Accelerating LSM and FWI with approximate Born Inversion

Jie Hou

TRIP 2015 Review Meeting

April 25, 2016

Approximate Inverse Operator

$$\bar{F}^{\dagger} \simeq W_{model}^{-1} \bar{F}^{T} W_{data}$$

(ten Kroode, 2012; Hou and Symes,2015)

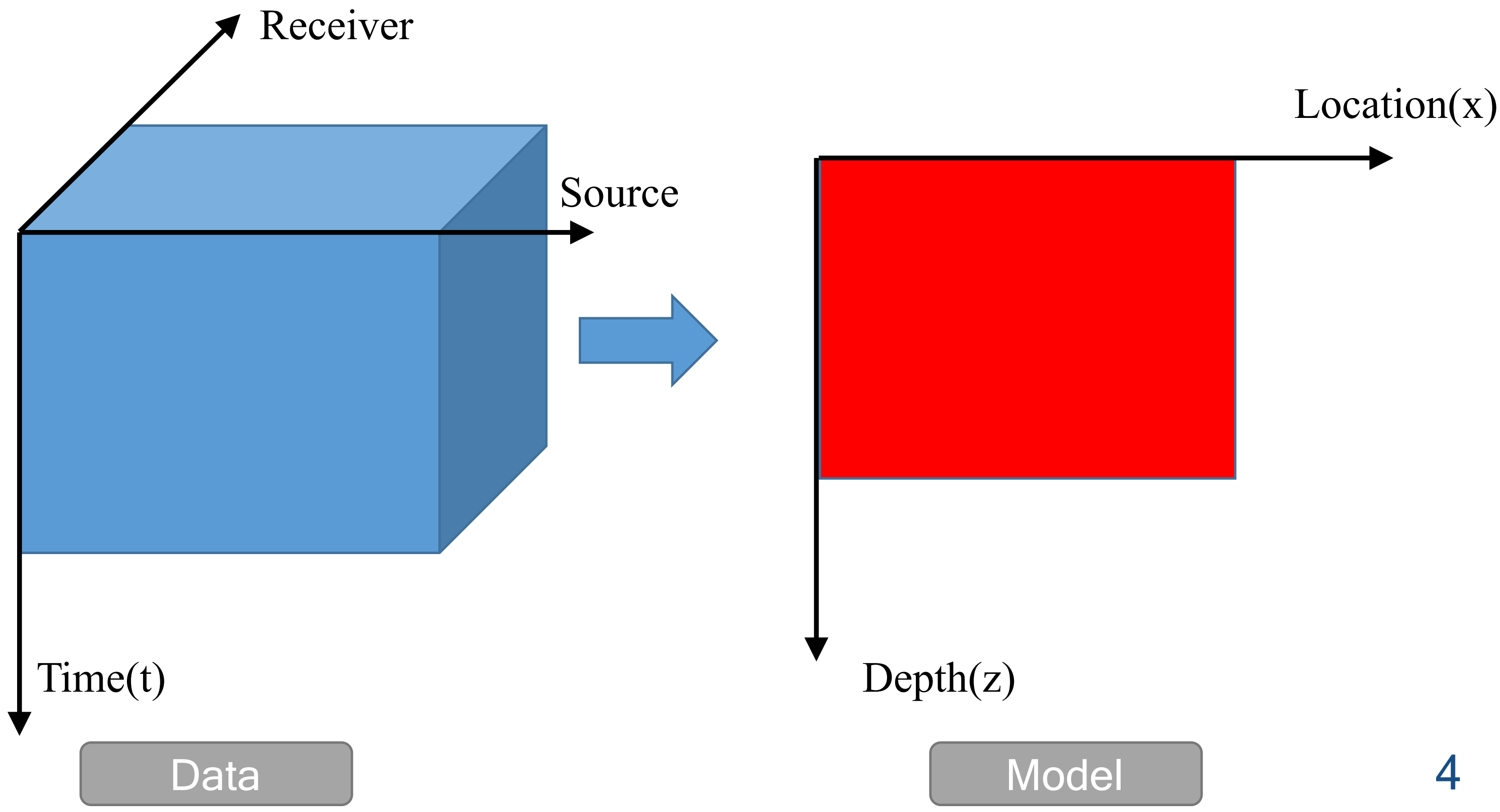
- $W_{model}^{-1} = 4v_0^5 |k_{xz}||k_{hz}| \quad W_{data} = I_t^4 D_{z_S} D_{z_R}$
- Derivation is based on High Frequency Approx.
- Implementation doesn't involve any ray tracing
- Invert the data even when velocity is wrong

Remarks

- Normal Operator $\bar{F}^T \bar{F}$ is order 0
  - Not change the frequency components
  - Weight Operators don't change the order
- Weight Operators add no appreciable cost
  - W_{model} involves a 3D Fourier transform
- Recover the physical model by stacking
- Subsurface offset extension is the key

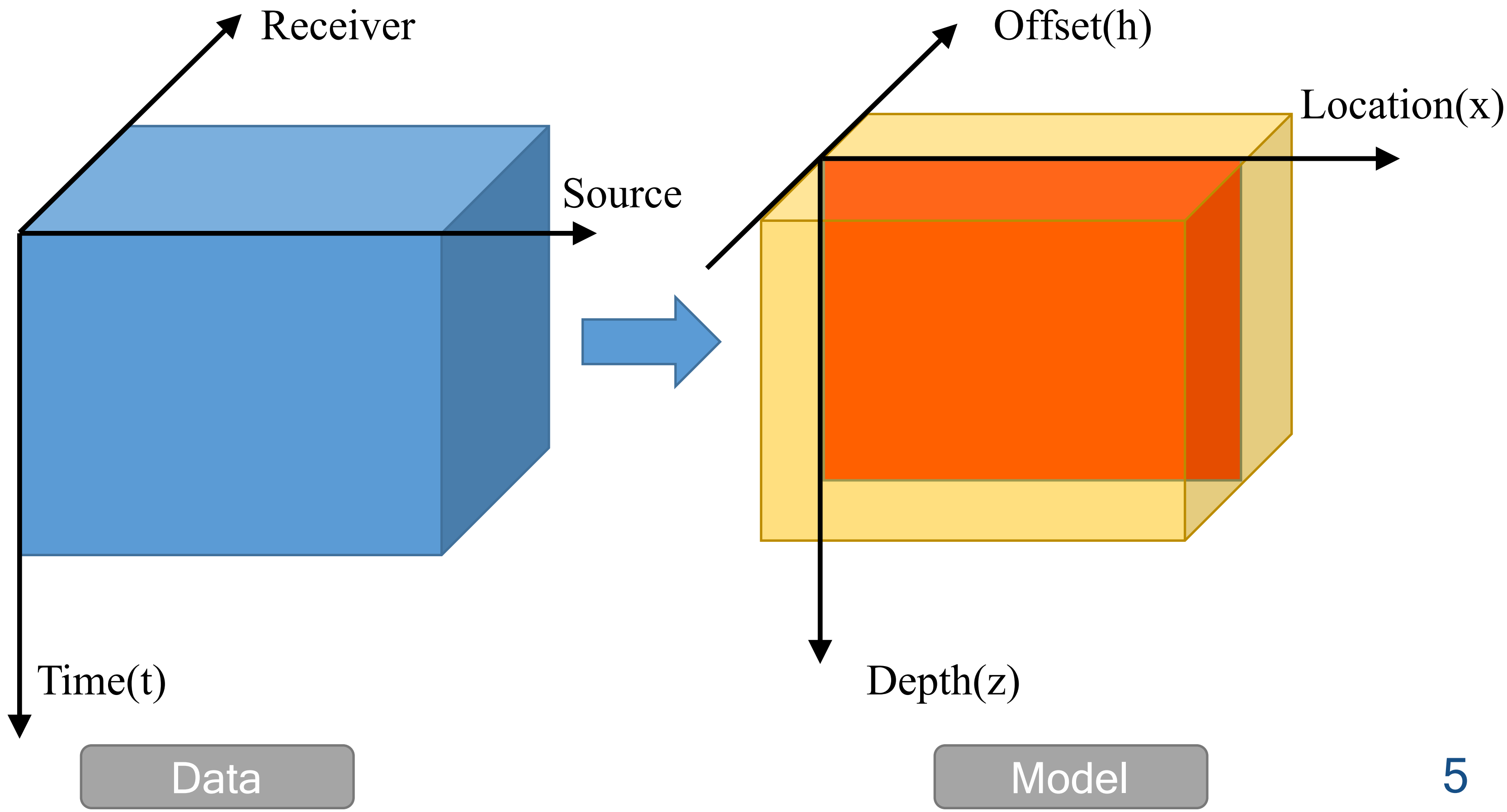
Is subsurface offset necessary?

- It is important when velocity is wrong.



Is subsurface offset necessary?

- It is important when velocity is wrong.



Is subsurface offset necessary?

- What if velocity is kinematically correct?
  - ✓ Subsurface offset is no more necessary
  - ? Will the approx. inverse operator still work

$$\bar{F}^{\dagger} \simeq W_{model}^{-1} \bar{F}^{T} W_{data}$$

Is subsurface offset necessary?

- What if velocity is kinematically correct?
  - ✓ Subsurface offset is no more necessary
  - ? Will the approx. inverse operator still work

$$\bar{F}^{\dagger} \simeq W_{model}^{-1} \bar{F}^{T} W_{data}$$

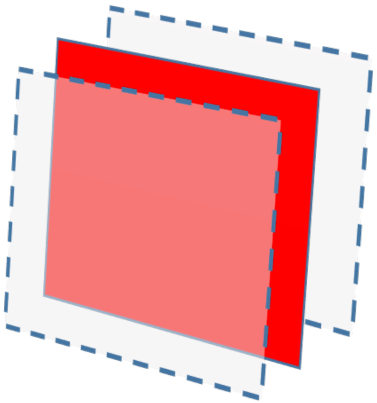
$$W_{model}^{-1} = 4 v_0^5 |k_{xz}||k_{hz}|$$

Is subsurface offset necessary?

- What if velocity is kinematically correct?
  - ✓ Subsurface offset is no more necessary
  - ? Will the approx. inverse operator still work

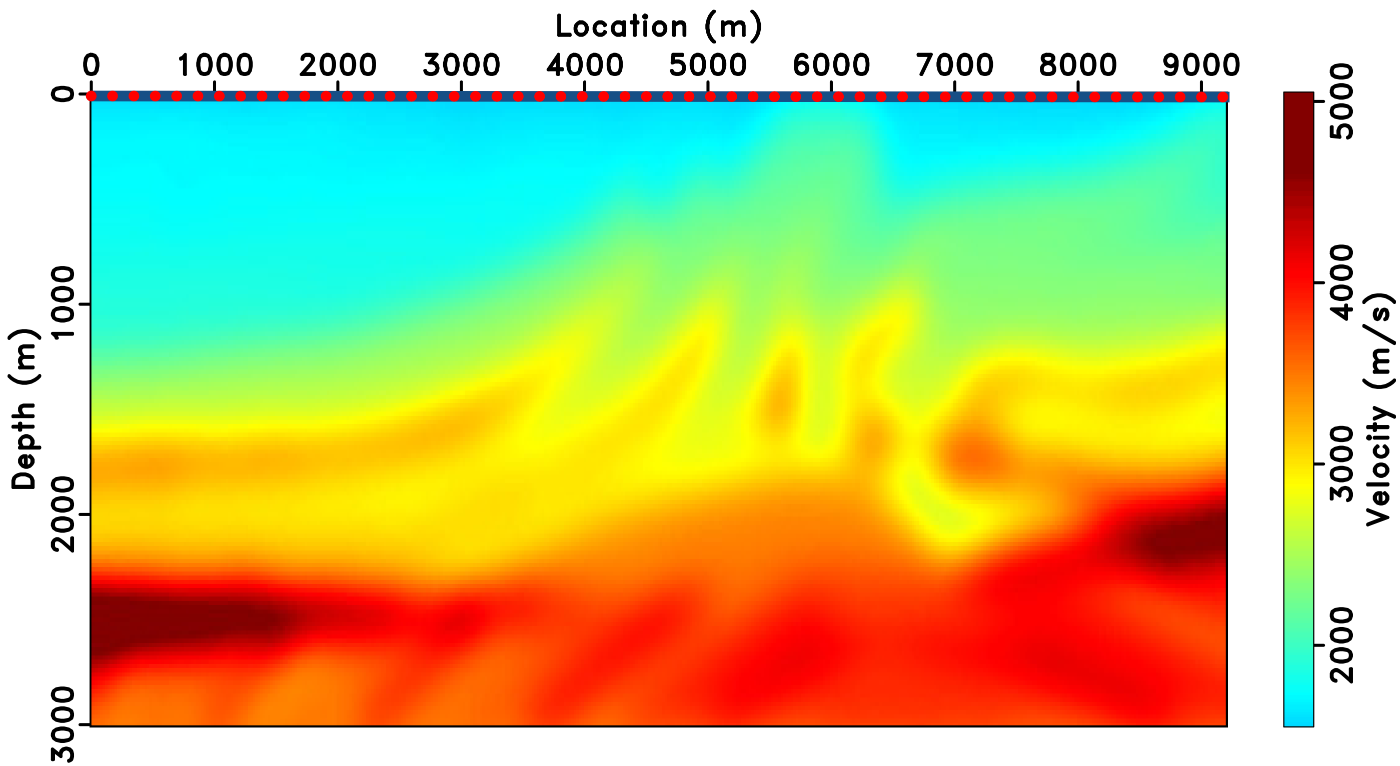
$$F^{\dagger} \simeq W_{model}^{-1} F^{T} W_{data}$$

