Bayesian physical reconstruction of initial conditions from large-scale structure surveys

paper by Jens Jasche & Benjamin D. Wandelt

Lorenzo Cavezza - 2130648

12/02/25

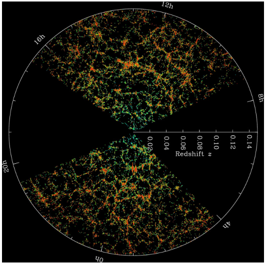
Table of contents

Introduction

LSS analysis

Standard Approach

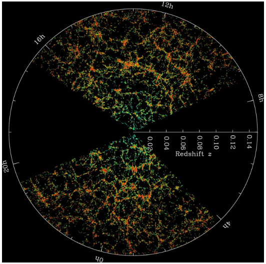
- Tracers' Surveys (Galaxy Surveys)



LSS analysis

Standard Approach

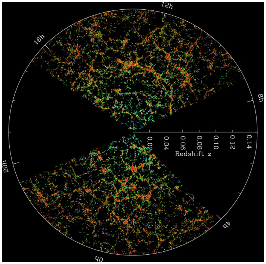
- Tracers' Surveys (Galaxy Surveys)
- Initial Conditions Constraining



LSS analysis

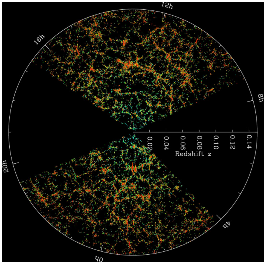
Standard Approach

- Tracers' Surveys (Galaxy Surveys)
- Initial Conditions Constraining
- Evolved Density Field Approximations



Standard Approach

- Tracers' Surveys (Galaxy Surveys)
- Initial Conditions Constraining
- Evolved Density Field Approximations



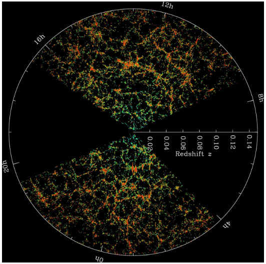
Proposed Alternative

- 2LPT dynamically evolved final density field prior

LSS analysis

Standard Approach

- Tracers' Surveys (Galaxy Surveys)
- Initial Conditions Constraining
- Evolved Density Field Approximations



Proposed Alternative

- 2LPT dynamically evolved final density field prior
- Hamiltonian Monte Carlo exploration of initial field conditioned on mock observations

Bayesian Framework

Given a model of **LSS** formation $G\left(a, \delta^{i}\right)$ (a *Second Order Lagrangian Perturbation Thoery* model (**2LPT**) in our case), we obtain the final density contrast δ^{f} prior as:

$$\mathcal{P}\left(\{\delta_l^f\}\right) = \int d\{\delta_l^i\}\mathcal{P}\left(\{\delta_l^i\}, \{\delta_l^f\} \,\middle|\, \mathbf{S}\right) =$$

The non - Gaussian Prior

Given a model of **LSS** formation $G\left(a,\delta^{i}\right)$ (a *Second Order Lagrangian Perturbation Thoery* model (**2LPT**) in our case), we obtain the final density contrast δ^{f} prior as:

$$\mathcal{P}\left(\{\delta_l^f\}\right) = \int d\{\delta_l^i\}\mathcal{P}\left(\{\delta_l^i\},\{\delta_l^f\}\,\middle|\,\mathbf{S}\right) =$$
$$= \int d\{\delta_l^i\}\mathcal{P}\left(\{\delta_l^f\}\,\middle|\,\{\delta_l^i\}\right)\mathcal{P}\left(\{\delta_l^i\}\,\middle|\,\mathbf{S}\right) =$$

The non - Gaussian Prior

Given a model of **LSS** formation $G\left(a, \delta^{i}\right)$ (a *Second Order Lagrangian Perturbation Thoery* model (**2LPT**) in our case), we obtain the final density contrast δ^{f} prior as:

$$\mathcal{P}\left(\{\delta_l^f\}\right) = \int d\{\delta_l^i\}\mathcal{P}\left(\{\delta_l^i\}, \{\delta_l^f\} \,\middle|\, \mathbf{S}\right) =$$

$$= \int d\{\delta_l^i\}\mathcal{P}\left(\{\delta_l^f\} \,\middle|\, \{\delta_l^i\}\right)\mathcal{P}\left(\{\delta_l^i\} \,\middle|\, \mathbf{S}\right) =$$

$$= \int d\{\delta_l^i\}\prod_l \delta^D\left(\delta_l^f - G\left(a, \delta^i\right)_l\right)\left(\frac{e^{-\frac{1}{2}\sum_{lm}\delta_l^i \mathbf{S}_{lm}^{-1}\delta_m^i}}{det(2\pi\mathbf{S})}\right)$$

The non - Gaussian Prior

Given a model of **LSS** formation $G\left(a, \delta^i\right)$ (a *Second Order Lagrangian Perturbation Thoery* model (**2LPT**) in our case), we obtain the final density contrast δ^f prior as:

$$\mathcal{P}\left(\{\delta^f_l\}\right) = \int d\{\delta^i_l\}\mathcal{P}\left(\{\delta^i_l\}, \{\delta^f_l\} \,\middle|\, \mathbf{S}\right) =$$
$$= \int d\{\delta^i_l\}\mathcal{P}\left(\{\delta^f_l\} \,\middle|\, \{\delta^i_l\}\right)\mathcal{P}\left(\{\delta^i_l\} \,\middle|\, \mathbf{S}\right) =$$
$$= \int d\{\delta^i_l\}\prod_l \delta^D\left(\delta^f_l - G\left(a, \delta^i\right)_l\right)\left(\frac{e^{-\frac{1}{2}\sum_{lm}\delta^i_l \mathbf{S}^{-1}_{lm}\delta^i_m}}{det(2\pi\mathbf{S})}\right)$$

with $\mathbf{S}$ representing the initial contrast covariance matrix.

