Amsterdam '14

Multilevel Acceleration Strategy for the Robust Estimation of Primaries by Sparse Inversion

Tim T.Y. Lin and Felix J. Herrmann

Amsterdam, *EAGE 2014*

SLIM
University of British Columbia

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (van Groenestijn and Verschuur, 2009)

true primary wavefield SRME-produced primary

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

$\mathbf{P}$	total up-going wavefield
$\mathbf{P_o}$	primary wavefield
$A(f)$	"matching" operator

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (van Groenestijn and Verschuur, 2009)

true primary wavefield — SRME-produced primary

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

SRMP

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

$\mathbf{P_o}$ primary wavefield

$A(f)$ "matching" operator

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (van Groenestijn and Verschuur, 2009)

adaptive subtraction

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

SRMP

$\mathbf{P}$	total up-going wavefield
$\mathbf{P}_\mathbf{o}$	primary wavefield
$A(f)$	"matching" operator

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (van Groenestijn and Verschuur, 2009)

true primary wavefield   SRME-produced primary

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

$\mathbf{P}$	total up-going wavefield
$\mathbf{P_o}$	primary wavefield
$A(f)$	“matching” operator

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (van Groenestijn and Verschuur, 2009)

recorded data predicted data from SRME

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

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

$\mathbf{P_o}$ primary wavefield

$A(f)$ "matching" operator

SLIM

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (van Groenestijn and Verschuur, 2009)

recorded data — predicted data from SRME

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

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

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

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (van Groenestijn and Verschuur, 2009)

recorded data — predicted data from SRME

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

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

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

$\mathbf{G}$ primary impulse response

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (van Groenestijn and Verschuur, 2009)

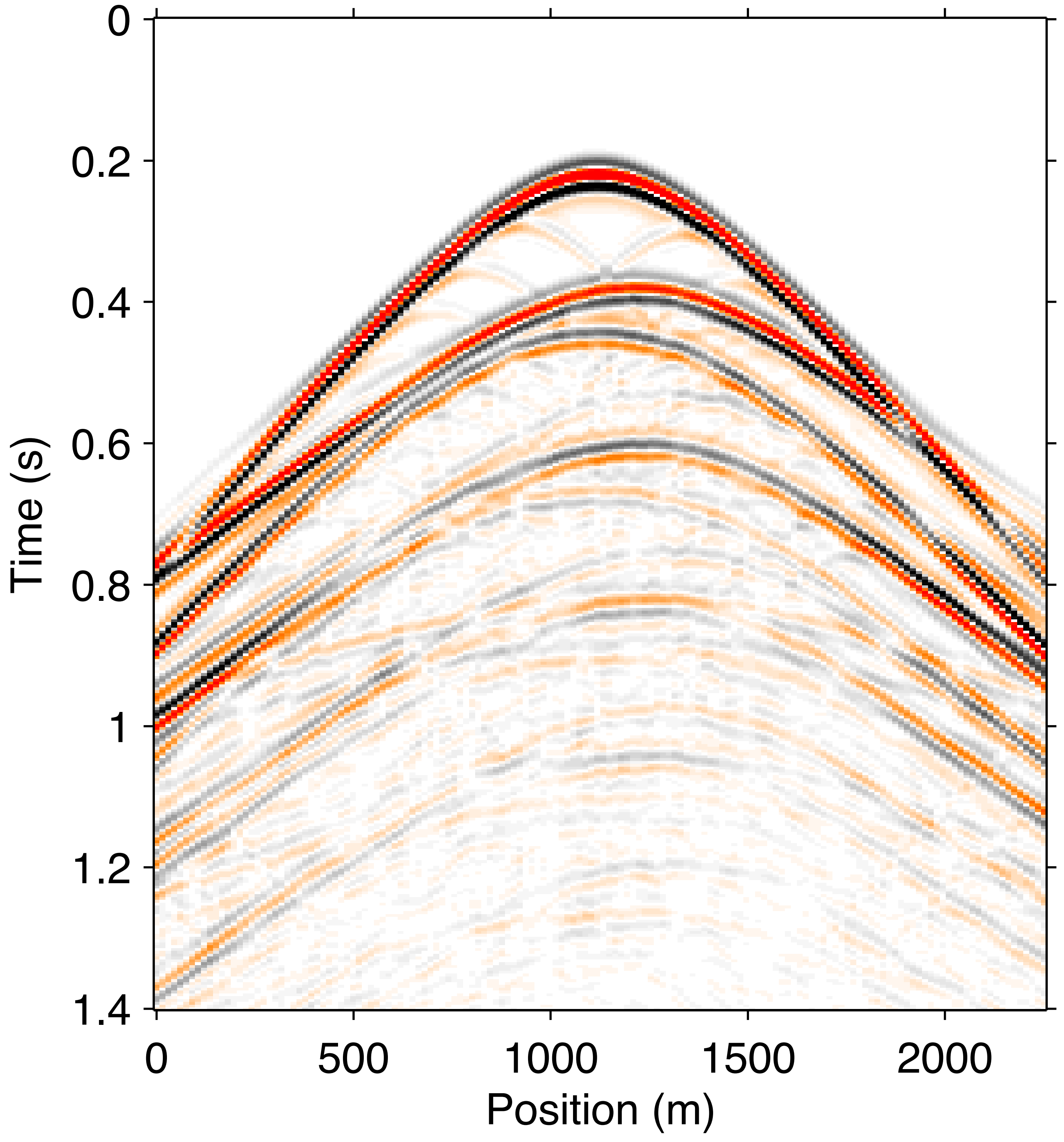
recorded data   predicted data from SRME

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

Inversion objective:

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

From SRME to Robust EPSI



Primaries by Sparse Inversion (van Groenestijn and Verschuur, 2009)

ʳded data predicted data from SRME

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

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

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (va

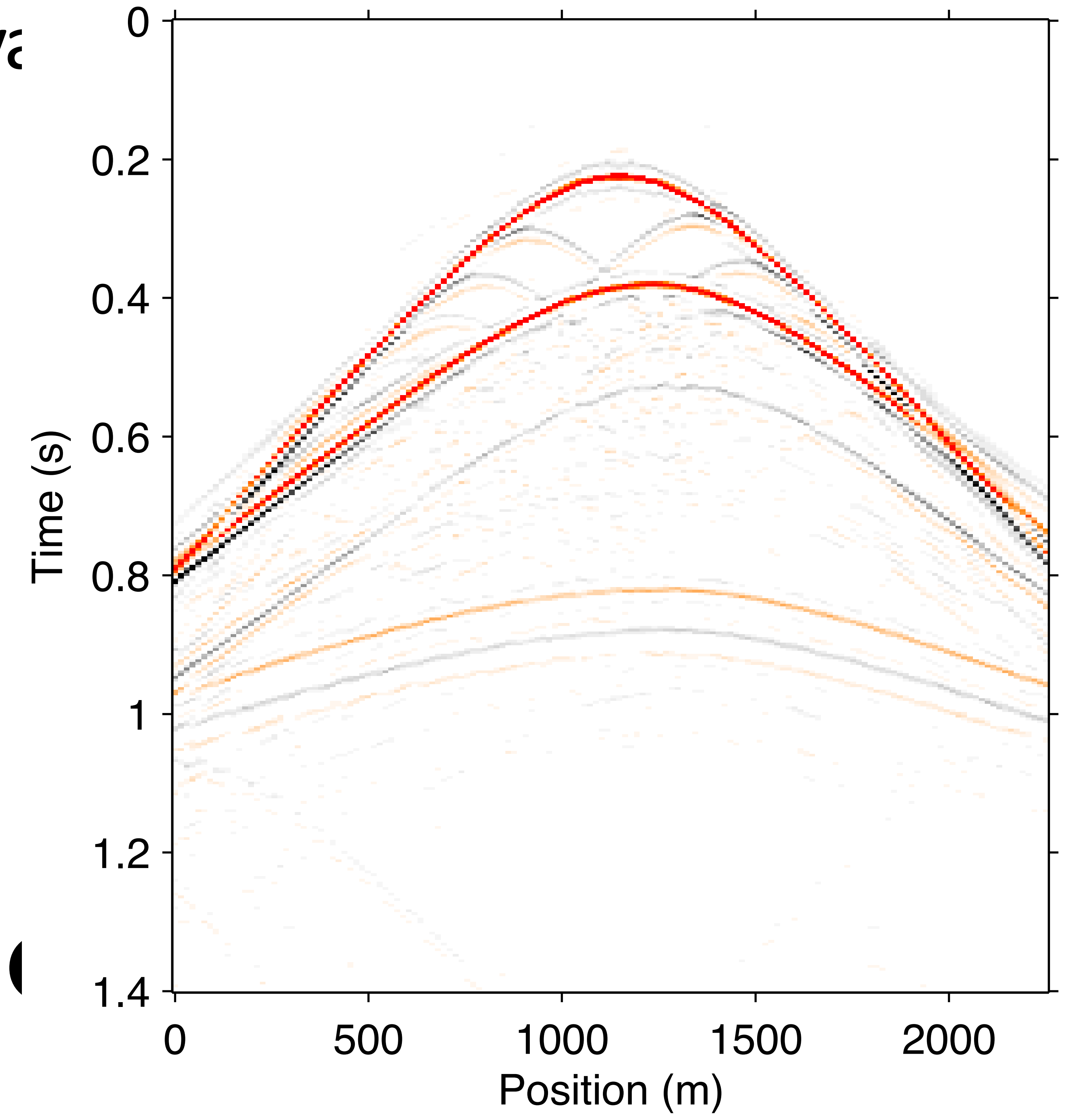
9)

recorded data predicted data from SRME

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

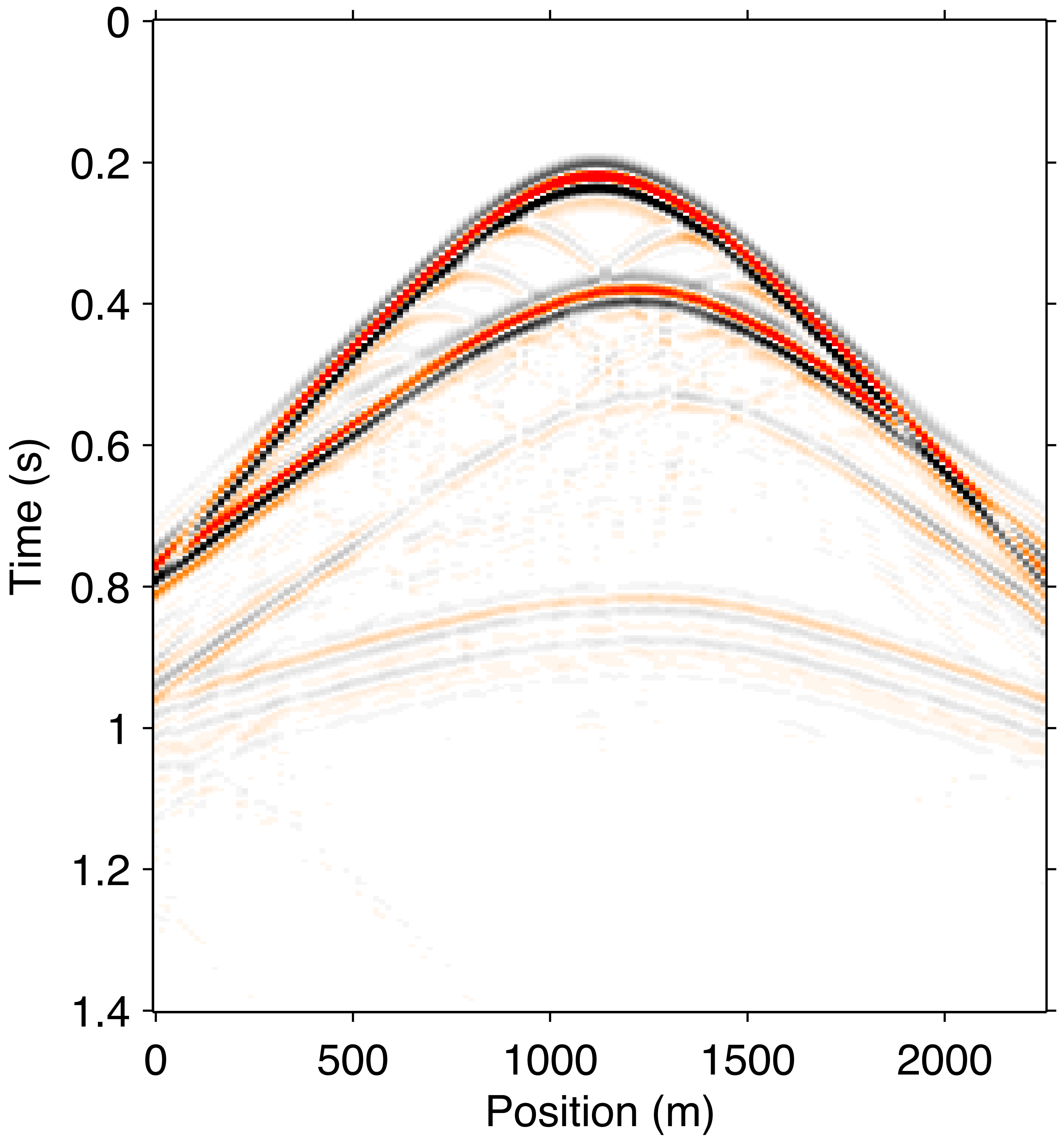
Inversion objective:

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



From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (va



9)

recorded data

predicted data from SRME

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

Inversion objective:

