Fully non-linear statistical analysis of Large scale structure data for wide and deep surveys

Institut d'Astrophysique de Paris (CNRS)

Guilhem Lavaux

Collaboration with

Jens Jasche (TUM/ExC)

Doogesh Kodi Ramanah (IAP/ILP) → poster S14.10

Tom Charnock (just joined @ IAP)
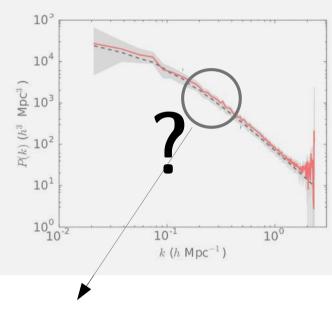
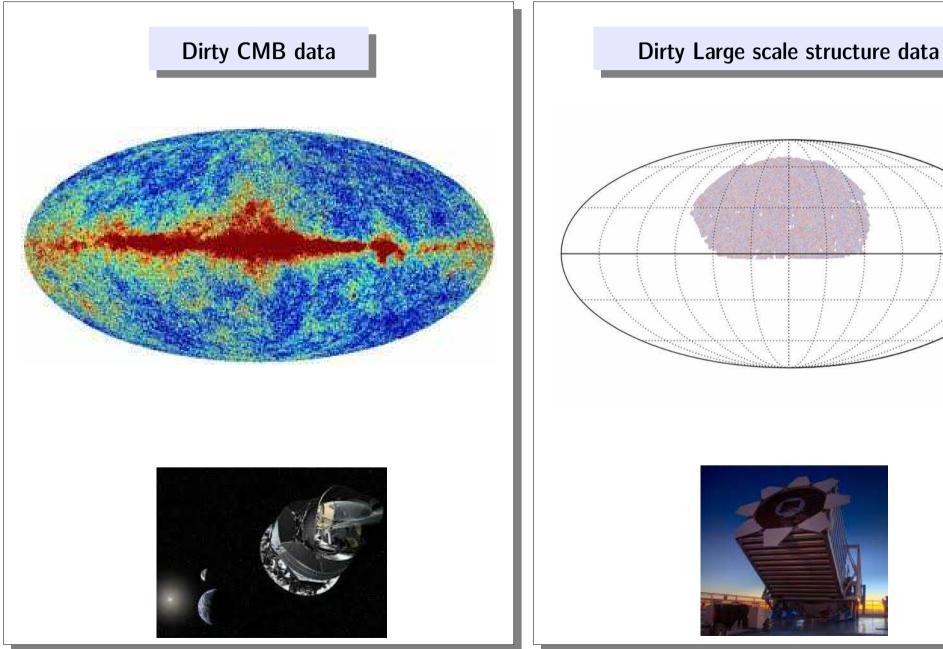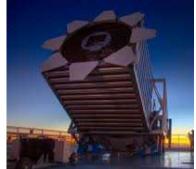# Cosmology with large scale structures
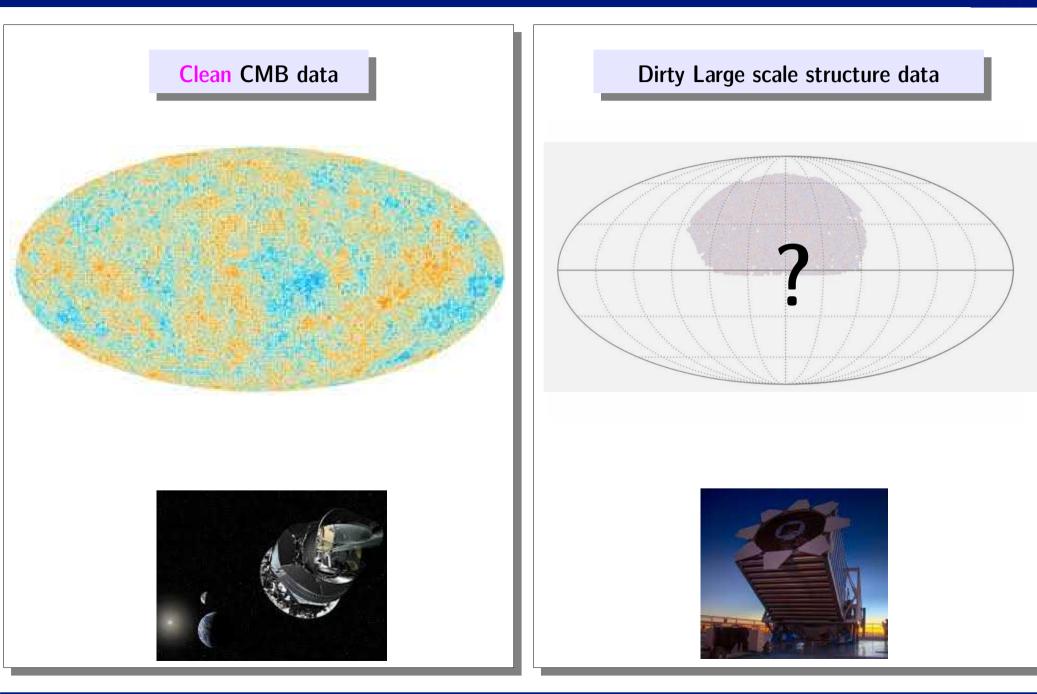

