Primary estimation with sparsity-promoting bi-convex optimization

Tim Tai-Yi Lin

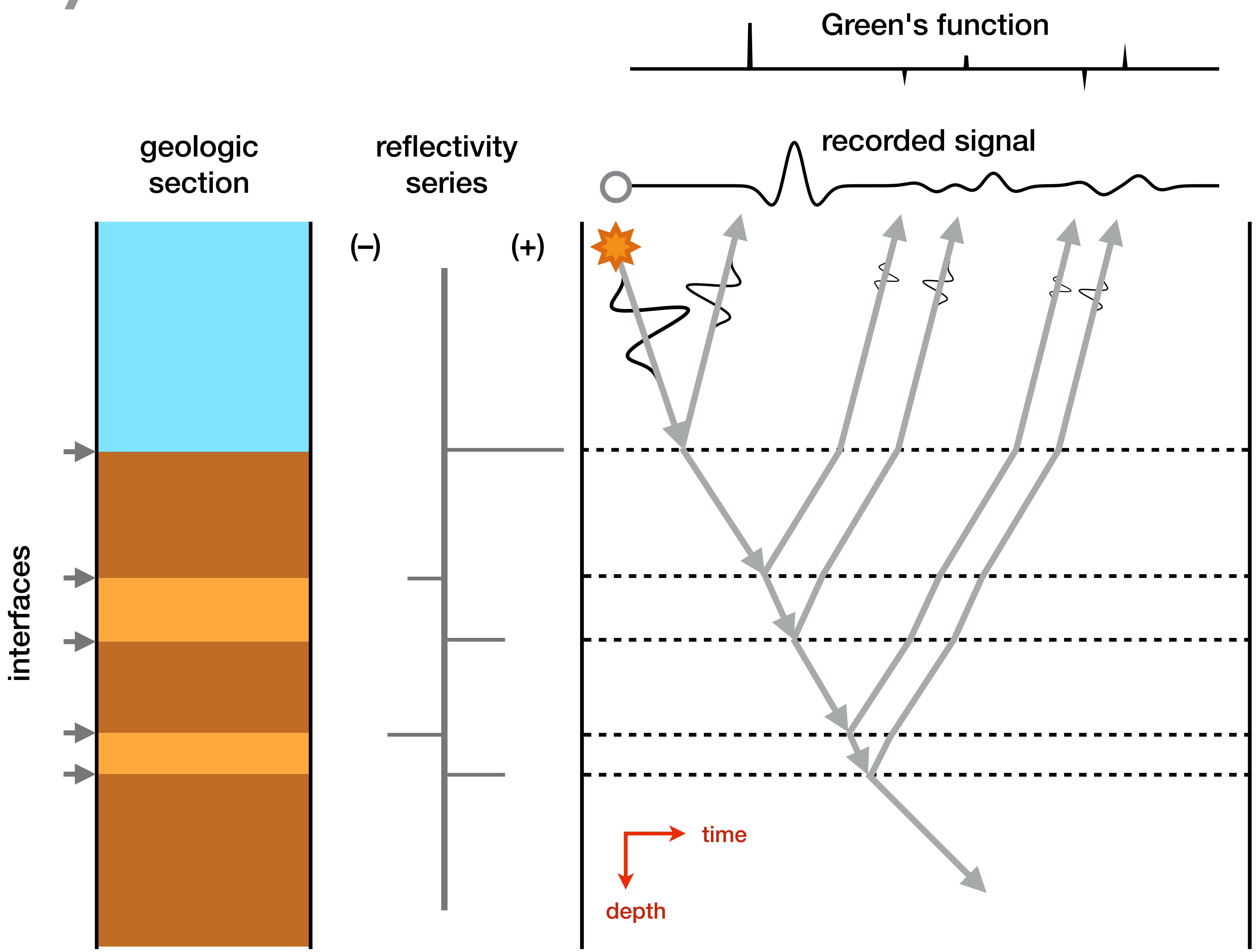
Ph.D. Final Oral Defence
October 1, 2015

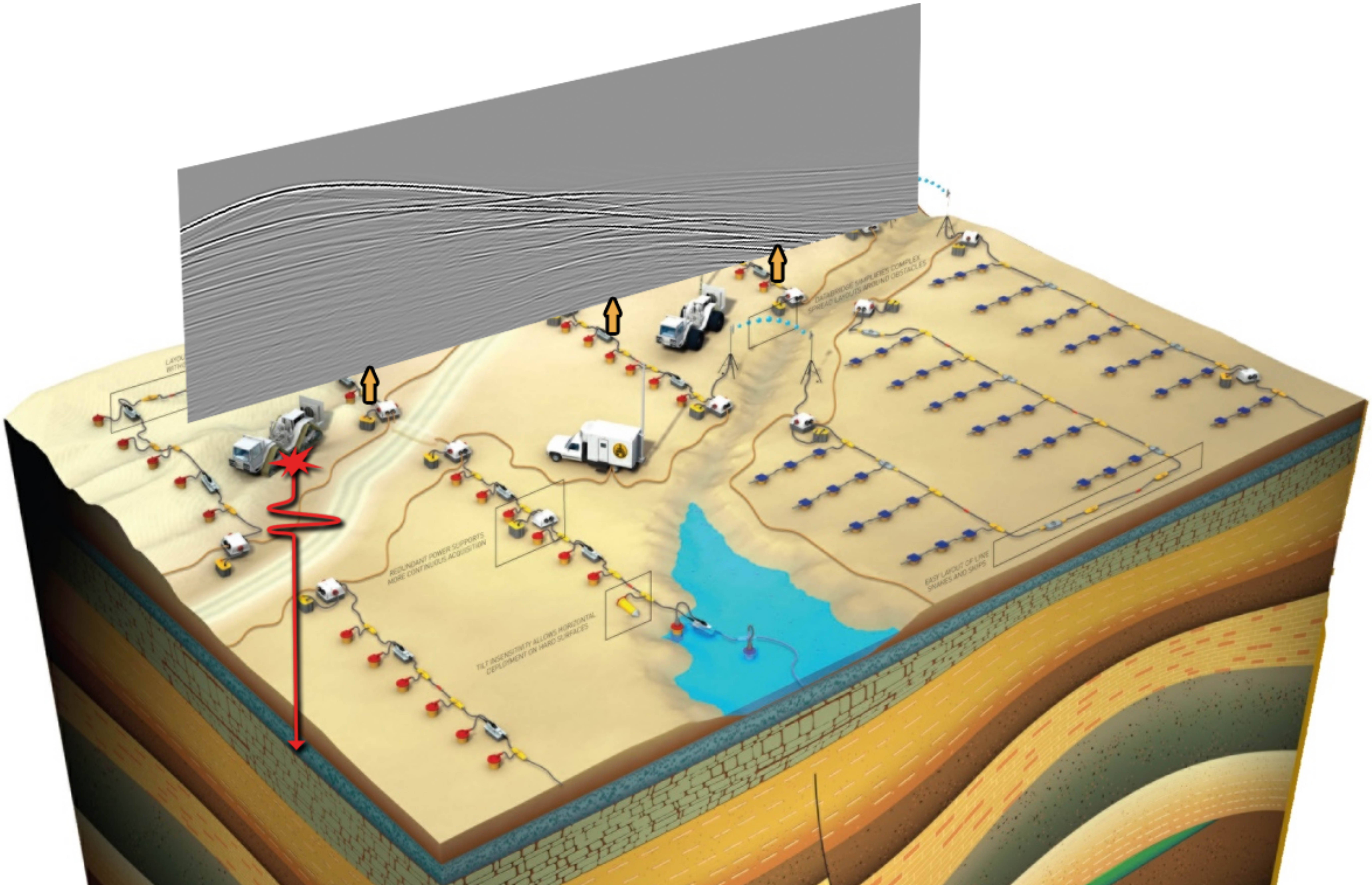
SLIM
University of British Columbia

Introduction

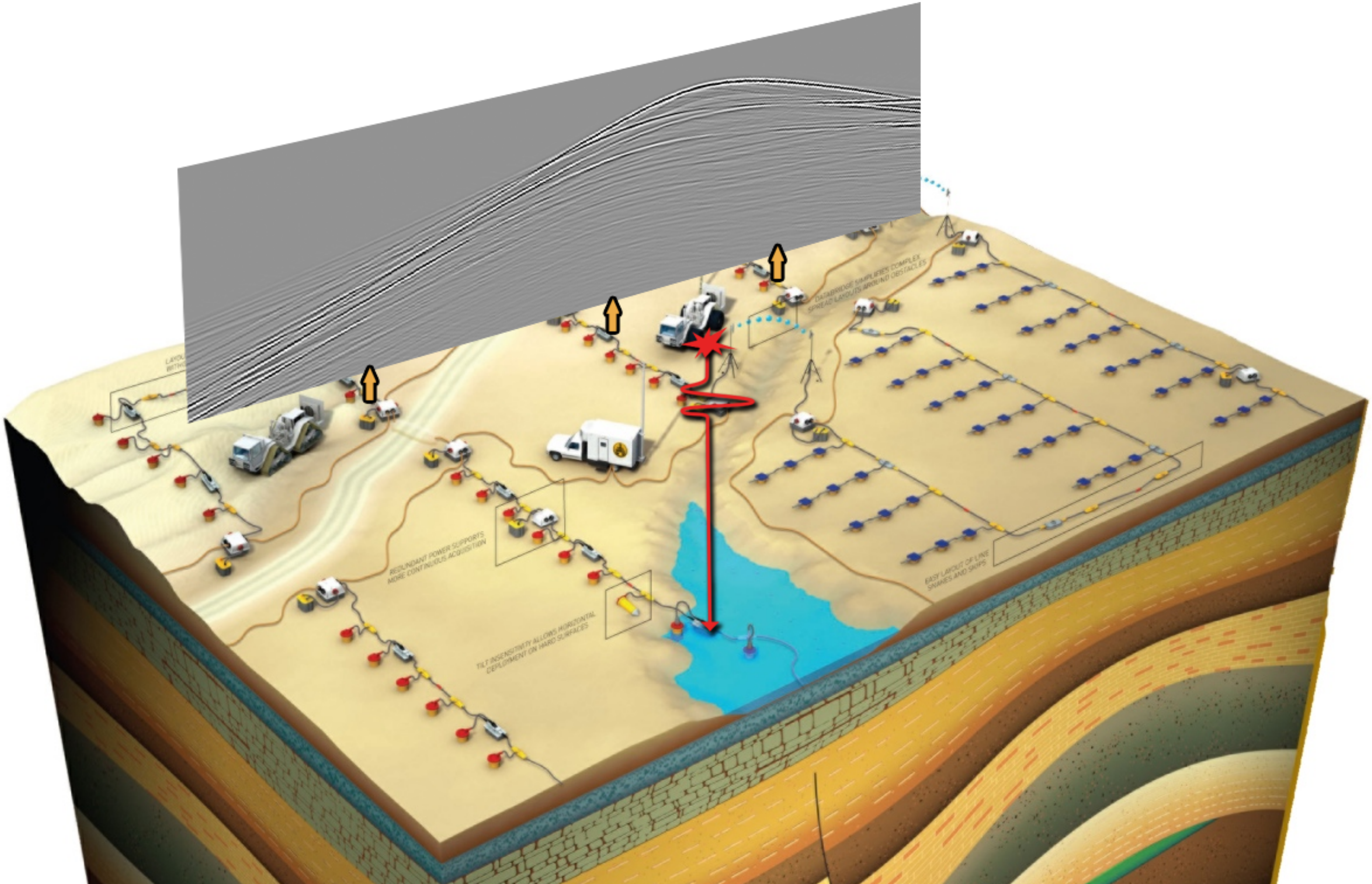
Active reflection seismic surveys

Reflectivity model

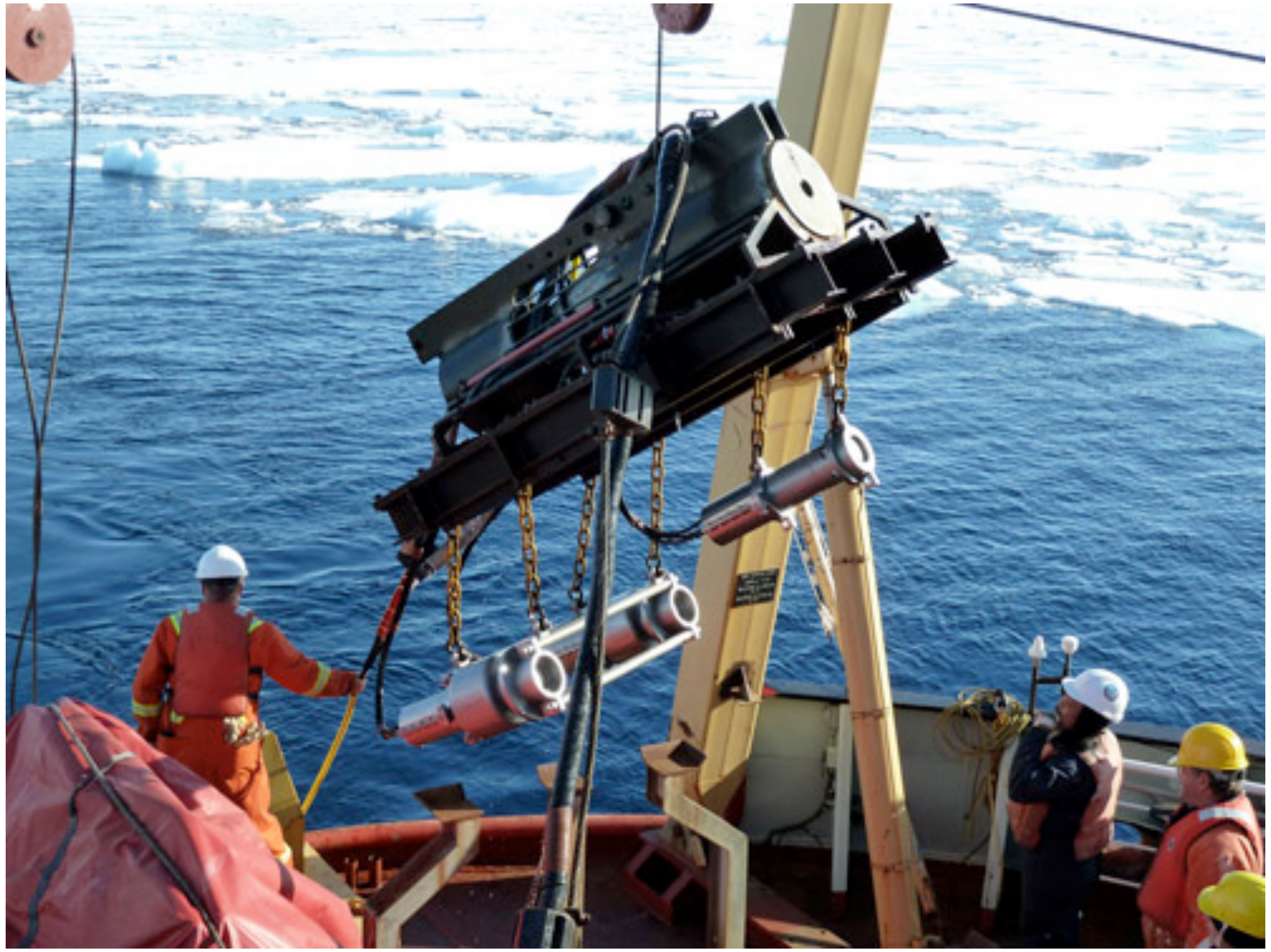




DATABRIDGE SIMPLIFIES COMPLEX
SPREAD LAYOUTS AROUND OBSTACLES
REDUNDANT POWER SUPPORTS
MORE CONTINUOUS ACQUISITION
EASY LAYOUT OF LINE
SNAKES AND SKIPS
TILT INSENSITIVITY ALLOWS HORIZONTAL
DEPLOYMENT ON HARD SURFACES


DATABRIDGE SIMPLIFIES COMPLEX SPREAD LAYOUTS AROUND OBSTACLES
REDUNDANT POWER SUPPORTS MORE CONTINUOUS ACQUISITION
TILT INSENSITIVITY ALLOWS HORIZONTAL DEPLOYMENT ON HARD SURFACES
EASY LAYOUT OF LINE SNAKES AND SKIPS

Marine seismic survey



Air gun array

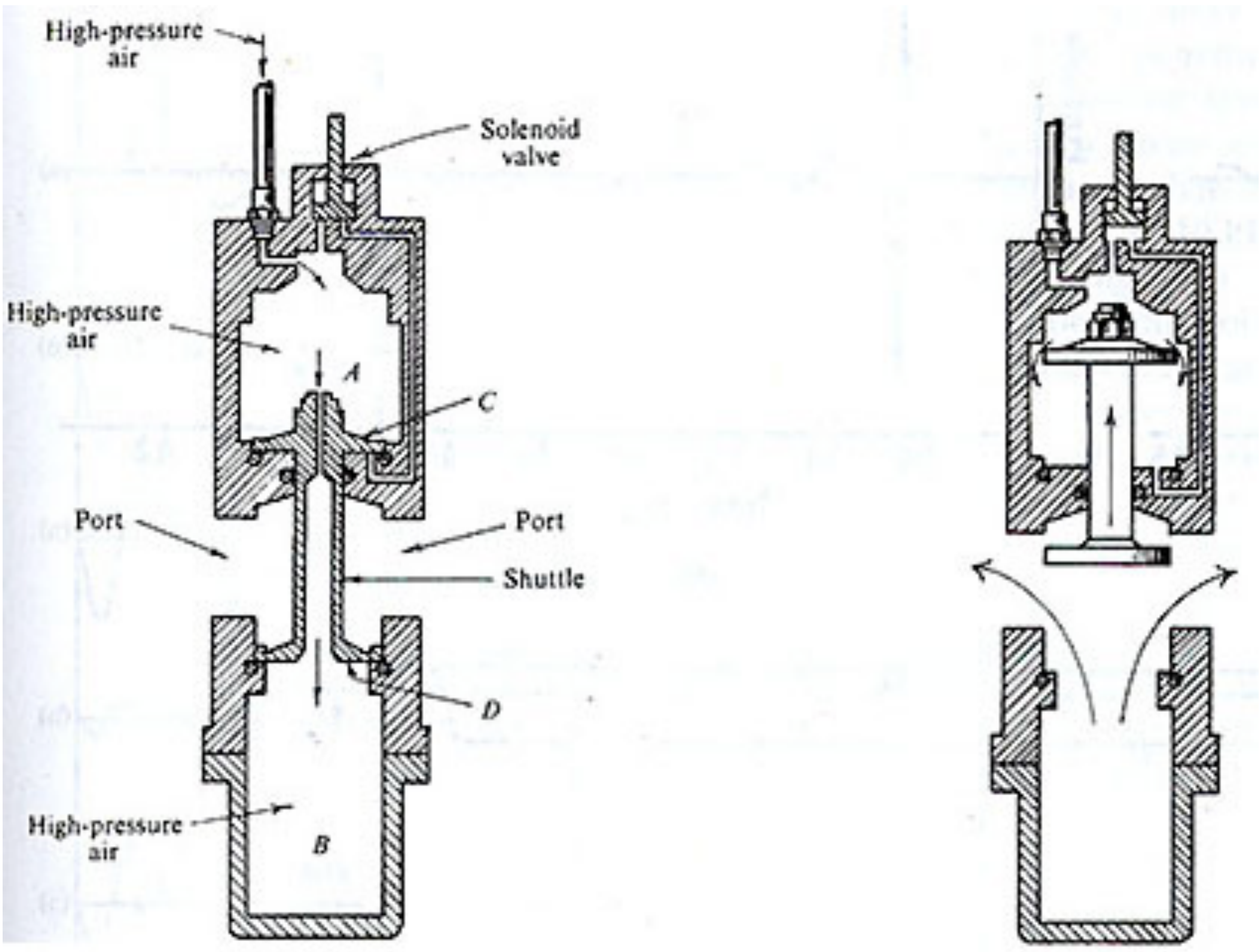
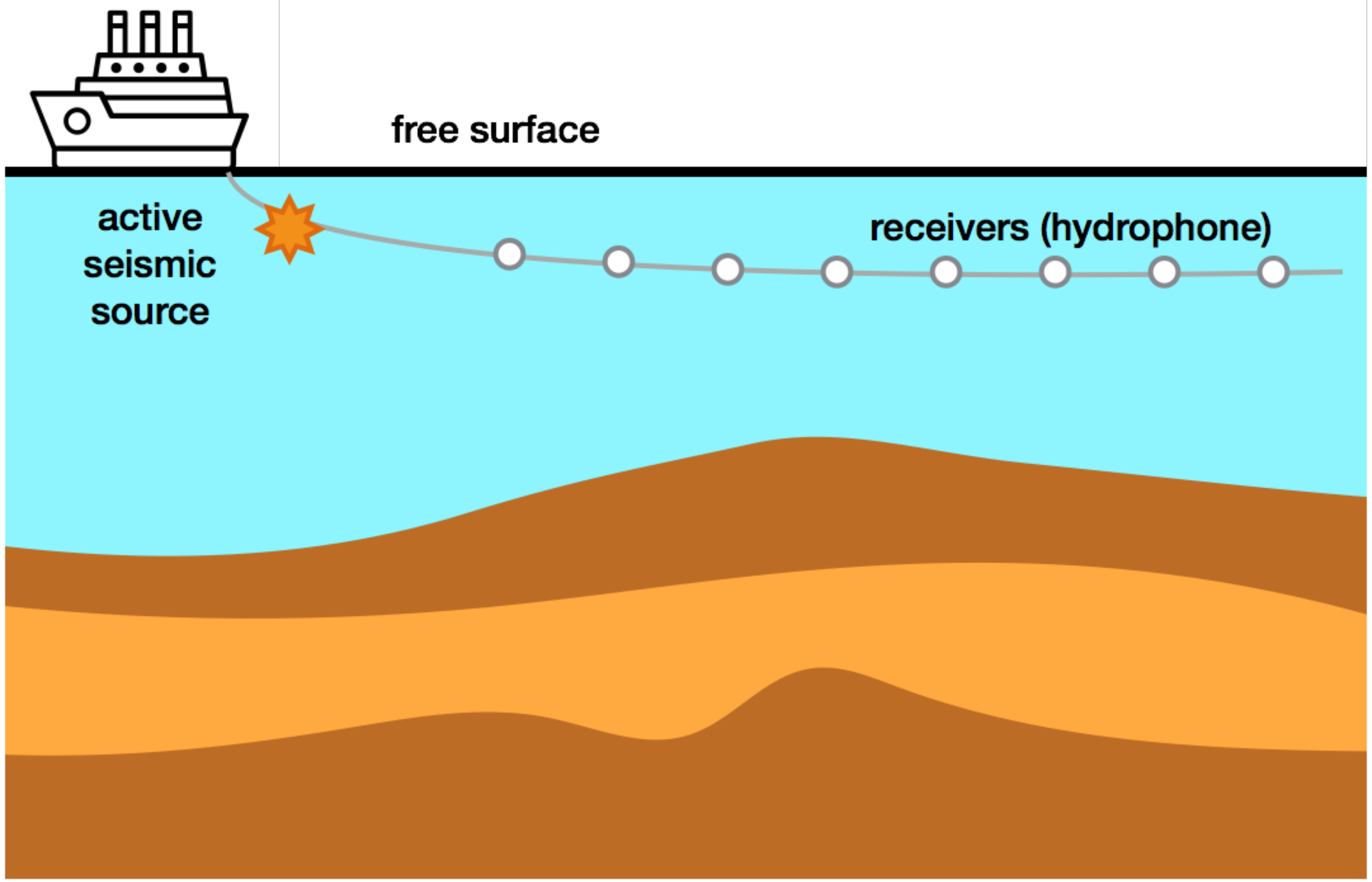


Figure 4. Schematic for air gun primed (left) and firing (right).

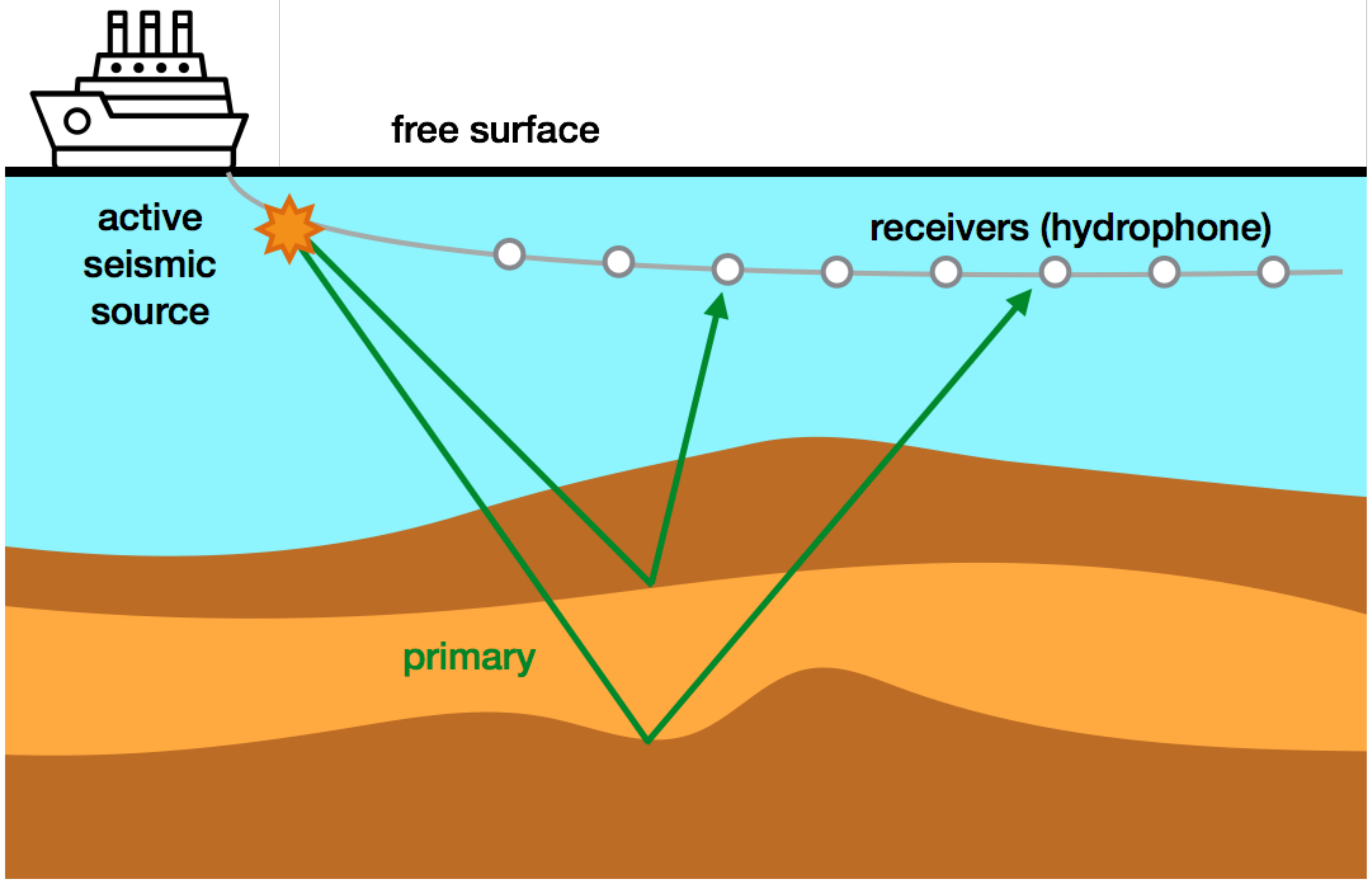
Marine seismic survey

Hydrophones

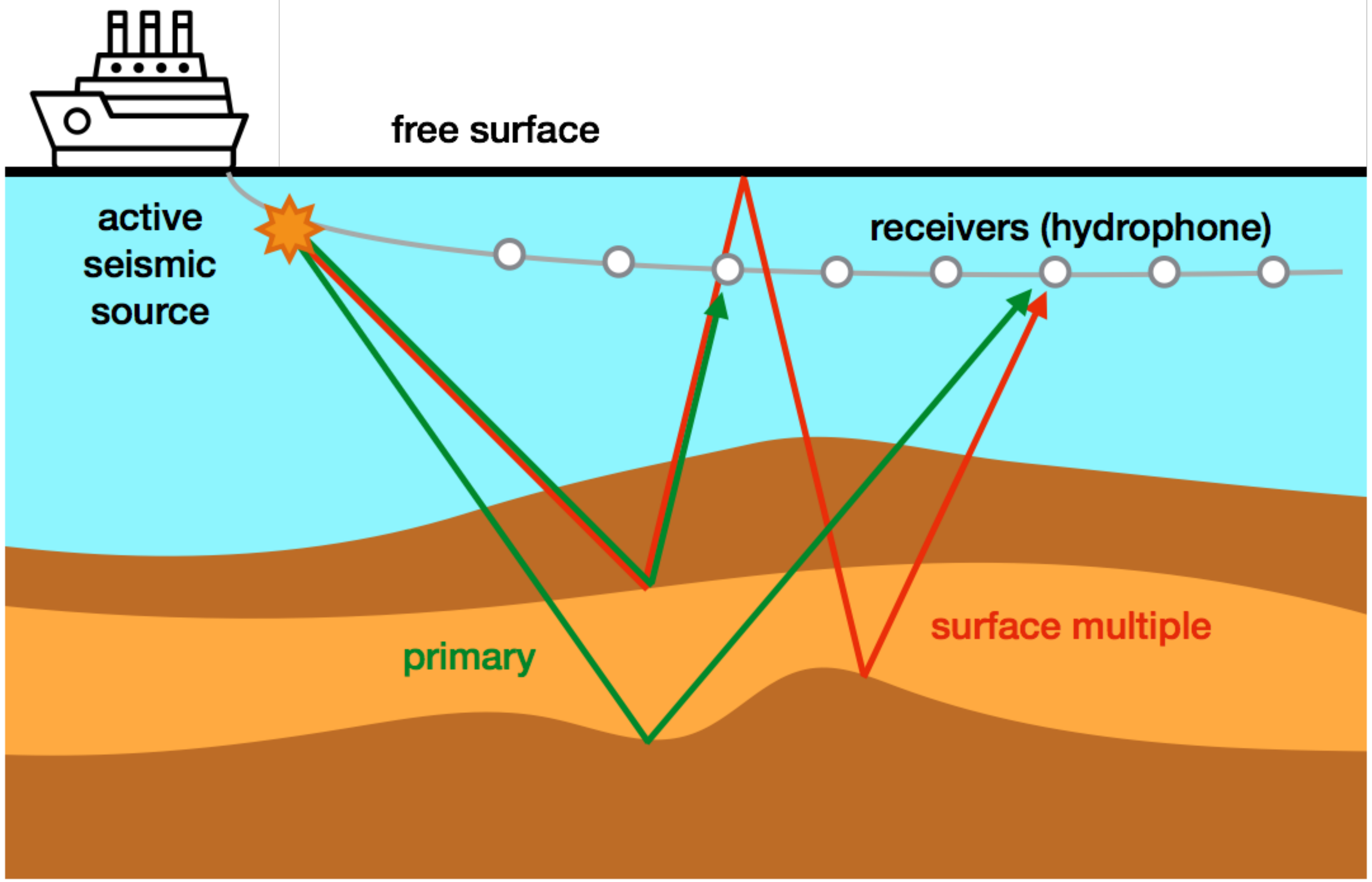
Marine seismic survey



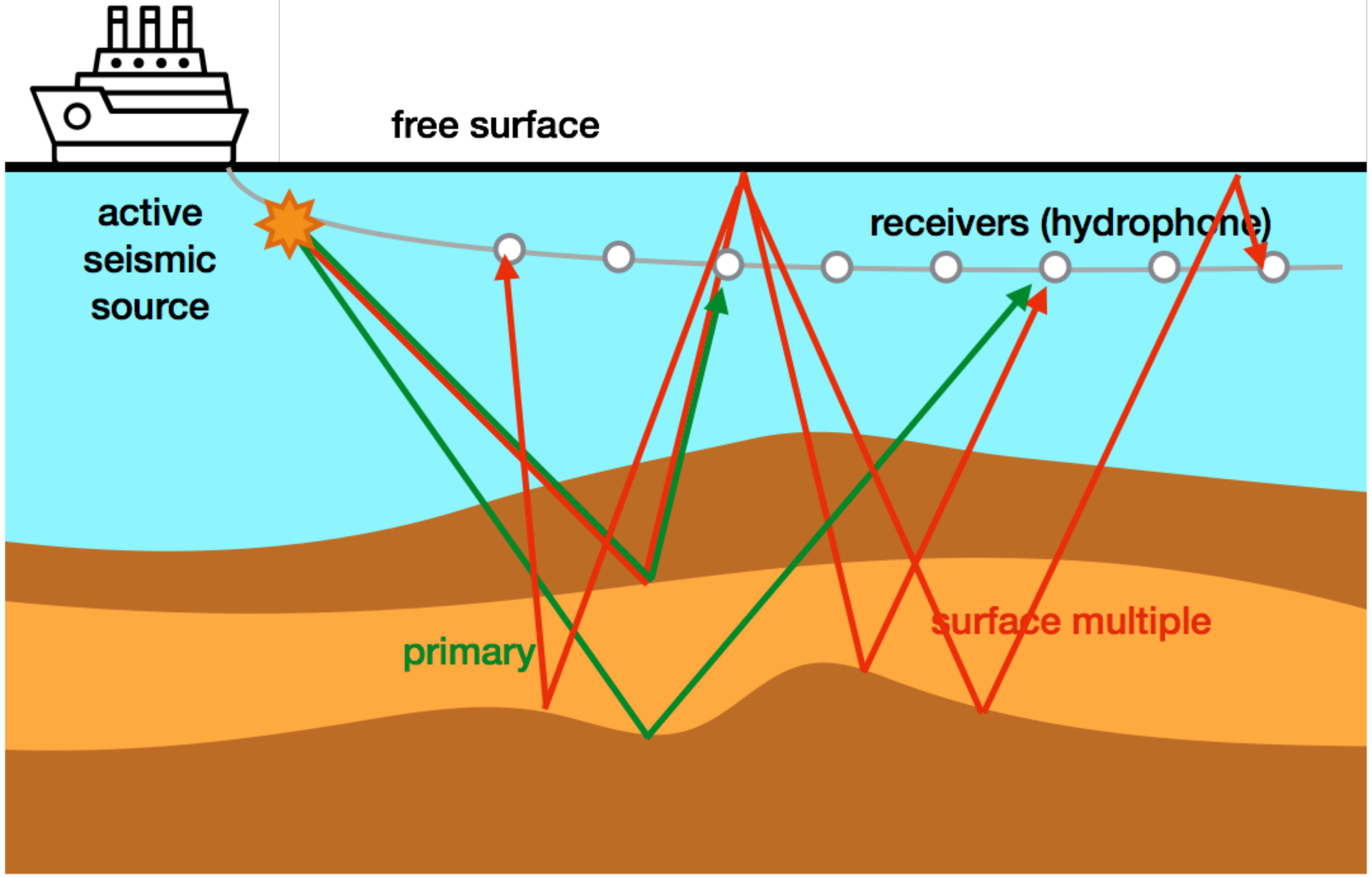
Marine seismic survey



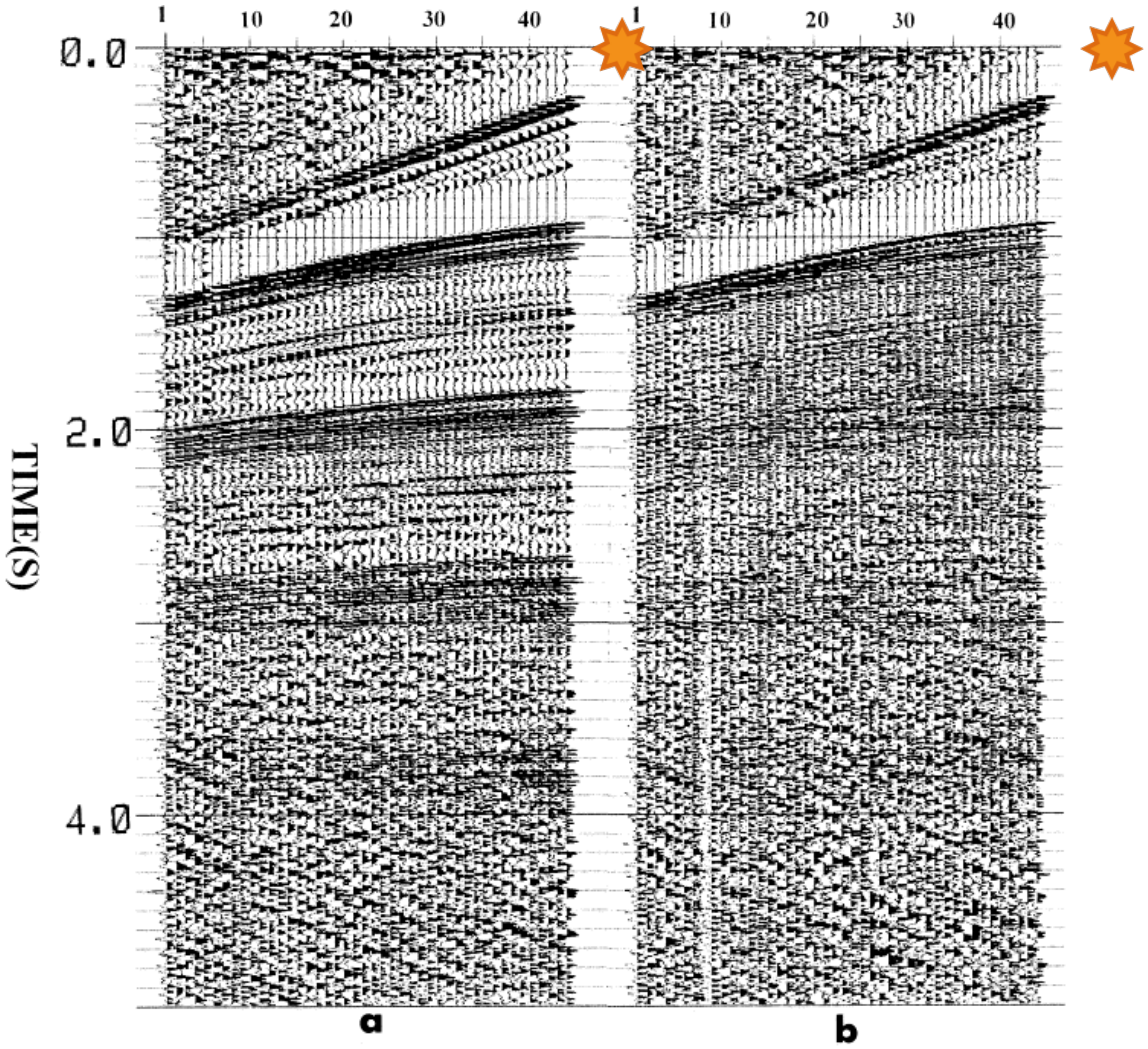
Marine seismic survey



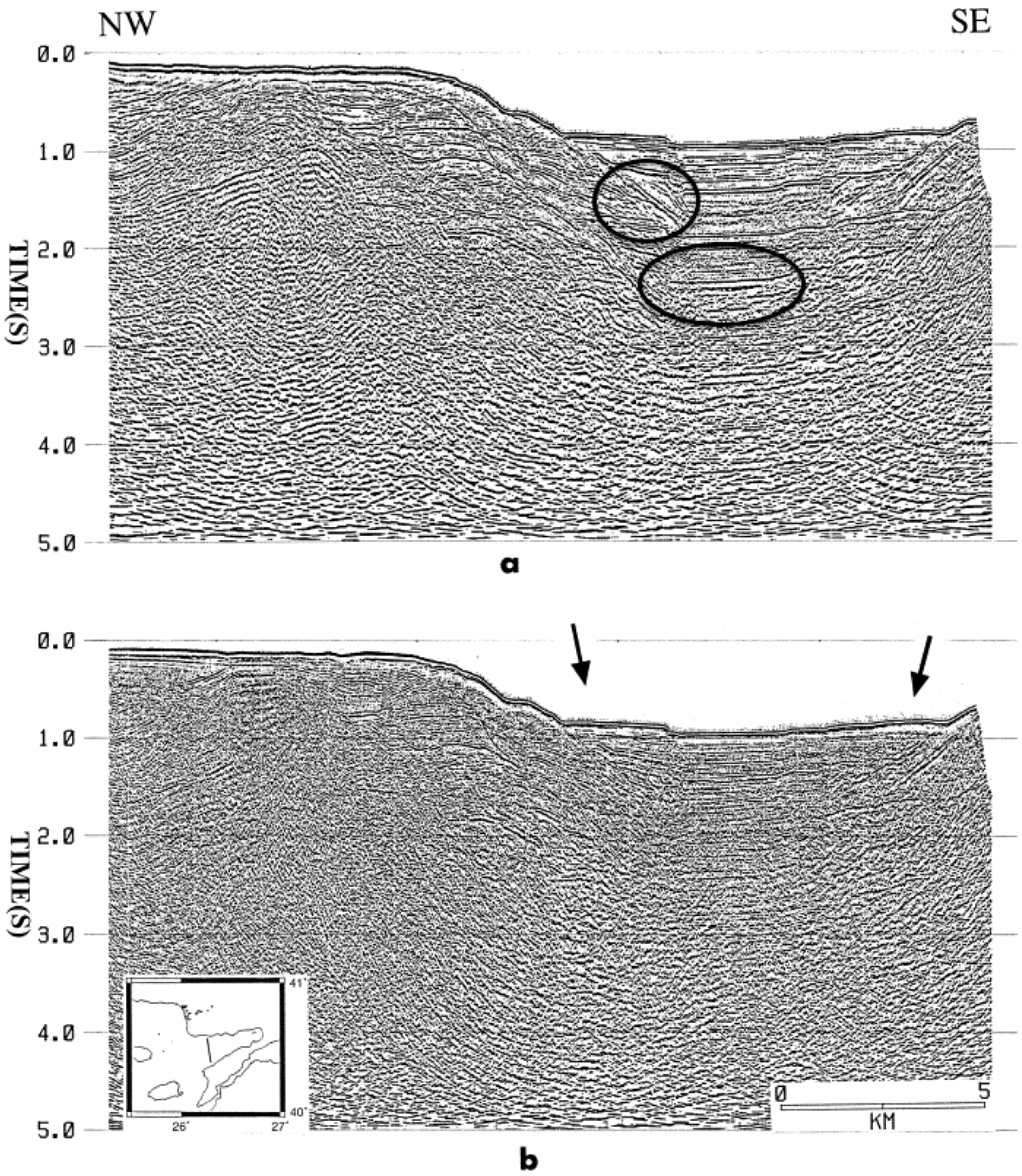
Marine seismic survey



Surface multiples

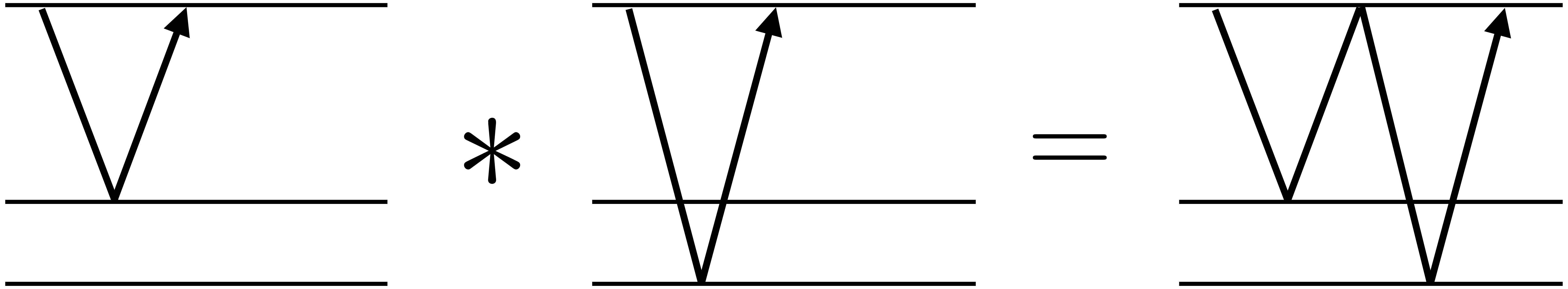


Surface multiples

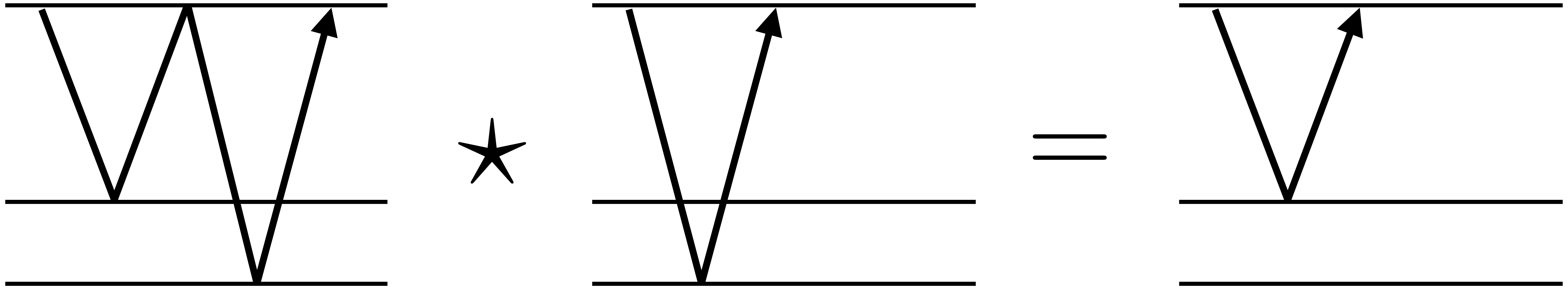


Modeling surface multiples by convolution

Concatenation of raypath under *convolution*



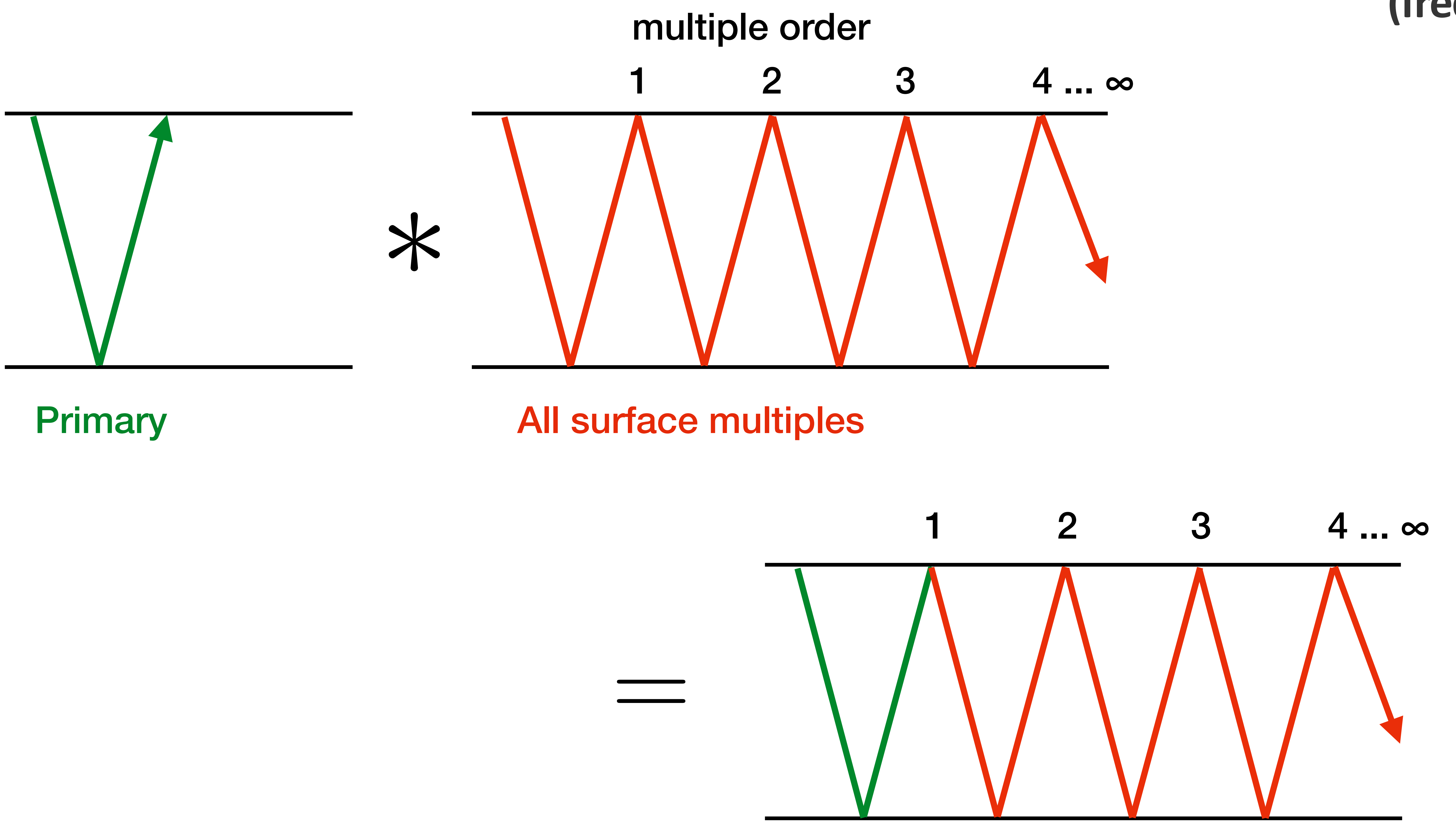
Subtraction of raypath under *cross-correlation*



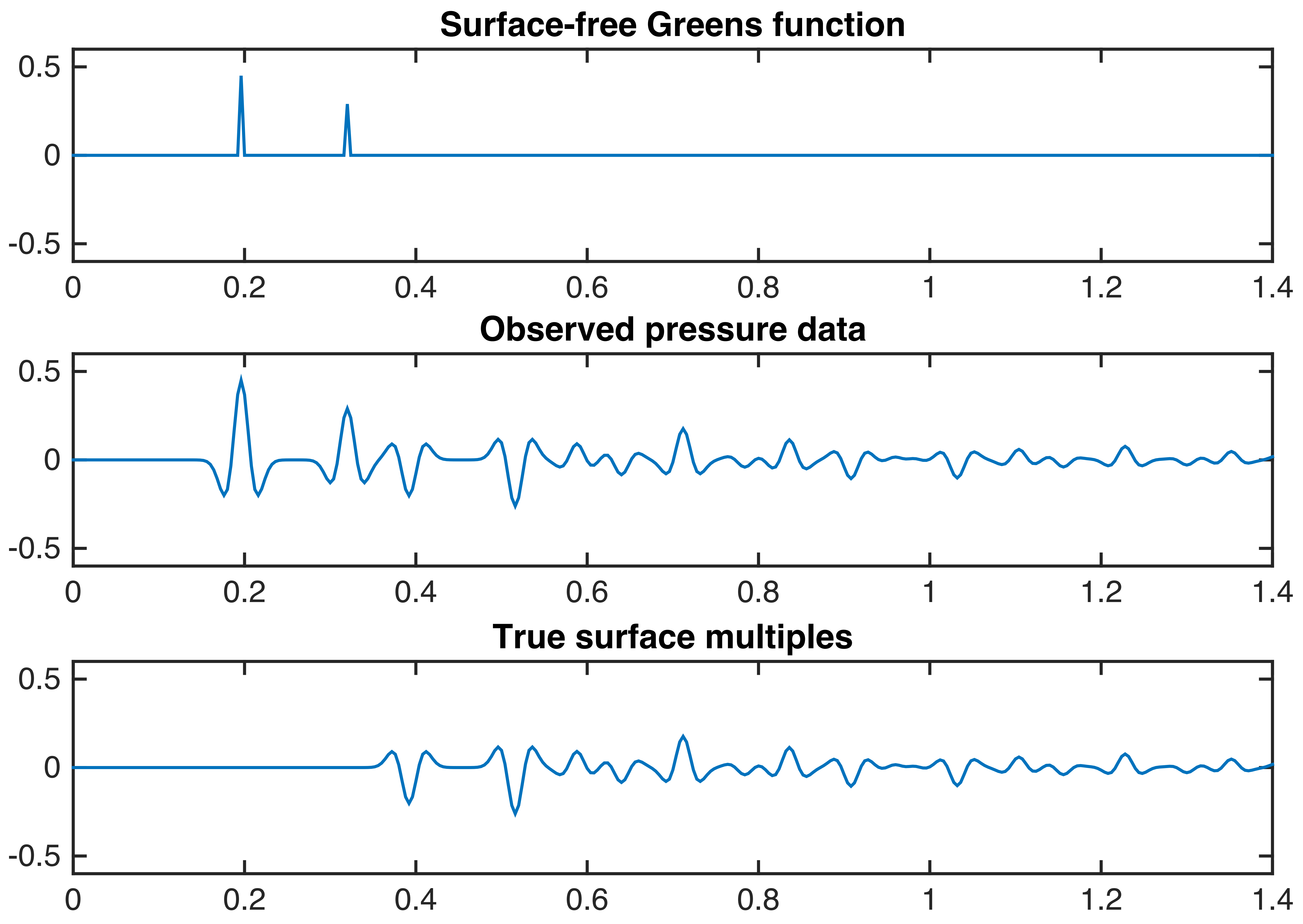
Modeling *all* surface multiples by convolution with *data*

$$p_m(\mathbf{x}, \omega; \mathbf{x}_{\mathrm{src}}) = \int_S g(\mathbf{x}, \omega; \mathbf{x}')r(\mathbf{x}, \mathbf{x}')p(\mathbf{x}', \omega; \mathbf{x}_{\mathrm{src}})\, \mathrm{d}\mathbf{x}'$$

(frequency domain)



Modeling *all* surface multiples by convolution with *data*



A multidimensional convolution model for seismic data

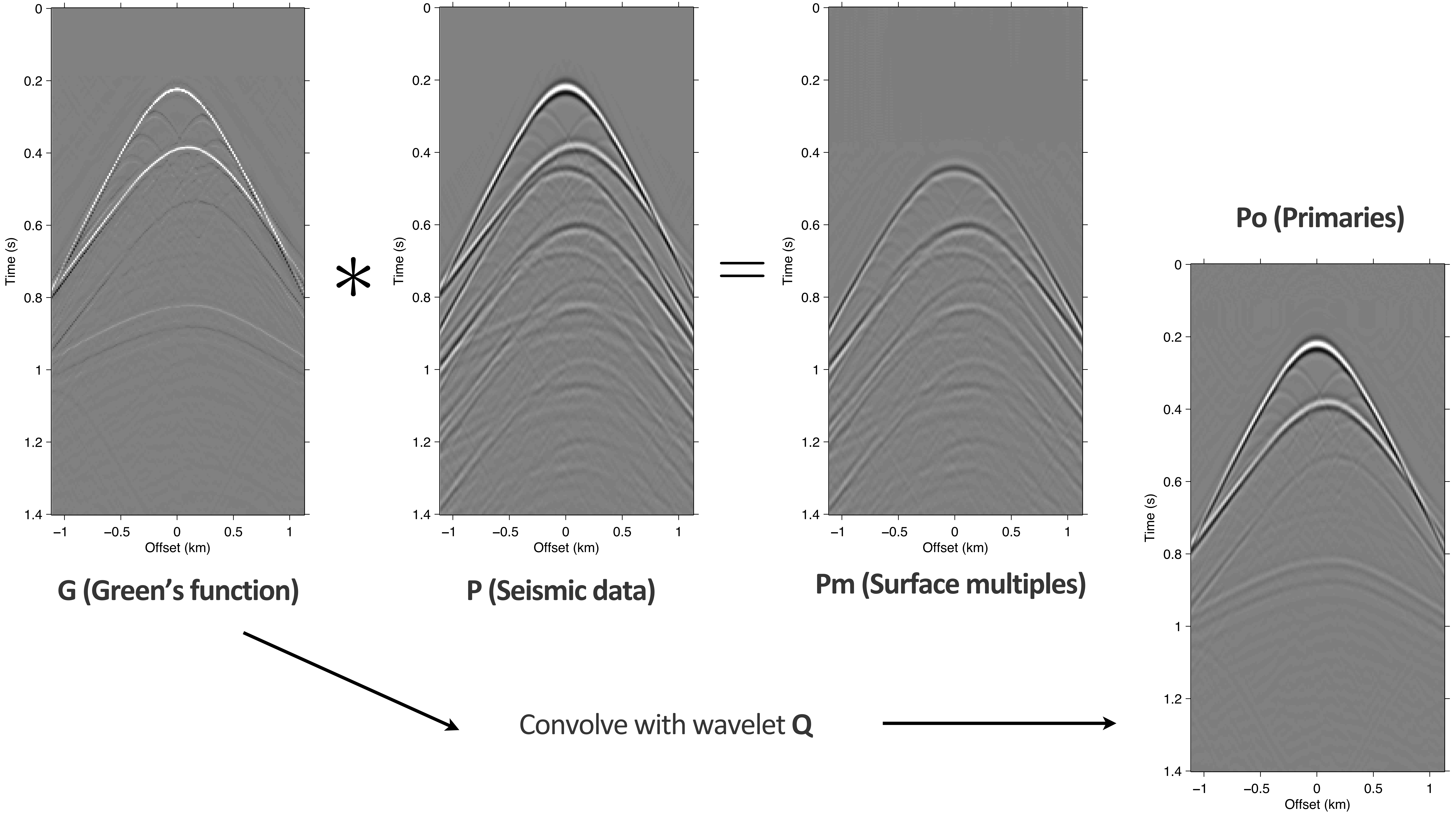
$$p_m(\mathbf{x}, \omega; \mathbf{x}_{\text{src}}) = \int_S g(\mathbf{x}, \omega; \mathbf{x}')r(\mathbf{x}, \mathbf{x}')p(\mathbf{x}', \omega; \mathbf{x}_{\text{src}})\, \mathrm{d}\mathbf{x}'$$

frequency domain

Primary model **Multiple model**

$$\begin{aligned} M(g, q; p) &:= \mathcal{F}_\omega^{-1}\left[g(\mathbf{x}, \omega; \mathbf{x}_{\text{src}})q(\omega) + \int_S g(\mathbf{x}, \omega; \mathbf{x}')r(\mathbf{x}, \mathbf{x}')p(\mathbf{x}', \omega; \mathbf{x}_{\text{src}})\, \mathrm{d}\mathbf{x}'\right] \\ &= p_o(\mathbf{x}, t; \mathbf{x}_{\text{src}}) + p_m(\mathbf{x}, t; \mathbf{x}_{\text{src}}) \\ &= p(\mathbf{x}, t; \mathbf{x}_{\text{src}}) \end{aligned}$$

Total data model


G (Green's function)
*
P (Seismic data)
=
Pm (Surface multiples)
Convolve with wavelet Q
Po (Primaries)
Time (s)
Offset (km)

A multidimensional convolution model for seismic data

SLIM

$$p_m(\mathbf{x},\omega;\mathbf{x}_{\text{src}}) = \int_S g(\mathbf{x},\omega;\mathbf{x}')r(\mathbf{x},\mathbf{x}')p(\mathbf{x}',\omega;\mathbf{x}_{\text{src}})\,\mathrm{d}\mathbf{x}'$$

frequency domain

Primary model **Multiple model**

$$\begin{aligned} M(g,q;p) &:= \mathcal{F}_\omega^{-1}\left[g(\mathbf{x},\omega;\mathbf{x}_{\text{src}})q(\omega) + \int_S g(\mathbf{x},\omega;\mathbf{x}')r(\mathbf{x},\mathbf{x}')p(\mathbf{x}',\omega;\mathbf{x}_{\text{src}})\,\mathrm{d}\mathbf{x}'\right] \\ &= p_o(\mathbf{x},t;\mathbf{x}_{\text{src}}) + p_m(\mathbf{x},t;\mathbf{x}_{\text{src}}) \\ &= p(\mathbf{x},t;\mathbf{x}_{\text{src}}) \end{aligned}$$

Total data model

Discretizes into

$$M(\mathbf{g},\mathbf{q};\mathbf{p}) = \mathbf{F}_\omega^{-1}\left[M_\omega(\mathbf{G},\mathbf{Q};\mathbf{P})\right]$$

$$M_\omega(\mathbf{G},\mathbf{Q};\mathbf{P}) := \mathbf{GQ} + \mathbf{GRP}$$

Underlying assumption: there exists **G** for all frequencies such that

$$\mathbf{P} \approx \mathbf{GQ} - \mathbf{GP}$$

$$\mathbf{P_o} = \mathbf{GQ}$$

"Data matrix" notation

$$\begin{array}{cc} & \texttt{source index} \\ \texttt{receiver index} & \left[\quad \mathbf{G} \quad\right] \end{array}$$

$\omega = \omega_a$ single fixed frequency

Estimation of Primaries by Sparse Inversion

(van Groenestijn and Verschuur, 2009)

recorded data — predicted data from convolution model

$$\mathbf{P} = \mathbf{QG} - \mathbf{GP}$$

$\mathbf{P}$	total up-going wavefield
$\mathbf{Q}$	down-going source signature
$\mathbf{G}$	primary impulse response

Estimation of Primaries by Sparse Inversion (EPSI)

(van Groenestijn and Verschuur, 2009)

recorded data predicted data from convolution model

$$\mathbf{P} = \mathbf{QG} - \mathbf{GP}$$

Inversion objective:

$$f(\mathbf{G}, \mathbf{Q}) = \frac{1}{2}\|\mathbf{P} - (\mathbf{QG} - \mathbf{GP})\|_2^2$$

Sparse regularization on **g** in physical (time) domain via hard-thresholding on gradients

Primary Estimation instead of Multiple Removal

EPSI

recorded data — predicted data from convolution model

$$\mathbf{P} = \mathbf{Q}\mathbf{G} - \mathbf{G}\mathbf{P}$$

$\mathbf{P}$ total up-going wavefield

$\mathbf{Q}$ down-going source signature

$\mathbf{G}$ primary impulse response

Primary Estimation instead of **Multiple Removal**

EPSI

recorded data — predicted data from convolution model

$$\mathbf{P} = \mathbf{QG} - \mathbf{GP}$$

$$\mathbf{P_o} = \mathbf{QG}$$
$$A(f) = -\mathbf{Q}^{-1}$$

$\mathbf{P}$ total up-going wavefield

$\mathbf{Q}$ down-going source signature

$\mathbf{G}$ primary impulse response

Primary Estimation instead of Multiple Removal

true primary wavefield

SRME-produced primary

SRME

$$\mathbf{P_o} = \mathbf{P} - A(f)\mathbf{P_o}\mathbf{P}$$

$$\mathbf{P_o} = \mathbf{QG}$$
$$A(f) = -\mathbf{Q}^{-1}$$

$\mathbf{P}$ total up-going wavefield

$\mathbf{Q}$ down-going source signature

$\mathbf{G}$ primary impulse response

Primary Estimation instead of **Multiple Removal**

true primary wavefield

SRME-produced primary

SRME

$$\mathbf{P_o} \approx \mathbf{P} - A(f)\boxed{\mathbf{PP}}$$

SRMP

$$\mathbf{P_o} = \mathbf{QG}$$
$$A(f) = -\mathbf{Q}^{-1}$$

$\mathbf{P}$ total up-going wavefield

$\mathbf{Q}$ down-going source signature

$\mathbf{G}$ primary impulse response

Primary Estimation instead of **Multiple Removal**

adaptive subtraction

$$\min_{A} \sum_{f} \|\mathbf{P} - A(f)\boxed{\mathbf{PP}}\|$$

SRMP

$$\mathbf{P_o} = \mathbf{QG}$$
$$A(f) = -\mathbf{Q}^{-1}$$

$\mathbf{P}$	total up-going wavefield
$\mathbf{Q}$	down-going source signature
$\mathbf{G}$	primary impulse response

Primary Estimation instead of **Multiple Removal**

recorded data predicted data from convolution model

$$\mathbf{P} = \mathbf{QG} - \mathbf{GP}$$

Inversion objective:

$$f(\mathbf{G}, \mathbf{Q}) = \frac{1}{2}\|\mathbf{P} - (\mathbf{QG} - \mathbf{GP})\|_2^2$$

Sparse regularization on **g** in physical (time) domain via hard-thresholding on gradients

Main contributions

1. **Robust EPSI**, a new formulation of the EPSI primary estimation problem which avoids *ad-hoc parameter adjustments* in favour of *self-tuning bi-convex optimization*

2. **Near-offset mitigation with scattering**, a correction to the multiple prediction in face of missing data, using Born scattering off the free surface.

