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
Why do we need Bayesian inference?

- Inference of signals = ill-posed problem
  - Noise
  - Incomplete observations
  - Systematics
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?”
“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
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
    - **No reduction of problem size possible**
  - Complex posterior distributions

- Numerical approximation
  - Dim > 4  MCMC

\[
\mathcal{P}(s|d) \rightarrow \mathcal{P}_N(s|d) = \frac{1}{N} \sum_{i=1}^{N} \delta^D(s - s_i)
\]
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 \implies \text{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
Parameter space exploration via Hamiltonian sampling

- interpret log-posterior as potential

\[ \psi(x) = -\ln(P(x)) \]
Parameter space exploration via Hamiltonian sampling

• interpret log-posterior as potential

\[ \psi(x) = -\ln(P(x)) \]

• 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
    \[ \psi(x) = -ln(P(x)) \]
  - introduce Gaussian auxiliary “momentum” variable
    \[ H = \sum_i \sum_j \frac{1}{2} p_i M^{-1}_{ij} p_j + \psi(x) \]
  - resultant joint posterior distribution of \( \chi \) and \( p \)
    \[ e^{-H} = P(|\chi_i|) e^{-\frac{1}{2} \sum_i \sum_j p_i M^{-1}_{ij} p_j} \]
  - separable in \( \chi \) and \( P \)
Hamiltonian sampling

- Parameter space exploration via Hamiltonian sampling
  - interpret log-posterior as potential
    \[ \psi(x) = -\ln(P(x)) \]
  - 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 \( p \)
    \[ e^{-H} = P(\{x_i\}) e^{-\frac{1}{2} \sum_i \sum_j p_i M_{ij}^{-1} p_j} \]
    - separable in \( x \) and \( p \)
    - marginalization over \( p \) yields again \( P(x) \)
IDEA: Use Hamiltonian dynamics to explore $e^{-H}$
IDEA: Use Hamiltonian dynamics to explore $e^{-H}$

- solve Hamiltonian system to obtain new sample

\[ \{x^i, p^i\} \rightarrow \left( \frac{dx_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = \frac{\partial H}{\partial x_i} = -\frac{\partial \psi(x)}{\partial x_i} \right) \rightarrow \{x^{i+1}, p^{i+1}\} \]
IDEA: Use Hamiltonian dynamics to explore $e^{-H}$

- solve Hamiltonian system to obtain new sample

$$\begin{align*}
\{x^i, p^i\} & \rightarrow \frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \\
\frac{dp_i}{dt} & = \frac{\partial H}{\partial x_i} = -\frac{\partial \psi(x)}{\partial x_i} \\
\end{align*}$$

$$\rightarrow \{x^{i+1}, p^{i+1}\}$$

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

- solve Hamiltonian system to obtain new sample

\[
\left\{ x^i, p^i \right\} \longrightarrow \begin{cases}
    \frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \\
    \frac{dp_i}{dt} = \frac{\partial H}{\partial x_i} = -\frac{\partial \psi(x)}{\partial x_i}
\end{cases} \longrightarrow \left\{ x^{i+1}, p^{i+1} \right\}
\]

Hamiltonian dynamics conserve the Hamiltonian $H$

- Metropolis acceptance probability is unity

\[
P_A = \min \left[ 1, \exp \left( - (H(\{x'_i\}, \{p'_i\}) - H(\{x_i\}, \{p_i\})) \right) \right]
\]
IDEA: Use Hamiltonian dynamics to explore $e^{-H}$

- solve Hamiltonian system to obtain new sample

\[
\begin{align*}
\{x^i, p^i\} & \rightarrow \left\{ \frac{dx_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = \frac{\partial H}{\partial x_i} = -\frac{\partial \psi(x)}{\partial x_i} \right\} \\
& \rightarrow \{x^{i+1}, p^{i+1}\}
\end{align*}
\]

Hamiltonian dynamics conserve the Hamiltonian $H$

- Metropolis acceptance probability is unity

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

- All samples are accepted
Hamiltonian sampling

Example: Wiener posterior = multivariate normal distribution

\[ \Psi = \frac{1}{2} \sum_{ij} x_i S_{ij}^{-1} x_j + \frac{1}{2} \sum_{ij} (x_i - d_i) N_{ij}^{-1} (x_j - d_j) \]
Example: Wiener posterior = multivariate normal distribution

\[ \Psi = \frac{1}{2} \sum_{ij} x_i S_{ij}^{-1} x_j + \frac{1}{2} \sum_{ij} (x_i - d_i) N_{ij}^{-1} (x_j - d_j) \]
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
Example: Wiener posterior = multivariate normal distribution

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

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

\[ = \sum_j A_{mj} x_j - B_m \]
Hamiltonian sampling

Example: Wiener posterior = multivariate normal distribution

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

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

\[ = \sum_j A_{mj} x_j - B_m \]

EOM:

coupled harmonic oscillator

\[ \frac{dx_m}{dt} = \sum_j M^{-1}_{mj} p_j \]

\[ \frac{dp_m}{dt} = -\sum_j A_{mj} x_j + B_m \]
Hamiltonian sampling

- How to set the Mass matrix?
How to set the Mass matrix?