Dirty CMB data
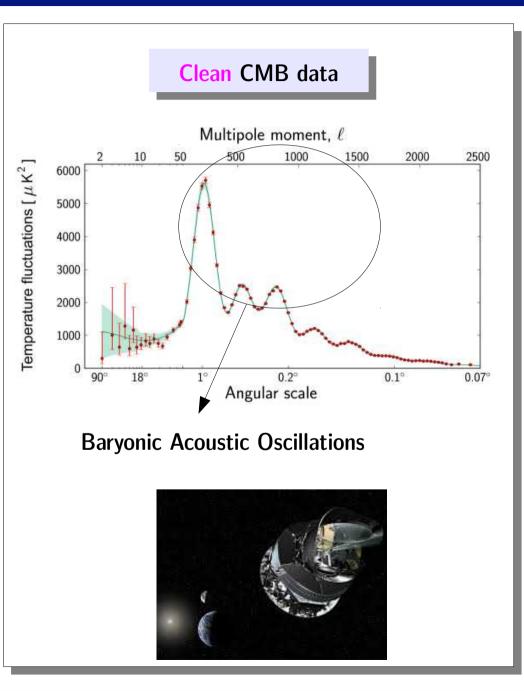
Dirty Large scale structure data

Clean CMB data

Dirty Large scale structure data

?

Cosmology with large scale structures

Baryonic Acoustic Oscillations

Baryonic Acoustic Oscillations

Principles of statistical reconstruction of large scale structure (LSS)

Algorithm for **RE**construction and **S**ampling

Bayesian **O**rigins **R**econstruction from **G**alaxies

ARES

Jasche et al. (2010), Jasche & Wandelt (2013)

Jasche & Lavaux (2014)
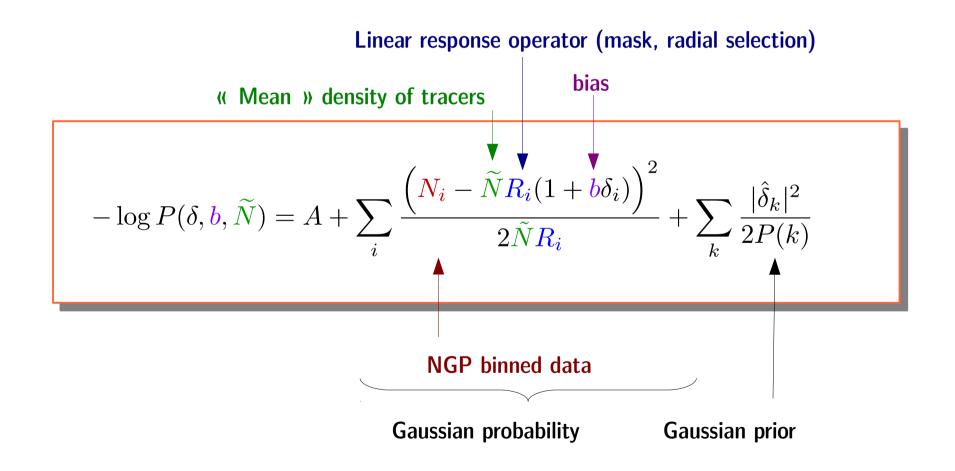
Jasche & Lavaux (2017, in press)