3. **A multigrid acceleration strategy** for REPSI, can *greatly reduce* REPSI computation time (1 order of magnitude in 2D, 2 order in 3D)

Chapter 3

Robust EPSI

Estimation of Primaries by Sparse Inversion (EPSI)

(van Groenestijn and Verschuur, 2009)

recorded data — predicted data from convolution model

$$\mathbf{P} = \mathbf{QG} - \mathbf{GP}$$

Inversion objective:

$$f(\mathbf{G}, \mathbf{Q}) = \frac{1}{2}\|\mathbf{P} - (\mathbf{QG} - \mathbf{GP})\|_2^2$$

Optimization form

In time domain (lower-case: whole dataset in time domain)

recorded data predicted data from SRME

$$\mathbf{p} = M(\mathbf{g}, \mathbf{q}; \mathbf{p})$$

$$M(\mathbf{g}, \mathbf{q}; \mathbf{p}) = \mathbf{F}_\omega^{-1}\left[M_\omega(\mathbf{G}, \mathbf{Q}; \mathbf{P})\right]$$

$$M_\omega(\mathbf{G}, \mathbf{Q}; \mathbf{P}) := \mathbf{GQ} + \mathbf{GRP}$$

Inversion objective:

$$f(\mathbf{g}, \mathbf{q}) = \frac{1}{2}\|\mathbf{p} - M(\mathbf{g}, \mathbf{q}; \mathbf{p})\|_2^2$$

Solving the EPSI problem

Linearizations

$$\mathbf{p} = M(\mathbf{g}, \mathbf{q}; \mathbf{p})$$

$$\mathbf{M}_{\tilde{q}} = \left(\frac{\partial M}{\partial \mathbf{g}}\right)_{\tilde{q}}$$

$$\mathbf{M}_{\tilde{g}} = \left(\frac{\partial \mathcal{M}}{\partial \mathbf{q}}\right)_{\tilde{g}}$$

In fact it is bilinear:

$$\mathbf{M}_{\tilde{q}}\mathbf{g} = M(\mathbf{g}, \tilde{\mathbf{q}}) \qquad \mathbf{M}_{\tilde{g}}\mathbf{q} = M(\mathbf{q}, \tilde{\mathbf{g}})$$

Solving the EPSI problem

SLIM

Linearizations

$$\mathbf{p} = M(\mathbf{g}, \mathbf{q}; \mathbf{p})$$

$$\mathbf{M}_{\tilde{q}} = \left(\frac{\partial M}{\partial \mathbf{g}}\right)_{\tilde{q}}$$

$$\mathbf{M}_{\tilde{g}} = \left(\frac{\partial \mathcal{M}}{\partial \mathbf{q}}\right)_{\tilde{g}}$$

Associated objectives:

$$f_{\tilde{q}}(\mathbf{g}) = \frac{1}{2}\|\mathbf{p} - \mathbf{M}_{\tilde{q}}\mathbf{g}\|_2^2 \qquad f_{\tilde{g}}(\mathbf{q}) = \frac{1}{2}\|\mathbf{p} - \mathbf{M}_{\tilde{g}}\mathbf{q}\|_2^2$$

Solving the EPSI problem

SLIM

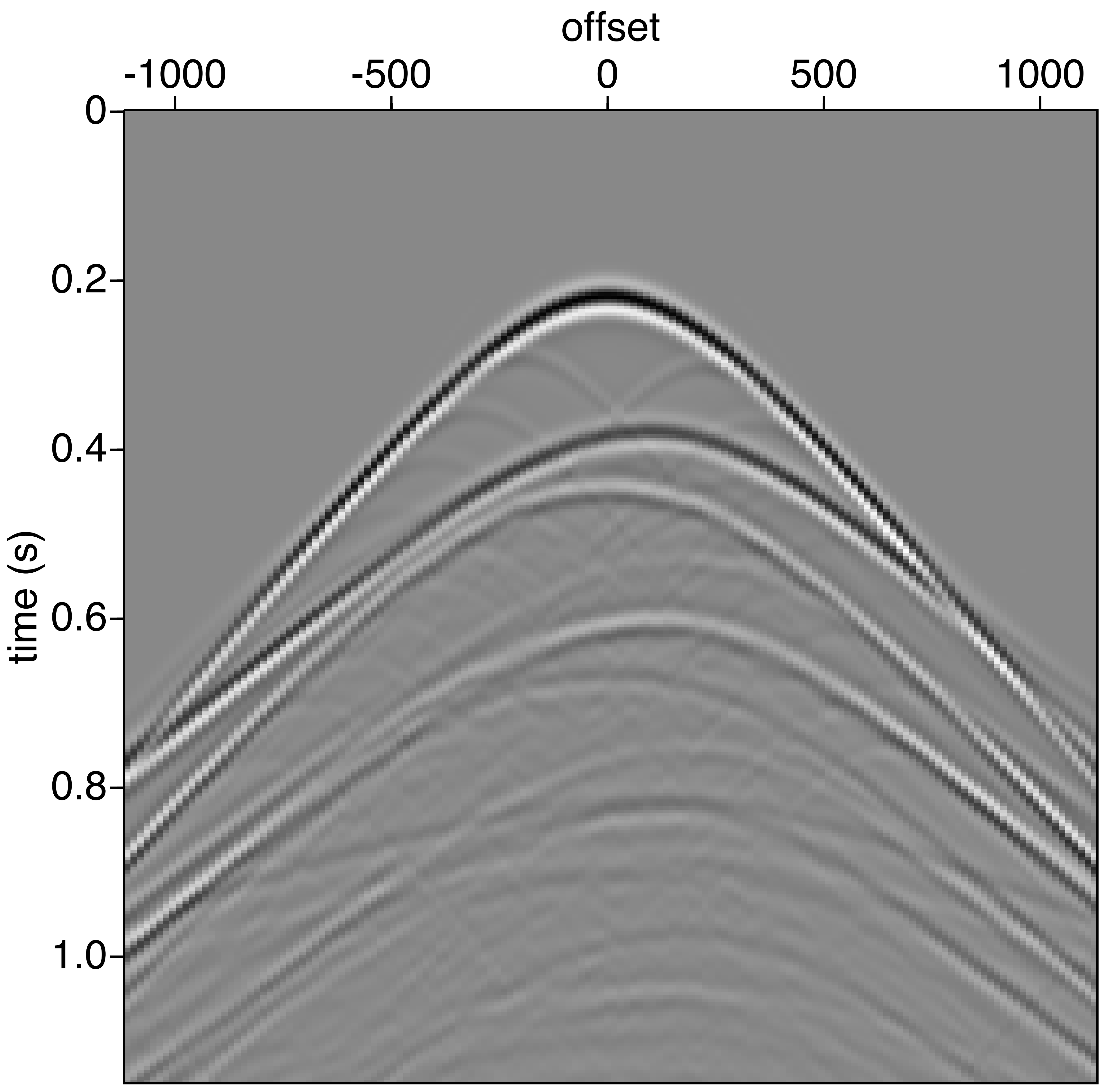
Do:

$$\mathbf{g}_{k+1} = \mathbf{g}_k + \alpha\mathcal{S}(\nabla f_{q_k}(\mathbf{g}_k))$$
$$\mathbf{q}_{k+1} = \mathbf{q}_k + \beta\nabla f_{g_{k+1}}(\mathbf{q}_k)$$

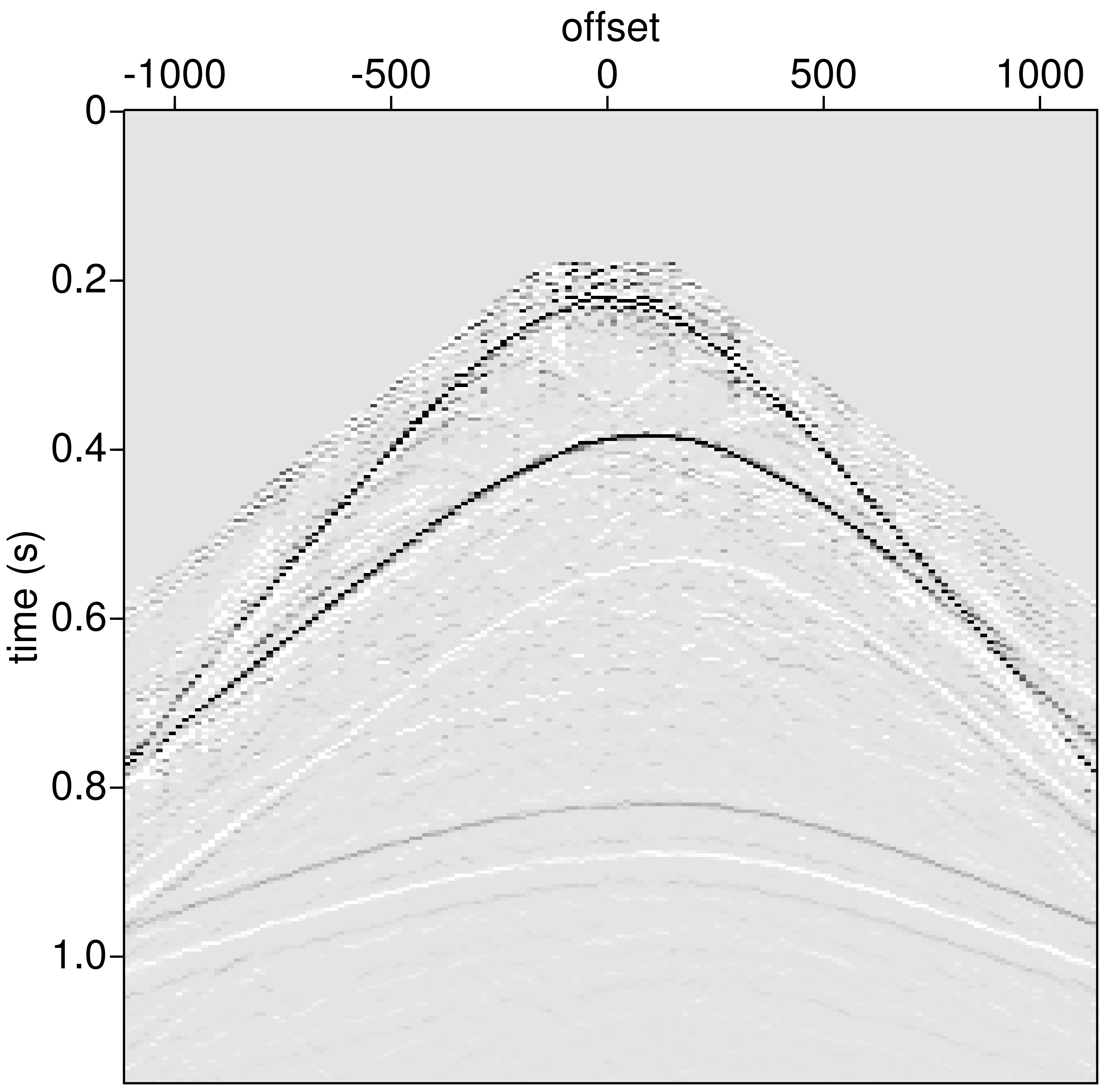
Gradient sparsity

$\mathcal{S}$: pick largest ρ elements per trace

Solving the EPSI problem



Data



EPSI Green’s function

Robust EPSI

L1-minimization approach to the EPSI problem

[Lin and Herrmann, 2013 *Geophysics*]

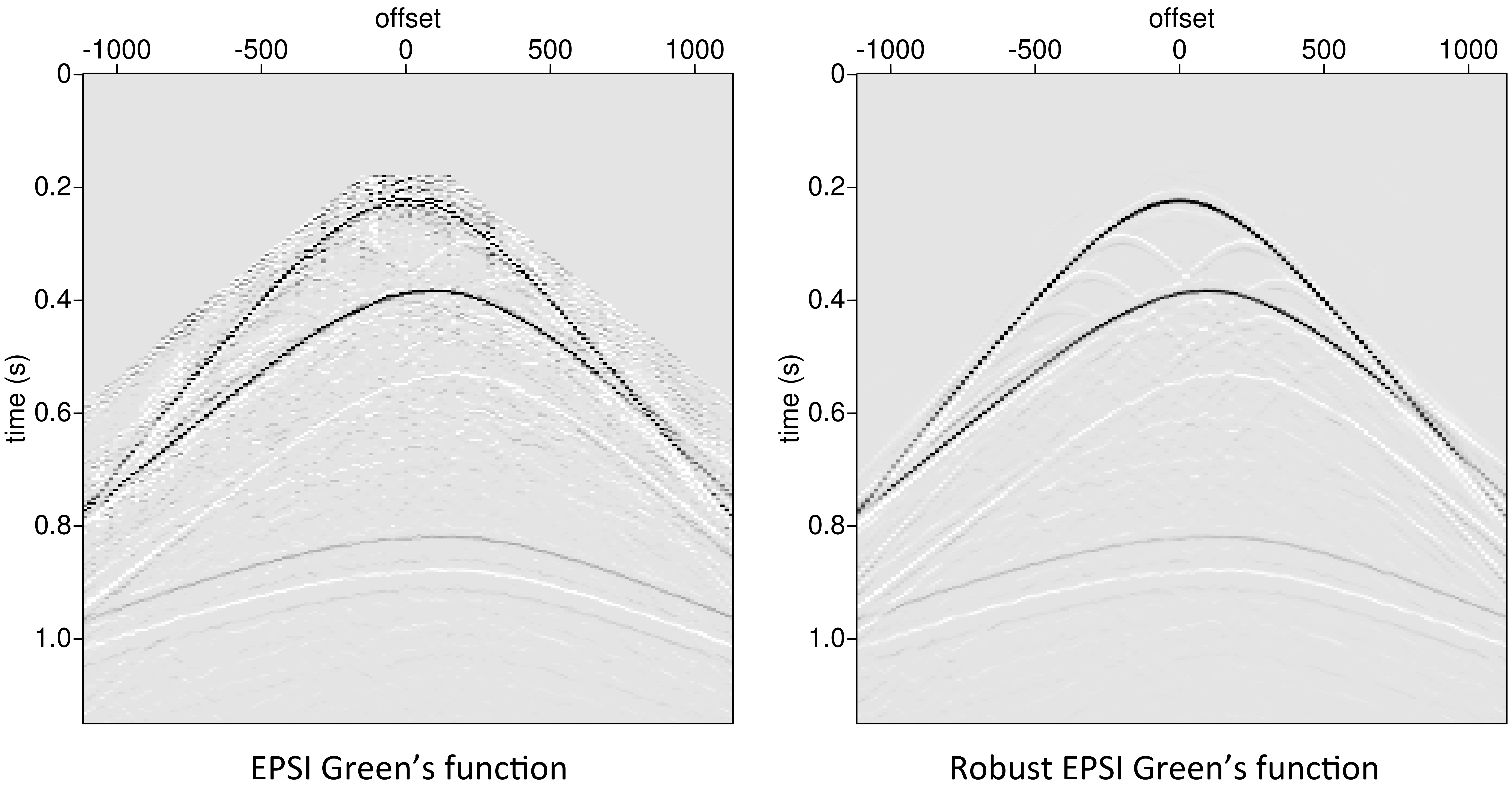
While $\|\mathbf{p} - \mathcal{M}(\mathbf{g}_k, \mathbf{q}_k)\|_2 > \sigma$

determine new τ_k from the Pareto curve

$$\mathbf{g}_{k+1} = \arg\min_{\mathbf{g}} \|\mathbf{p} - \mathbf{M}_{q_k}\mathbf{g}\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \leq \tau_k$$

$$\mathbf{q}_{k+1} = \arg\min_{\mathbf{q}} \|\mathbf{p} - \mathbf{M}_{g_{k+1}}\mathbf{q}\|_2$$

Robust EPSI
L1-minimization approach to the EPSI problem



Robust EPSI
L1-minimization approach to the EPSI problem

[Lin and Herrmann, 2013 *Geophysics*]

While $\|\mathbf{p} - \mathcal{M}(\mathbf{g}_k, \mathbf{q}_k)\|_2 > \sigma$

determine new τ_k from the Pareto curve

$$\mathbf{g}_{k+1} = \underset{\mathbf{g}}{\arg\min} \|\mathbf{p} - \mathbf{M}_{q_k}\mathbf{g}\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \leq \tau_k$$

$$\mathbf{q}_{k+1} = \underset{\mathbf{q}}{\arg\min} \|\mathbf{p} - \mathbf{M}_{g_{k+1}}\mathbf{q}\|_2$$

Robust EPSI
L1-minimization approach to the EPSI problem

[Lin and Herrmann, 2013 *Geophysics*]

While $\|\mathbf{p} - \mathcal{M}(\mathbf{g}_k, \mathbf{q}_k)\|_2 > \sigma$

determine new τ_k from the Pareto curve

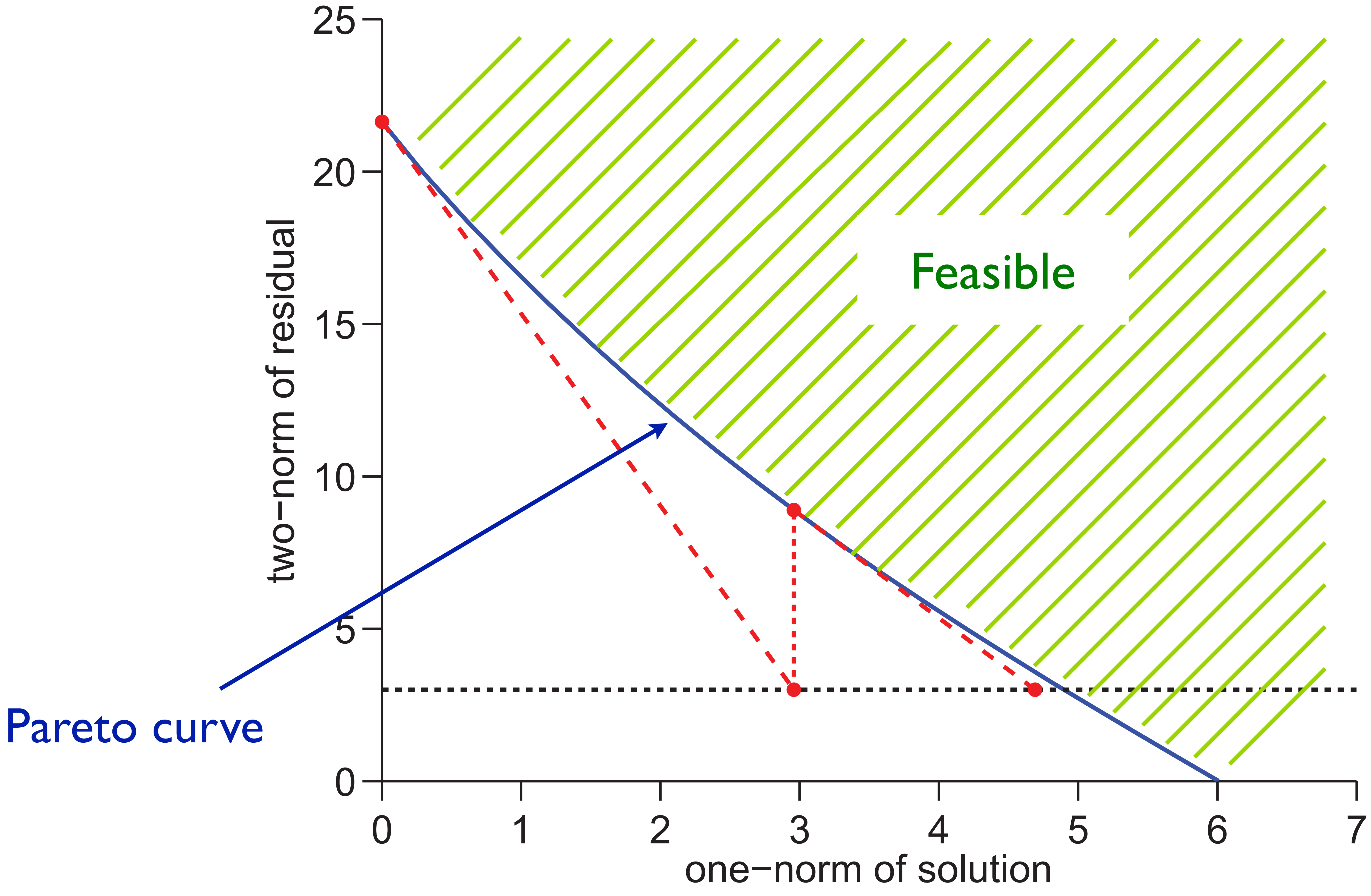
$$\mathbf{g}_{k+1} = \arg\min_{\mathbf{g}} \|\mathbf{p} - \mathbf{M}_{q_k}\mathbf{g}\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \leq \tau_k$$

(Lasso problem, solve with SPG part of SPGL1 until Pareto curve reached)

$$\mathbf{q}_{k+1} = \arg\min_{\mathbf{q}} \|\mathbf{p} - \mathbf{M}_{g_{k+1}}\mathbf{q}\|_2$$

(wavelet matching, solve with LSQR)

The Pareto curve

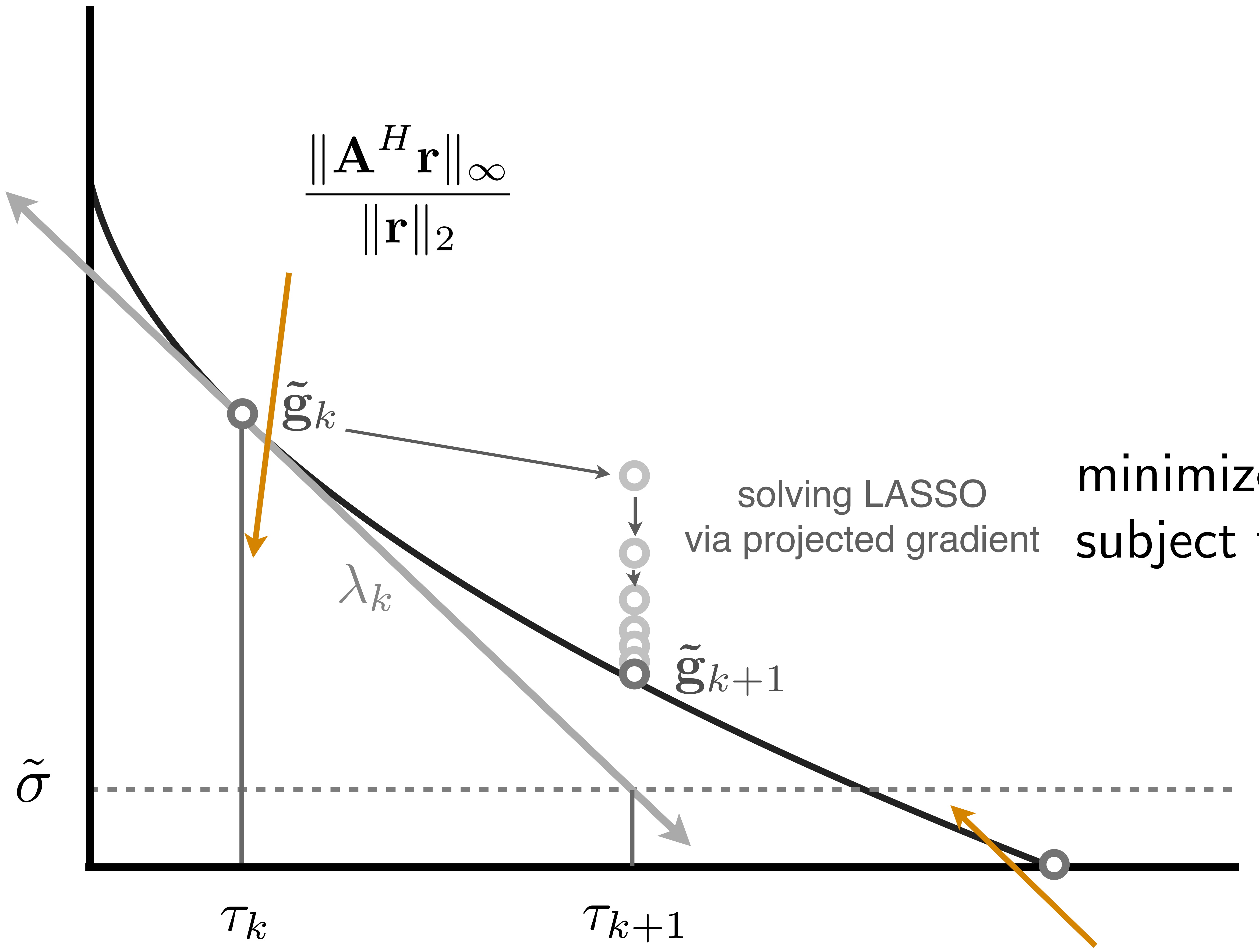


(van den Berg, Friedlander, 2008)

Choosing Tau from the Pareto curve

minimize $\quad \|x\|_1$
subject to $\quad \|Ax - b\|_2 \leq \sigma$

Look at the solution space and the line of optimal solutions (Pareto curve)



minimize $\quad \|Ax - b\|_2$
subject to $\quad \|x\|_1 \leq \tau$

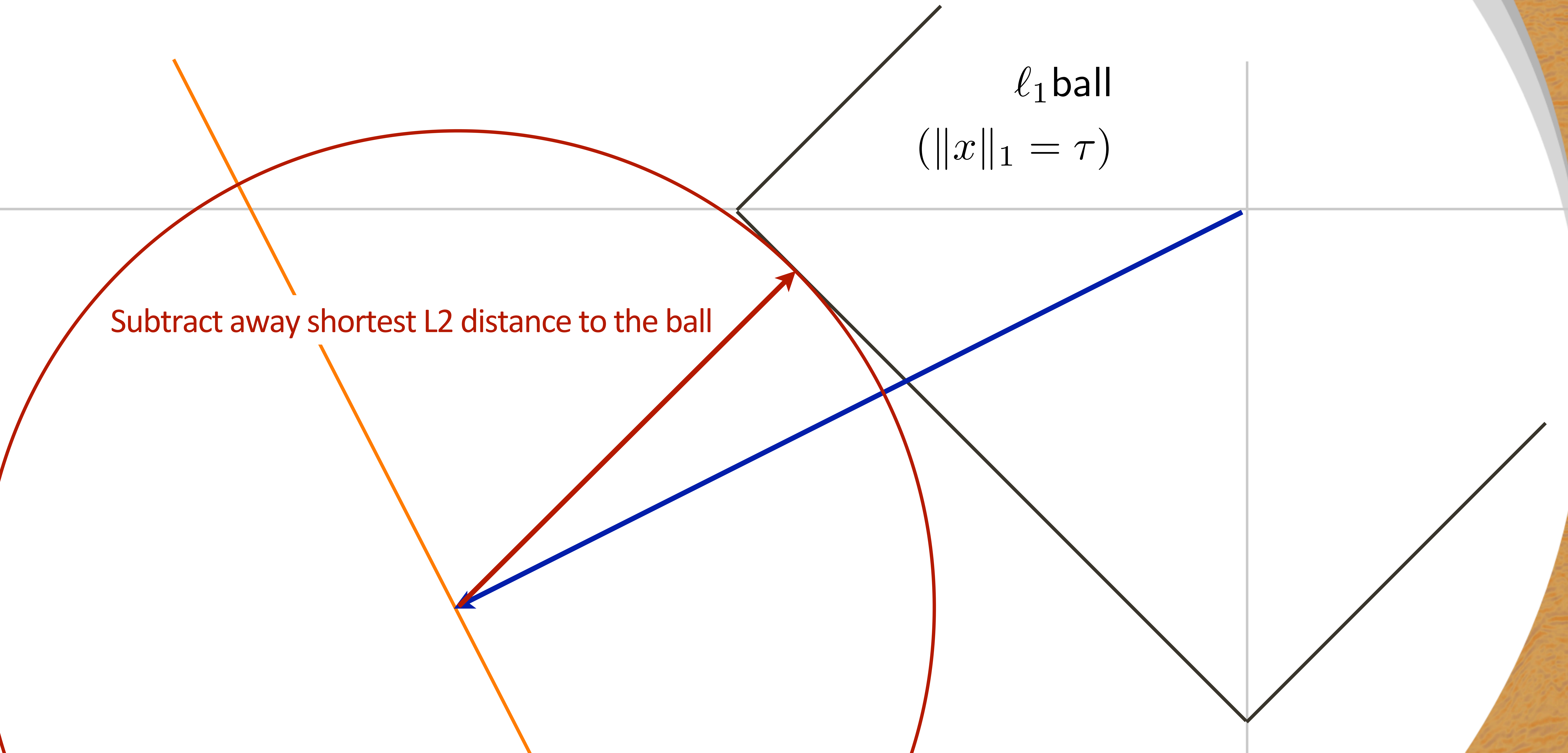
feasible solution with smallest $\|x\|_1$

Solving the LASSO subproblems

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|_2 \\ \text{subject to} & \|x\|_1 \leq \tau \end{array}$$

Solving the LASSO subproblems

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|_2 \\ \text{subject to} & \|x\|_1 \leq \tau \end{array}$$



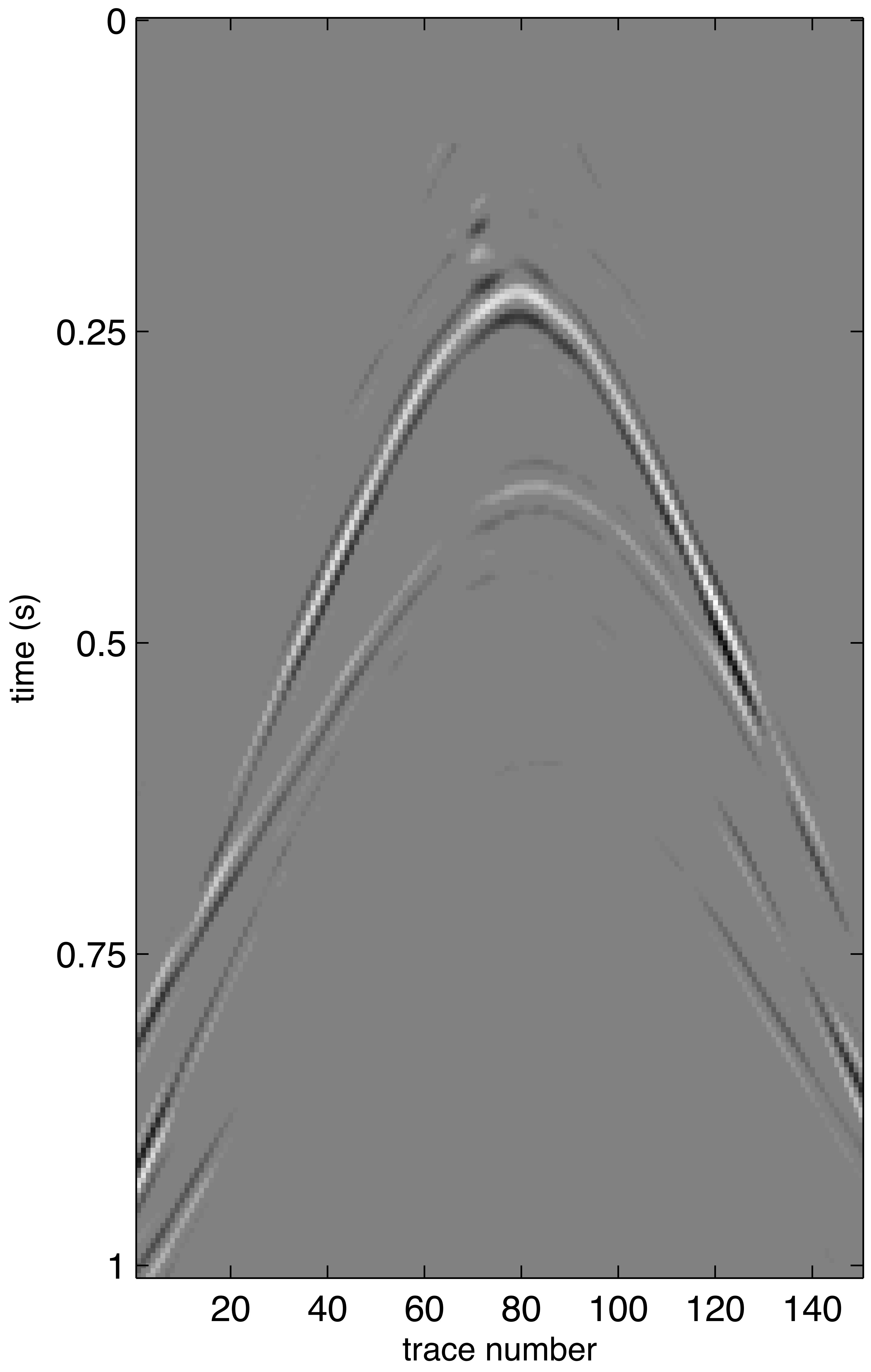
Solving the LASSO subproblems

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|_2 \\ \text{subject to} & \|x\|_1 \leq \tau \end{array}$$

L1 projection and sparsity

variable g at beginning of LASSO

$$\mathbf{g}_{k+1} = \arg\min_{\mathbf{g}} \|\mathbf{p} - \mathbf{M}_{q_k}\mathbf{g}\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \le \tau_k$$

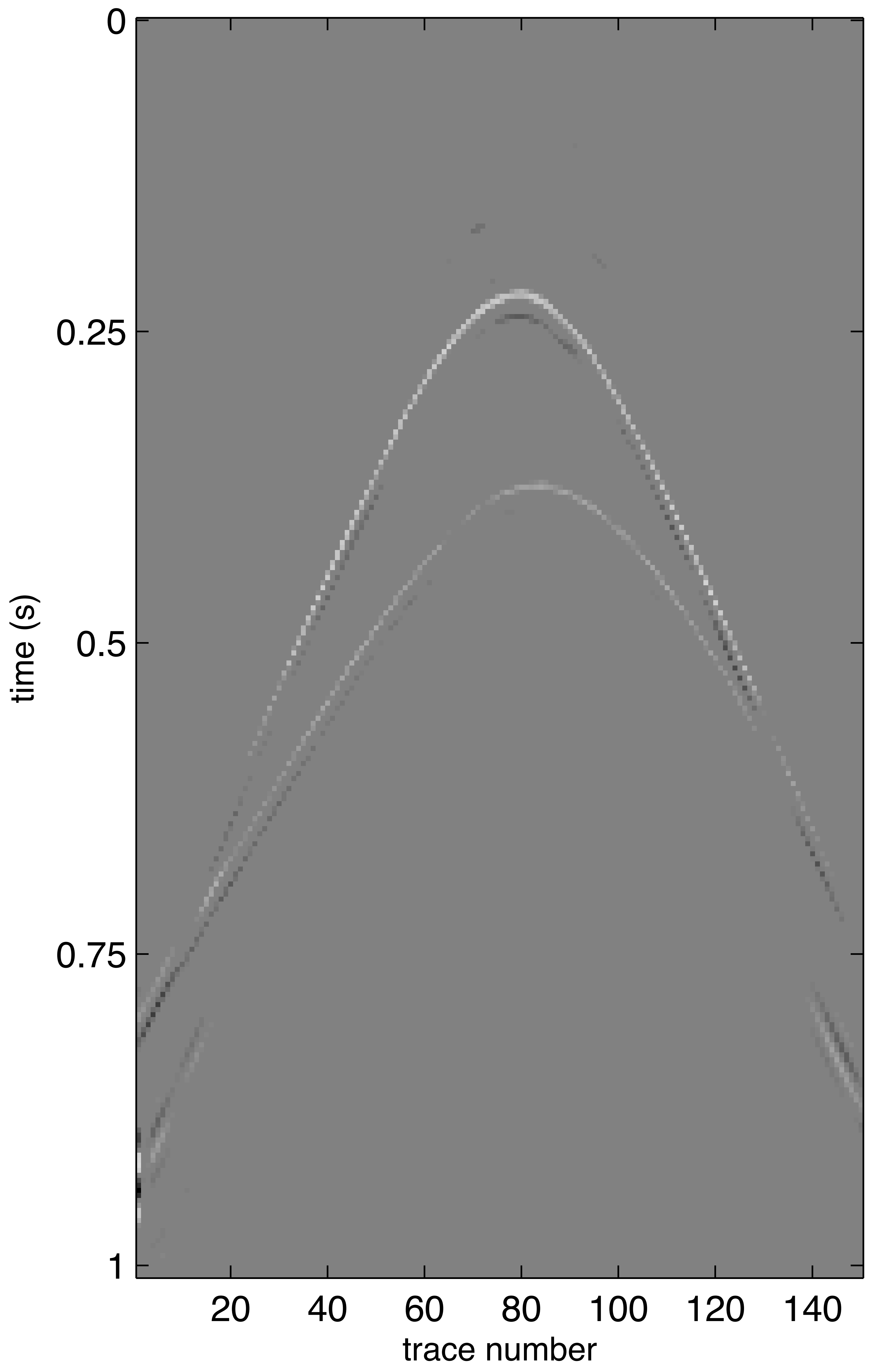


L1 projection and sparsity

variable g at end of LASSO

$$\mathbf{g}_{k+1} = \arg\min_{\mathbf{g}} \|\mathbf{p} - \mathbf{M}_{q_k}\mathbf{g}\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \leq \tau_k$$

Emits "deconvolved" solution



Robust EPSI

L1-minimization approach to the EPSI problem

[Lin and Herrmann, 2013 *Geophysics*]

While $\|\mathbf{p} - \mathcal{M}(\mathbf{g}_k, \mathbf{q}_k)\|_2 > \sigma$

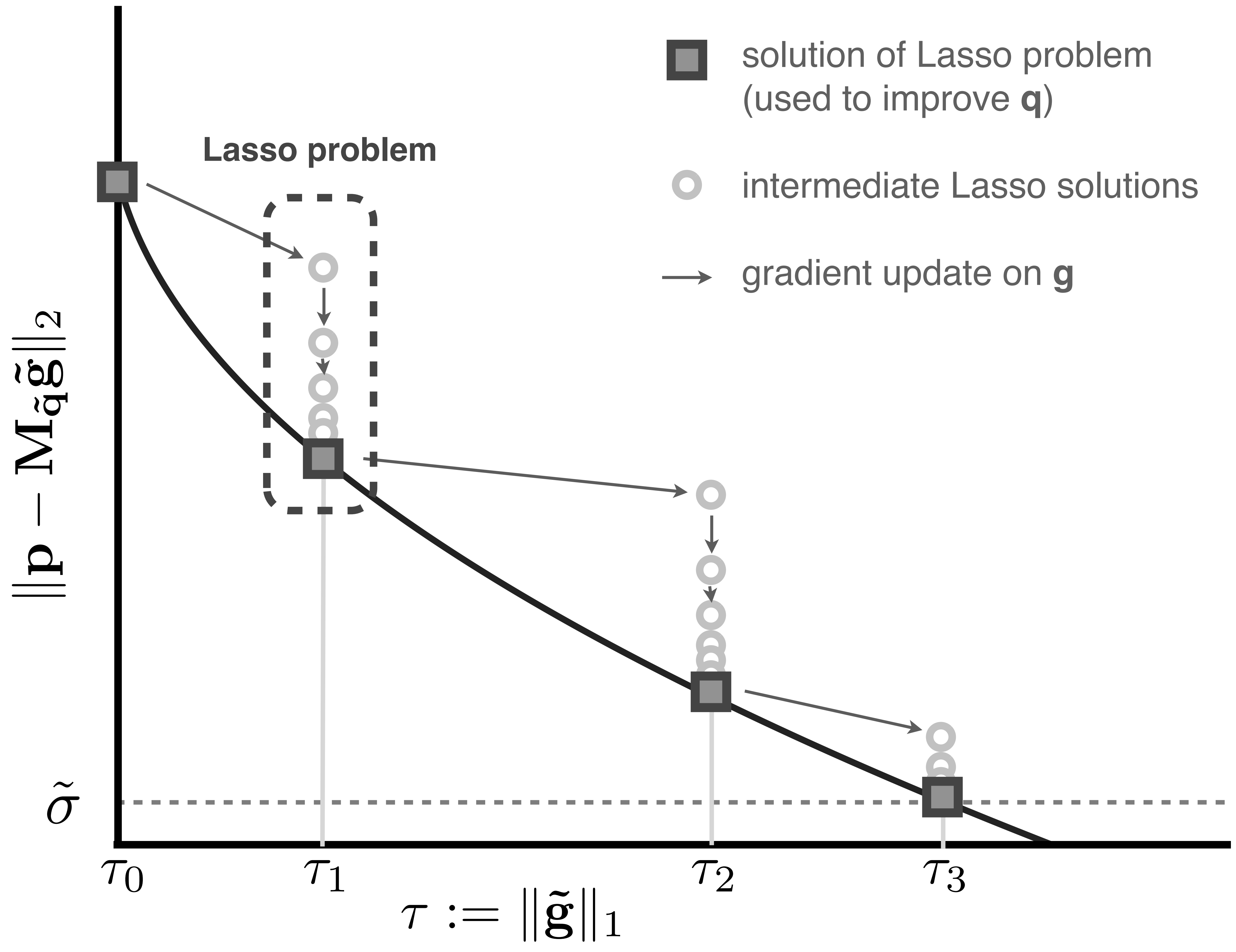
determine new τ_k from the Pareto curve

$$\mathbf{g}_{k+1} = \arg\min_{\mathbf{g}} \|\mathbf{p} - \mathbf{M}_{q_k}\mathbf{g}\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \leq \tau_k$$

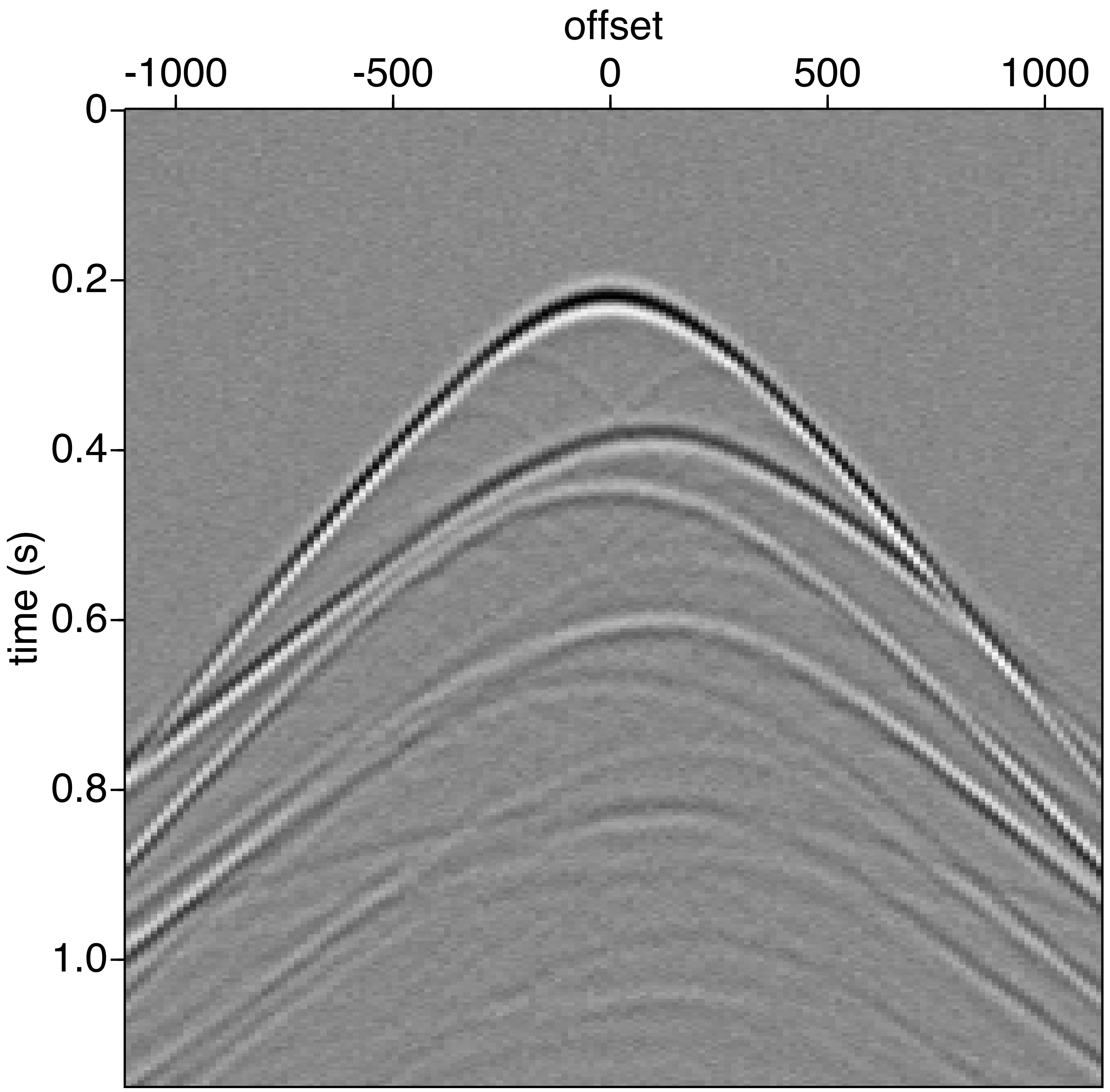
$$\mathbf{q}_{k+1} = \arg\min_{\mathbf{q}} \|\mathbf{p} - \mathbf{M}_{g_{k+1}}\mathbf{q}\|_2$$

Overall REPSI solution path

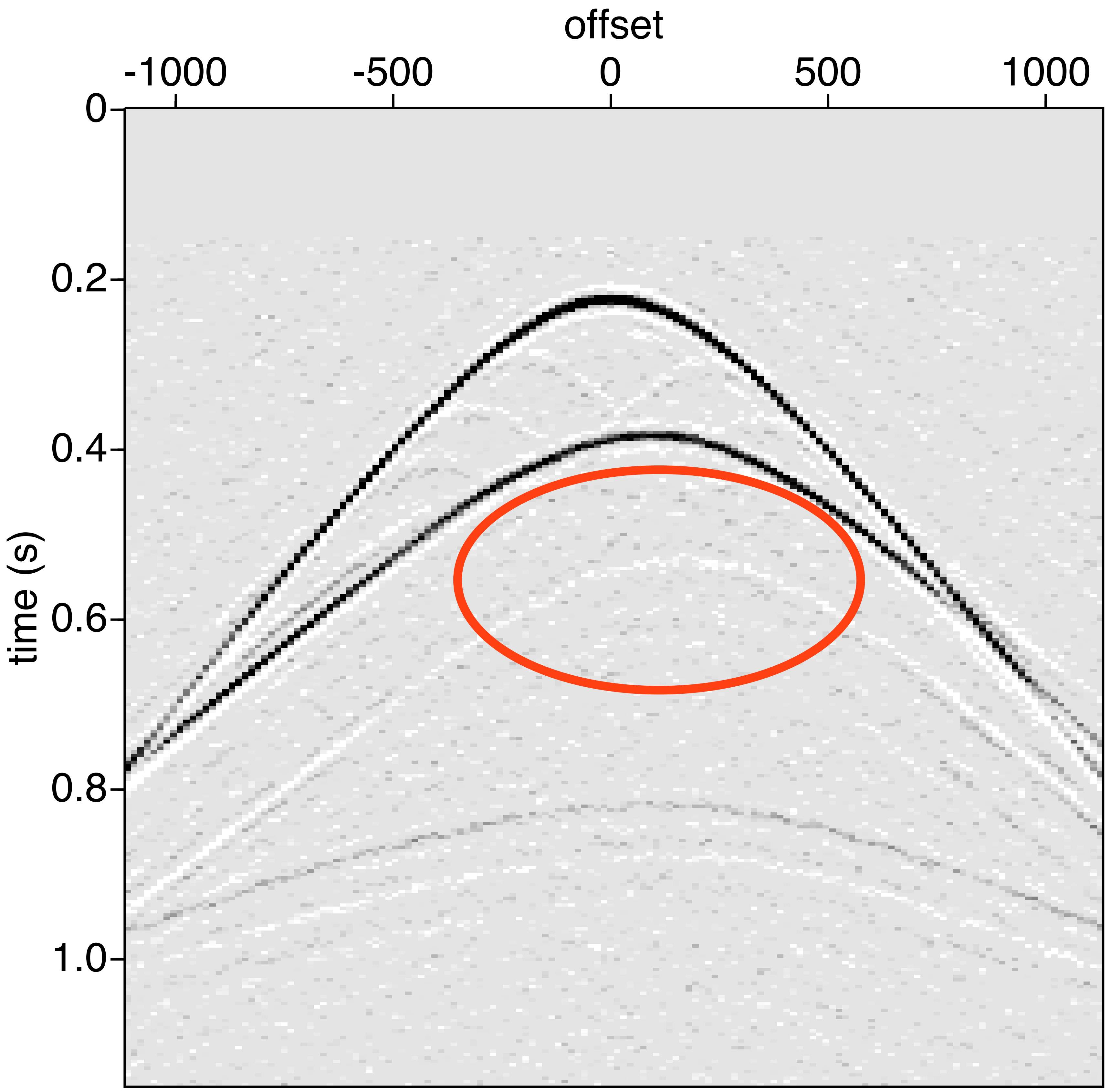
$$\begin{array}{ll}\text{minimize} & \|x\|_1 \\ \text{subject to} & \|Ax - b\|_2 \le \sigma\end{array}$$



Robustness under noise

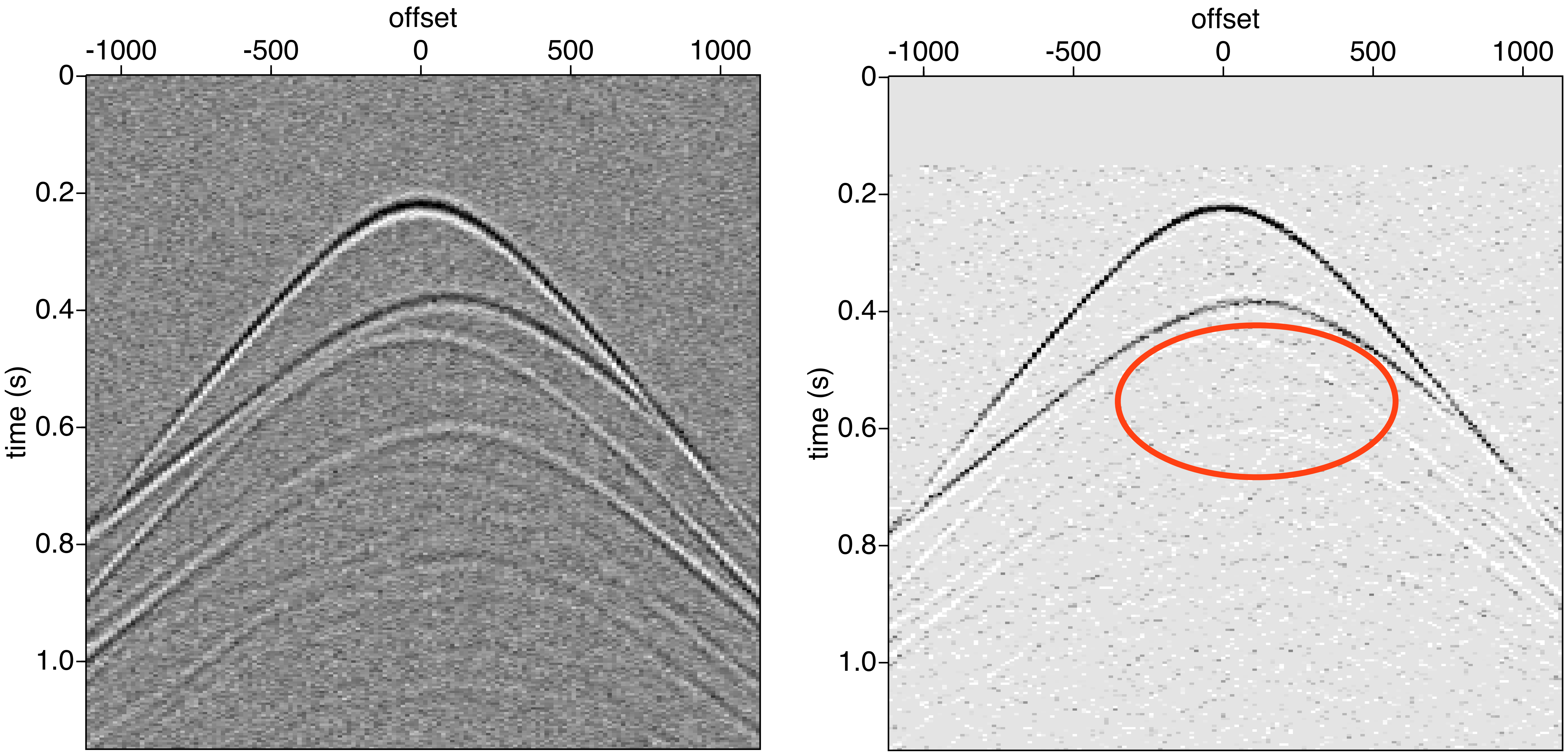


Data + 40% noise (SNR 7)



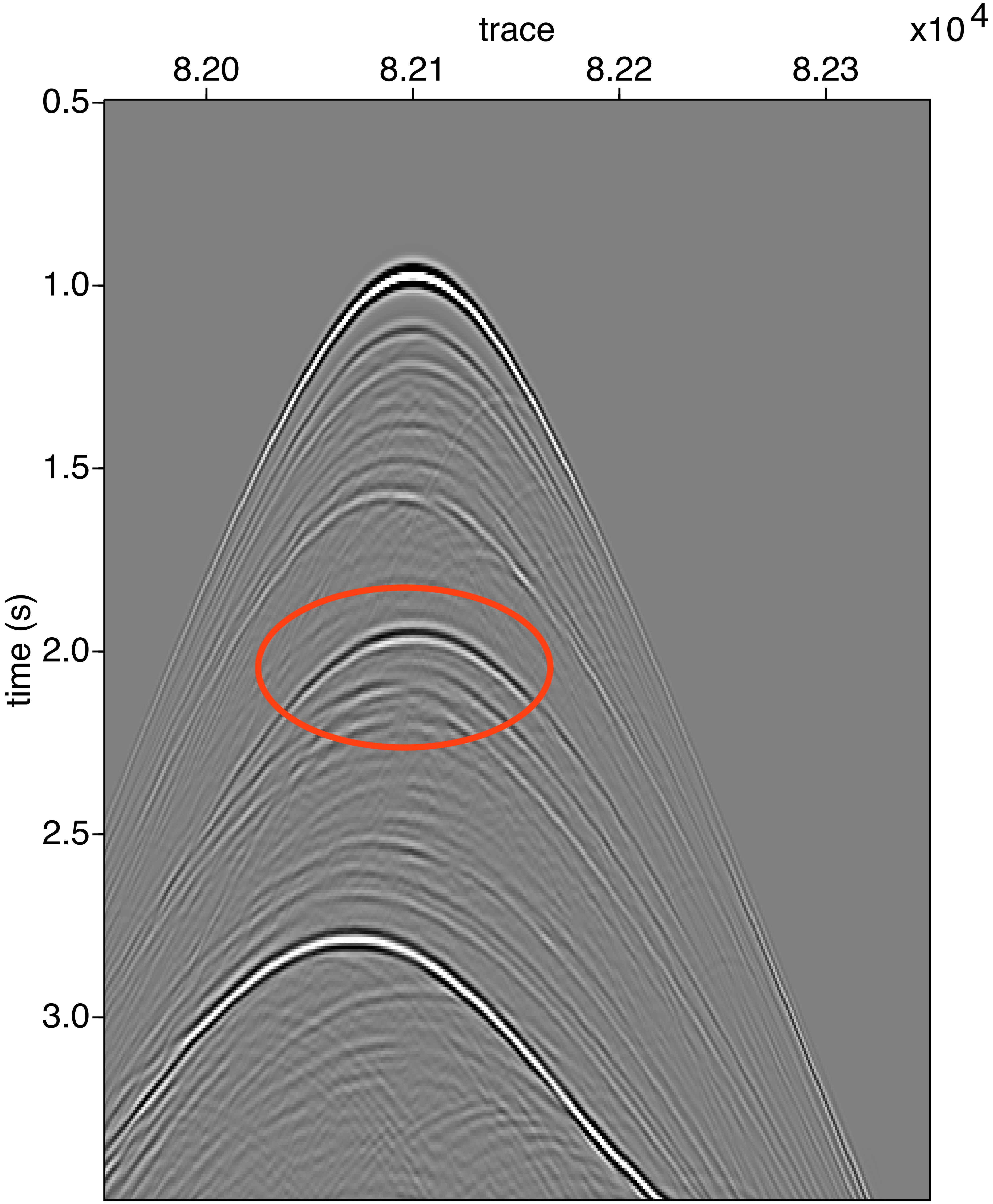
Robust EPSI IR

Robustness under noise



Data + 100% noise (SNR 0)