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

From SRME to Robust EPSI

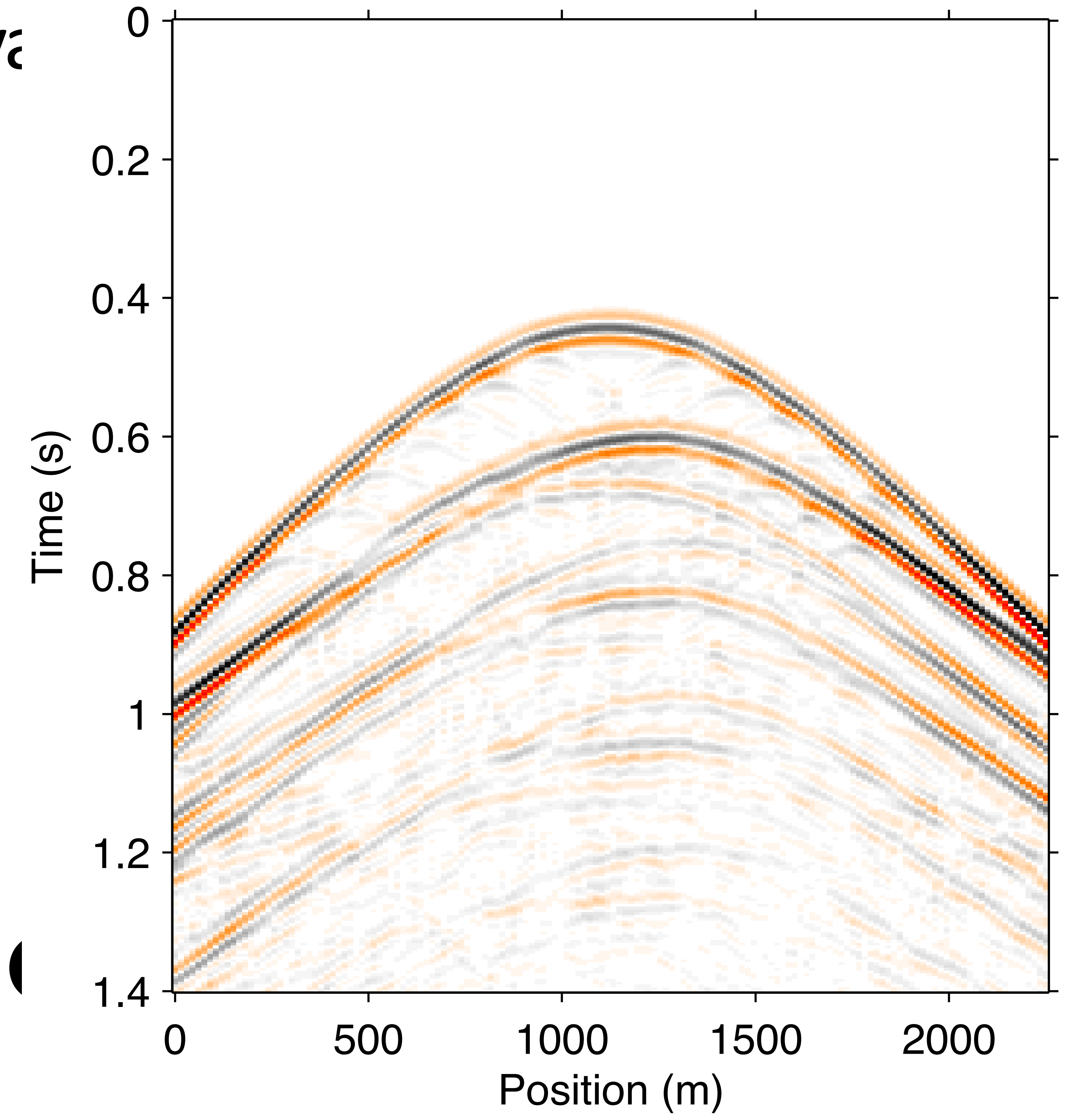
Based on **Estimation of Primaries by Sparse Inversion** (va ...9)

recorded data predicted data from SRME

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

Inversion objective:

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



From SRME to Robust EPSI

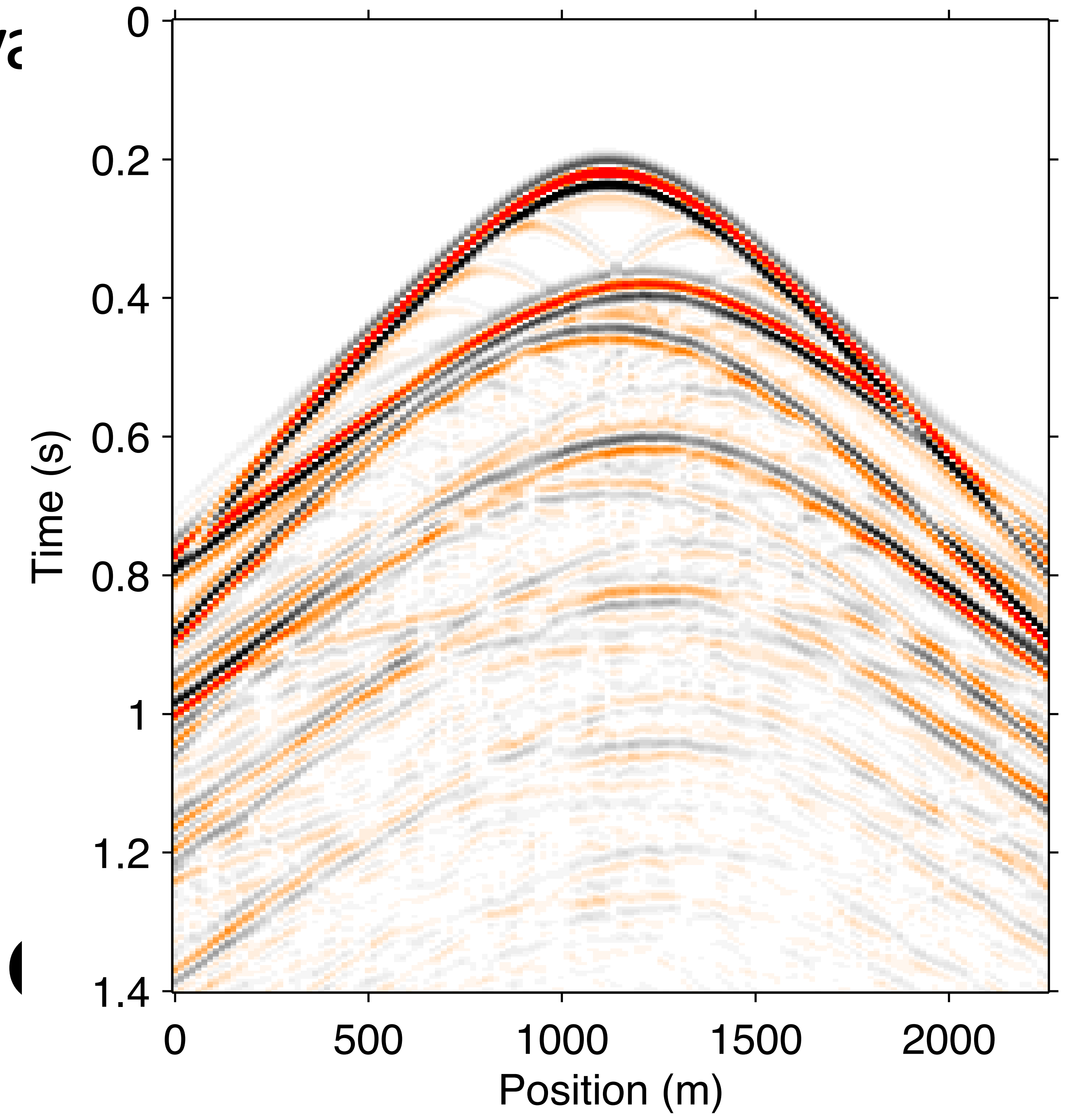
Based on **Estimation of Primaries by Sparse Inversion** (va 9)

recorded data predicted data from SRME

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

Inversion objective:

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



From SRME to Robust EPSI

Two ways to obtain the final primary wavefield

"Direct" Primary "Conservative" Primary

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

Inversion objective:

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

SLIM

From SRME to Robust EPSI

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

recorded data predicted data from SRME

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

$$\mathcal{M}(\mathbf{g}, \mathbf{q}) := \mathcal{F}_{\mathrm{t}}^{\dagger} \mathrm{BlockDiag}_{\omega_1 \cdots \omega_{nf}} [(q(\omega)\mathbf{I} - \mathbf{P})^{\dagger} \otimes \mathbf{I}] \mathcal{F}_{\mathrm{t}} \mathbf{g}$$

Inversion objective:

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

From SRME to Robust EPSI

Based on **Estimation of Primaries by Sparse Inversion** (van Groenestijn and Verschuur, 2009)

recorded data predicted data from SRME

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

Inversion objective:

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

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

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

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

Solving the EPSI problem

Linearizations

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

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

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

In fact it is bilinear:

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

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

[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

Emits sparse, or "deconvolved" solution

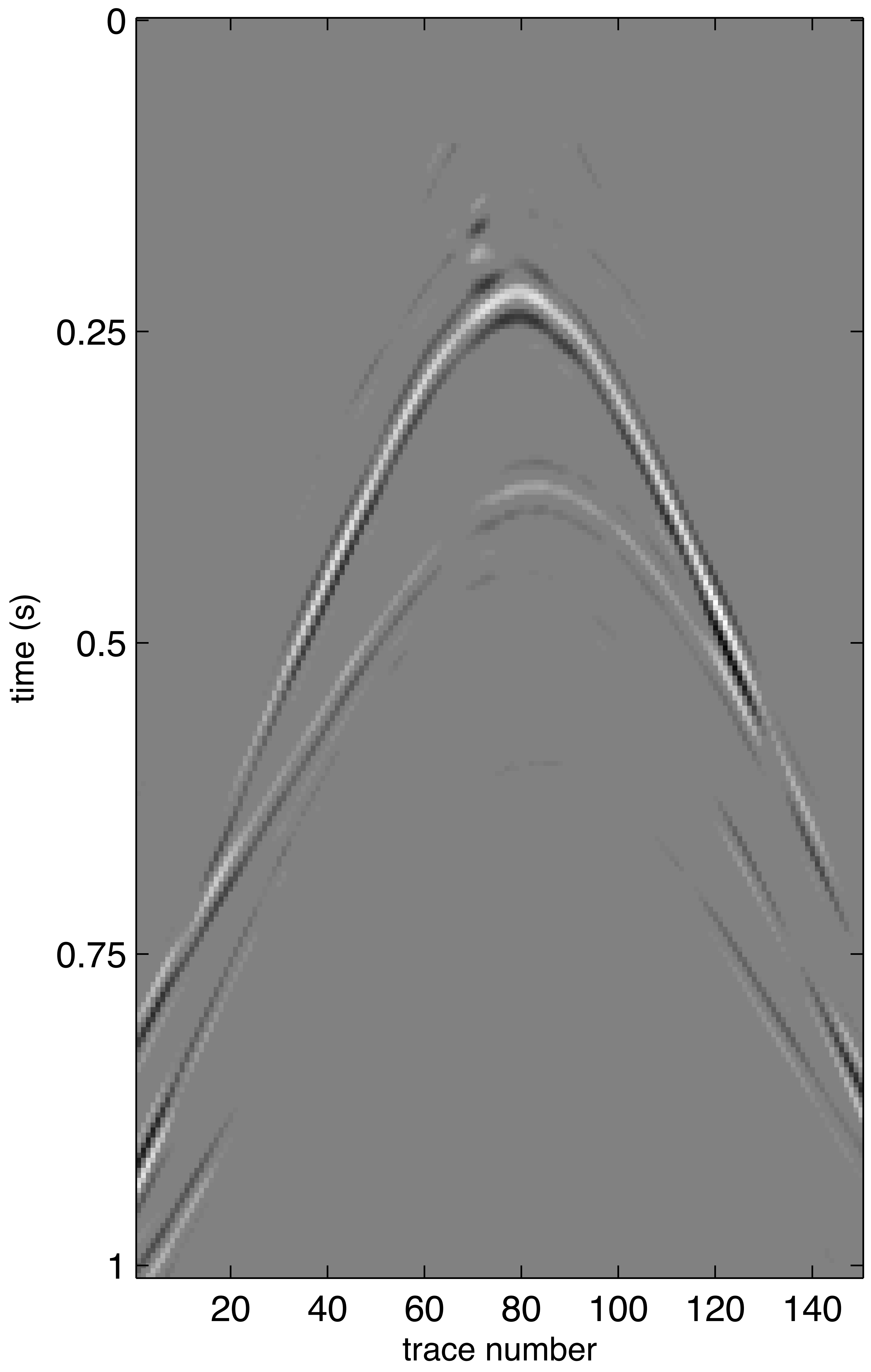
$$\mathbf{g}_{k+1} = \arg\min_{\mathbf{g}} \|\mathbf{p} - \mathbf{M}_{q_k}\mathbf{g}\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \leq \tau_k$$