BORG

Lavaux & Jasche (2016)

Jasche & Lavaux (2017 in prep)

Lavaux & Jasche (2017 in prep)

Statistical reconstruction of LSS

Galaxies are counted in 3d cells

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

Poisson probability on large scales
Noise is signal dependent!

Mean density

$$\lambda_p = R_p \bar{N}(1 + \delta_{\text{galaxies},p})$$

Linear response

Selection

$$1 + \delta_{\text{galaxies},p} = f(1 + \delta_{\text{matter},p})$$

Local differentiable dependency on matter field

e.g. bias function

$$\delta_{\text{matter},p} = \mathcal{M}_p(\{\delta_{\text{ic},p}\})$$
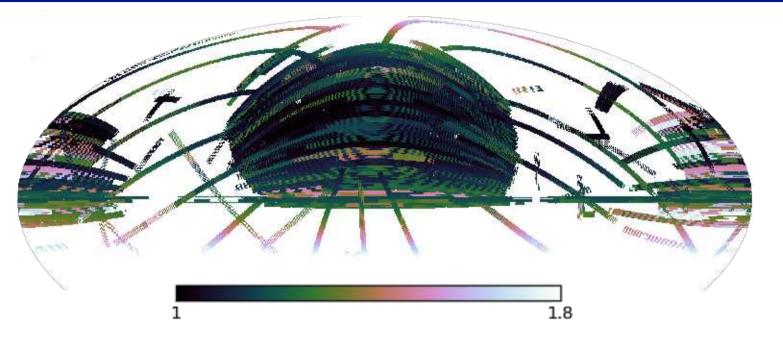
forward model for dynamics

ARES3: posterior for the linear model

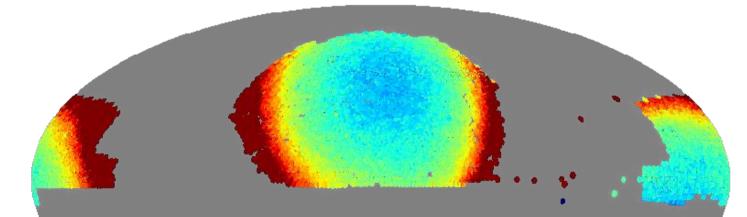$$P(N_{obs}|\lambda) \propto \exp(-(N_{obs} - \lambda)^2/(2\lambda)) \quad f(\delta) = b\delta \qquad \mathcal{M}_p(\{\delta\}) = \delta_p$$

Linear response operator (mask, radial selection)

bias

« Mean » density of tracers

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

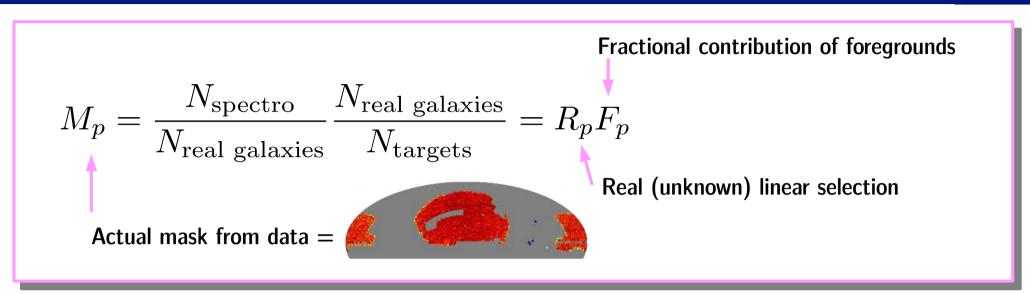NGP binned data

Gaussian probability   Gaussian prior

Airmass
(absorption)