$$W_{model}^{-1} = 4v_0^5 |k_{xz}||k_{hz}|$$



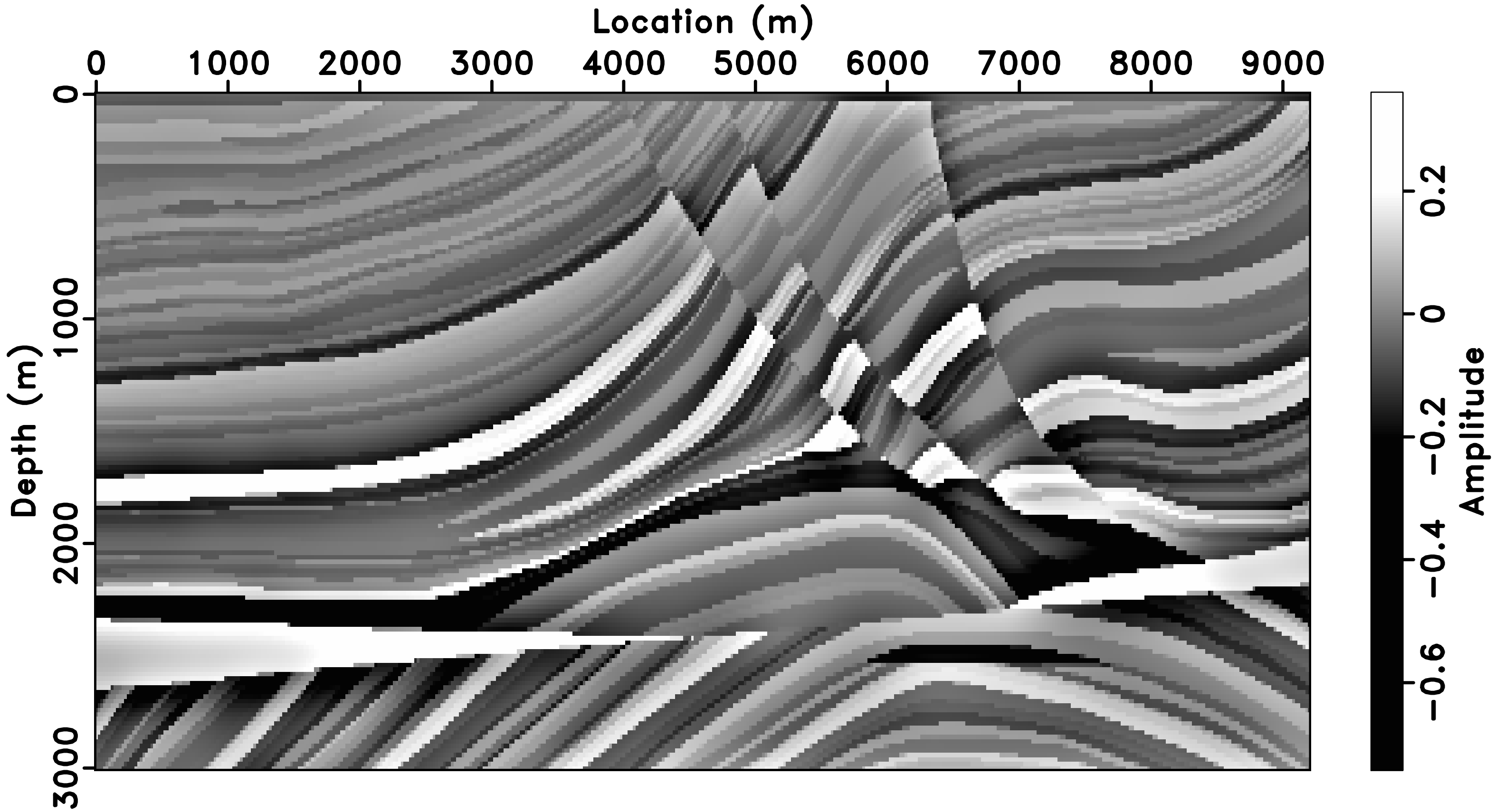
Born inversion

- 2-8 finite difference, 231 shots & 461 receivers
- 2.5-5-20-25Hz Bandpass wavelet
- 2ms time sample, 20m grid interval



