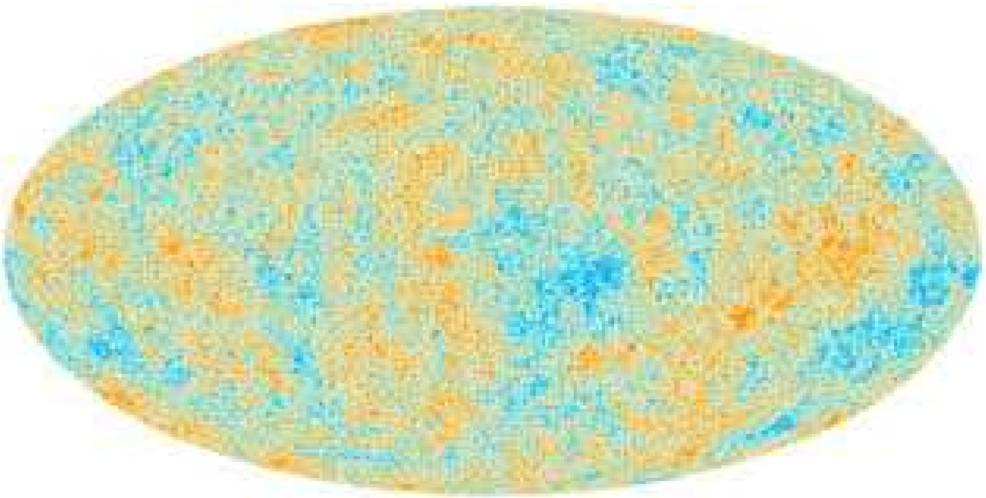


Dirty Large scale structure data



Cosmology with large scale structures

Clean CMB data



Dirty Large scale structure data

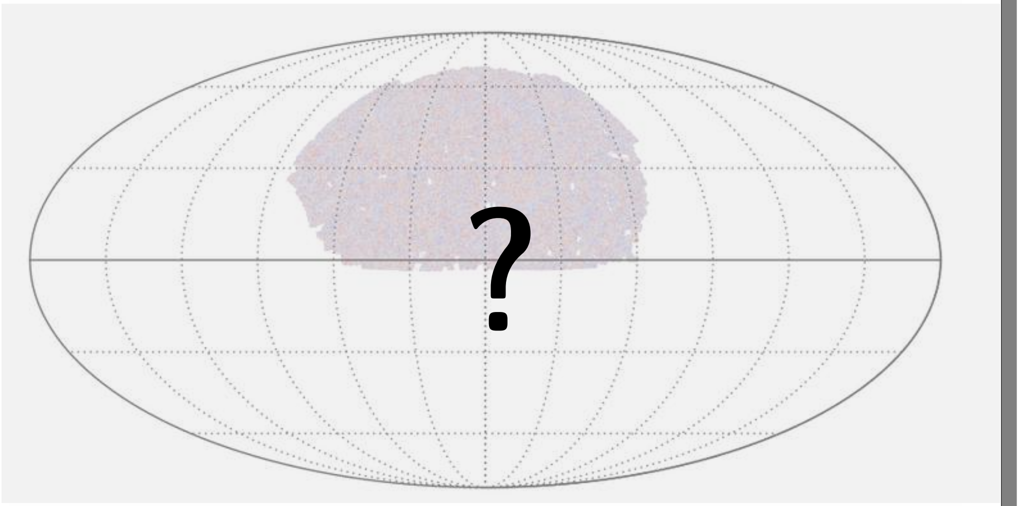
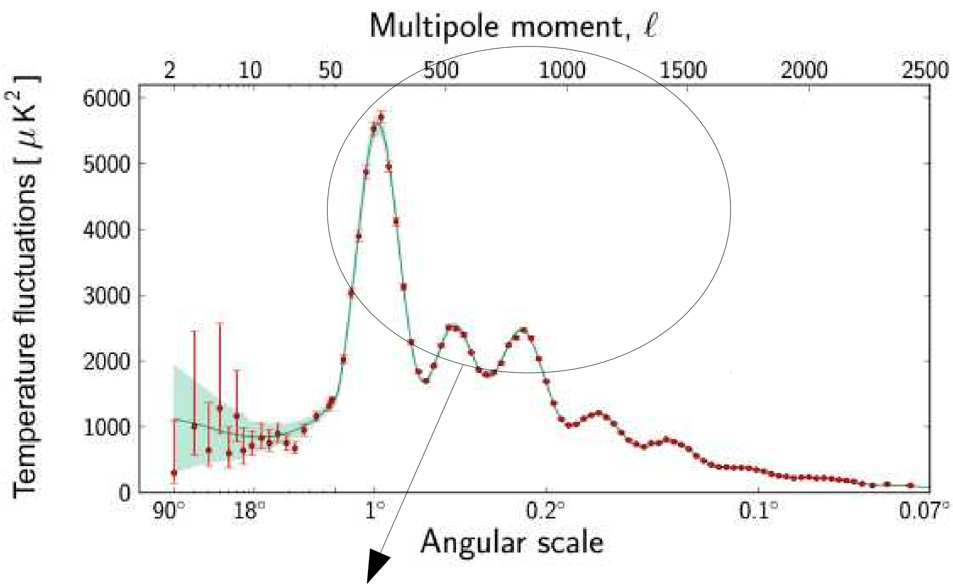


Image credit: Planck collaboration, SDSS

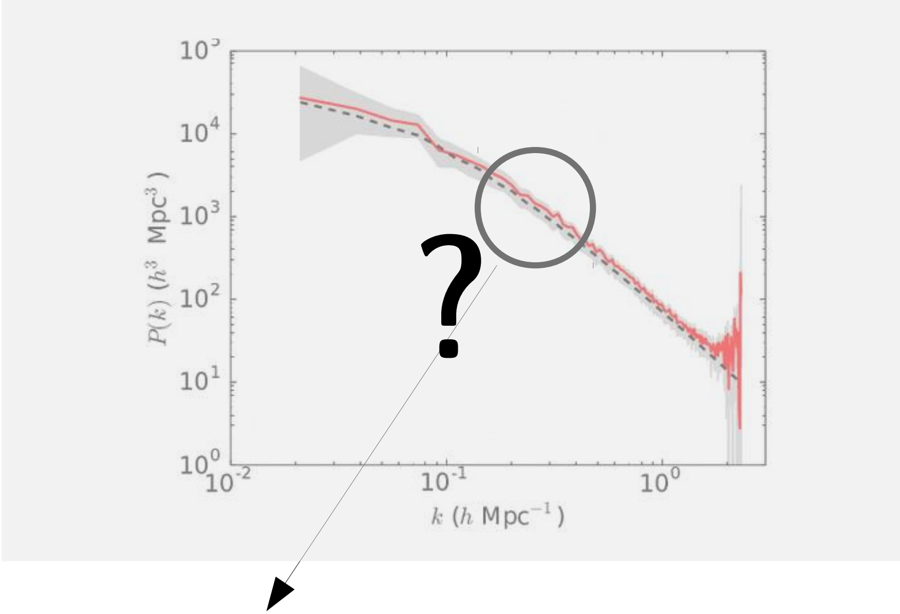
Cosmology with large scale structures

Clean CMB data



Baryonic Acoustic Oscillations

Dirty Large scale structure data



Baryonic Acoustic Oscillations

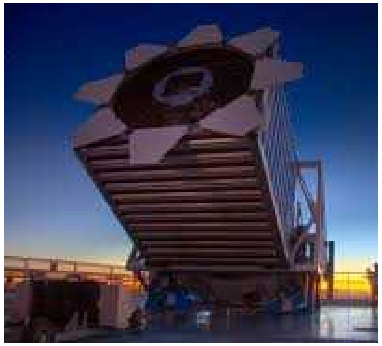


Image credit: Planck collaboration, SDSS

Principles of statistical reconstruction of large scale structure (LSS)

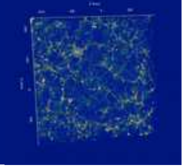
Algorithm for REconstruction and Sampling

Bayesian Origins Reconstruction from Galaxies

ARES

BORG

Statistical reconstruction of LSS



Galaxies are counted in 3d cells

$$P({\color{red}N_{obs}}|\lambda) \propto \lambda^{\color{red}N_{obs}} \exp(-\lambda)$$

Poisson probability on large scales

Noise is signal dependent!

Mean density

$$\lambda_p = {\color{blue}R_p}{\color{green}\bar{N}}(1 + \delta_{\mathrm{galaxies},p})$$

Selection

Linear response

$$1 + \delta_{\mathrm{galaxies},p} = {\color{purple}f}(1 + \delta_{\mathrm{matter},p})$$

e.g. bias function

Local differentiable dependency on matter field

$$\delta_{\mathrm{matter},p} = \mathcal{M}_p(\{\delta_{\mathrm{ic},p}\})$$

forward model for dynamics

Principles of statistical reconstruction of large scale structure (LSS)

Dealing with Foregrounds

Algorithm for REconstruction and Sampling

ARES

ARES3: posterior for the linear model

$$P(N_{obs}|\lambda) \propto \exp((N_{obs} - \lambda)^2/\lambda) \qquad f(\delta) = \delta \qquad \mathcal{M}_p(\{\delta\}) = \delta_p$$

Linear response operator (mask, radial selection)

bias

« Mean » density of tracers

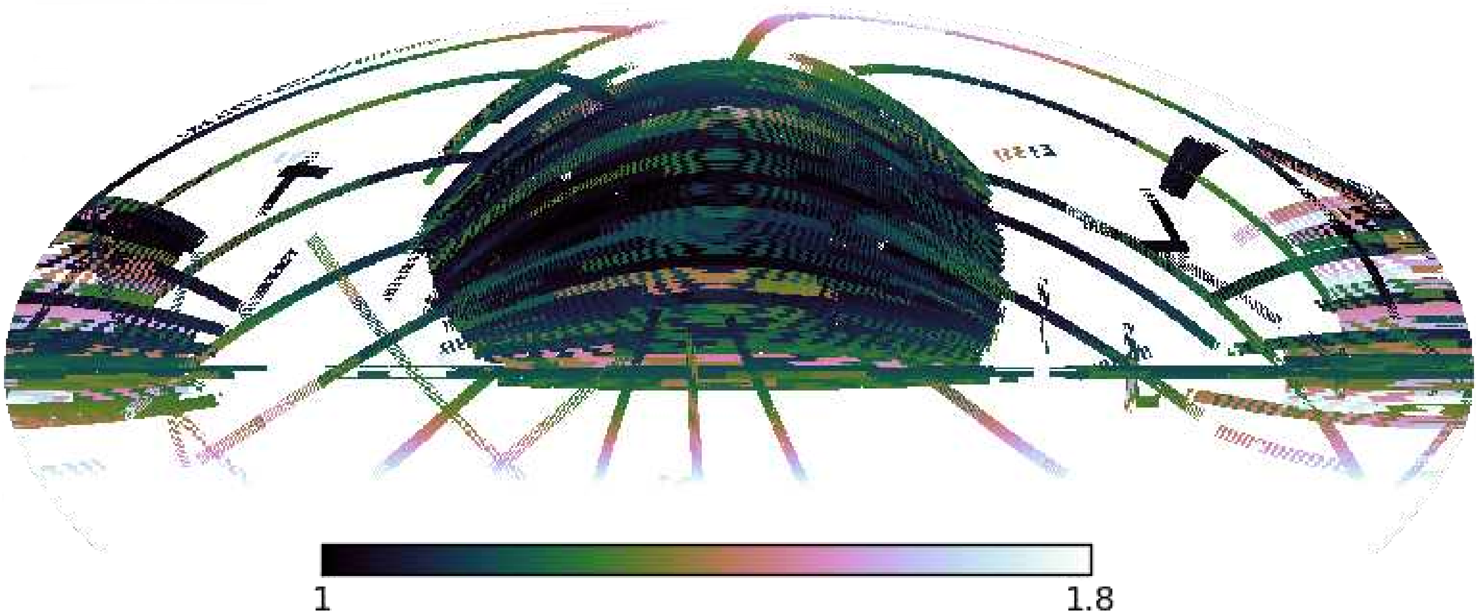
$$-\log P(\delta, b, \tilde{N}) = A + \sum_i \frac{\left(N_i - \tilde{N} R_i (1 + b\delta_i)\right)^2}{2\tilde{N} R_i} + \sum_k \frac{|\hat{\delta}_k|^2}{2P(k)}$$

