Tomographic full waveform inversion: Practical and computationally feasible approach

SEP147-13

Ali Almomin*, Biondo Biondi

May 22, 2012

Outline

- Introduction
- Theory
  - Scale separation
  - Scale mixing
- Synthetic examples
  - Gaussian model
  - Marmousi model
- Conclusions

Introduction

- FWI objective function:

$$J(\mathbf{v}) = \frac{1}{2}\left\|\mathbf{F}(\mathbf{v}) - \mathbf{d}_{\text{obs}}\right\|_2^2$$

- Simultaneous inversion of all scales (high resolution)
- Far from convex
- Requires very small errors in initial model

Introduction

- EFWI objective function:

$$J\left(\mathbf{v}\left(h\right)\right) = \frac{1}{2}\left\|\mathbf{F}\left(\mathbf{v}\left(h\right)\right) - \mathbf{d}_{\text{obs}}\right\|_2^2$$

- High resolution
- Fits the data easily
- Energy can be at any subsurface offset
- Not physical
- Very expensive

Introduction

- TFWI objective function:

$$J(\mathbf{v}(h)) = \frac{1}{2}\left\|\mathbf{F}(\mathbf{v}(h)) - \mathbf{d}_{\text{obs}}\right\|_2^2 + \frac{1}{2}\left\|\mathbf{E}(\mathbf{v}(h))\right\|_2^2$$

- High resolution
- Energy slowly moves towards zero offset
- Physical model
- Even more expensive (slower convergence)

Introduction

- Cost comparison to WEMVA:
- TFWI
  - Convolution along h every propagation step
  - Convolution along h every scattering/imaging
- WEMVA
  - Scalar multiplication every propagation step
  - Convolution along h every scattering/imaging

Introduction

- Problems so far:
  - Computational cost
  - Stability
  - It is very impractical to allow negative values and zeros in the velocity model

- Need a cheaper method without sacrificing the accuracy (mostly)

Outline

- Introduction
- Theory
  - Scale separation
  - Scale mixing
- Synthetic examples
  - Gaussian model
  - Marmousi model
- Conclusions

Scale separation

- When the model is not correct, a certain behavior with subsurface offset is observed:
  - The smooth components (long wavelength) are located mostly around the zero offset
  - The rough components (short wavelength) extend to large offsets

Scale separation

- Separate velocity into two components:

$$\mathbf{v} \approx \mathbf{b} + \mathbf{r}$$

- $\mathbf{b}$ = background (long wavelength)
- $\mathbf{r}$ = Born scattering potential (short wavelength)

- Assuming first order scattering data or "primaries" (Born approximation)

Scale separation

$$\mathbf{d} = \mathbf{F}(\mathbf{v})$$

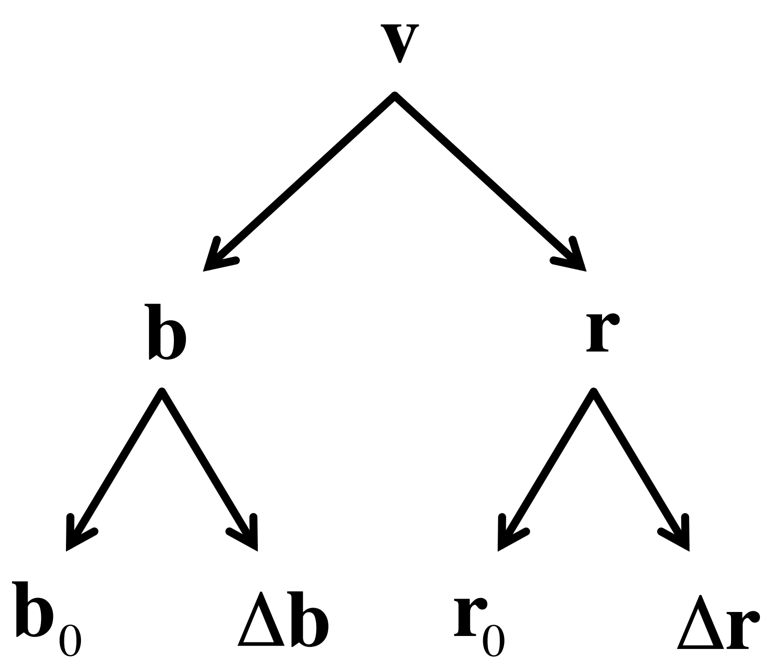
$$\downarrow$$

$$\mathbf{d} = \mathbf{L}(\mathbf{b})\ \mathbf{r}$$

$$\downarrow$$

$$\Delta\mathbf{d} = \mathbf{L}(\mathbf{b}_0)\ \Delta\mathbf{r}$$

$$\Delta\mathbf{d} = \mathbf{T}(\mathbf{b}_0, \mathbf{r}_0)\ \Delta\mathbf{b}$$



Scale separation

- Several approximations:

$$\mathbf{v}(h) = \mathbf{b}(h) + \mathbf{r}(h)$$

Scale separation

- Several approximations:

$$\mathbf{v}(h) = \mathbf{b}(h) + \mathbf{r}(h)$$

$$\mathbf{v}(h) = \mathbf{b}(h_{\text{small}}) + \mathbf{r}(h)$$

Scale separation

- Several approximations:

$$\mathbf{v}(h) = \mathbf{b}(h) + \mathbf{r}(h)$$

$$\mathbf{v}(h) = \mathbf{b}(h_{\text{small}}) + \mathbf{r}(h)$$

$$\mathbf{v}(h) = \mathbf{b} + \mathbf{r}(h)$$

Scale separation

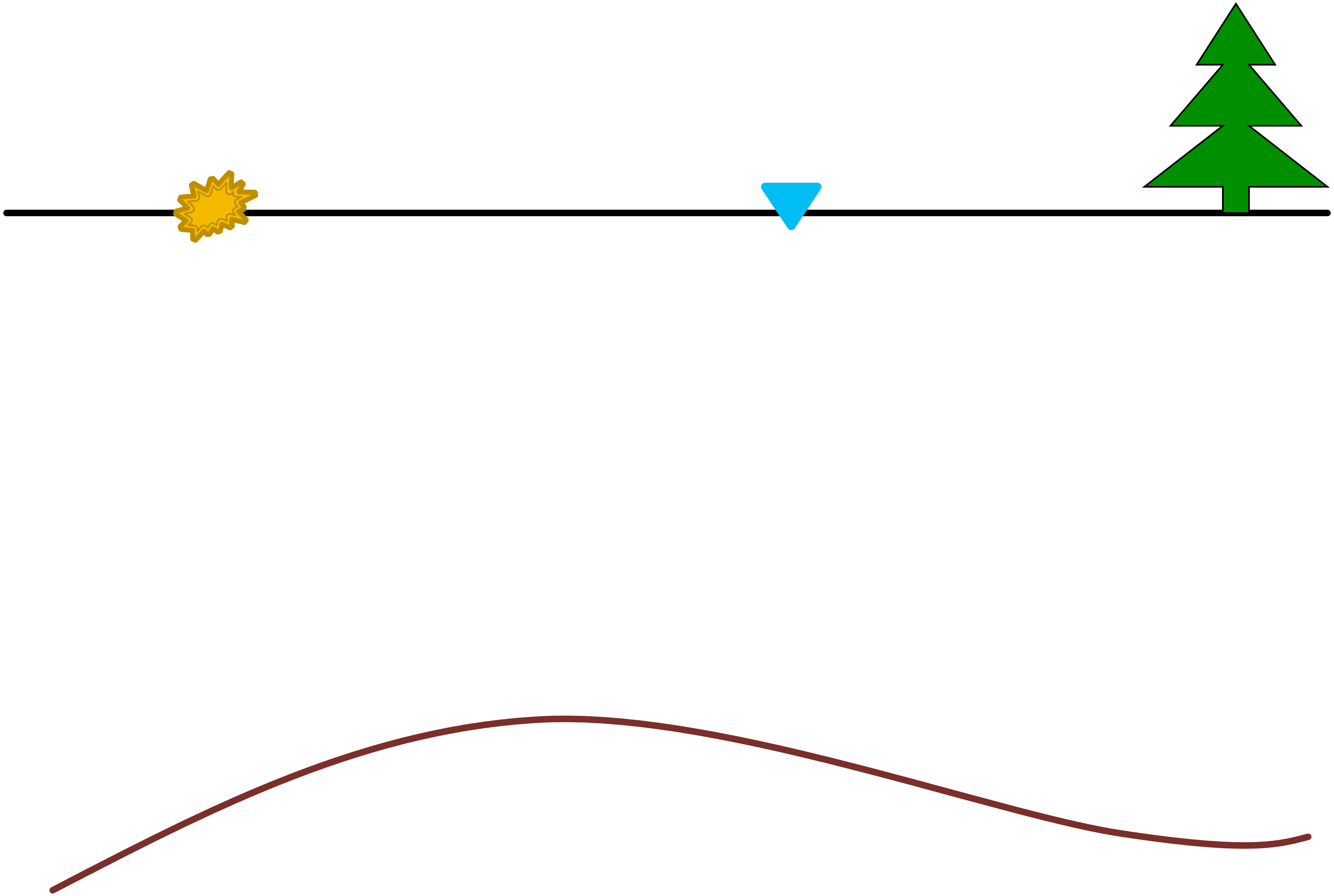
- ETFWI objective function:

$$J(\mathbf{b},\mathbf{r}) = \frac{1}{2}\left\|\mathbf{L}(\mathbf{b})\ \mathbf{r} - \mathbf{d}_{obs}\right\|_2^2 + \frac{1}{2}\left\|\mathbf{E}\ \mathbf{r}\right\|_2^2$$

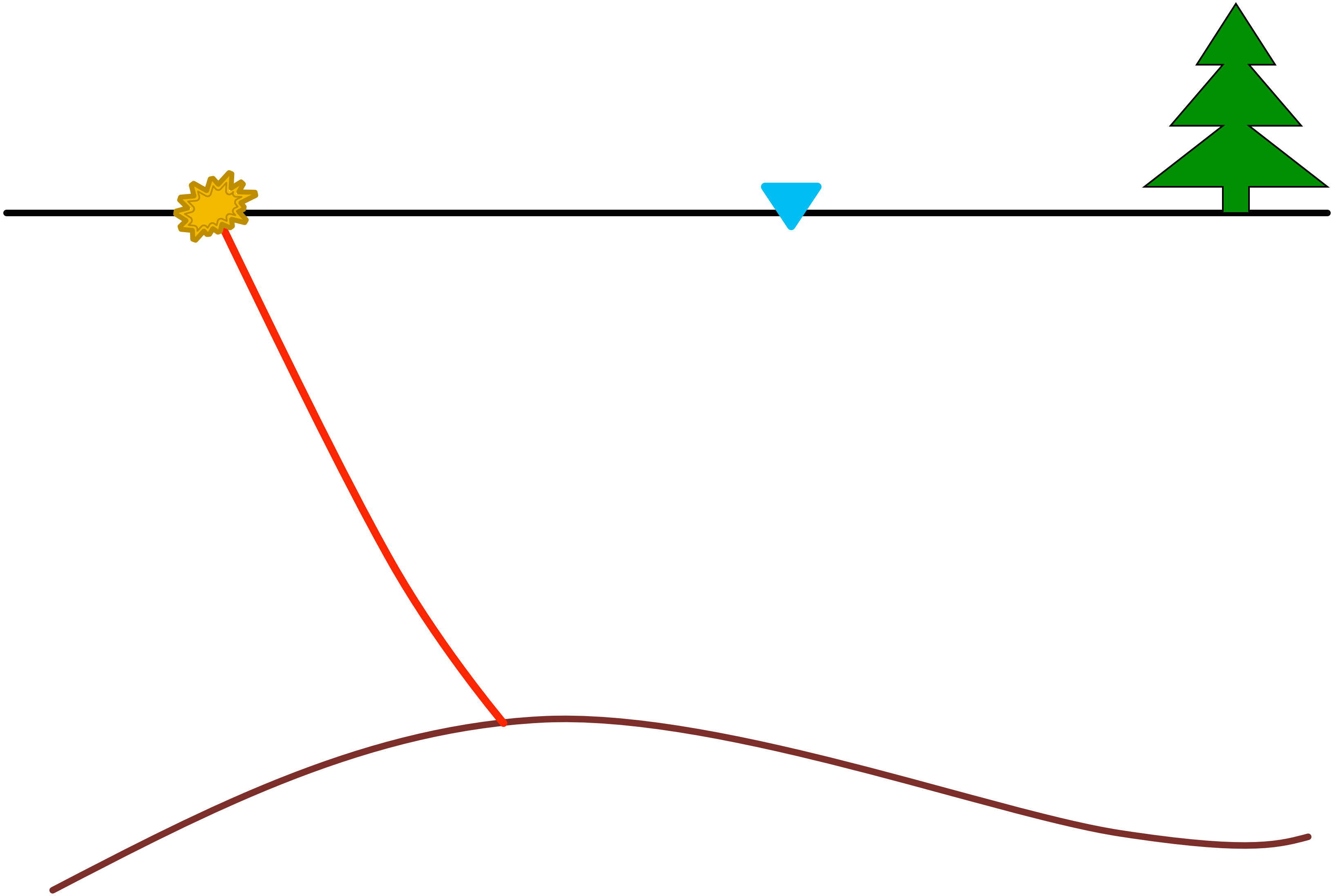
- ETFWI gradients:

$$\frac{\partial J}{\partial \mathbf{r}} = \mathbf{L}^*(\mathbf{b}_0)\Delta\mathbf{d} \qquad \frac{\partial J}{\partial \mathbf{b}} = \left(\frac{\partial \mathbf{L}}{\partial \mathbf{b}}\mathbf{r}_0\right)^* \Delta\mathbf{d} = \mathbf{T}^*(\mathbf{b}_0,\mathbf{r}_0)\Delta\mathbf{d}$$