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

L1 projection and sparsity

variable g at beginning of LASSO

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

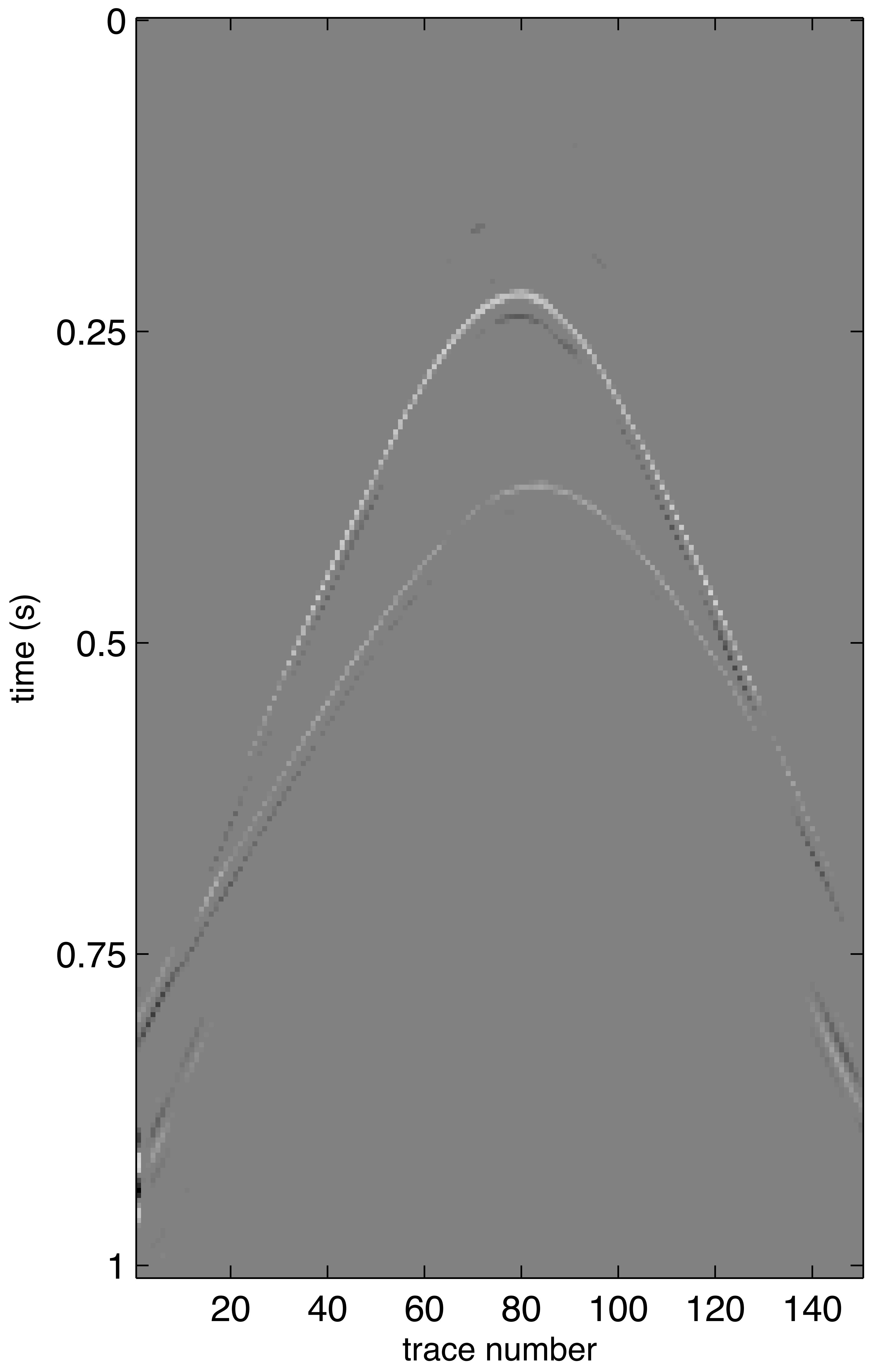


L1 projection and sparsity

variable g at end of LASSO

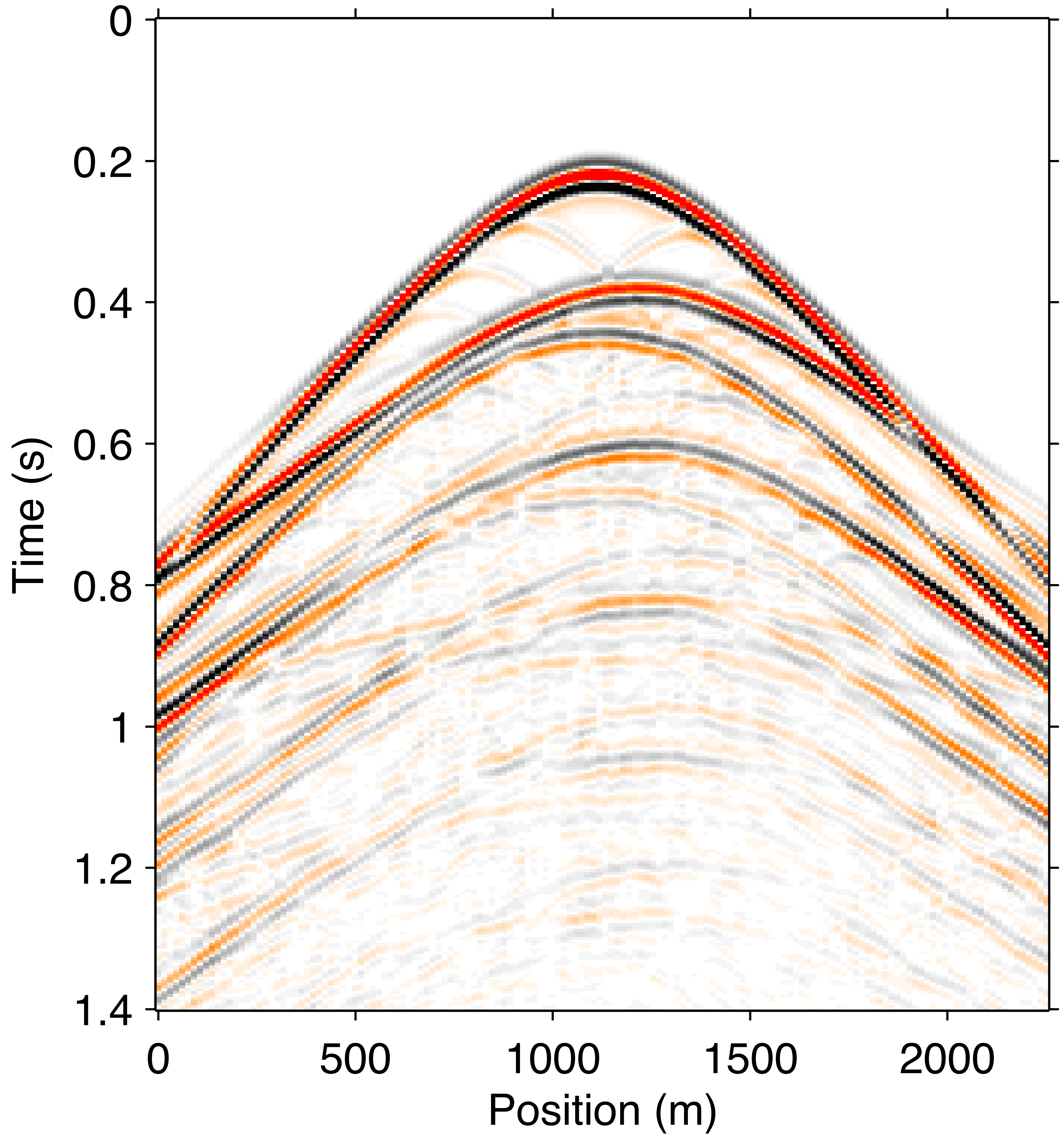
$$\mathbf{g}_{k+1} = \arg\min_{\mathbf{g}} \|\mathbf{p} - \mathbf{M}_{q_k}\mathbf{g}\|_2 \text{ s.t. } \|\mathbf{g}\|_1 \leq \tau_k$$

