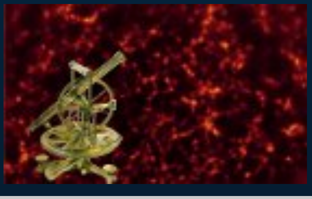
Bayesian Large Scale Structure inference

Jens Jasche

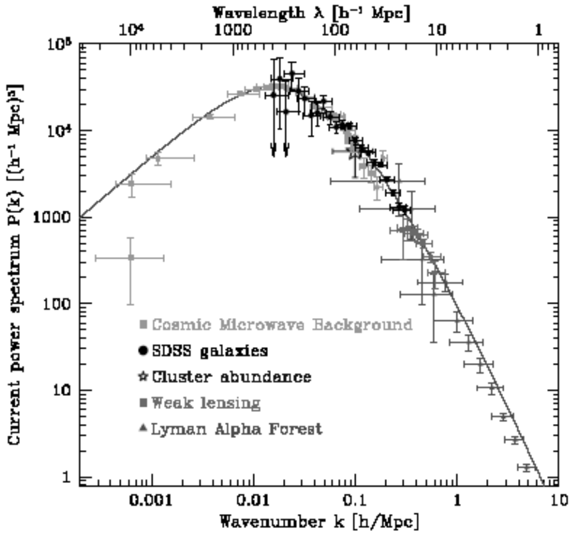
Argelander-Institut für Astronomie

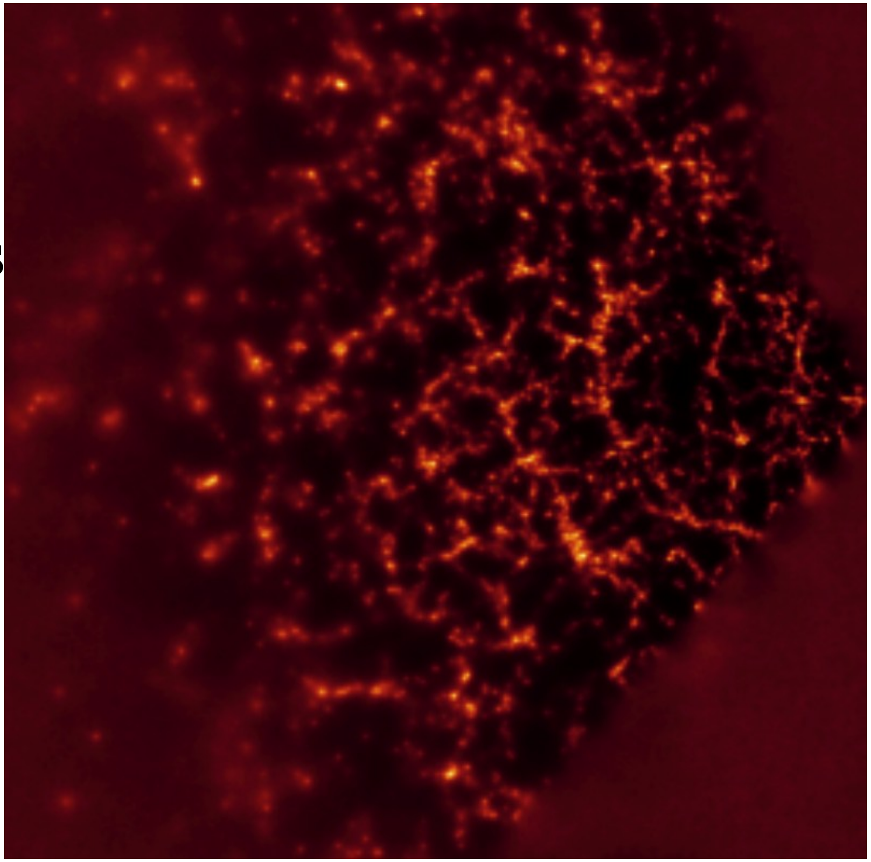
Heidelberg, 20 September 2011

Motivation



- Goal: 3D cosmography
  - Map matter distribution
  - Quantify statistical properties