Robust EPSI IR



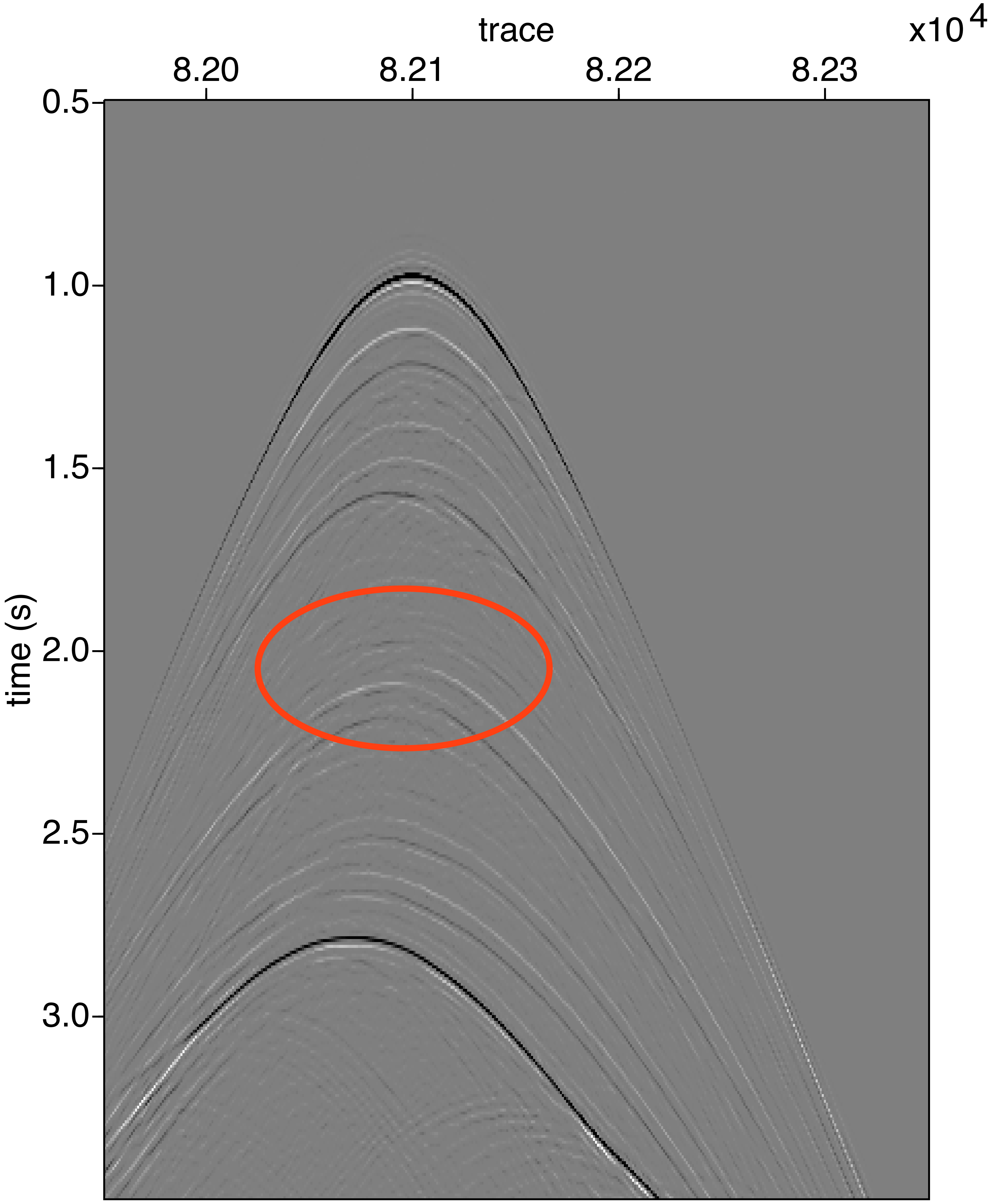
Pluto1.5

data

elastic FD modeling

muted

no deghosting



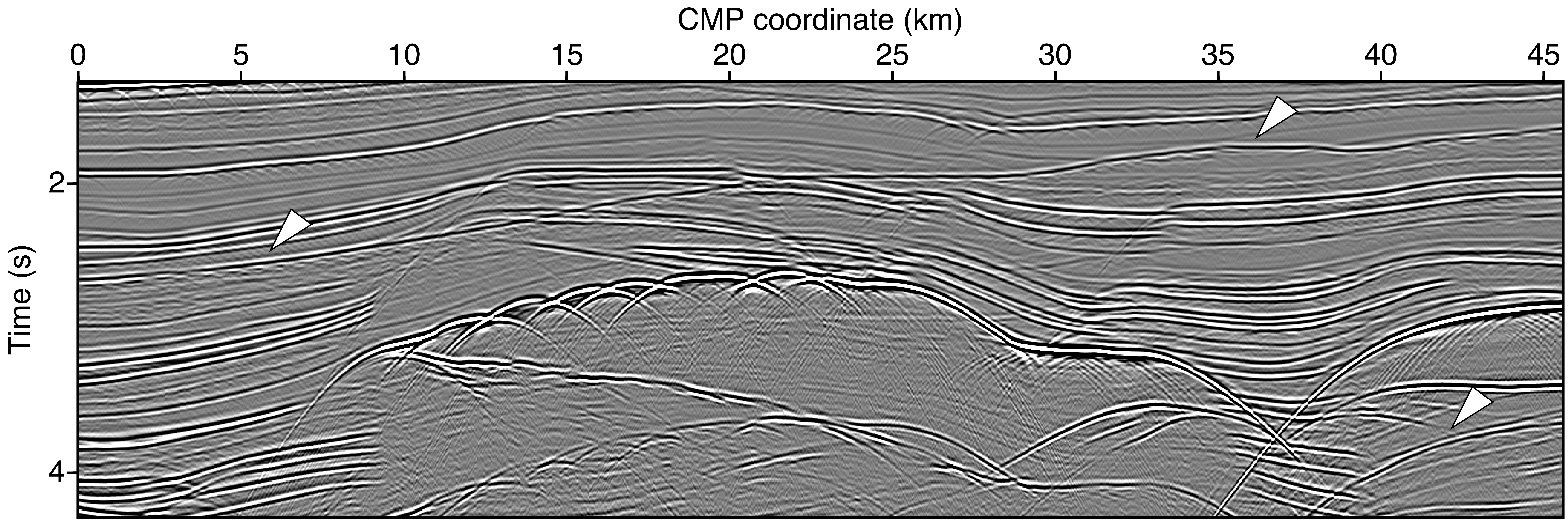
Pluto1.5

REPSI Primary IR

80 total gradient iters

Pluto1.5 **Data**

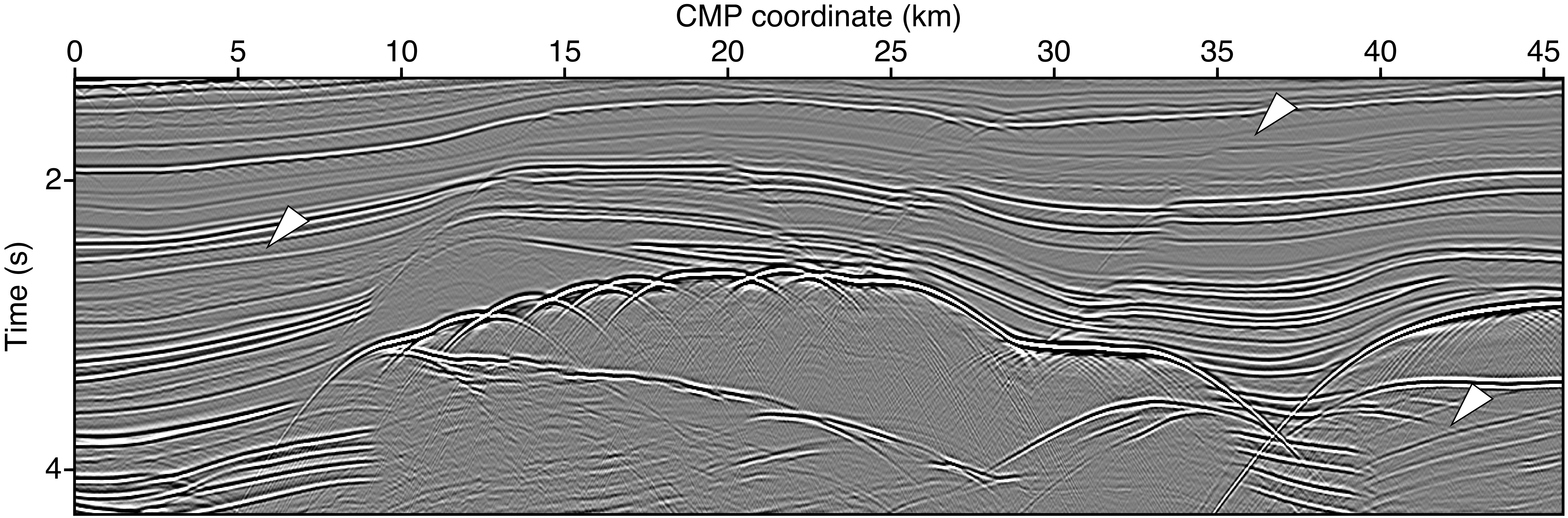
NMO-corrected stacks



SLIM

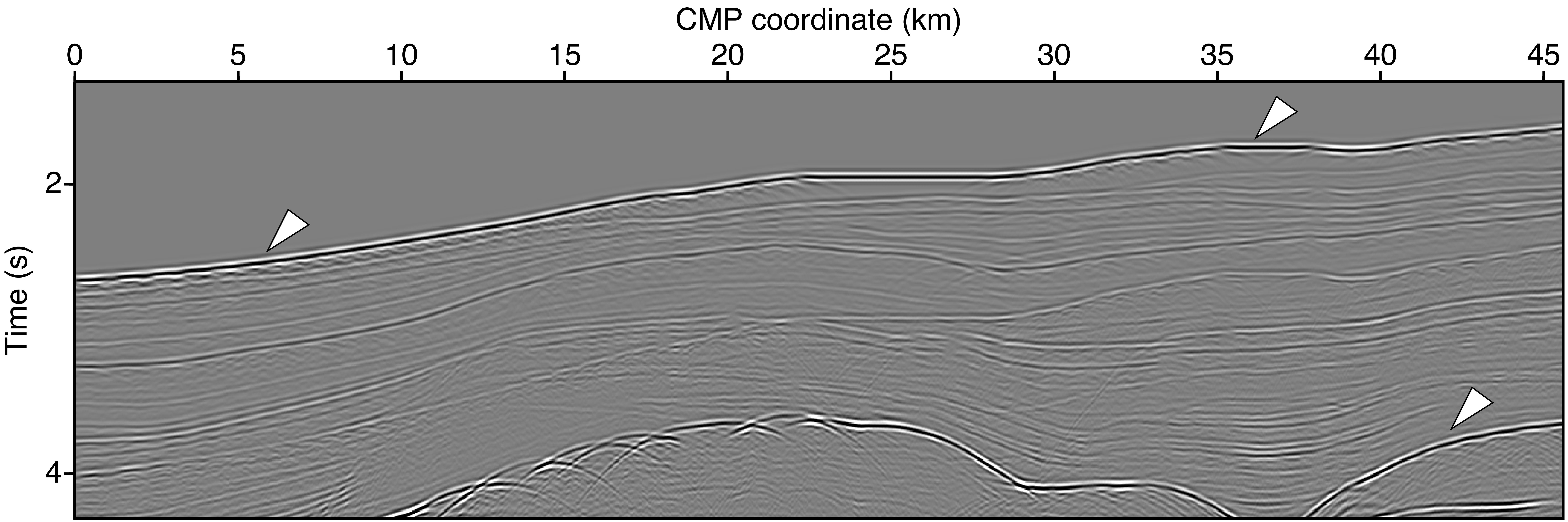
Pluto1.5 **REPSI primary**

80 total gradient iters



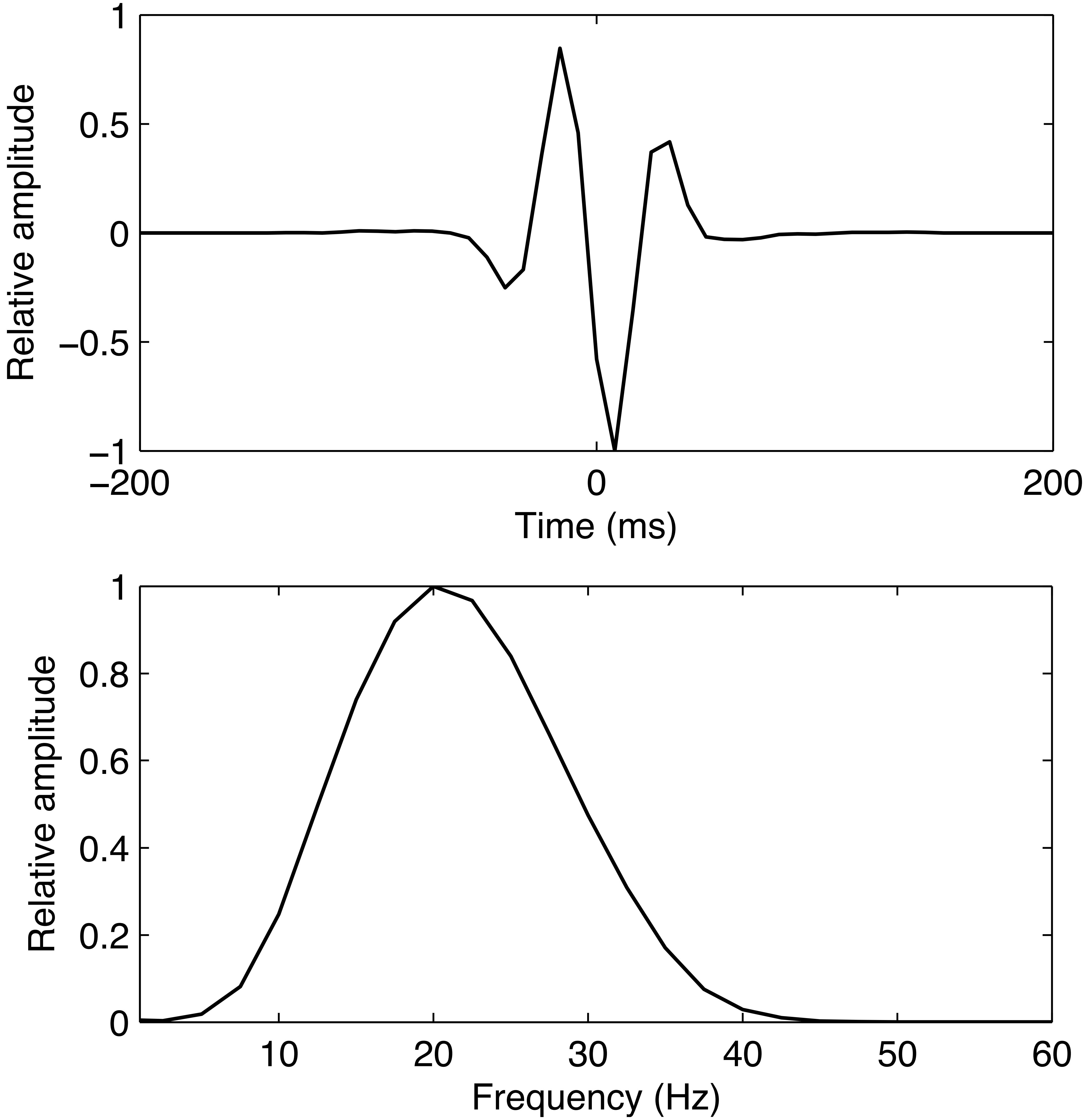
Pluto1.5 **REPSI multiple model**

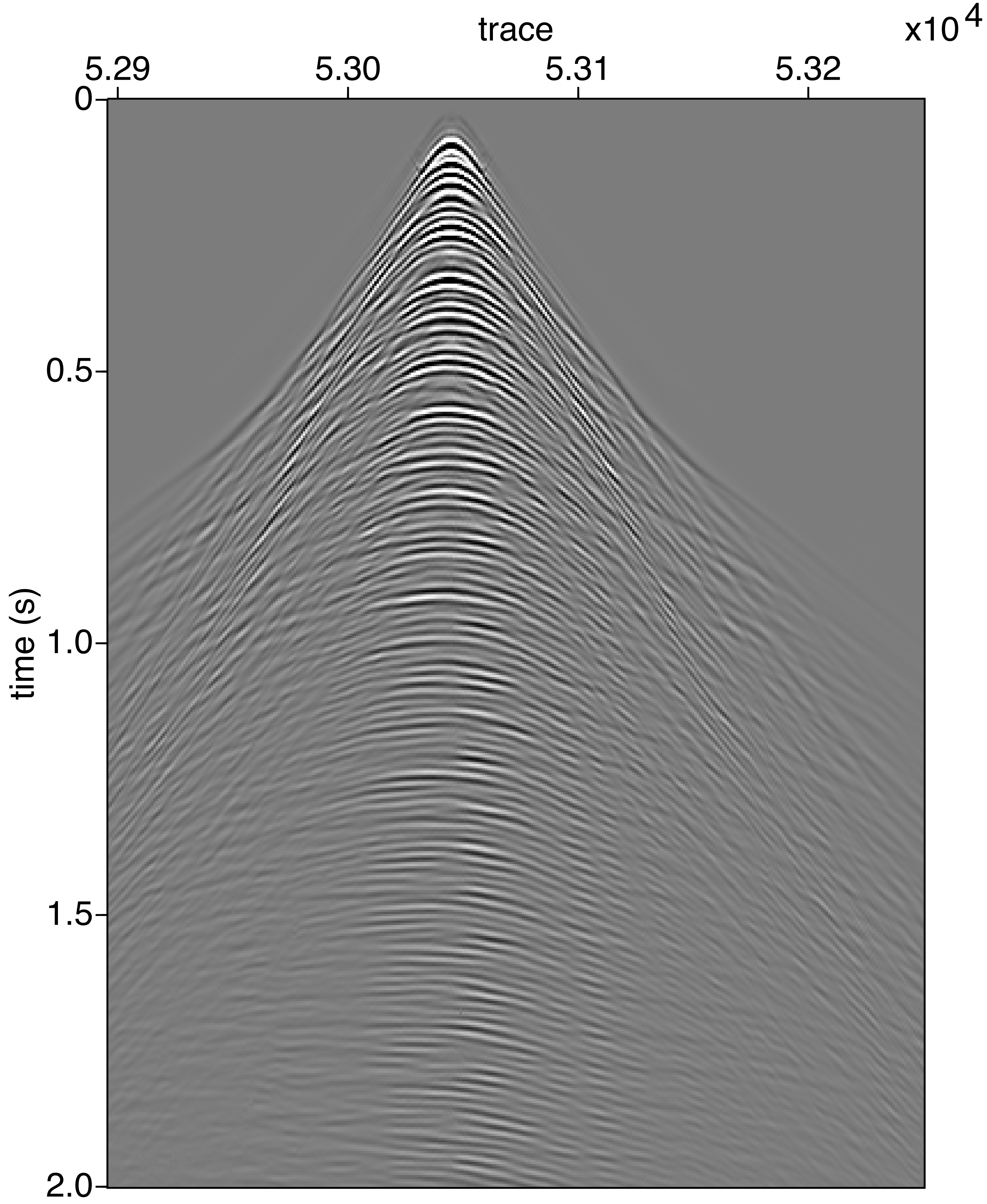
80 total gradient iters



REPSI source signature model

total 8 updates

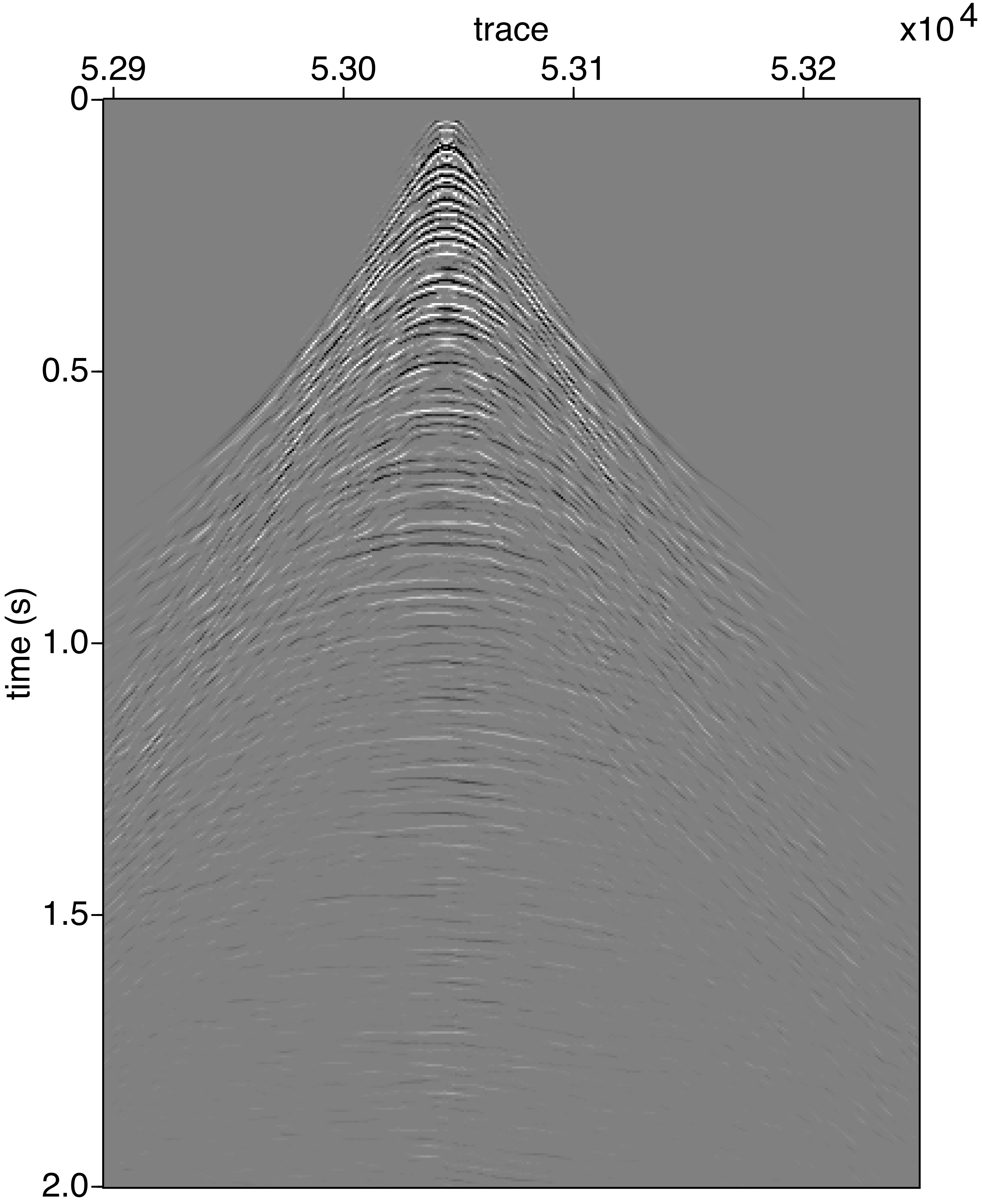




Gulf of Suez

data

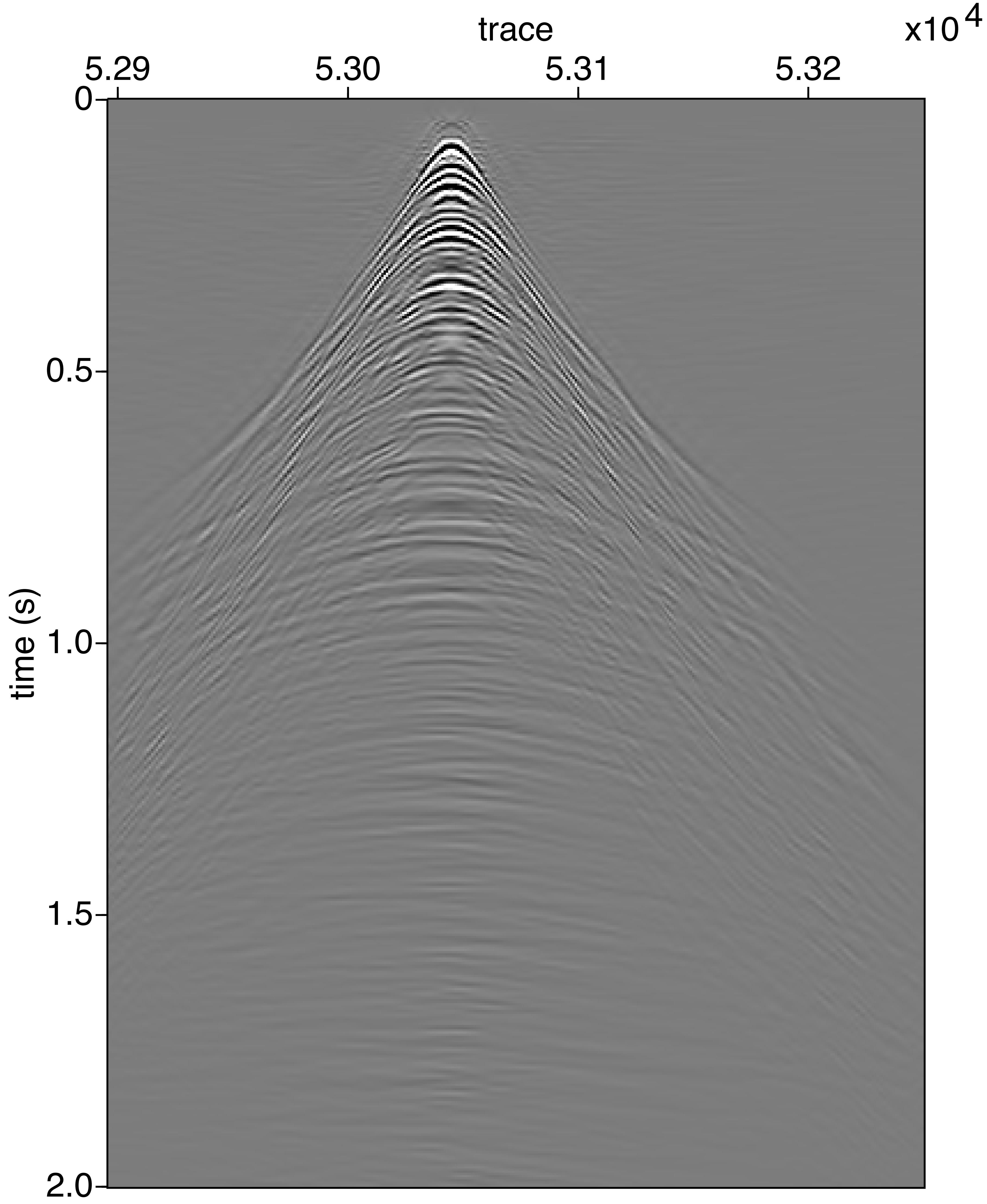
- shot gather
- reciprocity applied
- interpolated, muted
- no deghosting



Gulf of Suez

REPSI Primary IR

80 total gradient iters



Gulf of Suez

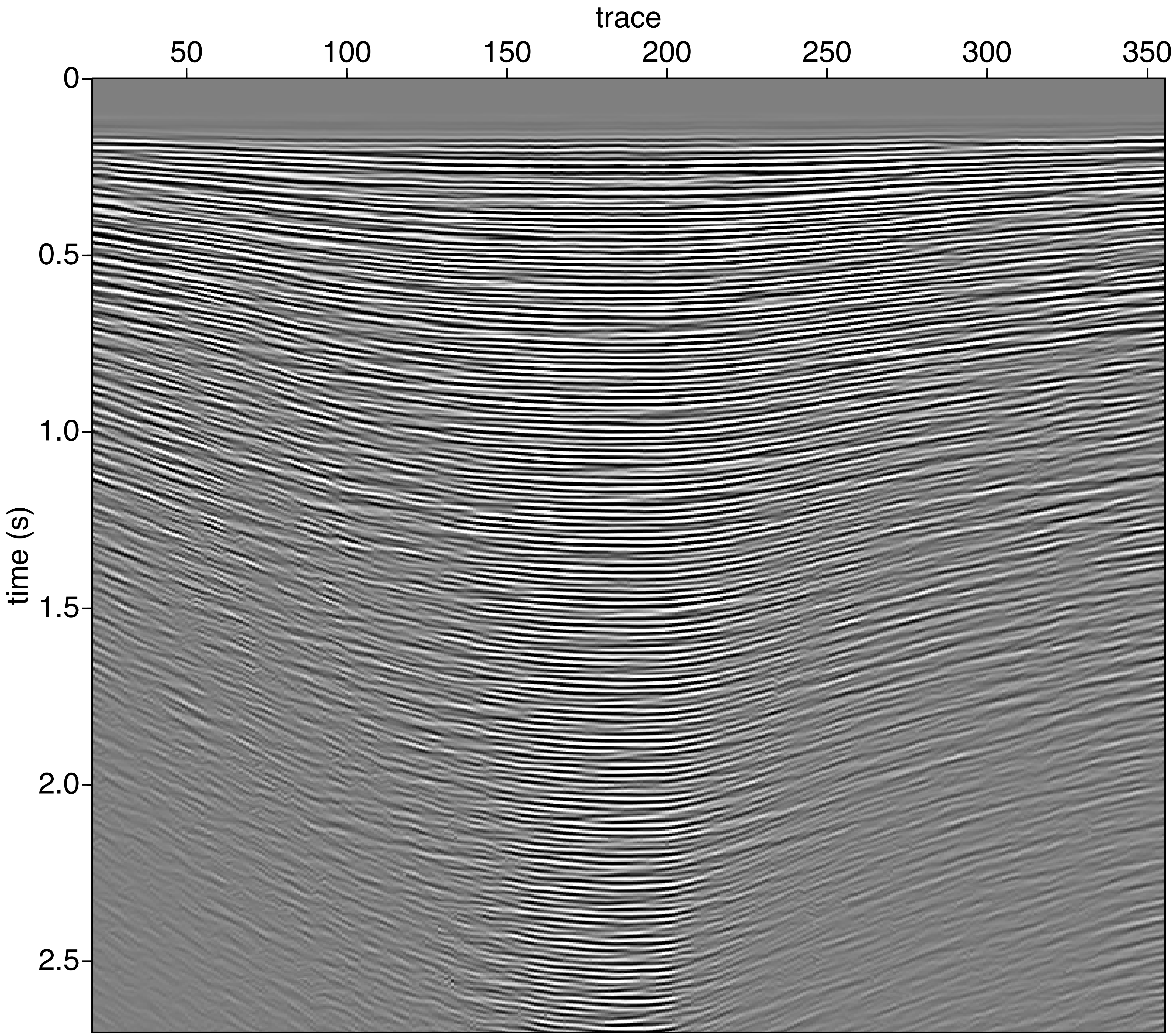
REPSI Primary IR

90 total gradient iters

Transform domain

2D Curvelet (src-rcv)

Spline a=3.0 DWT (time)

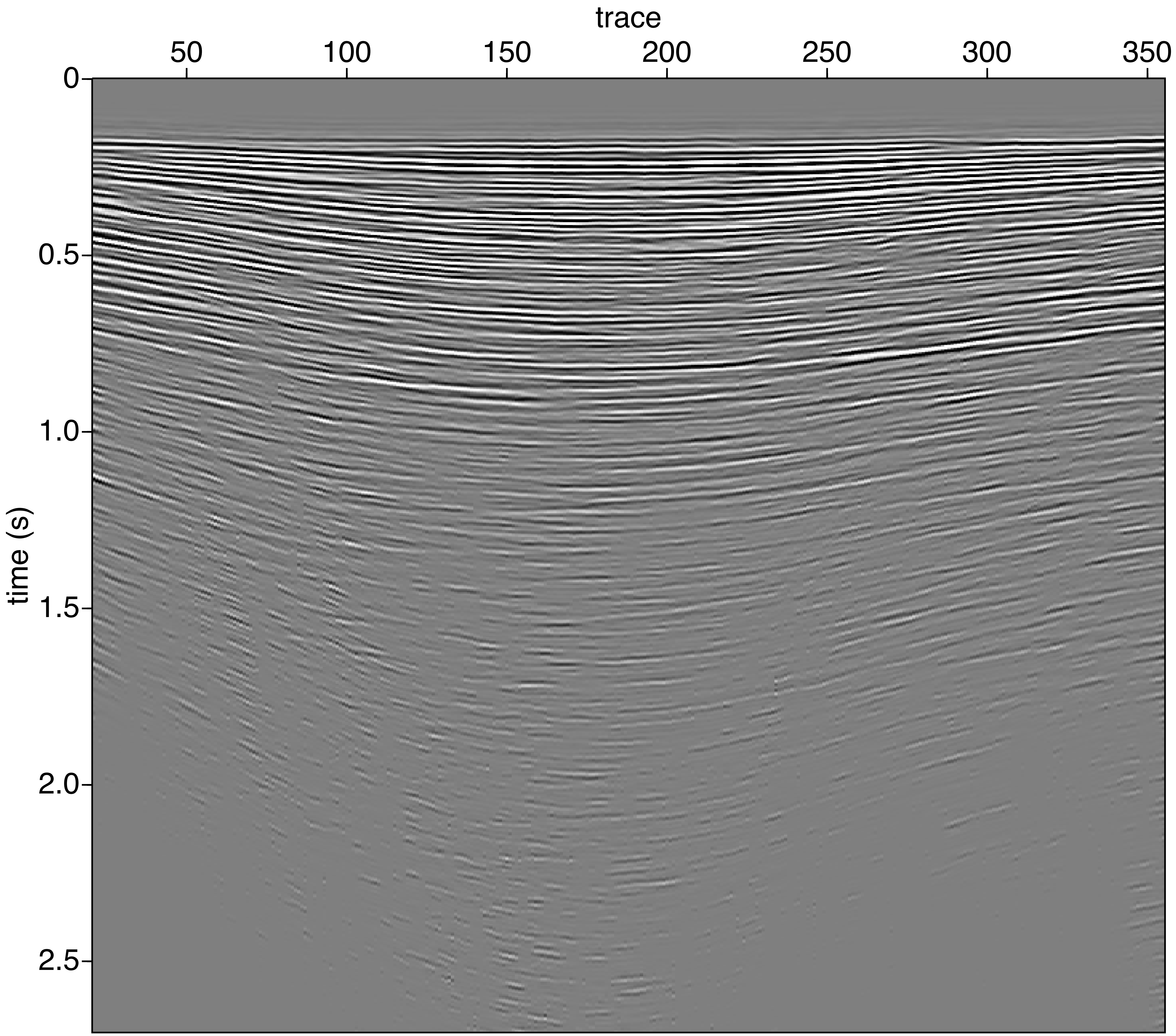


Input data

Gulf of Suez

data

NMO-corrected stack

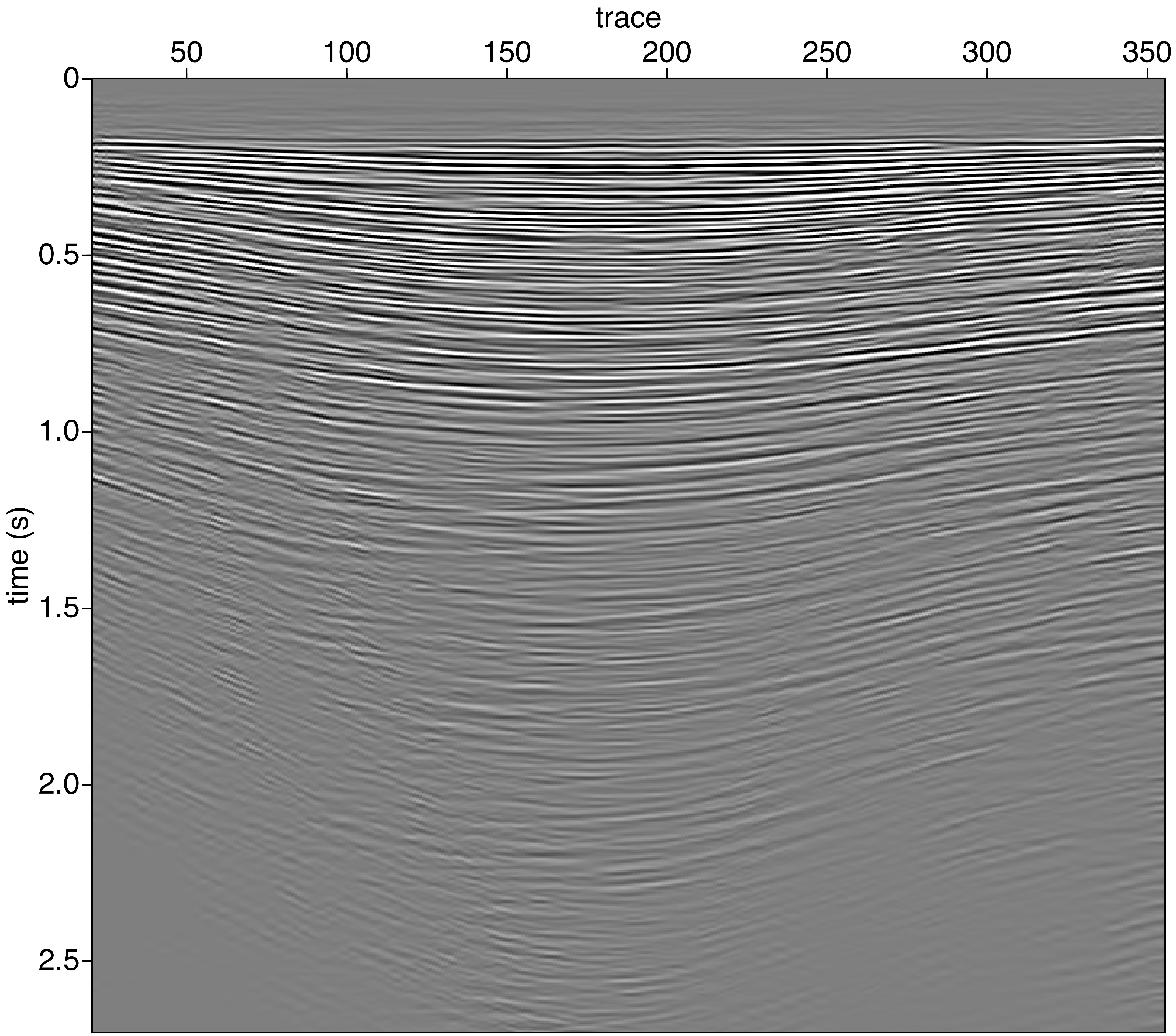


REPSI primary wavefield

Gulf of Suez

data

NMO-corrected stack



REPSI + transform domain primary wavefield

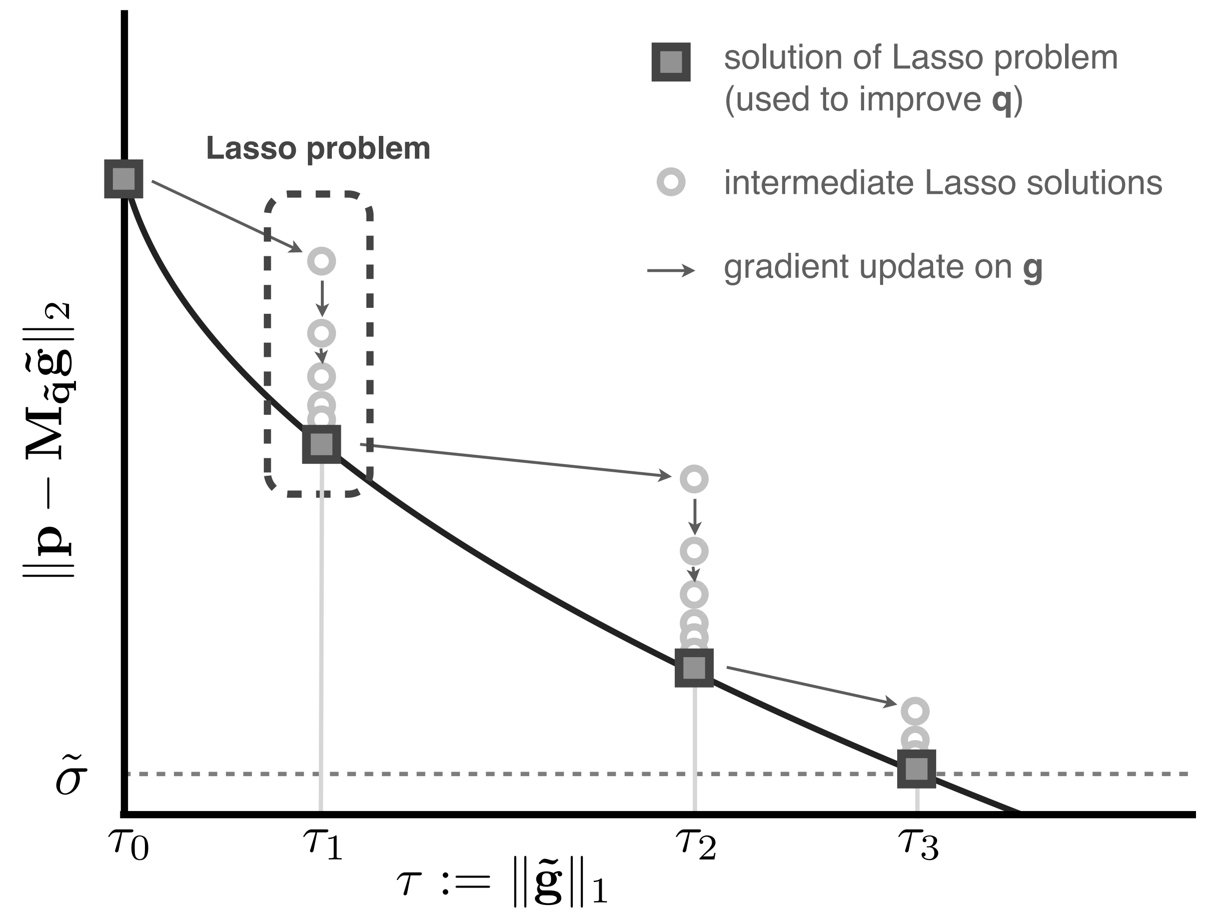
Gulf of Suez

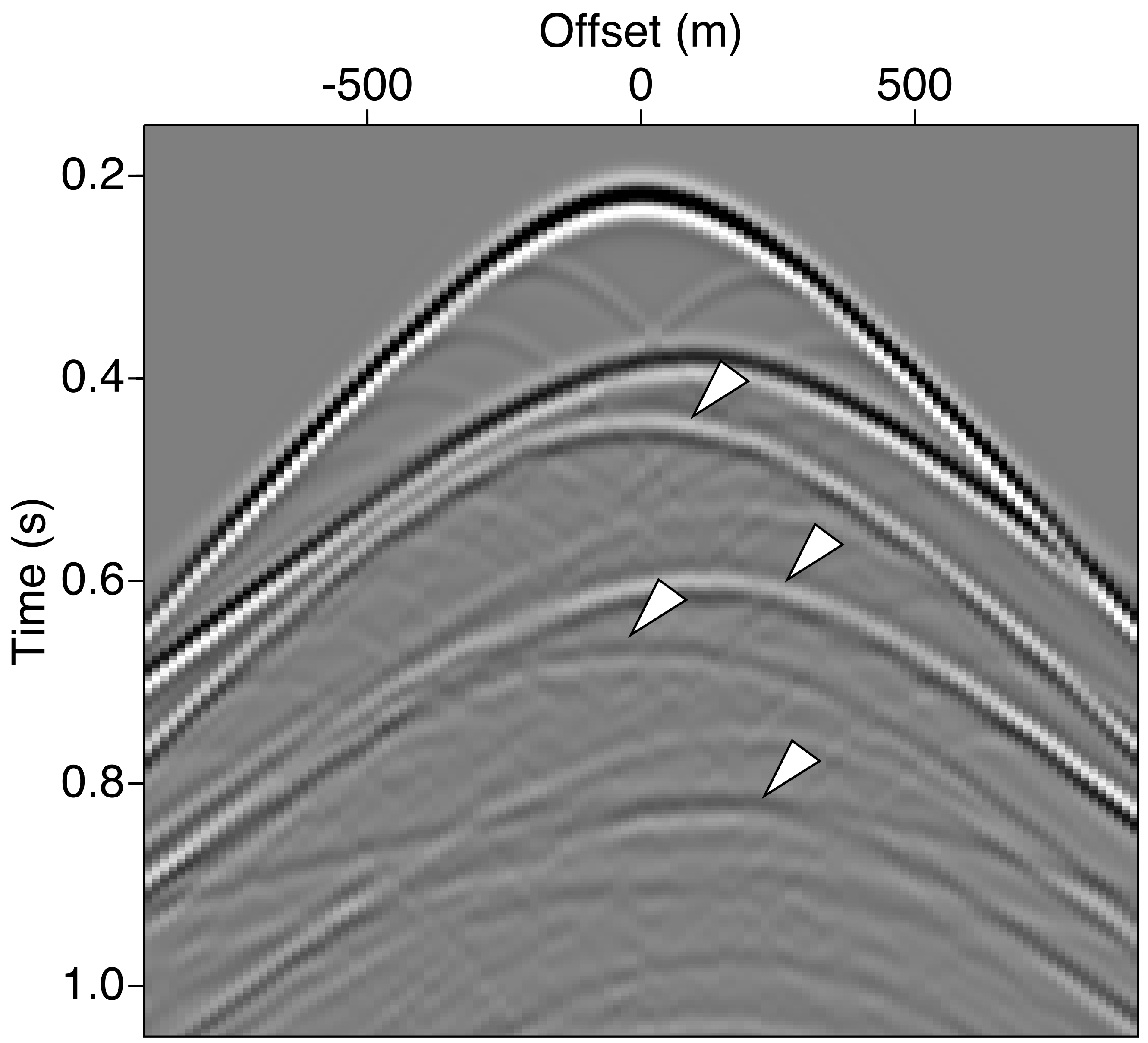
data

NMO-corrected stack

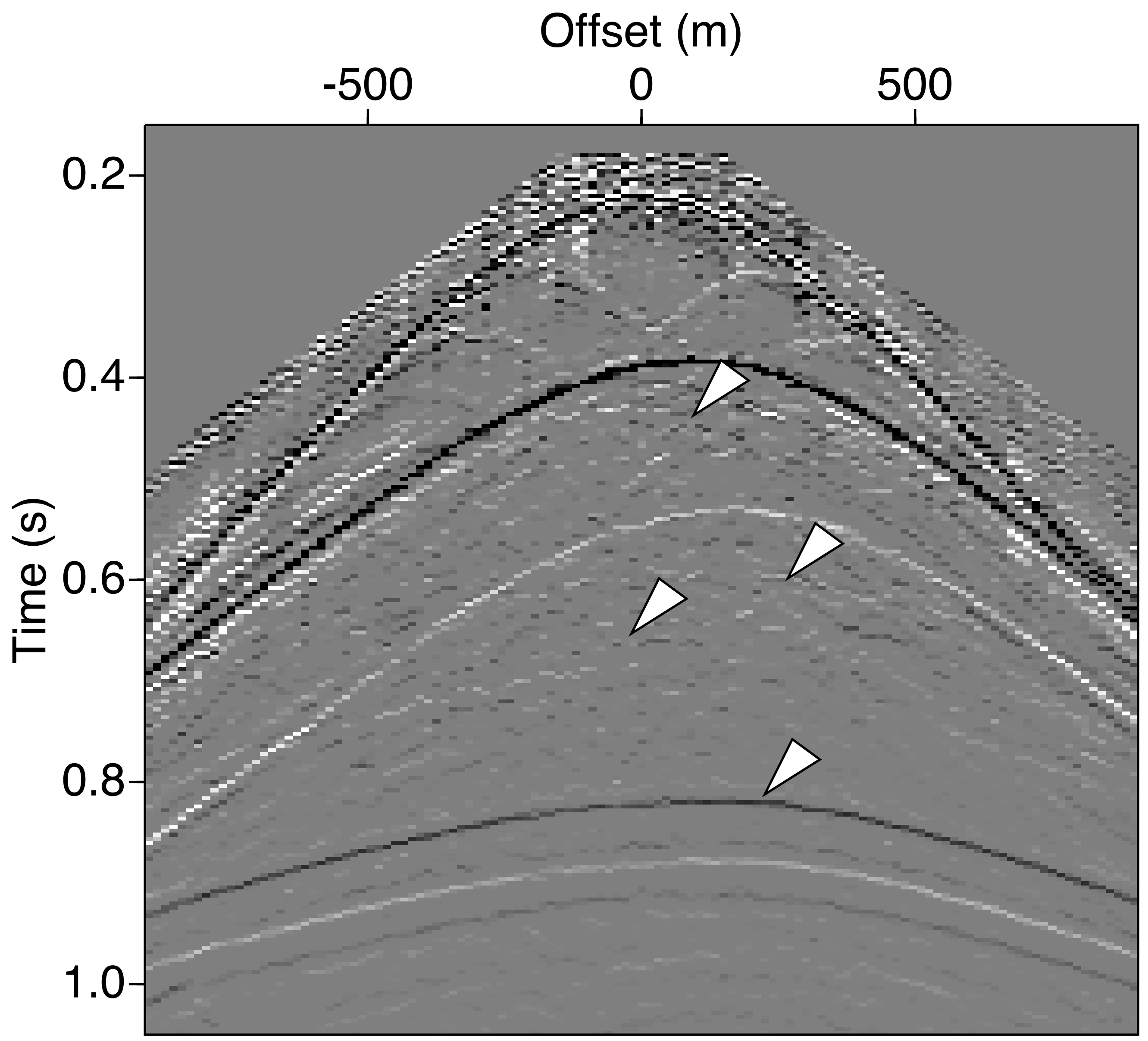
Summary: Robust EPSI

New formulation of the EPSI primary estimation problem which avoids *ad-hoc parameter adjustments* in favour of *self-tuning bi-convex optimization*

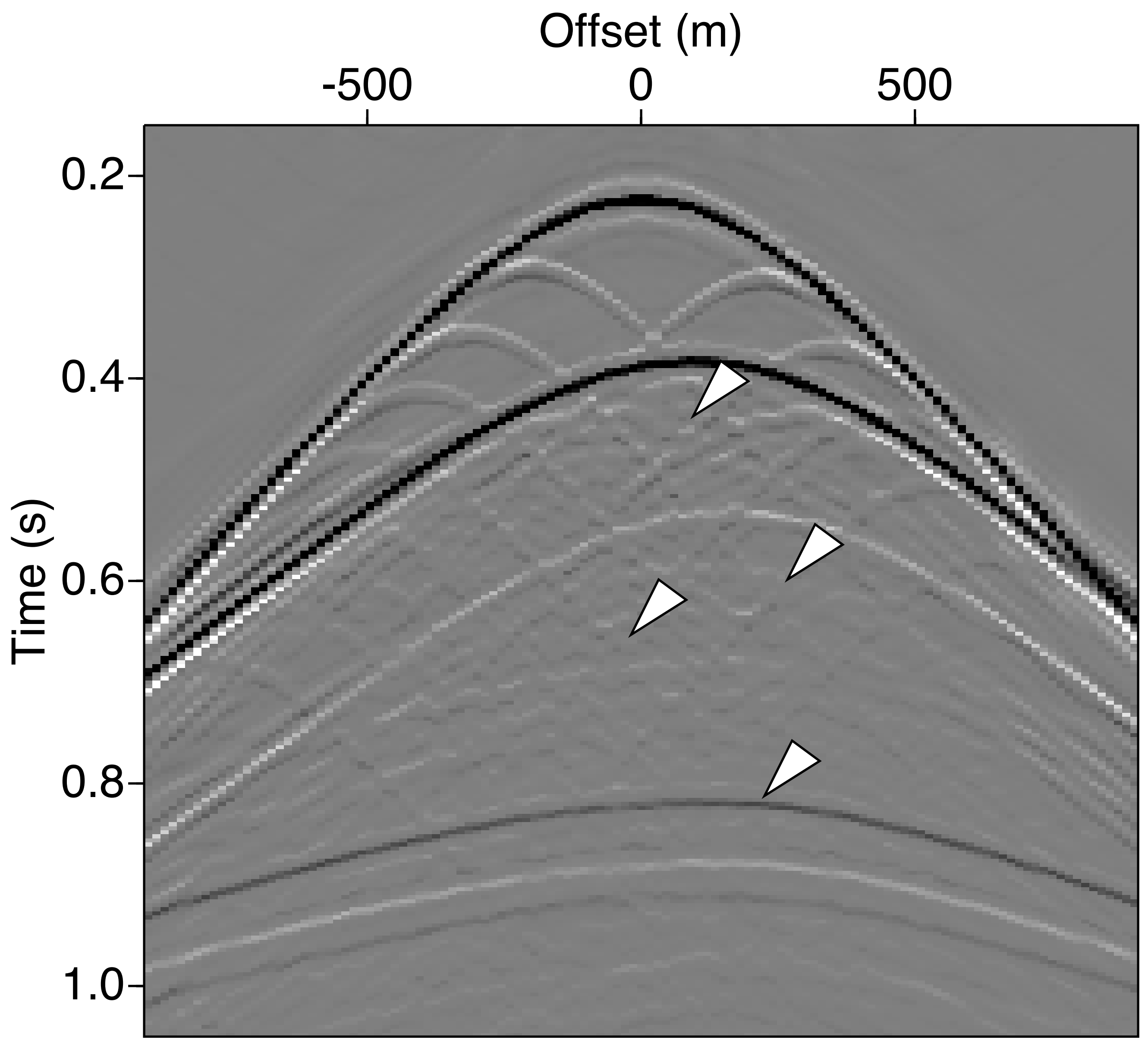




Data



EPSI Solution



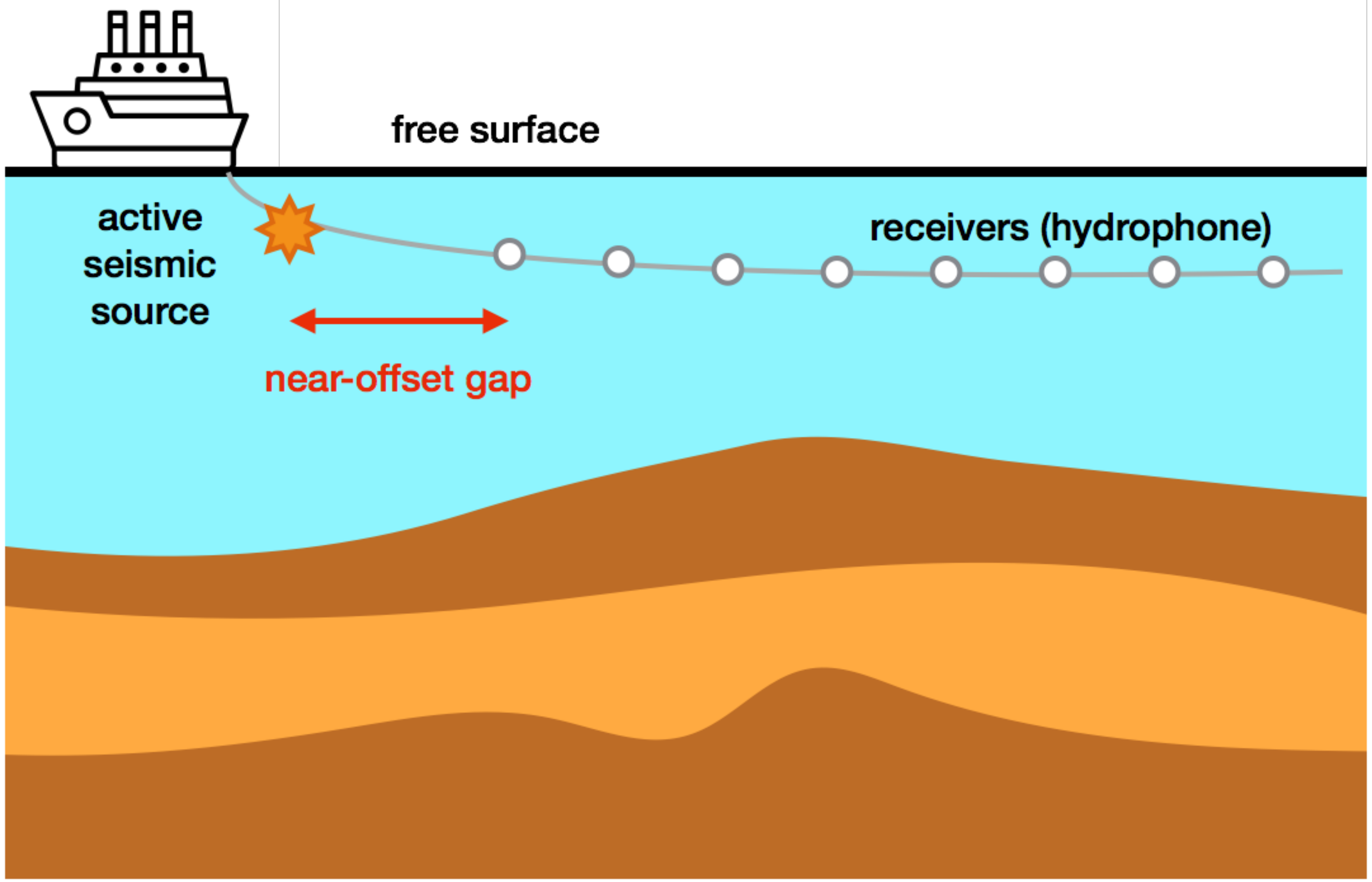
Robust EPSI Solution

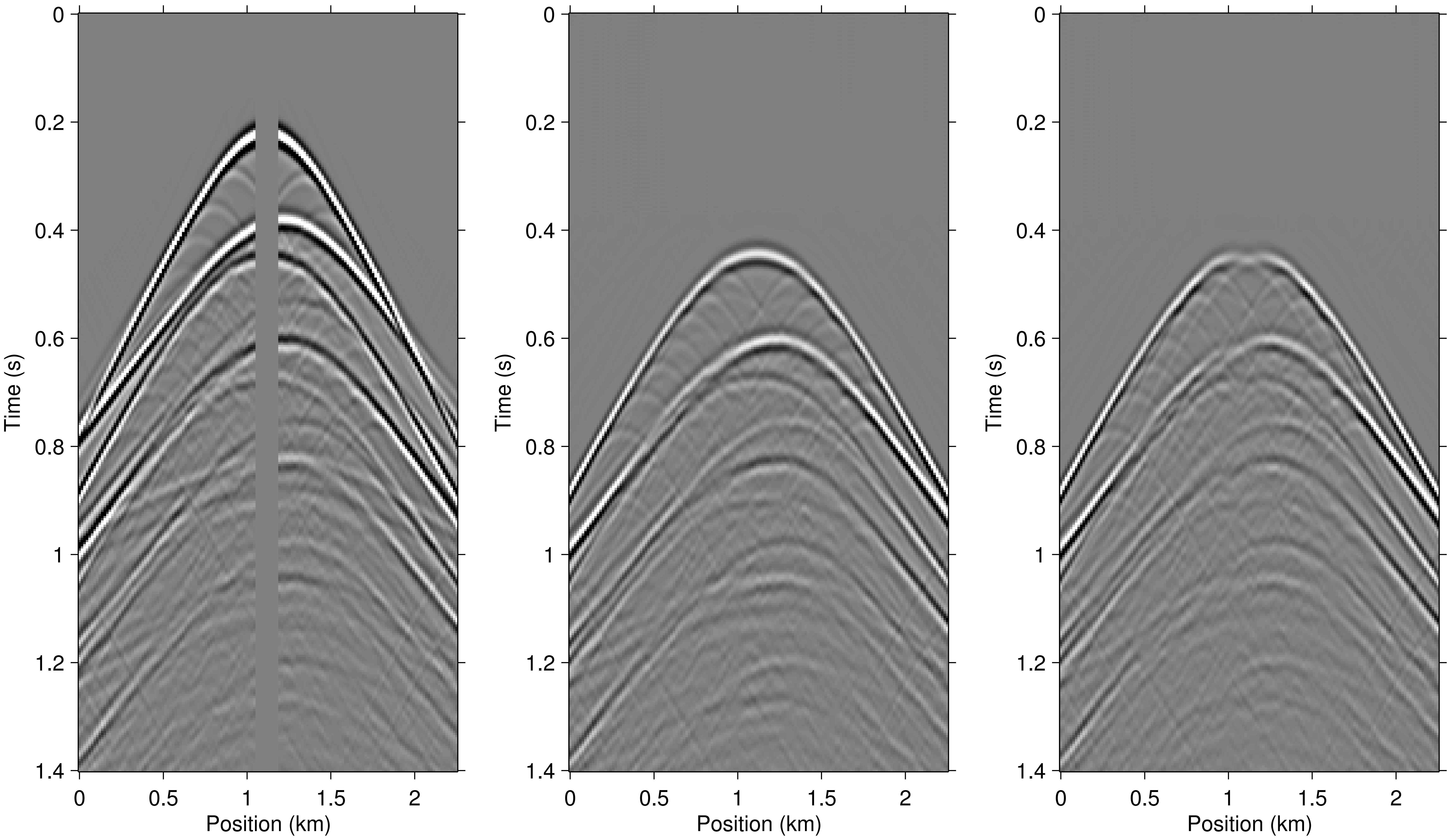
So what happens...

...when there are gaps in your data?

