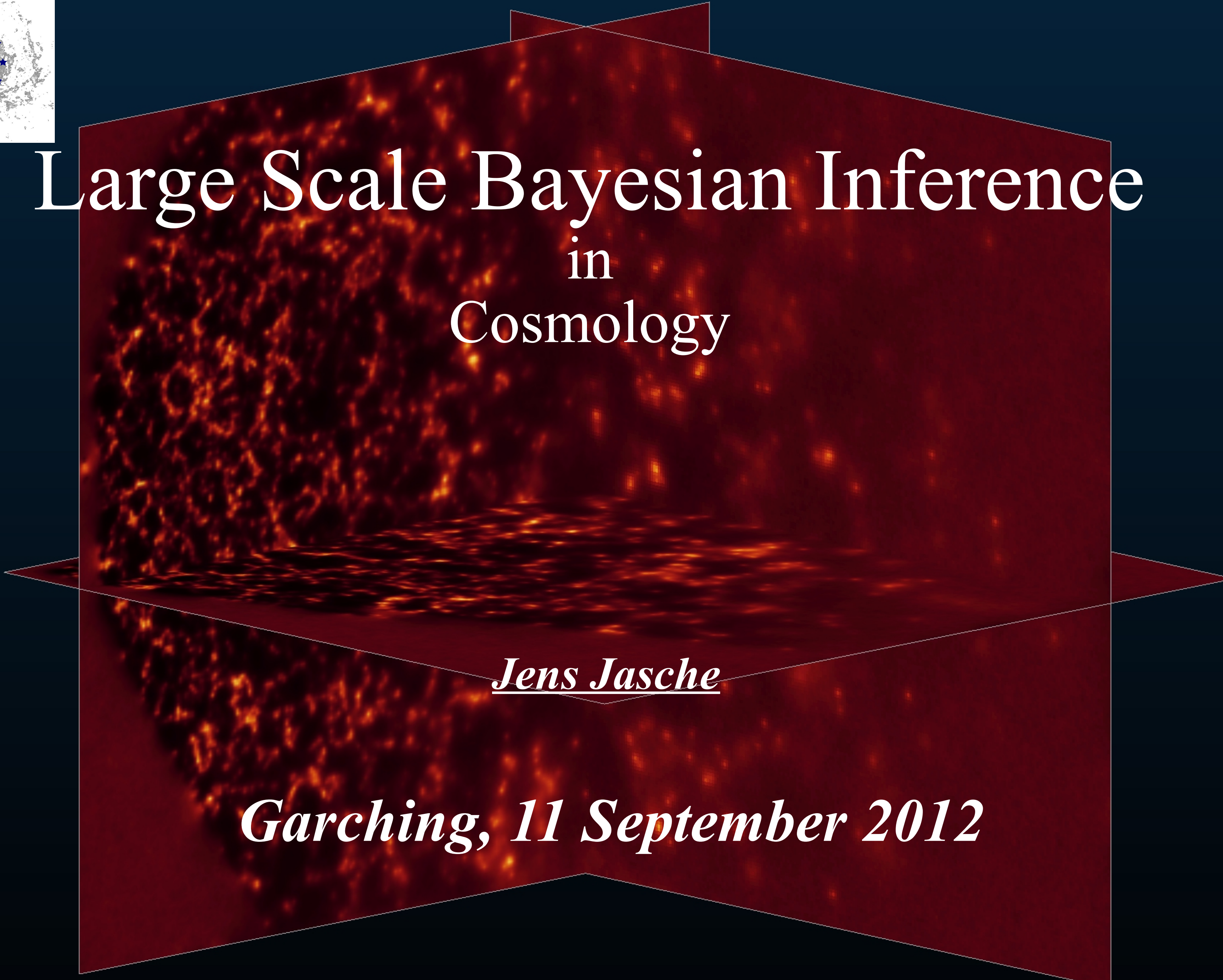
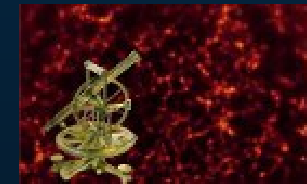
Large Scale Bayesian Inference
in
Cosmology



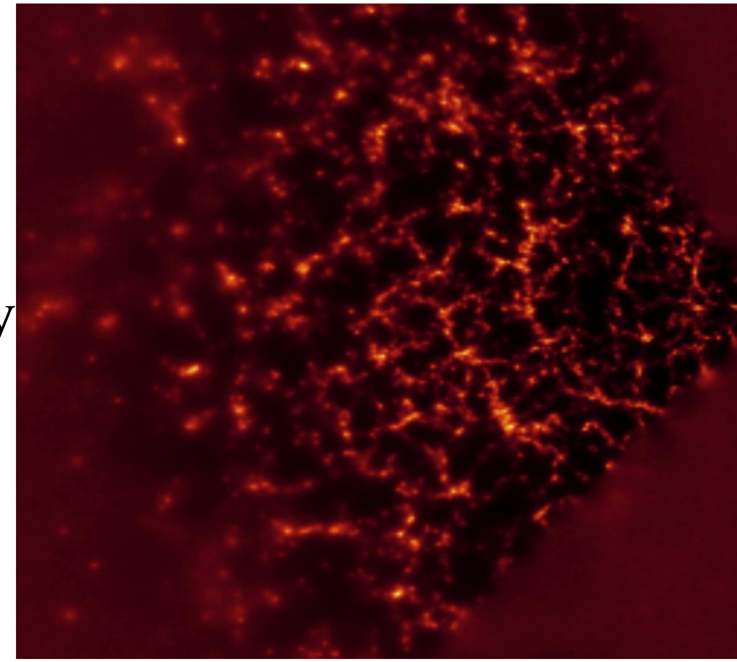
Jens Jasche

Garching, 11 September 2012

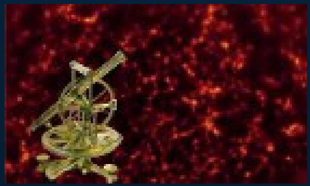
Introduction



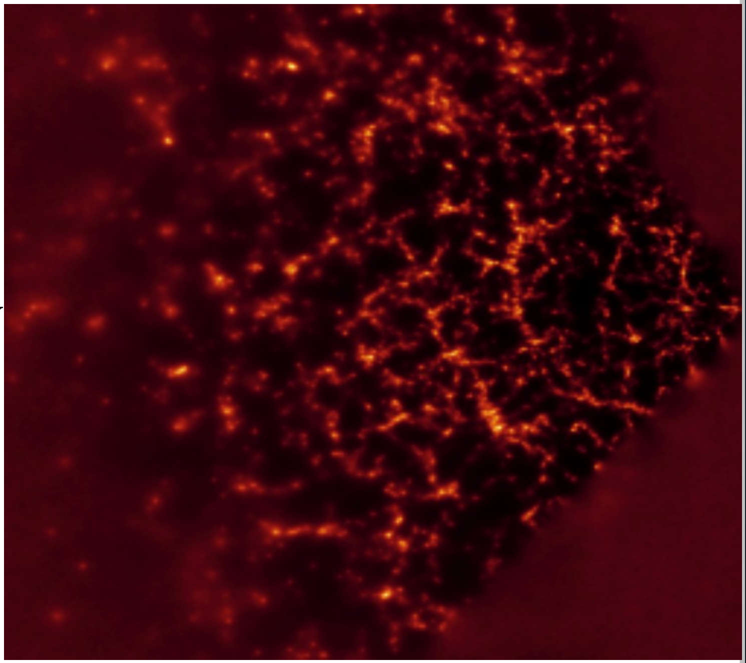
- Cosmography
  - 3D density and velocity fields
  - Power-spectra, bi-spectra
  - Dark Energy, Dark Matter, Gravity
  - Cosmological parameters



Introduction



- Cosmography
  - 3D density and velocity fields
  - Power-spectra, bi-spectra
  - Dark Energy, Dark Matter, Gravity
  - Cosmological parameters



- Large Scale Bayesian inference
  - High dimensional (~ 10^7 parameters)
  - State-of-the-art technology
  - On the verge of numerical feasibility

Introduction

- Why do we need Bayesian inference?

