Uncertainty quantification for Wavefield-Reconstruction Inversion using a positive-definite approximated Hessian

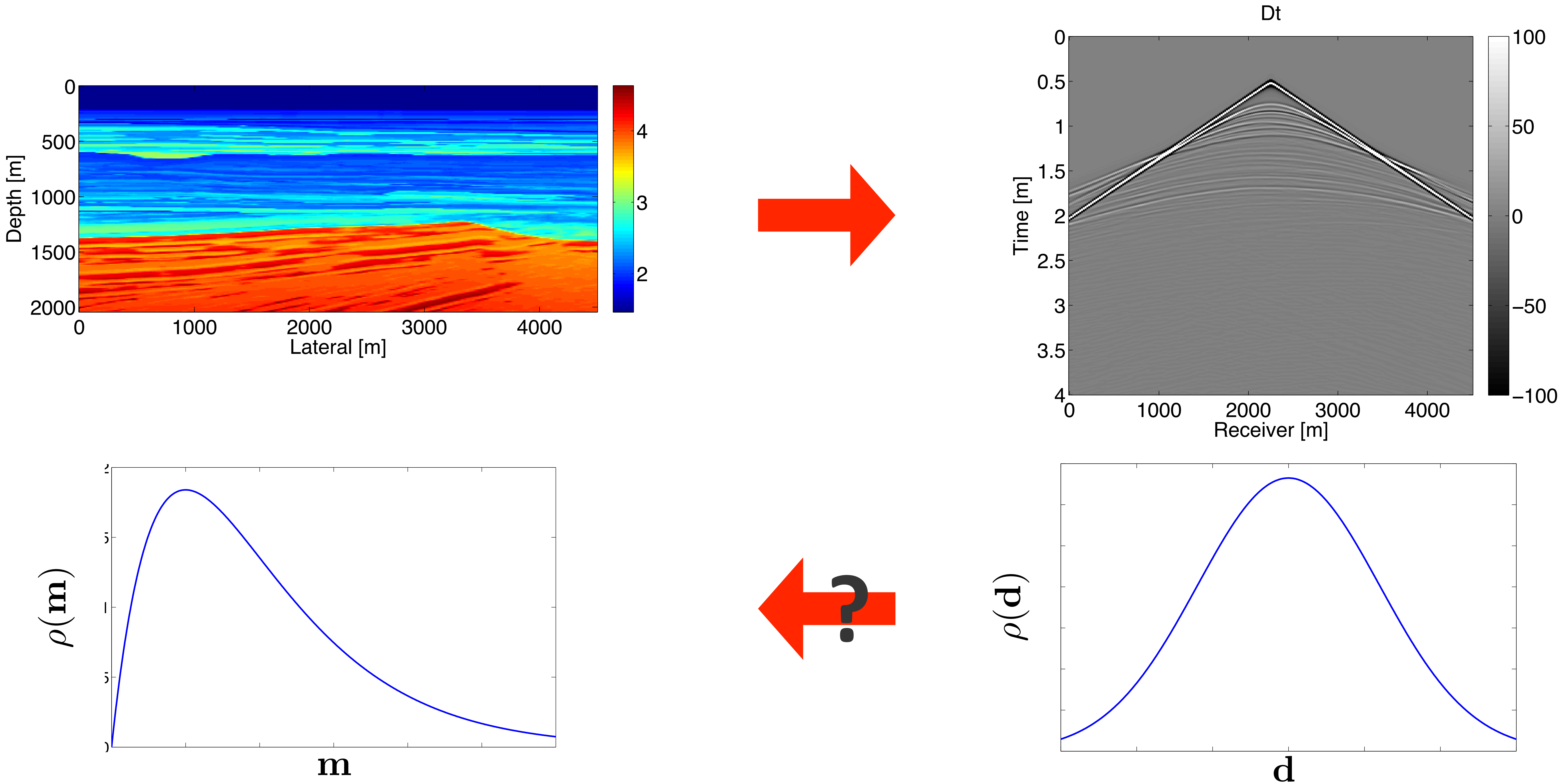
Zhilong Fang, Chia-Ying Lee, Curt Da Silva, Felix J. Herrmann and Rachel Kuske





SLIM
University of British Columbia

Motivation



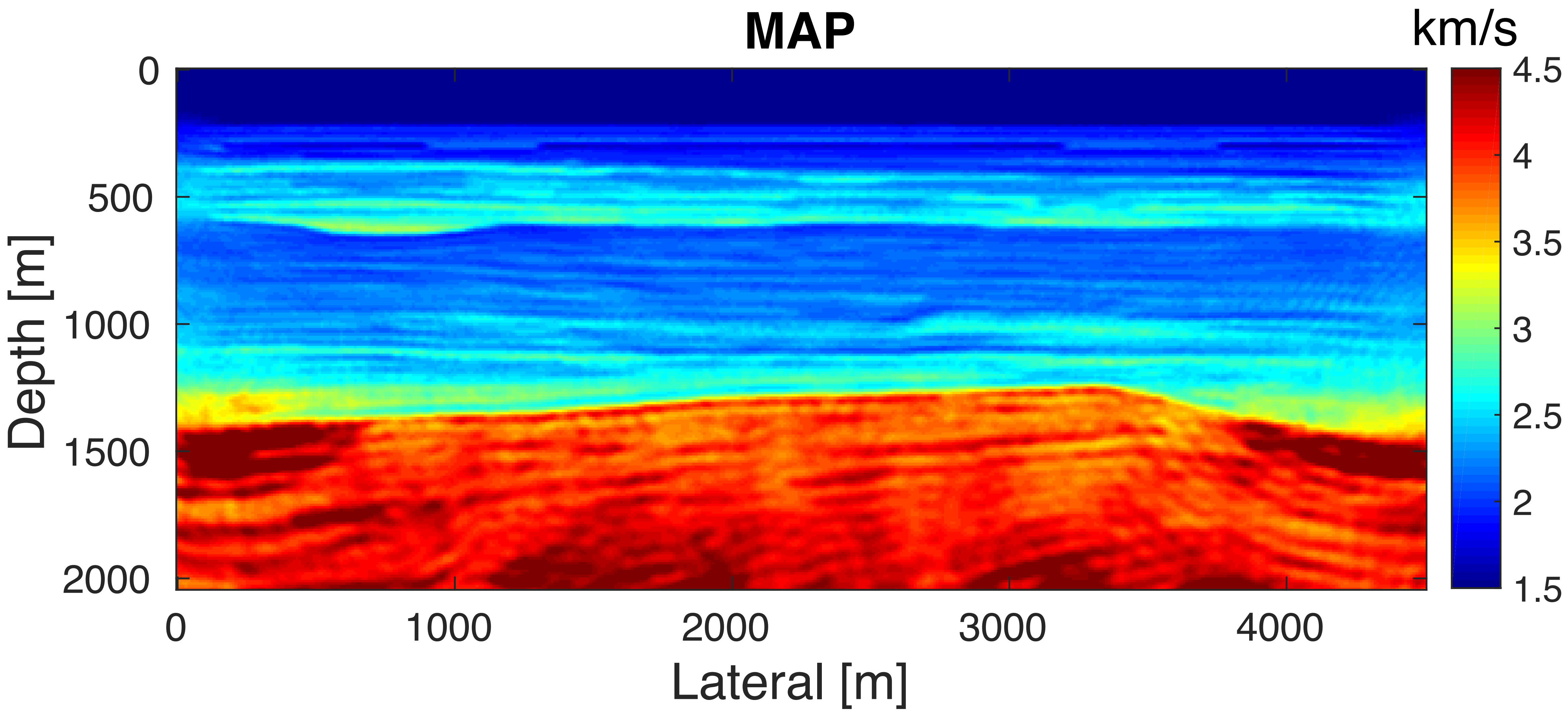
Distribution of model

Distribution of data

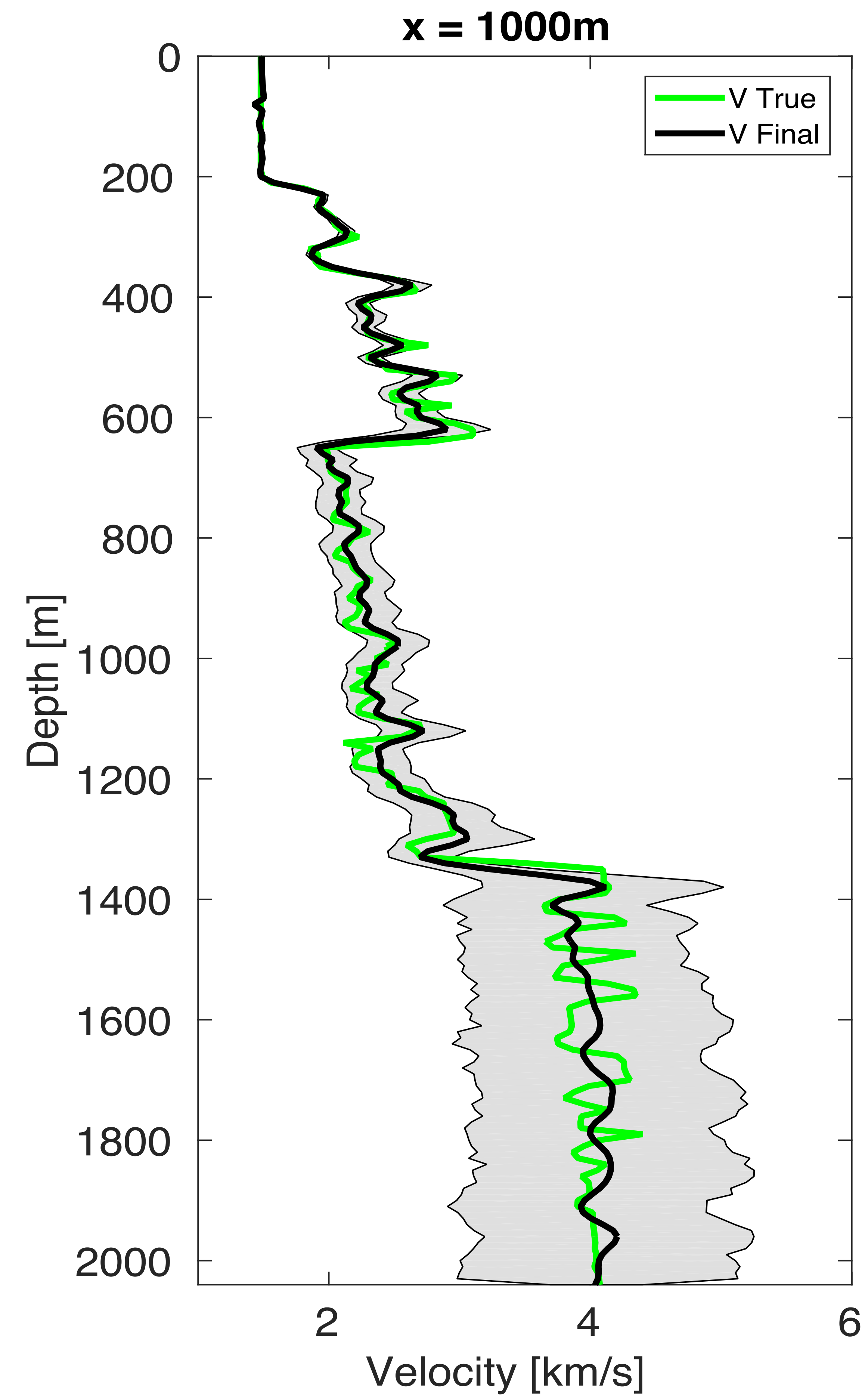
Motivation

Uncertainty of veolocity, density and elastic parameters → Uncertainty of porosity, saturation and permeability → Uncertainty of oil/gas volume