This can be done through sampling of $\mathcal{P}\left(\{\delta^f_l\}, \{\delta^i_l\} \,\middle|\, \mathbf{S}\right)$ and discarding the δ^i_l realization

The LSS Likelihood

We can model the observed tracers, namely galaxies, through an *inhomogeneous* **Poisson** *Likelihood*:

$$\mathcal{P}\left(\{N_k^g\} \,\middle|\, \{\lambda_k\}\right) = \prod_k \frac{(\lambda_k)^{N_k^g} e^{-\lambda_k}}{N_k^g!}$$

The LSS Likelihood

We can model the observed tracers, namely galaxies, through an *inhomogeneous* **Poisson** *Likelihood*:

$$\mathcal{P}\left(\{N_k^g\} \,\middle|\, \{\lambda_k\}\right) = \prod_k \frac{(\lambda_k)^{N_k^g} e^{-\lambda_k}}{N_k^g!}$$

with

- N_k^g: observed galaxy number at position x_k
- λ_k: expected galaxy number at x_k:

$$\lambda_k = \lambda_k(\delta) = R_k \overline{N} \left(1 + B(\delta^f)_k\right)$$

The LSS Likelihood

We can model the observed tracers, namely galaxies, through an *inhomogeneous* **Poisson** *Likelihood*:

$$\mathcal{P}\left(\{N_k^g\} \,\middle|\, \{\lambda_k\}\right) = \prod_k \frac{(\lambda_k)^{N_k^g} e^{-\lambda_k}}{N_k^g!}$$

with

- N_k^g: observed galaxy number at position x_k
- λ_k: expected galaxy number at x_k:

$$\lambda_k = \lambda_k(\delta) = R_k \overline{N} \left(1 + B(\delta^f)_k\right)$$

  where:

  - R_k: survey geometries and selection functions linear operator
  - $B(x)_k$: non linear and non local bias operator

The 2LPT - Poisson Posterior

The joint posterior of the final and initial contrast densities is:

$$\mathcal{P}\left(\left\{\delta_l^f\right\}, \left\{\delta_l^i\right\} \,\middle|\, \left\{N_i\right\}, \mathbf{S}\right) =$$

The joint posterior of the final and initial contrast densities is:

$$\mathcal{P}\left(\left\{\delta_l^f\right\}, \left\{\delta_l^i\right\} \,\middle|\, \left\{N_i\right\}, \mathbf{S}\right) =$$

$$= \prod_k \frac{\left(\lambda_k\left(\delta^f\right)\right)^{N_k^g} e^{-\lambda_k\left(\delta^f\right)}}{N_k^g!} \prod_l \delta^D\left(\delta_l^f - G\left(a, \delta^i\right)_l\right) \left(\frac{e^{-\frac{1}{2}\sum_{lm}\delta_l^i \mathbf{S}_{lm}^{-1}\delta_m^i}}{det(2\pi\mathbf{S})}\right)$$

The 2LPT - Poisson Posterior

The joint posterior of the final and initial contrast densities is:

$$\mathcal{P}\left(\left\{\delta_l^f\right\},\left\{\delta_l^i\right\}\,\middle|\,\left\{N_i\right\},\mathbf{S}\right)=$$

$$=\prod_k \frac{\left(\lambda_k\left(\delta^f\right)\right)^{N_k^g} e^{-\lambda_k\left(\delta^f\right)}}{N_k^g!}\prod_l \delta^D\left(\delta_l^f - G\left(a,\delta^i\right)_l\right)\left(\frac{e^{-\frac{1}{2}\sum_{lm}\delta_l^i \mathbf{S}_{lm}^{-1}\delta_m^i}}{det(2\pi\mathbf{S})}\right)$$

which leads to the marginalized initial density posterior:

$$\mathcal{P}\left(\left\{\delta_l^i\right\}\,\middle|\,\left\{N_i\right\},\mathbf{S}\right)=\prod_k \frac{\left(\lambda_k\left(G\left(a,\delta^i\right)\right)\right)^{N_k^g} e^{-\lambda_k\left(G\left(a,\delta^i\right)\right)}}{N_k^g!}\left(\frac{e^{-\frac{1}{2}\sum_{lm}\delta_l^i \mathbf{S}_{lm}^{-1}\delta_m^i}}{det(2\pi\mathbf{S})}\right)$$