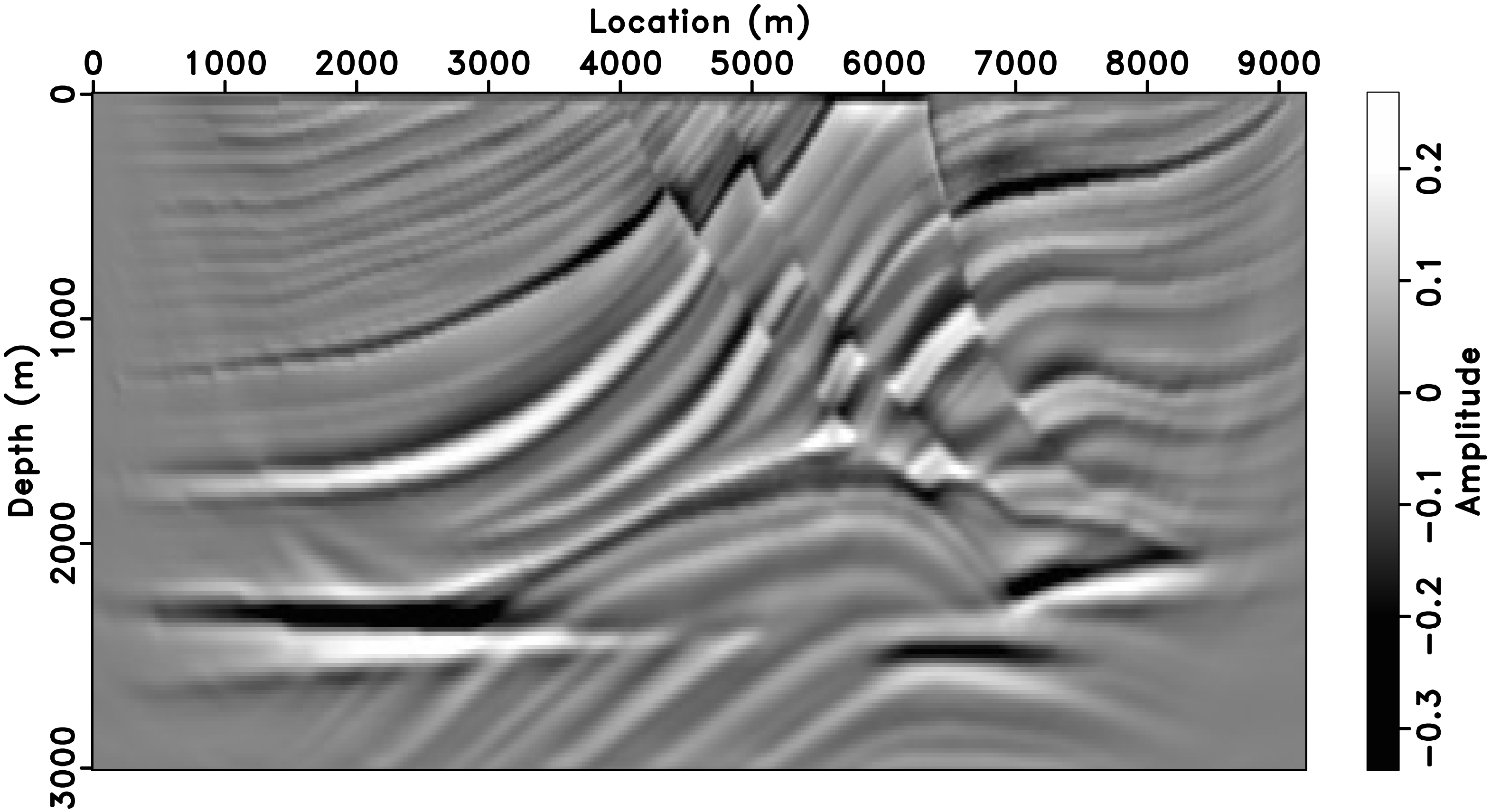
Smooth Background Model

Born inversion



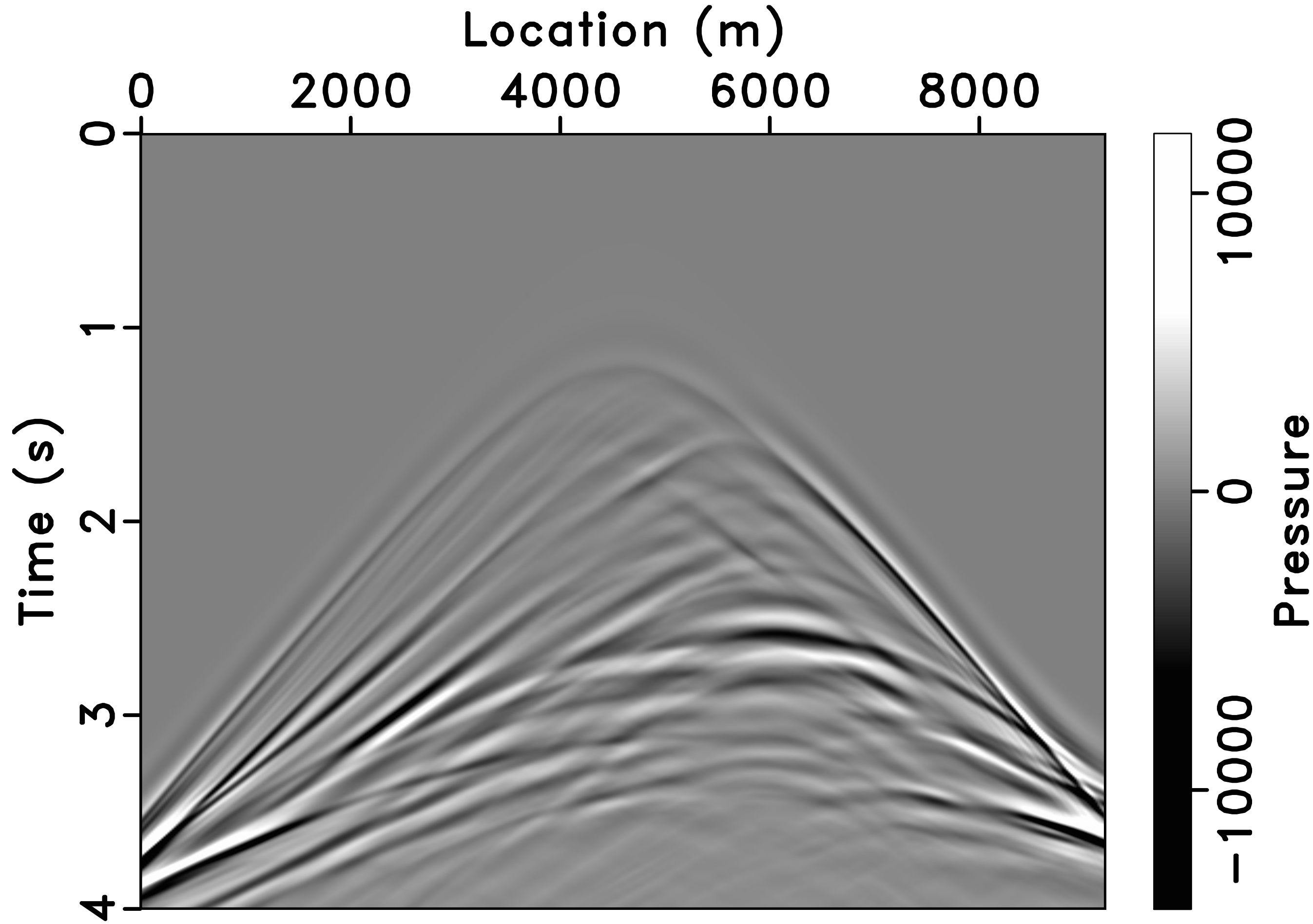
Reflectivity Model

Born inversion



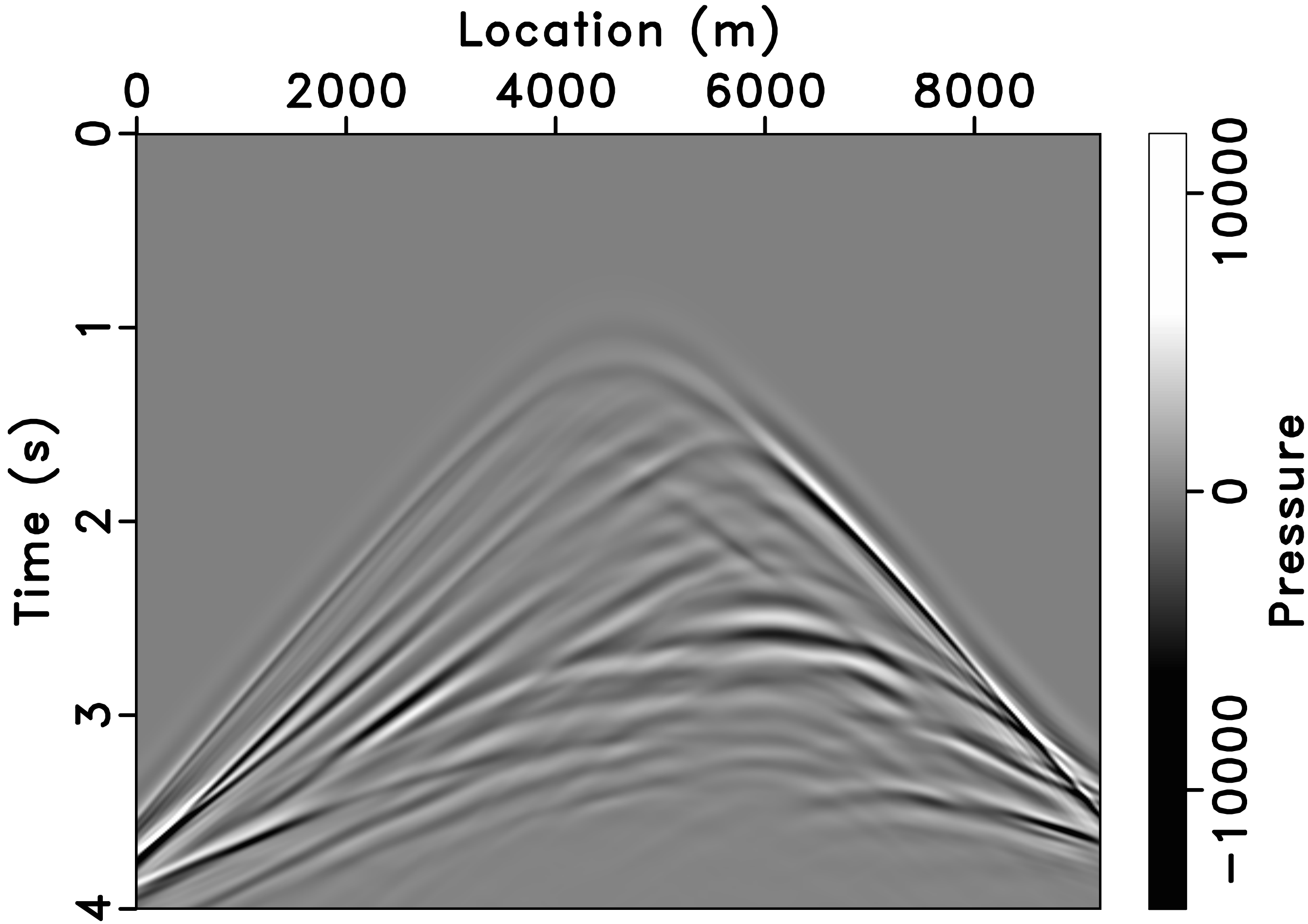
Recovered Model

Born inversion



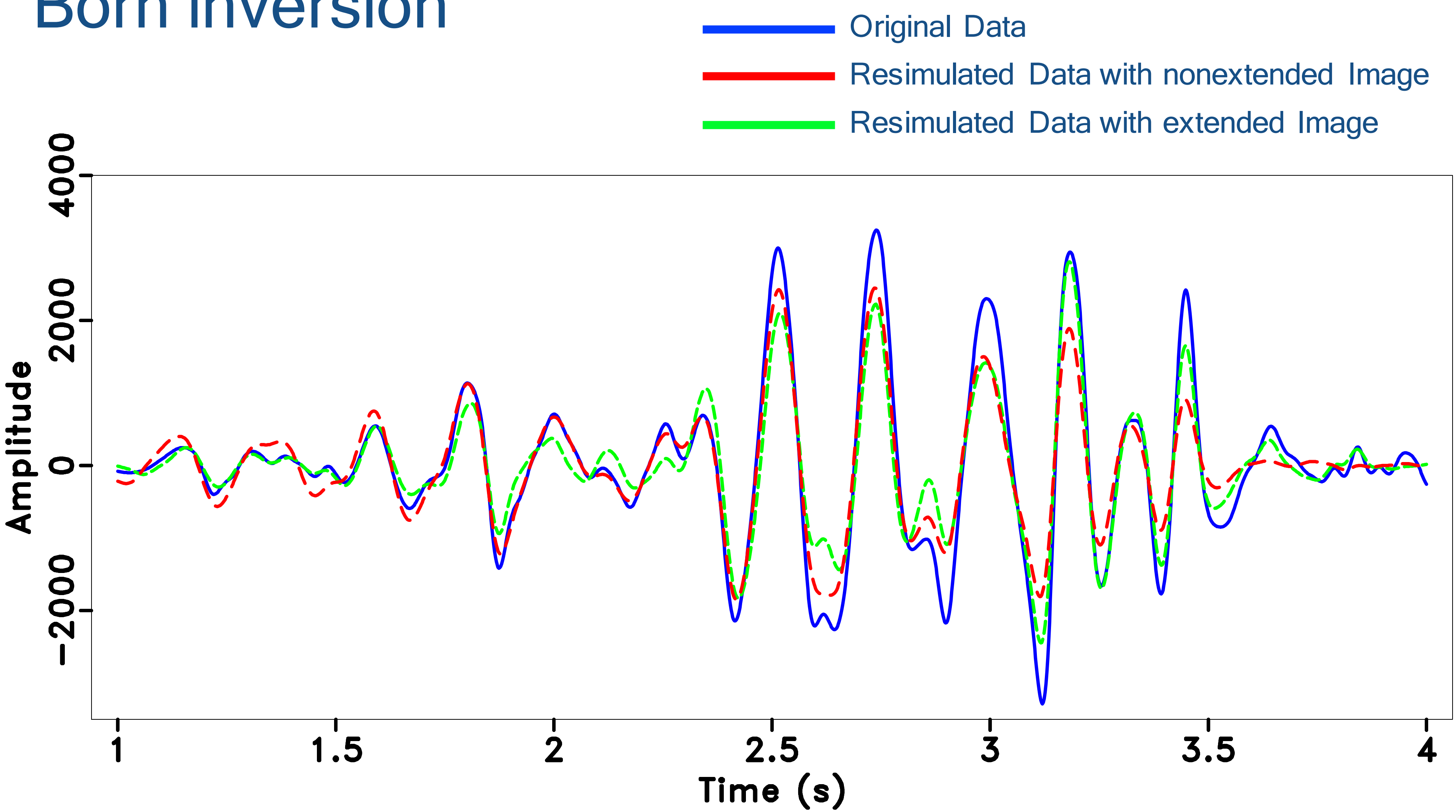
Original Data

Born inversion



Resimulated Data

Born inversion



Single Trace Comparison

LSM – Theory

Least Squares Migration (LSM) seeks reflectivity model to minimize :

$$J_{LS} = \frac{1}{2}||Fm - d||^2$$

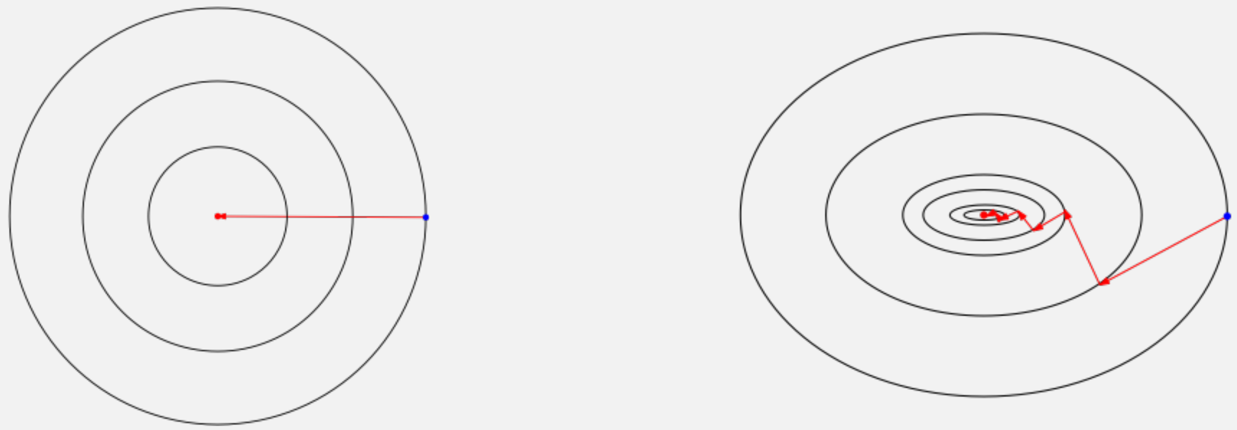
It is equivalent to solve

$$F^T F m = F^T d$$

Goal : Accelerate the convergence of LSM

LSM – Theory

Convergence of an optimization problem



- Make the normal operator close to Identity

$$F^T F m = F^T d$$

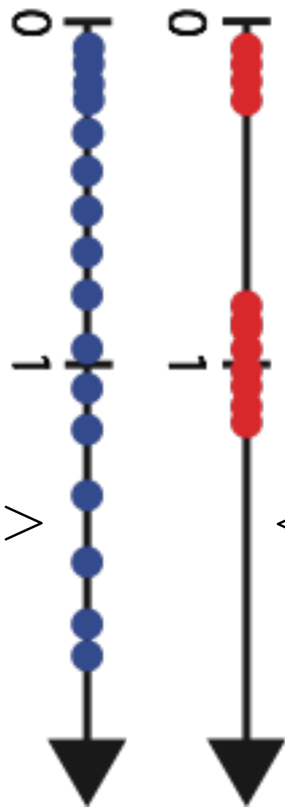
LSM – CG vs. WCG

$$F^T F m = F^T d \iff F^\dagger F m = F^\dagger d$$

$$< m, m >= m^T m$$

$$< d, d >= d^T d$$

$$< m, F^T d >=< d, Fm >$$



Eigenvalue

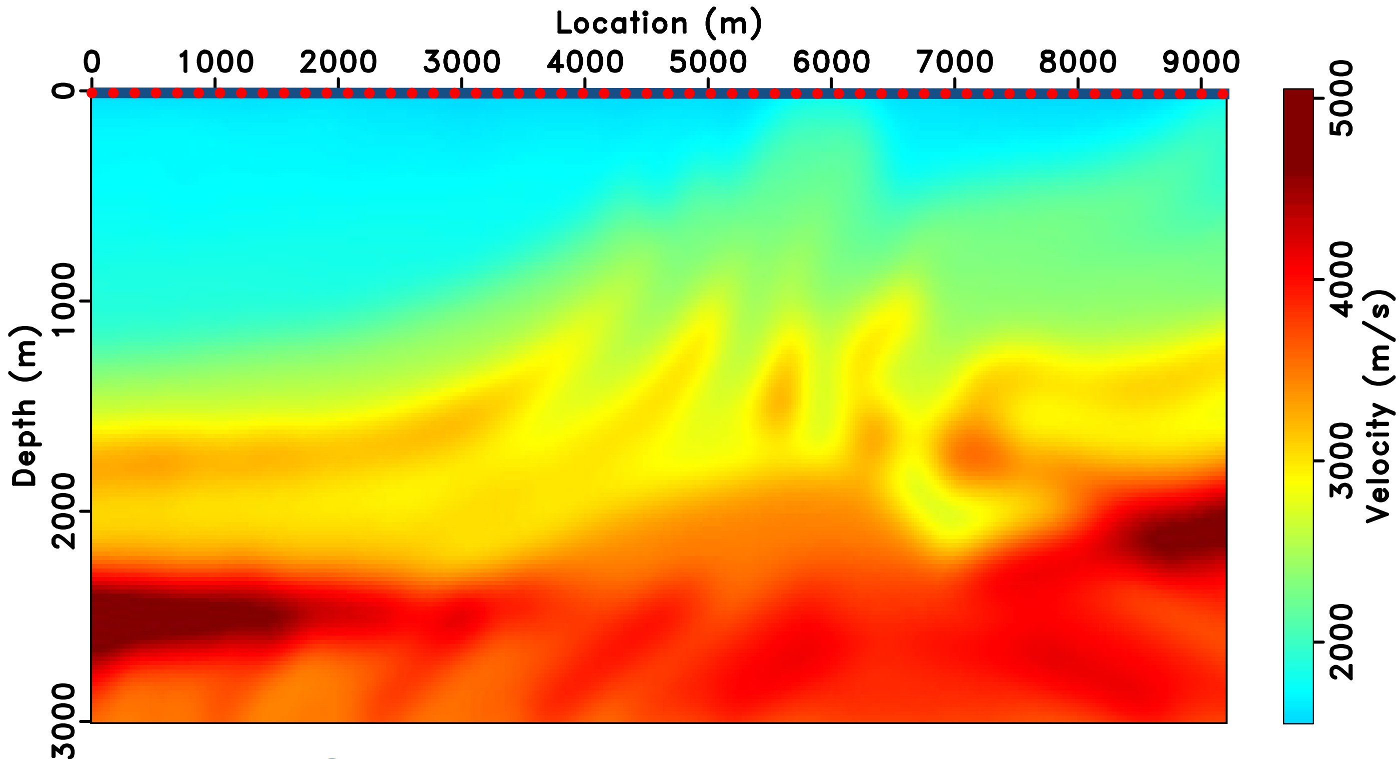
$$< m, m >_{model}= m^T W_{model} m$$

$$< d, d >_{data}= d^T W_{data} d$$

$$< m, F^\dagger d >_{model}=< d, Fm >_{data}$$

LSM – Numerical Example I

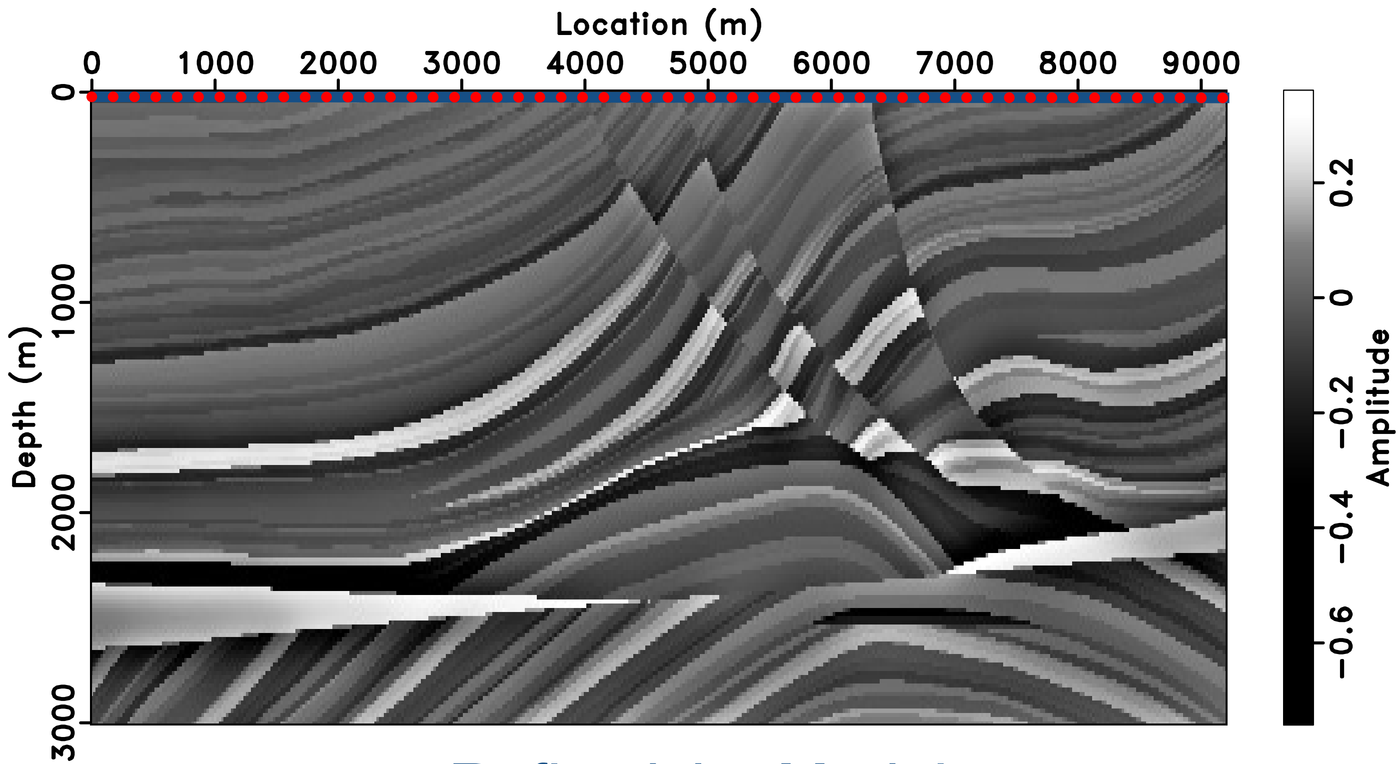
- 2-8 finite difference, 231 shots & 461 receivers
- 2.5-5-20-25Hz Bandpass wavelet
- 2ms time sample, 20m grid interval