Emits "deconvolved" solution



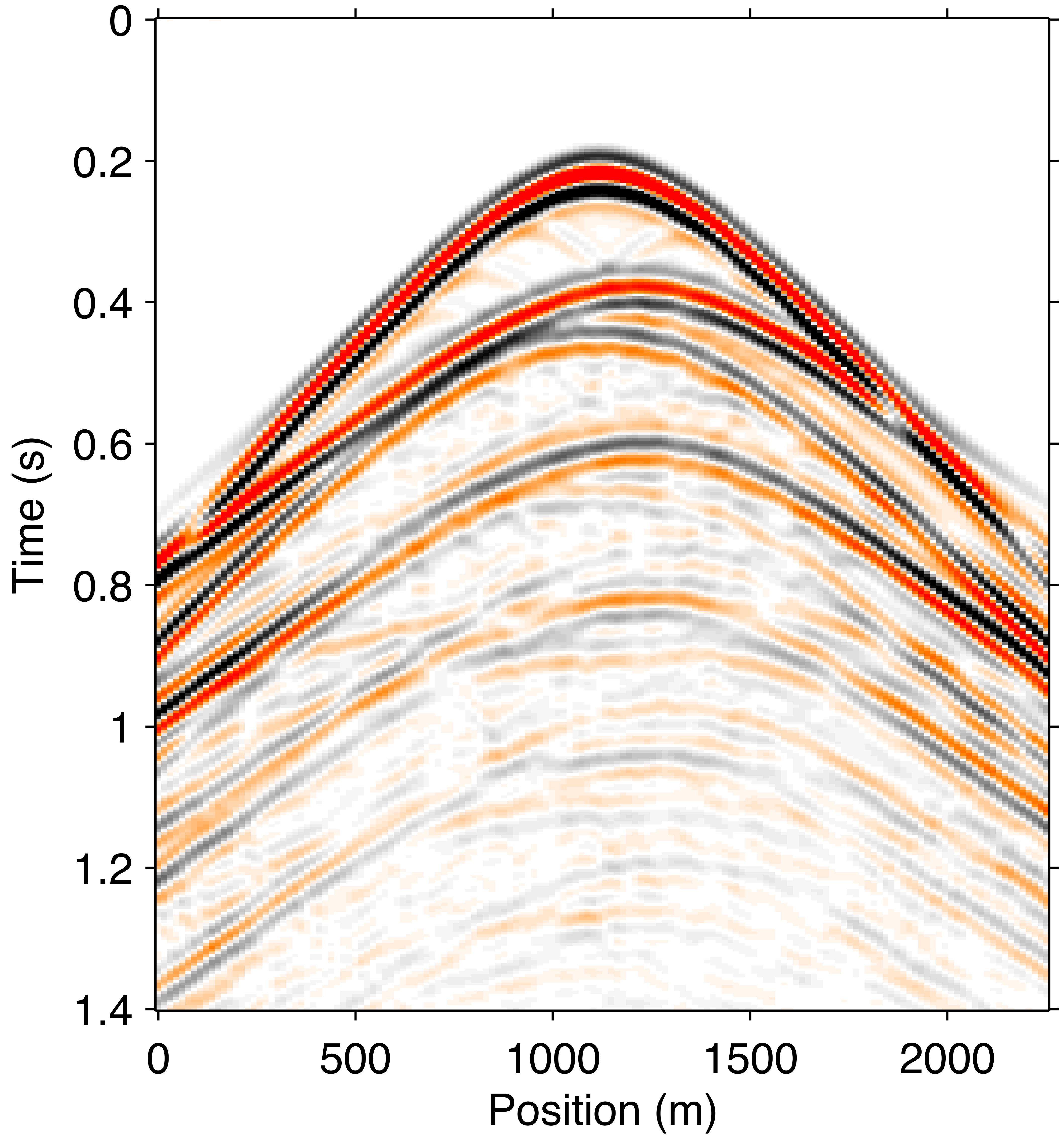
SLIM

Motivation: G tolerates lowpass filtering



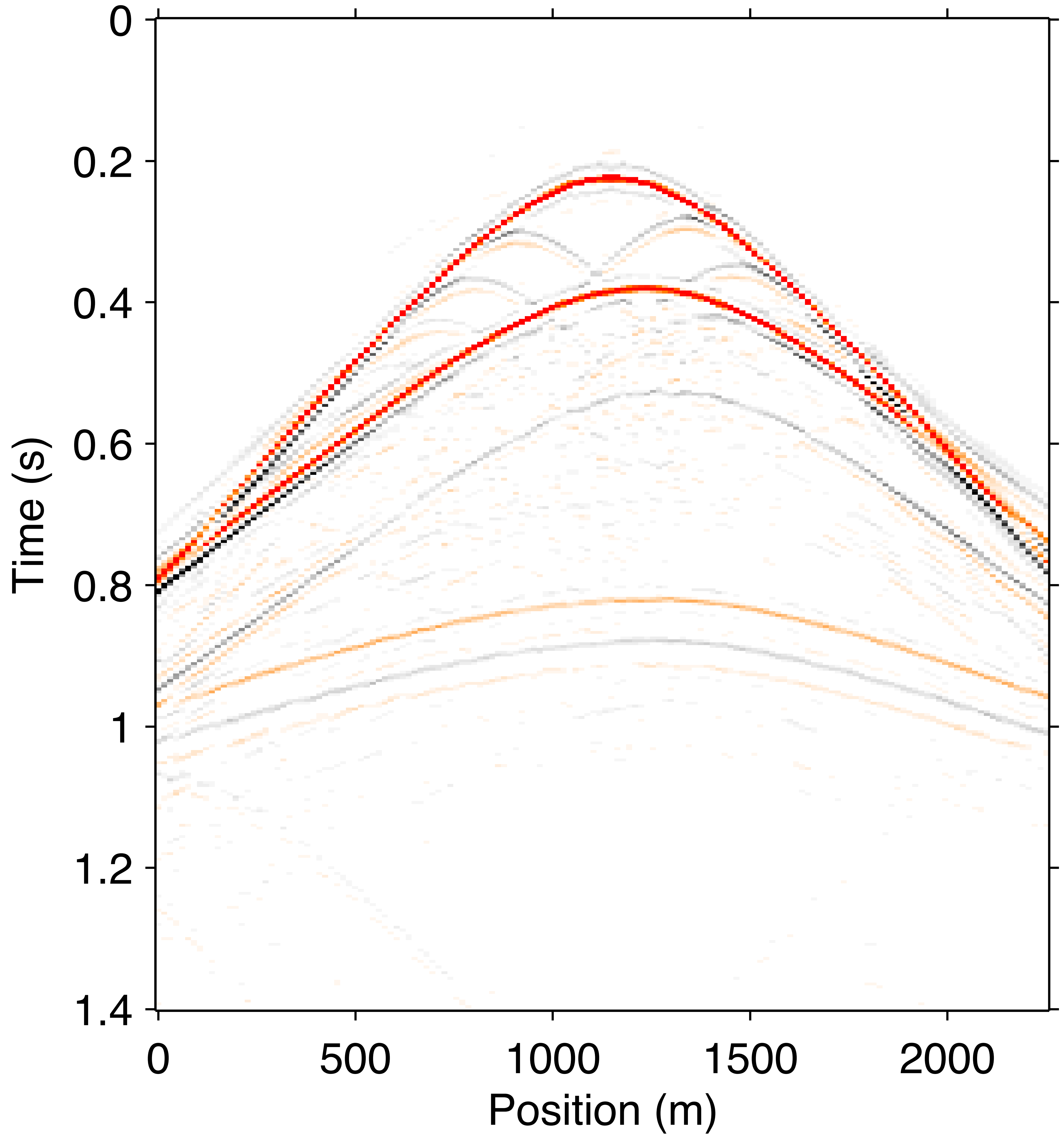
Data
modeled with Ricker 30Hz

Motivation: G tolerates lowpass filtering



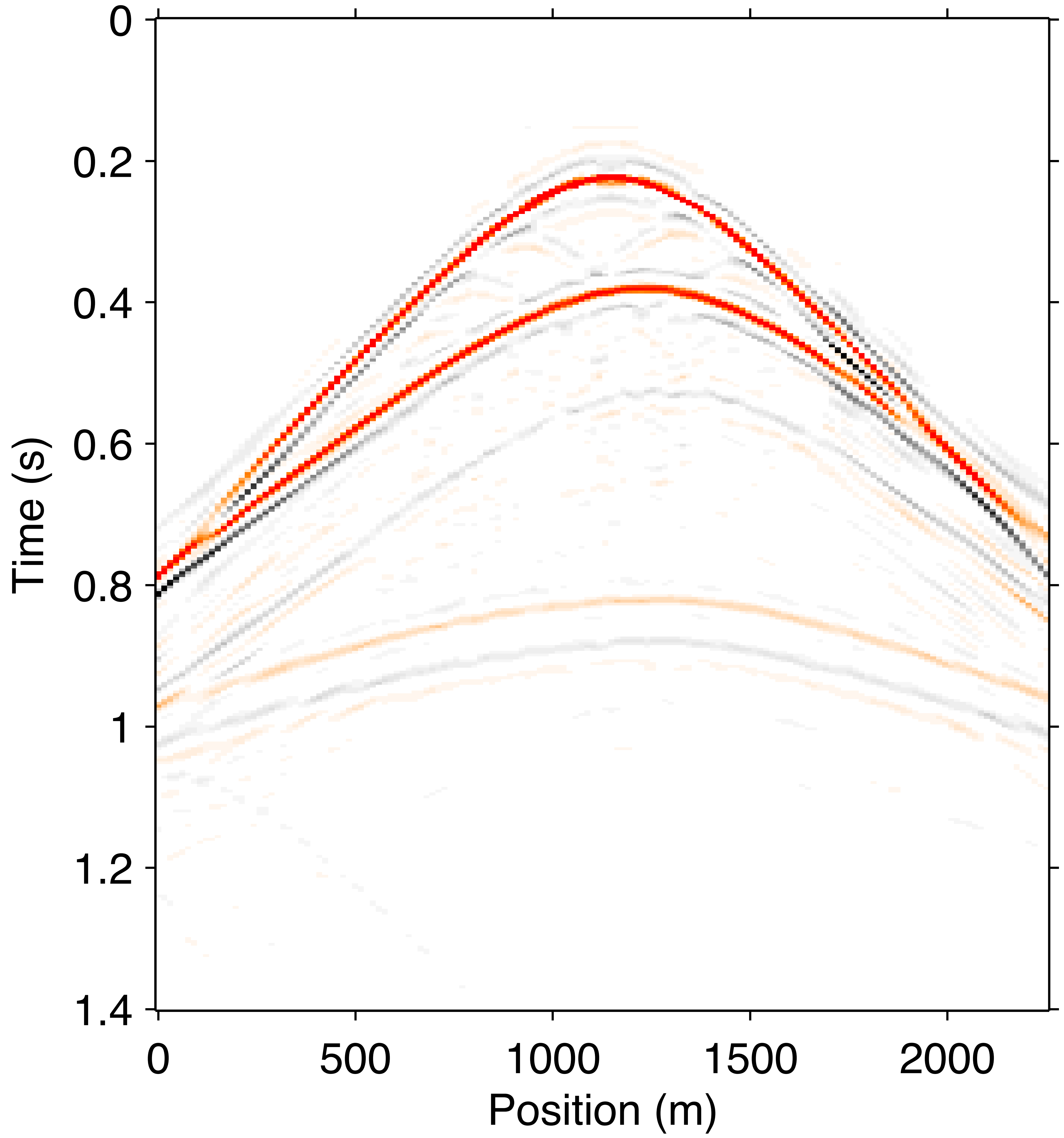
Lowpassed Data
modeled with Ricker 30Hz
lowpass at 40Hz
(25-order, zero-phase, Hann window)

Motivation: G tolerates lowpass filtering



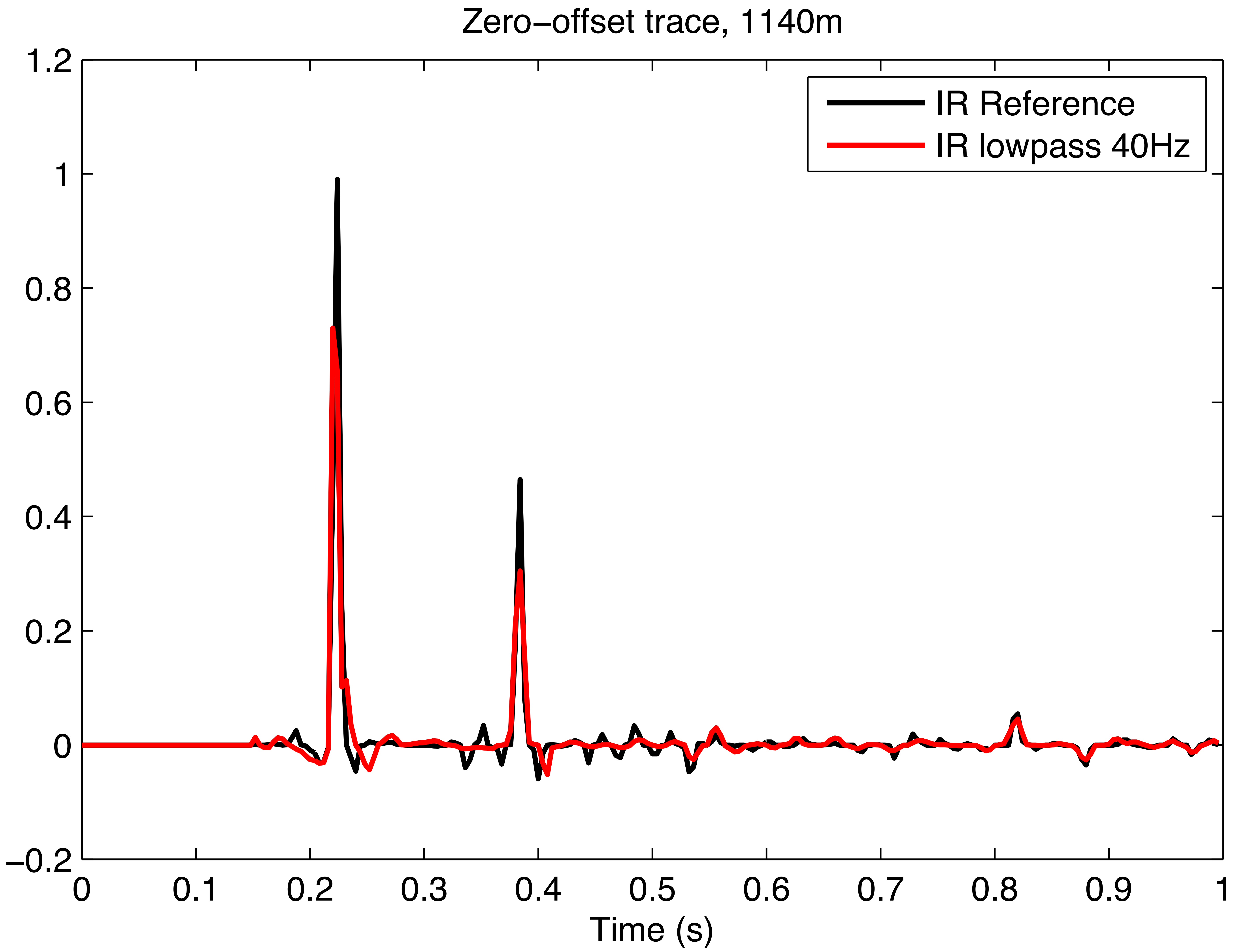
Reference REPSI primary IR
from original data