NGP binned data

Gaussian probability

Gaussian prior

Some foregrounds for SDSS3

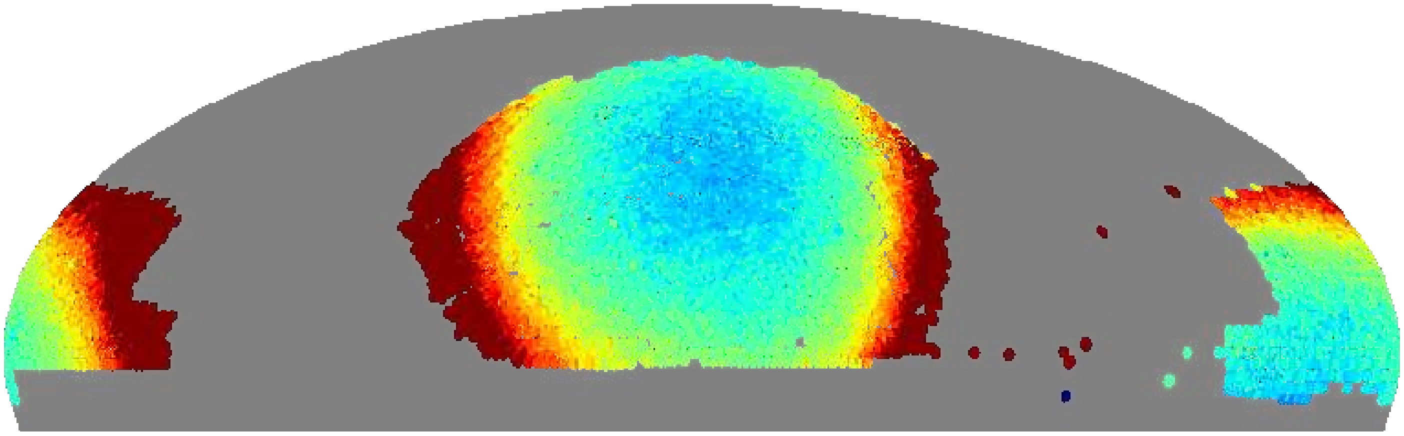


Airmass

(absorption)

Star density

(contamination and absorption)



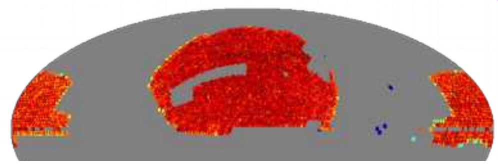
Foregrounds: 1st order correction

Fractional contribution of foregrounds

$$M_p = \frac{N_{\text{spectro}}}{N_{\text{real galaxies}}} \frac{N_{\text{real galaxies}}}{N_{\text{targets}}} = R_p F_p$$

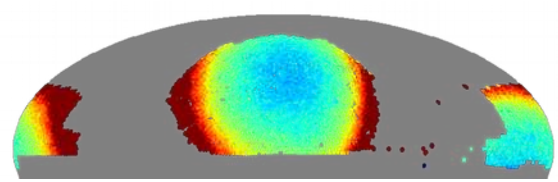
Real (unknown) linear selection

Actual mask from data =



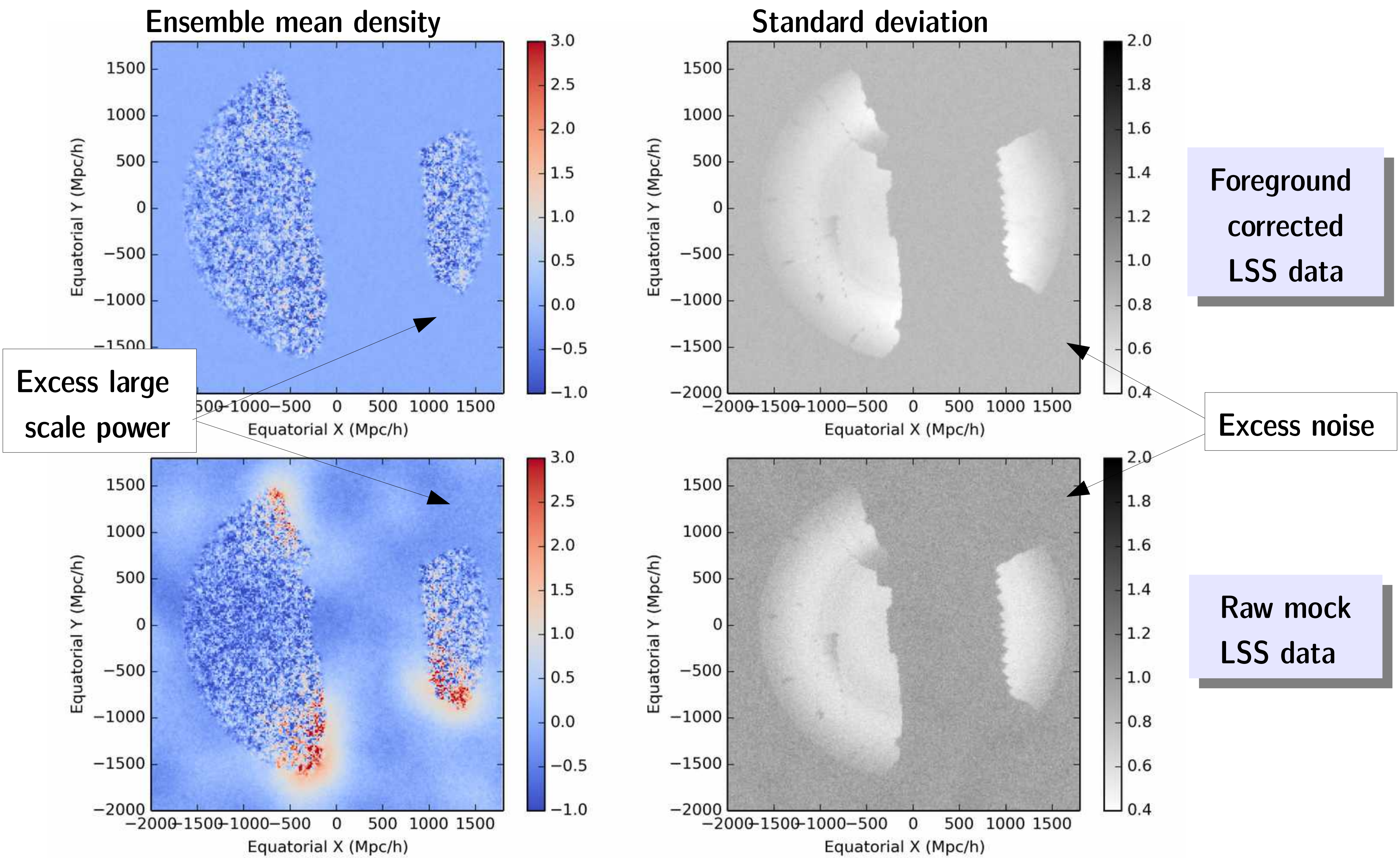
$$R_p = \underbrace{\left[\prod_{i=1}^{N_{\text{fg}}} (1 - \alpha_i G_{i,p})\right]}_{1/F_p} M_p$$