Introduction

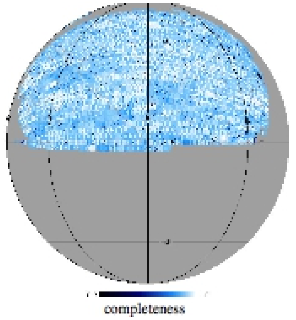
- Why do we need Bayesian inference?
  - Inference of signals = ill-posed problem

Introduction

- Why do we need Bayesian inference?
  - Inference of signals = ill-posed problem
    - Noise

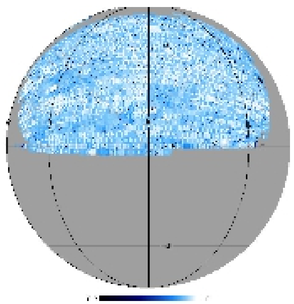
Introduction

- Why do we need Bayesian inference?
  - Inference of signals = ill-posed problem
    - Noise
    - Incomplete observations



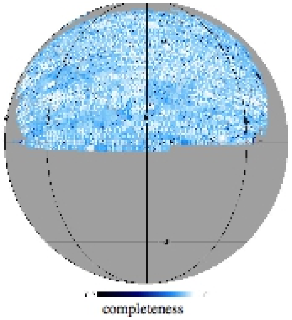
Introduction

- Why do we need Bayesian inference?
  - Inference of signals = ill-posed problem
    - Noise
    - Incomplete observations
    - Systematics



Introduction

- Why do we need Bayesian inference?
  - Inference of signals = ill-posed problem
    - Noise
    - Incomplete observations
    - Systematics



➡ No unique recovery possible!!!

Introduction

“**What are the possible signals compatible with observations?**”

Introduction

“What are the possible signals compatible with observations?”

- Object of interest: Signal posterior distribution

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

Introduction

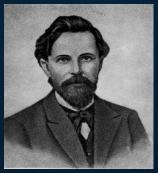
"What are the possible signals compatible with observations?"

- Object of interest: Signal posterior distribution

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

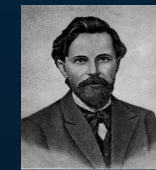
- We can do science!
  - Model comparison
  - Parameter studies
  - Report statistical summaries
  - Non-linear, Non-Gaussian error propagation

Markov Chain Monte Carlo



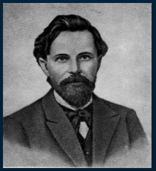
- Problems:
  - High dimensional (~10^7 parameter)

Markov Chain Monte Carlo



- Problems:
  - High dimensional (~10^7 parameter)
  - A large number of correlated parameters
    - No reduction of problem size possible

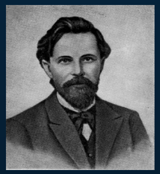
Markov Chain Monte Carlo



- Problems:
  - High dimensional (~10^7 parameter)
  - A large number of correlated parameters
    - No reduction of problem size possible
  - Complex posterior distributions

- Numerical approximation

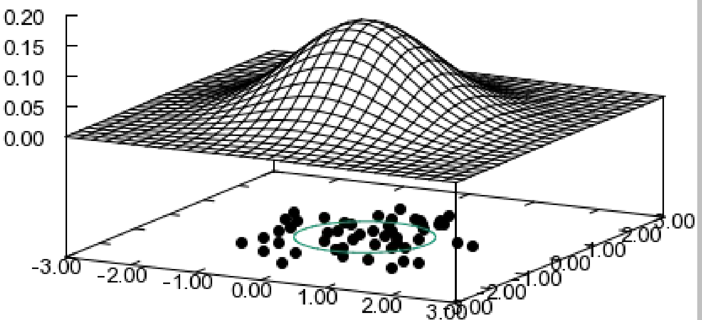
Markov Chain Monte Carlo



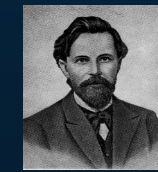
- Problems:
  - High dimensional (~10^7 parameter)
  - A large number of correlated parameters
    - <u>No reduction of problem size possible</u>
  - Complex posterior distributions

- Numerical approximation
  - Dim > 4 → MCMC

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



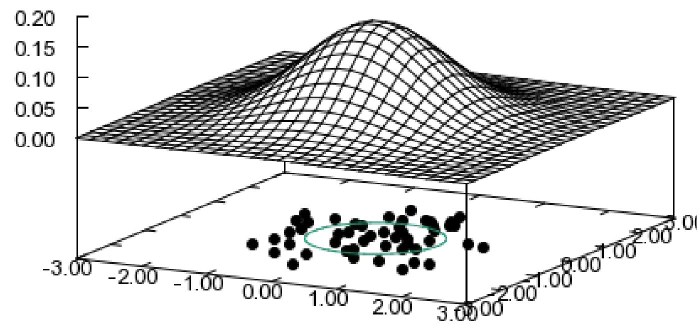
Markov Chain Monte Carlo



- Problems:
  - High dimensional (~10^7 parameter)
  - A large number of correlated parameters
    - No reduction of problem size possible
  - Complex posterior distributions

- Numerical approximation
  - Dim > 4 → MCMC

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



  - Metropolis-Hastings

Hamiltonian sampling

- Parameter space exploration via Hamiltonian sampling

Hamiltonian sampling

- Parameter space exploration via Hamiltonian sampling
  - interpret log-posterior as potential

$$\psi(x) = -ln(\mathcal{P}(x))$$

Hamiltonian sampling

- Parameter space exploration via Hamiltonian sampling
  - interpret log-posterior as potential

$$\psi(x) = -ln(\mathcal{P}(x))$$

  - introduce Gaussian auxiliary “momentum” variable

$$H = \sum_i \sum_j \frac{1}{2} p_i M_{ij}^{-1} p_j + \psi(x)$$

Hamiltonian sampling

- Parameter space exploration via Hamiltonian sampling
  - interpret log-posterior as potential

$$\psi(x) = -ln(\mathcal{P}(x))$$

  - introduce Gaussian auxiliary "momentum" variable

$$H = \sum_i \sum_j \frac{1}{2} p_i M_{ij}^{-1} p_j + \psi(x)$$

  - resultant joint posterior distribution of x and p

$$e^{-H} = \mathcal{P}(\{x_i\})\, e^{-\frac{1}{2} \sum_i \sum_j p_i M_{ij}^{-1} p_j}$$

    - separable in x and p

Hamiltonian sampling

- Parameter space exploration via Hamiltonian sampling
  - interpret log-posterior as potential

$$\psi(x) = -ln(\mathcal{P}(x))$$

  - introduce Gaussian auxiliary "momentum" variable

$$H = \sum_i \sum_j \frac{1}{2} p_i M_{ij}^{-1} p_j + \psi(x)$$

  - resultant joint posterior distribution of x and p

$$e^{-H} = \mathcal{P}(\{x_i\})\, e^{-\frac{1}{2}\sum_i \sum_j p_i M_{ij}^{-1} p_j}$$

    - separable in x and p
    - marginalization over p yields again $\mathcal{P}(x)$

Hamiltonian sampling

- IDEA: Use Hamiltonian dynamics to explore e^{-H}

Hamiltonian sampling

- IDEA: Use Hamiltonian dynamics to explore e^{-H}
  - solve Hamiltonian system to obtain new sample

$$\{x^i, p^i\} \Rightarrow \boxed{\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} &= \frac{\partial H}{\partial x_i} = -\frac{\partial \psi(x)}{\partial x_i} \end{aligned}} \Rightarrow \{x^{i+1}, p^{i+1}\}$$

Hamiltonian sampling

- IDEA: Use Hamiltonian dynamics to explore e^{-H}
  - solve Hamiltonian system to obtain new sample

$$\{x^i, p^i\} \Rightarrow \boxed{\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} &= \frac{\partial H}{\partial x_i} = -\frac{\partial \psi(x)}{\partial x_i} \end{aligned}} \Rightarrow \{x^{i+1}, p^{i+1}\}$$

- Hamiltonian dynamics conserve the Hamiltonian H

Hamiltonian sampling

- IDEA: Use Hamiltonian dynamics to explore e^{-H}
  - solve Hamiltonian system to obtain new sample

$$\{x^i, p^i\} \Rightarrow \boxed{\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} &= \frac{\partial H}{\partial x_i} = -\frac{\partial \psi(x)}{\partial x_i} \end{aligned}} \Rightarrow \{x^{i+1}, p^{i+1}\}$$

