Uncertainty Quantification for Wave-field Reconstruction Inversion

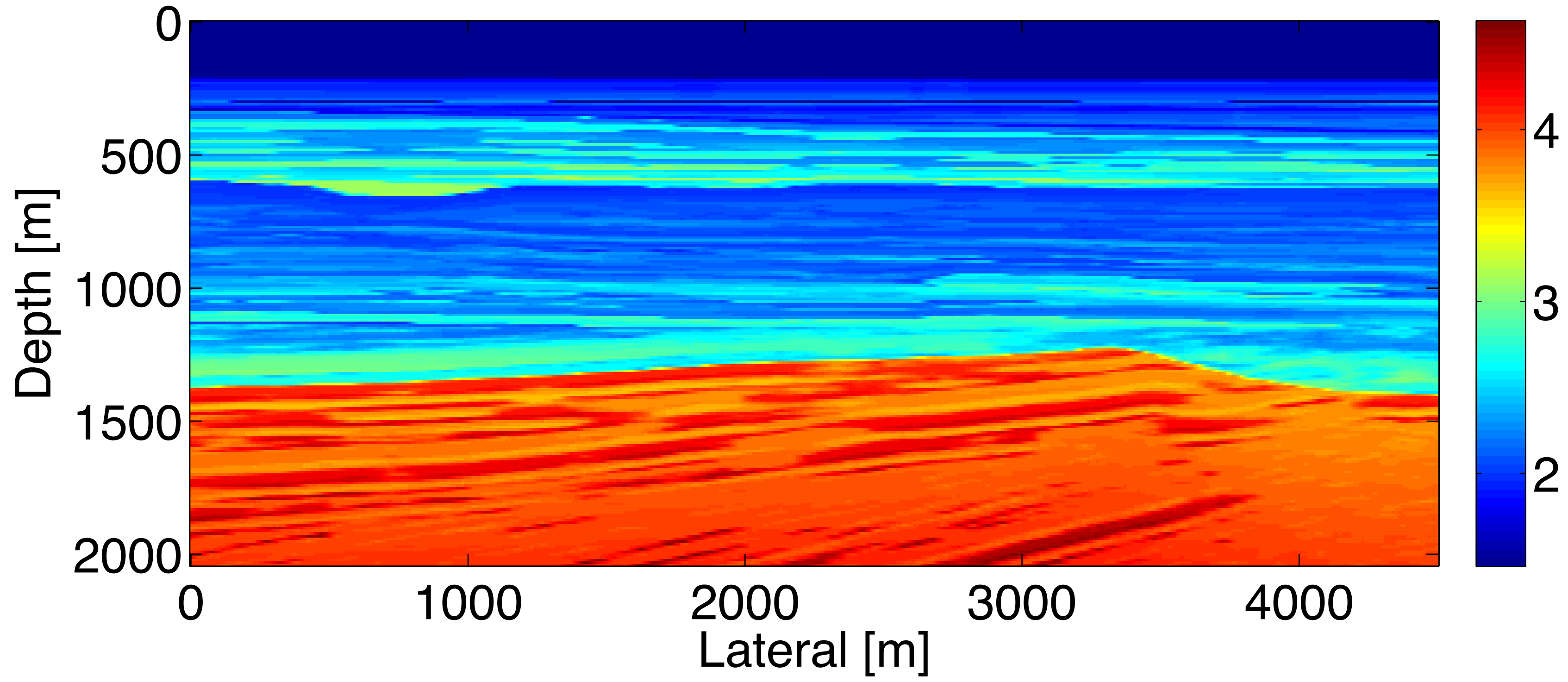
Zhilong Fang, Chia-Ying Lee, Curt Da Silva, Felix Herrmann and Rachel Kuske

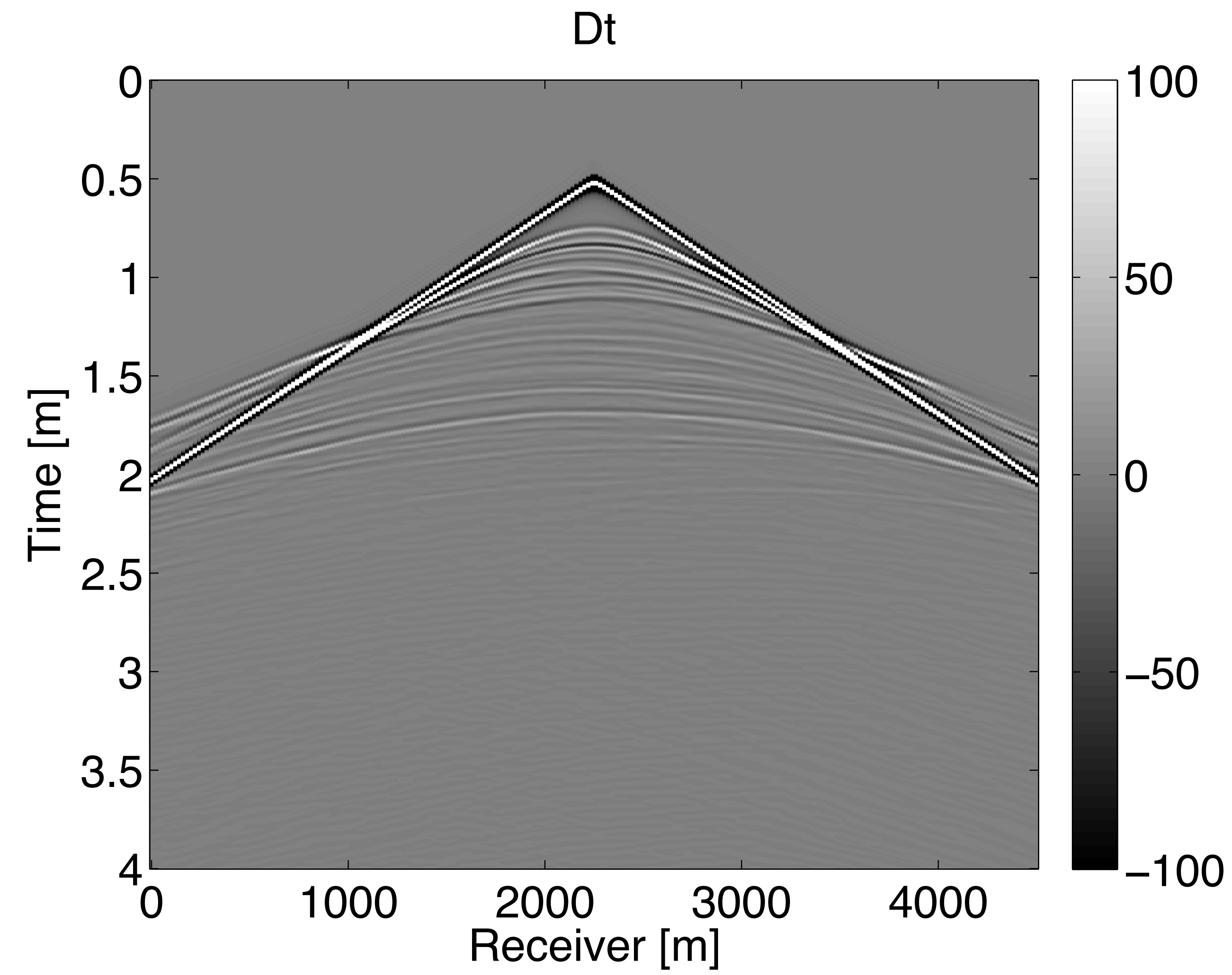


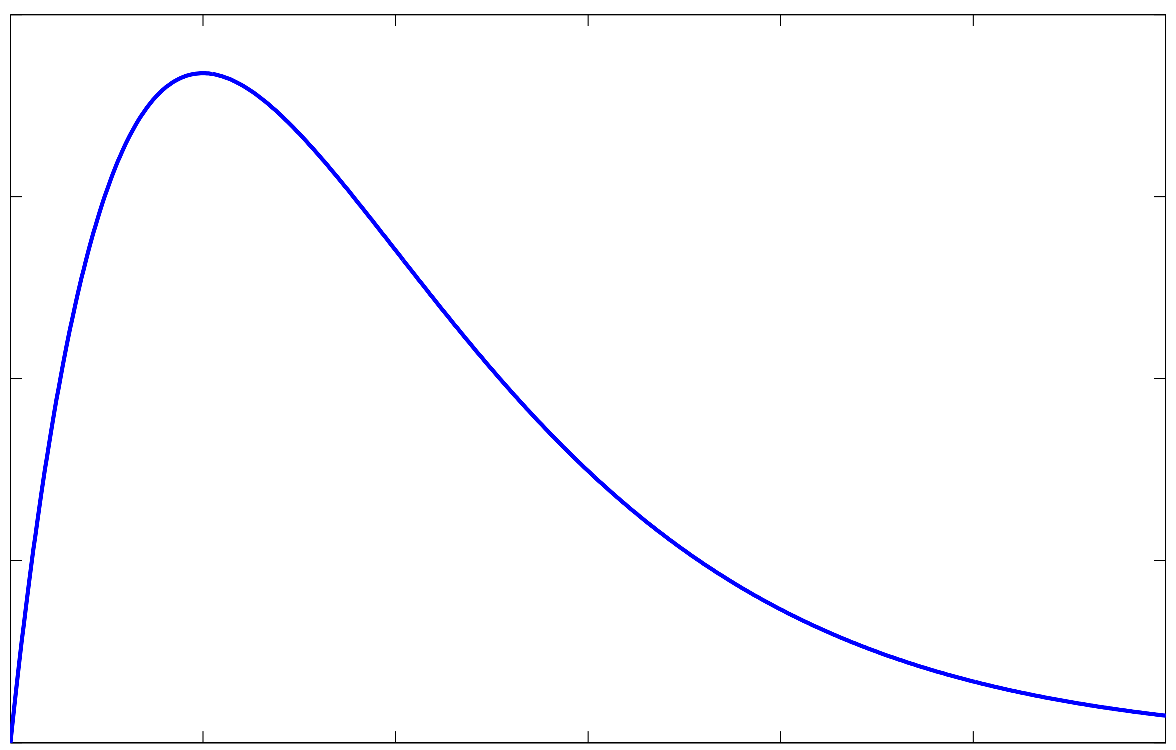


SLIM
University of British Columbia

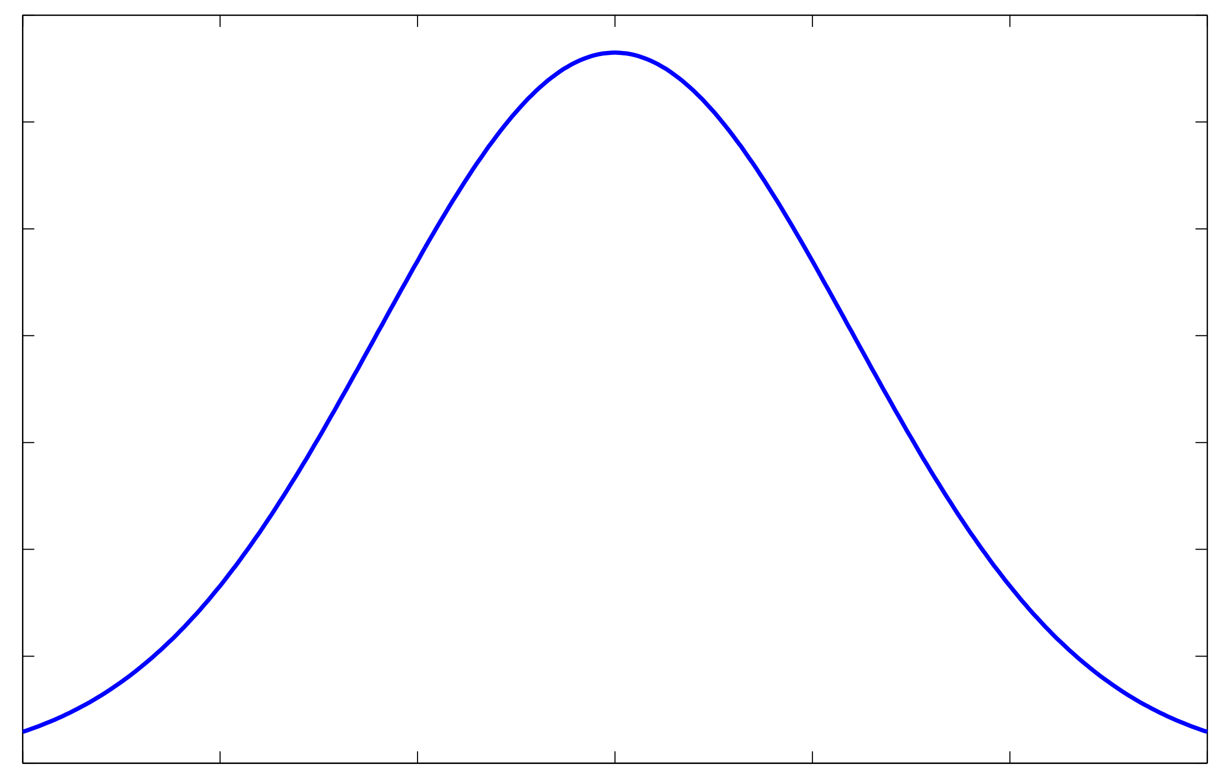
Motivation







Distribution of Model

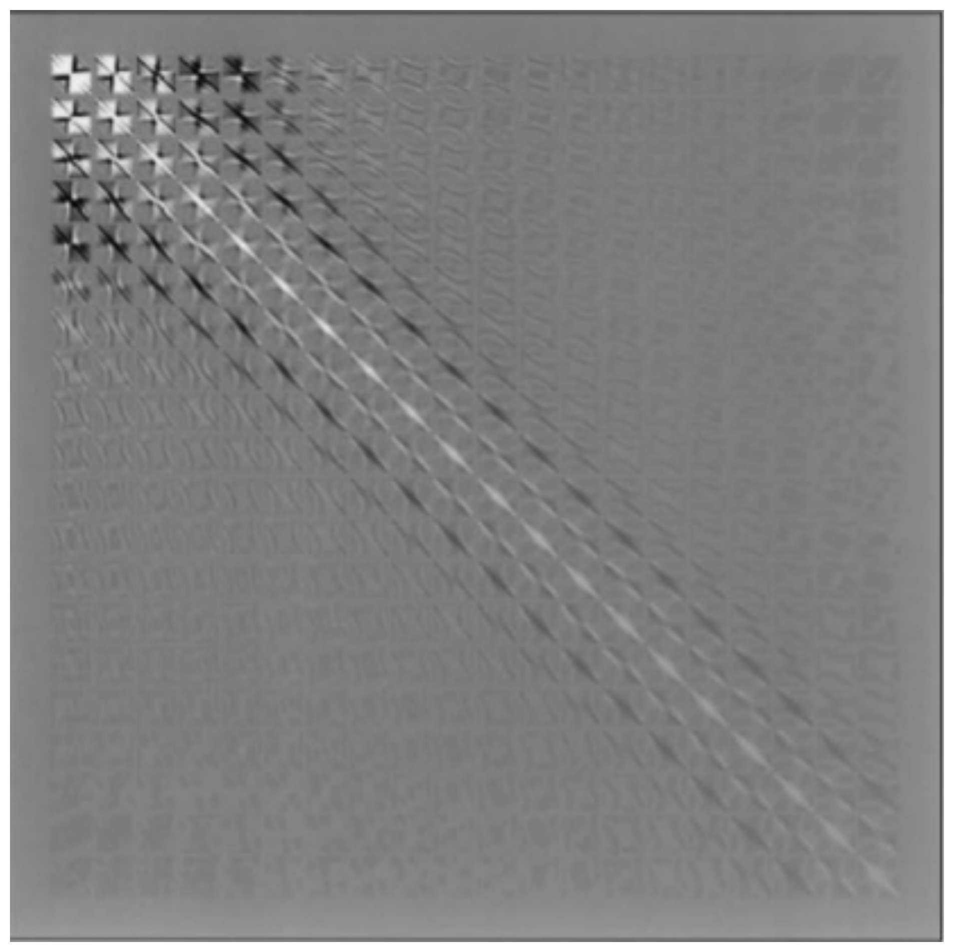


Distribution of Data

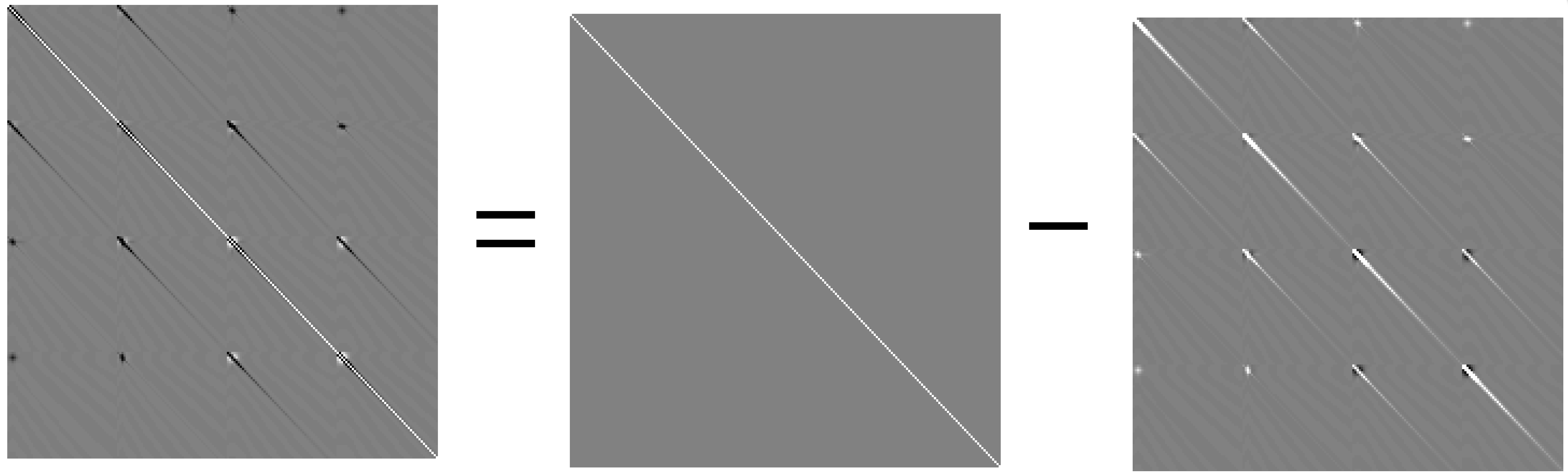
Motivation

FWI	WRI
Strongly nonlinear	“bilinear”
Dense Gauss-Newton Hessian	Sparse Gauss-Newton Hessian

Motivation



Gauss-Newton Hessian of FWI



Hessian of WRI

Goal

- Set up a reasonable distribution for the model given observed data.
- Derive efficient method to calculate/estimate the distribution.
- Generate different statistical parameters of the model to quantify the uncertainty.

SLIM

[van Leeuwen, T and Herrmann, F J , 2013]
[J. Virieux and S. Operto, 2009]
[Tarantola, 1984]

Deterministic FWI vs WRI

Full-Waveform Inversion (FWI):

$$\min_{\mathbf{m}} \|\mathbf{P}\mathbf{A}(\mathbf{m})^{-1}\mathbf{q} - \mathbf{d}_{\mathrm{obs}}\|_2^2$$

Wave-field Reconstruction Inversion (WRI):

$$\min_{\mathbf{m},\mathbf{u}} \|\mathbf{P}\mathbf{u} - \mathbf{d}_{\mathrm{obs}}\|_2^2 + \lambda^2\|\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{q}\|_2^2$$

[Tarantola, 1982]

Statistical FWI vs WRI

FWI:

$$\rho_{\text{post}}(\mathbf{m}) \propto \exp\left(\underbrace{-\|\mathbf{P}\mathbf{A}(\mathbf{m})^{-1}\mathbf{q} - \mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}}}_{\text{Likelihood}} \underbrace{- \|\mathbf{m} - \mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}}}_{\text{Prior}}\right)$$

WRI:

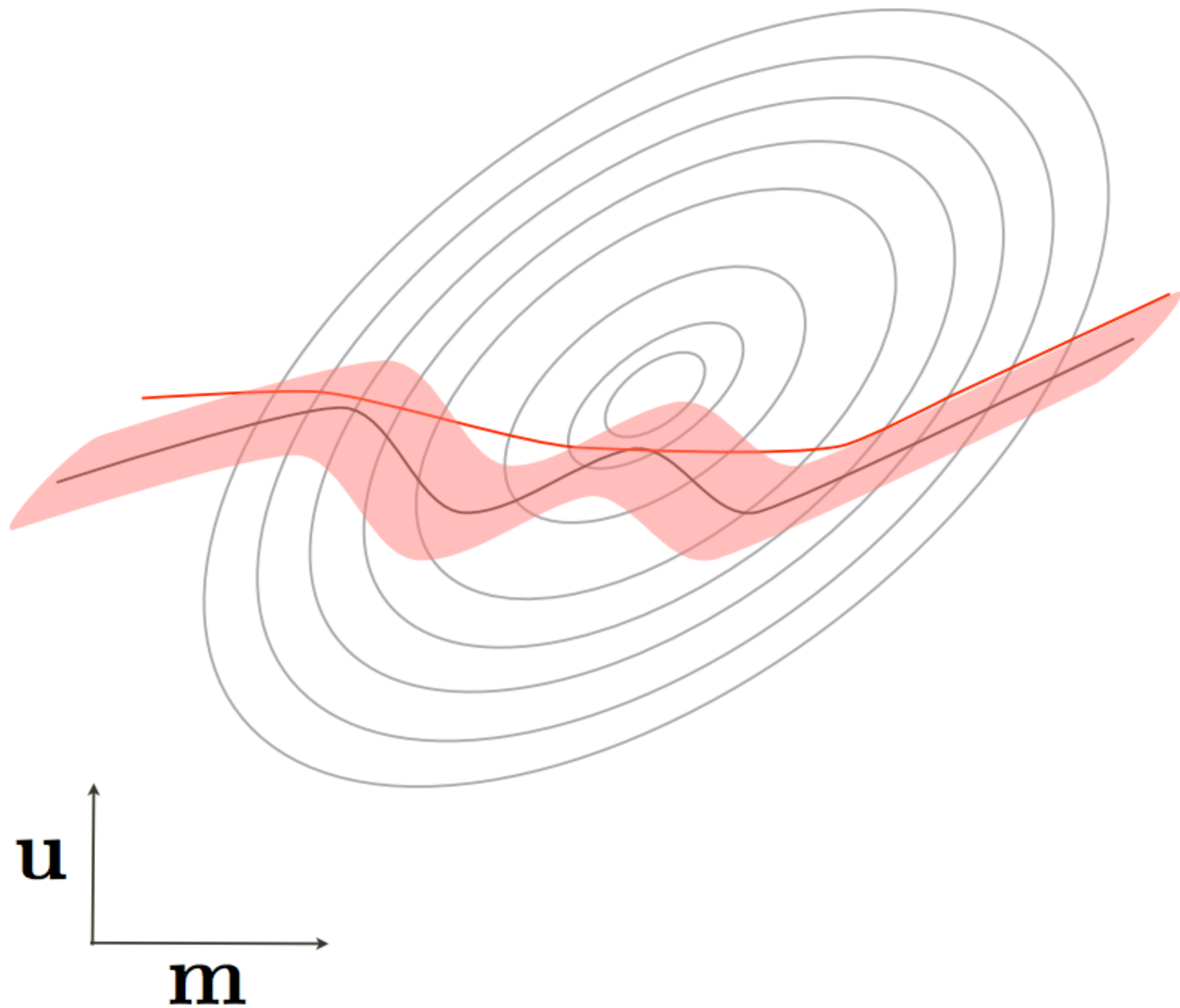
$$\rho_{\text{post}}(\mathbf{m},\mathbf{u}) \propto \exp(\underbrace{-\|\mathbf{P}\mathbf{u} - \mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}}}_{\text{Likelihood}}$$

$$\underbrace{-\lambda^2\|\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{q}\|^2_{\mathbf{\Sigma}^{-1}_{\text{pde}}} - \|\mathbf{m} - \mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}}}_{\text{Prior}})$$

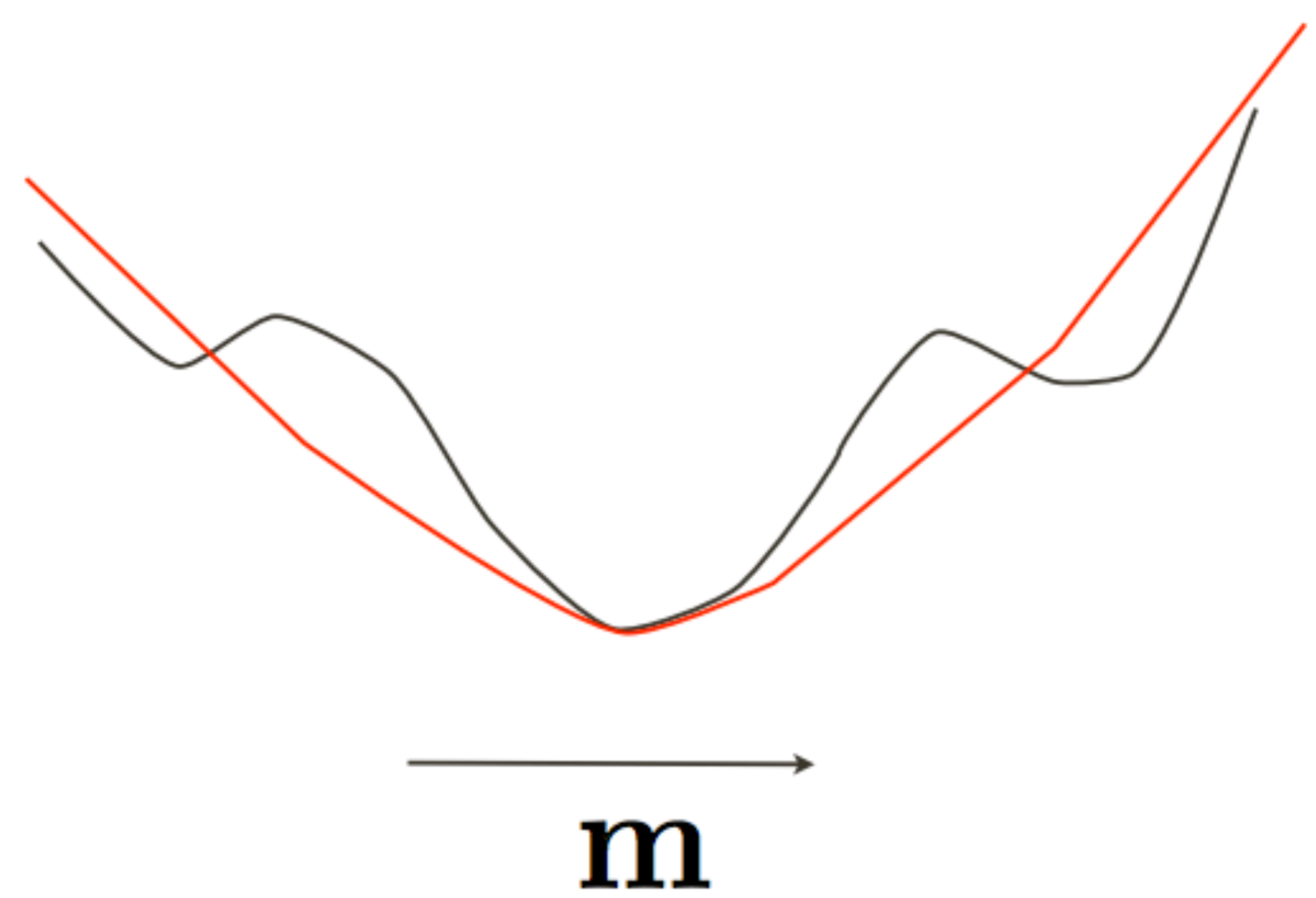
WRI vs FWI

[van Leeuwen, T and Herrmann, F J , 2013]