The 2LPT - Poisson Posterior

The joint posterior of the final and initial contrast densities is:

$$\mathcal{P}\left(\left\{\delta_l^f\right\},\left\{\delta_l^i\right\} \,\middle|\, \left\{N_i\right\},\mathbf{S}\right) =$$

$$= \prod_k \frac{\left(\lambda_k\left(\delta^f\right)\right)^{N_k^g} e^{-\lambda_k\left(\delta^f\right)}}{N_k^g!} \prod_l \delta^D\left(\delta_l^f - G\left(a,\delta^i\right)_l\right)\left(\frac{e^{-\frac{1}{2}\sum_{lm}\delta_l^i \mathbf{S}_{lm}^{-1}\delta_m^i}}{det(2\pi\mathbf{S})}\right)$$

which leads to the marginalized initial density posterior:

$$\mathcal{P}\left(\left\{\delta_l^i\right\} \,\middle|\, \left\{N_i\right\},\mathbf{S}\right) = \prod_k \frac{\left(\lambda_k\left(G\left(a,\delta^i\right)\right)\right)^{N_k^g} e^{-\lambda_k\left(G\left(a,\delta^i\right)\right)}}{N_k^g!}\left(\frac{e^{-\frac{1}{2}\sum_{lm}\delta_l^i \mathbf{S}_{lm}^{-1}\delta_m^i}}{det(2\pi\mathbf{S})}\right)$$

Benefits:

- Forward Model Evaluation
- All uncertaintes accounted (survey geometry, selection biases, galaxy distribution noise and cosmic variance)

Hamiltonian Monte Carlo

- Why *HMC*?

- Why *HMC*?

$\implies$ $2 \cdot 10^6$ parameters to infer (box's voxels) $\implies$ very low acceptance in random walk based MC (*curse of dimensionality*)

Hamiltonian Monte Carlo

- Why *HMC*?

$\Longrightarrow$ $2 \cdot 10^6$ parameters to infer (box's voxels) $\Longrightarrow$ very low acceptance in random walk based MC (*curse of dimensionality*)

$\Longrightarrow$ *HMC* yields an acceptance of nearly 1 (integration errors)

Hamiltonian Monte Carlo

- Why *HMC*?

$\Longrightarrow$ $2 \cdot 10^6$ parameters to infer (box's voxels) $\Longrightarrow$ very low acceptance in random walk based MC (*curse of dimensionality*)

$\Longrightarrow$ *HMC* yields an acceptance of nearly 1 (integration errors)

- *HMC* moves smoothly inside *typical volumes* by exploiting well defined equations of motion through the *Hamiltonian*:

$$H = \sum_i \sum_j \frac{1}{2} p_i \mathbf{M}_{ij}^{-1} p_j + \psi(x) = \sum_i \sum_j \frac{1}{2} p_i \mathbf{M}_{ij}^{-1} p_j - \ln \mathcal{P}(x)$$

Hamiltonian Monte Carlo

- Why *HMC*?

$\implies$ $2 \cdot 10^6$ parameters to infer (box's voxels) $\implies$ very low acceptance in random walk based MC (*curse of dimensionality*)

$\implies$ *HMC* yields an acceptance of nearly 1 (integration errors)

- *HMC* moves smoothly inside *typical volumes* by exploiting well defined equations of motion through the *Hamiltonian*:

$$H = \sum_i \sum_j \frac{1}{2} p_i \mathbf{M}_{ij}^{-1} p_j + \psi(x) = \sum_i \sum_j \frac{1}{2} p_i \mathbf{M}_{ij}^{-1} p_j - \ln \mathcal{P}(x)$$

$$\implies e^{-H} = \mathcal{P}(\{x_i\})\, exp\left(-\frac{1}{2} \sum_i \sum_j p_i \mathbf{M}_{ij}^{-1} p_j\right)$$

Hamiltonian Monte Carlo

- Why *HMC*?

$\Longrightarrow$ $2 \cdot 10^6$ parameters to infer (box's voxels) $\Longrightarrow$ very low acceptance in random walk based MC (*curse of dimensionality*)

$\Longrightarrow$ *HMC* yields an acceptance of nearly 1 (integration errors)

- *HMC* moves smoothly inside *typical volumes* by exploiting well defined equations of motion through the *Hamiltonian*:

$$H = \sum_i \sum_j \frac{1}{2} p_i \mathbf{M}_{ij}^{-1} p_j + \psi(x) = \sum_i \sum_j \frac{1}{2} p_i \mathbf{M}_{ij}^{-1} p_j - \ln \mathcal{P}(x)$$

$$\Longrightarrow \; e^{-H} = \mathcal{P}(\{x_i\})\, exp\left(-\frac{1}{2} \sum_i \sum_j p_i \mathbf{M}_{ij}^{-1} p_j\right)$$

- Momenta randomly drawn and easily marginalized as nuisance parameters. After dynamical evolution the *acceptance probability* is

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

LSS Equations of Motion

We get the Equations of Motion through *Hamilton's equations*:

LSS Equations of Motion

We get the Equations of Motion through *Hamilton's equations*:

$$\frac{d\delta_n^i}{dt} = \frac{\partial H}{\partial p_i} = \sum_j \mathbf{M}_{nj}^{-1} p_j \qquad \text{with} \qquad \mathbf{M}_{nj} = \mathbf{S}_{nj}^{-1} - \delta_{nj}^K D^1 \left. \frac{\partial \mathbf{J}_j(s)}{\partial s_j} \right|_{s_j=\xi_j}$$

LSS Equations of Motion

We get the Equations of Motion through *Hamilton's equations*:

$$\frac{d\delta_n^i}{dt} = \frac{\partial H}{\partial p_i} = \sum_j \mathbf{M}_{nj}^{-1} p_j \qquad \text{with} \qquad \mathbf{M}_{nj} = \mathbf{S}_{nj}^{-1} - \delta_{nj}^K D^1 \left. \frac{\partial \mathbf{J}_j(s)}{\partial s_j} \right|_{s_j = \xi_j}$$

The Hamiltonian mass optimizes the integrator's stability . Vector **J** accounts for the gradient of the first order approximate "*Zeldovich*" potential.