Motivation

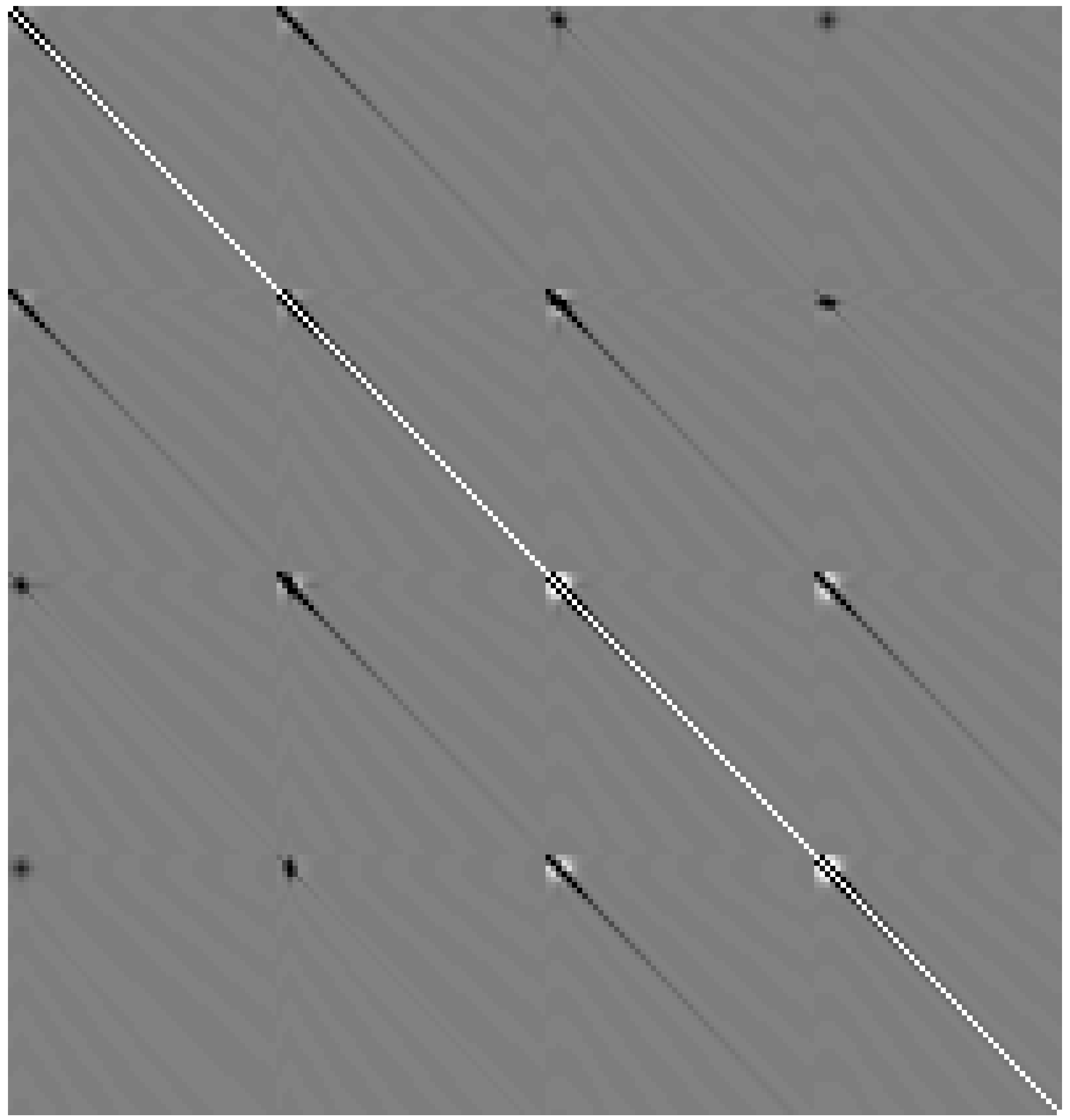


Inversion result



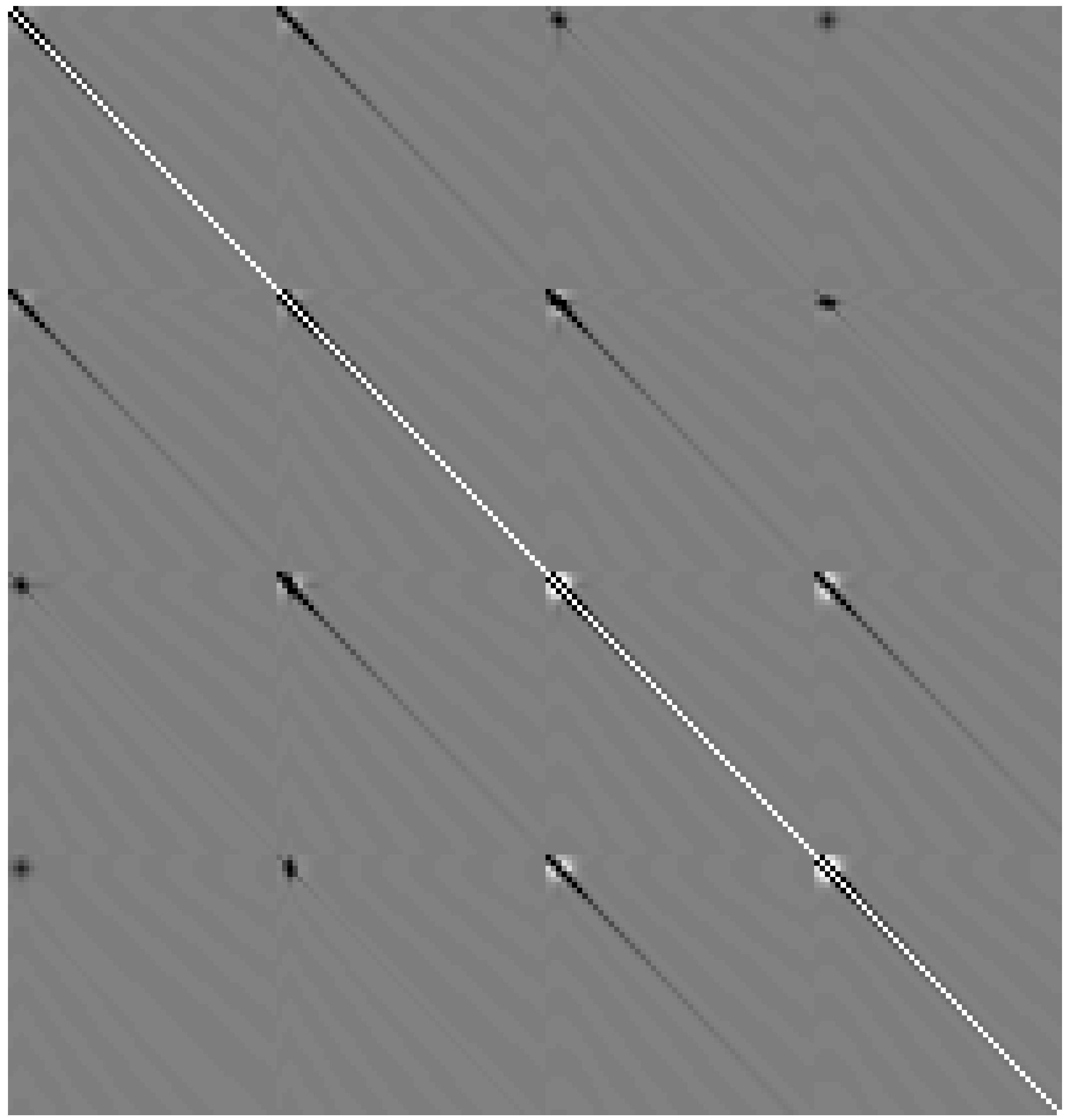
Confidence Interval

Motivation



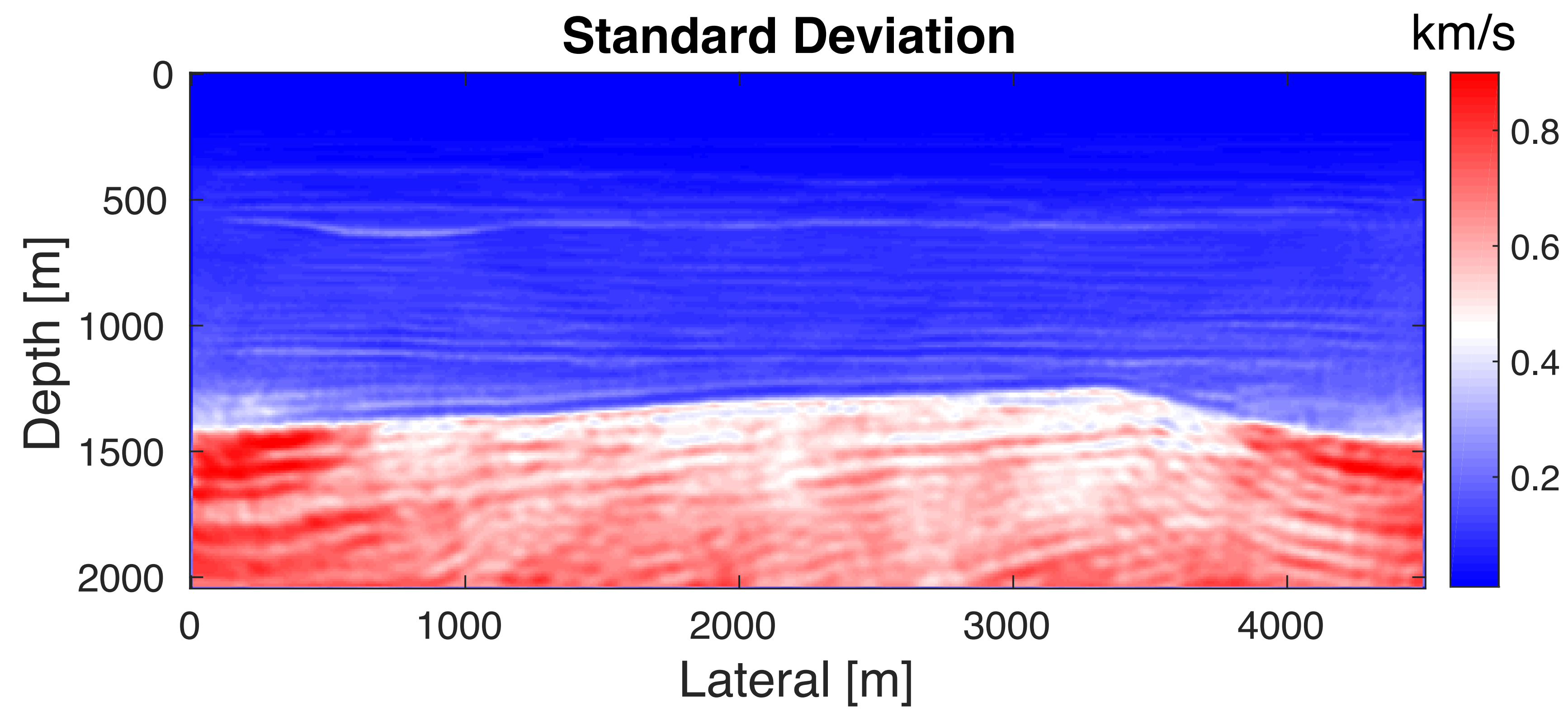
Hessian

Motivation



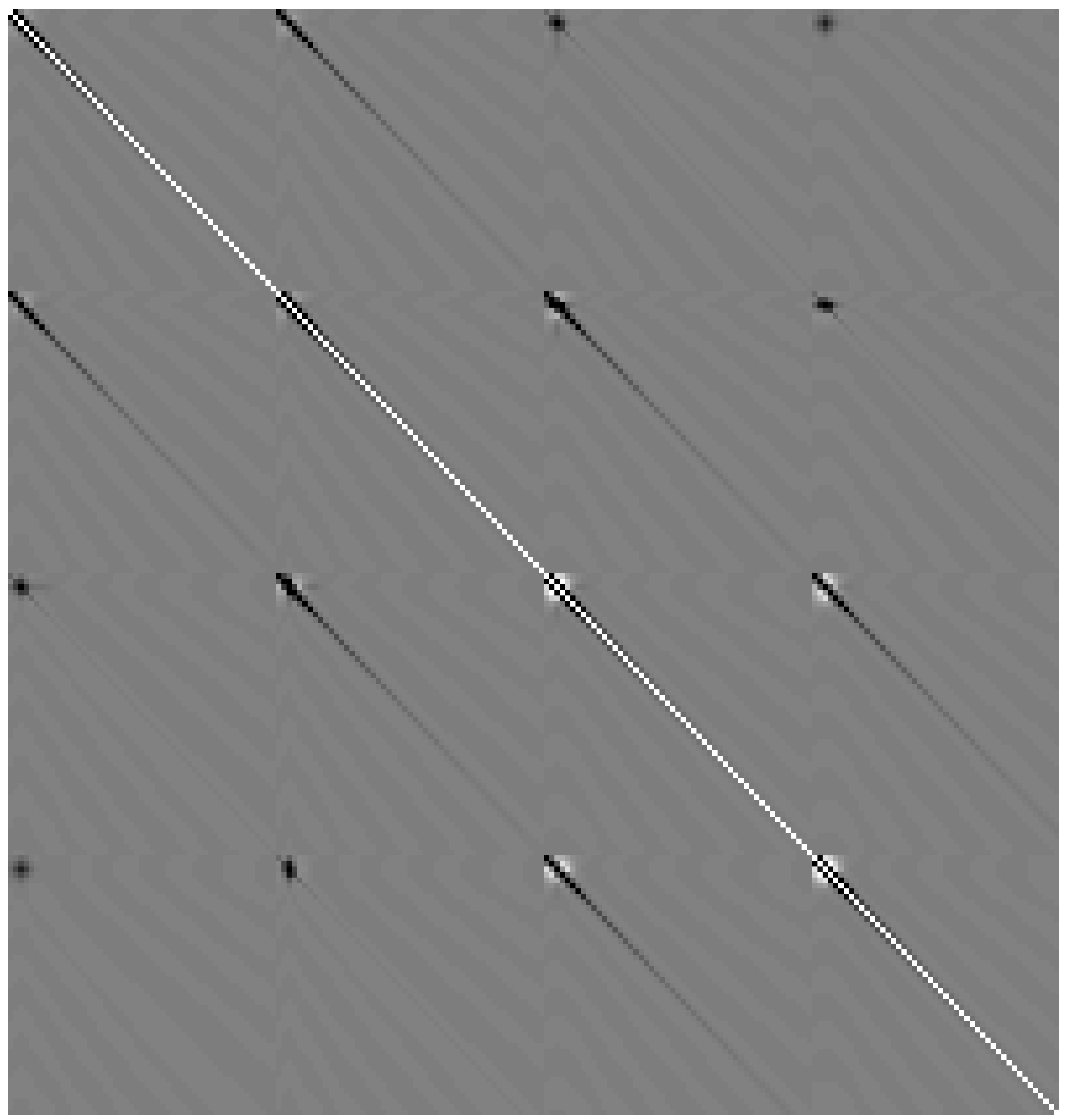
Hessian

statistical inversion



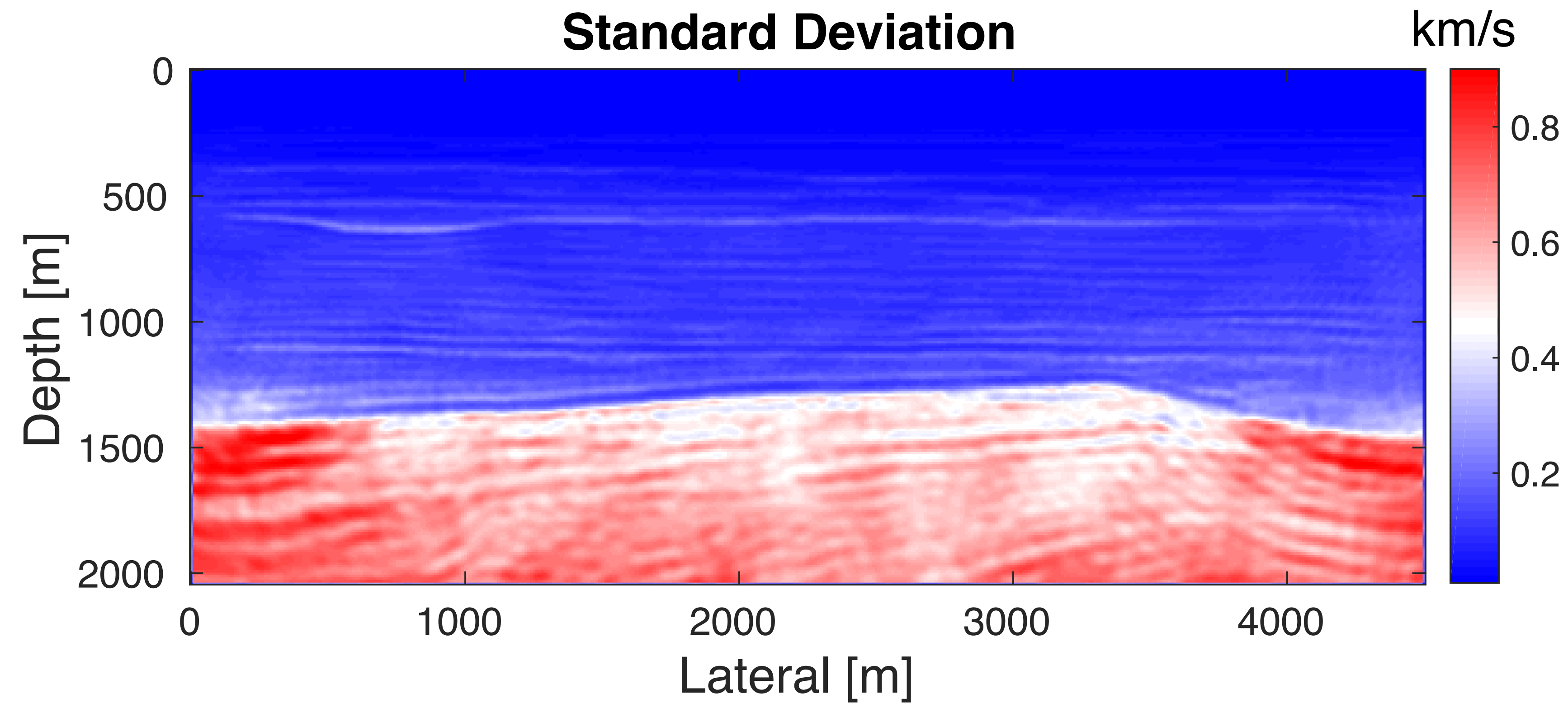
standard deviation

Motivation



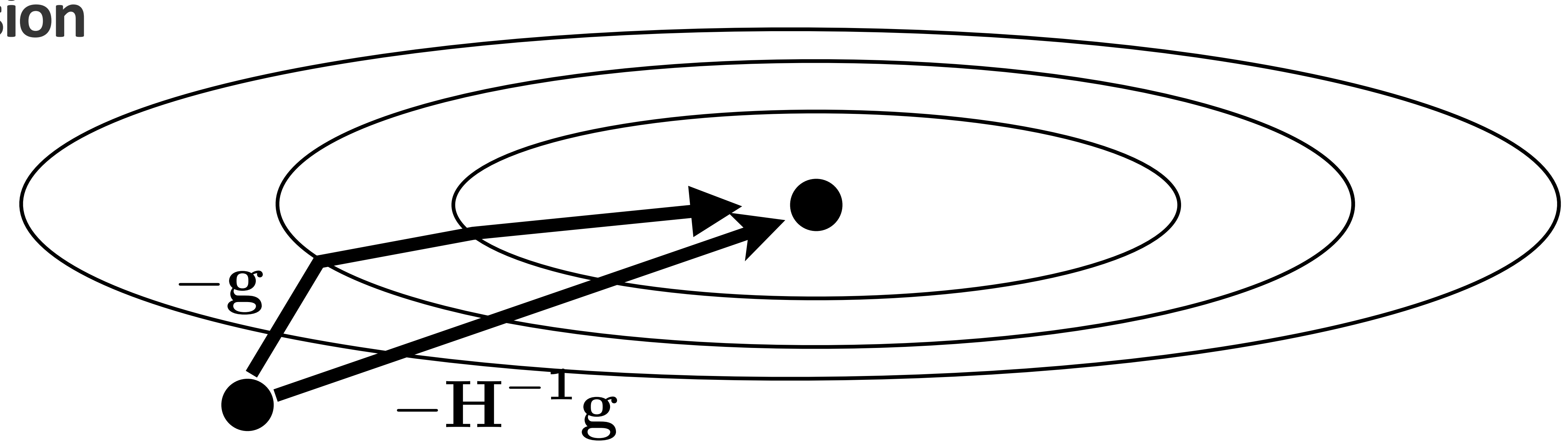
Hessian

statistical
inversion



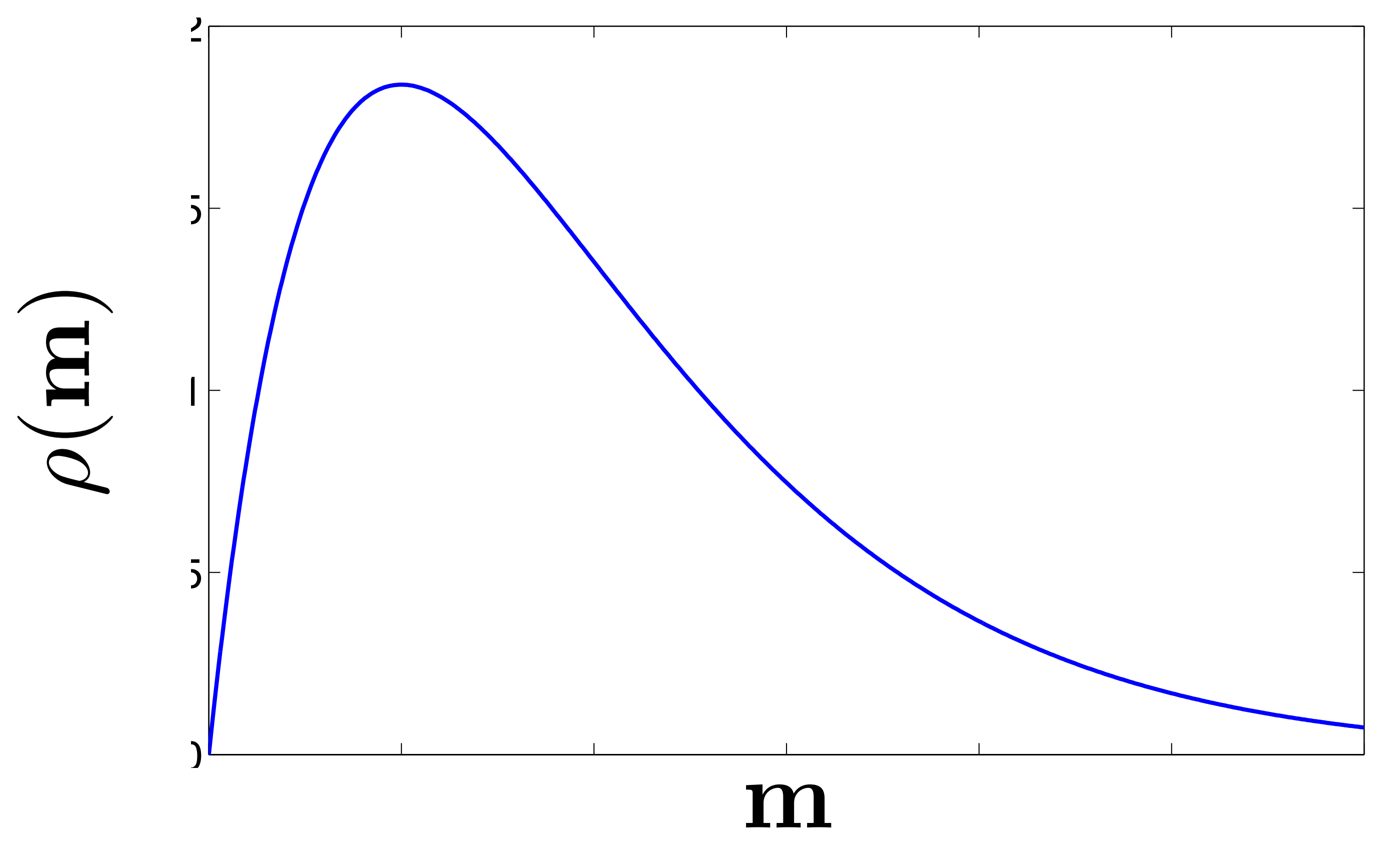
standard deviation

deterministic
inversion

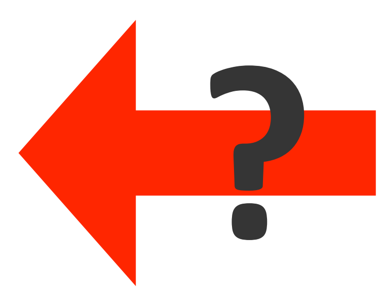


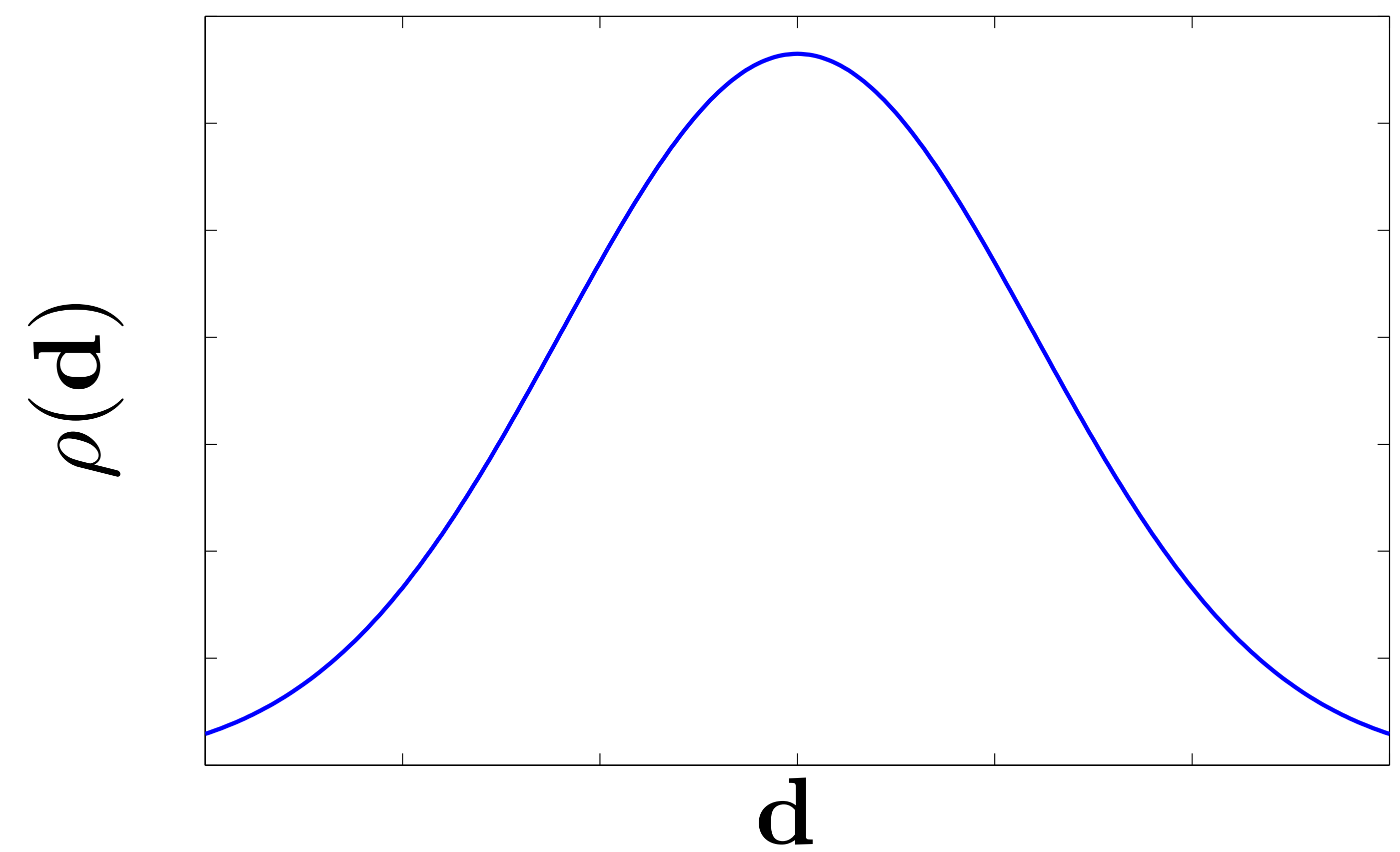
Newton type method

Bayesian theory



Distribution of model





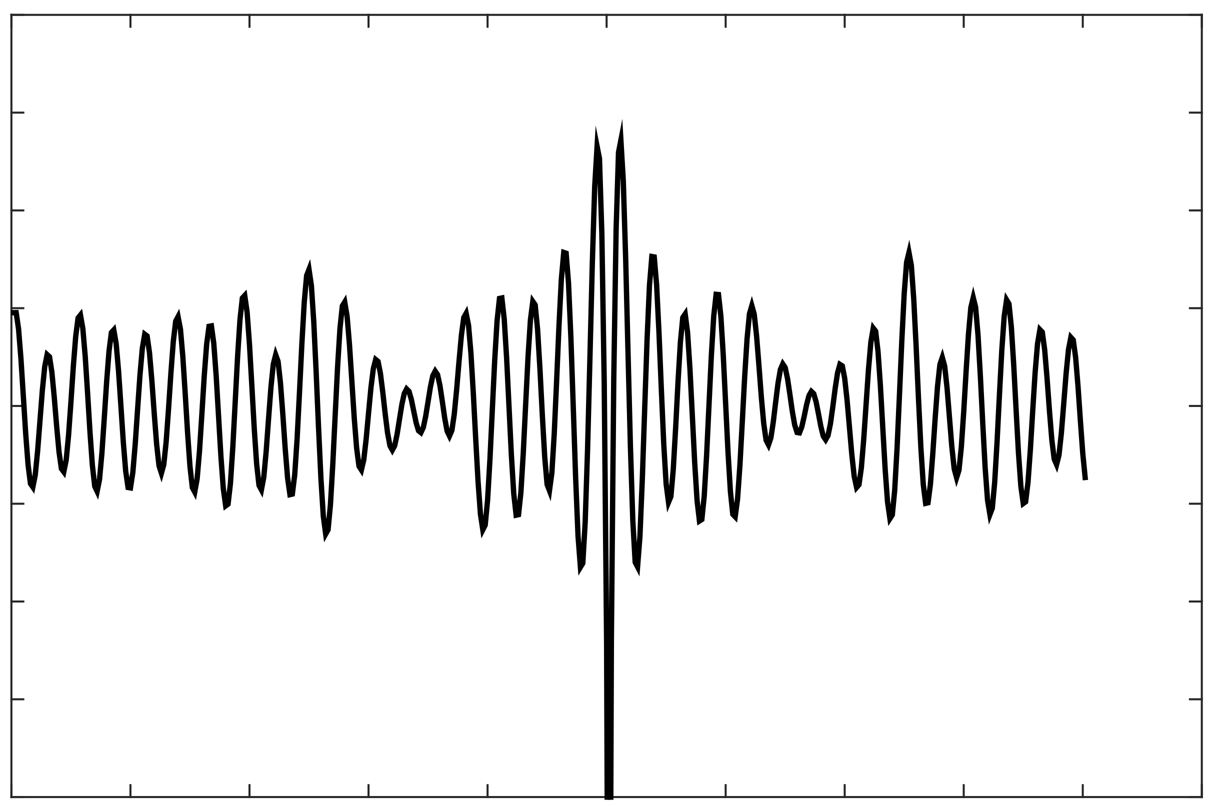
Distribution of data

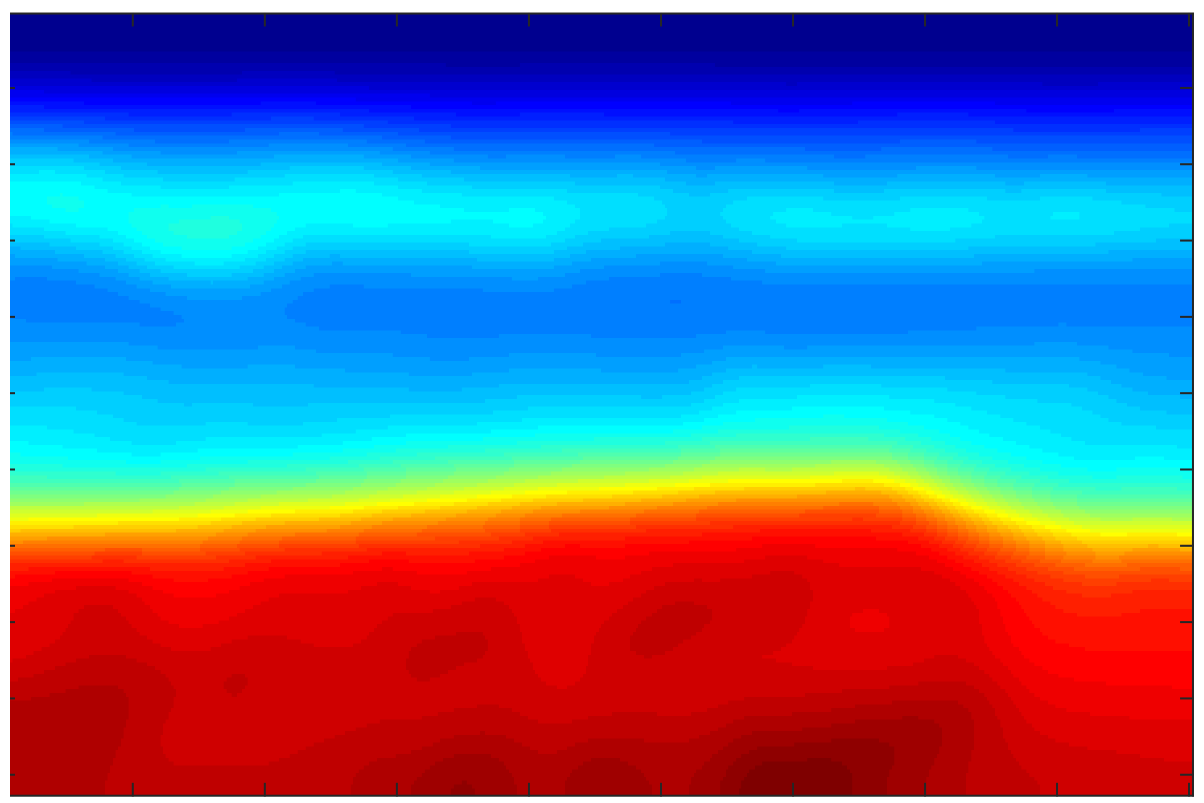
SLIM

Bayesian theory

Bayesian inference:

$$\rho_{\text{post}}(\mathbf{m}) \propto \rho_{\text{like}}(\mathbf{m}|\mathbf{d}_{obs})\rho_{\text{prior}}(\mathbf{m})$$



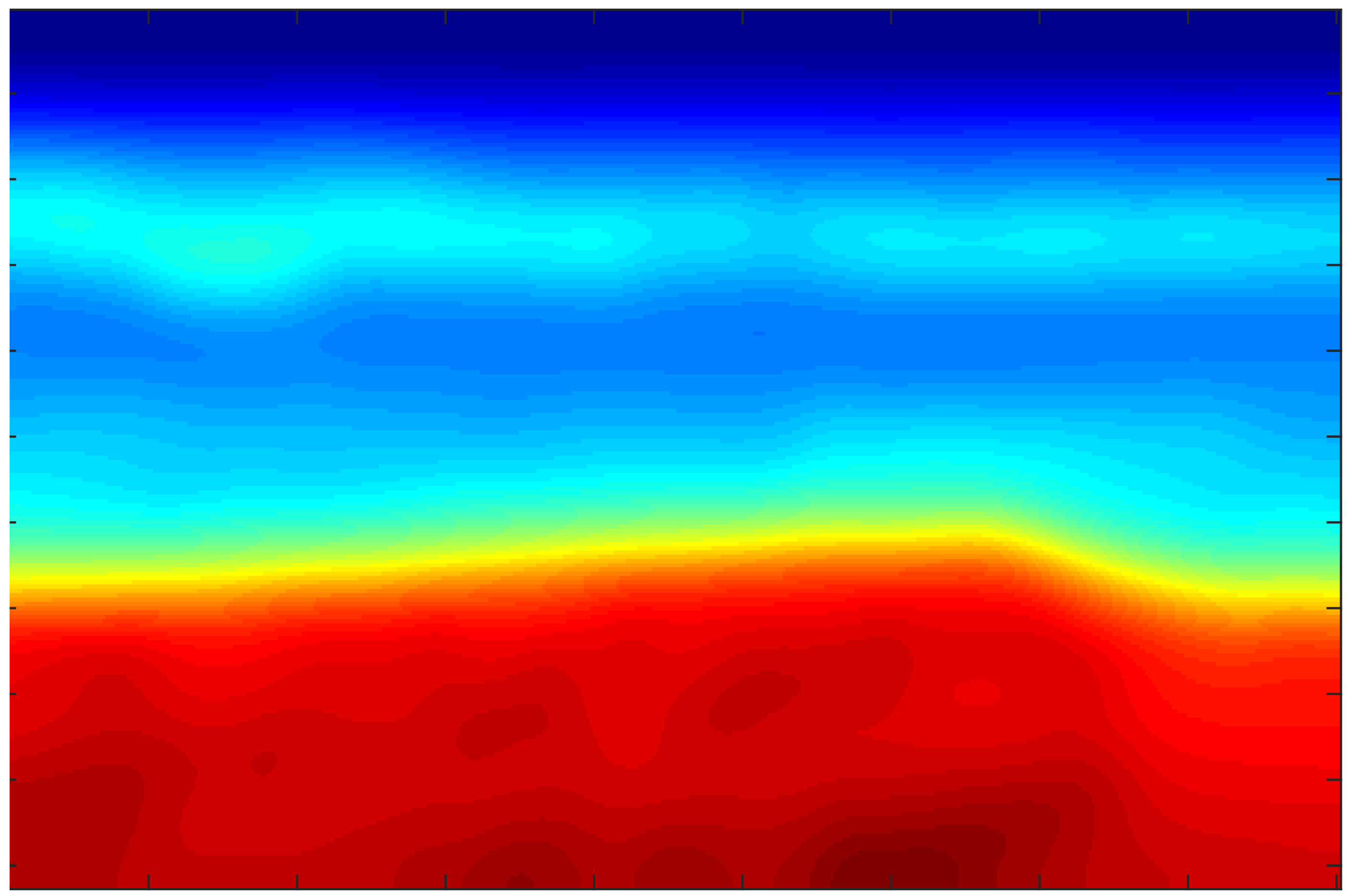


Bayesian theory

Prior information

Bayesian theory

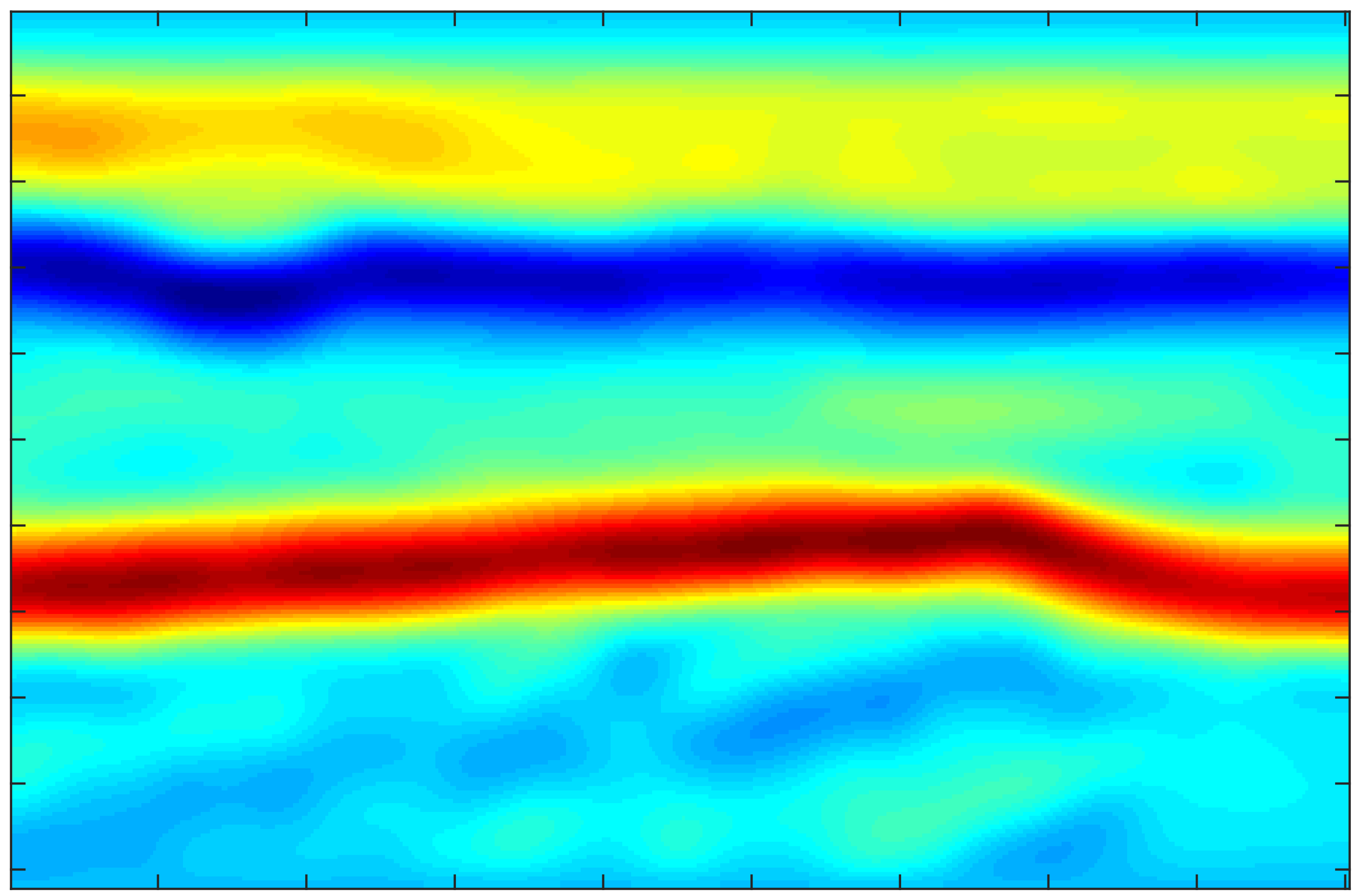
Prior information



$$\rho_{\mathrm{prior}}(\mathbf{m}) \propto \exp\left(-\|\mathbf{m} - \mathbf{m}_{\mathrm{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\mathrm{prior}}}\right)$$

Bayesian theory

Prior information

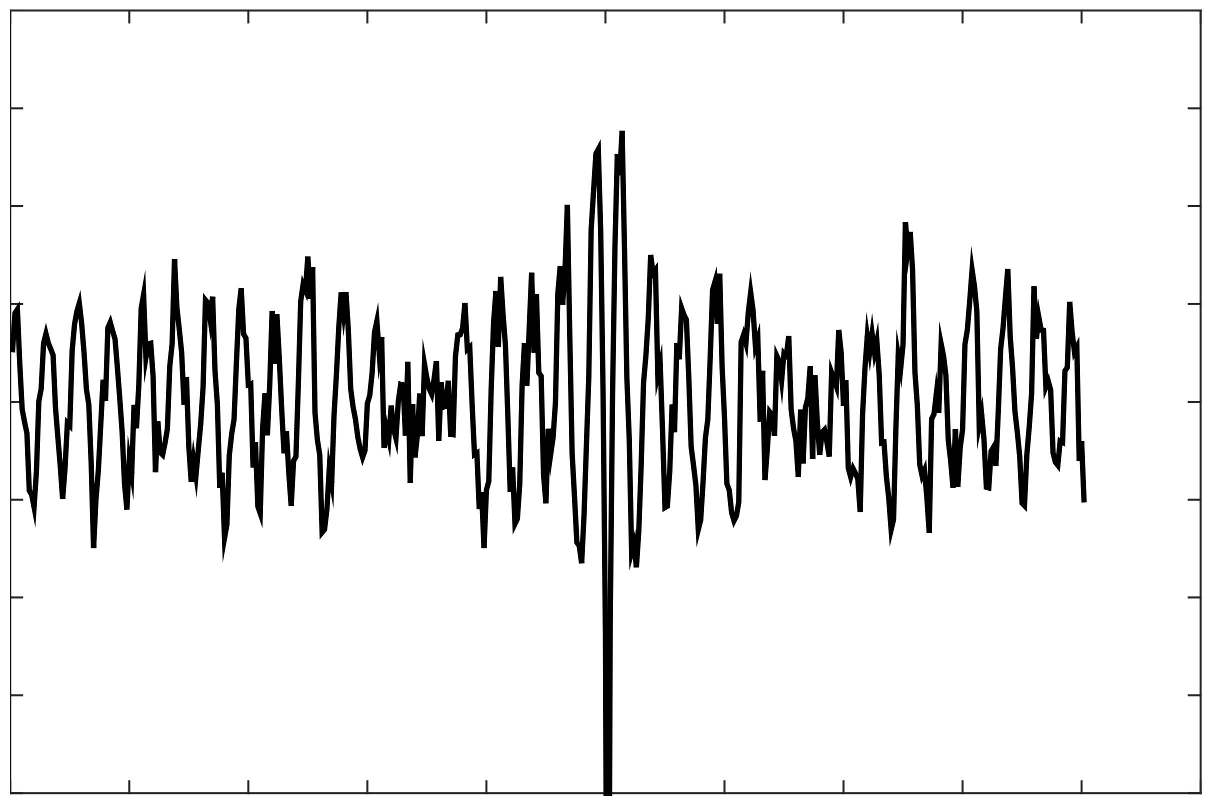


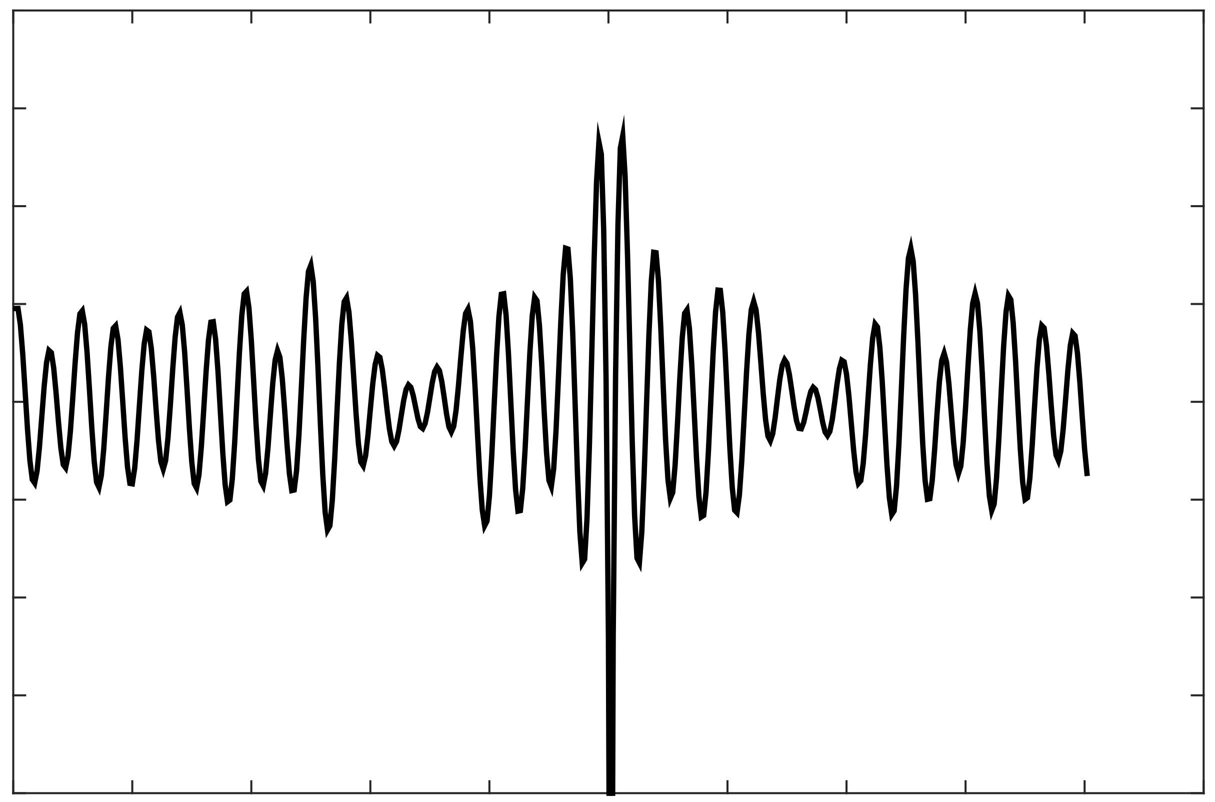
$$\rho_{\text{prior}}(\mathbf{m}) \propto \exp\left(-\|\mathbf{m}-\mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}}\right)$$

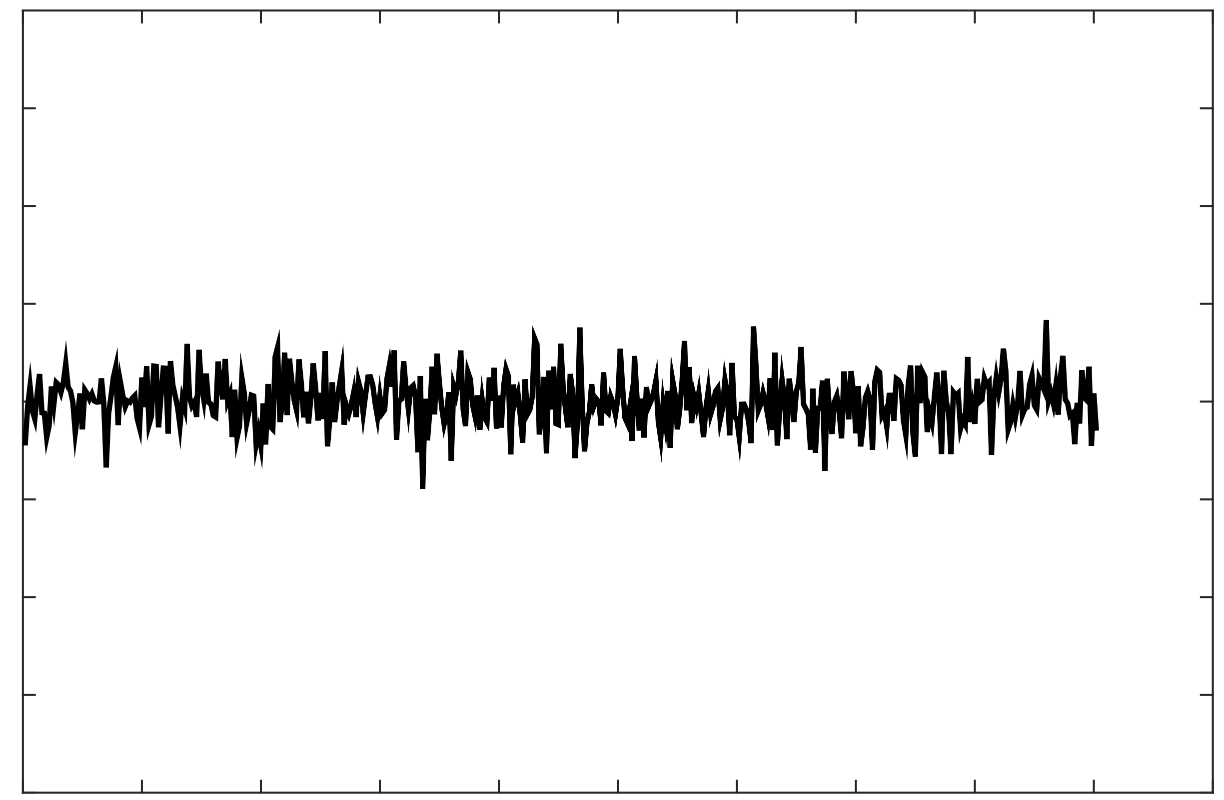
$$\rho_{\text{prior}}(\mathbf{m}) \propto \exp\left(-\|\nabla\mathbf{m}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}}\right)$$

Bayesian theory

Gaussian distribution assumption:



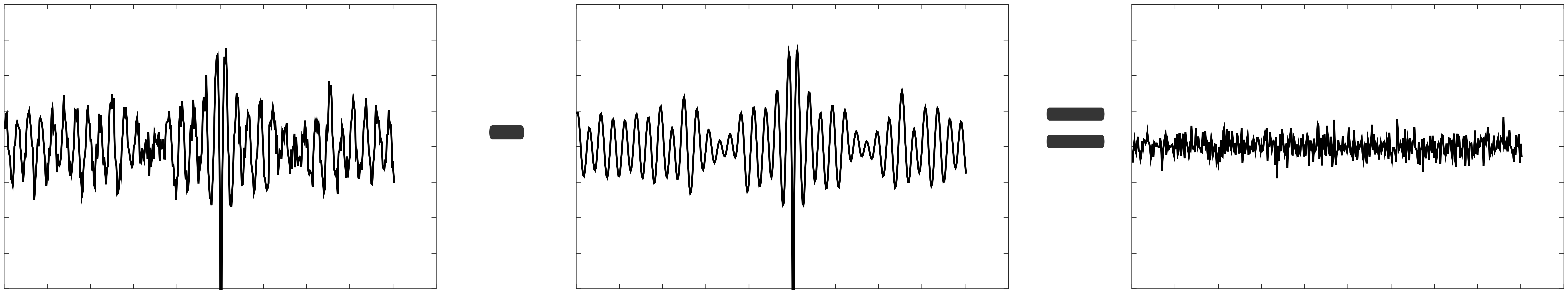




$$\mathbf{d}_{\text{obs}} = f(\mathbf{m}) + \text{noise}$$

Bayesian theory

Gaussian distribution assumption:



$\mathbf{d}_{\text{obs}}$ $f(\mathbf{m})$ noise

$$\rho_{\text{like}}(\mathbf{m}|\mathbf{d}_{\text{obs}}) \propto \exp\left(-\|f(\mathbf{m})-\mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}}\right)$$

Bayesian theory

Posterior distribution of FWI:

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

Bayesian theory

Posterior distribution of FWI:

$$\rho_{\text{post}}(\mathbf{m}) \propto \exp\left(-\underline{\|\mathbf{PA}(\mathbf{m})^{-1}\mathbf{q} - \mathbf{d}_{\text{obs}}\|}^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}} - \|\mathbf{m} - \mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}}\right)$$

Strong nonlinearity
Many local minima

SLIM

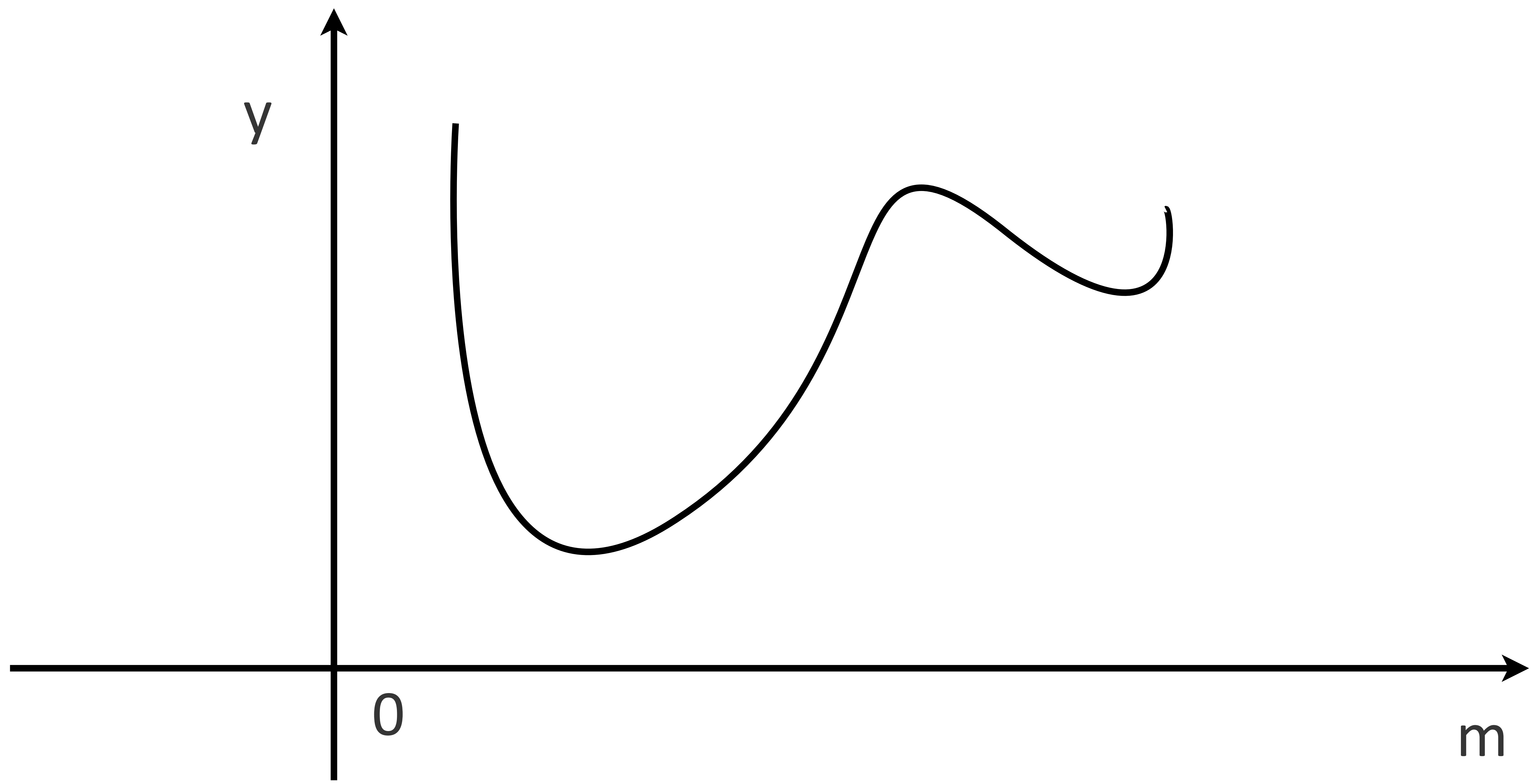
Bayesian theory

Posterior distribution of FWI:

$$\rho_{\text{post}}(\mathbf{m}) \propto \exp\left(-\underline{\|\mathbf{PA}(\mathbf{m})^{-1}\mathbf{q} - \mathbf{d}_{\text{obs}}\|}^2_{\mathbf{\Sigma}_{\text{noise}}^{-1}} - \|\mathbf{m} - \mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}_{\text{prior}}^{-1}}\right)$$

Strong nonlinearity

Many local minima

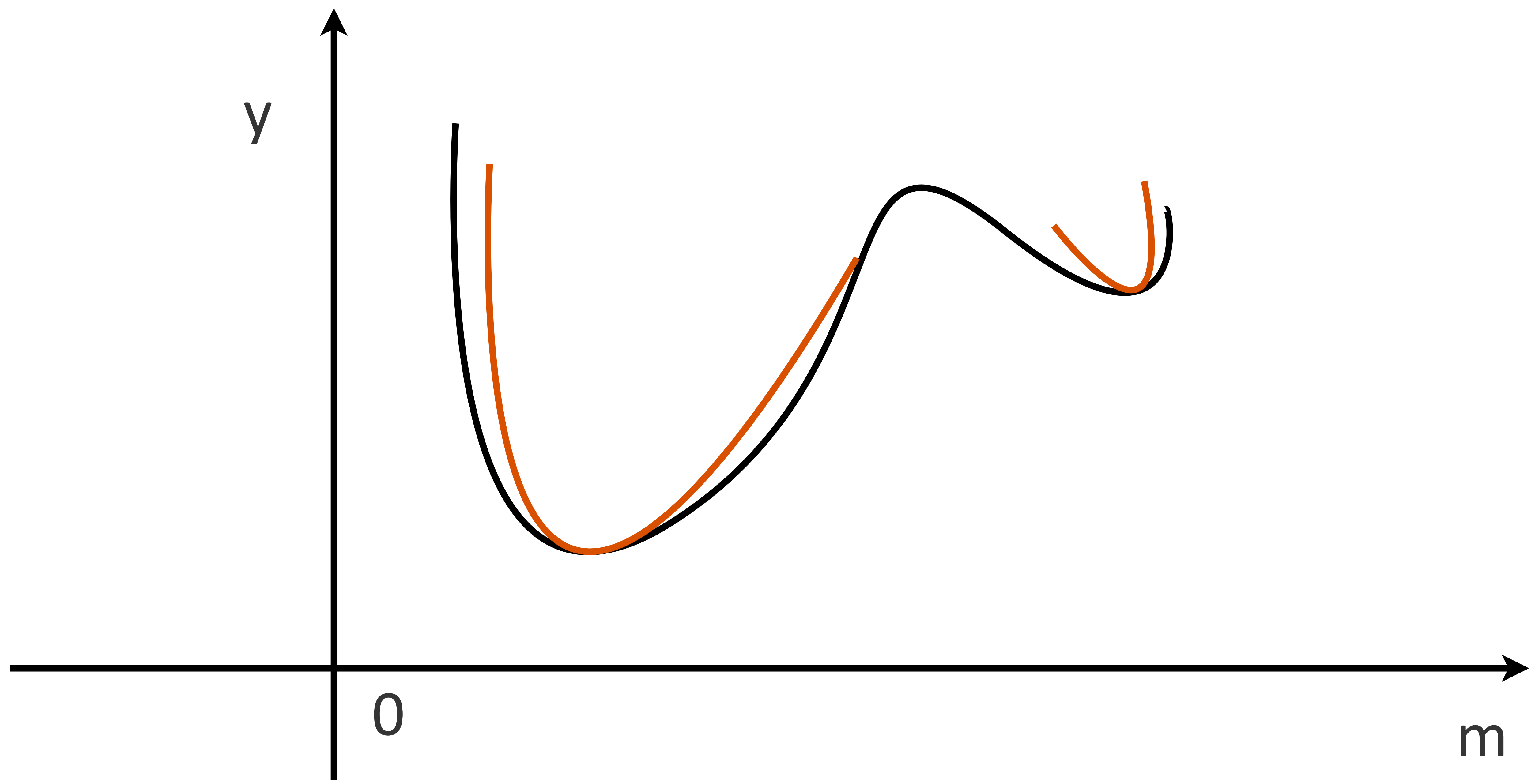


Bayesian theory

Posterior distribution of FWI:

$$\rho_{\mathrm{post}}(\mathbf{m}) \propto \exp\left(-\|\underline{\mathbf{PA}(\mathbf{m})^{-1}\mathbf{q} - \mathbf{d}_{\mathrm{obs}}}\|^2_{\mathbf{\Sigma}^{-1}_{\mathrm{noise}}} - \|\mathbf{m} - \mathbf{m}_{\mathrm{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\mathrm{prior}}}\right)$$

Strong nonlinearity
Many local minima



SLIM

Bayesian and WRI

Posterior distribution of WRI:

$$\rho_{\text{post}}(\mathbf{m}) \propto$$

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

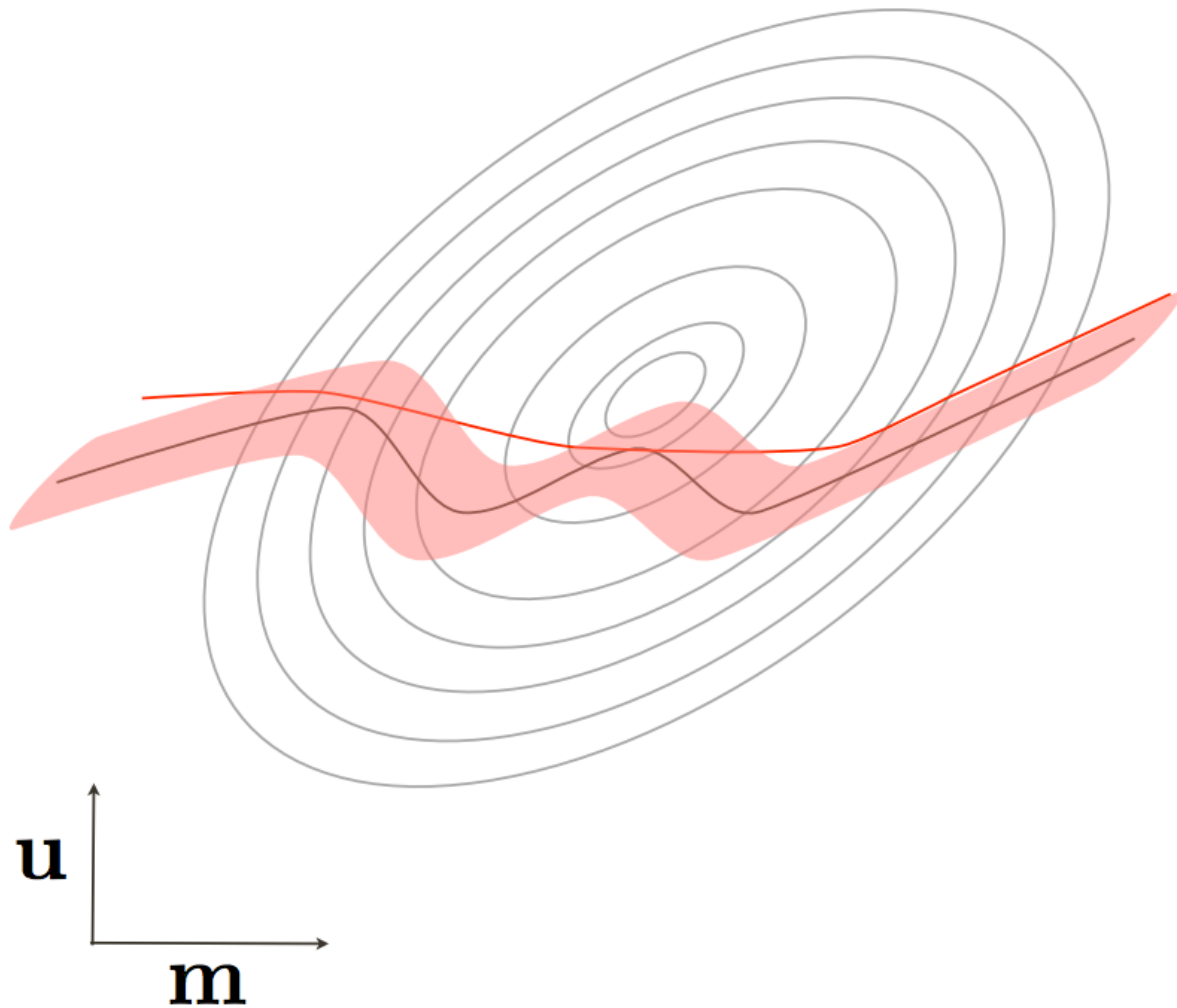
where,

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

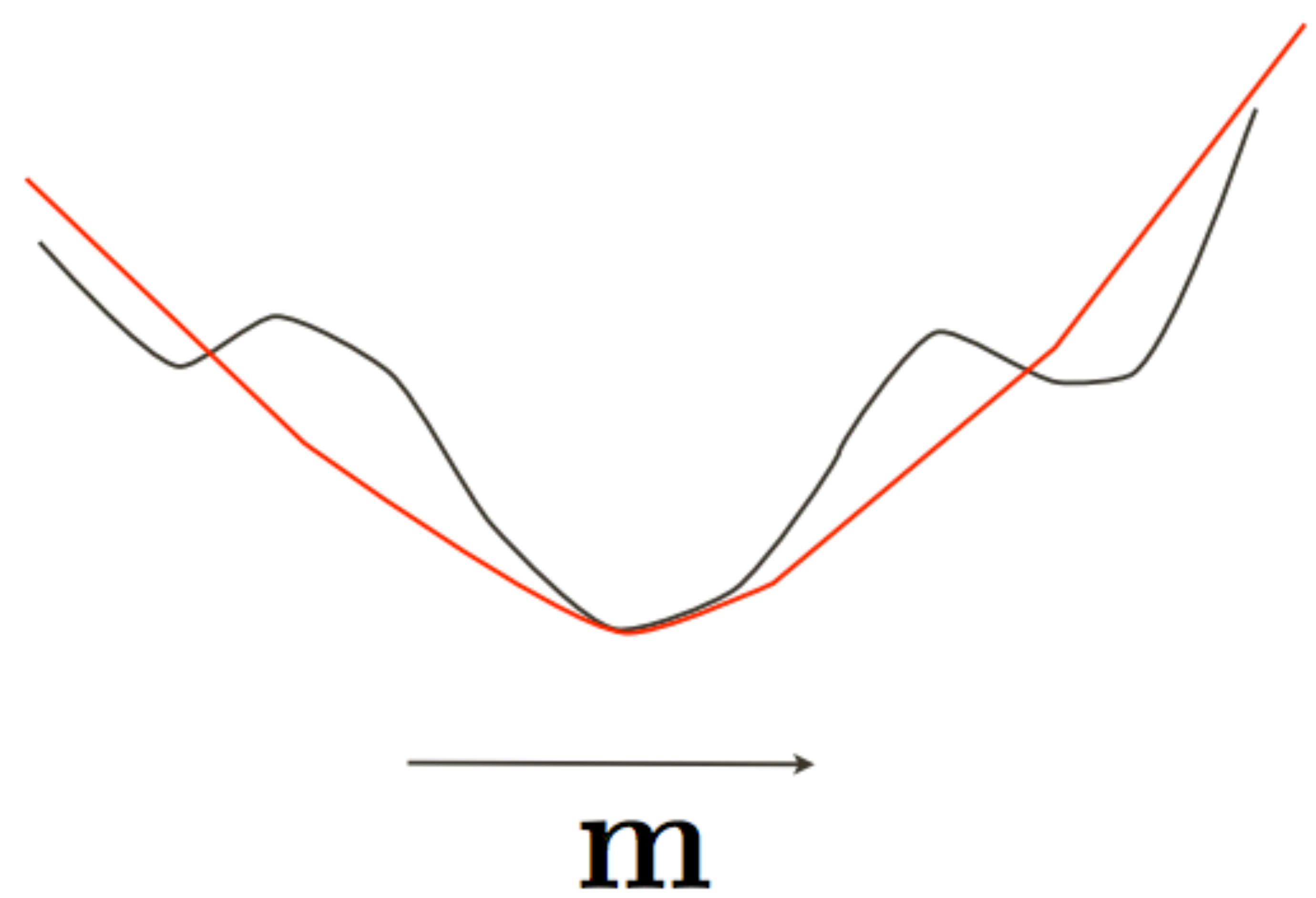
WRI vs FWI

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

Larger # of degrees of freedom



"more convex"