Chapter 4

Mitigating data gaps with scattering terms


free surface
active
seismic
source
receivers (hydrophone)
near-offset gap

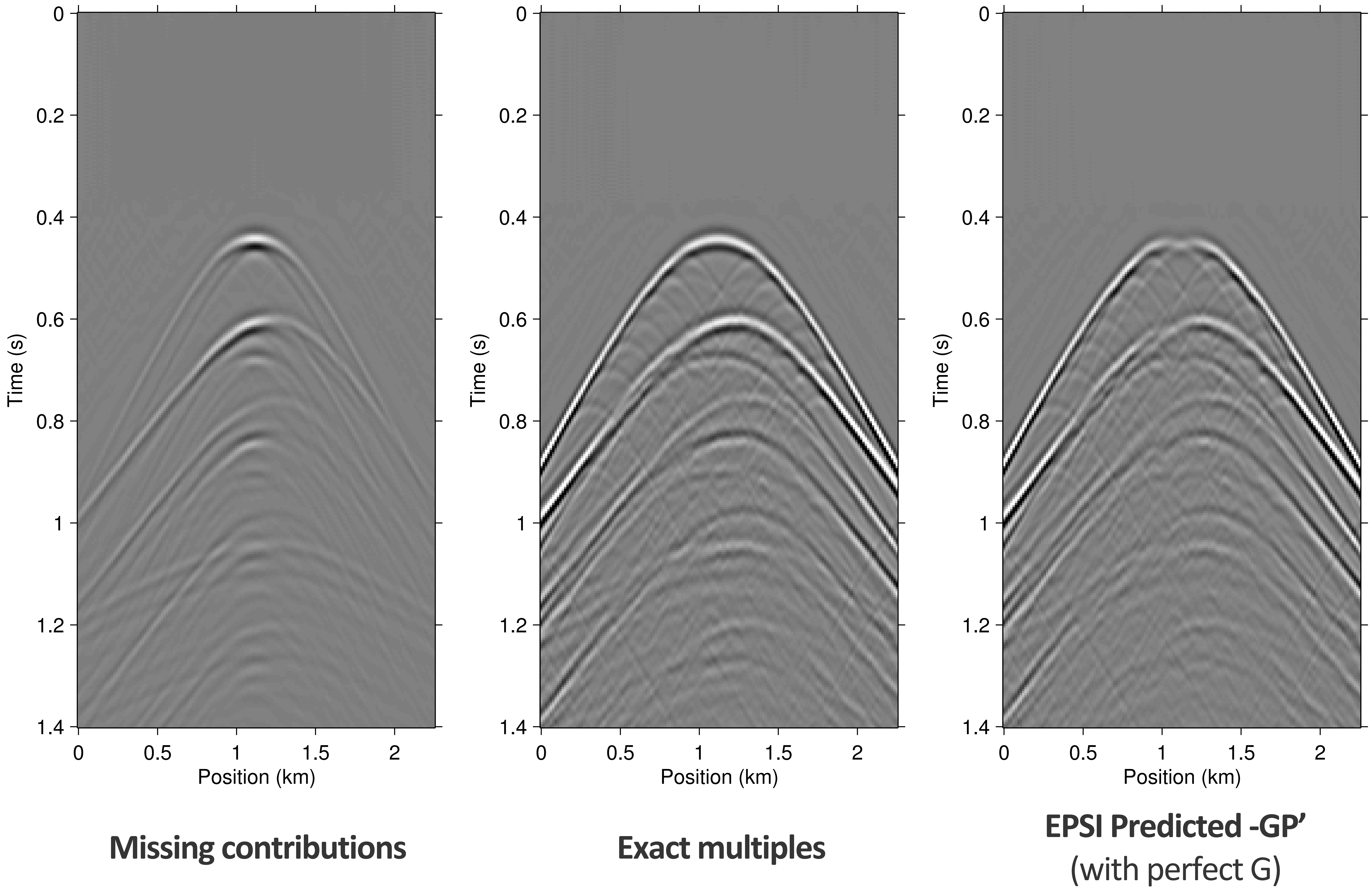


Data with missing traces P’

Exact multiples

EPSI Predicted -GP’
(with perfect G)

SLIM


Time (s)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
0
0.5
1
1.5
2
Position (km)
Missing contributions
Exact multiples
EPSI Predicted -GP'
(with perfect G)

Main idea

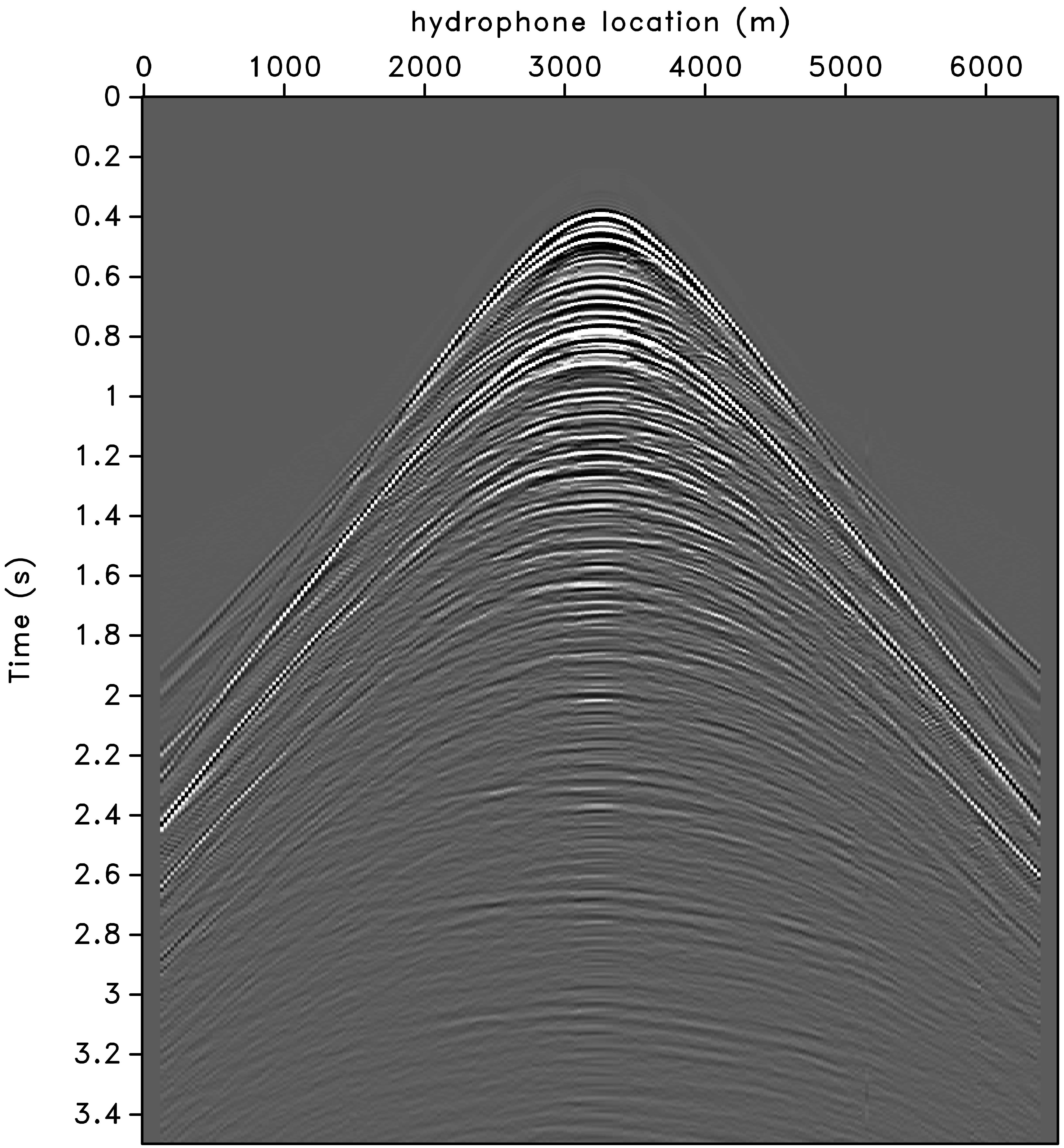
Augment the convolutional model

$$\mathbf{P} = \mathbf{GQ} + \mathbf{GRP}$$

to account for the missing contribution

Trace Mask

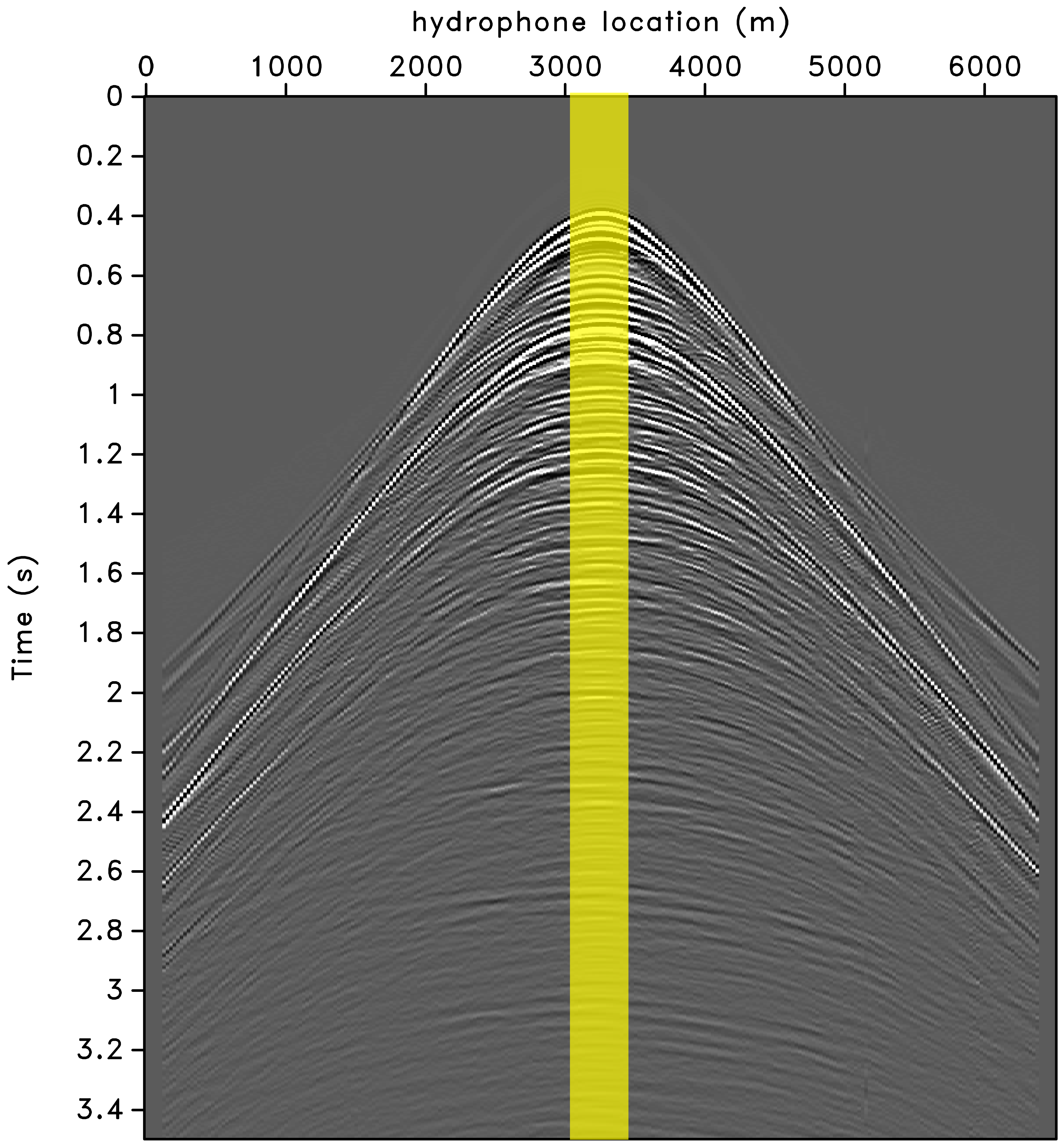
Masking operator $\mathbf{K}$



Trace Mask

Bisects wavefield data to unknown/ uncertain traces

(e.g., near-offset)

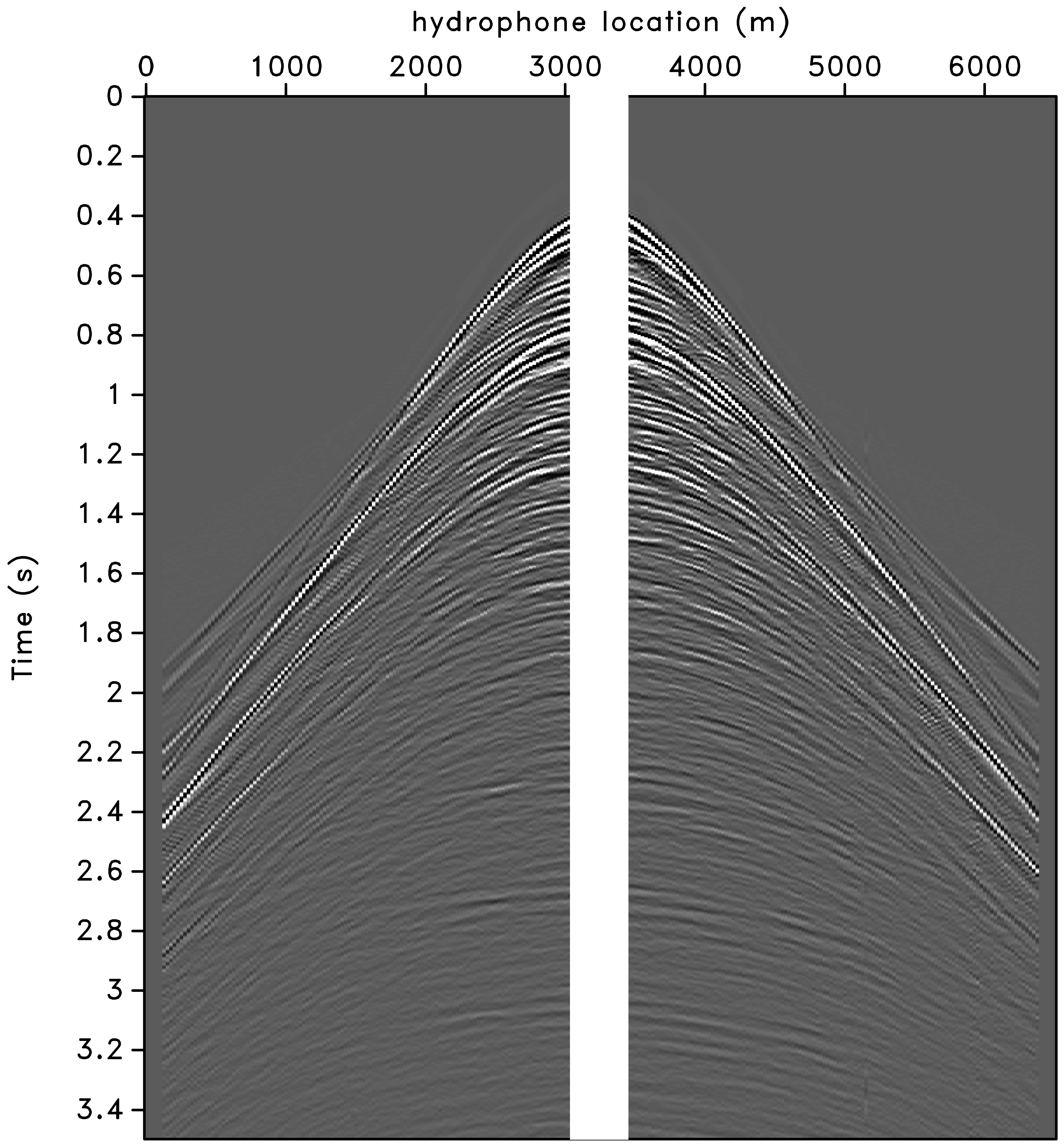


Trace Mask

Masking operator $\mathbf{K}$

Time domain: $\mathbf{K}\mathbf{p}$

Frequency slices: $\mathbf{K} \circ \mathbf{P}$

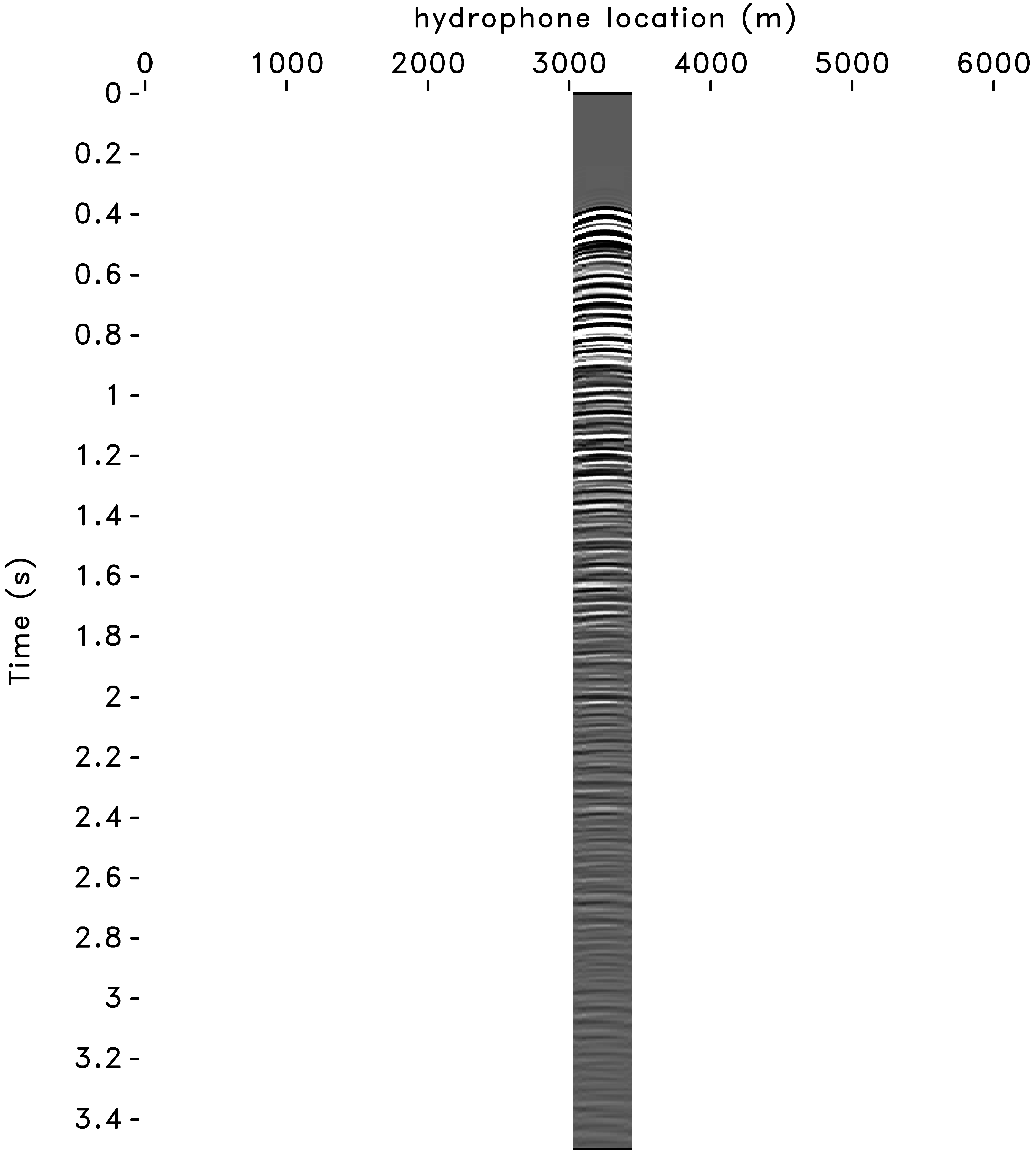


Trace Mask

Complement of Masking operator $\mathbf{K_c}$

Time domain: $\mathbf{K_c p}$

Frequency slices: $\mathbf{K_c} \circ \mathbf{P}$



Bisected data variables $\mathbf{P}' + \mathbf{P}'' = \mathbf{P}$

Known data traces: $\mathbf{P}' := \mathbf{K} \circ \mathbf{P}$

Unknown data traces: $\mathbf{P}'' := \mathbf{K_c} \circ \mathbf{P}$

Bisected data variables $\mathbf{P}' + \mathbf{P}'' = \mathbf{P}$

Known data traces: $\mathbf{P}' := \mathbf{K} \circ \mathbf{P}$

Unknown data traces: $\mathbf{P}'' := \mathbf{K_c} \circ \mathbf{P}$

$$= \mathbf{K_c} \circ \left(\mathbf{GQ} + \mathbf{RGP}' + \mathbf{RGP}''\right)$$

Bisected data variables $\mathbf{P}' + \mathbf{P}'' = \mathbf{P}$

Known data traces: $\mathbf{P}' := \mathbf{K} \circ \mathbf{P}$

Trace stencil

Unknown data traces: $\mathbf{P}'' := \boxed{\mathbf{K_c}} \circ \mathbf{P}$

$$= \mathbf{K_c} \circ \left(\mathbf{GQ} + \mathbf{RGP}' + \mathbf{RGP}''\right)$$

Bisected data variables $\mathbf{P}' + \mathbf{P}'' = \mathbf{P}$

Known data traces: $\mathbf{P}' := \mathbf{K} \circ \mathbf{P}$

Trace stencil

Unknown data traces: $\boxed{\mathbf{P}''} := \boxed{\mathbf{K_c}} \circ \mathbf{P}$

$$= \mathbf{K_c} \circ \left(\mathbf{GQ} + \mathbf{RGP}' + \mathbf{RG}\boxed{\mathbf{P}''}\right)$$

Recursively defined

Modify the modeling operator

$$\mathcal{M}(\mathbf{G}, \mathbf{Q}; \mathbf{P}') = \mathbf{G}\mathbf{Q} + \mathbf{R}\mathbf{G}\mathbf{P}' + \mathbf{R}\mathbf{G}\mathbf{P}''$$

Modify the modeling operator

$$\mathcal{M}(\mathbf{G},\mathbf{Q};\mathbf{P}') = \mathbf{G}\mathbf{Q} + \mathbf{R}\mathbf{G}\mathbf{P}' + \mathbf{R}\mathbf{G}\boxed{\mathbf{P}''}$$

Tilde: modified with higher-order terms

$$\widetilde{\mathcal{M}}(\mathbf{G},\mathbf{Q};\mathbf{P}') = \mathbf{G}\mathbf{Q} + \mathbf{R}\mathbf{G}\mathbf{P}' + \mathbf{R}\mathbf{G}\boxed{\mathbf{K_c} \circ (\mathbf{G}\mathbf{Q} + \mathbf{R}\mathbf{G}\mathbf{P}')} + \mathcal{O}(\mathbf{G}^3)$$

SLIM

SLIM

Modify the modeling operator

$$\widetilde{\mathcal{M}}(\mathbf{G},\mathbf{Q};\mathbf{P}') = \mathbf{GQ} + \mathbf{RGP}' \\ + \mathbf{RGK_c} \circ (\mathbf{GQ} + \mathbf{RGP}') \\ + \mathcal{O}(\mathbf{G}^3)$$

Modify the modeling operator

$$
\begin{aligned}
\widetilde{\mathcal{M}}(\mathbf{G}, \mathbf{Q}; \mathbf{P}') = & \mathbf{G}\mathbf{Q} + \mathbf{R}\mathbf{G}\mathbf{P}' \\
& + \mathbf{R}\boxed{\mathbf{G}}\mathbf{K_c} \circ (\boxed{\mathbf{G}}\mathbf{Q} + \mathbf{R}\boxed{\mathbf{G}}\mathbf{P}') \\
& + \mathcal{O}(\mathbf{G}^3)
\end{aligned}
$$

Modify the modeling operator

$$\widetilde{\mathcal{M}}(\mathbf{G},\mathbf{Q};\mathbf{P}') = \mathbf{GQ} + \mathbf{RGP}' + \mathbf{RGK_c} \circ (\mathbf{GQ} + \mathbf{RGP}') + \mathcal{O}(\mathbf{G}^3)$$

2nd Order autoconvolution term (of G)

Modify the modeling operator

Trace mask over all modeled wavefield

2nd Order autoconvolution term

$$
\begin{aligned}
\widetilde{\mathcal{M}}(\mathbf{G}, \mathbf{Q}; \mathbf{P}') = \boxed{\mathbf{K} \circ} \big[& \mathbf{G}\mathbf{Q} + \mathbf{R}\mathbf{G}\mathbf{P}' \\
+ \; & \mathbf{R}\mathbf{G}\mathbf{K_c} \circ (\mathbf{G}\mathbf{Q} + \mathbf{R}\mathbf{G}\mathbf{P}') \\
+ \; & \mathcal{O}(\mathbf{G}^3) \big]
\end{aligned}
$$

Modify the modeling operator

$$\widetilde{\mathcal{M}}(\mathbf{G},\mathbf{Q};\mathbf{P}') = \mathbf{K} \circ \Big[\mathbf{GQ} + \mathbf{RGP}'$$

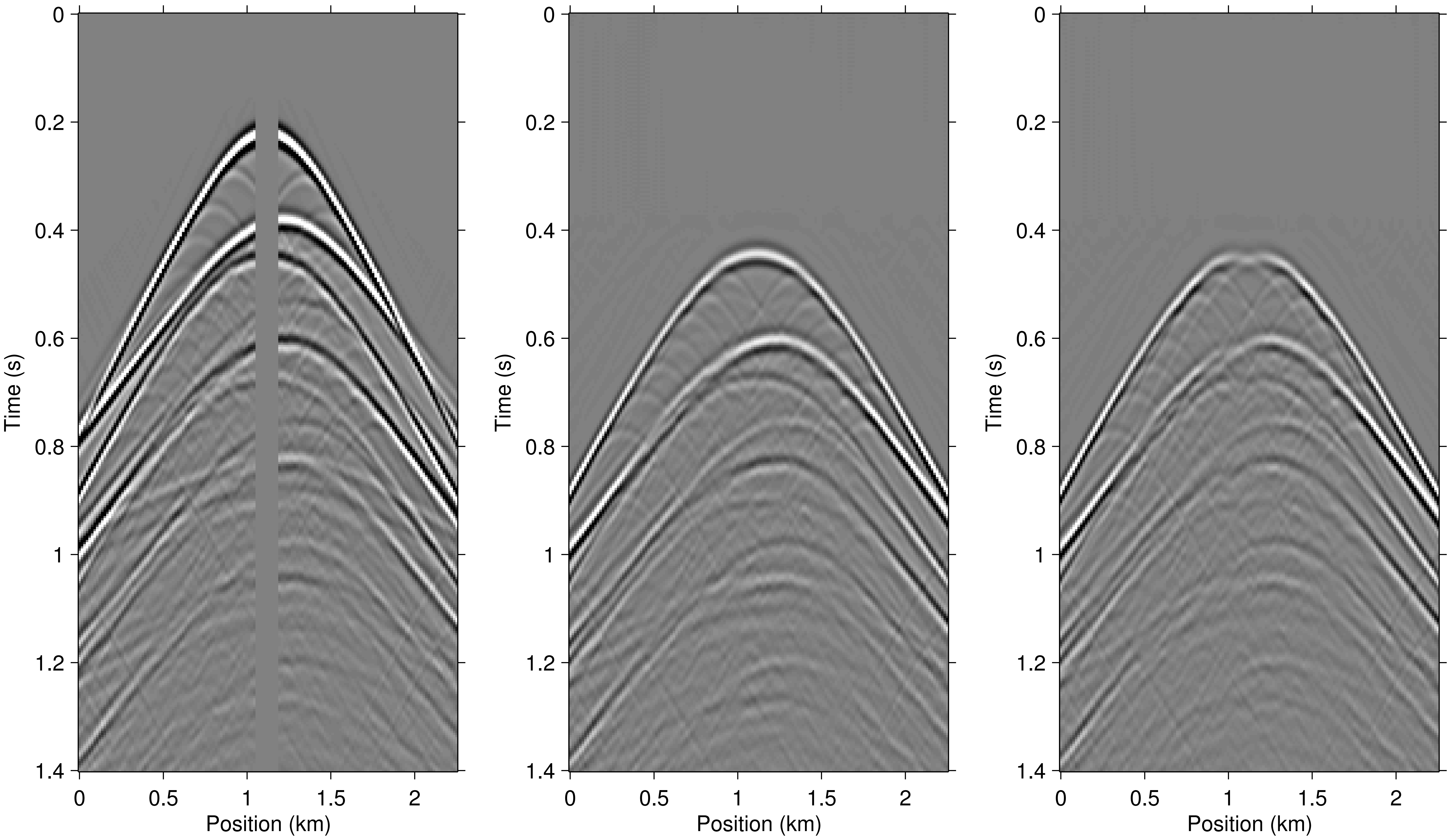
2nd Order autoconvolution term $\quad + \ \mathbf{RGK_c} \circ (\mathbf{GQ} + \mathbf{RGP}')$

3rd Order autoconvolution term $\quad + \ \mathbf{RGK_c} \circ (\mathbf{RGK_c} \circ (\mathbf{GQ} + \mathbf{RGP}'))$

$$+ \ \mathcal{O}(\mathbf{G}^4)\Big]$$

$$:= \mathbf{K} \circ \sum_{n=0}^{\infty} (\mathbf{RGK_c}\circ)^n \, (\mathbf{GQ} + \mathbf{RGP}')\,.$$

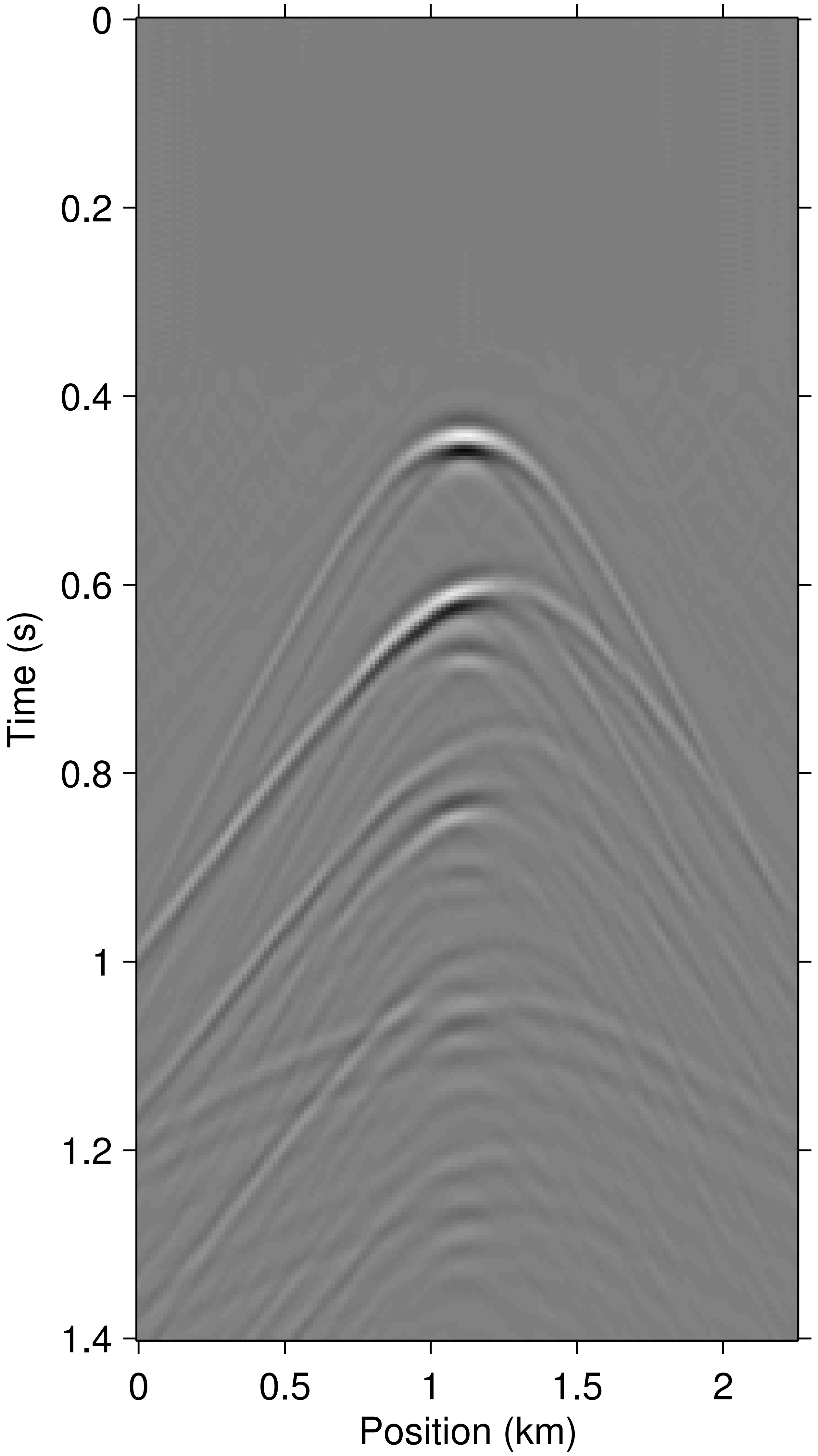
What these terms look like



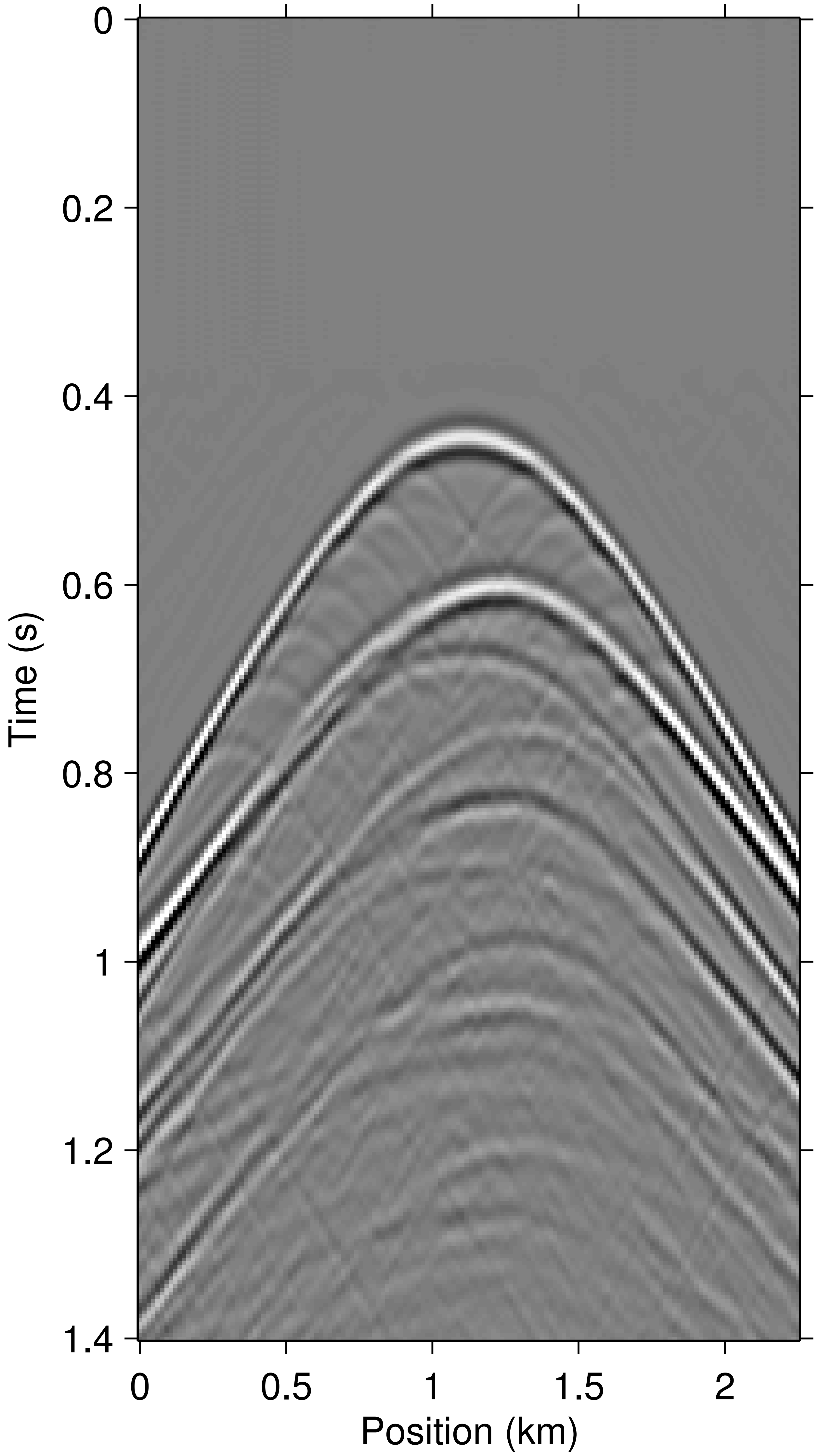
Data with missing traces P’

Exact multiples

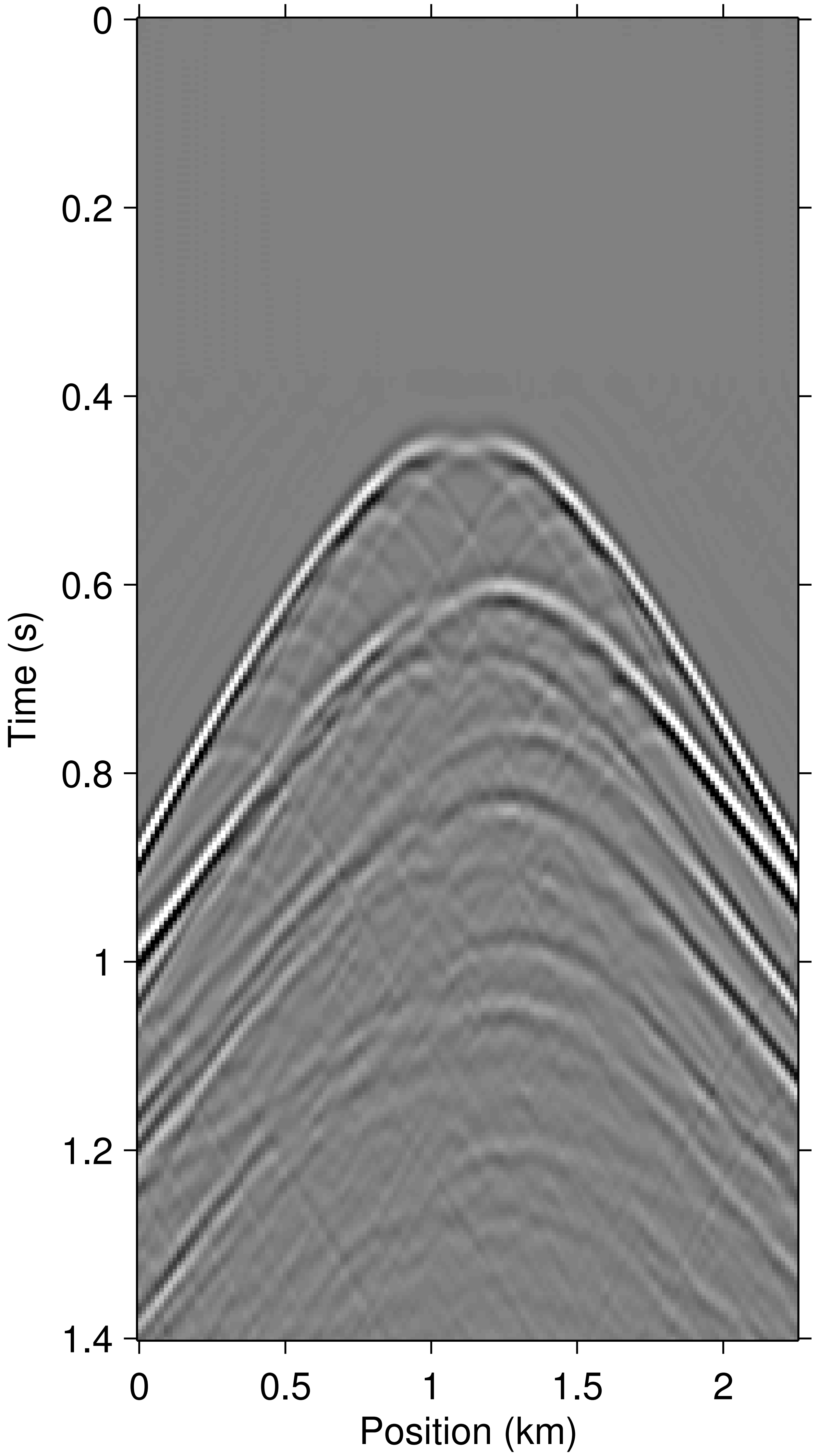
EPSI Predicted -GP’
(with perfect G)



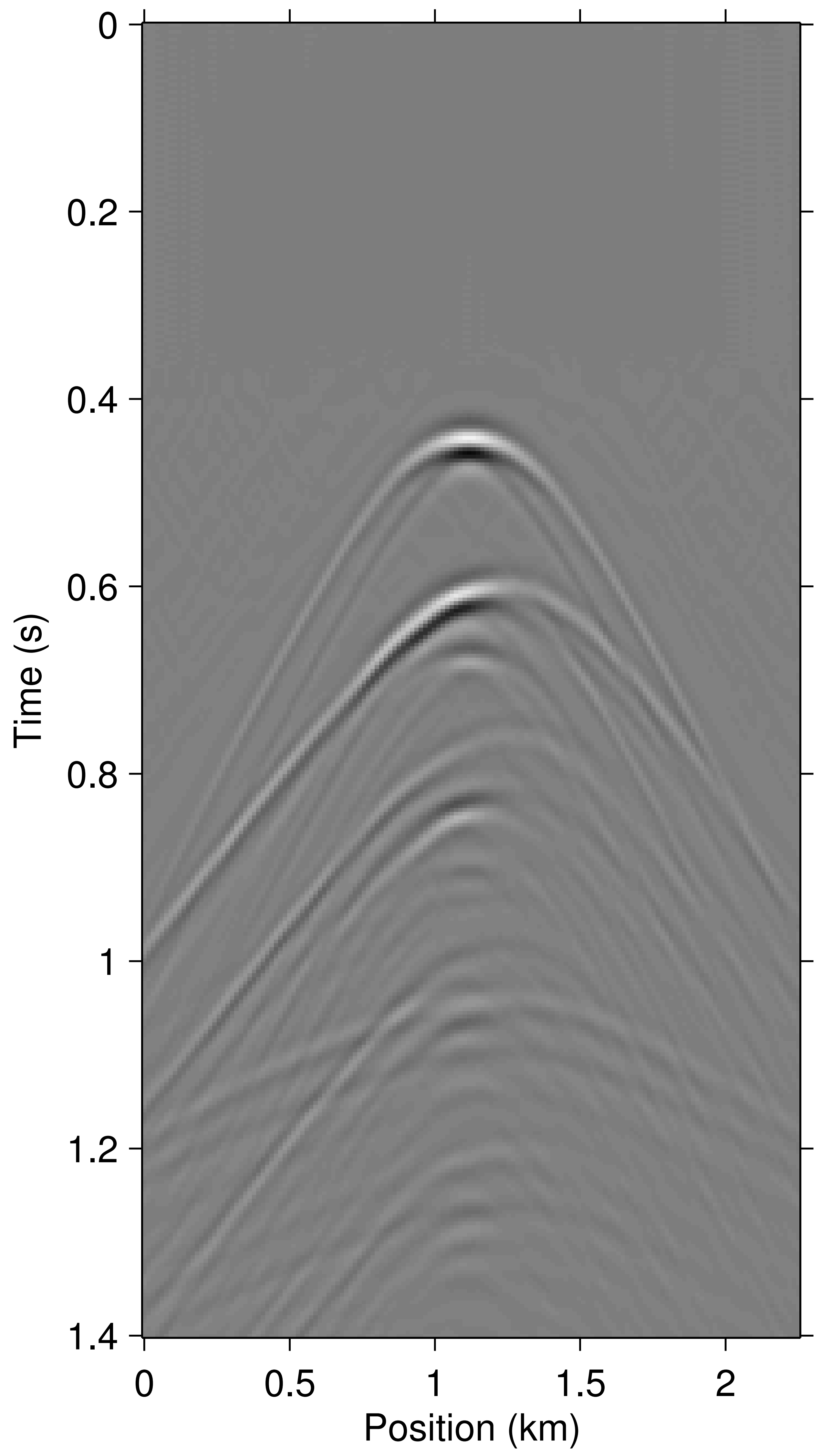
Missing contributions



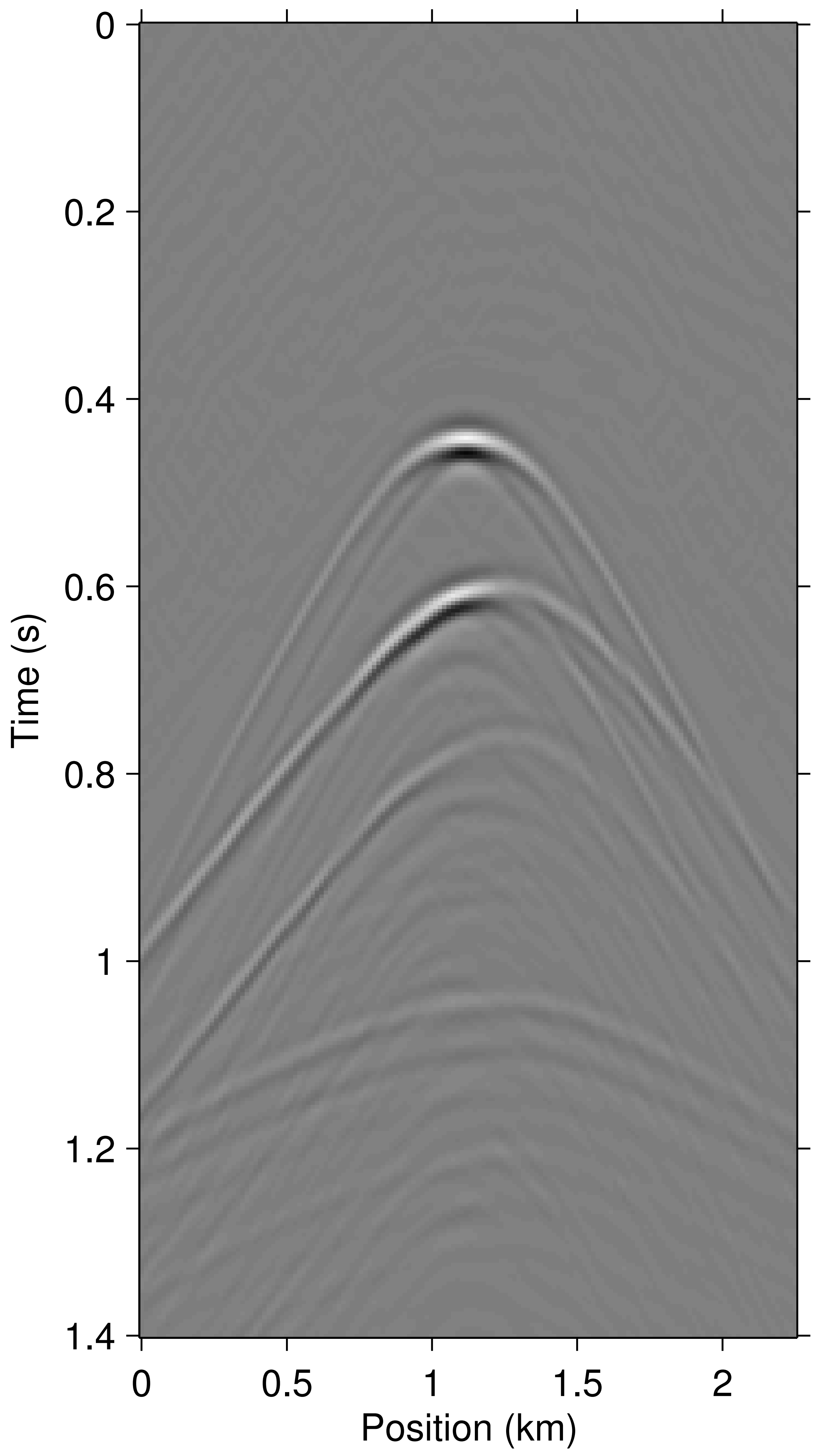
Exact multiples



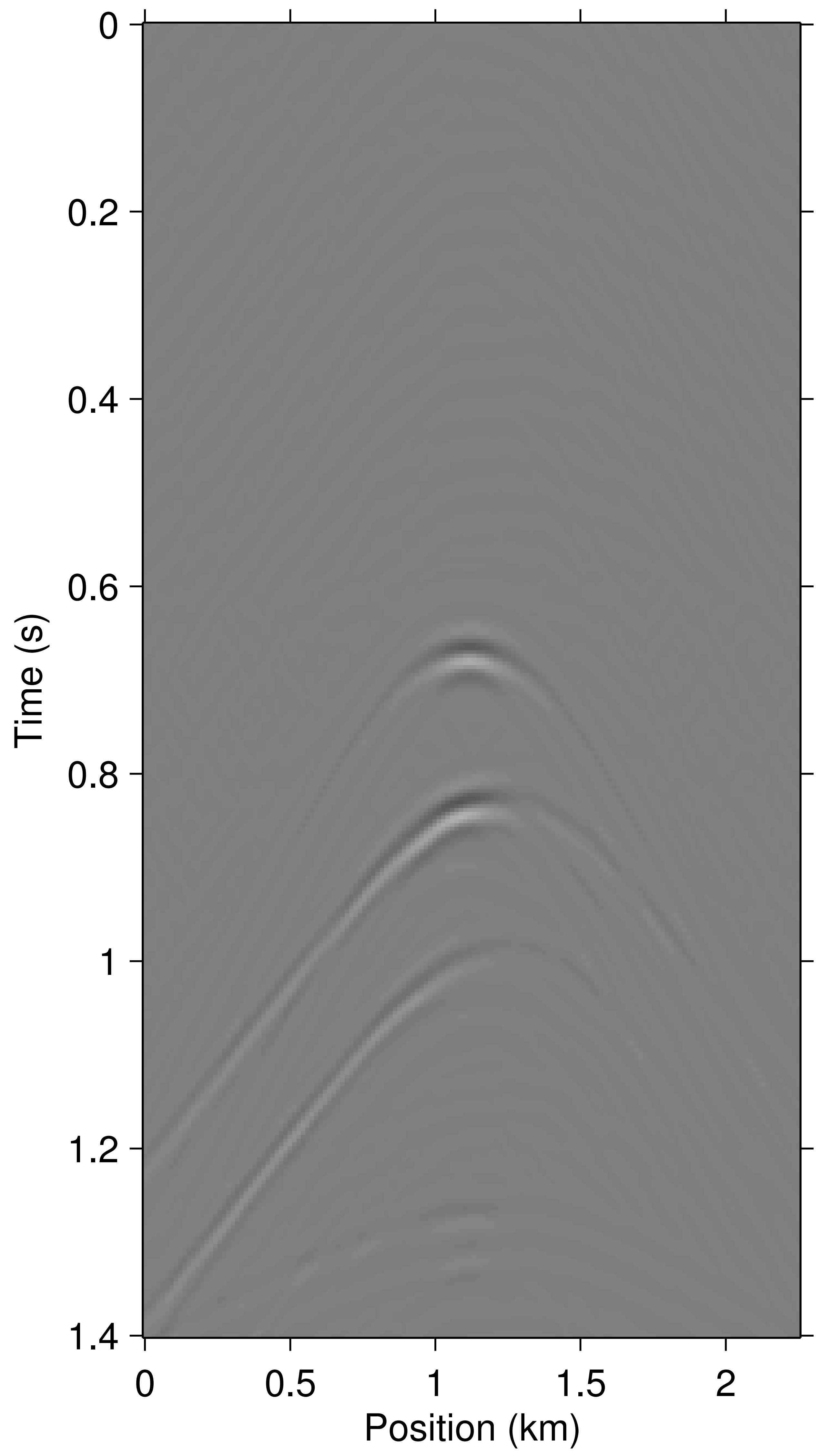
EPSI Predicted -GP’
(with perfect G)



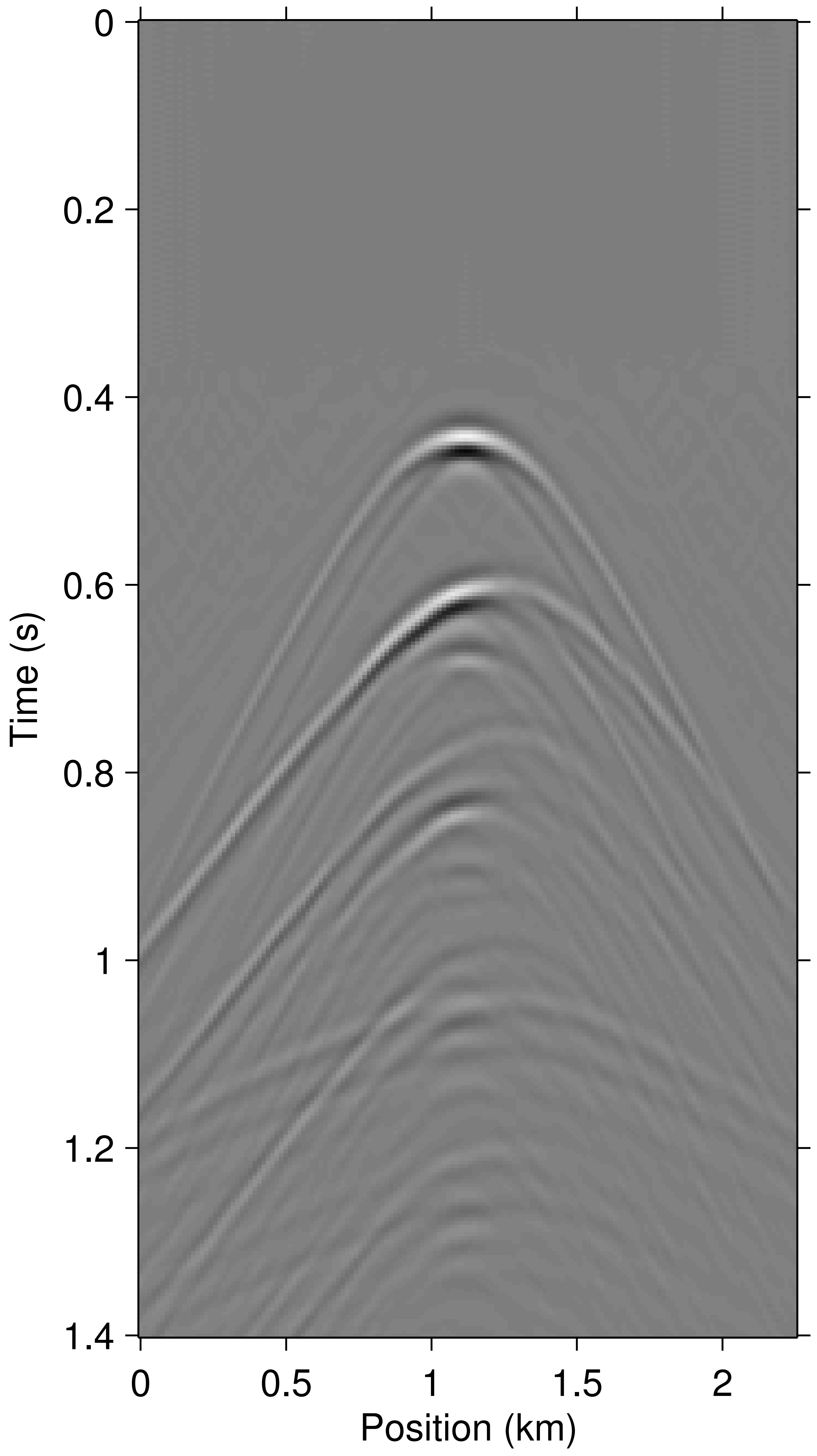
Missing contributions



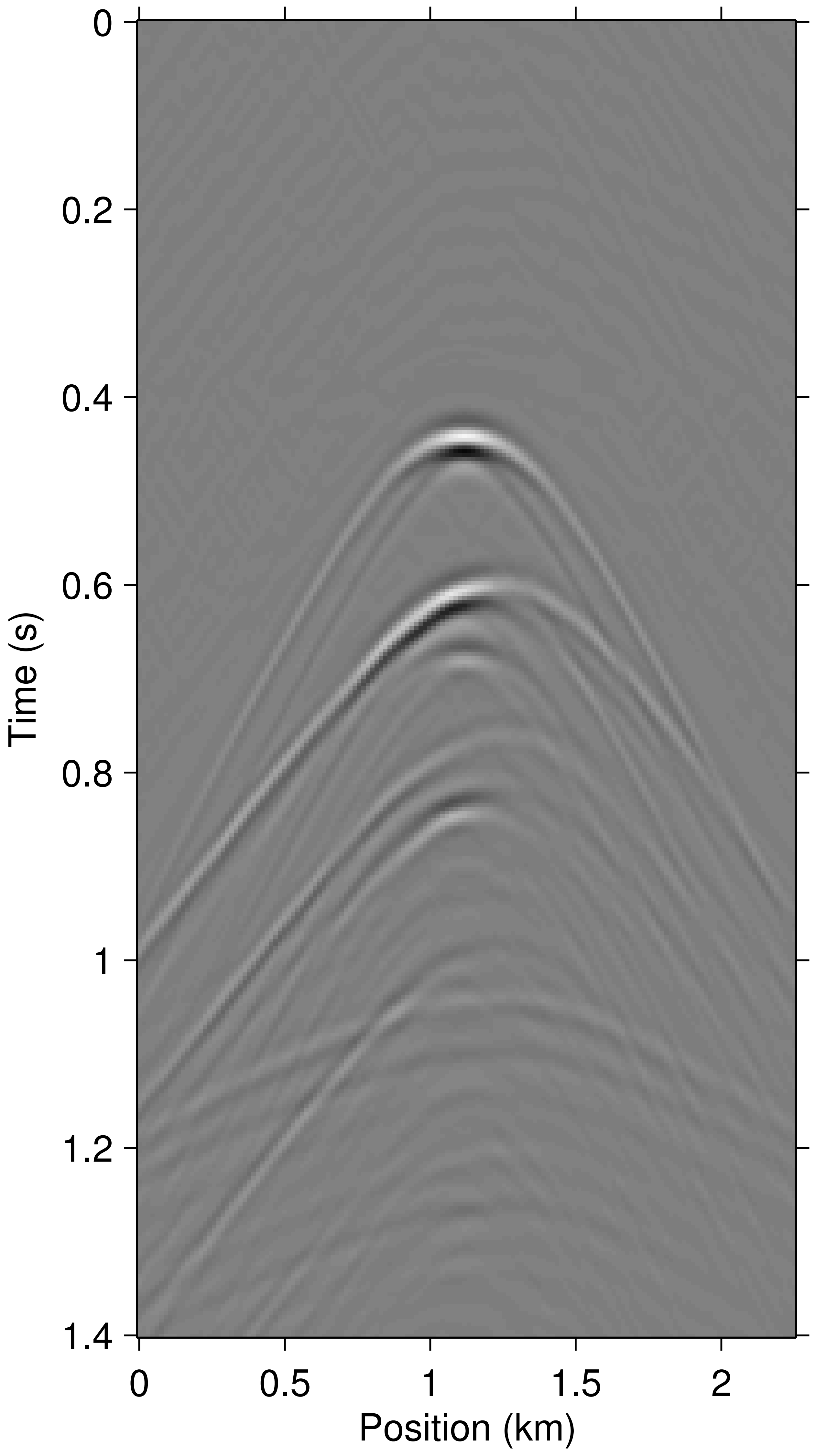
2nd order term



3rd order term

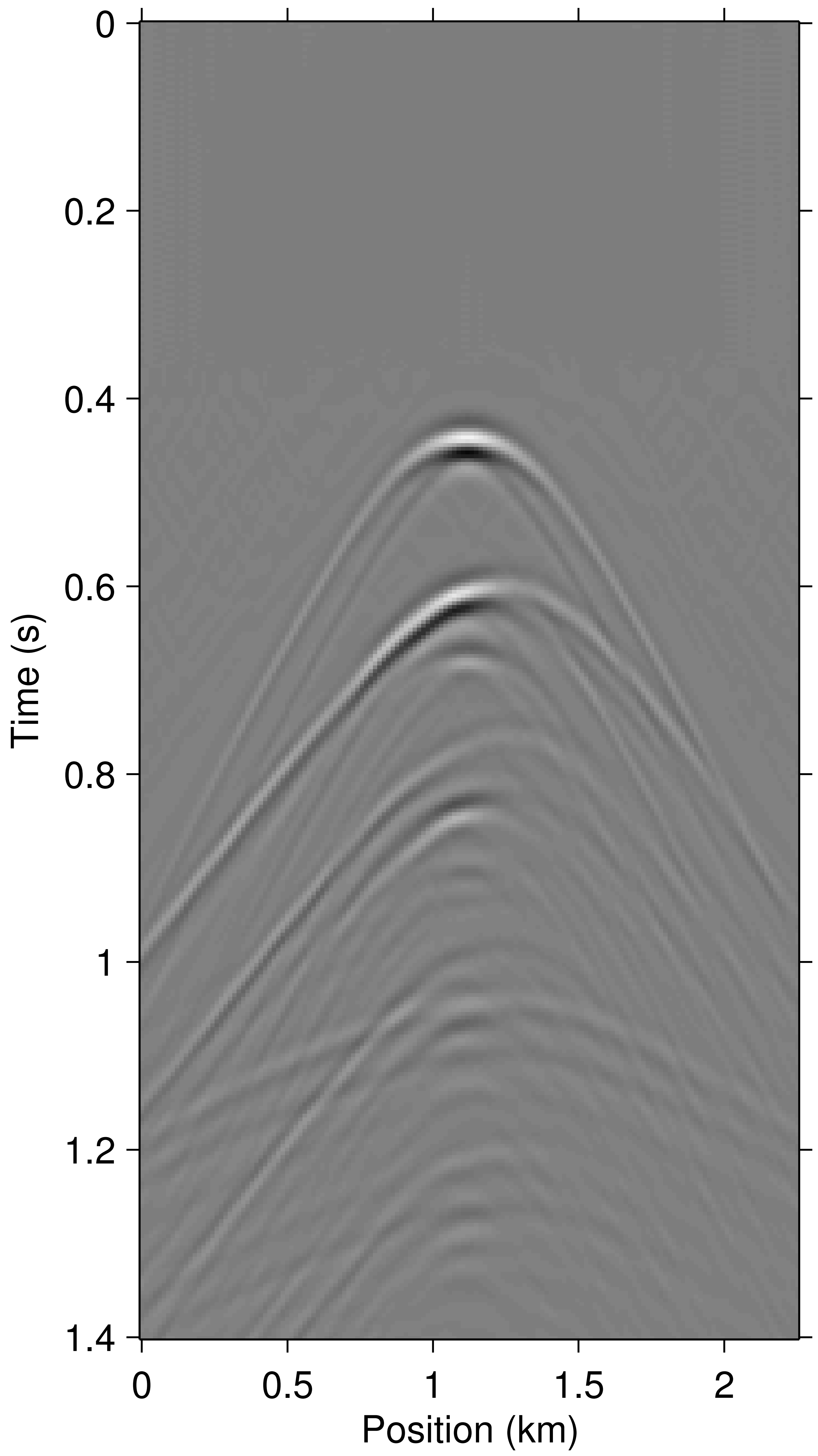


Missing contributions

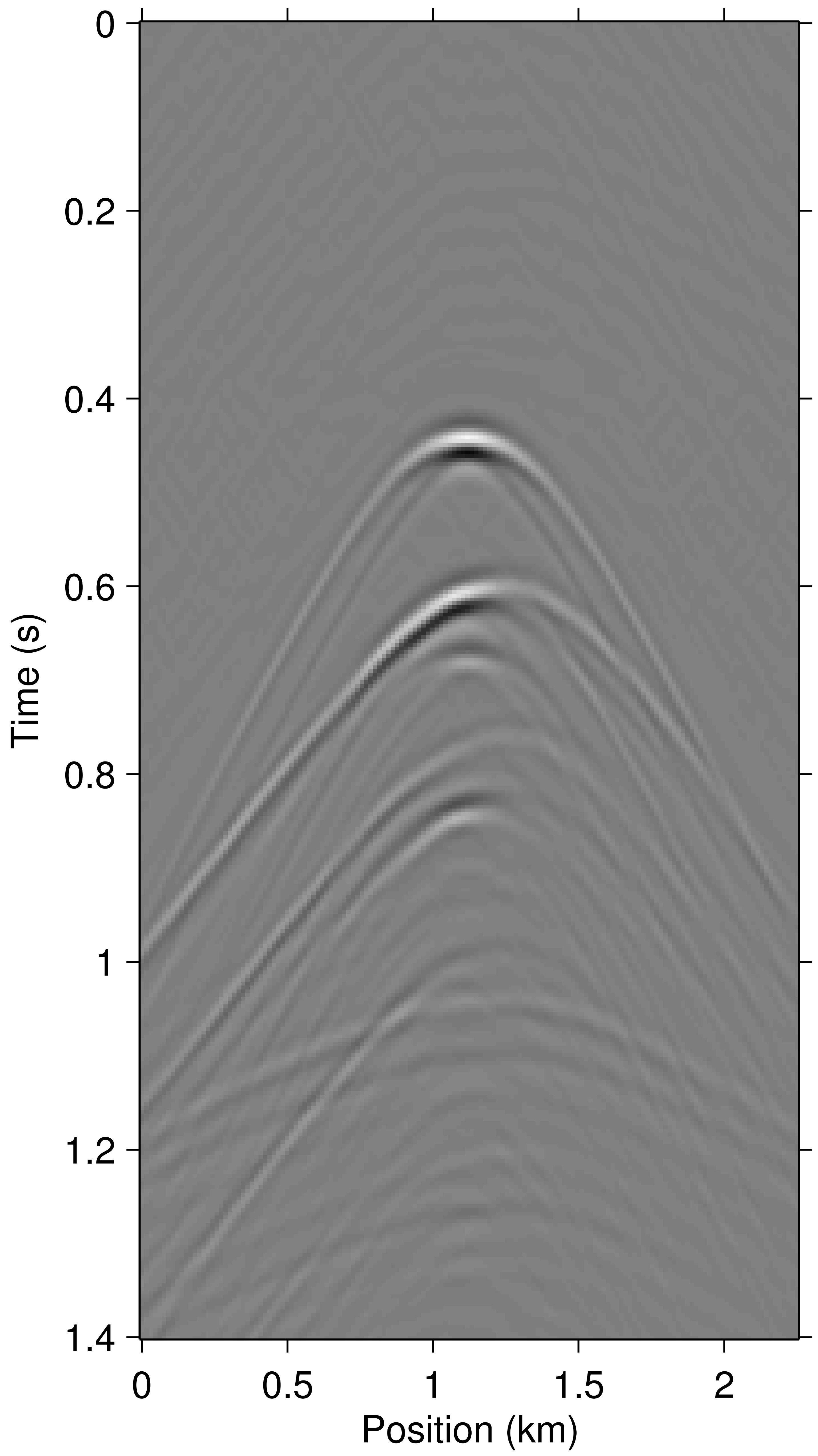


2nd order term + 3rd order

Total missing contrib.