LSS Equations of Motion

We get the Equations of Motion through *Hamilton's equations*:

$$\frac{d\delta_n^i}{dt} = \frac{\partial H}{\partial p_i} = \sum_j \mathbf{M}_{nj}^{-1} p_j \qquad \text{with} \qquad \mathbf{M}_{nj} = \mathbf{S}_{nj}^{-1} - \delta_{nj}^K D^1 \left. \frac{\partial \mathbf{J}_j(s)}{\partial s_j} \right|_{s_j=\xi_j}$$

The Hamiltonian mass optimizes the integrator's stability . Vector **J** accounts for the gradient of the first order approximate "*Zeldovich*" potential.We define the potential of the Hamiltonian system explicitely:

$$\Psi\left(\{\delta_l^i\}\right) = -\ln\left(\mathcal{P}\left(\{\delta_l^i \,\middle|\, \{N_i\}, \mathbf{S}\}\right)\right)$$

LSS Equations of Motion

We get the Equations of Motion through *Hamilton's equations*:

$$\frac{d\delta_n^i}{dt} = \frac{\partial H}{\partial p_i} = \sum_j \mathbf{M}_{nj}^{-1} p_j \qquad \text{with} \qquad \mathbf{M}_{nj} = \mathbf{S}_{nj}^{-1} - \delta_{nj}^K D^1 \left. \frac{\partial \mathbf{J}_j(s)}{\partial s_j} \right|_{s_j=\xi_j}$$

The Hamiltonian mass optimizes the integrator's stability . Vector **J** accounts for the gradient of the first order approximate "*Zeldovich*" potential.We define the potential of the Hamiltonian system explicitely:

$$\Psi\left(\{\delta_l^i\}\right) = -\ln\left(\mathcal{P}\left(\{\delta_l^i \,\middle|\, \{N_i\}, \mathbf{S}\}\right)\right) = \Psi_{prior}\left(\{\delta_l^i\}\right) + \Psi_{likelihood}\left(\{\delta_l^i\}\right) =$$

LSS Equations of Motion

We get the Equations of Motion through *Hamilton's equations*:

$$\frac{d\delta^i_n}{dt} = \frac{\partial H}{\partial p_i} = \sum_j \mathbf{M}^{-1}_{nj} p_j \qquad \text{with} \qquad \mathbf{M}_{nj} = \mathbf{S}^{-1}_{nj} - \delta^K_{nj} D^1 \left. \frac{\partial \mathbf{J}_j(s)}{\partial s_j} \right|_{s_j=\xi_j}$$

The Hamiltonian mass optimizes the integrator's stability . Vector **J** accounts for the gradient of the first order approximate *"Zeldovich"* potential.We define the potential of the Hamiltonian system explicitely:

$$\Psi\left(\{\delta^i_l\}\right) = -\ln\left(\mathcal{P}\left(\{\delta^i_l \,\middle|\, \{N_i\}, \mathbf{S}\}\right)\right) = \Psi_{prior}\left(\{\delta^i_l\}\right) + \Psi_{likelihood}\left(\{\delta^i_l\}\right) =$$

$$= \frac{1}{2}\sum_{lm} \delta^i_l \mathbf{S}^{-1}_{lm} \delta^i_m + \sum_k \left[R_k \bar{N}\left(1 + G\left(a, \delta^i\right)_k\right) - N_k \ln\left(R_k \bar{N}\left(1 + G\left(a, \delta^i\right)_k\right)\right)\right]$$

LSS Equations of Motion

We get the Equations of Motion through *Hamilton's equations*:

$$\frac{d\delta_n^i}{dt} = \frac{\partial H}{\partial p_i} = \sum_j \mathbf{M}_{nj}^{-1} p_j \qquad \text{with} \qquad \mathbf{M}_{nj} = \mathbf{S}_{nj}^{-1} - \delta_{nj}^K D^1 \left.\frac{\partial \mathbf{J}_j(s)}{\partial s_j}\right|_{s_j=\xi_j}$$

The Hamiltonian mass optimizes the integrator's stability . Vector **J** accounts for the gradient of the first order approximate *"Zeldovich"* potential.We define the potential of the Hamiltonian system explicitely:

$$\Psi\left(\{\delta_l^i\}\right) = -\ln\left(\mathcal{P}\left(\{\delta_l^i \,\middle|\, \{N_i\}, \mathbf{S}\}\right)\right) = \Psi_{prior}\left(\{\delta_l^i\}\right) + \Psi_{likelihood}\left(\{\delta_l^i\}\right) =$$

$$= \frac{1}{2}\sum_{lm} \delta_l^i \mathbf{S}_{lm}^{-1} \delta_m^i + \sum_k \left[R_k \bar{N}\left(1 + G\left(a, \delta^i\right)_k\right) - N_k \ln\left(R_k \bar{N}\left(1 + G\left(a, \delta^i\right)_k\right)\right)\right]$$

$$\implies \frac{dp_n^i}{dt} = -\frac{\partial \Psi_{prior}\left(\{\delta_l^i\}\right)}{\partial \delta_p^i} - \frac{\partial \Psi_{likelihood}\left(\{\delta_l^i\}\right)}{\partial \delta_p^i} =$$