SLIM

Bayesian and WRI

Mean and covariance of the model:

$$\mathbb{E}(\mathbf{m}) = \int \mathbf{m}\rho_{\text{post}}(\mathbf{m})d\mathbf{m}$$

$$\text{Cov}(\mathbf{m}) = \int (\mathbb{E}(\mathbf{m}) - \mathbf{m})^2 \rho_{\text{post}}(\mathbf{m})d\mathbf{m}$$

Quantify the uncertainty

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

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 $\boxed{\text{Huge computational cost!!!}}$

Quantify the uncertainty

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

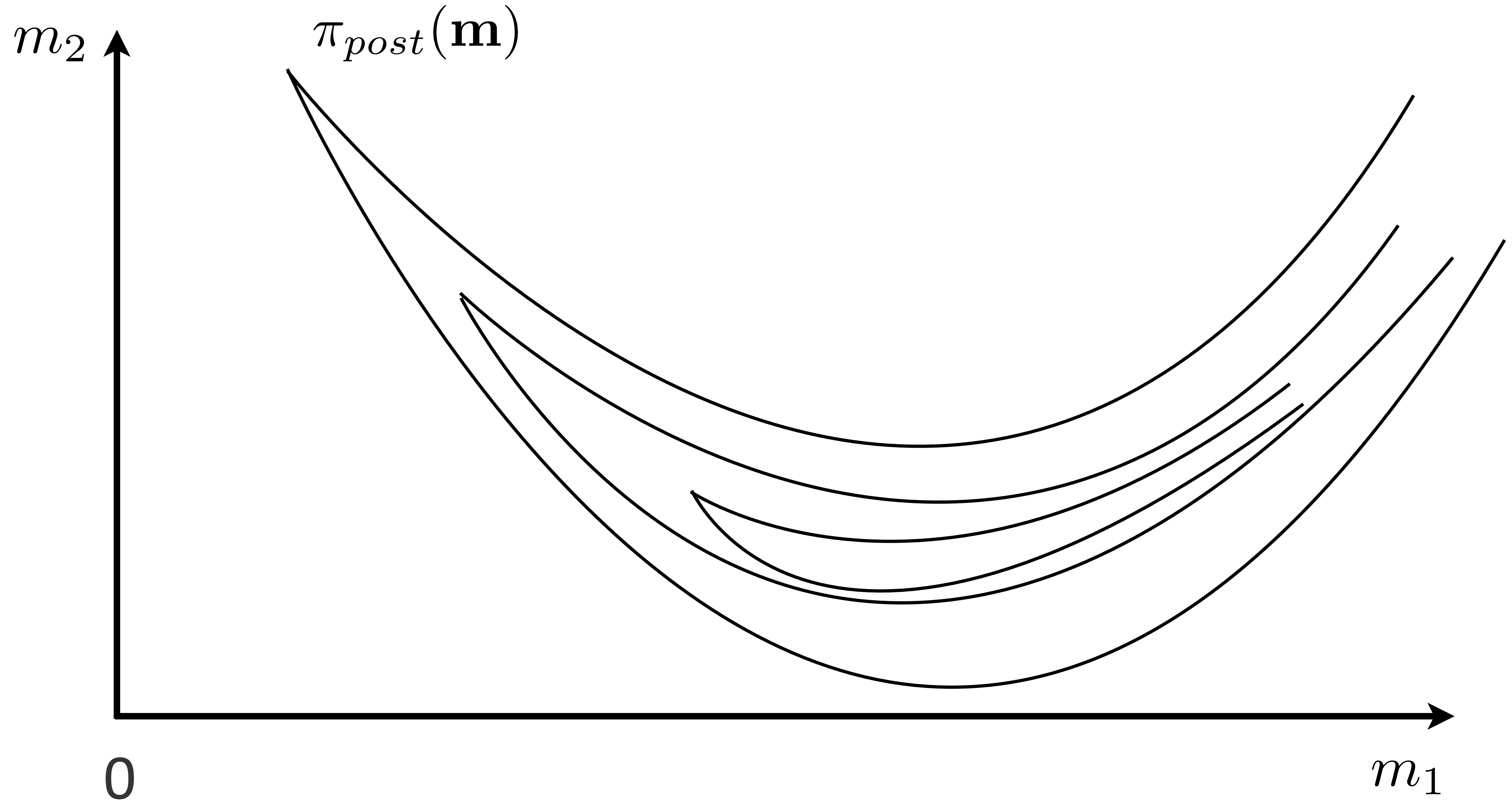
Solution:

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

SLIM

McMC method

Metropolis-Hasting Method:



At sample $\mathbf{m}_k$

Draw sample $\mathbf{y}$ from the proposal distribution $\tilde{\pi}_k(\mathbf{m})$

if $\min(1, \dfrac{\pi_{post}(\mathbf{y})\tilde{\pi}_y(\mathbf{m}_k)}{\pi_{post}(\mathbf{m}_k)\tilde{\pi}_k(\mathbf{y})}) > \alpha$

    set $\mathbf{m}_{k+1} = \mathbf{y}$

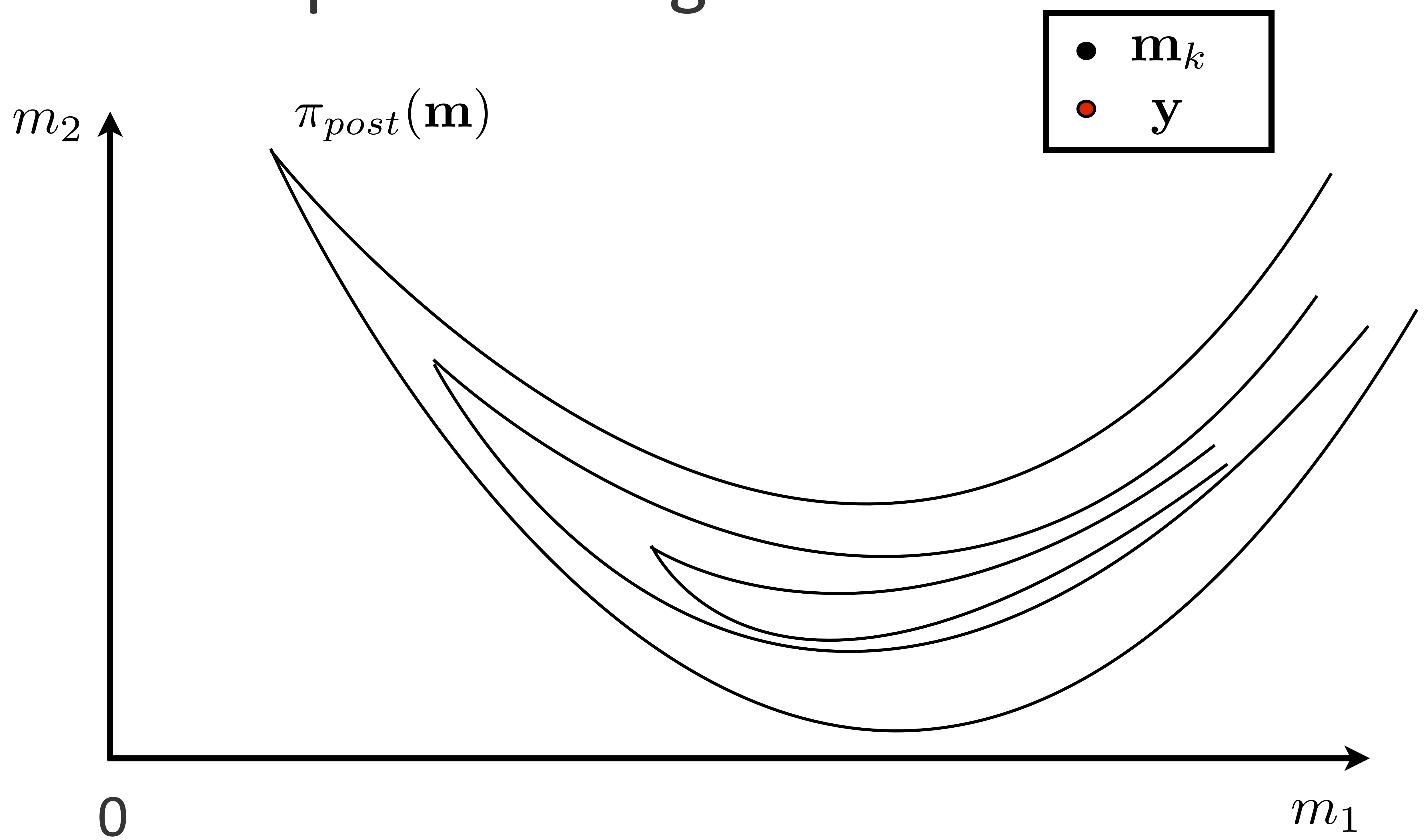
else

    regenerate $\mathbf{y}$

end

McMC method

Metropolis-Hasting Method:



At sample $\mathbf{m}_k$

Draw sample $\mathbf{y}$ from the proposal distribution $\tilde{\pi}_k(\mathbf{m})$

if $\min(1, \dfrac{\pi_{post}(\mathbf{y})\tilde{\pi}_y(\mathbf{m}_k)}{\pi_{post}(\mathbf{m}_k)\tilde{\pi}_k(\mathbf{y})}) > \alpha$

  set $\mathbf{m}_{k+1} = \mathbf{y}$

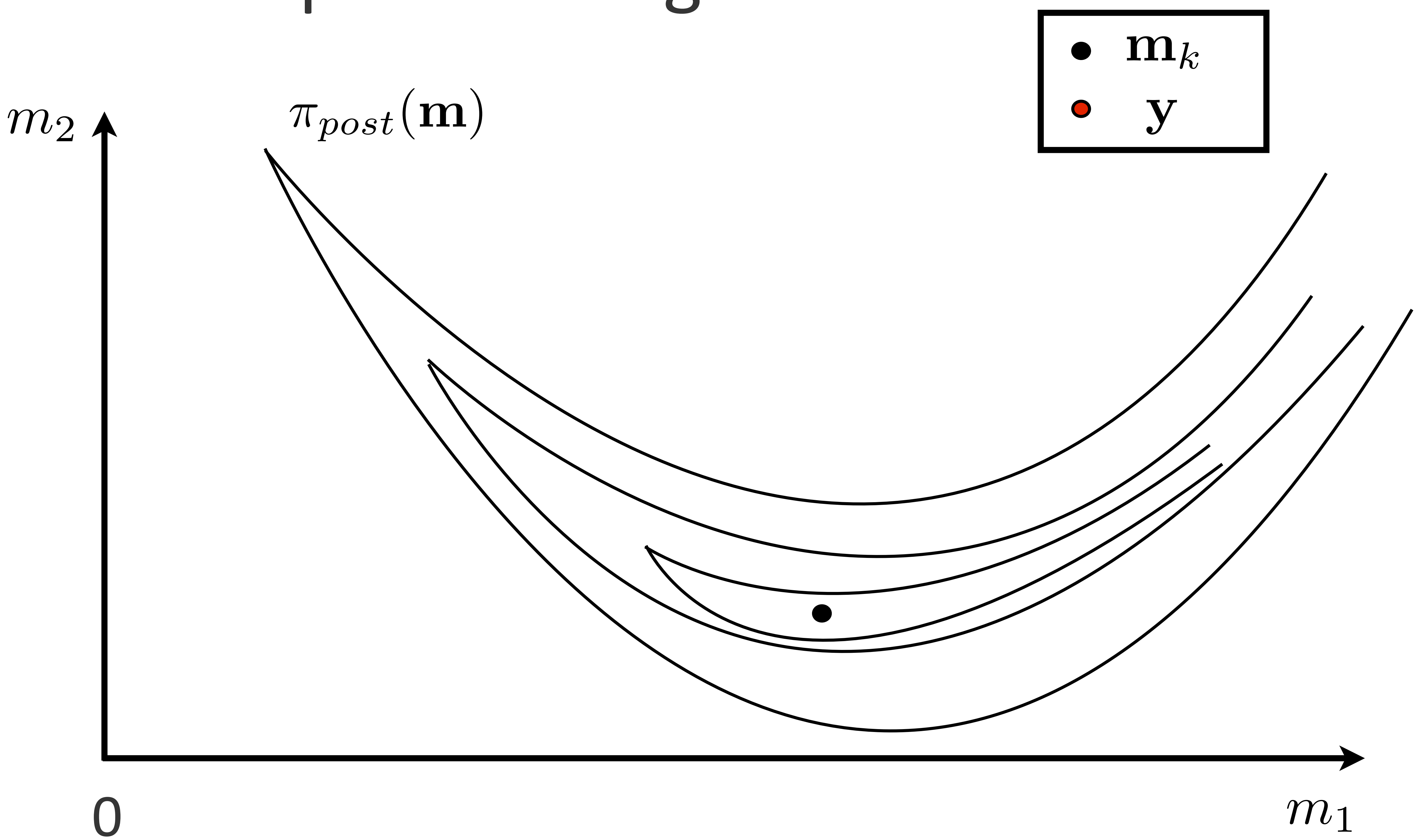
else

  regenerate $\mathbf{y}$

end

McMC method

Metropolis-Hasting Method:



At sample $\mathbf{m}_k$

Draw sample $\mathbf{y}$ from the proposal distribution $\tilde{\pi}_k(\mathbf{m})$

if $\min(1, \dfrac{\pi_{post}(\mathbf{y})\tilde{\pi}_y(\mathbf{m}_k)}{\pi_{post}(\mathbf{m}_k)\tilde{\pi}_k(\mathbf{y})}) > \alpha$

  set $\mathbf{m}_{k+1} = \mathbf{y}$

else

  regenerate $\mathbf{y}$

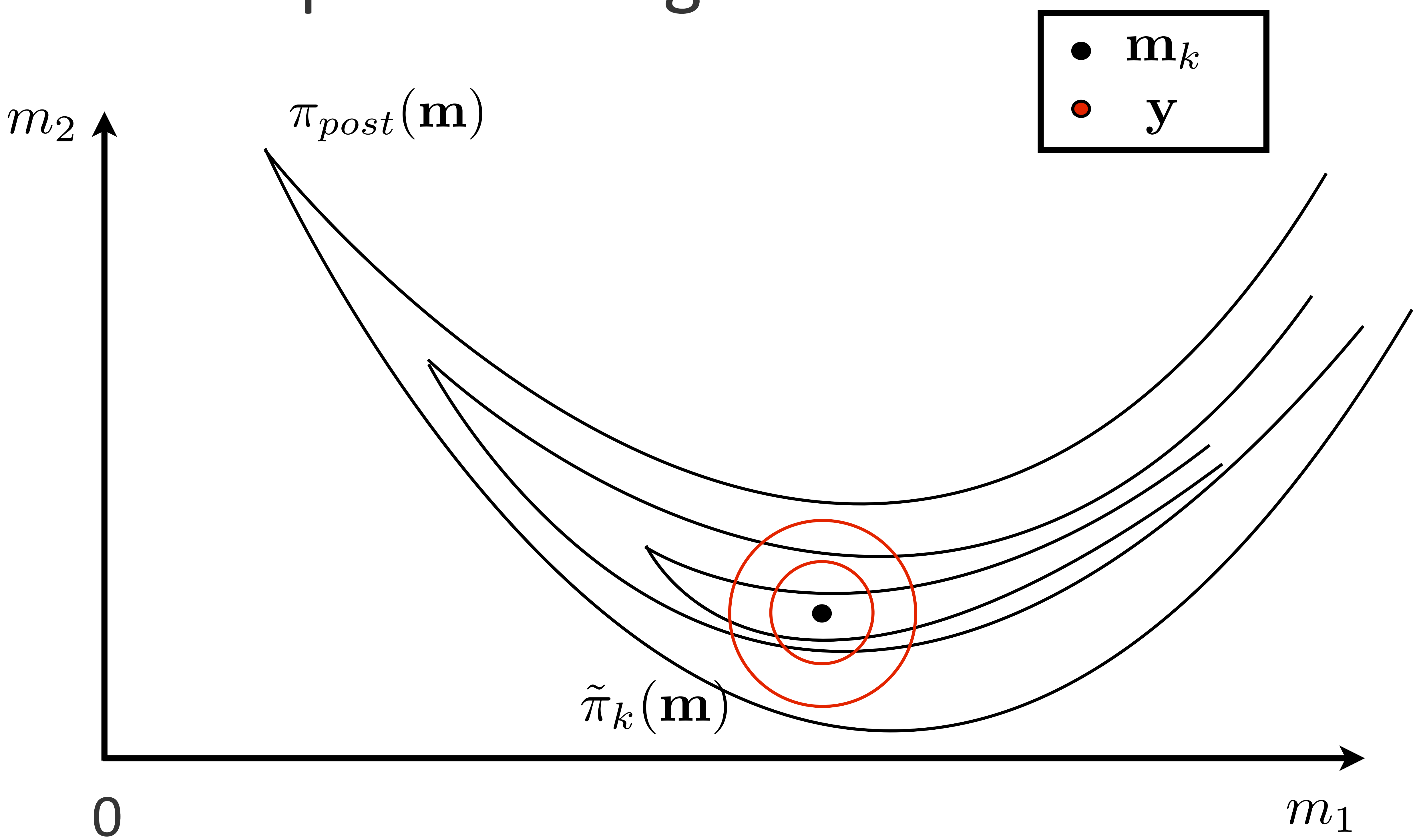
end

[D. Higdon, *et al*, 2010]

SLIM

McMC method

Metropolis-Hasting Method:



At sample $\mathbf{m}_k$

Draw sample $\mathbf{y}$ from the proposal distribution $\tilde{\pi}_k(\mathbf{m})$

if $\min(1, \dfrac{\pi_{post}(\mathbf{y})\tilde{\pi}_y(\mathbf{m}_k)}{\pi_{post}(\mathbf{m}_k)\tilde{\pi}_k(\mathbf{y})}) > \alpha$

    set $\mathbf{m}_{k+1} = \mathbf{y}$

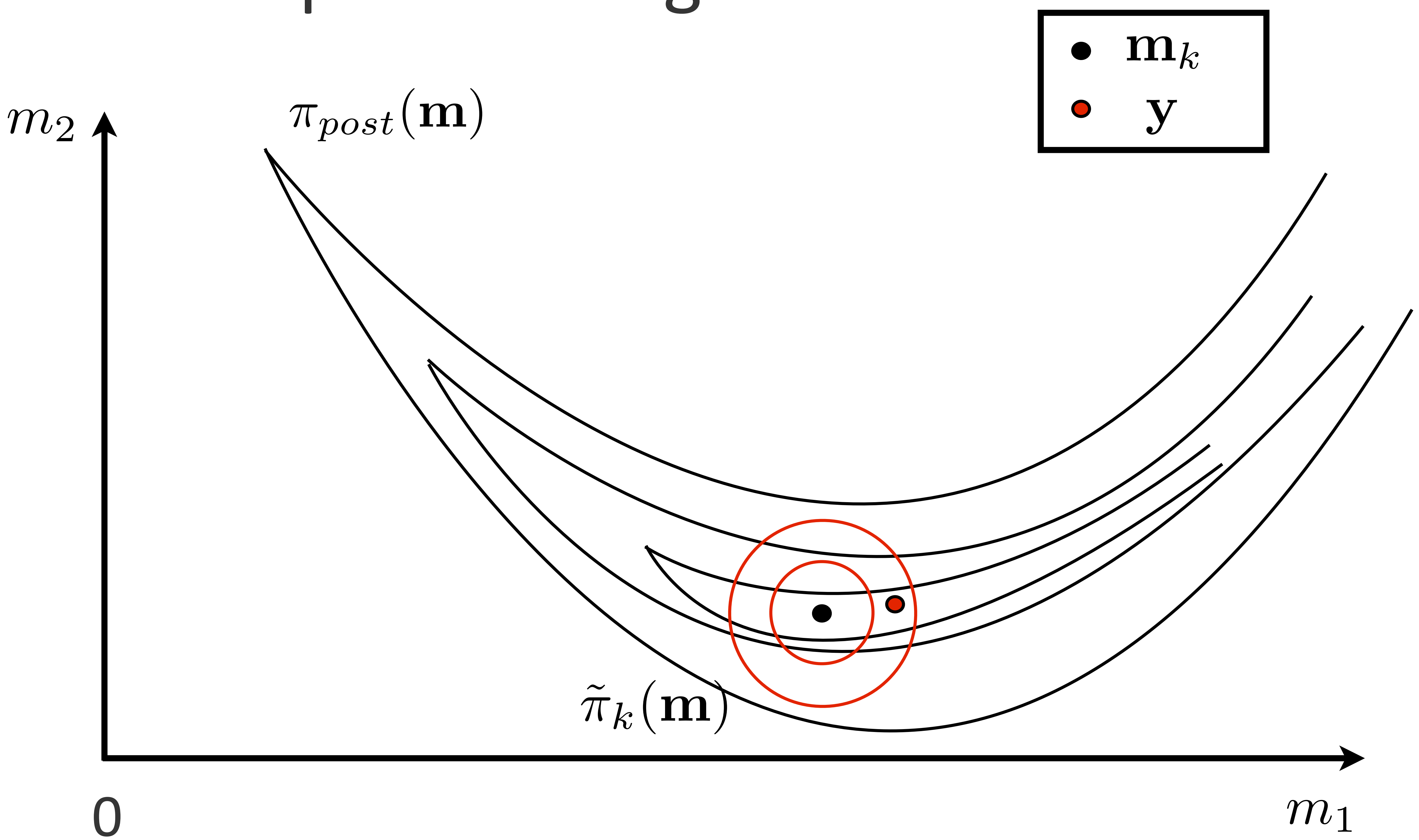
else

    regenerate $\mathbf{y}$

end

McMC method

Metropolis-Hasting Method:



At sample $\mathbf{m}_k$

Draw sample $\mathbf{y}$ from the proposal distribution $\tilde{\pi}_k(\mathbf{m})$

if $\min(1, \frac{\pi_{post}(\mathbf{y})\tilde{\pi}_y(\mathbf{m}_k)}{\pi_{post}(\mathbf{m}_k)\tilde{\pi}_k(\mathbf{y})}) > \alpha$

  set $\mathbf{m}_{k+1} = \mathbf{y}$

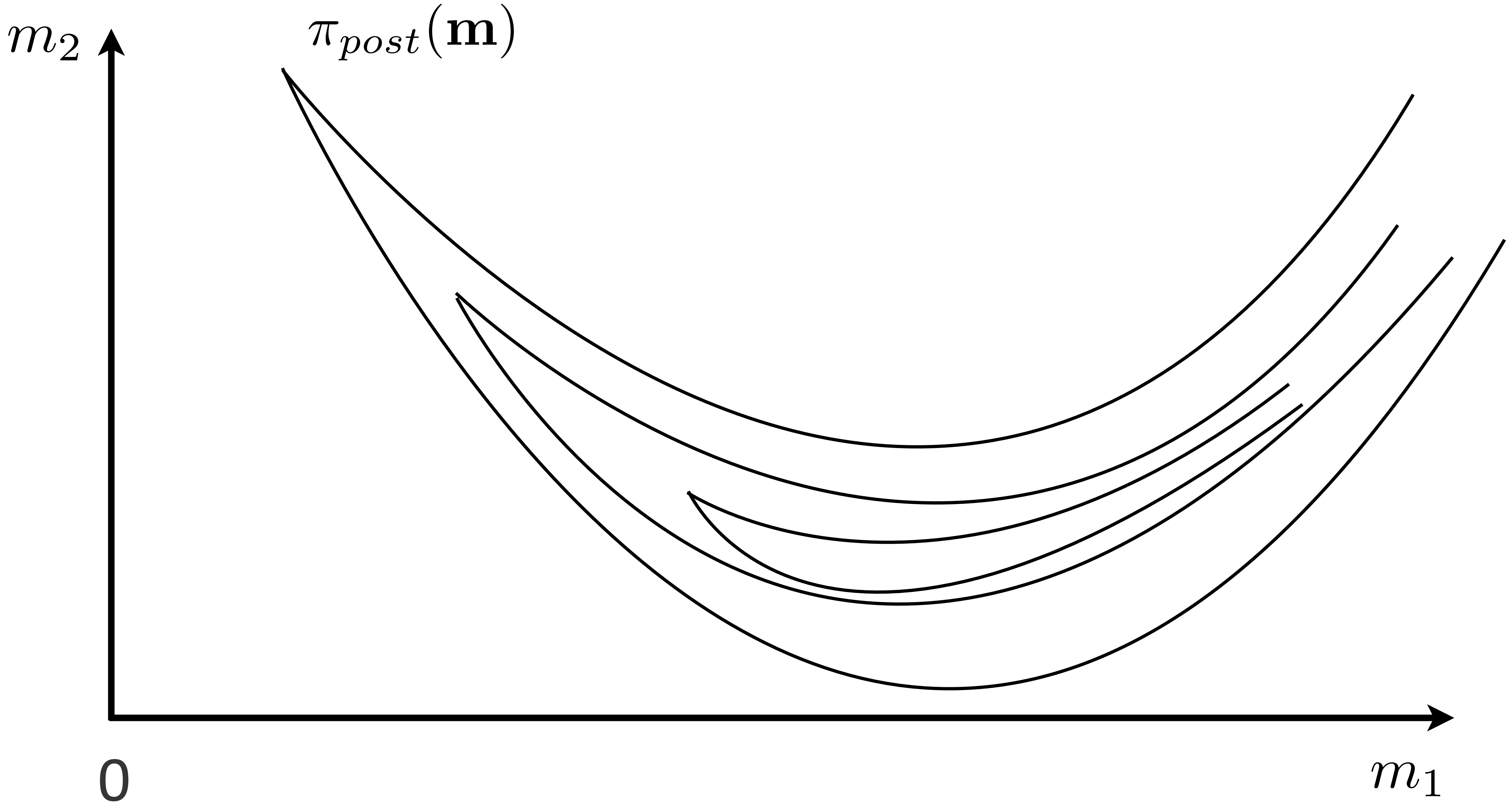
else

  regenerate $\mathbf{y}$

end

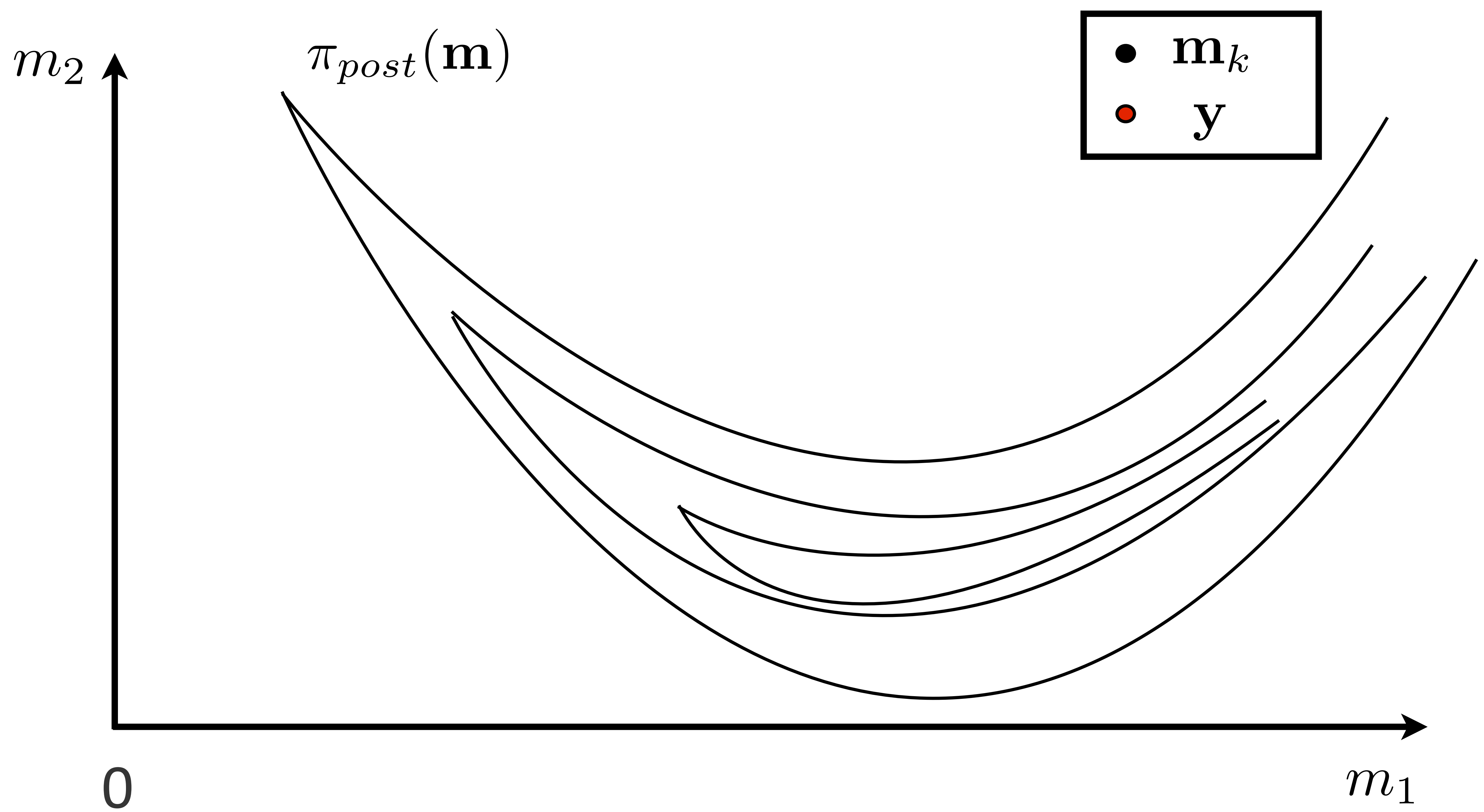
McMC method

Random Walk: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k, \alpha\mathbf{I})$



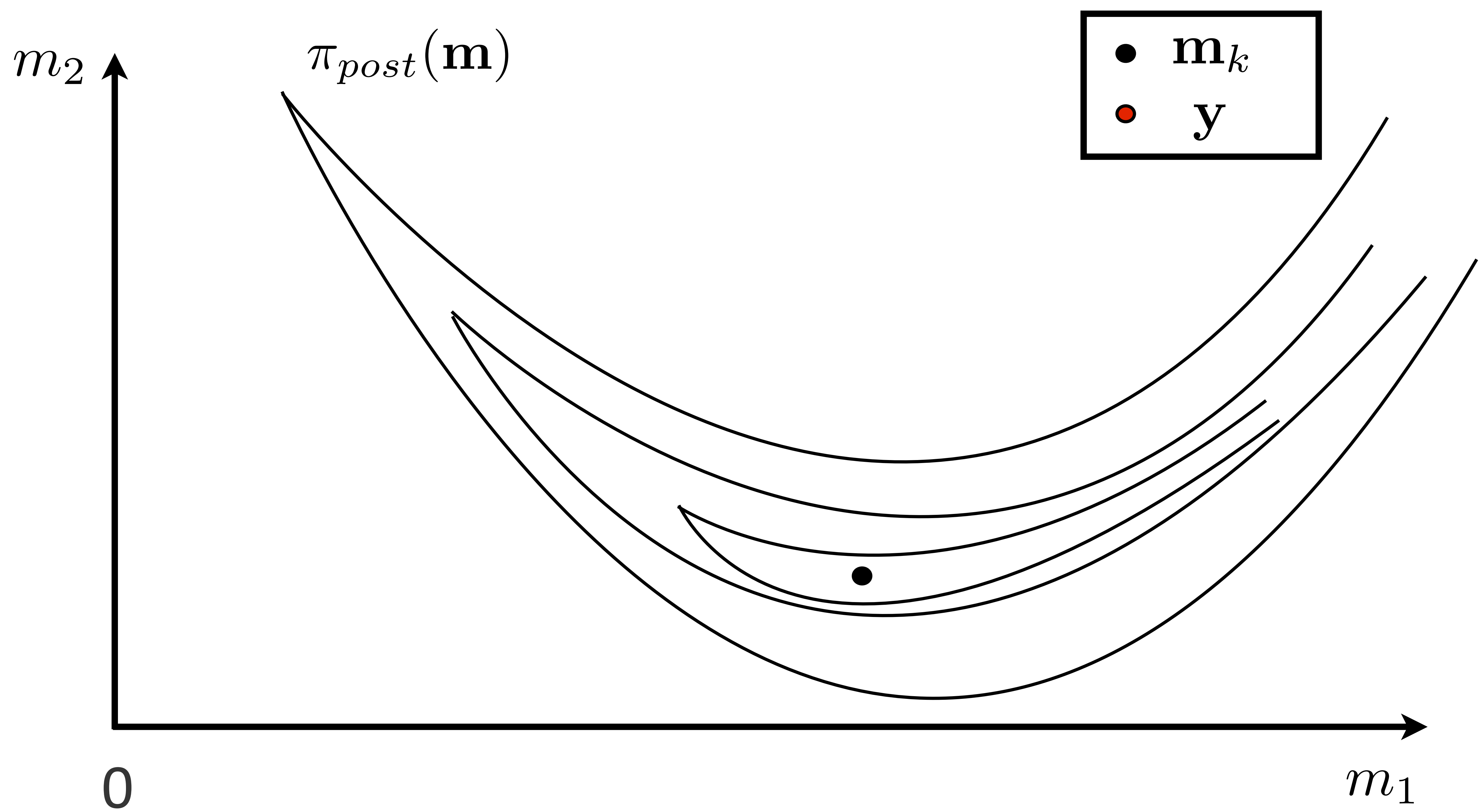
McMC method

Random Walk: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k, \alpha\mathbf{I})$



McMC method

Random Walk: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k, \alpha\mathbf{I})$

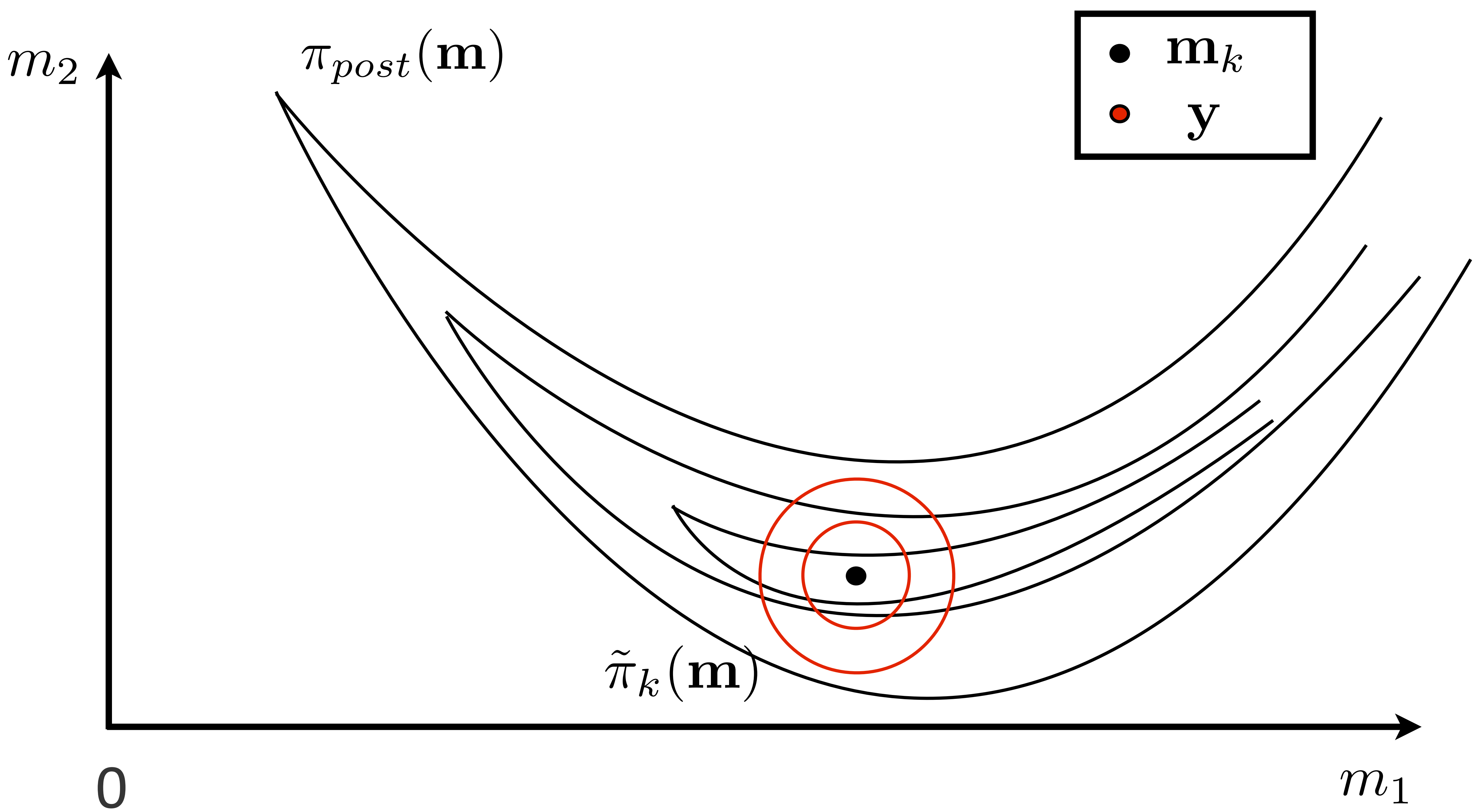


[D. Higdon, *et al*, 2010]