Smooth Background Model

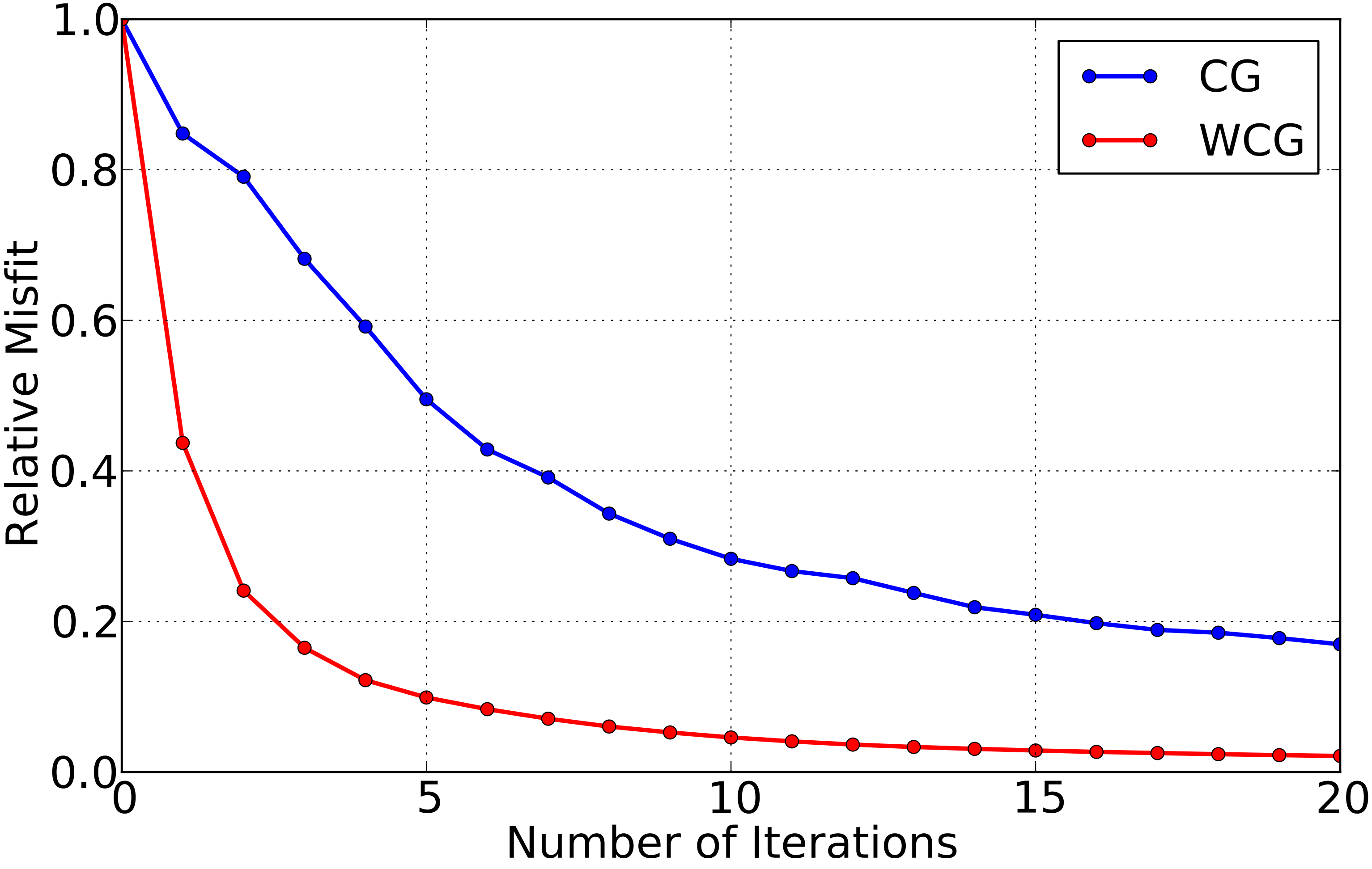
LSM – Numerical Example I

- 2-8 finite difference, 231 shots & 461 receivers
- 2.5-5-20-25Hz Bandpass wavelet
- 2ms time sample, 20m grid interval