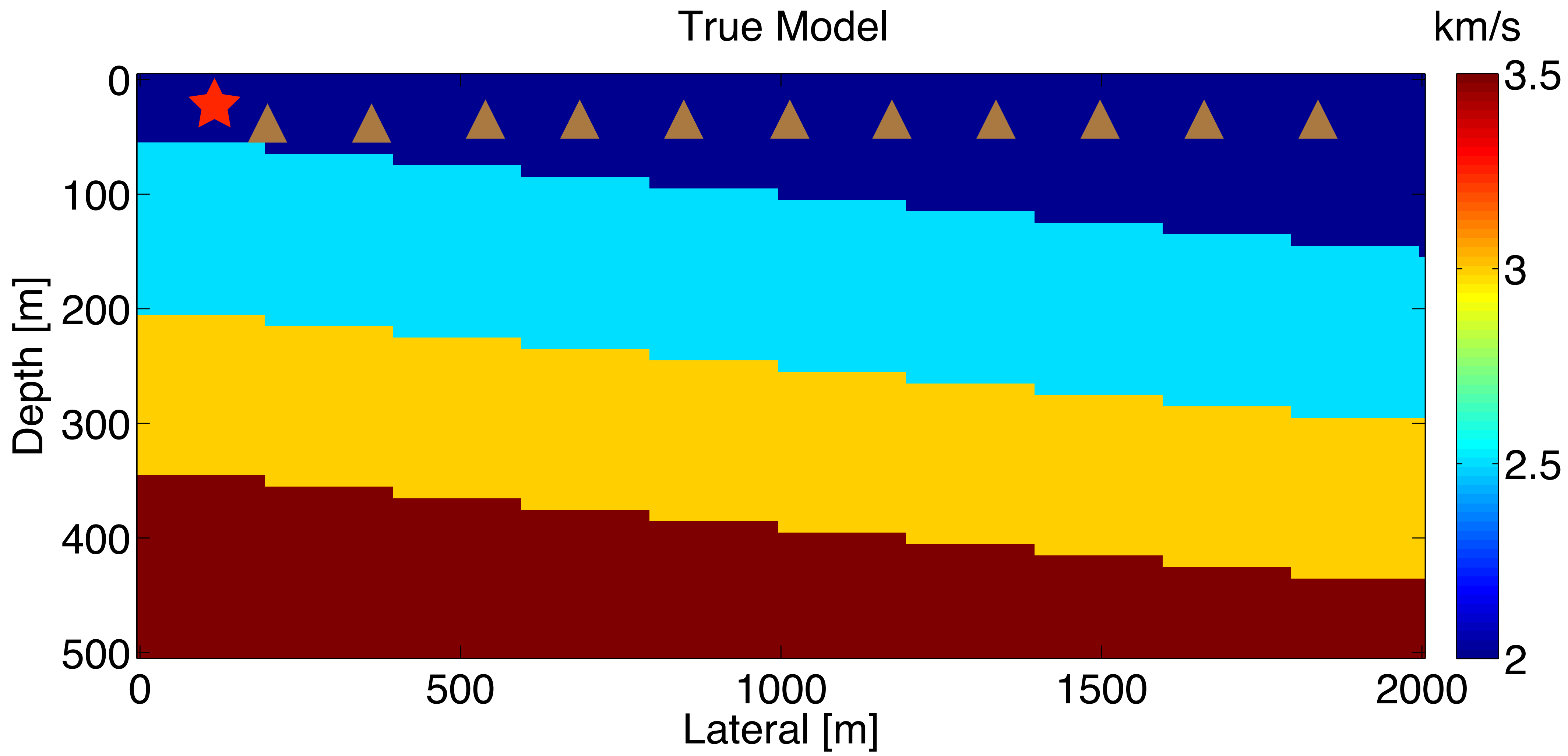
Larger # of degrees of freedom



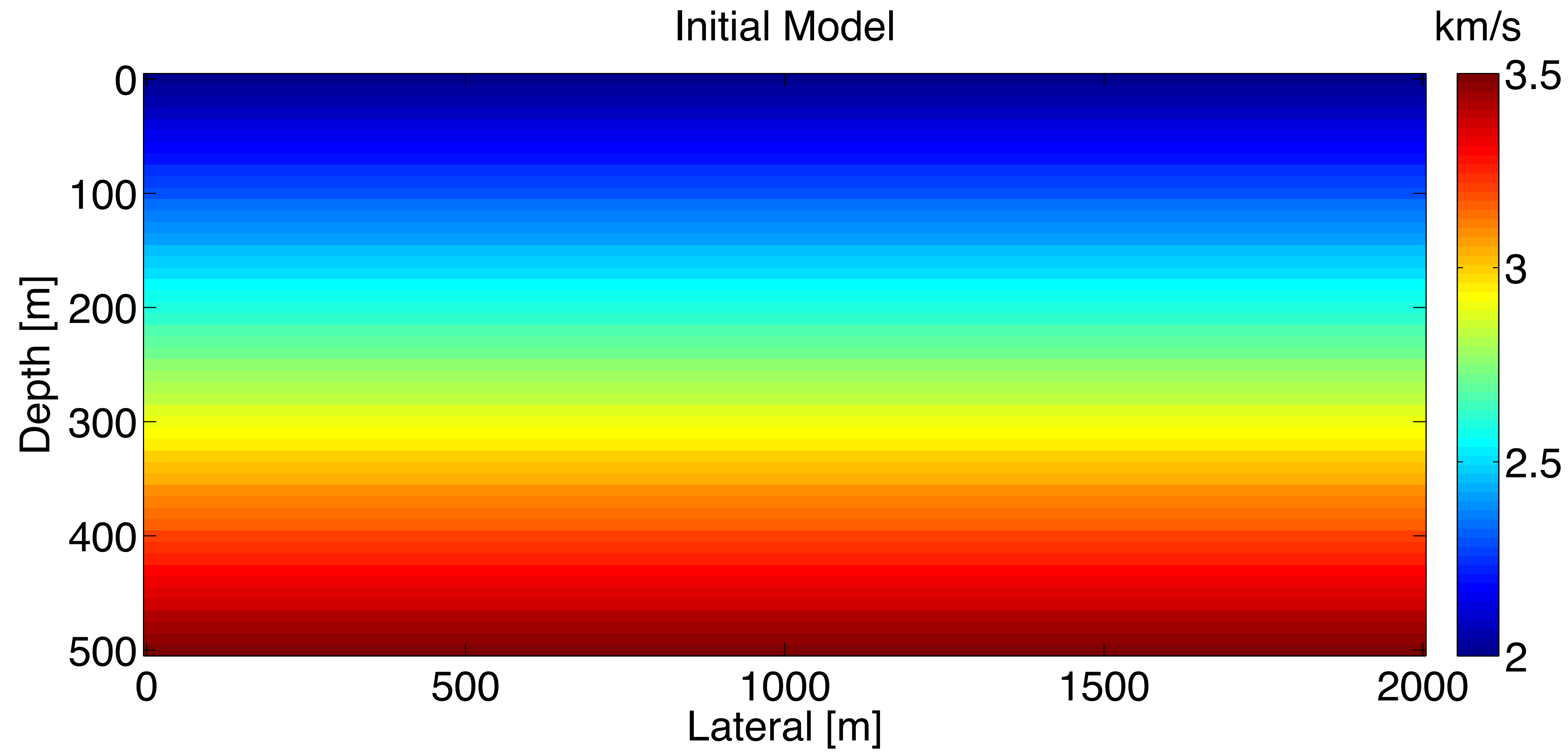
"more convex"



Simple test

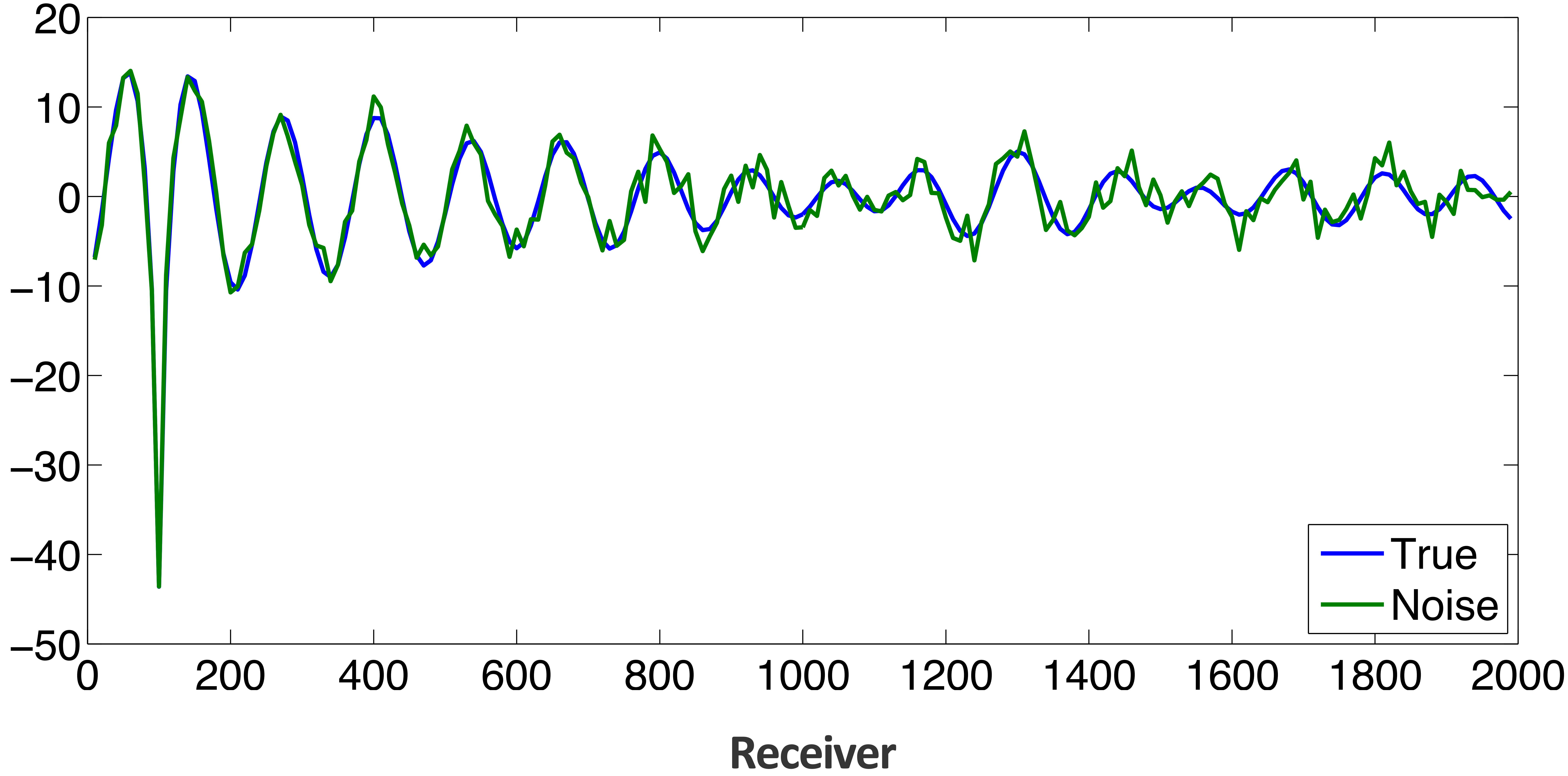


True Model
km/s
3.5
3
2.5
2
0
100
200
300
400
500
Depth [m]
0
500
1000
1500
2000
Lateral [m]



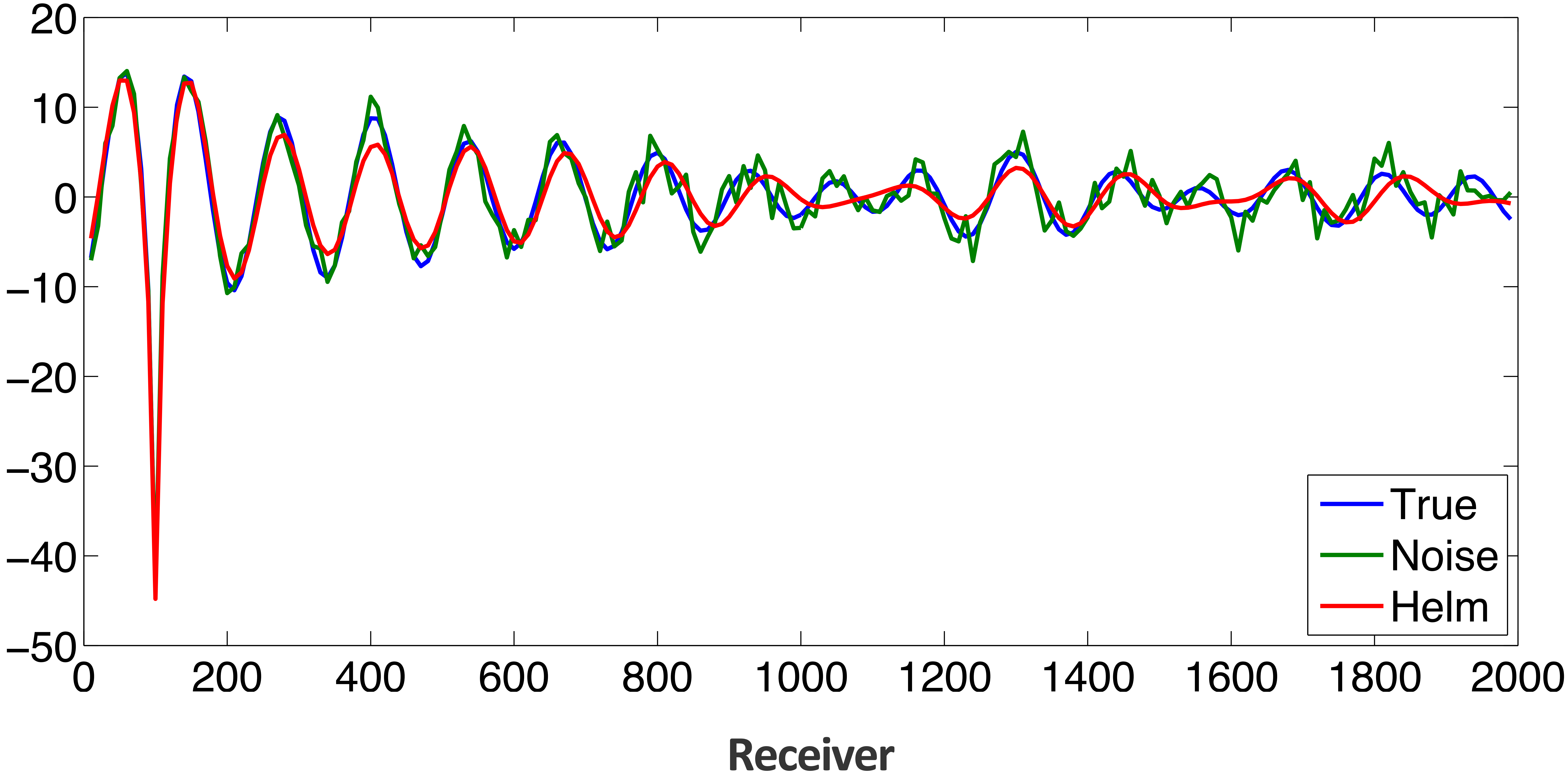
Initial Model
km/s
3.5
3
2.5
2
0
100
200
300
400
500
Depth [m]
0
500
1000
1500
2000
Lateral [m]

Data

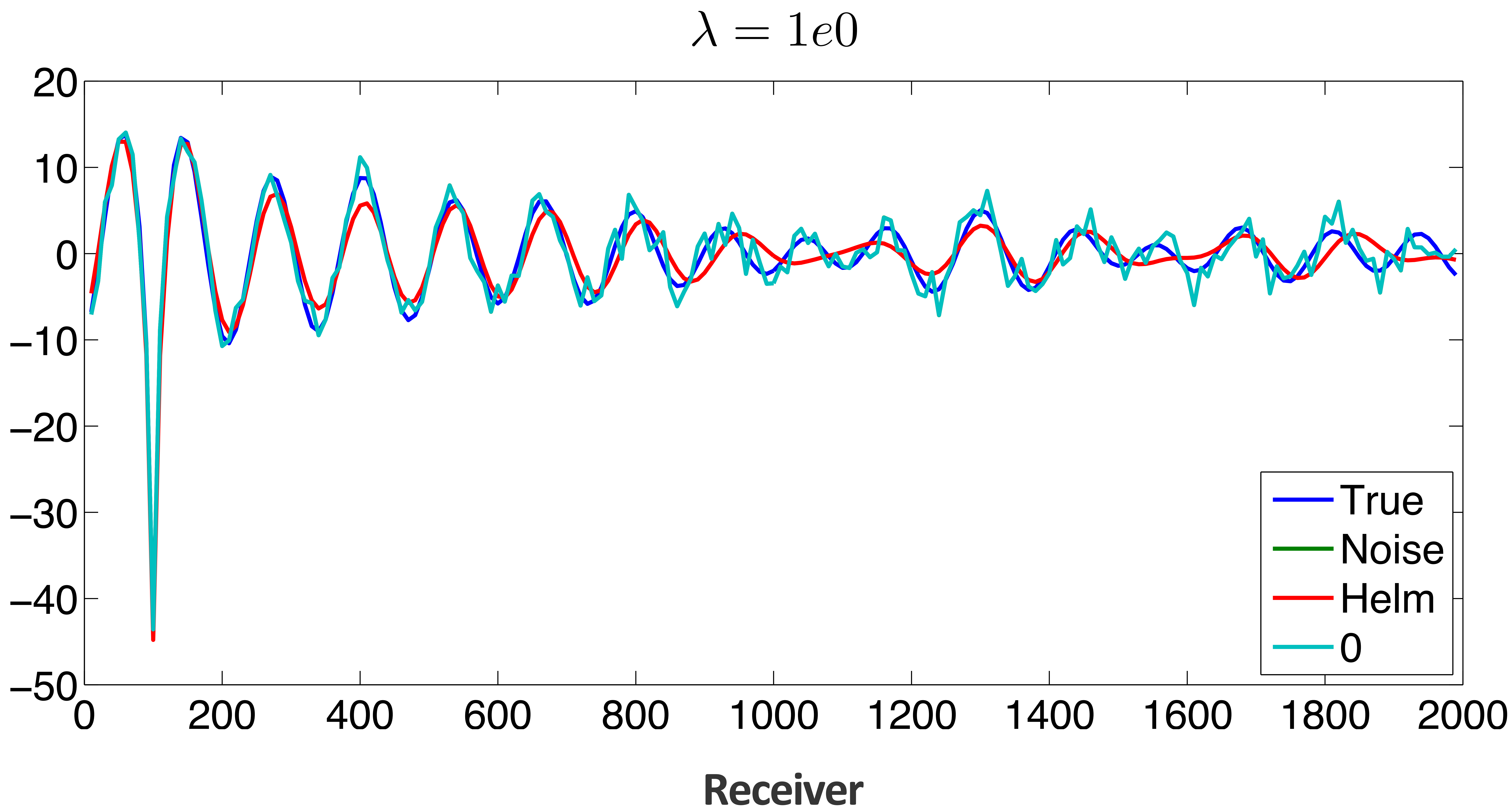


20
10
0
−10
−20
−30
−40
−50
0
200
400
600
800
1000
1200
1400
1600
1800
2000
True
Noise
Receiver