Template =



Free parameter

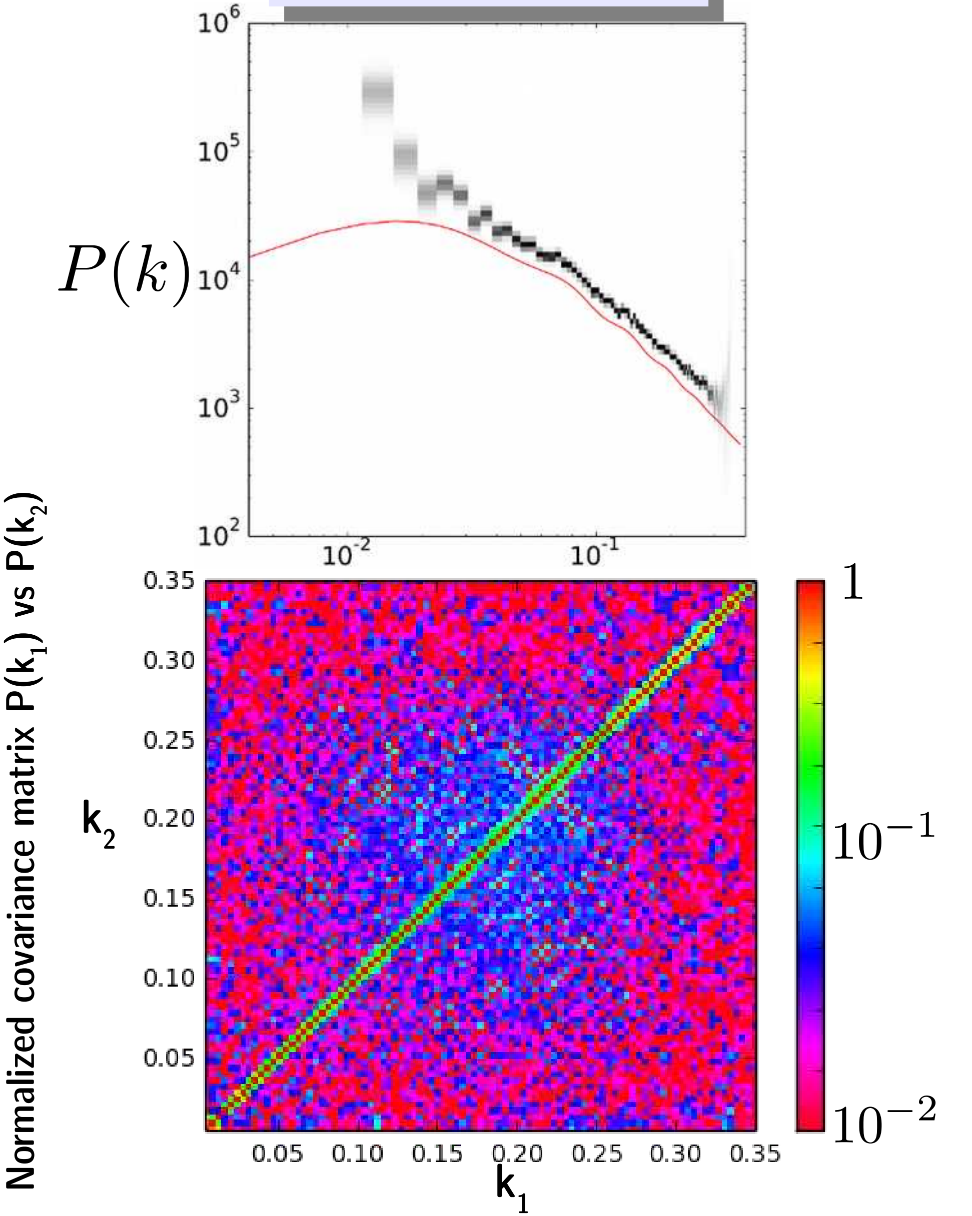
Example on a mock SDSS3



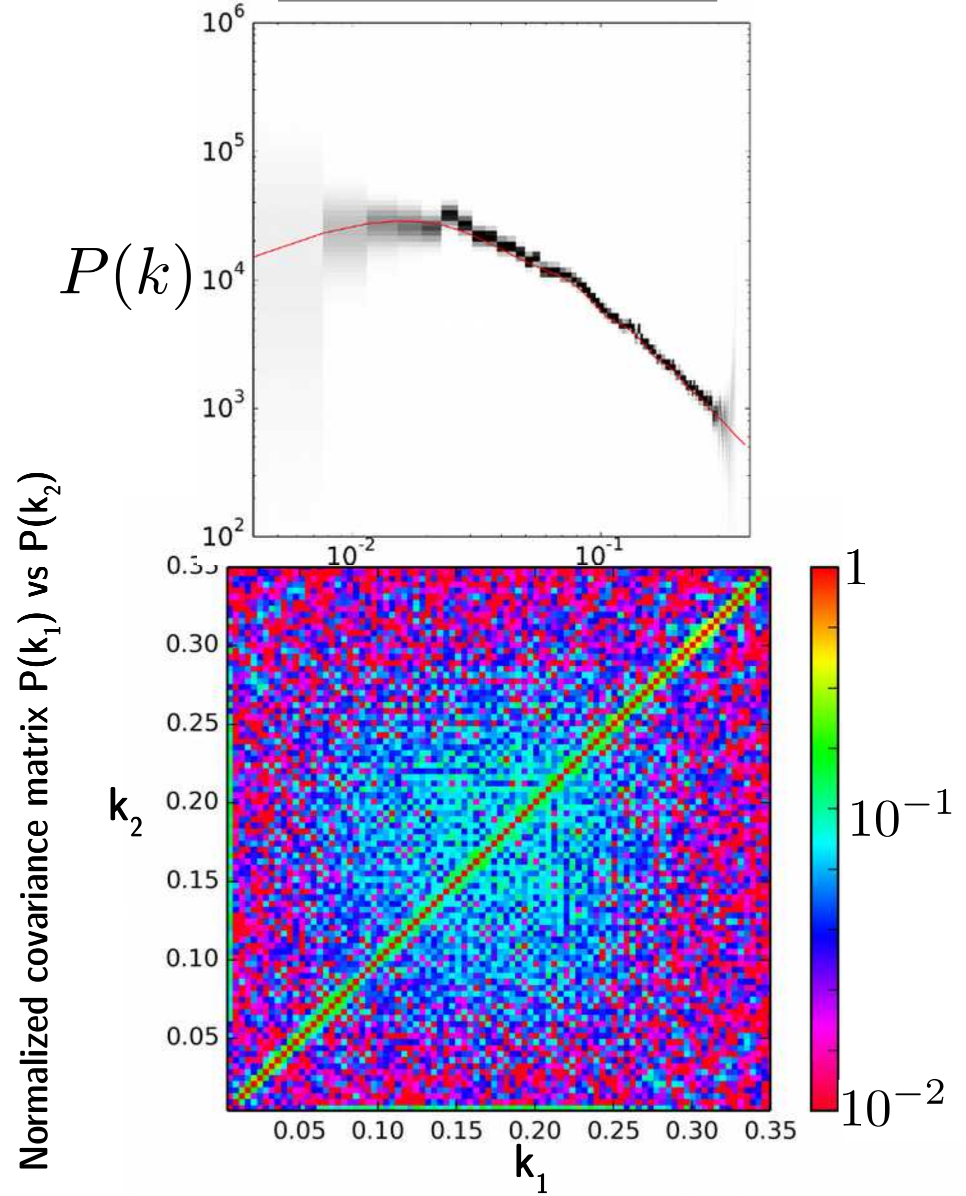
Jasche & Lavaux (submitted A&A, 2017)

Powerspectrum (un)corrected

Unclean LSS data



Cleaned LSS data



Jasche & Lavaux (submitted A&A, 2017)

Principles of statistical reconstruction of large scale structure (LSS)

Dealing with Foregrounds

Dealing with non-linearities of LSS

Bayesian Origins Reconstruction from Galaxies

BORG

BORG3: the non-linear model

Full poisson model:

$$P({\color{red}N_{obs}}|\rho_g) \propto \exp(-\rho_g)\rho_g^{{\color{red}N_{obs}}}$$

Broken power law model for bias:

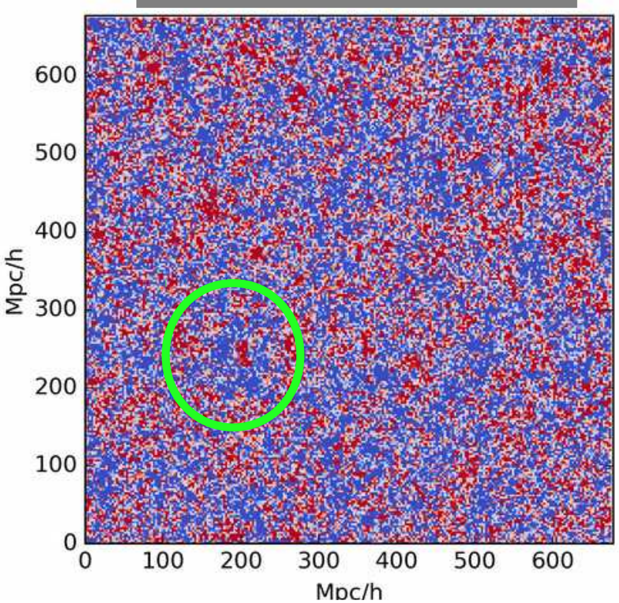
$${\color{purple}f}(\delta) = A(1+\delta)^{\alpha}\exp\left[-\left(\frac{1+\delta}{\rho_0}\right)^{-\epsilon}\right] - 1$$

Non-linear physical forward model:

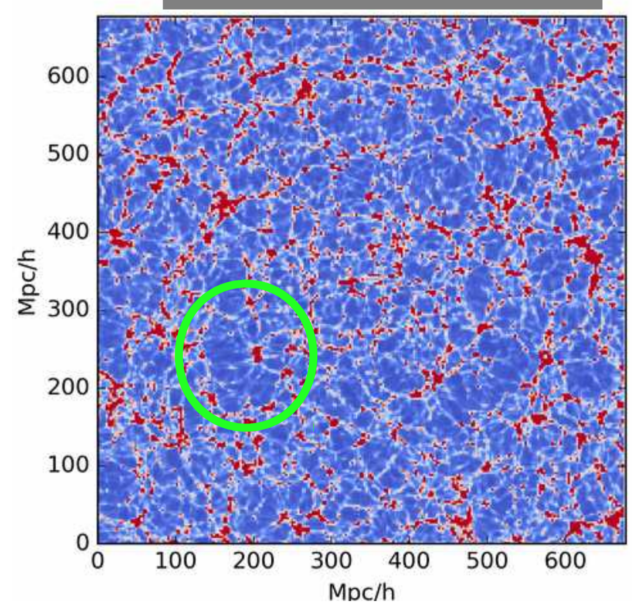
$$\mathcal{M} = \left\{ \begin{array}{l} \text{L. P. T.} \\ \text{2 L. P. T.} \\ \text{Particle mesh} \end{array} \right\} + \text{Redshift space distortion}$$

L.P.T. = Lagrangian Perturbation Theory

Eulerian linear theory



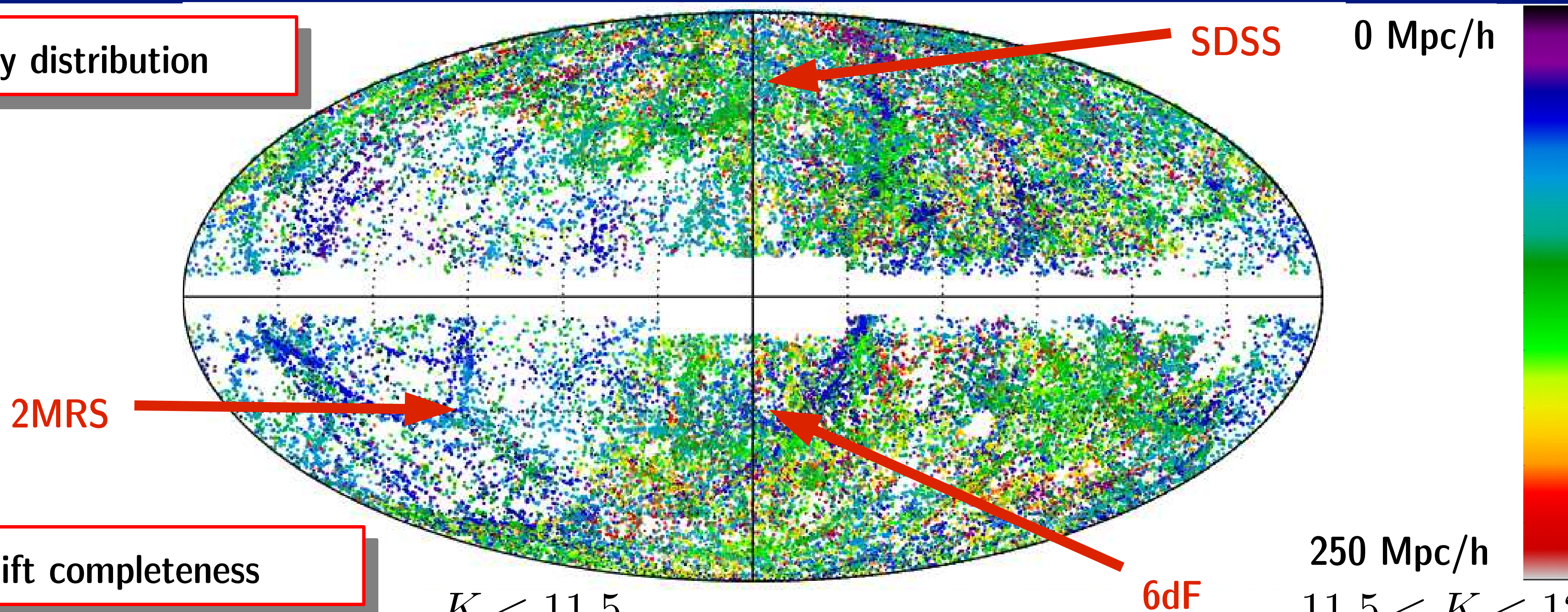
Particle mesh



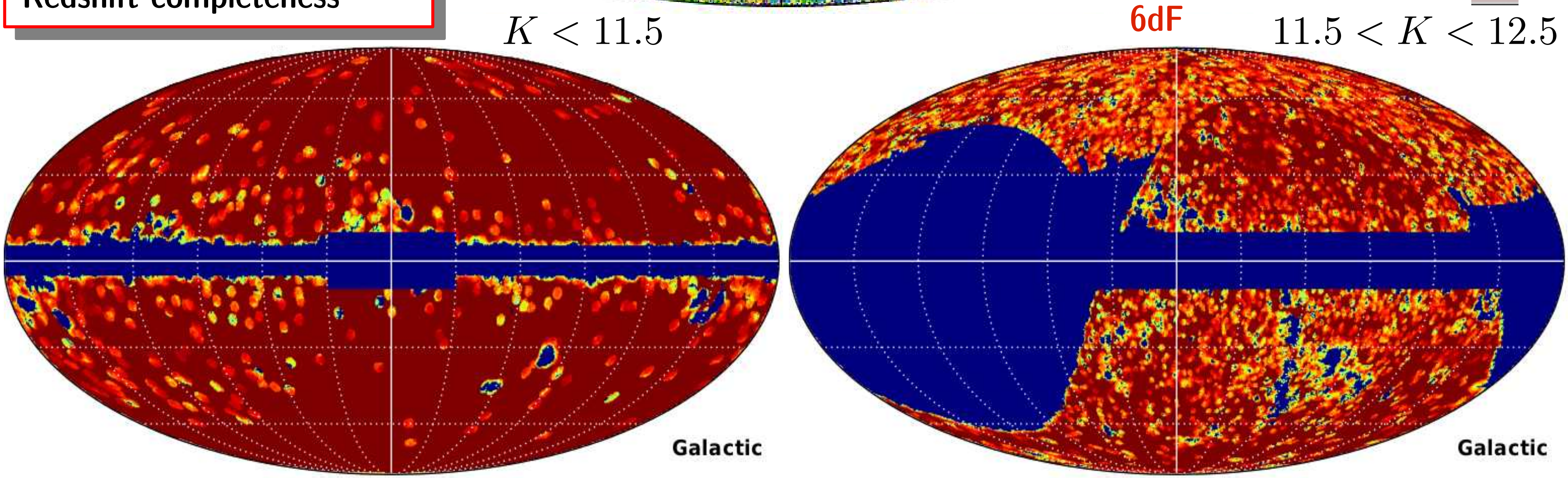
Jasche et al. (2010); Jasche & Lavaux (2015); Jasche & Lavaux (2017, in prep.)