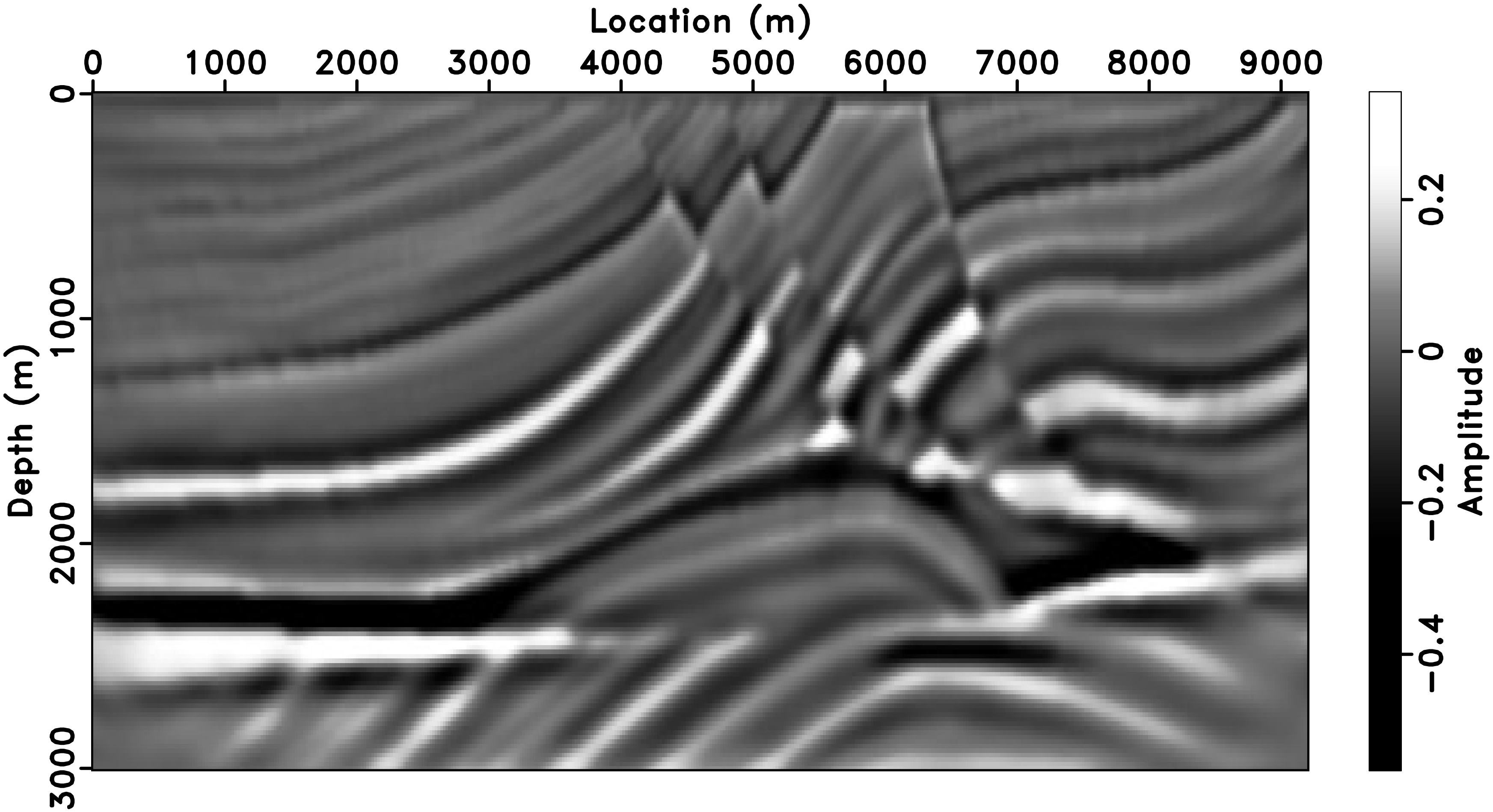


Reflectivity Model

LSM – Convergence Curve

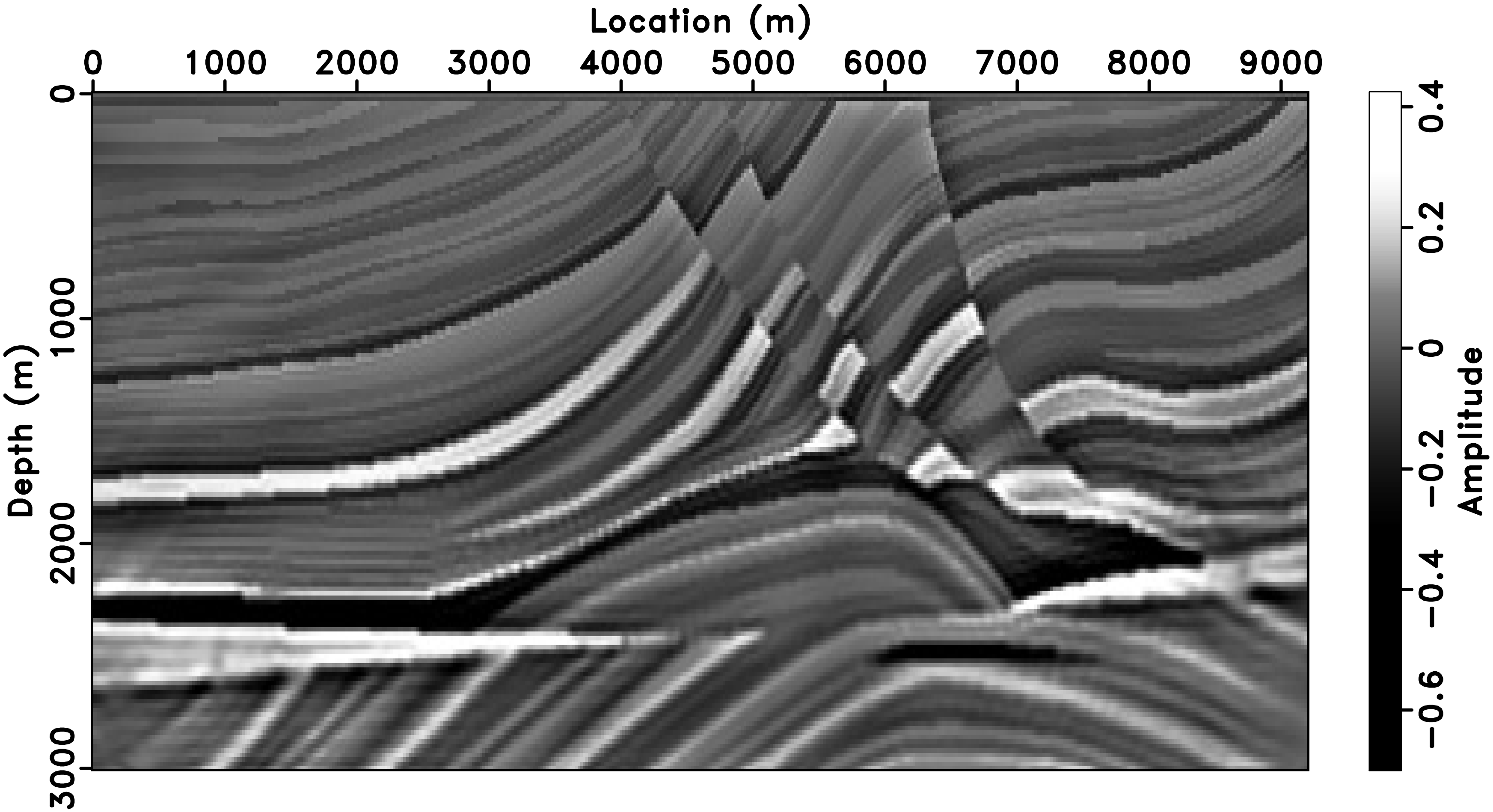


LSM – Numerical Example I



20 CG iteration result

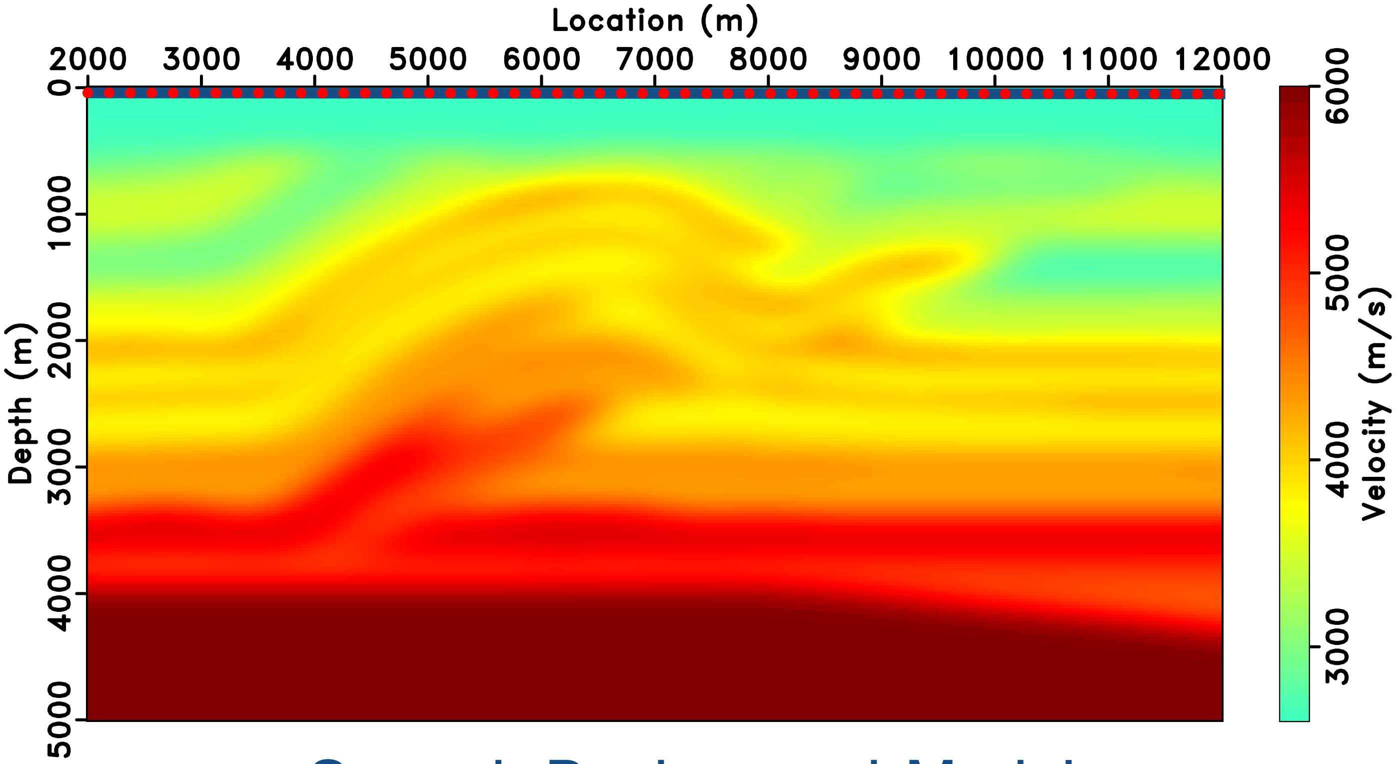
LSM – Numerical Example I



20 WCG iteration result

LSM – Numerical Example II

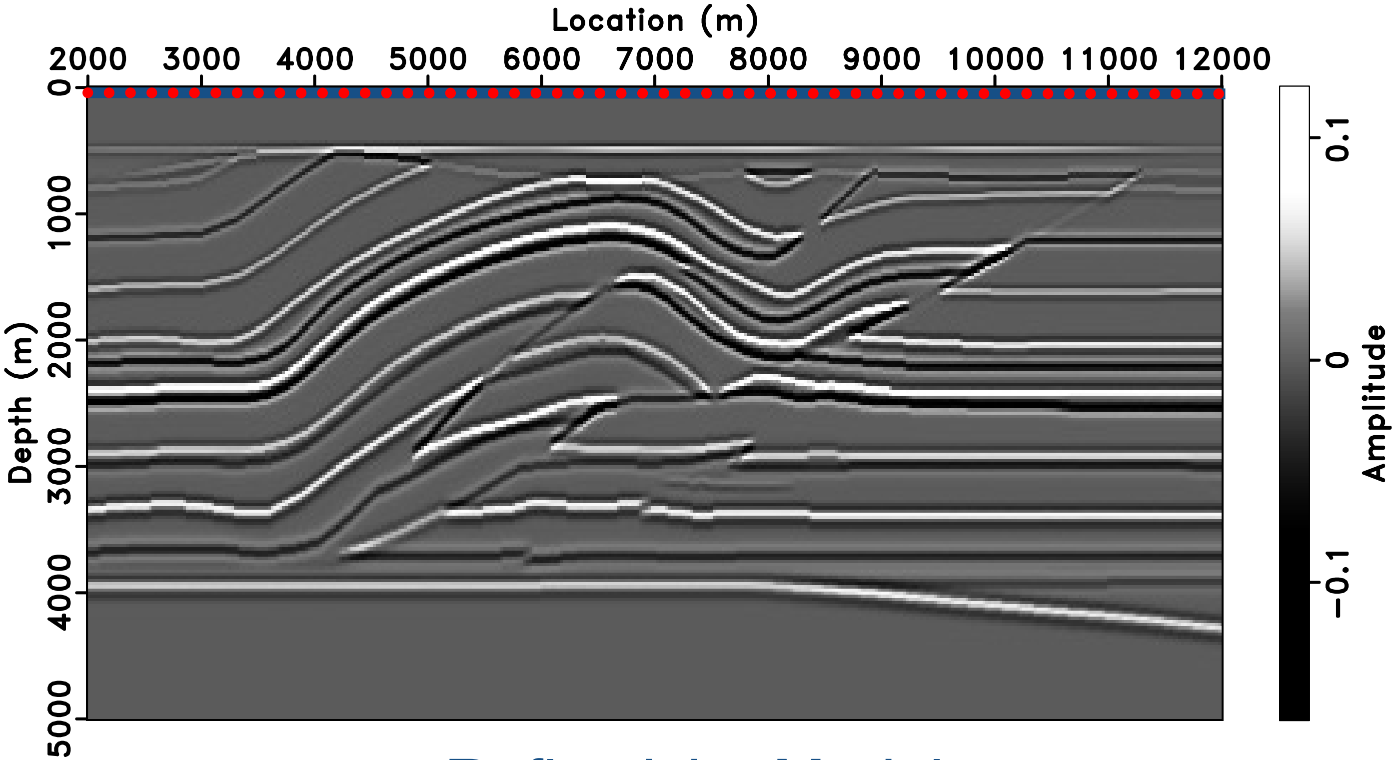
- 2-8 finite difference, 201 shots & 401 receivers
- 2.5-5-20-25Hz Bandpass wavelet
- 2ms time sample, 25m grid interval



Smooth Background Model

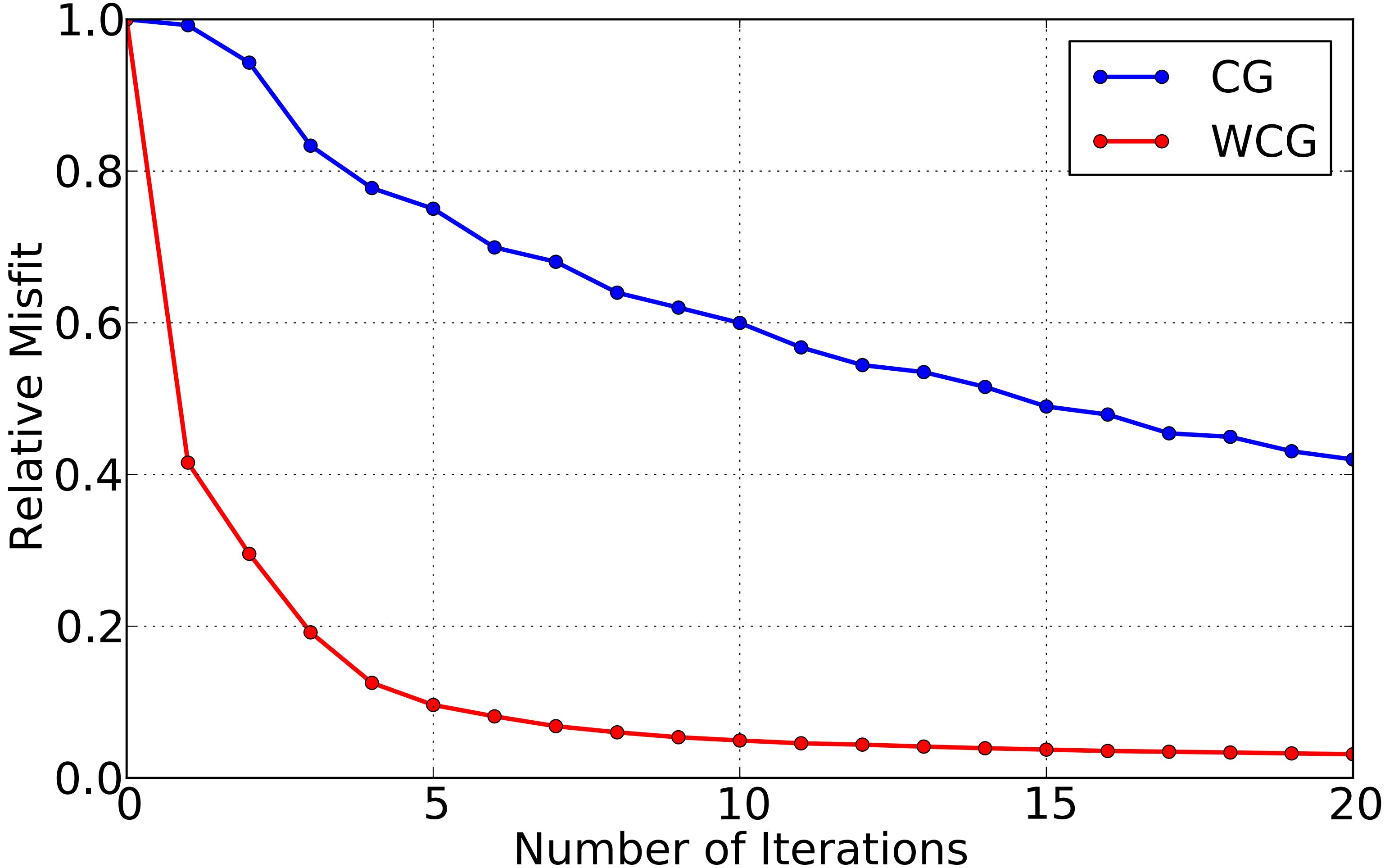
LSM – Numerical Example II

- 2-8 finite difference, 201 shots & 401 receivers
- 2.5-5-20-25Hz Bandpass wavelet
- 2ms time sample, 25m grid interval