Motivation: G tolerates lowpass filtering

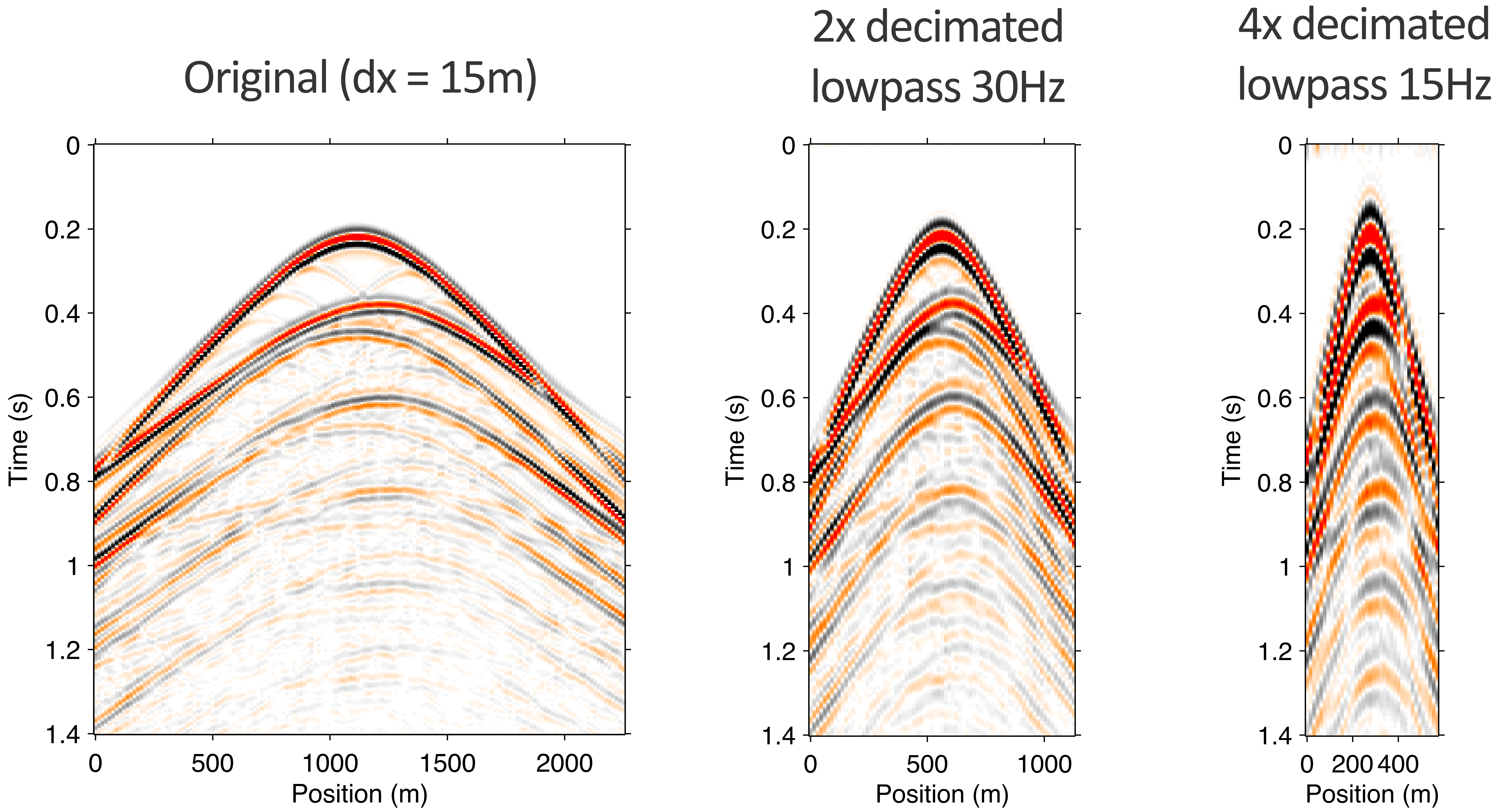


REPSI primary IR
from low-passed data @ 40Hz

Motivation: G tolerates lowpass filtering

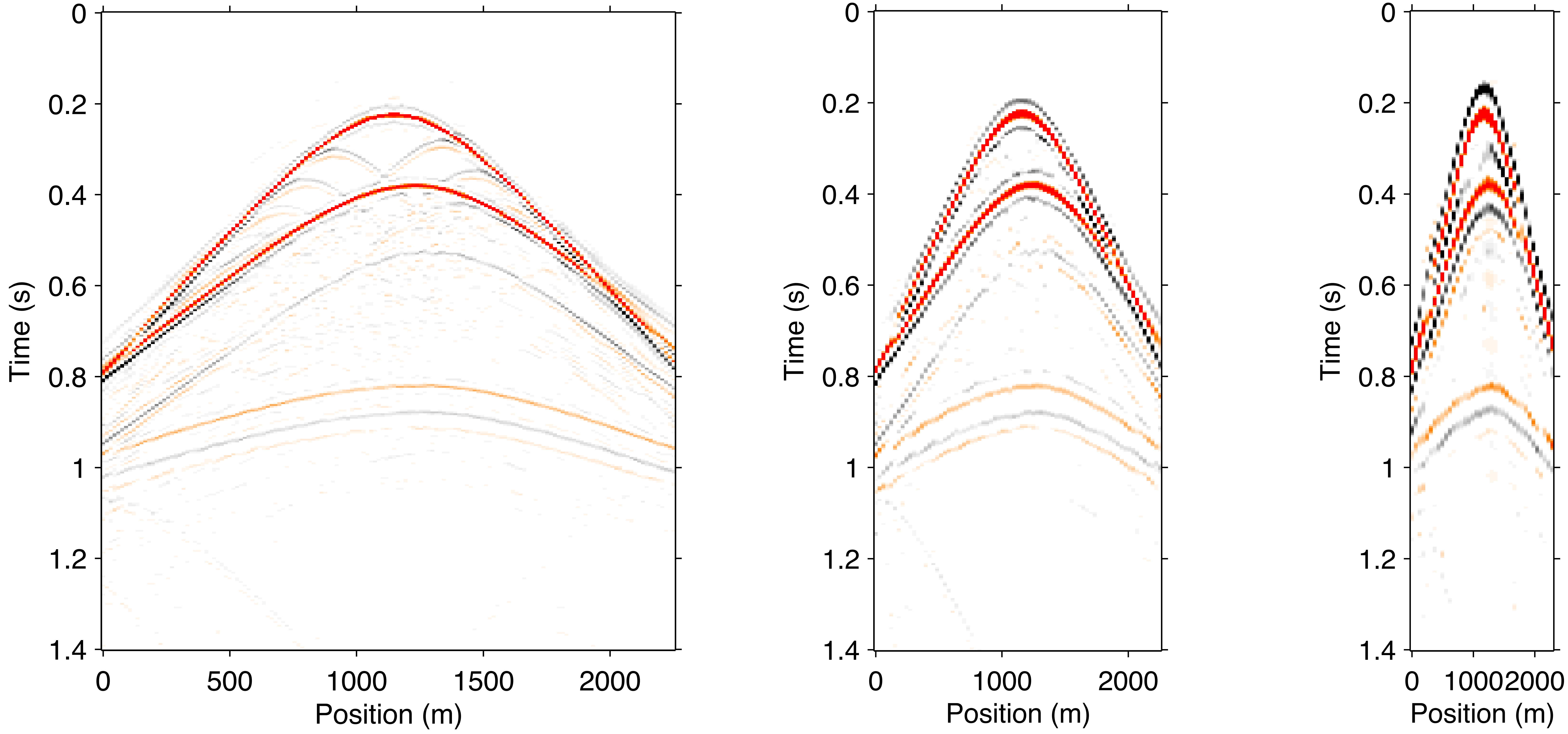


Lowpass data permits coarser sampling w/o aliasing



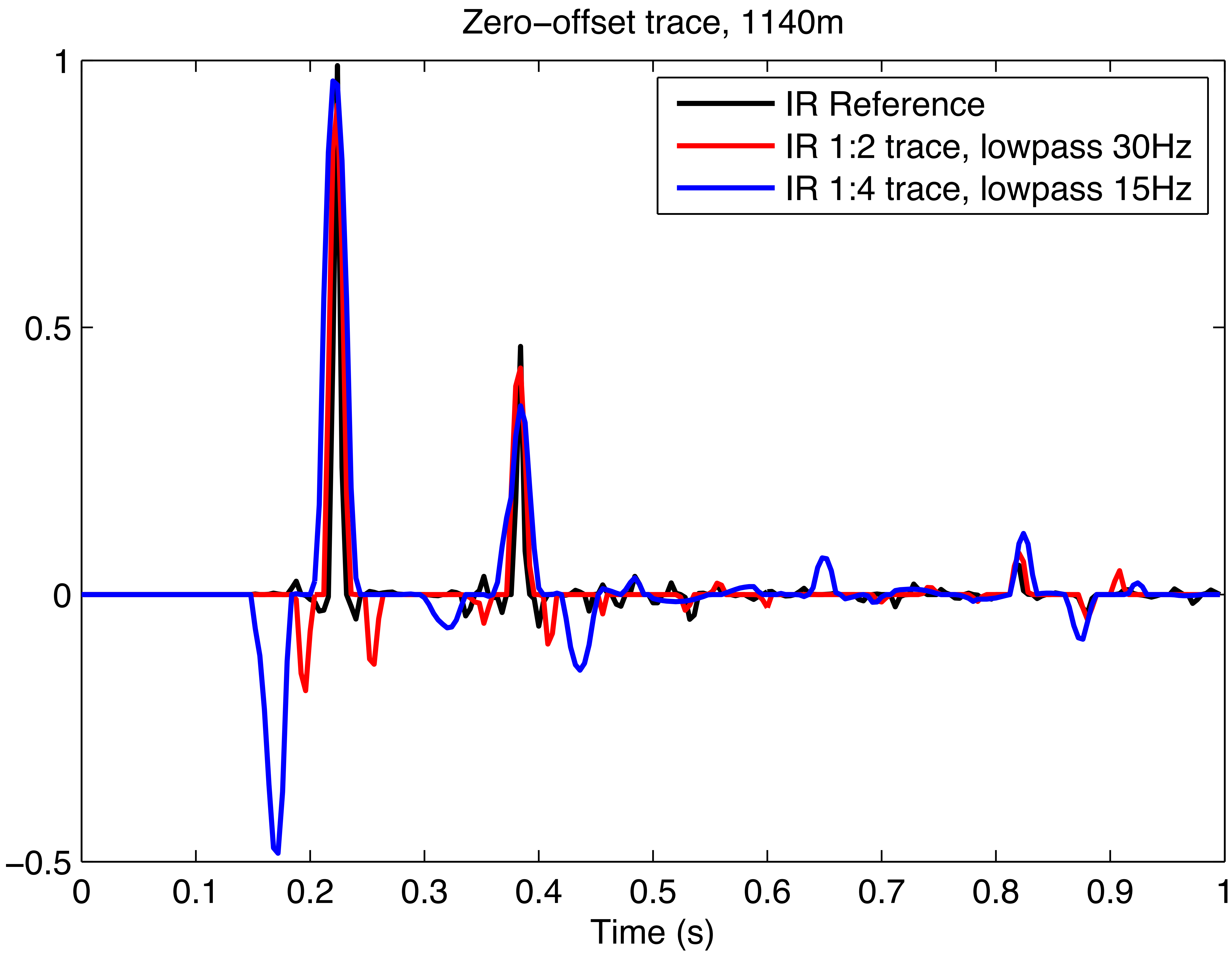
Lowpass data permits coarser sampling w/o aliasing

Impulse response solutions



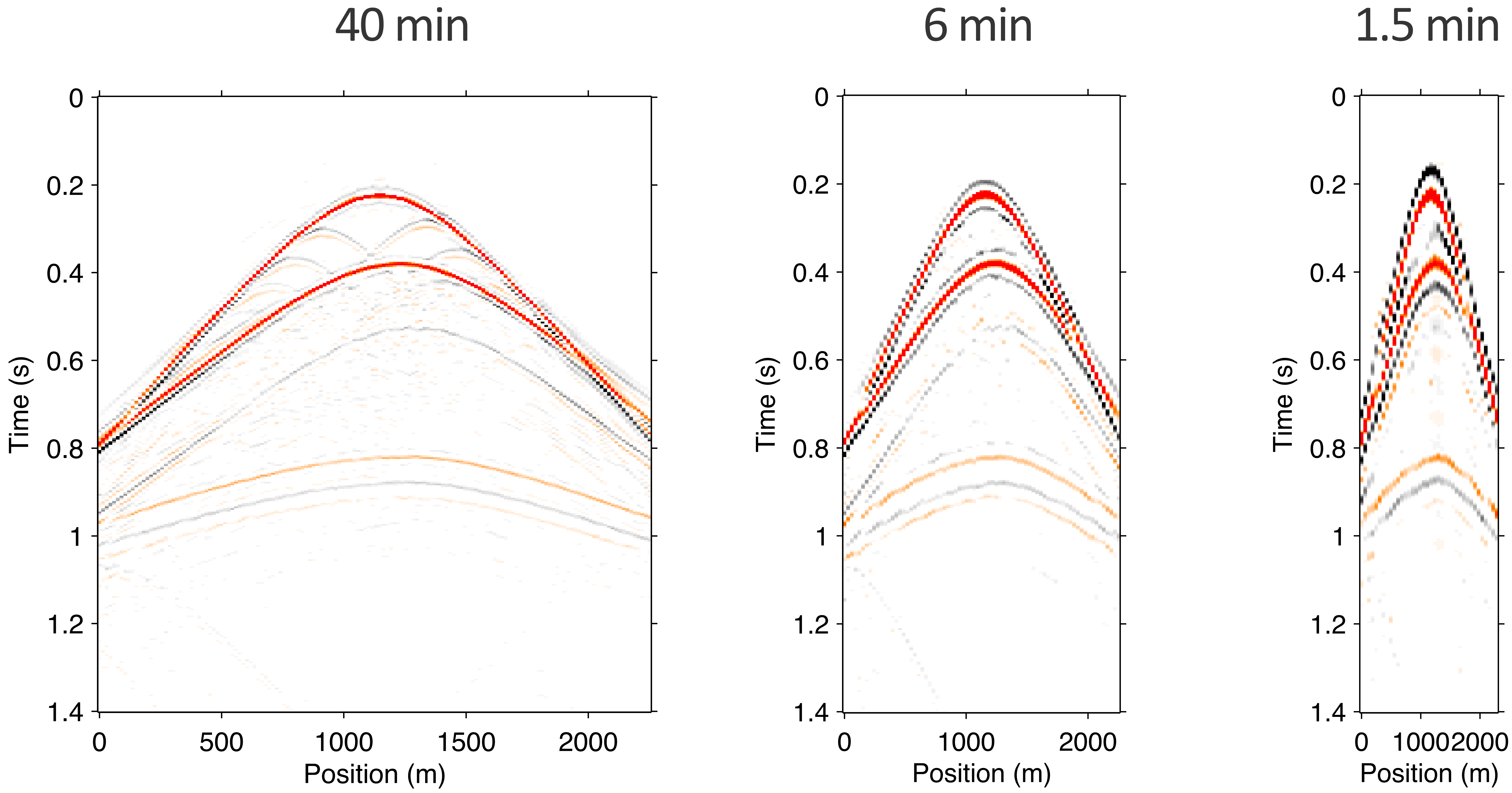
SLIM

Lowpass data permits coarser sampling w/o aliasing



SLIM

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



Multilevel strategy for EPSI

warm-start fine-scale problem with coarse-scale solutions

Idea: Warm-start with coarse-scale solutions

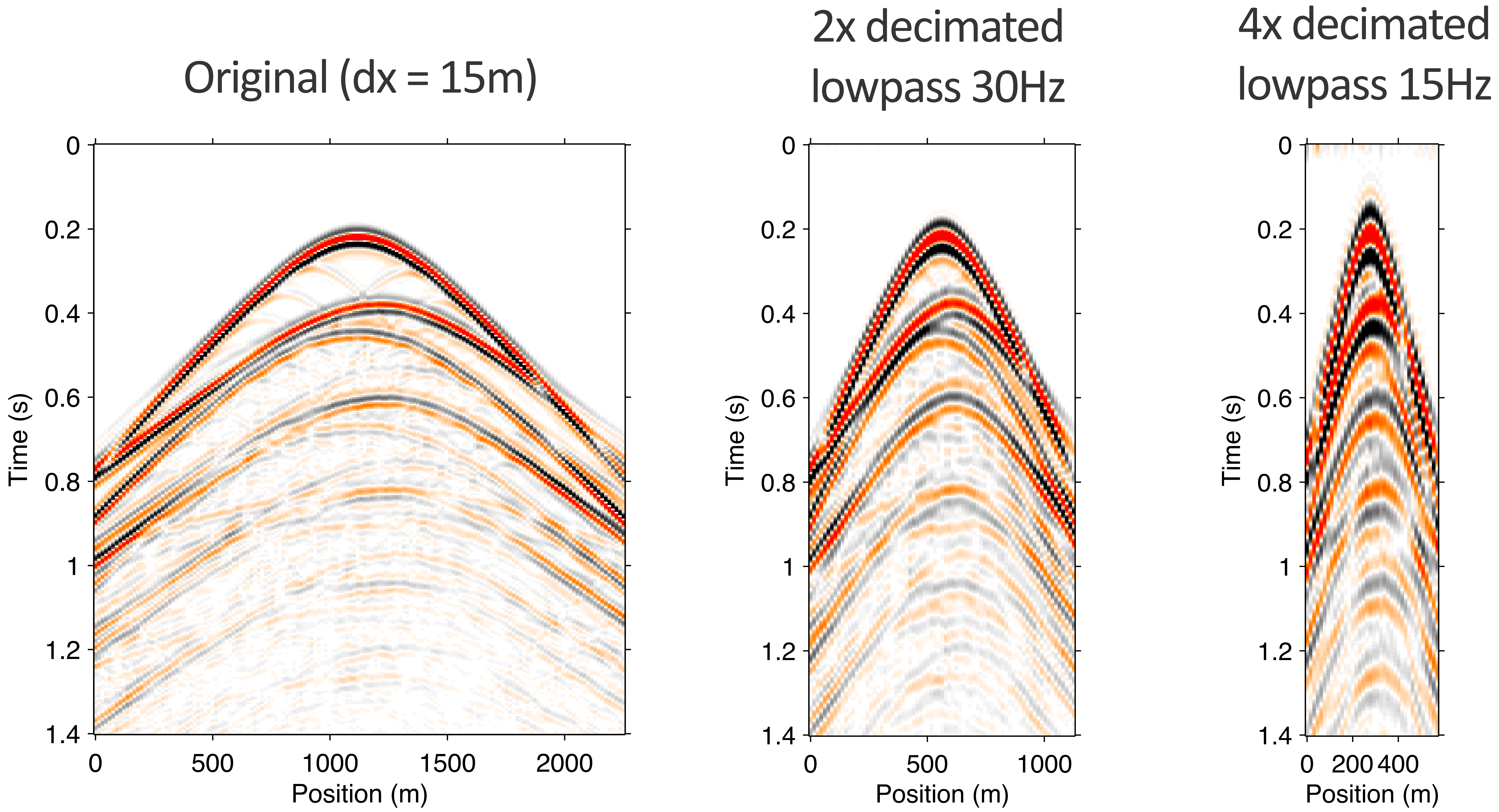
EPSI takes **70-100 iterations** to converge (each iteration is doing 2 SRME multiple prediction), can we make it **FASTER**?