Data: the 2M++ compilation

Galaxy distribution



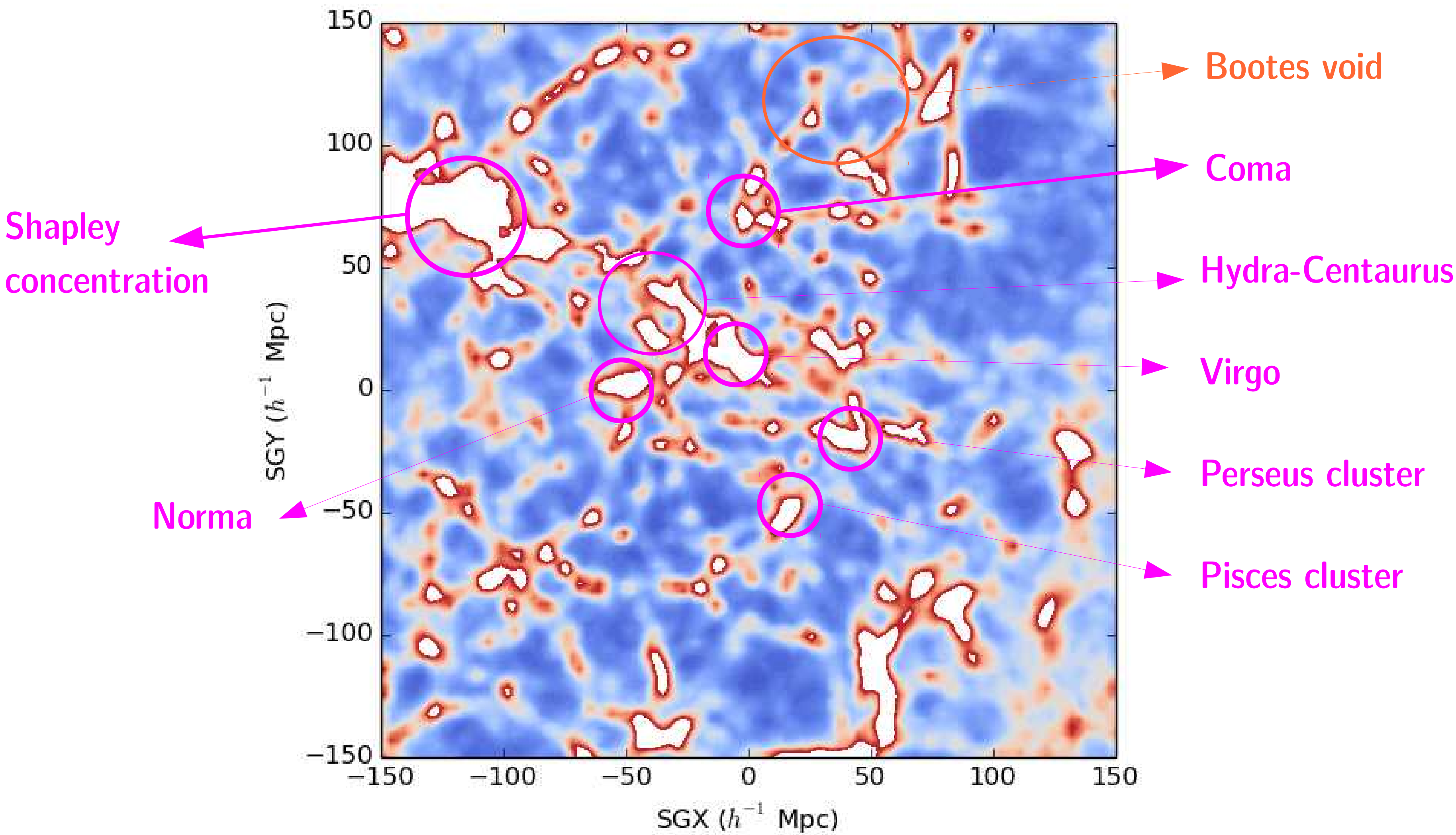
Redshift completeness



BORG3 density field

Supergalactic plane, final density field, no smoothing

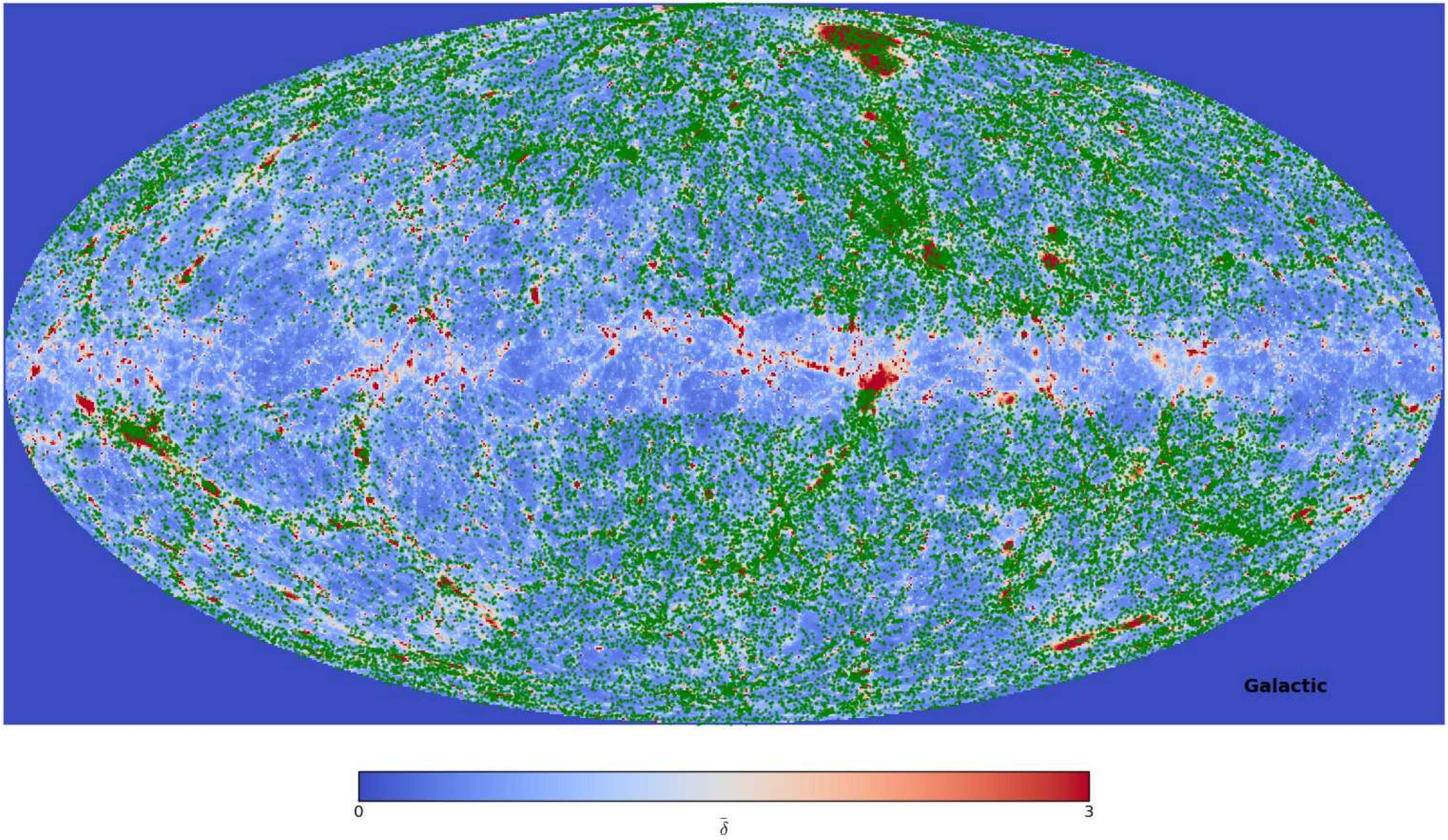
PRELIMINARY



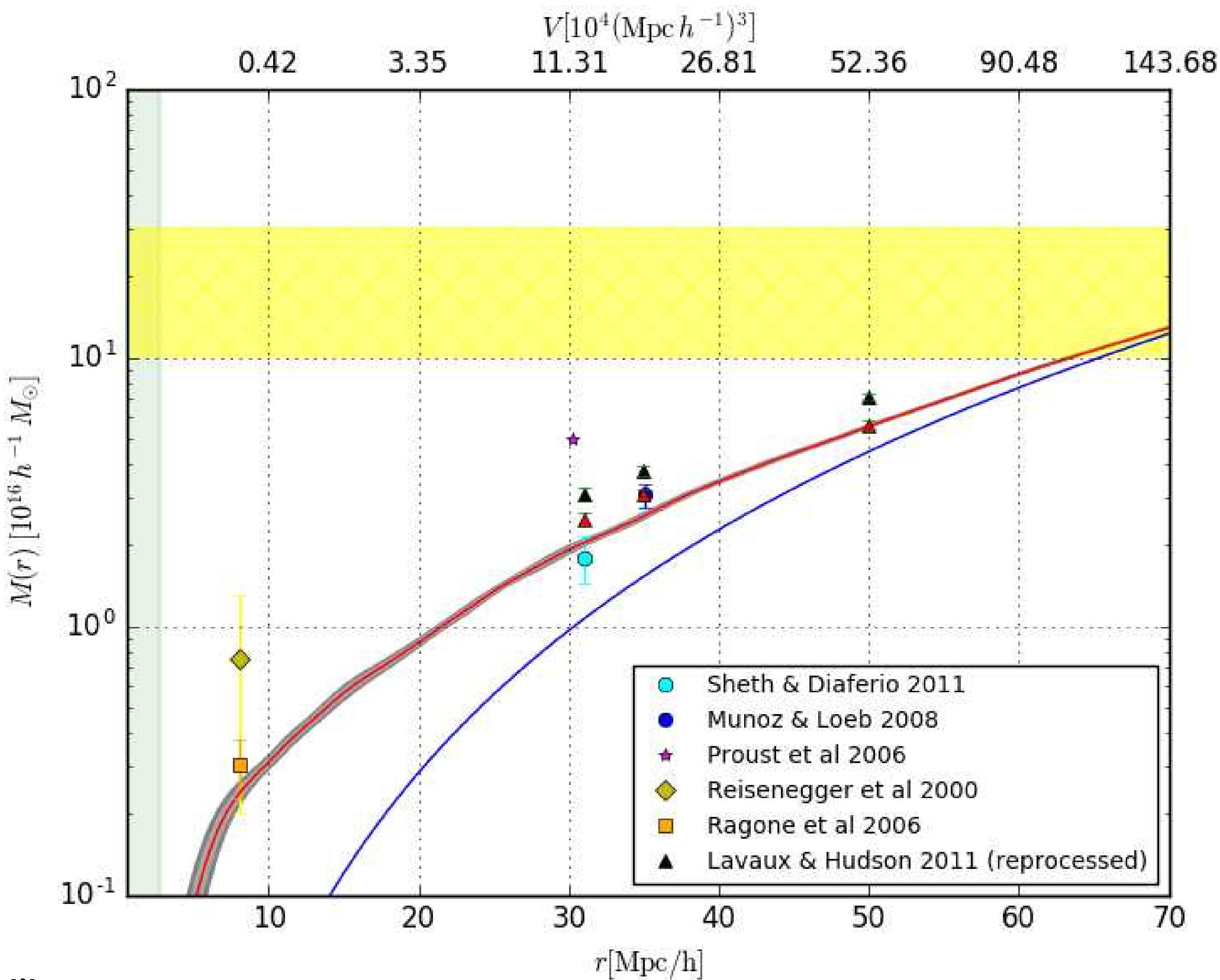
2M++, mean final matter density field

Supergalactic plane

BORG density vs Galaxy density



Shapley mass profile (PM)