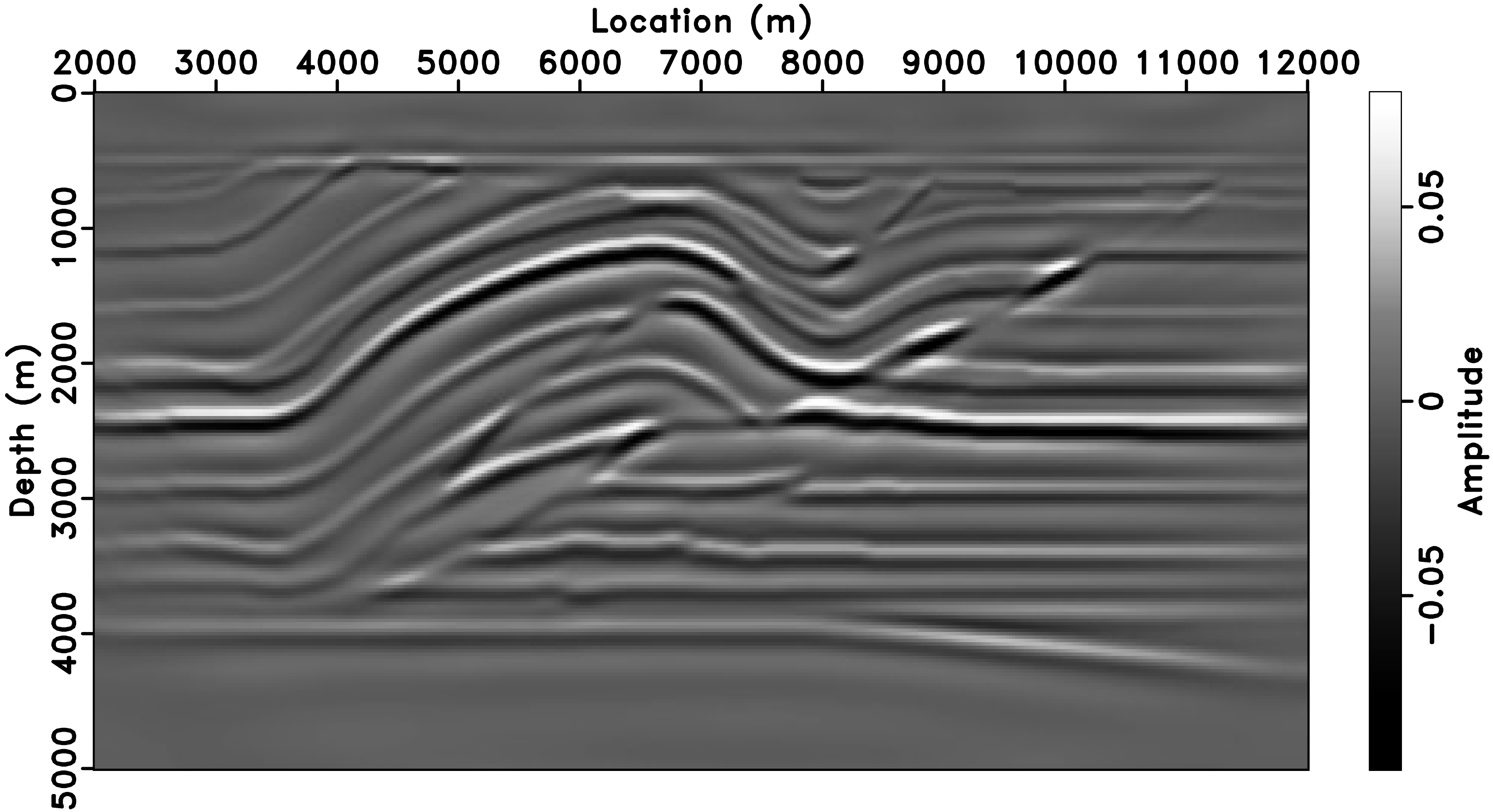


Reflectivity Model

LSM – Convergence Curve

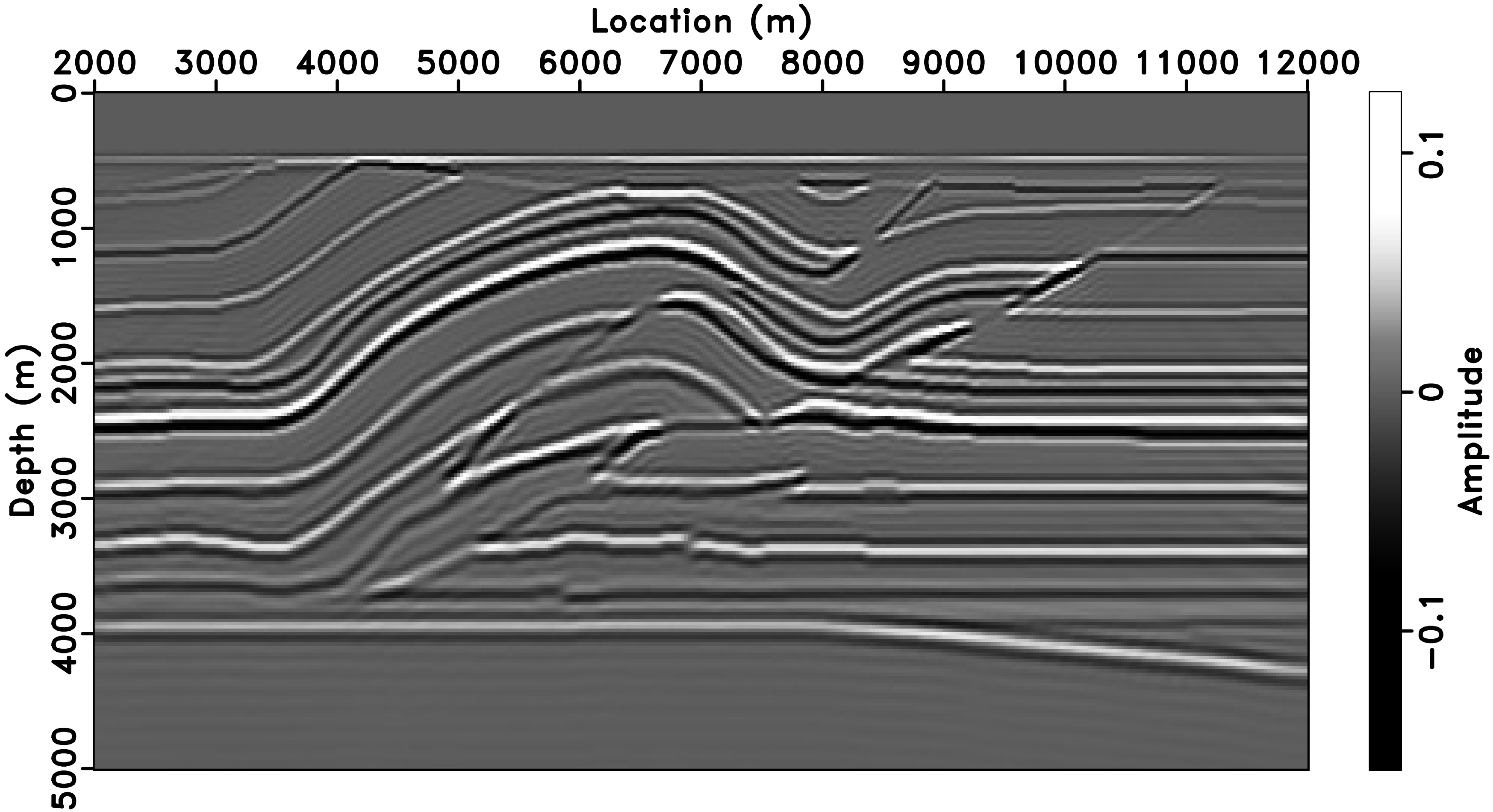


LSM – Numerical Example II



20 CG iteration result

LSM – Numerical Example II



20 WCG iteration result

FWI – Theory

Full Waveform Inversion (FWI) seeks velocity model to minimize :

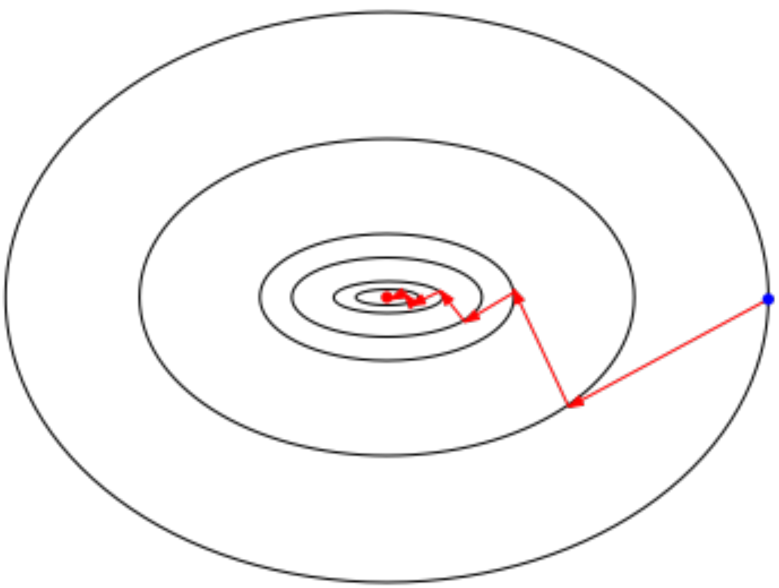
$$J_{LS} = \frac{1}{2}||\mathcal{F}[m] - d||^2$$

- Nonlinear → Local Minimal (Cycle Skipping)
- Large Scale → Local Optimization Method
- Ill-posed → Slow Convergence Rate

FWI – Optimization

Iterative Method :

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha \mathbf{p}_k$$



Steepest Descent Method :

- Negative Gradient direction

$$\mathbf{p}_k = -F^T(\mathcal{F}[\mathbf{m}_k] - d)$$

- Easy, cheap and works

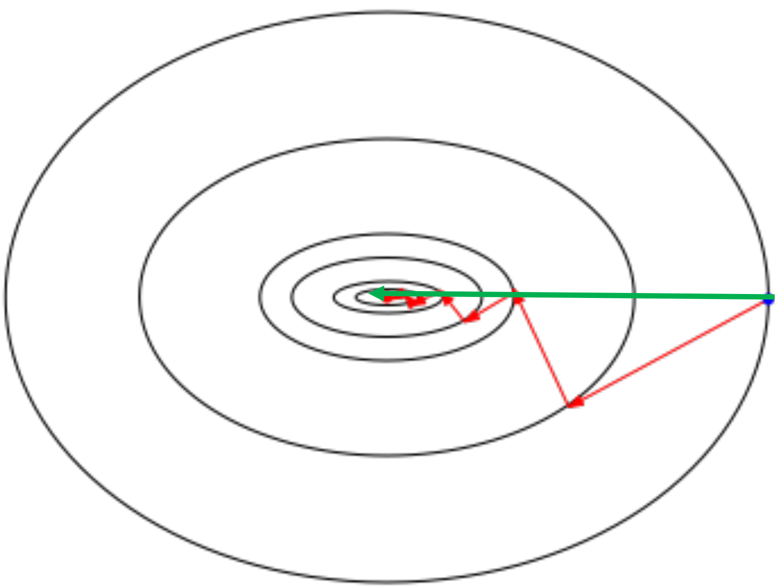
But

- Slow convergence
- Gradient changes rapidly
- Only uses 1st order approximation

FWI – Optimization

Iterative Method :

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha \mathbf{p}_k$$



Converge in one step if objective function is convex quadratic

Newton's Method (L. Métivier et al.,2014)

- Newton Update direction

$$\mathbf{p}_k = -\mathbf{H}_k^{-1} \nabla_{\mathbf{m}} J$$

- Use curvature information
- Fast convergence

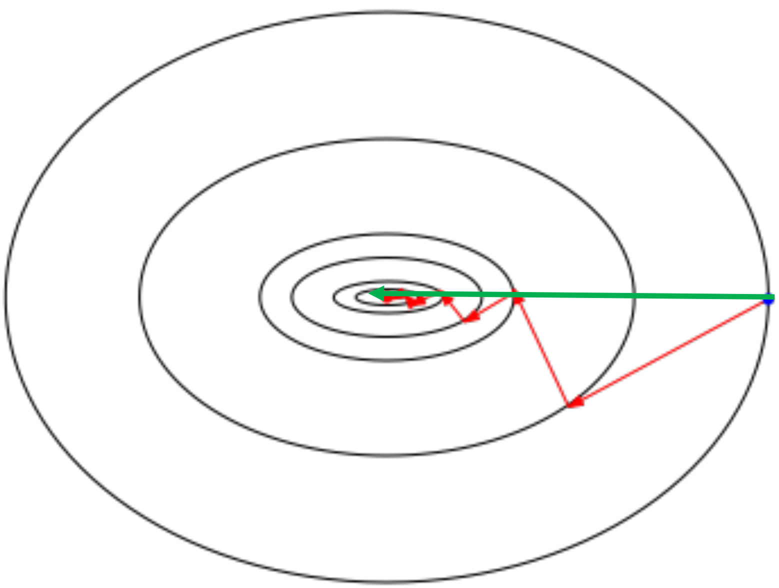
But

- Hessian hard to invert
- Expensive

FWI – Optimization

Iterative Method :

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha \mathbf{p}_k$$



Converge in one step if objective function is convex quadratic

Newton's Method (L. Métivier et al.,2014)

- Newton Update direction

$$F^T F \mathbf{p}_k + D^2 \mathcal{F}^T(\mathbf{p}_k, \mathcal{F}[\mathbf{m}_k] - \mathbf{d}) = -\mathbf{g}_k$$

- Use curvature information
- Fast convergence

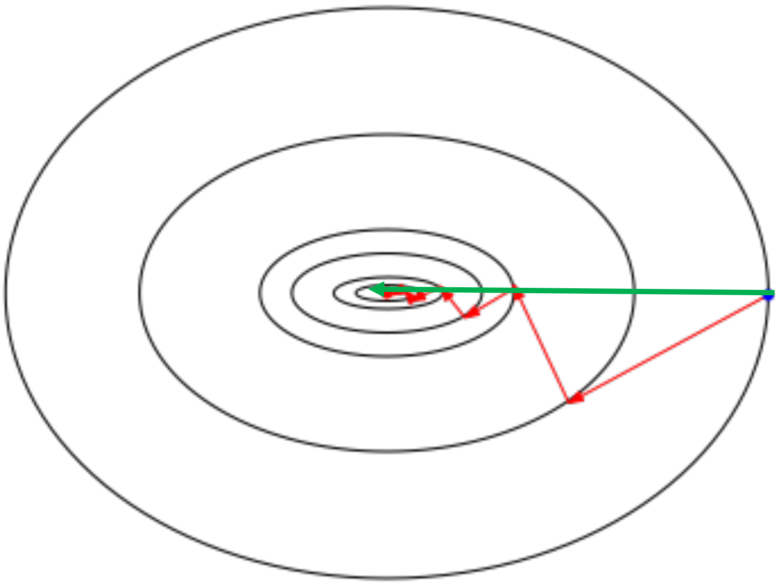
But

- Hessian hard to invert
- Expensive

FWI – Optimization

Iterative Method :

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha \mathbf{p}_k$$



Converge in one step if objective function is convex quadratic

Gauss-Newton Method

(I Epanomeritakis et al.,2009)

- Gauss-Newton Update direction

$$F^T F \mathbf{p}_k = -\mathbf{g}_k$$

- ≈ Newton's method near optimum

But

- Still expensive

FWI – Optimization

Iterative Method :

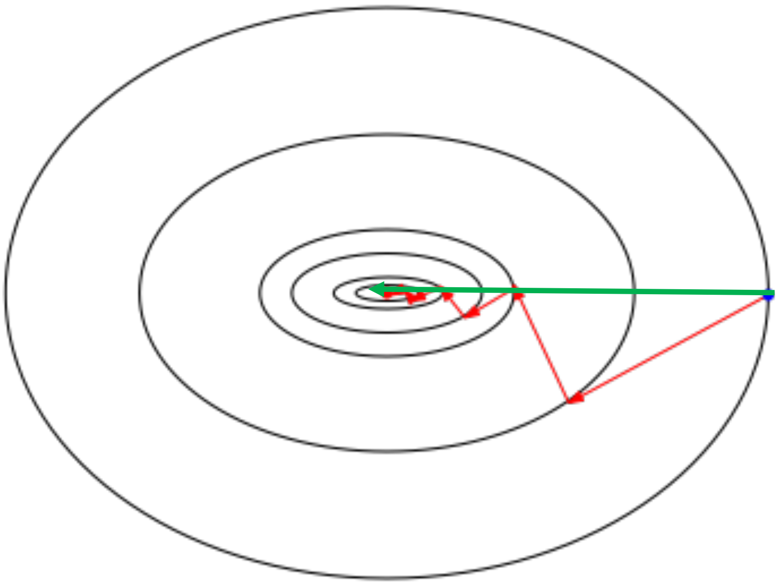
$$\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha \mathbf{p}_k$$

Gauss-Newton Method

(I Epanomeritakis et al.,2009)