SLIM

McMC method

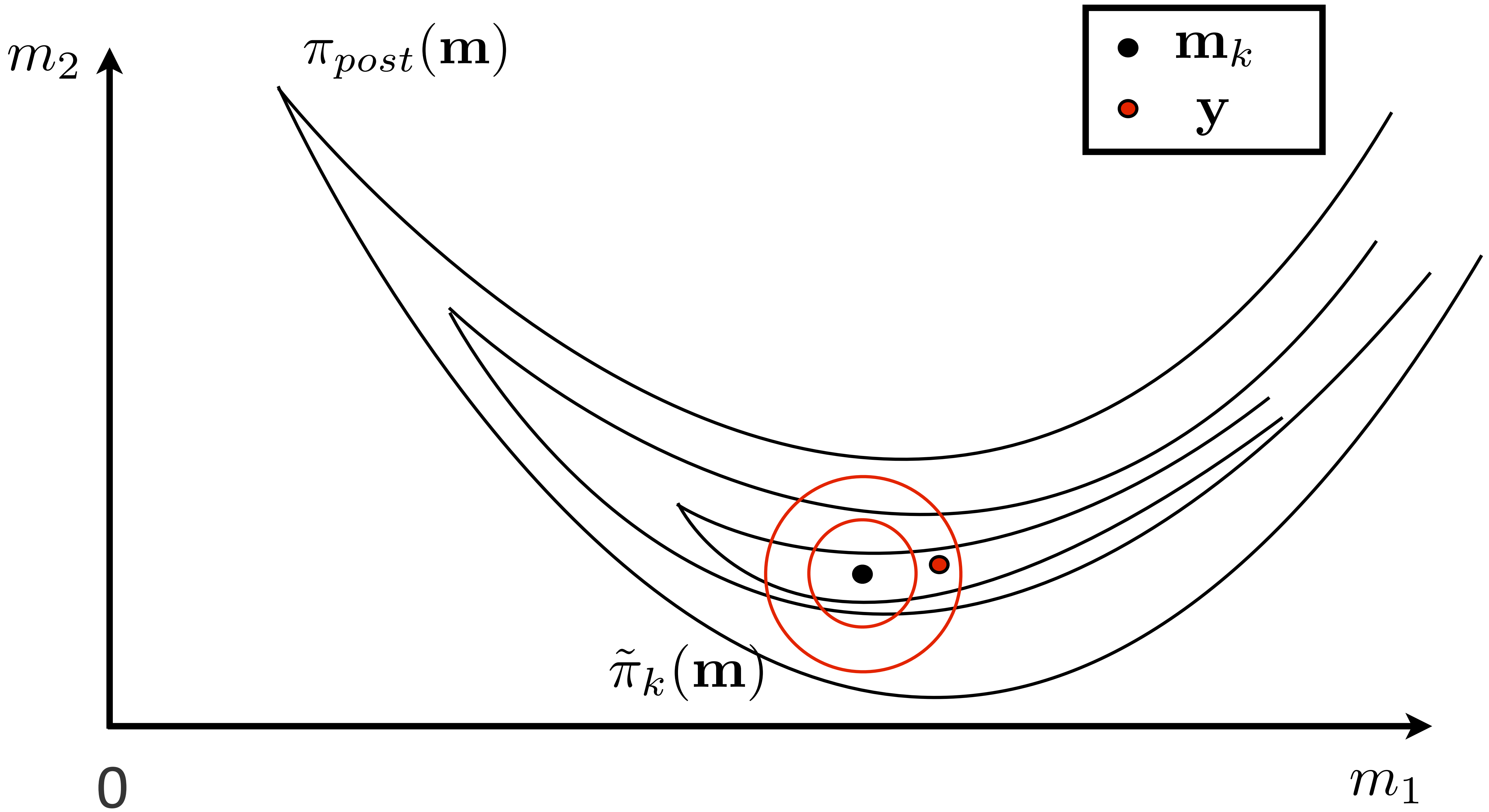
Random Walk: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k, \alpha\mathbf{I})$



Wednesday, 28 October, 15

McMC method

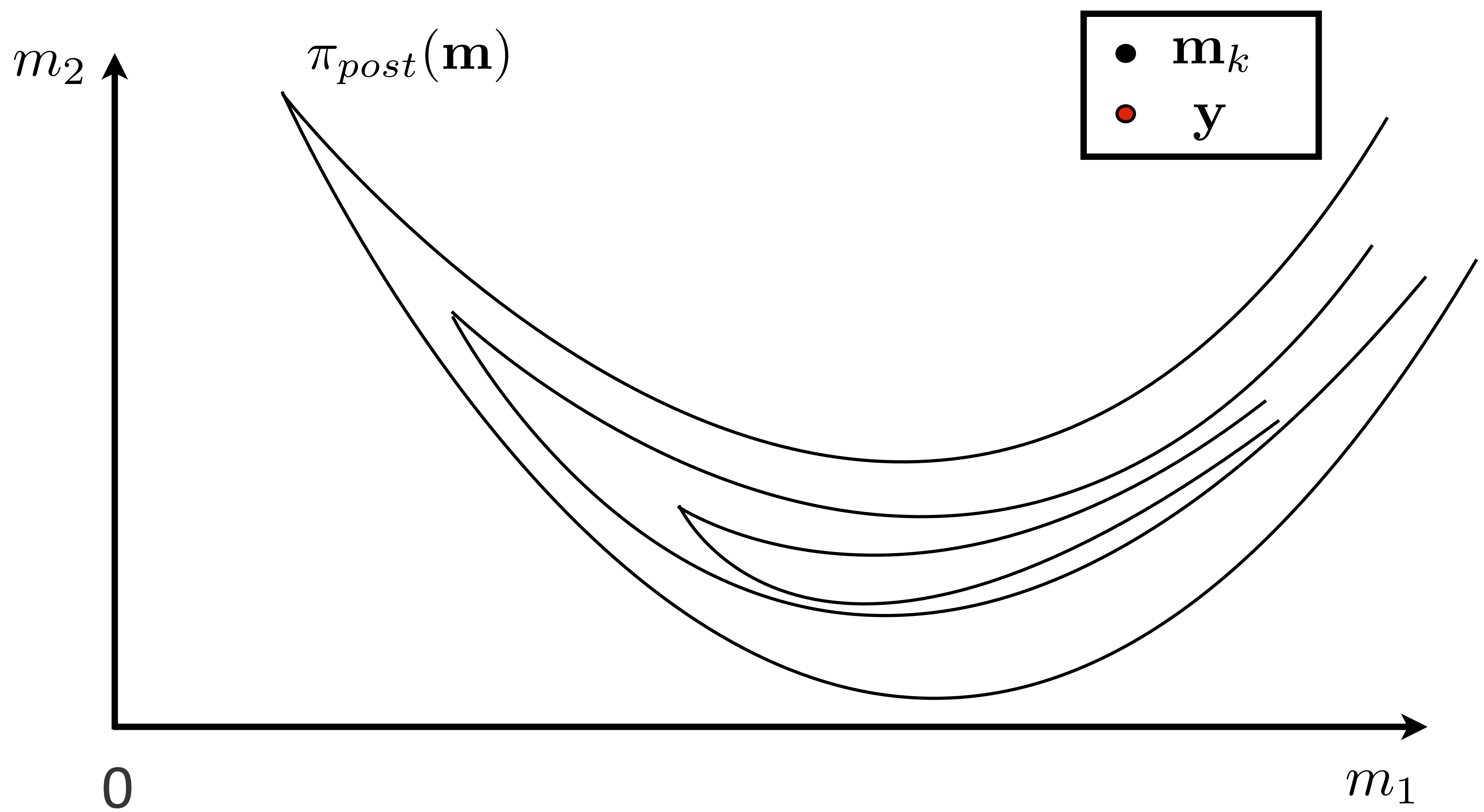
Random Walk: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k, \alpha\mathbf{I})$



SLIM

McMC method

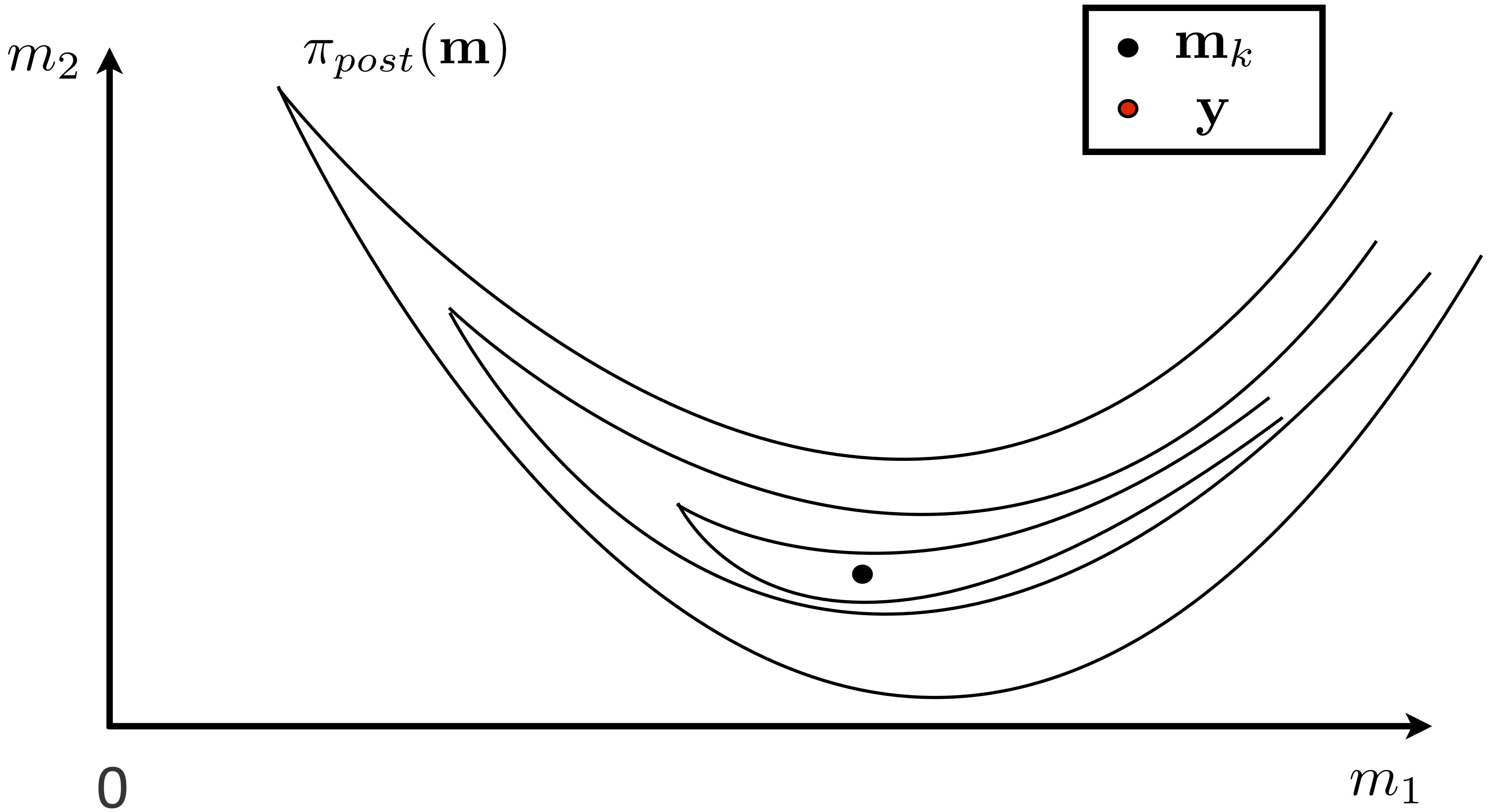
Langevin McMC method: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{L}\mathbf{g}_k, \mathbf{L})$



SLIM

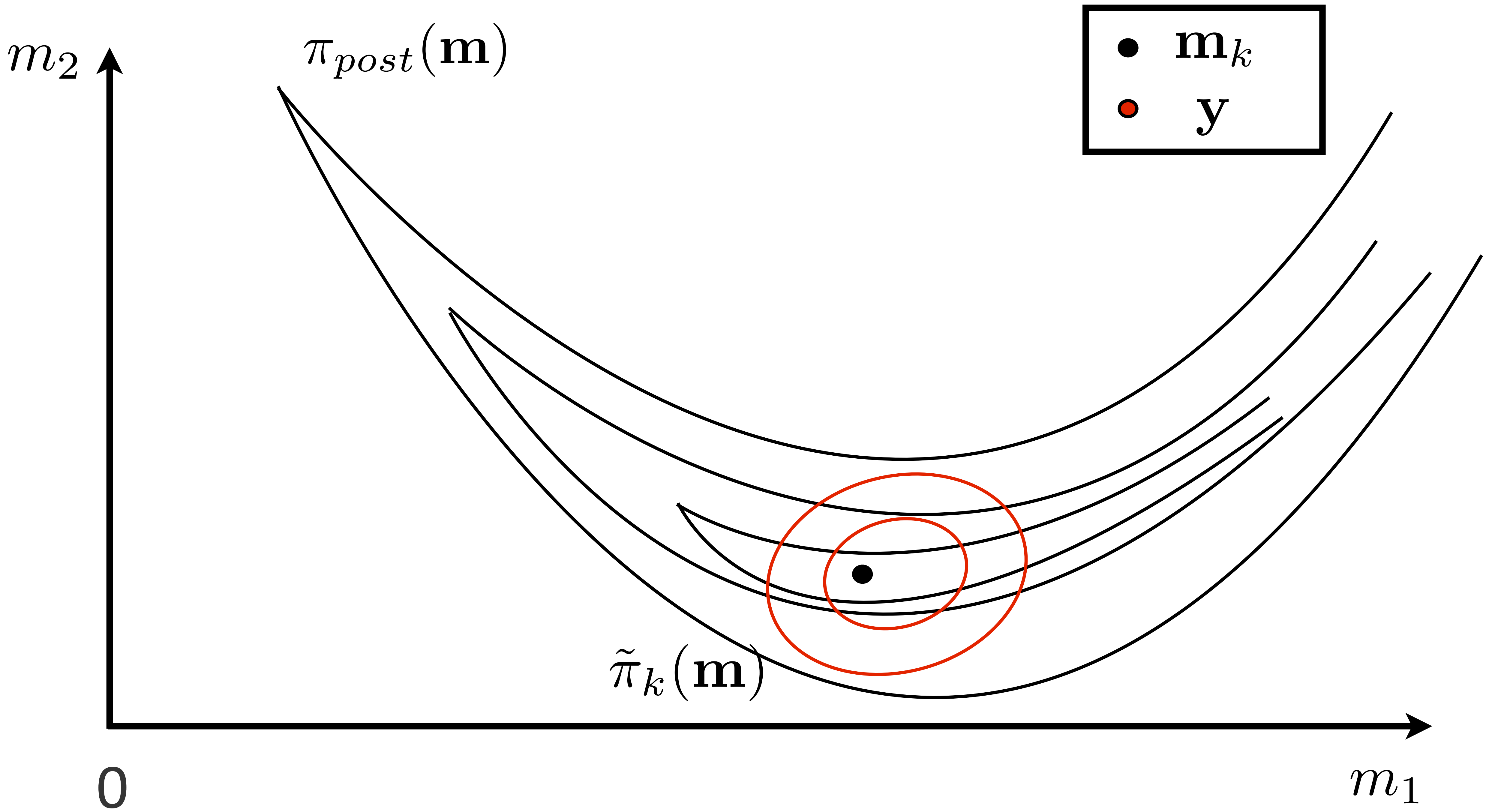
McMC method

Langevin McMC method: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{L}\mathbf{g}_k, \mathbf{L})$



McMC method

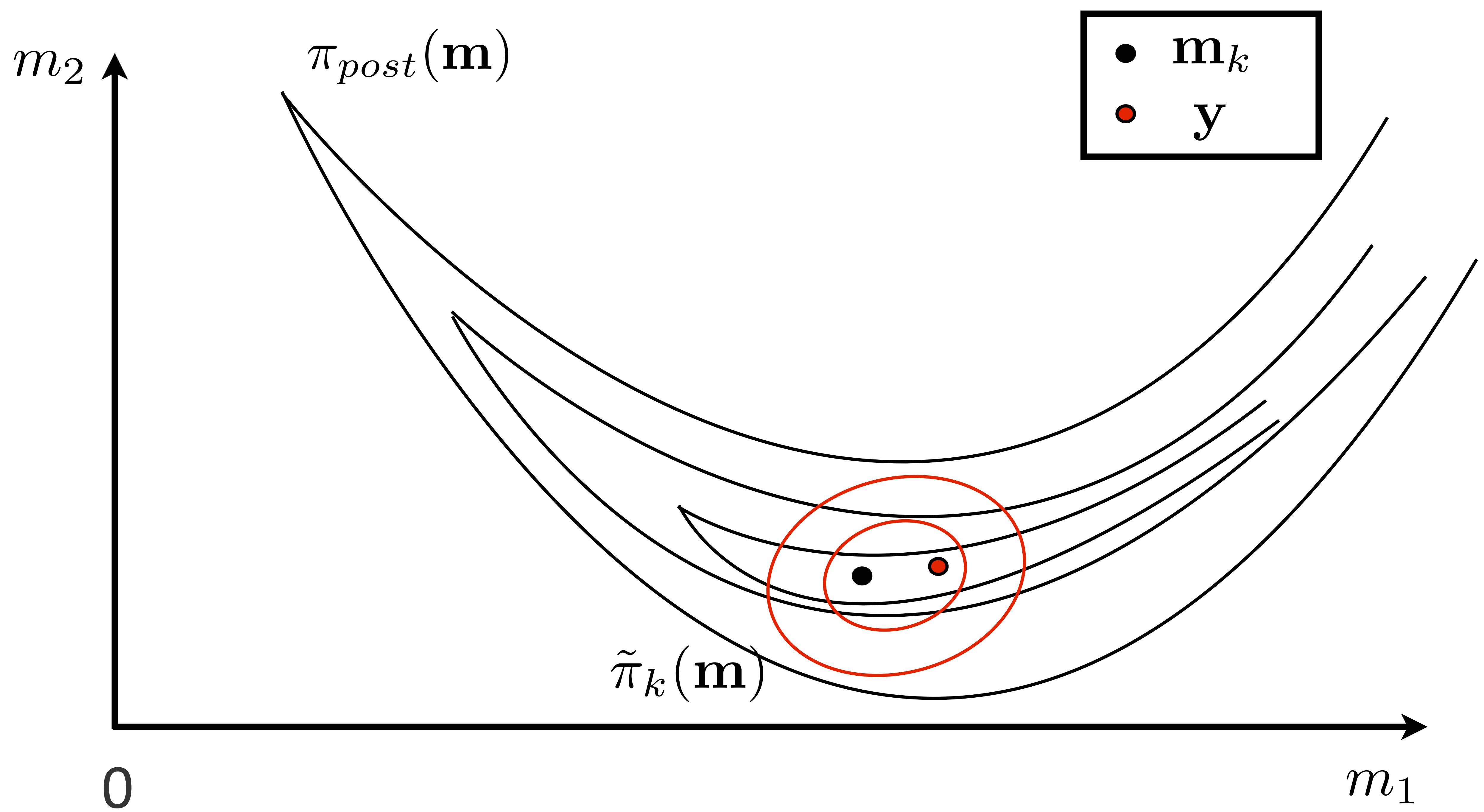
Langevin McMC method: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{L}\mathbf{g}_k, \mathbf{L})$



SLIM

McMC method

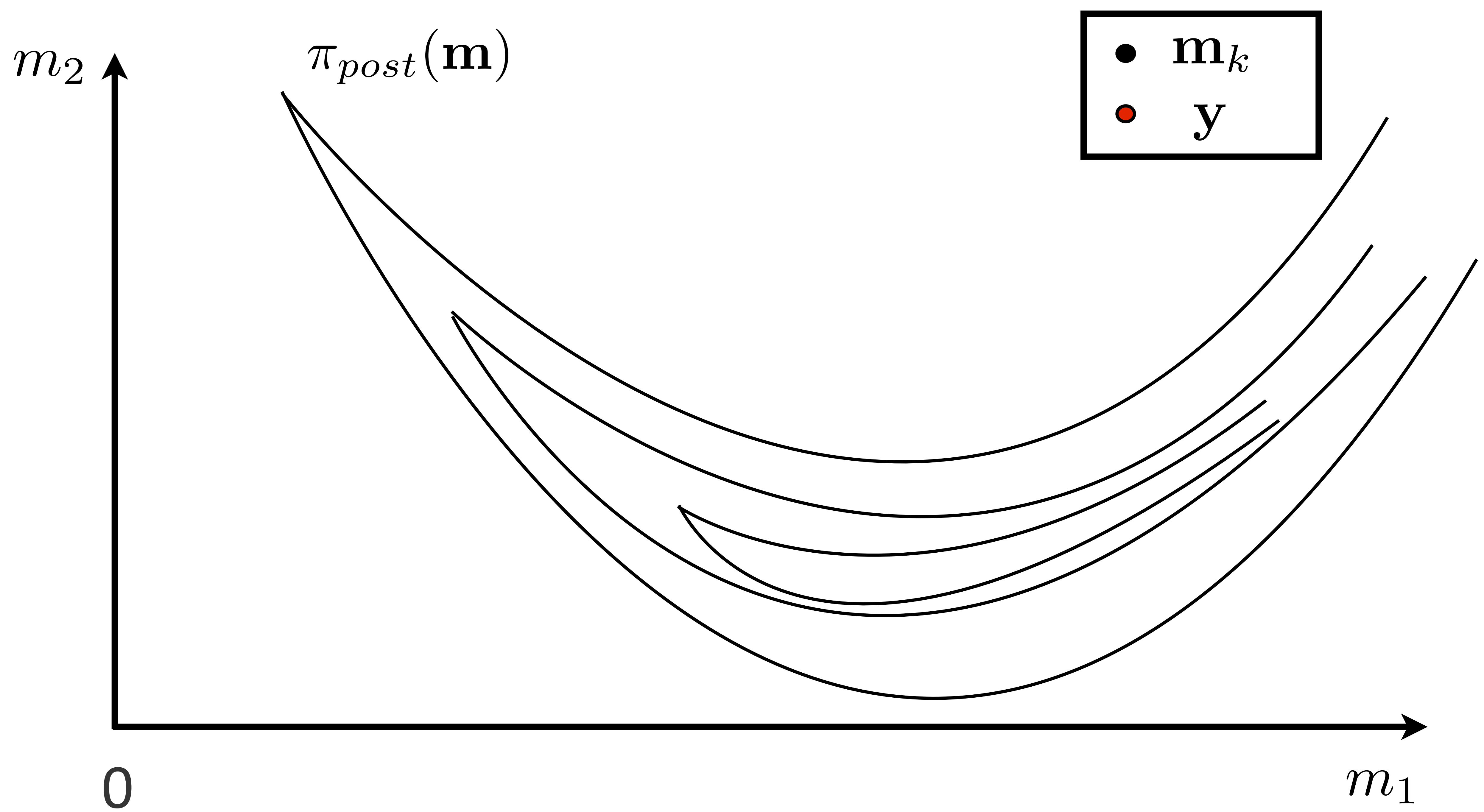
Langevin McMC method: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{L}\mathbf{g}_k, \mathbf{L})$



SLIM

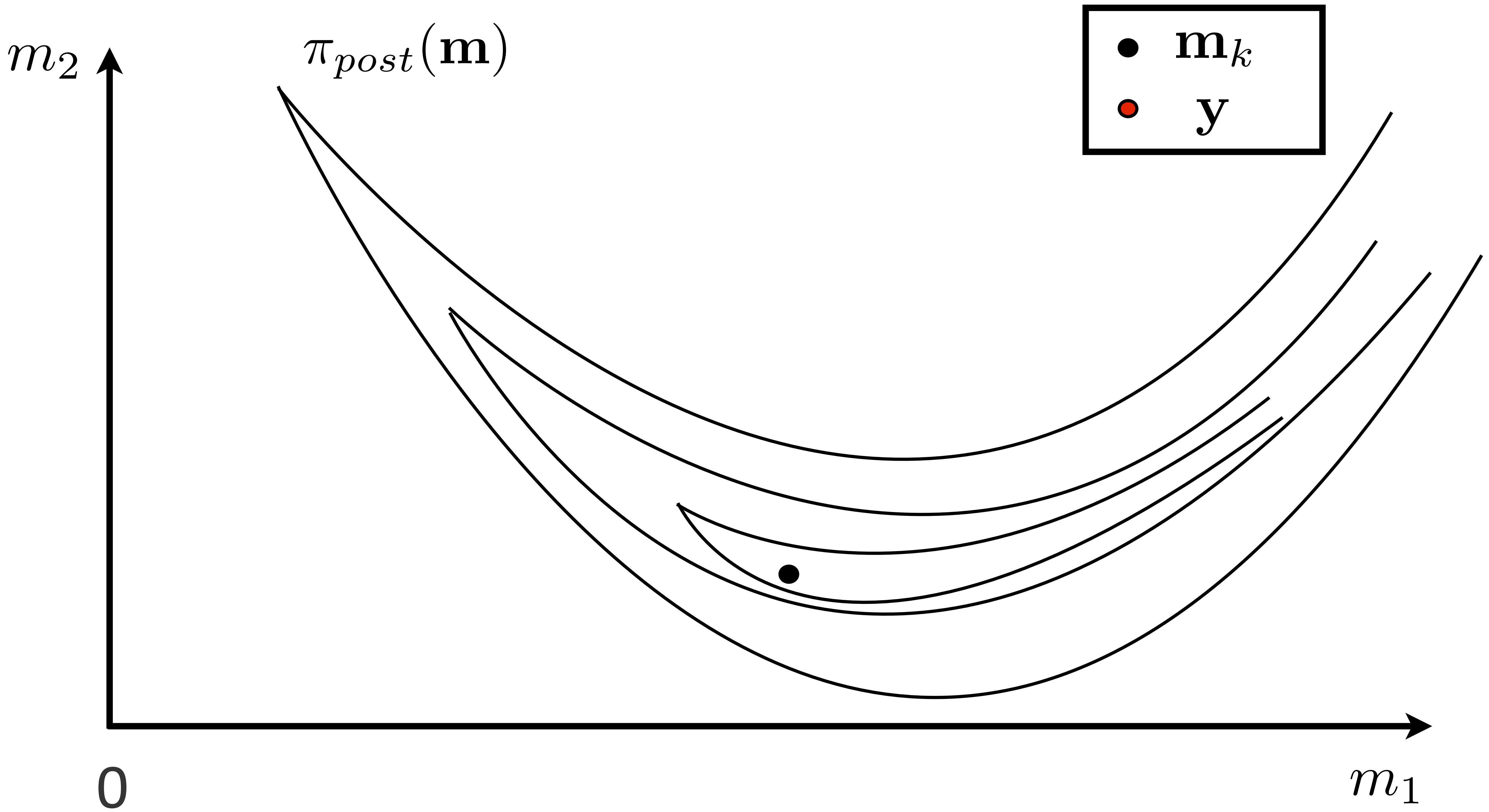
McMC method

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



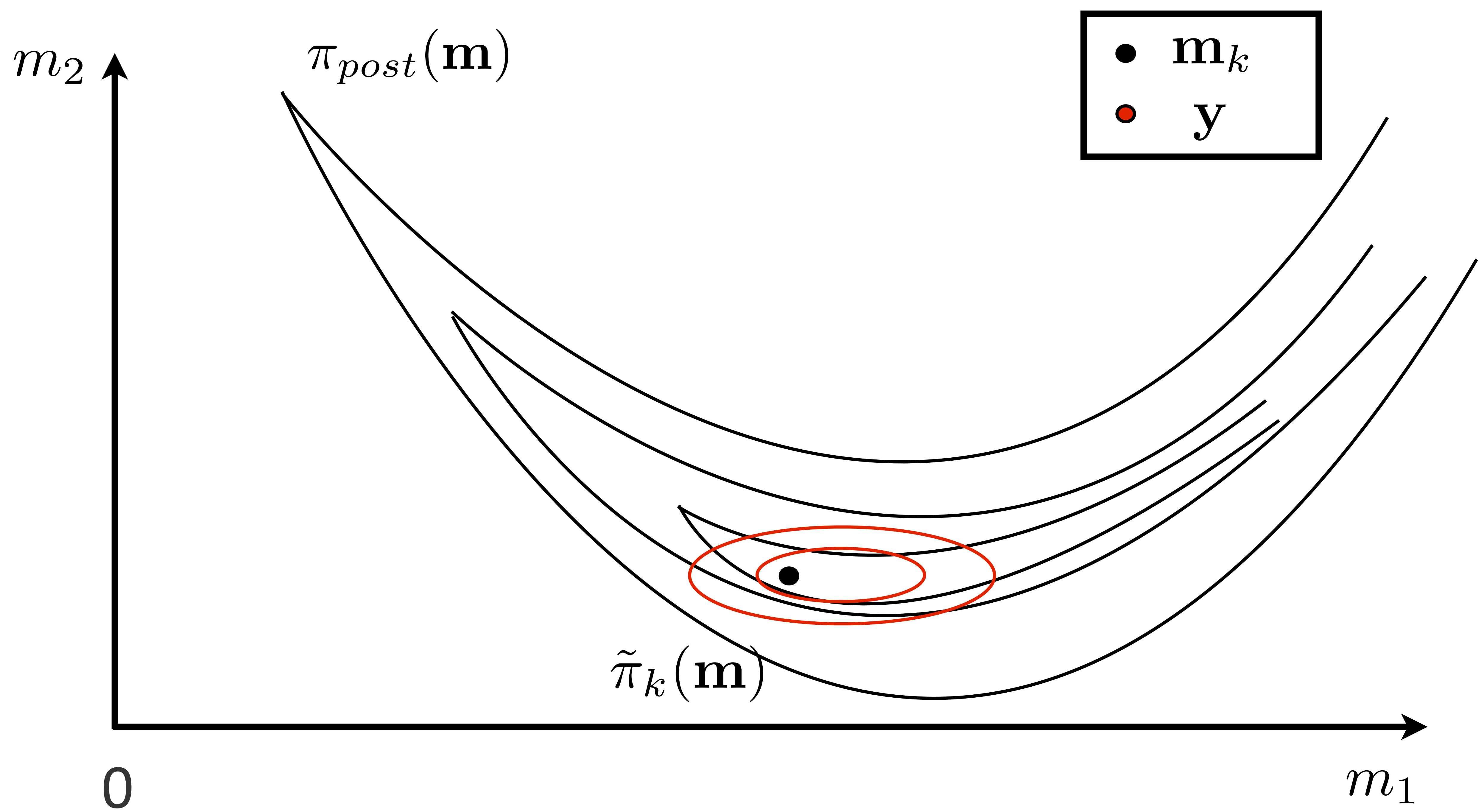
McMC method

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

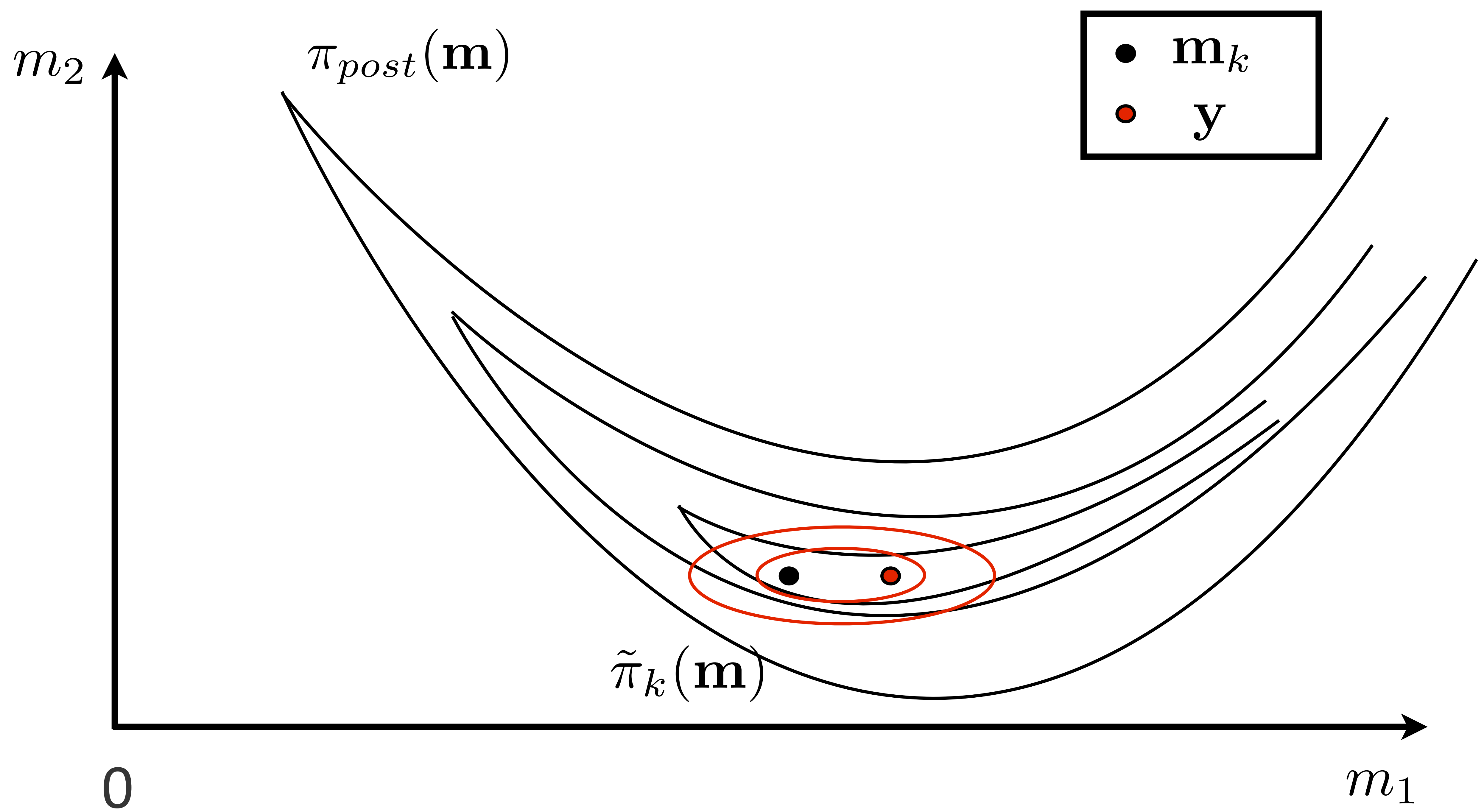
Newton Type McMC: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{H}_k^{-1}\mathbf{g}_k, \mathbf{H}_k^{-1})$



[James Martin *et al*, 2012]

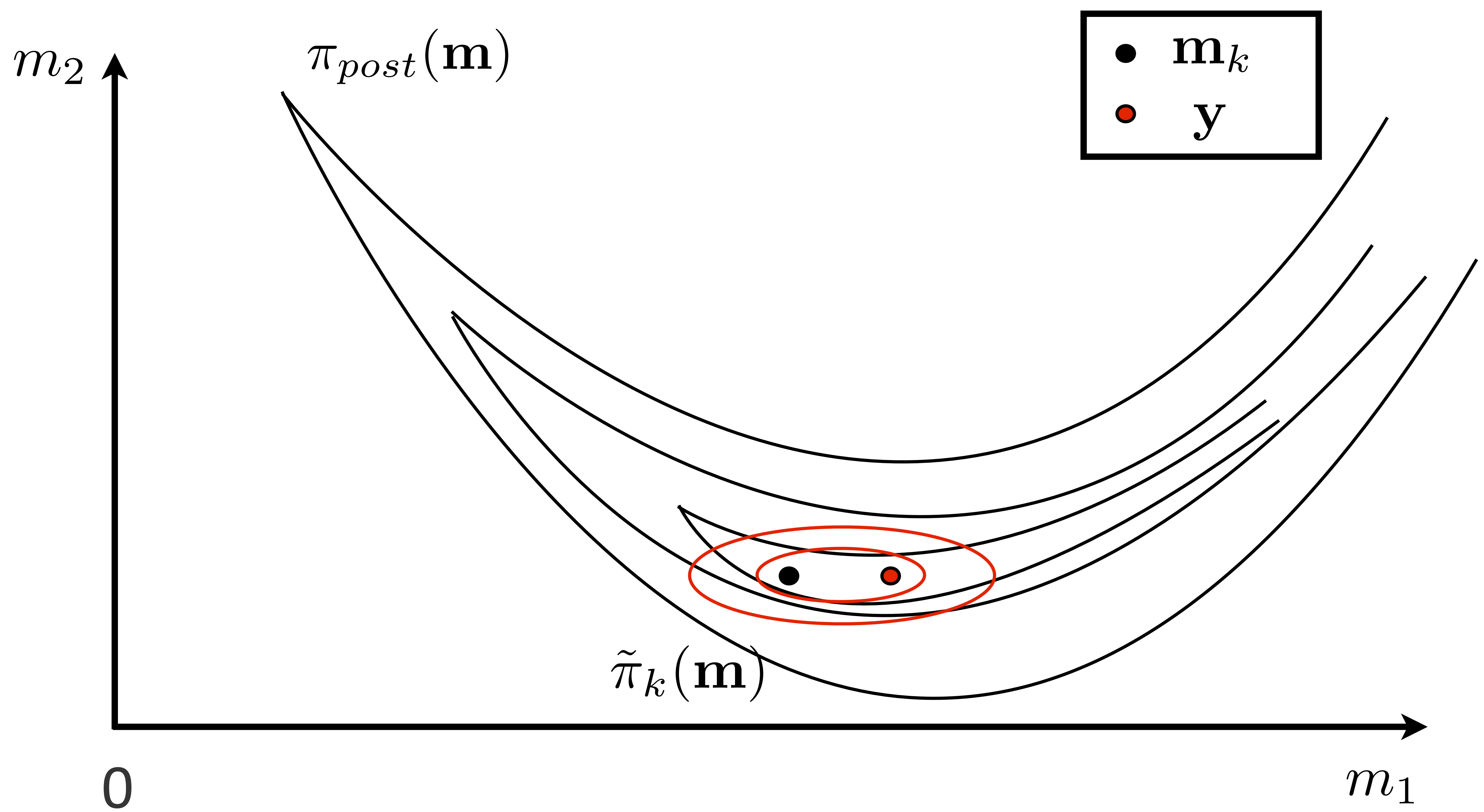
McMC method

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



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 $\rho_{\mathrm{post}}(\mathbf{m})$