1   1.8
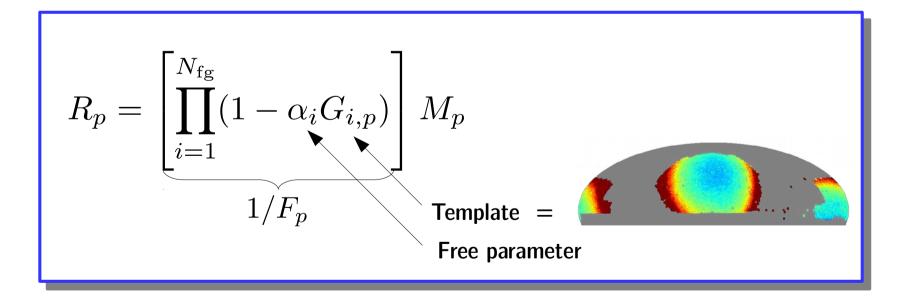
Star density

(contamination and
absorption)
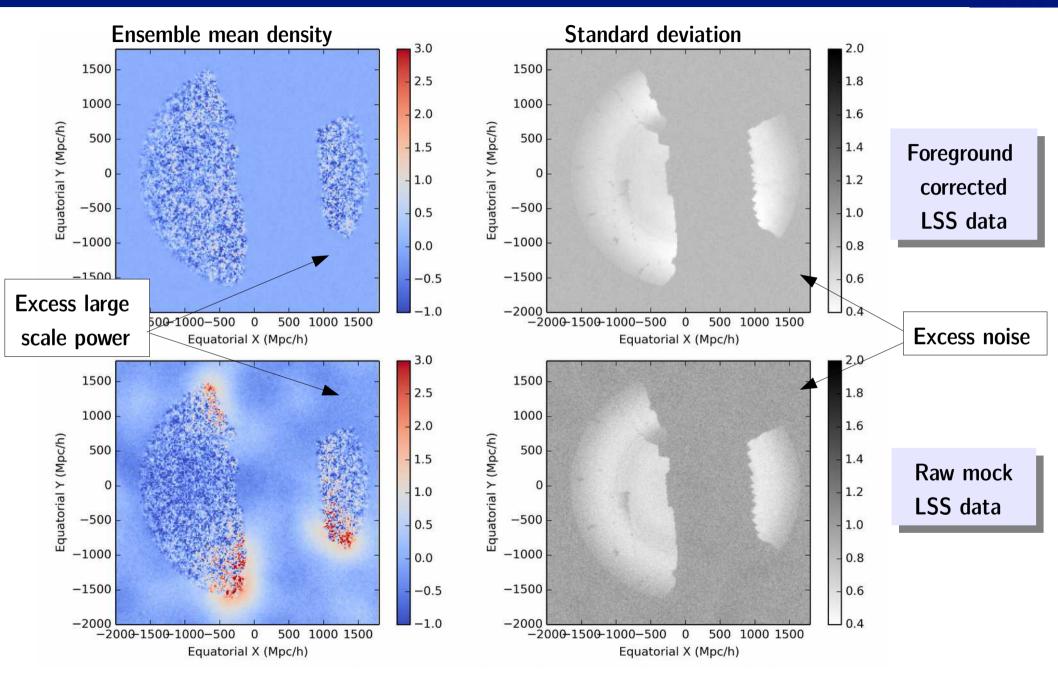
Foregrounds: 1ˢᵗ order correction
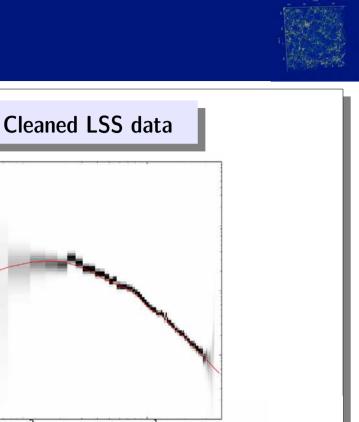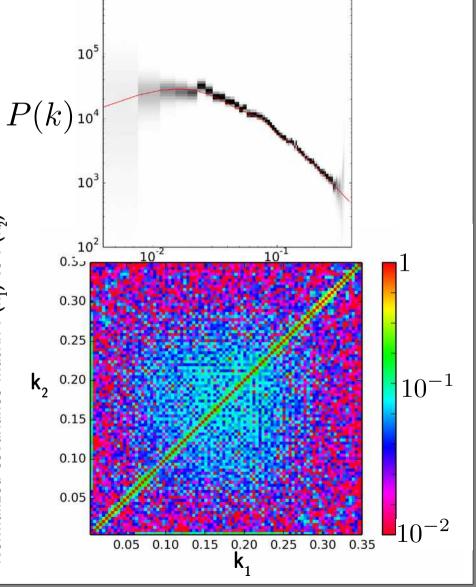
Fractional contribution of foregrounds