Since decimated datasets solve much faster, we interpolate its (slightly inaccurate) G for the initial estimate to full problem

Previous Q is *discarded*

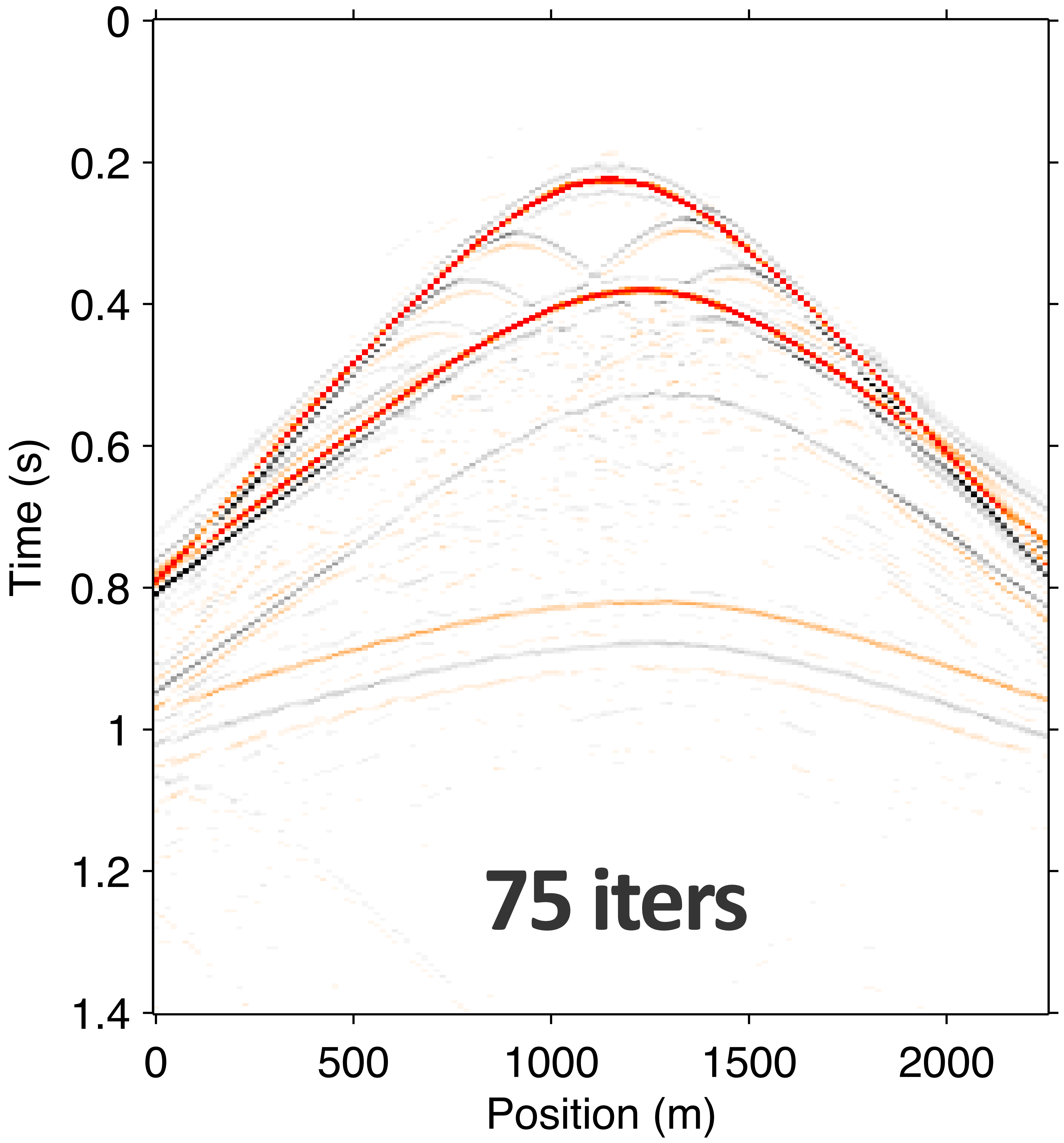
Interpolation method of G not important, just can't alias. Simple constant NMO (i.e., at water velocity) + linear interpolation works fine

Warm-starting/continuation from coarse solution
Example

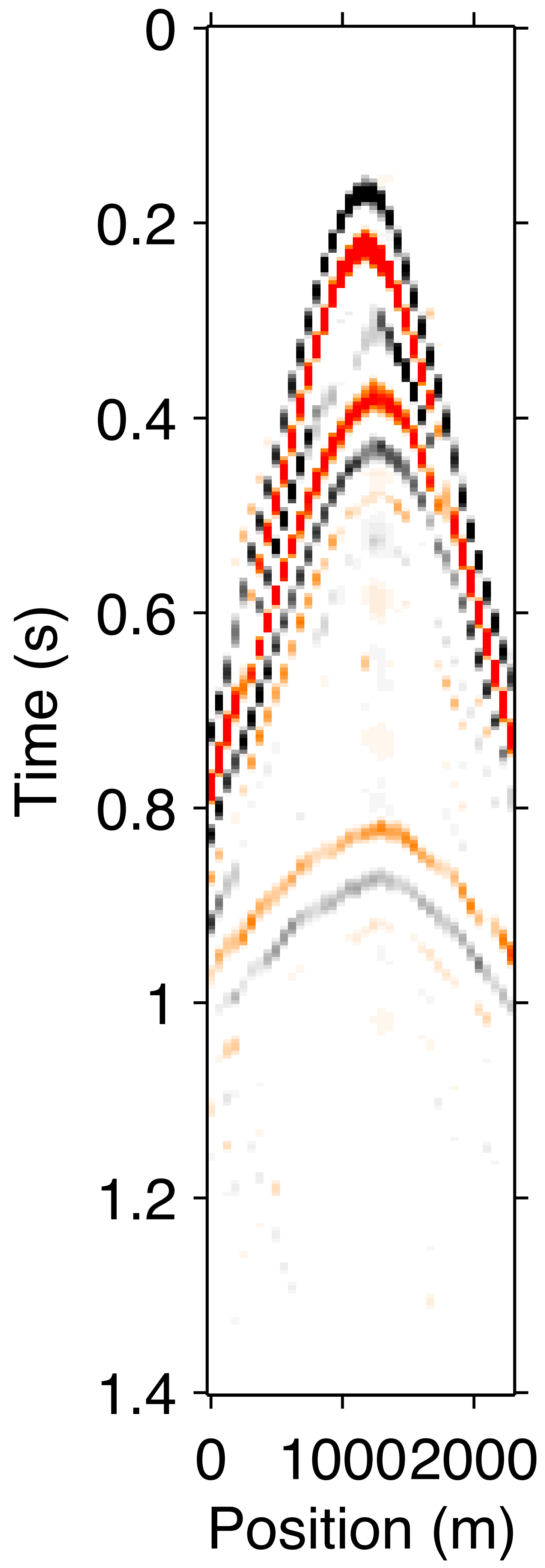


Warm-starting/continuation from coarse solution
Example

Solution of full data



Solution of 4x decimated data

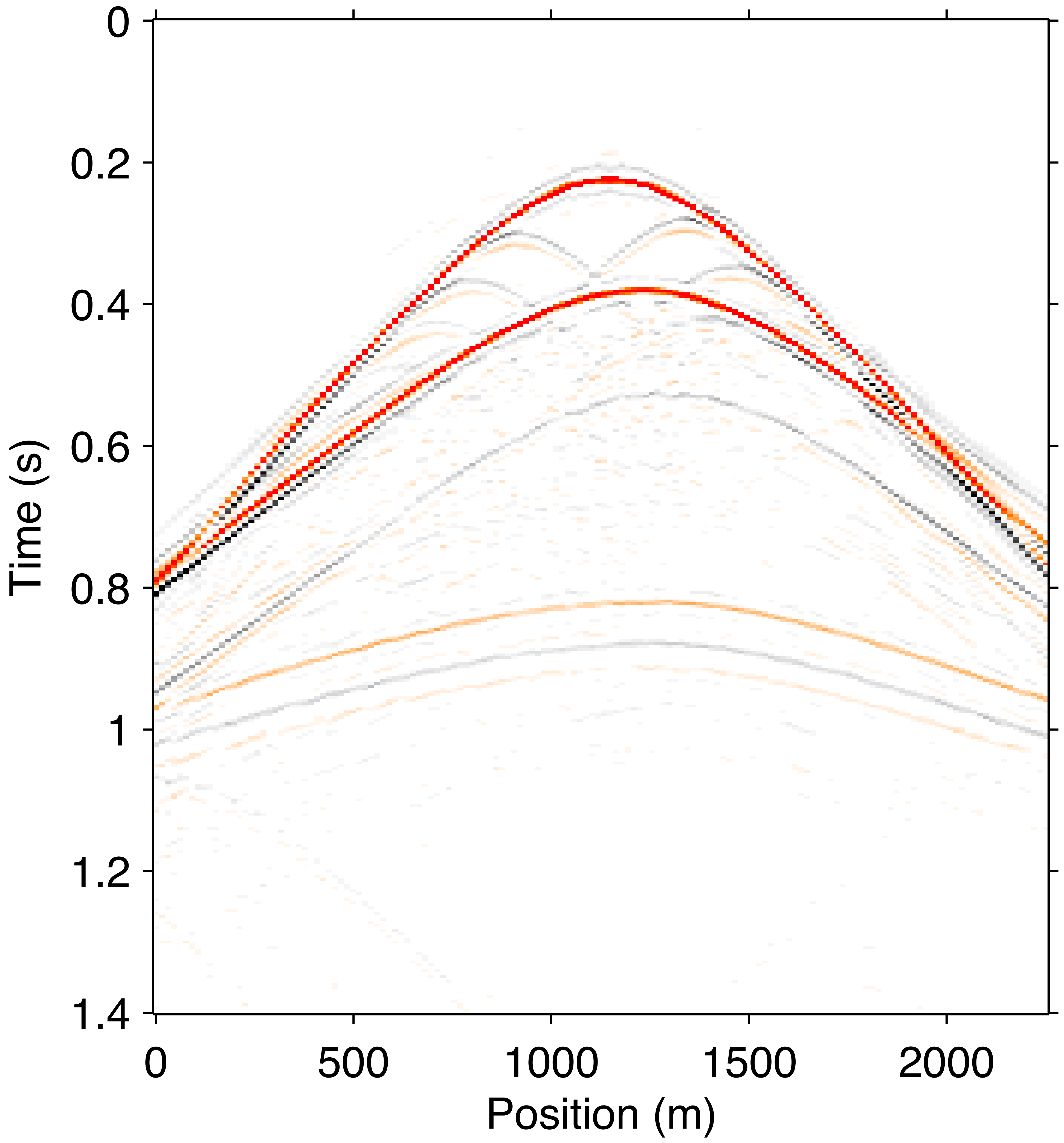


60 iters

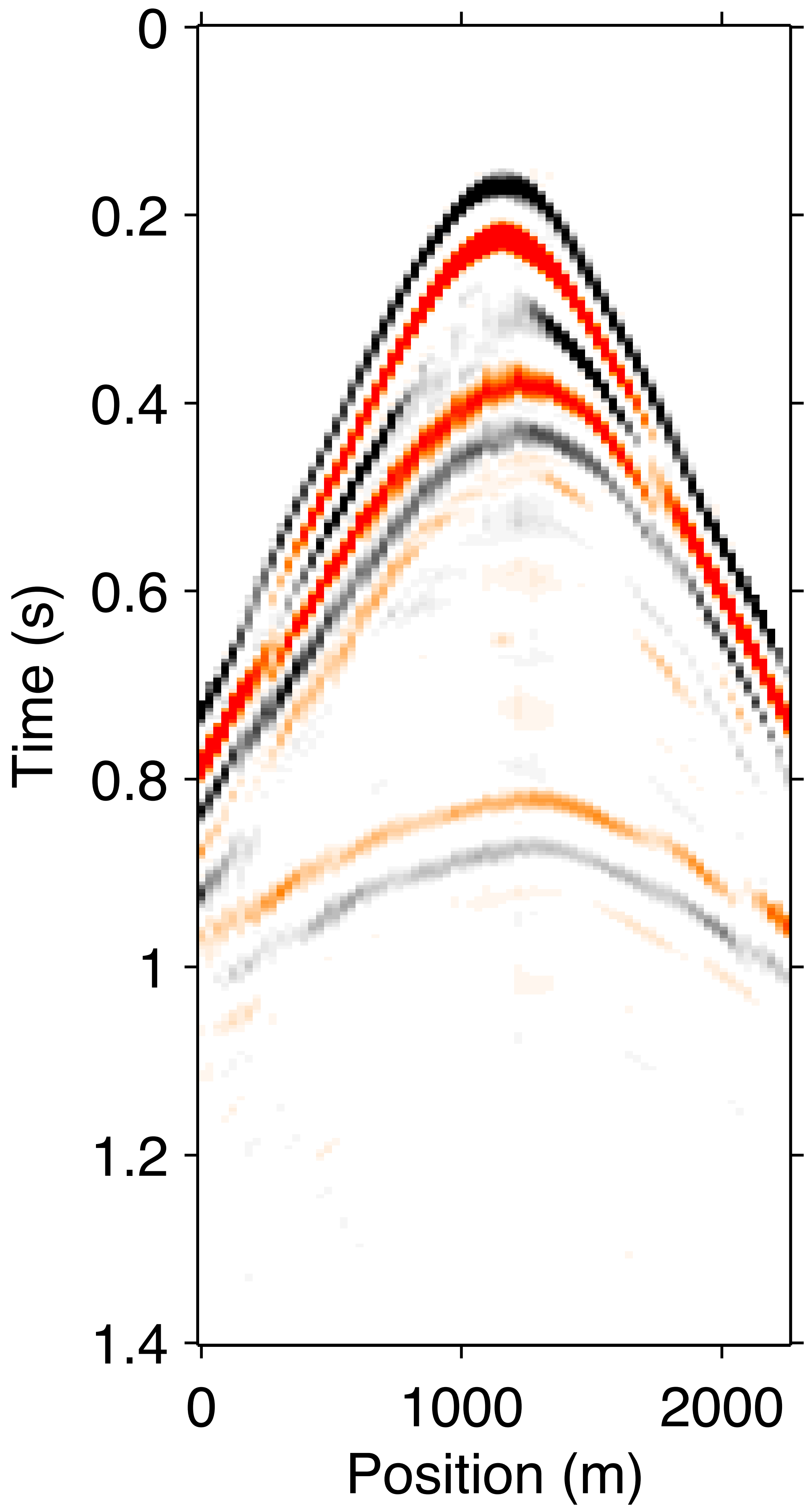
Warm-starting/continuation from coarse solution