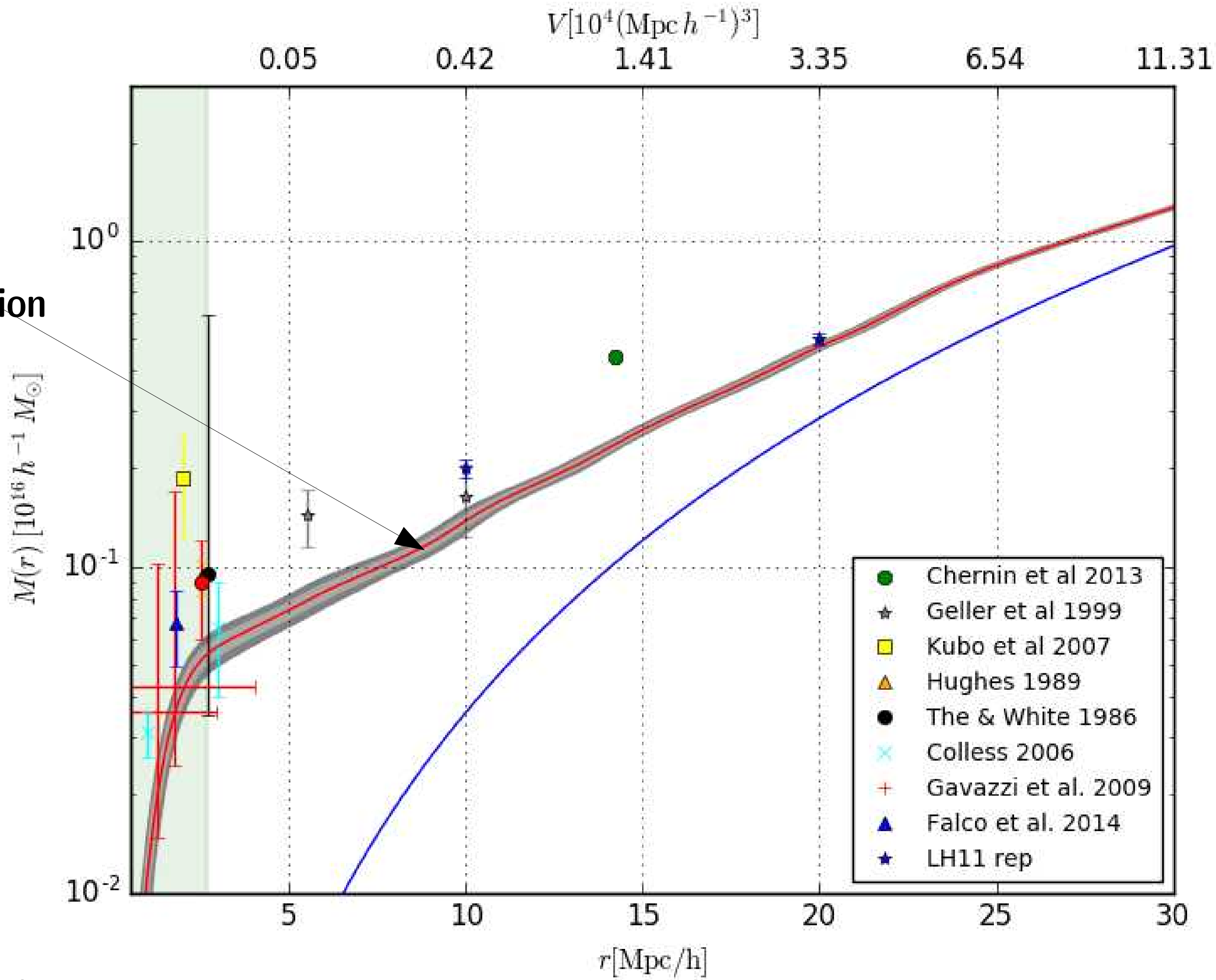
Mean

68% probability

95% probability

In detail: Coma mass profile (PM)

Image of the posterior distribution

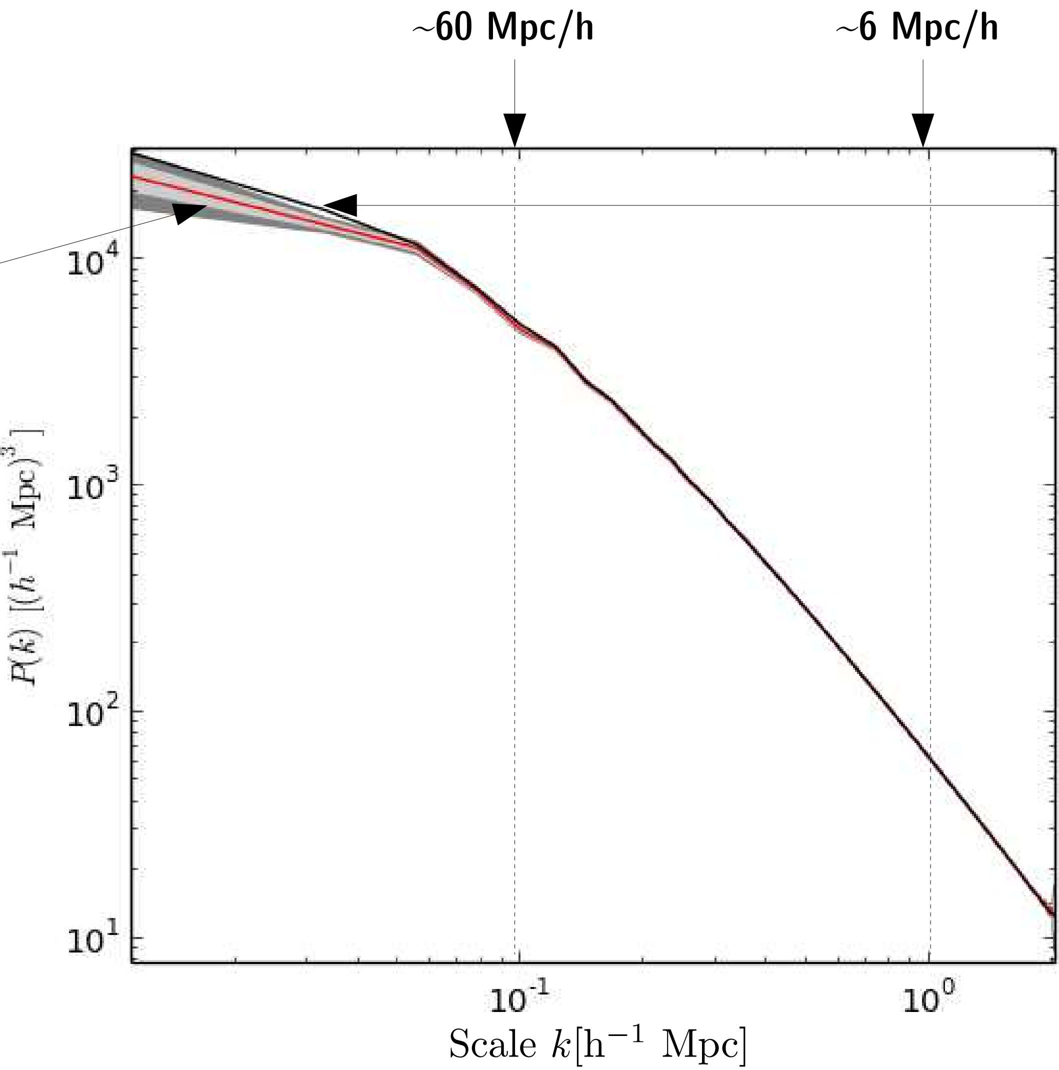


Mean

68% probability

95% probability

Power spectrum of initial conditions ?



Prior values of the power spectrum

Image of the posterior distribution

Mean

68% probability

95% probability

Jasche & Lavaux, 2017, in prep.

Conclusion / Perspective

Model works (> 16 million parameters)

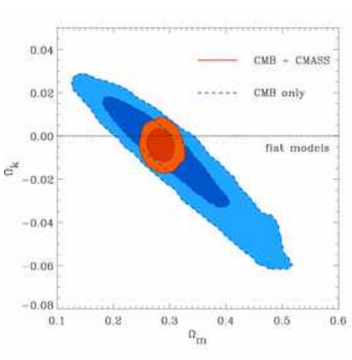
LCDM still rocks

Biases can be alleviated or at least identified

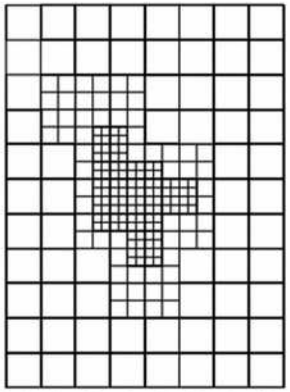
Distance survey and spectroscopic surveys are converging