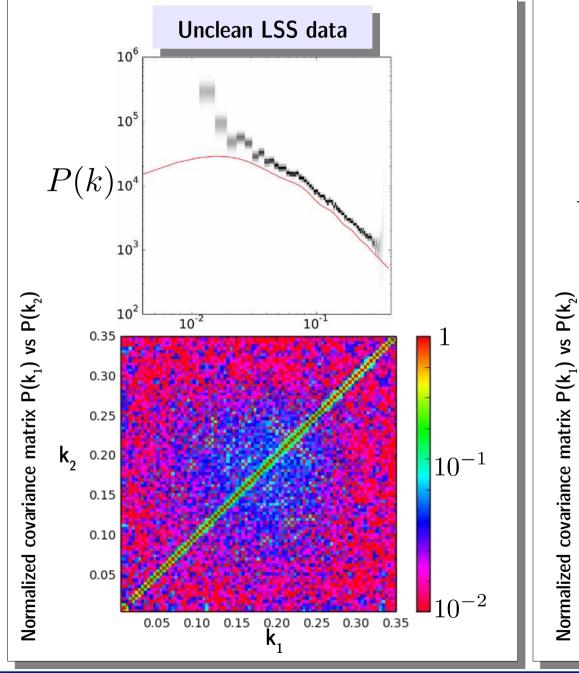$$M_p = \frac{N_{\text{spectro}}}{N_{\text{real galaxies}}} \frac{N_{\text{real galaxies}}}{N_{\text{targets}}} = R_p F_p$$