- Hamiltonian dynamics conserve the Hamiltonian H
  - Metropolis acceptance probability is unity

$$\mathcal{P}_A = min\left[1, exp(-(H(\{x_i'\},\{p_i'\}) - H(\{x_i\},\{p_i\}))\right]$$

Hamiltonian sampling

- IDEA: Use Hamiltonian dynamics to explore e^{-H}
  - solve Hamiltonian system to obtain new sample

$$\{x^i, p^i\} \Rightarrow \boxed{\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} = \frac{\partial H}{\partial x_i} &= -\frac{\partial \psi(x)}{\partial x_i} \end{aligned}} \Rightarrow \{x^{i+1}, p^{i+1}\}$$

- Hamiltonian dynamics conserve the Hamiltonian H
  - Metropolis acceptance probability is unity

$$\mathcal{P}_A = min[1, exp(-(H(\{x_i'\}, \{p_i'\}) - H(\{x_i\}, \{p_i\}))]$$

  - All samples are accepted

Hamiltonian sampling

- Example: Wiener posterior = multivariate normal distribution

$$\Psi = \frac{1}{2}\sum_{ij} x_i S^{-1}_{ij} x_j + \frac{1}{2}\sum_{ij} (x_i - d_i) N^{-1}_{ij} (x_j - d_j)$$

Hamiltonian sampling

- Example: Wiener posterior = multivariate normal distribution

$$\Psi = \boxed{\frac{1}{2}\sum_{ij} x_i S_{ij}^{-1} x_j} + \frac{1}{2}\sum_{ij}(x_i - d_i)N_{ij}^{-1}(x_j - d_j)$$

Prior

Hamiltonian sampling

- Example: Wiener posterior = multivariate normal distribution

$$\Psi = \frac{1}{2}\sum_{ij} x_i S^{-1}_{ij} x_j + \frac{1}{2}\sum_{ij}(x_i - d_i) N^{-1}_{ij} (x_j - d_j)$$

Likelihood

Hamiltonian sampling

- Example: Wiener posterior = multivariate normal distribution

$$\Psi = \frac{1}{2}\sum_{ij} x_i S^{-1}_{ij} x_j + \frac{1}{2}\sum_{ij}(x_i - d_i) N^{-1}_{ij}(x_j - d_j)$$

$$\begin{aligned}\frac{\partial \Psi}{\partial x_m} &= \sum_j \left[S^{-1}_{mj} + N^{-1}_{mj}\right] x_j - \sum_{ij} N^{-1}_{mj} d_j \\ &= \sum_j A_{mj} x_j - B_m\end{aligned}$$

Hamiltonian sampling

- Example: Wiener posterior = multivariate normal distribution

$$\Psi = \frac{1}{2}\sum_{ij} x_i S^{-1}_{ij} x_j + \frac{1}{2}\sum_{ij}(x_i - d_i)N^{-1}_{ij}(x_j - d_j)$$

$$\frac{\partial \Psi}{\partial x_m} = \sum_j \left[S^{-1}_{mj} + N^{-1}_{mj}\right] x_j - \sum_{ij} N^{-1}_{mj} d_j$$

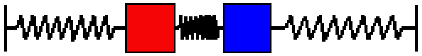
$$= \sum_j A_{mj} x_j - B_m$$

$$\frac{\mathrm{dx_m}}{\mathrm{dt}} = \sum_j M^{-1}_{mj} p_j$$

$$\frac{\mathrm{dp_m}}{\mathrm{dt}} = -\sum_j A_{mj} x_j + B_m$$

EOM:

coupled harmonic oscillator



Hamiltonian sampling

- How to set the Mass matrix?

Hamiltonian sampling

- How to set the Mass matrix?
  - Large number of tunable parameter

Hamiltonian sampling

- How to set the Mass matrix?
  - Large number of tunable parameter
  - Determines efficiency of sampler

Hamiltonian sampling

- How to set the Mass matrix?
  - Large number of tunable parameter
  - Determines efficiency of sampler

$$
\begin{aligned}
\frac{d^2 x_i}{dt^2} &= -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + \sum_l M_{il}^{-1} B_l \\
&= -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + D_m
\end{aligned}
$$

Hamiltonian sampling

- How to set the Mass matrix?
  - Large number of tunable parameter
  - Determines efficiency of sampler

$$\frac{d^2 x_i}{dt^2} = -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + \sum_l M_{il}^{-1} B_l$$
$$= -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + D_m$$

$$\boxed{M_{ij} = A_{ij}}$$

  - Mass matrix aims at decoupling the system

Hamiltonian sampling

- How to set the Mass matrix?
  - Large number of tunable parameter
  - Determines efficiency of sampler

$$\frac{d^2 x_i}{dt^2} = -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + \sum_l M_{il}^{-1} B_l$$

$$= -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + D_m$$

$$\boxed{M_{ij} = A_{ij}}$$

  - Mass matrix aims at decoupling the system
  - In practice: use diagonal approximation

Hamiltonian sampling

- How to set the Mass matrix?
  - Large number of tunable parameter
  - Determines efficiency of sampler

$$\frac{d^2 x_i}{dt^2} = -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + \sum_l M_{il}^{-1} B_l$$
$$= -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + D_m$$

$$\boxed{M_{ij} = A_{ij}}$$

  - Mass matrix aims at decoupling the system
  - In practice: use diagonal approximation
  - The quality of approximation determines sampler efficiency

Hamiltonian sampling

- How to set the Mass matrix?
  - Large number of tunable parameter
  - Determines efficiency of sampler

$$\frac{d^2 x_i}{dt^2} = -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + \sum_l M_{il}^{-1} B_l$$

$$= -\sum_l M_{il}^{-1} \sum_j A_{lj} x_j + D_m$$

$$\boxed{M_{ij} = A_{ij}}$$

  - Mass matrix aims at decoupling the system
  - In practice: use diagonal approximation
  - The quality of approximation determines sampler efficiency
  - Non-Gaussian case: Taylor expand to find Mass matrix

HMC in action

- Inference of non-linear density fields in cosmology
  - Non-linear density field
    - Log-normal prior

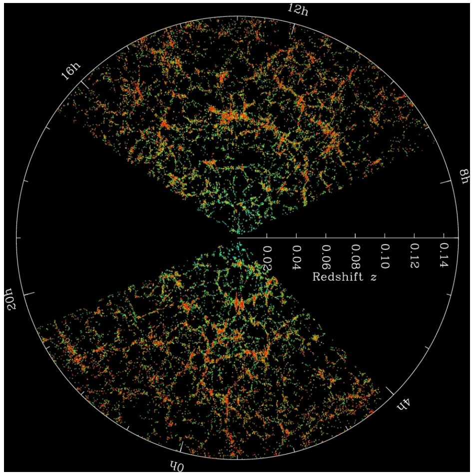
      See e.g. Coles & Jones (1991), Kayo et al. (2001)

Jasche, Kitaura (2010)

HMC in action

- Inference of non-linear density fields in cosmology
  - Non-linear density field
    - Log-normal prior

      See e.g. Coles & Jones (1991), Kayo et al. (2001)
  - Galaxy distribution
    - Poisson likelihood
    - Signal dependent noise



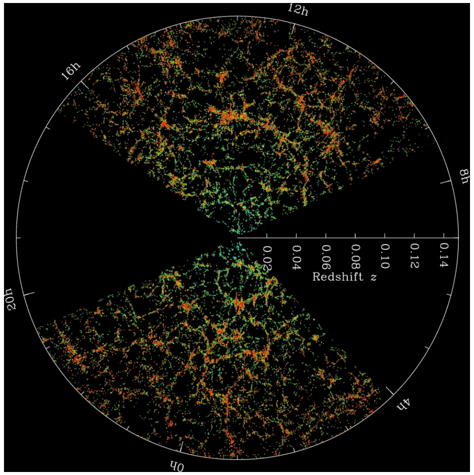
Credit: M. Blanton and the Sloan Digital Sky Survey

Jasche, Kitaura (2010)