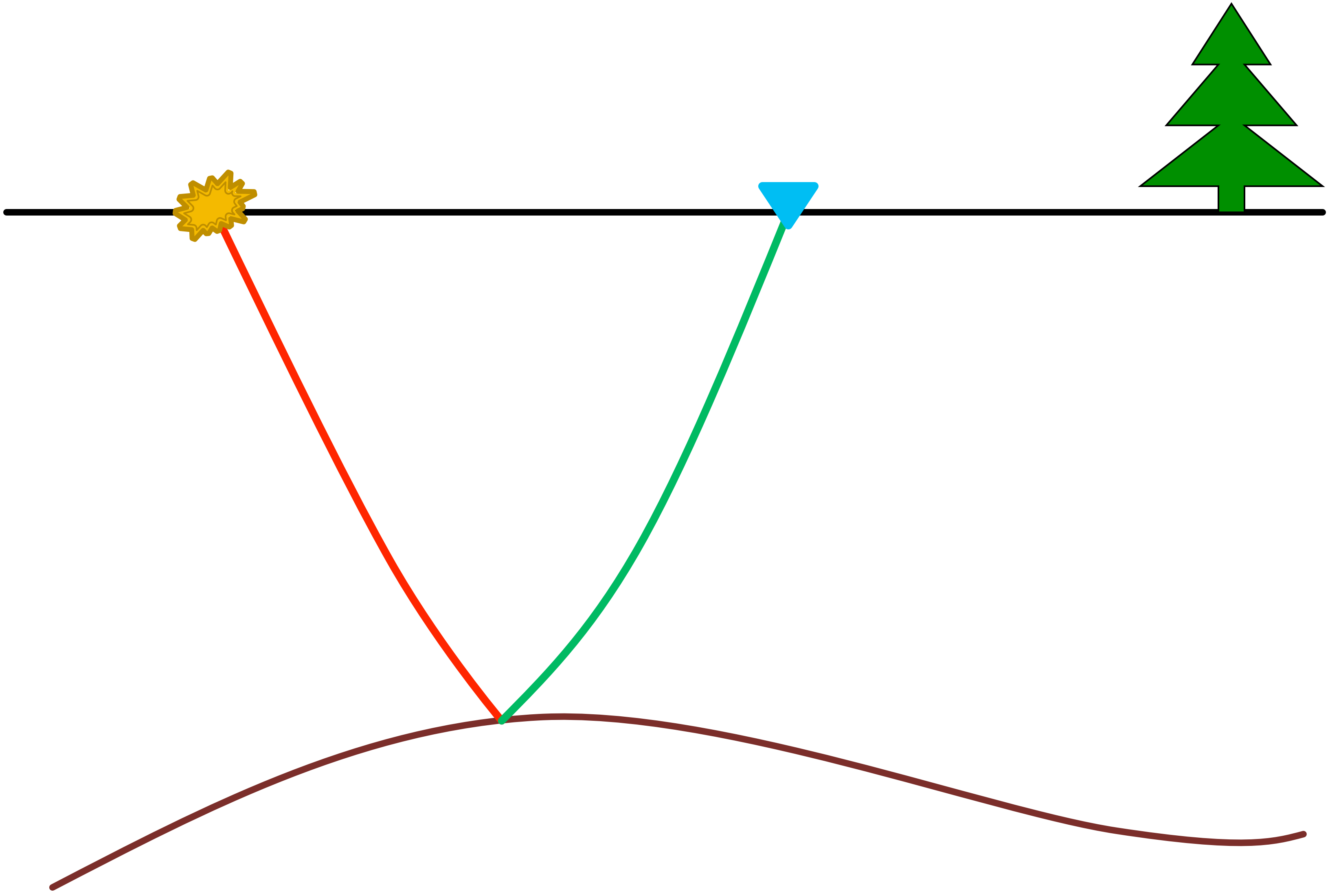
$\mathbf{L}$ = Born modeling op $\qquad$ $\mathbf{T}$ = Tomographic op