Real (unknown) linear selection

Actual mask from data =

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

$1/F_p$

Template =

Free parameter

Example on a mock SDSS3

Ensemble mean density

Standard deviation

Excess large scale power

Excess noise

Foreground corrected LSS data

Raw mock LSS data

Unclean LSS data

Cleaned LSS data
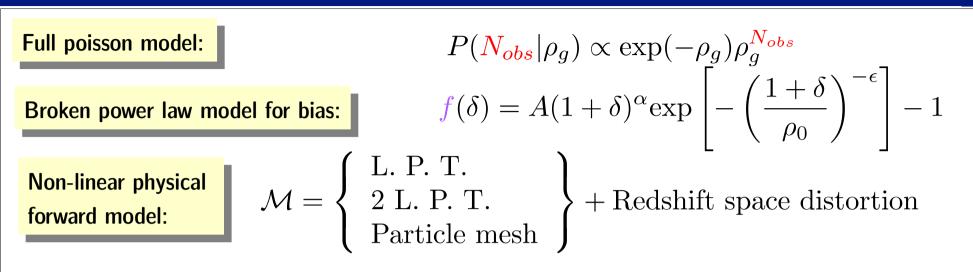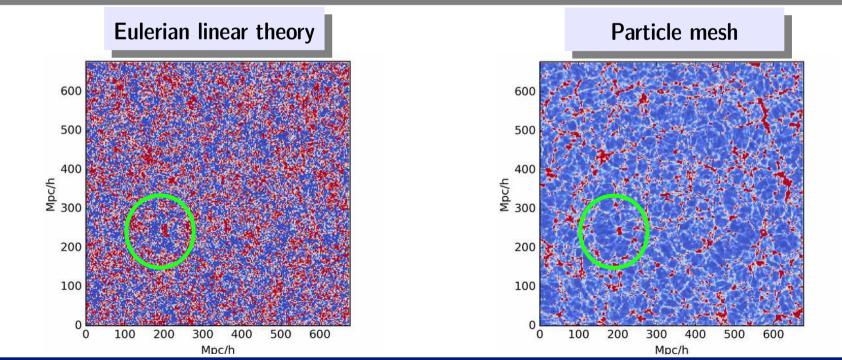
BORG3: the non-linear model

Full poisson model:

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

Broken power law model for bias:

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

Non-linear physical forward model:

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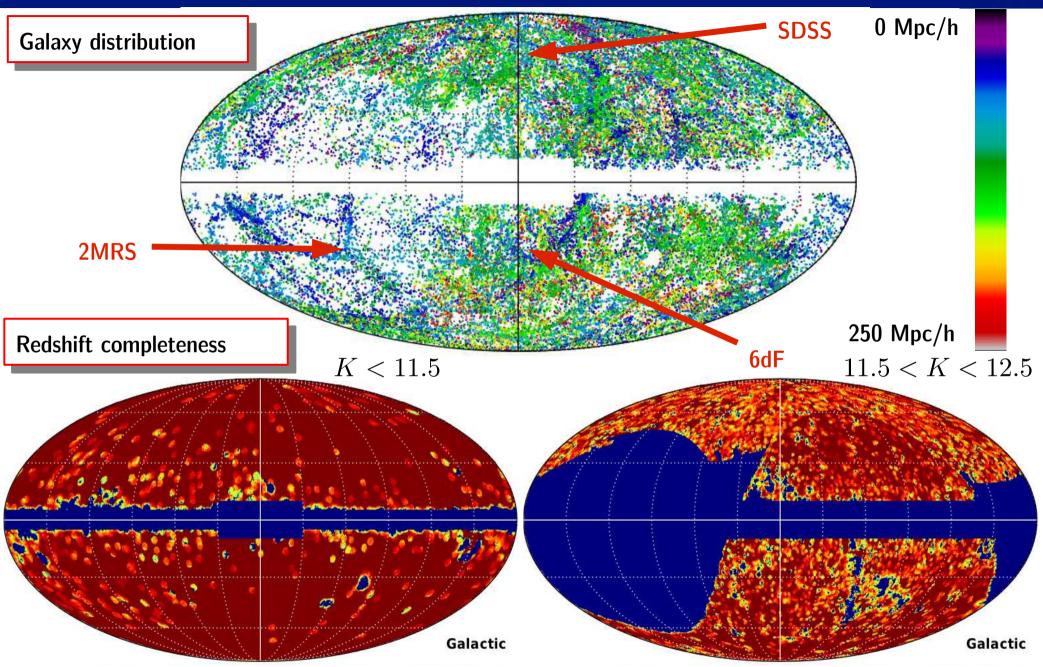
L.P.T. = Lagrangian Perturbation Theory
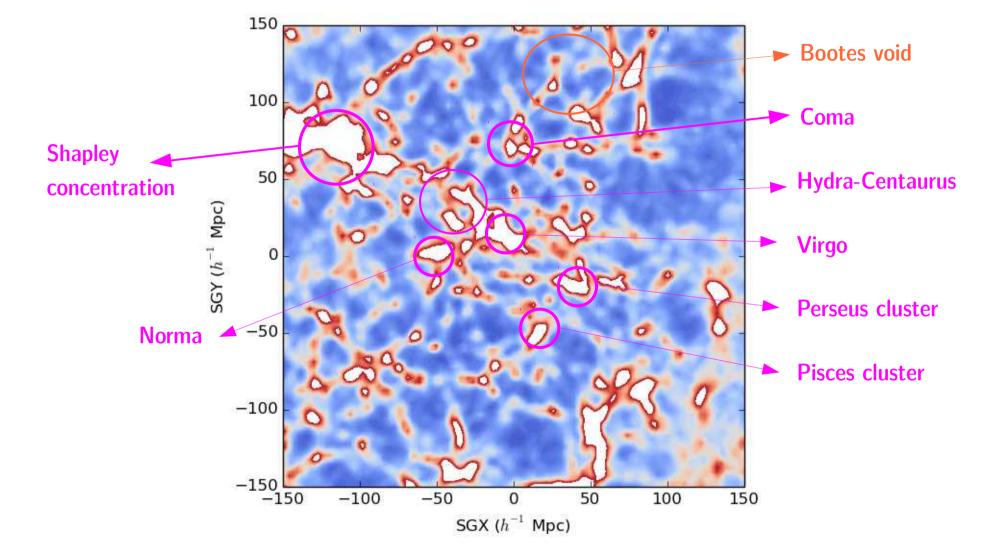
Eulerian linear theory