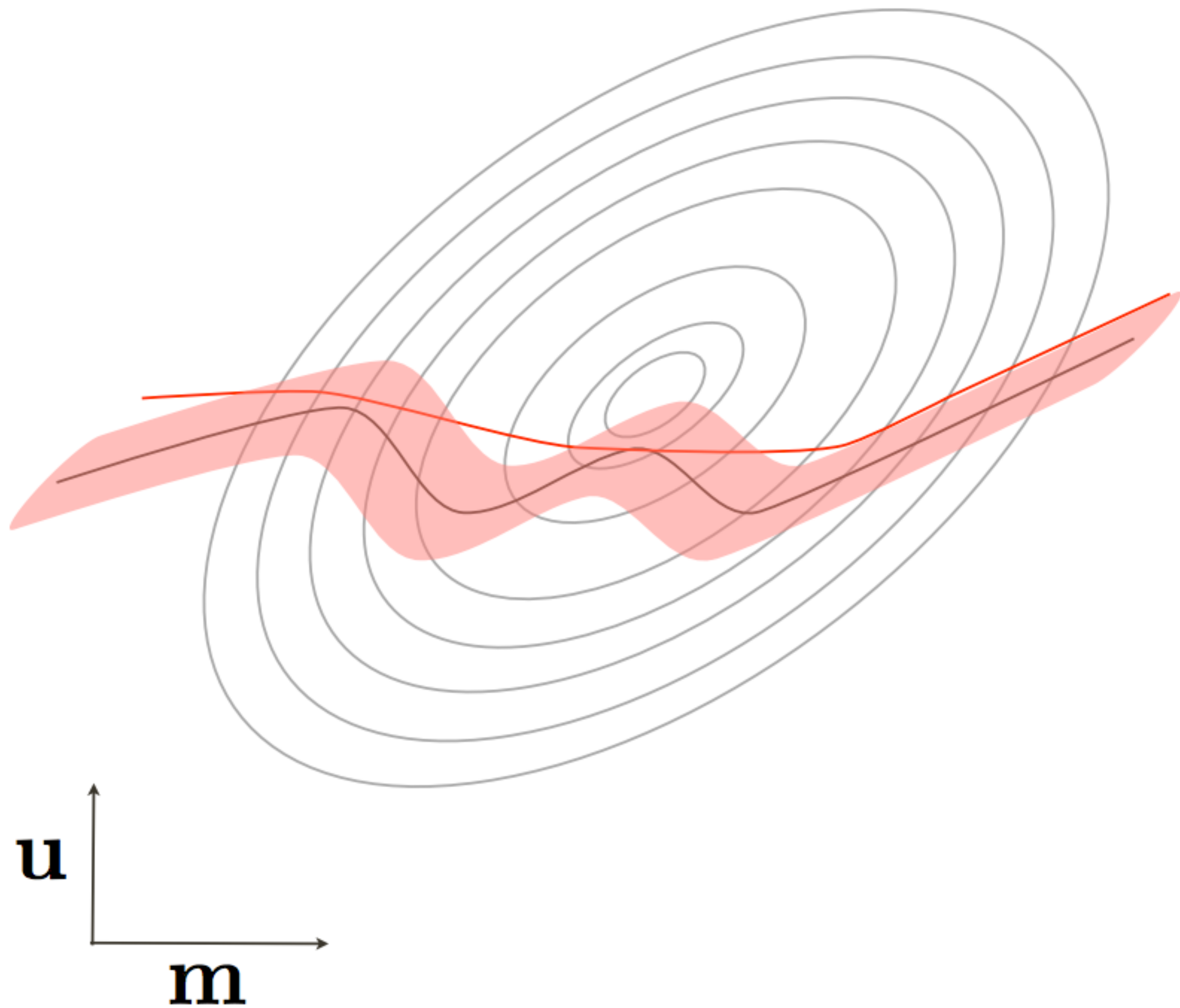
Solution:

- Integrate the posterior distribution
- McMC method to sample the posterior distribution
  - Advantage: the true uncertainty can be quantified
  - Disadvantage: huge computational cost
- Use an approximate distribution to quantify the uncertainty

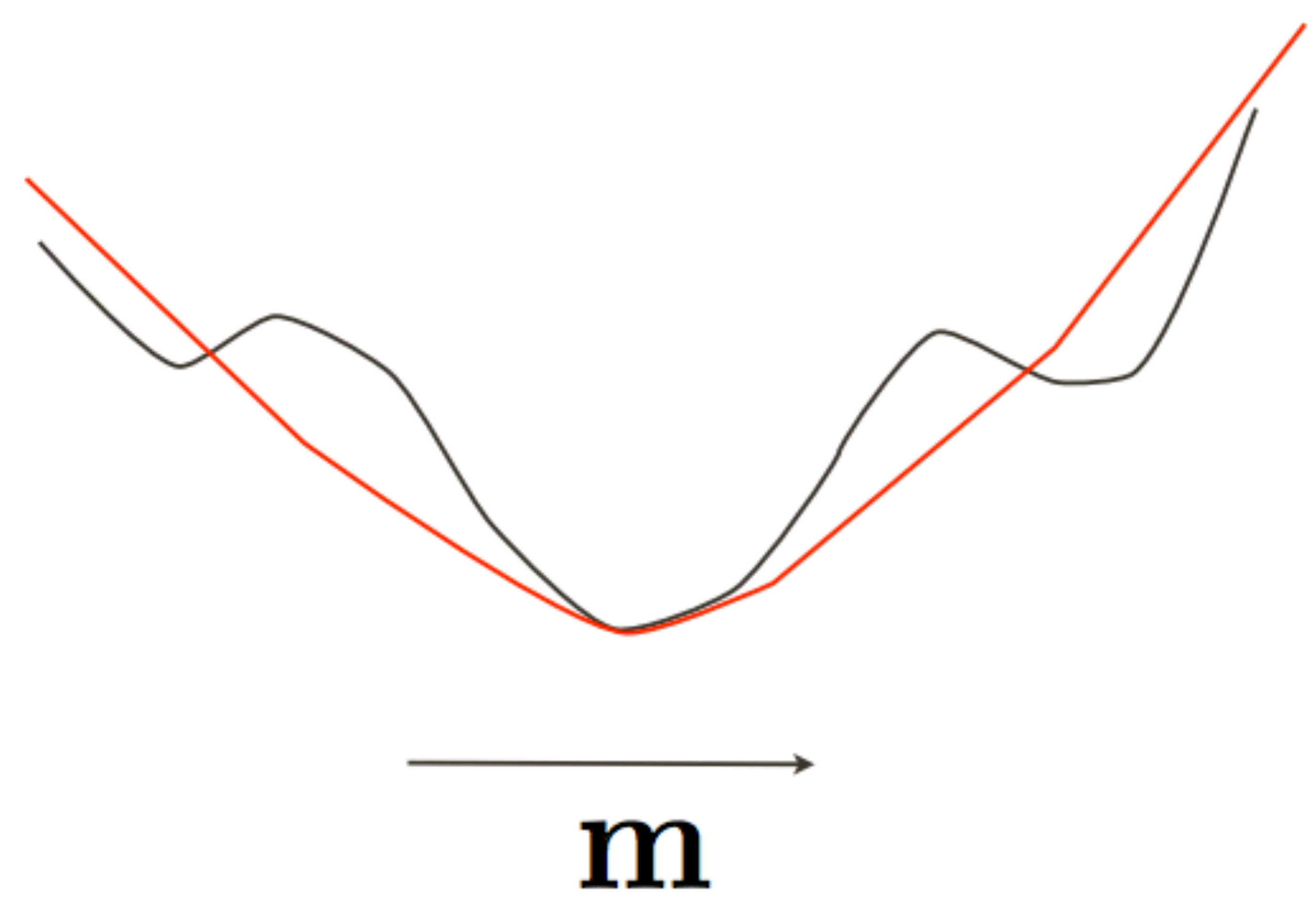
WRI vs FWI

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

Larger # of degrees of freedom

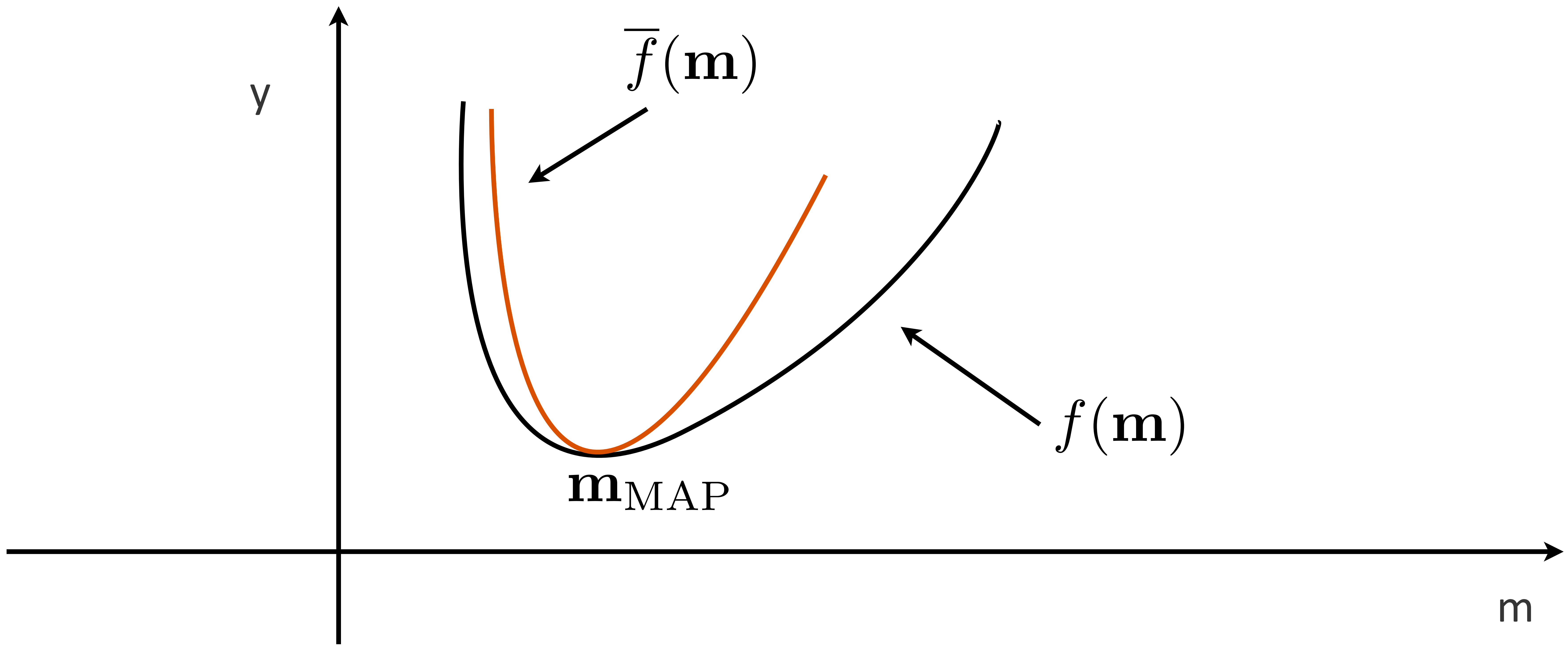


"more convex"



Quadratic approximation

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



Hessian of WRI

Misfit function of WRI:

$$f(\mathbf{m}, \mathbf{u}(\mathbf{m})) = \|\mathbf{Pu} - \mathbf{d}\|^{\mathbf{2}} + \lambda^{\mathbf{2}}\|\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{q}\|^{\mathbf{2}}$$

Hessian of WRI:

$$\mathbf{H} = \frac{\partial^{\mathbf{2}}\mathbf{f}}{\partial\mathbf{m}^{\mathbf{2}}} - \frac{\partial^{\mathbf{2}}\mathbf{f}}{\partial\mathbf{m}\partial\mathbf{u}}\frac{\partial^{\mathbf{2}}\mathbf{f}}{\partial\mathbf{u}^{\mathbf{2}}}^{-\mathbf{1}}\frac{\partial^{\mathbf{2}}\mathbf{f}}{\partial\mathbf{u}\partial\mathbf{m}}$$

Hessian of WRI

$$\mathbf{H} = \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m}^2} - \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m} \partial \mathbf{u}} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u}^2}^{-1} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u} \partial \mathbf{m}}$$

Dense and not positive-definite

where,

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{u}^2} = \mathbf{P^T P} + \lambda^2 \mathbf{A^T A}$$

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{m}^2} = \lambda^2 \left(\frac{\partial \mathbf{A u}}{\partial \mathbf{m}}\right)^{\mathbf{T}} \left(\frac{\partial \mathbf{A u}}{\partial \mathbf{m}}\right)$$

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{u} \partial \mathbf{m}} = \lambda^2 \left(\frac{\partial \mathbf{A^T (A u - q)}}{\partial \mathbf{m}}\right) + \lambda^2 \mathbf{A^T} \left(\frac{\partial \mathbf{A u}}{\partial \mathbf{m}}\right)$$

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{m} \partial \mathbf{u}} = \lambda^2 \left(\frac{\partial \mathbf{A^T (A u - q)}}{\partial \mathbf{m}}\right)^{\mathbf{T}} + \lambda^2 \left(\frac{\partial \mathbf{A u}}{\partial \mathbf{m}}\right)^{\mathbf{T}} \mathbf{A}$$

Hessian of WRI

$$\mathbf{H} = \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m}^2} - \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m} \partial \mathbf{u}} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u}^2}^{-1} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u} \partial \mathbf{m}}$$

Dense and not positive-definite

where,

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{u}^2} = \mathbf{P}^{\mathbf{T}}\mathbf{P} + \lambda^{\mathbf{2}}\mathbf{A}^{\mathbf{T}}\mathbf{A}$$

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{m}^2} = \lambda^{\mathbf{2}}\left(\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{m}}\right)^{\mathbf{T}}\left(\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{m}}\right)$$

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{u} \partial \mathbf{m}} = \lambda^{\mathbf{2}}\left(\frac{\partial \mathbf{A}^{\mathbf{T}}(\mathbf{A}\mathbf{u} - \mathbf{q})}{\partial \mathbf{m}}\right) + \lambda^{\mathbf{2}}\mathbf{A}^{\mathbf{T}}\left(\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{m}}\right)$$

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{m} \partial \mathbf{u}} = \lambda^{\mathbf{2}}\left(\frac{\partial \mathbf{A}^{\mathbf{T}}(\mathbf{A}\mathbf{u} - \mathbf{q})}{\partial \mathbf{m}}\right)^{\mathbf{T}} + \lambda^{\mathbf{2}}\left(\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{m}}\right)^{\mathbf{T}}\mathbf{A}$$

$$\boxed{\begin{array}{l} \lambda \to \infty \text{ or } \mathbf{m} \to \mathbf{m}_t \\ \mathbf{A}\mathbf{u} - \mathbf{q} \to 0 \end{array}}$$

Hessian of WRI

$$\mathbf{H} = \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m}^2} - \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m} \partial \mathbf{u}} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u}^2}^{-1} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u} \partial \mathbf{m}}$$

Dense and not positive-definite

where,

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{u}^2} = \mathbf{P}^{\mathbf{T}}\mathbf{P} + \lambda^2 \mathbf{A}^{\mathbf{T}}\mathbf{A}$$

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{m}^2} = \lambda^2 \left(\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{m}}\right)^{\mathbf{T}} \left(\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{m}}\right)$$

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{u} \partial \mathbf{m}} = \xcancel{\lambda^2 \left(\frac{\partial \mathbf{A}^{\mathbf{T}}(\mathbf{A}\mathbf{u} - \mathbf{q})}{\partial \mathbf{m}}\right)} + \lambda^2 \mathbf{A}^{\mathbf{T}} \left(\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{m}}\right)$$

$$\frac{\partial^2 \mathbf{f}}{\partial \mathbf{m} \partial \mathbf{u}} = \xcancel{\lambda^2 \left(\frac{\partial \mathbf{A}^{\mathbf{T}}(\mathbf{A}\mathbf{u} - \mathbf{q})}{\partial \mathbf{m}}\right)^{\mathbf{T}}} + \lambda^2 \left(\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{m}}\right)^{\mathbf{T}} \mathbf{A}$$

$$\boxed{\begin{array}{l} \lambda \to \infty \text{ or } \mathbf{m} \to \mathbf{m}_t \\ \mathbf{A}\mathbf{u} - \mathbf{q} \to 0 \end{array}}$$

Positive-definite approximation of Hessian

Approximated Hessian:

$$
\begin{aligned}
\mathbf{H} &= \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m}^2} - \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m}\partial \mathbf{u}} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u}^2}^{-1} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u}\partial \mathbf{m}} \\
&\approx \lambda^2 \mathbf{G}^T\mathbf{G} - \lambda^4 \mathbf{G}^T \mathbf{A}(\lambda^2 \mathbf{A}^T \mathbf{A} + \mathbf{P}^T\mathbf{P})^{-1}\mathbf{A}^T\mathbf{G} \\
&= \lambda^2 \mathbf{G}^T\left(\frac{1}{\lambda^2}\mathbf{A}^{-T}\mathbf{P}^T \underbrace{(\mathbf{I} + \frac{1}{\lambda^2}\mathbf{P}\mathbf{A}^{-1}\mathbf{A}^{-T}\mathbf{P}^T)^{-1}}_{n_r^2} \underbrace{\mathbf{P}\mathbf{A}^{-1}}_{n_r n_g}\right)\underbrace{\mathbf{G}}_{\text{diagonal}}
\end{aligned}
$$

Positive-definite approximation of Hessian

Approximated Hessian:

$$
\begin{aligned}
\mathbf{H} &= \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m}^2} - \frac{\partial^2 \mathbf{f}}{\partial \mathbf{m} \partial \mathbf{u}} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u}^2}^{-1} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{u} \partial \mathbf{m}} \\
&\approx \lambda^2 \mathbf{G}^T \mathbf{G} - \lambda^4 \mathbf{G}^T \mathbf{A} (\lambda^2 \mathbf{A}^T \mathbf{A} + \mathbf{P}^T \mathbf{P})^{-1} \mathbf{A}^T \mathbf{G} \\
&= \lambda^2 \mathbf{G}^T \Big(\frac{1}{\lambda^2} \mathbf{A}^{-T} \mathbf{P}^T \underbrace{\Big(\mathbf{I} + \frac{1}{\lambda^2} \mathbf{P} \mathbf{A}^{-1} \mathbf{A}^{-T} \mathbf{P}^T\Big)}_{\mathbf{S}}{}^{-1} \underbrace{\mathbf{P} \mathbf{A}^{-1} \Big) \mathbf{G}}_{\mathbf{W}}
\end{aligned}
$$

Positive-definite approximation of Hessian

$$\mathbf{H} = \sum_{n_f, n_s} \mathbf{W}^{\mathbf{T}} \mathbf{S} \mathbf{W}$$

Positive-definite approximation of Hessian

$$\mathbf{H} = \sum_{n_f, n_s} \mathbf{W}^{\mathbf{T}} \, \mathbf{S} \, \mathbf{W}$$

Computational cost:

$n_f * (n_s + n_r)$

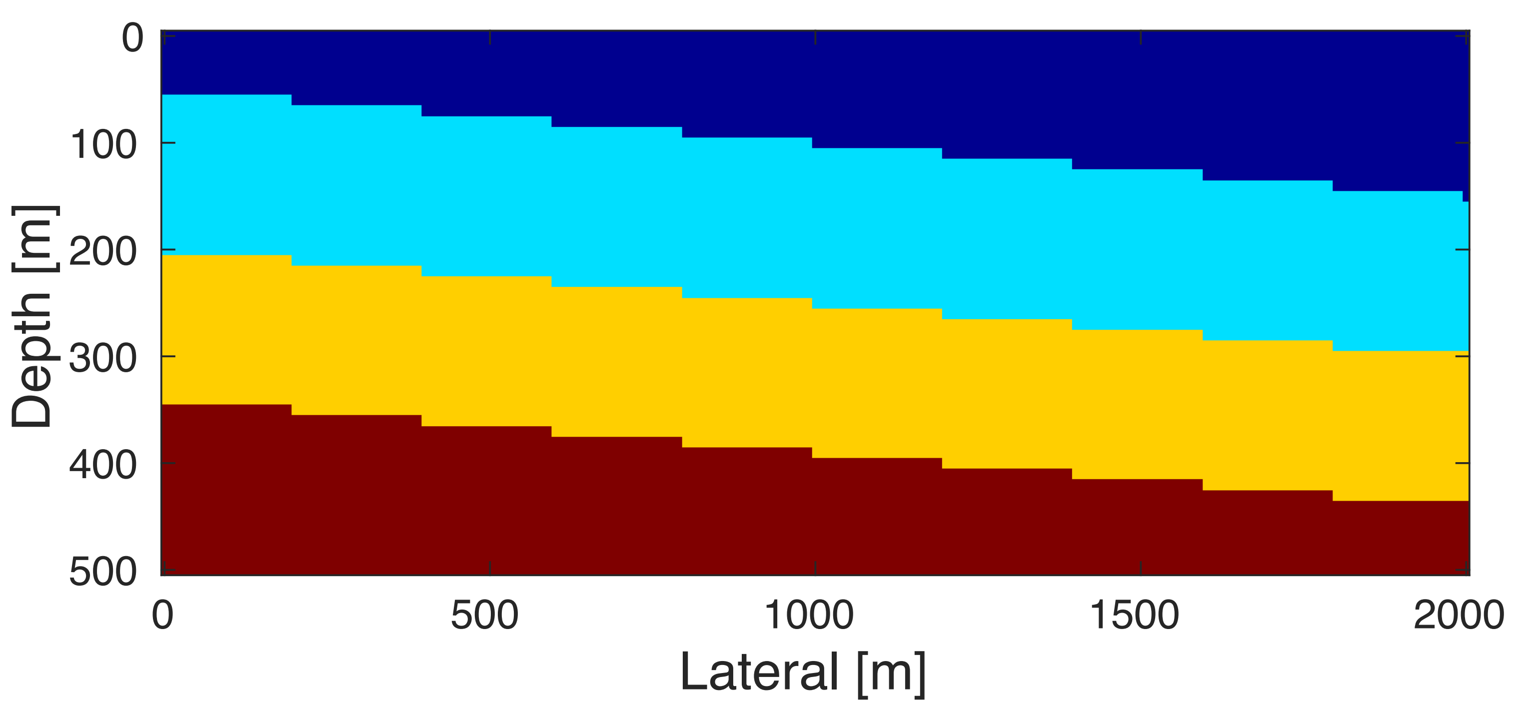
storage cost:

$n_f * n_g * (n_s + n_r)$

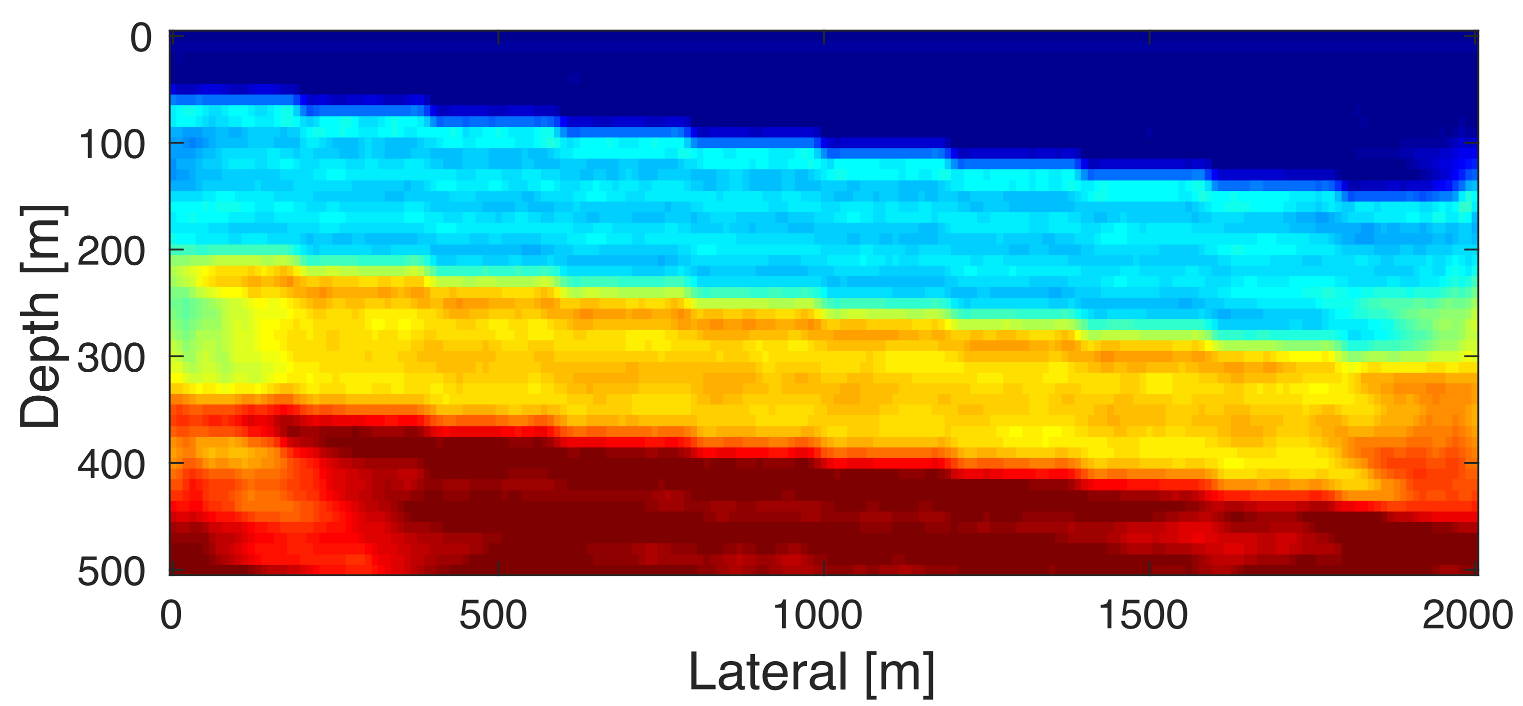
Positive-definite approximation of Hessian

Evaluate the Hessian with three different models:

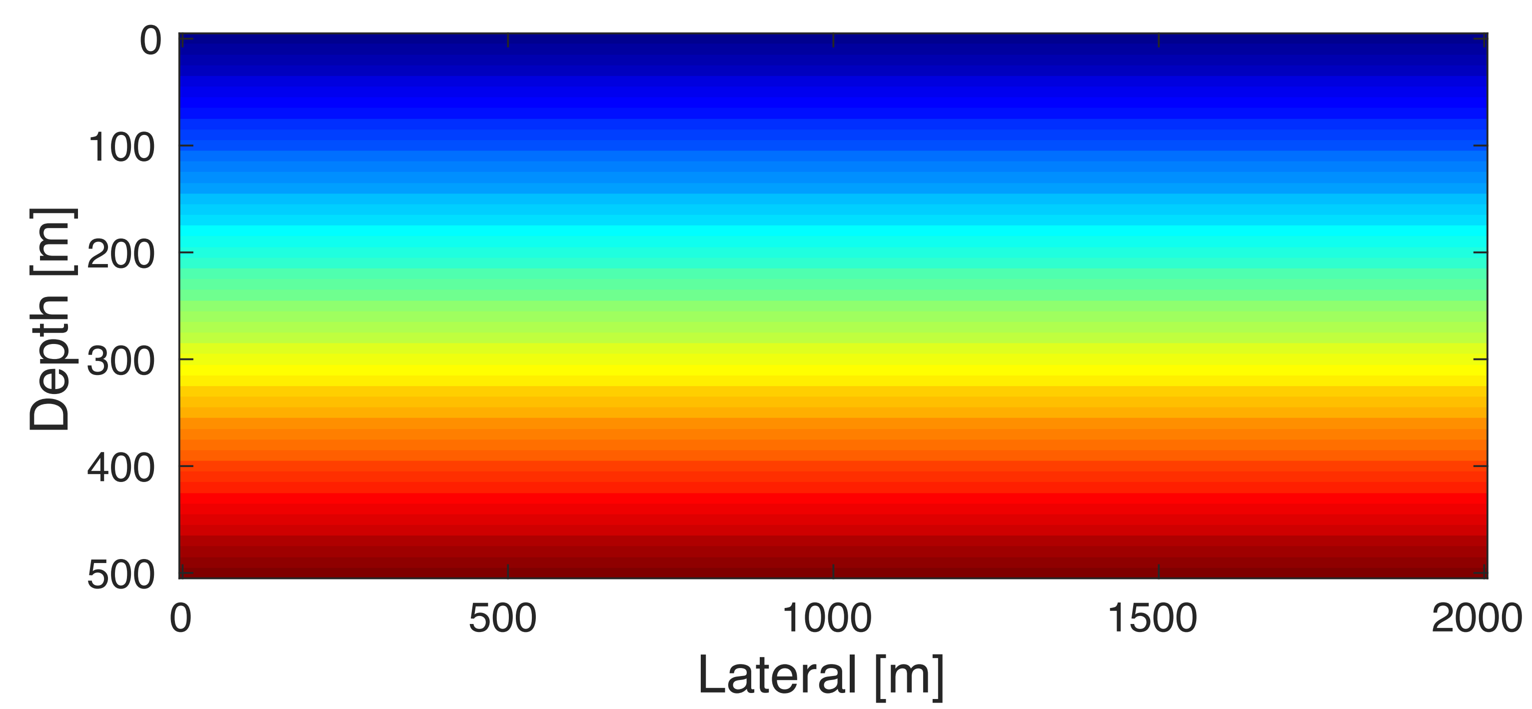
- True model
- Model from inversion
- Initial model



$\mathbf{m}_t$



$\mathbf{m}_n$



$\mathbf{m}_0$

Positive-definite approximation of Hessian

Evaluate the Hessian with three different model:

- True model
- Model from inversion
- Initial model

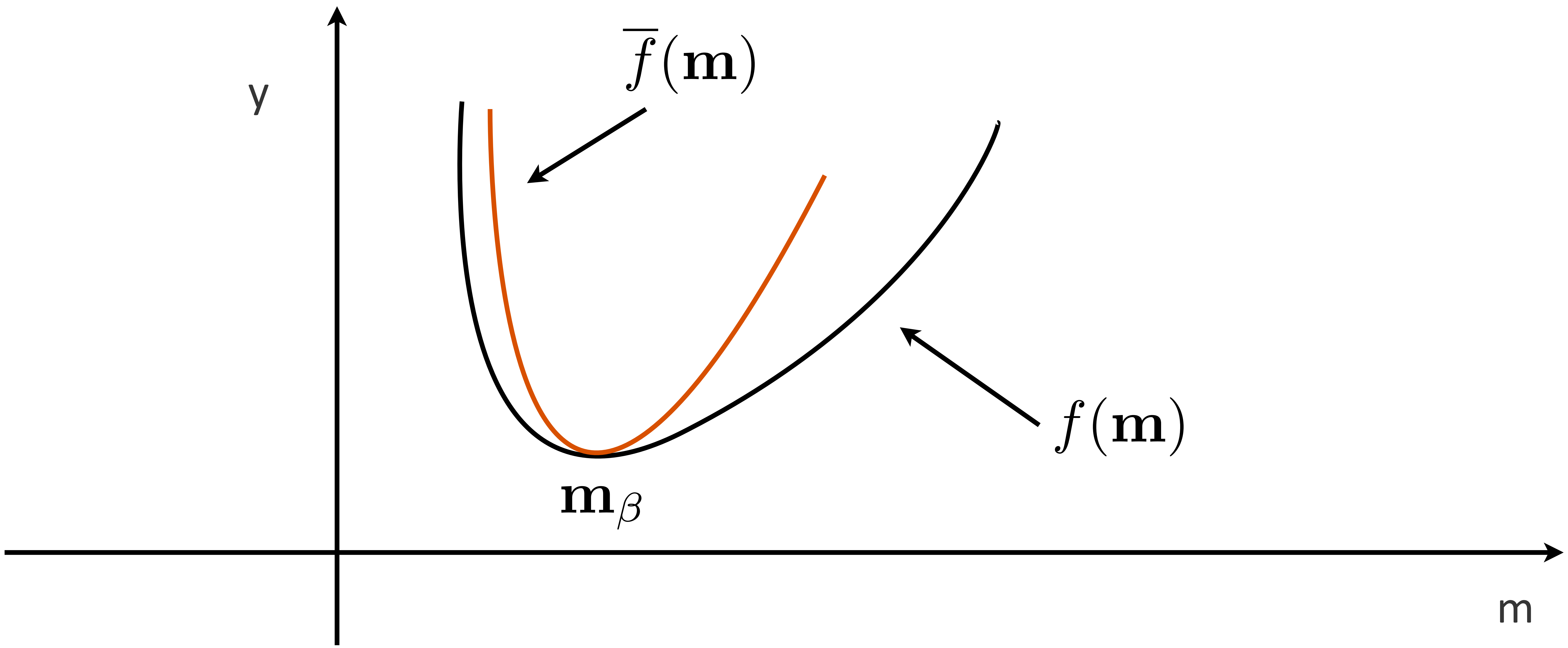
Evaluate the Hessian with three different λ choices:

- $\lambda = 1e0$
- $\lambda = 1e3$
- $\lambda = 1e6$

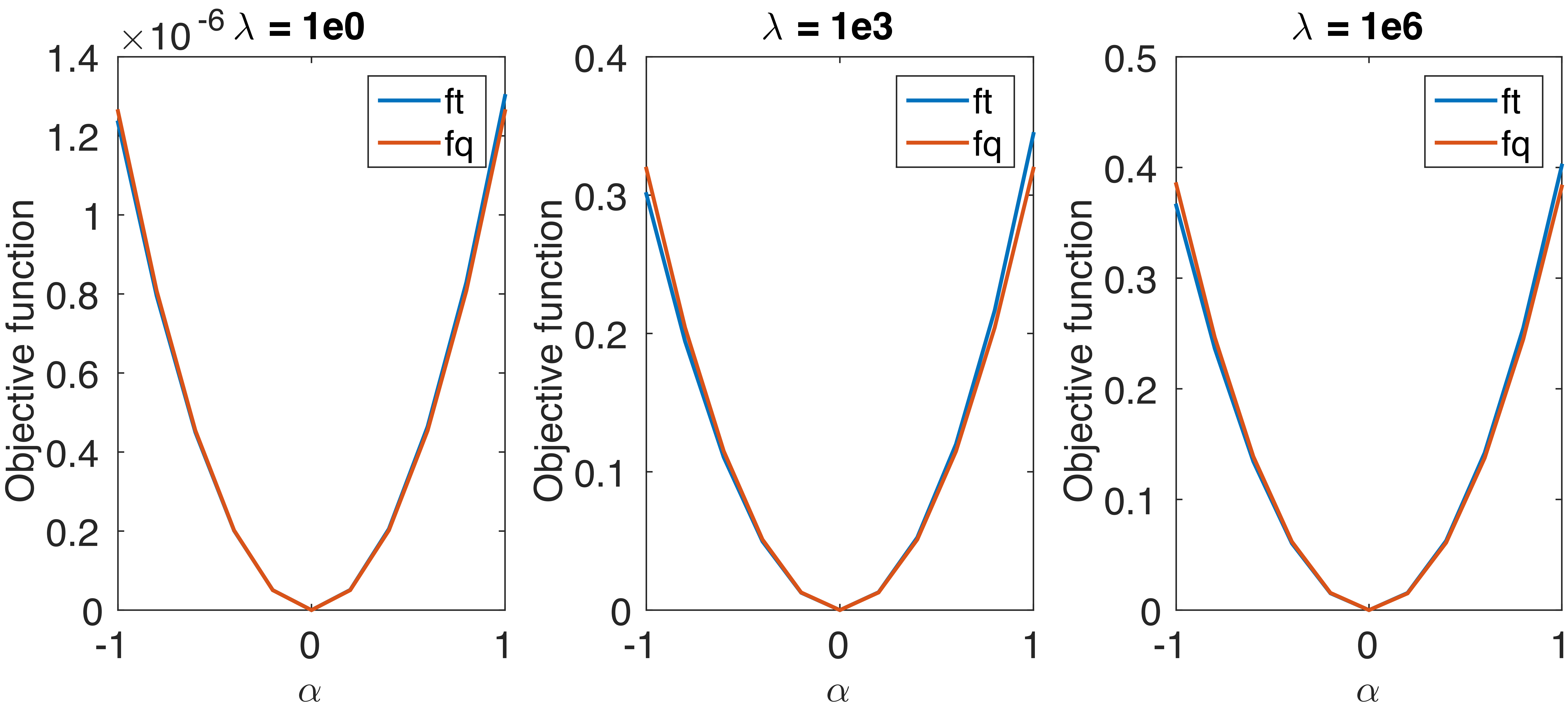
Quadratic approximation

$$f(\mathbf{m}) \approx f(\mathbf{m}_\beta) + \mathbf{g}^T(\mathbf{m} - \mathbf{m}_\beta) + \frac{1}{2}(\mathbf{m} - \mathbf{m}_\beta)^T\mathbf{H}(\mathbf{m} - \mathbf{m}_\beta) := \overline{f}(\mathbf{m})$$

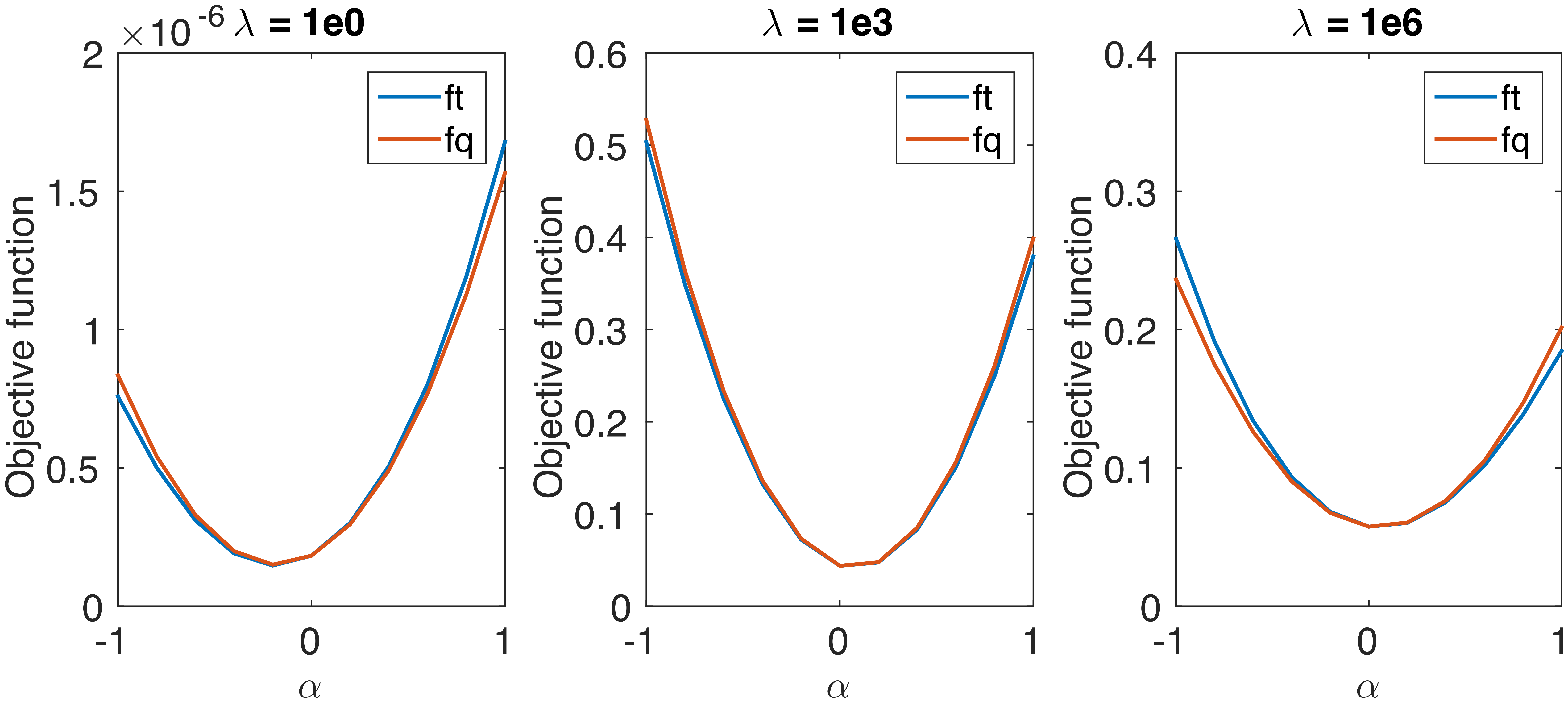
$$\mathbf{m} = \mathbf{m}_\beta + \alpha \mathrm{d}\mathbf{m},\ \beta \in \{t, n, 0\}$$