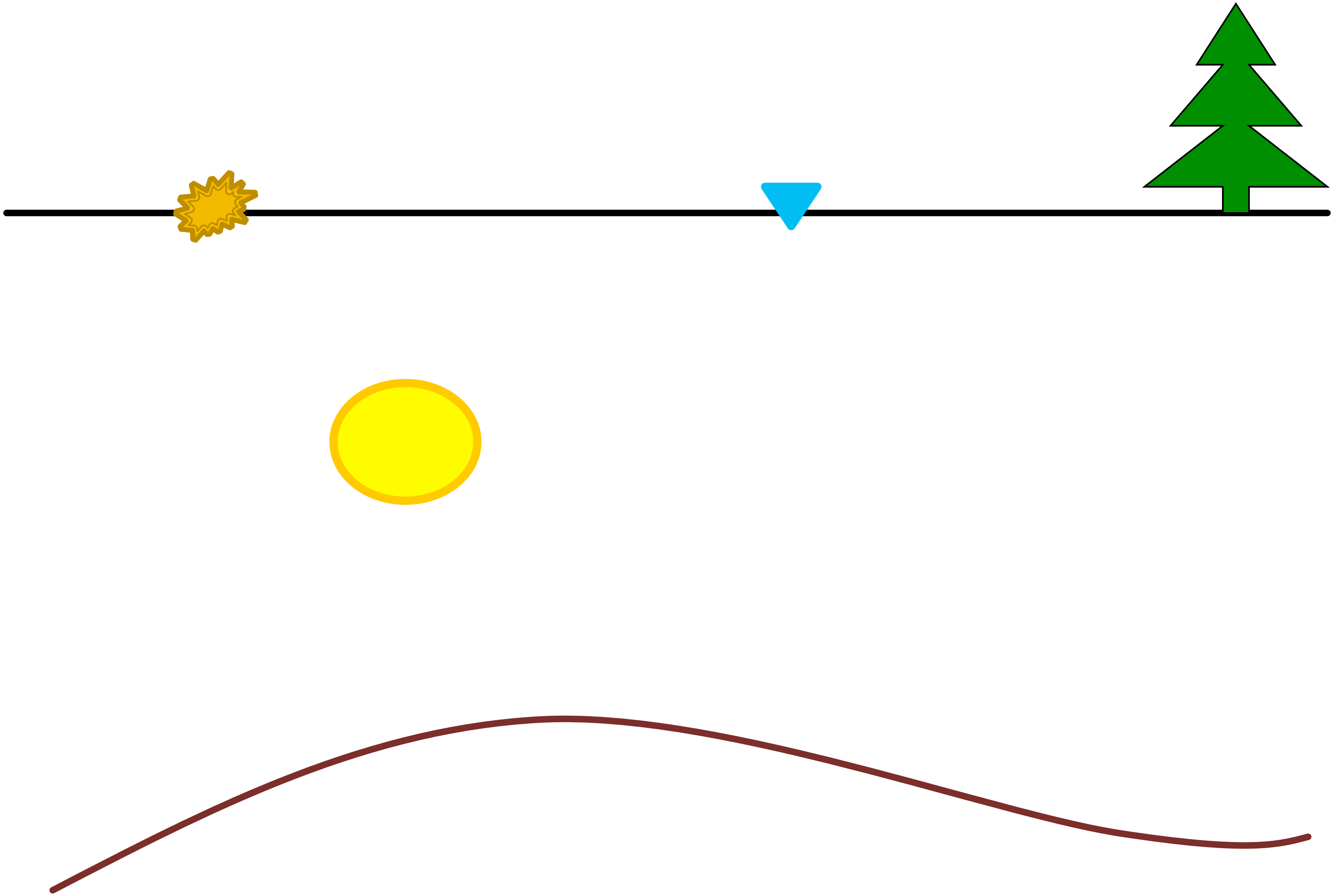
Born Operator



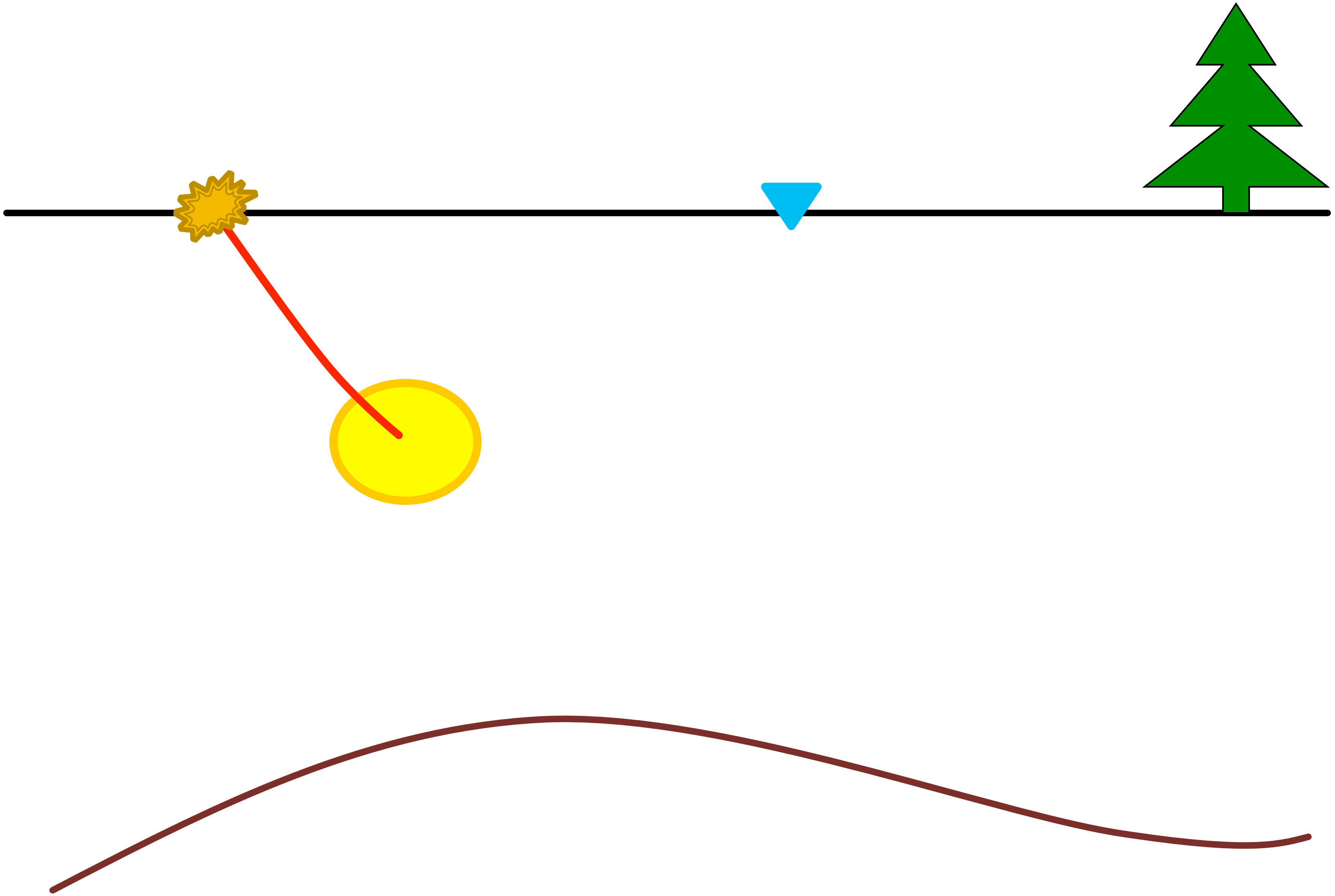
Born Operator



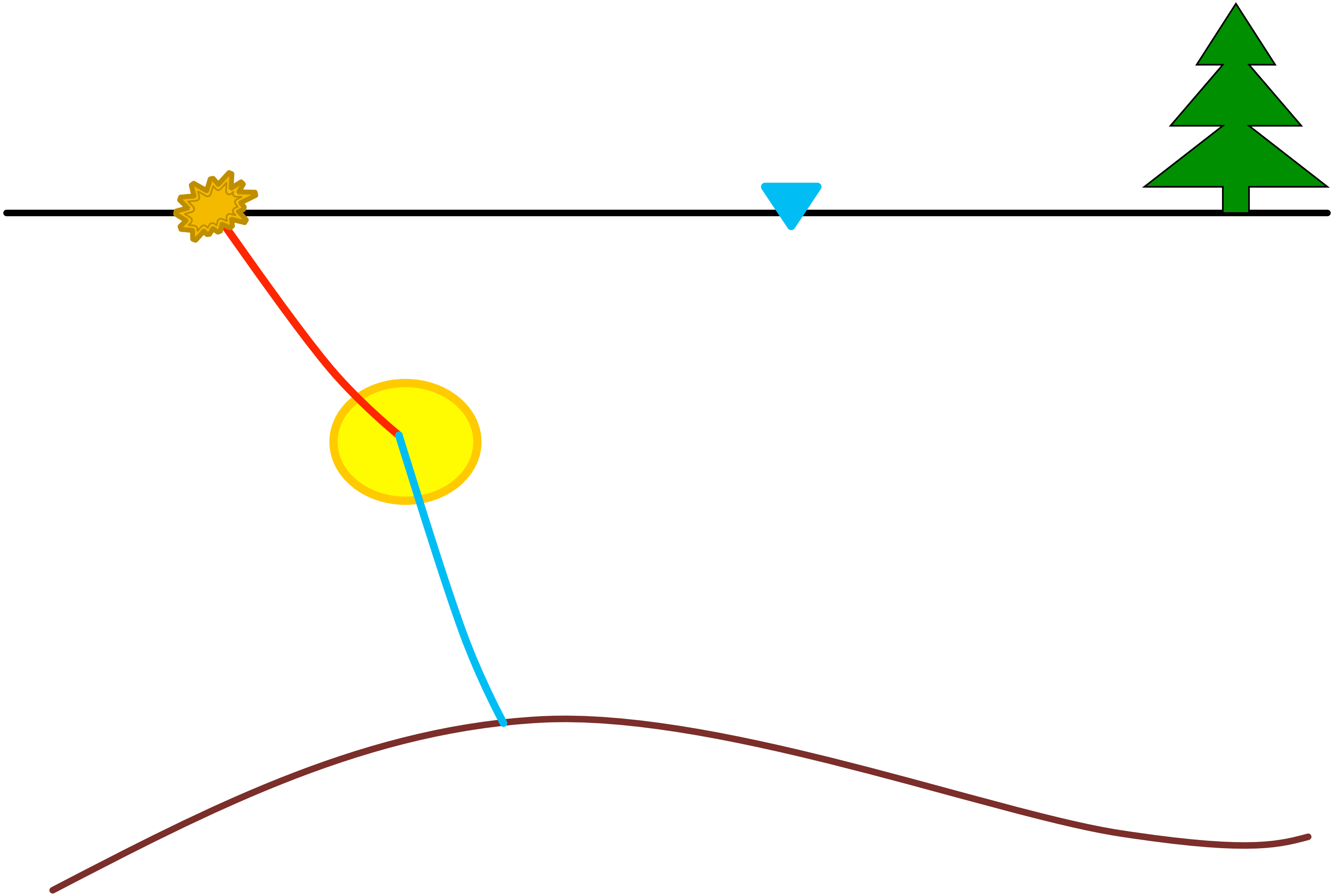
Born Operator



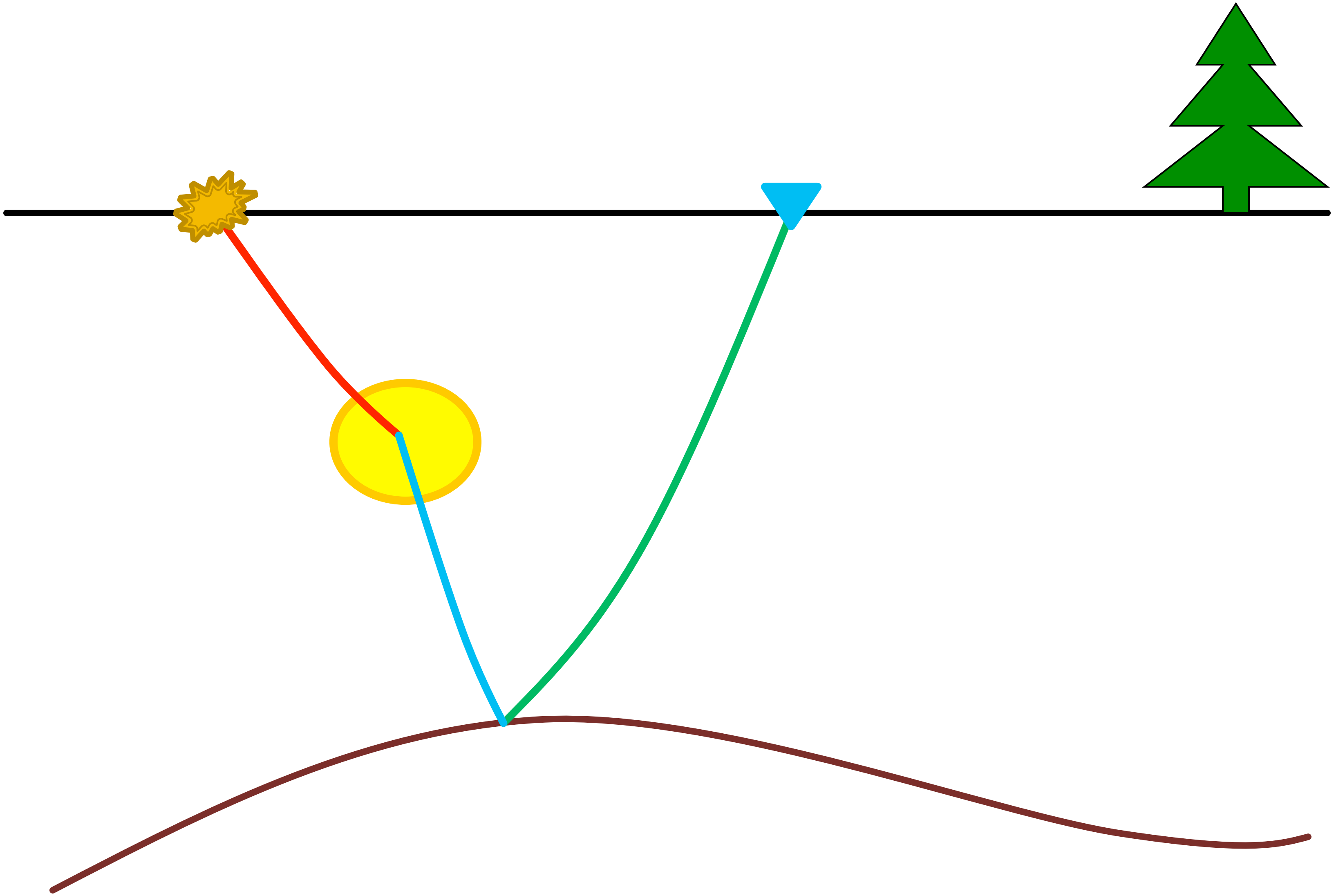
Tomographic Operator



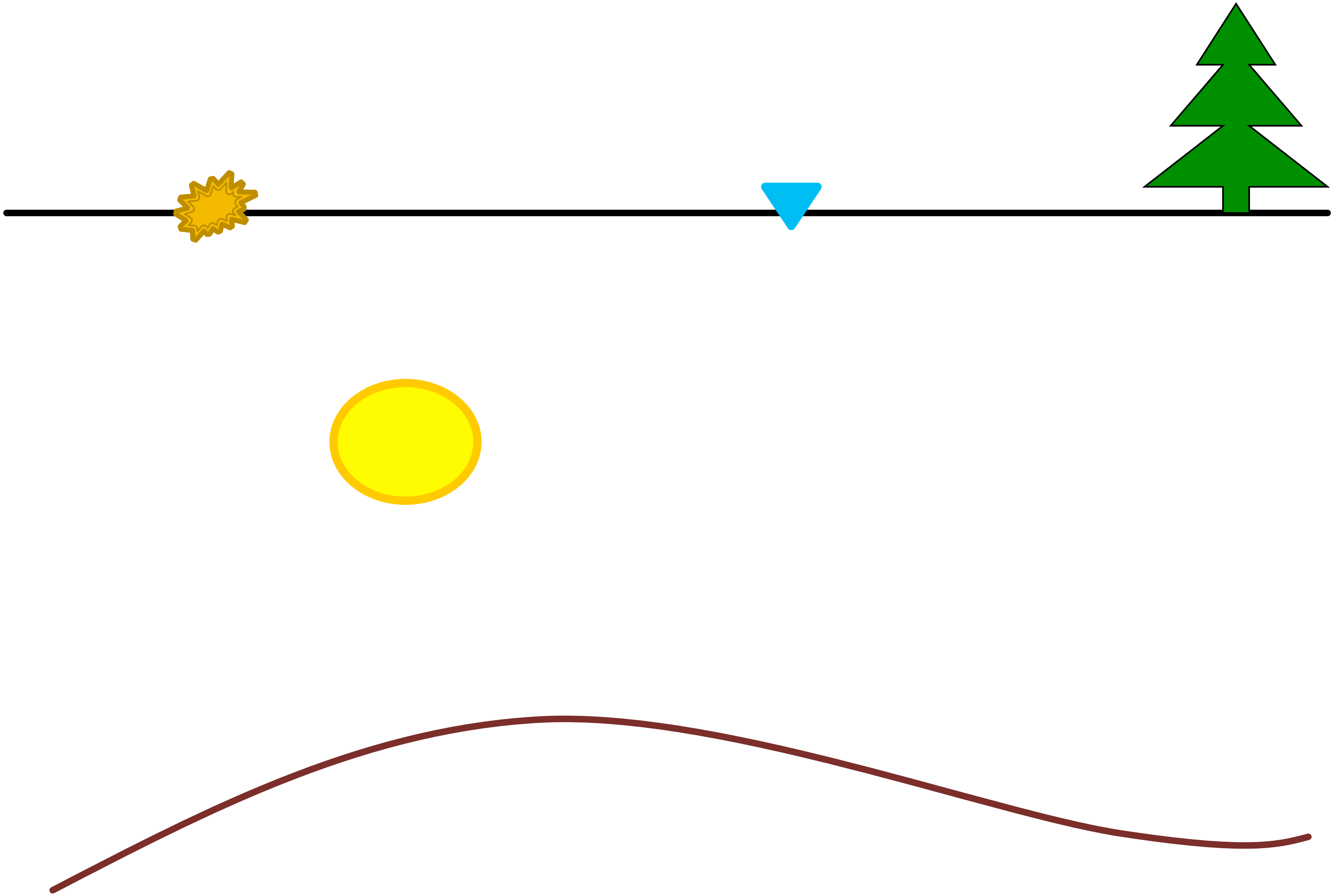
Tomographic Operator



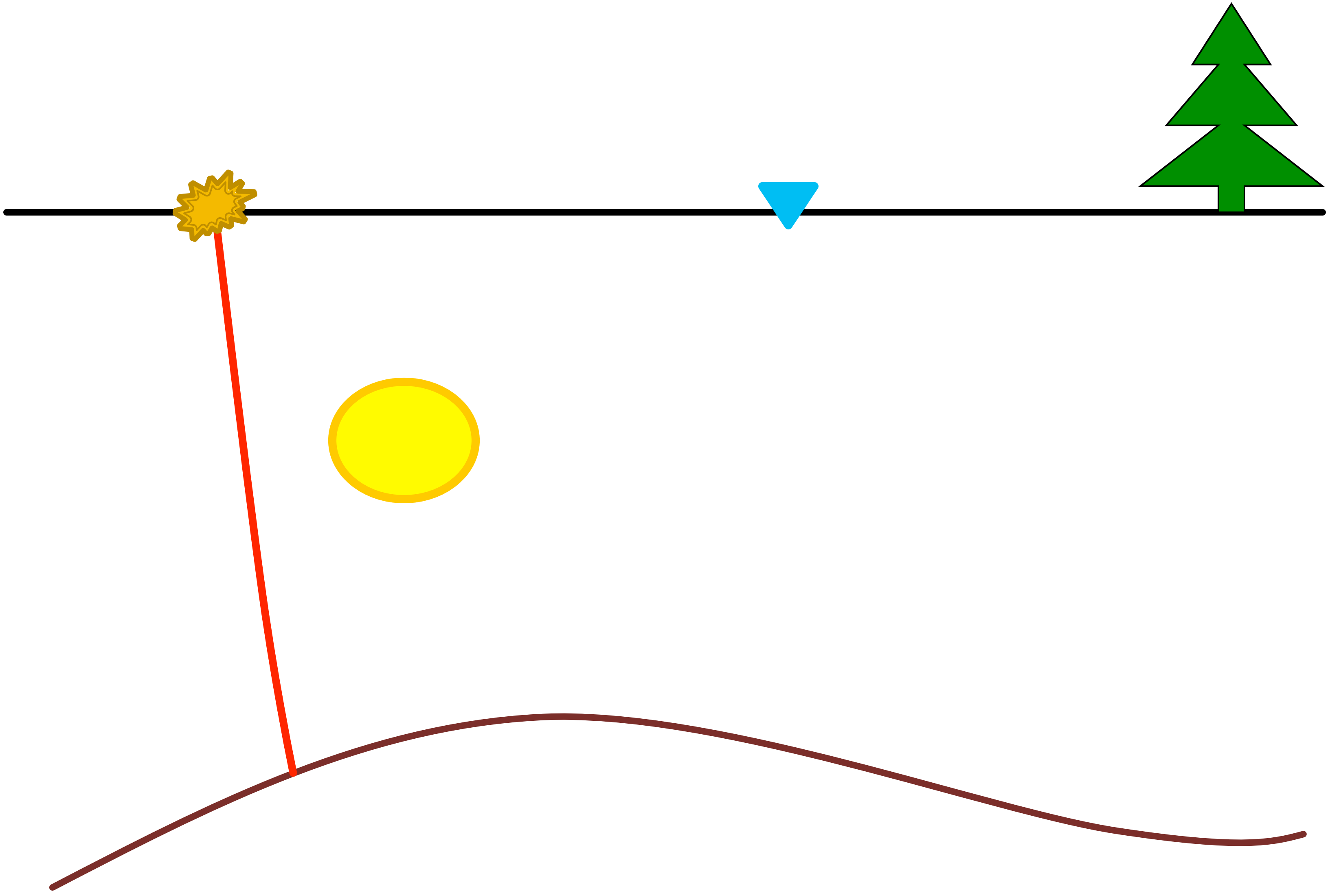
Tomographic Operator



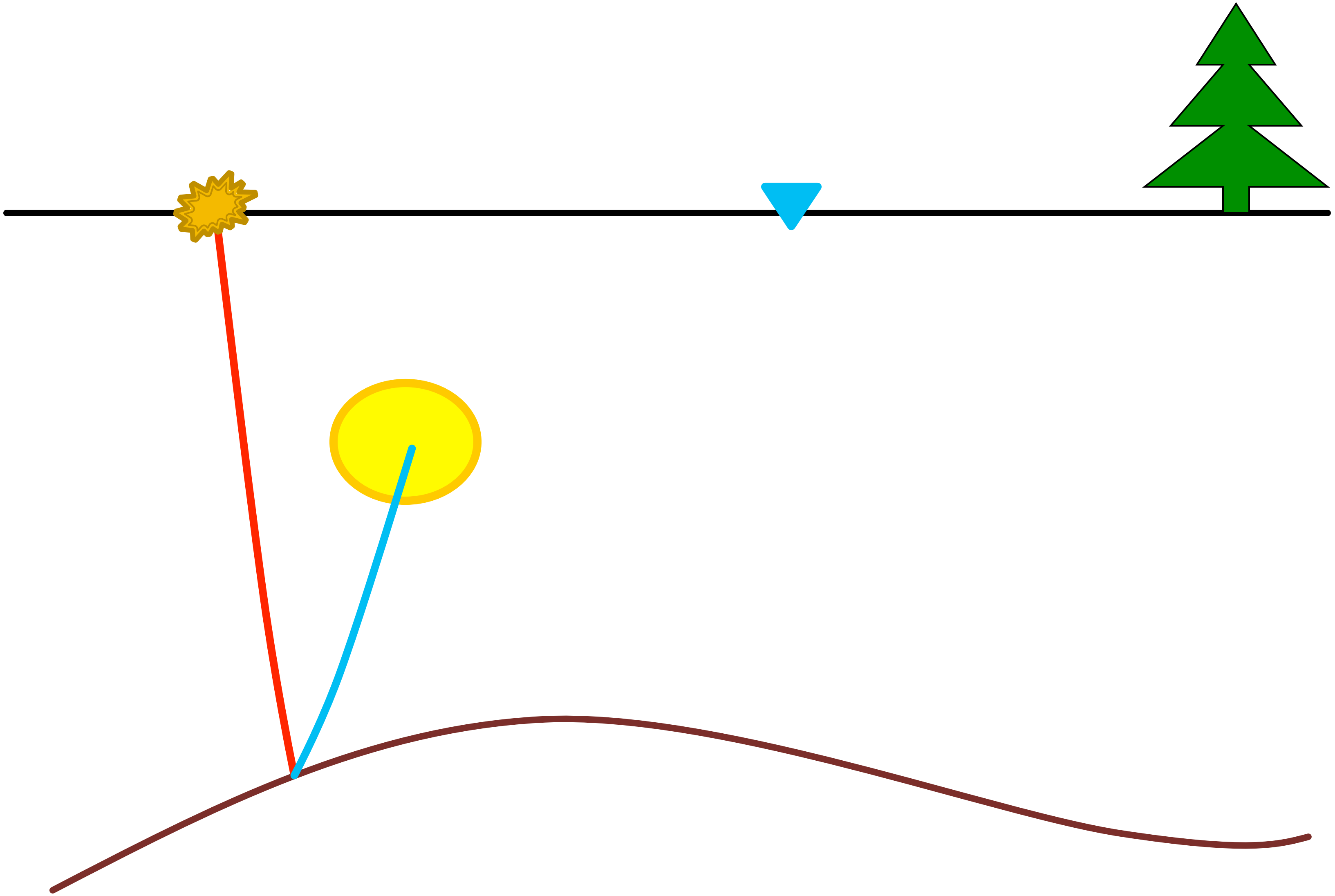
Tomographic Operator



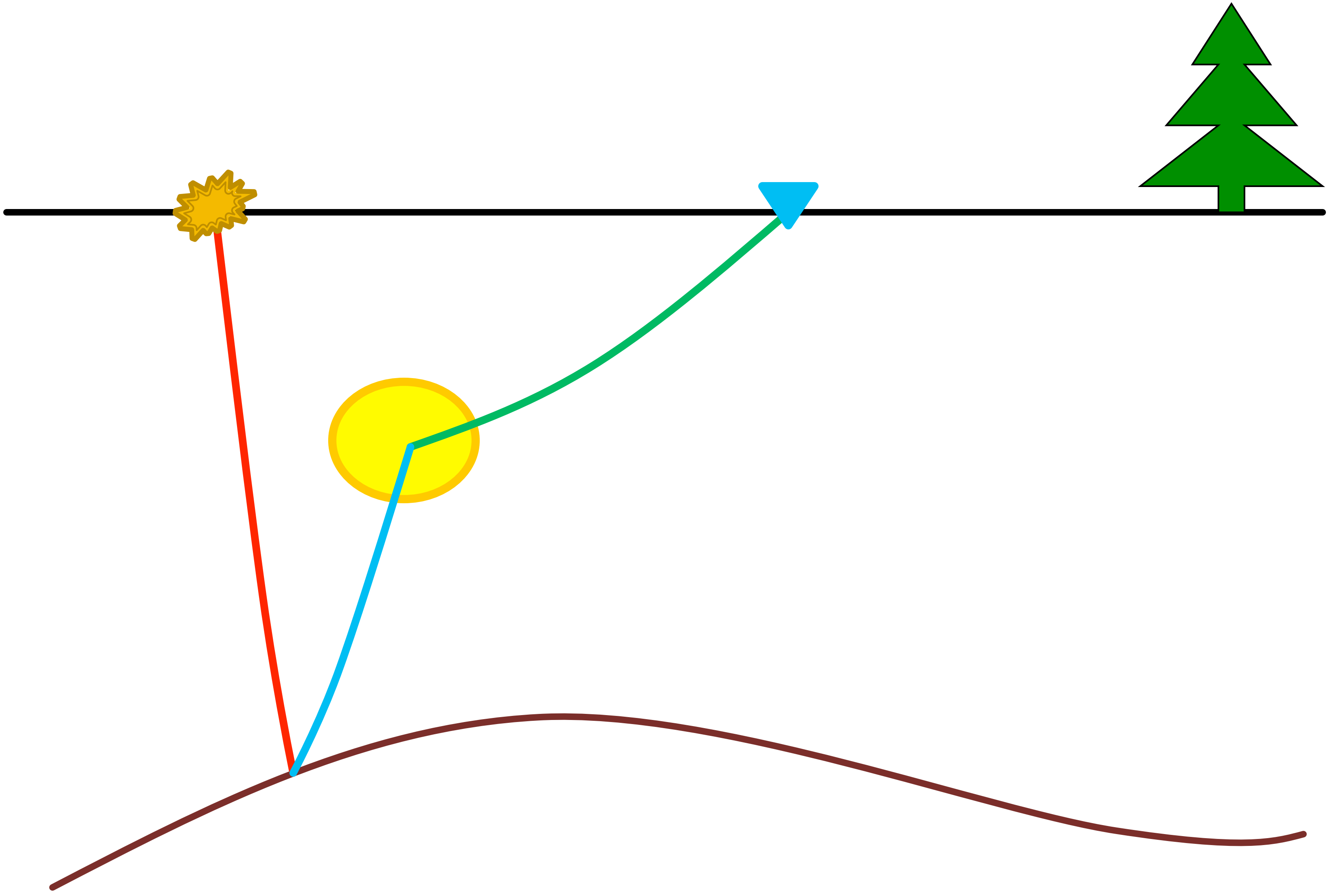
Tomographic Operator



Tomographic Operator



Tomographic Operator



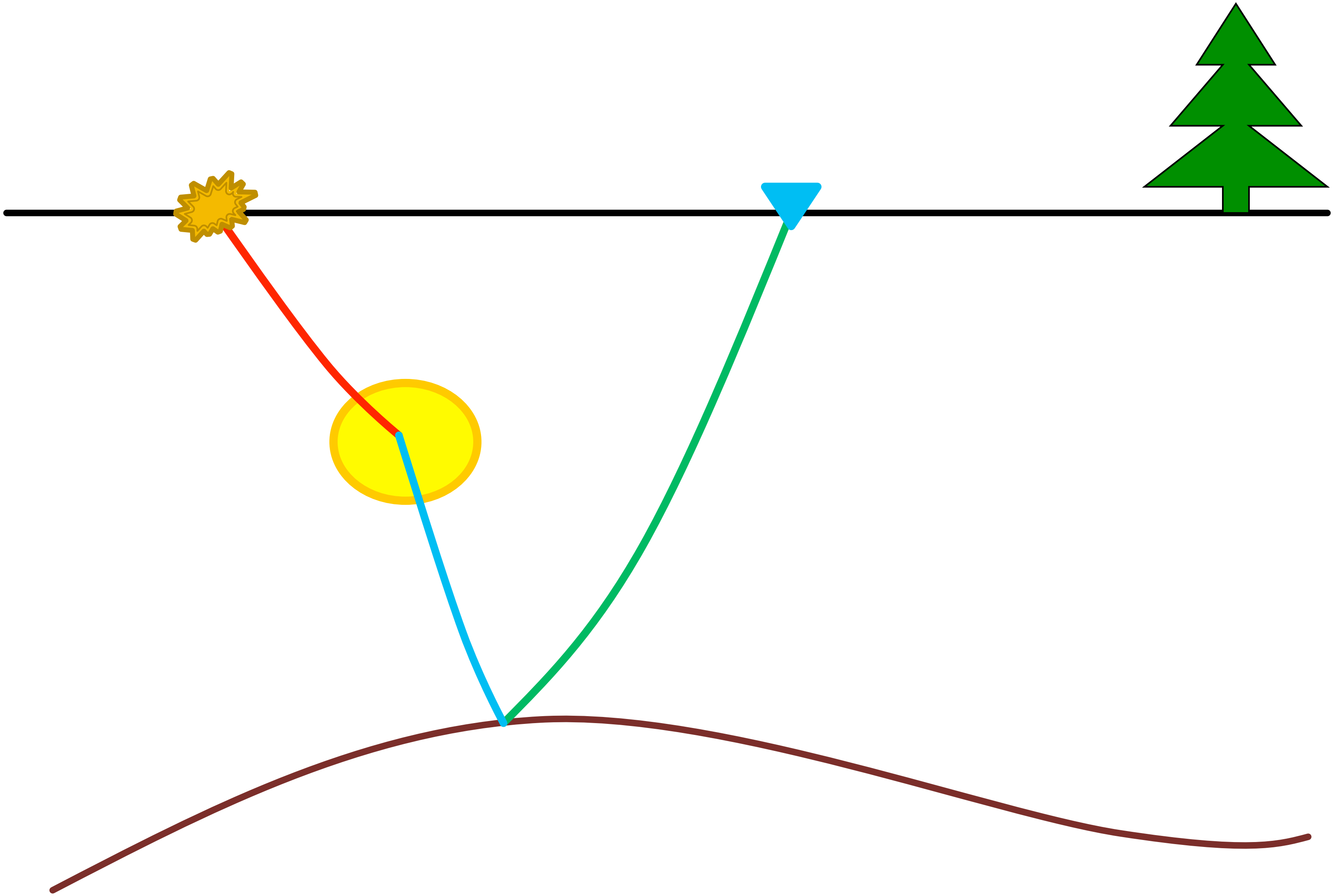
Tomographic Operator

Scale separation

- Relationship of tomographic operator to WEMVA operator:

$$\Delta \mathbf{v}_{\mathrm{ETFWI}} = \sum \mathbf{U}\ \mathbf{r}\ \Delta \mathbf{d}^{*}$$

$$\Delta \mathbf{v}_{\mathrm{WEMVA}} = \sum \mathbf{U}\ \Delta \mathbf{I}\ \mathbf{d}^{*}$$



Tomographic vs. WEMVA

Scale separation

- Relationship of tomographic operator to WEMVA operator:

$$\Delta \mathbf{v}_{\mathrm{ETFWI}} = \sum \mathbf{U}\ \mathbf{r}\ \Delta \mathbf{d}^{*}$$

$$\Delta \mathbf{v}_{\mathrm{WEMVA}} = \sum \mathbf{U}\ \Delta \mathbf{I}\ \mathbf{d}^{*}$$

- Adjoint artifacts
- Accurate matching

Scale mixing

- We achieved cost cutting
- We achieved an additional degree of freedom
  - Can potentially handle variations in density and AVO effects
- Did we achieve the same accuracy?
  - Only using "primaries"