- Gauss-Newton Update direction

$$\mathbf{p}_k = -(F^T F)^{-1} F^T (\mathcal{F}[\mathbf{m}_k] - \mathbf{d})$$



Converge in one step if objective function is convex quadratic

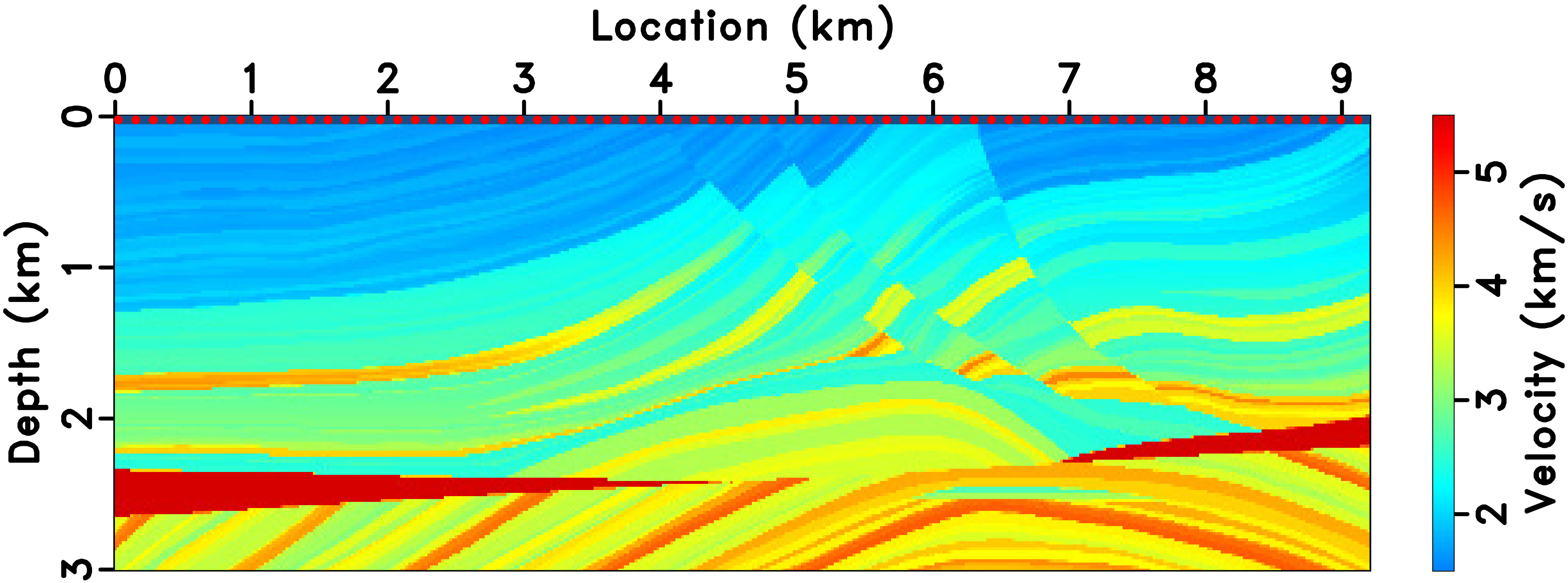
But

- Still expensive
- Approximate with Born inversion

$$\mathbf{p}_k = -F^\dagger (\mathcal{F}[\mathbf{m}_k] - \mathbf{d})$$

FWI – Numerical Example

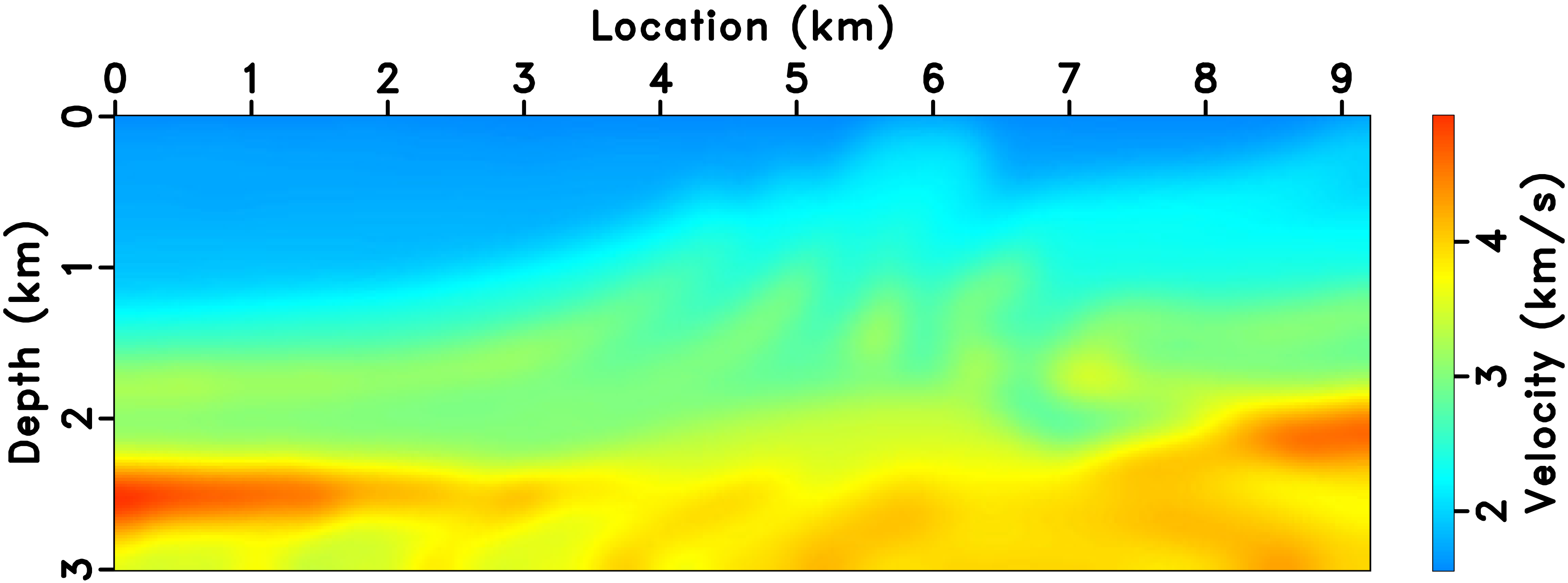
- 2-8 finite difference, 231 shots & 461 receivers
- 2.5-5-20-25Hz Bandpass wavelet
- 2ms time sample, 20m grid interval



True Model

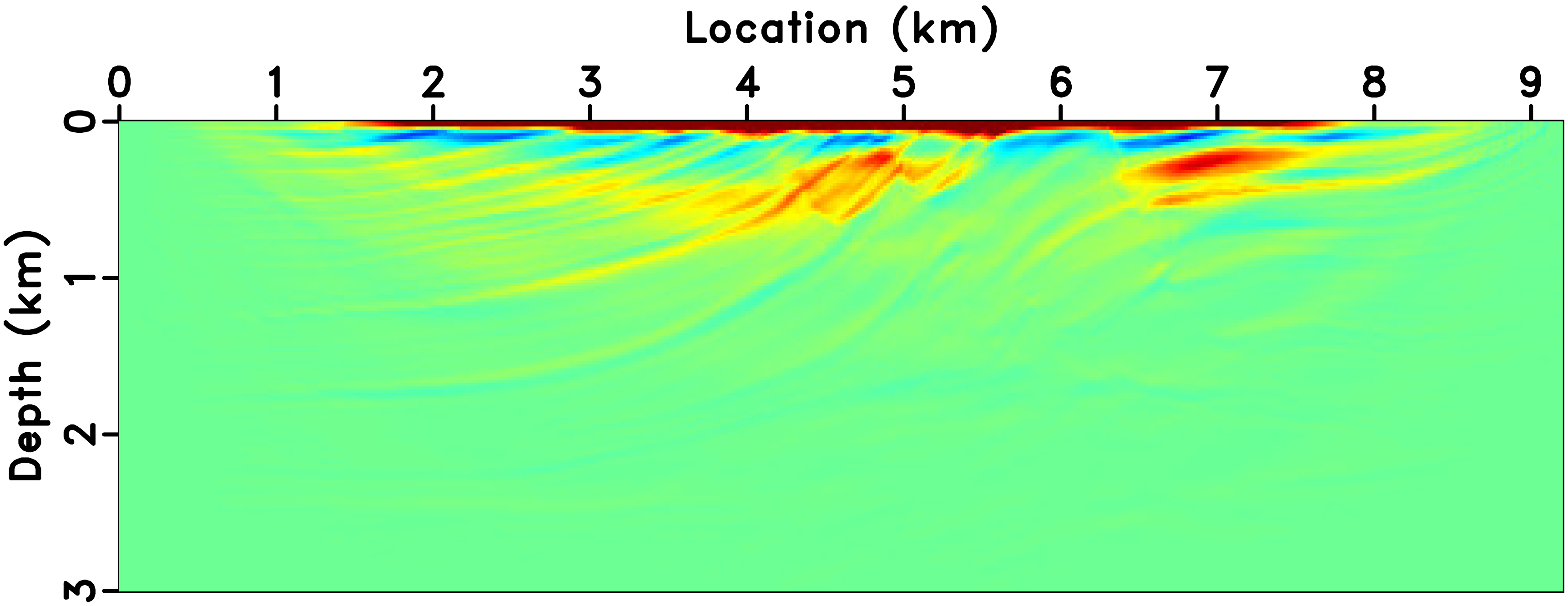
FWI – Numerical Example

- 2-8 finite difference, 231 shots & 461 receivers
- 2.5-5-20-25Hz Bandpass wavelet
- 2ms time sample, 20m grid interval