Tegmark et al. (2004)

Jasche et al. (2010)

Motivation

- No ideal Observations in reality

Motivation

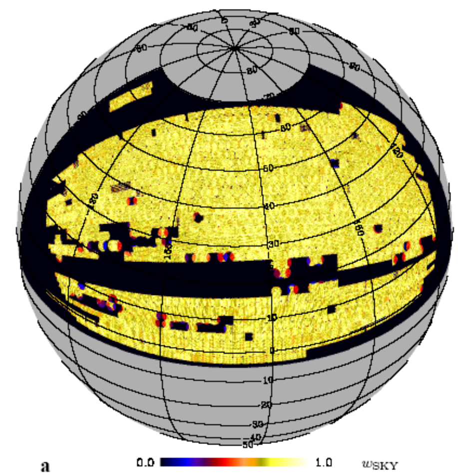
- No ideal Observations in reality
  - Statistical uncertainties

Noise, cosmic variance etc.

Motivation

- No ideal Observations in reality
  - Statistical uncertainties
  - Systematics

Noise, cosmic variance etc.



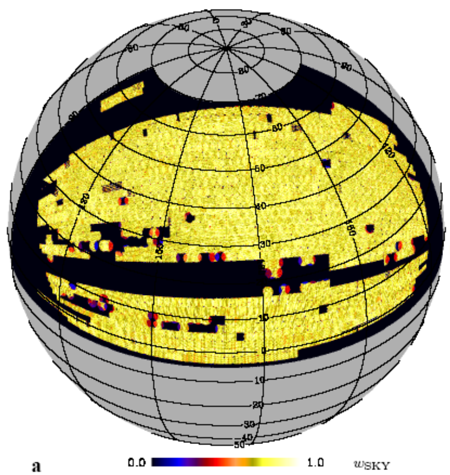
Kitaura et al. (2009)

Survey geometry, selection effects, biases

Motivation

- No ideal Observations in reality
  - Statistical uncertainties
  - Systematics

Noise, cosmic variance etc.

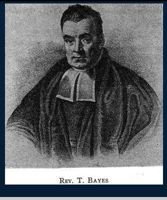


Kitaura et al. (2009)

Survey geometry, selection effects, biases

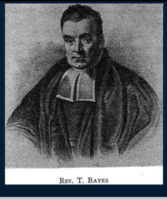
➡ No unique recovery possible!!!

Bayesian Approach



“Which are the possible signals compatible with the observations?”

Bayesian Approach

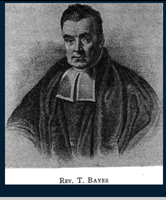


"Which are the possible signals compatible with the observations?"

- Object of interest: Signal posterior distribution

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

Bayesian Approach



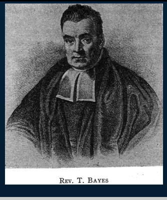
“Which are the possible signals compatible with the observations?”

- Object of interest: Signal posterior distribution

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

Prior

Bayesian Approach



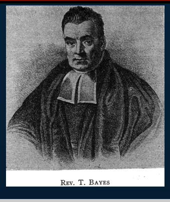
"Which are the possible signals compatible with the observations?"

- Object of interest: Signal posterior distribution

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

Likelihood

Bayesian Approach



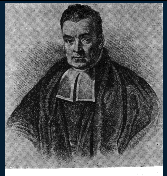
“Which are the possible signals compatible with the observations?”

- Object of interest: Signal posterior distribution

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

Evidence

Bayesian Approach



Rev. T. Bayes

"Which are the possible signals compatible with the observations?"