  - Not completely simultaneous inversion

Scale mixing

- ETFWI gradients:

$$\frac{\partial J}{\partial \mathbf{r}} = \mathbf{L}^{*}\left(\mathbf{b}_0\right)\Delta\mathbf{d} \qquad \frac{\partial J}{\partial \mathbf{b}} = \mathbf{T}^{*}\left(\mathbf{b}_0, \mathbf{r}_0\right)\Delta\mathbf{d}$$

- The two models are indirectly connected:
  - Data residuals (of next iteration)
  - Reflectivity
- Not directly connected in model space

Scale mixing

- There is an important influence in model space between $\mathbf{r}$ and $\mathbf{b}$
  - Data-constrained model components
    - Both parameters can share the same model components at different frequencies/offsets
  - Null model space
    - A component in the null space of one parameter might by constrained by the other parameter
    - This applies to both signal and noise

Scale mixing

- Scale mixing can be done by radial tapering in Fourier domain:

$$\mathbf{s_b} = \mathbf{C_b}\left(\mathbf{g_b} + \mathbf{g_r}\left(h = 0\right)\right)$$

$$\mathbf{s_r} = \mathbf{C_r}\left(\mathbf{g_b} + \mathbf{g_r}\right)$$

$$\mathbf{C_b} + \mathbf{C_r} = \mathbf{I}$$

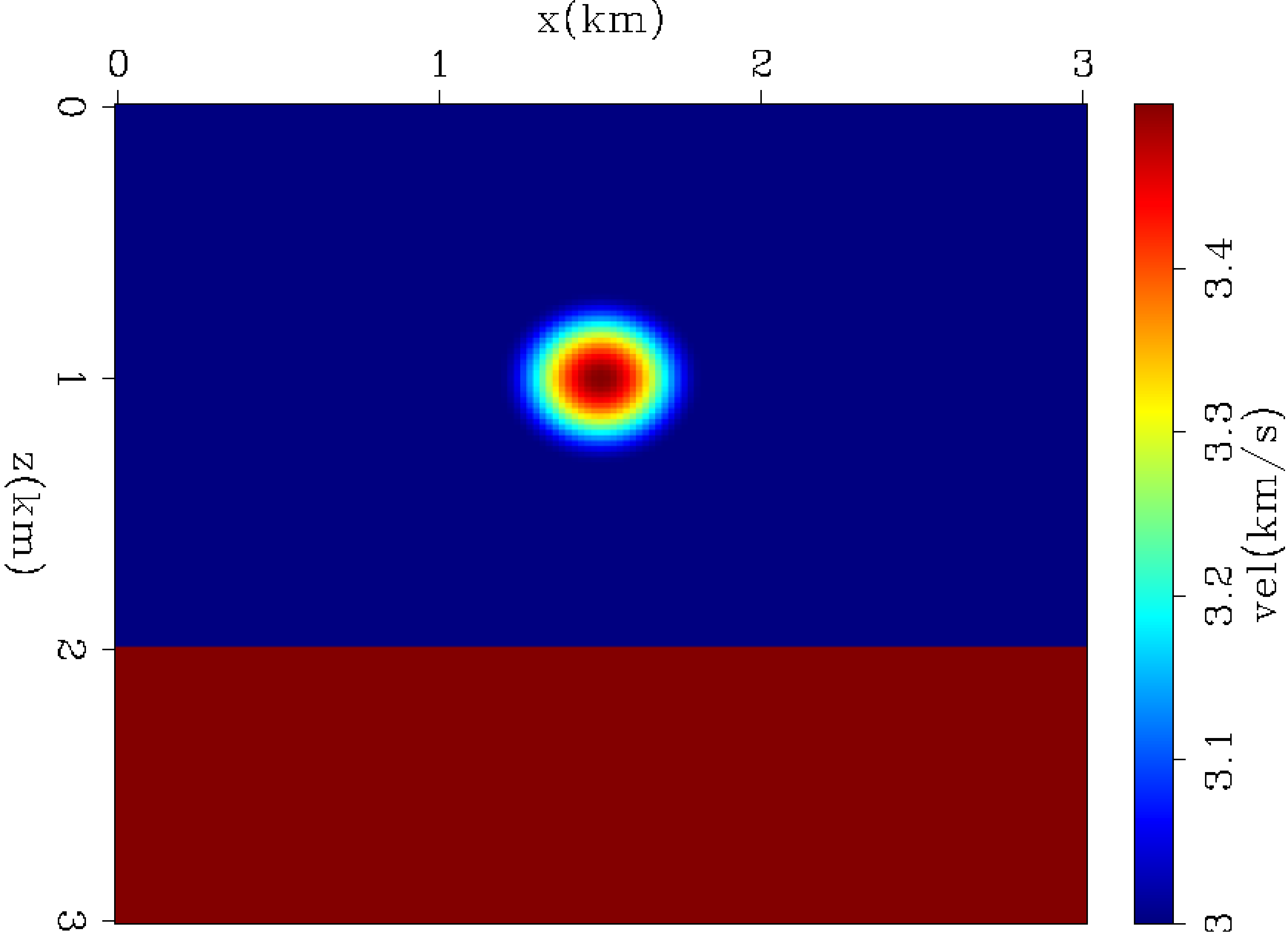
- Both gradients need to have the same units

Outline

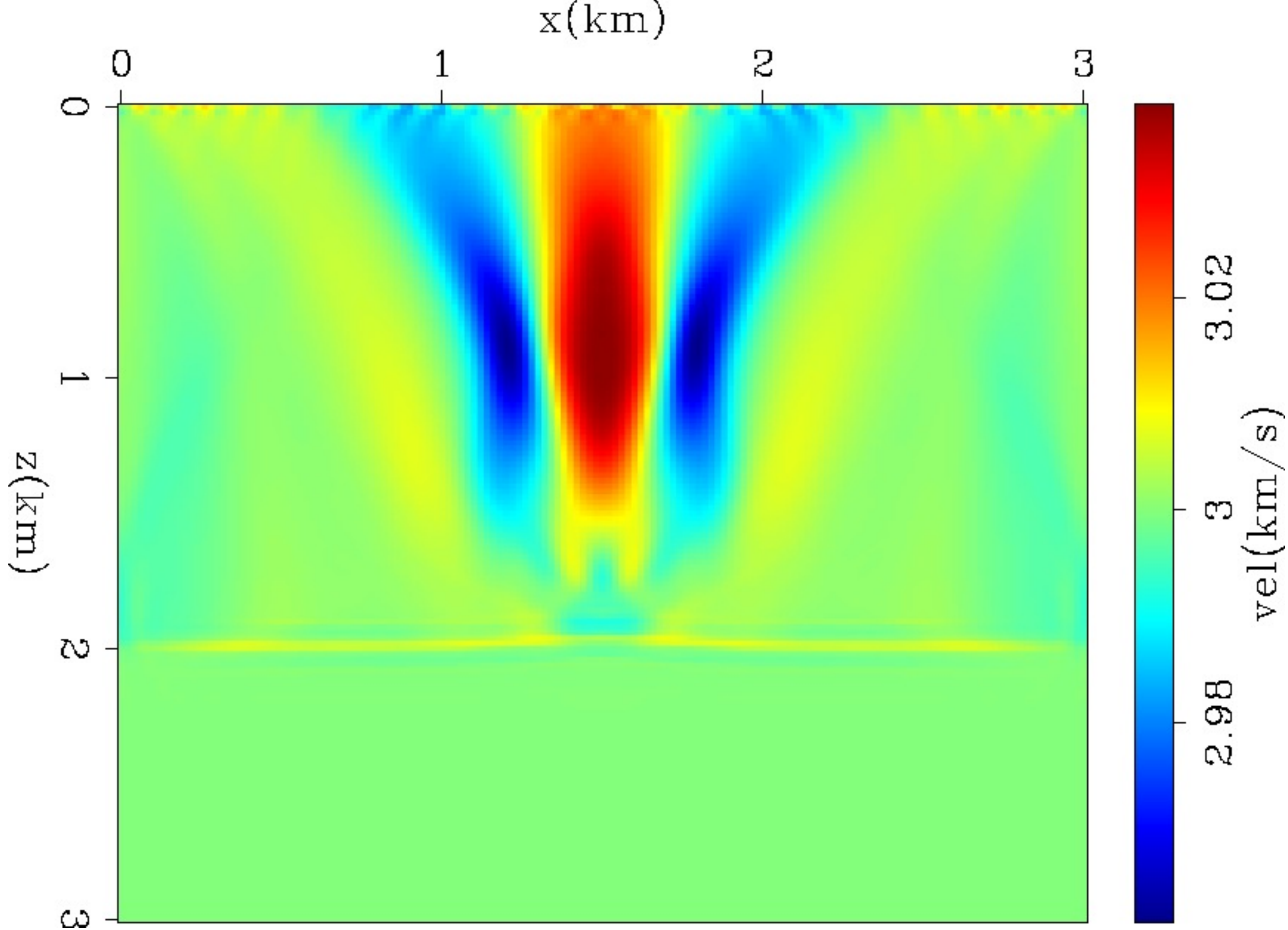
- Introduction
- Theory
  - Scale separation
  - Scale mixing
- Synthetic examples
  - Gaussian model
  - Marmousi model
- Conclusions