HMC in action

- Inference of non-linear density fields in cosmology
  - Non-linear density field
    - Log-normal prior

      See e.g. Coles & Jones (1991), Kayo et al. (2001)
  - Galaxy distribution
    - Poisson likelihood
    - Signal dependent noise



Credit: M. Blanton and the Sloan Digital Sky Survey

Problem: Non-Gaussian sampling in high dimensions

➡ **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

Multiple Block Sampling

- What if the HMC is not an option?

$$A, B \curvearrowleft \mathcal{P}(A, B)$$

Multiple Block Sampling

- What if the HMC is not an option?

$$A, B \curvearrowleft \mathcal{P}(A, B)$$

  - Problem: Design of “good” proposal distributions

Multiple Block Sampling

- What if the HMC is not an option?

$$A, B \curvearrowleft \mathcal{P}(A, B)$$

  - Problem: Design of “good” proposal distributions
  - High rejection rates

Multiple Block Sampling

- What if the HMC is not an option?

$$A, B \curvearrowleft \mathcal{P}(A, B)$$

  - Problem: Design of "good" proposal distributions
  - High rejection rates

- Multiple block sampling (see e.g. Hastings (1997))

Multiple Block Sampling

- What if the HMC is not an option?

$$A, B \curvearrowleft \mathcal{P}(A, B)$$

  - Problem: Design of "good" proposal distributions
  - High rejection rates

- Multiple block sampling (see e.g. Hastings (1997))
  - Break down into subproblems

$$A^{i+1} \curvearrowleft \mathcal{P}(A|B^{i})$$
$$B^{i+1} \curvearrowleft \mathcal{P}(B|A^{i+1})$$

Multiple Block Sampling

- What if the HMC is not an option?

$$A, B \curvearrowleft \mathcal{P}(A, B)$$

  - Problem: Design of "good" proposal distributions
  - High rejection rates

- Multiple block sampling (see e.g. Hastings (1997))
  - Break down into subproblems

$$A^{i+1} \curvearrowleft \mathcal{P}(A|B^{i})$$
$$B^{i+1} \curvearrowleft \mathcal{P}(B|A^{i+1})$$

  - simplifies design of conditional proposal distributions

Multiple Block Sampling

- What if the HMC is not an option?

$$A, B \curvearrowleft \mathcal{P}(A, B)$$

  - Problem: Design of “good” proposal distributions
  - High rejection rates

- Multiple block sampling (see e.g. Hastings (1997))
  - Break down into subproblems

$$A^{i+1} \curvearrowleft \mathcal{P}(A|B^{i})$$
$$B^{i+1} \curvearrowleft \mathcal{P}(B|A^{i+1})$$

  - simplifies design of conditional proposal distributions
  - Average acceptance rate is higher

Multiple Block Sampling

- What if the HMC is not an option?

$$A, B \curvearrowleft \mathcal{P}(A, B)$$

  - Problem: Design of “good” proposal distributions
  - High rejection rates

- Multiple block sampling (see e.g. Hastings (1997))
  - Break down into subproblems

$$A^{i+1} \curvearrowleft \mathcal{P}(A|B^{i})$$
$$B^{i+1} \curvearrowleft \mathcal{P}(B|A^{i+1})$$

Serial processing only!

  - simplifies design of conditional proposal distributions
  - Average acceptance rate is higher
  - Requires serial processing

Multiple Block Sampling

- Can we “boost” block sampling?

$$\mathcal{P}(A, B) = \int \mathrm{d}C \, \mathcal{P}(A, B, C)$$

Multiple Block Sampling

- Can we "boost" block sampling?

$$\mathcal{P}(A, B) = \int \mathrm{d}C\, \mathcal{P}(A, B, C)$$

  - Sometimes it is easier to explore full joint the PDF

$$A, B, C \curvearrowleft \mathcal{P}(A, B, C)$$

Multiple Block Sampling

- Can we "boost" block sampling?

$$\mathcal{P}(A, B) = \int \mathrm{d}C\, \mathcal{P}(A, B, C)$$

  - Sometimes it is easier to explore full joint the PDF

$$A, B, C \curvearrowleft \mathcal{P}(A, B, C)$$

  - Block sampler:

$$\begin{aligned} A^{i+1} &\curvearrowleft \mathcal{P}(A|B^{i}, C^{i}) \\ B^{i+1} &\curvearrowleft \mathcal{P}(B|A^{i+1}, C^{i}) \\ C^{i+1} &\curvearrowleft \mathcal{P}(C|A^{i+1}, B^{i+1}) \end{aligned}$$

Multiple Block Sampling

- Can we "boost" block sampling?

$$\mathcal{P}(A, B) = \int \mathrm{d}C\, \mathcal{P}(A, B, C)$$

  - Sometimes it is easier to explore full joint the PDF

$$A, B, C \curvearrowleft \mathcal{P}(A, B, C)$$

  - Block sampler:

$$\begin{aligned} A^{i+1} &\curvearrowleft \mathcal{P}(A|B^i, C^i) = \mathcal{P}(A|C^i) \\ B^{i+1} &\curvearrowleft \mathcal{P}(B|A^{i+1}, C^i) = \mathcal{P}(B|C^i) \\ C^{i+1} &\curvearrowleft \mathcal{P}(C|A^{i+1}, B^{i+1}) \end{aligned}$$

Multiple Block Sampling

- Can we "boost" block sampling?

$$\mathcal{P}(A, B) = \int \mathrm{d}C\,\mathcal{P}(A, B, C)$$

  - Sometimes it is easier to explore full joint the PDF

$$A, B, C \curvearrowleft \mathcal{P}(A, B, C)$$

  - Block sampler:

$$\begin{aligned} A^{i+1} &\curvearrowleft \mathcal{P}(A|B^i, C^i) = \boxed{\mathcal{P}(A|C^i)} \\ B^{i+1} &\curvearrowleft \mathcal{P}(B|A^{i+1}, C^i) = \boxed{\mathcal{P}(B|C^i)} \\ C^{i+1} &\curvearrowleft \mathcal{P}(C|A^{i+1}, B^{i+1}) \end{aligned}$$

process in parallel!

  - Permits efficient sampling for numerical expensive posteriors

Photometric redshift sampling

- Photometric surveys
  - millions of galaxies (~10^7 - 10^8)

Photometric redshift sampling

- Photometric surveys
  - millions of galaxies (~10^7 - 10^8)
  - low redshift accuracy (~ 100 Mpc along LOS)

Photometric redshift sampling

- Photometric surveys
  - millions of galaxies (~10^7 - 10^8)
  - low redshift accuracy (~ 100 Mpc along LOS)
  - → Infer accurate redshifts: $\{z_p\} \curvearrowleft \mathcal{P}(\{z_p\}|d)$

Photometric redshift sampling

- Photometric surveys
  - millions of galaxies (~10^7 - 10^8)
  - low redshift accuracy (~ 100 Mpc along LOS)
  - → Infer accurate redshifts: $\{z_p\} \curvearrowleft \mathcal{P}(\{z_p\}|d)$
  - Rather sample from joint distribution:

$$\{z_p\}, \delta \curvearrowleft \mathcal{P}(\delta, \{z_p\}|d)$$

Photometric redshift sampling

- Photometric surveys
  - millions of galaxies (~10^7 - 10^8)
  - low redshift accuracy (~ 100 Mpc along LOS)
  - → Infer accurate redshifts: $\{z_p\} \curvearrowleft \mathcal{P}(\{z_p\}|d)$
  - Rather sample from joint distribution:

$$\{z_p\}, \delta \curvearrowleft \mathcal{P}(\delta, \{z_p\}|d)$$

  - Block sampler:

$$\begin{aligned}
z_1^{i+1} &\curvearrowleft \mathcal{P}(z_1|\delta^i, d) \\
&\vdots \\
z_N^{i+1} &\curvearrowleft \mathcal{P}(z_N|\delta^i, d) \\
\delta^{i+1} &\curvearrowleft \mathcal{P}(\delta|\{z_p\}^{i+1}, d)
\end{aligned}$$

