Bayesian reconstruction of the Cosmological Large Scale Structure

Francisco-Shu Kitaura

Leibniz Institute for Astrophysics (AIP)

May 31, 2011

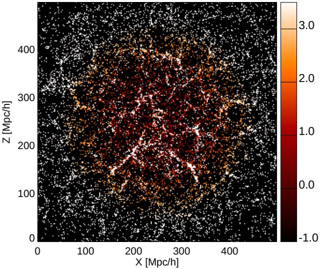
Max Planck Institute for Astrophysics

SISSA

SCUOLA NORMALE SUPERIORE

Introduction

- the cosmological large-scale structure encodes a wealth of information about the evolution and origin of the Universe
- the data are plagged by many observational effects (mask, selection function, bias ...)
- statistical treatment is necessary
- compare observations with theory
- study structure formation



⟶ matter field $\delta_{\rm M}$?

⟶ pec. vel. field $\mathbf{v}$?

⟶ grav. pot. $\Phi_{\rm Grav}$?

Bayes theorem: the posterior

$$\mathcal{P}(\mathbf{s}|\mathbf{d},\mathbf{p}) = \frac{\mathcal{P}(\mathbf{s}|\mathbf{p})\mathcal{P}(\mathbf{d}|\mathbf{s},\mathbf{p})}{\int \mathrm{d}\mathbf{s}\,\mathcal{P}(\mathbf{s}|\mathbf{p})\mathcal{P}(\mathbf{d}|\mathbf{s},\mathbf{p})}, \quad (1)$$

Posterior=prior×likelihood/evidence

⟶ clear representation of the assumptions
⟶ can be easily extended to nonlinear cases

Bayes theorem: the posterior

$$\mathcal{P}(\mathbf{s}|\mathbf{d},\mathbf{p}) = \frac{\mathcal{P}(\mathbf{s}|\mathbf{p})\mathcal{P}(\mathbf{d}|\mathbf{s},\mathbf{p})}{\int \mathrm{d}\mathbf{s}\,\mathcal{P}(\mathbf{s}|\mathbf{p})\mathcal{P}(\mathbf{d}|\mathbf{s},\mathbf{p})}, \tag{1}$$

Posterior=prior×likelihood/evidence

$\longrightarrow$ clear representation of the assumptions
$\longrightarrow$ can be easily extended to nonlinear cases

Bayesian inference steps

- Definition of the prior: knowledge of the underlying signal
- Definition of the likelihood: nature of the observed data
- Linking the prior to the likelihood: link signal to the data
- Bayes theorem: the posterior
- Maximization of the posterior: MAP
- Sampling the posterior: MCMC

Bayesian inference steps

- Definition of the prior: knowledge of the underlying signal
- Definition of the likelihood: nature of the observed data
- Linking the prior to the likelihood: link signal to the data
- Bayes theorem: the posterior
- Maximization of the posterior: MAP
- Sampling the posterior: MCMC

Bayesian inference steps

- Definition of the prior: knowledge of the underlying signal
- Definition of the likelihood: nature of the observed data
- Linking the prior to the likelihood: link signal to the data
- Bayes theorem: the posterior
- Maximization of the posterior: MAP
- Sampling the posterior: MCMC

Bayesian inference steps

- Definition of the prior: knowledge of the underlying signal
- Definition of the likelihood: nature of the observed data
- Linking the prior to the likelihood: link signal to the data
- Bayes theorem: the posterior
- Maximization of the posterior: MAP
- Sampling the posterior: MCMC

Bayesian inference steps

- Definition of the prior: knowledge of the underlying signal
- Definition of the likelihood: nature of the observed data
- Linking the prior to the likelihood: link signal to the data
- Bayes theorem: the posterior
- Maximization of the posterior: MAP
- Sampling the posterior: MCMC

Bayesian inference steps

- Definition of the prior: knowledge of the underlying signal
- Definition of the likelihood: nature of the observed data
- Linking the prior to the likelihood: link signal to the data
- Bayes theorem: the posterior
- Maximization of the posterior: MAP
- Sampling the posterior: MCMC

Physical motivation: model I FK & Raul Angulo in prep

- Lagrangian perturbation theory: displacement field

$$\mathbf{x} = \mathbf{q} + \mathbf{\Psi} \,. \tag{2}$$

- Mass conservation:

$$\rho(\mathbf{x}, t)\mathrm{d}\mathbf{x} = \langle \rho(t_i) \rangle \mathrm{d}\mathbf{q} \,. \tag{3}$$

- The inverse of the Jacobean leads to the overdensity field:

$$1 + \delta(\mathbf{x}(\mathbf{q}, t)) = \mathbf{J}(\mathbf{q}, t)^{-1} \,, \tag{4}$$

with

$$\mathbf{J}(\mathbf{q}, t) = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{q}} \right| \,. \tag{5}$$