Modeled with two terms



Main Result

Just one or two of these terms is enough to account for missing traces

Modify the modeling operator

$$
\begin{aligned}
\widetilde{\mathcal{M}}(\mathbf{G},\mathbf{Q};\mathbf{P}') =& \mathbf{K}\circ\Big[\mathbf{G}\mathbf{Q}+\mathbf{R}\mathbf{G}\mathbf{P}' \\
&+\ \mathbf{R}\mathbf{G}\mathbf{K_c}\circ(\mathbf{G}\mathbf{Q}+\mathbf{R}\mathbf{G}\mathbf{P}') \\
&+\ \mathbf{R}\mathbf{G}\mathbf{K_c}\circ(\mathbf{R}\mathbf{G}\mathbf{K_c}\circ(\mathbf{G}\mathbf{Q}+\mathbf{R}\mathbf{G}\mathbf{P}')) \\
&+\ \mathcal{O}(\mathbf{G}^4)\Big] \\
:=& \mathbf{K}\circ\sum_{n=0}^{\infty}(\mathbf{R}\mathbf{G}\mathbf{K_c}\circ)^n(\mathbf{G}\mathbf{Q}+\mathbf{R}\mathbf{G}\mathbf{P}')\,.
\end{aligned}
$$

2nd Order scattering term

3rd Order scattering term

Solution strategy

(dealing with non-linear modeling operator)

Autoconvolving Robust EPSI
Accounting for unknown data with G

While $\|\mathbf{p} - \widetilde{\mathcal{M}}(\mathbf{g}_k, \mathbf{q}_k)\|_2 > \sigma$

determine new τ_k from the Pareto curve

$$\mathbf{g}_{k+1} = \arg\min_{\mathbf{g}} \|\mathbf{p} - \widetilde{\mathcal{M}}(\mathbf{g}, \mathbf{q_k})\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \leq \tau_k$$

$$\mathbf{q}_{k+1} = \arg\min_{\mathbf{q}} \|\mathbf{p} - \widetilde{\mathbf{M}}_{g_{k+1}}\mathbf{q}\|_2$$

Strategy 1: Re-linearization
Using G from previous iter in higher-order terms

While $\|\mathbf{p} - \widetilde{\mathcal{M}}(\mathbf{g}_k, \mathbf{q}_k)\|_2 > \sigma$

determine new τ_k from the Pareto curve

$$\mathbf{g}_{k+1} = \arg\min_{\mathbf{g}} \|\mathbf{p} - \boxed{\widetilde{\mathbf{M}}_{q_k}}\mathbf{g}\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \leq \tau_k$$

fix at $\mathbf{g_k}$ for autoconv terms

$$\mathbf{q}_{k+1} = \arg\min_{\mathbf{q}} \|\mathbf{p} - \widetilde{\mathbf{M}}_{g_{k+1}}\mathbf{q}\|_2$$

Strategy 2: Modified Gauss-Newton
Obtain Jacobian using G from previous iter

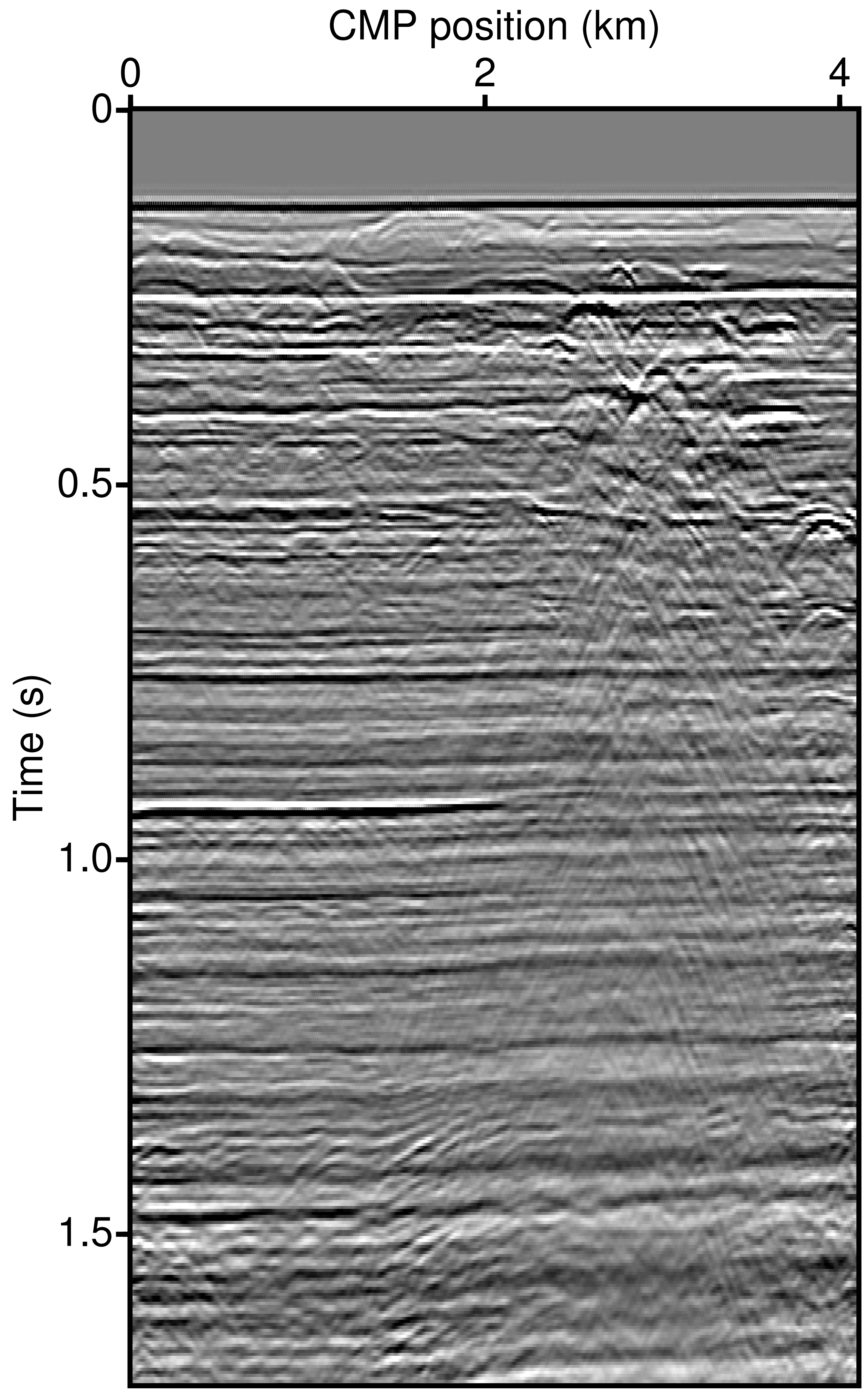
While $\|\mathbf{p} - \widetilde{\mathcal{M}}(\mathbf{g}_k, \mathbf{q}_k)\|_2 > \sigma$

determine new τ_k from the Pareto curve

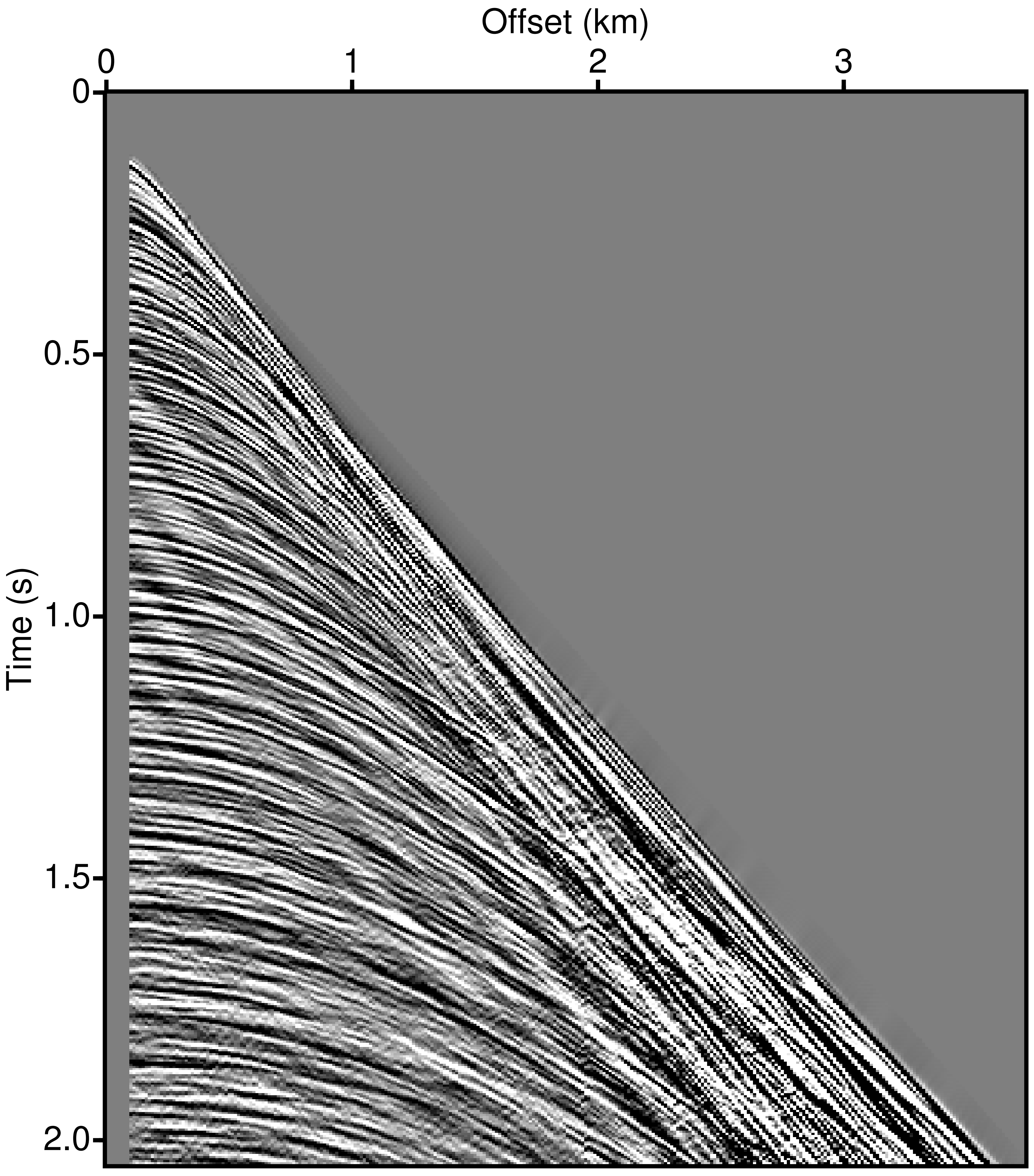
$$\mathbf{g}_{k+1} = \mathbf{g}_k + \arg\min_{\Delta\mathbf{g}} \|\mathbf{r}_k - \partial_{(g_k, q_k)}\widetilde{\mathcal{M}}\Delta\mathbf{g}\|_2 \text{ s.t. } \|\Delta\mathbf{g}\|_1 \leq \tau_k$$

$$\mathbf{q}_{k+1} = \arg\min_{\mathbf{q}} \|\mathbf{p} - \widetilde{\mathbf{M}}_{g_{k+1}}\mathbf{q}\|_2$$

Field data example
North Sea dataset



Data



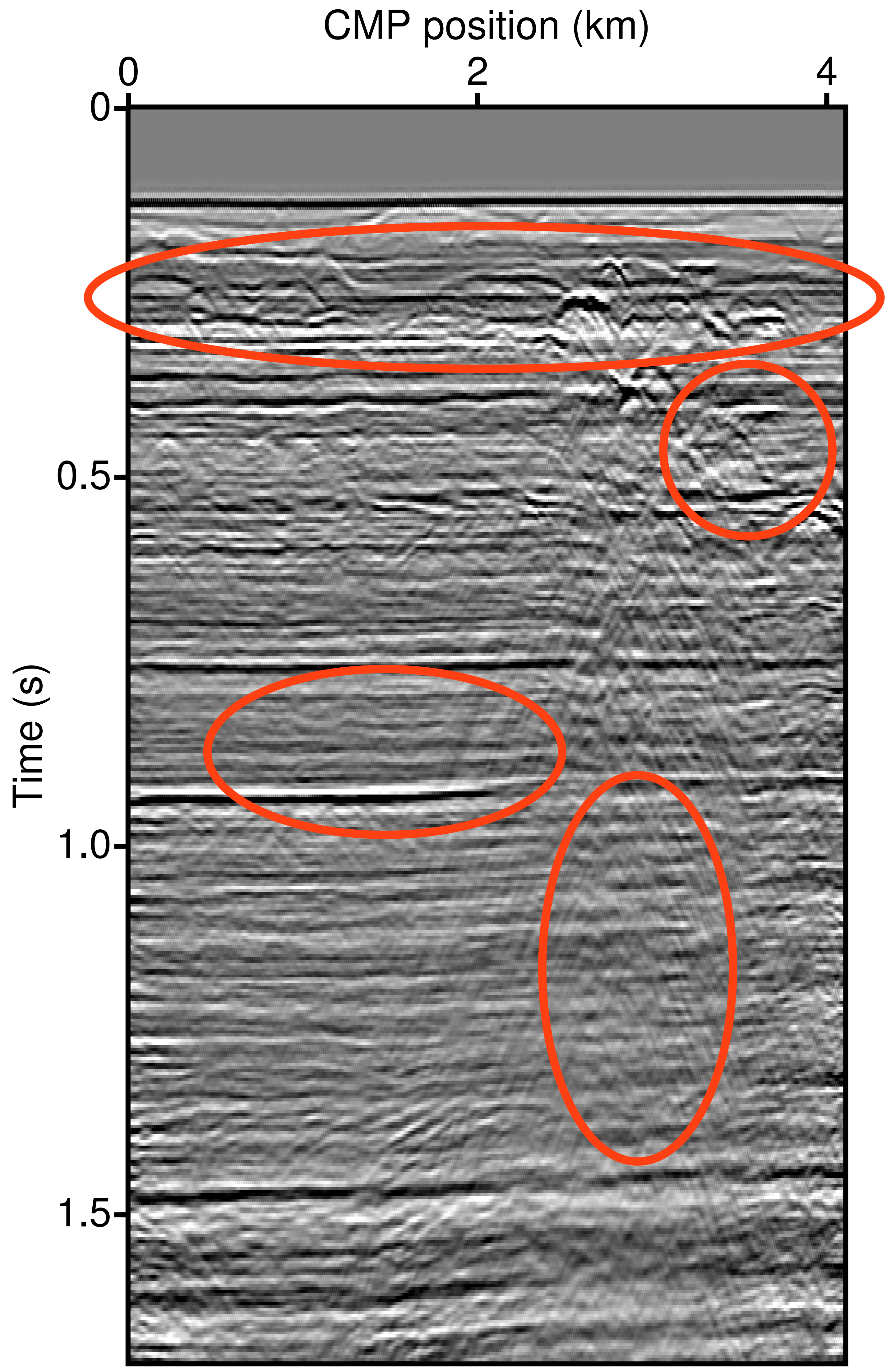
Shot

North Sea dataset

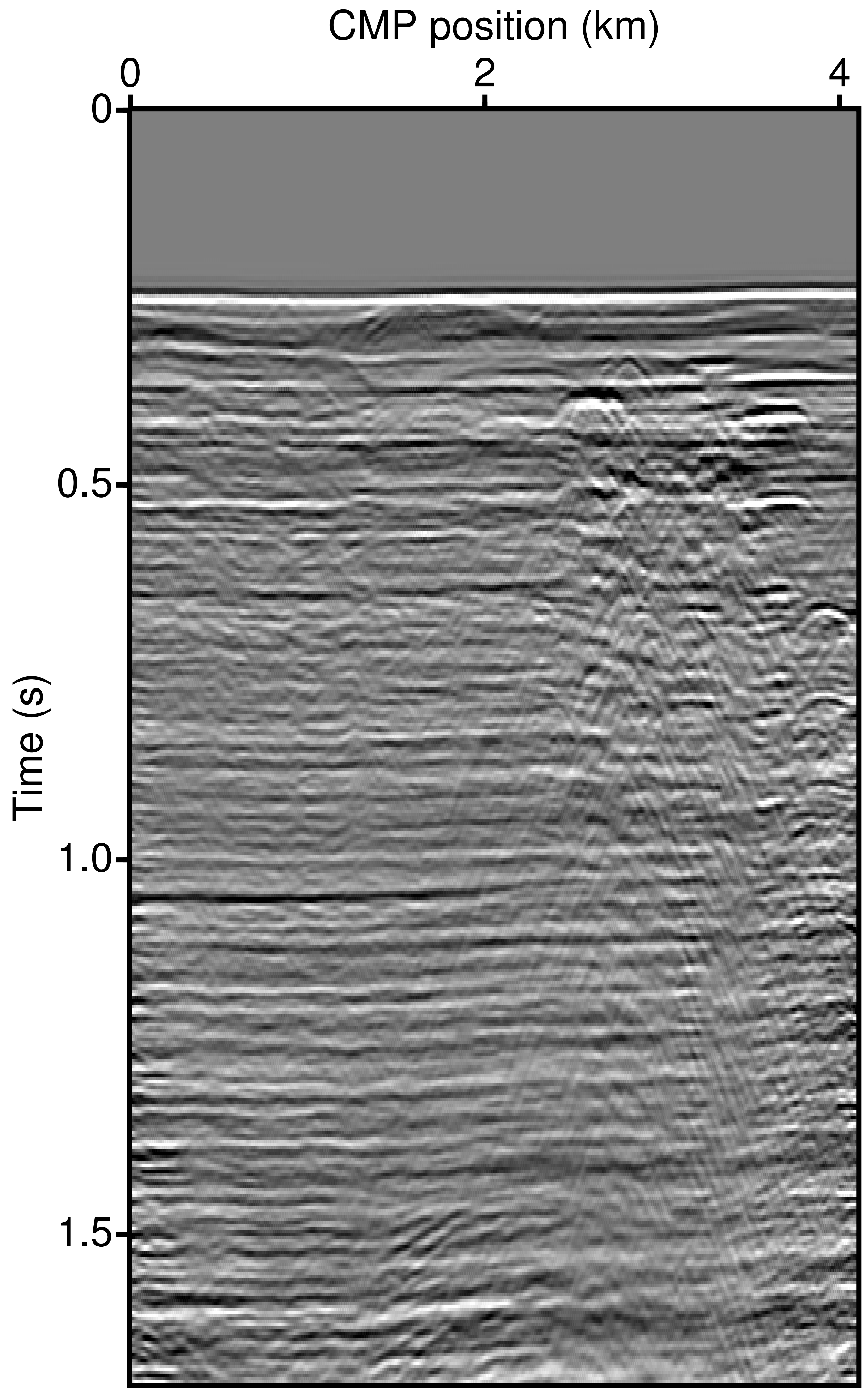
100m near-offset

regularized to 12.5m dx and 4km fixed-spread from streamer

4ms sampling

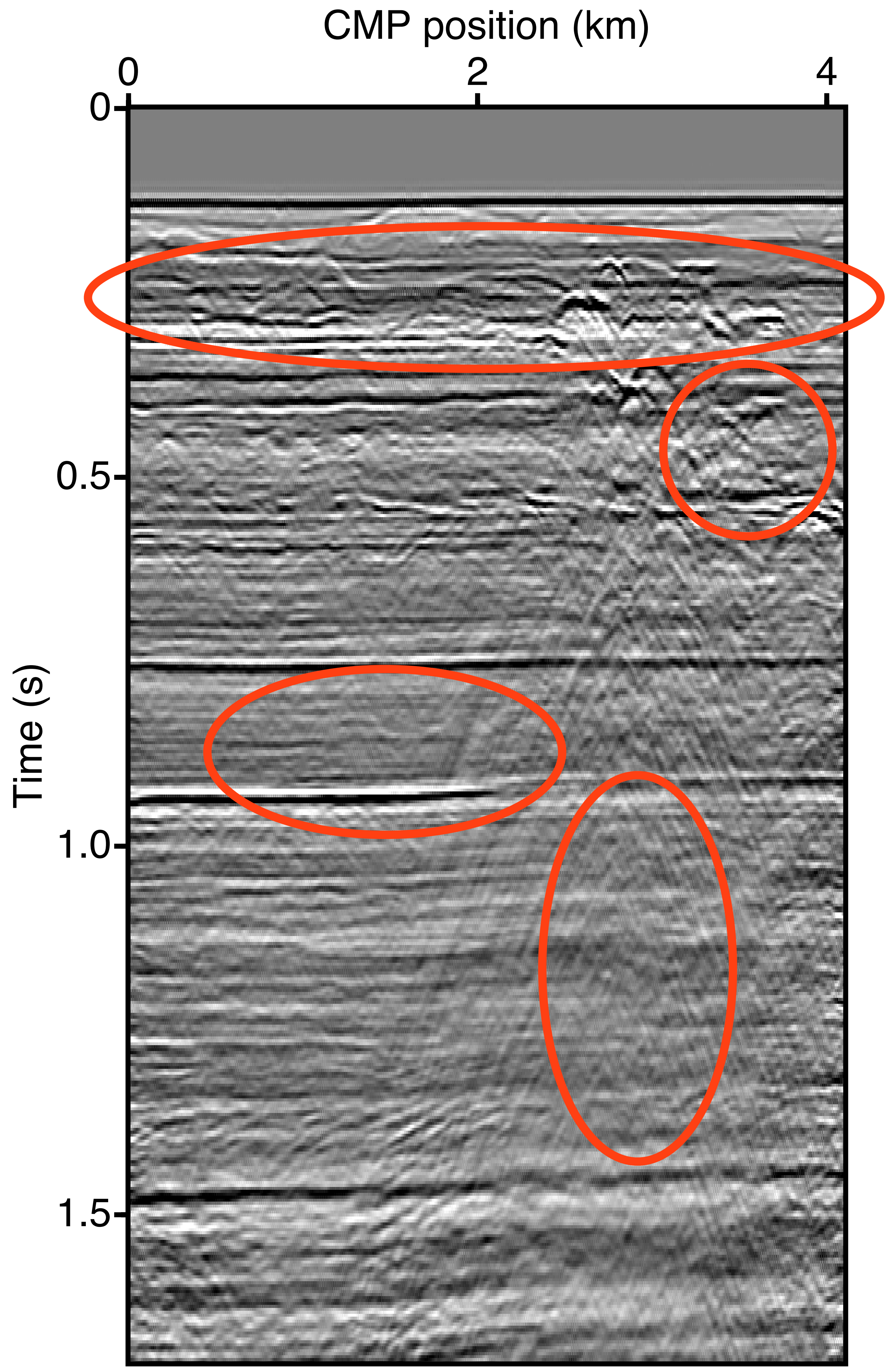


Conservative primary

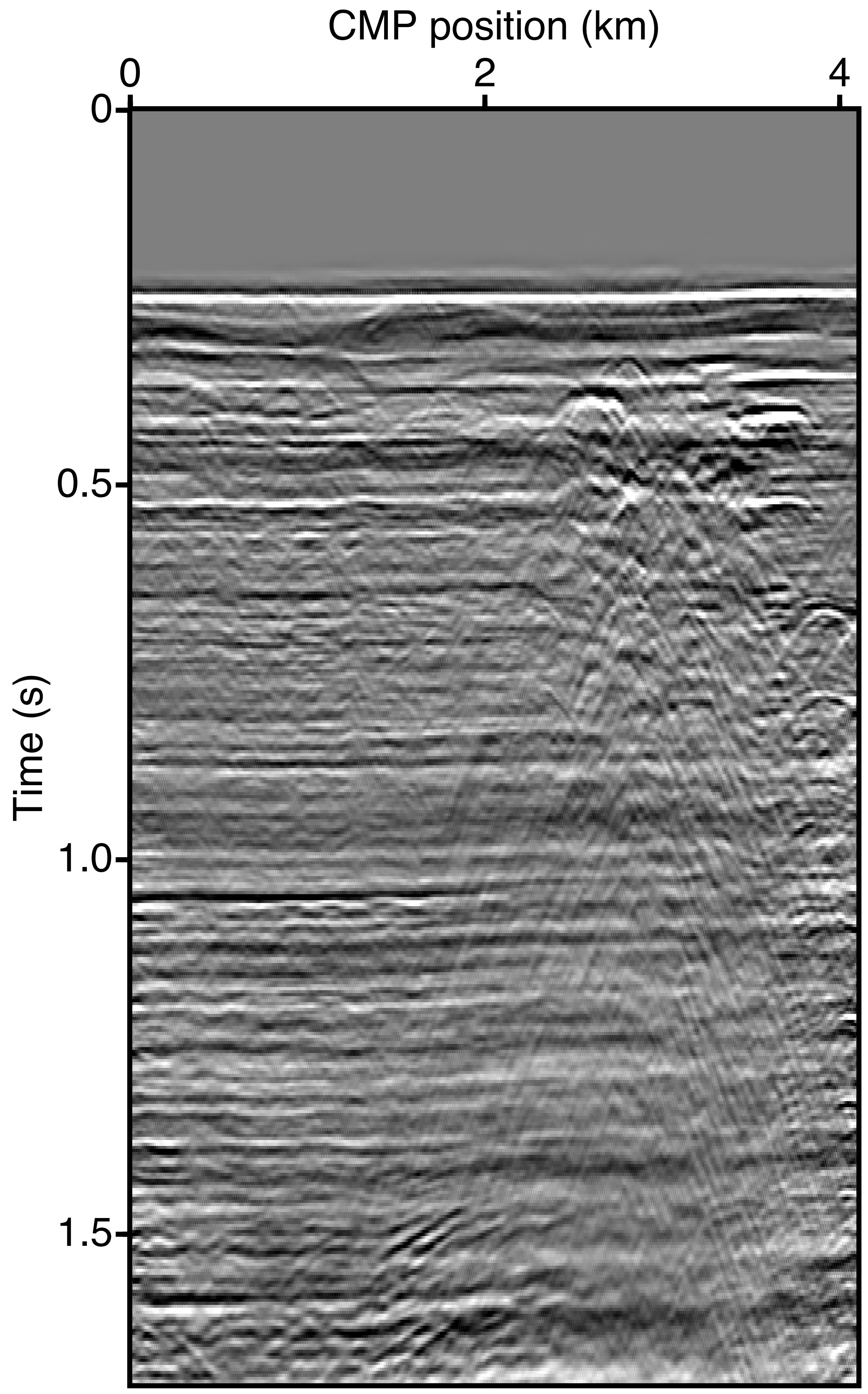


Multiple

NMO stack
Parabolic Radon Interp

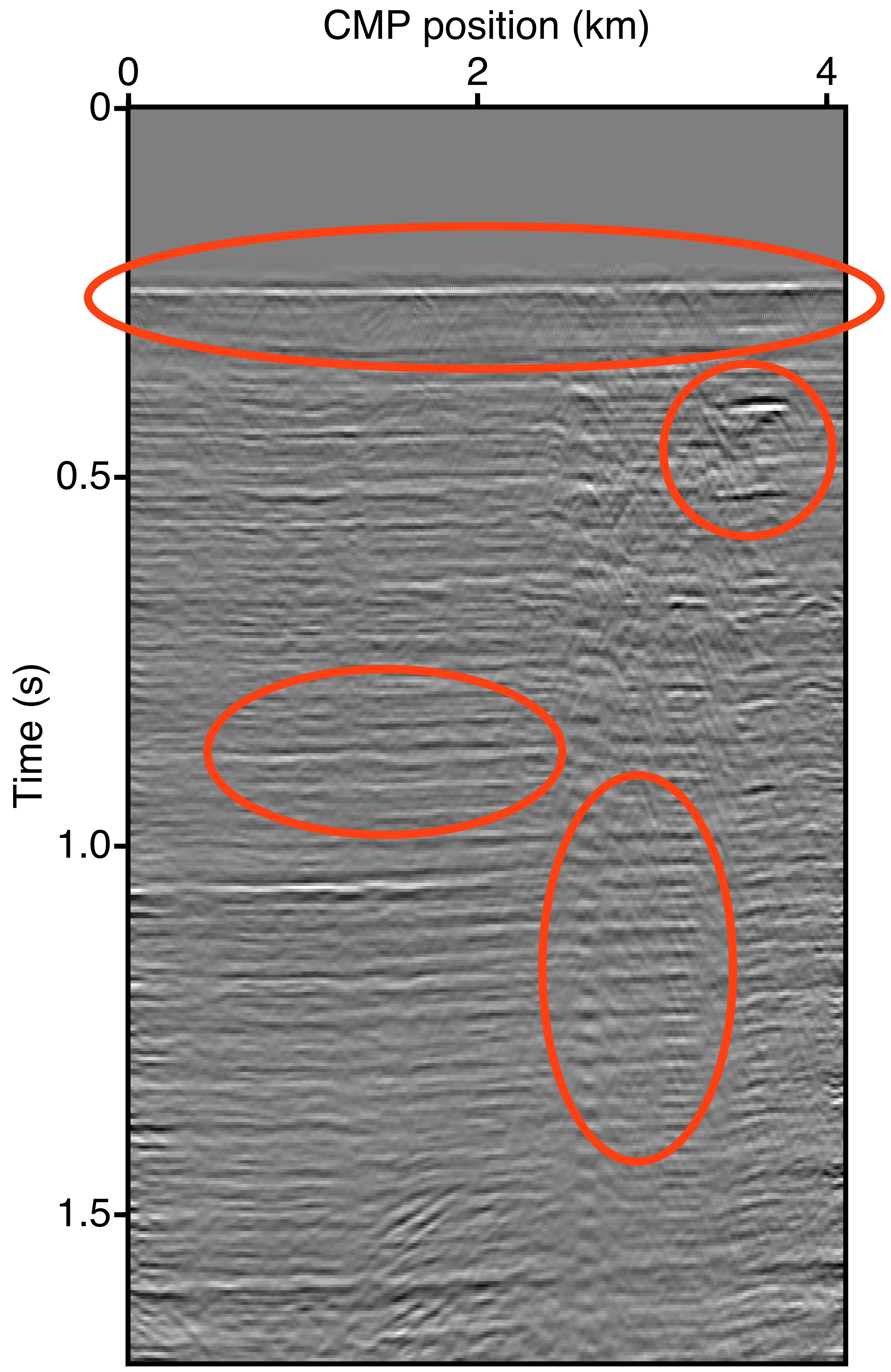


Conservative primary

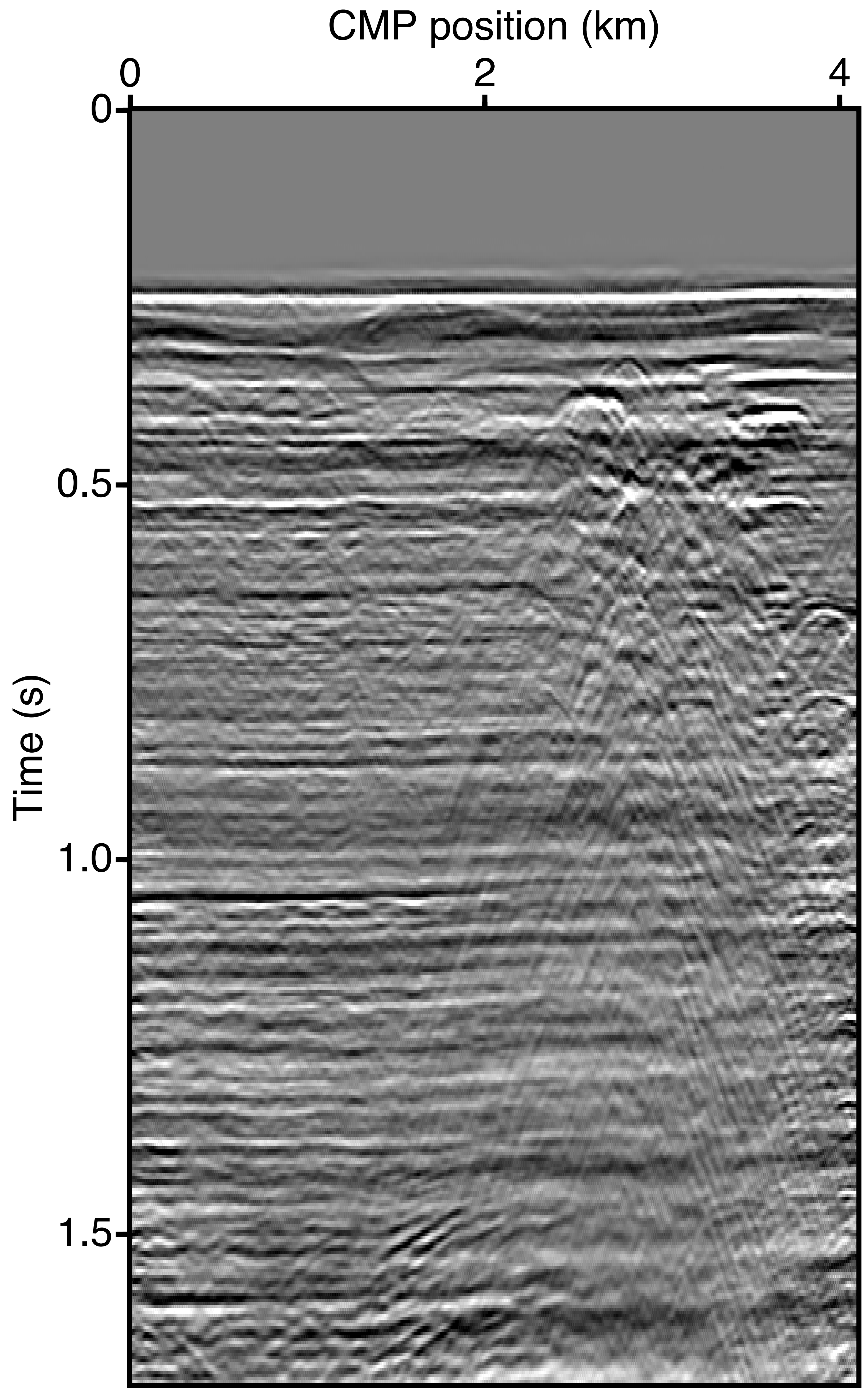


Multiple

NMO stack
Re-linearization
Using 3rd Order terms



Difference from Radon interp.

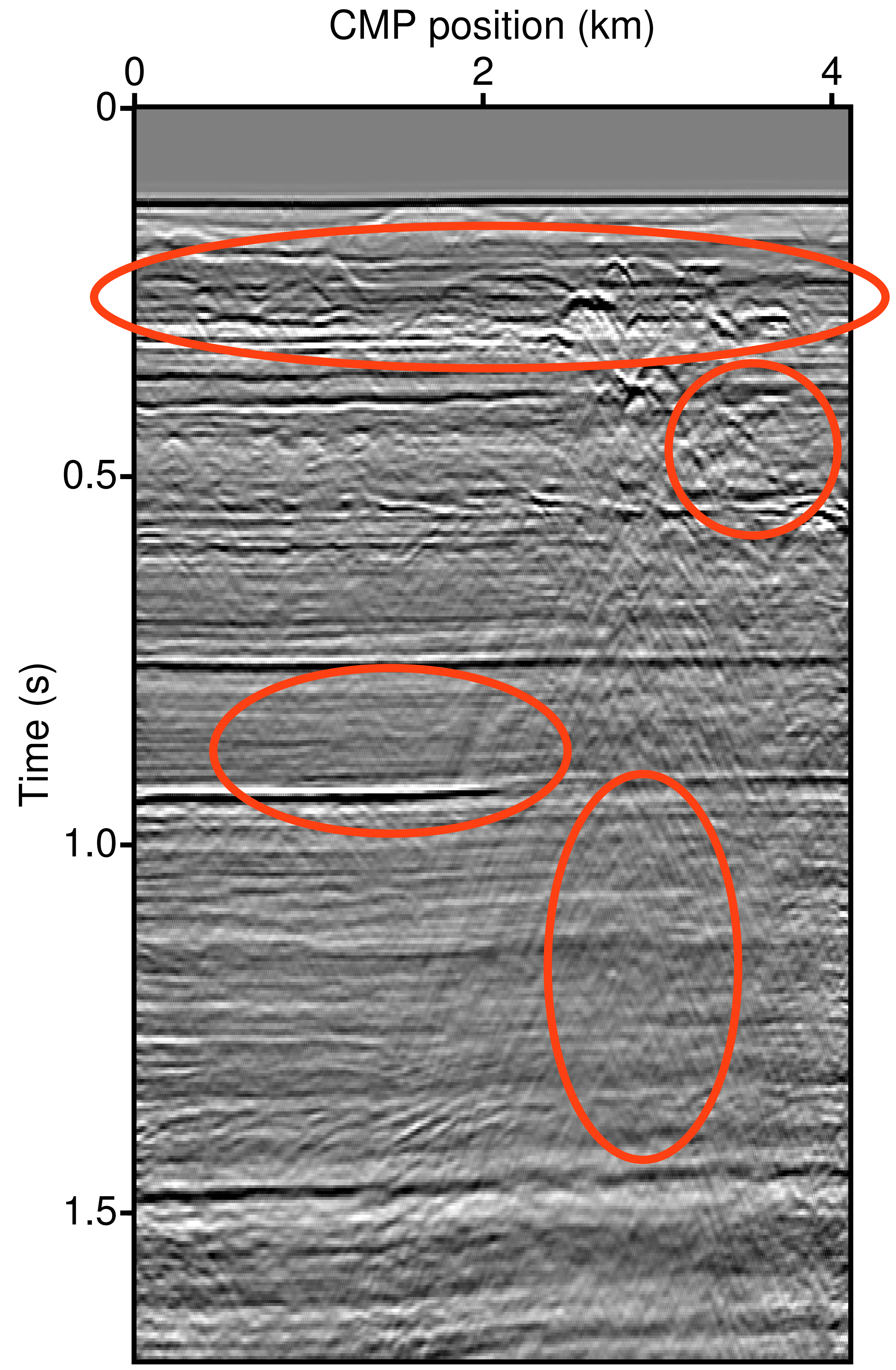


Multiple

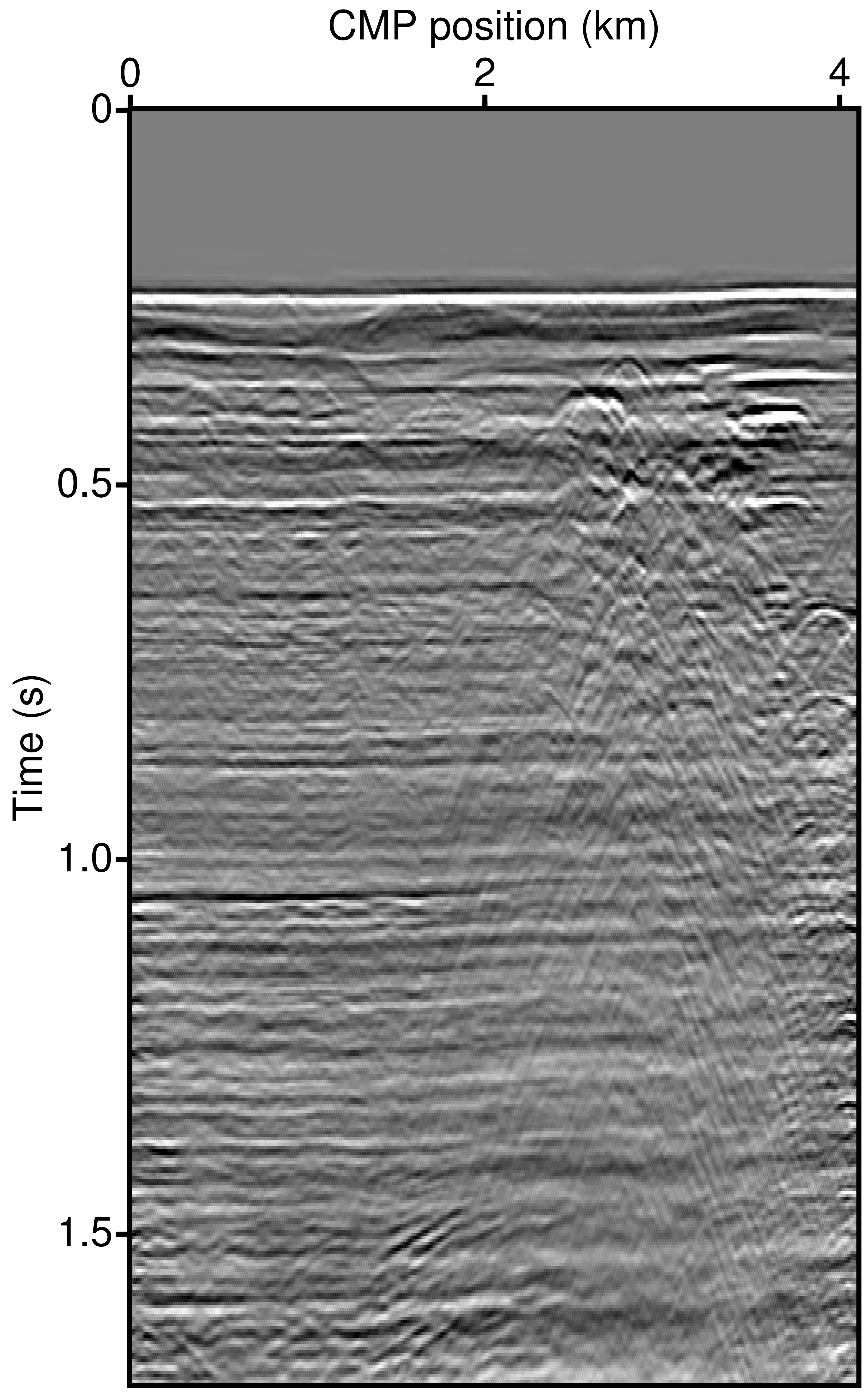
NMO stack

Re-linearization

Using 3rd Order terms

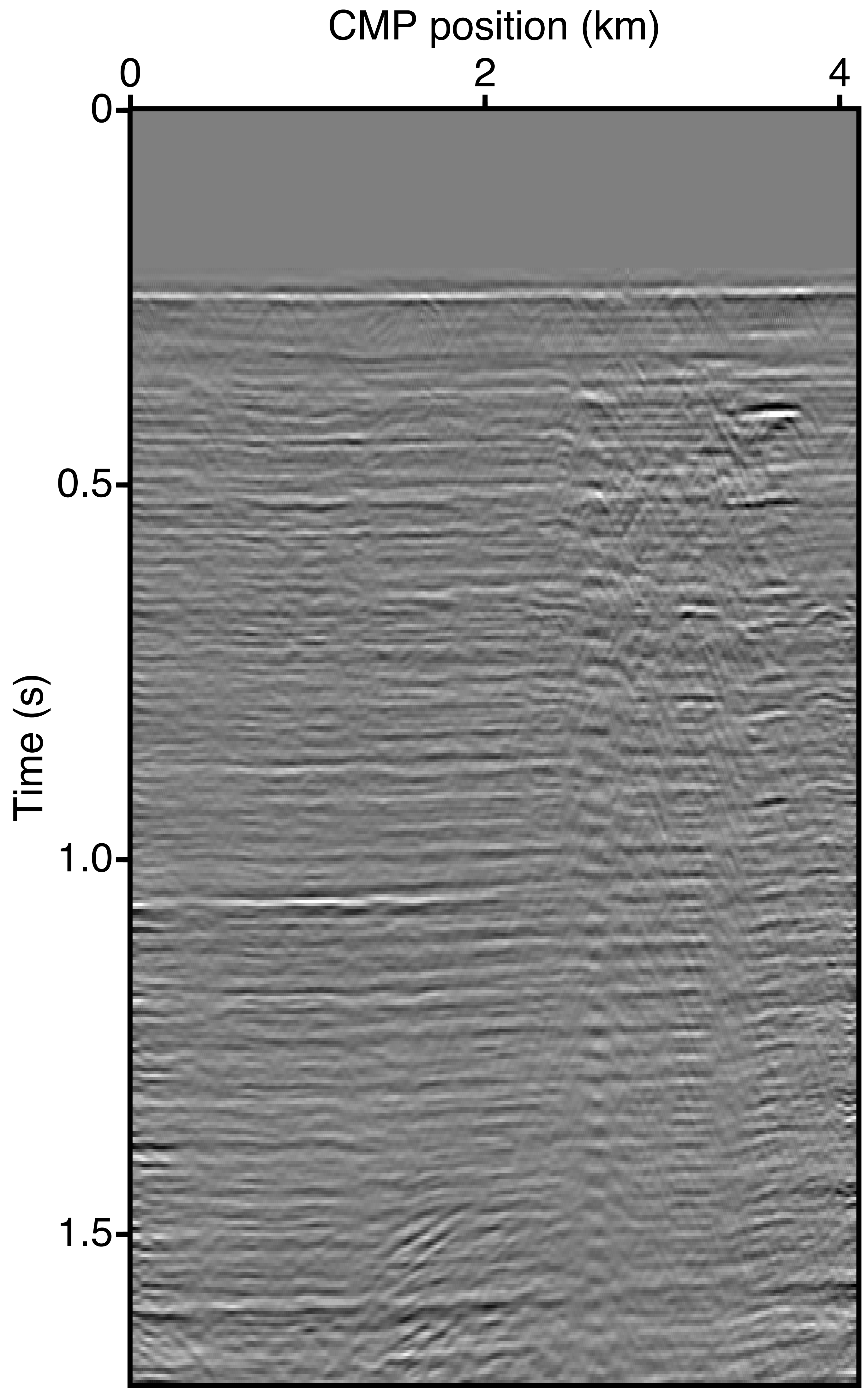


Conservative primary

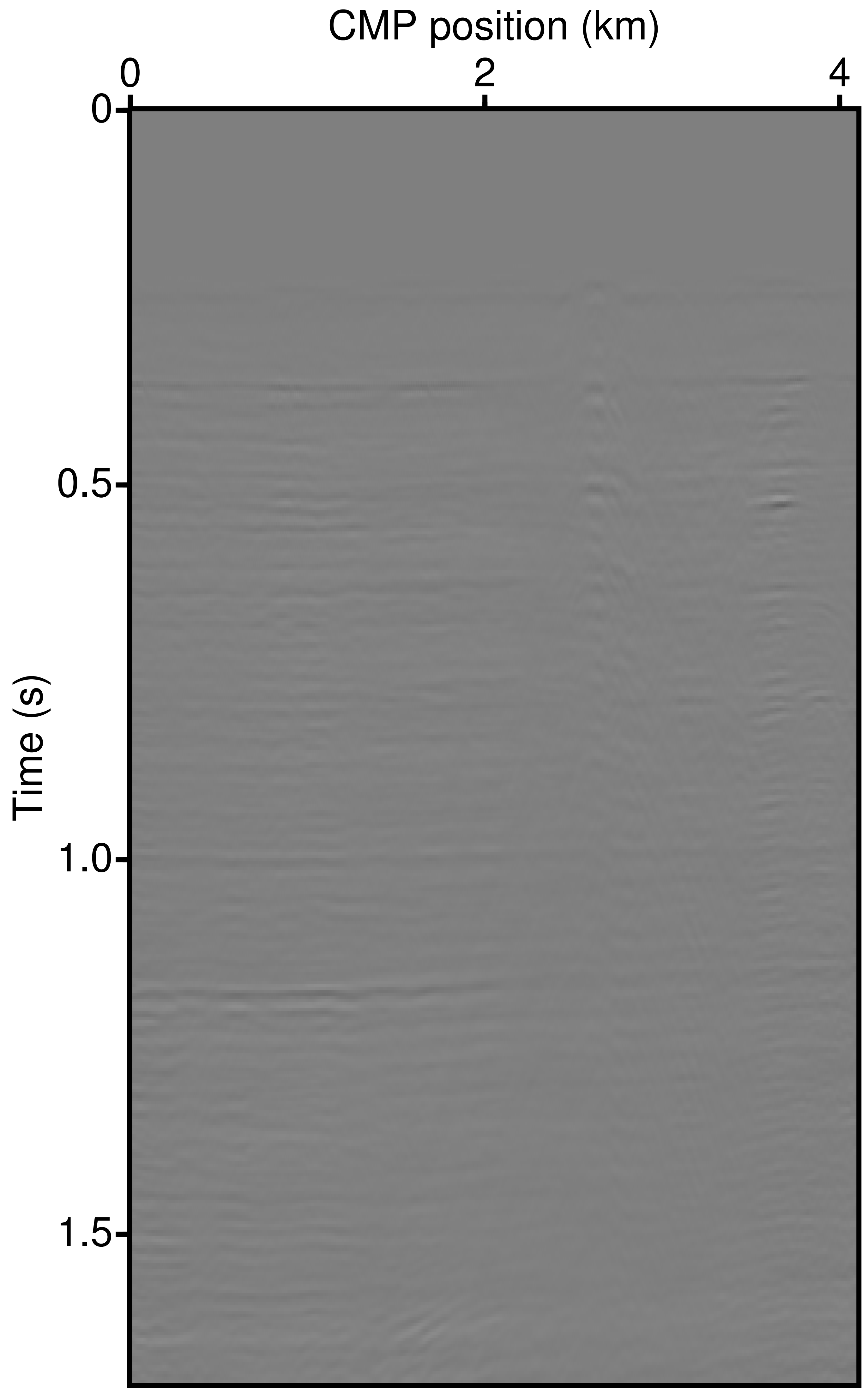


Multiple

NMO stack
Re-linearization
Using 2nd Order terms

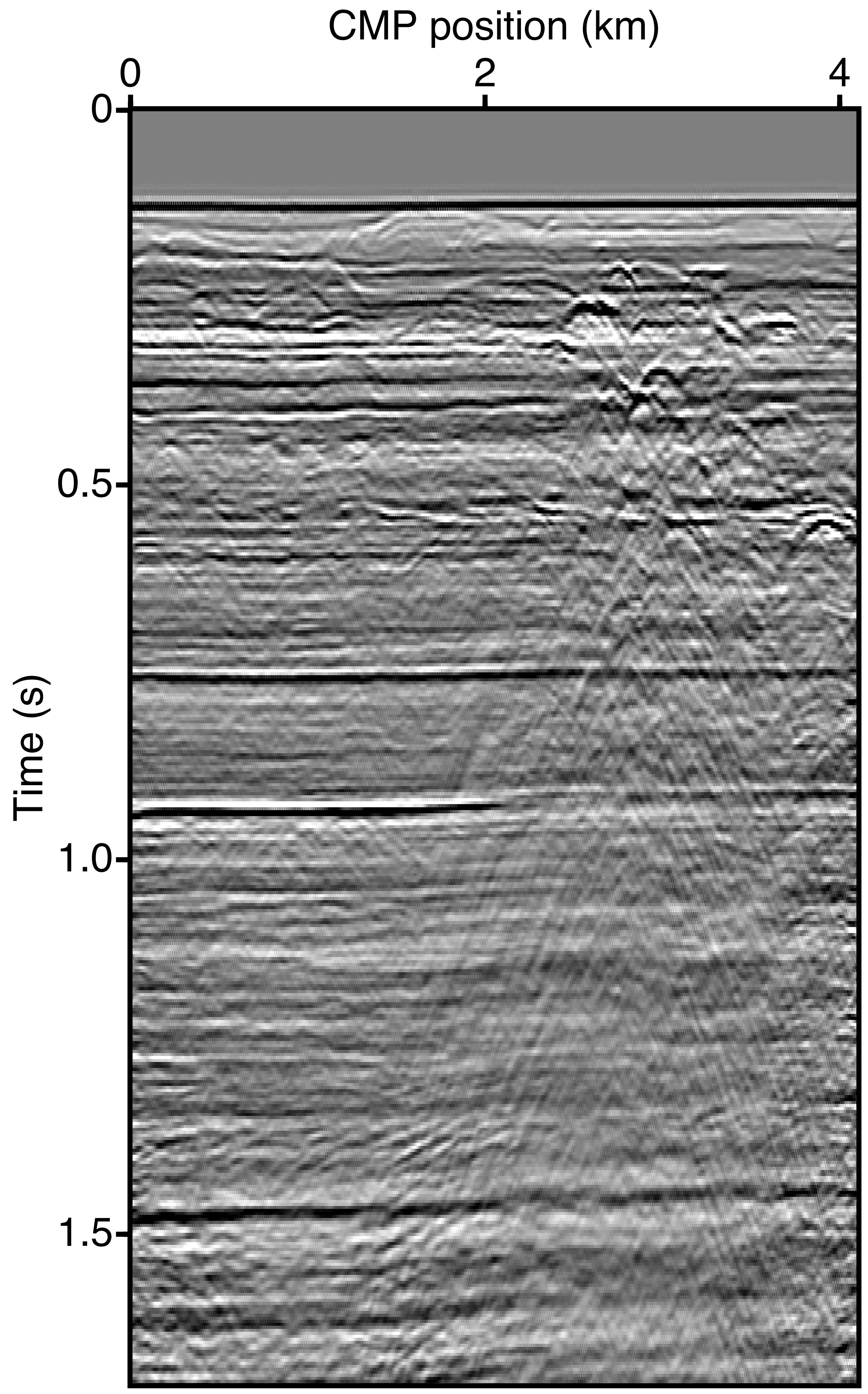


Radon interp - Re-linearization 2nd order

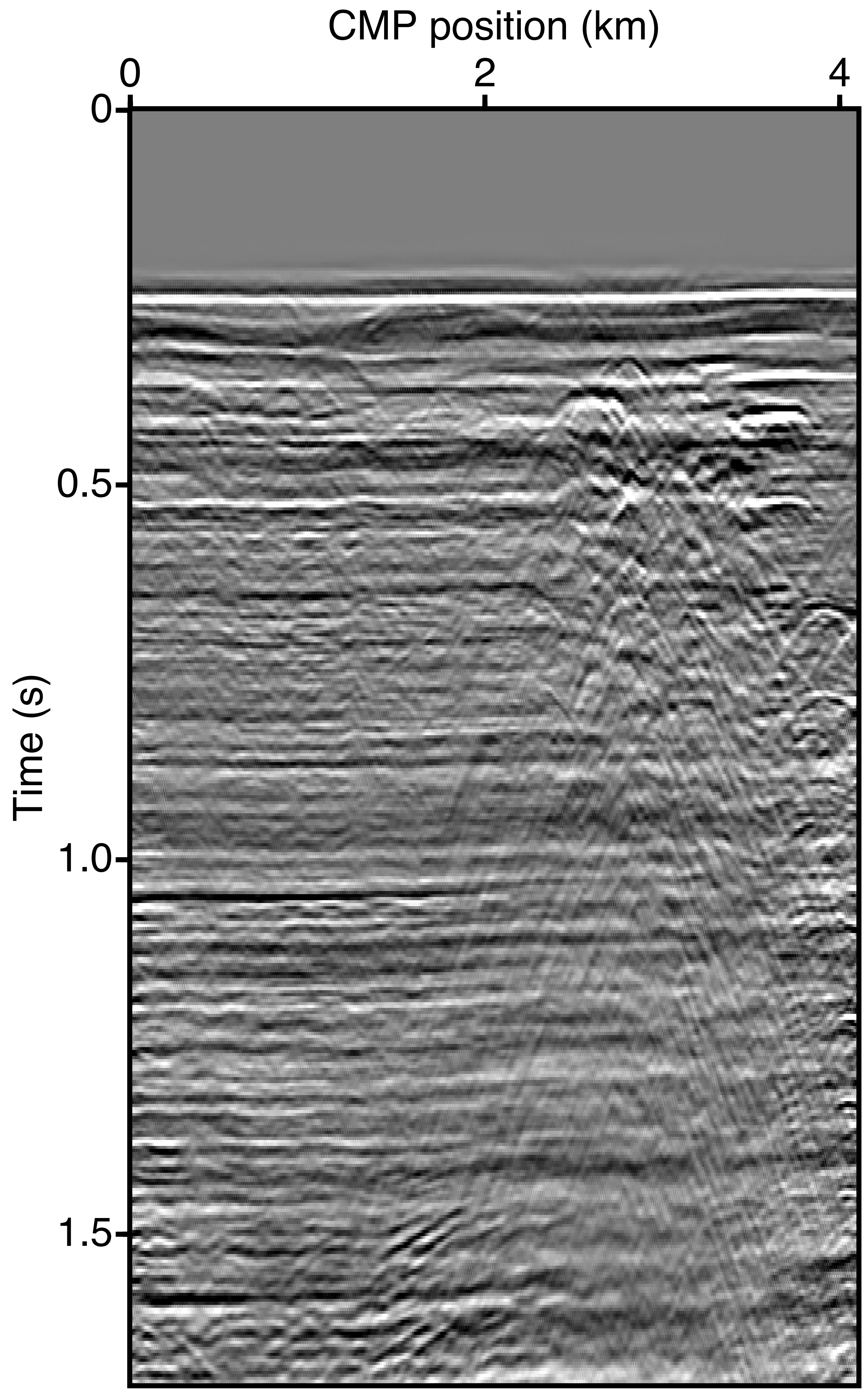


Re-linearization 3rd - 2nd order

NMO stack
Difference plots

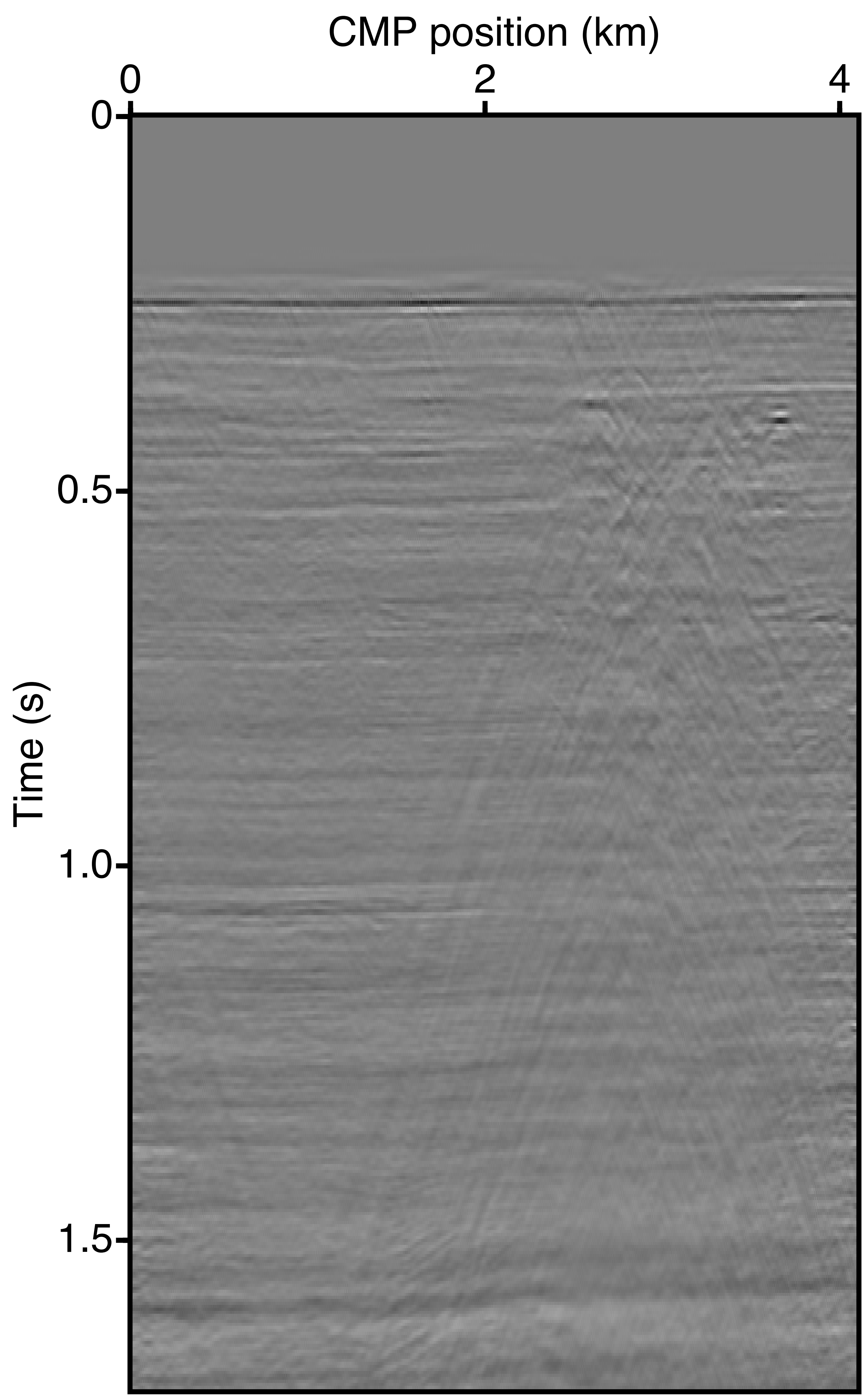


Conservative primary

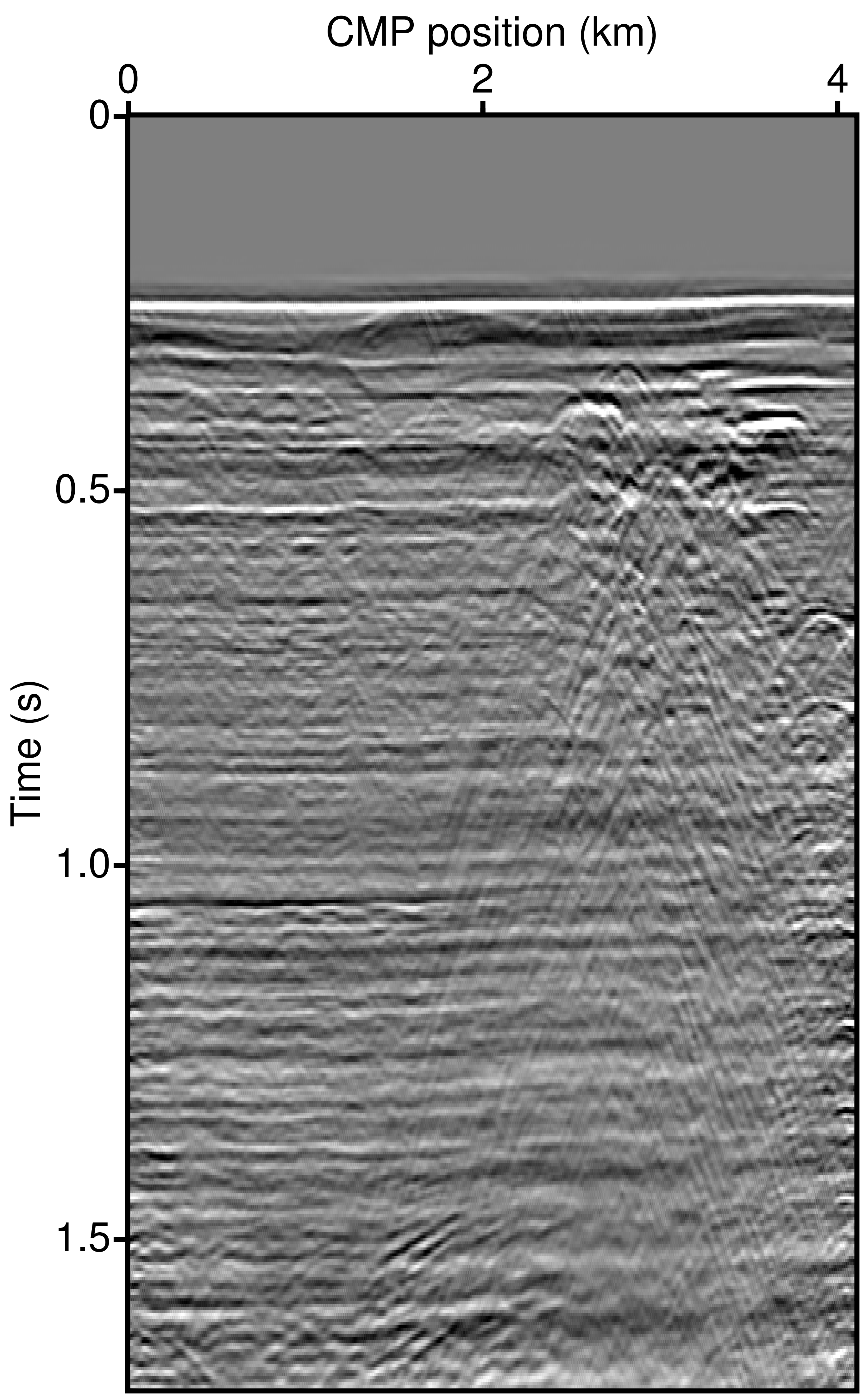


Multiple

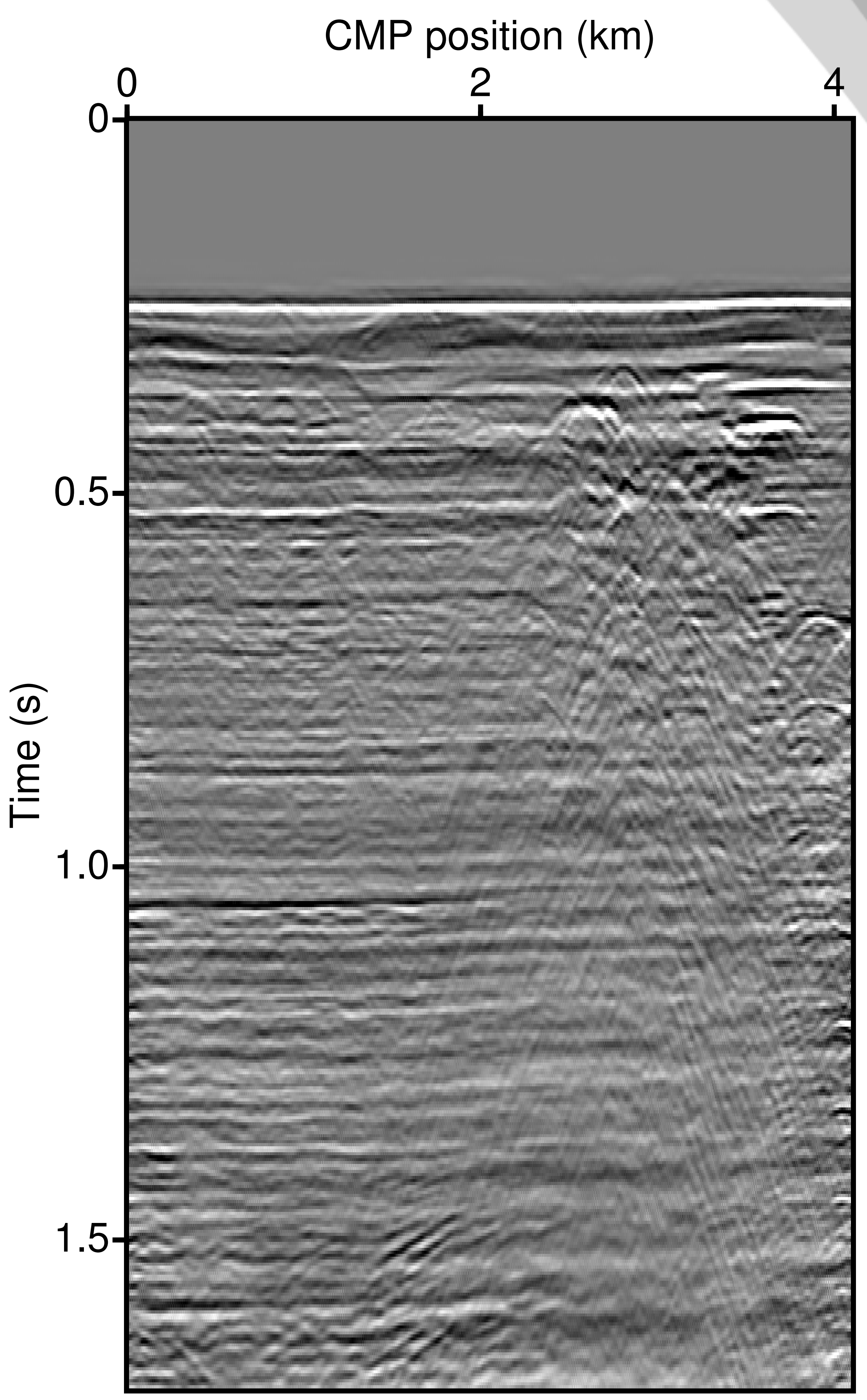
NMO stack
Modified Gauss-Newton Using 3rd Order terms