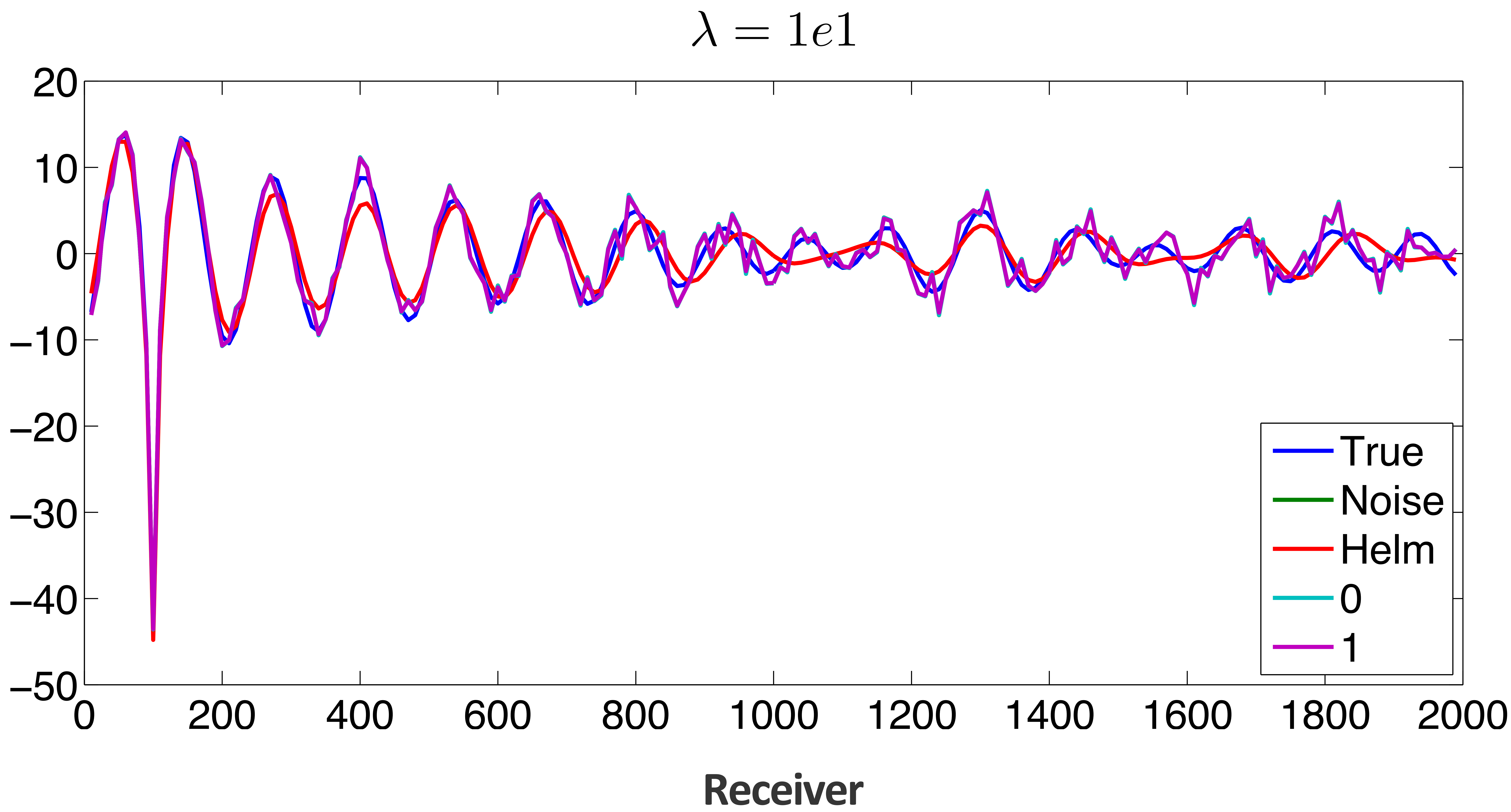
FWI data with initial model



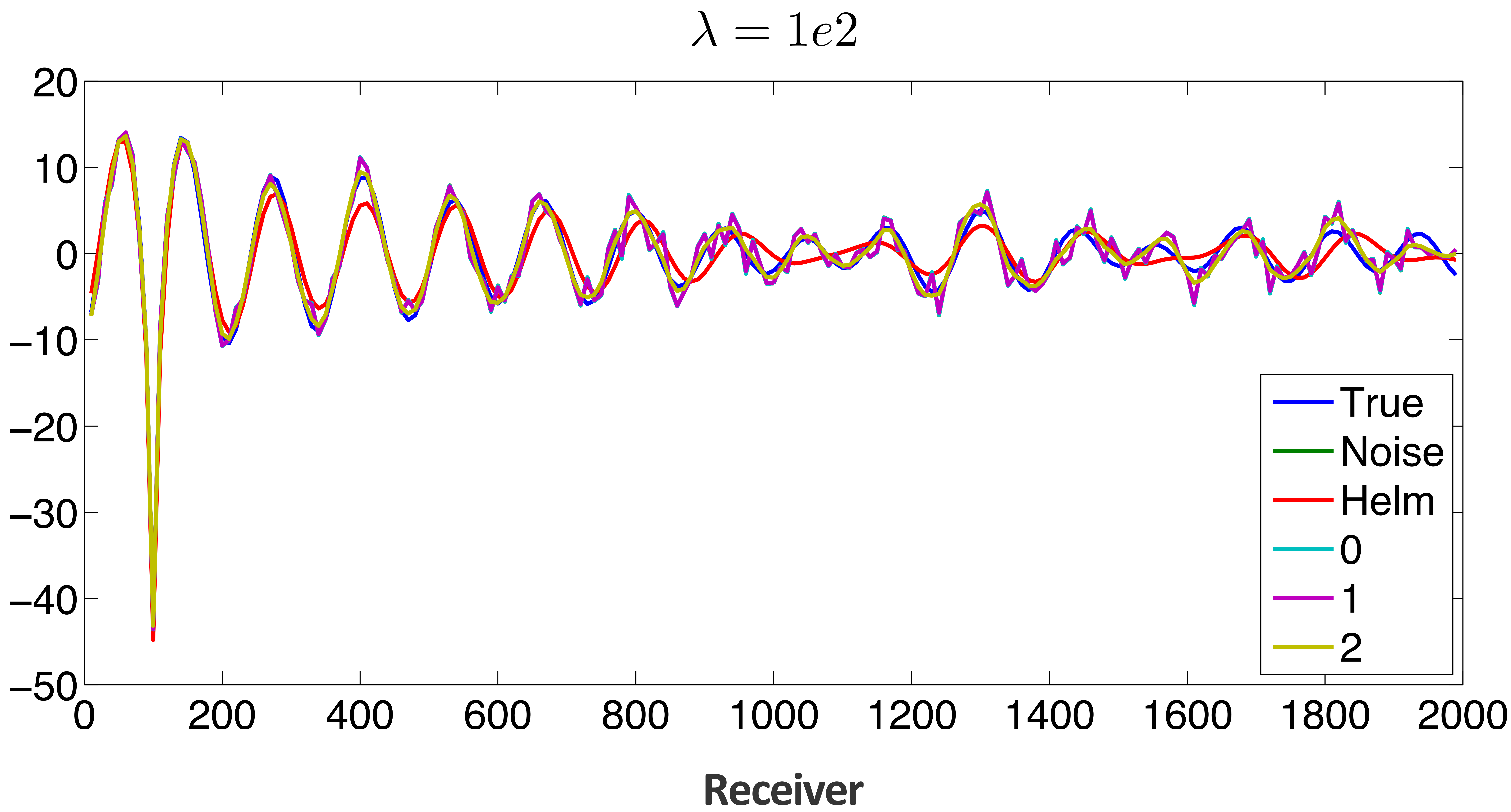
WRI data with initial model



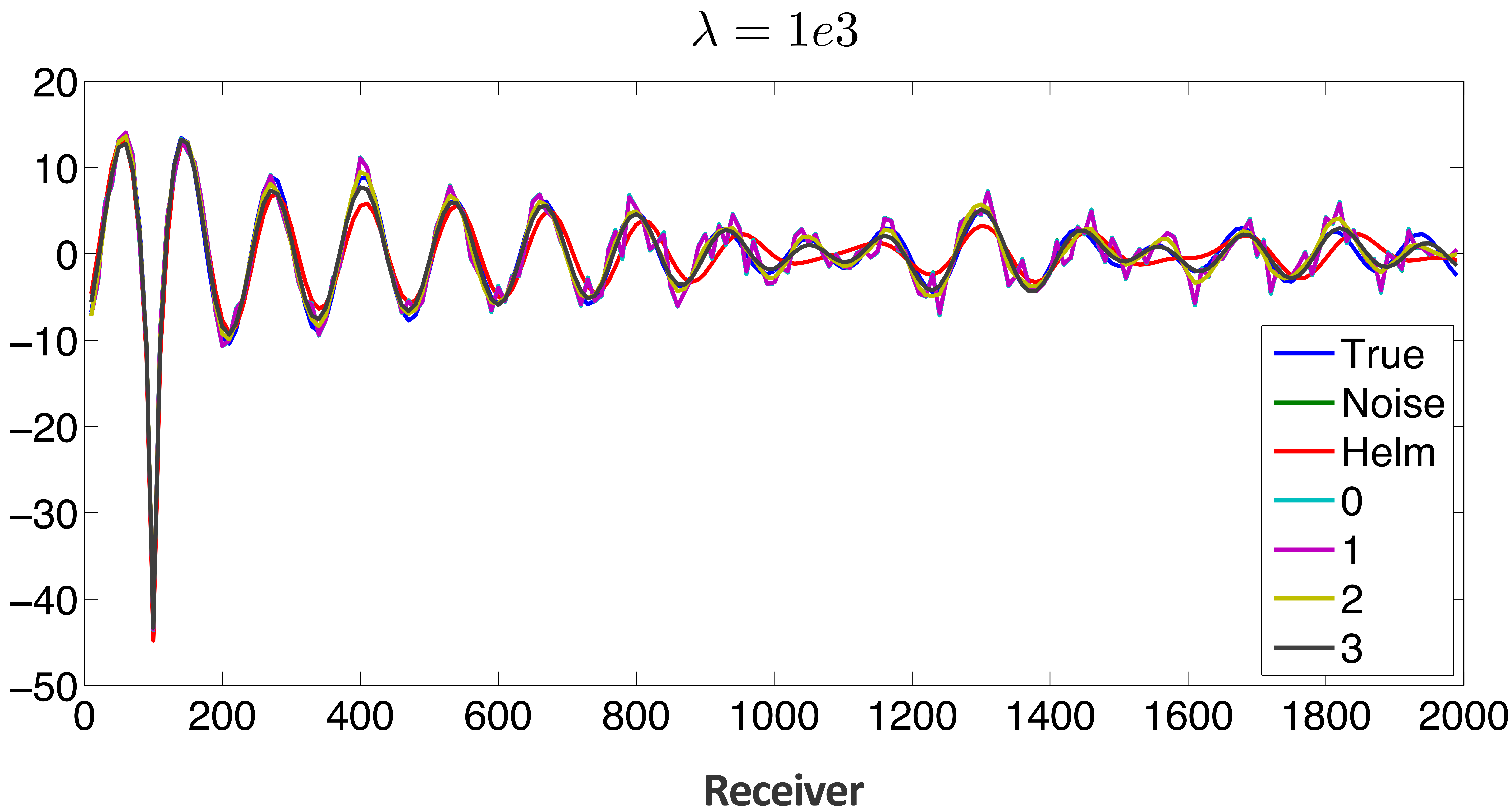
WRI data with initial model



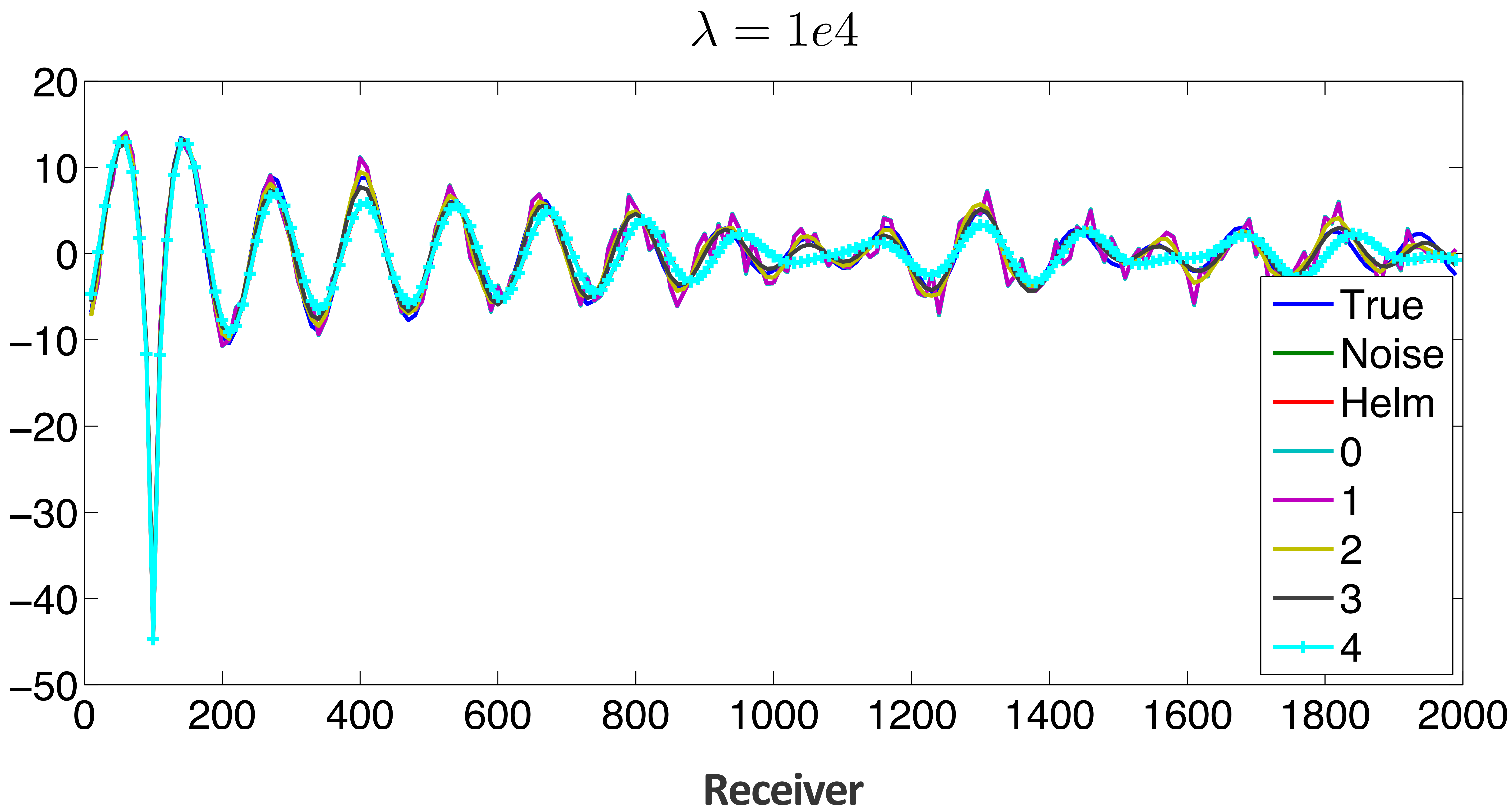
WRI data with initial model



WRI data with initial model



WRI data with initial model



Posterior Distribution of WRI

Marginal distribution of **m**:

$$\rho_{\text{post}}(\mathbf{m}) \propto \int \rho_{\text{post}}(\mathbf{m},\mathbf{u})\mathrm{d}\mathbf{u}$$

$$= (2\pi)^{N_{\text{pu}}/2}|\boldsymbol{\Sigma}_{\mathbf{u}}|^{1/2}\exp(-\|\mathbf{P}\overline{\mathbf{u}}(\mathbf{m})-\mathbf{d}_{\text{obs}}\|^2_{\boldsymbol{\Sigma}^{-1}_{\text{noise}}} -\lambda^2\|\mathbf{A}(\mathbf{m})\overline{\mathbf{u}}(\mathbf{m})-\mathbf{q}\|^2_{\boldsymbol{\Sigma}^{-1}_{\text{pde}}} - \|\mathbf{m}-\mathbf{m}_{\text{prior}}\|^2_{\boldsymbol{\Sigma}^{-1}_{\text{prior}}})$$

where

$$\overline{\mathbf{u}} = \arg\min_{\mathbf{u}} \| \begin{pmatrix} \lambda\boldsymbol{\Sigma}^{-1/2}_{\text{pde}}\mathbf{A} \\ \boldsymbol{\Sigma}^{-1/2}_{\text{noise}}\mathbf{P} \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda\boldsymbol{\Sigma}^{-1/2}_{\text{pde}}\mathbf{q} \\ \boldsymbol{\Sigma}^{-1/2}_{\text{noise}}\mathbf{d}_{\text{obs}} \end{pmatrix} \|^2$$

$$|\boldsymbol{\Sigma}_{\mathbf{u}}| = \det((\lambda^2\mathbf{A}^T\boldsymbol{\Sigma}^{-1}_{\text{pde}}\mathbf{A}+\mathbf{P}^T\boldsymbol{\Sigma}^{-1}_{\text{noise}}\mathbf{P})^{-1})$$

determine

Posterior Distribution of WRI

Posterior distribution w/ variable projection:

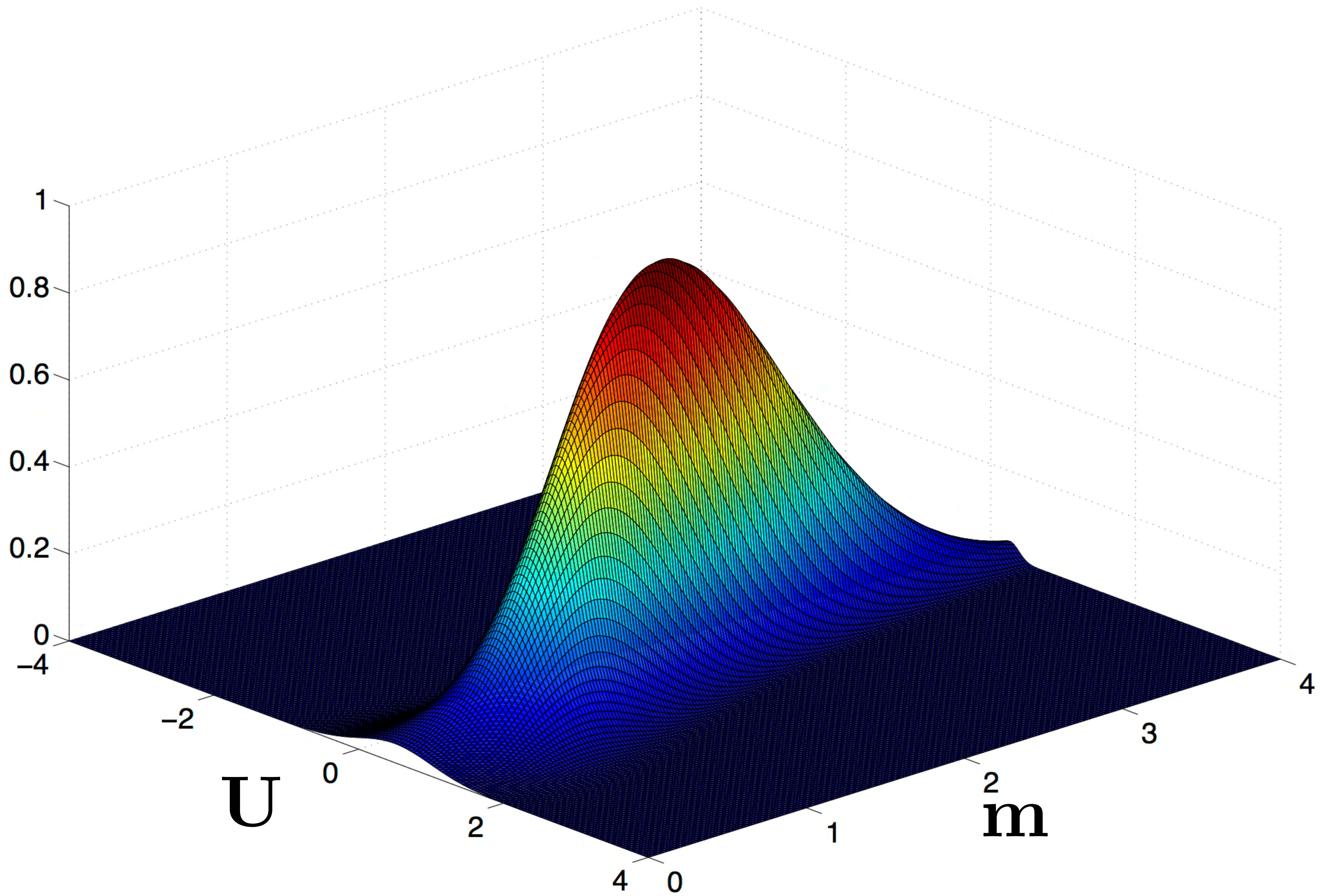
$$\rho_{\text{post}}(\mathbf{m}) \approx \rho_{\text{post}}(\mathbf{m}, \overline{\mathbf{u}}(\mathbf{m}))$$

$$\propto \exp\left(-\|\mathbf{P}\overline{\mathbf{u}}(\mathbf{m}) - \mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}} - \lambda^2\|\mathbf{A}(\mathbf{m})\overline{\mathbf{u}}(\mathbf{m}) - \mathbf{q}\|^2_{\mathbf{\Sigma}^{-1}_{\text{pde}}} - \|\mathbf{m} - \mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}}\right)$$

where

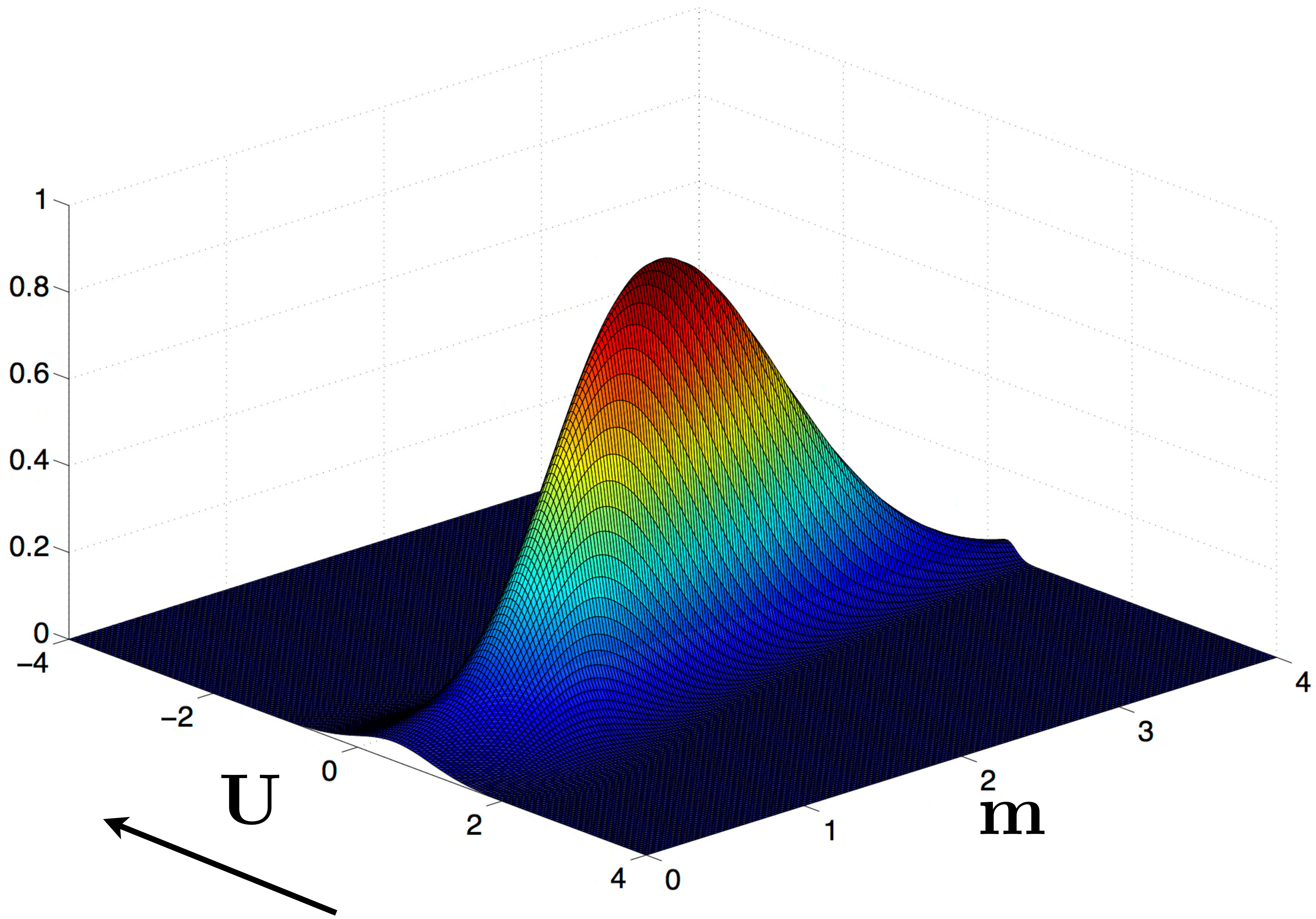
$$\overline{\mathbf{u}} = \arg\min_{\mathbf{u}} \| \begin{pmatrix} \lambda\mathbf{\Sigma}^{-1/2}_{\text{pde}}\mathbf{A} \\ \mathbf{\Sigma}^{-1/2}_{\text{noise}}\mathbf{P} \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda\mathbf{\Sigma}^{-1/2}_{\text{pde}}\mathbf{q} \\ \mathbf{\Sigma}^{-1/2}_{\text{noise}}\mathbf{d}_{\text{obs}} \end{pmatrix} \|^2$$

Marginal distribution vs Var-Prj distribution



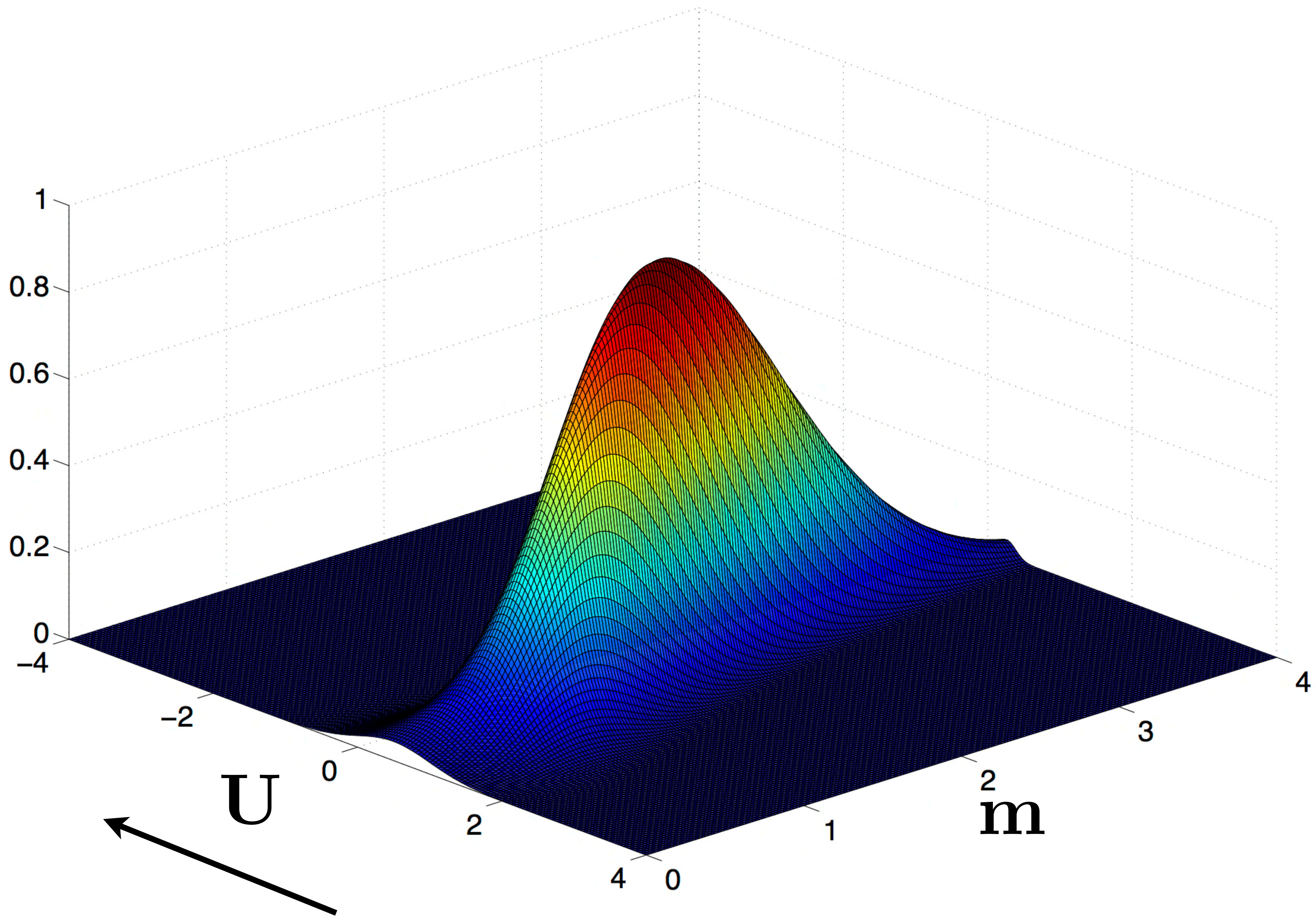
Joint Distribution

Marginal distribution vs Var-Prj distribution



Joint Distribution

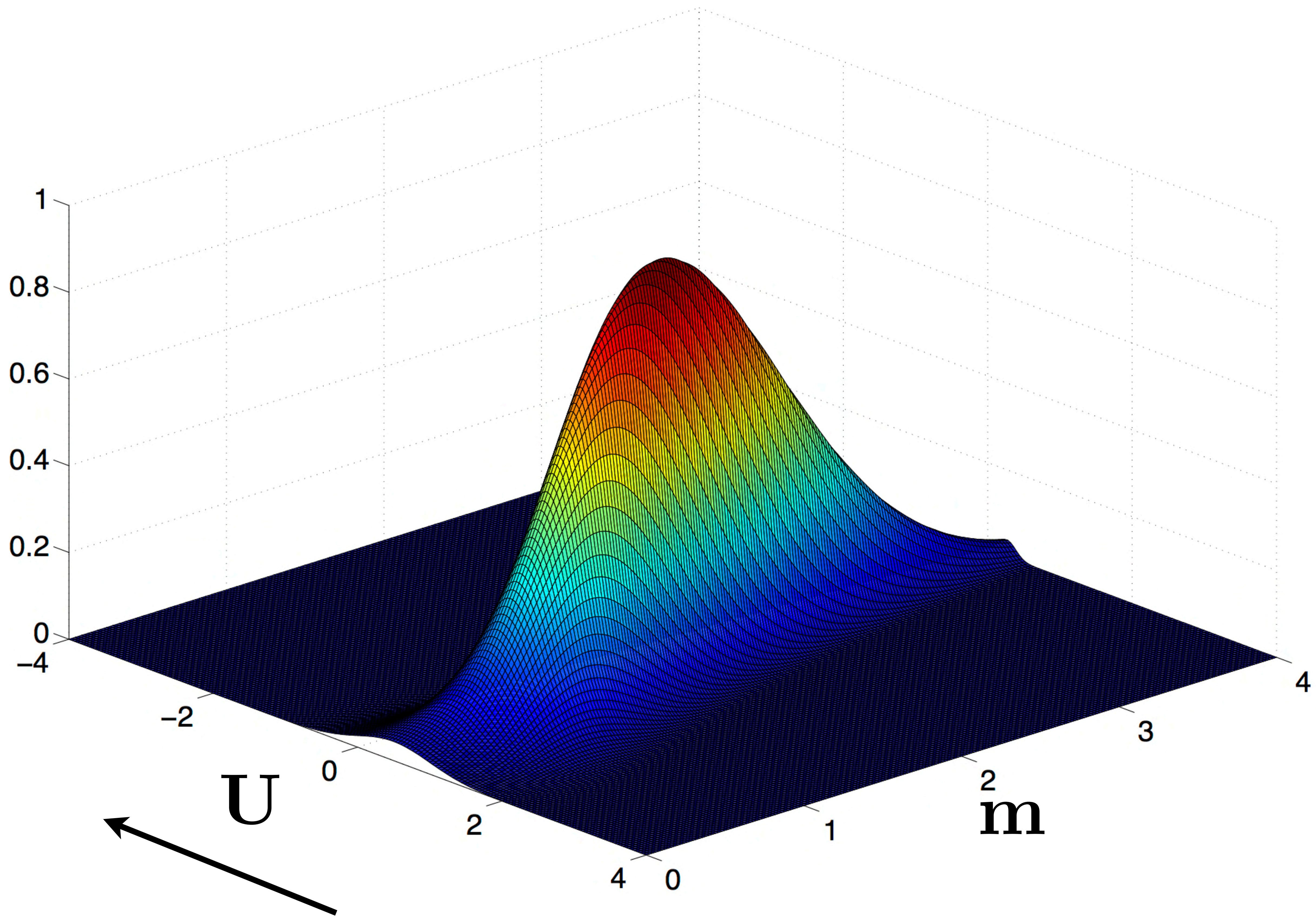
Marginal distribution vs Var-Prj distribution