Example

Solution of full data

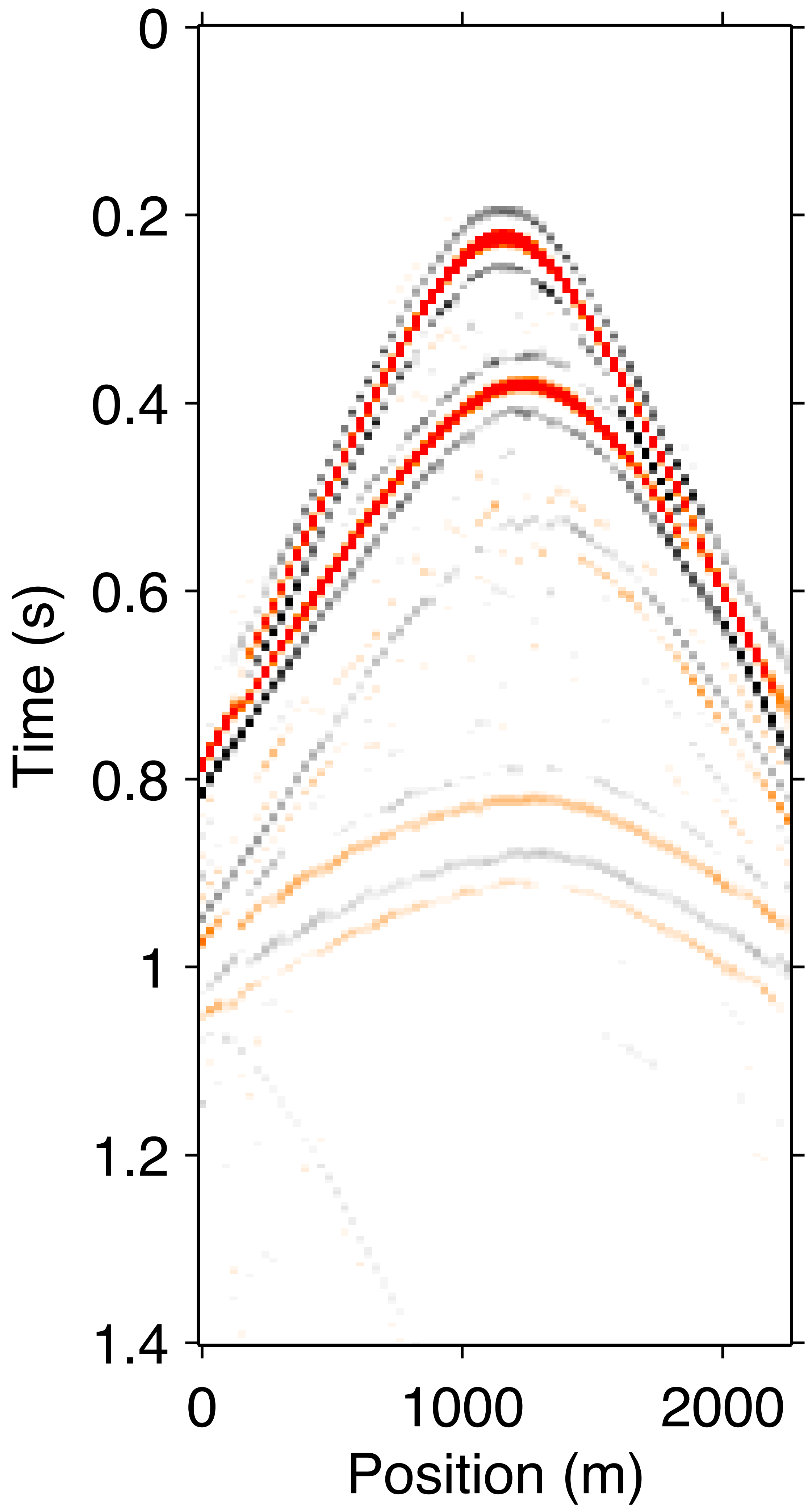


Solution of 4x decimated data
1600m/s NMO, linear interp 2x

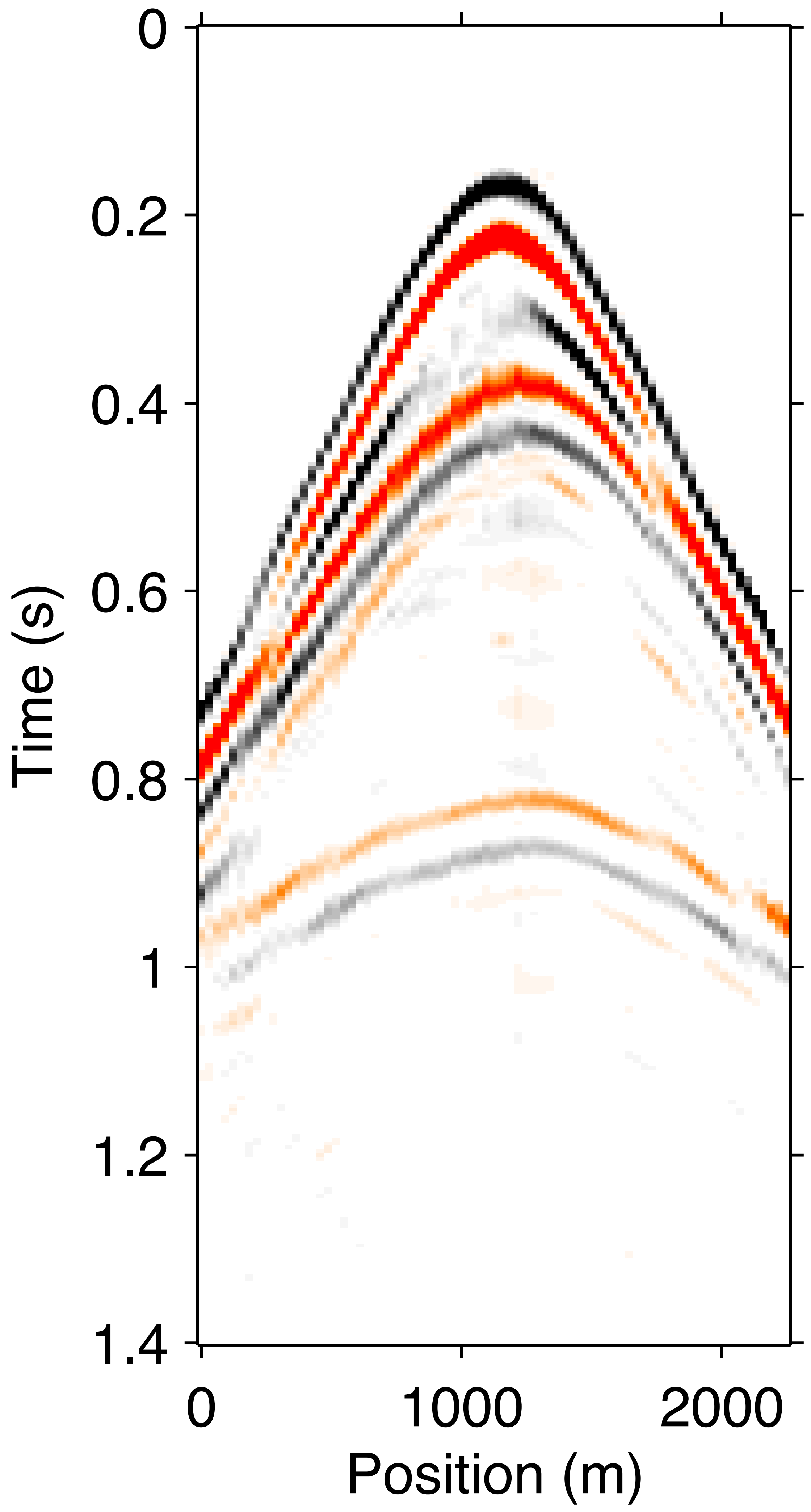


Warm-starting/continuation from coarse solution
Example

Solution of 2x decimated data

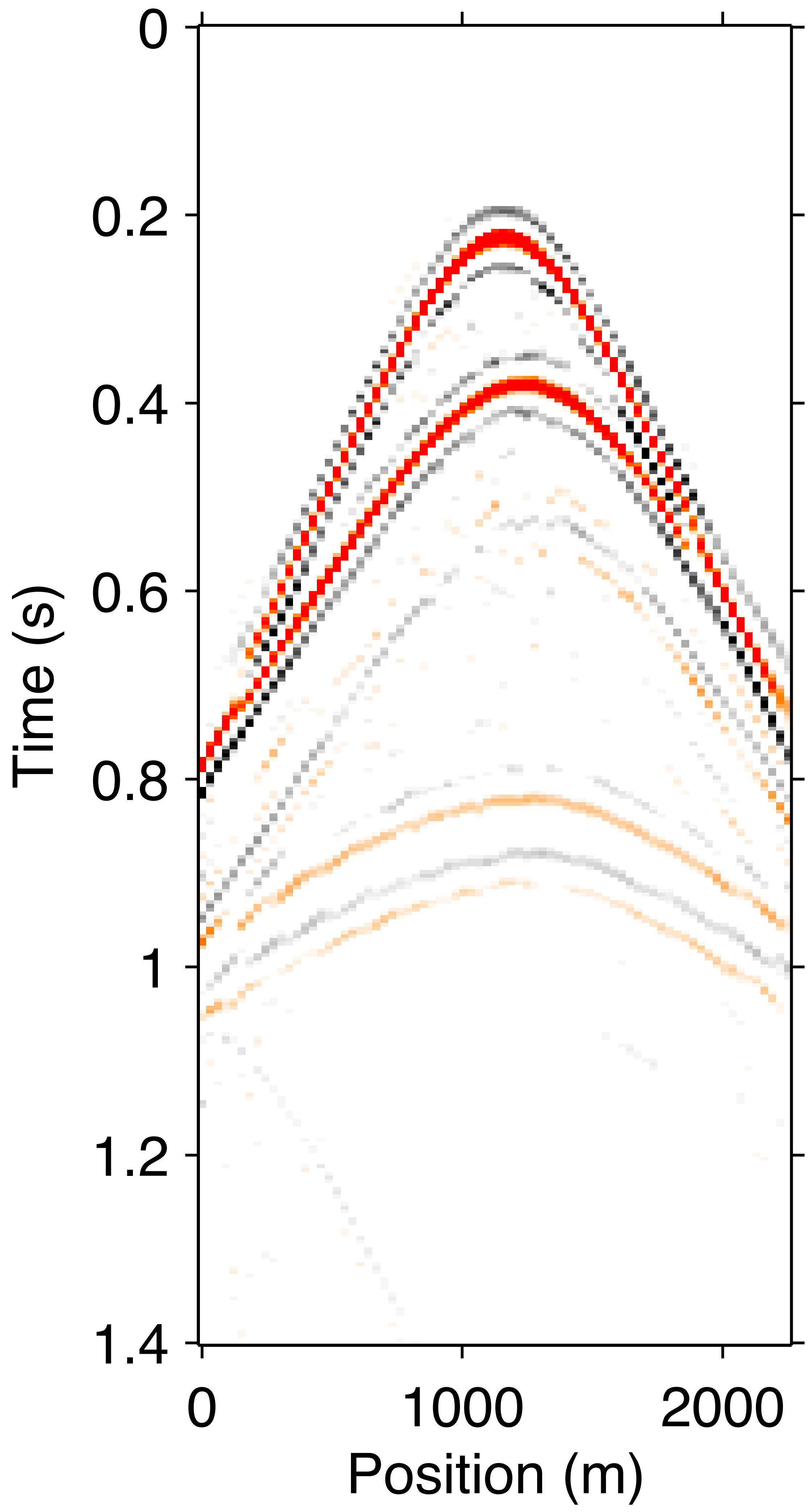


Solution of 4x decimated data
1600m/s NMO, linear interp 2x

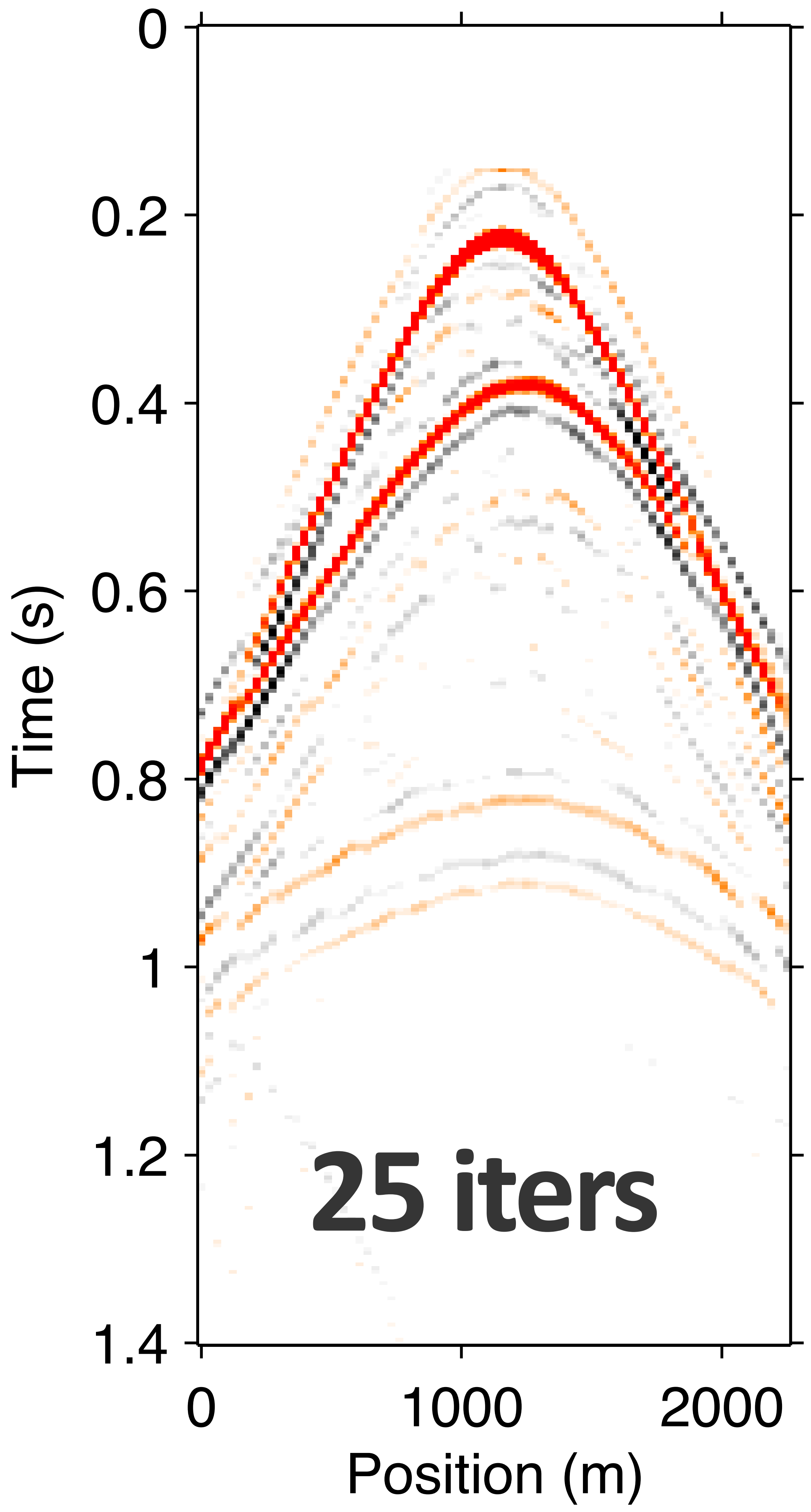