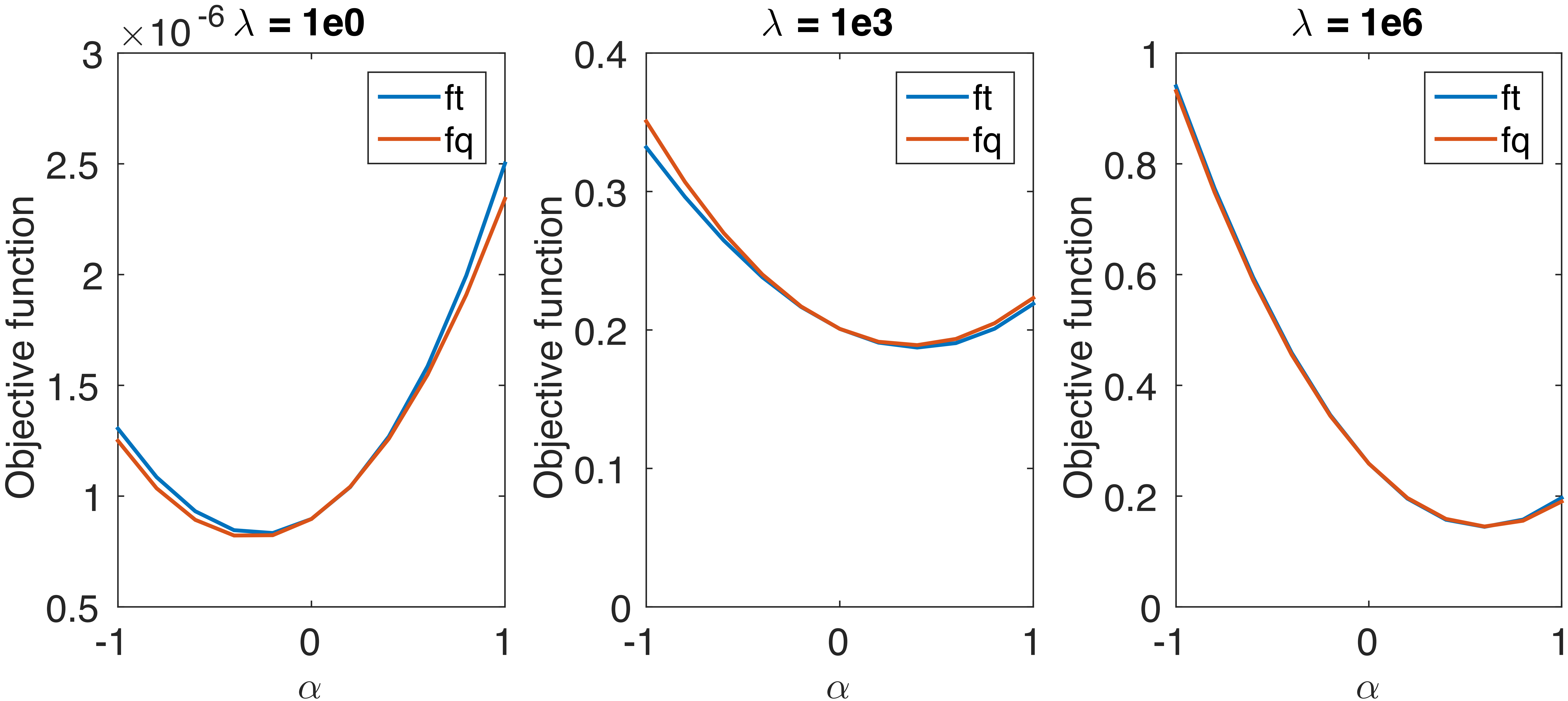
Quadratic approximation with approximated Hessian at m_t



Quadratic approximation with approximated Hessian at m_n



Quadratic approximation with approximated Hessian at m_0



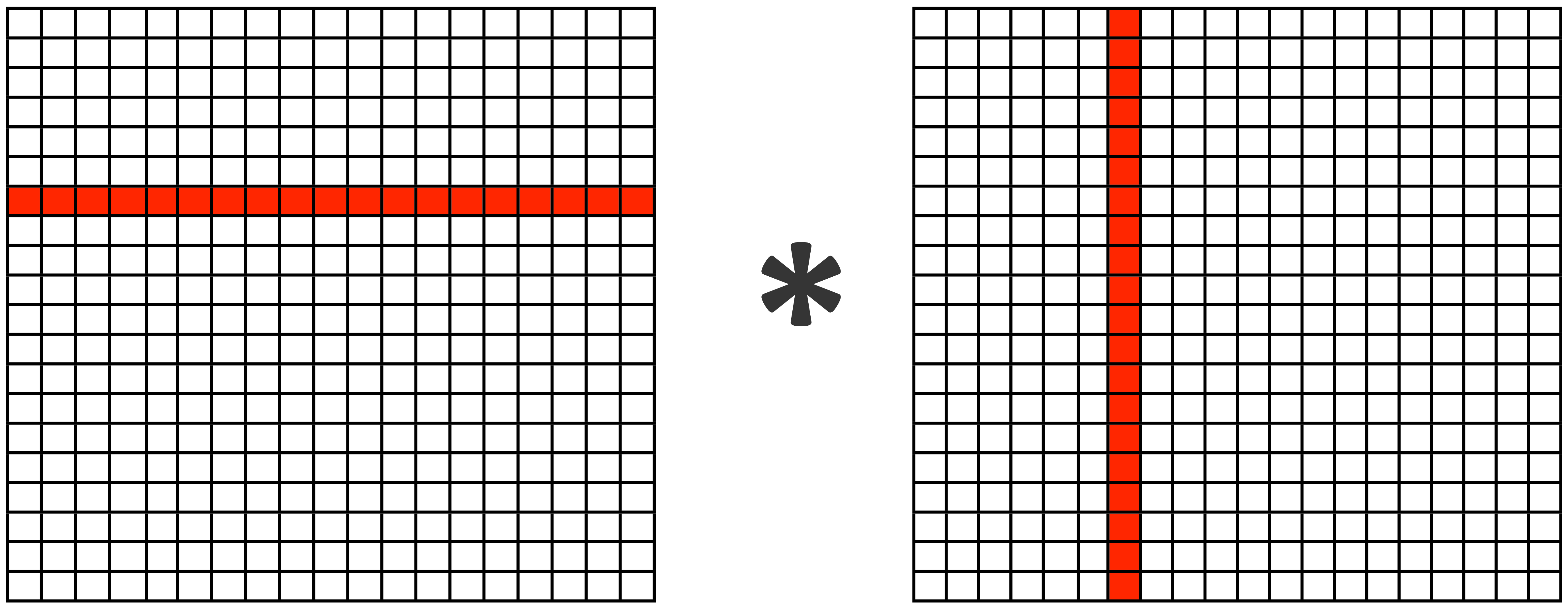
Diagonal part of the Hessian

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

Diagonal part of the Hessian:

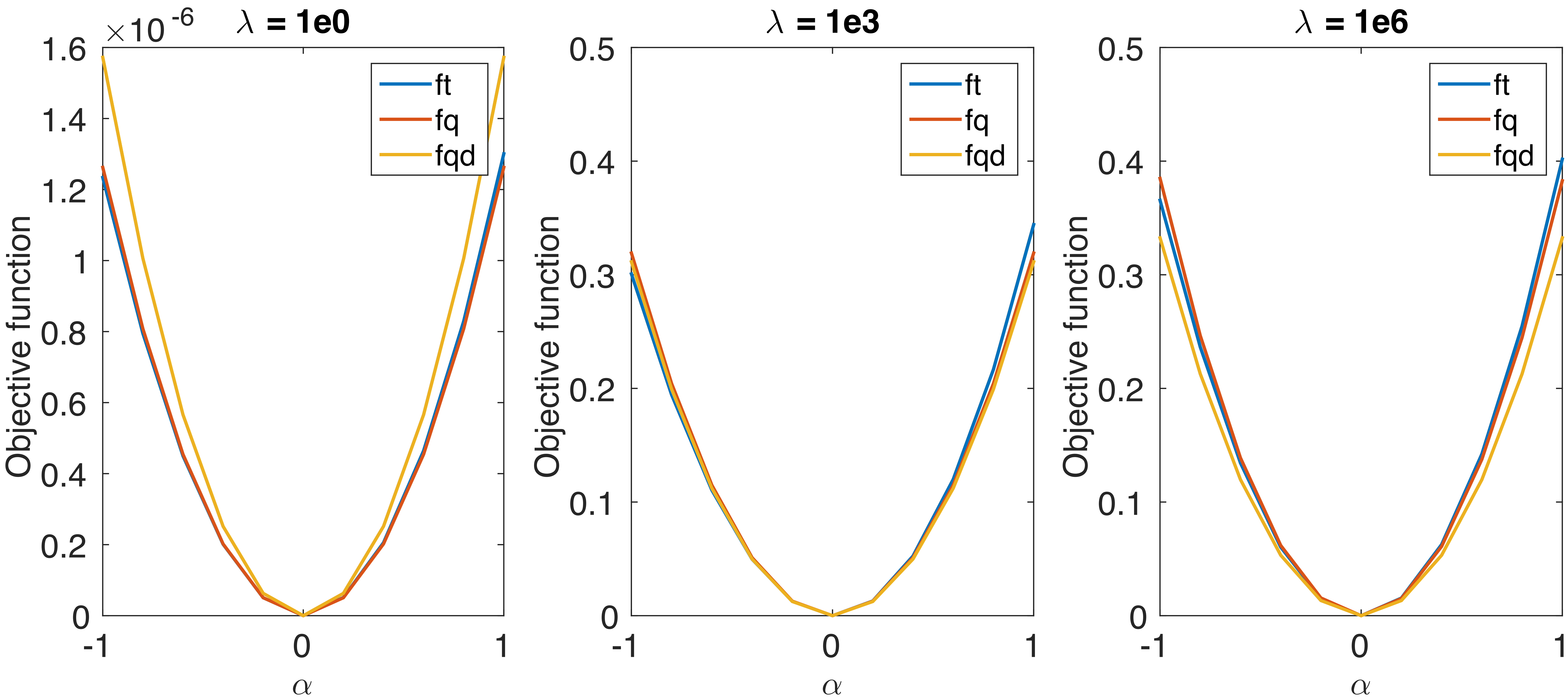
$$h_{i,i} = \mathbf{B}(:,i)^\mathbf{T} * \mathbf{B}(:,i)$$

$h_{i,i}$ =

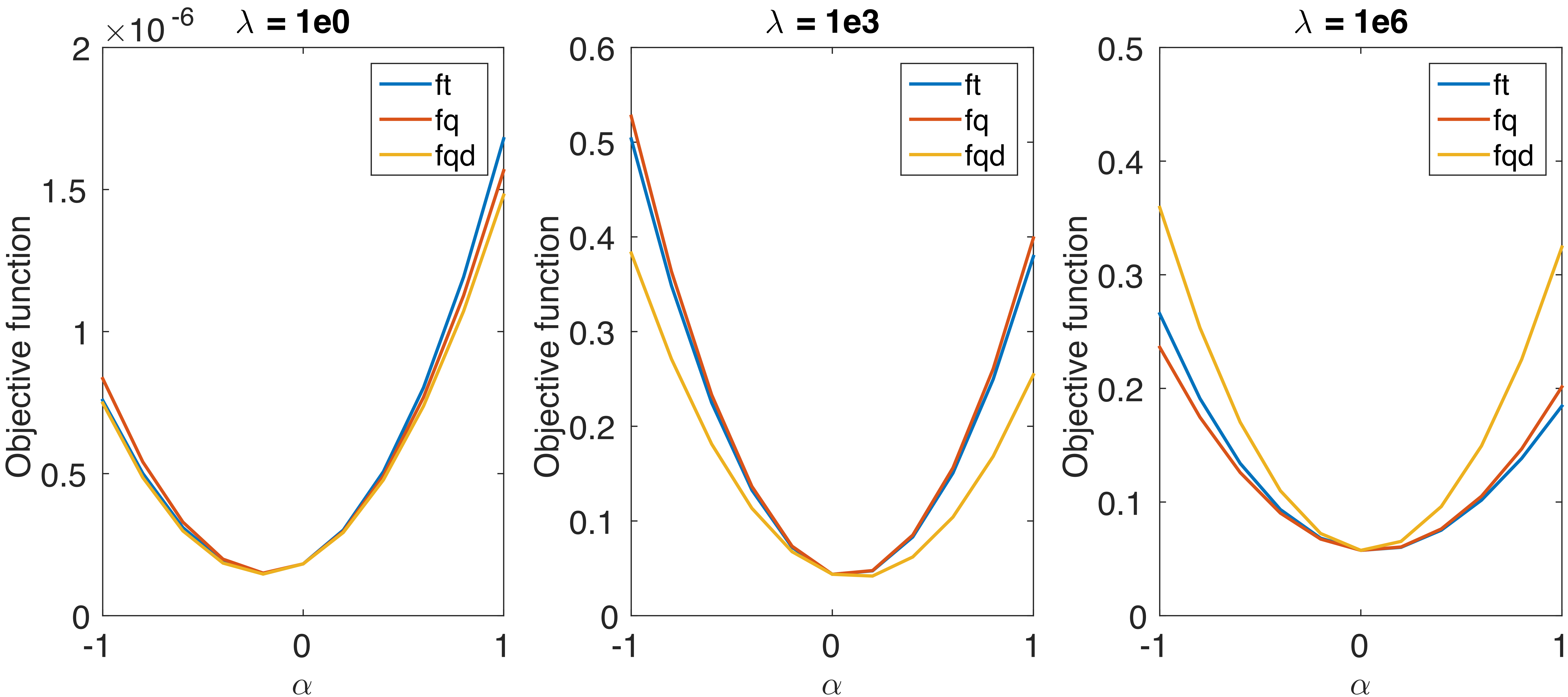


$\mathbf{B}^T$ $\quad$ $\mathbf{B}$

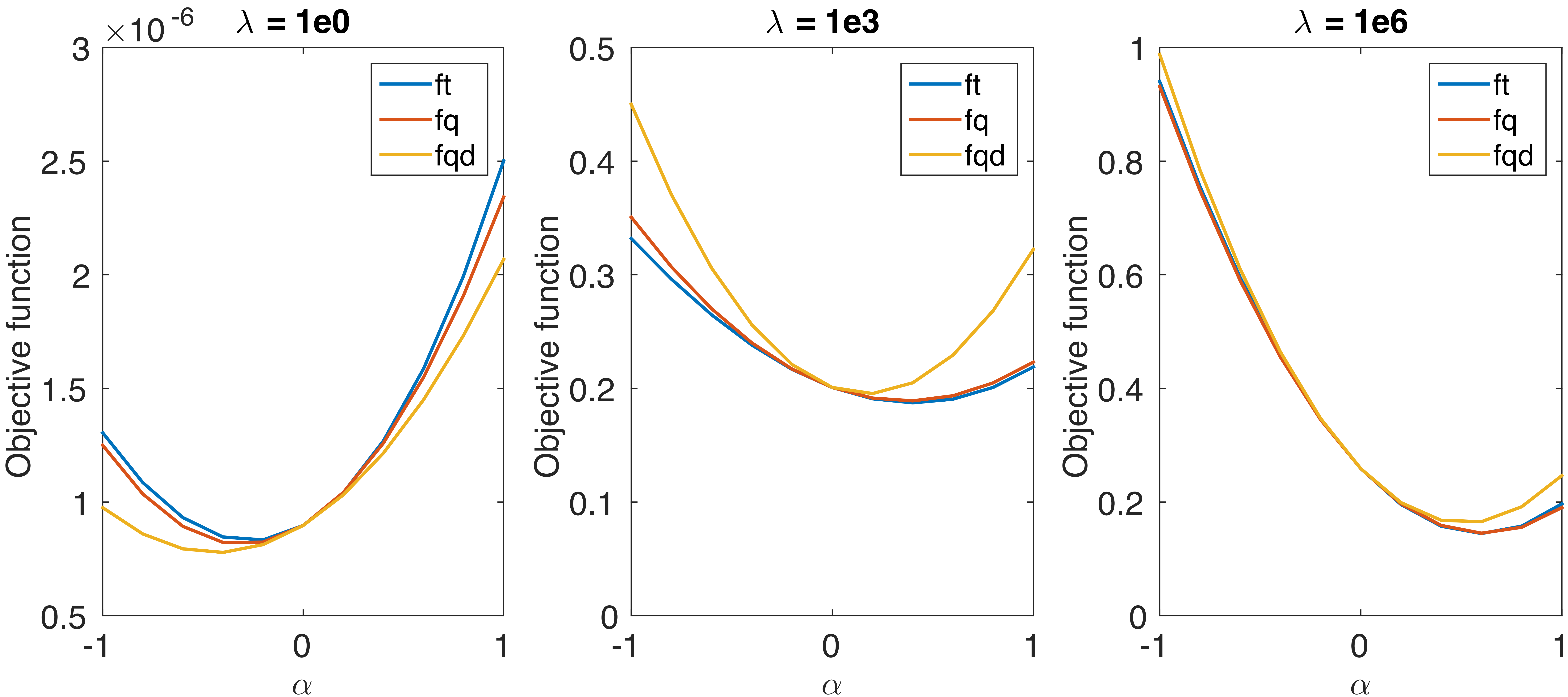
Quadratic approximation with diagonal part of approximated Hessian at m_t



Quadratic approximation with diagonal part of approximated Hessian at m_n



Quadratic approximation with diagonal part of approximated Hessian at m_0



Positive-definite approximation of Hessian

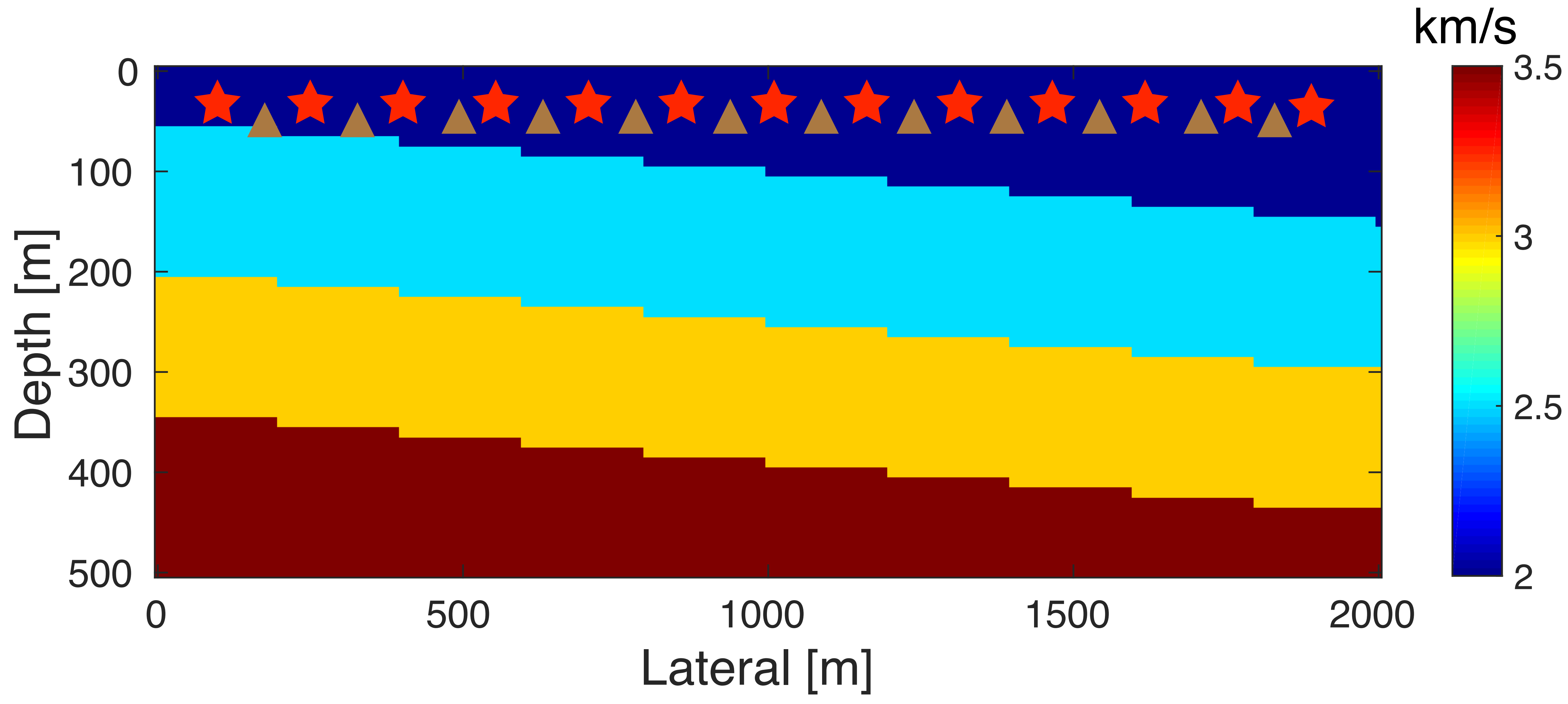
The application of the approximated Hessian and its diagonal part:

- Invert the MAP point with Newton type method (PDE free and matrix free)
- Quantify the uncertainty of the inversion result (standard deviation)

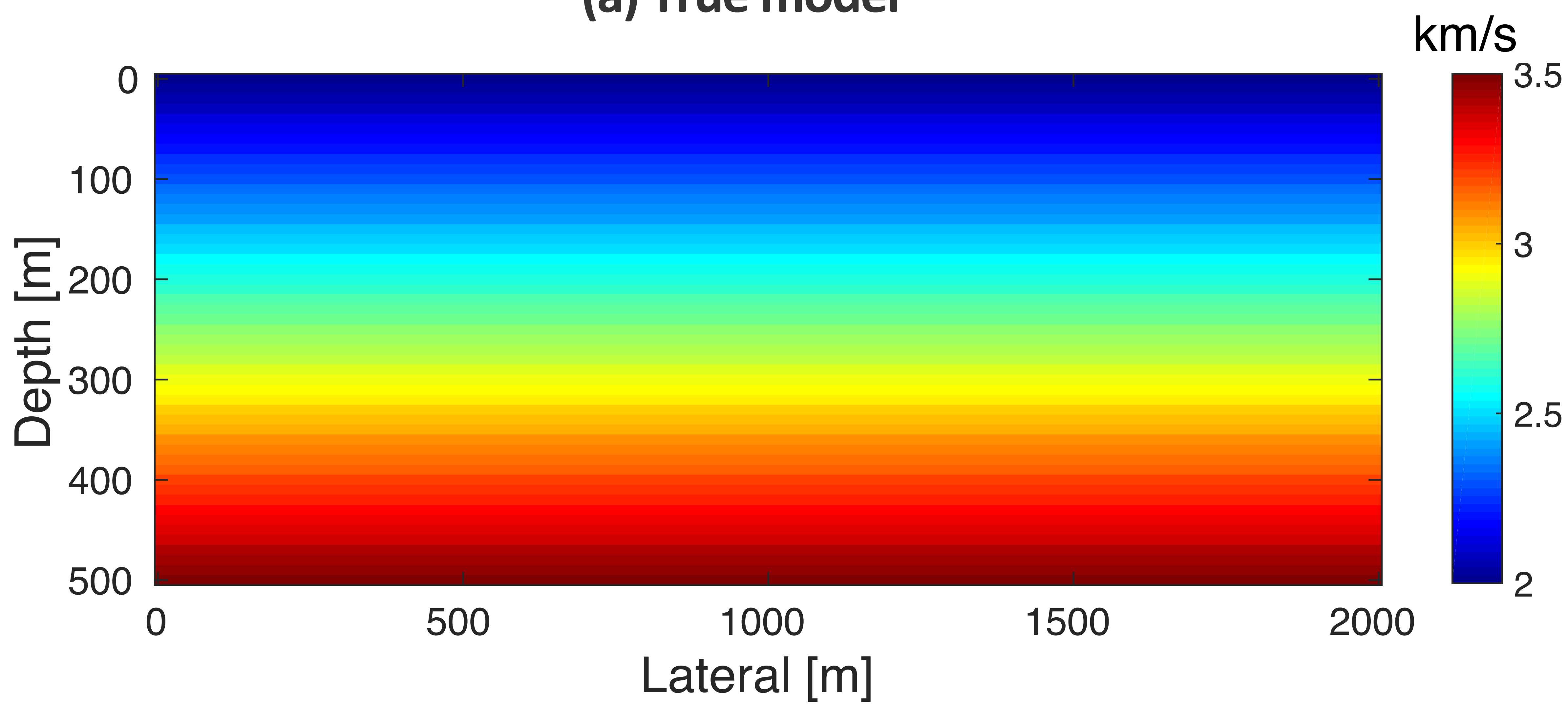
Workflow
– uncertainty quantification

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

Numerical experiment (Full acquisition)



(a) True model

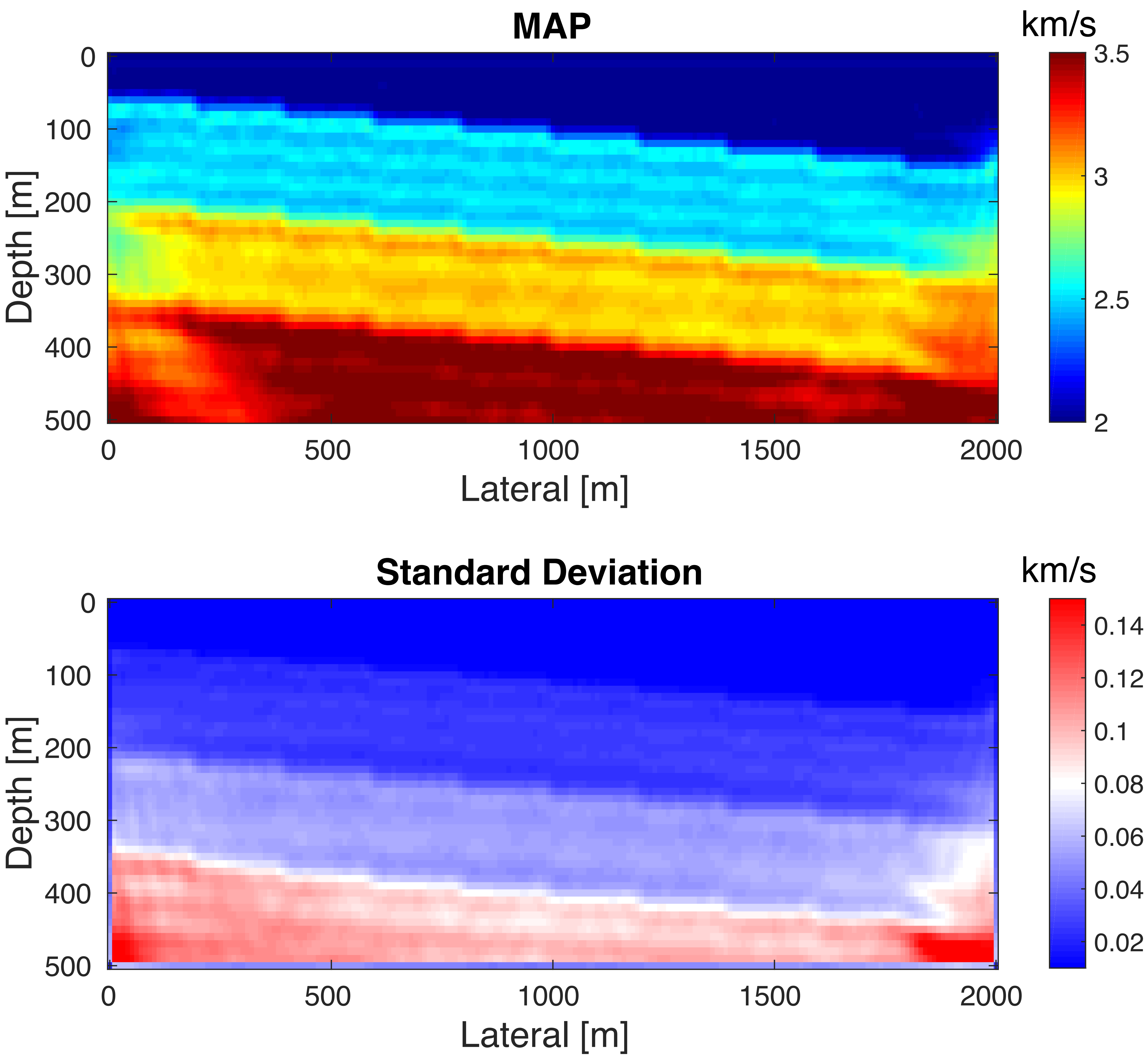


(b) Initial model

Model size: 500m x 2000m
Source spacing: 80m
Receiver spacing: 20m
Fixed spread 2km
Frequency : 10-31 Hz

Standard deviation of data noise: 0.1
Standard deviation of pde: 0.1
Standard deviation of model: 1
lambda: 1

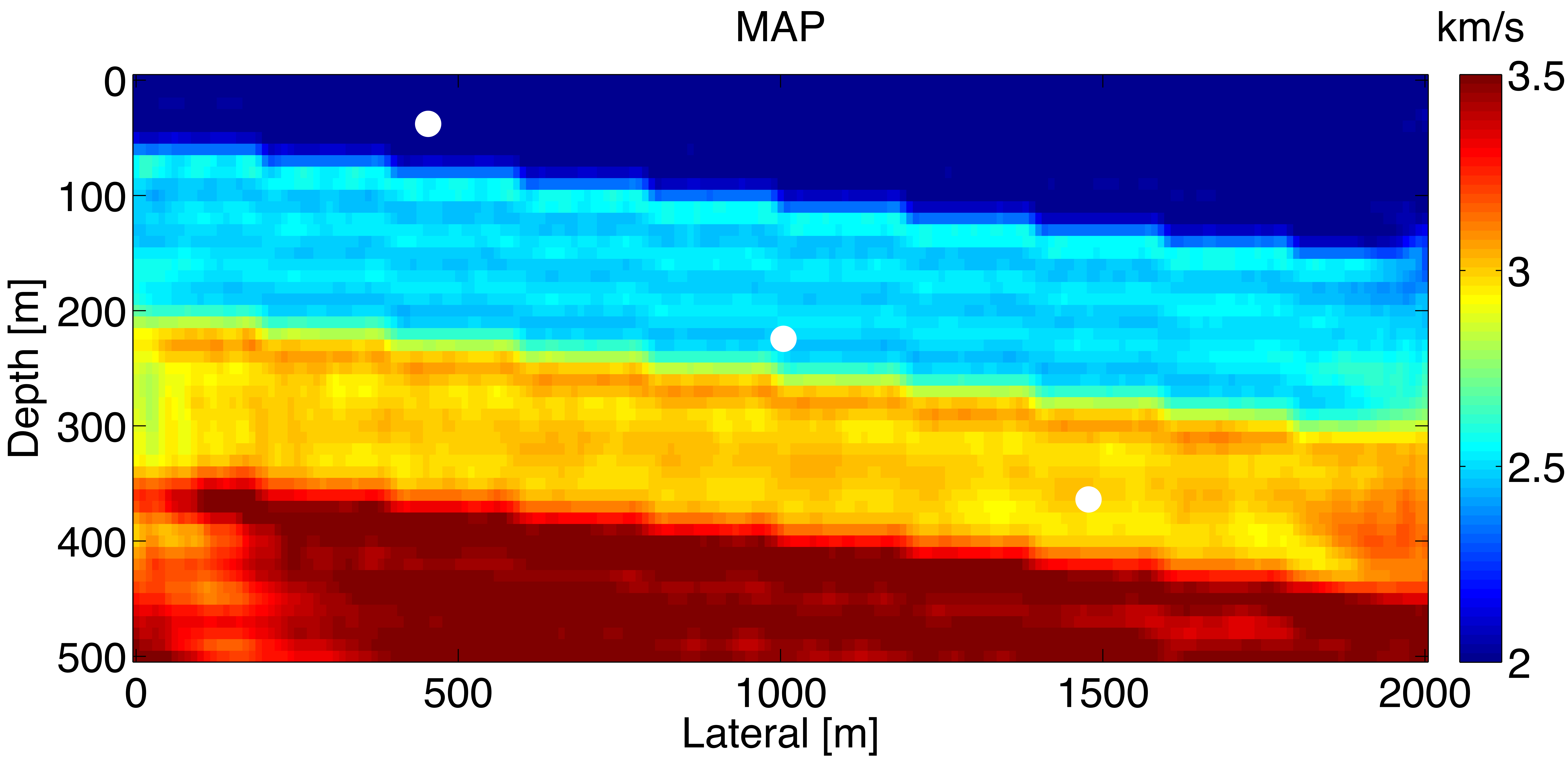
Numerical experiment (Full acquisition)



a) Maximum a posteriori estimate

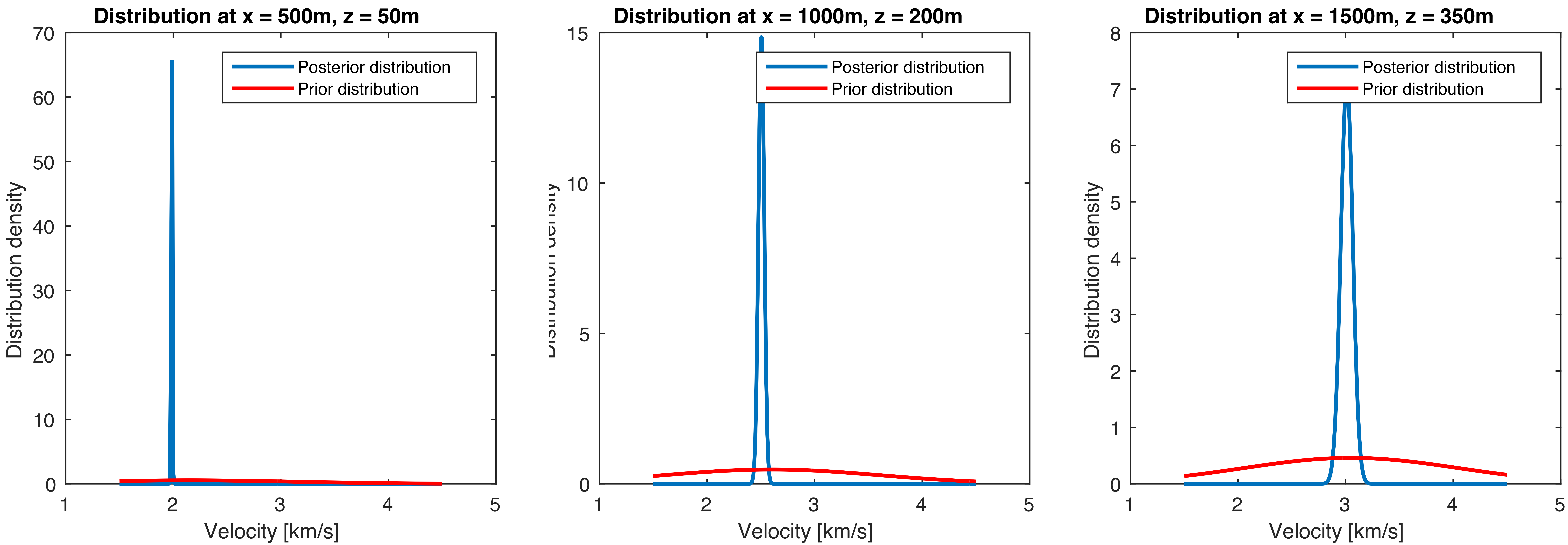
b) The standard deviation

Posterior distribution



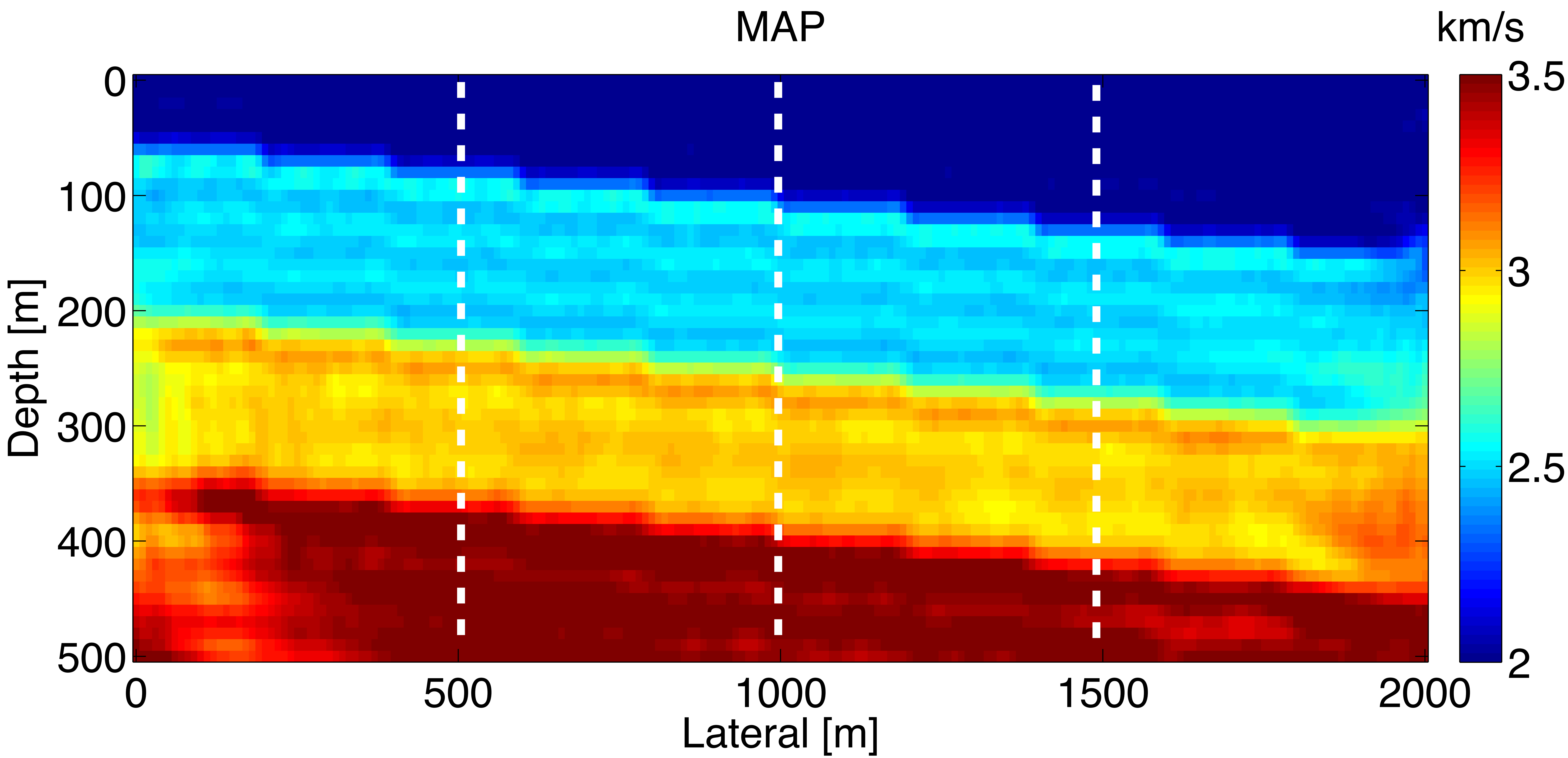
Maximum a posterior

Posterior distribution vs Prior distribution



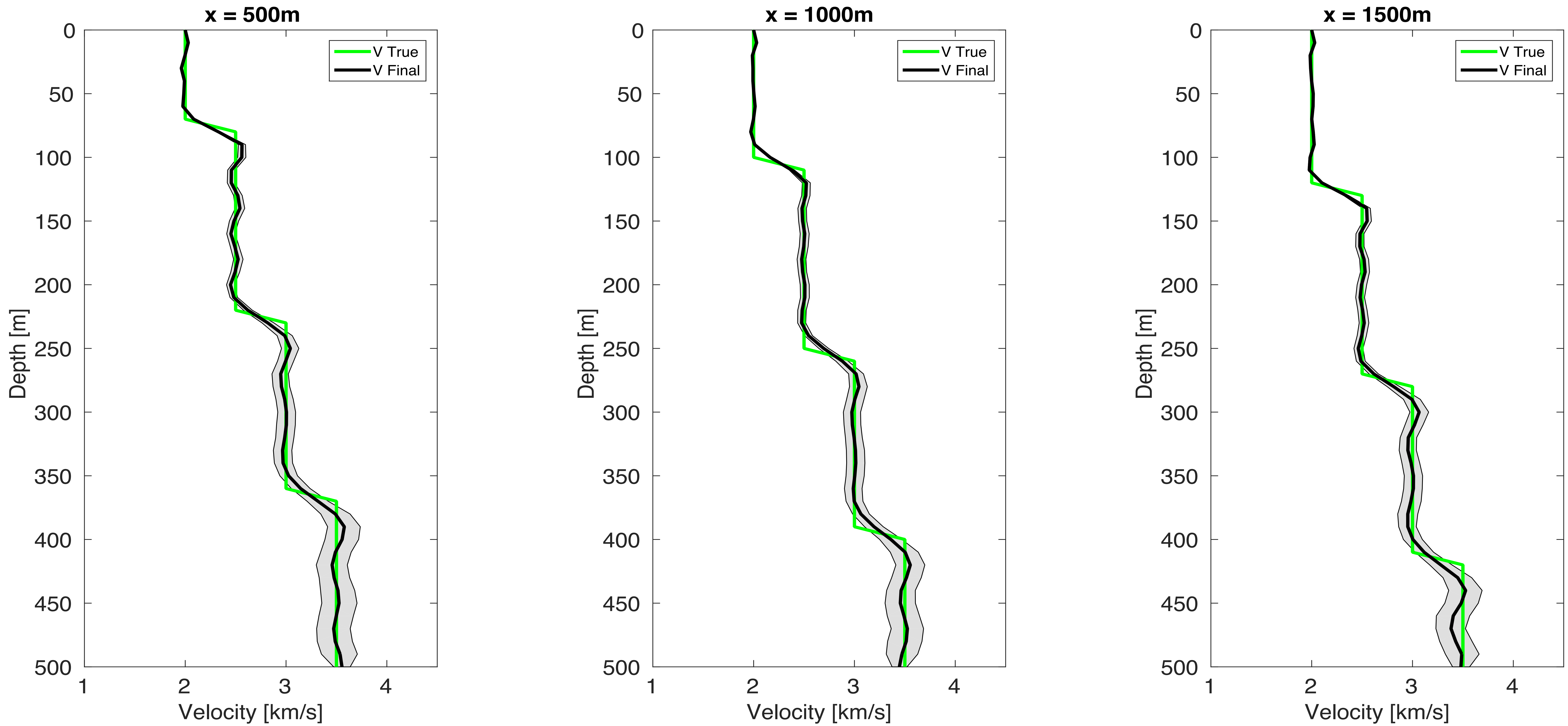
(a) **(b)** **(c)**

Confidence intervals (90%)