Physical motivation: model I FK & Raul Angulo in prep

- Lagrangian perturbation theory: displacement field

$$\mathbf{x} = \mathbf{q} + \boldsymbol{\Psi}\,. \tag{2}$$

- Mass conservation:

$$\rho(\mathbf{x}, t)\mathrm{d}\mathbf{x} = \langle\rho(t_i)\rangle\mathrm{d}\mathbf{q}\,. \tag{3}$$

- The inverse of the Jacobean leads to the overdensity field:

$$1 + \delta(\mathbf{x}(\mathbf{q}, t)) = \mathbf{J}(\mathbf{q}, t)^{-1}\,, \tag{4}$$

with

$$\mathbf{J}(\mathbf{q}, t) = \left|\frac{\partial\mathbf{x}}{\partial\mathbf{q}}\right|\,. \tag{5}$$

Physical motivation: model I FK & Raul Angulo in prep

- Lagrangian perturbation theory: displacement field

$$\mathbf{x} = \mathbf{q} + \mathbf{\Psi}\,. \tag{2}$$

- Mass conservation:

$$\rho(\mathbf{x}, t)\mathrm{d}\mathbf{x} = \langle\rho(t_i)\rangle\mathrm{d}\mathbf{q}\,. \tag{3}$$

- The inverse of the Jacobean leads to the overdensity field:

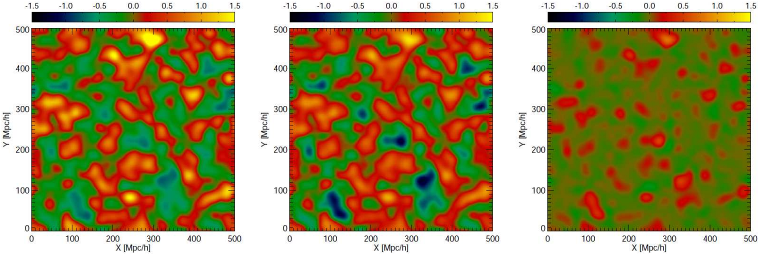
$$1 + \delta(\mathbf{x}(\mathbf{q}, t)) = \mathbf{J}(\mathbf{q}, t)^{-1}\,, \tag{4}$$

with

$$\mathbf{J}(\mathbf{q}, t) = \left|\frac{\partial \mathbf{x}}{\partial \mathbf{q}}\right|\,. \tag{5}$$

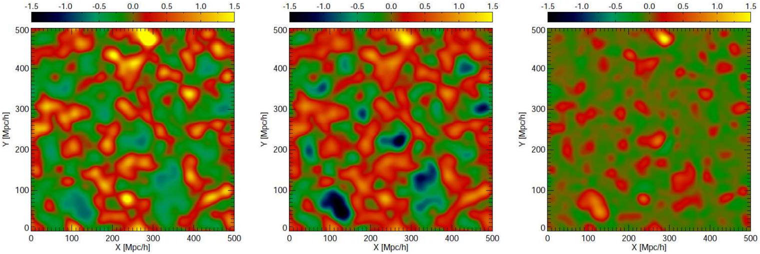
Testing model I with MR

Forward relation from $z = 127$ to $z = 0$:



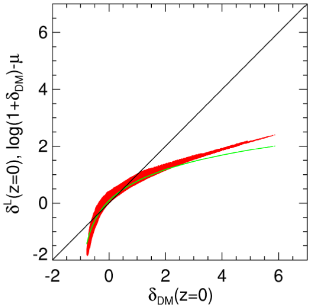
Testing model I with MR

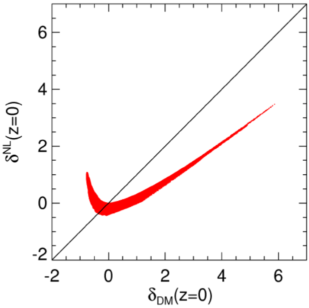
Inverse relation from $z = 0$ to $z = 127$:



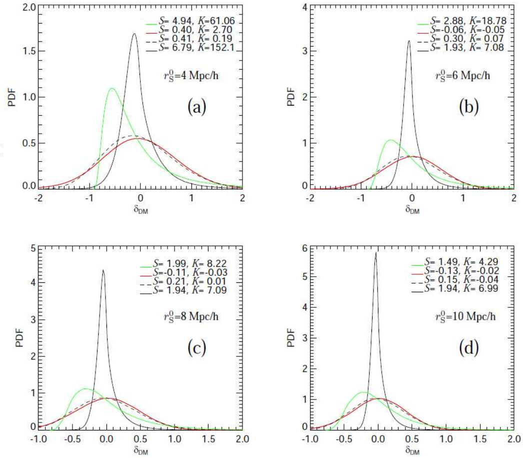
Testing model I and II with MR FK & Raul Angulo in prep

Cell-to-cell comparison:

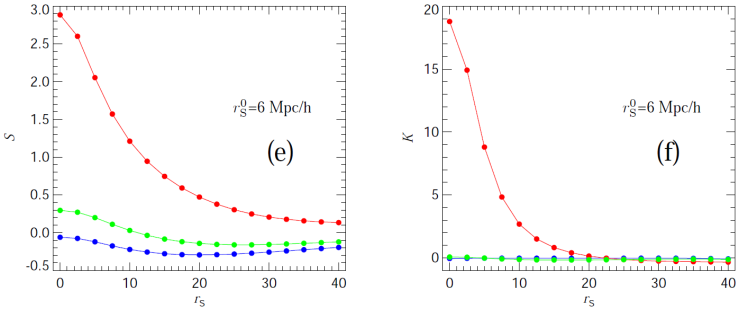




Testing model I and II with MR: matter statistics

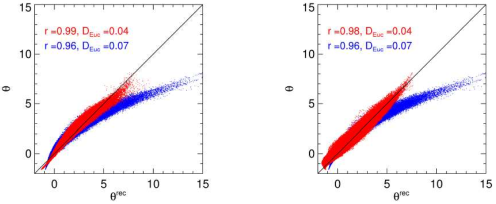


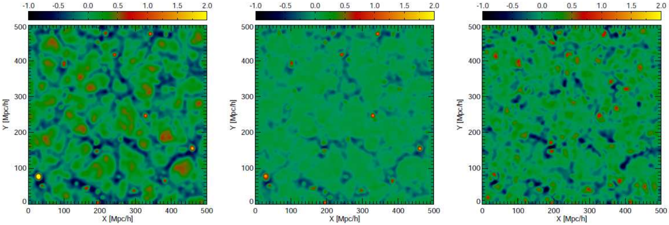
Testing model I and II with MR: matter statistics



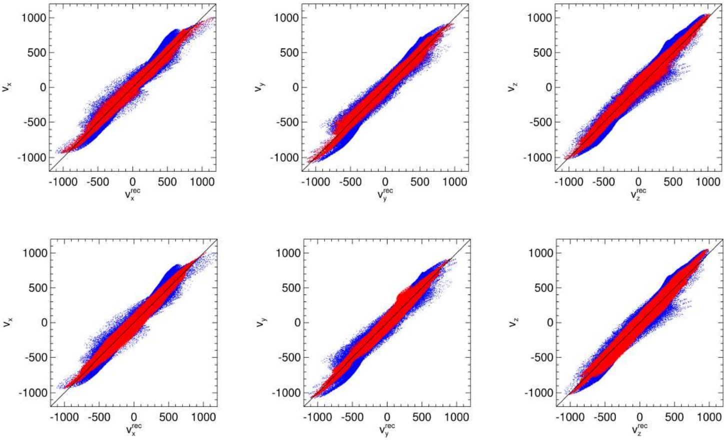
Testing model I and II with MR: peculiar motions