We get the Equations of Motion through *Hamilton's equations*:

$$\frac{d\delta_n^i}{dt} = \frac{\partial H}{\partial p_i} = \sum_j \mathbf{M}_{nj}^{-1} p_j \qquad \text{with} \qquad \mathbf{M}_{nj} = \mathbf{S}_{nj}^{-1} - \delta_{nj}^K D^1 \left. \frac{\partial \mathbf{J}_j(s)}{\partial s_j} \right|_{s_j=\xi_j}$$

The Hamiltonian mass optimizes the integrator's stability . Vector $\mathbf{J}$ accounts for the gradient of the first order approximate "*Zeldovich*" potential.We define the potential of the Hamiltonian system explicitely:

$$\Psi\left(\{\delta_l^i\}\right) = -\ln\left(\mathcal{P}\left(\{\delta_l^i \,\middle|\, \{N_i\}, \mathbf{S}\}\right)\right) = \Psi_{prior}\left(\{\delta_l^i\}\right) + \Psi_{likelihood}\left(\{\delta_l^i\}\right) =$$

$$= \frac{1}{2}\sum_{lm} \delta_l^i \mathbf{S}_{lm}^{-1} \delta_m^i + \sum_k \left[R_k \bar{N}\left(1 + G\left(a, \delta^i\right)_k\right) - N_k \ln\left(R_k \bar{N}\left(1 + G\left(a, \delta^i\right)_k\right)\right)\right]$$

$$\implies \frac{dp_n^i}{dt} = -\frac{\partial \Psi_{prior}\left(\{\delta_l^i\}\right)}{\partial \delta_p^i} - \frac{\partial \Psi_{likelihood}\left(\{\delta_l^i\}\right)}{\partial \delta_p^i} =$$
$$= -\sum_j \mathbf{S}_{pj}^{-1} \delta_j^i + D^1 J_n - D^2 \sum_{a>b}\left(\tau_n^{aabb} - \tau_n^{bbaa} - 2\tau_n^{abab}\right)$$

where the τ_m^{abcd} accounts for the gradient of the second order lagrangian potential.

The Leapfrog Method

As an integration scheme the Leapfrog Method is used:
$\Longrightarrow$ simplest integration method that conserves the energy and phase space area and is time reversible

The Leapfrog Method

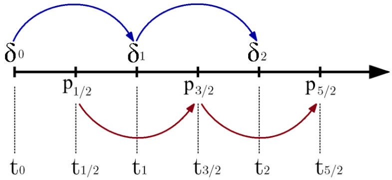
As an integration scheme the Leapfrog Method is used:

$\Longrightarrow$ simplest integration method that conserves the energy and phase space area and is time reversible

$$p_m\left(t+\frac{\epsilon}{2}\right) = p_m(t) - \frac{\epsilon}{2}\frac{\partial\psi\left(\{\delta^i_k\}\right)}{\delta^i_l}\Bigg|_{\delta^i_m(t)}$$

$$\delta^i_m\left(t+\epsilon\right) = \delta^i_m(t) - \frac{\epsilon}{m_i}p_m\left(t+\frac{\epsilon}{2}\right)$$

$$p_m\left(t+\epsilon\right) = p_m\left(t+\frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\frac{\partial\psi\left(\{\delta^i_k\}\right)}{\delta^i_l}\Bigg|_{\delta^i_m(t+\epsilon)}$$



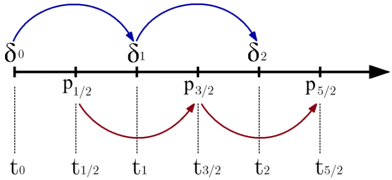
The Leapfrog Method

As an integration scheme the Leapfrog Method is used:
$\Longrightarrow$ simplest integration method that conserves the energy and phase space area and is time reversible

$$p_m\left(t+\tfrac{\epsilon}{2}\right) = p_m(t) - \tfrac{\epsilon}{2}\left.\frac{\partial\psi\left(\{\delta^i_k\}\right)}{\delta^i_l}\right|_{\delta^i_m(t)}$$

$$\delta^i_m\left(t+\epsilon\right) = \delta^i_m(t) - \tfrac{\epsilon}{m_i}\, p_m\left(t+\tfrac{\epsilon}{2}\right)$$

$$p_m\left(t+\epsilon\right) = p_m\left(t+\tfrac{\epsilon}{2}\right) - \tfrac{\epsilon}{2}\left.\frac{\partial\psi\left(\{\delta^i_k\}\right)}{\delta^i_l}\right|_{\delta^i_m(t+\epsilon)}$$



which translates to:

$$p_m(t+\epsilon) = p_m(t) - \tfrac{\epsilon}{2}\left(\left.\frac{\partial\Psi(\delta^i)}{\partial\delta^i_i}\right|_{\delta^i(t)} + \left.\frac{\partial\Psi(\delta^i)}{\partial\delta^i_m}\right|_{\delta^i(t+\epsilon)}\right)$$

$$\delta_m(t+\epsilon) = \delta_m(t) + \epsilon\sum_j M^{-1}_{mj}\, p_j(t) - \frac{\epsilon^2}{2}\sum_j M^{-1}_{mj}\left.\frac{\partial\Psi(\delta^i)}{\partial\delta^i_j}\right|_{\delta^i(t)}$$