Gaussian model

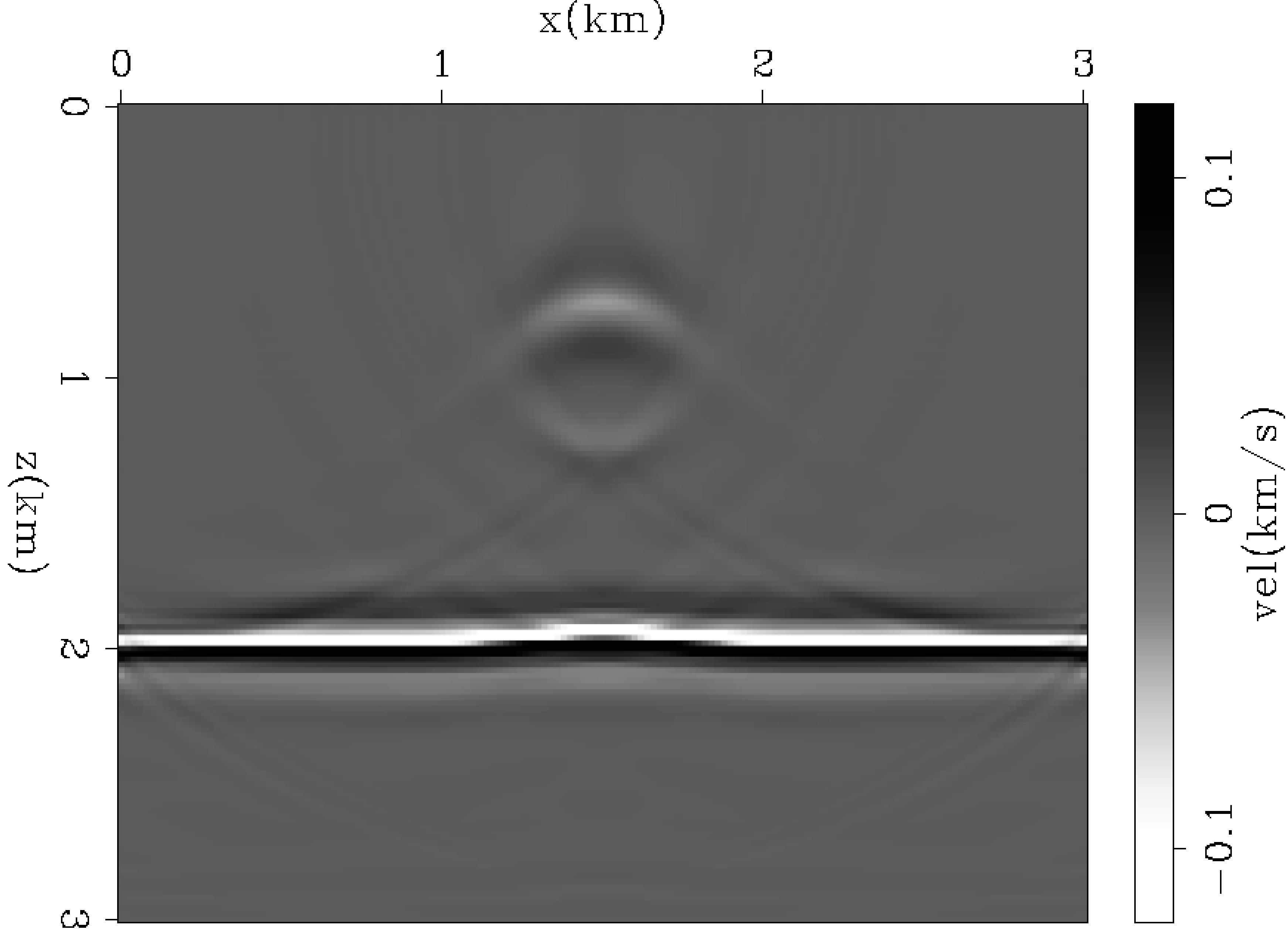
- Gaussian model
  - Model size is 3 km x 3 km
  - Ricker wavelet of 15Hz
  - Maximum offset of 1.5km
  - Source spacing 100m
  - Receiver spacing 20m