Maximum a posterior

Confidence intervals (90%)



(a) (b) (c)

Confidence intervals

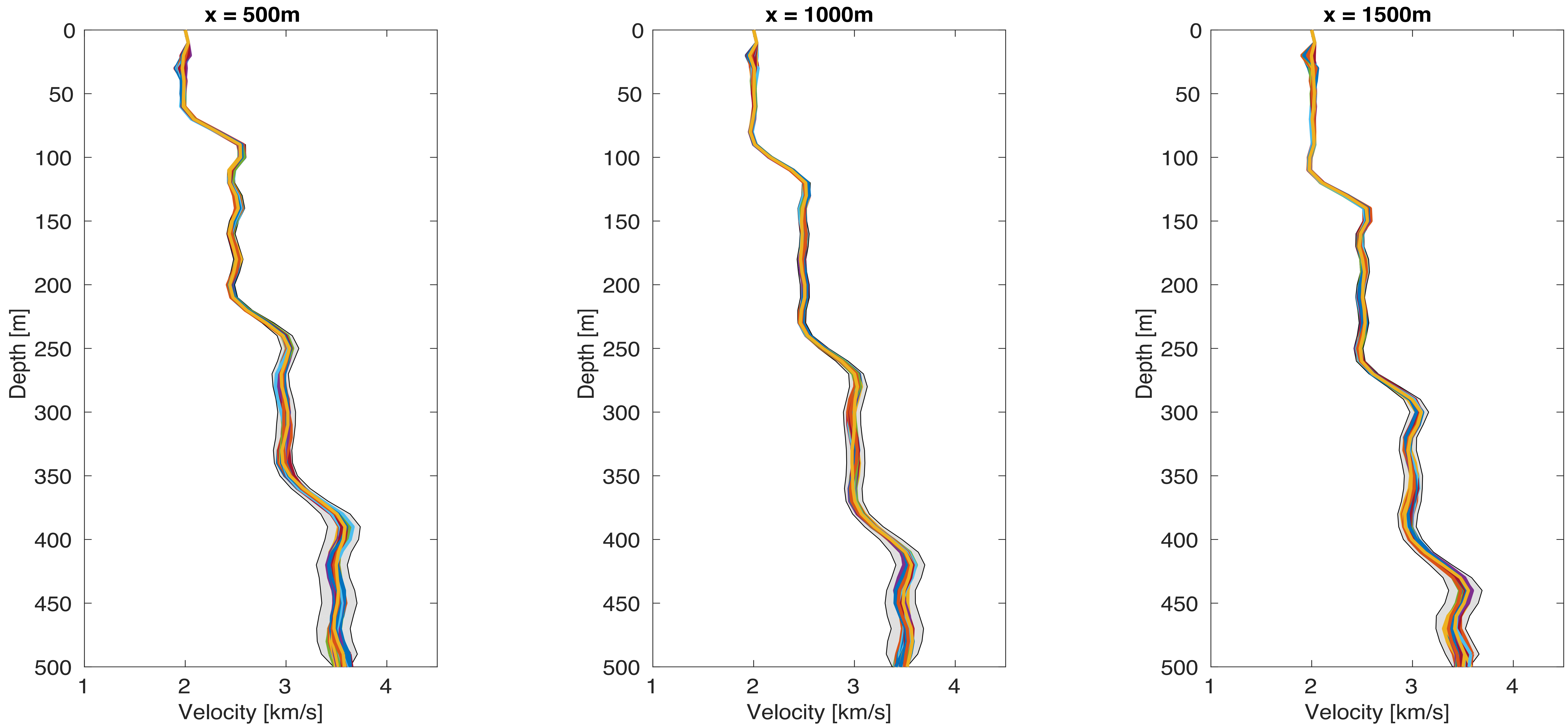
100 random realizations of data:

$$\mathbf{d}_i = \mathbf{F}(\mathbf{m}_{\mathrm{t}}) + \epsilon_i, \epsilon_i \sim \mathcal{N}(0, \sigma^2\mathbf{I})$$

Inversion results correspond to these data:

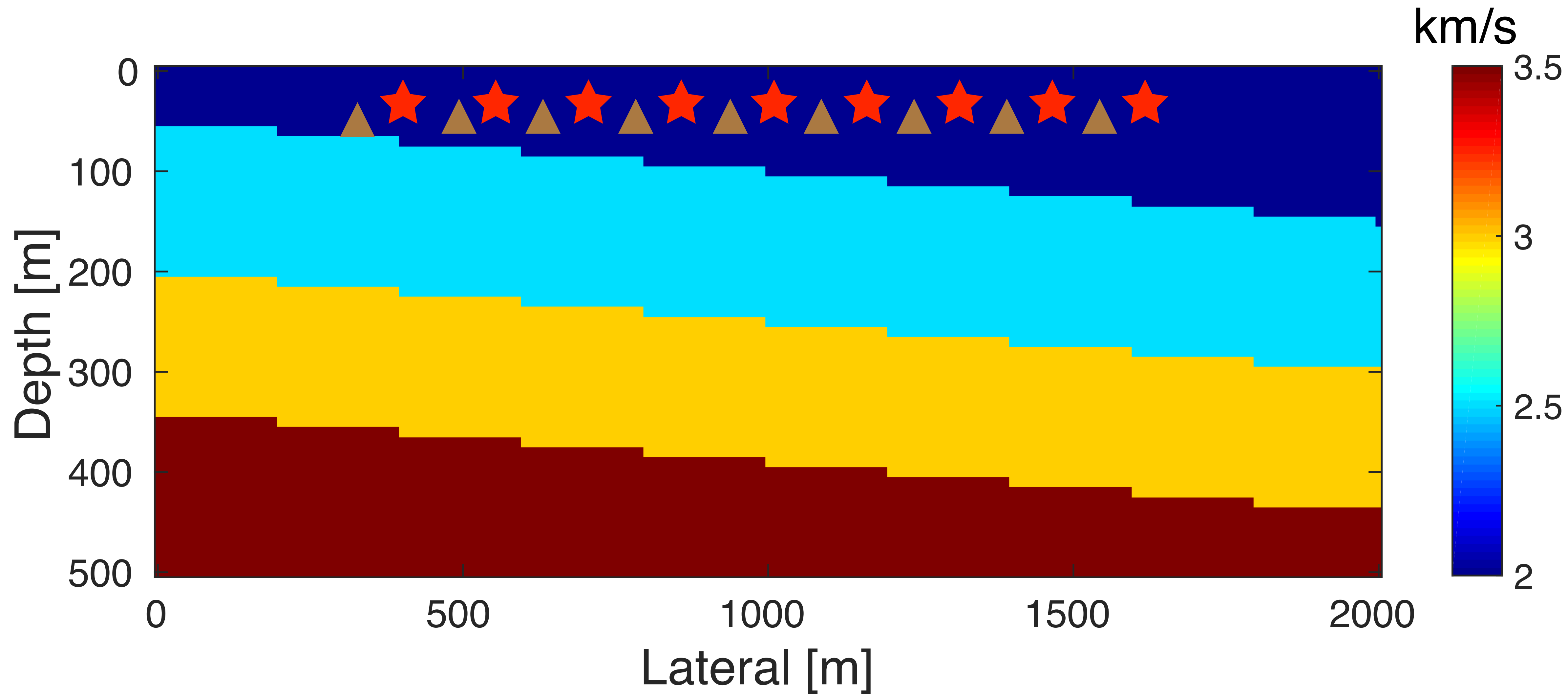
$$\mathbf{d}_i \rightarrow \mathbf{m}_{\mathrm{i}}$$

Confidence intervals (90%)
— Inversion results of 100 random realizations

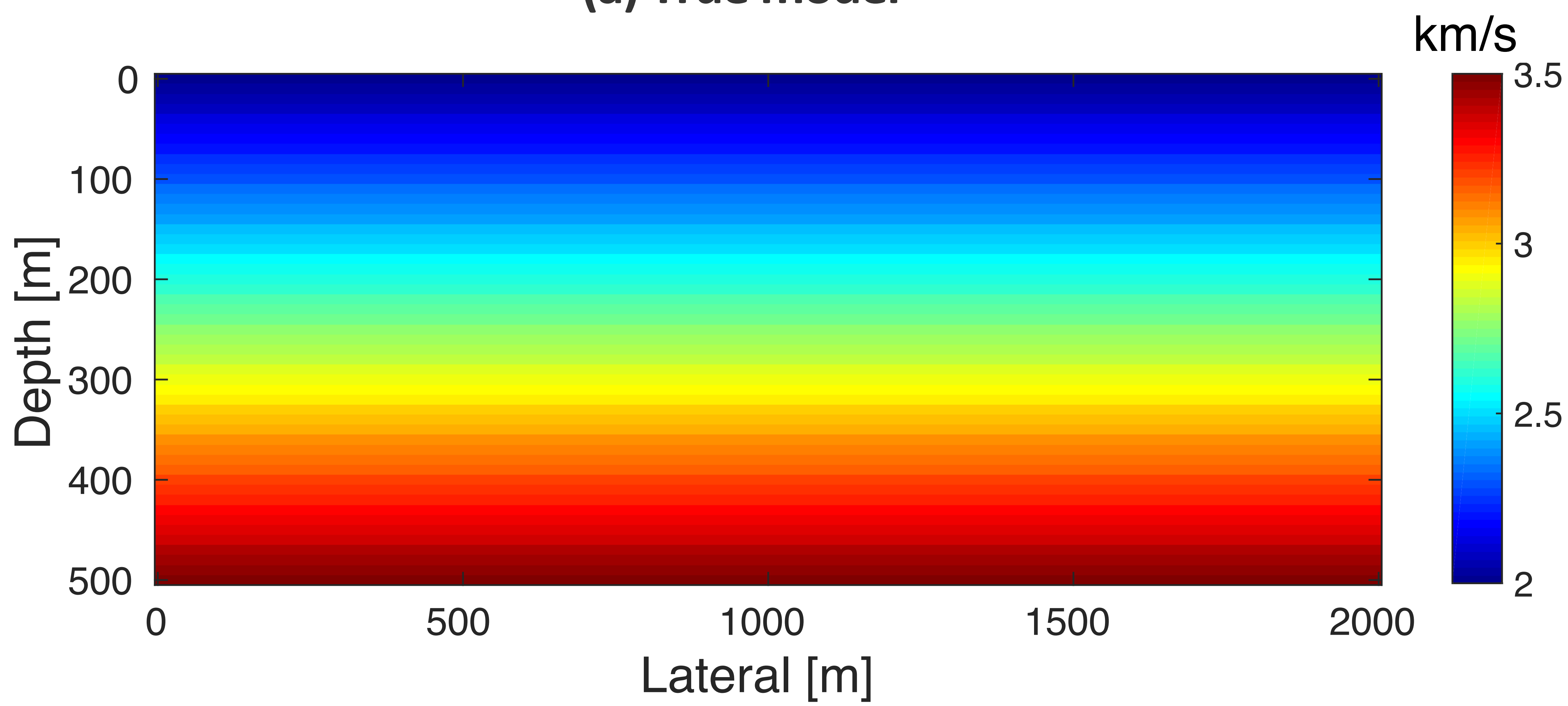


(a) (b) (c)

Numerical experiment (Partial acquisition)



(a) True model

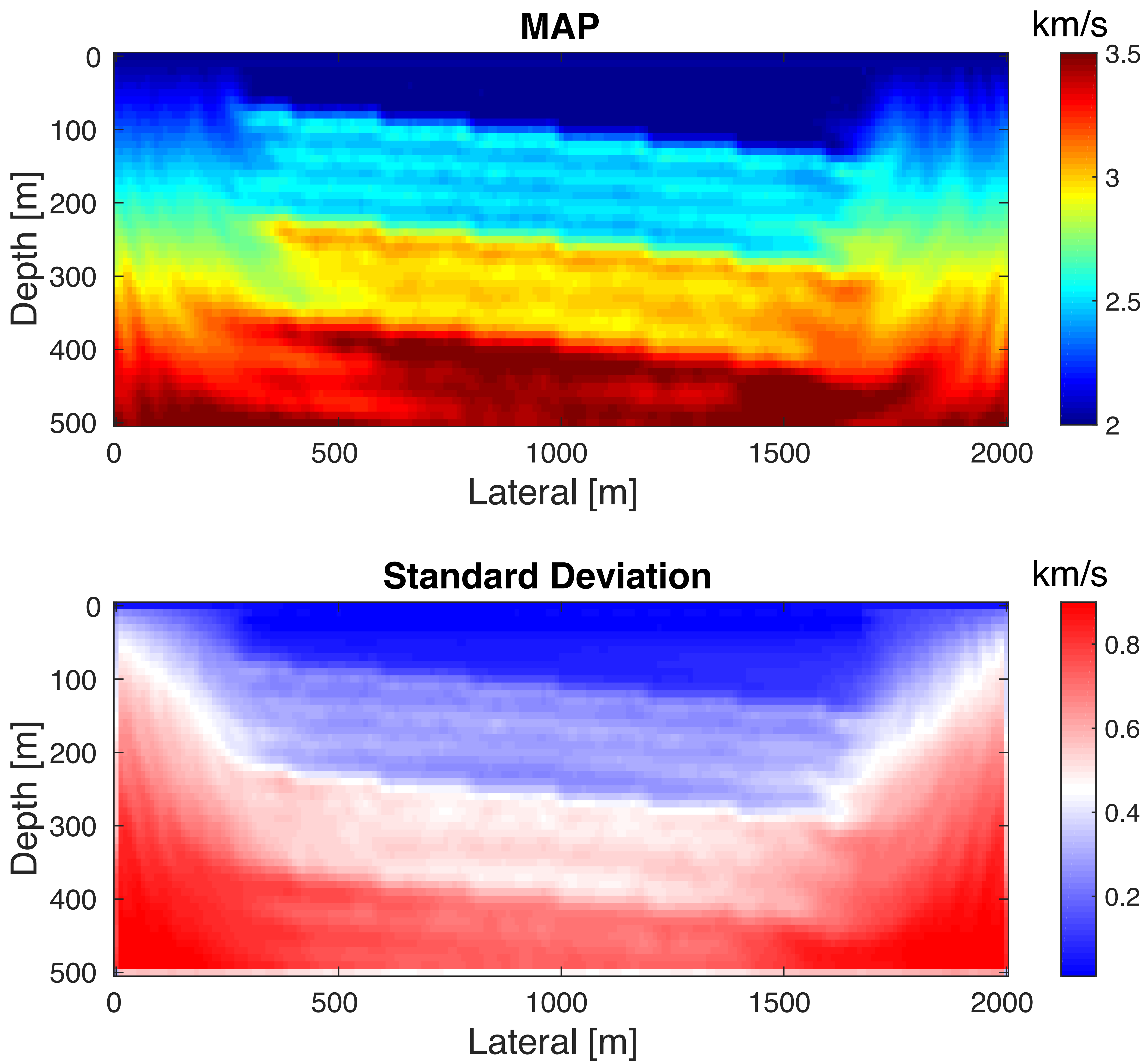


(b) Initial model

Model size: 500m x 2000m
Source spacing: 80m
Receiver spacing: 20m
Fixed spread 2km
Frequency : 10-31 Hz

Standard deviation of data noise: 0.1
Standard deviation of pde: 0.1
Standard deviation of model: 1
lambda: 1

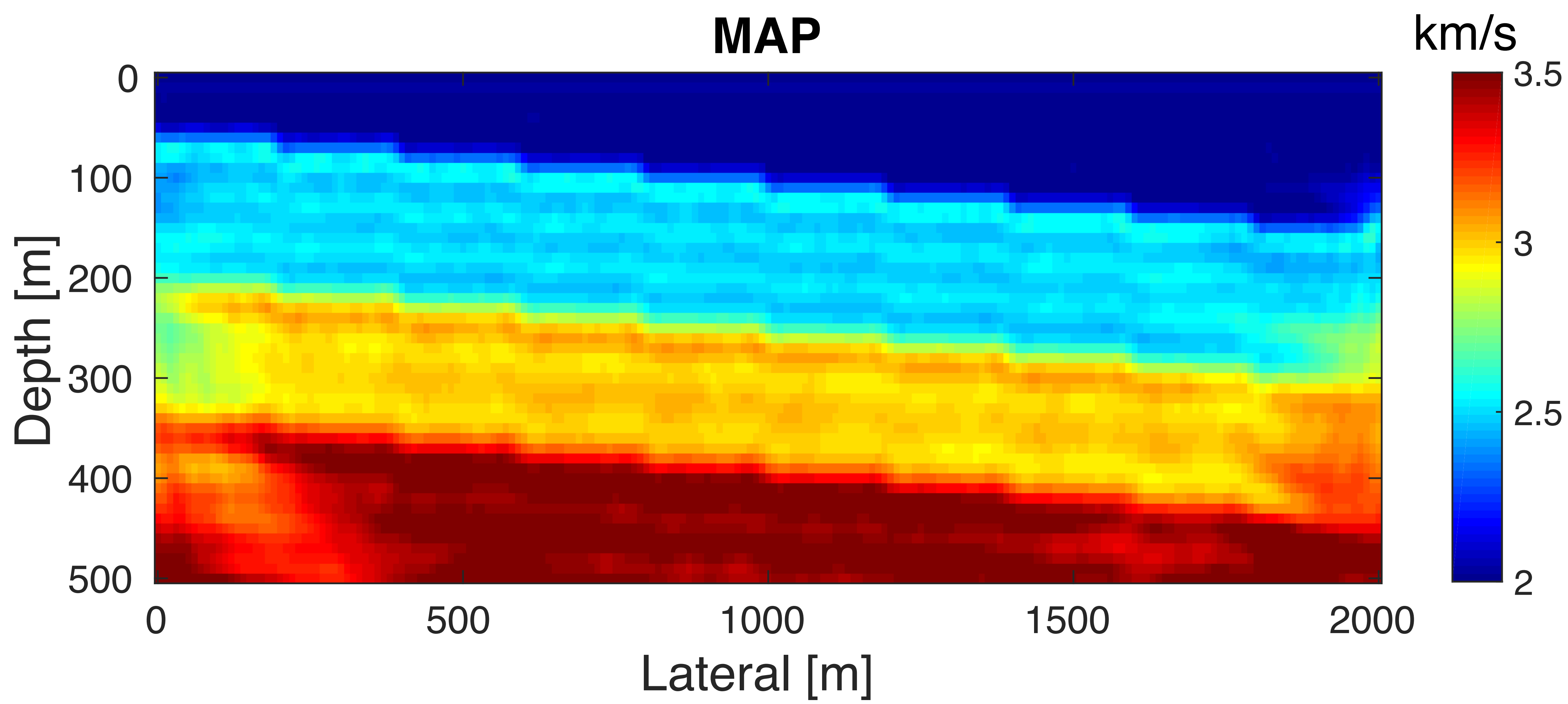
Numerical experiment (Partial acquisition)