FK, Raul Angulo, Yehuda Hoffman & Stefan Gottloeber in prep

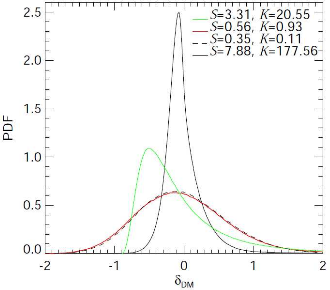


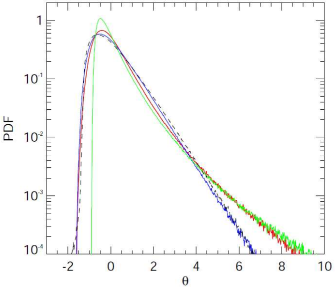


Testing model I and II with MR: peculiar motions

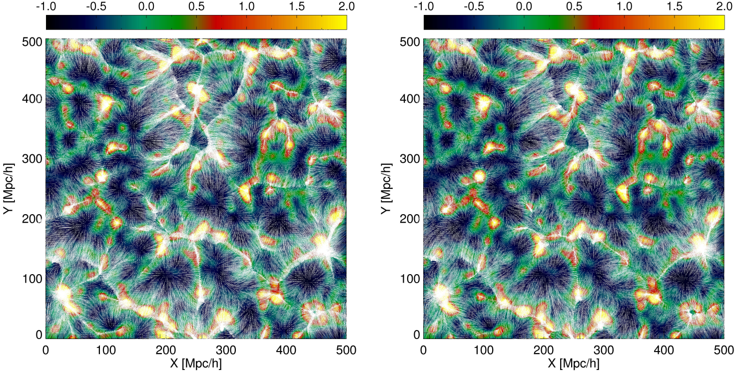


Testing model I and II with MR: peculiar motions

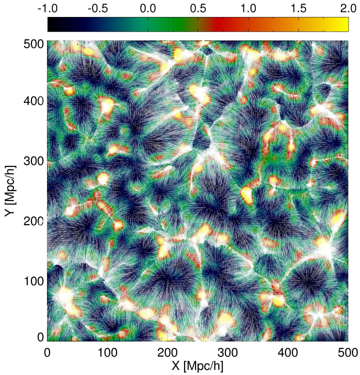


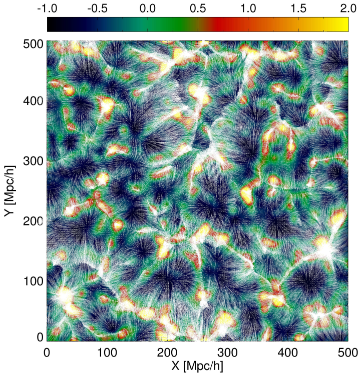


Testing model I and II with MR: peculiar motions



Testing model I and II with MR: peculiar motions





Non-Gaussian expansions

- relax the Gaussian-Lognormal assumption
- skewed Gaussian Edgeworth expansion (univariate case: Juszkiewicz, Bouchet & Colombi 93)
- skewed lognormal model with the 1D Edgeworth exansion (univariate case: Colombi 94)

Non-Gaussian expansions: multivariate case FK 2010

- Let me introduce here the multivariate case:

$$\Phi_i \equiv \ln \rho_i - \langle \ln \rho \rangle = s_i - \mu_i, \;\; \nu_i \equiv \sum_j S_{ij}^{-1/2} \Phi_i \,, \qquad (6)$$

$s_i = \ln(\rho_i / \langle \rho \rangle) = \ln(1 + \delta_{\mathrm{M}i})$

- Multidimensional Edgeworth expansion

$$P(\Phi) = G(\nu) \big[1 + \frac{1}{3!} \sum_{i'j'k'} \langle \Phi_{i'} \Phi_{j'} \Phi_{k'} \rangle \sum_{ijk} S_{ii'}^{-1/2} S_{jj'}^{-1/2} S_{kk'}^{-1/2} h_{ijk}(\nu)$$
$$+ \frac{1}{4!} \sum_{i'j'k'l'} \langle \Phi_{i'} \Phi_{j'} \Phi_{k'} \Phi_{l'} \rangle \sum_{ijkl} S_{ii'}^{-1/2} S_{jj'}^{-1/2} S_{kk'}^{-1/2} S_{ll'}^{-1/2} h_{ijkl}(\nu) + \ldots \big] \,, \qquad (7)$$

Non-Gaussian expansions: multivariate case FK 2010

- Let me introduce here the multivariate case:

$$\Phi_i \equiv \ln\rho_i - \langle \ln\rho \rangle = s_i - \mu_i,\ \ \nu_i \equiv \sum_j S_{ij}^{-1/2}\Phi_i\,, \tag{6}$$

$s_i = \ln(\rho_i/\langle\rho\rangle) = \ln(1+\delta_{\mathrm{M}i})$

- Multidimensional Edgeworth expansion

$$P(\mathbf{\Phi}) = G(\boldsymbol{\nu})\big[1 + \frac{1}{3!}\sum_{i'j'k'}\langle\Phi_{i'}\Phi_{j'}\Phi_{k'}\rangle\sum_{ijk} S_{ii'}^{-1/2}S_{jj'}^{-1/2}S_{kk'}^{-1/2}h_{ijk}(\boldsymbol{\nu})$$
$$+\frac{1}{4!}\sum_{i'j'k'l'}\langle\Phi_{i'}\Phi_{j'}\Phi_{k'}\Phi_{l'}\rangle\sum_{ijkl} S_{ii'}^{-1/2}S_{jj'}^{-1/2}S_{kk'}^{-1/2}S_{ll'}^{-1/2}h_{ijkl}(\boldsymbol{\nu}) + \dots\big]\,, \tag{7}$$

Lognormal model

- 0th order: lognormal model ($G(\boldsymbol{\nu}) \rightarrow \mathcal{P}(\boldsymbol{\delta}_{\rm M}|\mathbf{S})$)
  $\langle \ln \rho_i \rangle = \ln\langle \rho \rangle + \mu_i$

$$\mathcal{P}(\boldsymbol{\delta}_{\rm M}|\mathbf{S}) = \frac{1}{\sqrt{(2\pi)^{N_{\rm cells}} {\rm det}(\mathbf{S})}} \prod_k \frac{1}{1+\delta_{{\rm M}k}} \times {\rm exp}\left(-\frac{1}{2}\sum_{ij}\left({\rm ln}(1+\delta_{{\rm M}i})-\mu_i\right) S^{-1}_{ij} \left({\rm ln}(1+\delta_{{\rm M}j})-\mu_j\right)\right), \quad (8)$$

  multidimensional implementation for matter field reconstructions:
  FK, Jasche & Metcalf 2009; applied to SDSS DR7: Jasche, FK, Li & Ensslin 2010