- Large number of tunable parameters

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.
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

\[
\frac{d^2 x_i}{dt^2} = - \sum_l M_{il}^{-1} \sum_j A_{lj} x_j + \sum_l M_{il}^{-1} B_l \\
= - \sum_l M_{il}^{-1} \sum_j A_{lj} x_j + D_m
\]
How to set the Mass matrix?

- Large number of tunable parameter
- Determines efficiency of sampler

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

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

\[ M_{ij} = A_{ij} \]

- Mass matrix aims at decoupling the system
Hamiltonian sampling

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

\[
\begin{align*}
\frac{d^2 x_i}{dt^2} &= - \sum_l M_{il}^{-1} \sum_j A_{ij} x_j + \sum_l M_{il}^{-1} B_l \\
&= - \sum_l M_{il}^{-1} \sum_j A_{ij} x_j + D_m \\
M_{ij} &= A_{ij}
\end{align*}
\]

- 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

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

\[ M_{ij} = A_{ij} \]

• 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

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

\[M_{ij} = A_{ij}\]

- 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

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?

\[ A, B \sim \mathcal{P}(A, B) \]
Multiple Block Sampling

- What if the HMC is not an option?

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

  - Problem: Design of “good” proposal distributions
Multiple Block Sampling

- What if the HMC is not an option?
  
  $A, B \sim \mathcal{P}(A, B)$
  
  - Problem: Design of “good” proposal distributions
  - High rejection rates
Multiple Block Sampling

What if the HMC is not an option?

\[ A, B \sim \mathcal{P}(A, B) \]

• 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?
  \[ A, B \sim \mathcal{P}(A, B) \]
  - Problem: Design of “good” proposal distributions
  - High rejection rates

- Multiple block sampling (see e.g. Hastings (1997))
  - Break down into subproblems
  \[ A_{i+1} \sim \mathcal{P}(A|B^i) \]
  \[ B_{i+1} \sim \mathcal{P}(B|A^{i+1}) \]
Multiple Block Sampling

- What if the HMC is not an option?
  
  \[ A, B \sim P(A, B) \]
  
  • Problem: Design of “good” proposal distributions
  • High rejection rates

- Multiple block sampling (see e.g. Hastings (1997))
  
  • Break down into subproblems
    
    \[ A^{i+1} \sim P(A|B^i) \]
    \[ B^{i+1} \sim P(B|A^{i+1}) \]
  
  • simplifies design of conditional proposal distributions
What if the HMC is not an option?

\[ A, B \sim \mathcal{P}(A, B) \]

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

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

- Break down into subproblems
  
  \[ A^{i+1} \sim \mathcal{P}(A|B^i) \]
  
  \[ B^{i+1} \sim \mathcal{P}(B|A^{i+1}) \]

- Simplifies design of conditional proposal distributions
- Average acceptance rate is higher
Multiple Block Sampling

- What if the HMC is not an option?
  \[ A, B \sim \mathcal{P}(A, B) \]
  - Problem: Design of “good” proposal distributions
  - High rejection rates

- Multiple block sampling (see e.g. Hastings (1997))
  - Break down into subproblems
    \[ A^{i+1} \sim \mathcal{P}(A|B^i) \]
    \[ B^{i+1} \sim \mathcal{P}(B|A^{i+1}) \]
    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?

\[ P(A, B) = \int dC \ P(A, B, C) \]
Can we “boost” block sampling?

\[ P(A, B) = \int dC P(A, B, C) \]

• Sometimes it is easier to explore full joint the PDF

\[ A, B, C \sim P(A, B, C) \]
Multiple Block Sampling

Can we “boost” block sampling?

\[ P(A, B) = \int dC \ P(A, B, C) \]

- Sometimes it is easier to explore full joint the PDF

\[ A, B, C \sim P(A, B, C) \]

- Block sampler:

\[
\begin{align*}
A^{i+1} &\sim P(A|B^i, C^i) \\
B^{i+1} &\sim P(B|A^{i+1}, C^i) \\
C^{i+1} &\sim P(C|A^{i+1}, B^{i+1})
\end{align*}
\]
Multiple Block Sampling

Can we “boost” block sampling?

\[ P(A, B) = \int dC \ P(A, B, C) \]

- Sometimes it is easier to explore full joint the PDF

\[ A, B, C \sim P(A, B, C) \]