Foreground contamination can be better assessed and corrected

Code scales for large surveys

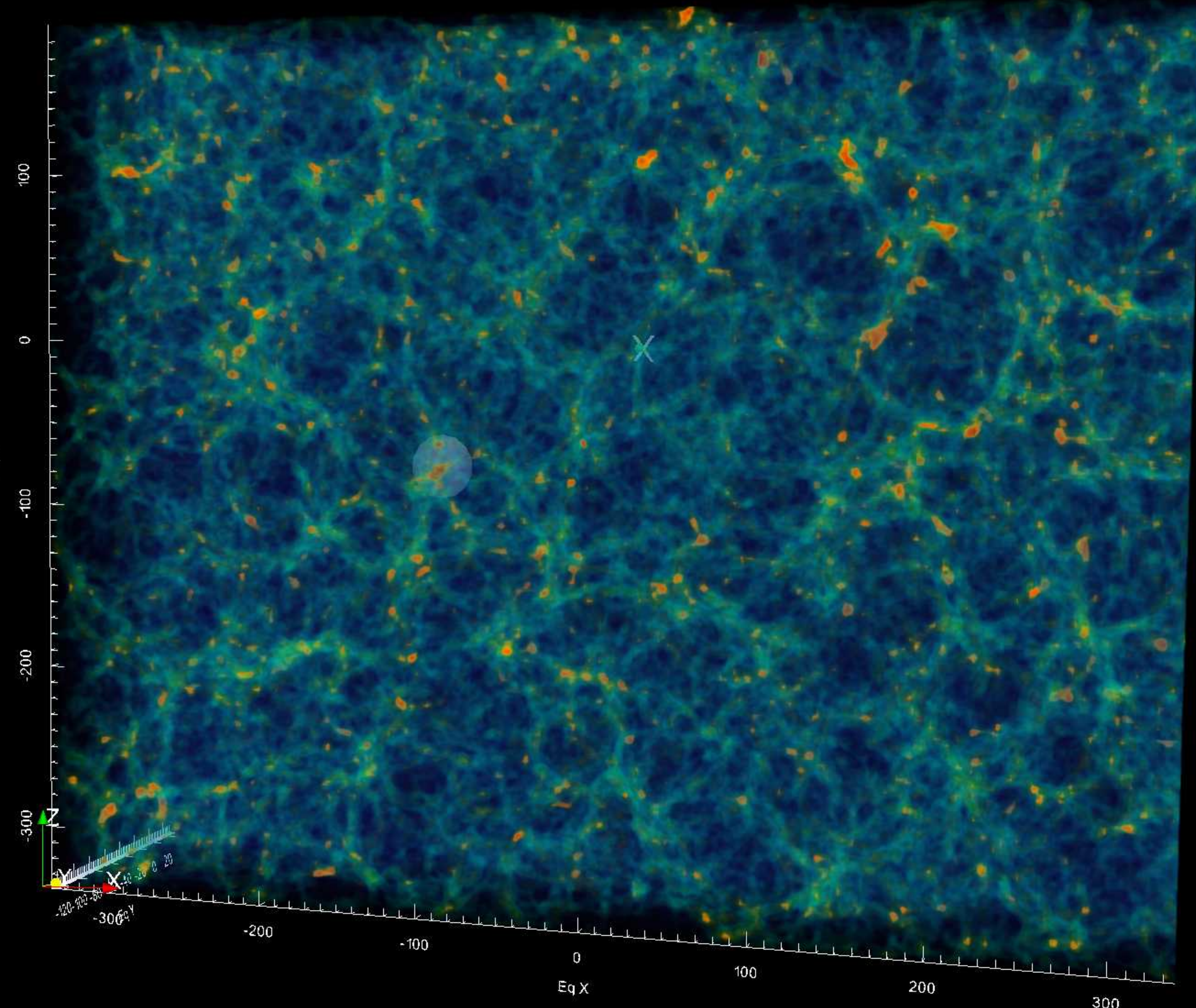


Liberate cosmology



Mesh refinement of initial conditions

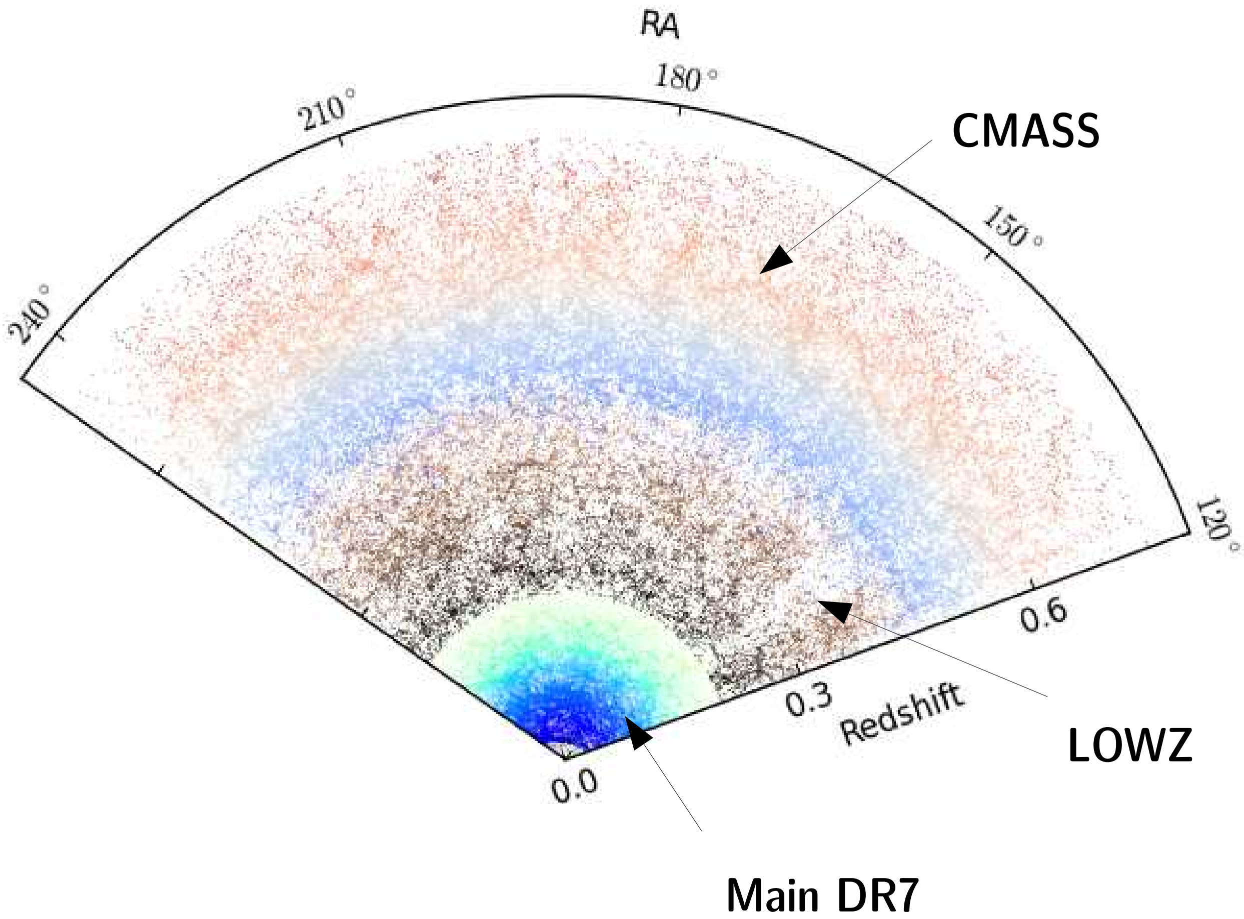
Improve sub-grid modeling of dynamics



Eq Z
100
0
-100
-200
-300
Eq X
-300
-200
-100
0
100
200
300
Z
X
Y

Additional material

Cosmology with large scale structures



Looks smooth

But tons of statistics and systematics are buried in the dataset

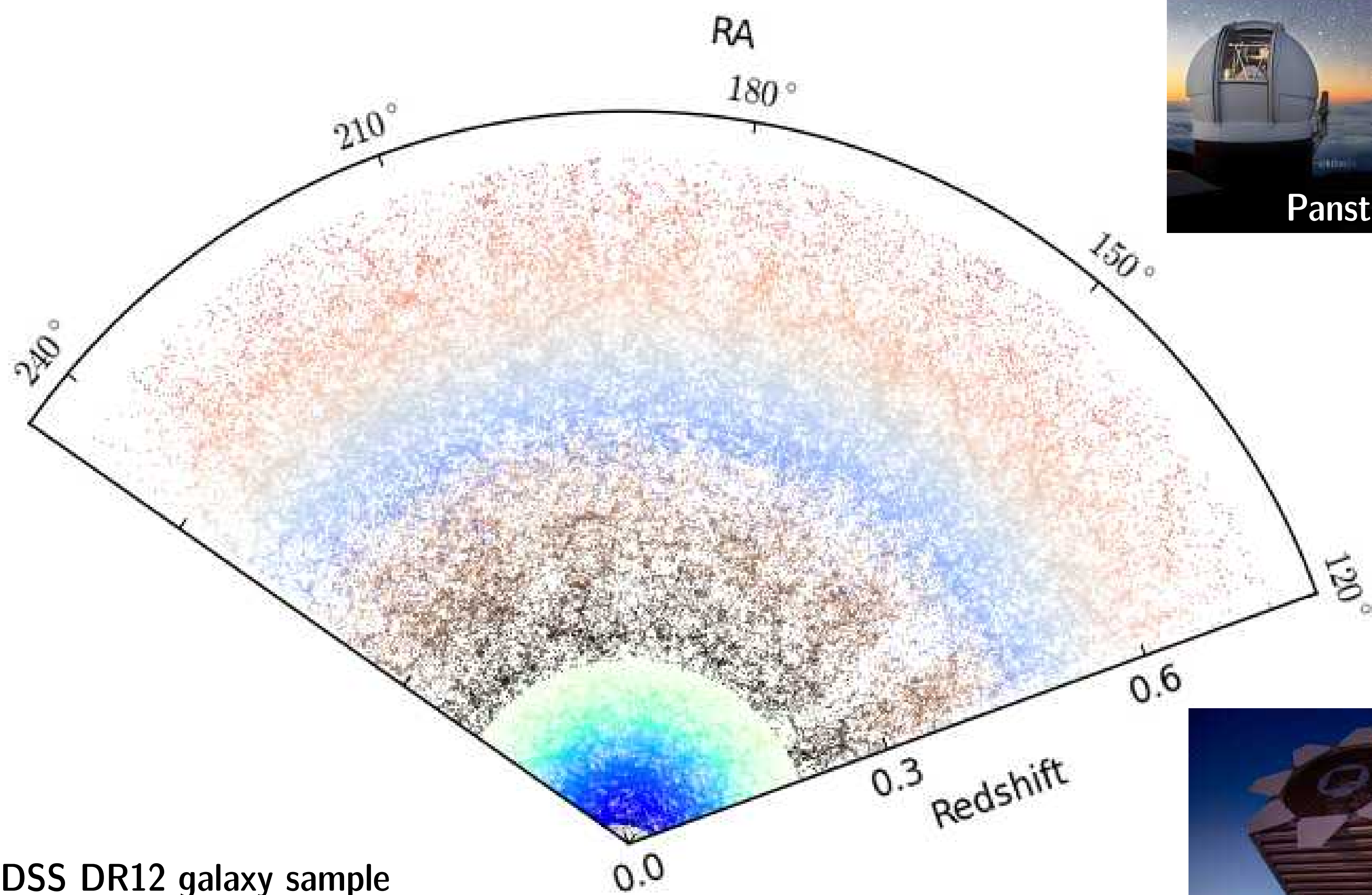
Cosmology hides both in the global geometry (Alcock-Paczynski) and the two point correlation of the underlying matter field

SDSS DR12 galaxy sample

~1.6 millions of galaxies

Cosmology with large scale structures

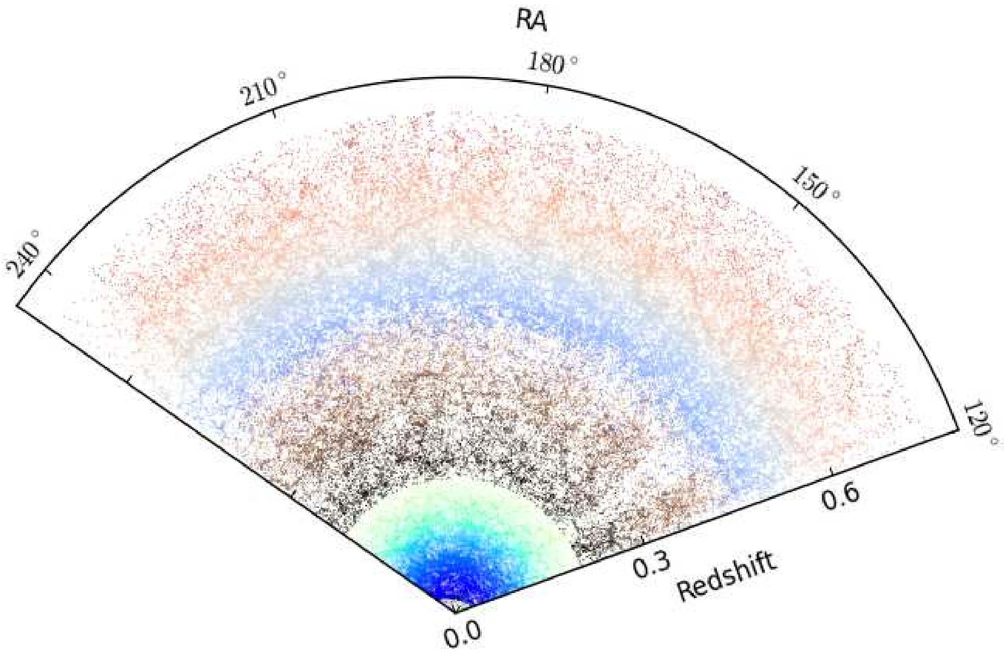
Aim: transform moderately processed photometric+spectroscopic data