- Object of interest: Signal posterior distribution

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

- We can do science!
  - Model comparison
  - Parameter studies
  - Report statistical summaries
  - Non-linear, Non-Gaussian error propagation

LSS Posterior

- Aim: inference of non-linear density fields
  - non-linear density field
    - lognormal See e.g. Coles & Jones (1991), Kayo et al. (2001)
  - "correct" noise treatment
    - full Poissonian distribution

- Problem: Non-Gaussian sampling in high dimensions
  - direct sampling not possible
  - usual Metropolis-Hastings algorithm inefficient

➡ **HADES** (HAmiltonian Density Estimation and Sampling)

Jasche, Kitaura (2010)

LSS inference with the SDSS

- Application of HADES to SDSS DR7
  - cubic, equidistant box with sidelength 750 Mpc
  - ~ 3 Mpc grid resolution
  - ~ 10^7 volume elements / parameters

Jasche, Kitaura, Li, Enßlin (2010)

LSS inference with the SDSS

- Application of HADES to SDSS DR7
  - cubic, equidistant box with sidelength 750 Mpc
  - ~ 3 Mpc grid resolution
  - ~ 10^7 volume elements / parameters

- Goal: provide a representation of the SDSS density posterior
  - to provide 3D cosmographic descriptions
  - to quantify uncertainties of the density distribution

Jasche, Kitaura, Li, Enßlin (2010)

LSS inference with the SDSS

- Application of HADES to SDSS DR7
  - cubic, equidistant box with sidelength 750 Mpc
  - ~ 3 Mpc grid resolution
  - ~ 10^7 volume elements / parameters

- Goal: provide a representation of the SDSS density posterior
  - to provide 3D cosmographic descriptions
  - to quantify uncertainties of the density distribution
    - 3 TB, 40,000 density samples

Jasche, Kitaura, Li, Enßlin (2010)

LSS inference with the SDSS

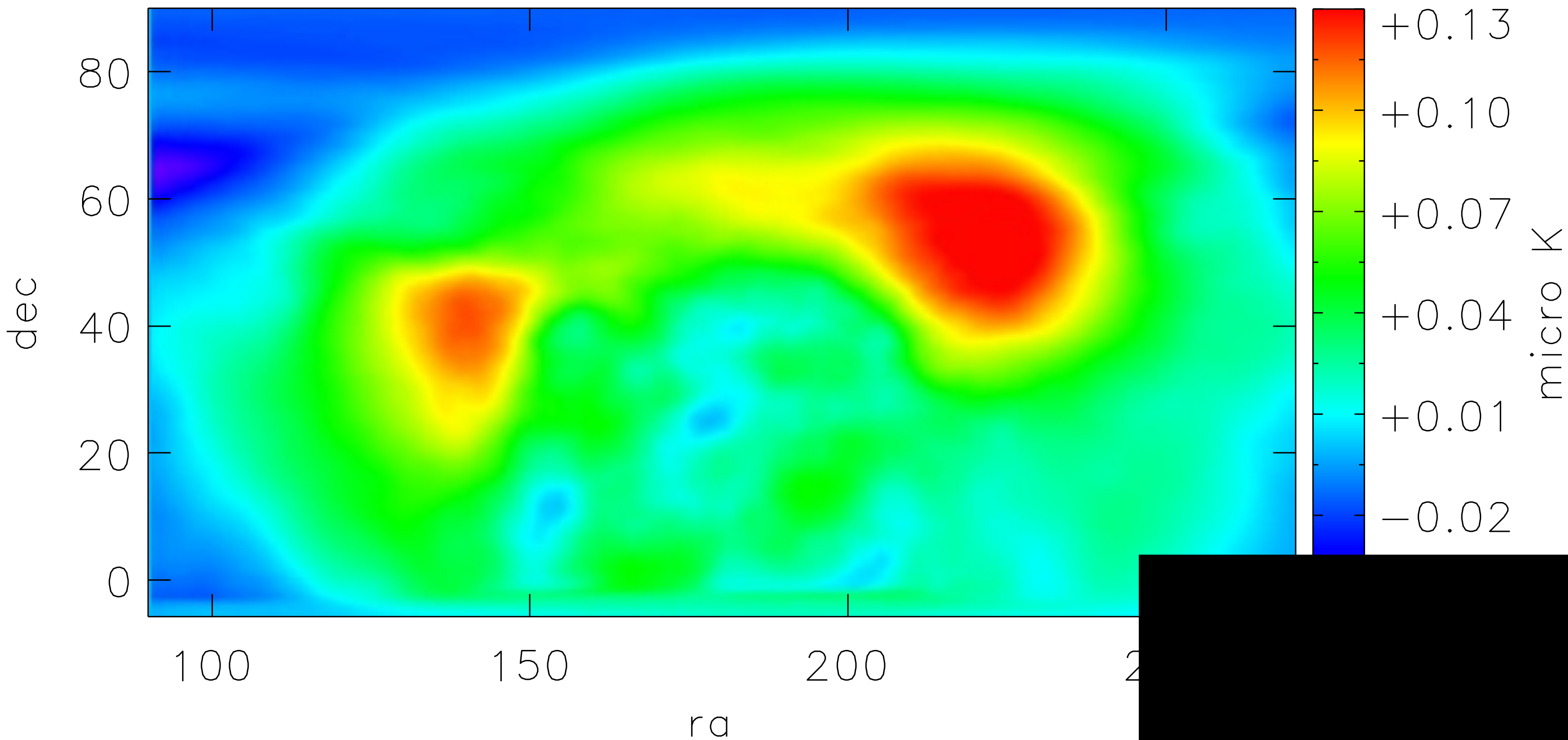
What are the possible density fields
Compatible with the observations ?