Photometric redshift sampling

- Photometric surveys
  - millions of galaxies (~10^7 - 10^8)
  - low redshift accuracy (~ 100 Mpc along LOS)
  - → Infer accurate redshifts: $\{z_p\} \curvearrowleft \mathcal{P}(\{z_p\}|d)$
  - Rather sample from joint distribution:

$$\{z_p\}, \delta \curvearrowleft \mathcal{P}(\delta, \{z_p\}|d)$$

  - Block sampler:

$$z_1^{i+1} \curvearrowleft \mathcal{P}(z_1|\delta^i, d)$$

$$\vdots$$

$$z_N^{i+1} \curvearrowleft \mathcal{P}(z_N|\delta^i, d)$$

Process in parallel!

$$\delta^{i+1} \curvearrowleft \mathcal{P}(\delta|\{z_p\}^{i+1}, d)$$

Photometric redshift sampling

- Photometric surveys
  - millions of galaxies (~10^7 - 10^8)
  - low redshift accuracy (~ 100 Mpc along LOS)
  - → Infer accurate redshifts: $\{z_p\} \curvearrowleft \mathcal{P}(\{z_p\}|d)$
  - Rather sample from joint distribution:

$$\{z_p\}, \delta \curvearrowleft \mathcal{P}(\delta, \{z_p\}|d)$$

  - Block sampler:

$$z_1^{i+1} \curvearrowleft \mathcal{P}(z_1|\delta^i, d)$$

$$\vdots$$

$$z_N^{i+1} \curvearrowleft \mathcal{P}(z_N|\delta^i, d)$$

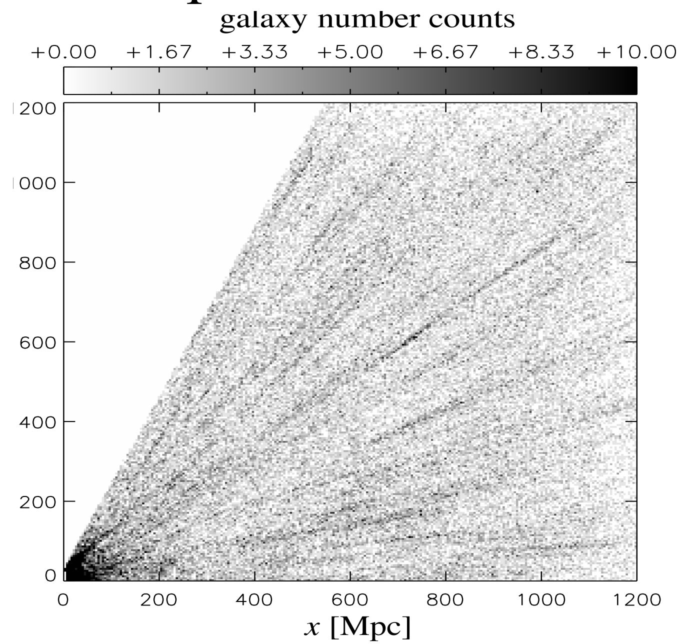
Process in parallel!

$$\delta^{i+1} \curvearrowleft \mathcal{P}(\delta|\{z_p\}^{i+1}, d)$$

HMC sampler!

Photometric redshift sampling

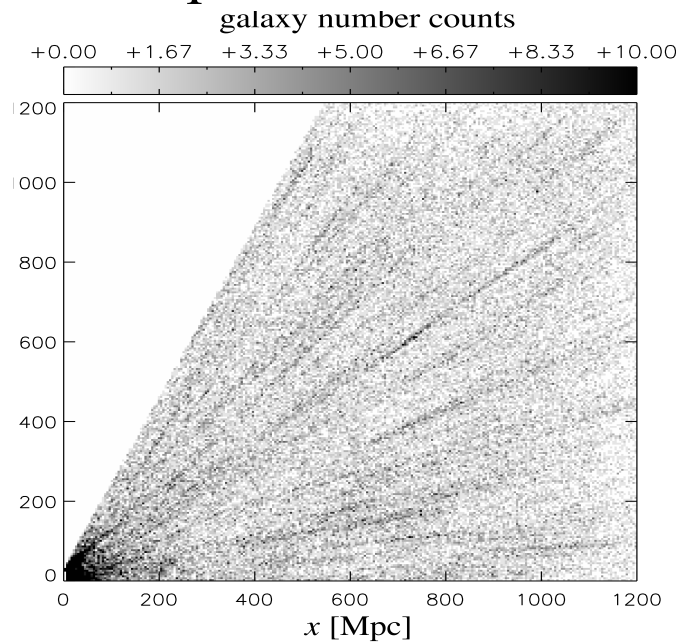
- Application to artificial photometric data



- ~ Noise, Systematics, Position uncertainty (~100 Mpc)
- ~ 10^7 density amplitudes /parameters
- ~ 2x10^7 radial galaxy positions / parameters

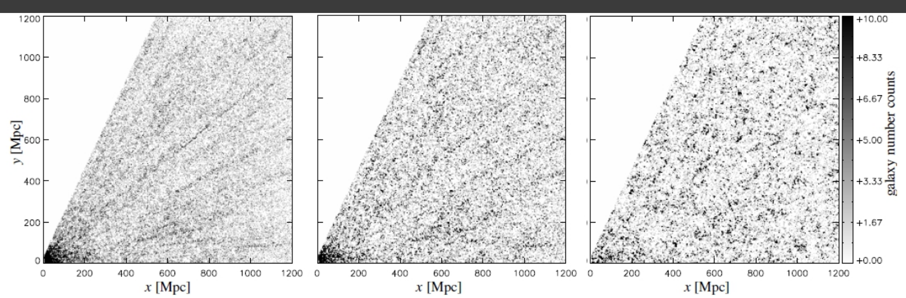
Photometric redshift sampling

- Application to artificial photometric data



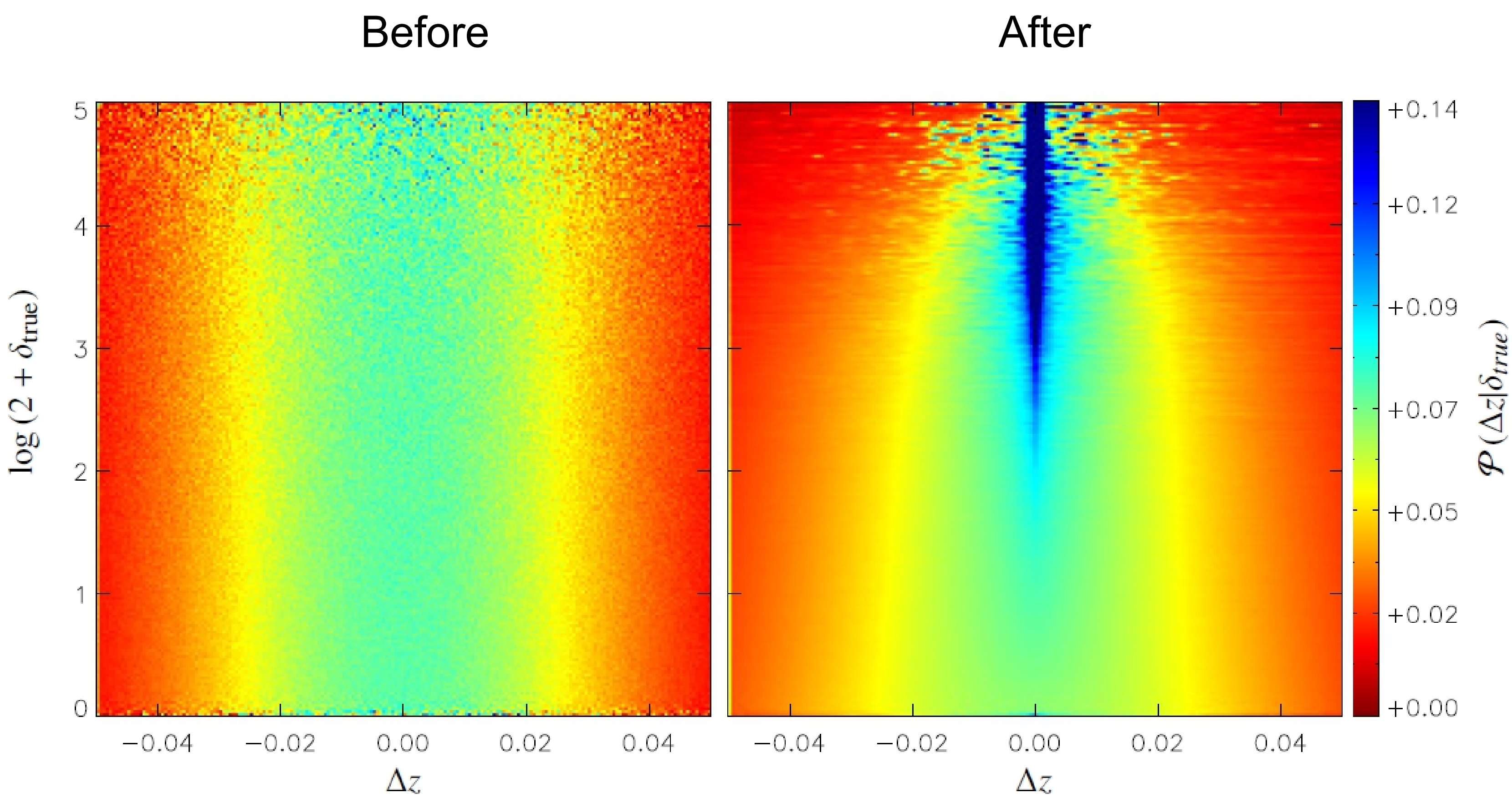
- ~ Noise, Systematics, Position uncertainty (~100 Mpc)
- ~ 10^7 density amplitudes /parameters
- ~ 2x10^7 radial galaxy positions / parameters
- **~ 3x10^7 parameters in total**