- when $\delta_{\rm M} \ll 1 \rightarrow$ Gauss distribution

Lognormal model

- 0th order: lognormal model ($G(\boldsymbol{\nu}) \rightarrow \mathcal{P}(\boldsymbol{\delta}_{\rm M}|\mathbf{S})$)
  $\langle \ln \rho_i \rangle = \ln \langle \rho \rangle + \mu_i$

$$\mathcal{P}(\boldsymbol{\delta}_{\rm M}|\mathbf{S}) = \frac{1}{\sqrt{(2\pi)^{N_{\rm cells}} {\rm det}(\mathbf{S})}} \prod_k \frac{1}{1+\delta_{{\rm M}k}} \times {\rm exp}\left(-\frac{1}{2}\sum_{ij}\left({\rm ln}(1+\delta_{{\rm M}i})-\mu_i\right) S^{-1}_{ij} \left({\rm ln}(1+\delta_{{\rm M}j})-\mu_j\right)\right) , \quad (8)$$

  multidimensional implementation for matter field reconstructions:
  FK, Jasche & Metcalf 2009; applied to SDSS DR7: Jasche, FK, Li & Ensslin 2010

- when $\delta_{\rm M} \ll 1 \rightarrow$ Gauss distribution

Definition of the likelihood

- Discreteness of the galaxy distribution: shot noise
- Galaxy selection function in a magnitude limited survey
- Link between underlying matter field and the galaxy field
- Discrete Press-Schechter: Borel distribution (Epstein 1983)
  Gravitothermal dynamics (Saslaw 1986, Itoh et al 1988, Sheth 1995)
- Non-Poissonian distribution (FK in prep)

$$\mathcal{P}(\{N_k^{\rm g}\}|\{\lambda_k\},\mathbf{Q}) = \prod_k \frac{\sum_j(\delta^{\rm K}_{k,j}-Q_{k,j})\lambda_j}{N_k^{\rm g}!} \times\left(\sum_l(\delta^{\rm K}_{k,l}-Q_{k,l})\lambda_l+\sum_m Q_{k,m}N_m^{\rm g}\right)^{N_k^{\rm g}-1} \times\exp\left(-\sum_n(\delta^{\rm K}_{k,n}-Q_{k,n})\lambda_n-\sum_o Q_{k,o}N_o^{\rm g}\right) \quad (9)$$

Define the likelihood: nature of the observable

- For a sparse sample: Poisson limit

$$\mathcal{P}(\{N_k^{\rm g}\}|\{\lambda_k\}) = \prod_k \frac{\lambda_k^{N_k^{\rm g}} \exp\left(-\lambda_k\right)}{N_k^{\rm g}!} \tag{10}$$

full treatment: FK & Ensslin 2008; FK, Jasche & Metcalf 2009

- Observation process: radial selection function and sky mask binomial process: *either we see the galaxy or not*

$$\lambda_i \equiv w_i \lambda_i' \tag{11}$$

treatment proposed in FK, Jasche, Li, Ensslin, Metcalf, Wandelt, Lemson & White 2009

- the correlation is encoded in the underlying density field in λ_k

Link between the prior and the likelihood

- galaxy bias

$$\delta_{\mathrm{g}i} = B(\boldsymbol{\delta}_{\mathrm{M}})_i, \tag{12}$$

- Fry & Gaztanaga 93 (generalized)

$$\delta_{\mathrm{g}i} = \sum_j B^1_{ij}\delta_{\mathrm{M}j} + \delta_{\mathrm{M}i}\sum_j B^2_{ij}\delta_{\mathrm{M}j} + \dots, \tag{13}$$

Response operator

- Response operator

$$\lambda_k = \lambda_k(\boldsymbol{\delta}_{\rm M}) = R(\boldsymbol{\delta}_{\rm g}(\boldsymbol{\delta}_{\rm M}))_k\,, \tag{14}$$

-

$$\lambda_k = w_k \bar{N}(1 + B(\boldsymbol{\delta}_{\rm M})_k)\,, \tag{15}$$

-

$$\lambda_k = w_k \bar{N}(1 + b\boldsymbol{\delta}_{{\rm M}k})\,, \tag{16}$$

-

Bayes theorem: the posterior

$$\mathcal{P}(\boldsymbol{\delta}_{\rm M}|\mathbf{N},\mathbf{S}) = \frac{\mathcal{P}(\boldsymbol{\delta}_{\rm M}|\mathbf{S})\mathcal{P}(\mathbf{N}|\boldsymbol{\lambda}(\boldsymbol{\delta}_{\rm M}))}{\int {\rm d}\boldsymbol{\delta}_{\rm M}\,\mathcal{P}(\boldsymbol{\delta}_{\rm M}|\mathbf{S})\mathcal{P}(\mathbf{N}_k|\boldsymbol{\lambda}(\boldsymbol{\delta}_{\rm M}))}, \tag{17}$$