LSS inference with the SDSS

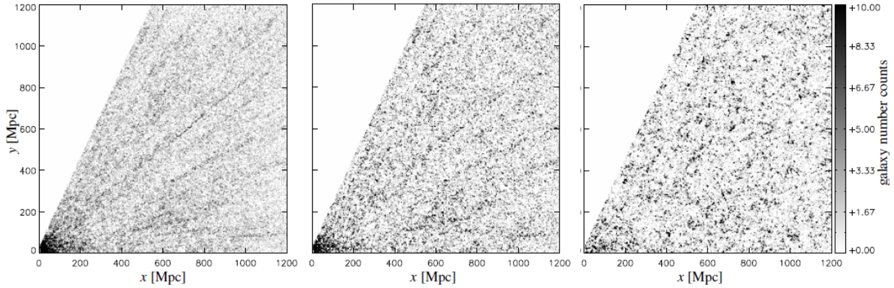
What are the possible density fields
Compatible with the observations ?

LSS inference with the SDSS

- ISW-templates



Photometric redshift sampling



Jasche, Wandelt (2011)

Summary & Conclusion

- Bayesian statistics answers the question:

"What is the probability distribution of possible signals compatible with the data?"

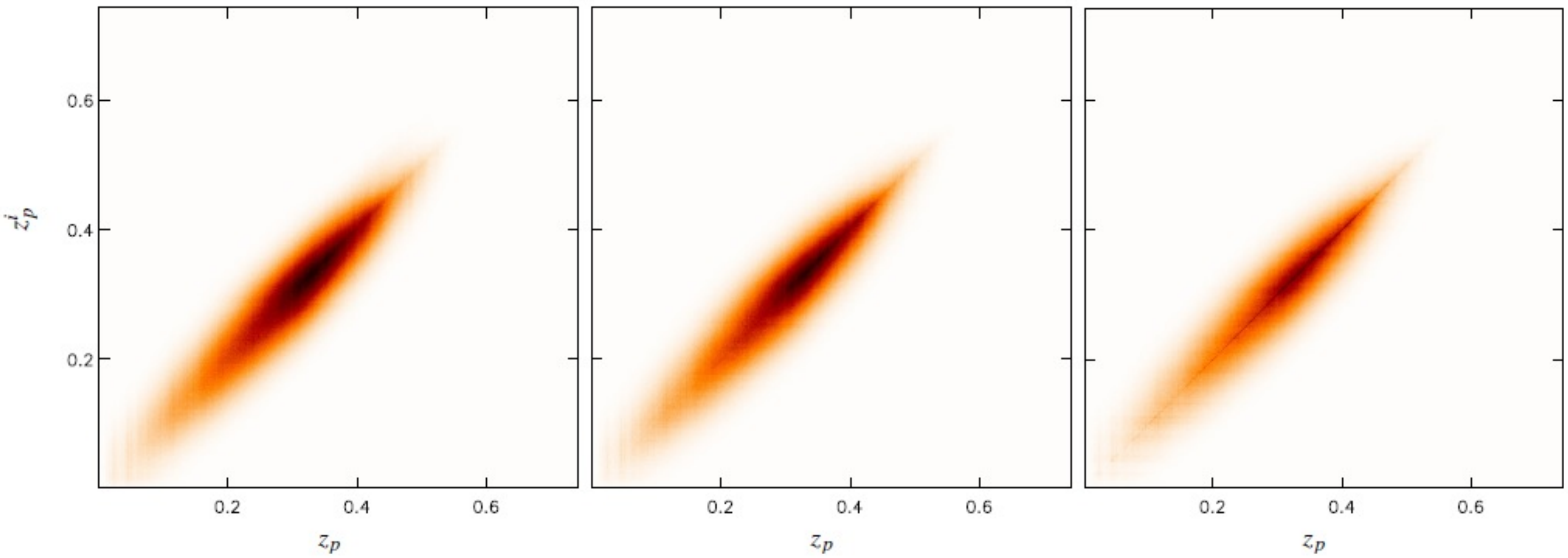
  - signal posterior distribution

- LSS inference framework: HADES
  - Correction of systematic effects, survey geometry, selection effectse
  - Precision non-linear density inference
  - accurate treatment of Poissonian shot noise
  - non-linear, non-Gaussian error propagation

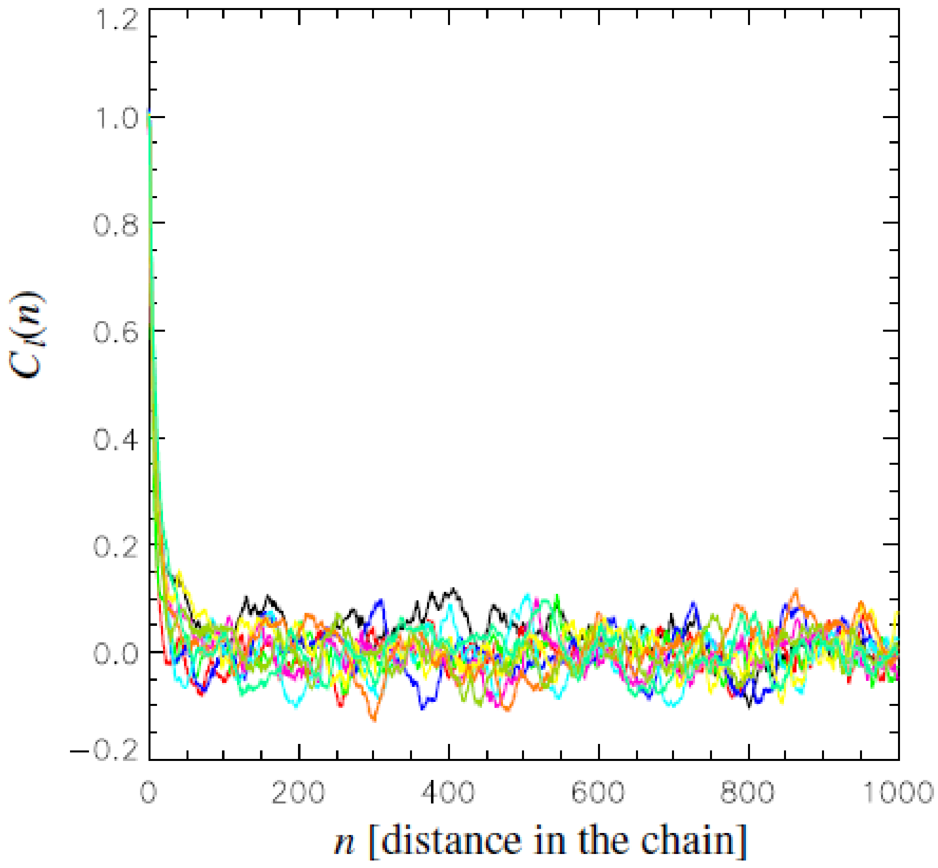
The End . . .