- Block sampler:

\[ A_{i+1} \sim P(A|B^i, C^i) = P(A|C^i) \]
\[ B_{i+1} \sim P(B|A_{i+1}, C^i) = P(B|C^i) \]
\[ C_{i+1} \sim P(C|A_{i+1}, B_{i+1}) \]
Multiple Block Sampling

- Can we “boost” block sampling?

\[ P(A, B) = \int dC P(A, B, C) \]

- Sometimes it is easier to explore full joint the PDF

\[ A, B, C \sim P(A, B, C) \]

- Block sampler:

\[
\begin{align*}
A^{i+1} &\sim P(A|B^i, C^i) = P(A|C^i) \\
B^{i+1} &\sim P(B|A^{i+1}, C^i) = P(B|C^i) \\
C^{i+1} &\sim P(C|A^{i+1}, B^{i+1})
\end{align*}
\]

- Permits efficient sampling for numerical expensive posteriors

process in parallel!
Photometric redshift sampling

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

- Photometric surveys
  - millions of galaxies (\( \sim 10^7 - 10^8 \))
  - low redshift accuracy (\( \sim 100 \text{ Mpc along LOS} \))
## Photometric redshift sampling

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

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

- Photometric surveys
  - millions of galaxies ($\sim 10^7 - 10^8$)
  - low redshift accuracy ($\sim 100$ Mpc along LOS)
  - Infer accurate redshifts: $\{z_p\} \sim P(\{z_p\}|d)$
  - Rather sample from joint distribution:
    - $\{z_p\}, \delta \sim P(\delta, \{z_p\}|d)$
  - Block sampler:
    - $z_1 \sim P(z_1|\delta^i, d)$
    - $\ldots$
    - $z_N \sim P(z_N|\delta^i, d)$
    - $\delta^{i+1} \sim P(\delta|\{z_p\}^{i+1}, d)$
Photometric redshift sampling

- Photometric surveys
  - millions of galaxies ($\sim 10^7 - 10^8$)
  - low redshift accuracy ($\sim 100$ Mpc along LOS)
  - Infer accurate redshifts: $\{z_p\} \sim \mathcal{P}(\{z_p\}|d)$
  - Rather sample from joint distribution:
    $$\{z_p\}, \delta \sim \mathcal{P}(\delta, \{z_p\}|d)$$
  - Block sampler:
    $$z_1^{i+1} \sim \mathcal{P}(z_1|\delta^i, d)$$
    $$\vdots$$
    $$z_N^{i+1} \sim \mathcal{P}(z_N|\delta^i, d)$$
    $$\delta^{i+1} \sim \mathcal{P}(\delta|\{z_p\}^{i+1}, d)$$

Process in parallel!
Photometric redshift sampling

- Photometric surveys
  - millions of galaxies (≈10^7 - 10^8)
  - low redshift accuracy (≈ 100 Mpc along LOS)
  - Infer accurate redshifts: \( \{z_p\} \sim \mathcal{P}(\{z_p\}|d) \)
  - Rather sample from joint distribution:
    \[ \{z_p\}, \delta \sim \mathcal{P}(\delta, \{z_p\}|d) \]
  - Block sampler:
    \[
    
    \begin{align*}
    z_1^{i+1} & \sim \mathcal{P}(z_1|\delta^i, d) \\
    \vdots & \\
    z_N^{i+1} & \sim \mathcal{P}(z_N|\delta^i, d) \\
    \delta^{i+1} & \sim \mathcal{P}(\delta|\{z_p\}^{i+1}, d)
    \end{align*}
    \]
    - Process in parallel!
    - HMC sampler!
Photometric redshift sampling

- Application to artificial photometric data

• ~ Noise, Systematics, Position uncertainty (~100 Mpc)
• ~ $10^7$ density amplitudes /parameters
• ~ $2 \times 10^7$ radial galaxy positions / parameters
Photometric redshift sampling

- Application to artificial photometric data

- ~ Noise, Systematics, Position uncertainty (~100 Mpc)
- ~ $10^7$ density amplitudes /parameters
- ~ $2 \times 10^7$ radial galaxy positions / parameters
- ~ $3 \times 10^7$ parameters in total
Photometric redshift sampling

Jasche, Wandelt (2012)
Deviation from the truth

Jasche, Wandelt (2012)
Deviation from the truth

\[ r(k_{th}) = \frac{\langle \delta_{true}^{k_{th} \cdot \langle \delta \rangle_{k_{th}}} \rangle}{\sqrt{\langle \delta_{true}^{k_{th}} \rangle^2} \sqrt{\langle \langle \delta \rangle_{k_{th}}^2 \rangle}} \]

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

Gravity
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 $\rightarrow$ primordial non-Gaussianity
  - 4D dynamic states $\rightarrow$ 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