Testing the Model

Mock Observations

δ_i Generation

Mock Observations

δ_i Generation

$\downarrow$

Expansion by D^+

Mock Observations

δ_i Generation

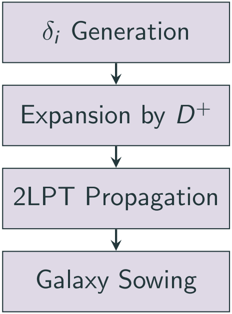
↓

Expansion by D^+

↓

2LPT Propagation

Mock Observations



Mock Observations

δ_i Generation

↓

Expansion by D^+

↓

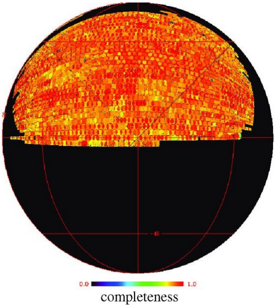
2LPT Propagation

↓

Galaxy Sowing

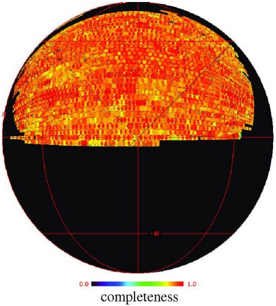
↓

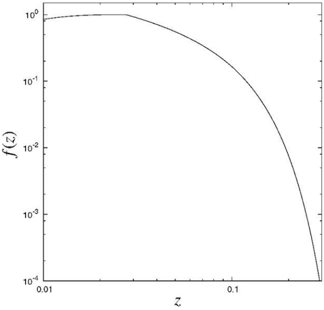
Sloan Mask



Mock Observations

δ_i Generation
↓
Expansion by D^+
↓
2LPT Propagation
↓
Galaxy Sowing
↓
Sloan Mask
↓
Schechter Selection





with: $f(z) = \frac{\int_{M_{min}(z)}^{M_{max}(z)} \Phi(M)dM}{\int_{M_{min}}^{M_{max}} \Phi(M)dM}$

where $\Phi(M)$ is the *Schechter* luminosity function.

Mock Observations

δ_i Generation

↓

Expansion by D^+

↓

2LPT Propagation

↓

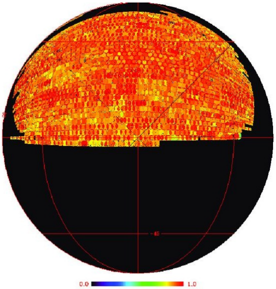
Galaxy Sowing

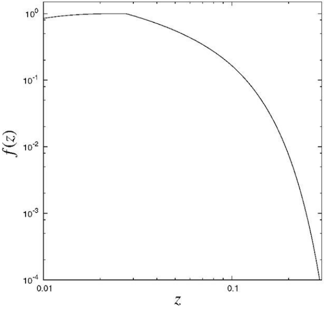
↓

Sloan Mask

↓

Schechter Selection





with: $f(z) = \frac{\int_{M_{min}(z)}^{M_{max}(z)} \Phi(M)dM}{\int_{M_{min}}^{M_{max}} \Phi(M)dM}$

where $\Phi(M)$ is the *Schechter* luminosity function. The linear operator used in Poisson Sampling is:

$$R_i = M_i F_i = M(\alpha_i, \delta_i) f^I(z_i)$$

Burn - In

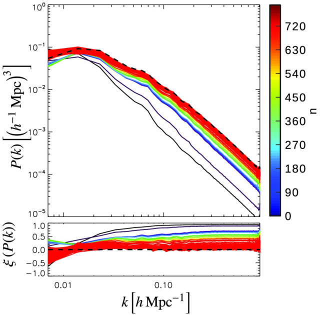
- Arbitrarily lower strength δ_l^i by one tenfold
- Quantify power spectra drift with:

$$\xi\left(P_i(K)\right) = 1 - \frac{P_i(k)}{P^0(k)}$$

Burn - In

- Arbitrarily lower strength δ_l^i by one tenfold
- Quantify power spectra drift with:

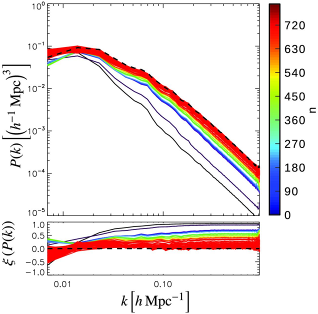
$$\xi\left(P_i(K)\right) = 1 - \frac{P_i(k)}{P^0(k)}$$