Thank you

Preliminary results



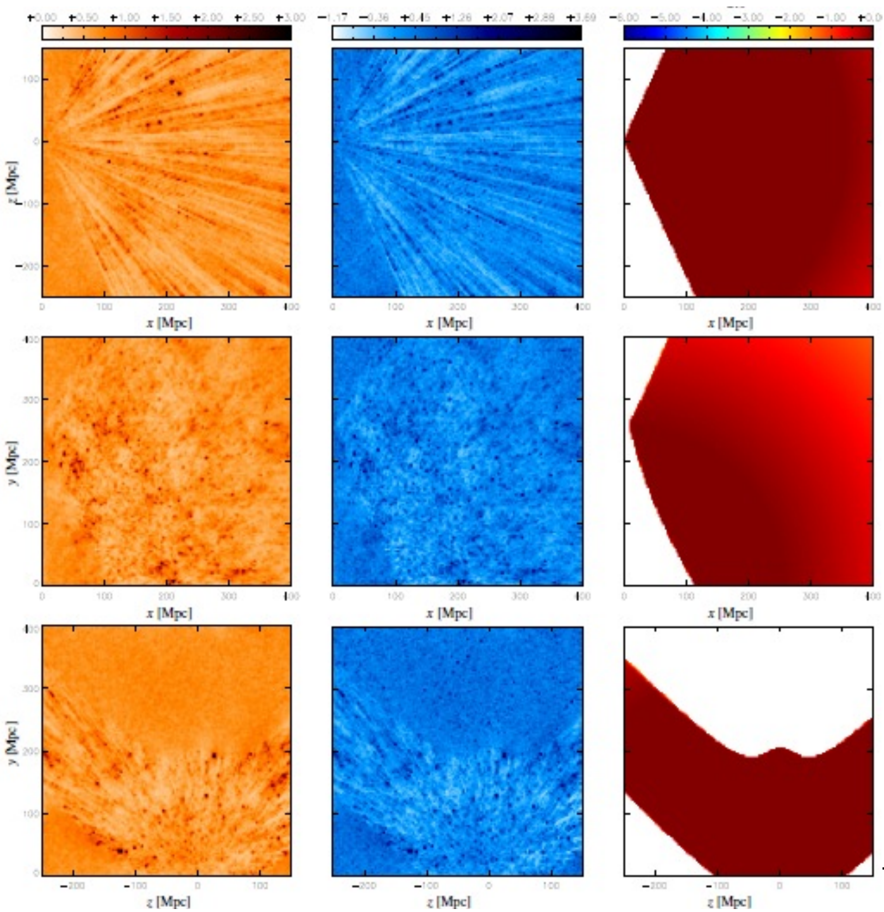
Preliminary results



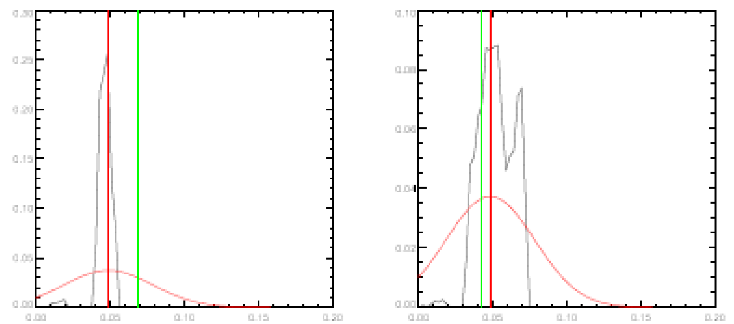
J. Jasche, Bayesian LSS Inference

Preliminary results

Density field



Marginalized redshift posteriors



SS Inference

The End …

LSS inference with the GAMA survey

By

Jens Jasche, Cristiano Porciani and Peder Norberg

J. Jasche, Bayesian LSS Inference

MCMC sampling

- Problem: How to evaluate the posterior distribution?
  - High dimensional parameter spaces (>10^6 volume elements)
  - Difficult functional shapes
  - Non-linear density inference
- Approximate posterior distribution