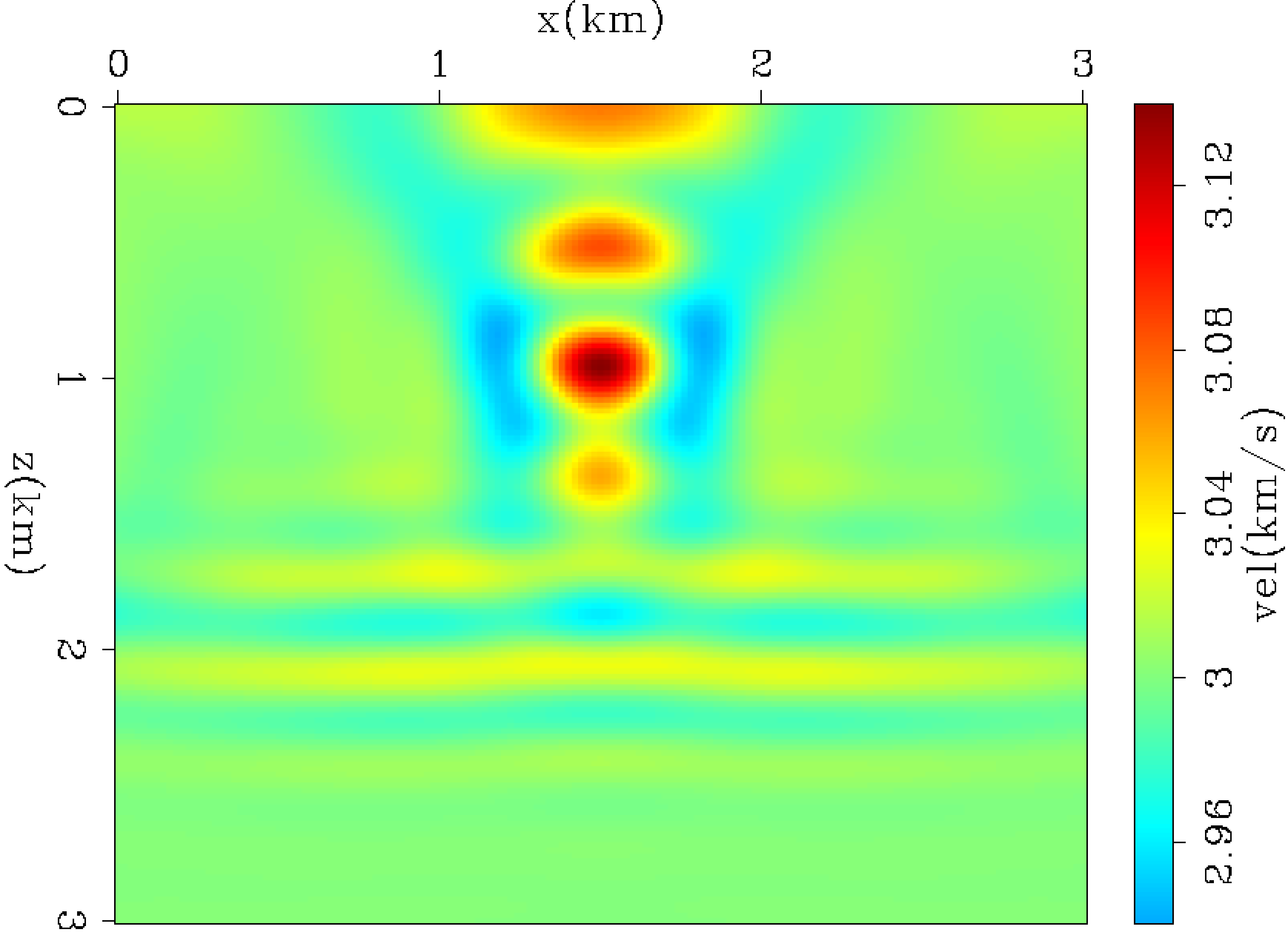
True Gaussian velocity



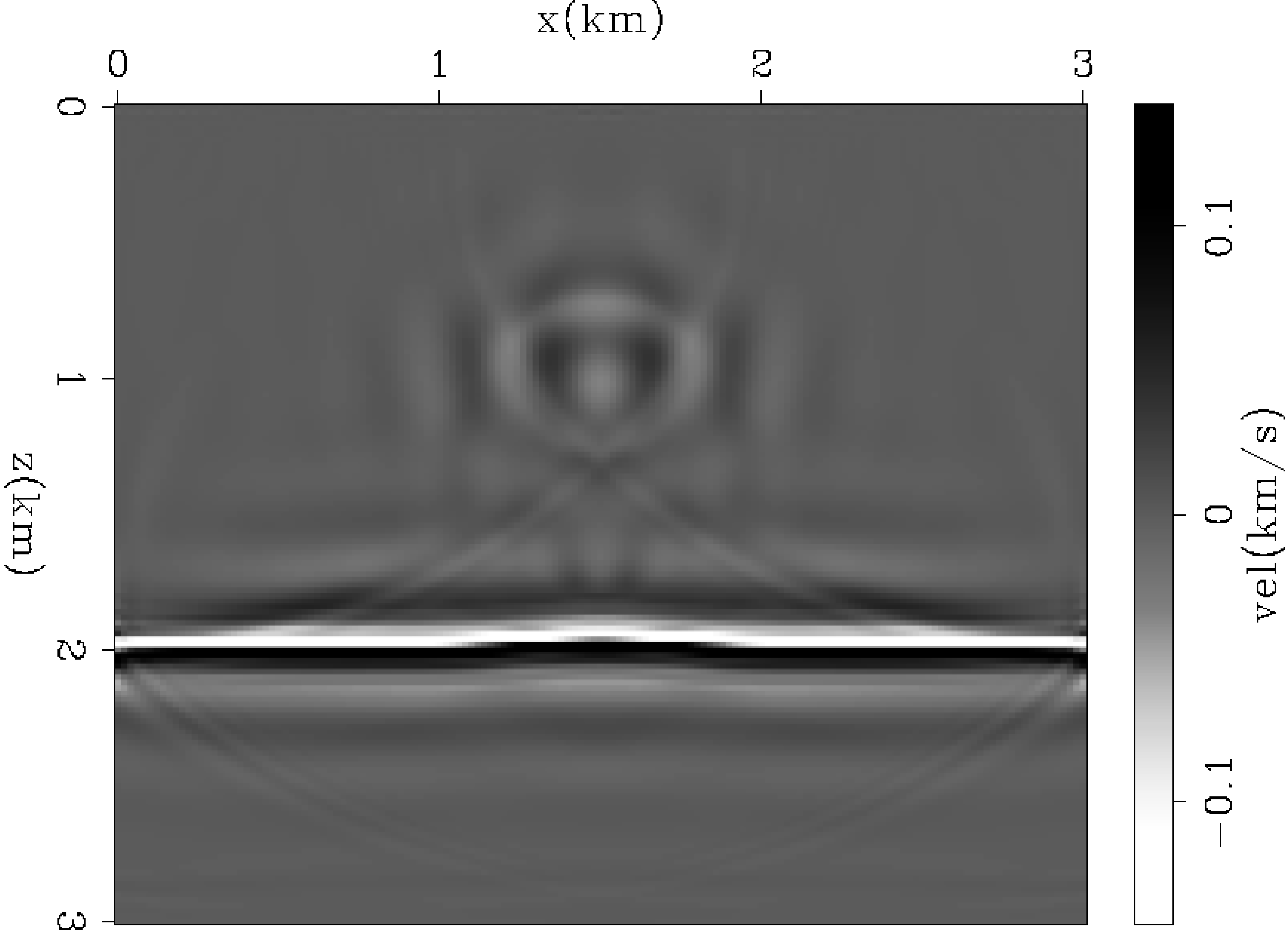
Inverted background



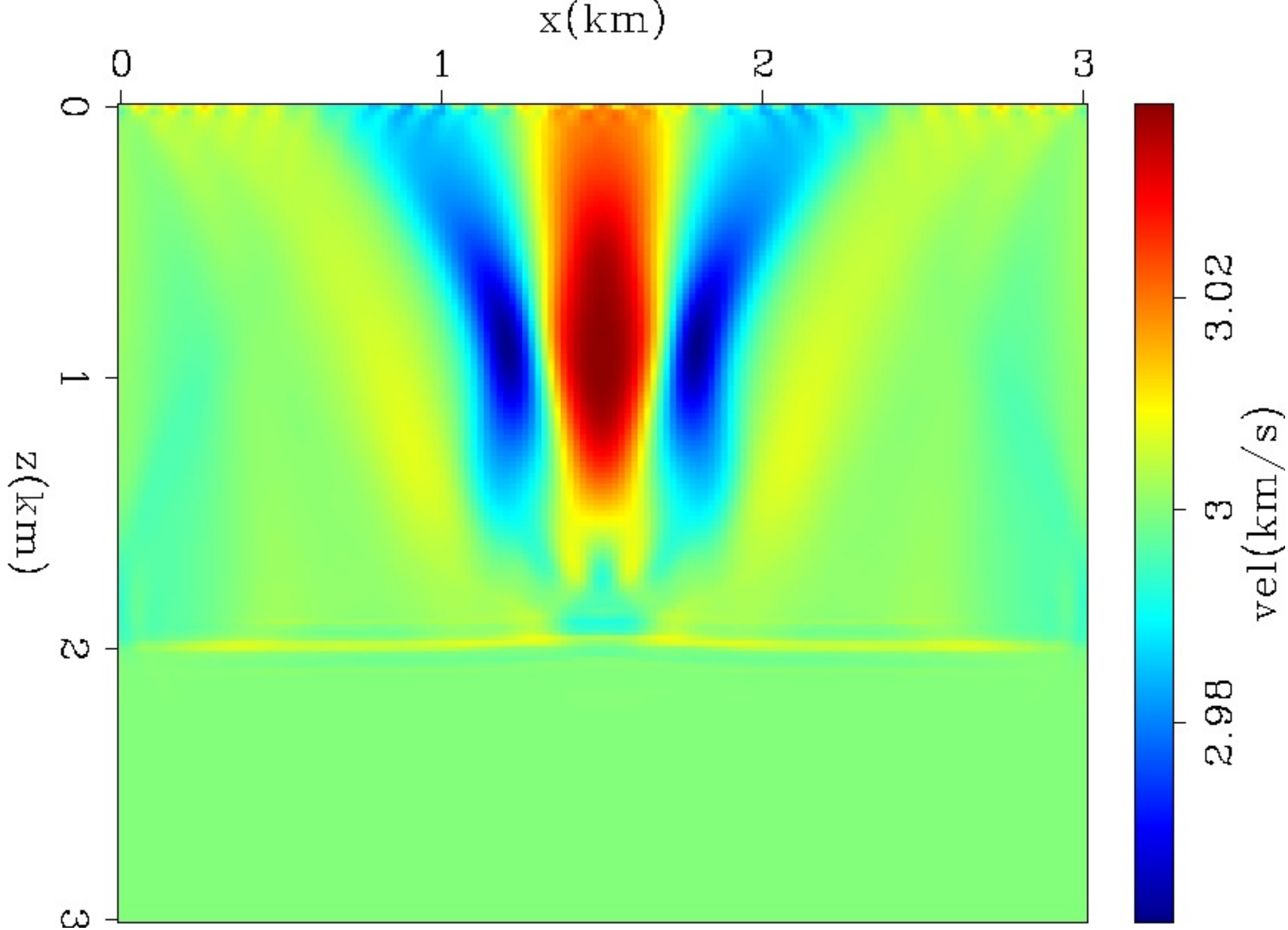
Inverted Born reflectivity



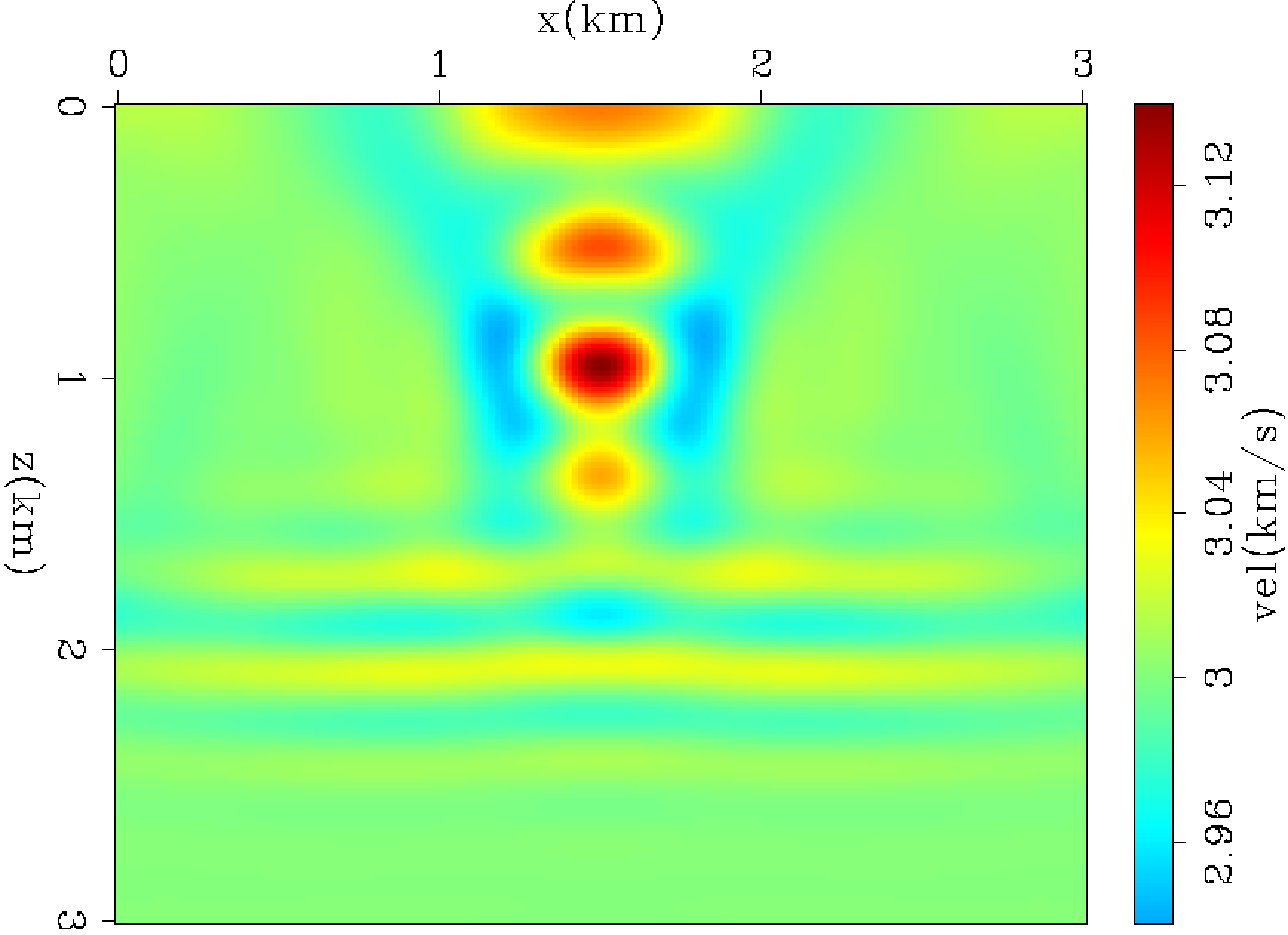
Inverted background with mixing



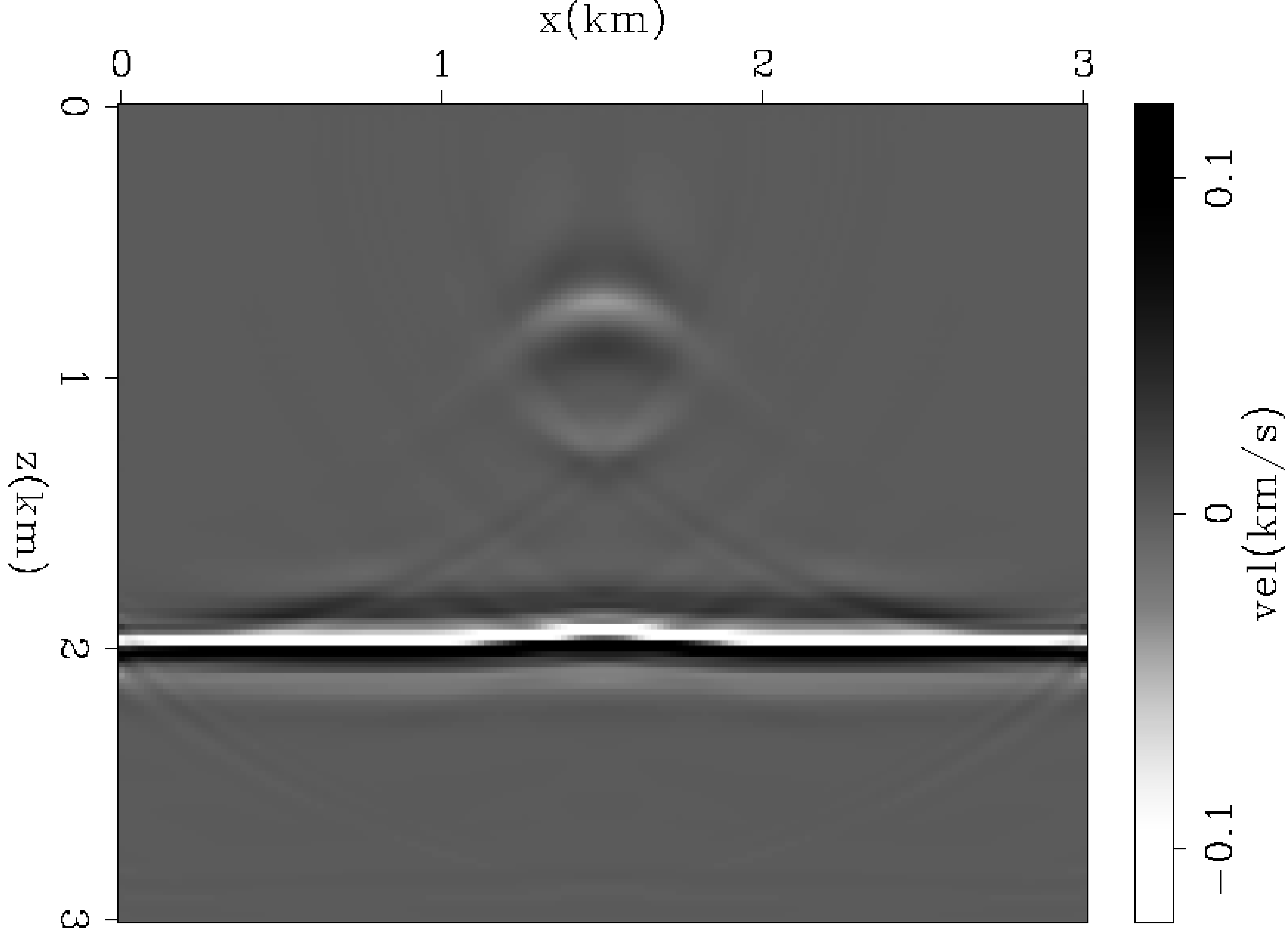
Inverted Born reflectivity with mixing



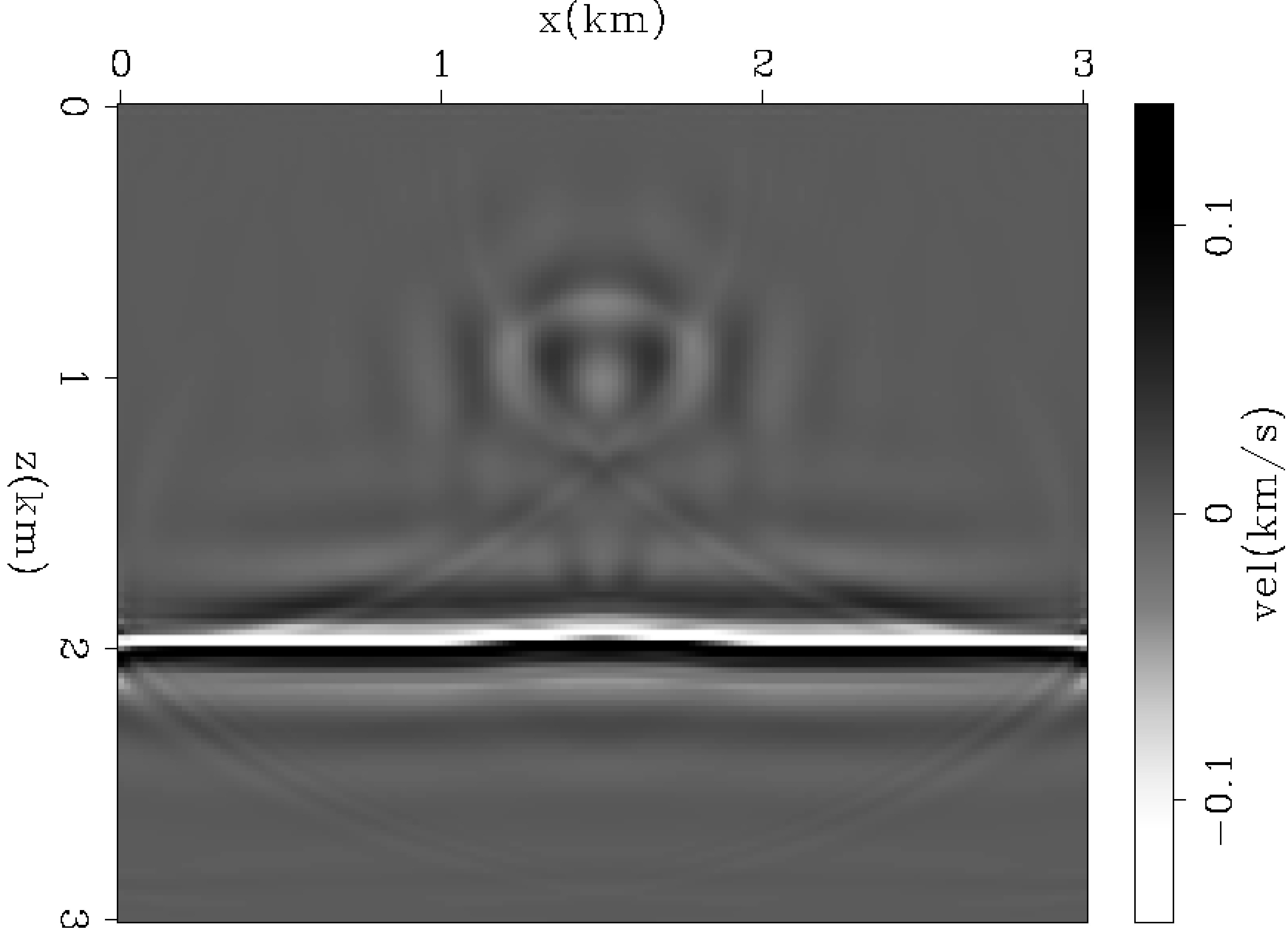
Inverted background



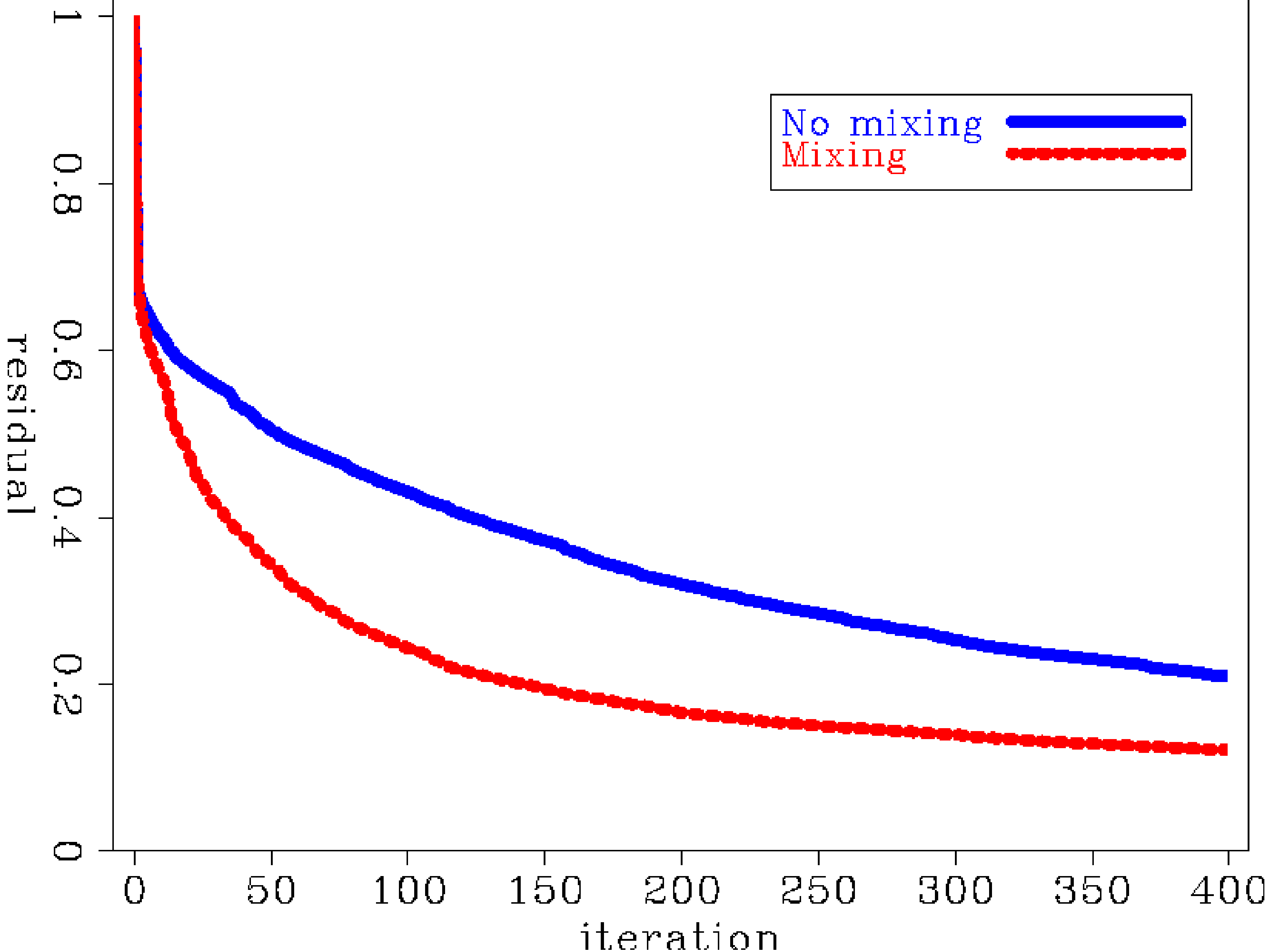
Inverted background with mixing