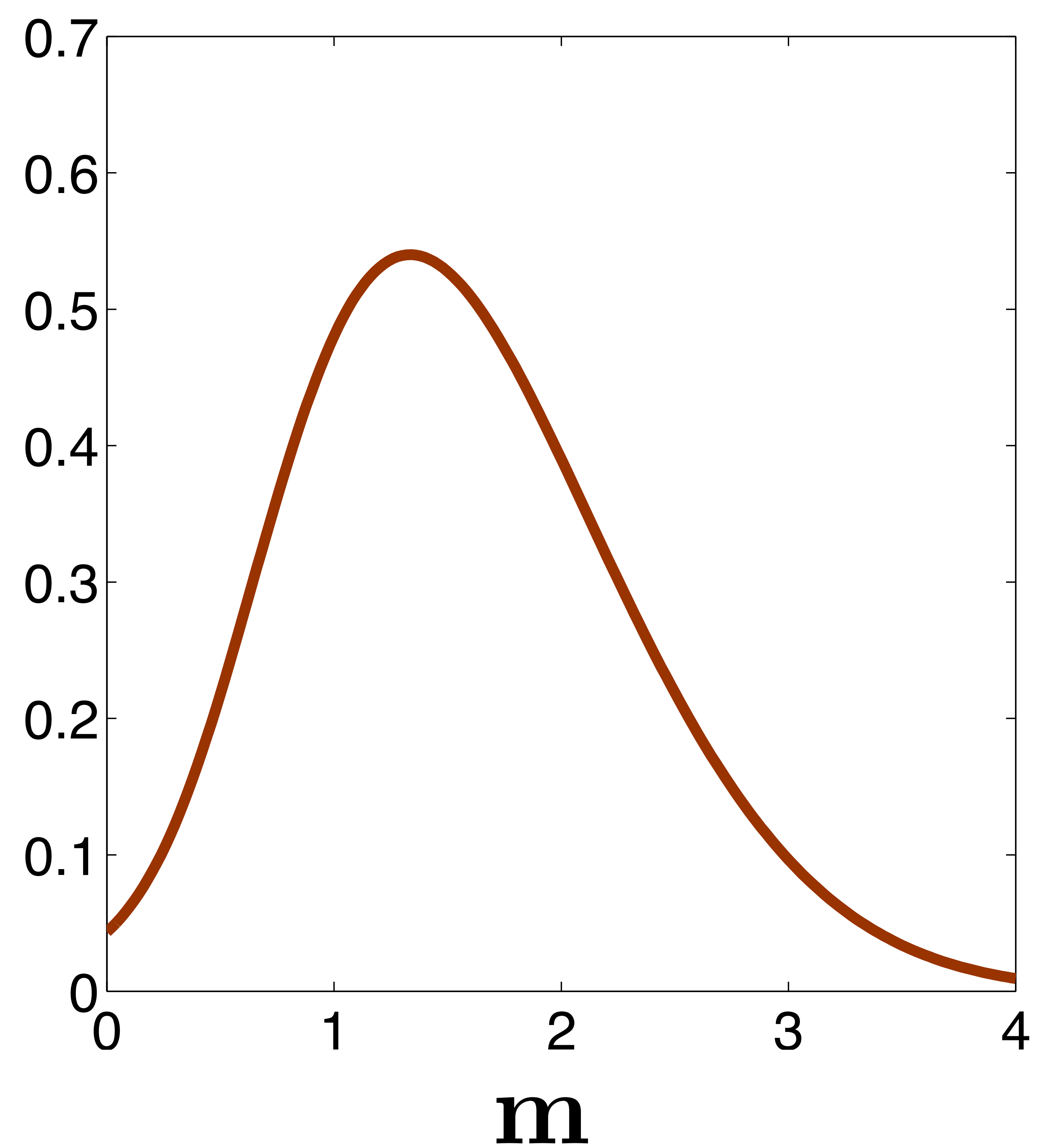
Joint Distribution

Marginal Distribution vs Var-Prj Distribution

Marginal distribution vs Var-Prj distribution

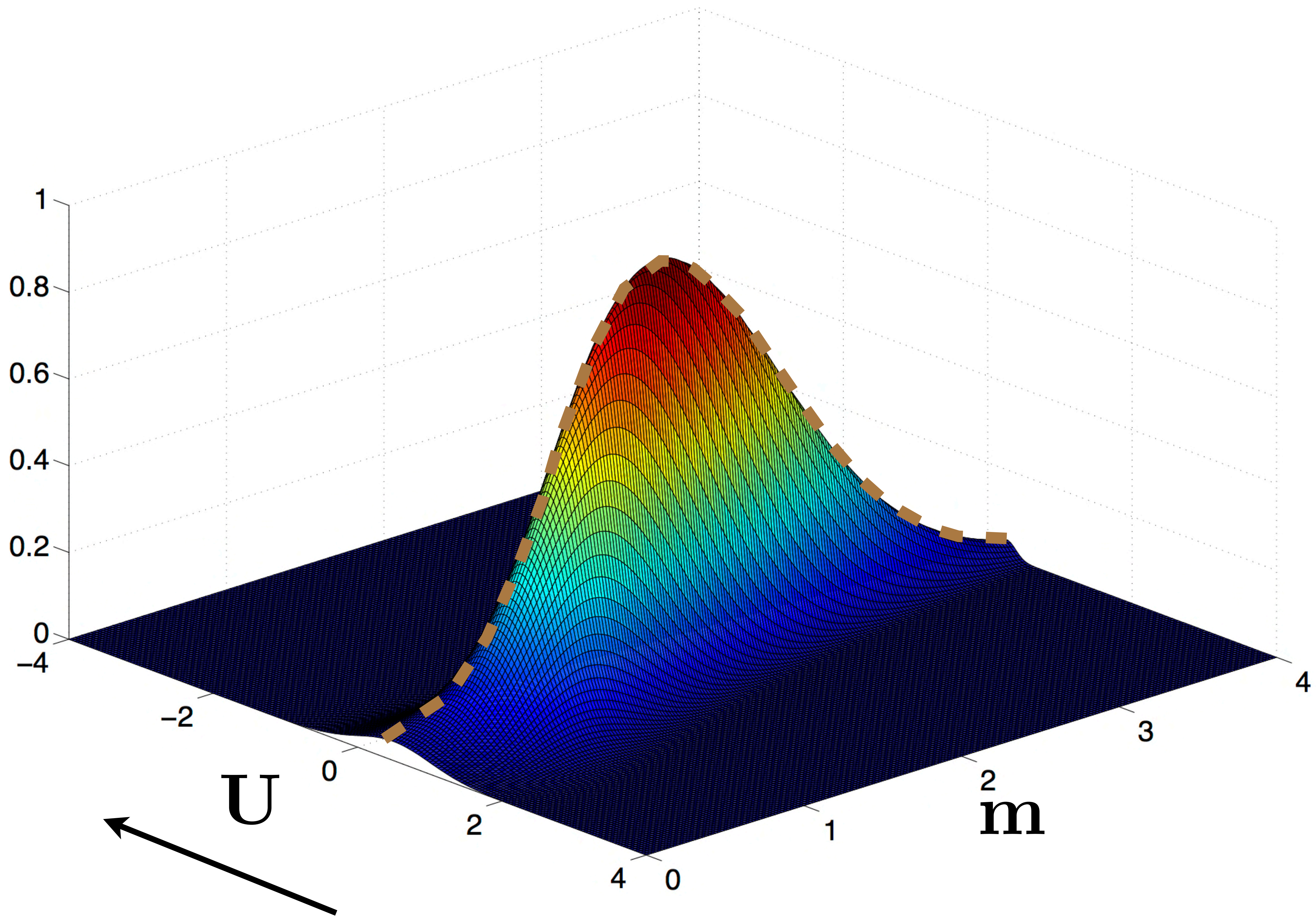


Joint Distribution

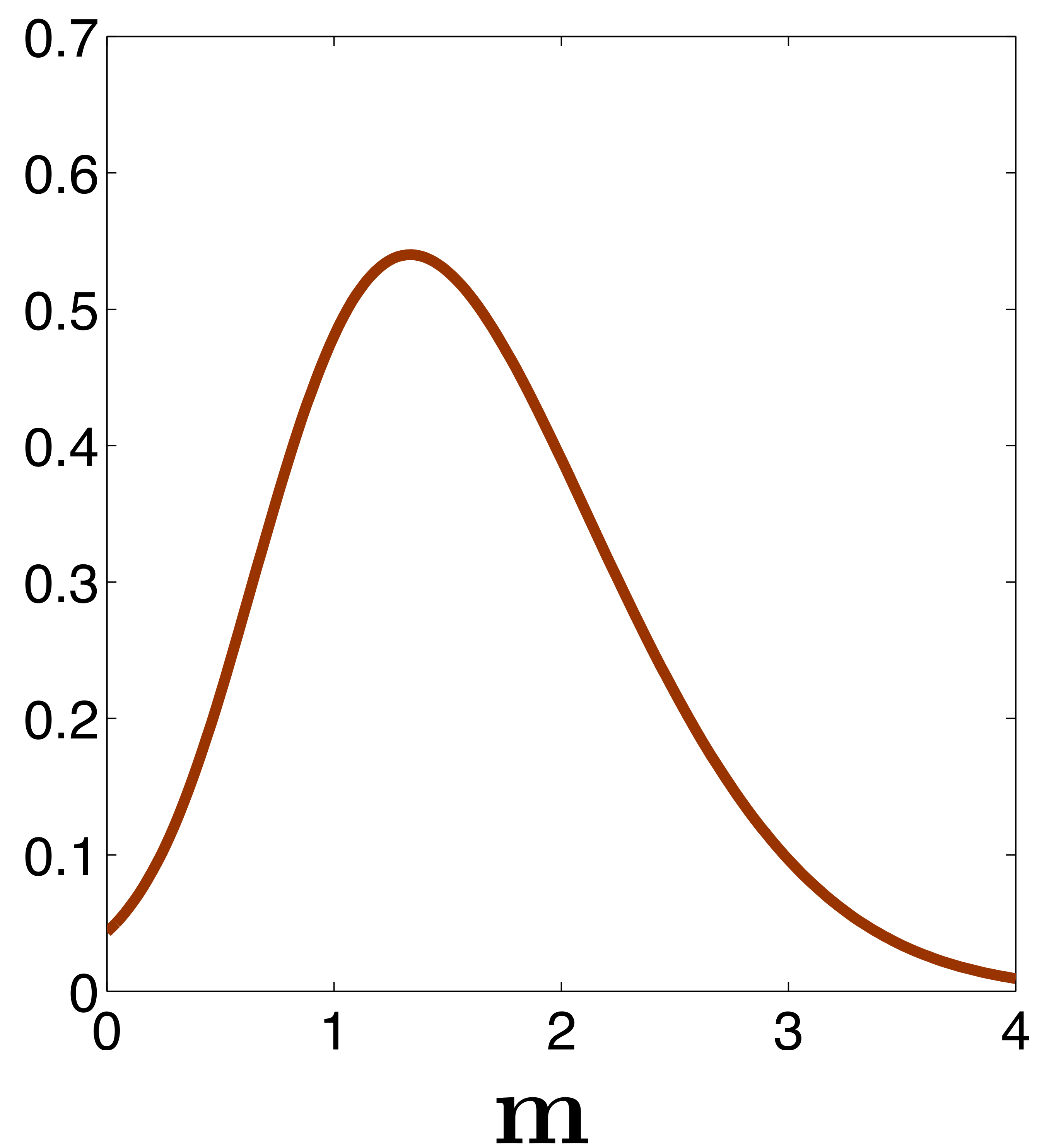


Marginal Distribution
vs Var-Prj Distribution

Marginal distribution vs Var-Prj distribution

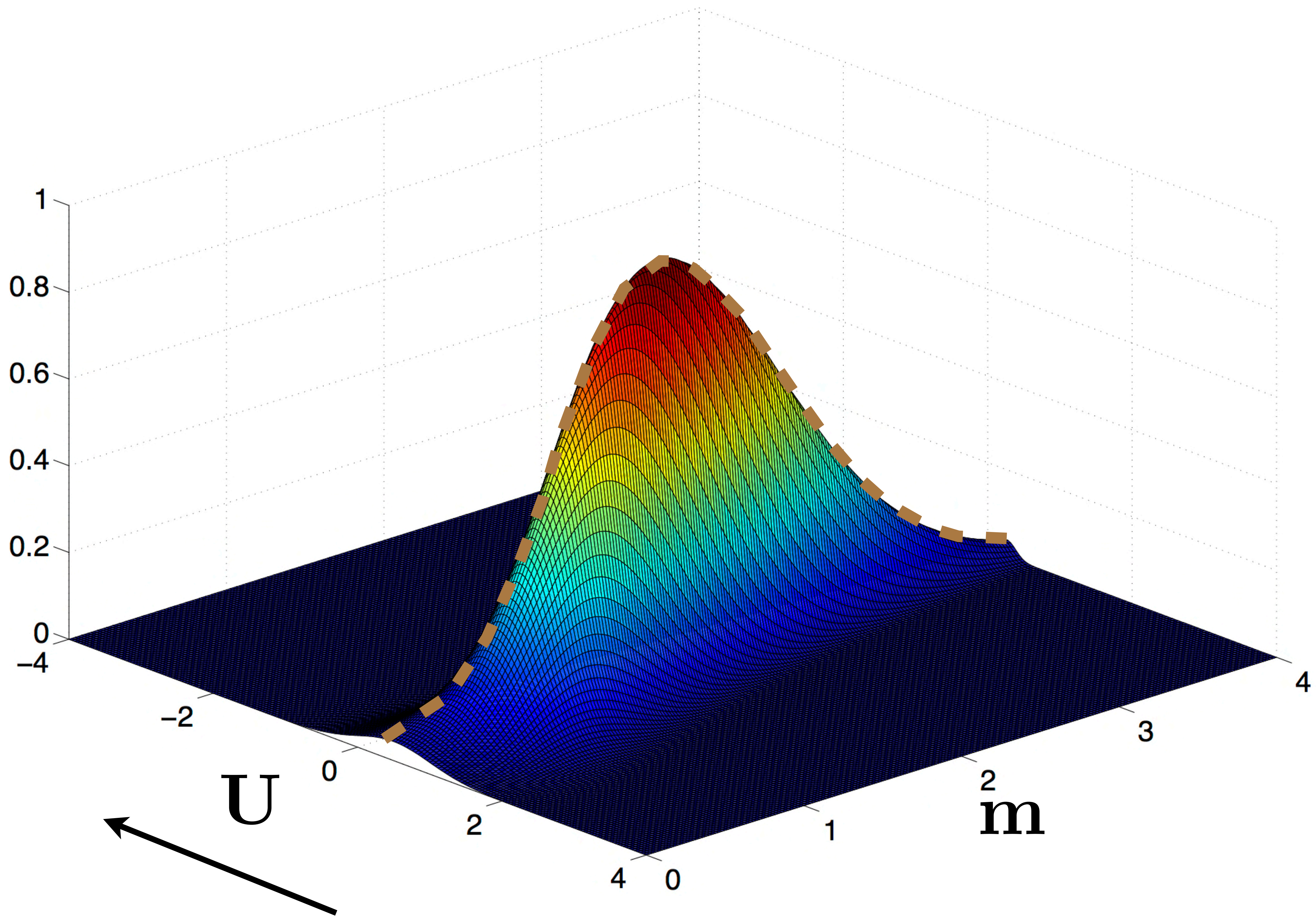


Joint Distribution

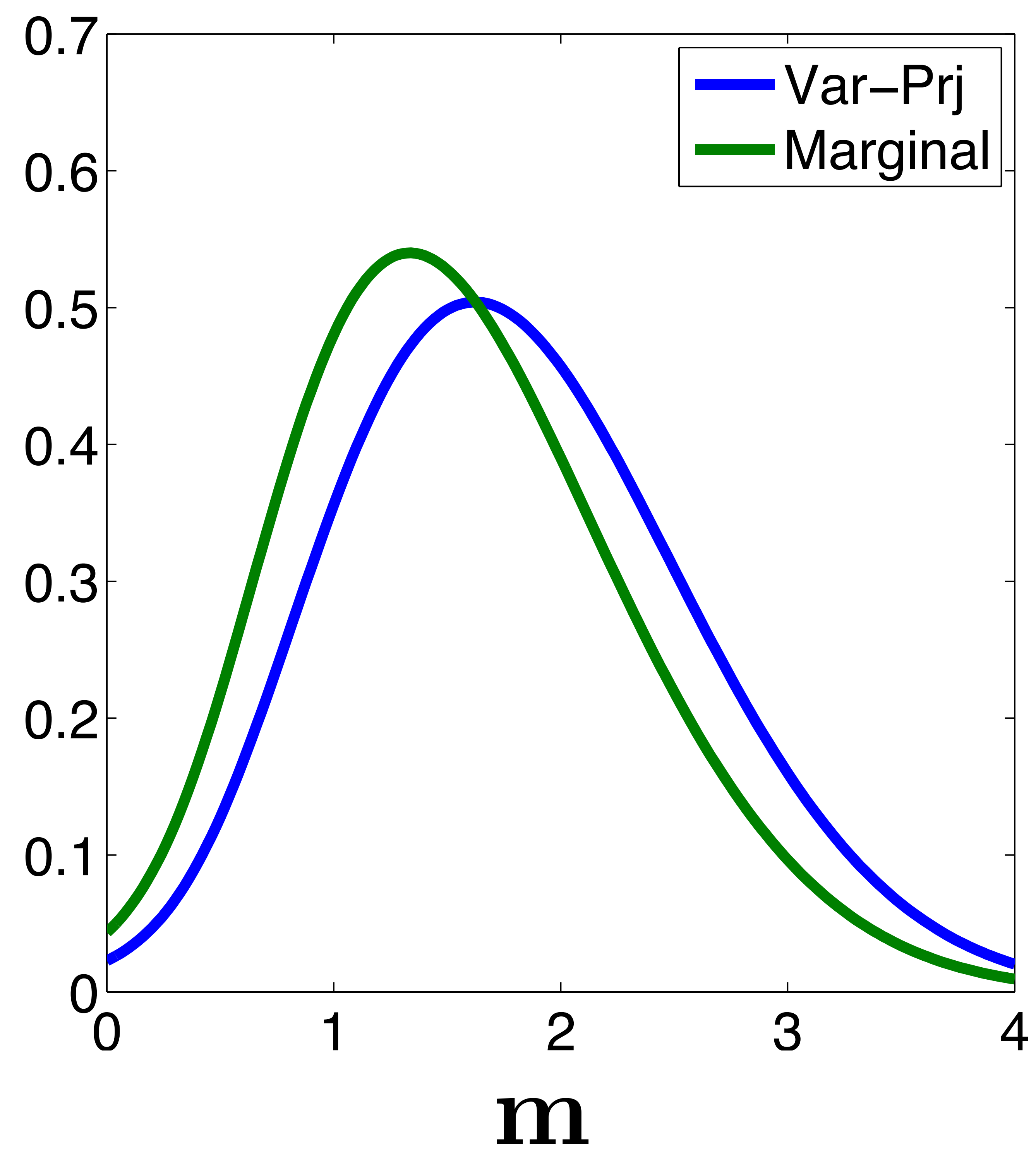


Marginal Distribution
vs Var-Prj Distribution

Marginal distribution vs Var-Prj distribution



Joint Distribution



Marginal Distribution vs Var-Prj Distribution

Quantify the uncertainty

Goal : Quantify the Uncertainty based on the posterior distribution $\rho_{\text{post}}(\mathbf{m})$

Solution:

- Integrate the posterior distribution.

Quantify the uncertainty

Goal : Quantify the Uncertainty based on the posterior distribution $\rho_{\text{post}}(\mathbf{m})$

Solution:

- Integrate the posterior distribution. Huge computational cost!!!

Quantify the uncertainty

Goal : Quantify the Uncertainty based on the posterior distribution $\rho_{\rm post}(\mathbf{m})$

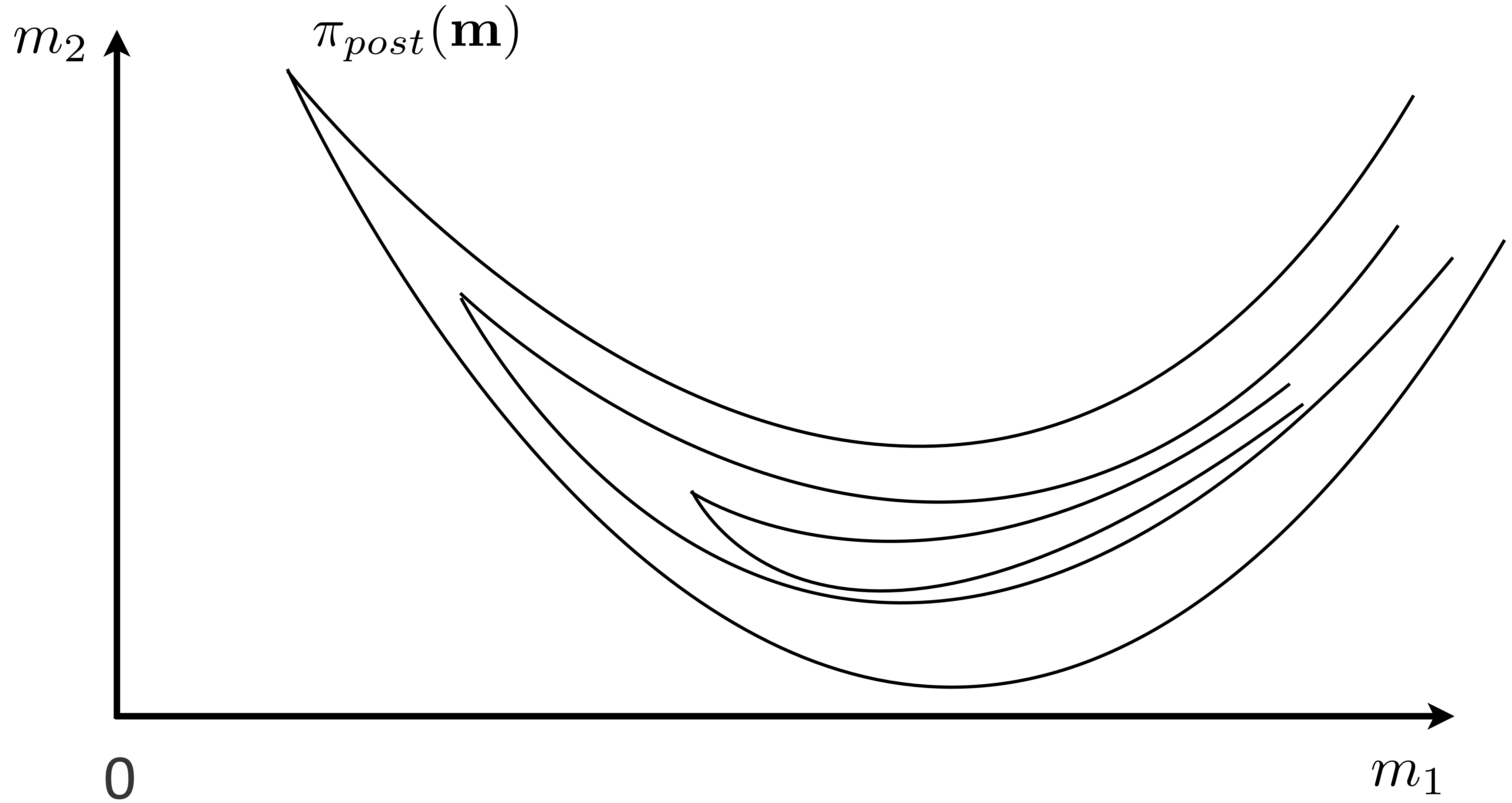
Solution:

- Integrate the posterior distribution.
- Use McMC method to sample the posterior distribution.

[James Martin *et al*, 2012]

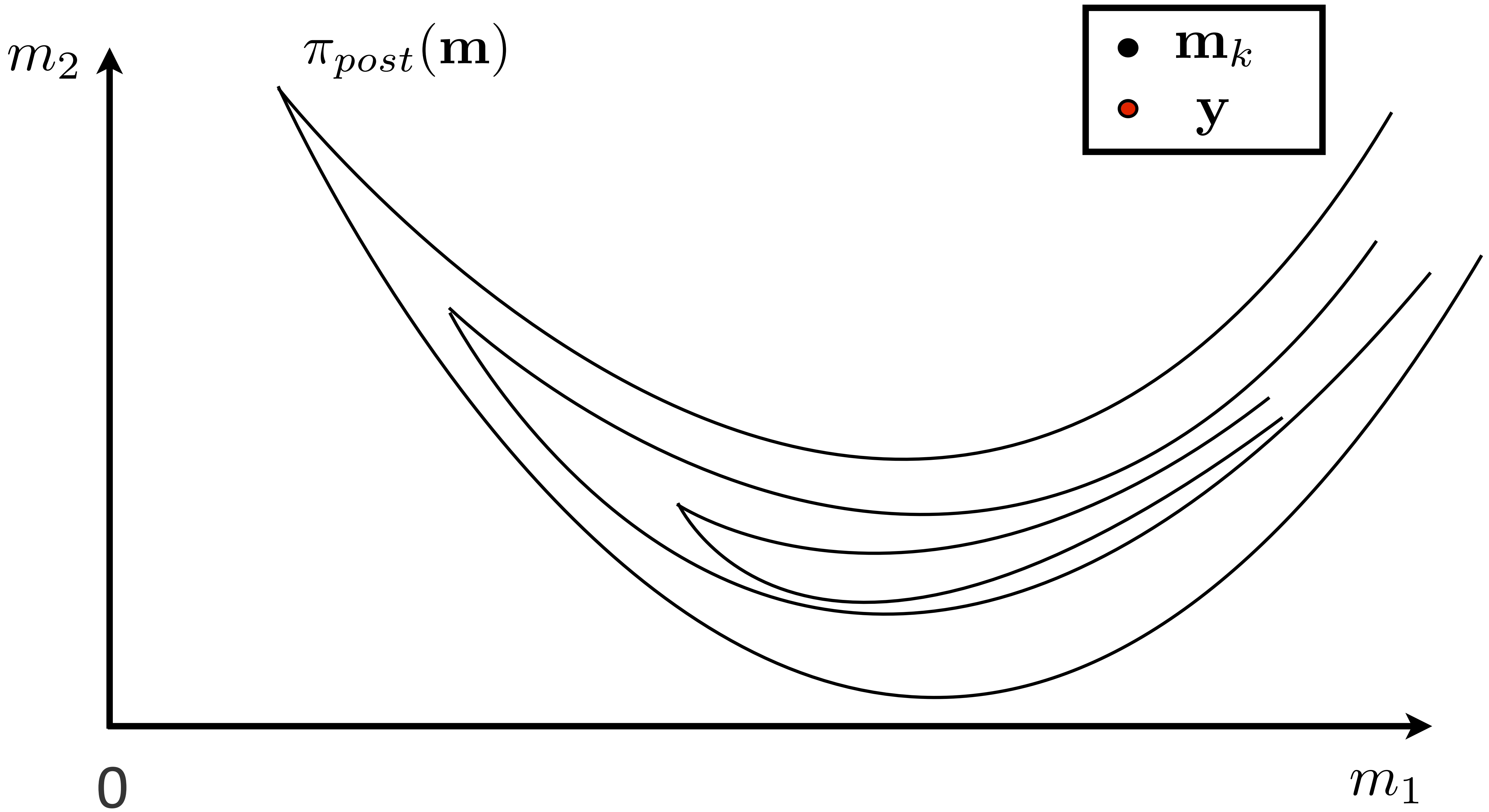
McMC method

Newton Type McMC: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{H}_k^{-1}\mathbf{g}_k, \mathbf{H}_k^{-1})$



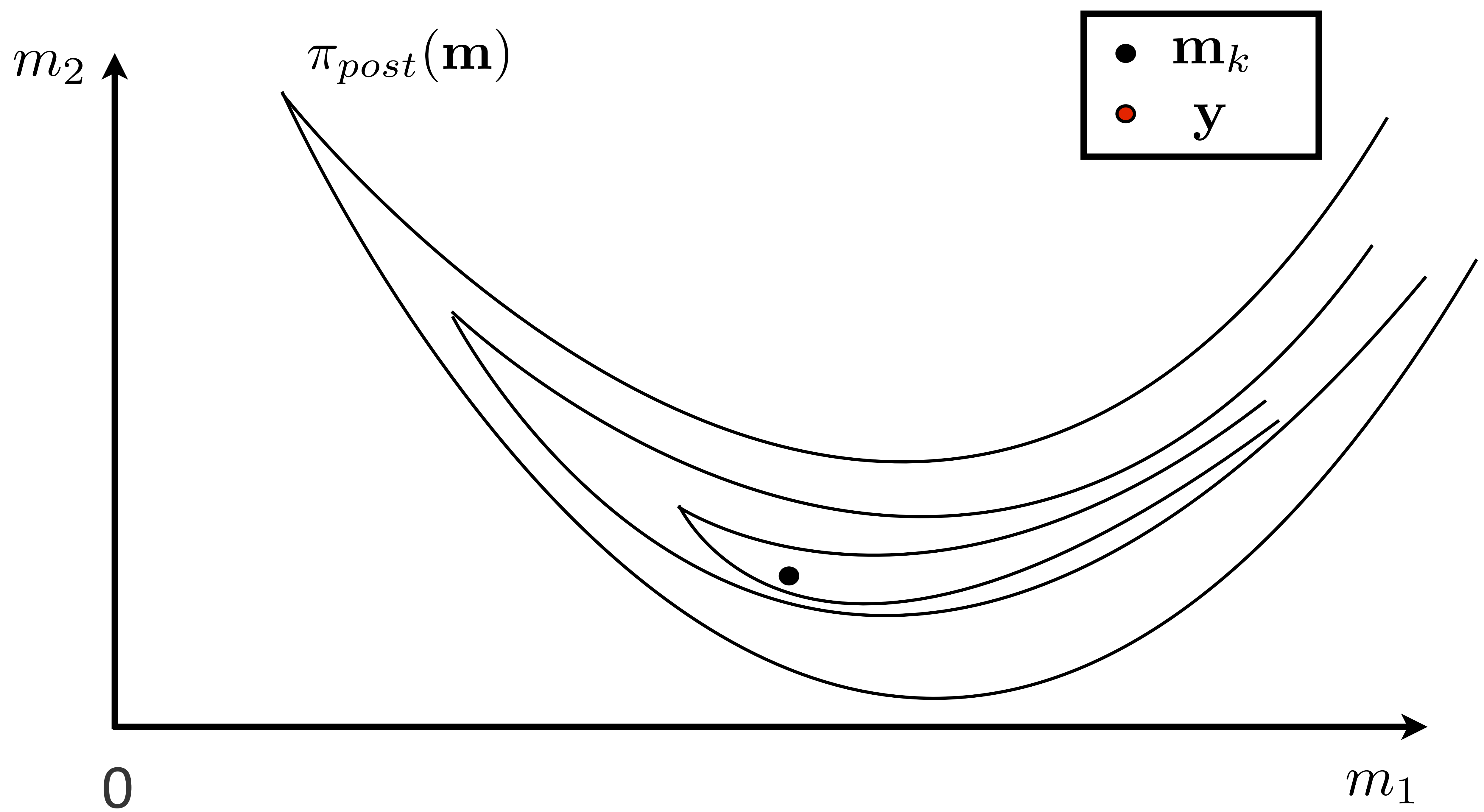
McMC method

Newton Type McMC: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{H}_k^{-1}\mathbf{g}_k, \mathbf{H}_k^{-1})$



McMC method

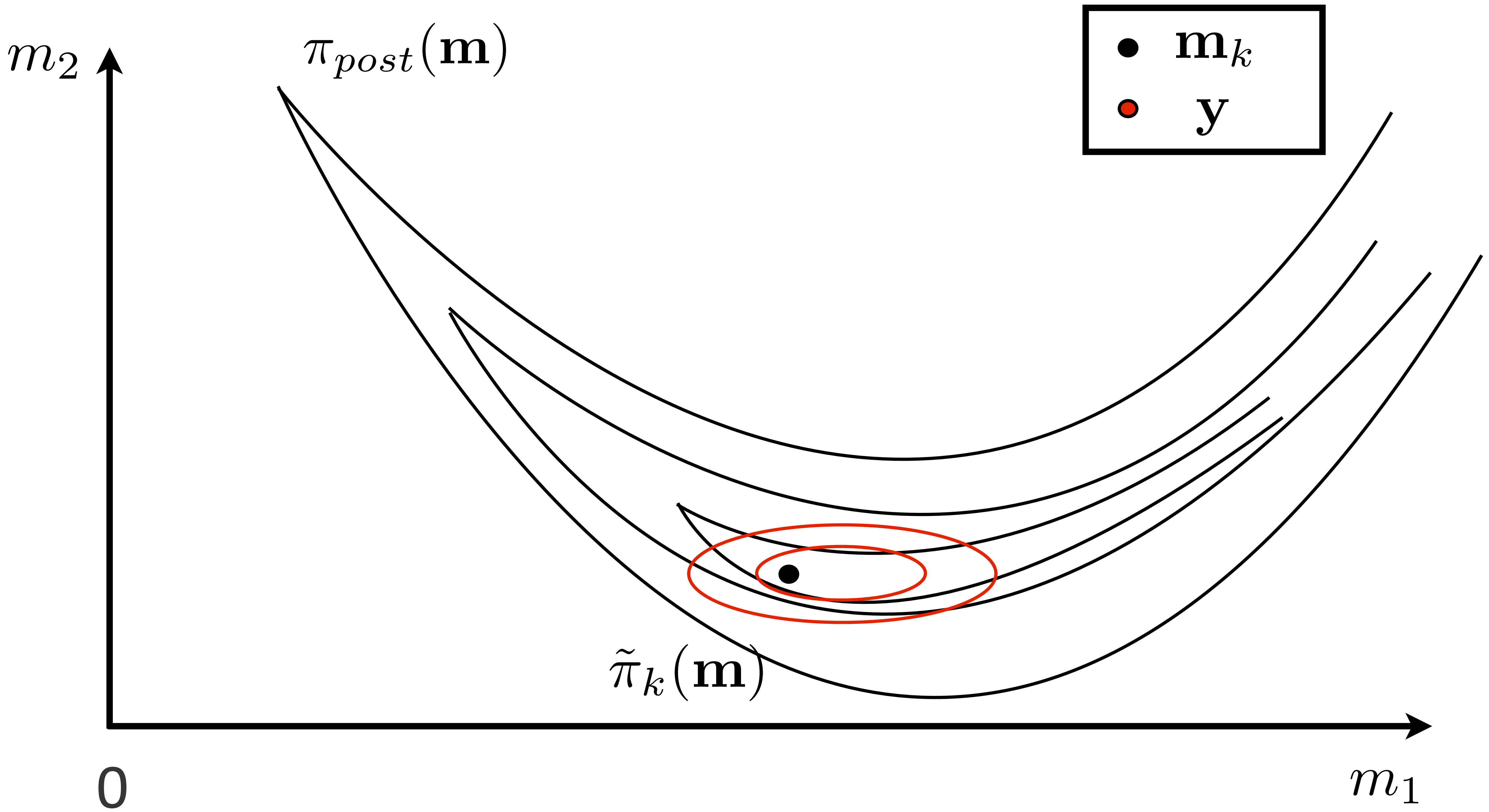
Newton Type McMC: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{H}_k^{-1}\mathbf{g}_k, \mathbf{H}_k^{-1})$