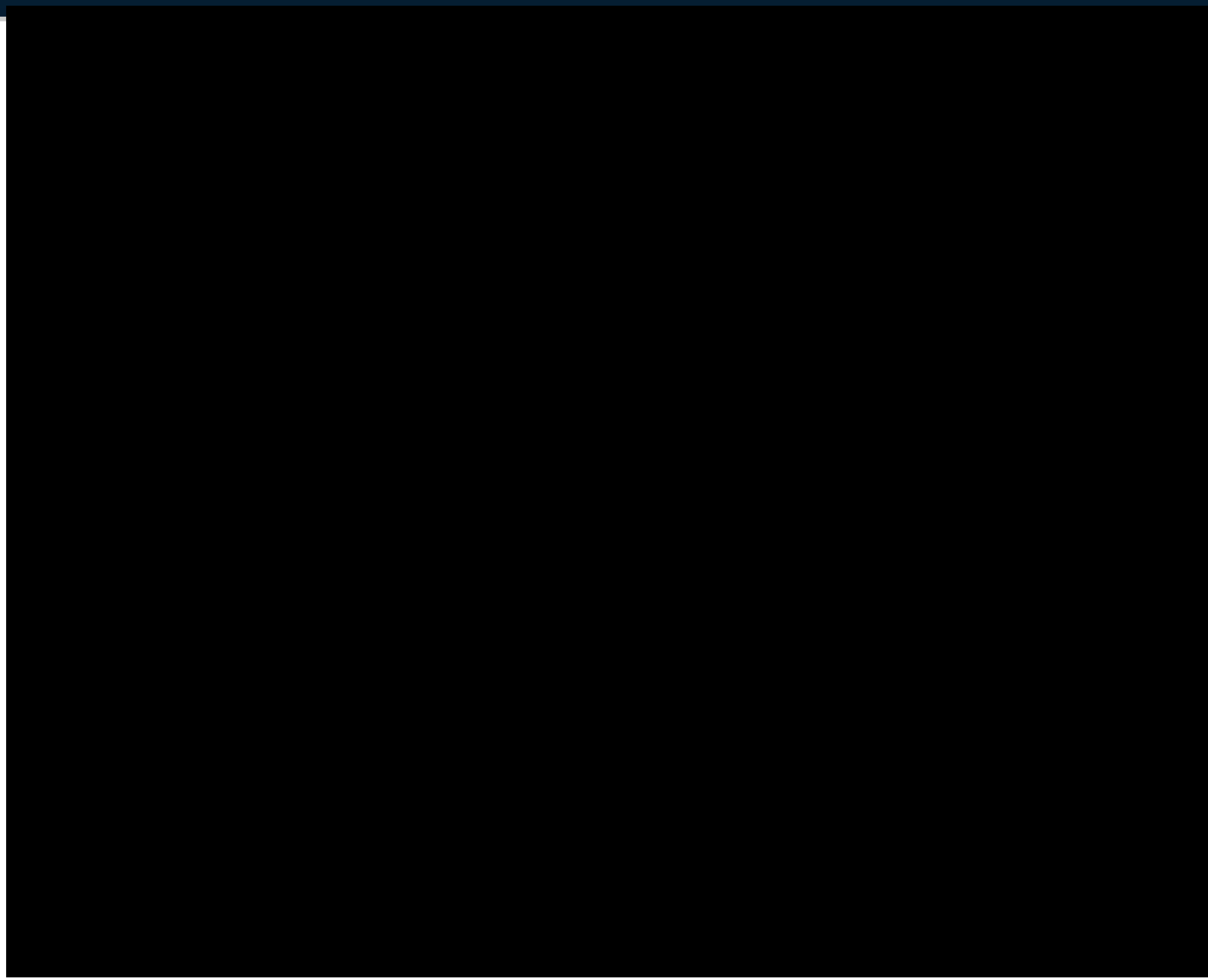
$$\mathcal{P}(s|d) = \frac{1}{N} \sum_{i=1}^{N} \delta^D(s - s_i)$$

➡ We need efficient MCMC algorithms!