Particle mesh

Data: the 2M++ compilation

Galaxy distribution

SDSS

2MRS

6dF

Redshift completeness

$K < 11.5$

$11.5 < K < 12.5$

0 Mpc/h

250 Mpc/h

Galactic

Galactic

BORG3 density field

Bootes void

Coma

Hydra-Centaurus

Virgo

Perseus cluster

Pisces cluster

Shapley concentration

Norma

2M++, mean final matter density field

Supergalactic plane
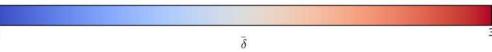
BORG density vs Galaxy density

Galactic

$\bar{\delta}$

0   3
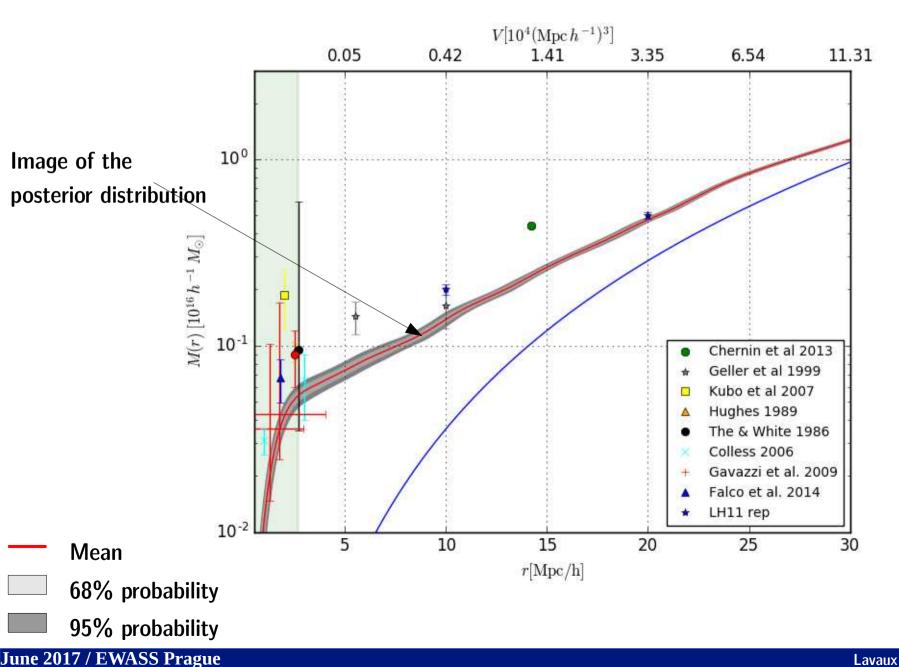
Conclusion / Perspective

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