[James Martin *et al*, 2012]

McMC method

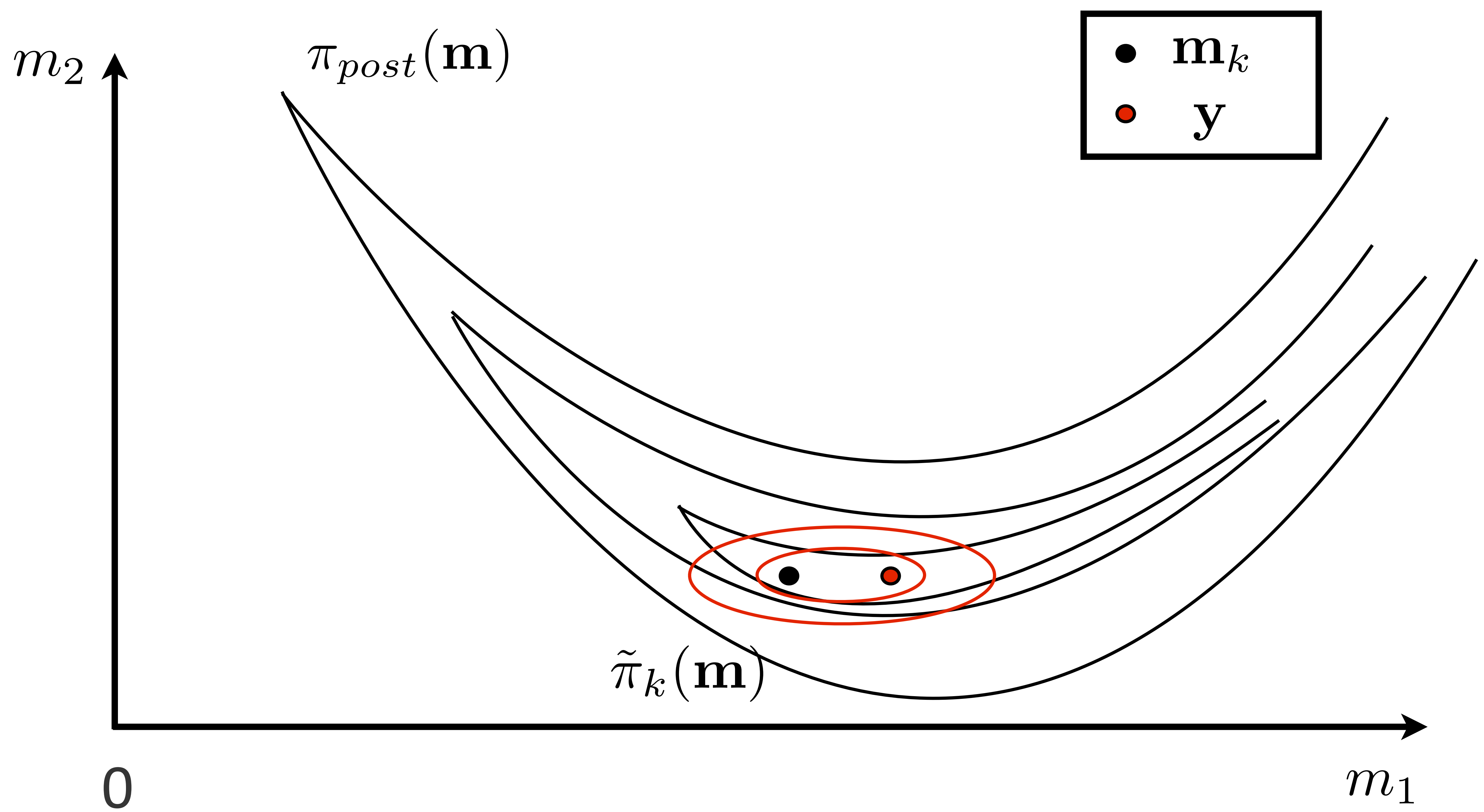
Newton Type McMC: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{H}_k^{-1}\mathbf{g}_k, \mathbf{H}_k^{-1})$



[James Martin *et al*, 2012]

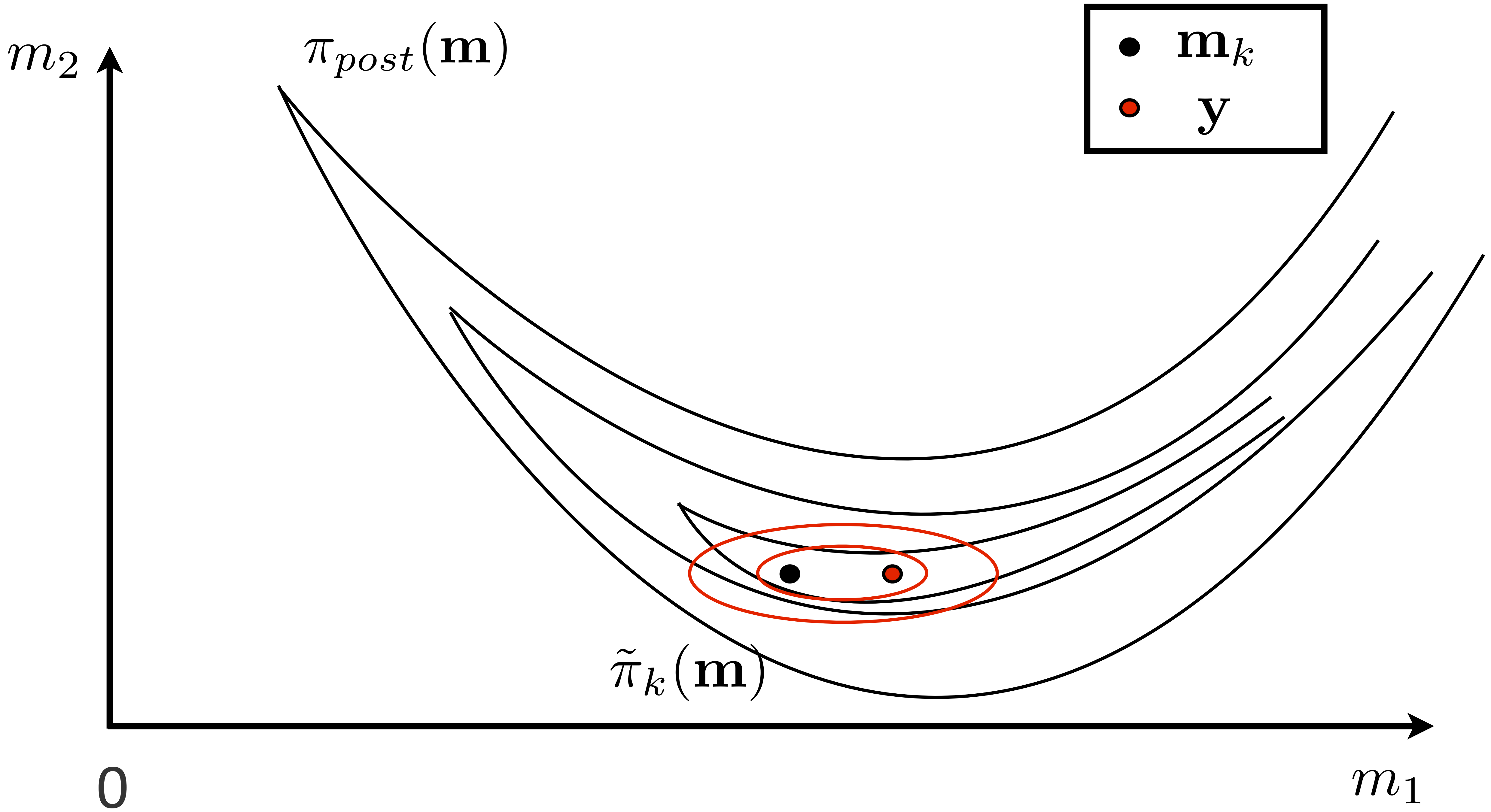
McMC method

Newton Type McMC: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{H}_k^{-1}\mathbf{g}_k, \mathbf{H}_k^{-1})$



McMC method

Newton Type McMC: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{H}_k^{-1}\mathbf{g}_k, \mathbf{H}_k^{-1})$



Computational cost:

1. **Low-rank approximation of the Hessian.**
2. **Number of PDE solvers ~ Number of samples.**

Quantify the uncertainty

Goal : Quantify the Uncertainty based on the posterior distribution $\rho_{\mathrm{post}}(\mathbf{m})$

Solution:

- Integrate the posterior distribution.

- McMC method to sample the posterior distribution.
  - Advantage: the true uncertainty can be quantified.
  - Disadvantage: Huge computational cost.

Quantify the uncertainty

Goal : Quantify the Uncertainty based on the posterior distribution $\rho_{\text{post}}(\mathbf{m})$

Solution:

- Integrate the posterior distribution.

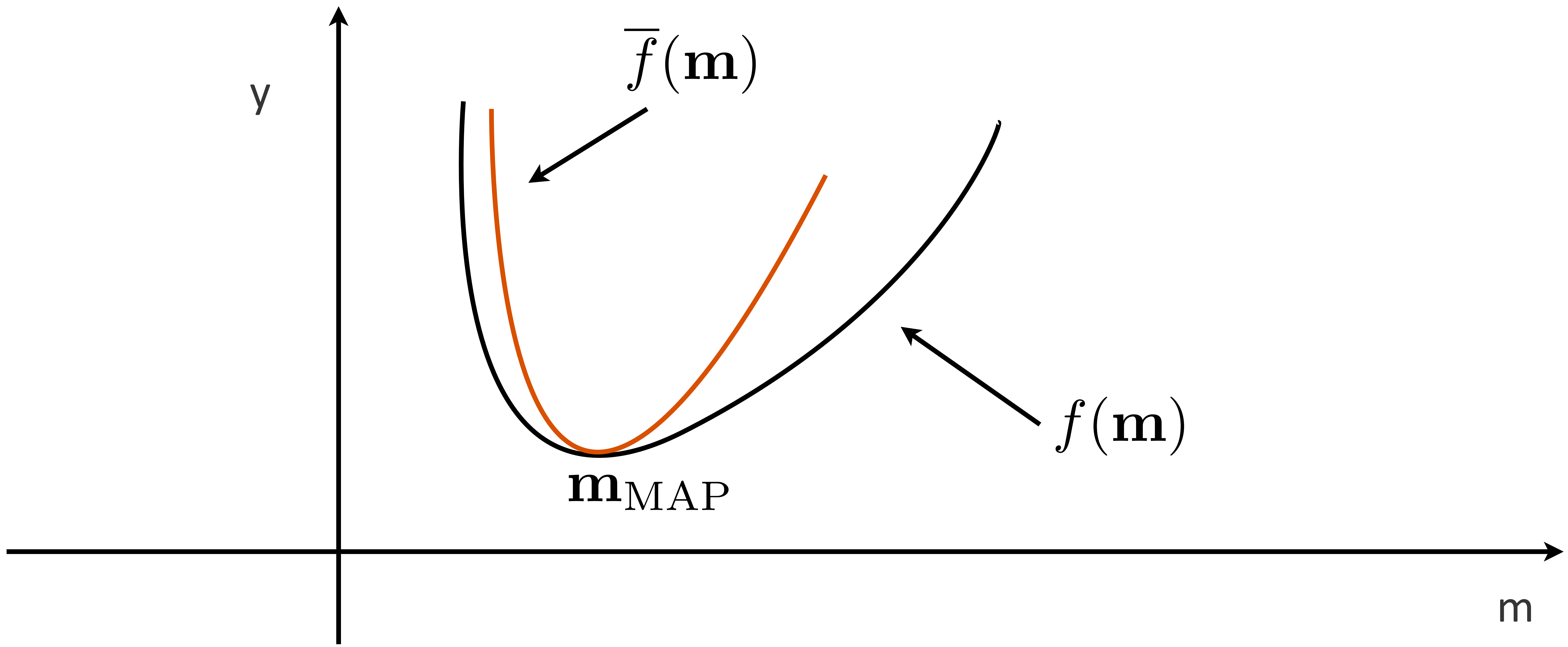
- McMC method to sample the posterior distribution.
  - Advantage: the true uncertainty can be quantified.
  - Disadvantage: Huge computational cost.

- Use an approximated distribution to quantify the uncertainty.

Quadratic approximation

$$f(\mathbf{m}) = \|\mathbf{P}\mathbf{u}(\mathbf{m}) - \mathbf{d}_{\mathrm{obs}}\|^2_{\mathbf{\Sigma}^{-1}_{\mathrm{noise}}} + \lambda^2\|\mathbf{A}(\mathbf{m})\mathbf{u}(\mathbf{m}) - \mathbf{q}\|^2_{\mathbf{\Sigma}^{-1}_{\mathrm{pde}}} + \|\mathbf{m} - \mathbf{m}_{\mathrm{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\mathrm{prior}}}$$

$$\approx f(\mathbf{m}_{\mathrm{MAP}}) + \mathbf{g}^T(\mathbf{m} - \mathbf{m}_{\mathrm{MAP}}) + \frac{1}{2}(\mathbf{m} - \mathbf{m}_{\mathrm{MAP}})^T\mathbf{H}(\mathbf{m} - \mathbf{m}_{\mathrm{MAP}})$$



Hessian of FWI

[R. Gerhard Pratt, Changsoo Shin and G.J. Hicks, 1998]

Full Hessian of FWI:

$$\mathbf{H} = \mathbf{H_{GN}} + \mathbf{H_2}$$

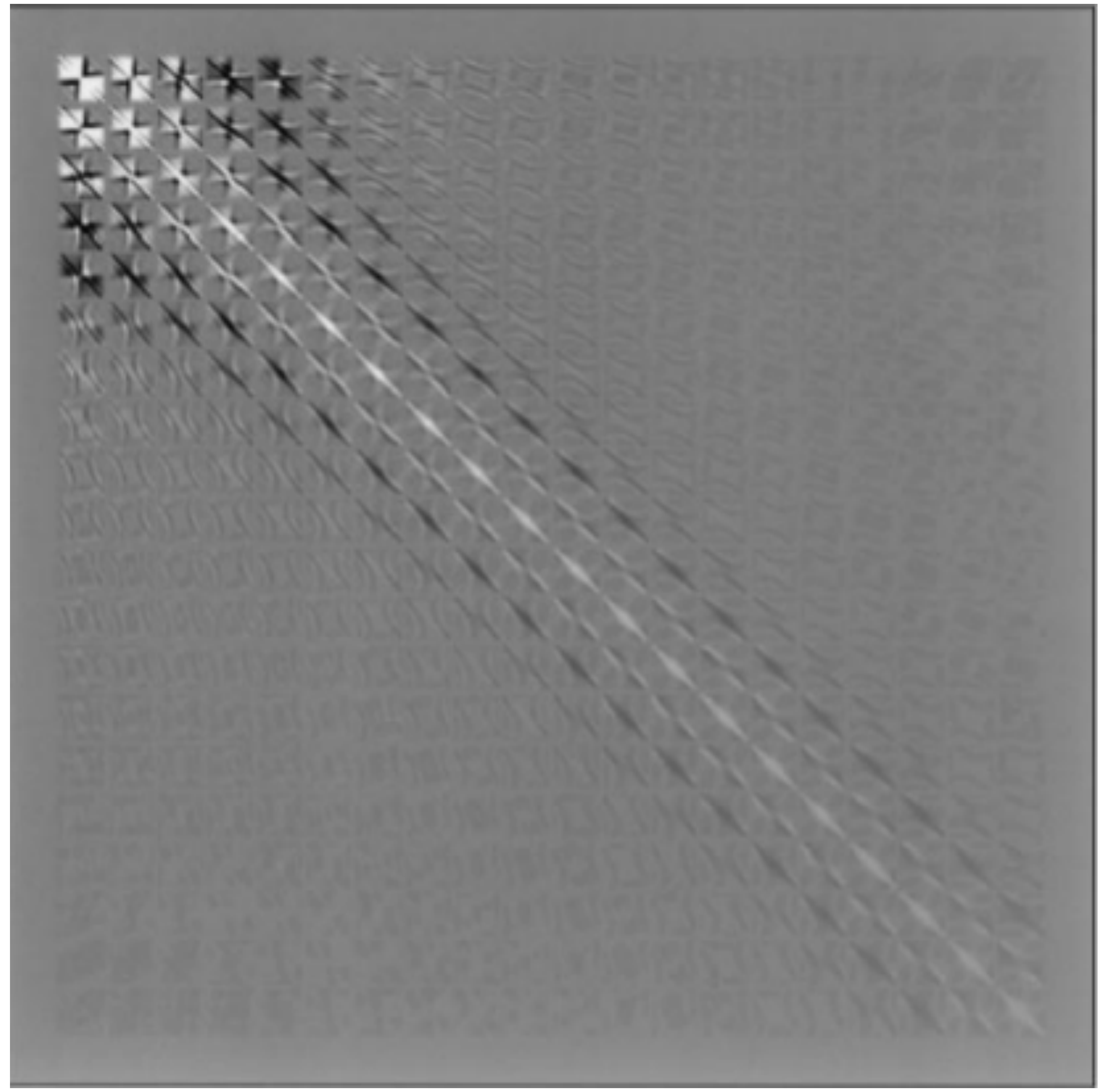
Gauss-Newton Hessian of FWI:

$$\mathbf{H_{GN}} = \mathbf{G^T A^{-T} P^T P A^{-1} G}$$

where

$$\mathbf{G} = \frac{\partial \mathbf{A(m)u}}{\partial \mathbf{m}}$$

dense



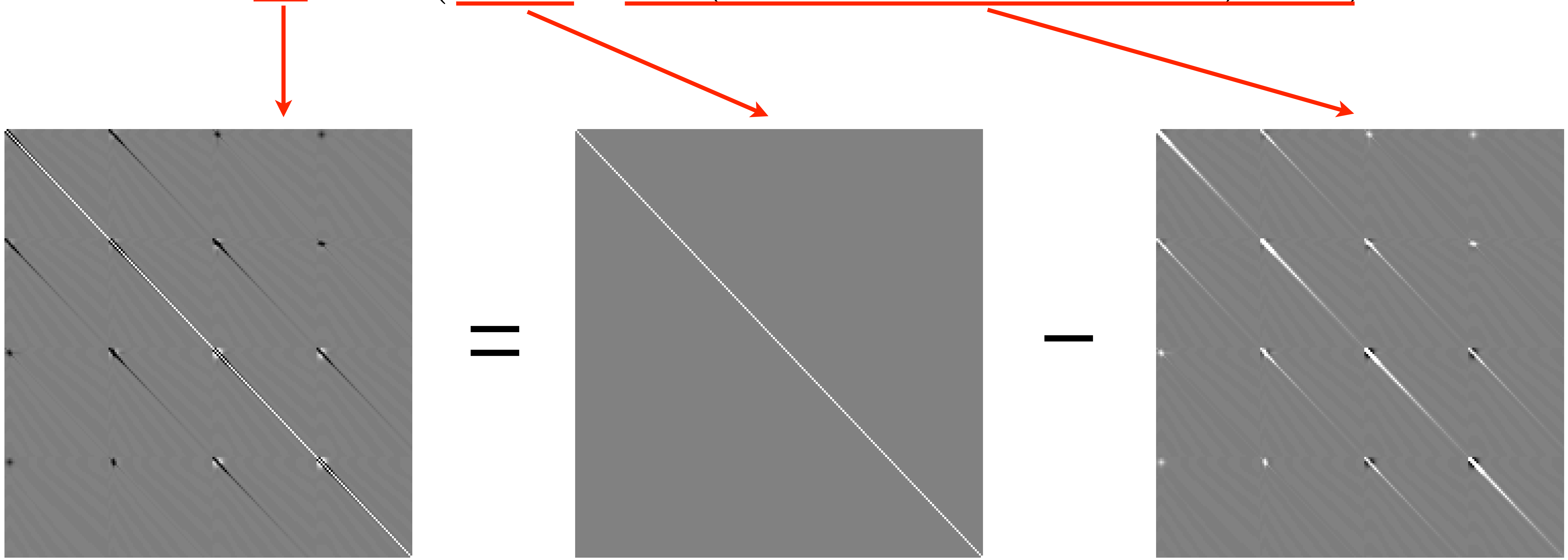
Gauss-Newton Hessian of FWI

Hessian of WRI

[van Leeuwen, T and Herrmann, F J , 2013]

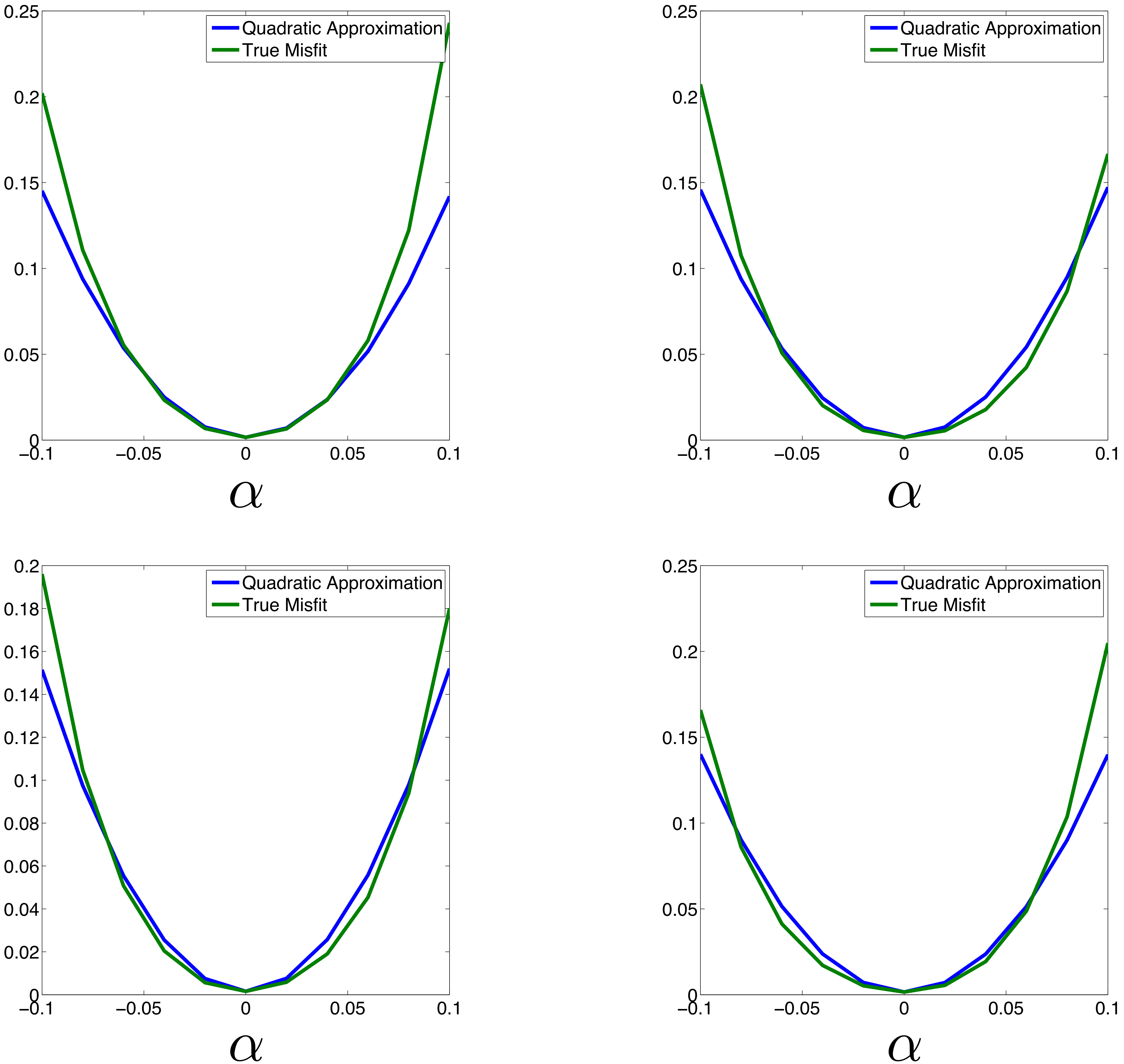
Hessian of WRI:

$$\mathbf{H} = \lambda^2(\mathbf{G}^T\mathbf{G} - \mathbf{G}^T(\mathbf{I} + \lambda^{-2}\mathbf{A}^{-T}\mathbf{P}^T\mathbf{P}\mathbf{A}^{-1})^{-1}\mathbf{G})$$



$$f(\mathbf{m}_t + \alpha \mathrm{d}\mathbf{m})$$

Different random directions dm



Uncertainty Quantification

Goal : Quantify the Uncertainty based on the posterior distribution $\rho_{\text{post}}(\mathbf{m})$

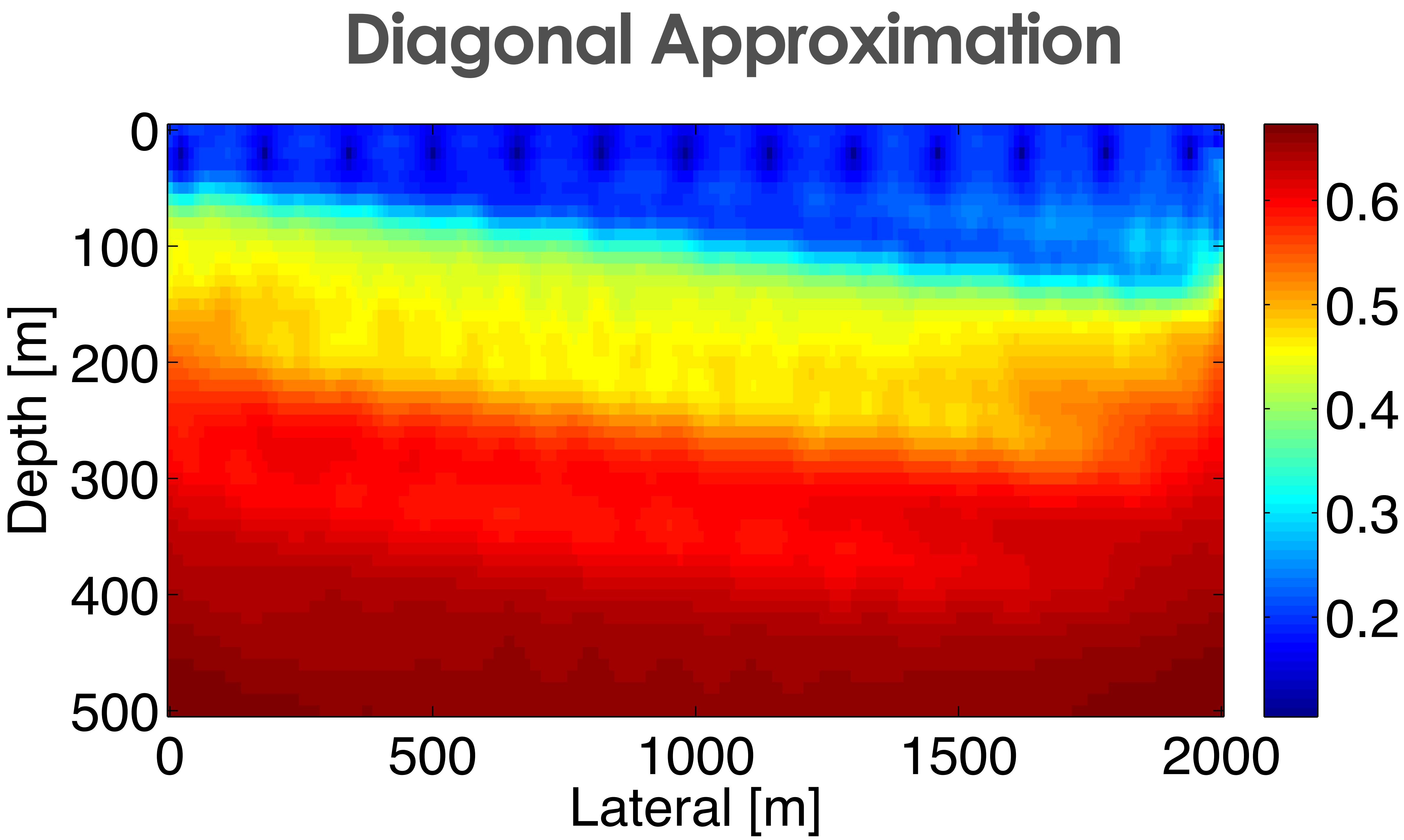
Solution:

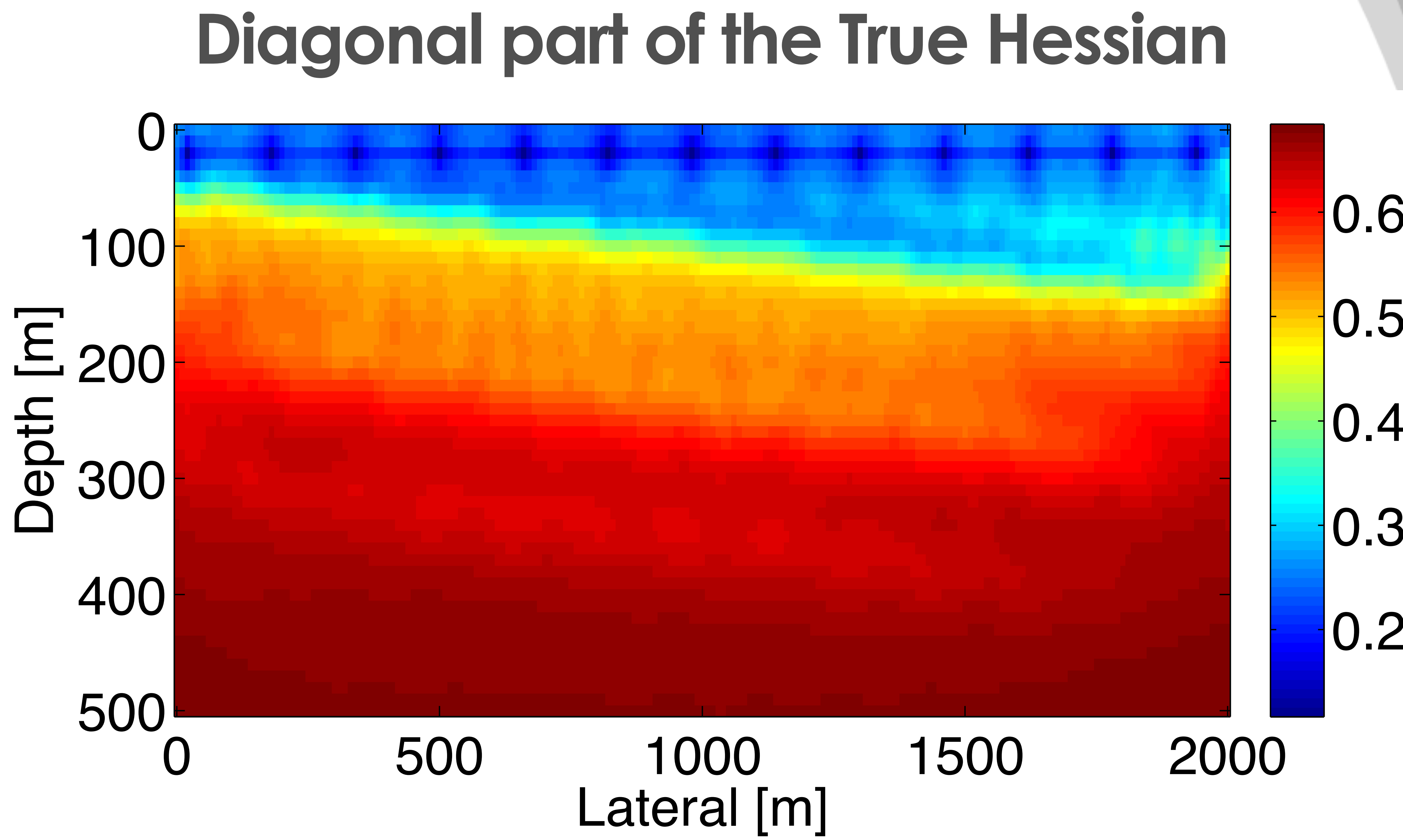
- Integrate the posterior distribution.

- McMC method to sample the posterior distribution.
  - Advantage: the true uncertainty can be quantified.
  - Disadvantage: Huge computational cost.

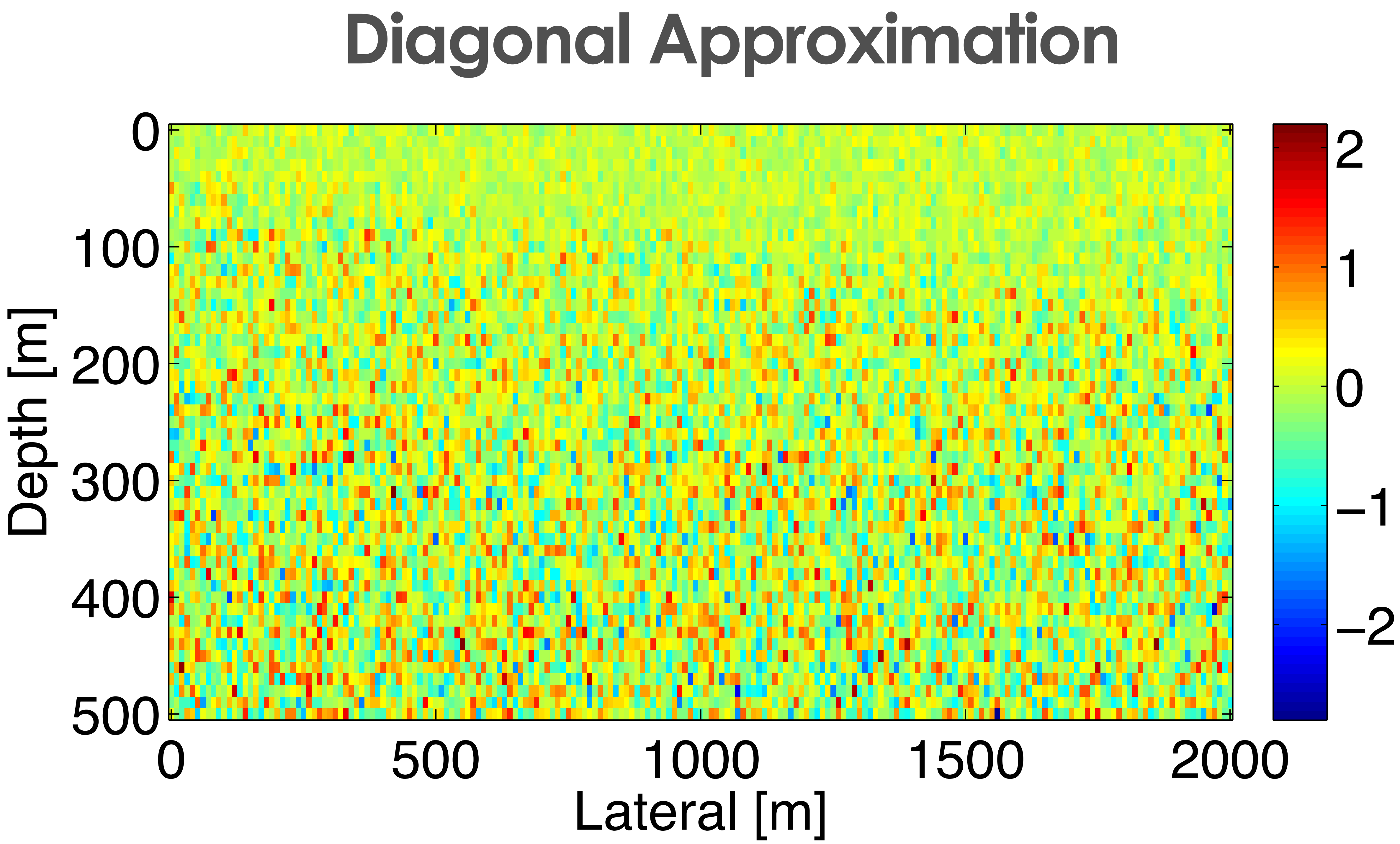
- Use an approximated distribution to quantify the uncertainty.
  - Quantify the uncertainty by estimating the diagonal part of the inverse of the Hessian
  - Advantage: cheap

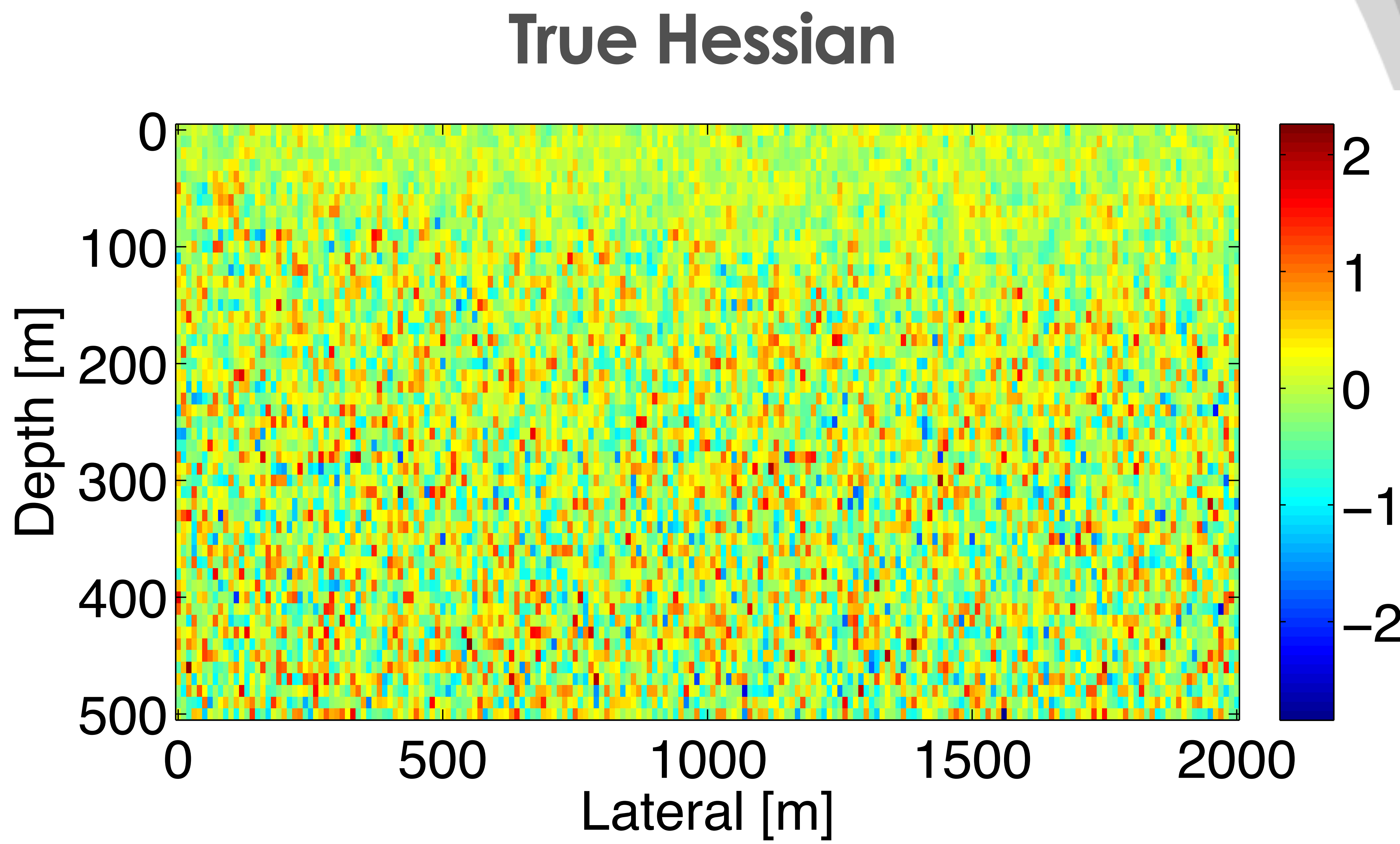
Diagonal approximation vs True Hessian
– comparision diagonal part



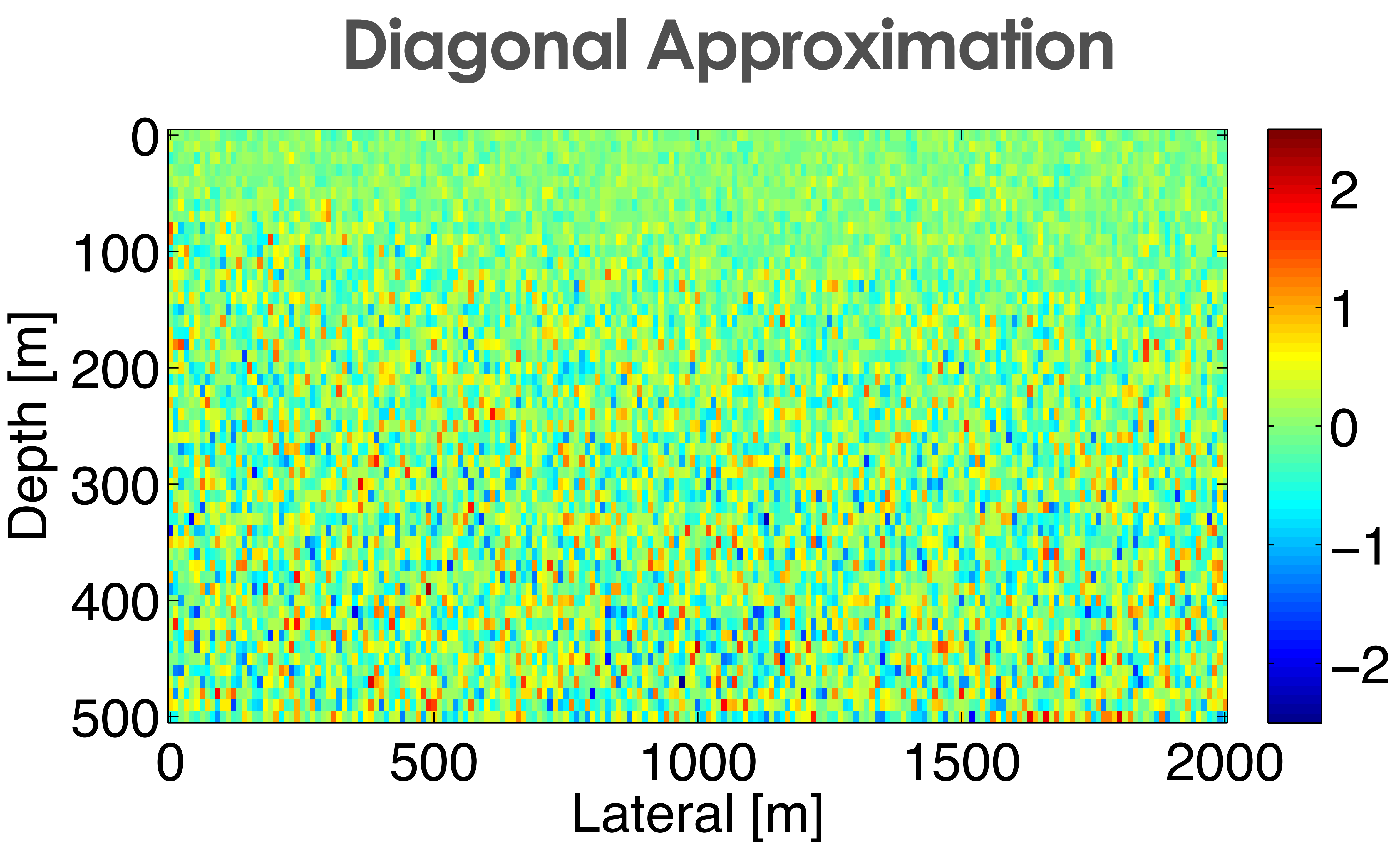


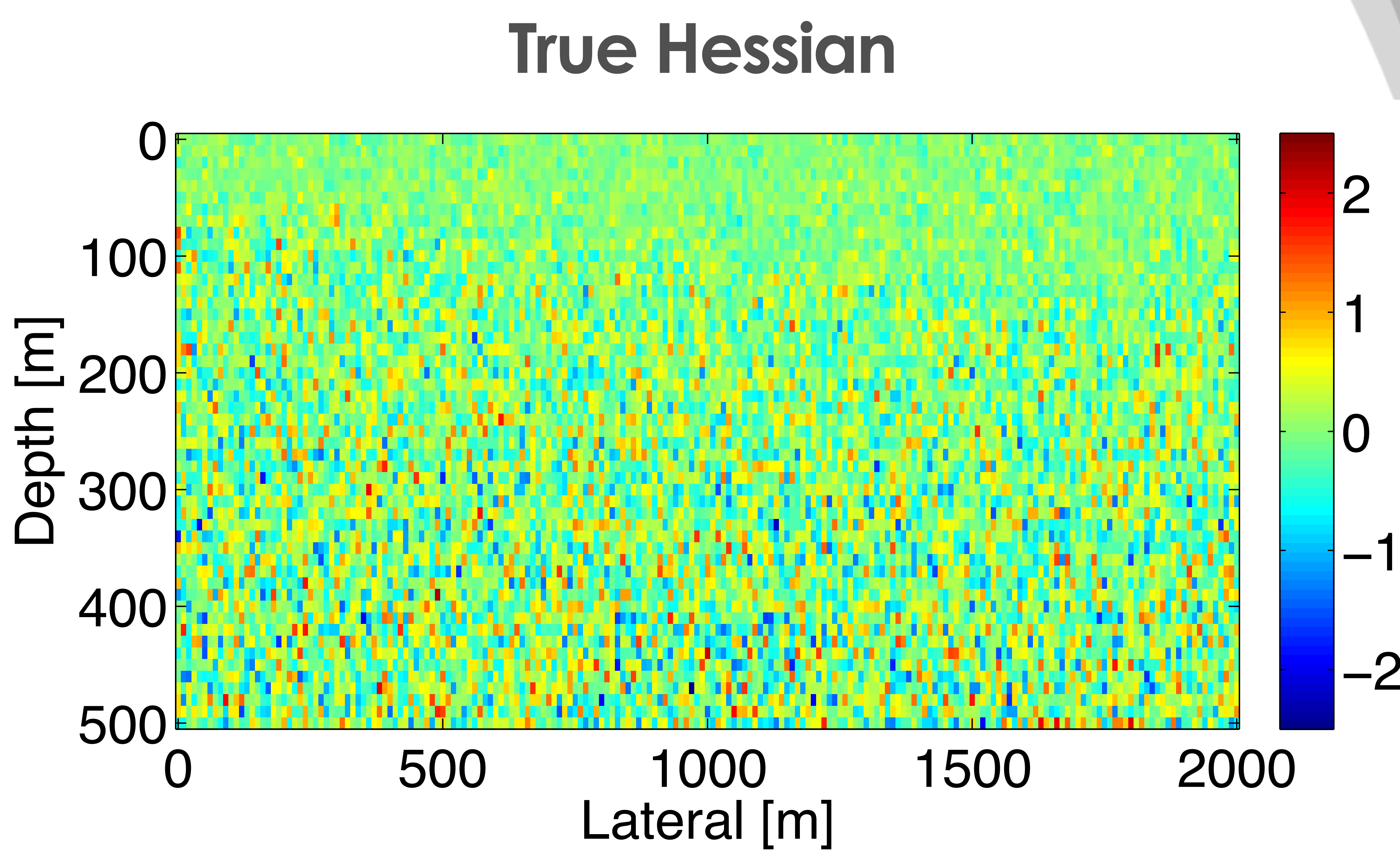
Diagonal approximation vs True Hessian
– comparison generation random perturbation



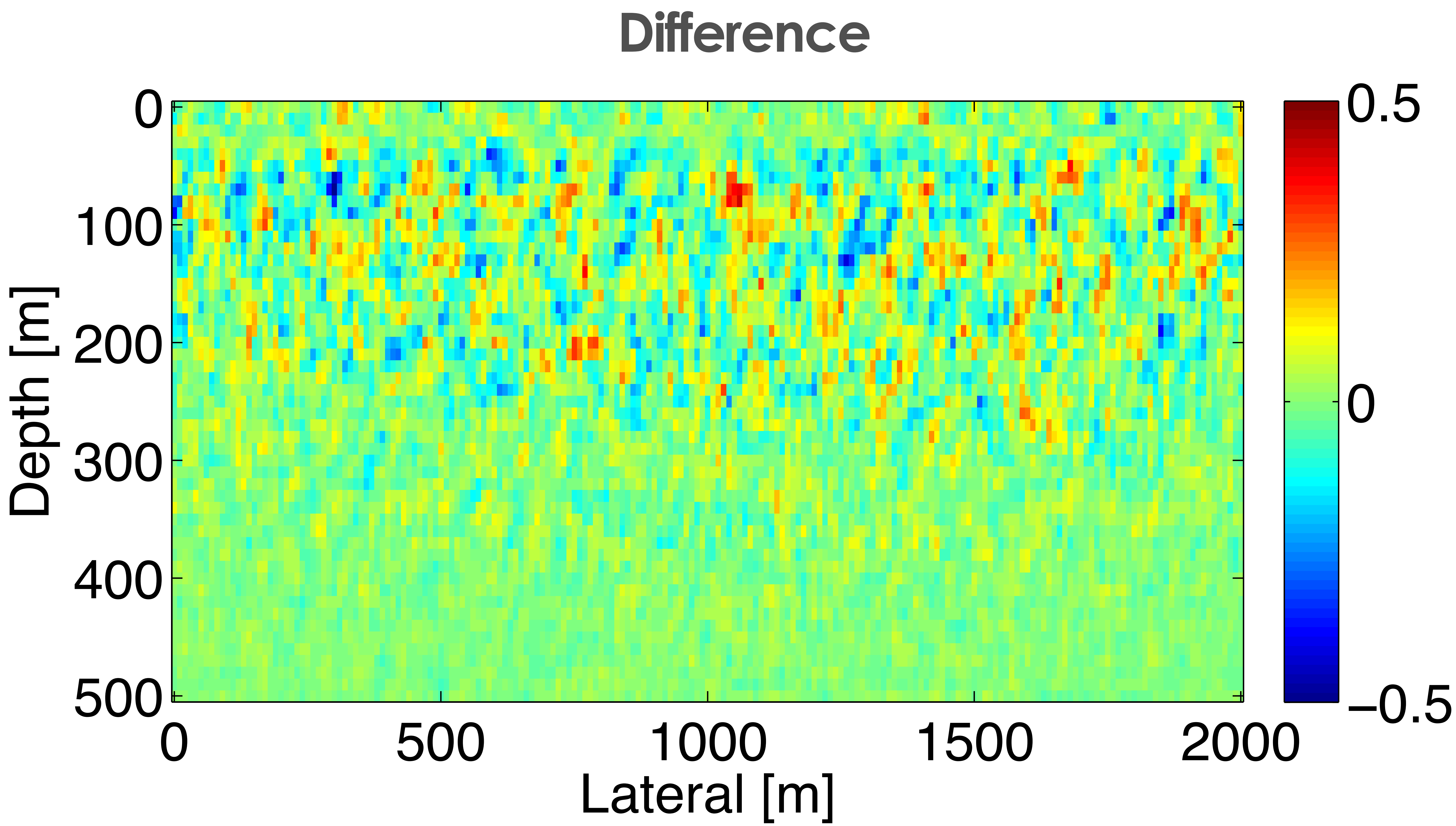


Diagonal approximation vs True Hessian
– comparison generation random perturbation

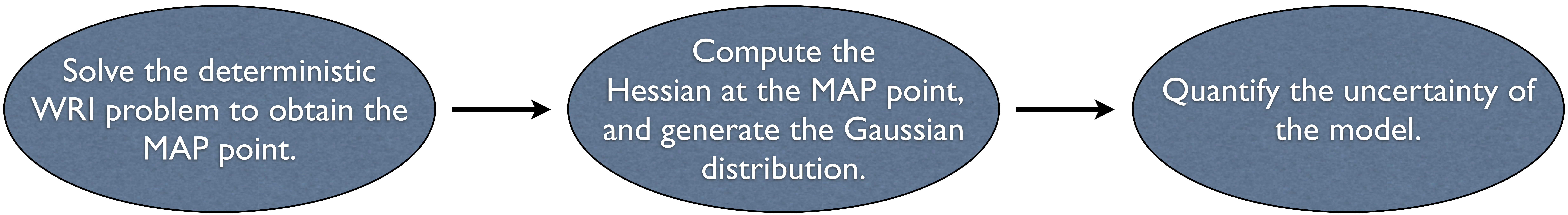




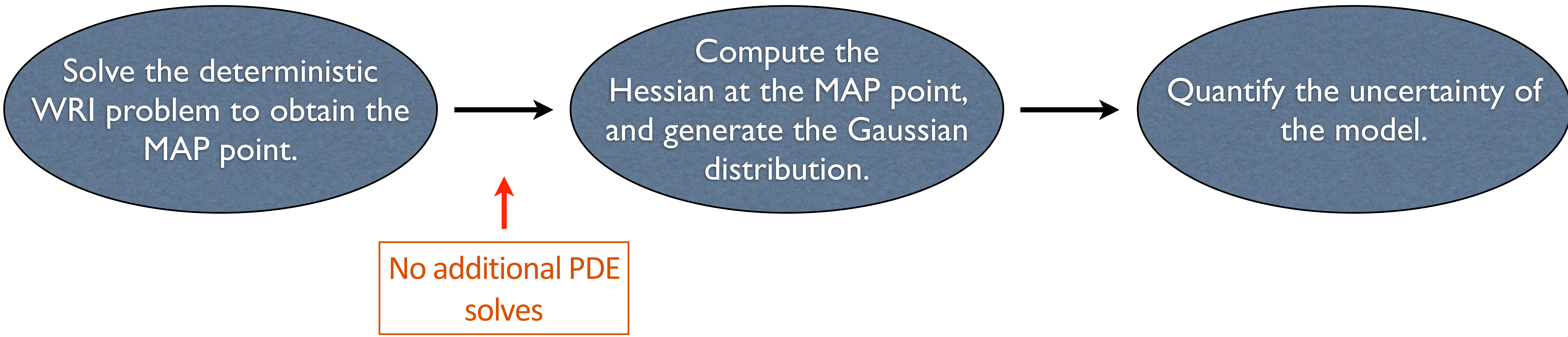
Diagonal Approximation vs True Hessian
– comparison generation random perturbation



Workflow

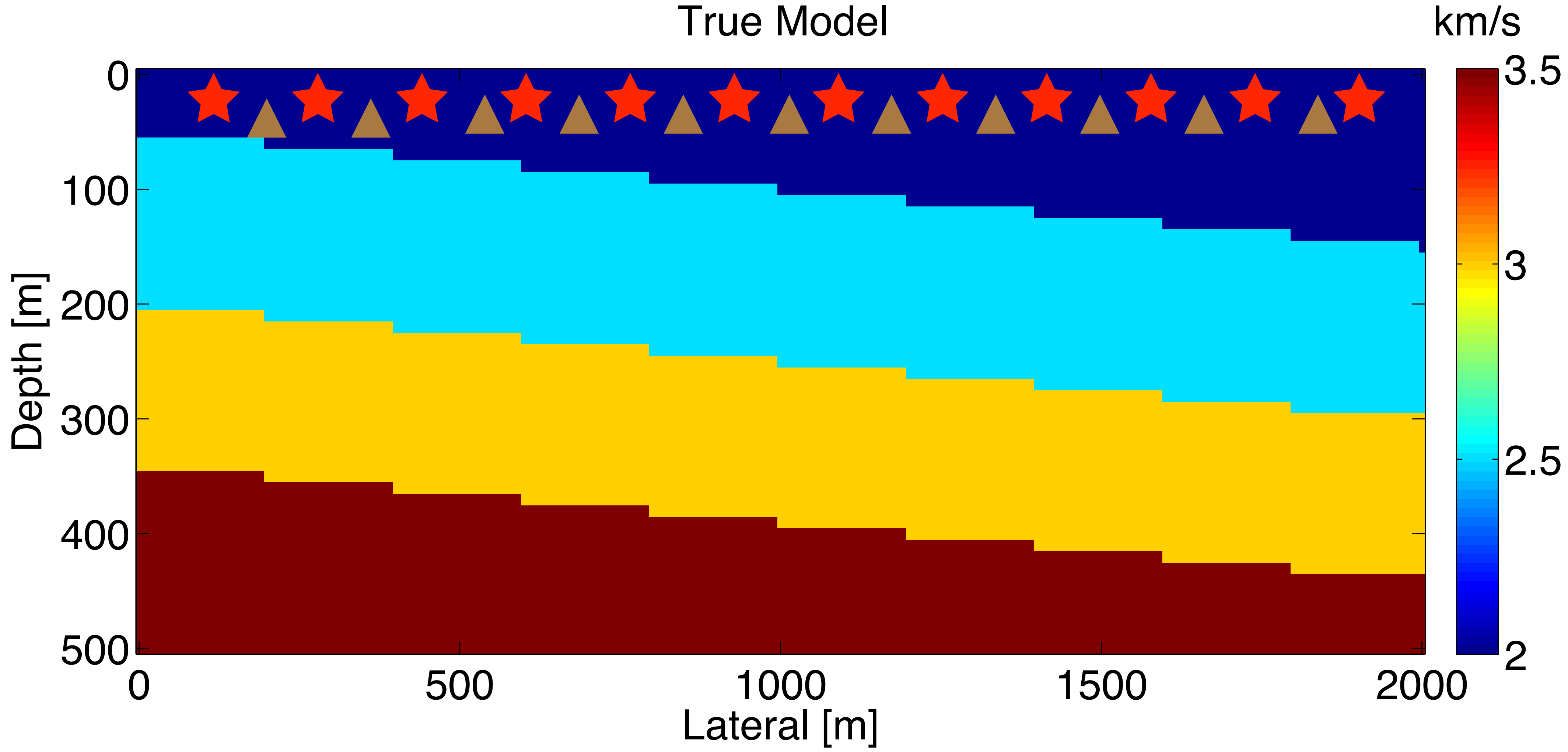

Solve the deterministic WRI problem to obtain the MAP point.
Compute the Hessian at the MAP point, and generate the Gaussian distribution.
Quantify the uncertainty of the model.