SDSS DR12 galaxy sample

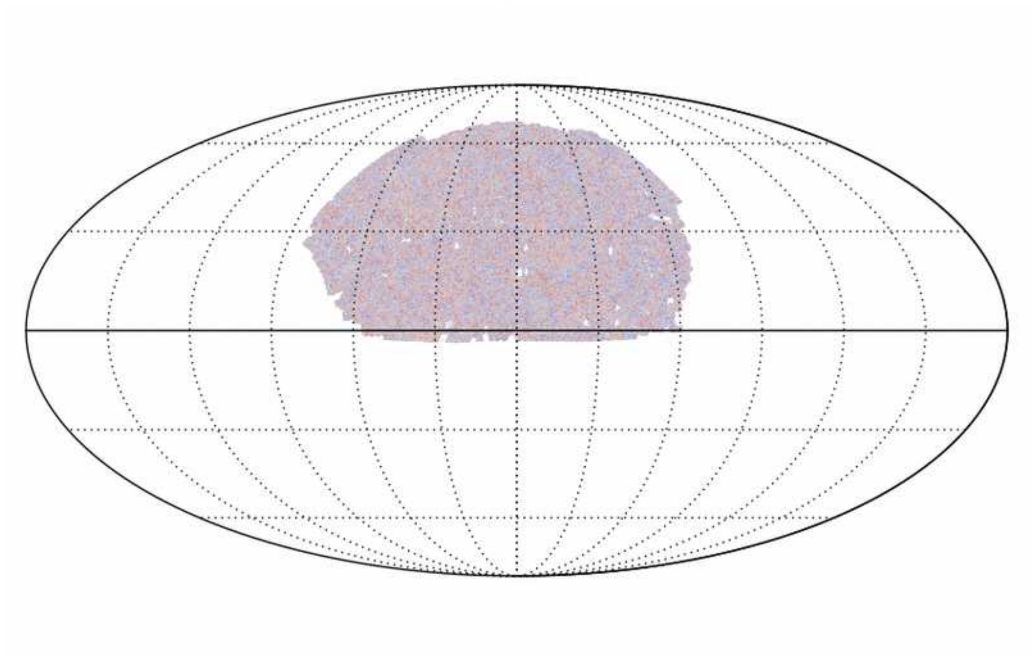
~1.6 millions of galaxies

Reminder on LSS data



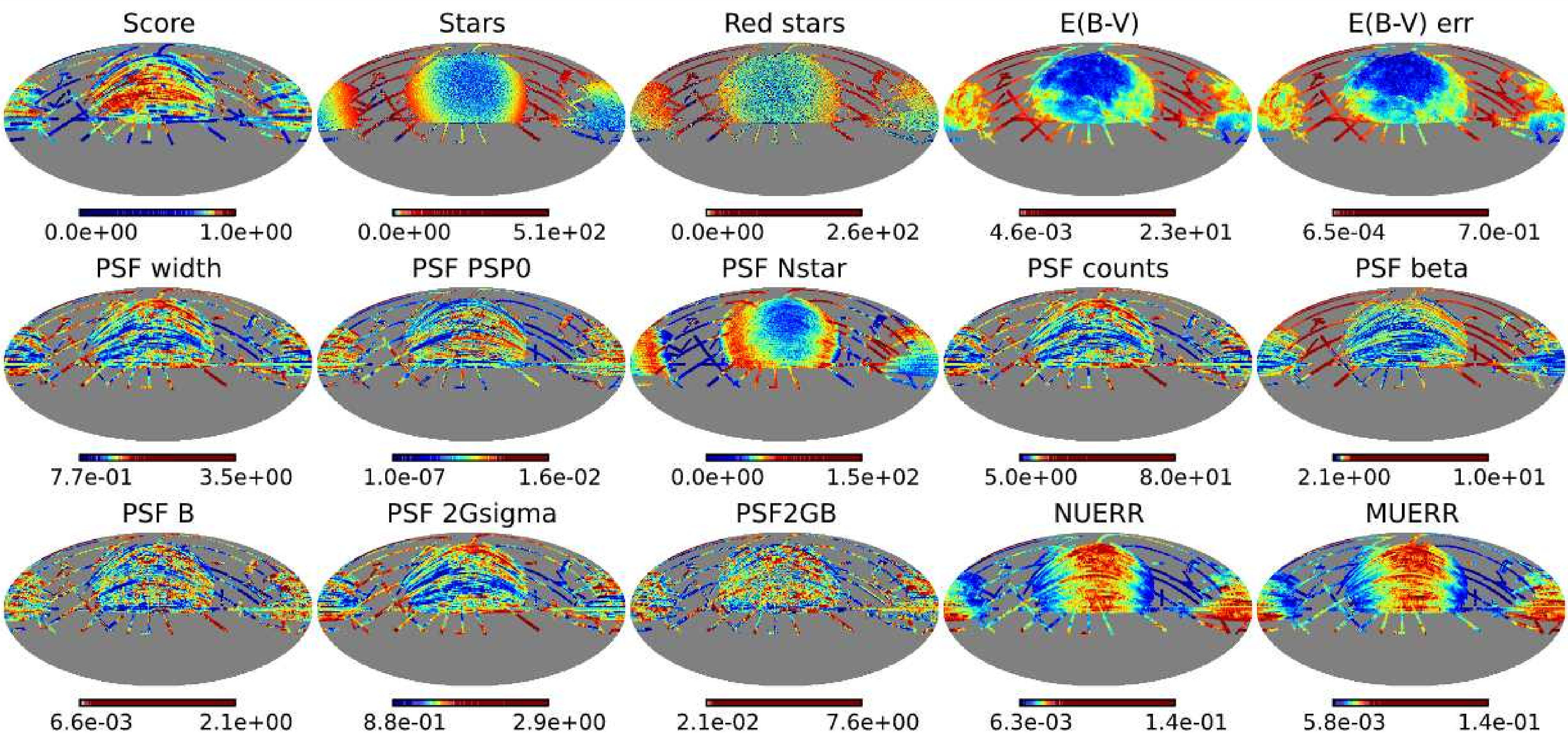
A wedge

vs.



Sky projection

LSS is full of foregrounds



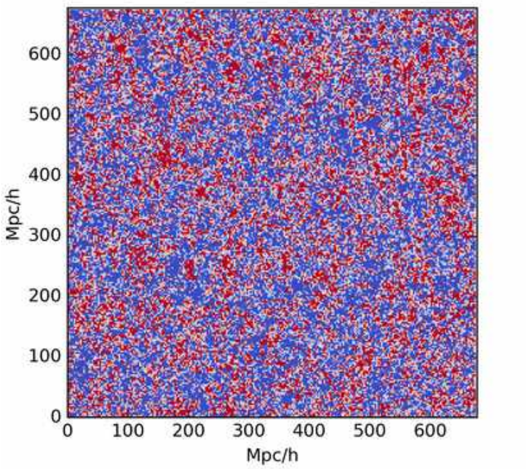
And 50 mores....

Leistedt & Peiris (2014)

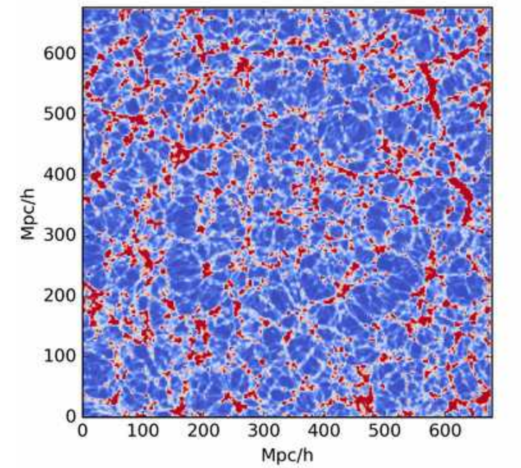
BORG: non-linear dynamics

$$\delta_{\mathrm{matter},p} = \mathcal{M}_p(\{\delta_{\mathrm{ic},p}\})$$

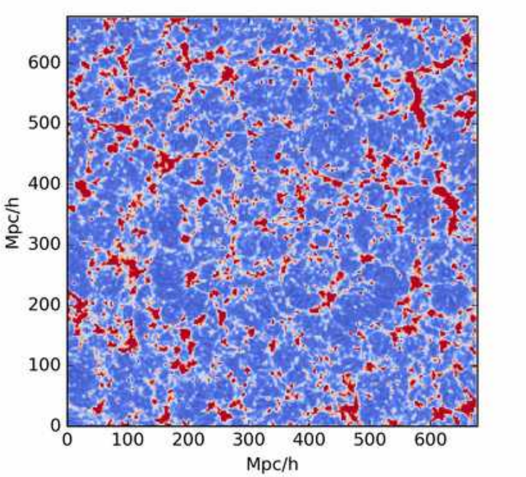
Eulerian linear perturbation theory: $\delta_{\mathrm{matter},p} = D_p \delta_{\mathrm{ic},p}$



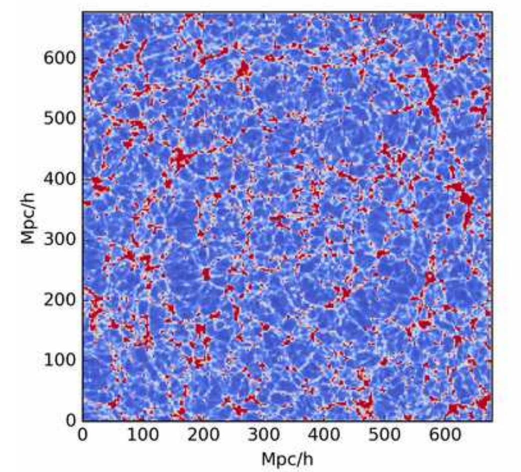
Lagrangian Perturbation Theory:
$\delta_{\mathrm{matter},p} = \Pi_p \circ S^{(1)} \circ \mathcal{D}(\{D^0_p \delta_{\mathrm{ic}_p}\})$



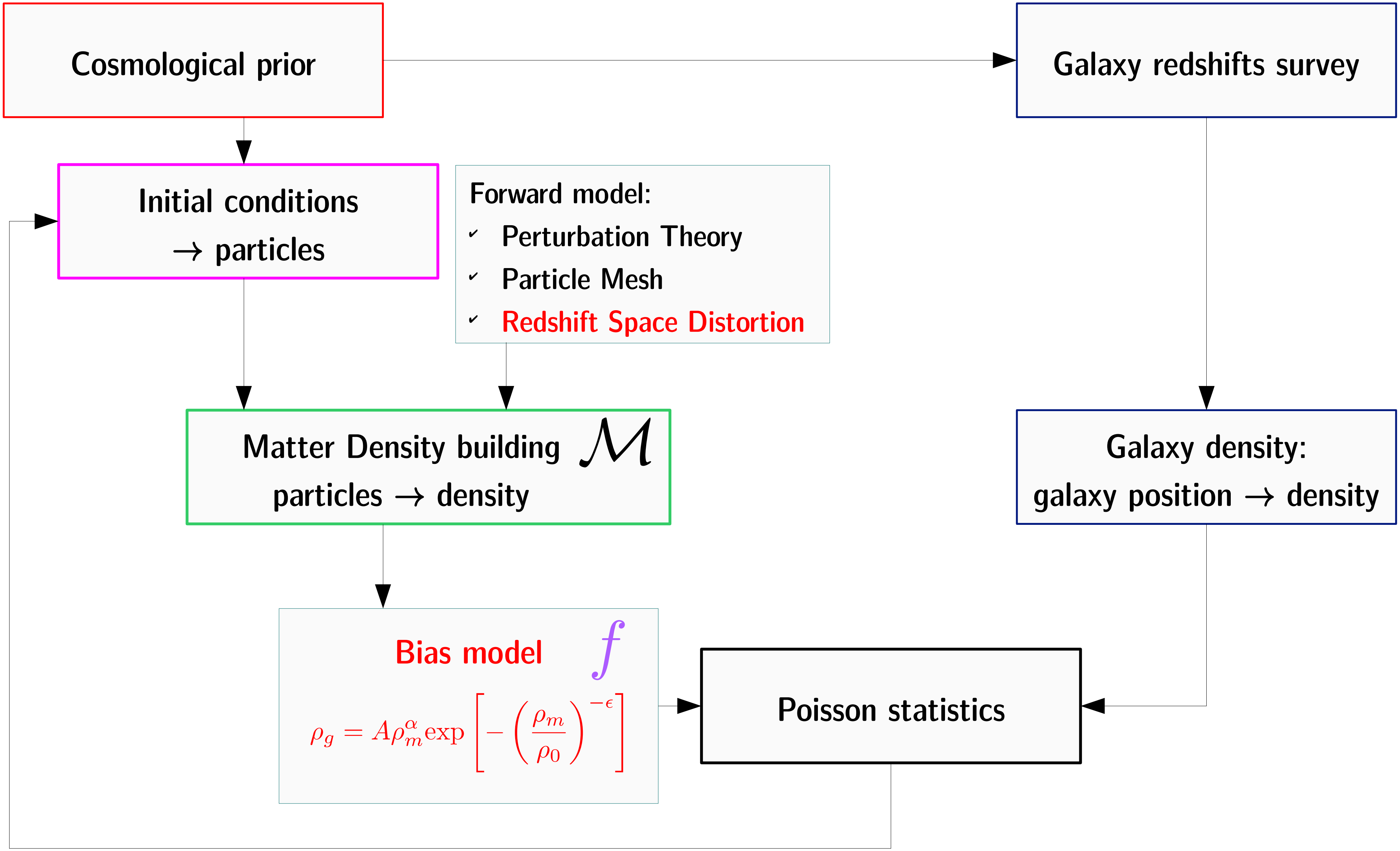
2nd order Lagrangian perturbation theory:
$\delta_{\mathrm{matter},p} = \Pi_p \circ S^{(2)} \circ \mathcal{D}(\{D^0_p \delta_{\mathrm{ic}_p}\})$