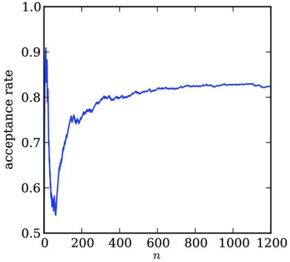


Burn - In

- Arbitrarily lower strength δ_l^i by one tenfold
- Quantify power spectra drift with:

$$\xi\left(P_i(K)\right) = 1 - \frac{P_i(k)}{P^0(k)}$$





- Acceptance takes a dip after manual drift
- Back to normal ($\sim$ 84%) after $\sim$ 600 epochs

Autocorrelation

Autocorrelation is computed to determine the number of independent samples and to assess good mixing:

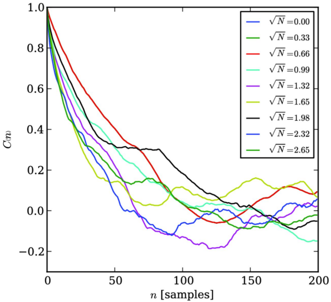
$$C(\delta)_n = \left\langle \frac{\delta^i - \langle\delta\rangle}{\sqrt{Var(\delta)}} \frac{\delta^{i+n} - \langle\delta\rangle}{\sqrt{Var(\delta)}} \right\rangle$$

Autocorrelation

Autocorrelation is computed to determine the number of independent samples and to assess good mixing:

$$C(\delta)_n = \left\langle \frac{\delta^i - \langle\delta\rangle}{\sqrt{Var(\delta)}} \frac{\delta^{i+n} - \langle\delta\rangle}{\sqrt{Var(\delta)}} \right\rangle$$

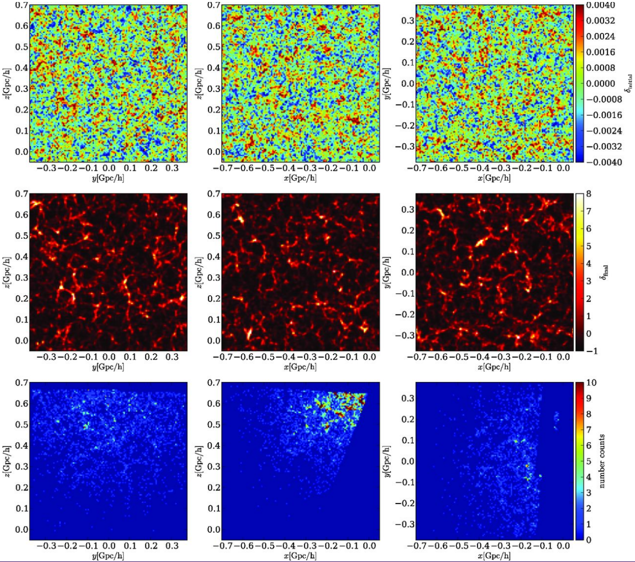
We see such correlations plotted from different signal-to-noise ratios $\sqrt{N}$



$$C(\delta)_n < 0.1 \text{ when } n \sim 200 \implies \text{correlation length is about 200}$$

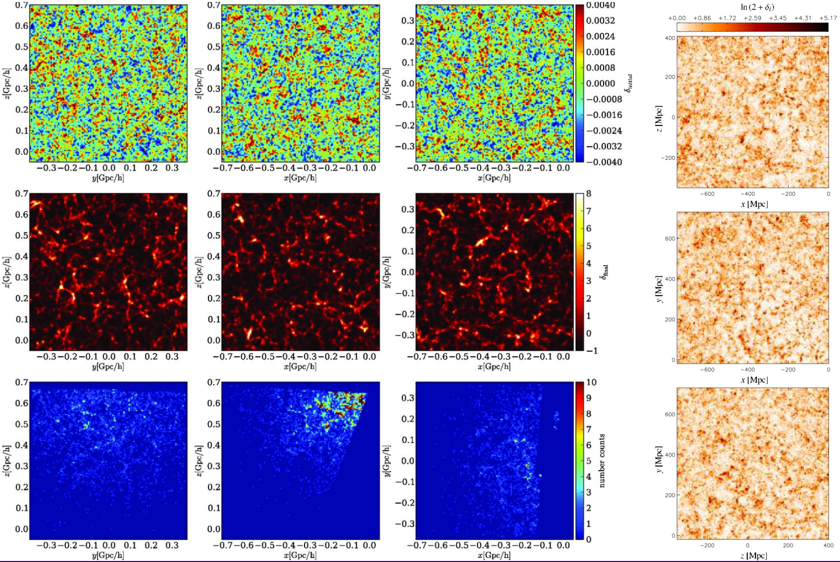
Inferred Density Fields

Samples from the sampled density fields and mock data:

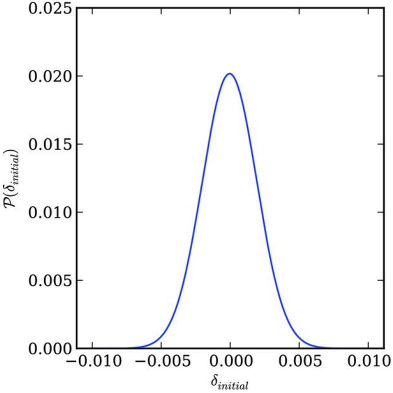


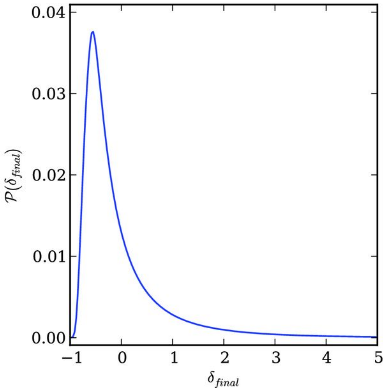
Inferred Density Fields

Samples from the sampled density fields and mock data:



One-point distribution of density contrasts:

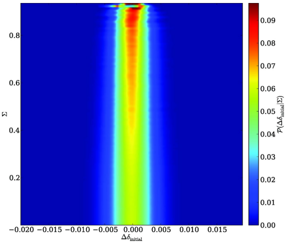




Accuracy

Inferred initial density accuracy is assessed with the posterior of $\Delta\delta_{initial} = \delta_{initial}^{true} - \delta_{initial}$ conditioned to a signal-to-noise ratio proxy:

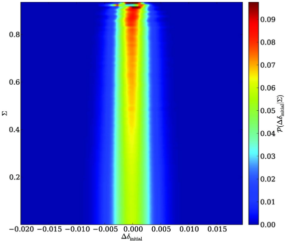
$$\Sigma = \frac{|\langle\delta_{initial}\rangle|}{\sqrt{\langle(\delta_{initial} - \langle\delta_{initial}\rangle)^2\rangle}}$$

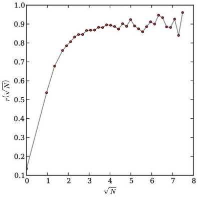


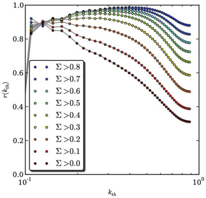
Accuracy

Inferred initial density accuracy is assessed with the posterior of $\Delta\delta_{initial} = \delta_{initial}^{true} - \delta_{initial}$ conditioned to a signal-to-noise ratio proxy:

$$\Sigma = \frac{|\langle\delta_{initial}\rangle|}{\sqrt{\langle(\delta_{initial} - \langle\delta_{initial}\rangle)^2\rangle}}$$





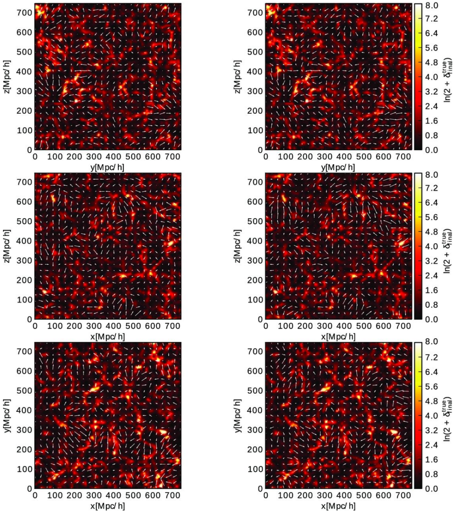


The correlation coefficient between inferred and real densities is computed as a function of $\sqrt{N}$ for the final and of k for different Σ for the initial:

$$r(k_x) = \frac{\langle\delta_0^x\langle\delta\rangle^x\rangle}{\sqrt{\langle(\delta_0^x)^2\rangle}\sqrt{\langle(\delta^x)^2\rangle}}$$

Inferred Dynamics

True vs inferred dynamics:



Conclusions

Conclusions

- Successfully implemented a forward propagation Bayesian dynamical LSS inference algorithm that correctly accounts for high order statistics in generating initial and final density fields through HMC and 2LPT conditioning on mock data

Conclusions

- Successfully implemented a forward propagation Bayesian dynamical LSS inference algorithm that correctly accounts for high order statistics in generating initial and final density fields through HMC and 2LPT conditioning on mock data
- Achieved low correlation length, short burn - in phase and good mixing, as well as impressive reconstruction capabilities in masked regions

Conclusions

- Successfully implemented a forward propagation Bayesian dynamical LSS inference algorithm that correctly accounts for high order statistics in generating initial and final density fields through HMC and 2LPT conditioning on mock data
- Achieved low correlation length, short burn - in phase and good mixing, as well as impressive reconstruction capabilities in masked regions
- High accuracy in final and initial inferred fields, in particular for high signal-to-noise regions and large scales

Conclusions

- Successfully implemented a forward propagation Bayesian dynamical LSS inference algorithm that correctly accounts for high order statistics in generating initial and final density fields through HMC and 2LPT conditioning on mock data
- Achieved low correlation length, short burn - in phase and good mixing, as well as impressive reconstruction capabilities in masked regions
- High accuracy in final and initial inferred fields, in particular for high signal-to-noise regions and large scales
- Faithful reconstruction of the final density velocity field

Conclusions

- Successfully implemented a forward propagation Bayesian dynamical LSS inference algorithm that correctly accounts for high order statistics in generating initial and final density fields through HMC and 2LPT conditioning on mock data
- Achieved low correlation length, short burn - in phase and good mixing, as well as impressive reconstruction capabilities in masked regions
- High accuracy in final and initial inferred fields, in particular for high signal-to-noise regions and large scales
- Faithful reconstruction of the final density velocity field
- Possible further improvements include better model errors and accuracy of 2LPT, real galaxy surveys tests, and incorporation of halo-mode-based galaxy bias models

Thank you for your attention