Photometric redshift sampling



Jasche, Wandelt (2012)

Deviation from the truth

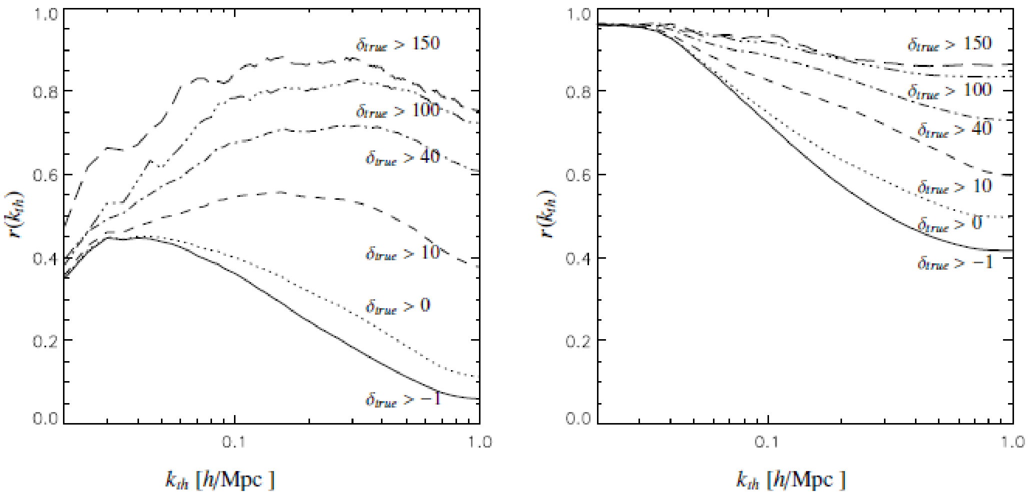


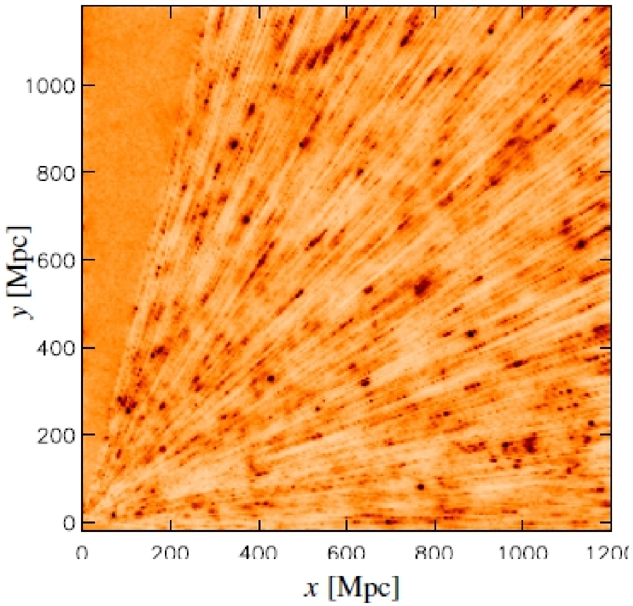
Jasche, Wandelt (2012)

Deviation from the truth

$$r(k_{th}) = \frac{\left\langle \delta_{true}^{k_{th}} \, \langle\delta\rangle^{k_{th}} \right\rangle}{\sqrt{\left\langle \left(\delta_{true}^{k_{th}}\right)^2 \right\rangle} \sqrt{\left\langle \left(\langle\delta\rangle^{k_{th}}\right)^2 \right\rangle}}$$

Raw data　　Density estimate





Jasche, Wandelt (2012)