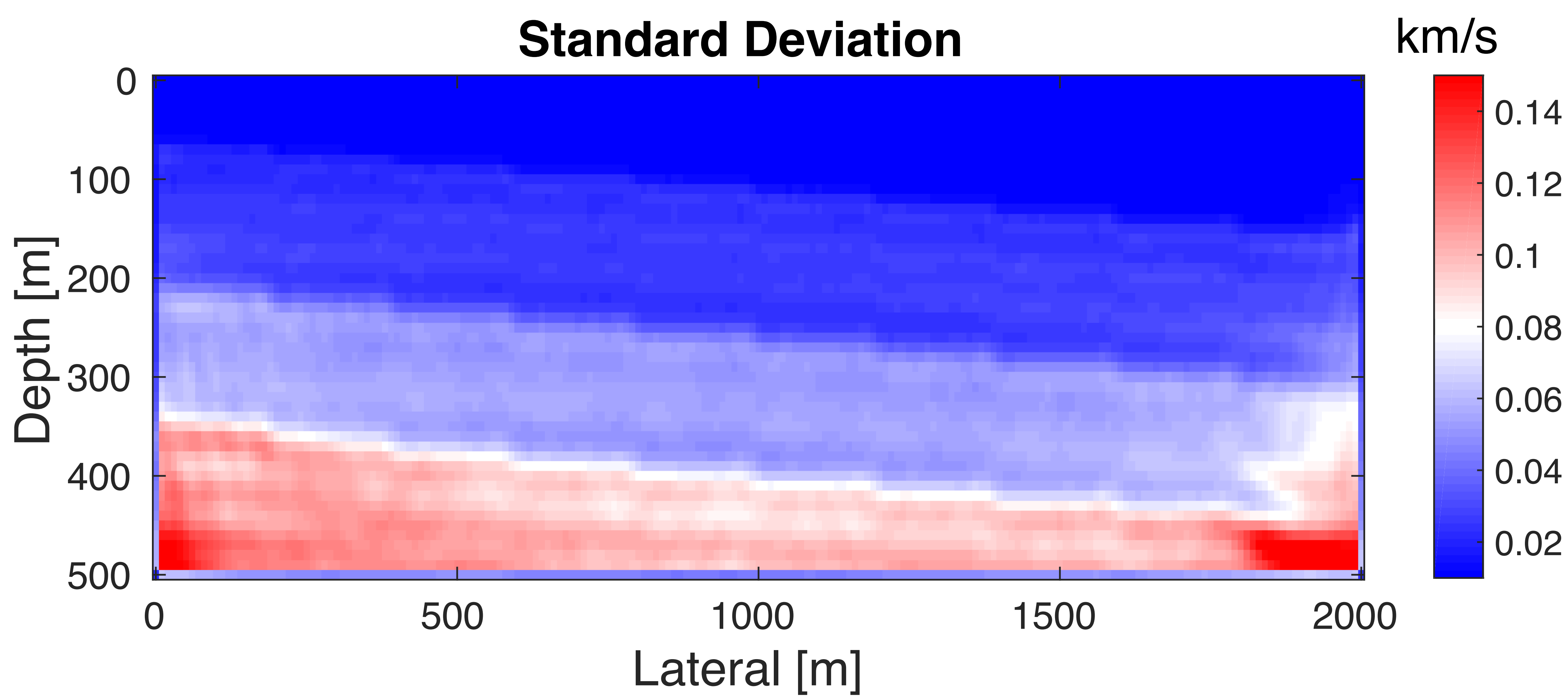
a) Maximum a posteriori estimate

b) The standard deviation

Full acquisition

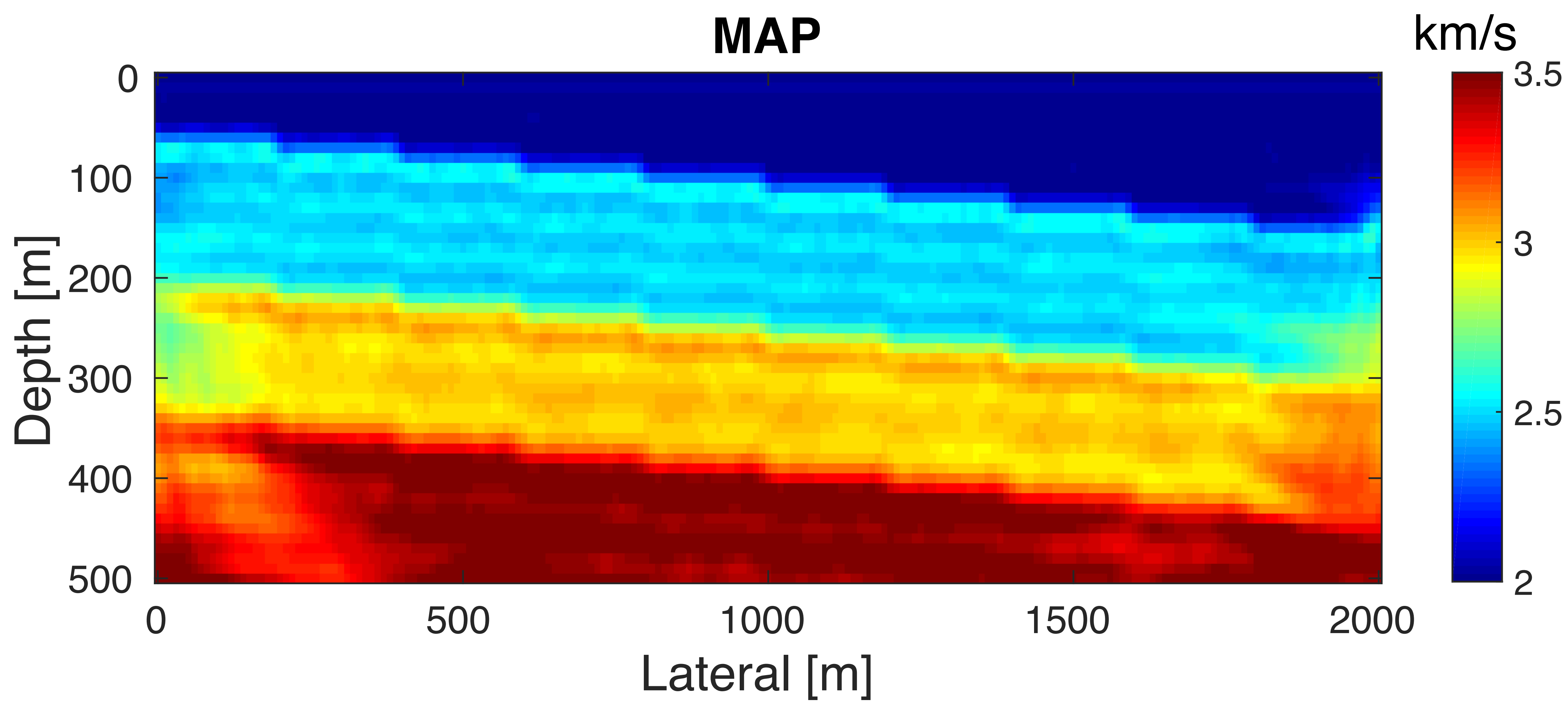


a) Maximum a posteriori estimate

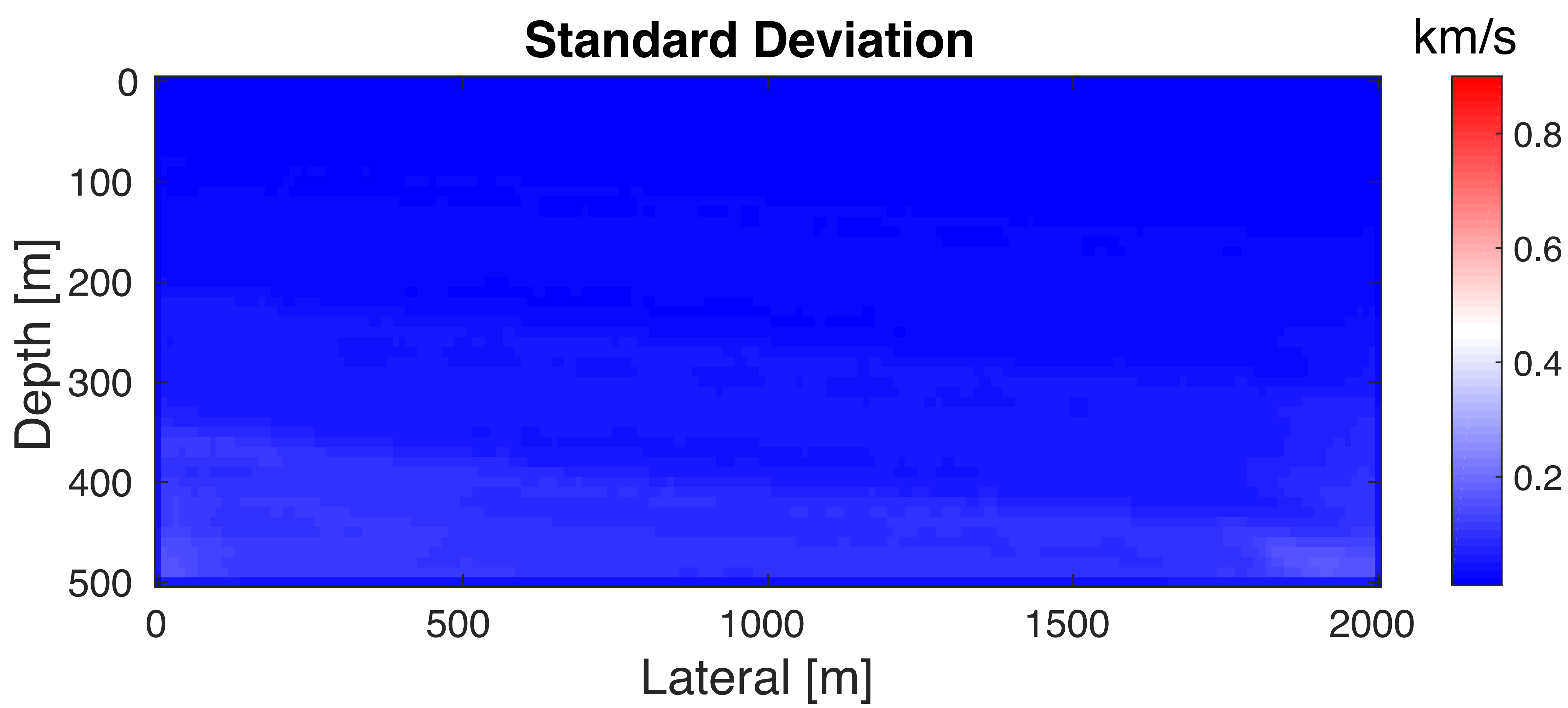


b) The standard deviation

Full acquisition

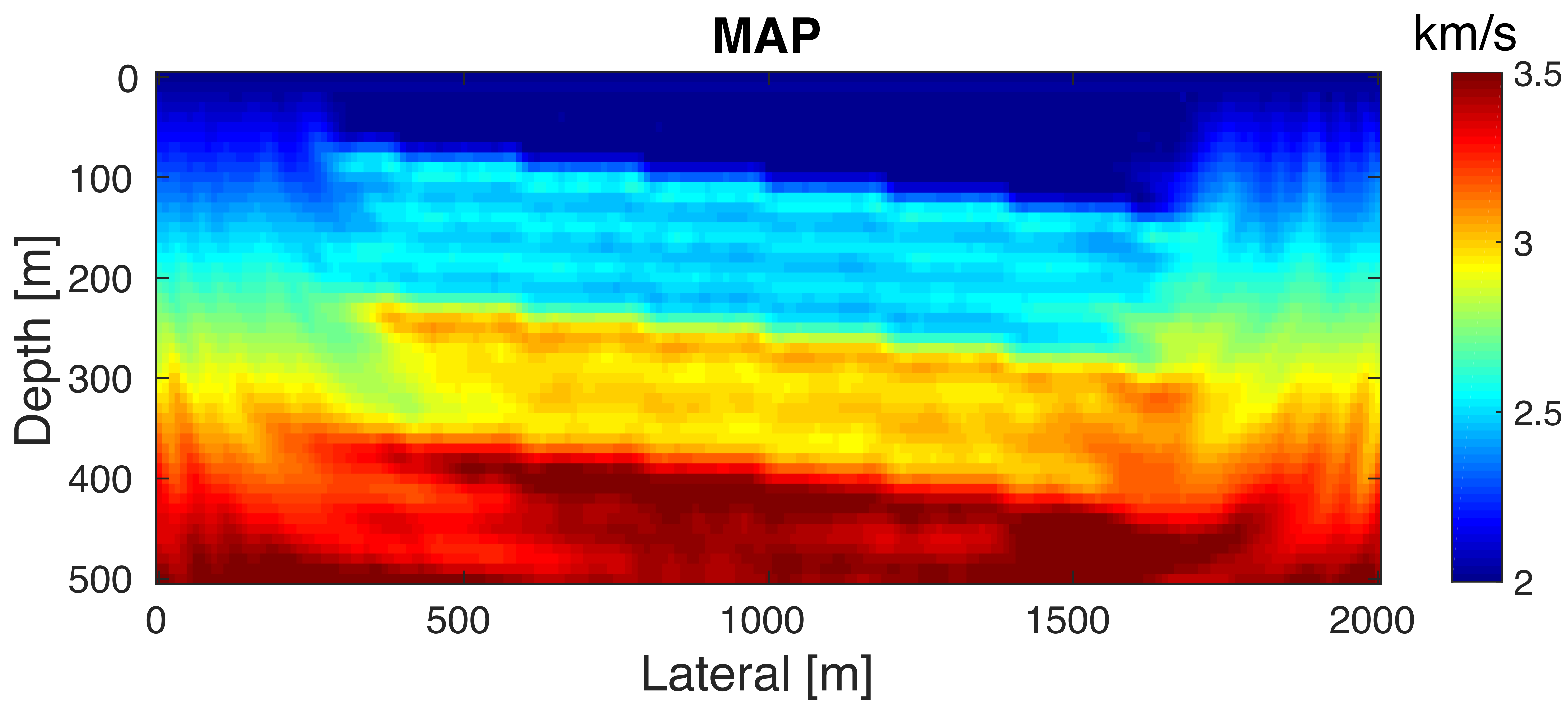


a) Maximum a posteriori estimate

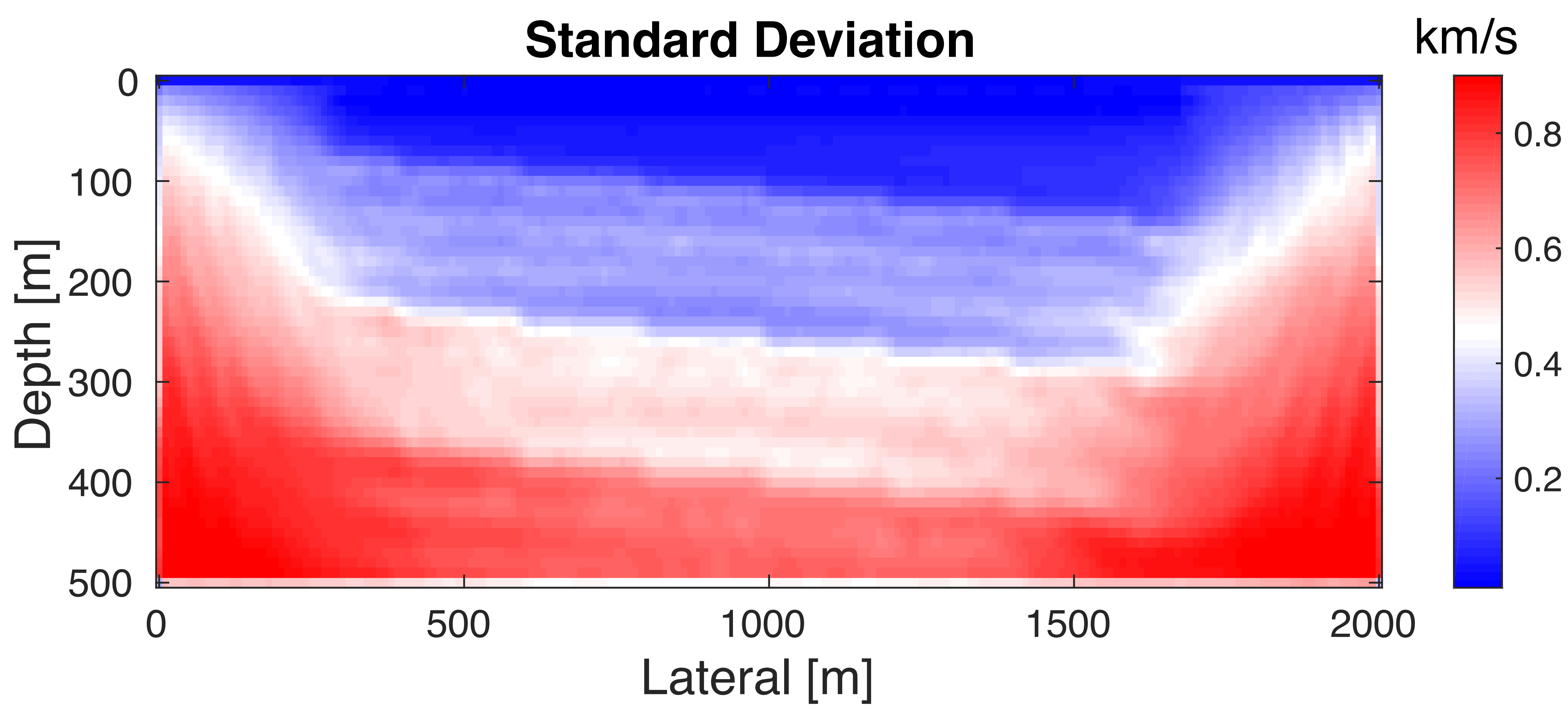


b) The standard deviation

Partial acquisition

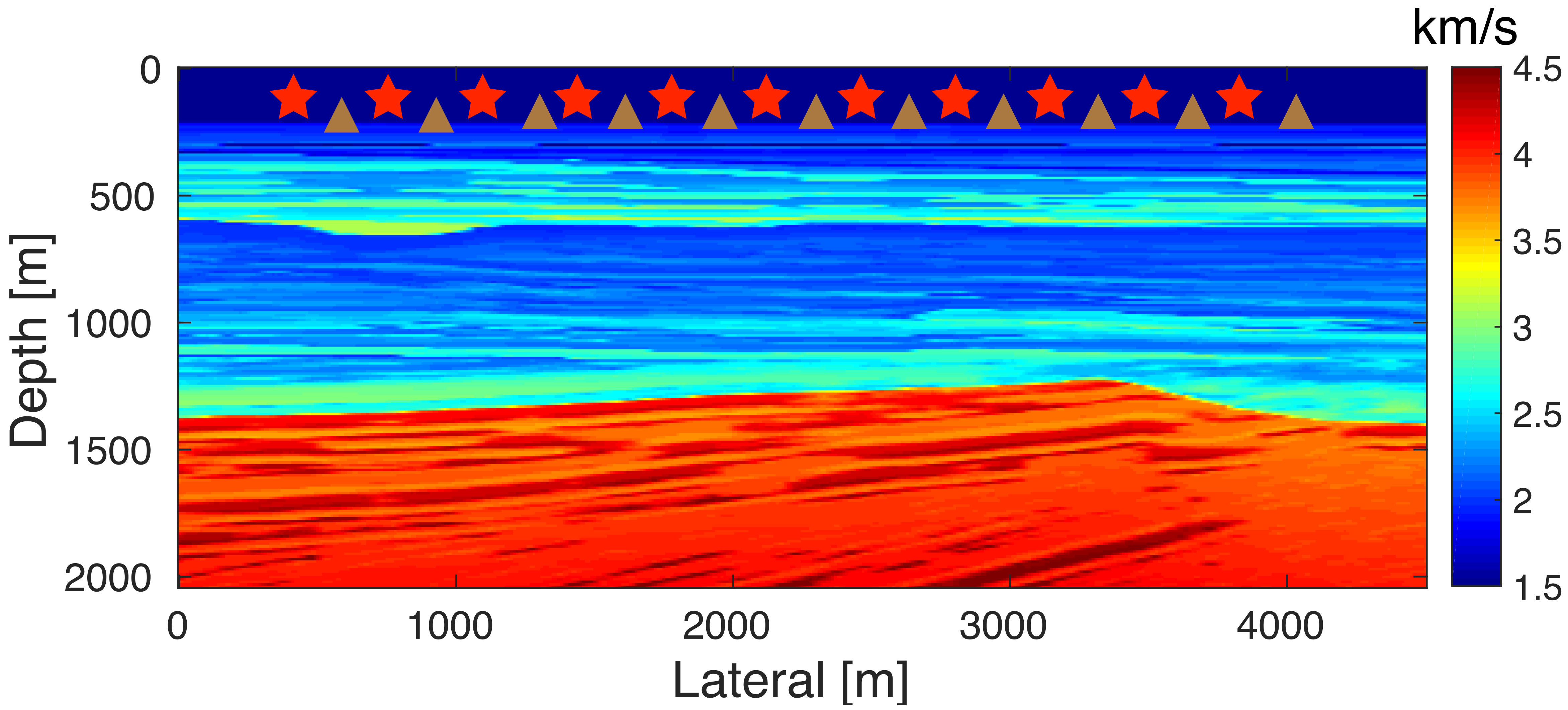


a) Maximum a posteriori estimate

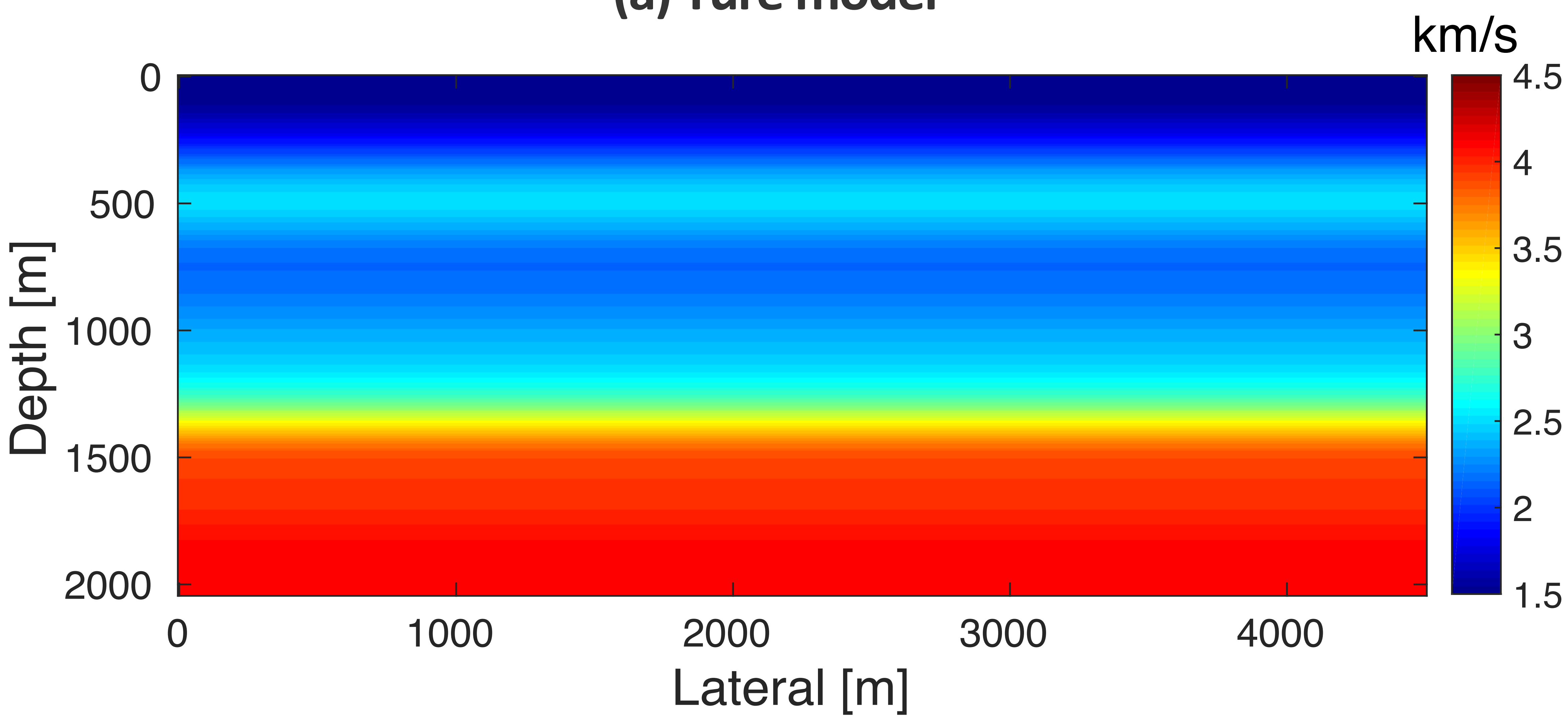


b) The standard deviation

BG model



(a) Ture model

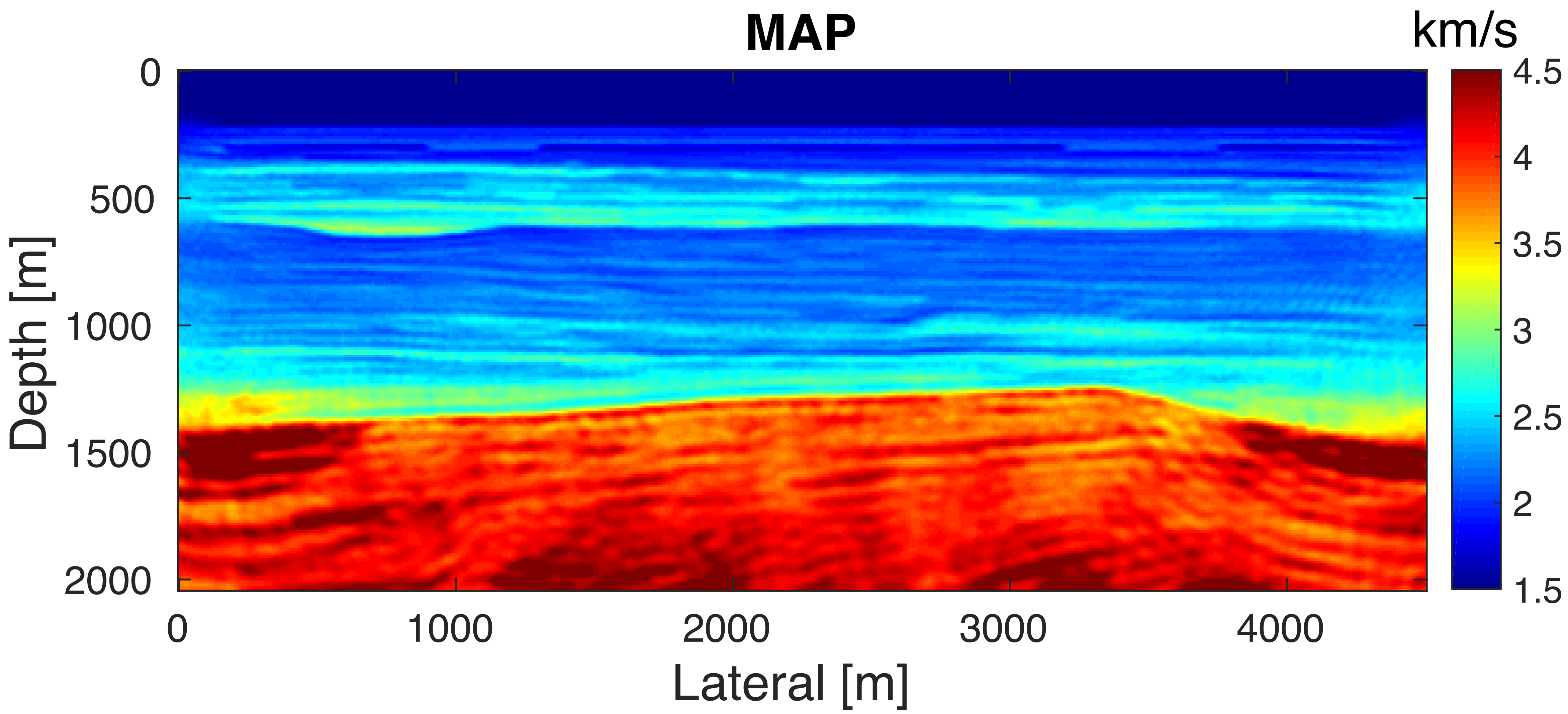


(b) Initial model

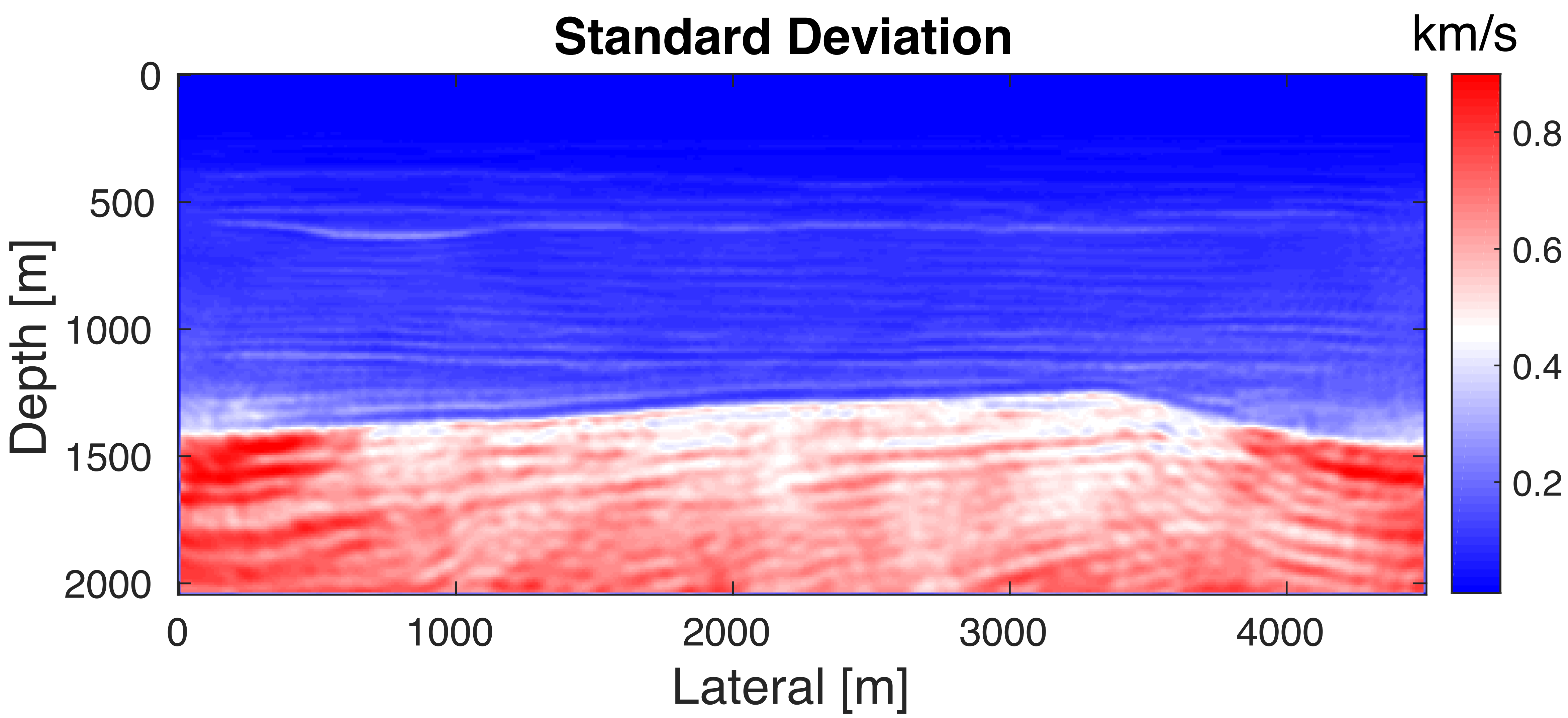
Model size: 2000m x 4500m
Source spacing: 50m
Receiver spacing: 10m
Fixed spread 4.5km
Frequency : 5~31 Hz

Standard deviation of data noise: 0.5
Standard deviation of pde: 0.5
Standard deviation of model: 1
lambda: 1

BG model

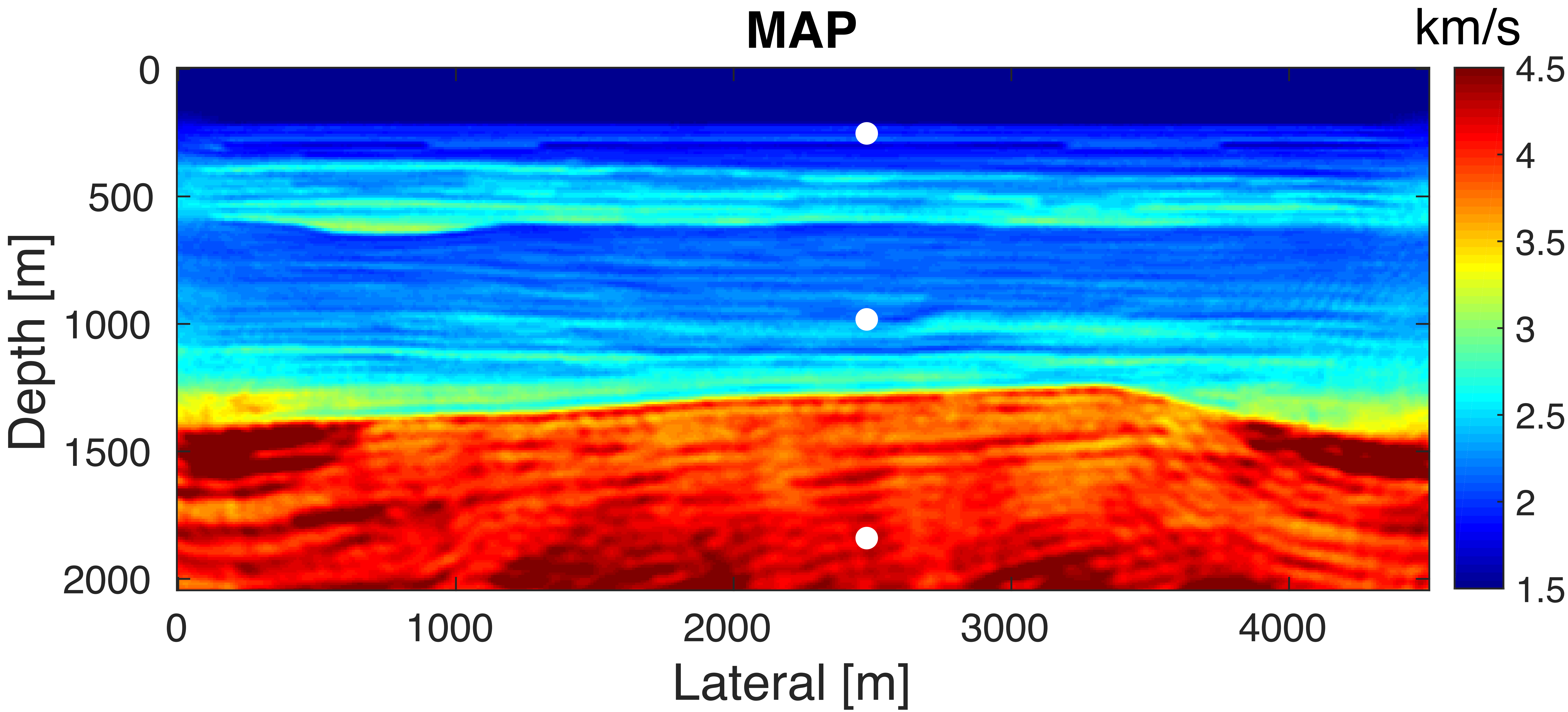


a) Maximum a posterior estimate



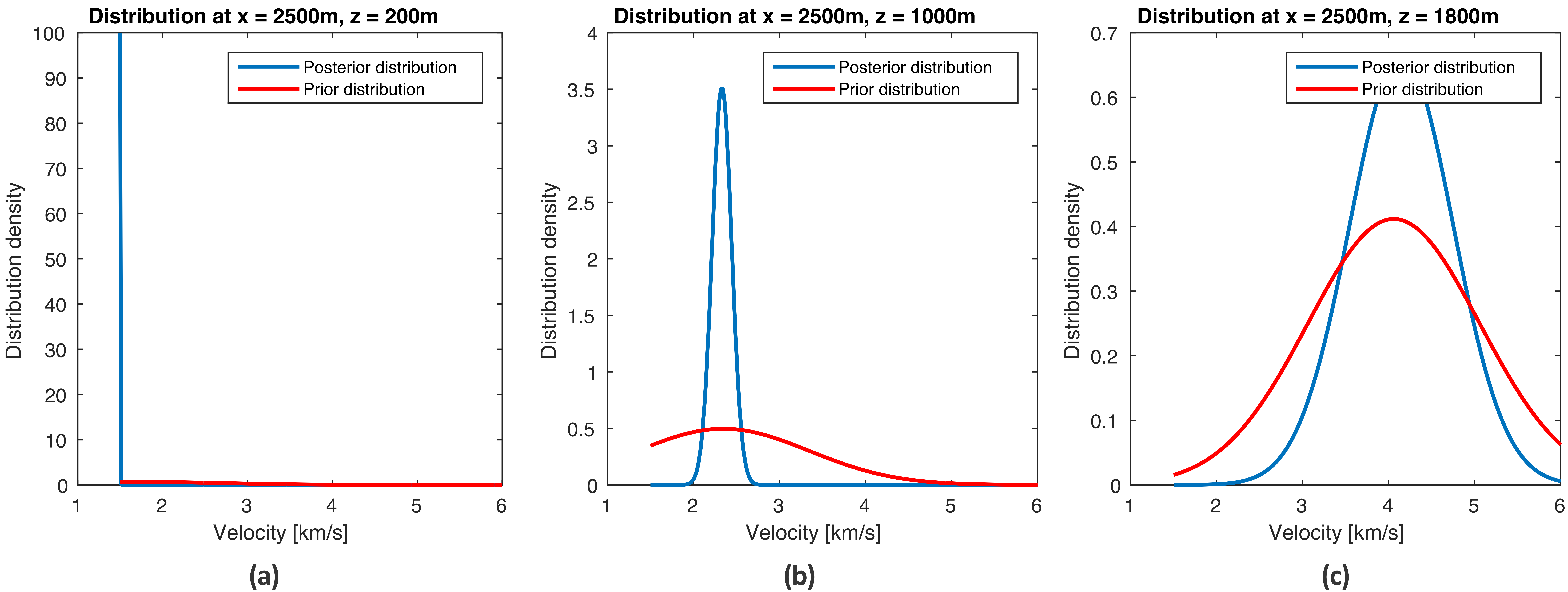
b) The standard deviation

Posterior distribution

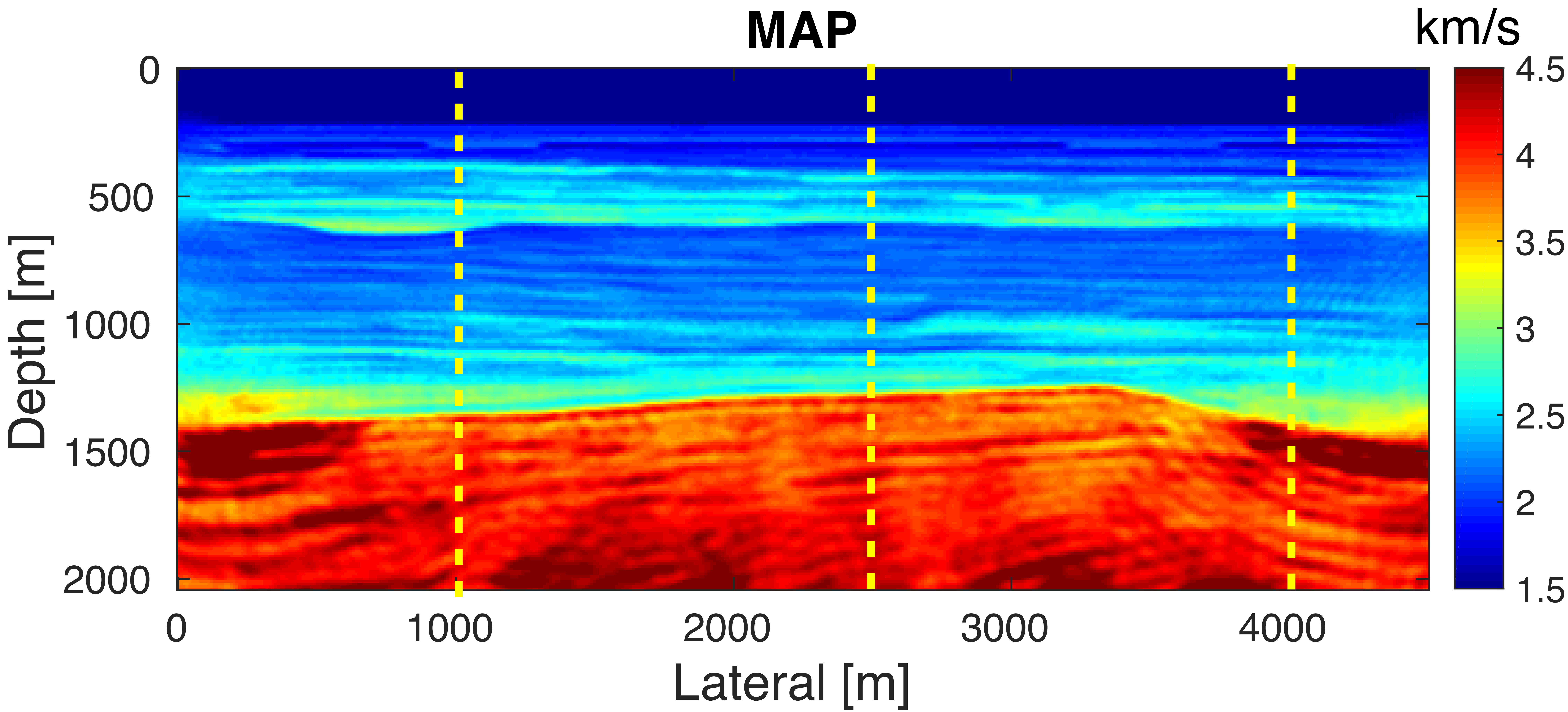


Maximum a posterior

Posterior distribution vs Prior distribution

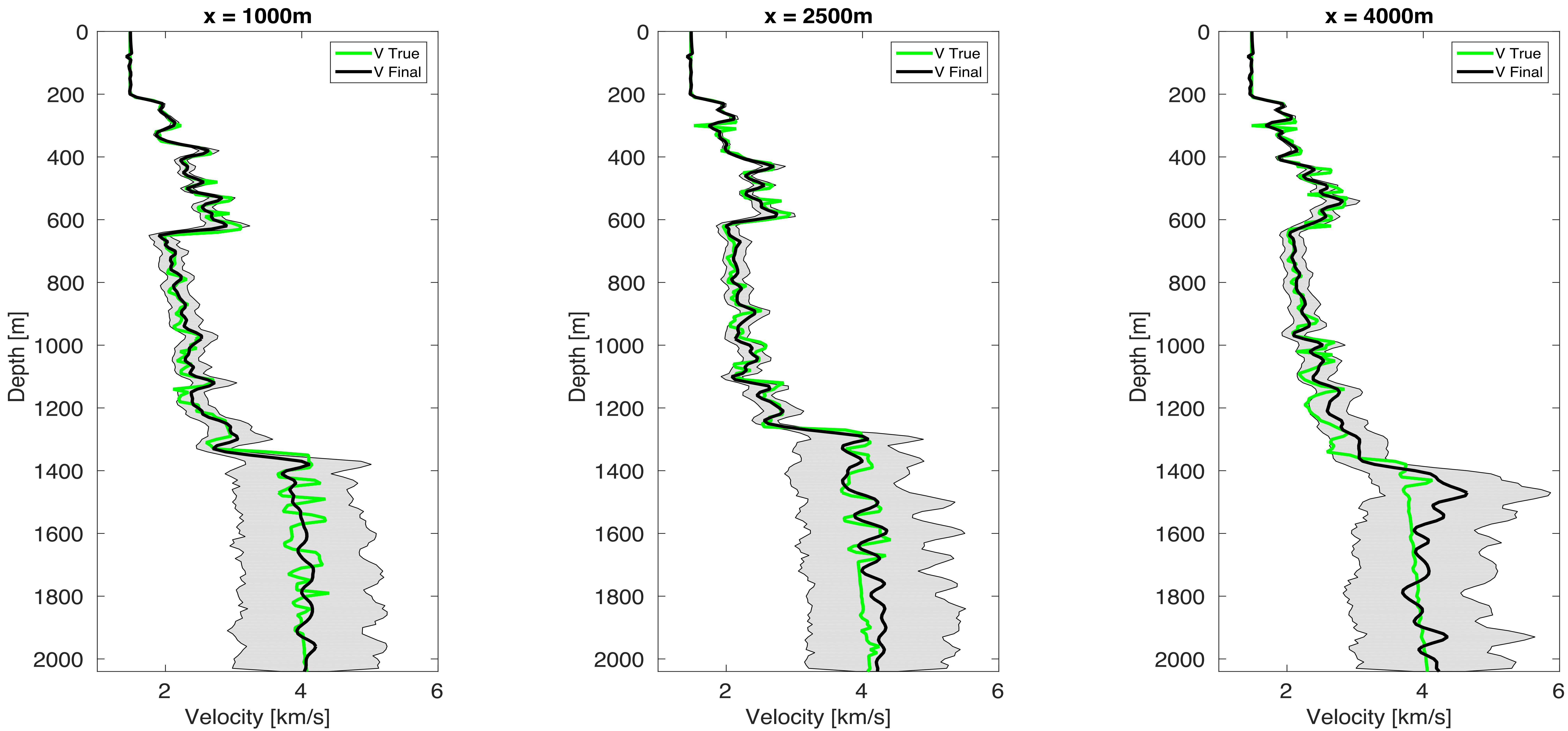


Confidence intervals (90%)



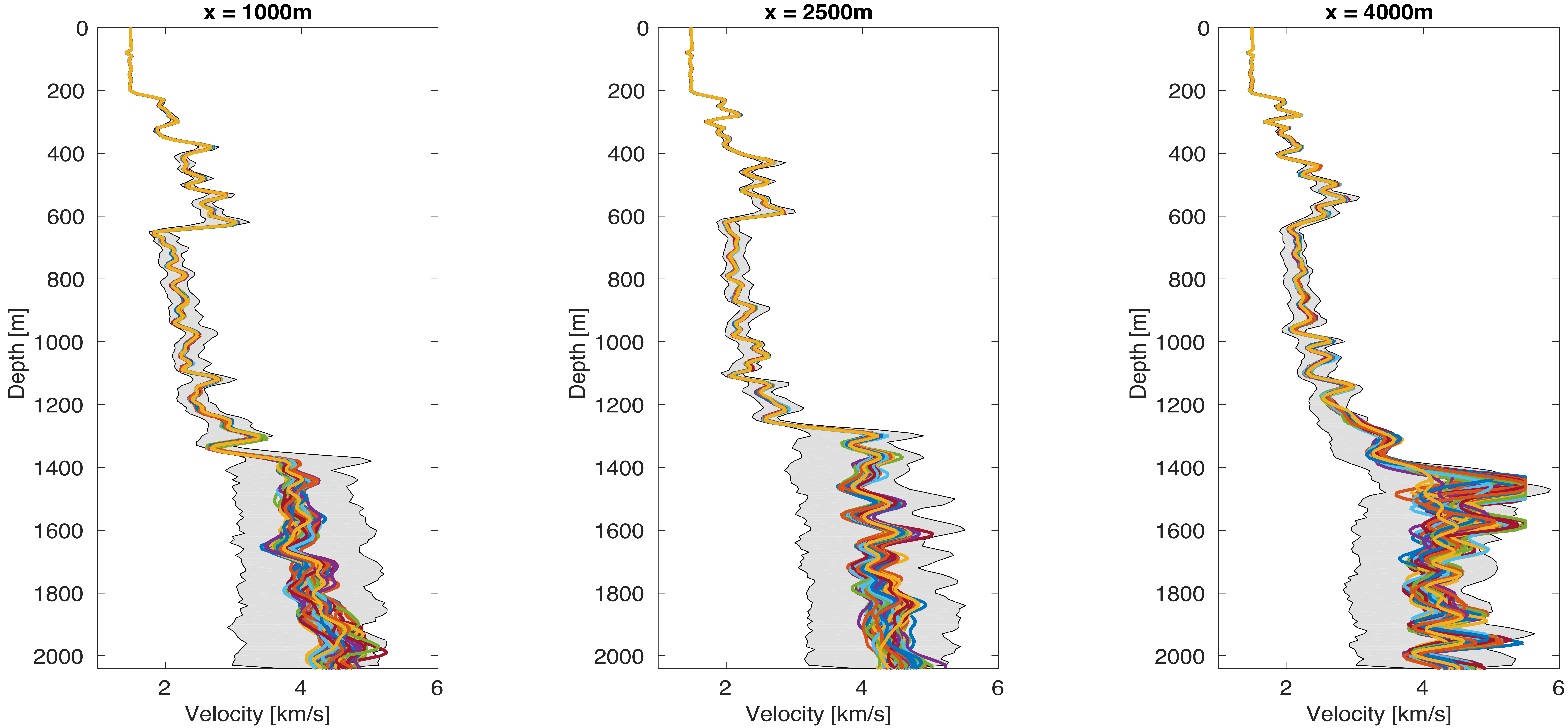
Maximum a posterior

Confidence intervals (90%)



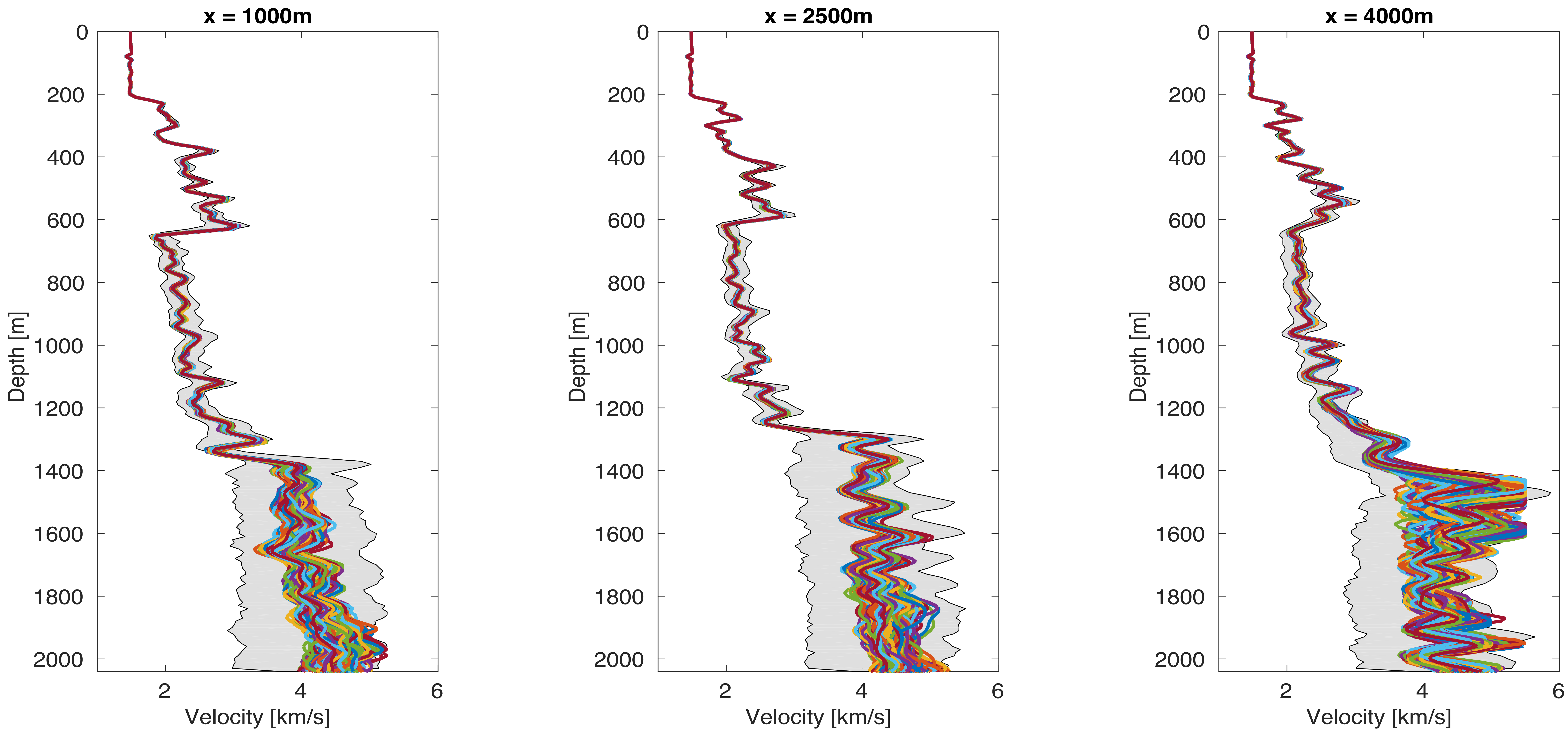
(a) (b) (c)

Confidence intervals (90%)
— Inversion results of 100 random realizations



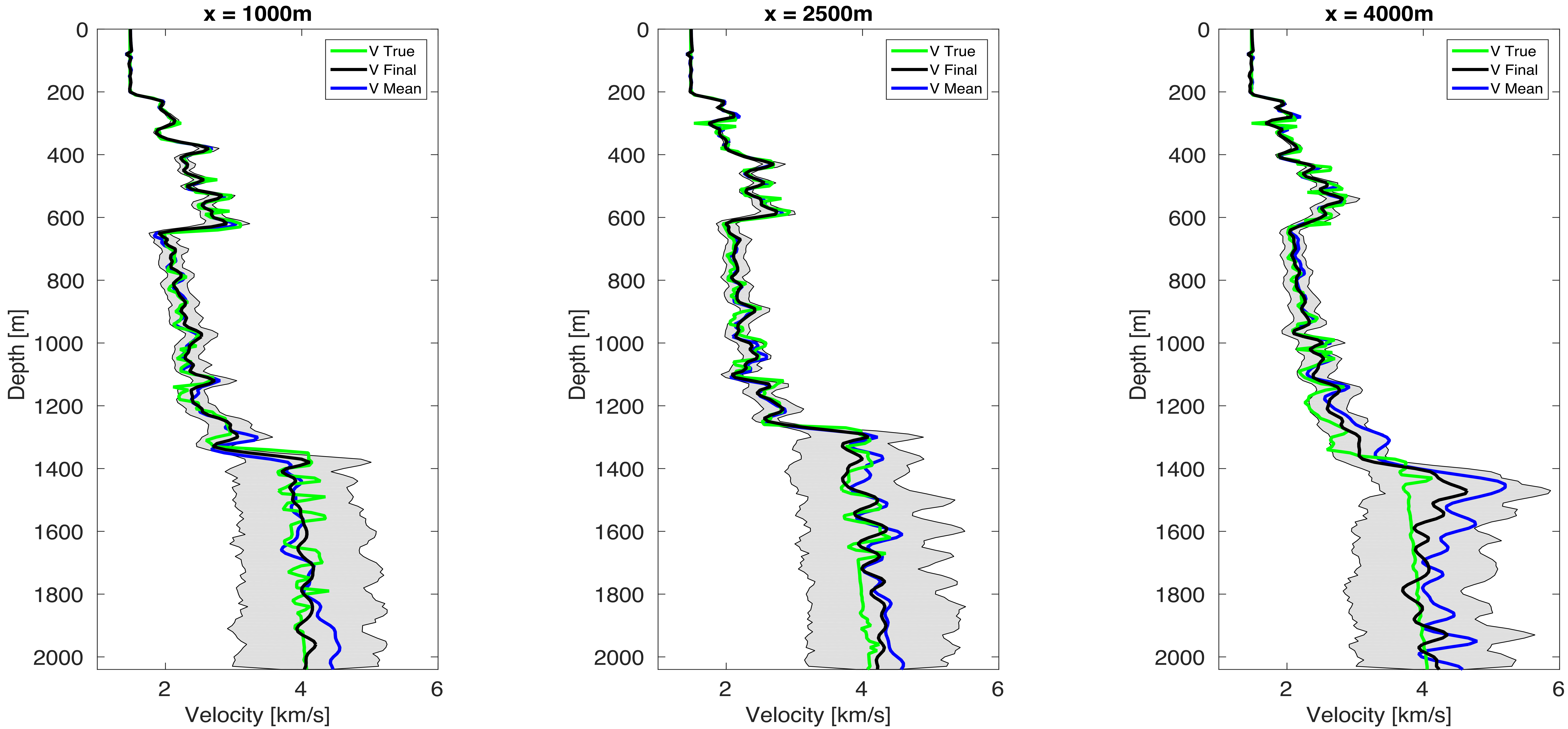
(a) (b) (c)

Confidence intervals (90%)
— Inversion results of 1000 random realizations



(a) (b) (c)

Confidence intervals (90%)
— Inversion results of 1000 random realizations



(a) (b) (c)

Conclusion

1. A positive-definite approximated Hessian of WRI is proposed
    - Matrix free
    - Easy to obtain the diagonal part

2. The quadratic function with the approximated Hessian is a good approximate of the original WRI misfit function

3. The results of inverting noisy data still lie in the confidence intervals
    - This gives us confidence in our confidence intervals

Future work

1. Apply the positive-definite approximated Hessian to the deterministic WRI.

2. Consider the density into the workflow and quantify the uncertainty of density.

3. Add the uncertainty quantification workflow into the soft-release of SINBAD.

Acknowledgements

Thank you for your attention !!!

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada via the Collaborative Research and Development Grant DNOISEII (375142--08). This research was carried out as part of the SINBAD II project which is supported by the following organizations: BG Group, BGP, CGG, Chevron, ConocoPhillips, DownUnder GeoSolutions, Hess, Petrobras, PGS, Schlumberger, Sub Salt Solutions and Woodside