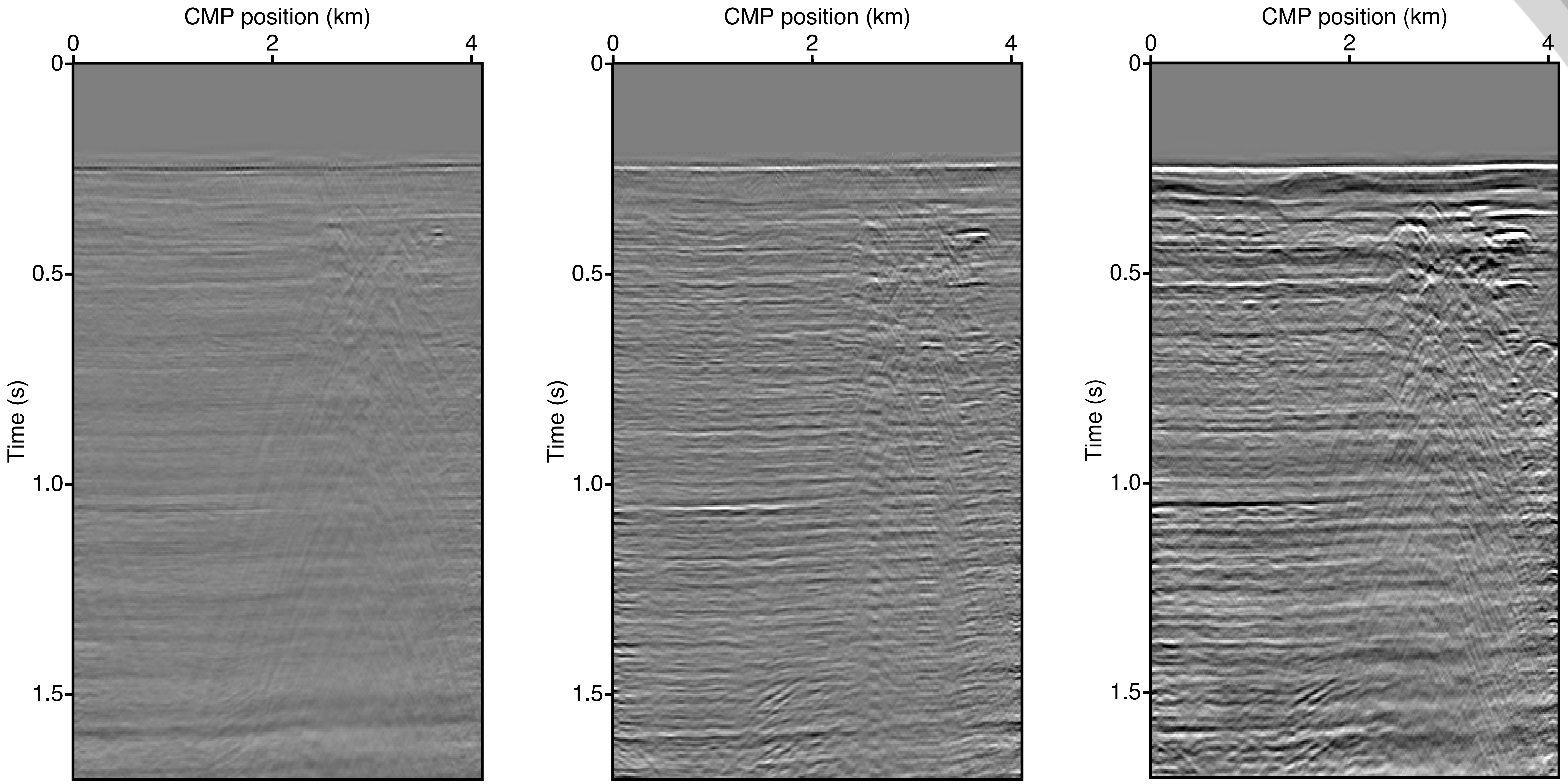
Diff: GN vs Relinearize



GN 3rd Order Multiple



Re-lin. 3rd Order Multiple



Diff: GN vs Relinearize **Diff: Radon interp vs Relinearize** **Re-lin. 3rd Order Multiple**

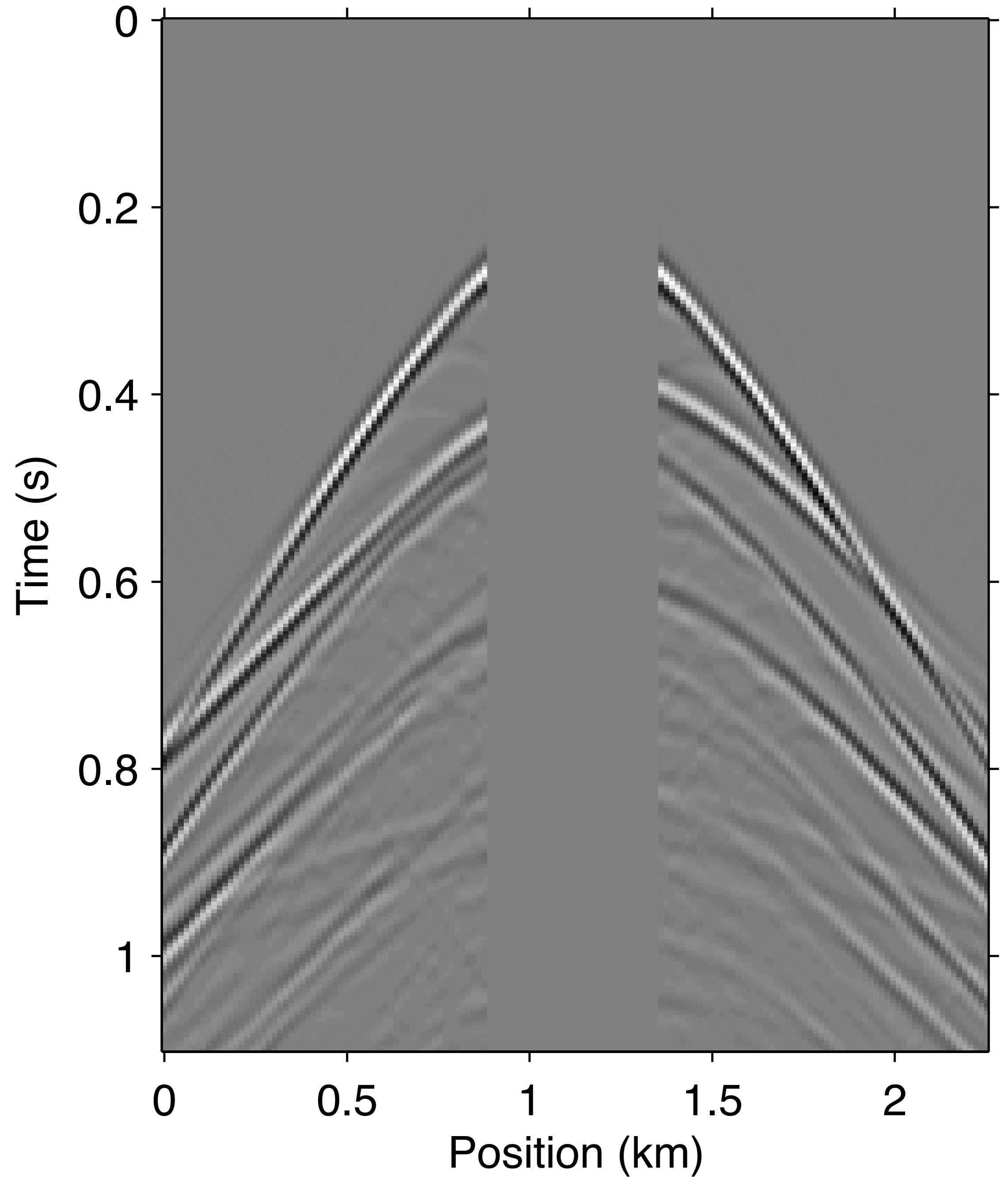
SLIM

Summary: Near-offset mitigation using scattering

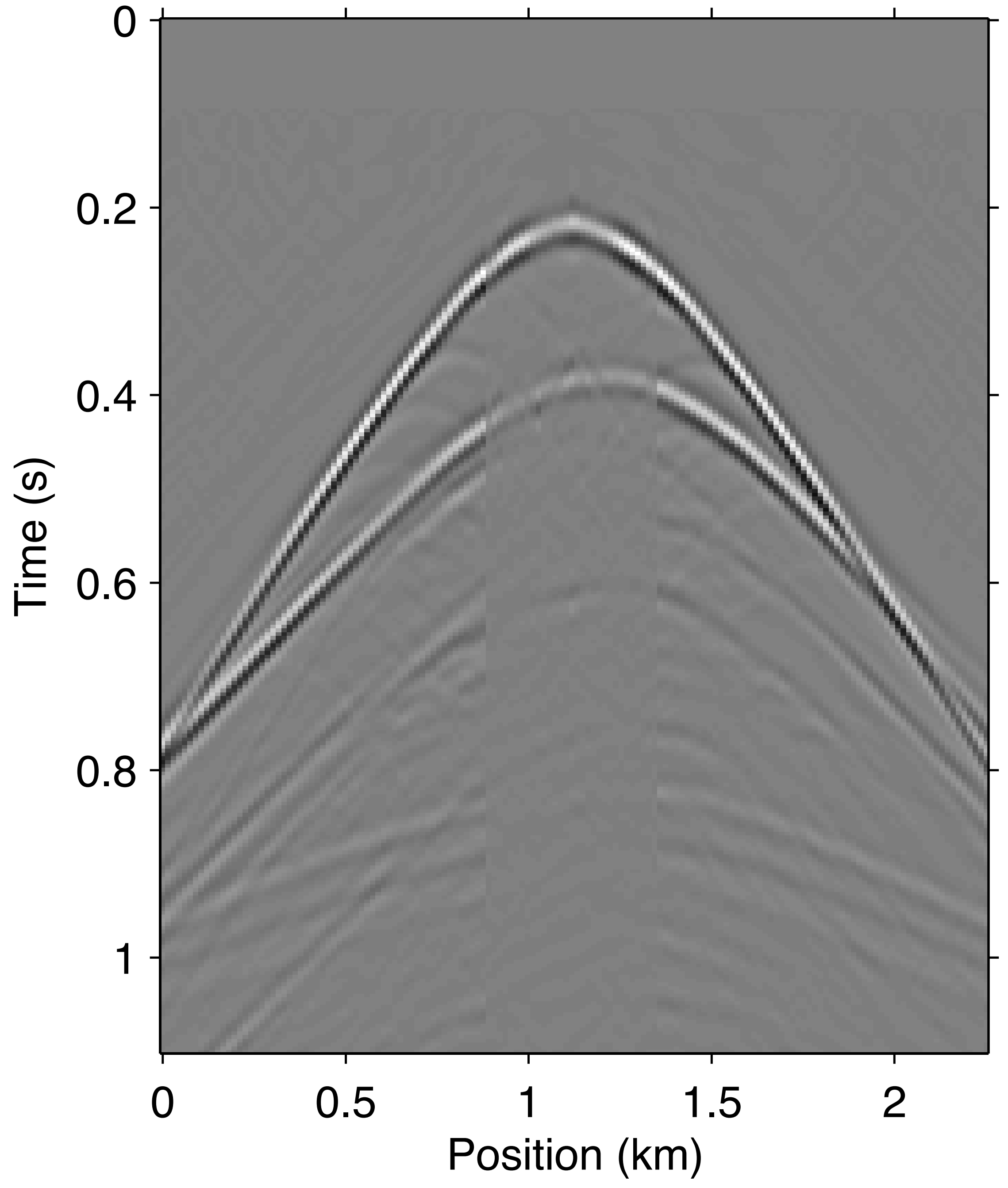
A correction to the multiple prediction in face of missing data using Born scattering off the free surface.

$$
\begin{aligned}
\widetilde{\mathcal{M}}(\mathbf{G},\mathbf{Q};\mathbf{P}') =& \mathbf{K}\circ\Big[\mathbf{G}\mathbf{Q}+\mathbf{R}\mathbf{G}\mathbf{P}' \\
\text{2nd Order Scattering} \quad & +\ \mathbf{R}\mathbf{G}\mathbf{K_c}\circ(\mathbf{G}\mathbf{Q}+\mathbf{R}\mathbf{G}\mathbf{P}') \\
\text{3rd Order Scattering} \quad & +\ \mathbf{R}\mathbf{G}\mathbf{K_c}\circ(\mathbf{R}\mathbf{G}\mathbf{K_c}\circ(\mathbf{G}\mathbf{Q}+\mathbf{R}\mathbf{G}\mathbf{P}')) \\
& +\ \mathcal{O}(\mathbf{G}^4)\Big] \\
:=& \mathbf{K}\circ\sum_{n=0}^{\infty}(\mathbf{R}\mathbf{G}\mathbf{K_c}\circ)^n(\mathbf{G}\mathbf{Q}+\mathbf{R}\mathbf{G}\mathbf{P}')\,.
\end{aligned}
$$

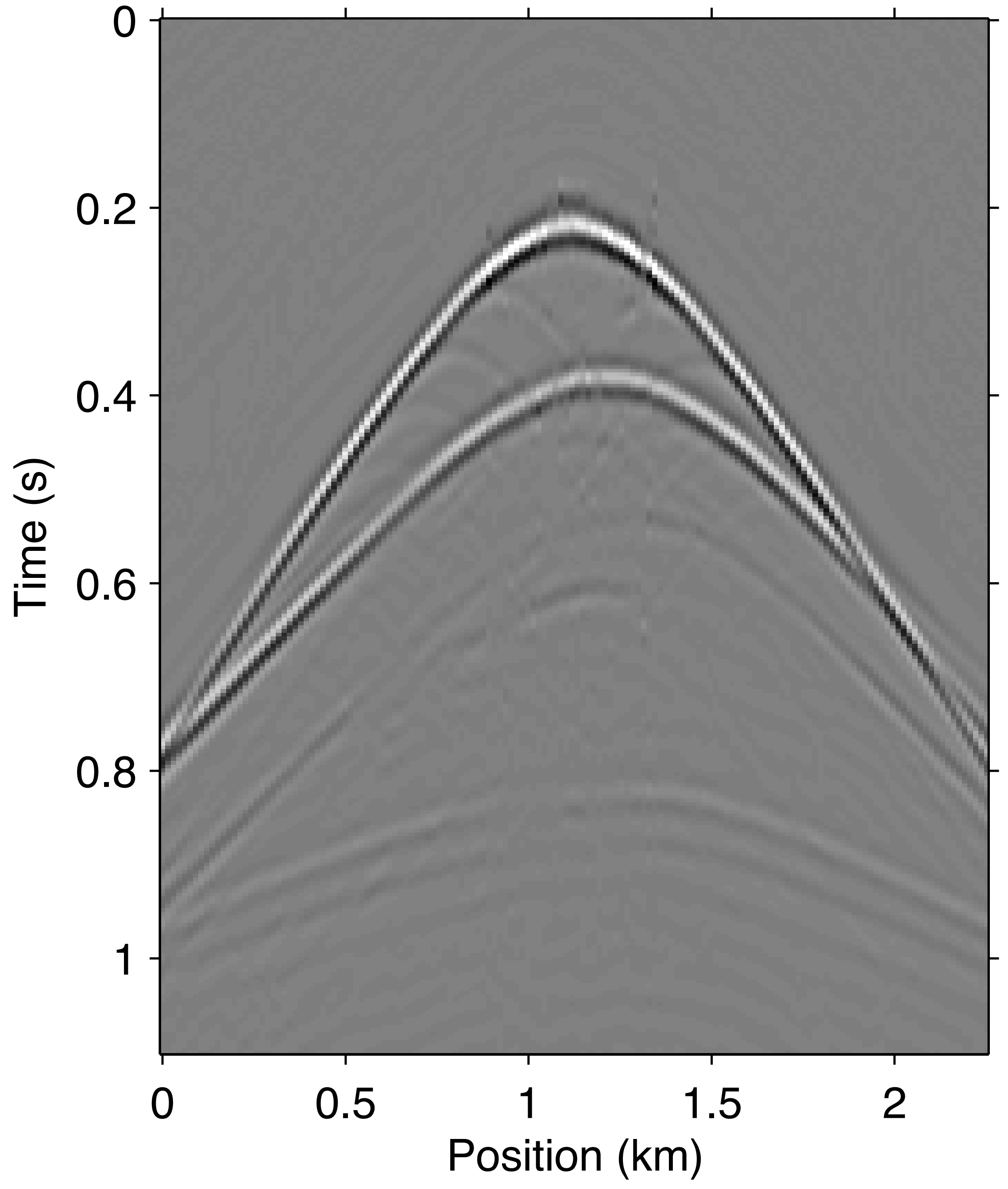
Data



Explicit Reconstruct.



Correction w/ 3rd Ord. Scattering



Chapter 5

Multi-grid acceleration strategy

Motivation

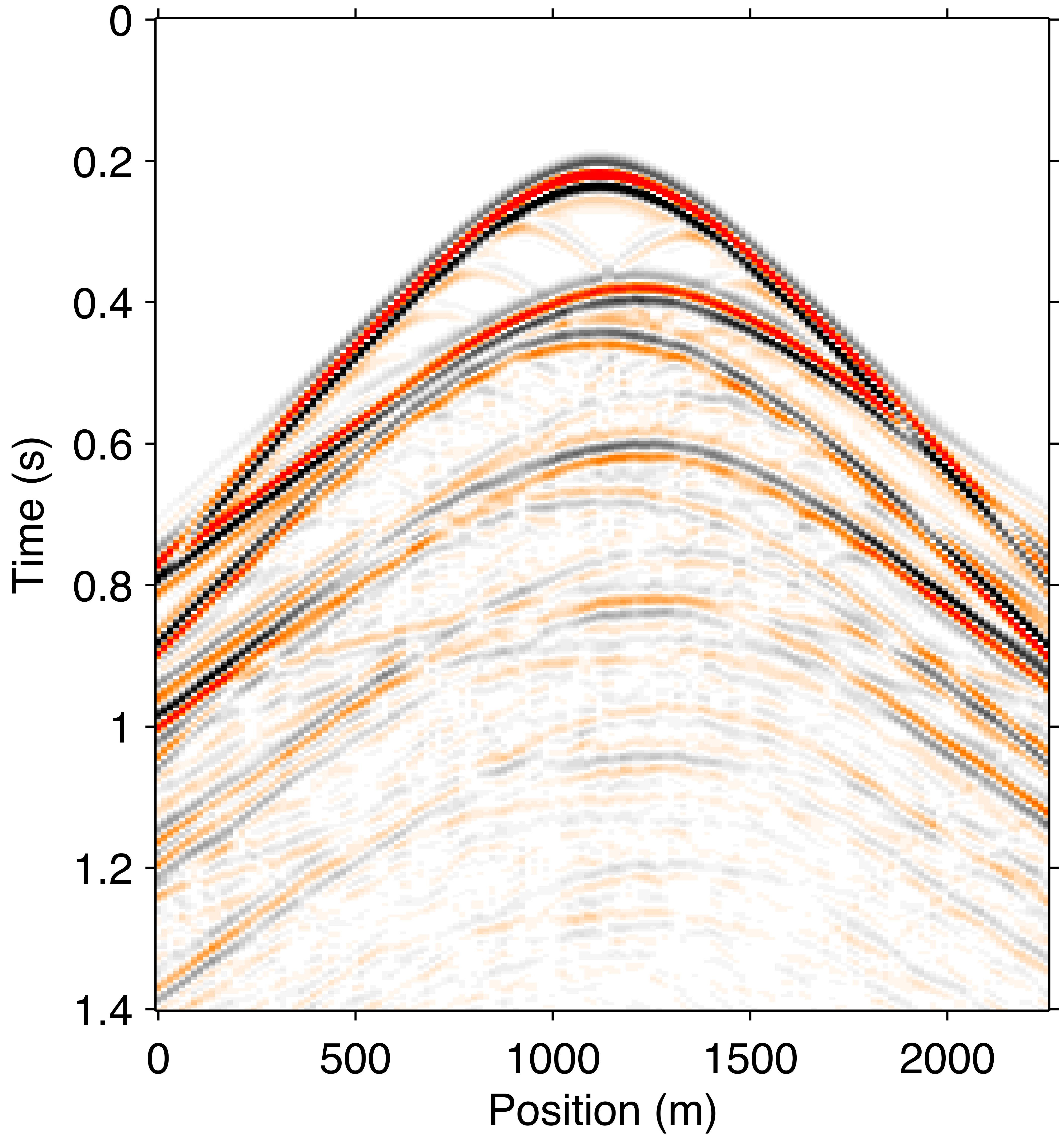
REPSI (and EPSI) is an inherently *expensive* method

- Not unusual for single gradient to take **days** to evaluate for modern 3D dataset

Data-matrix size exceeds 1,000,000-by-1,000,000 for typical 3D datasets, 2000 time samples

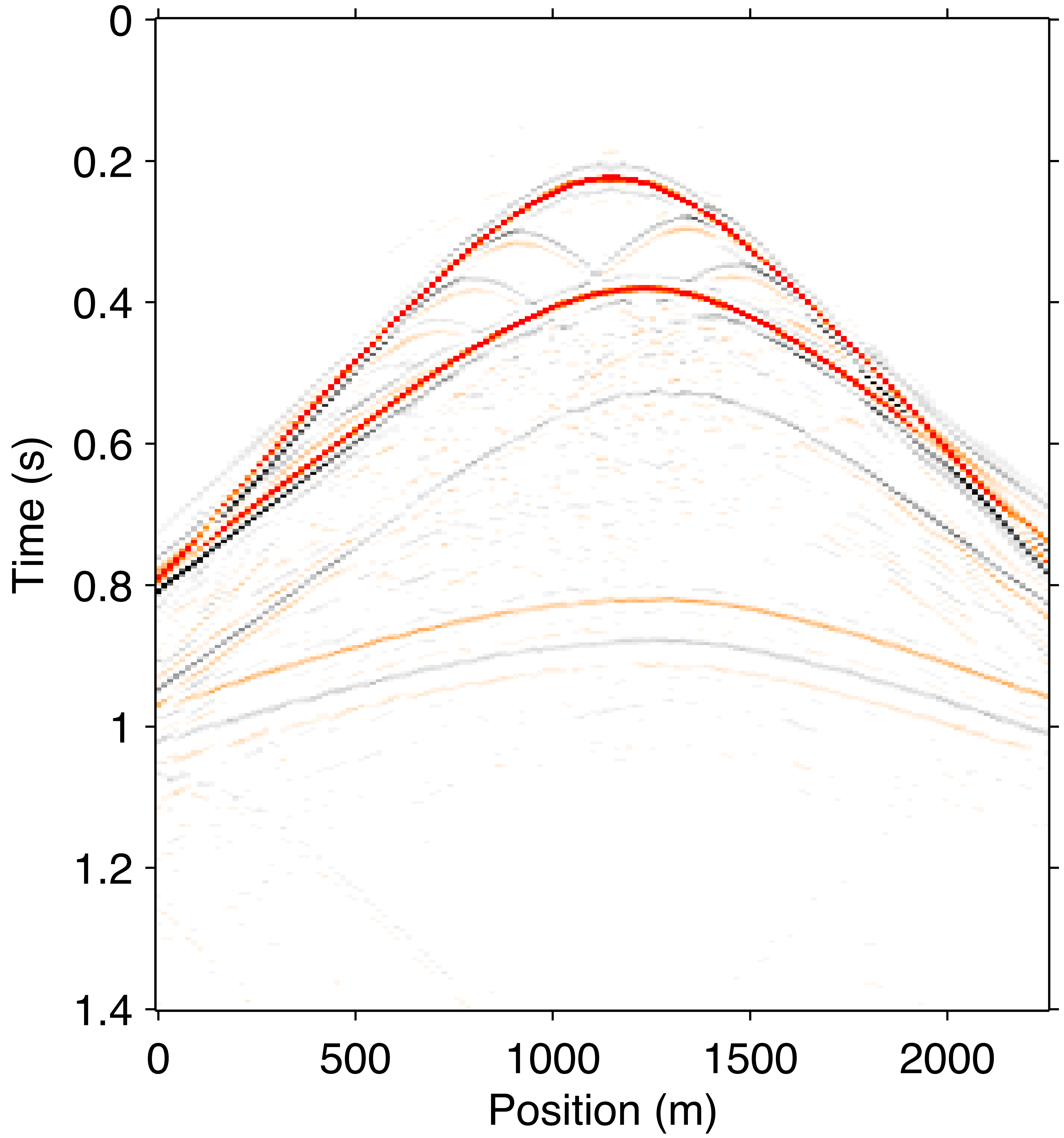
Does all iterations have to be done on the whole wavefield? Can we borrow ideas from multigrid methods?

Motivation: G tolerates lowpass filtering



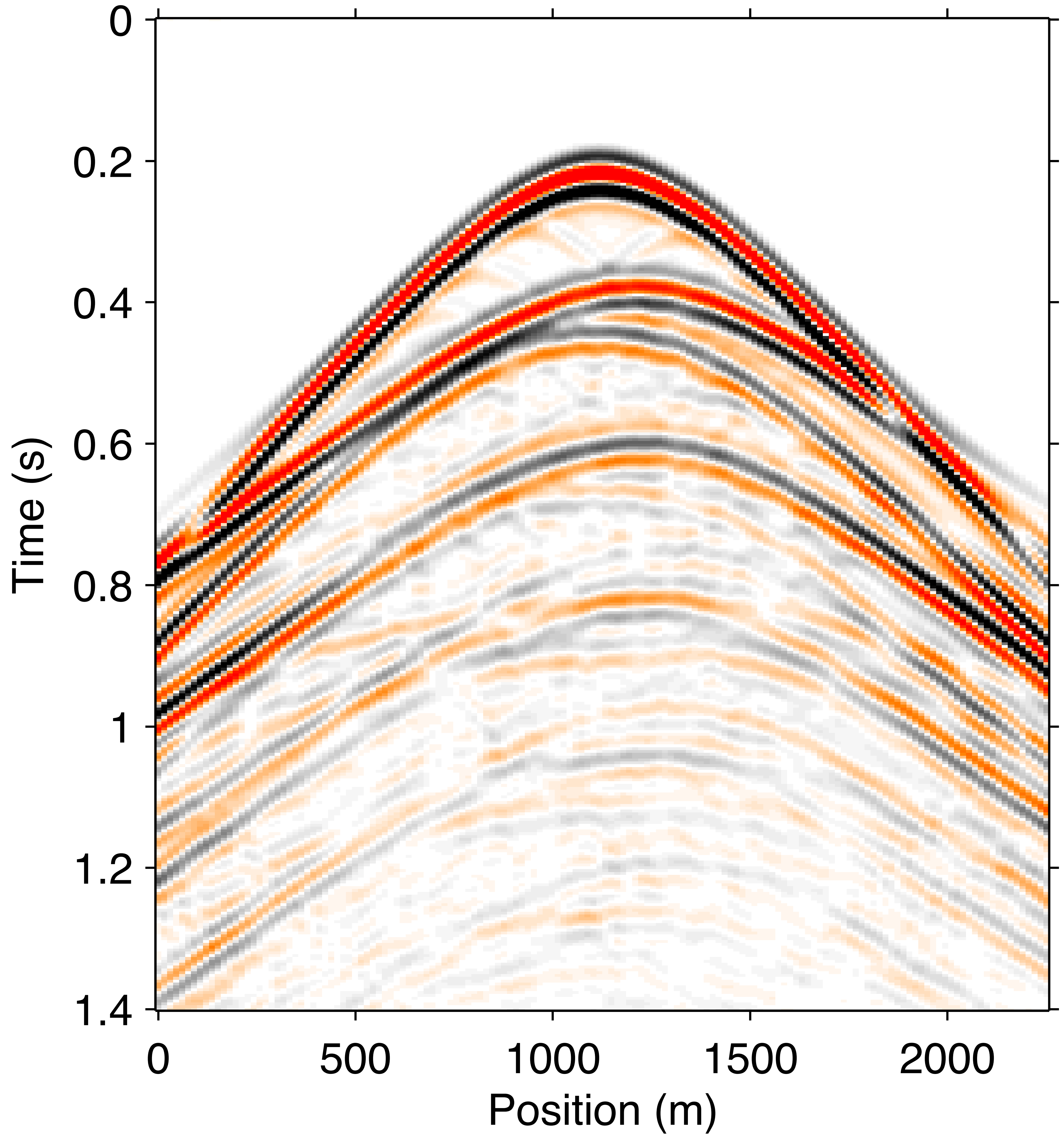
Data
modeled with Ricker 30Hz

Motivation: G tolerates lowpass filtering



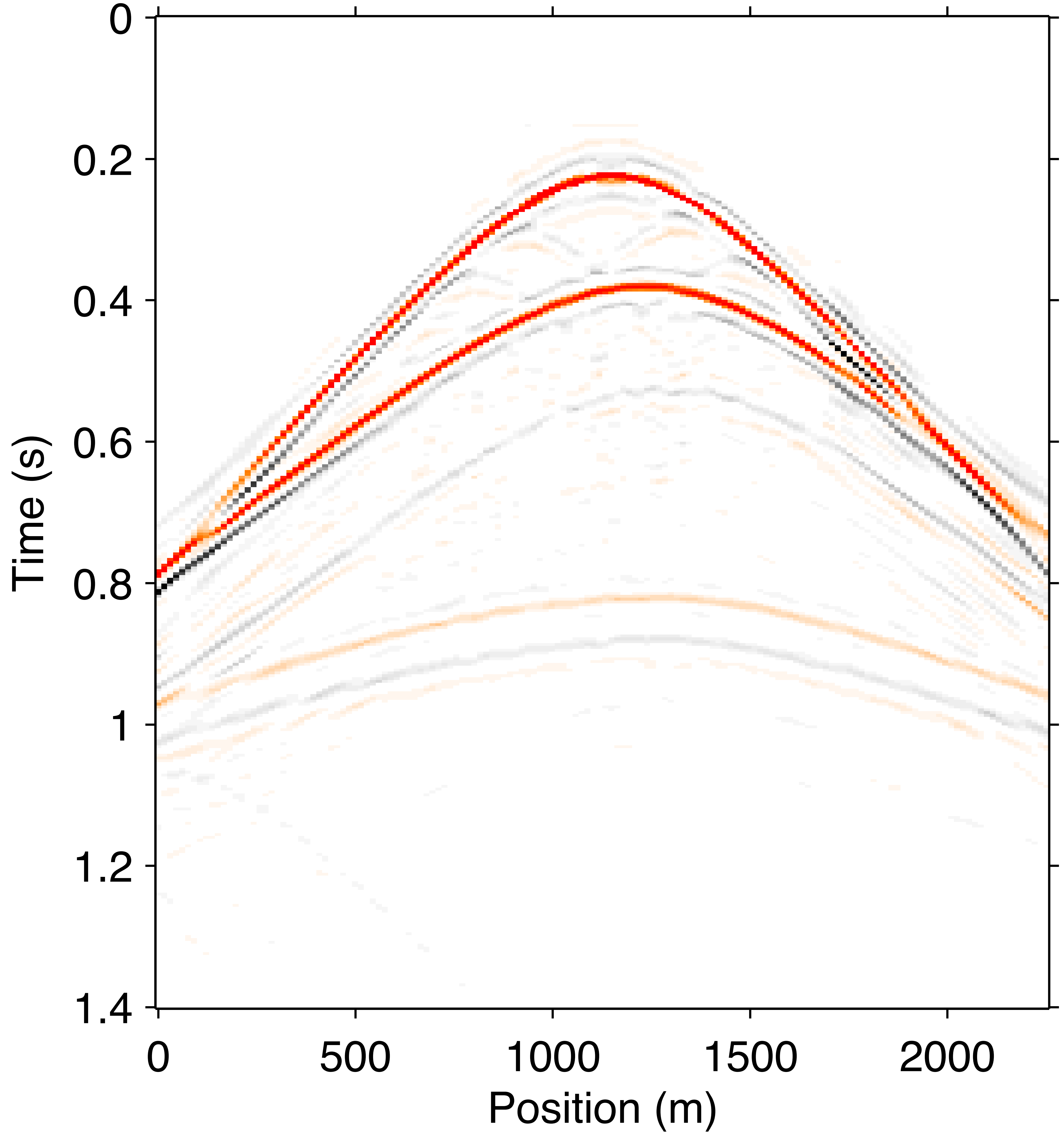
Reference REPSI primary IR
from original data

Motivation: G tolerates lowpass filtering



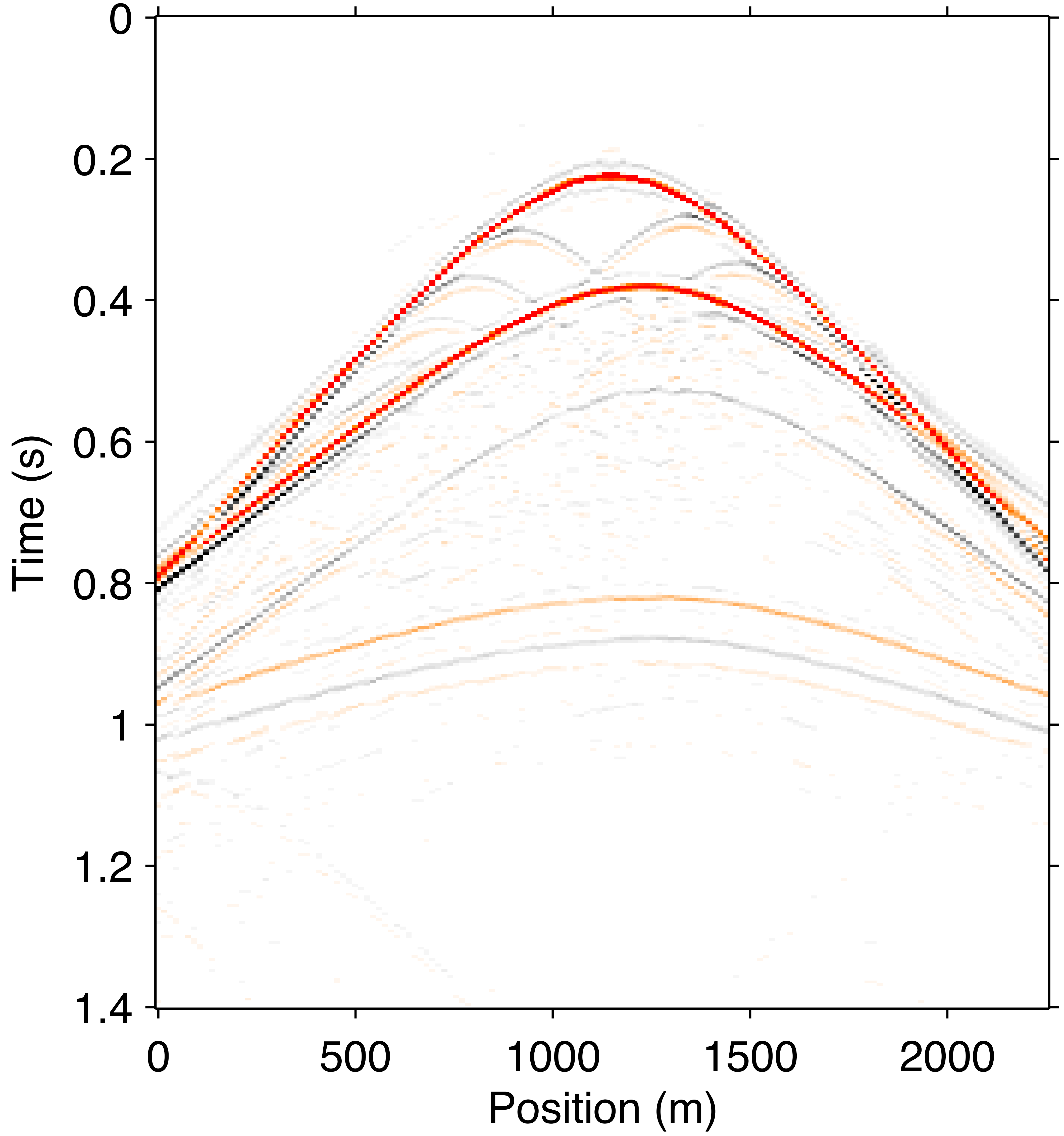
Lowpassed Data
modeled with Ricker 30Hz
lowpass at 40Hz
(25-order, zero-phase, Hann window)

Motivation: G tolerates lowpass filtering



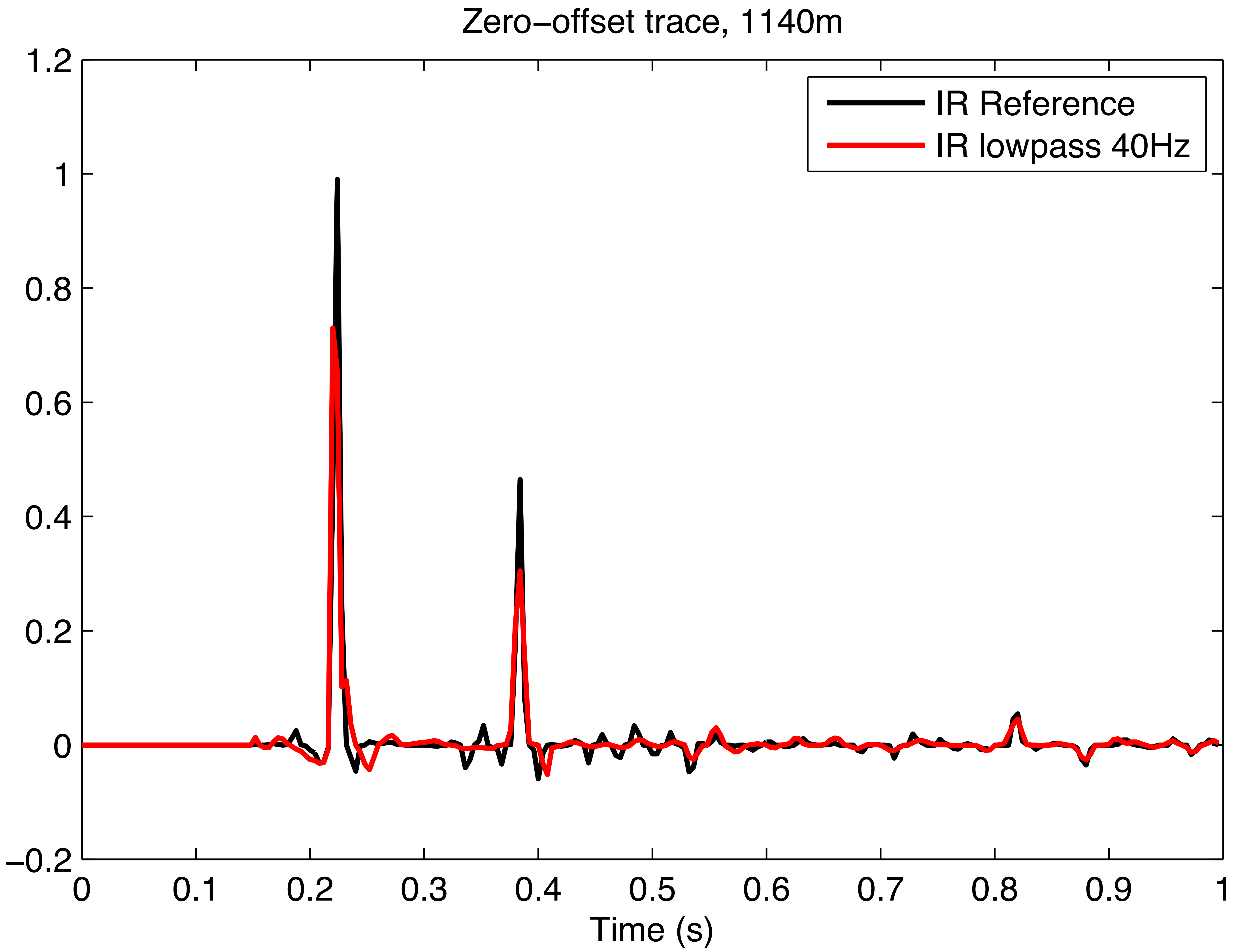
REPSI primary IR
from low-passed data @ 40Hz

Motivation: G tolerates lowpass filtering



Reference REPSI primary IR
from original data

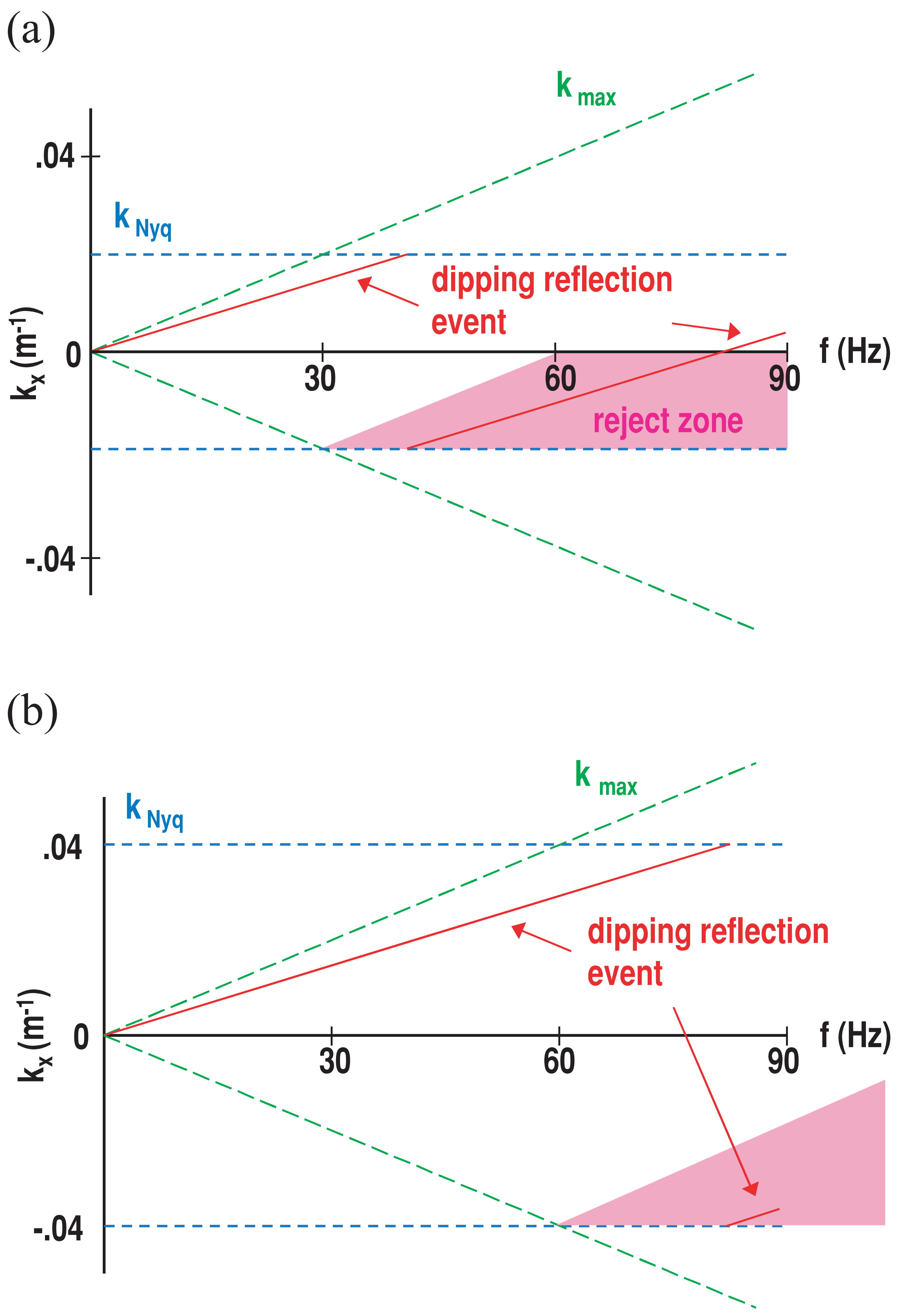
Motivation: G tolerates lowpass filtering

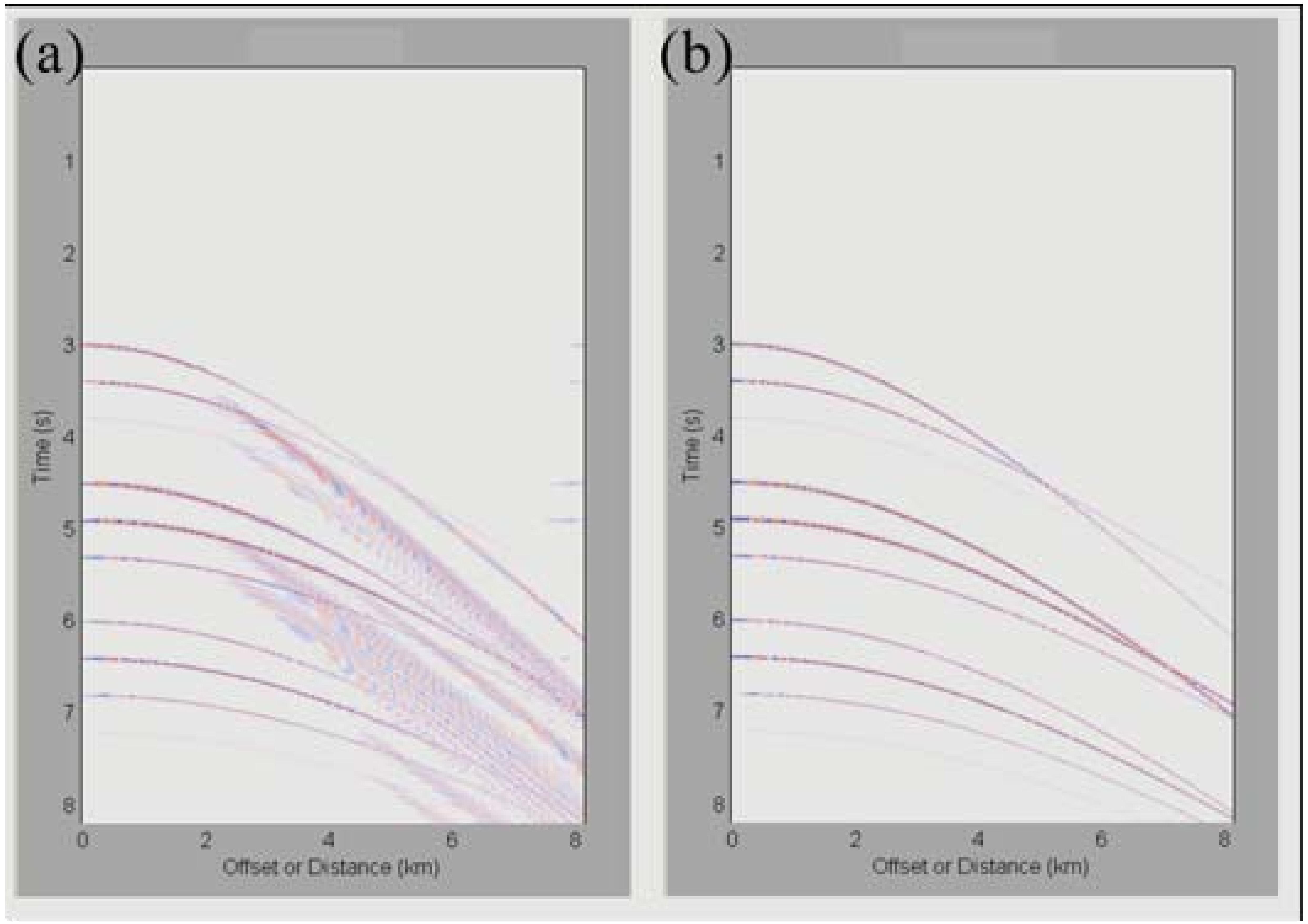


Why? Grid subsampling results in spatial aliasing

"The impact of field-survey characteristics on surface-related multiple attenuation"

Dragoset, Moore, Kostov 2006



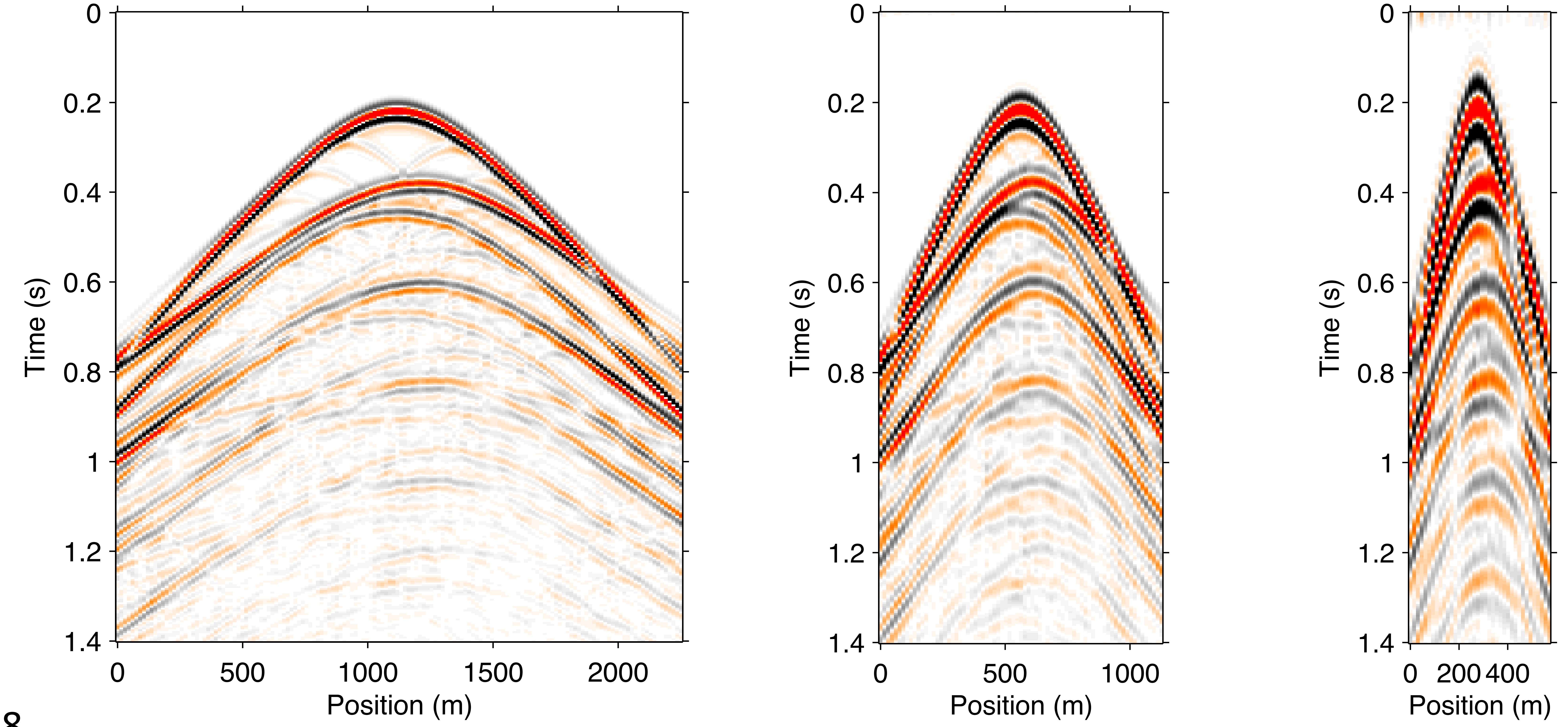


Lowpass data permits coarser sampling w/o aliasing

Original (dx = 15m)

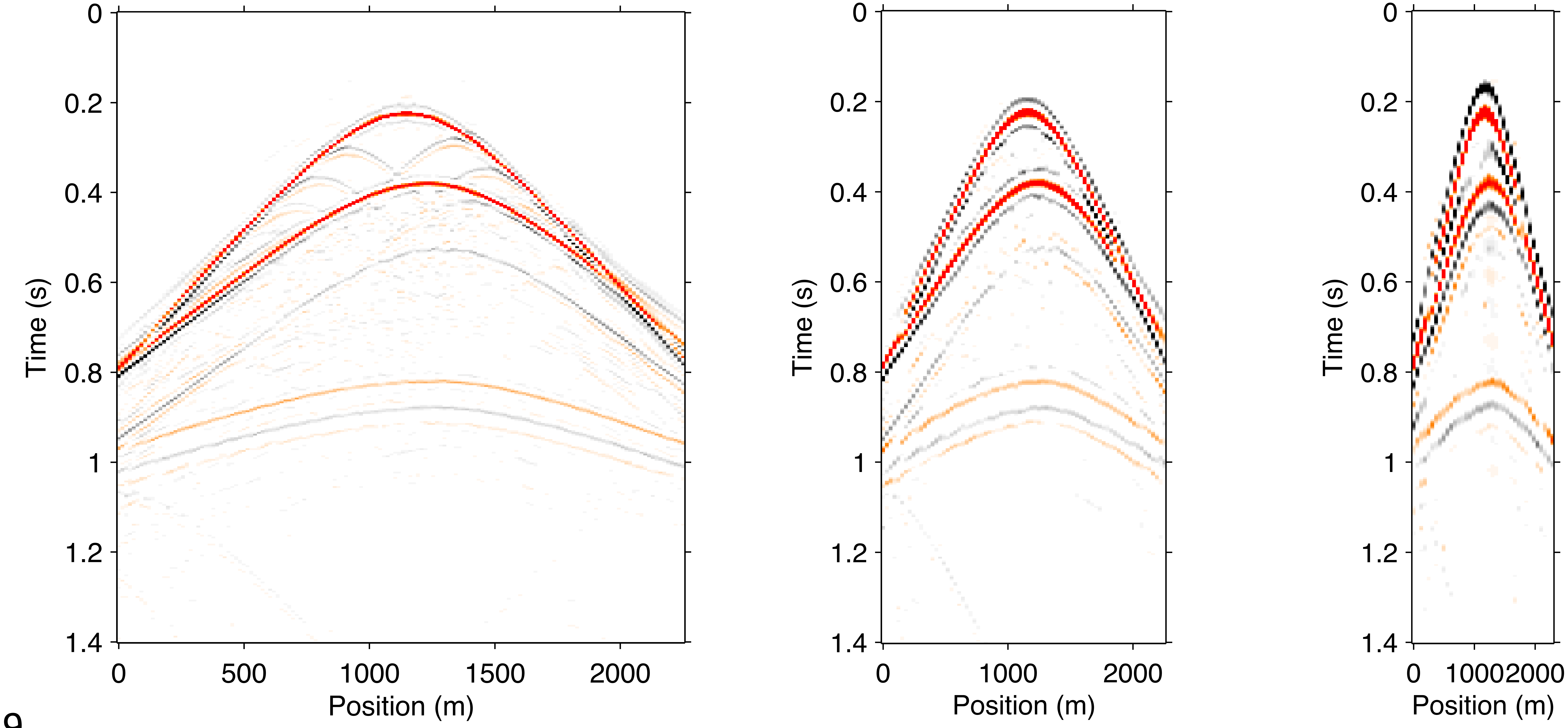
2x decimated
lowpass 30Hz

4x decimated
lowpass 15Hz

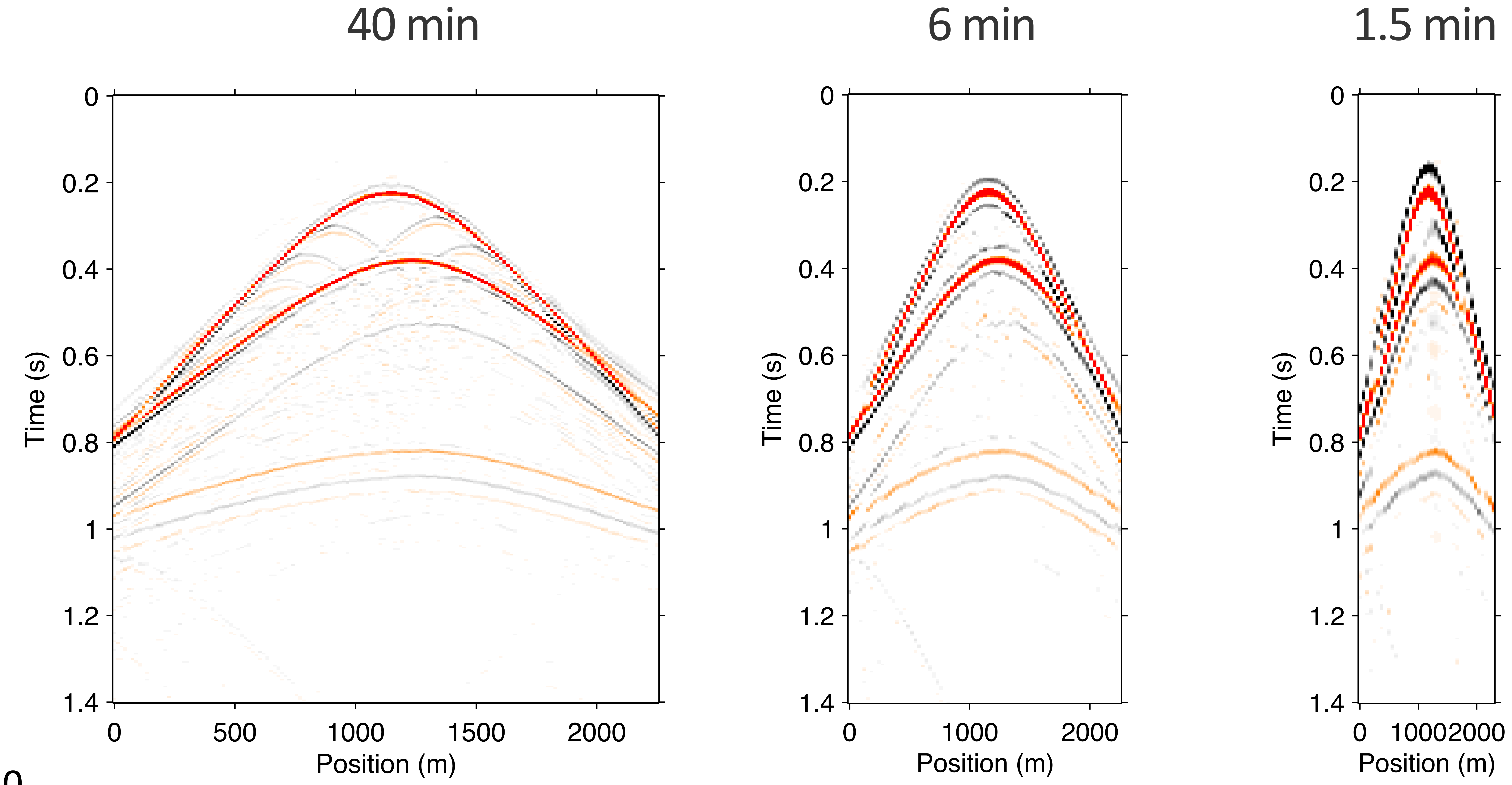


Lowpass data permits coarser sampling w/o aliasing

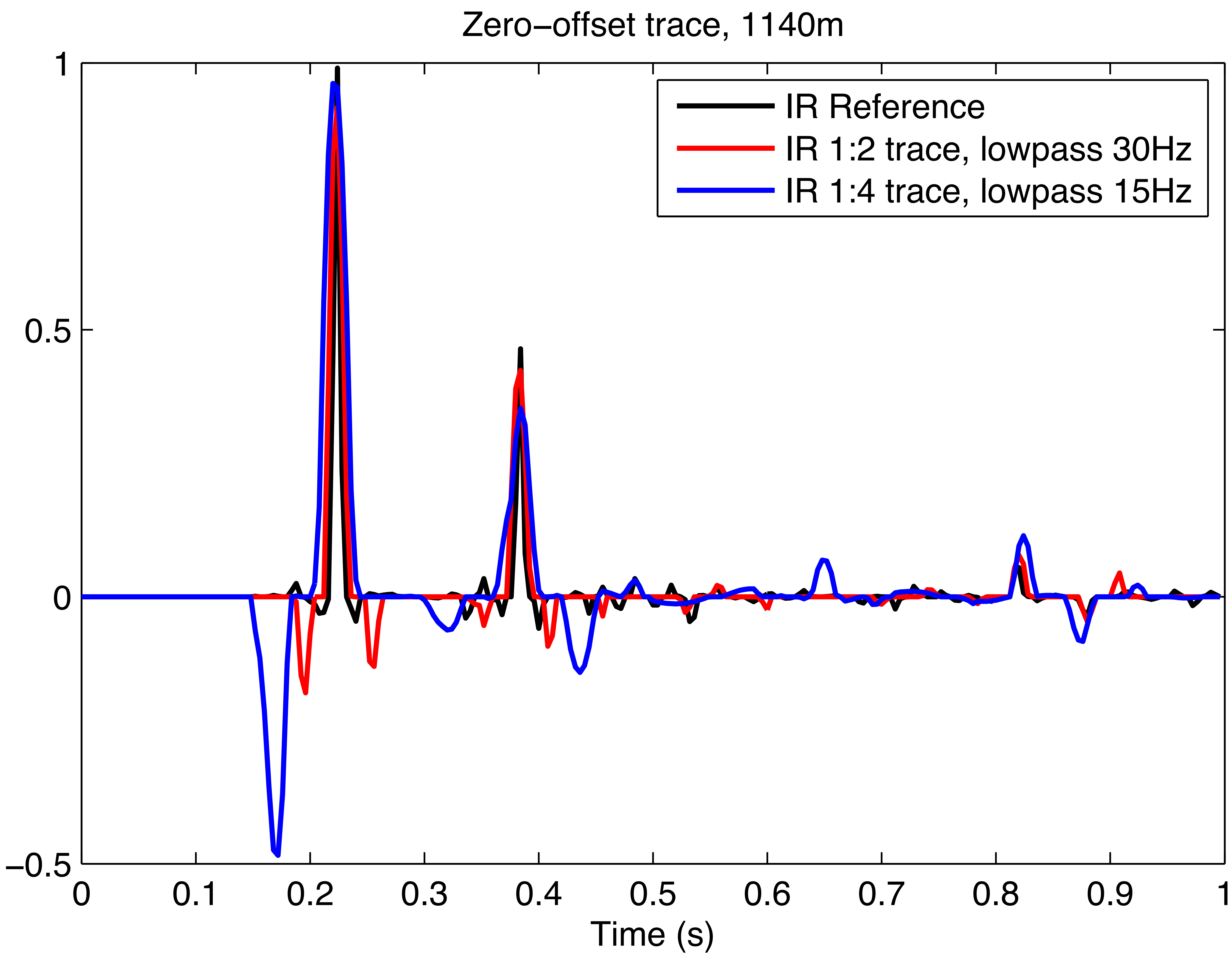
Impulse response solutions



Lowpass data permits coarser sampling w/o aliasing
(much faster!)