Bayes theorem: the posterior

$$
\begin{aligned}
&\mathcal{P}(\boldsymbol{\delta}_{\rm M}|\mathbf{N},\mathbf{S}) \\
&\propto \prod_l \frac{1}{1+\delta_{{\rm M}l}} \exp\left(-\frac{1}{2}\sum_{ij}\left(\ln\left(1+\delta_{{\rm M}i}\right)-\mu_i\right) S^{-1}_{ij}\left(\ln\left(1+\delta_{{\rm M}j}\right)-\mu_j\right)\right) \\
&\times[1+\frac{1}{3!}\sum_{i'j'k'}\langle\Phi_{i'}\Phi_{j'}\Phi_{k'}\rangle\sum_{ijk} S^{-1/2}_{ii'}S^{-1/2}_{jj'}S^{-1/2}_{kk'}h_{ijk}(\boldsymbol{\nu}) \\
&+\frac{1}{4!}\sum_{i'j'k'l'}\langle\Phi_{i'}\Phi_{j'}\Phi_{k'}\Phi_{l'}\rangle\sum_{ijkl} S^{-1/2}_{ii'}S^{-1/2}_{jj'}S^{-1/2}_{kk'}S^{-1/2}_{ll'}h_{ijkl}(\boldsymbol{\nu})+\dots] \\
&\times\prod_k \frac{\left(w_k\bar{N}(1+b\boldsymbol{\delta}_{{\rm M}k})\right)^{N_k}\exp\left(-w_k\bar{N}(1+b\boldsymbol{\delta}_{{\rm M}k})\right)}{N_k!}, \qquad (18)
\end{aligned}
$$

Bayes theorem: the posterior

Maximum a posteriori

- Let us define the *energy* $E(\mathbf{s})$

$$E(\mathbf{s}) \equiv -\ln\left(\mathcal{P}\left(\mathbf{s}|\mathbf{N},\mathbf{S}\right)\right), \tag{19}$$

- MAP

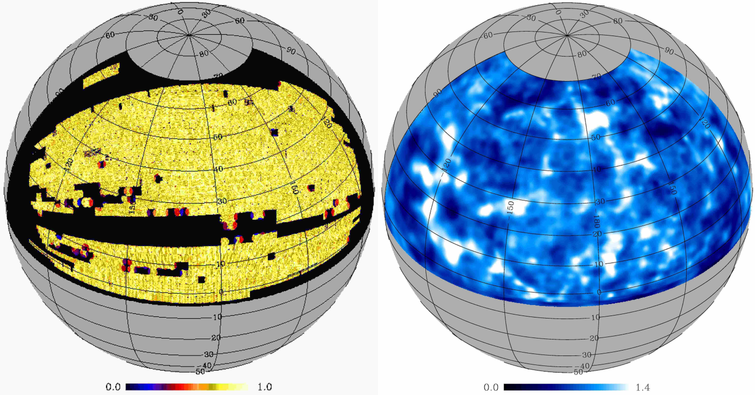
$$\frac{\partial E(\mathbf{s})}{\partial s_l} = 0, \tag{20}$$

- Krylov conjugate gradient schemes (FK & Ensslin 2008; Jasche, FK, Wandelt, Ensslin 2009; FK, Jasche, & Metcalf 2009)

$$s_i^{j+1} = s_i^j - \sum_k T_{ik}\frac{\partial E(\mathbf{s})}{\partial s_k}, \tag{21}$$

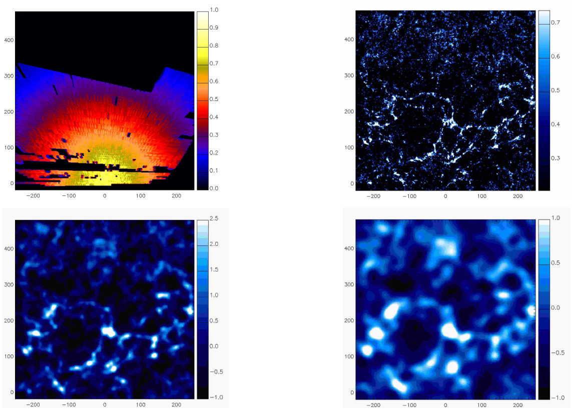
Results: Wiener filter reconstruction of the SDSS DR6

Wiener-filter with the ARGO code: FK, Jasche, Li, Ensslin, Metcalf, Wandelt, Lemson & White 2009



Results: detection of a super-void in the SDSS DR6

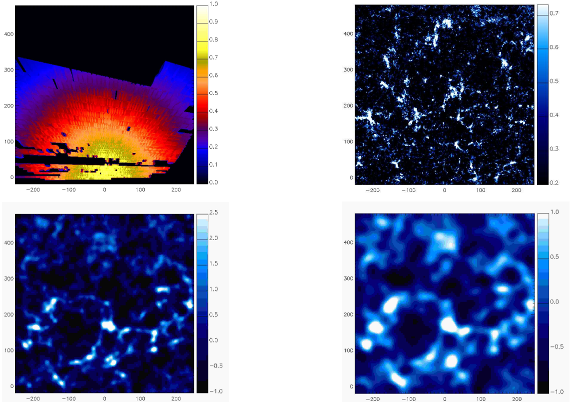
(about 250.000 galaxies from the main sample)