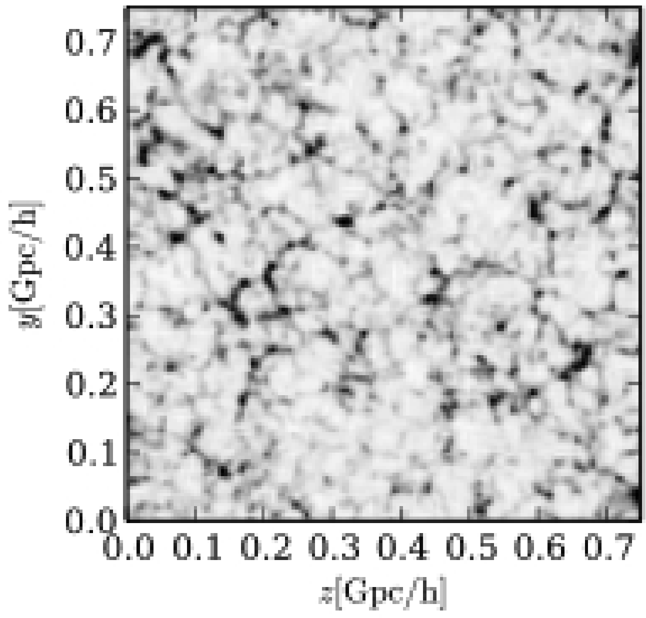
4D physical inference

- Physical motivation

4D physical inference

- Physical motivation
  - Complex final state

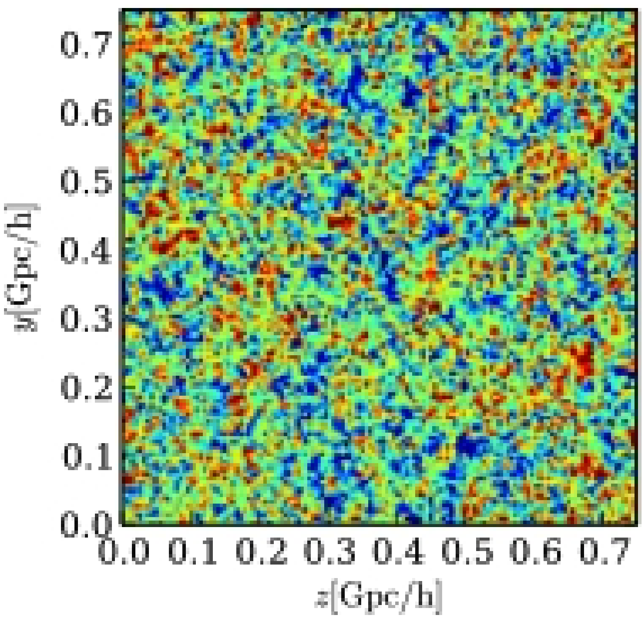
Final state



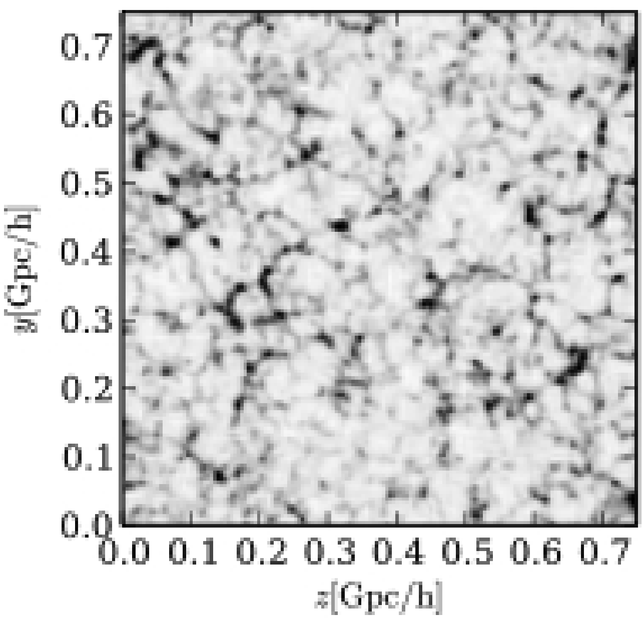
4D physical inference

- Physical motivation
  - Complex final state
  - Simple initial state

Initial state



Final state

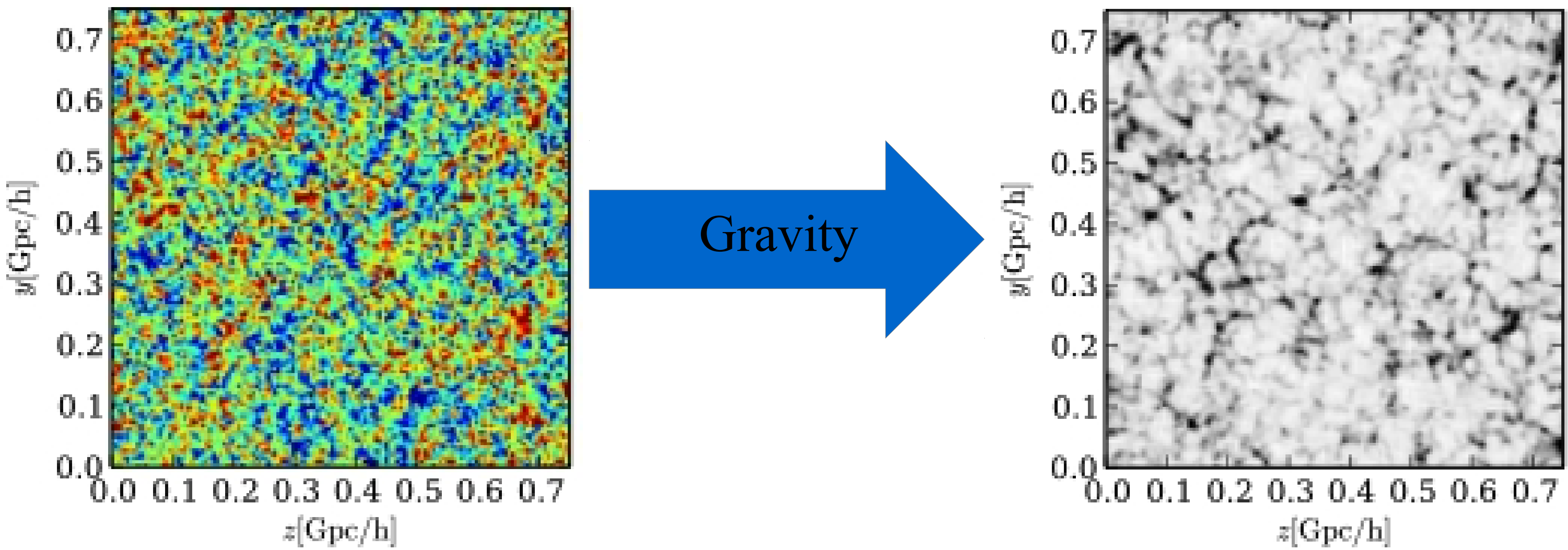


4D physical inference

- Physical motivation
  - Complex final state
  - Simple initial state

Initial state

Final state

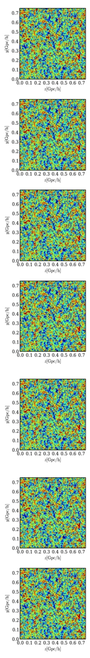


4D physical inference

- The ideal scenario:
  - We need a **very very very** large computer!

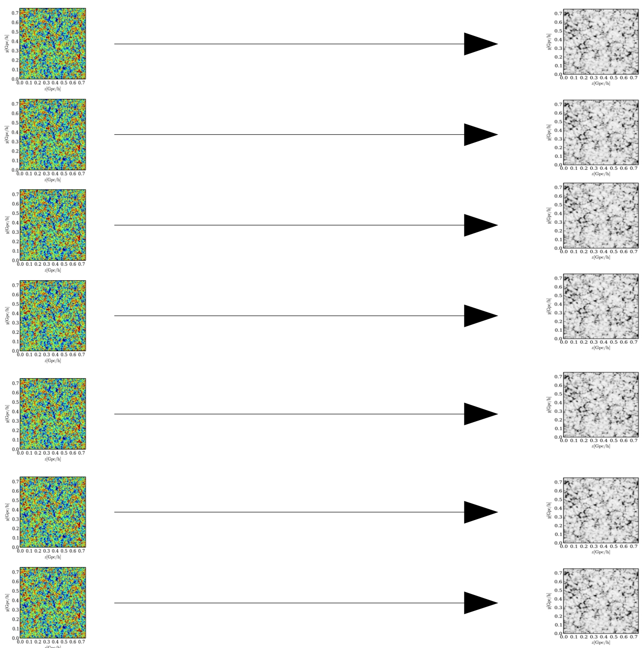
4D physical inference

- The ideal scenario:
  - We need a **very very very** large computer!



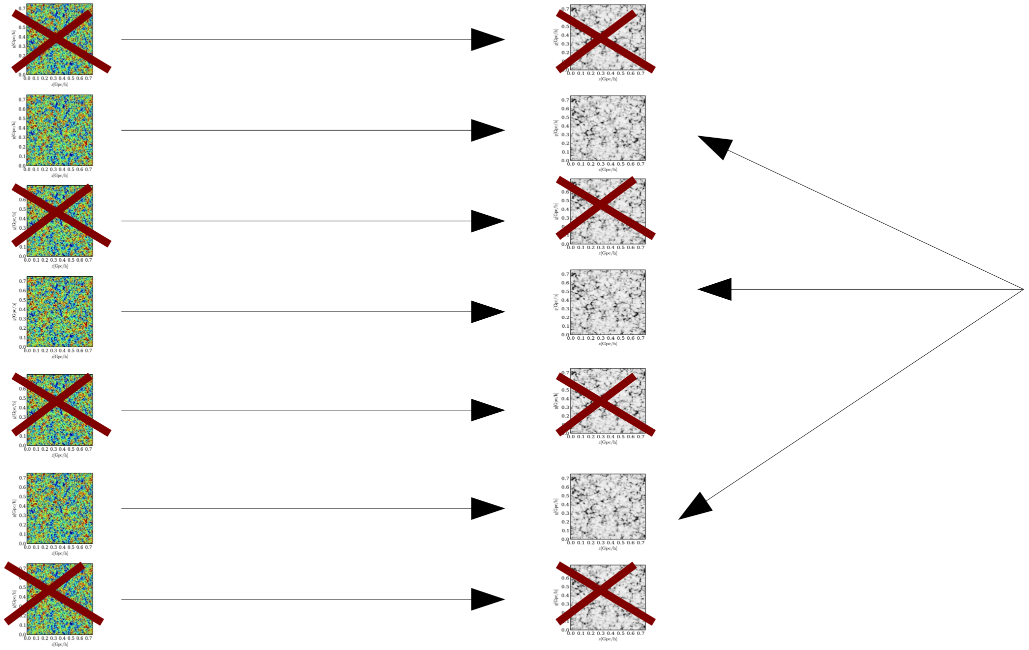
4D physical inference

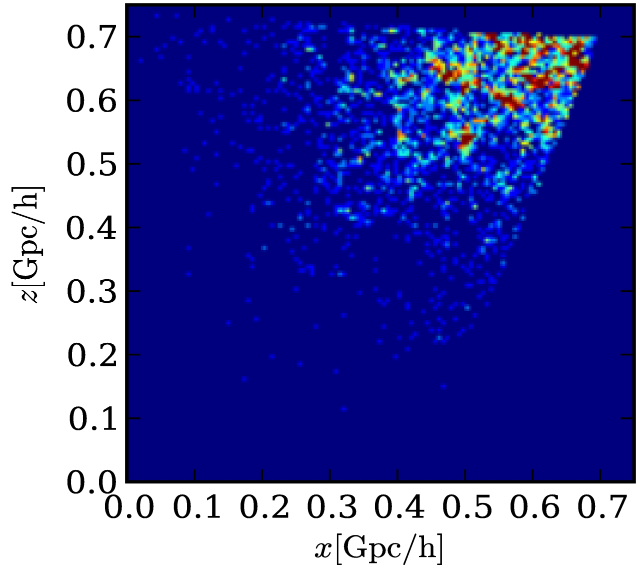
- The ideal scenario:
  - We need a **very** **very** **very** large computer!