Lowpass data permits coarser sampling w/o aliasing



Multilevel strategy for EPSI

warm-start fine-scale problem *(slow)* with coarse-scale solutions *(fast)*

Significant speedup from bootstrapping (in 2D)

Per-iteration FLOPs cost (one forward/adjoint): $n = n_{\mathrm{rcv}} = n_{\mathrm{src}}$

$$\text{Cost}(n) = \mathcal{O}(2n_t n^2 \log n_t) + \mathcal{O}(n_f n^3)$$

2 times FFT

computing MCG & sum in FX

$$\text{Cost}\left(\frac{1}{2}n\right) = \frac{1}{4}\mathcal{O}(2n_t n^2 \log n_t) + \frac{1}{8}\mathcal{O}(n_f n^3)$$

$$\text{Cost}\left(\frac{1}{4}n\right) = \frac{1}{16}\mathcal{O}(2n_t n^2 \log n_t) + \frac{1}{64}\mathcal{O}(n_f n^3)$$

Significant speedup from bootstrapping (in 2D)

Per-iteration FLOPs cost (one forward/adjoint): $n = n_{\mathrm{rcv}} = n_{\mathrm{src}}$

$$\text{Cost}(n) = \mathcal{O}(2n_t n^2 \log n_t) + \mathcal{O}(n_f n^3)$$

2 times FFT

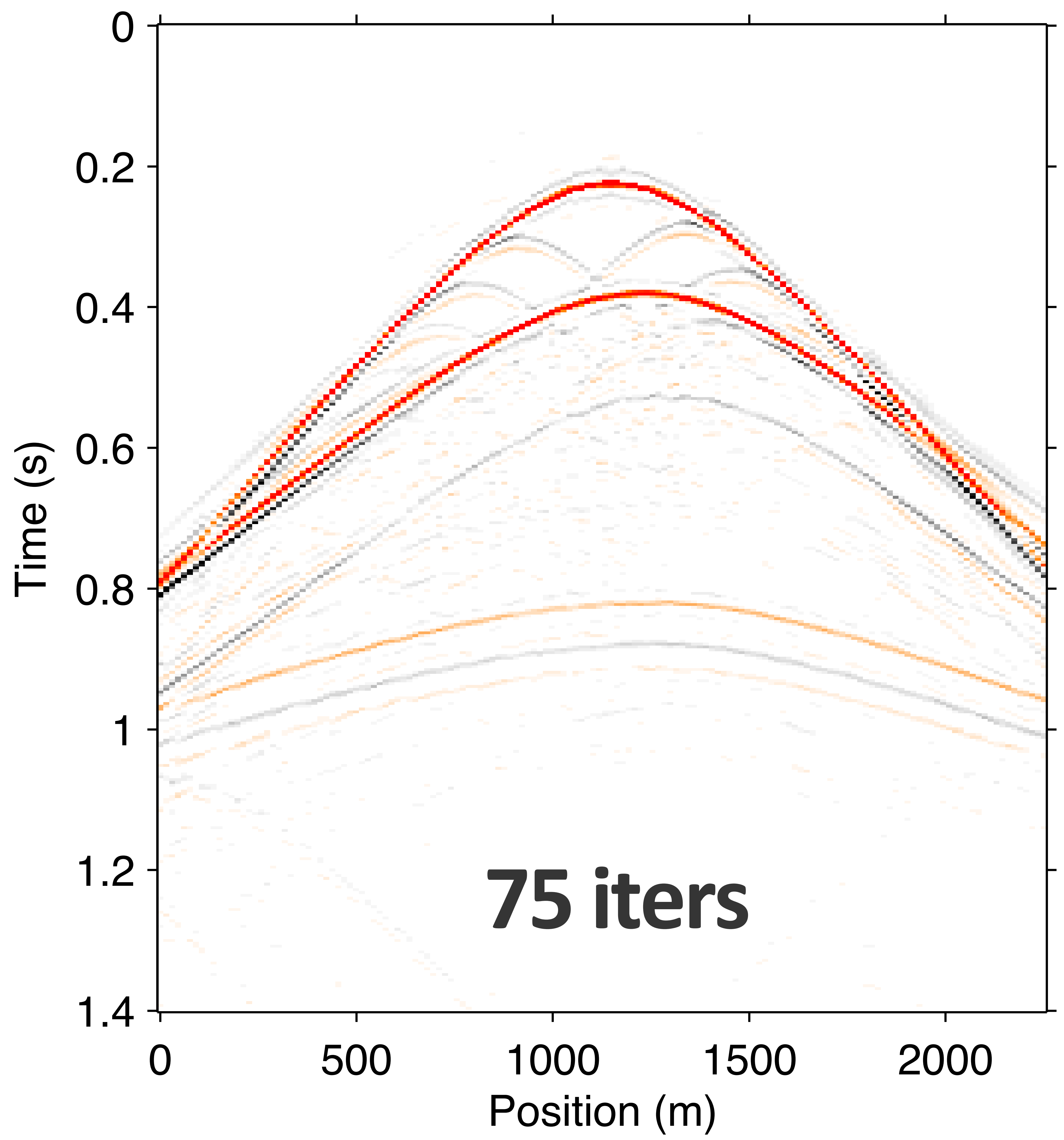
computing MCG & sum in FX

$$\text{Cost}\left(\frac{1}{2}n, \frac{1}{2}n_f\right) = \frac{1}{4}\mathcal{O}(2n_t n^2 \log n_t) + \frac{1}{16}\mathcal{O}(n_f n^3)$$

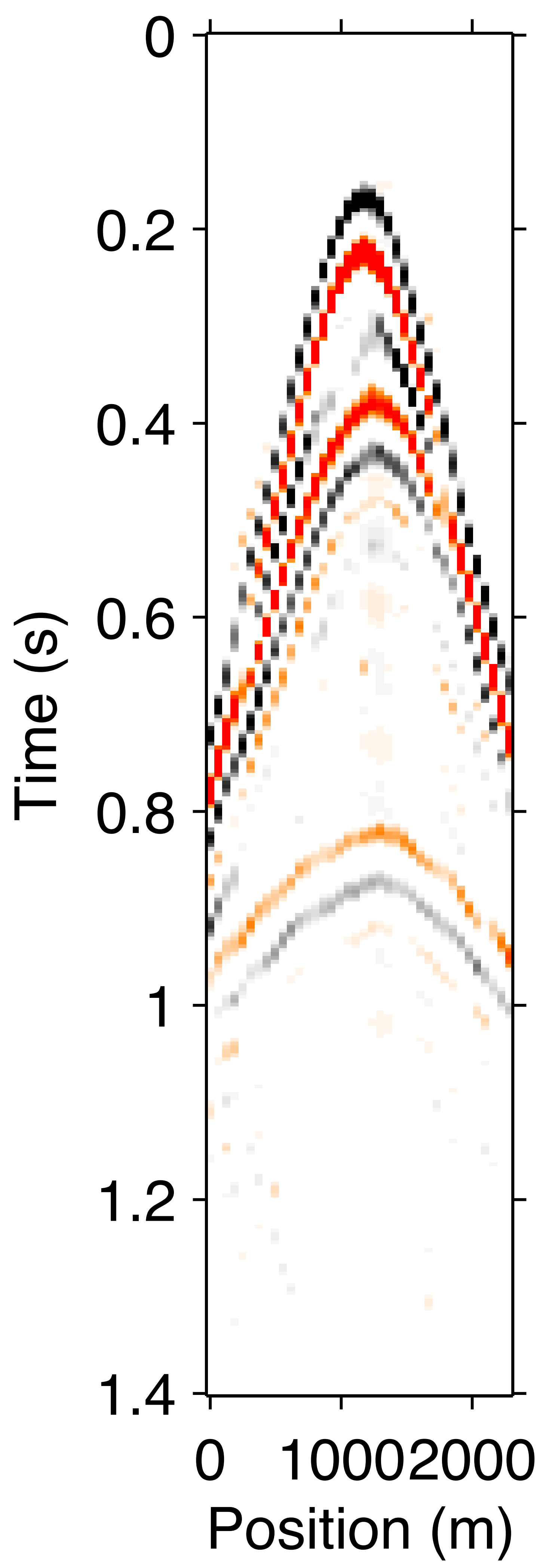
$$\text{Cost}\left(\frac{1}{4}n, \frac{1}{4}n_f\right) = \frac{1}{16}\mathcal{O}(2n_t n^2 \log n_t) + \frac{1}{128}\mathcal{O}(n_f n^3)$$

Warm-starting/continuation from coarse solution
Example

Solution of full data



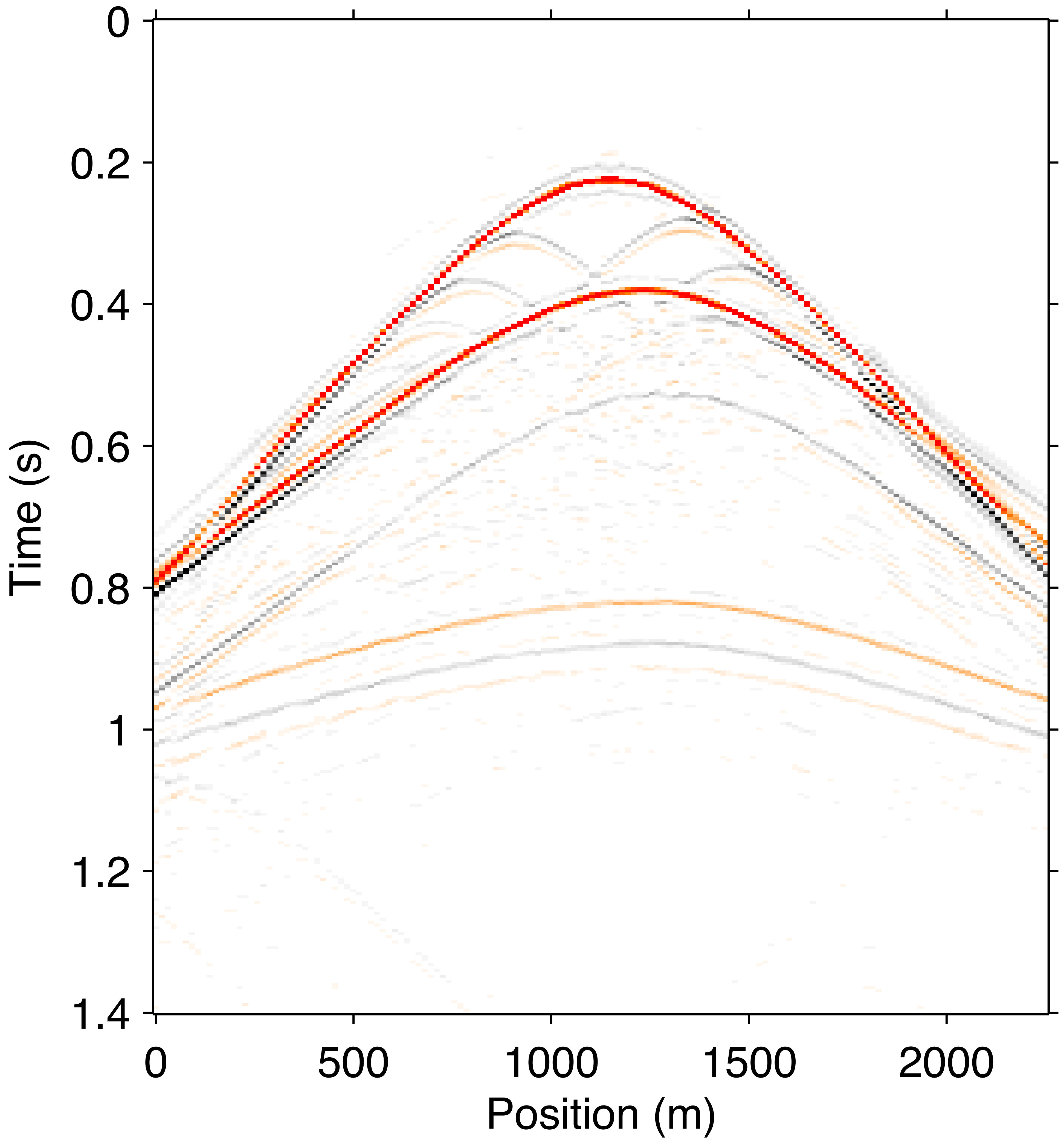
Solution of 4x decimated data



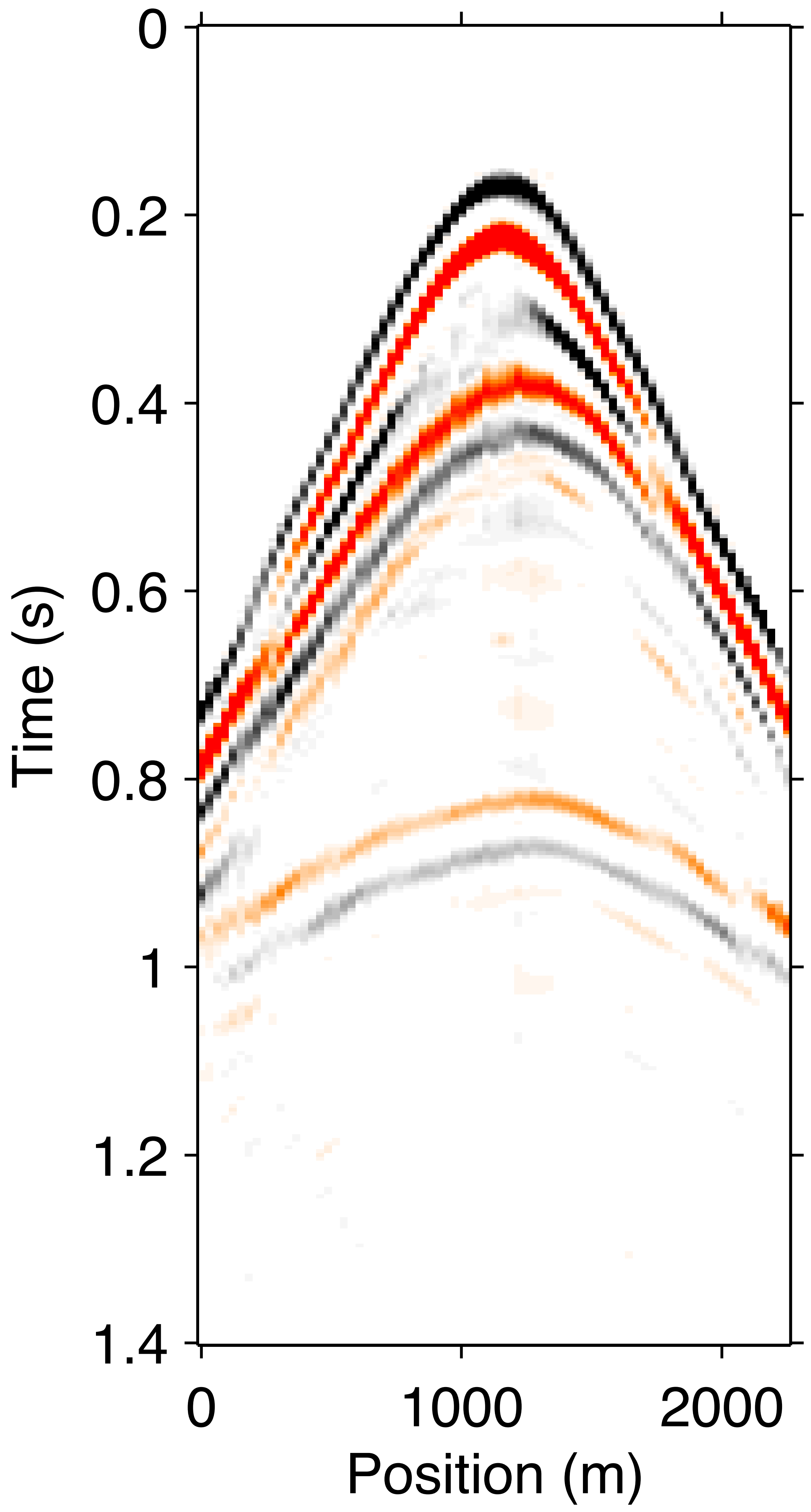
60 iters

Warm-starting/continuation from coarse solution
Example

Solution of full data



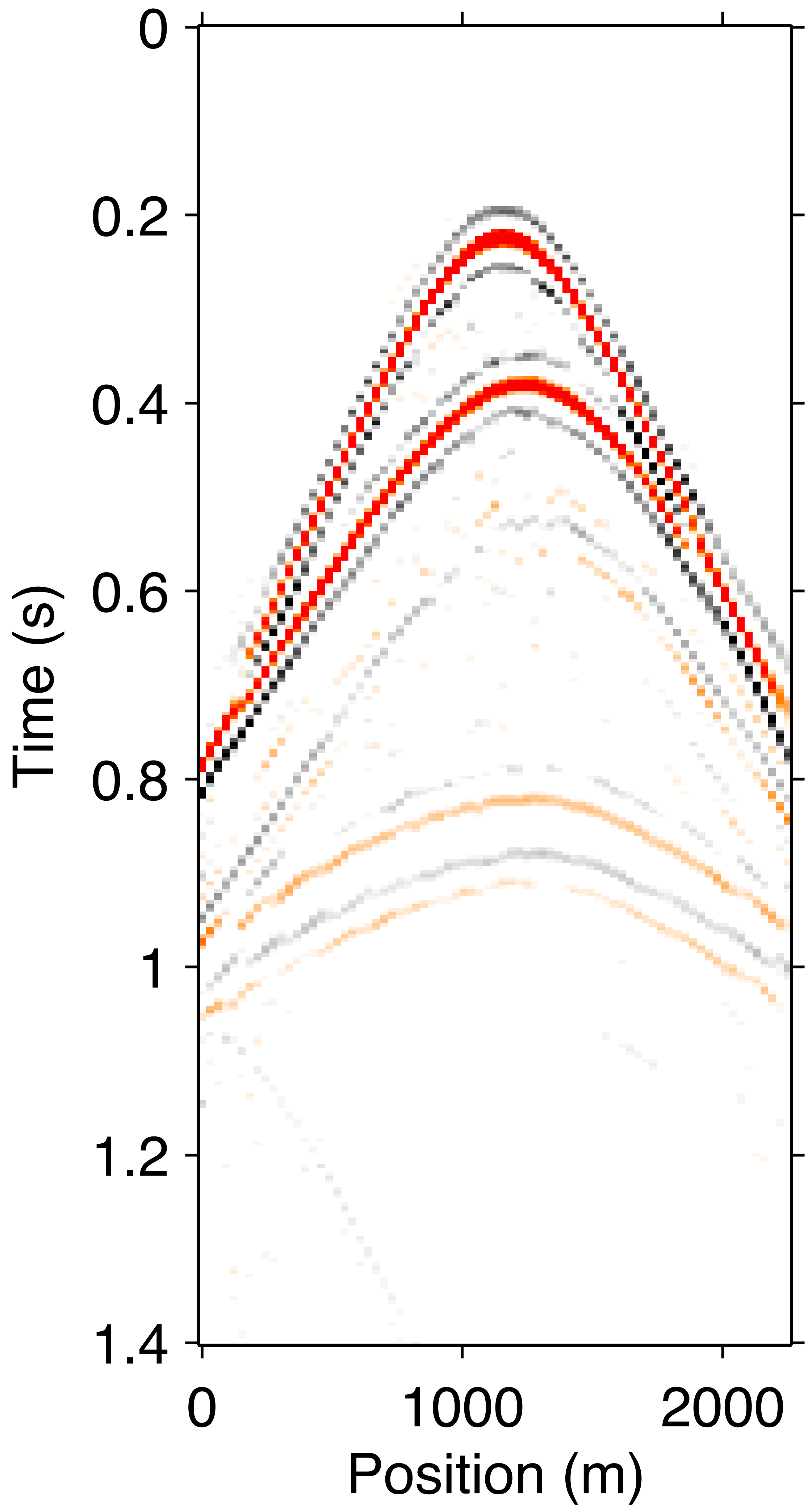
Solution of 4x decimated data
1600m/s NMO, linear interp 2x



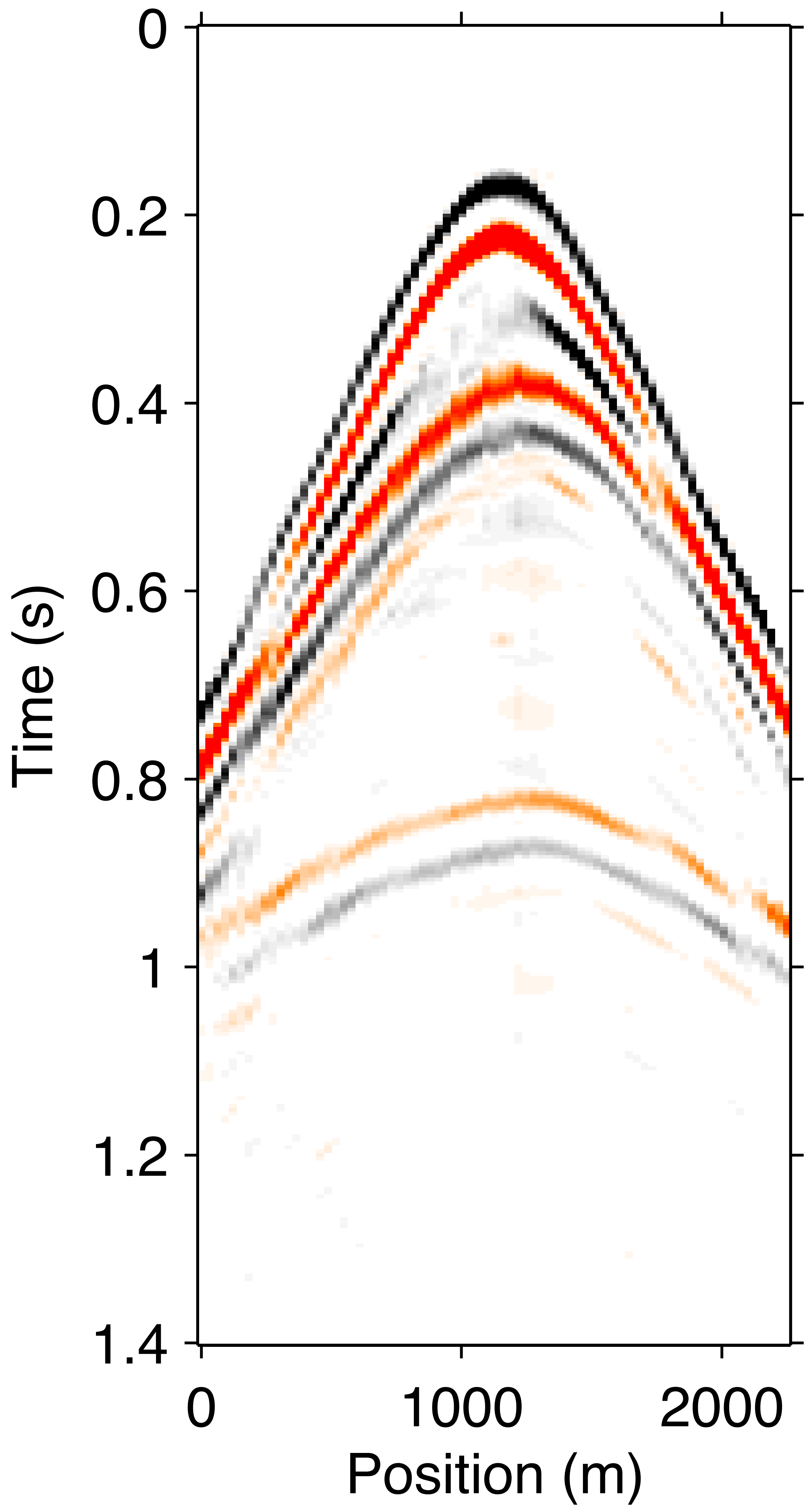
Prolongation

Warm-starting/continuation from coarse solution
Example

Solution of 2x decimated data



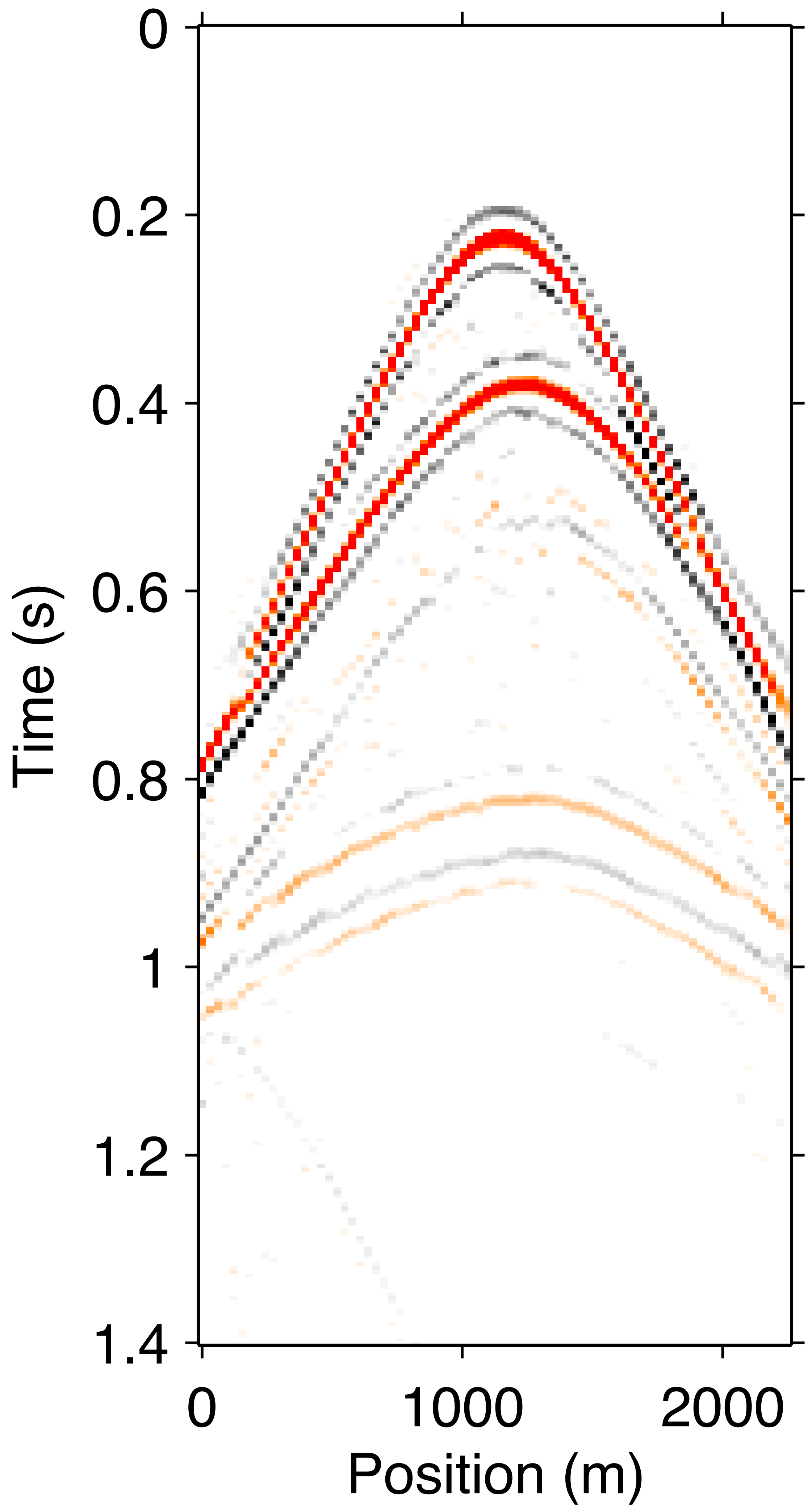
Solution of 4x decimated data
1600m/s NMO, linear interp 2x



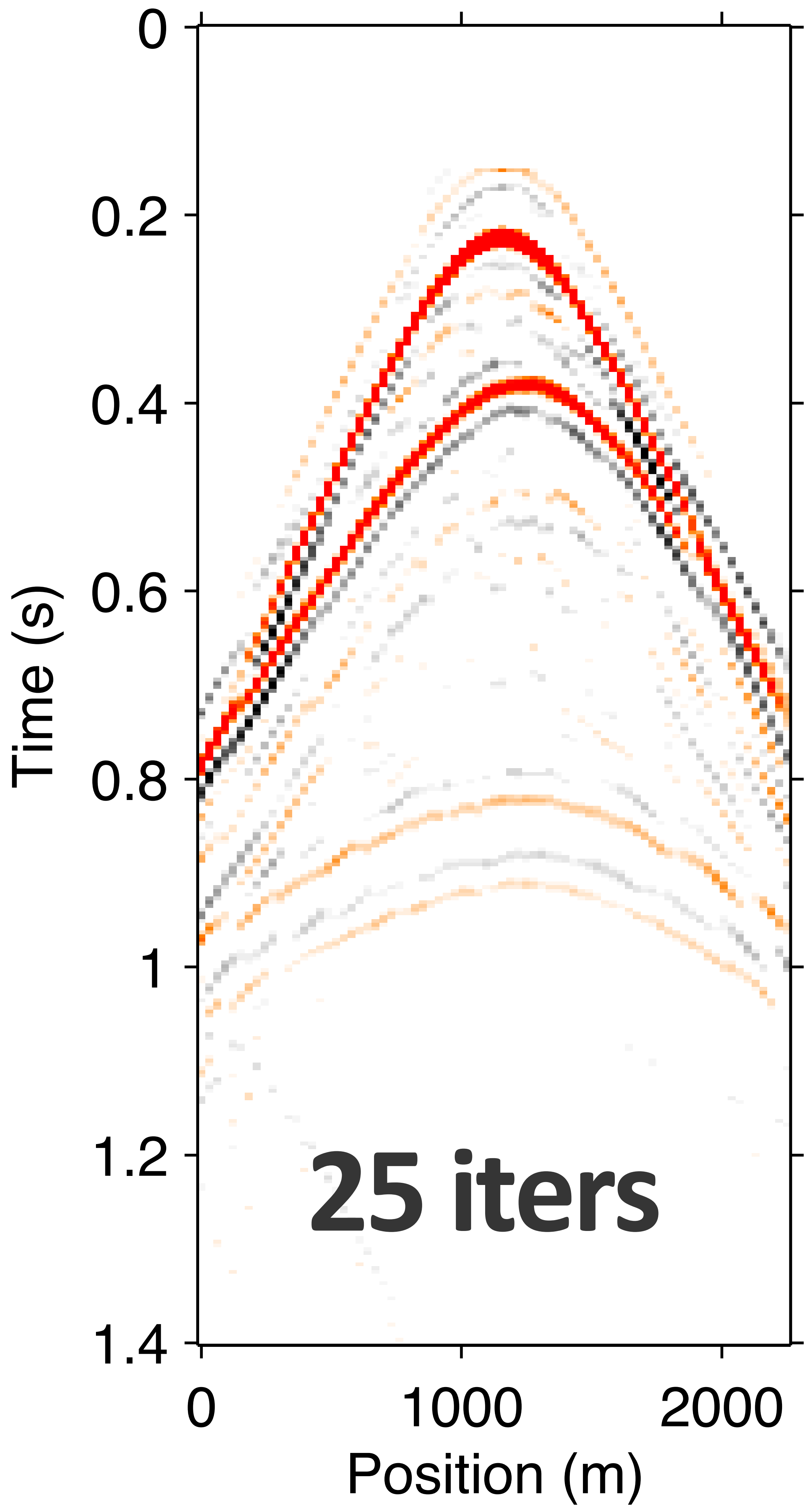
Solve

Warm-starting/continuation from coarse solution
Example

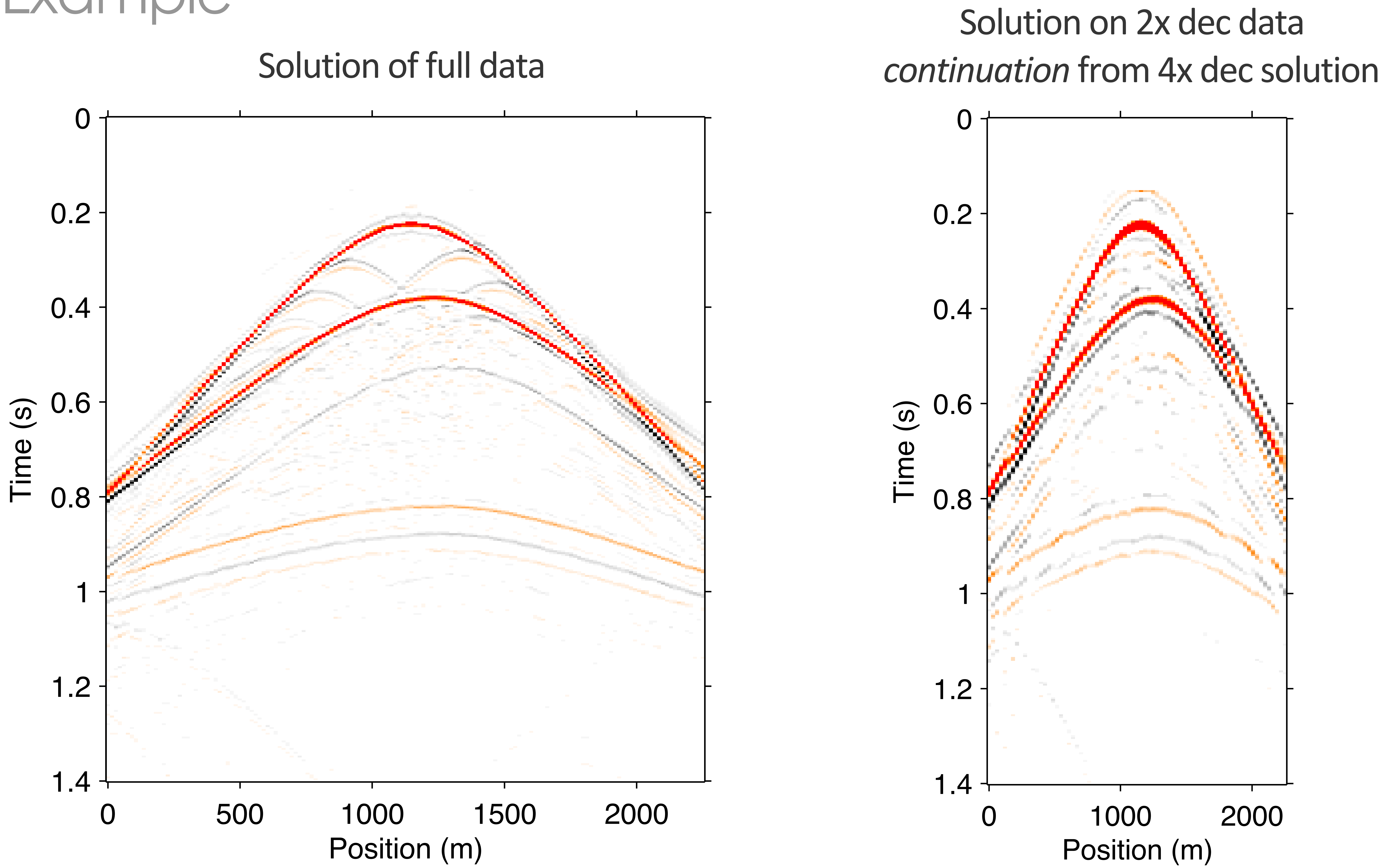
Solution of 2x decimated data



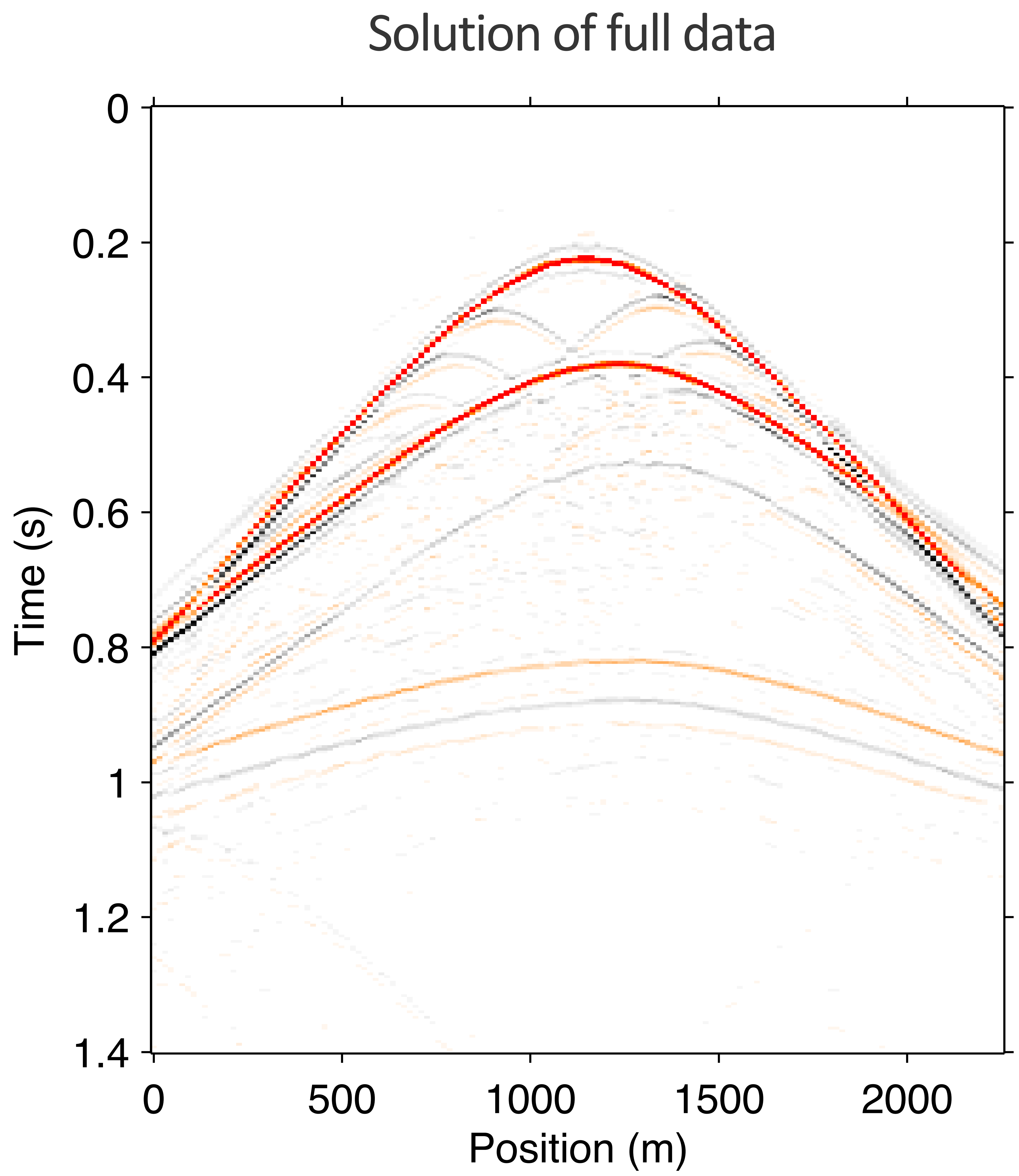
Solution on 2x dec data
continuation from 4x dec solution

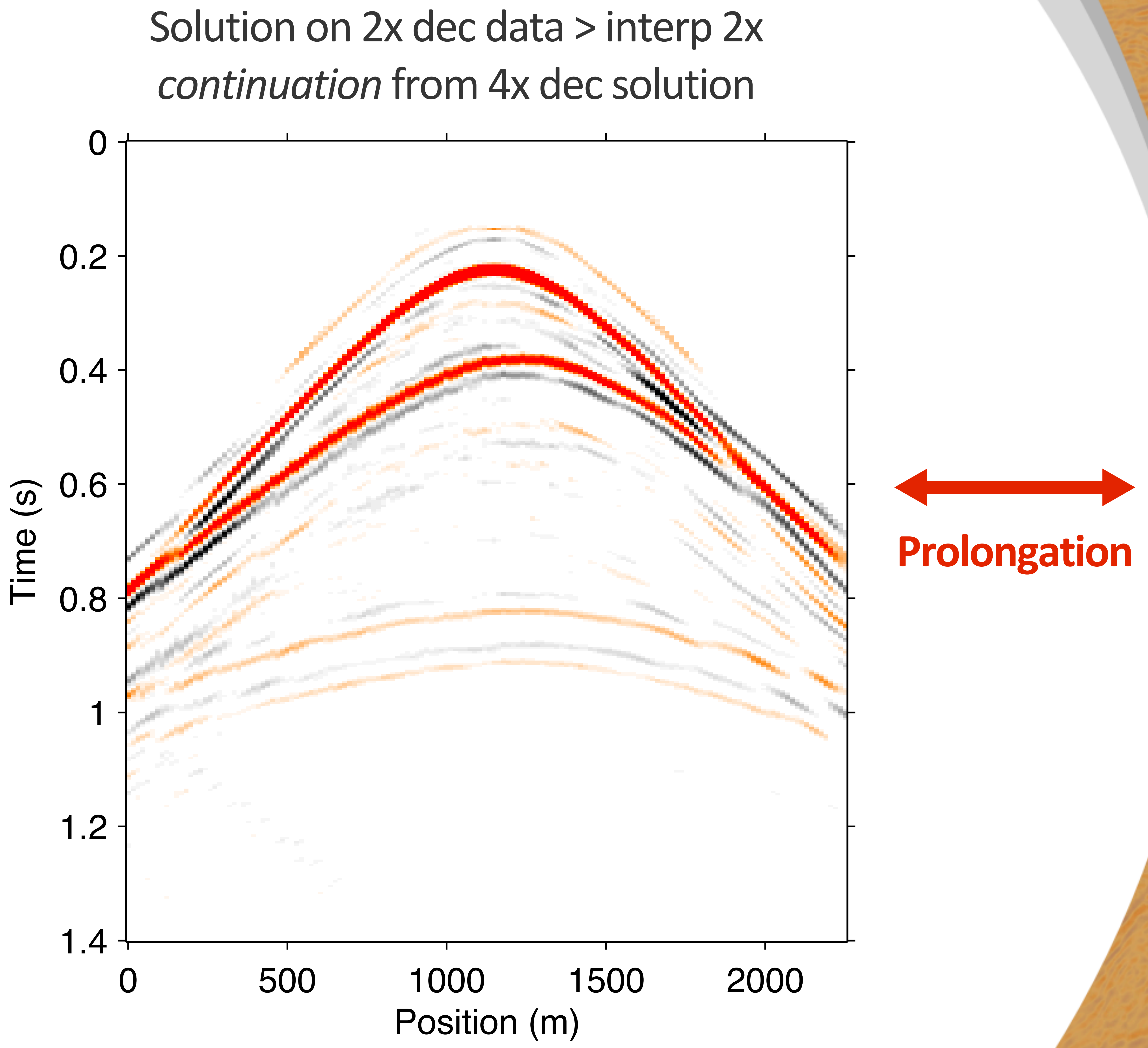


Warm-starting/continuation from coarse solution
Example



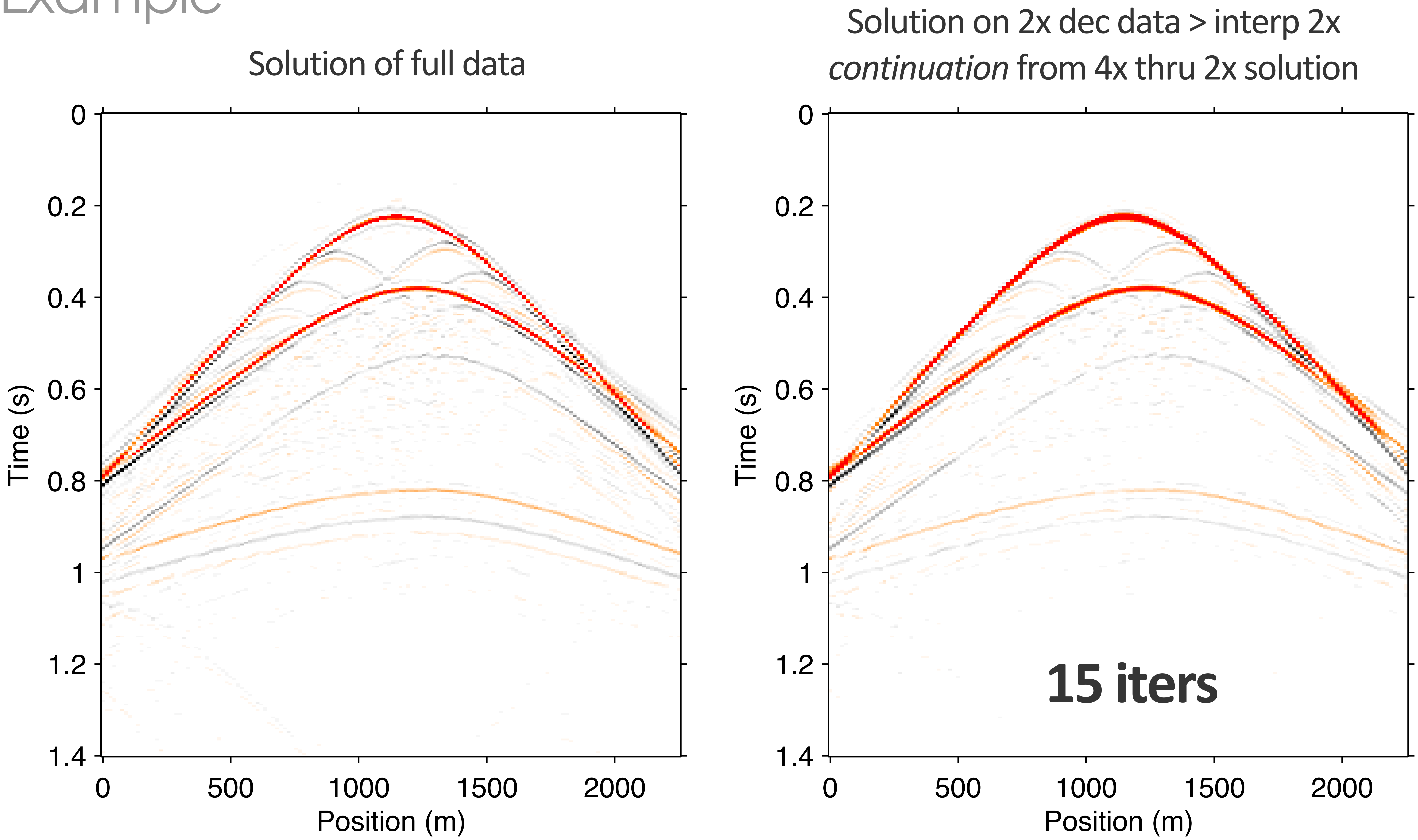
Warm-starting/continuation from coarse solution
Example





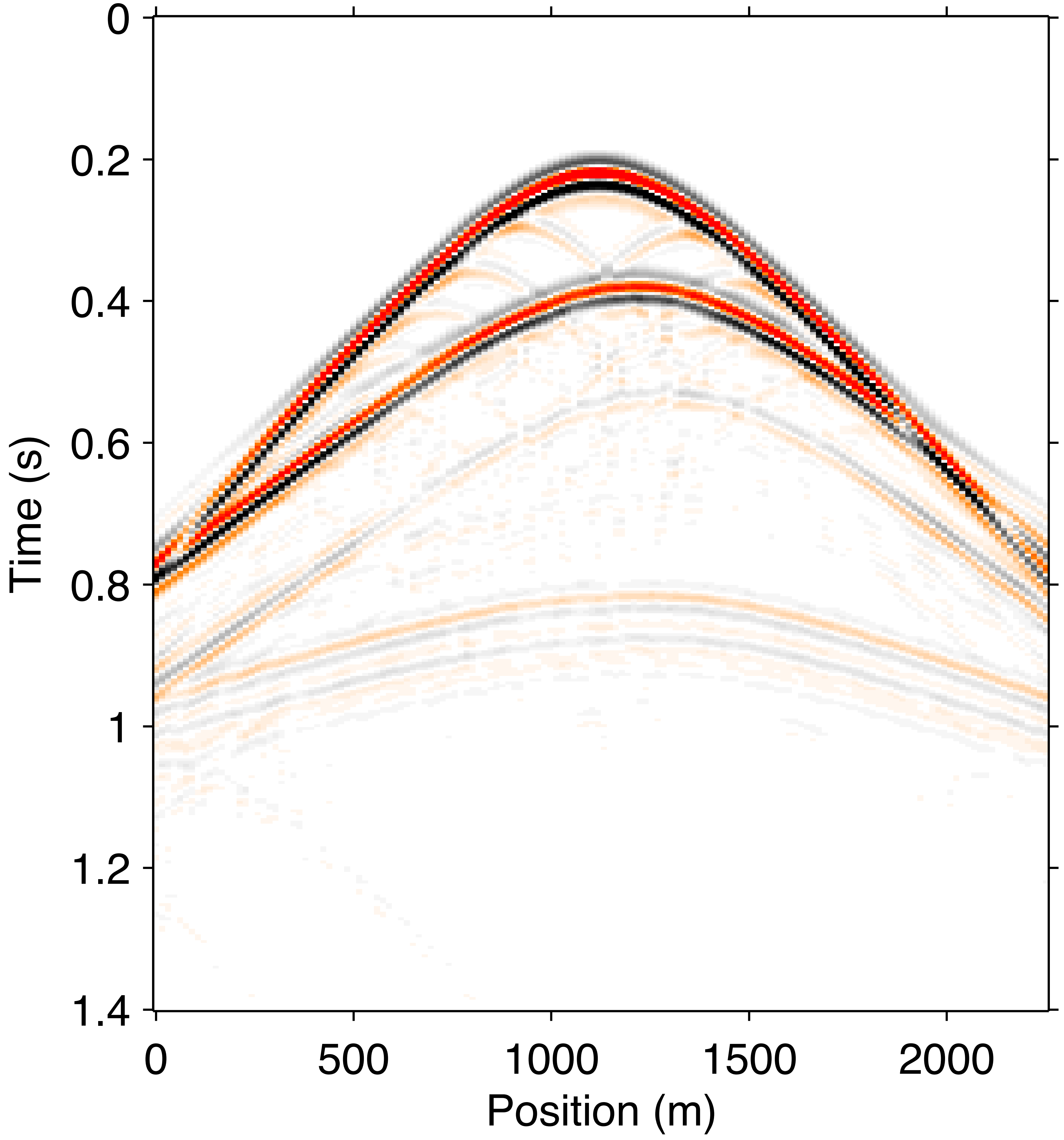
Warm-starting/continuation from coarse solution

Example



Solve

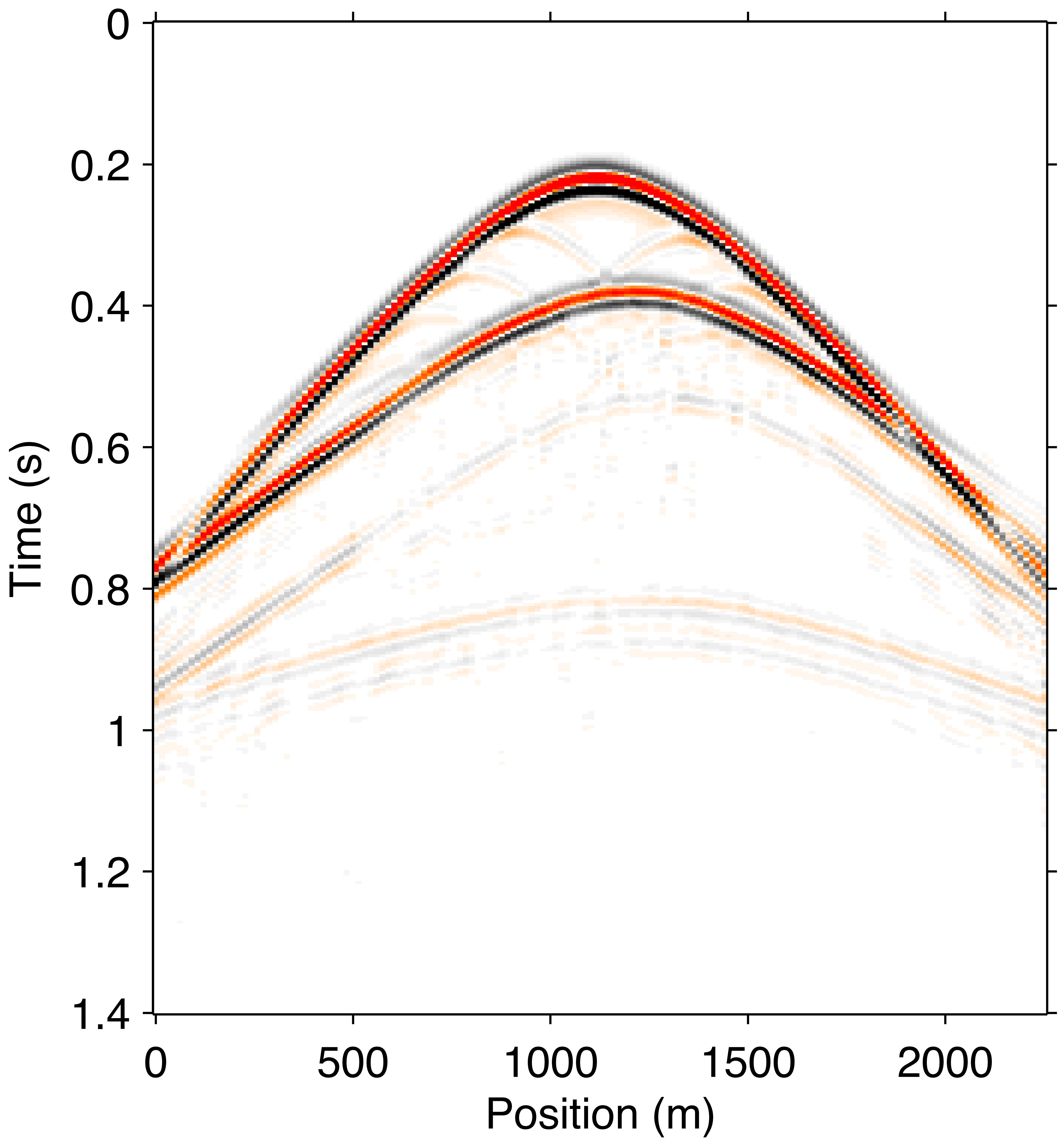
Warm-starting/continuation from coarse solution
Example



Direct Primary

Solved with plain algorithm from finest scale data

Warm-starting/continuation from coarse solution
Example

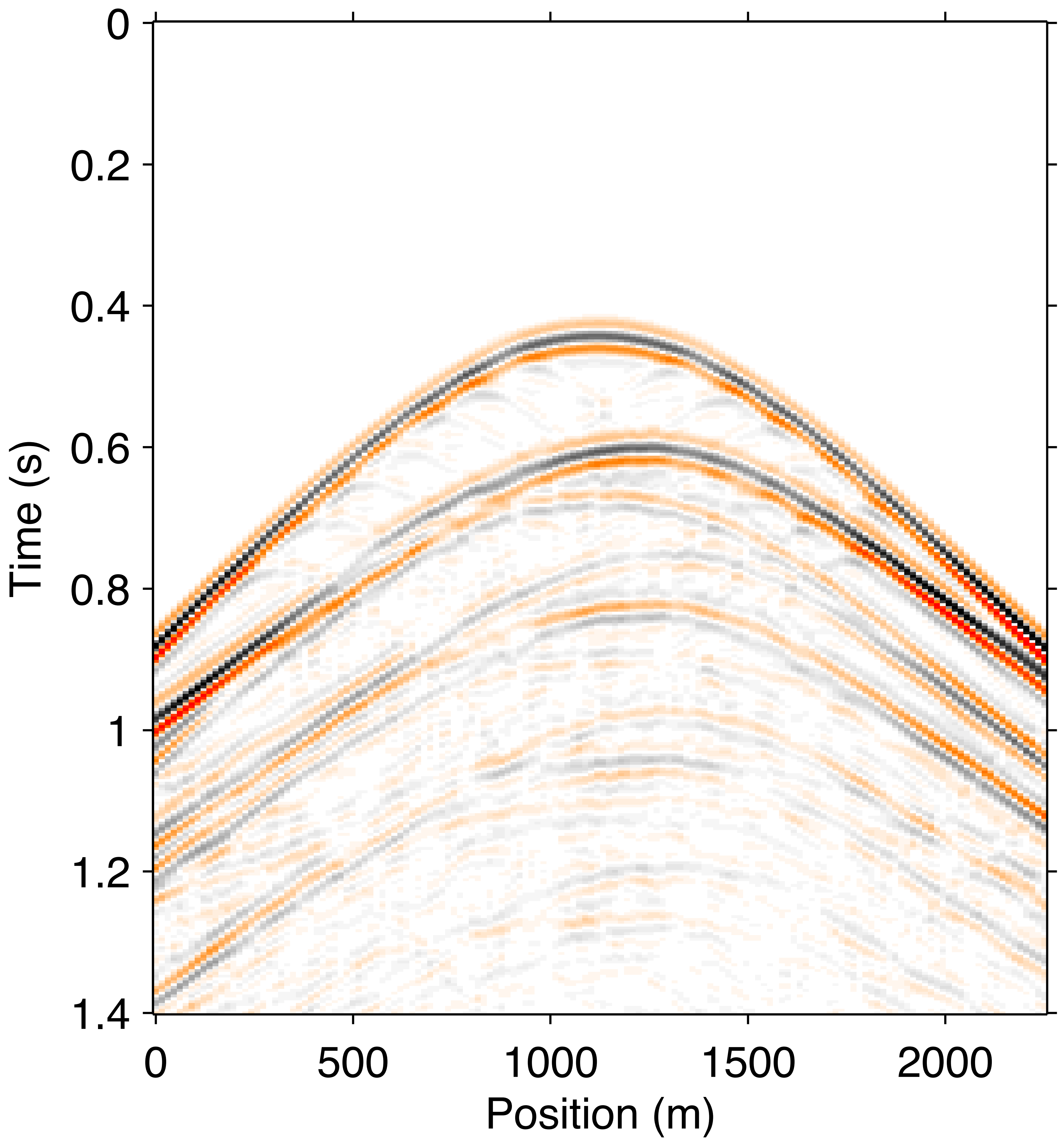


Direct Primary

Solved with spatial sampling continuation

dx = 60m > 30m > 15m

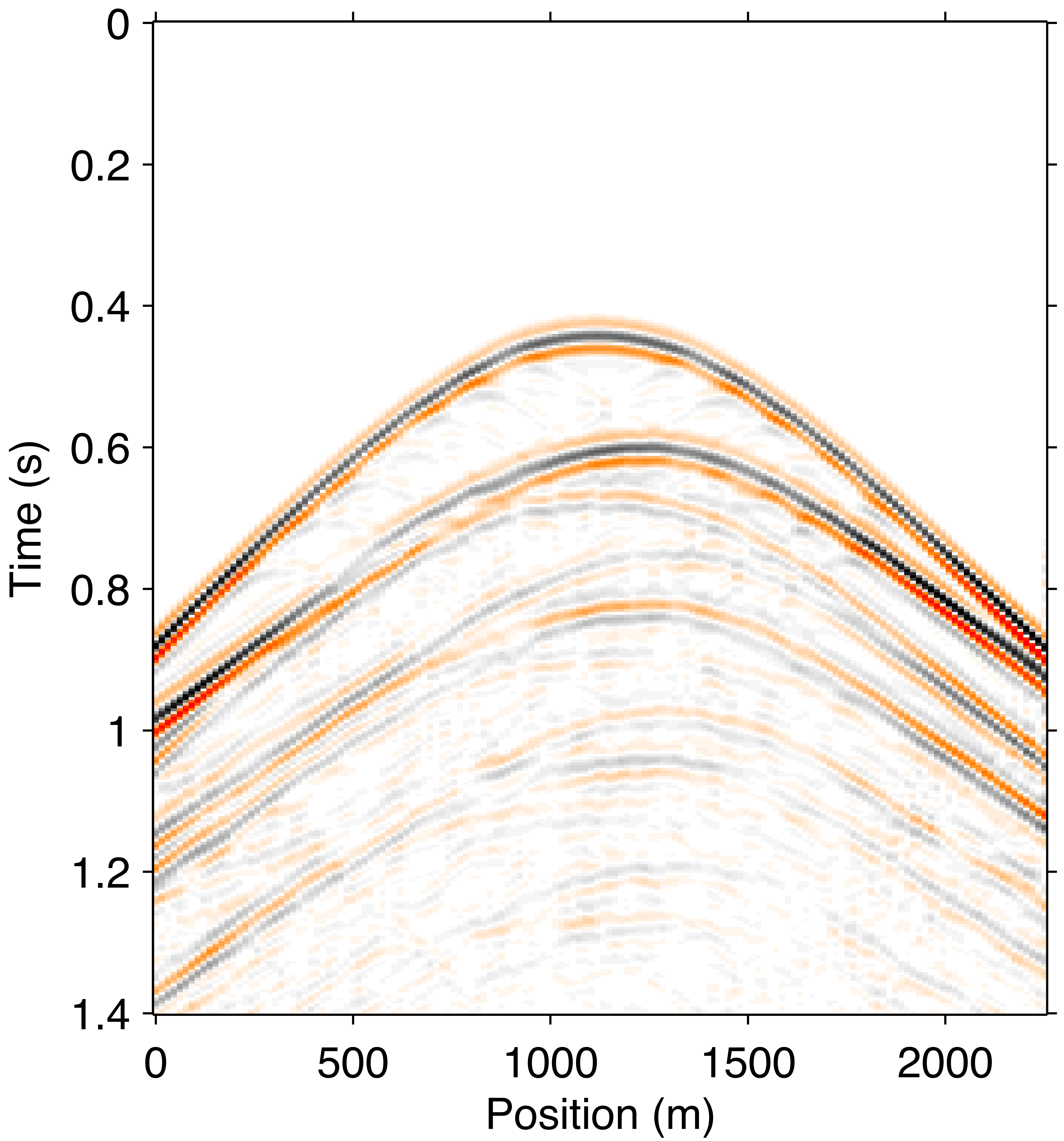
Warm-starting/continuation from coarse solution
Example



Predicted Surface Multiple

Solved with plain algorithm from finest scale data

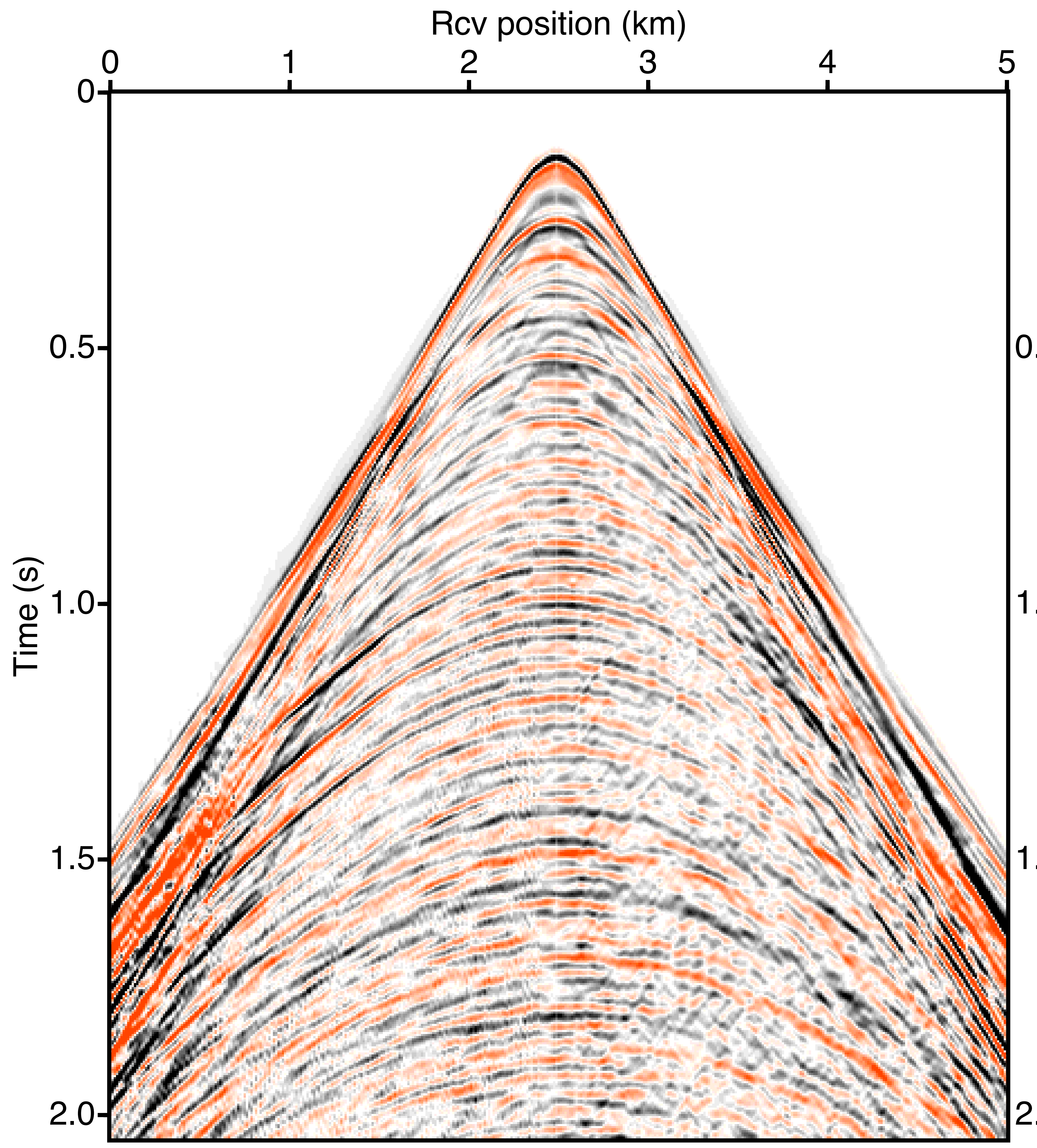
Warm-starting/continuation from coarse solution
Example