FK, Jasche, Li, Ensslin, Metcalf, Wandelt, Lemson & White 2009

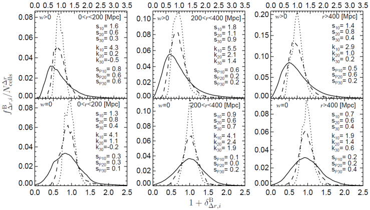
Results: detection of a super-void in the SDSS DR6

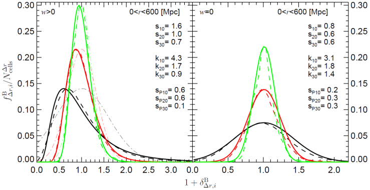
(about 250.000 galaxies from the main sample)→ cluster prediction



FK, Jasche, Li, Ensslin, Metcalf, Wandelt, Lemson & White 2009

Results: matter statistics in the SDSS DR6

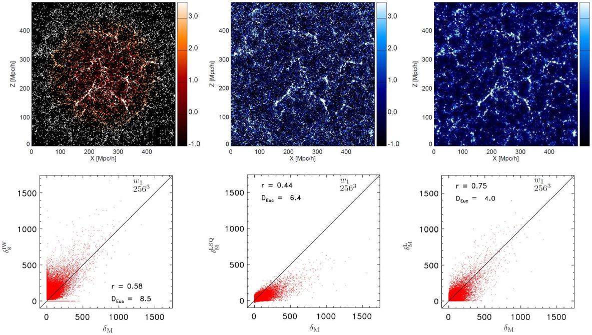




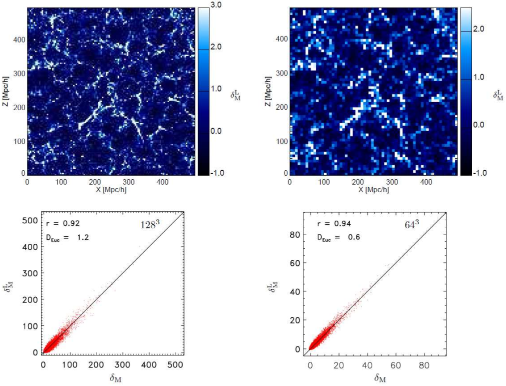
FK, Jasche, Li, Ensslin, Metcalf, Wandelt, Lemson & White 2009

Results: lognormal filter against Wiener filter and inverse weighting

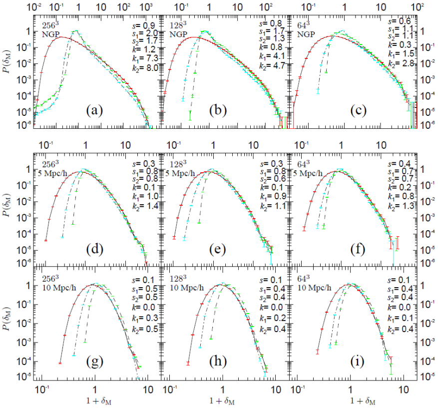
tests with the Millenium Run including selection function effects (about 350.000 mock galaxies)



Results: lognormal filter against Wiener filter and inverse weighting



Results: matter statistics in the lognormal reconstruction



FK, Jasche & Metcalf 2009 (upgrade of the ARGO code)

Sampling the posterior

- Hamiltonian sampling (Taylor et al 2010, Jasche & FK 2010, FK, Simona Gallerani & Andrea Ferrara 2010)

$$H(\mathbf{s},\mathbf{p}) = K(\mathbf{p}) + E(\mathbf{s}), \tag{22}$$

- kinetic term with a given mass as the variance for the momenta

$$K(\mathbf{p}) = \frac{1}{2}\mathbf{p}^{\dagger}\mathbf{M}^{-1}\mathbf{p}, \tag{23}$$

- Marginalization over the momenta

$$P(\mathbf{s},\mathbf{p}) = \frac{e^{-H}}{Z_H} = \frac{e^{-K}}{Z_K}\frac{e^{-E}}{Z_E} = P(\mathbf{p})P(\mathbf{s}), \tag{24}$$

Sampling the posterior

- Hamiltonian evolution equations: $(\mathbf{s},\mathbf{p}) \rightarrow (\mathbf{s}',\mathbf{p}')$

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\frac{\partial H}{\partial \mathbf{s}} = -\frac{\partial E}{\partial \mathbf{s}}, \tag{25}$$

$$\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p}, \tag{26}$$

- Metropolis-Hastings acceptance step

$$p_a = \min(1, e^{-\delta H}), \tag{27}$$

$\delta H = H(\mathbf{s}',\mathbf{p}') - H(\mathbf{s},\mathbf{p}) \rightarrow$ we do not care about the evidence!