4D physical inference

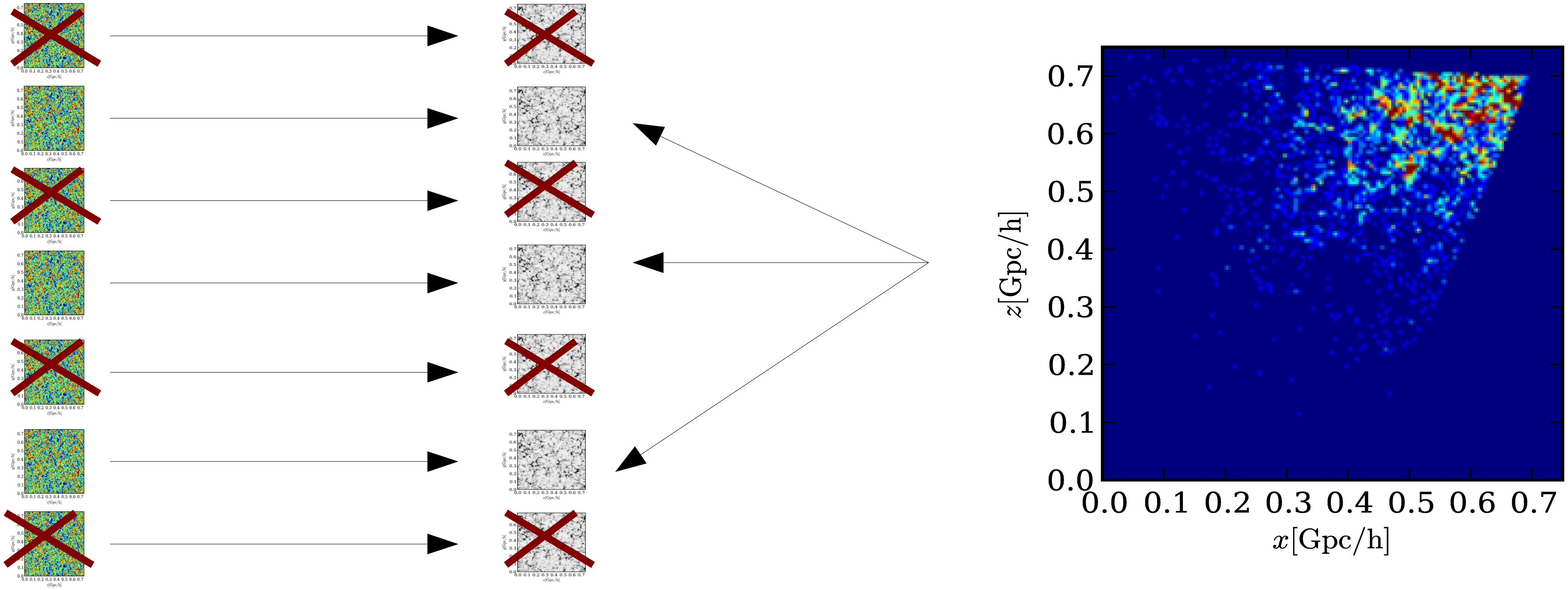
- The ideal scenario:
  - We need a **very very very** large computer!





4D physical inference

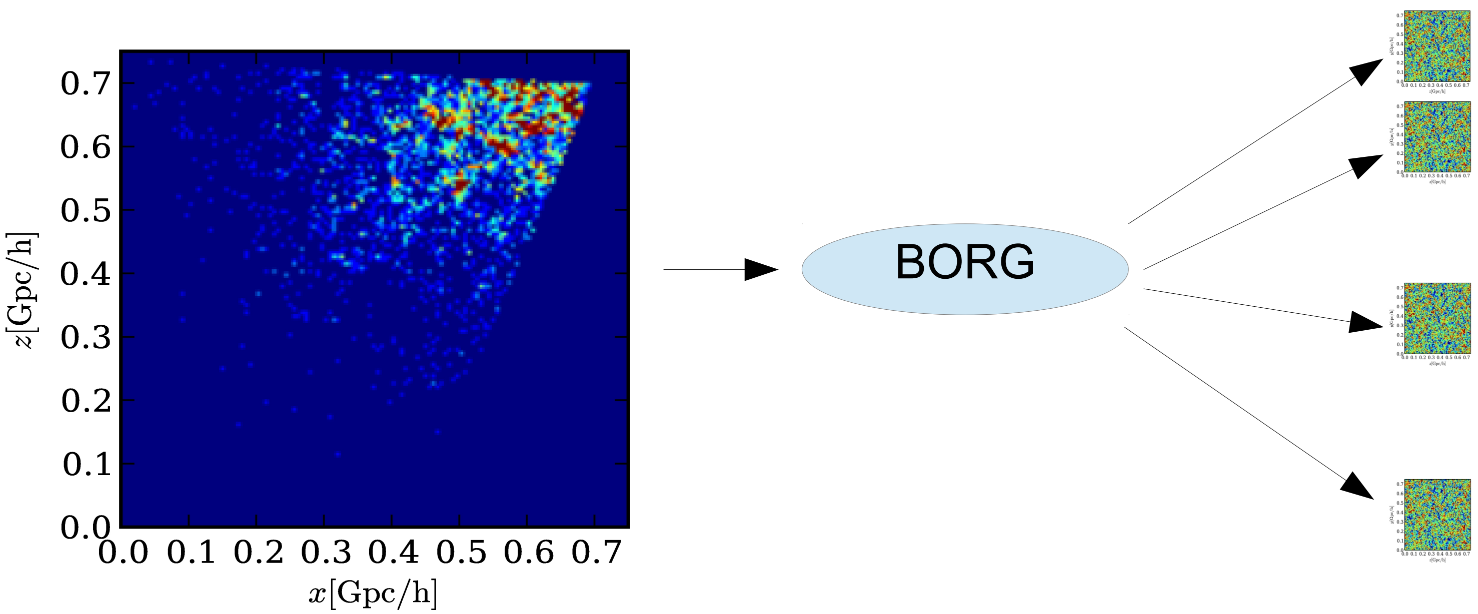
- The ideal scenario:
  - We need a **very very very** large computer!



Not practical! Even with approximations!!!!

4D physical inference

- BORG (Bayesian Origin Reconstruction from Galaxies)
  - HMC
  - Second order Lagrangian perturbation theory



Jasche, Wandelt (2012)

4D physical inference

4D physical inference

- Cosmological applications:
  - Higher order statistics → primordial non-Gaussianity
  - 4D dynamic states → Dark Energy, ISW, kSZ
  - Physically joint analysis of data at different cosmic Epochs

Summary & Conclusion

- Large scale Bayesian inference
  - Inference in high dimensions from incomplete observations
    - Noise, systematic effects, survey geometry, selection effects, biases
  - Need to quantify uncertainties → **explore posterior distribution**
    - Markov Chain Monte Carlo methods
    - Hamiltonian sampling (exploit symmetries, decouple system)
    - Multiple block sampling (break down into subproblems)
- 3 high dimensional examples (>10^7 parameter)
  - Nonlinear density inference
  - Photometric redshift and density inference
  - 4D physical inference

The End . . .

Thank you