Inverted Born reflectivity



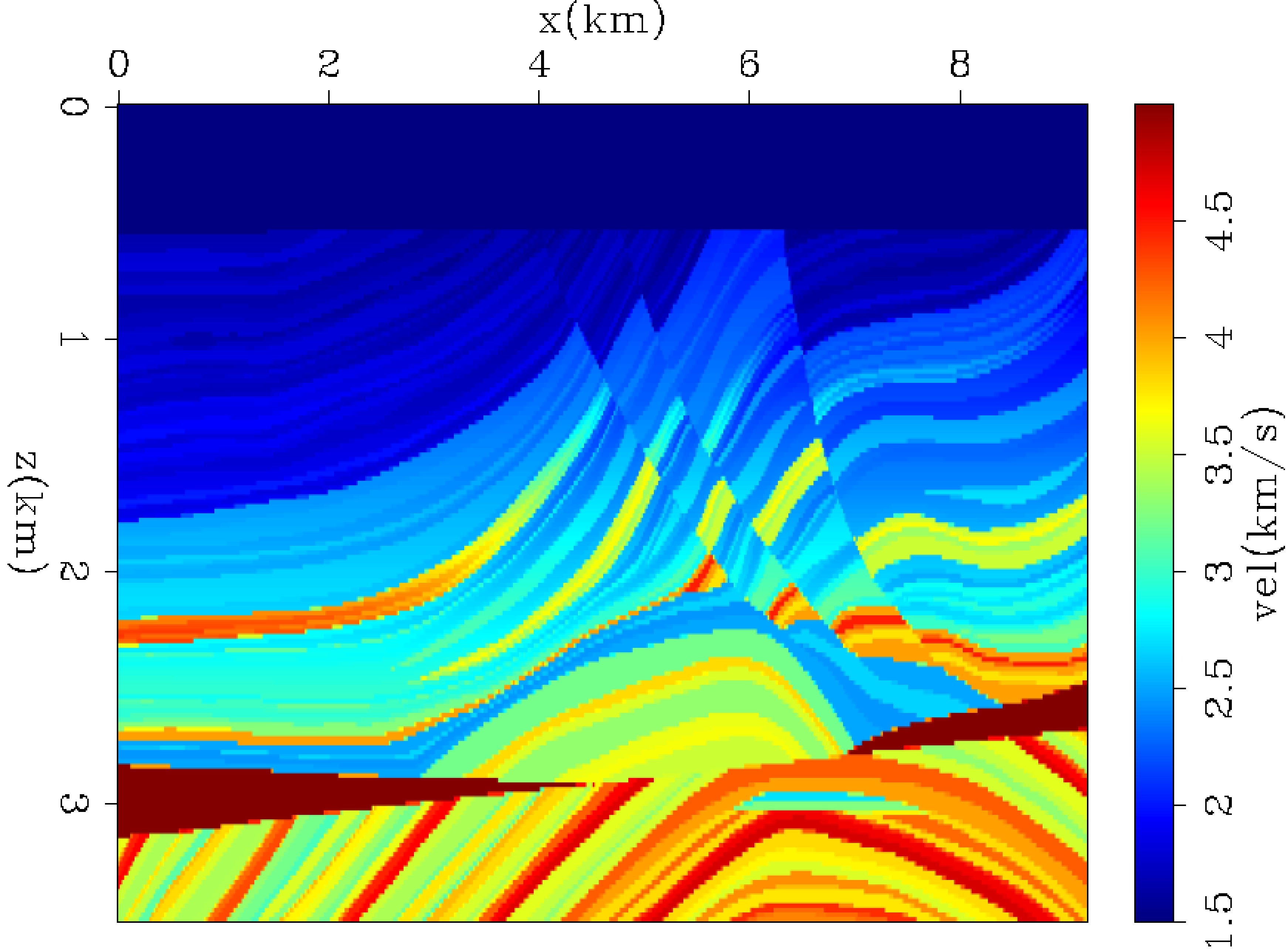
Inverted Born reflectivity with mixing



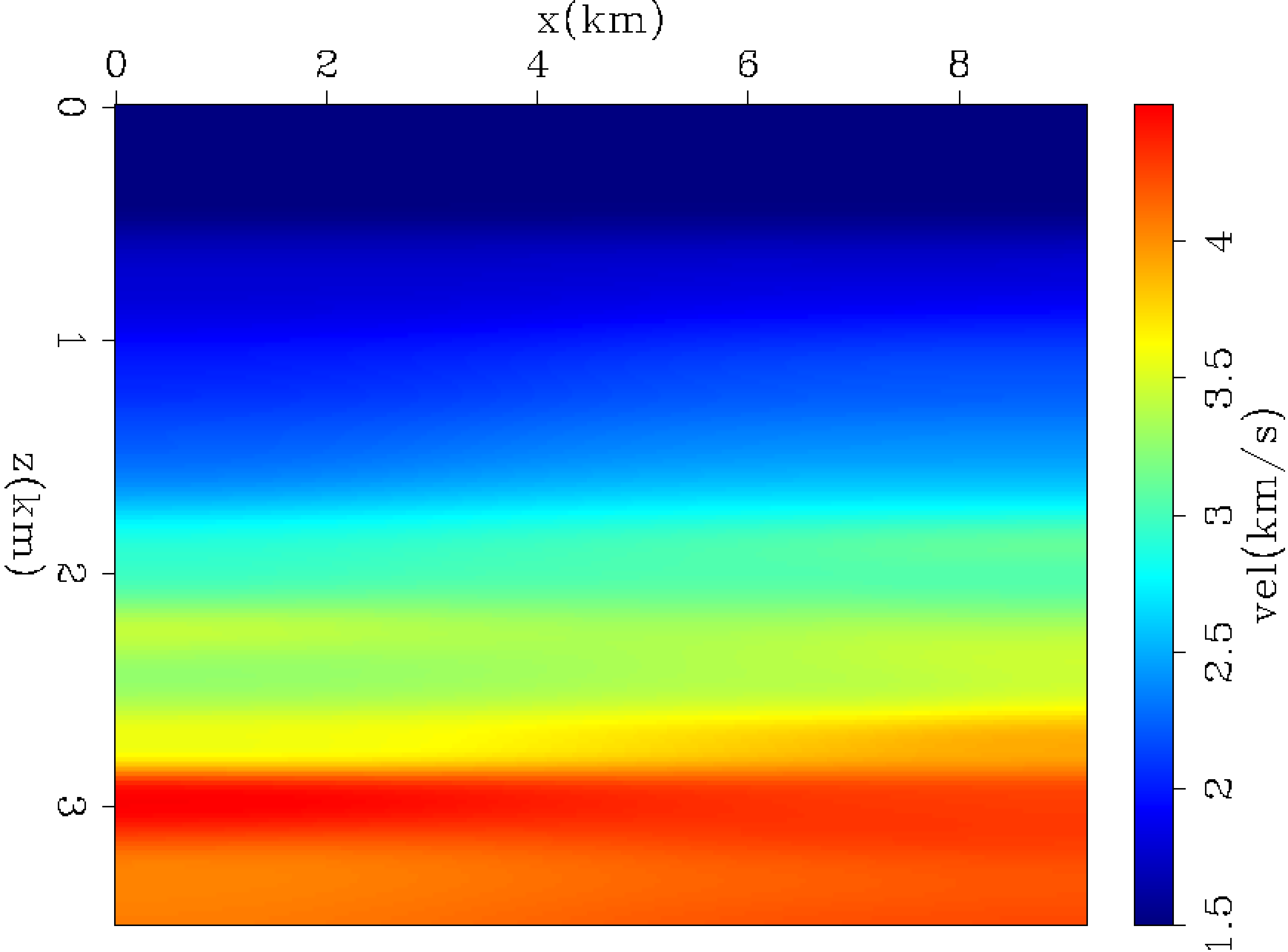
No mixing
Mixing
residual
iteration
0
0.2
0.4
0.6
0.8
1
0
50
100
150
200
250
300
350
400

Marmousi model

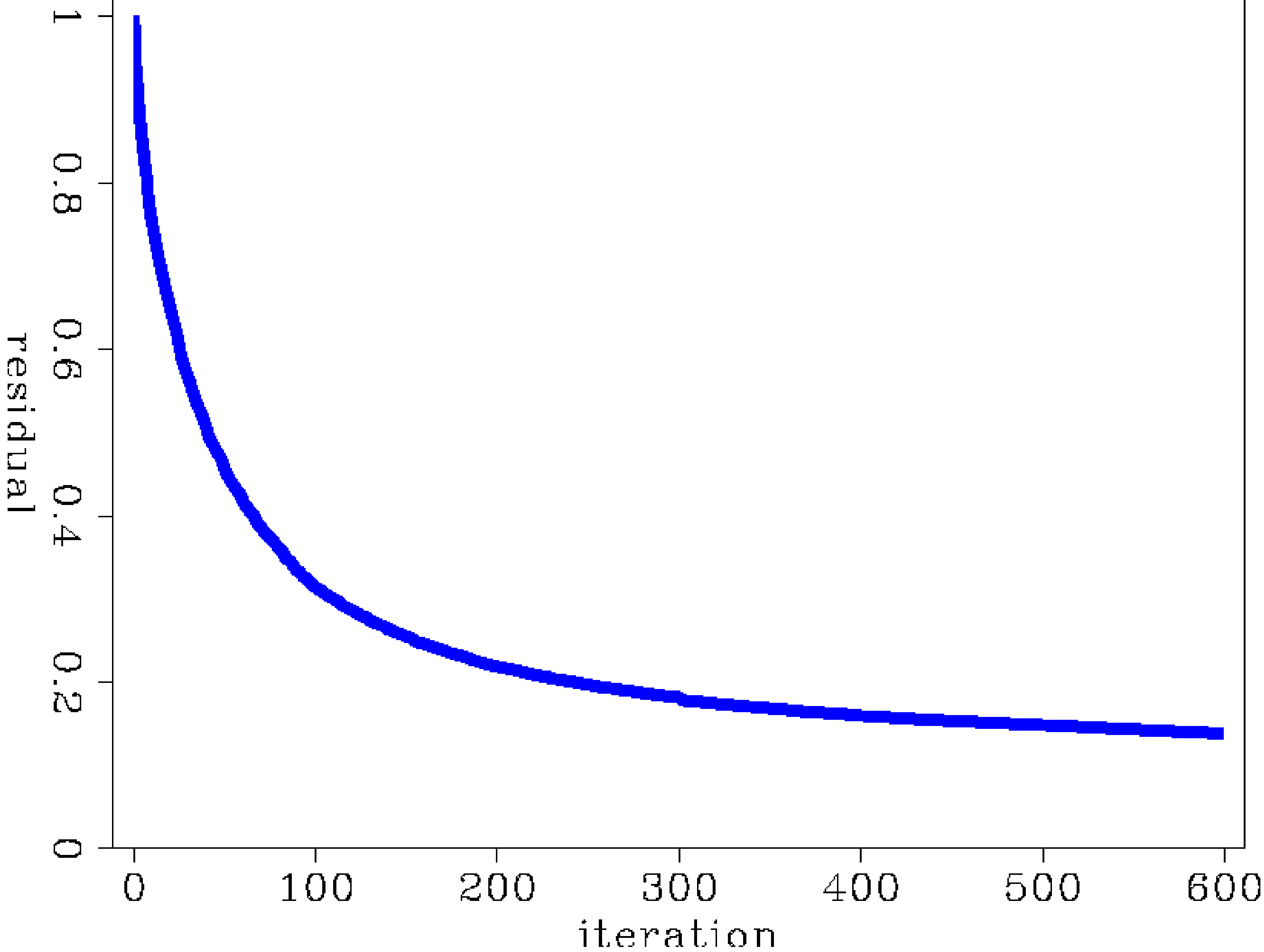
- Modified Marmousi model
  - Model size 3.5 km x 9.2 km
  - Ricker wavelet of 15 Hz
  - Fixed receiver spread, complete coverage
  - Source spacing 100m
  - Receiver spacing 20m