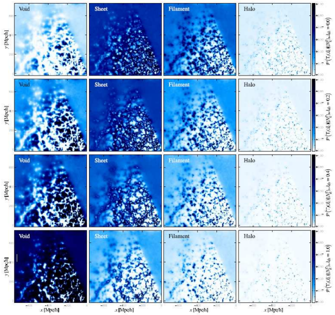
Results: matter field reconstruction of the SDSS DR7 with Hamiltonian sampling and the lognormal prior

- deformation tensor of the grav. Pot. Φ:

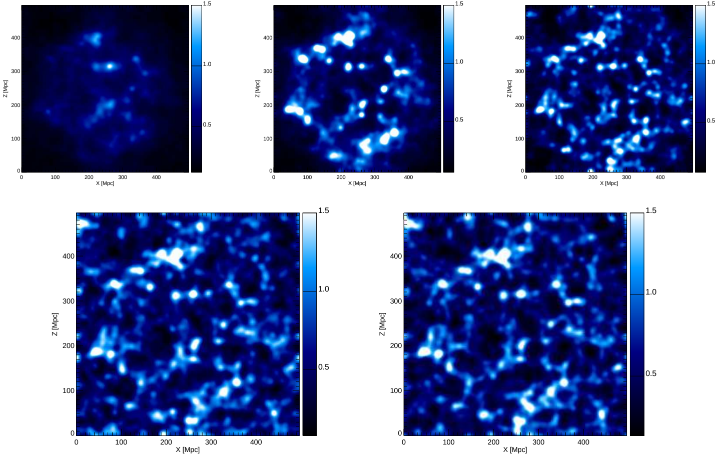
$$T_{ij} = \frac{\partial^2 \Phi}{\partial x_i \partial x_j} \qquad (28)$$

- eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$, > 0: contraction, < 0: expansion
- classification: void: all< 0, sheet $1\lambda > 0$, filament $2\lambda > 0$, halo: $3\lambda > 0$ (with threshold see Forero-Romero et al 2009)

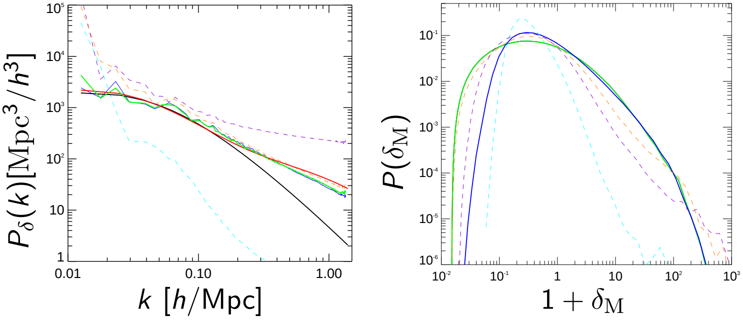
Jasche, FK, Li & Ensslin 2010



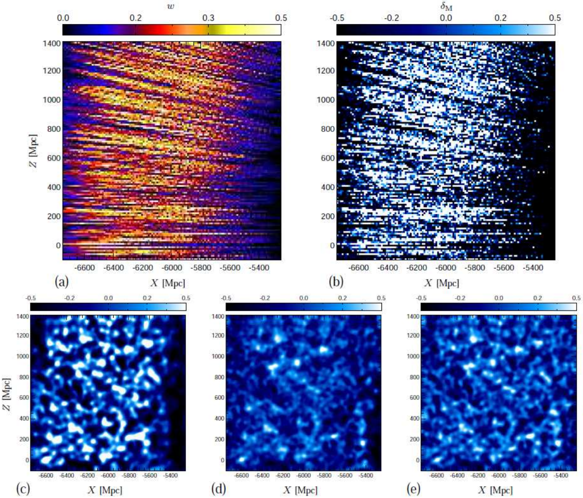
Skewed matter statistics: FK in prep



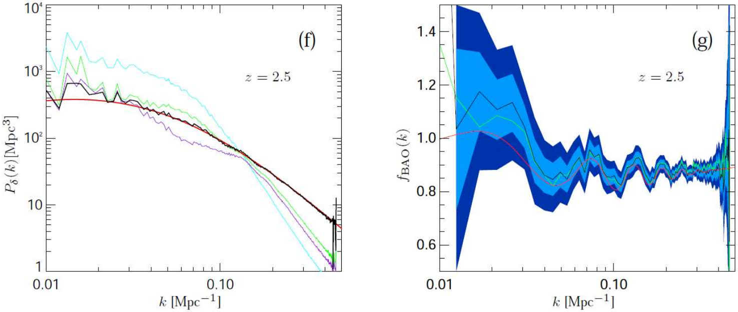
Skewed matter statistics



Lyman alpha forest 3D



Lyman alpha forest 3D



FK, Simona Gallerani & Andrea Ferrara 2010

Simona Gallerani, FK & Andrea Ferrara 2010

Conclusions

- There is a need to compare observations with theory as precisely as possible.
- Observations are plagued by many uncertainties which require a statistical treatment.
- The Bayesian approach is flexible and clear.
- We have shown that we can deal with complex models in this framework.
- There is a lot to do in Cosmology!