The End …

Thank you

LSS inference with the GAMA survey



Extensions of the sampler

- Multiple block Metropolis Hastings sampling

$$(\xi_1, \xi_2^j, ..., \xi_M^j) \curvearrowleft \mathcal{P}(\xi_1, \xi_2^j, ..., \xi_M^j | d)$$

1) $\xi_1^{(j+1)} \curvearrowleft \mathcal{P}(\xi_1 | \xi_2^j, ..., \xi_M^j, d)$
2) $\xi_2^{(j+1)} \curvearrowleft \mathcal{P}(\xi_2 | \xi_1^{j+1}, \xi_3^j, ..., \xi_M^j, d)$

.

.

.

M) $\xi_M^{(j+1)} \curvearrowleft \mathcal{P}(\xi_M | \xi_1^{j+1}, \xi_2^{j+1}, ..., \xi_{M-1}^{j+1})$

Extensions of the sampler

- Multiple block Metropolis Hastings sampling

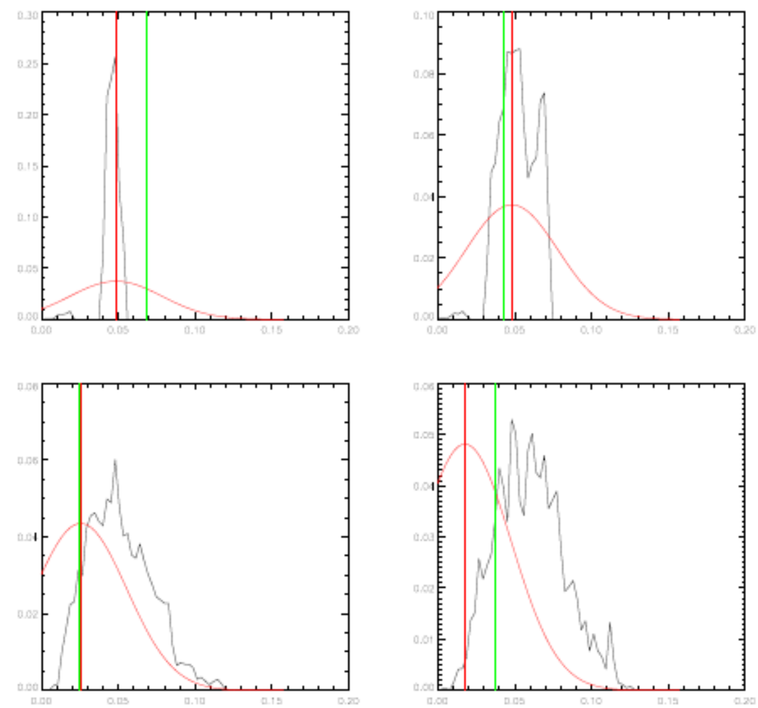
$$(\xi_1, \xi_2^j, ..., \xi_M^j) \curvearrowleft \mathcal{P}(\xi_1, \xi_2^j, ..., \xi_M^j | d)$$

$$
\begin{array}{ll}
1) & \xi_1^{(j+1)} \curvearrowleft \mathcal{P}(\xi_1 | \xi_2^j, ..., \xi_M^j, d) \\
2) & \xi_2^{(j+1)} \curvearrowleft \mathcal{P}(\xi_2 | \xi_1^{j+1}, \xi_3^j, ..., \xi_M^j, d) \\
. & \\
. & \\
. & \\
\mathrm{M}) & \xi_M^{(j+1)} \curvearrowleft \mathcal{P}(\xi_M | \xi_1^{j+1}, \xi_2^{j+1}, ..., \xi_{M-1}^{j+1})
\end{array}
$$

➡ Example redshift sampling

Redshift sampling

Preliminary Results



J. Jasche, Bayesian LSS Inference

Redshift sampling