Workflow

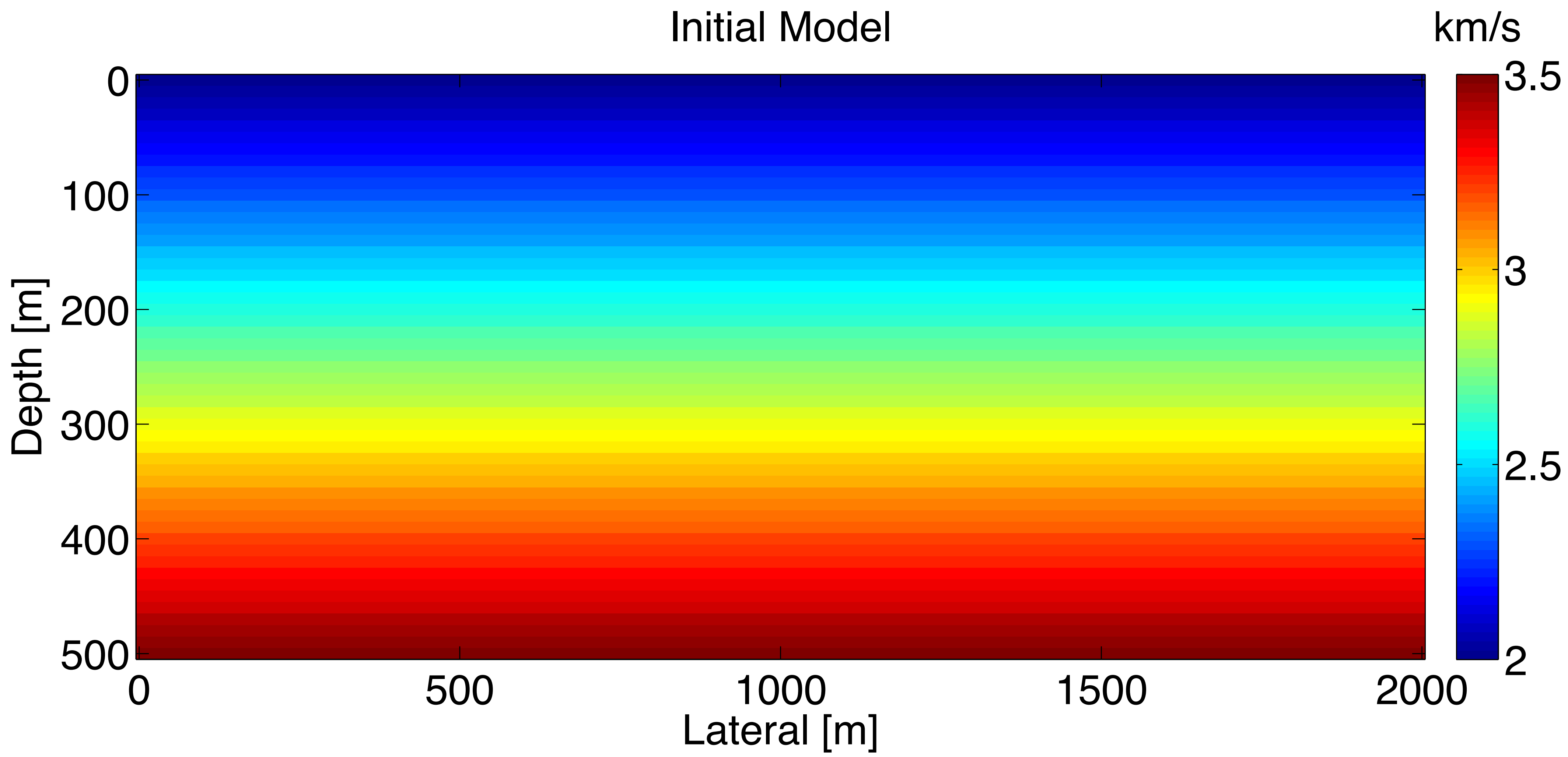


Solve the deterministic WRI problem to obtain the MAP point.
Compute the Hessian at the MAP point, and generate the Gaussian distribution.
Quantify the uncertainty of the model.
No additional PDE solves

Numerical experiment



(a)

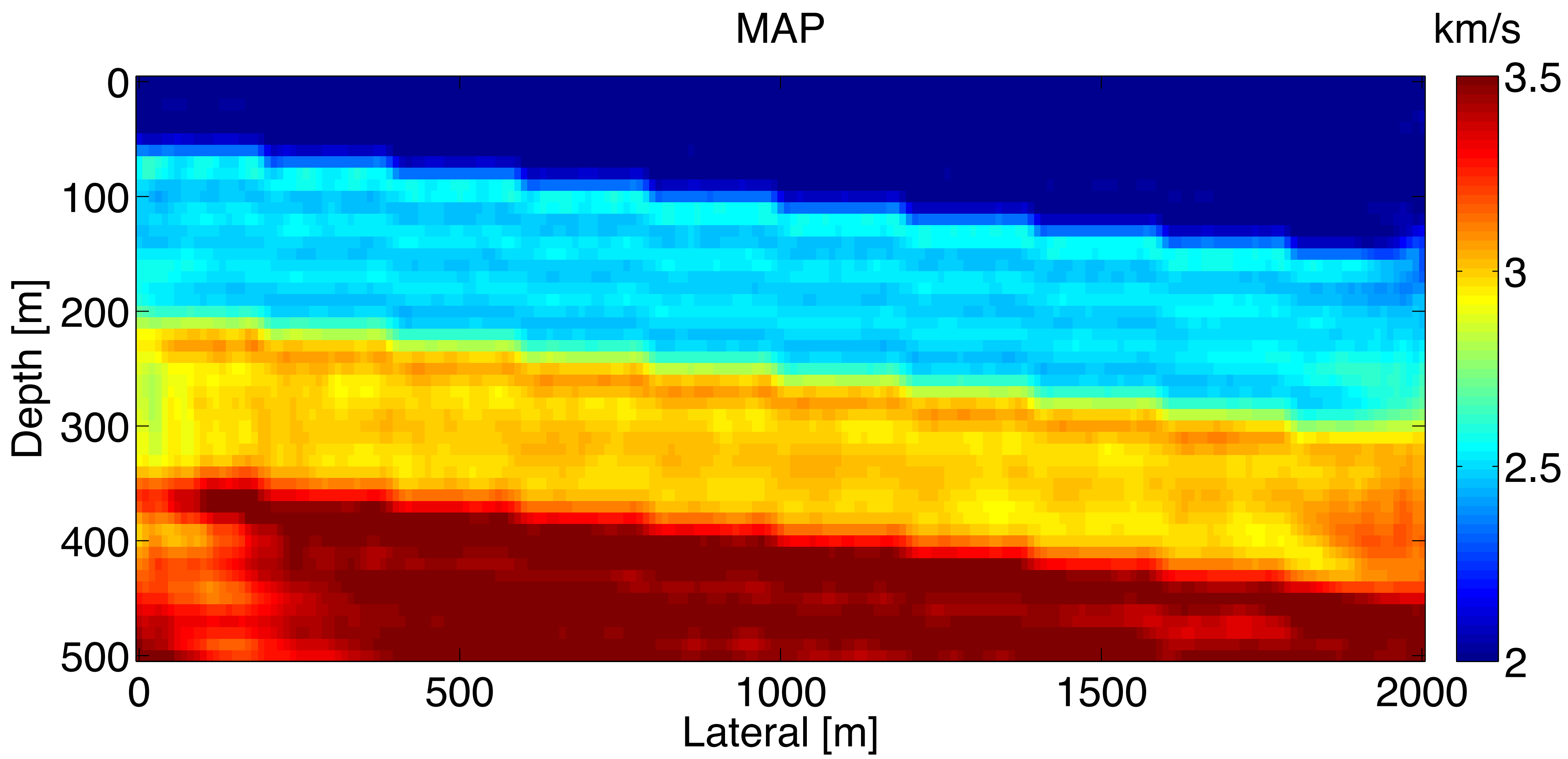


(b)

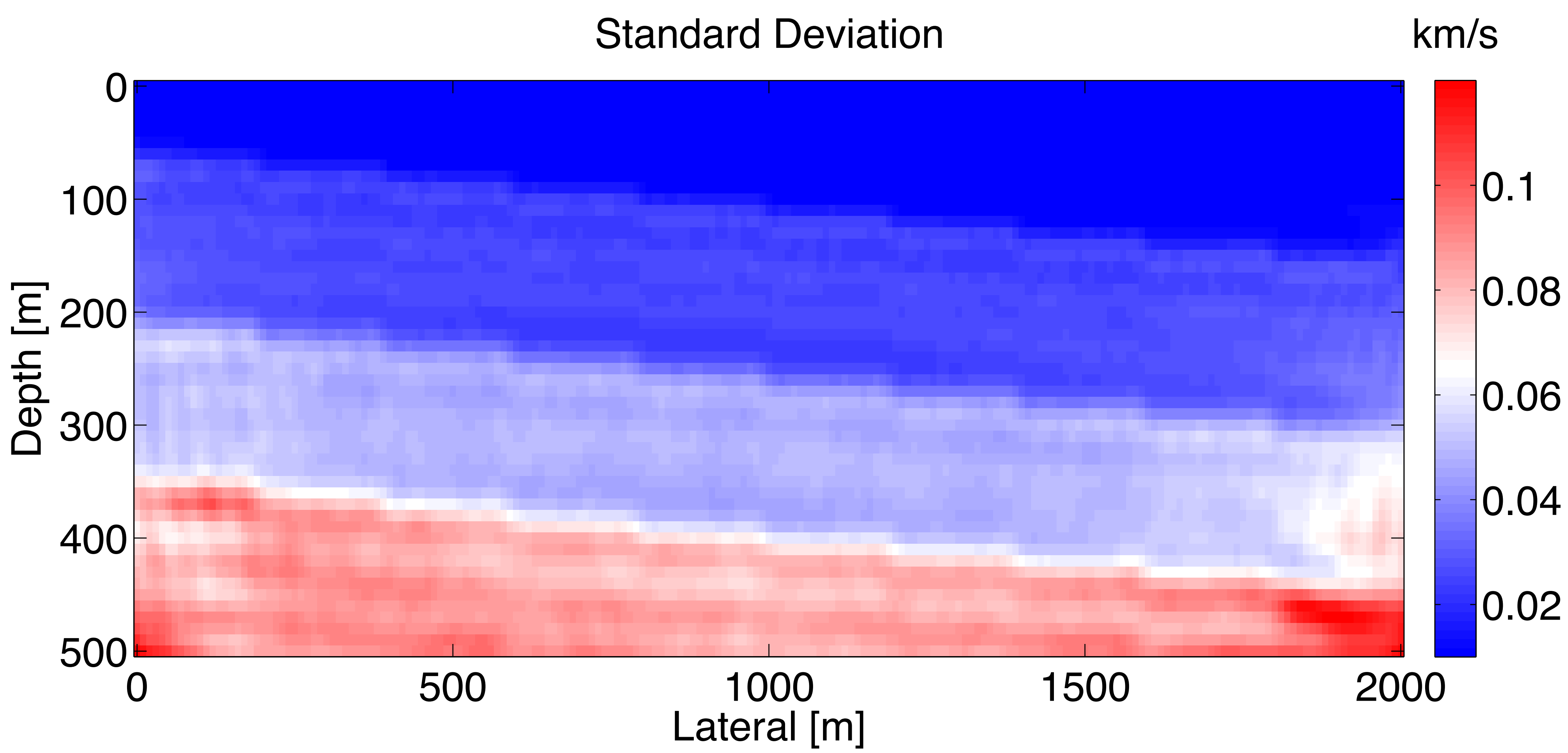
Model size: 500m x 2000m
Source spacing: 80m
Receiver spacing: 20m
Fixed spread 2km
Frequency : 10-30 Hz