Particle mesh:
$\delta_{\mathrm{matter},p} = \Pi_p(\circ \mathcal{S})^n \circ S^{(1)} \circ \mathcal{D}(\{D^0_p \delta_{\mathrm{ic}_p}\})$



BORG3 model

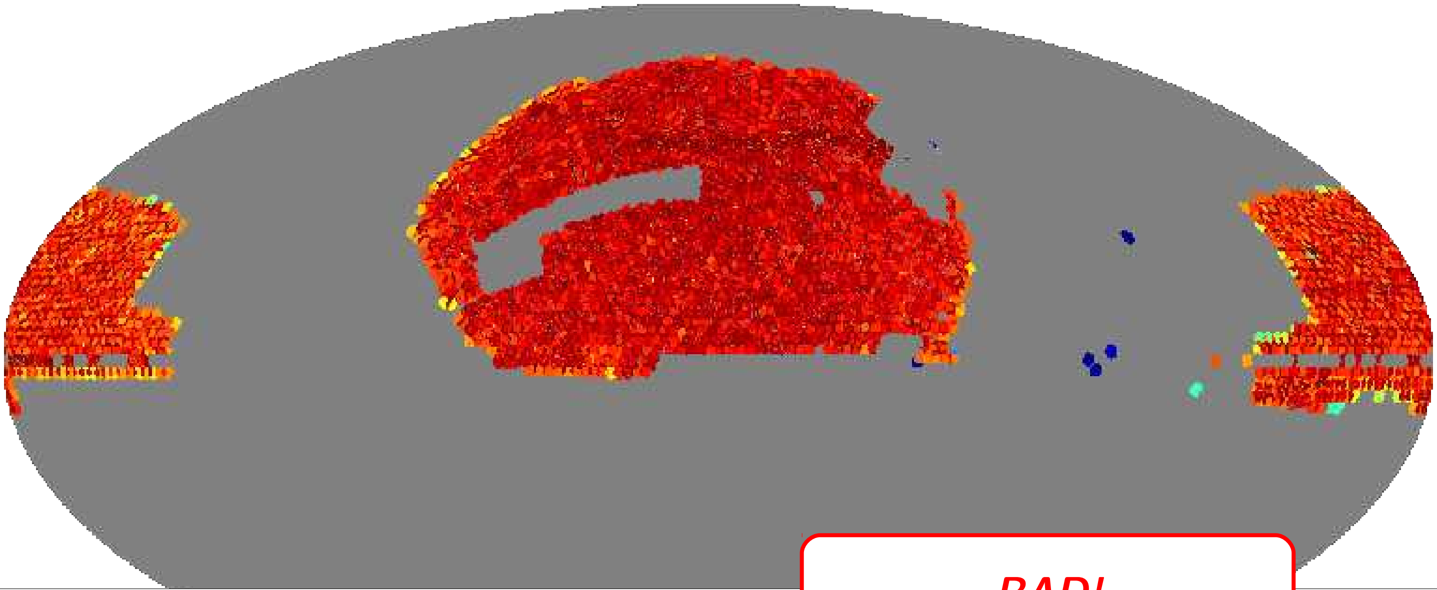


MPI + OpenMP parallel, exact supersampling, entire code rewriting

Lavaux & Jasche, 2017, in prep.

Jasche & Lavaux, 2017, in prep.

Completeness estimate



LOWZ survey ($z < 0.4$)

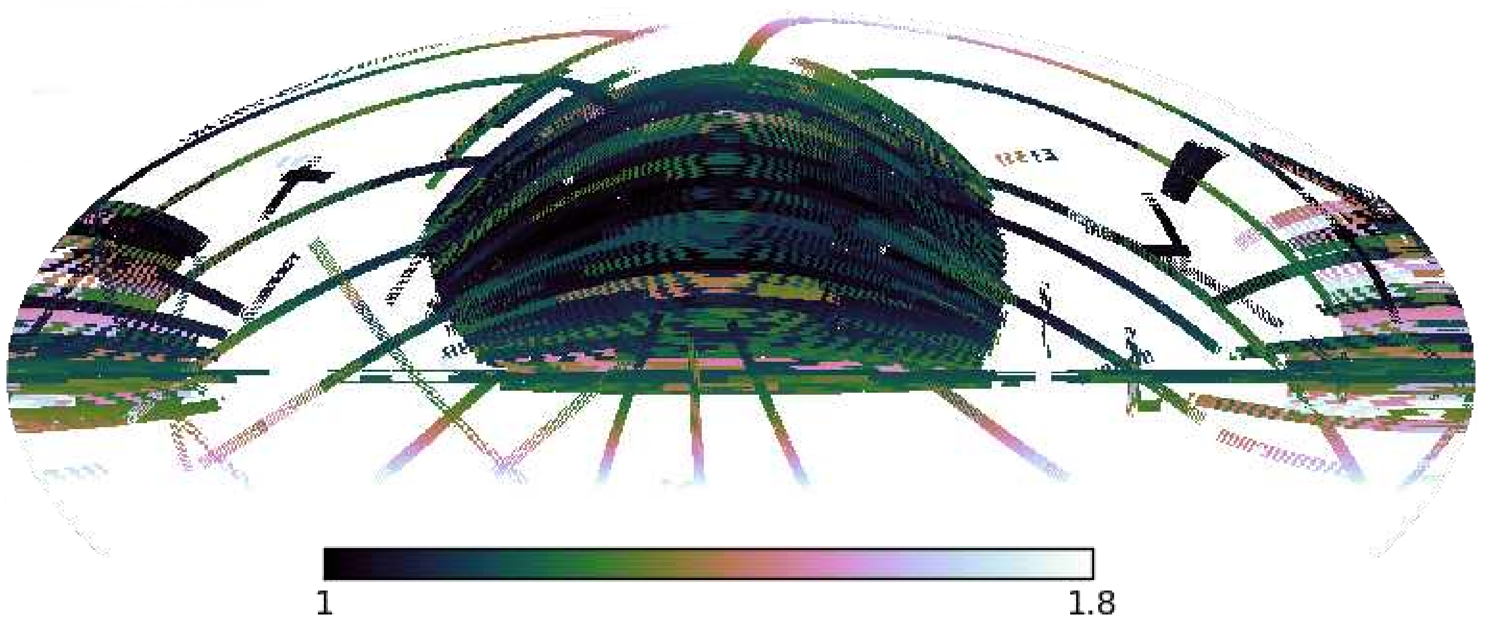
BAD!

$$R = \frac{\text{number of spectroscopically acquired galaxies}}{\text{number of wide band target galaxies}} = \frac{N_{spectro}}{N_{targets}}$$

+ spectroscopy contaminated by physical fiber collisions, spectral confusion

Lavaux & Jasche (to be submitted, 2017)

Some foregrounds for SDSS3

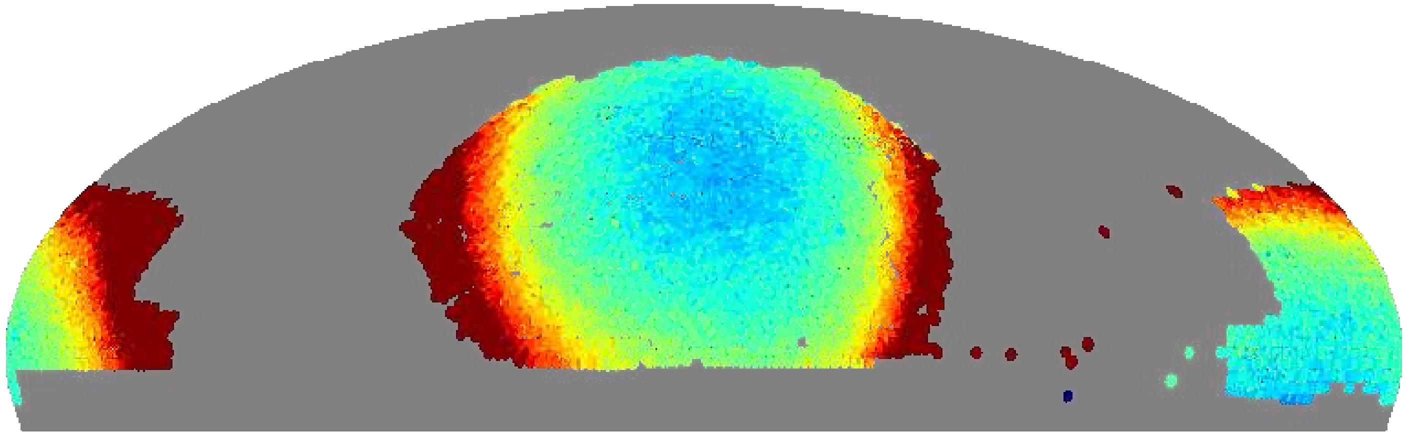


Airmass

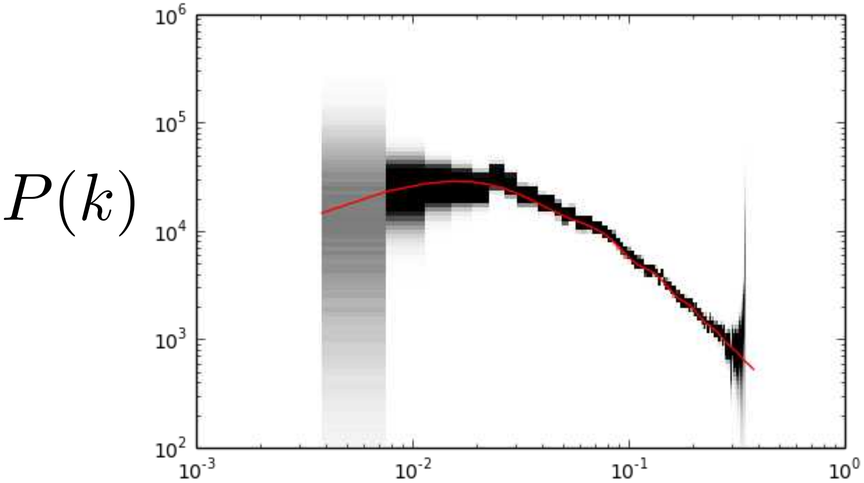
(absorption)

Star density

(contamination and absorption)



Foregrounds / Powerspectrum correlation



$$R_p = \underbrace{\left[\prod_{i=1}^{N_{\mathrm{fg}}} (1-\alpha_i G_{i,p})\right]}_{1/F_p} M_p$$