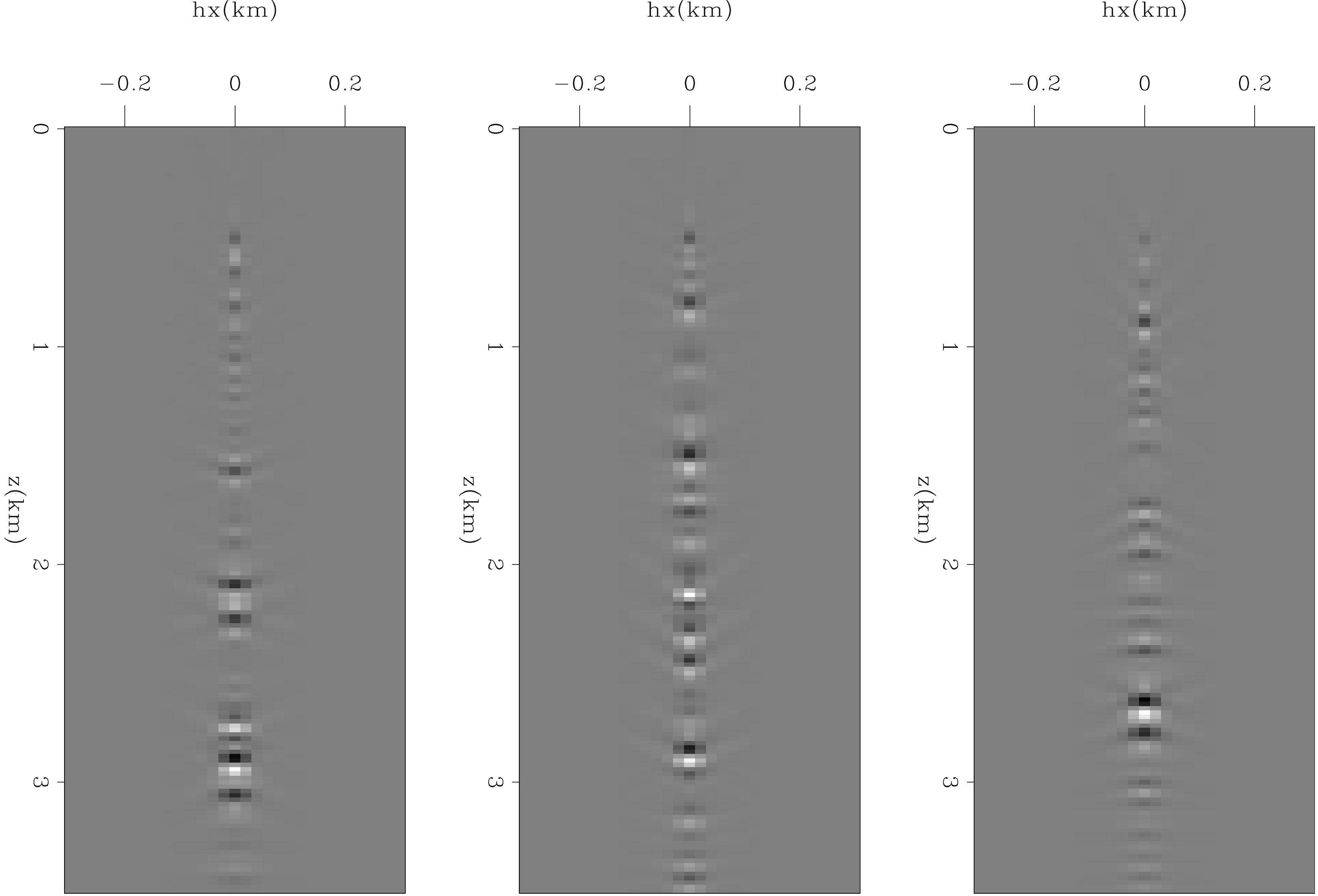


True Marmousi velocity

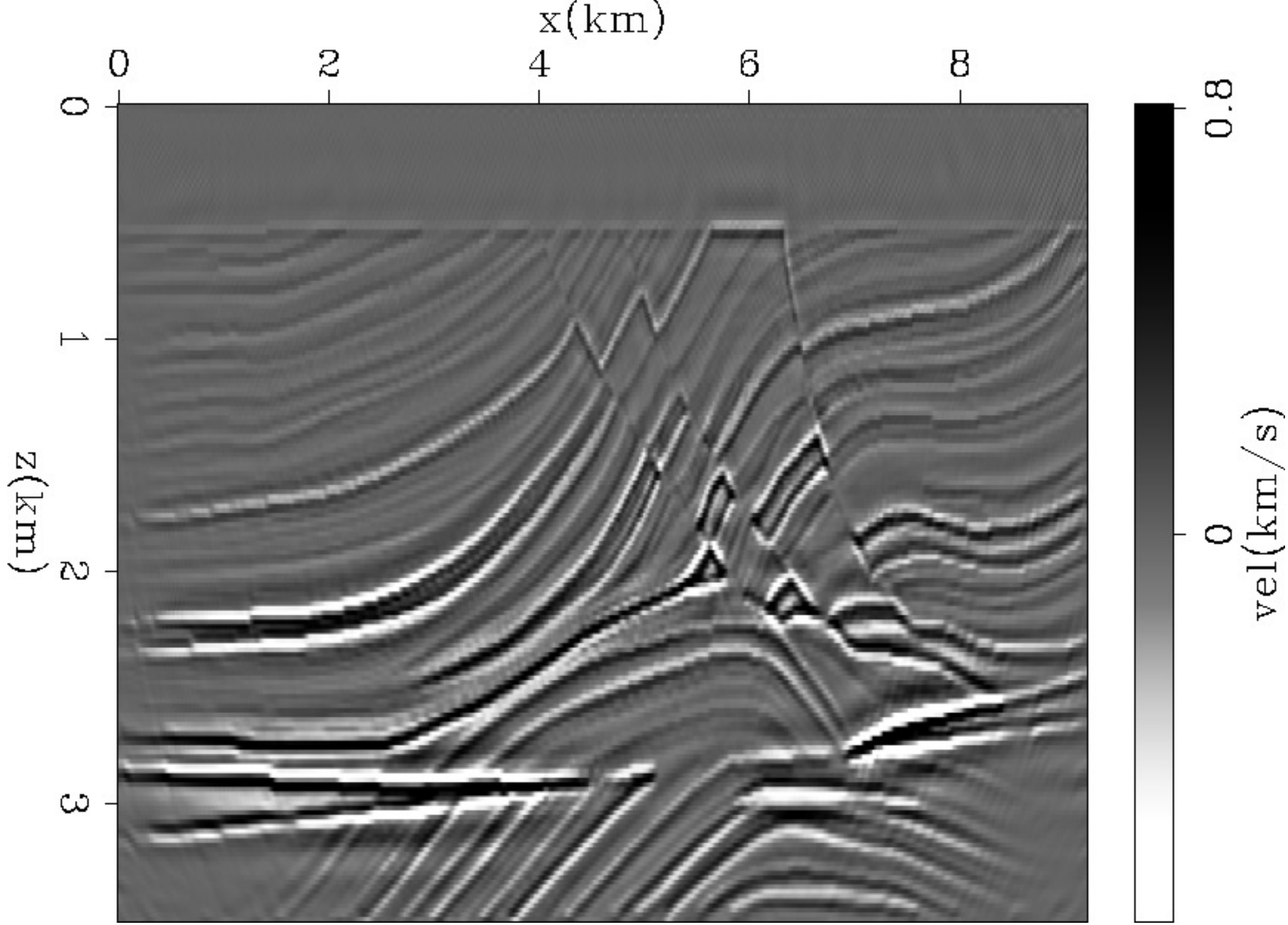


Initial velocity

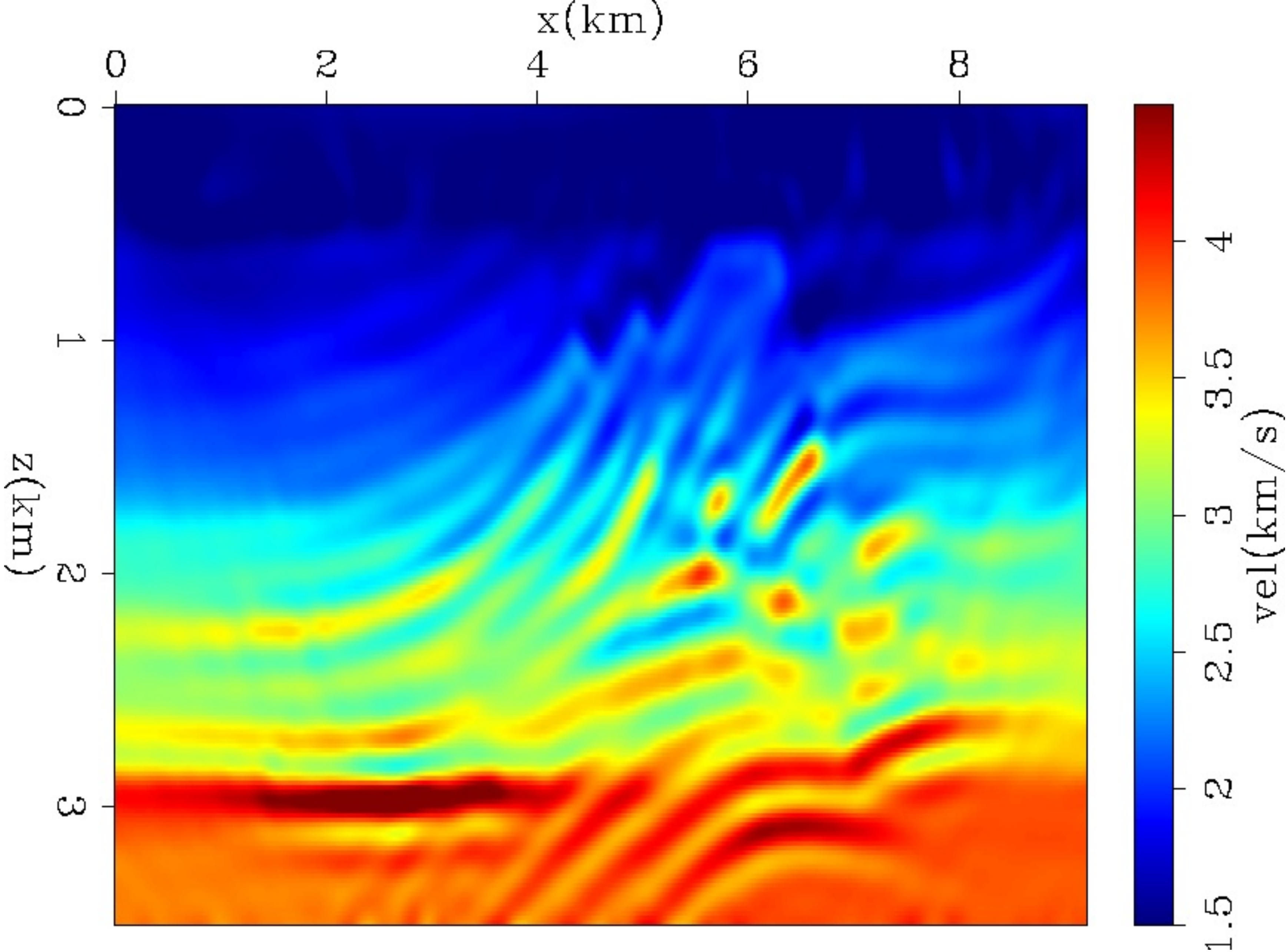

residual
1
0.8
0.6
0.4
0.2
0
0
100
200
300
400
500
600
iteration



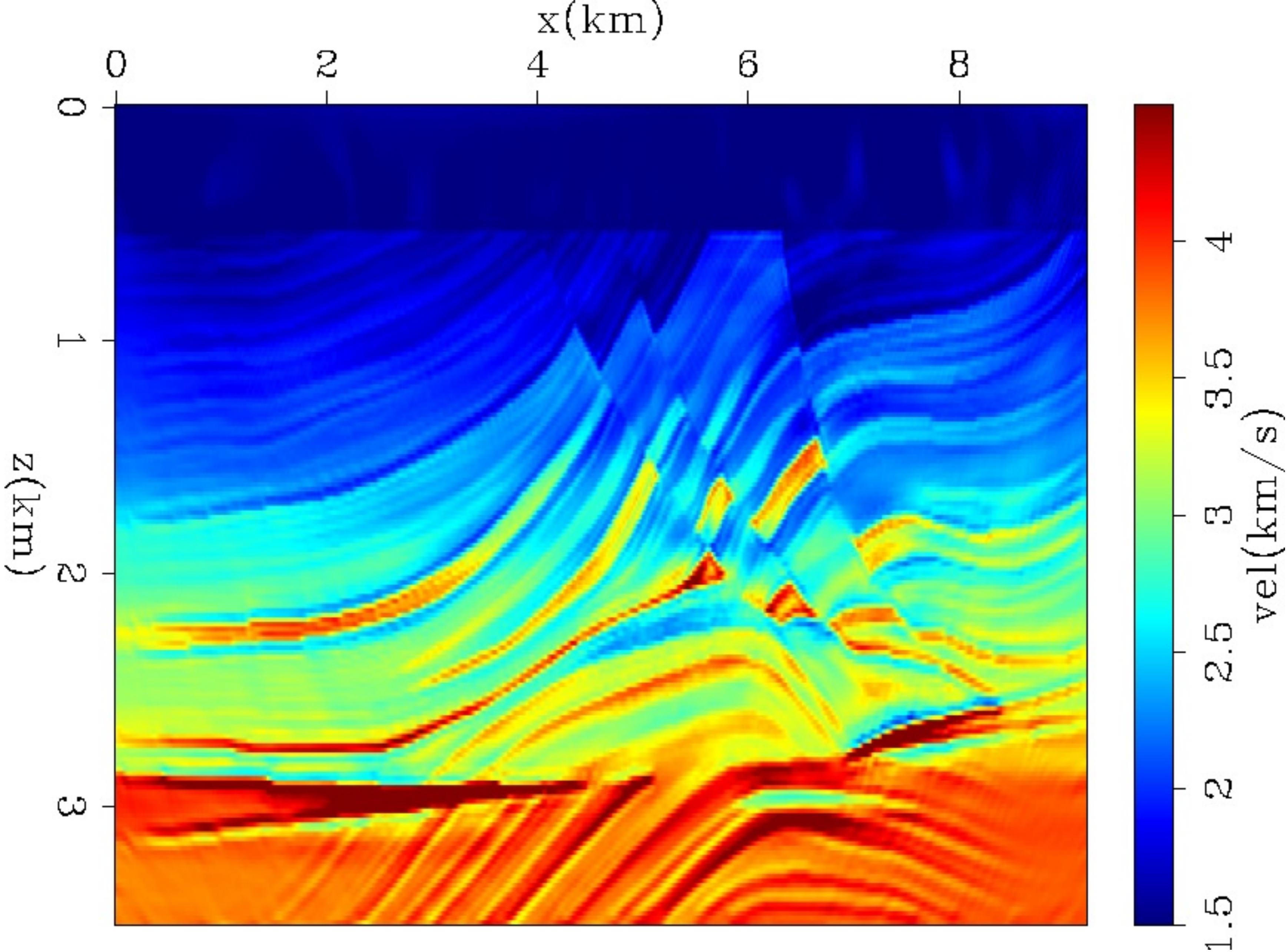
ODCIGs of inverted Born reflectivity



Inverted Born reflectivity with mixing



Inverted background with mixing



Total Inverted velocity with mixing

Outline

- Introduction
- Theory
  - Scale separation
  - Scale mixing
- Synthetic examples
  - Gaussian model
  - Marmousi model
- Conclusions

Conclusions

- TFWI is robust against initial model errors and gives high resolution estimates
- TWFI suffers from cost and practical issues
- By using Born approximation, we cut the cost by separating the model into background and reflectivity

Conclusions

- We only use primaries of the data
- The model separation is hinders the simultaneous inversion of scales
- We regain the high resolution results by scale mixing of parameters in Fourier domain

Future work

- Precondition the gradients
- Improve the scale mixing by using better filters and more information
- Relax or eliminate the primary data assumption
- Implement in 3D

Acknowledgment

- SEP sponsors for financial support
- Saudi Aramco for supporting my studies

Thanks!