Standard deviation of data noise: 0.5
Standard deviation of PDE: 0.5
lambda: 1

Simple Model

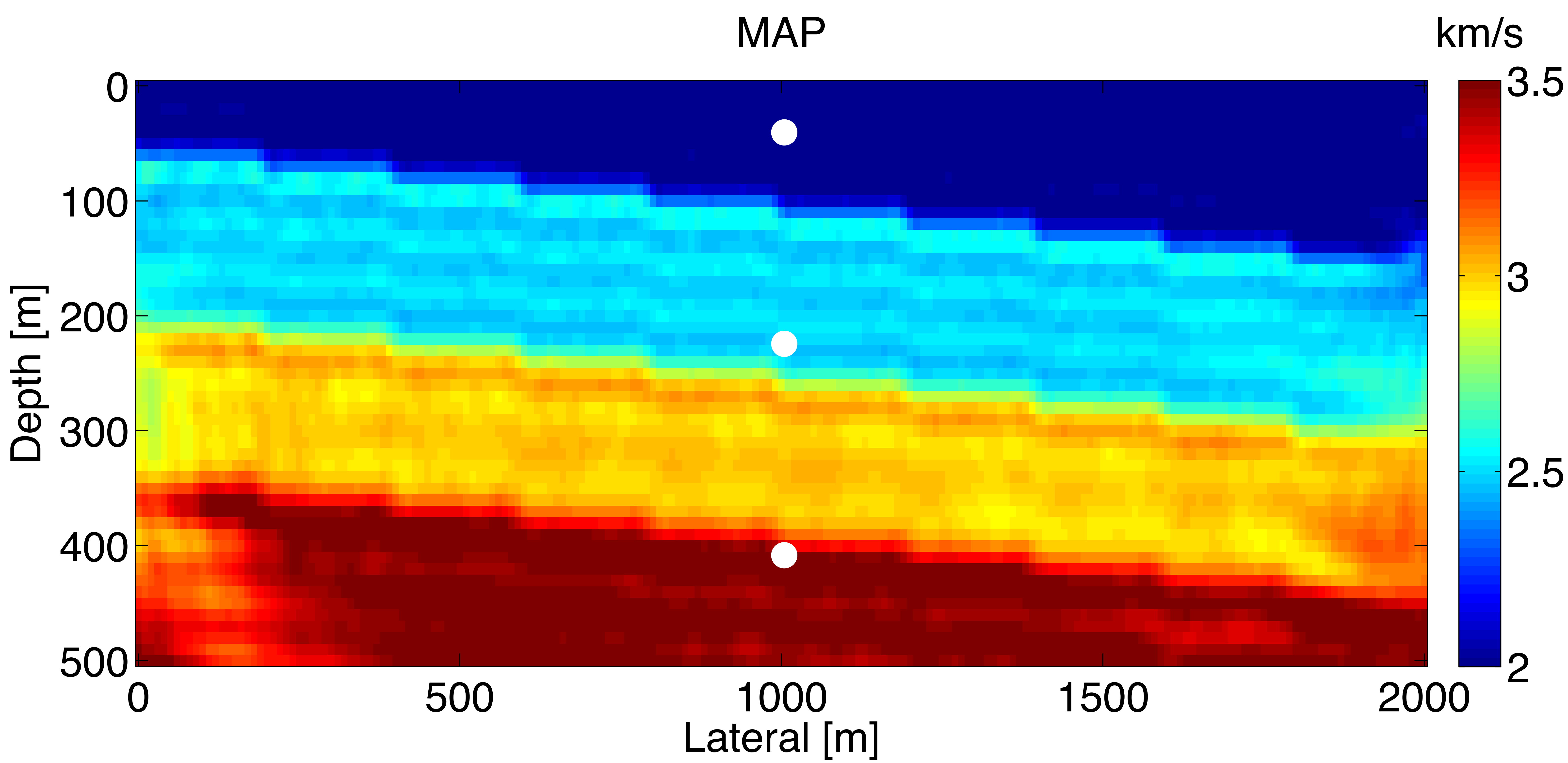


a). Maximum a posterior point



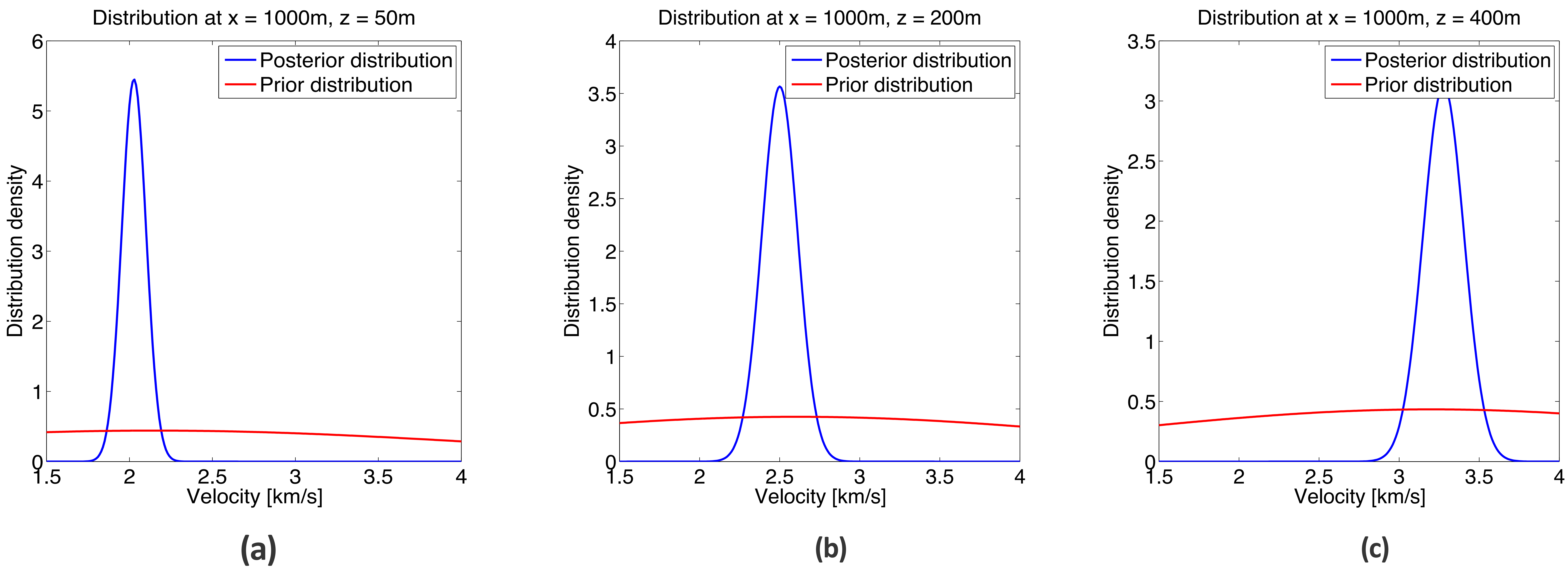
b). The standard deviation

Posterior Distribution

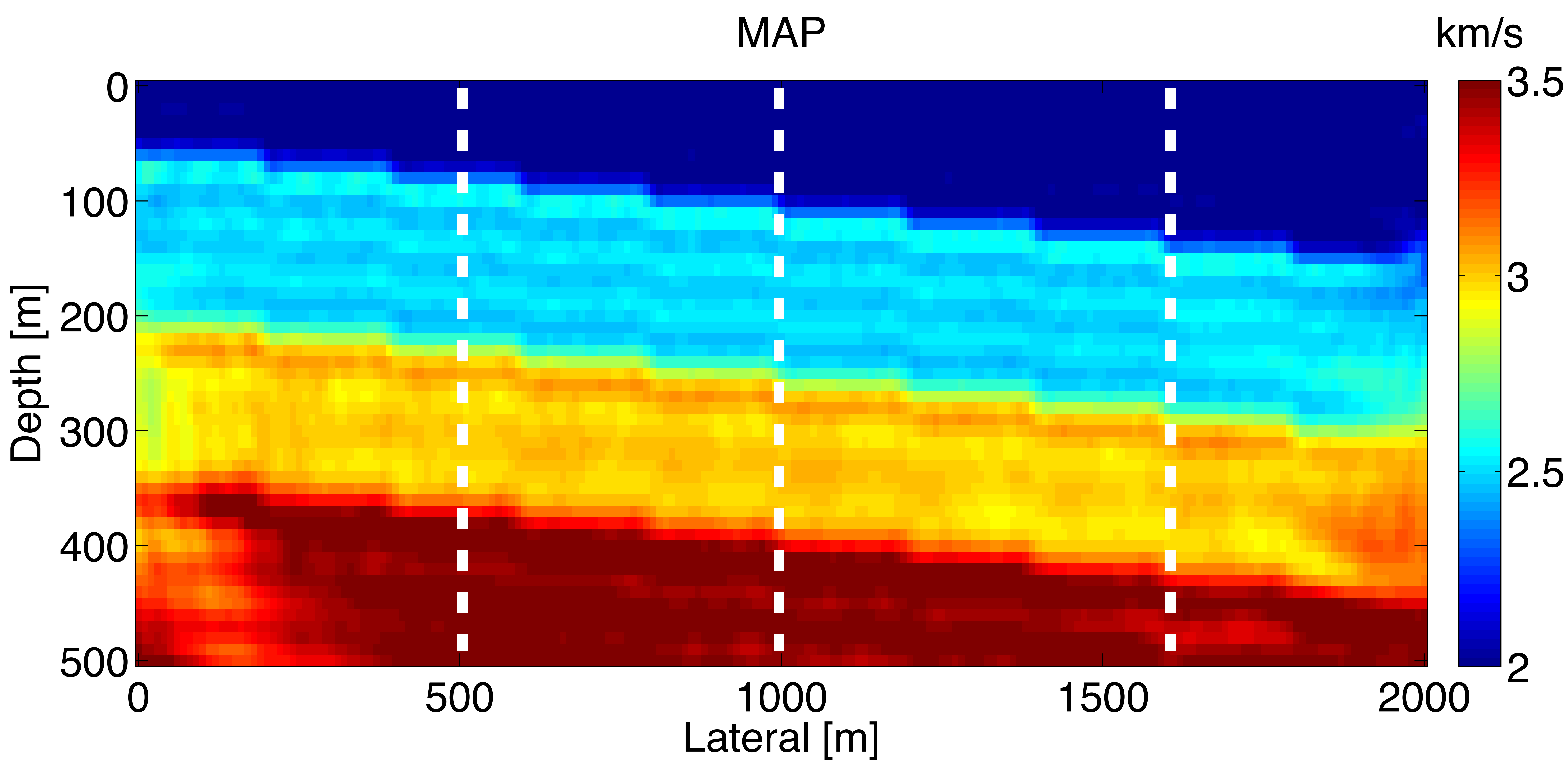


Maximum a posterior

Posterior Distribution



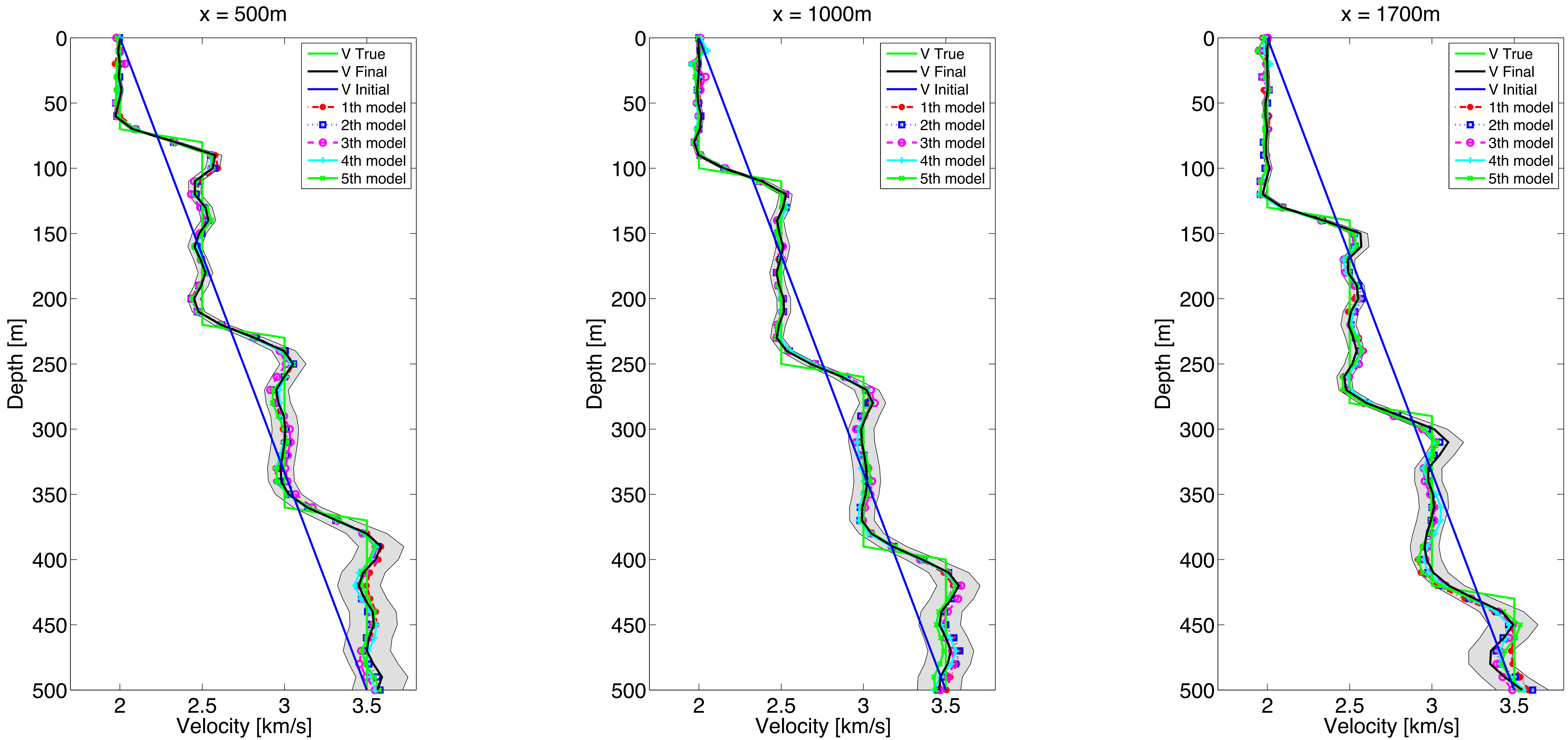
Confidence Interval



Maximum a posterior

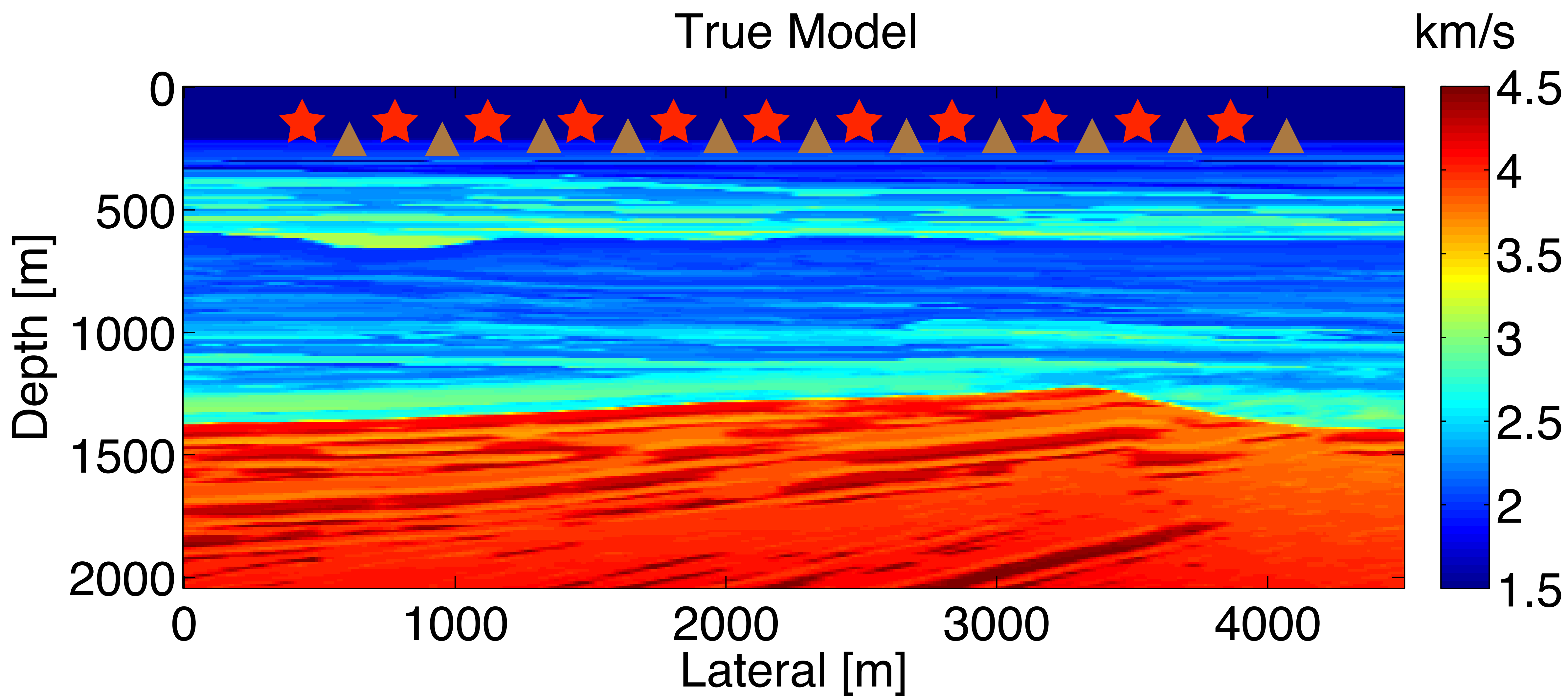
Confidence Interval

$\alpha = 0.95$

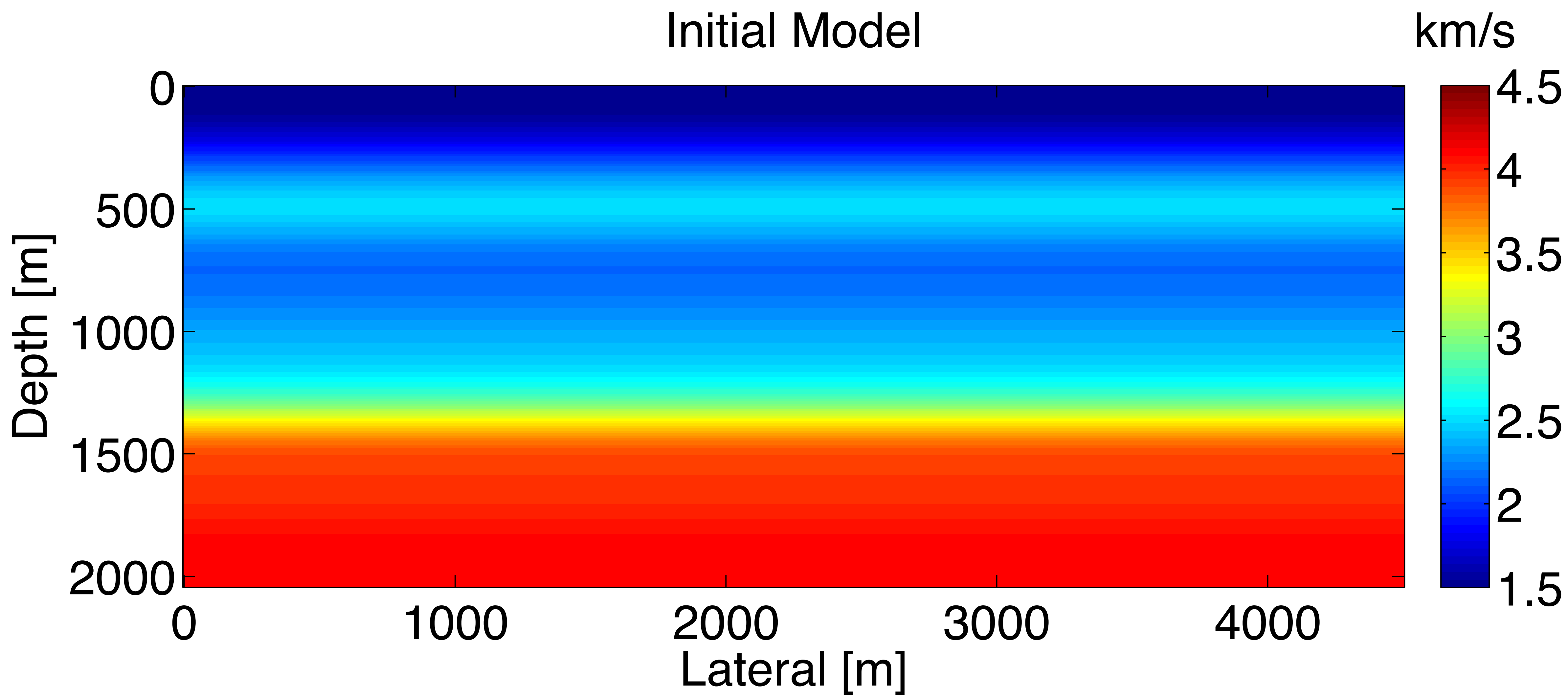


(a) **(b)** **(c)**

BG Model



(a)

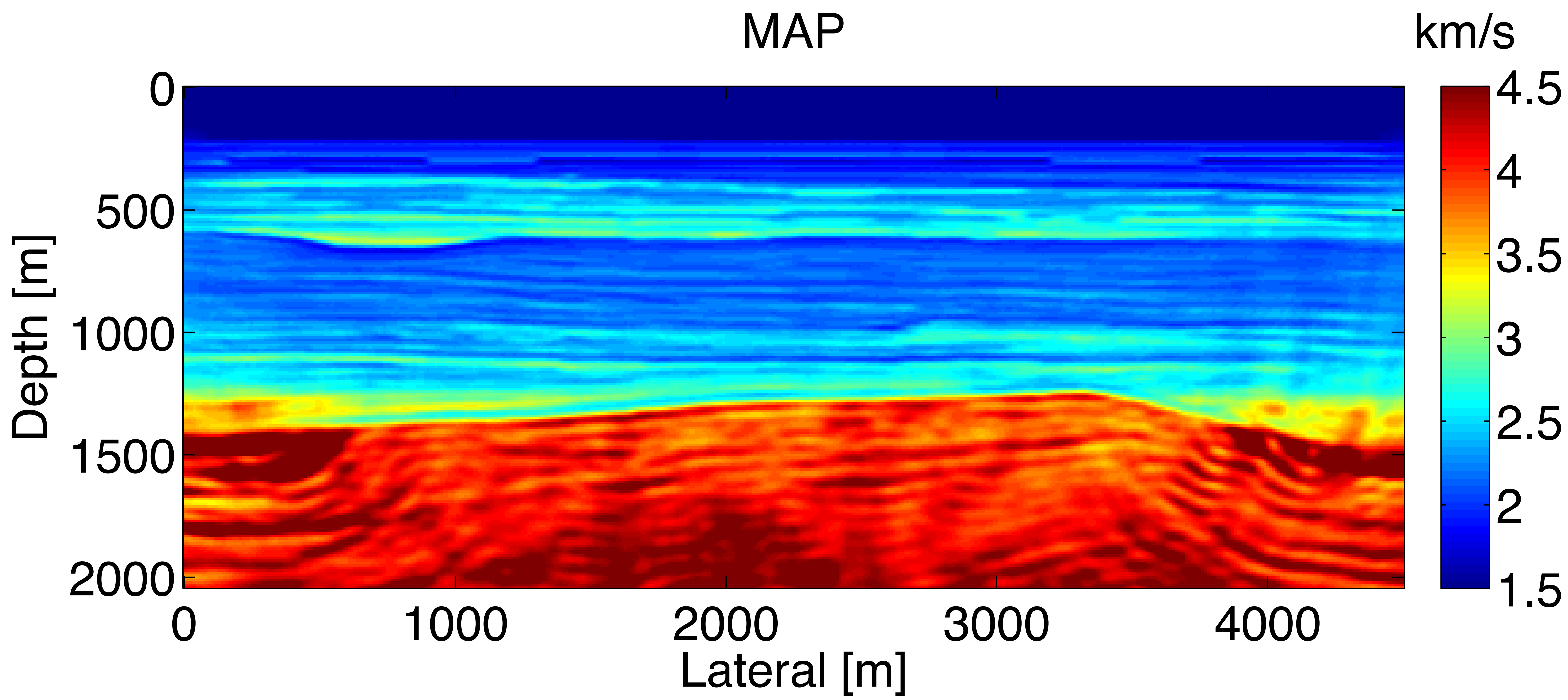


(b)

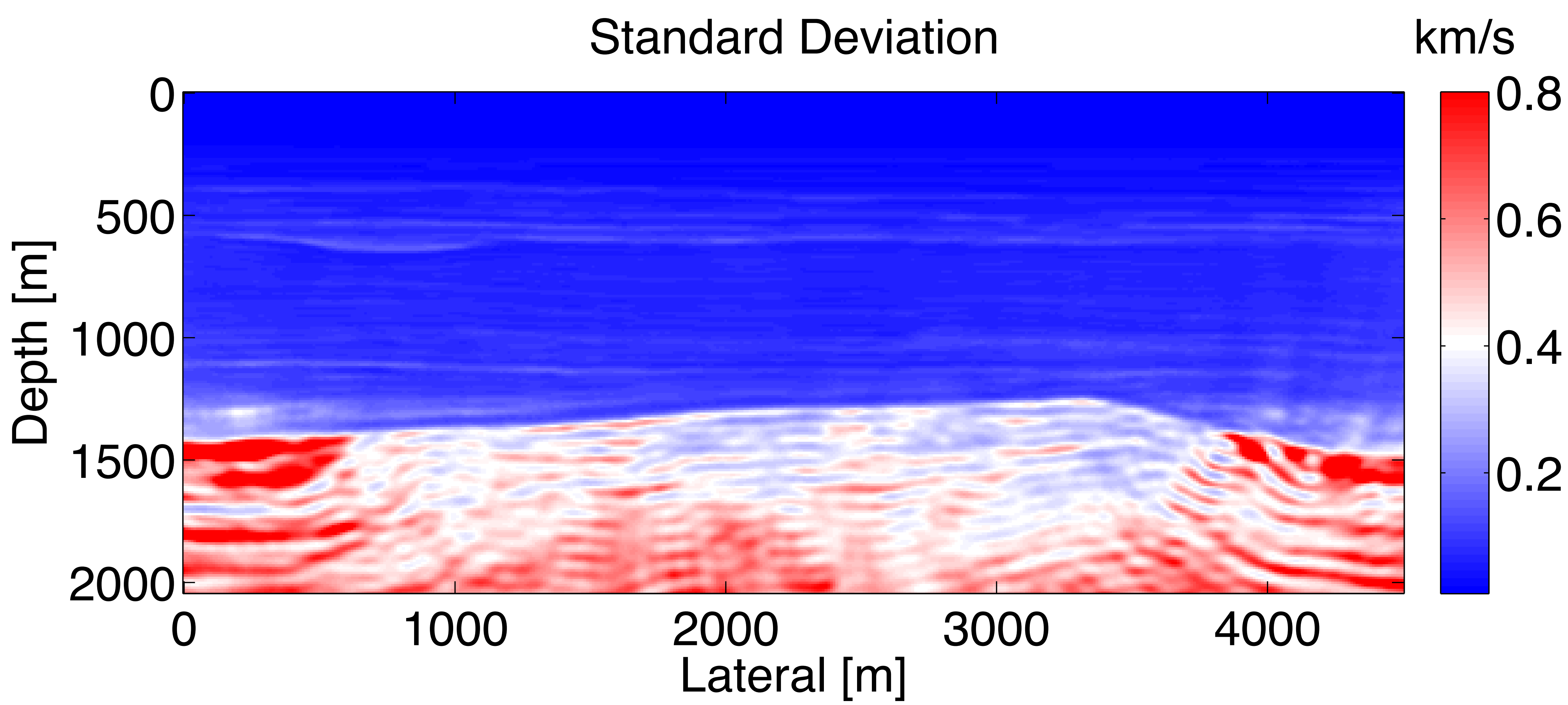
Model size: 2000m x 4500m
Source spacing: 50m
Receiver spacing: 10m
Fixed spread 4.5km
Frequency : 5~31 Hz

Standard deviation of data noise: 0.5
Standard deviation of PDE: 0.5
lambda: 1

BG Model

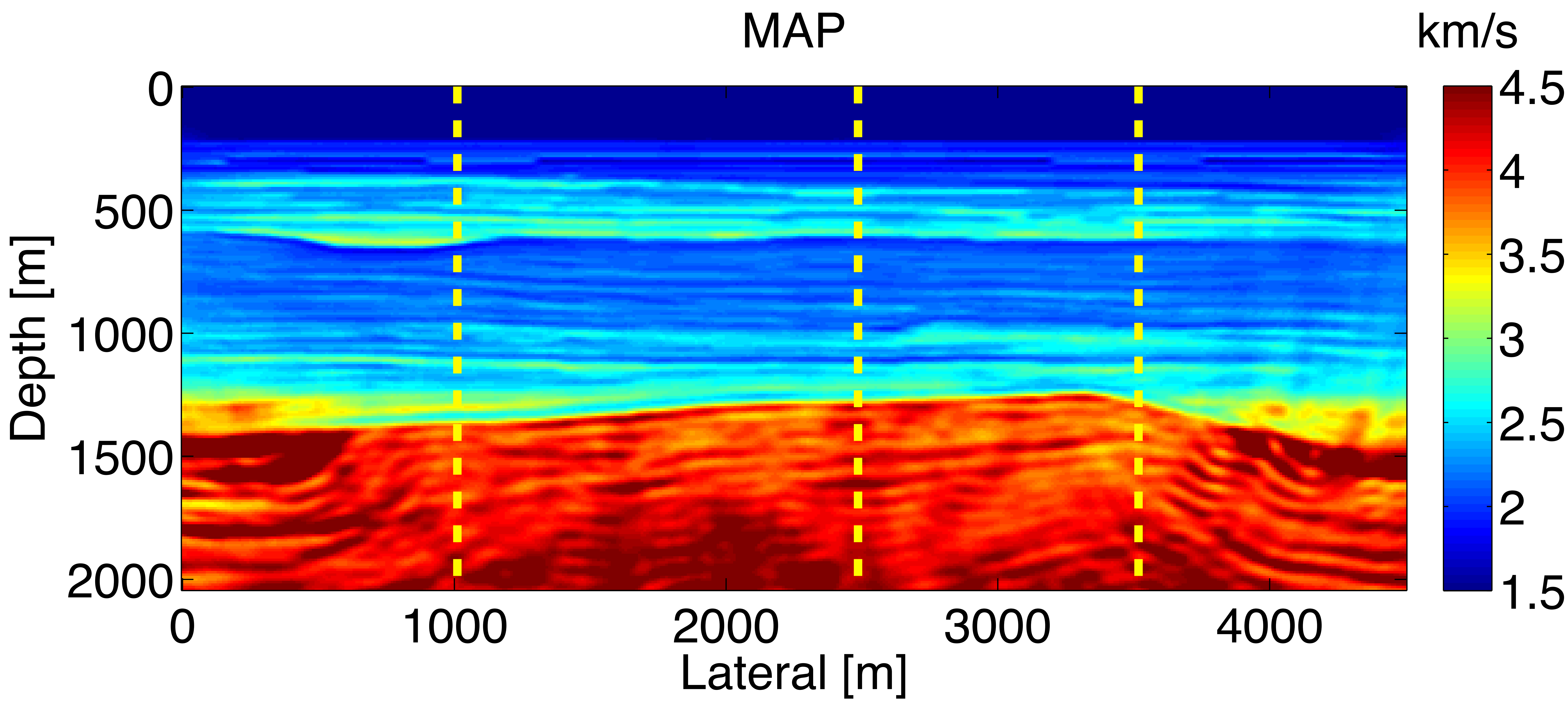


a). Maximum a posterior point



b). The standard deviation

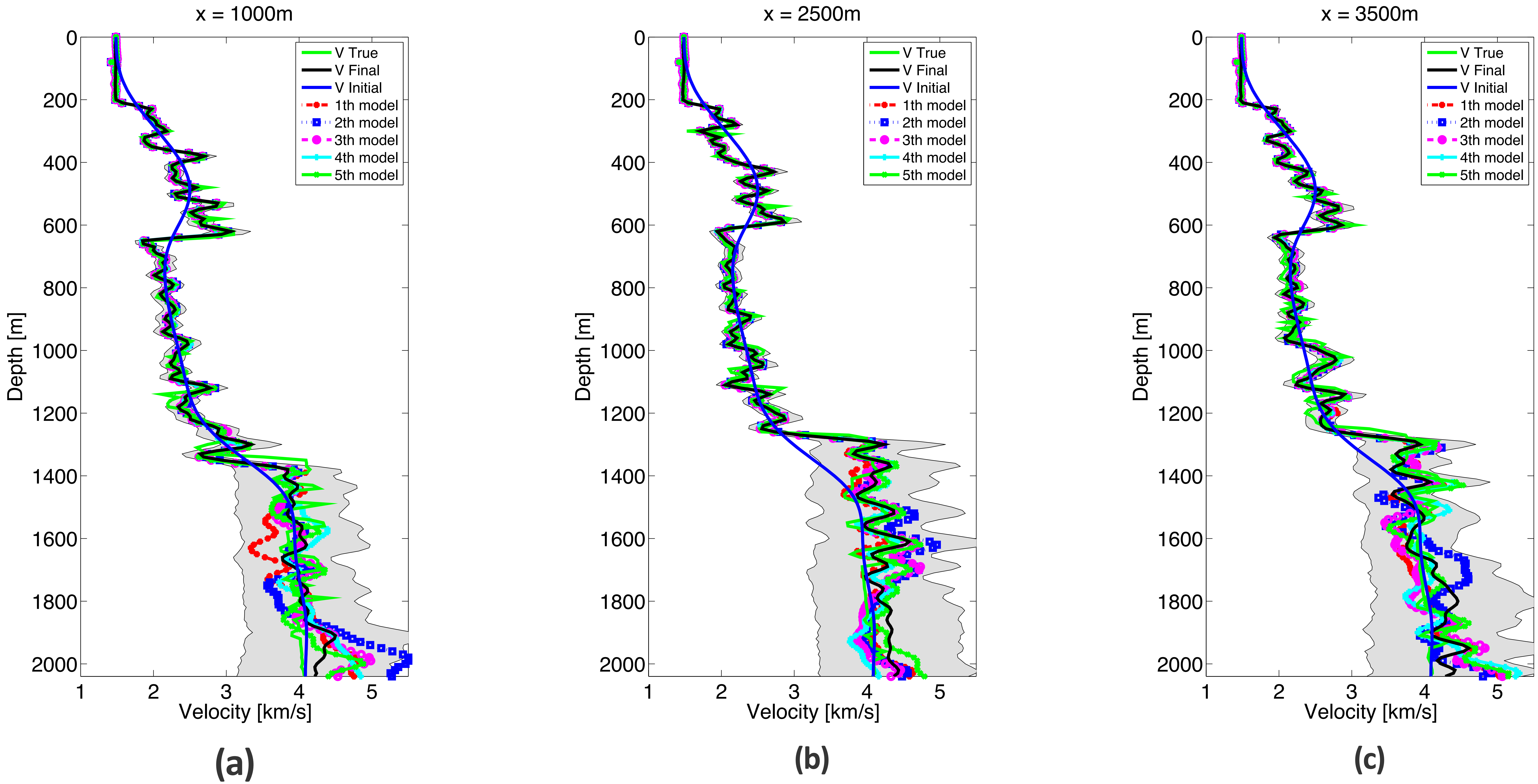
Confidence Interval



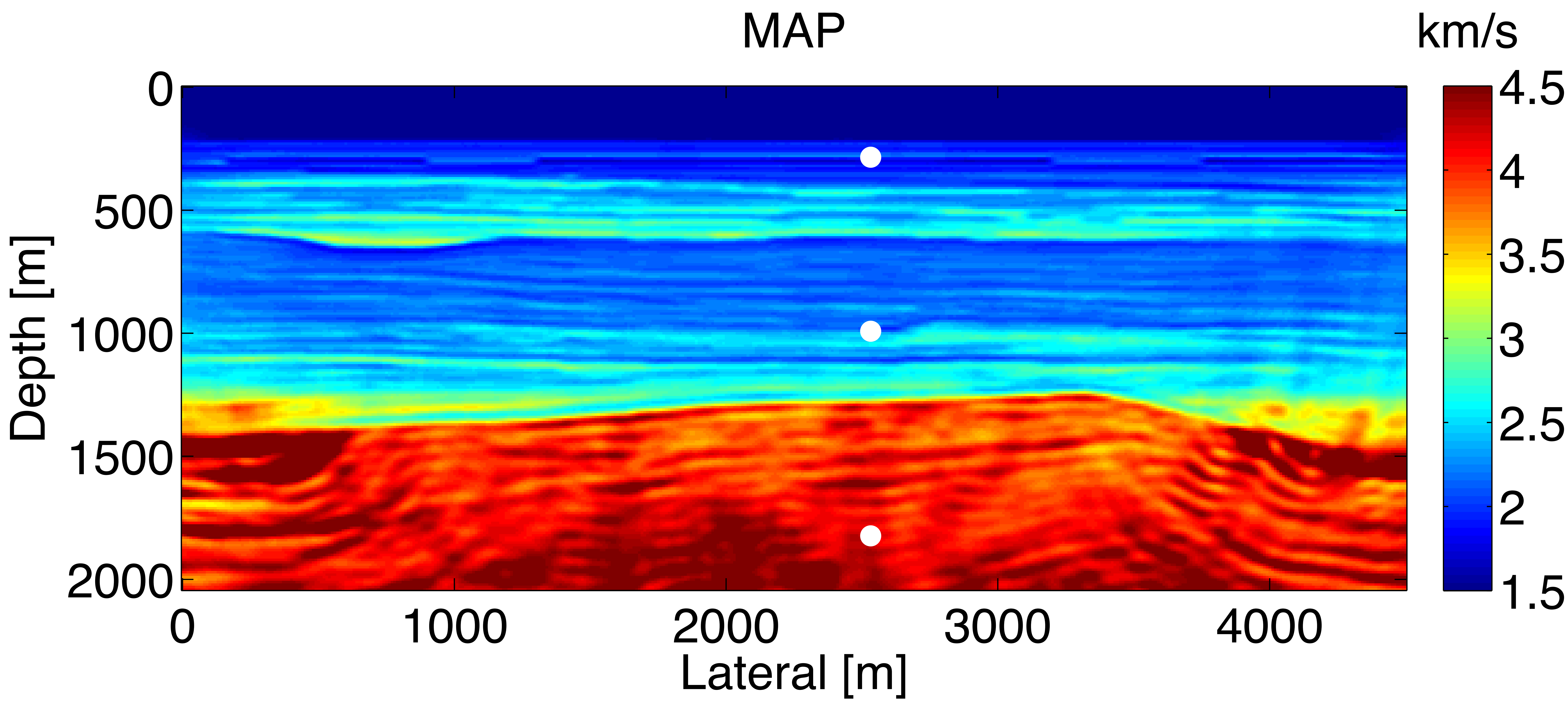
Maximum a posterior

Confidence Interval

$\alpha = 0.95$

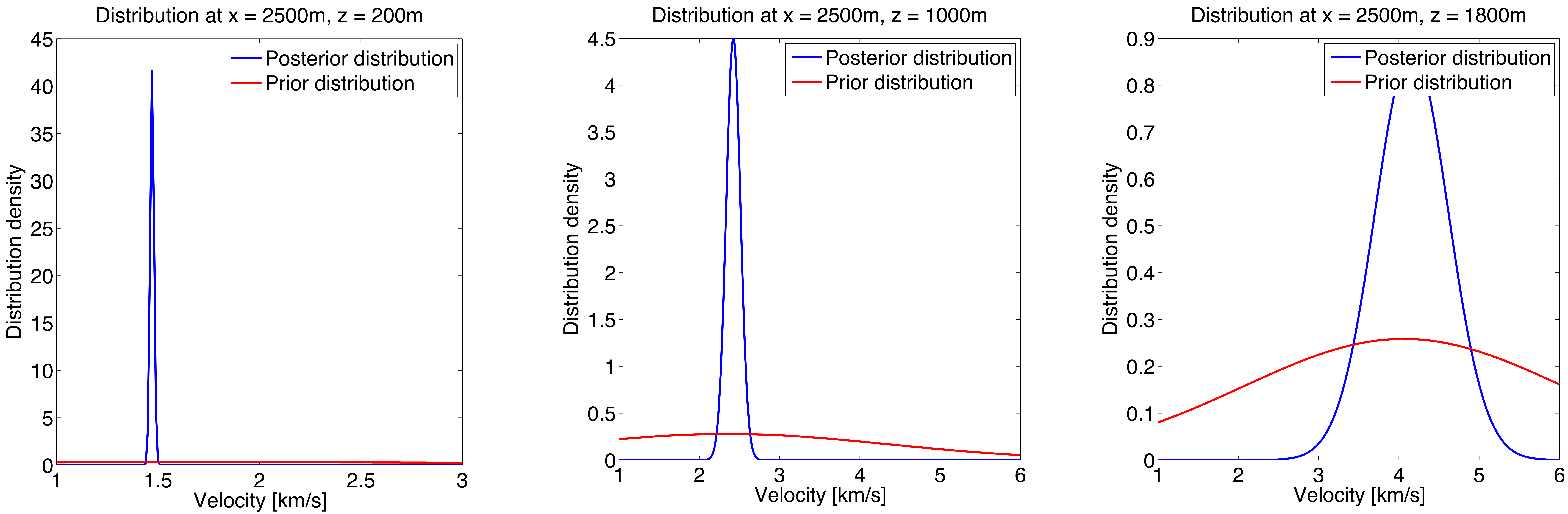


Posterior Distribution



Maximum a posterior

Posterior Distribution vs Prior Distribution



(a) **(b)** **(c)**

Summary

1. The negative logarithm of the posterior distribution of WRI can be well approximated by a quadratic approximation at the MAP.

2. The diagonal Gauss-Newton Hessian can be used to approximate the true Hessian to quantify the uncertainty.

3. Leads to a computationally feasible algorithm to conduct Uncertainty Quantification.

Future Work

1. Incorporate spatial structure of earth models in priors.

2. Study the effect of λ to the posterior distribution.

3. More thorough theoretical analysis of UQ for WRI.

4. Extension to 3D.

SLIM

Acknowledgements

Thank you for your attention !!

SINBAD

NSERC CRSNG

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, CGG, Chevron, ConocoPhillips, ION, Petrobras, PGS, Statoil, Total SA, Sub Salt Solutions, WesternGeco, and Woodside.