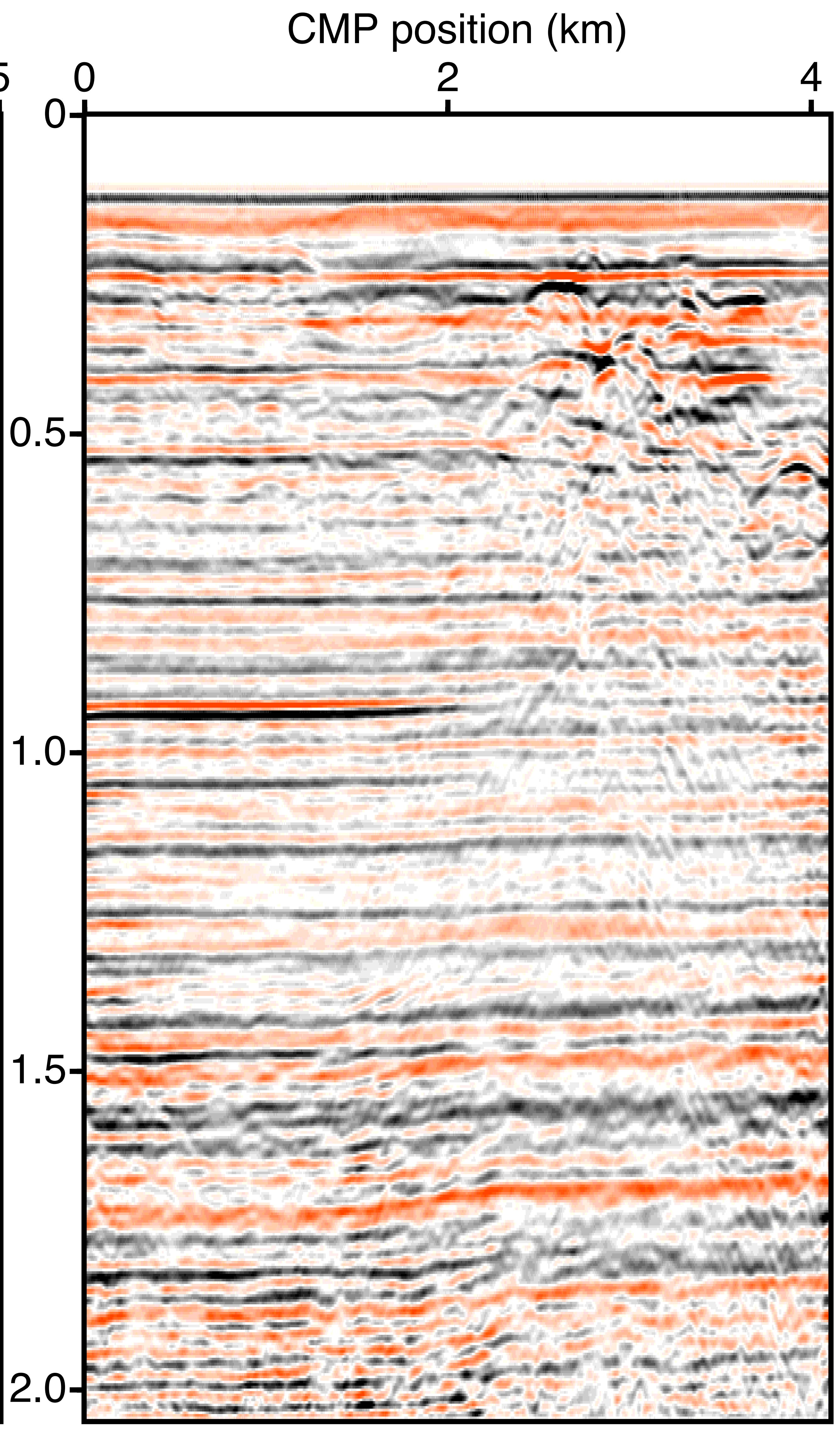
Predicted Surface Multiple

Solved with spatial sampling continuation

dx = 60m > 30m > 15m



Shot gather



NMO-corrected stack

North sea data

Shot gather and stack

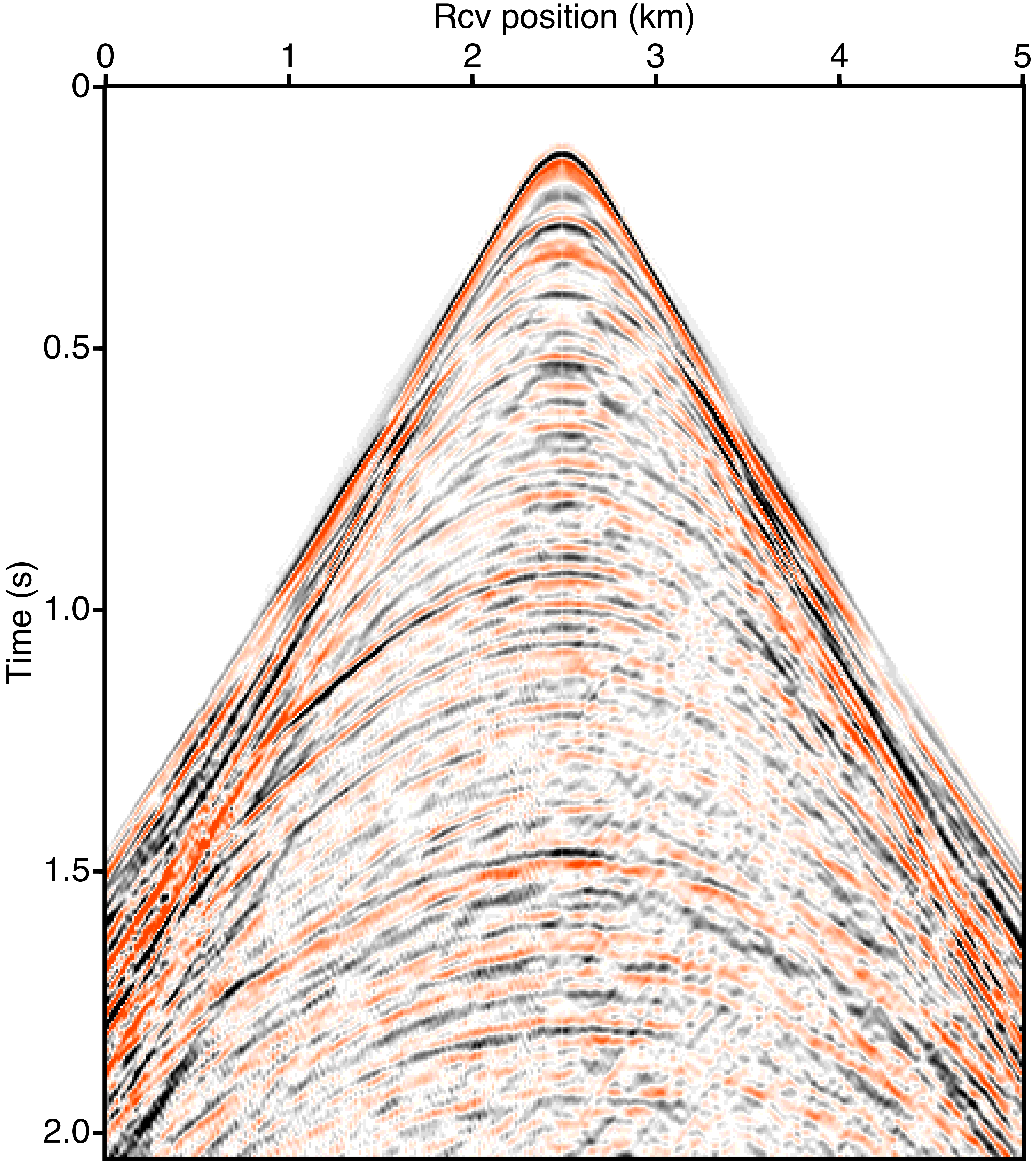
Streamer data *(regularized to fixed-spread data)*

401 source and reciever

12.5 m spatial grid

4 ms time sampling

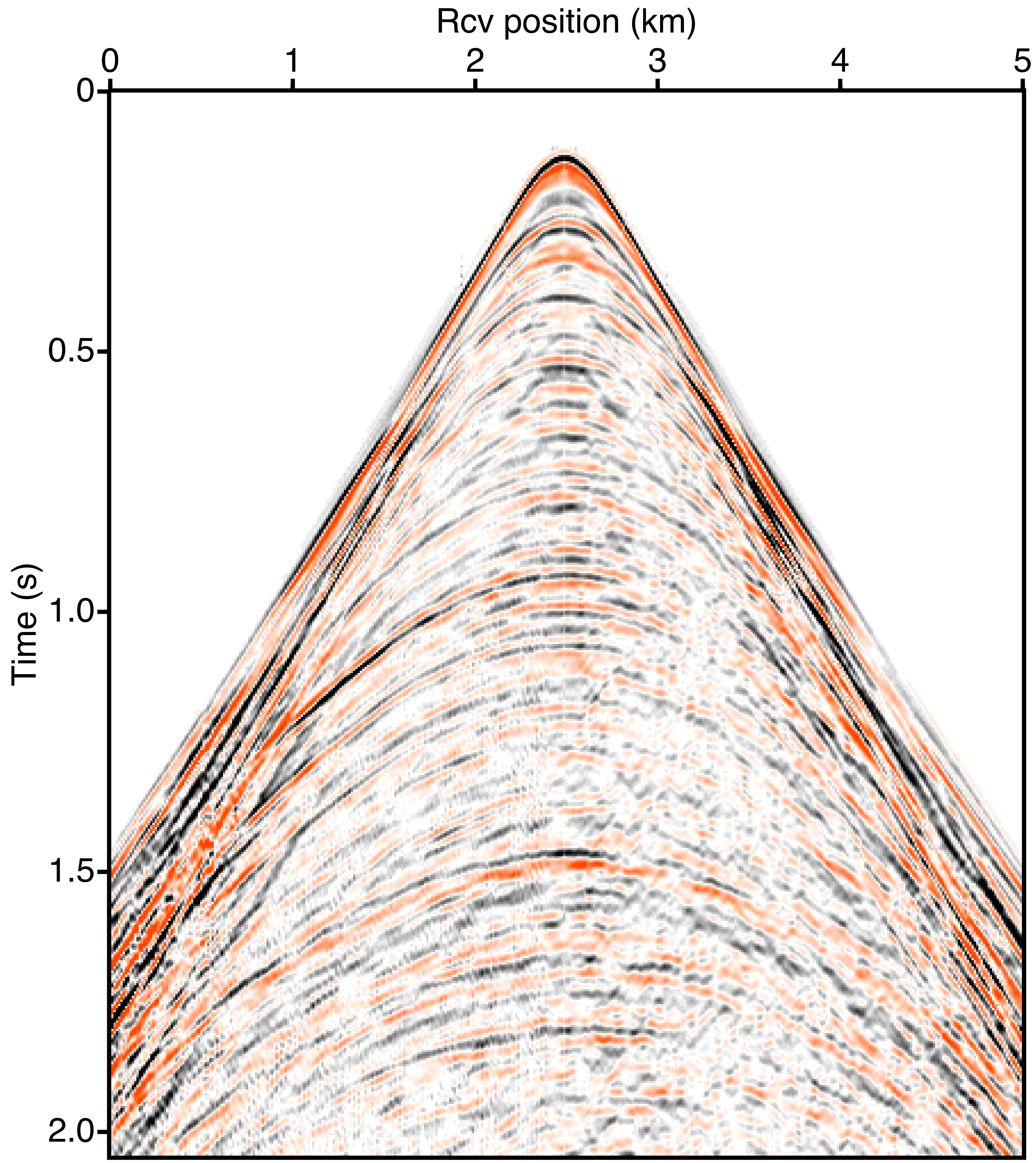
Solution wavefield comparison



Direct Primary

Solved with plain algorithm from finest scale data

Solution wavefield comparison

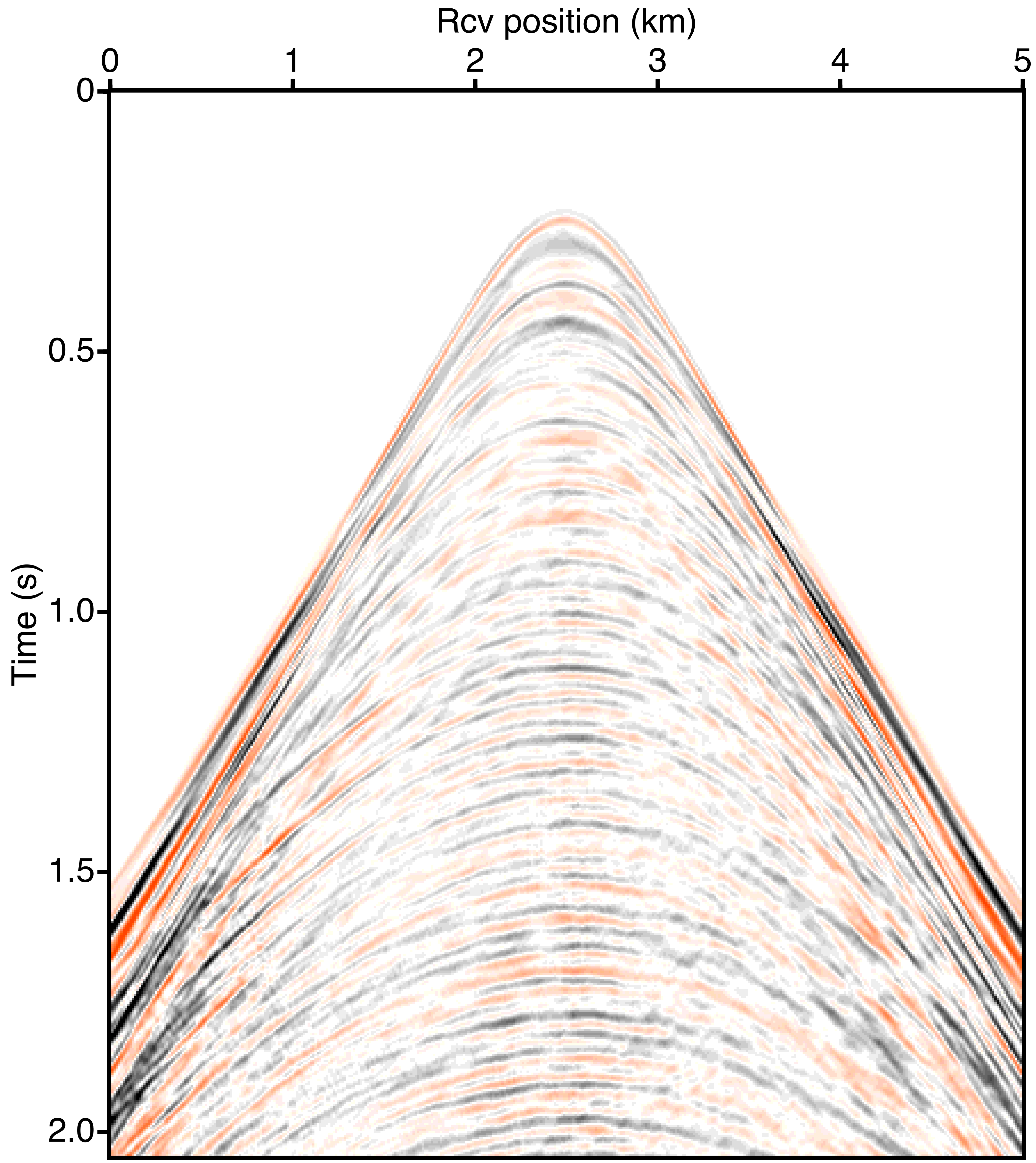


Direct Primary

Solved with spatial sampling continuation

dx = 50m > 25m > 12.5m

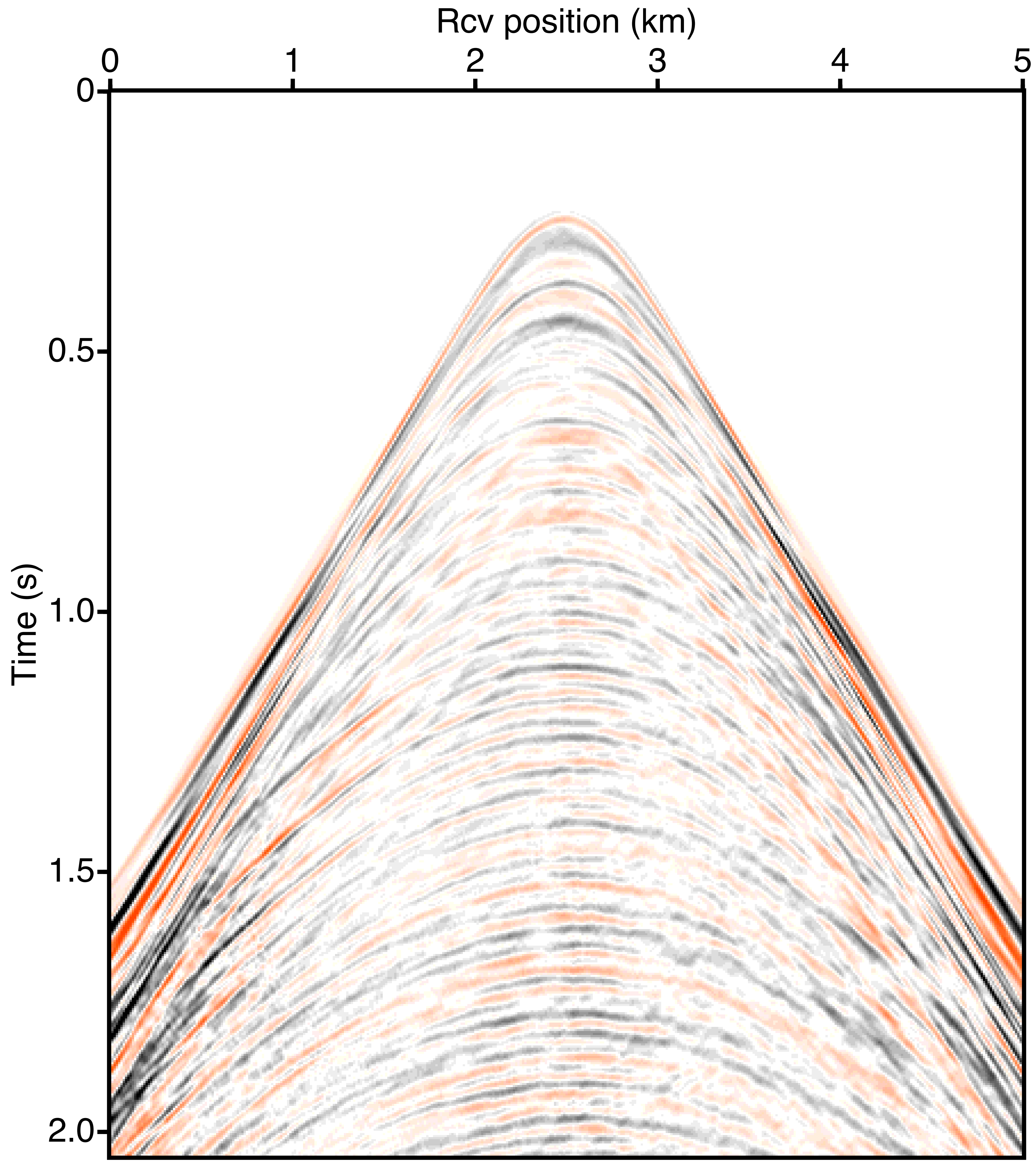
Solution multiple comparison



Predicted Surface Multiple

Solved with plain algorithm from finest scale data

Solution multiple comparison



Predicted Surface Multiple

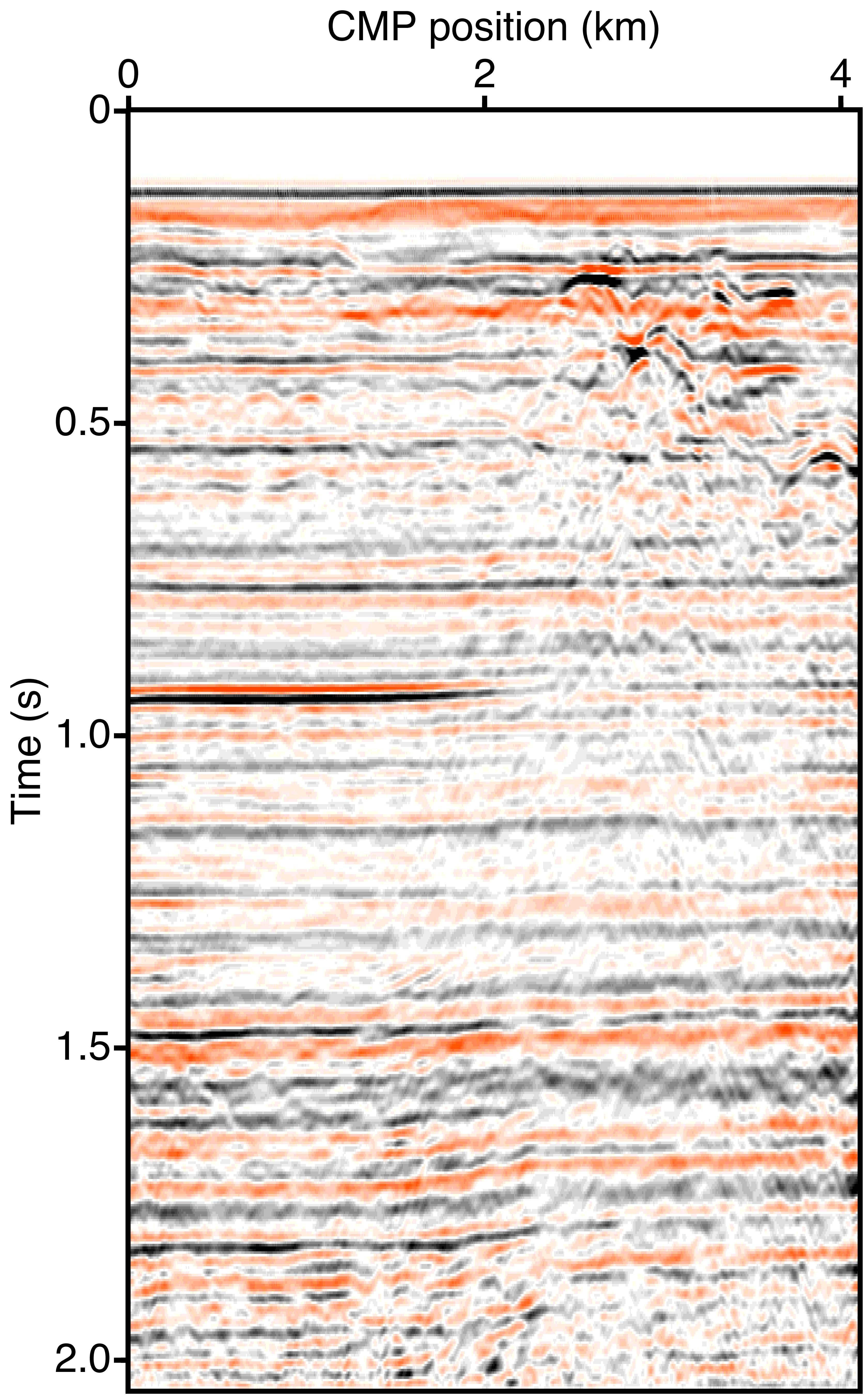
Solved with spatial sampling continuation

dx = 50m > 25m > 12.5m

Solution stack comparison

REPSI Primaries NMO Stack

Solved with plain algorithm from finest scale data

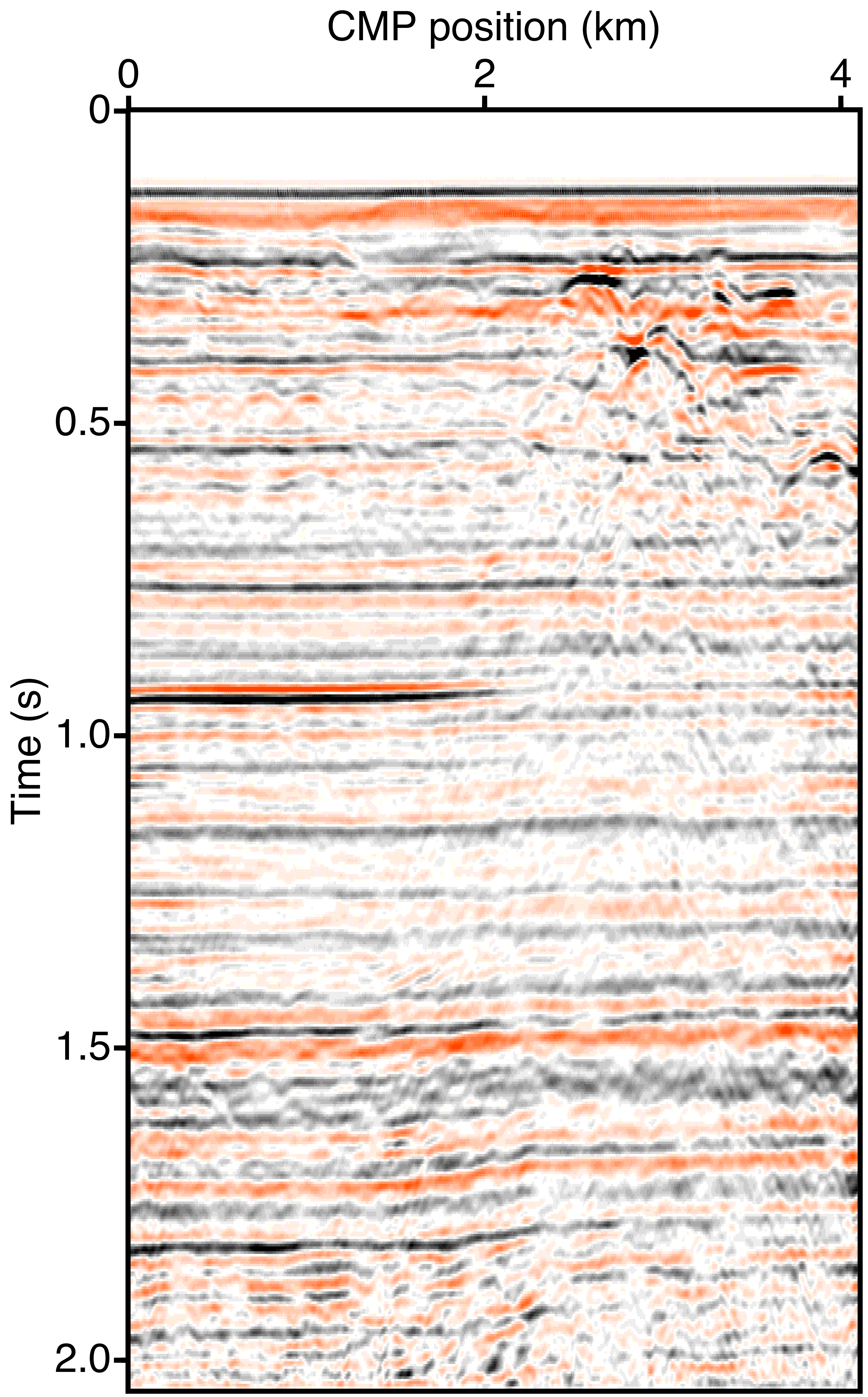


Solution stack comparison

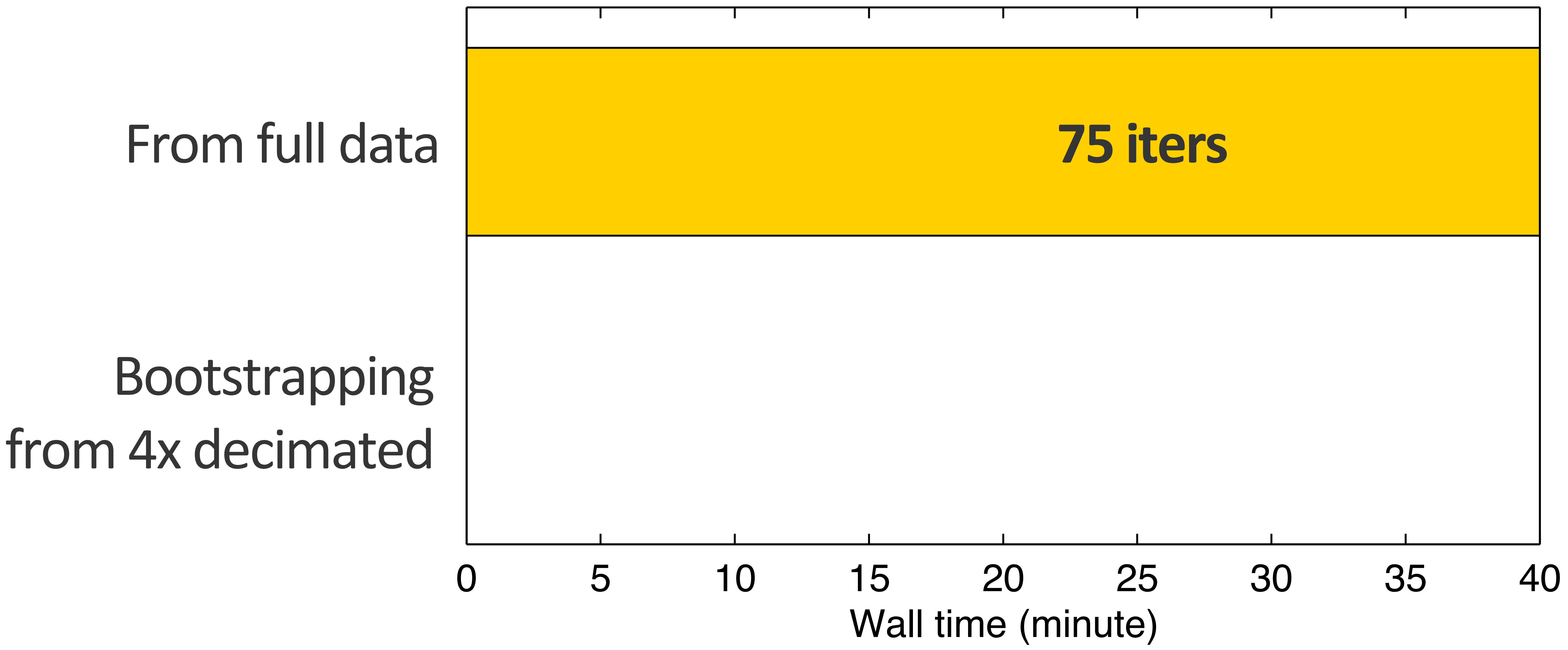
REPSI Primaries NMO Stack

Solved with spatial sampling continuation

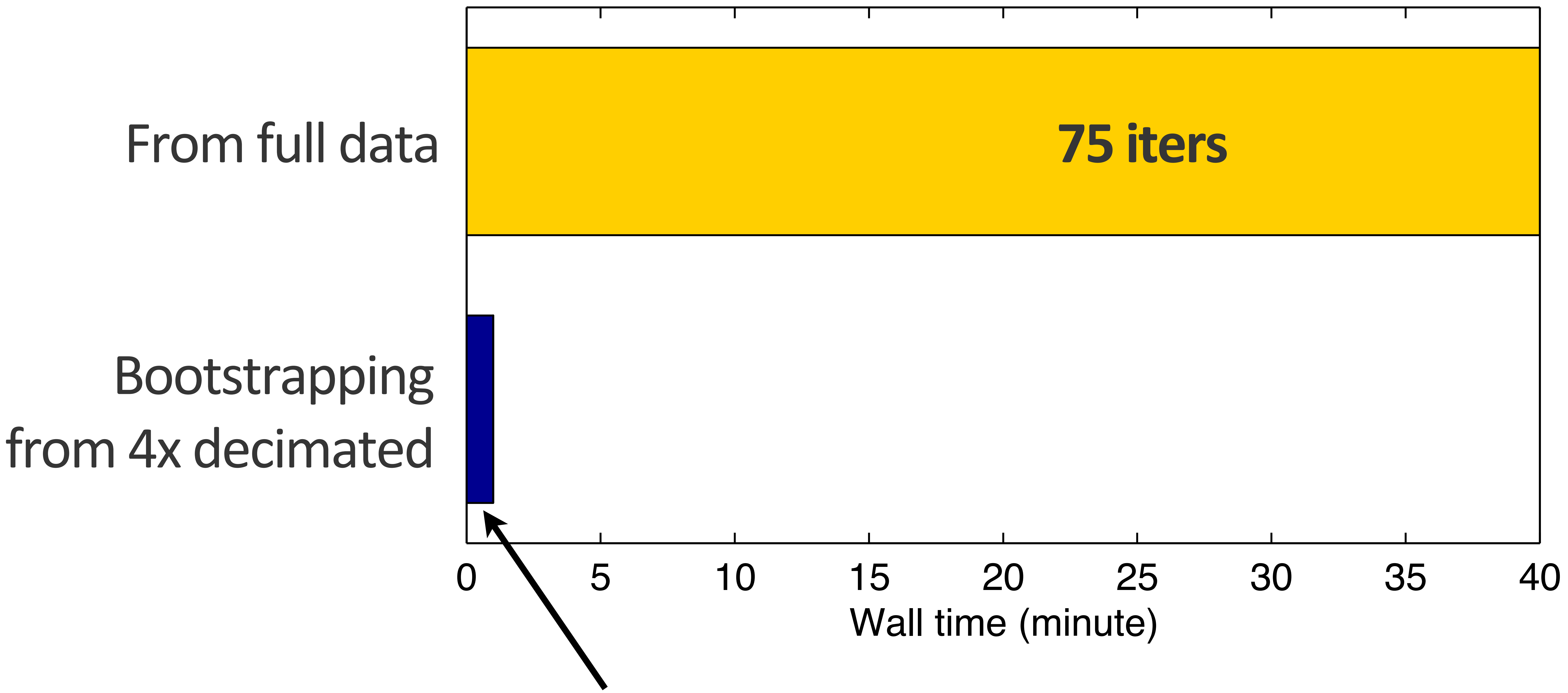
dx = 50m > 25m > 12.5m



Significant speedup from bootstrapping
Wall times

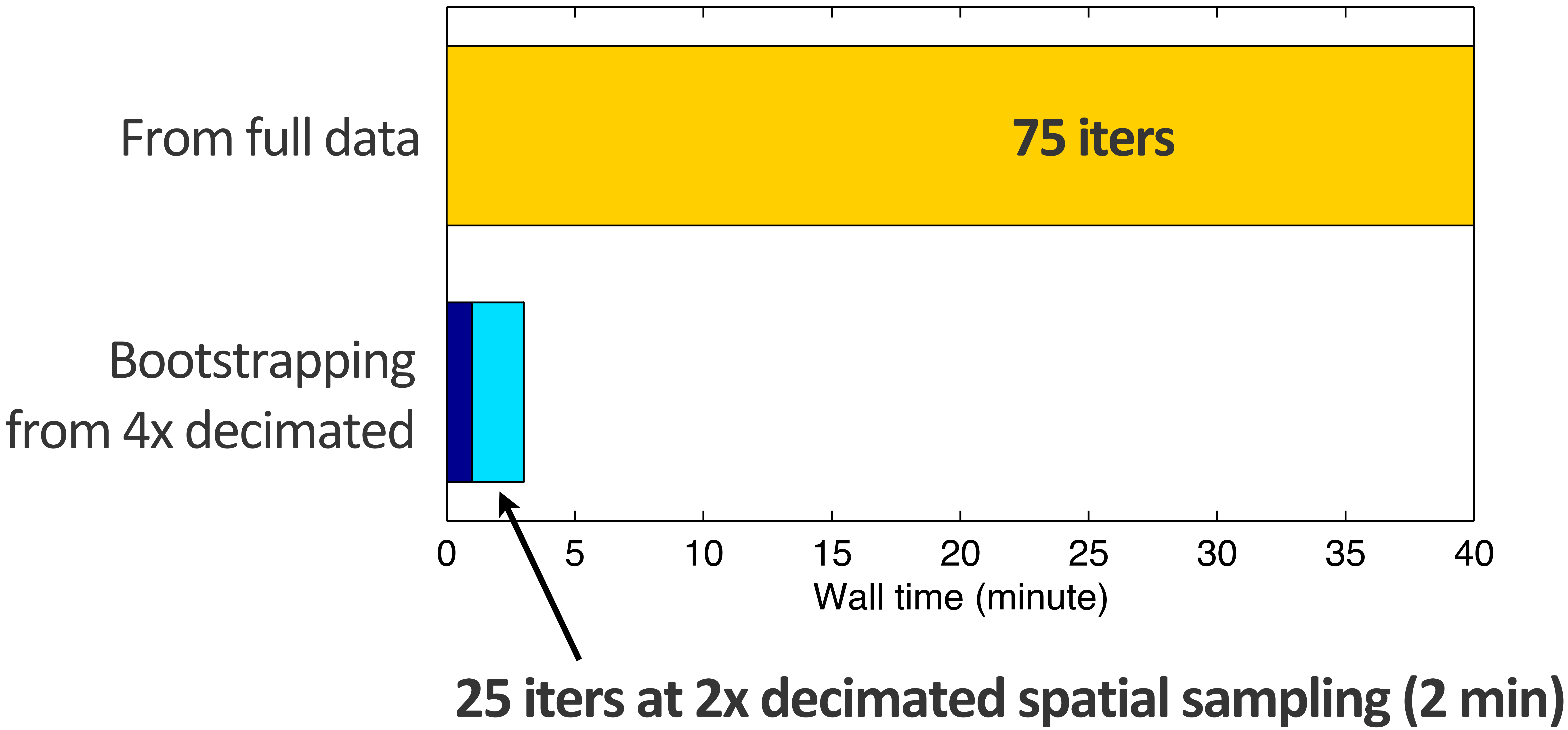


Significant speedup from bootstrapping
Wall times

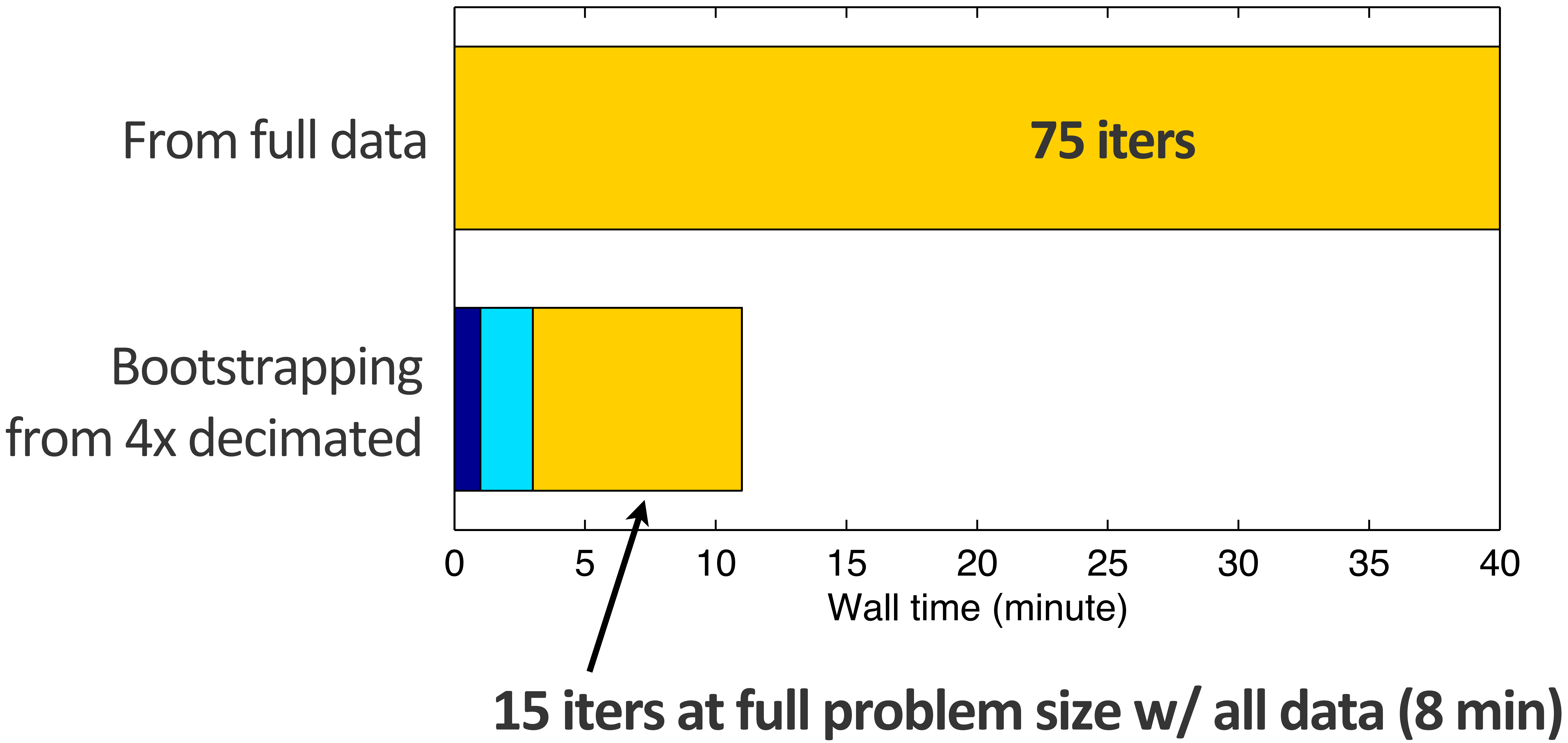


60 iters at 4x decimated spatial sampling (1 min)

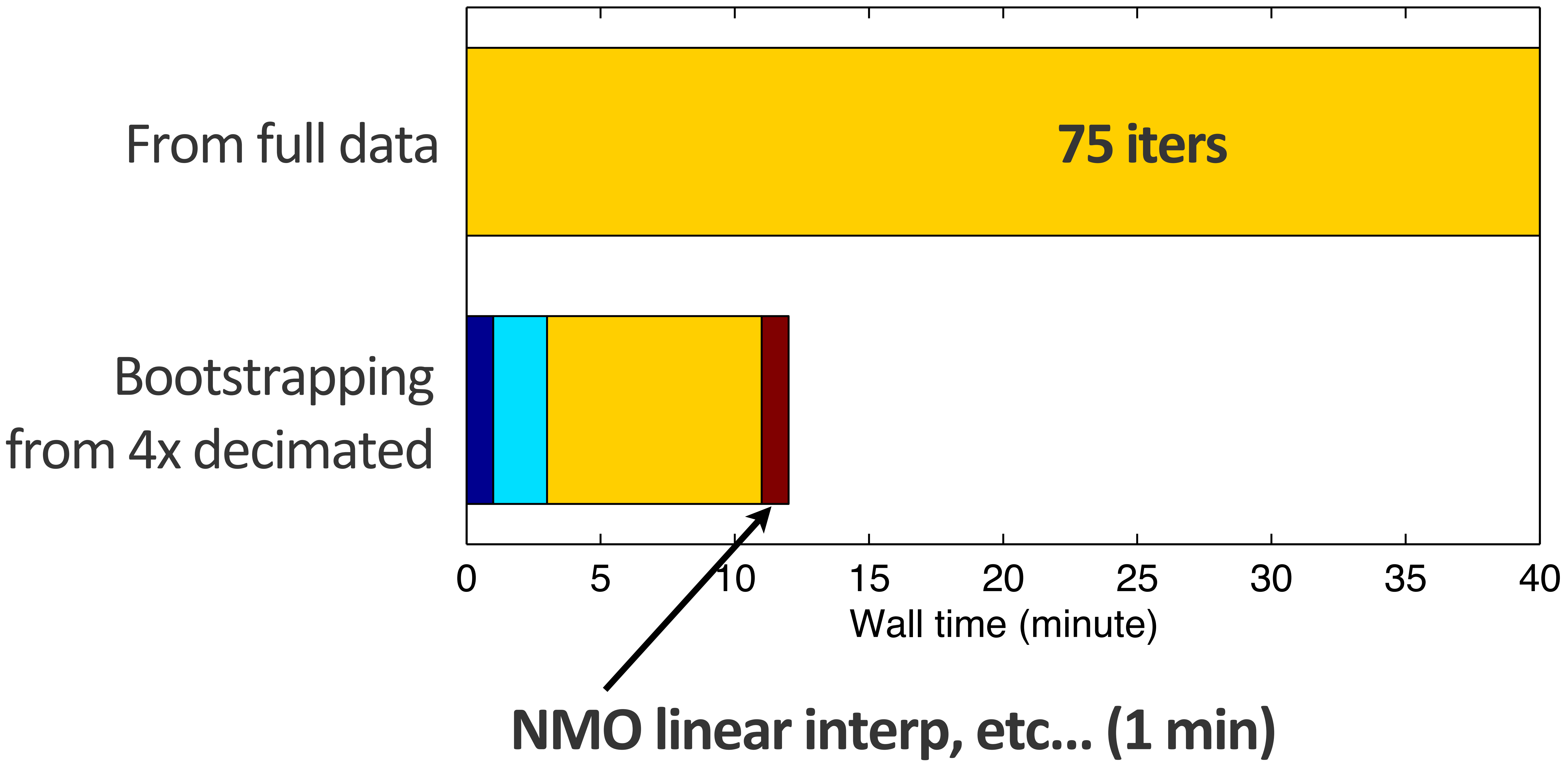
Significant speedup from bootstrapping
Wall times



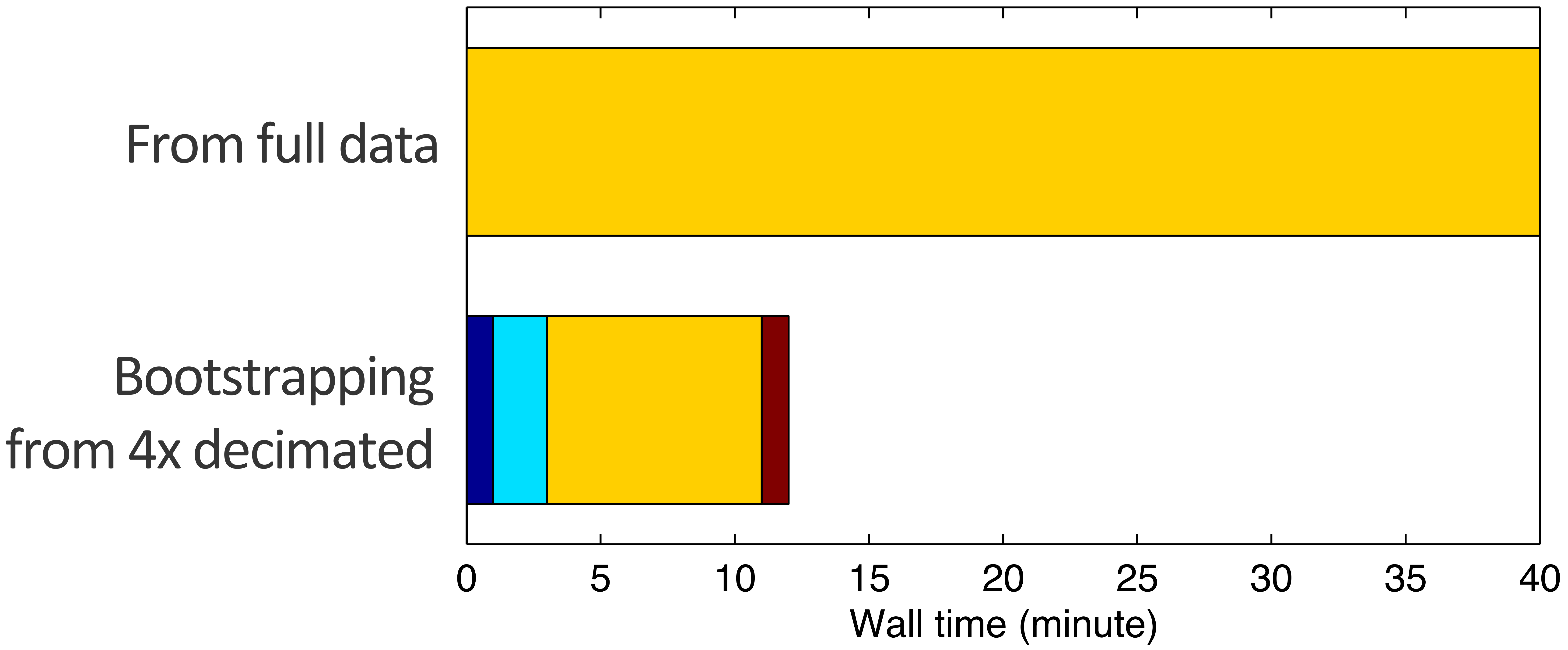
Significant speedup from bootstrapping
Wall times



Significant speedup from bootstrapping
Wall times

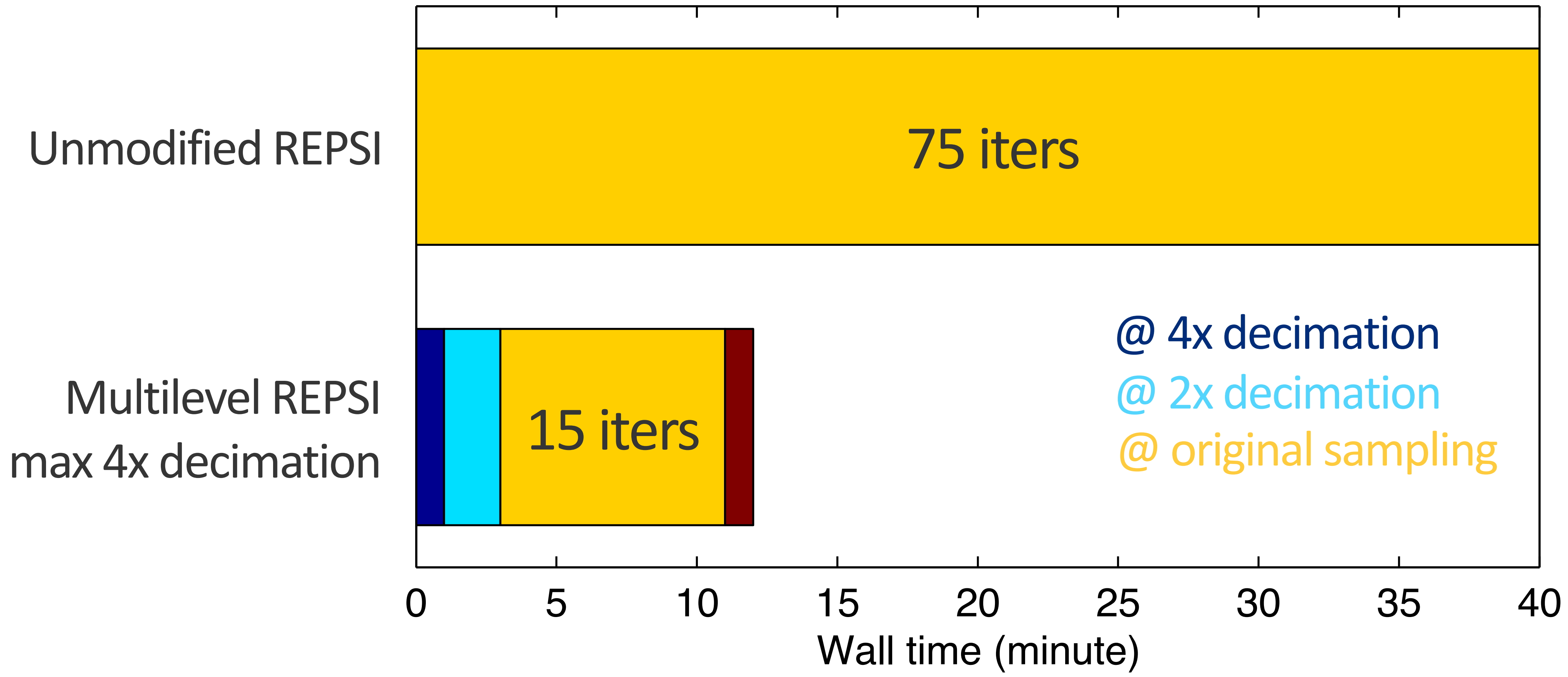


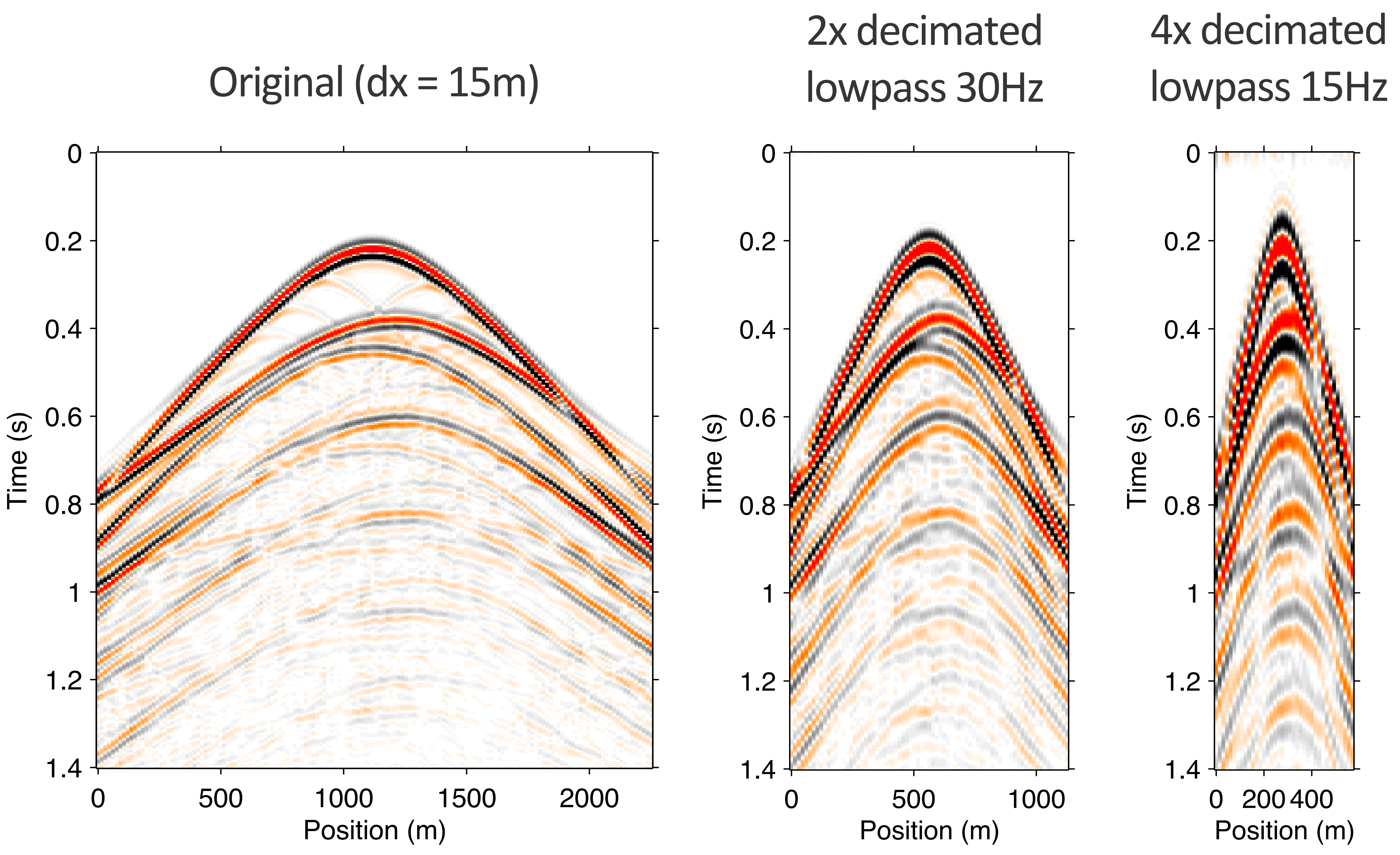
Significant speedup from bootstrapping
Wall times



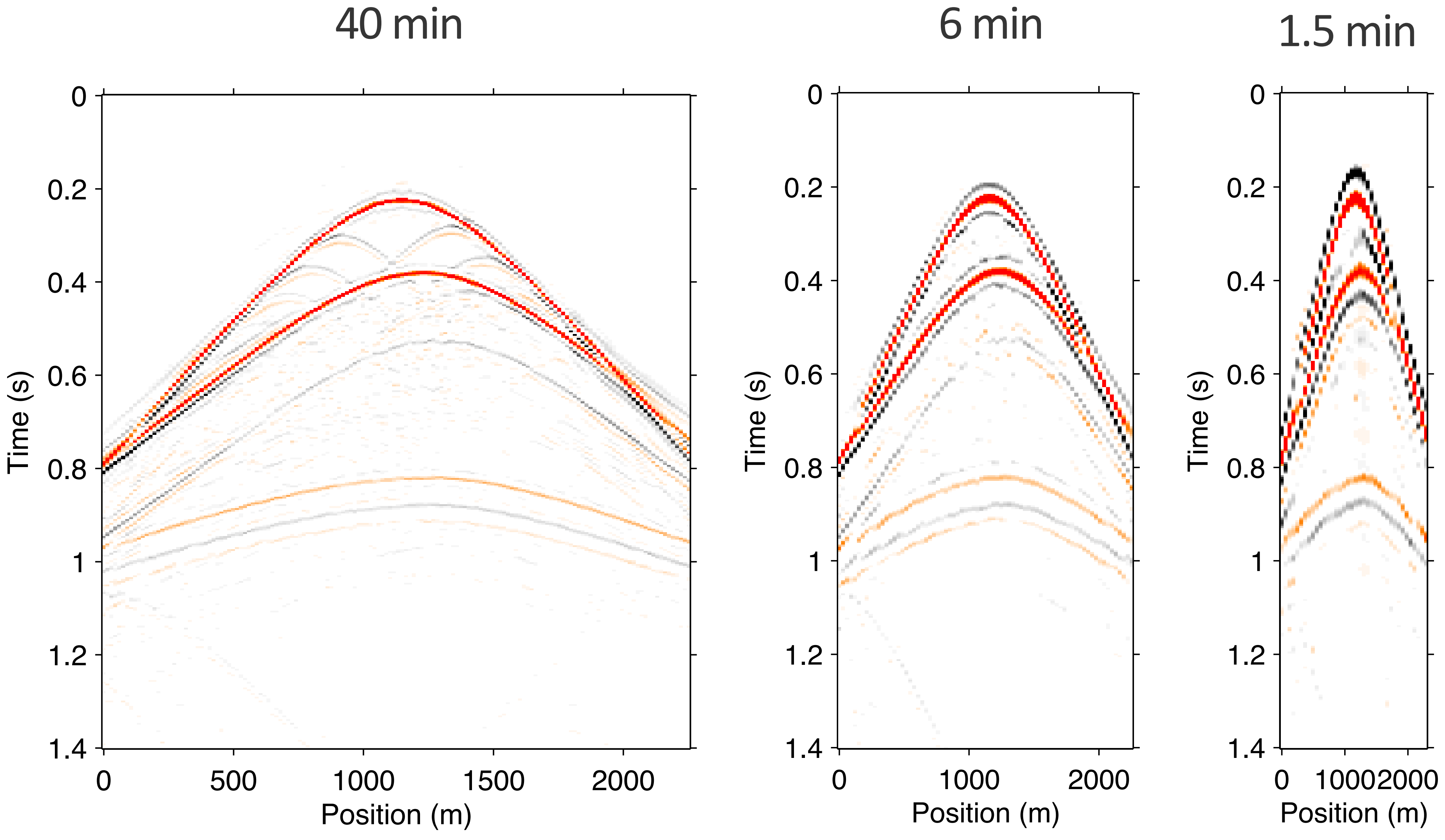
Summary: Multigrid acceleration strategy

A multigrid iterative bootstrapping strategy for REPSI, can *greatly reduce* REPSI computation time (1 order of magnitude in 2D, 2 order in 3D)





Data **mapped to coarse grids**



Solution **solved at coarse grids**

Main contributions

1. **Robust EPSI**, a new formulation of the EPSI primary estimation problem which avoids *ad-hoc parameter adjustments* in favour of *self-tuning bi-convex optimization*

2. **Near-offset mitigation with scattering**, a correction to the multiple prediction in face of missing data, using Born scattering off the free surface.

3. **A multigrid acceleration strategy** for REPSI, can *greatly reduce* REPSI computation time (1 order of magnitude in 2D, 2 order in 3D)

Main Conclusions

1. The EPSI problem can be effectively described as a dual-variable, bi-convex optimization problem, solvable using a Pareto root-finding approach that transforms it into a series of Lasso problems

2. The EPSI forward modeling operator can be augmented with a truncated scattering series to mitigate errors introduced by gaps in the data, with the subsequent non-linearities due to the scattering terms dealt with using either Gauss-Newton or Gauss-Sidel approaches with preconditioning.

3. A simple multigrid-inspired continuation strategy can significantly decrease the computational cost needed for Robust EPSI.

Other contributions

- An analysis formulation of Curvelet-based seismic data interpolation using cosparsity theory (Greedy Analysis Pursuit)
- A fast, scalable method for L1-norm projection on distributed array (part of SPGL1-SLIM)
- **pSPOT**, a simple, powerful, and composable abstraction for separable linear operations on distributed multidimensional numeric arrays *(with Thomas Lai)*
- Shot separation of simultaneous-source data using basis pursuit in Curvelet frame *(with Haneet Wason)*
- A novel L1/L2-norm regularization deconvolution scheme for REPSI *(with Ernie Esser and Rongrong Wang)*

Summary of past publications

Journal Papers

- *Robust estimation of primaries by sparse inversion via one-norm minimization,*
  Tim T. Y. Lin and Felix J. Herrmann, Geophysics 78, R133 (2013), DOI:10.1190/GEO2012-0097.1

- *Estimation of primaries by sparse inversion with scattering-based multiple predictions for data with large gaps,*
  Tim T. Y. Lin and Felix J. Herrmann, submitted for publication to Geophysics

- *Compressed wavefield extrapolation,*
  Tim T. Y. Lin and Felix J. Herrmann, Geophysics 72, SM77 (2007), DOI:10.1190/1.2750716

(with other authors)

- *Source estimation with multiples — fast ambiguity-resolved seismic imaging,*
  Ning Tu, Aleksandr Y. Aravkin, Tristan van Leeuwen, Tim Lin, and Felix J. Herrmann, submitted for publication

- *Simultaneous-source marine acquisition with compressive sampling matrices,*
  Hassan Mansour, Haneet Wason, Tim T. Y. Lin, Felix J. Herrmann, Geophysical Prospecting 60, 648-662 (2012), DOI:10.1111/j.1365-2478.2012.01075.x

- *Compressive simultaneous full-waveform simulation,*
  Felix J. Herrmann, Yogi A. Erlangga, and Tim T. Y. Lin, Geophysics 74, A35 (2009), DOI:10.1190/1.3115122

Summary of past publications

Conference proceedings

- *Mitigating data gaps in the estimation of primaries by sparse inversion without data reconstruction,*
  Tim T. Y. Lin and Felix J. Herrmann, SEG Annual Meeting 2014
- *Multilevel acceleration strategy for the robust estimation of primaries by sparse inversion,*
  Tim T. Y. Lin and Felix J. Herrmann, 76th EAGE Conference & Exhibition, 2014
- *Dense shot-sampling via time-jittered marine sources,*
  Tim T. Y. Lin, Haneet Wason, and Felix J. Herrmann, SEG Annual Meeting Workshop 2013
- *Cosparse seismic data interpolation,*
  Tim T. Y. Lin and Felix J. Herrmann, 75th EAGE Conference & Exhibition, 2013
- *Robust source signature deconvolution and the estimation of primaries by sparse inversion,*
  Tim T. Y. Lin and Felix J. Herrmann, SEG Annual Meeting Workshop 2011
- *Estimating primaries by sparse inversion in a curvelet-like representation domain,*
  Tim T. Y. Lin and Felix J. Herrmann, 73rd EAGE Conference & Exhibition, 2011
- *Sparsity-promoting migration from surface-related multiples,*
  Tim T. Y. Lin, Ning Tu, and Felix J. Herrmann, SEG Annual Meeting 2010
- *Stabilized estimation of primaries via sparse inversion,*
  Tim T. Y. Lin and Felix J. Herrmann, 72nd EAGE Conference & Exhibition, 2010
- *Compressive simultaneous full-waveform simulation,*
  Felix J. Herrmann, Yogi A. Erlangga, and Tim T. Y. Lin, SEG Annual Meeting 2009
- *Designing Simultaneous Acquisitions with Compressive Sensing,*
  Tim T. Y. Lin and Felix J. Herrmann, 71st EAGE Conference & Exhibition, 2009
- *Interpolating solutions of the Helmholtz equation with compressed sensing,*
  Tim T. Y. Lin, Evgeniy Lebed, Yogi A. Erlangga, and Felix J. Herrmann, SEG Annual Meeting 2008
- *Compressed wavefield extrapolation with curvelets,*
  Tim T. Y. Lin and Felix J. Herrmann, SEG Annual Meeting 2007

Thank you!

To my advisor, committee, and everyone at SLIM...
I really appreciated my time here