Initial Model

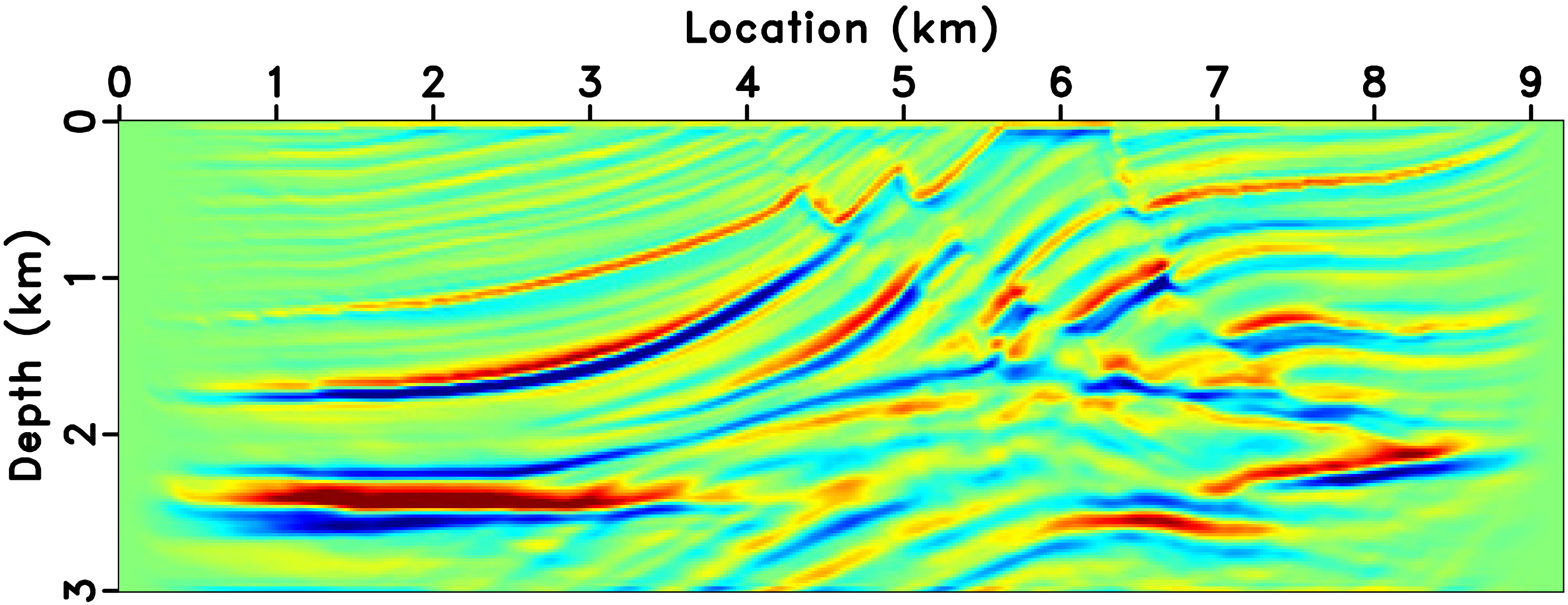
FWI – Numerical Example



Gradient at first step

$$F^T(\mathcal{F}[\mathbf{m}_0] - \mathbf{d})$$

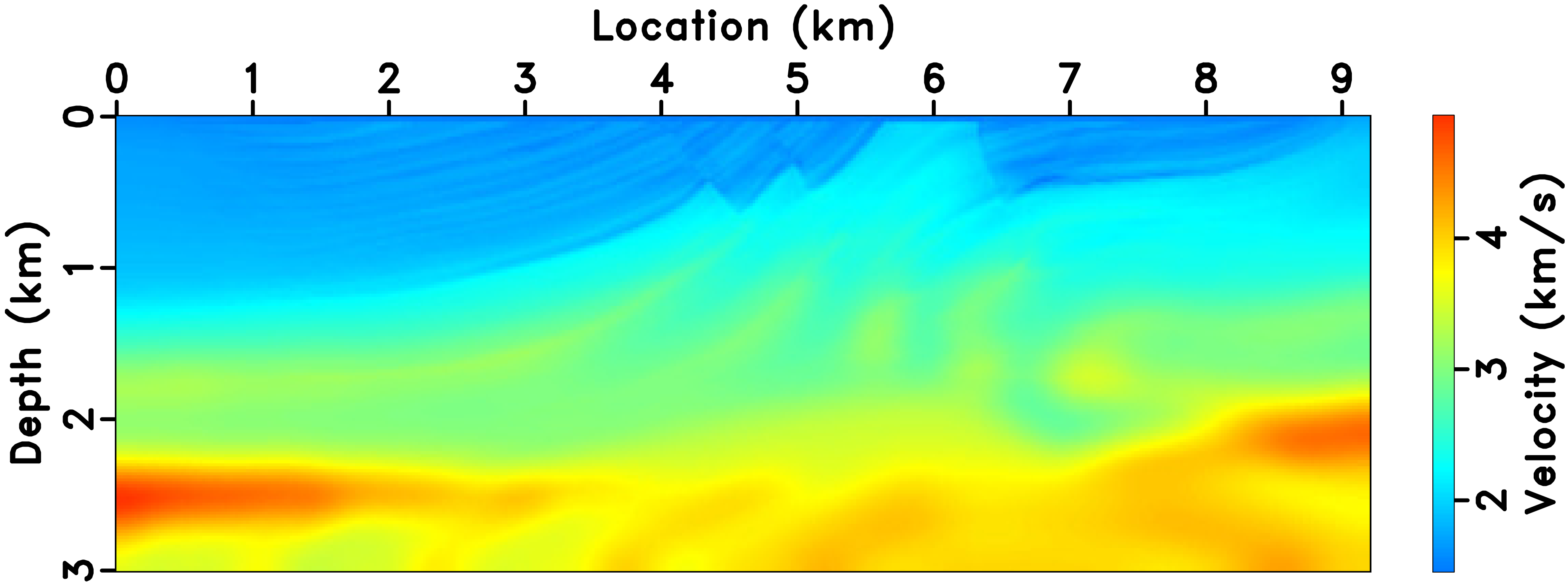
FWI – Numerical Example



Preconditioned Gradient at first step

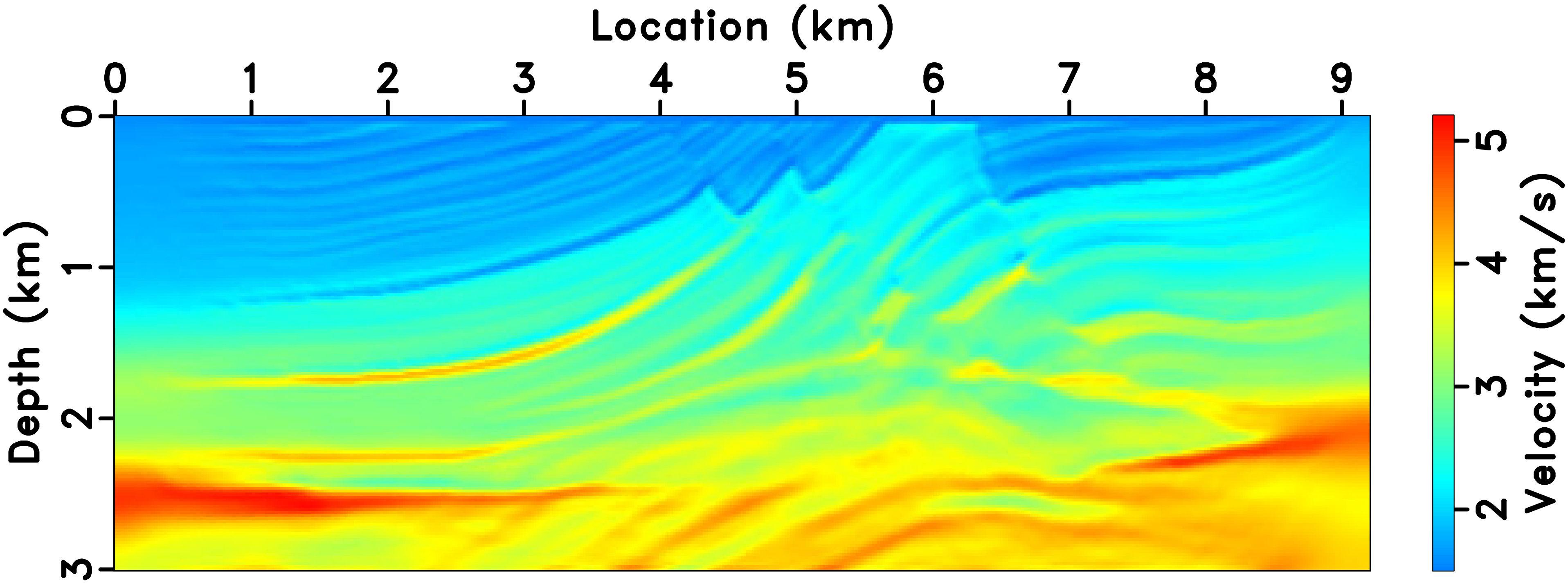
$$F^{\dagger}(\mathcal{F}[\mathbf{m}_0] - \mathbf{d})$$

FWI – Numerical Example



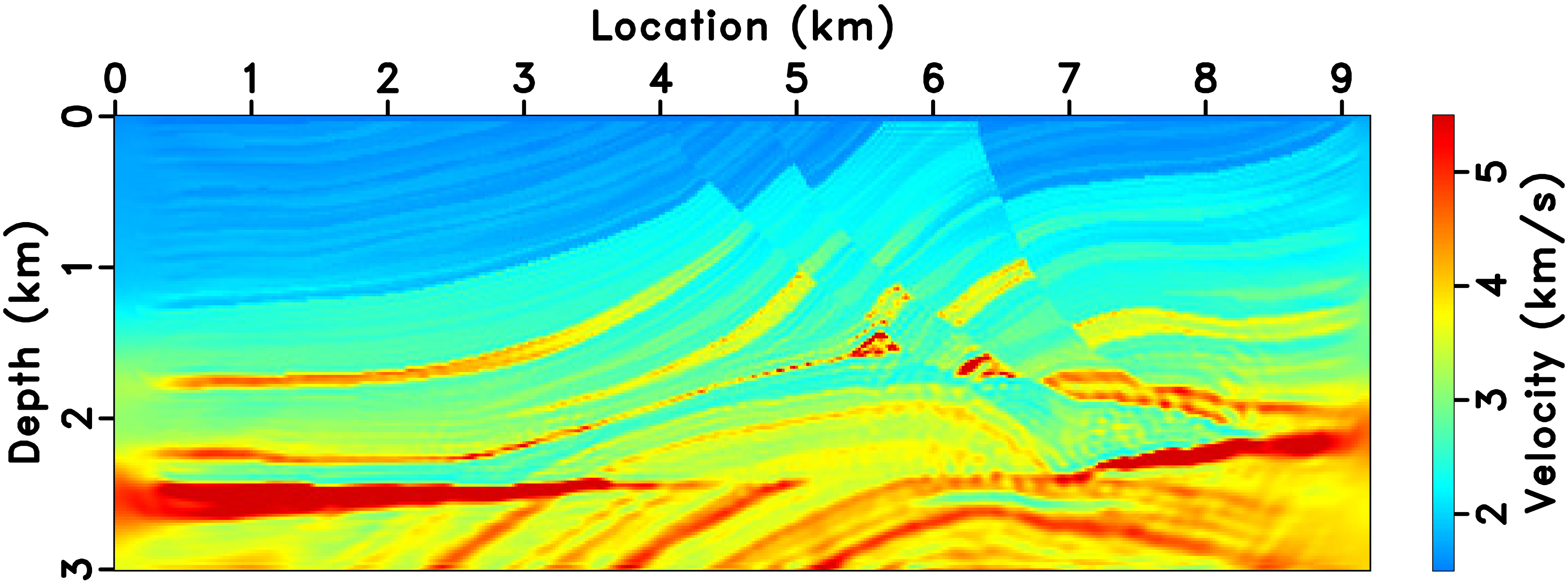
40 iteration steepest descent result

FWI – Numerical Example



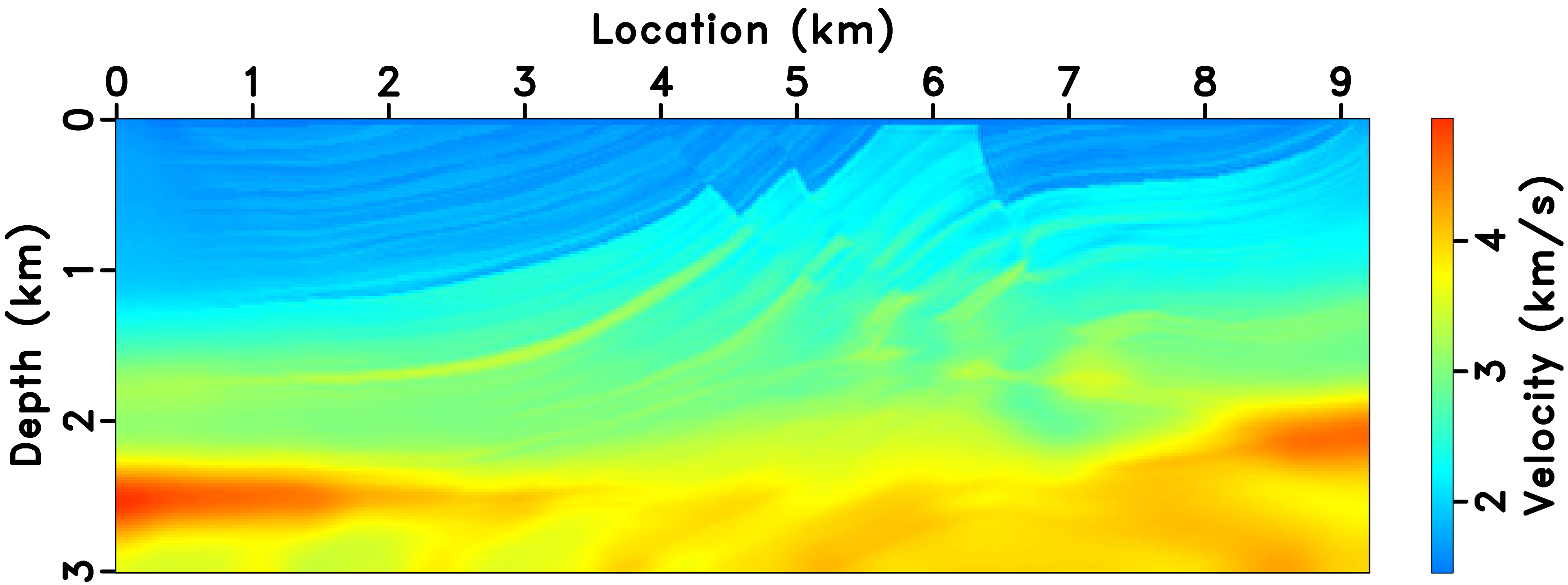
1 iteration approximate Gauss-Newton result

FWI – Numerical Example



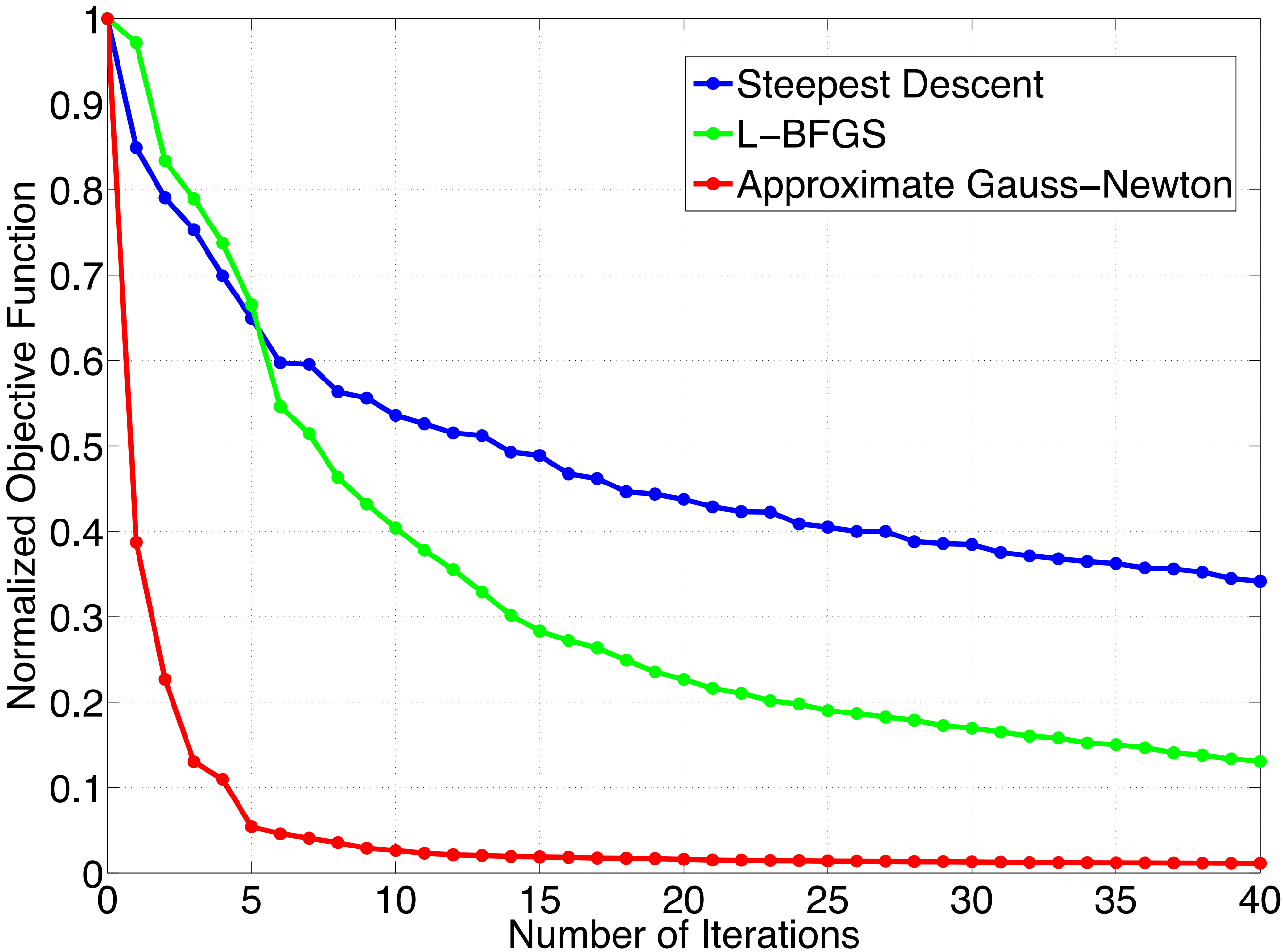
40 iteration approximate Gauss-Newton result

FWI – Numerical Example



40 iteration L-BFGS result

FWI – Convergence Curve



Summary

- Approximate Born inversion works even without subsurface offset
- Accelerate LSM with WCG by defining weighted norms
- Accelerate FWI with approximate Gauss-Newton by preconditioning the gradient

Acknowledgement

- Fons ten Kroode for inspiring our work
- Jon Sheiman, Henning Kuehl, Peng Shen
- Shell International Exploration & Production
- TRIP members and sponsors
- TACC and RCSG for HPC resources
- Thank you for listening