Warm-starting/continuation from coarse solution
Example

Solution of 2x decimated data



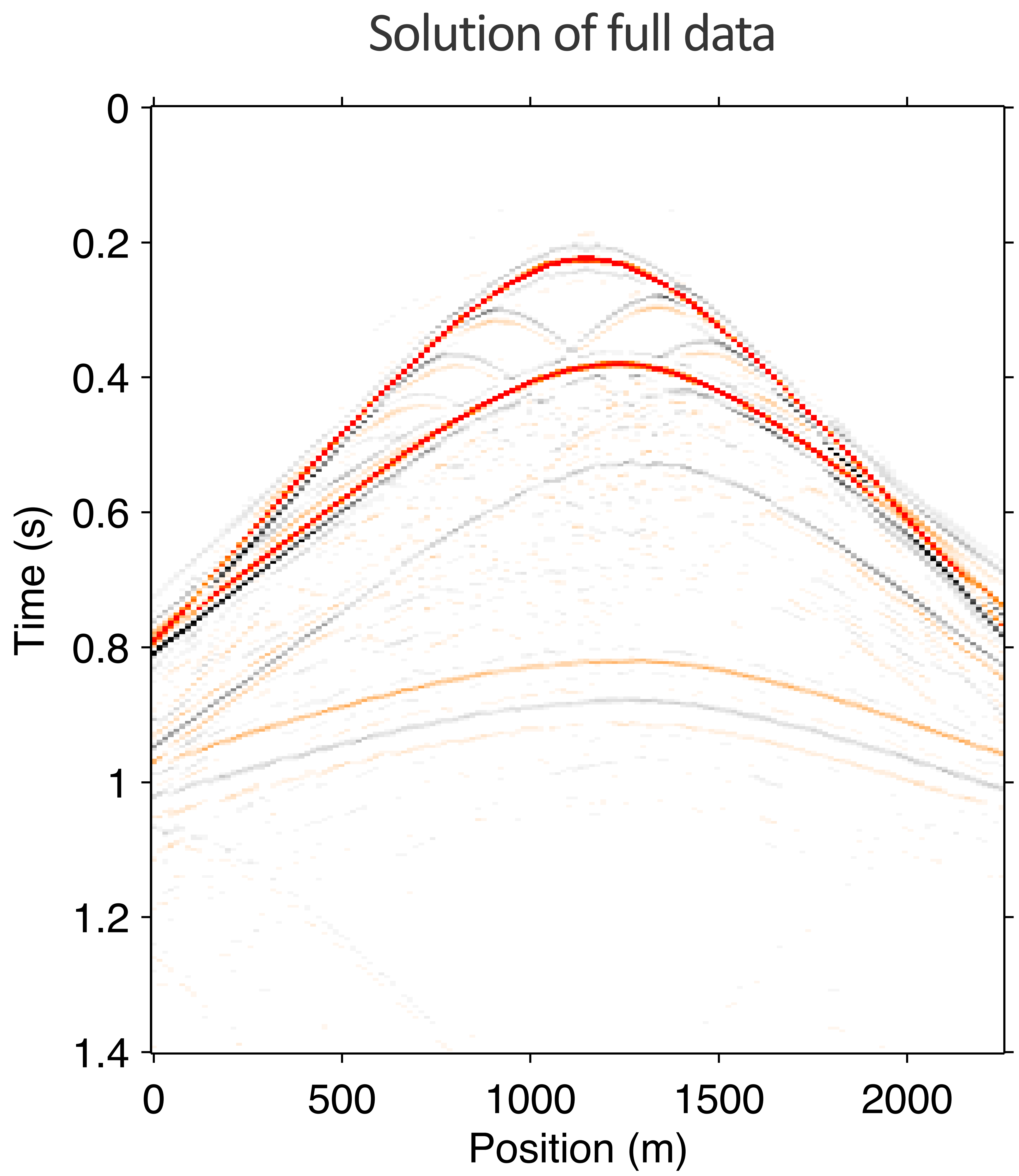
Solution on 2x dec data
continuation from 4x dec solution

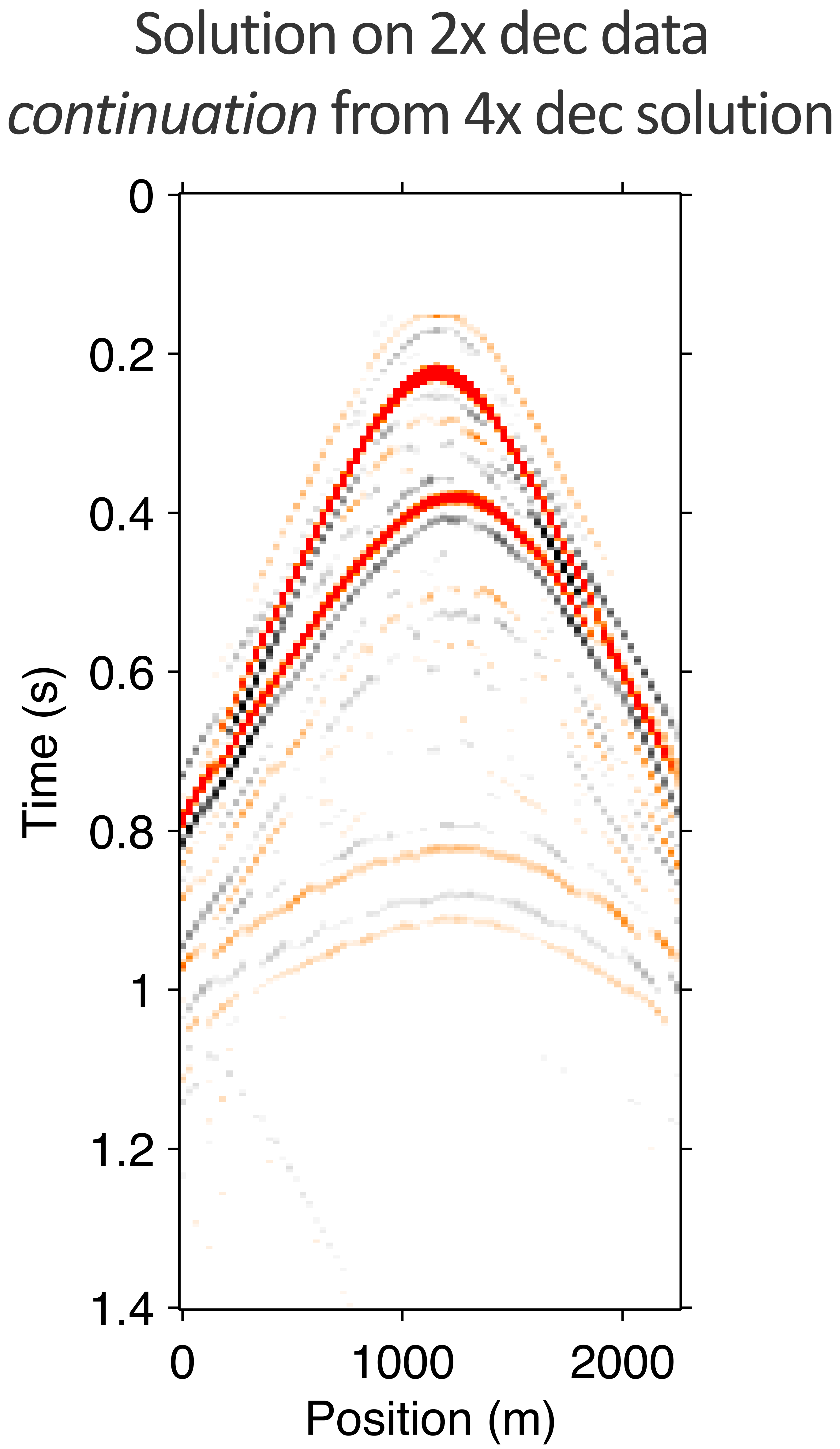


SLIM

Warm-starting/continuation from coarse solution

Example

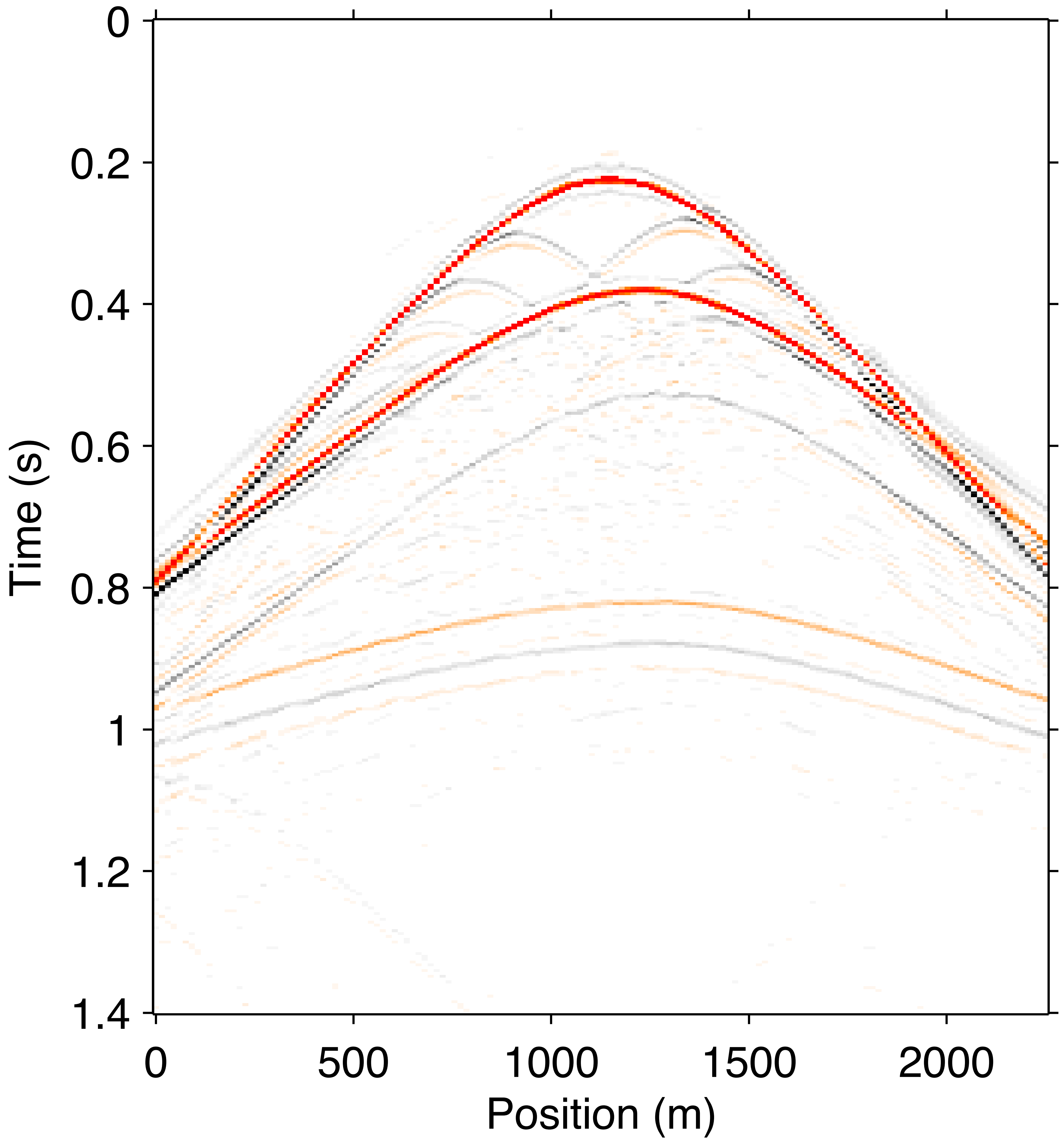




Warm-starting/continuation from coarse solution

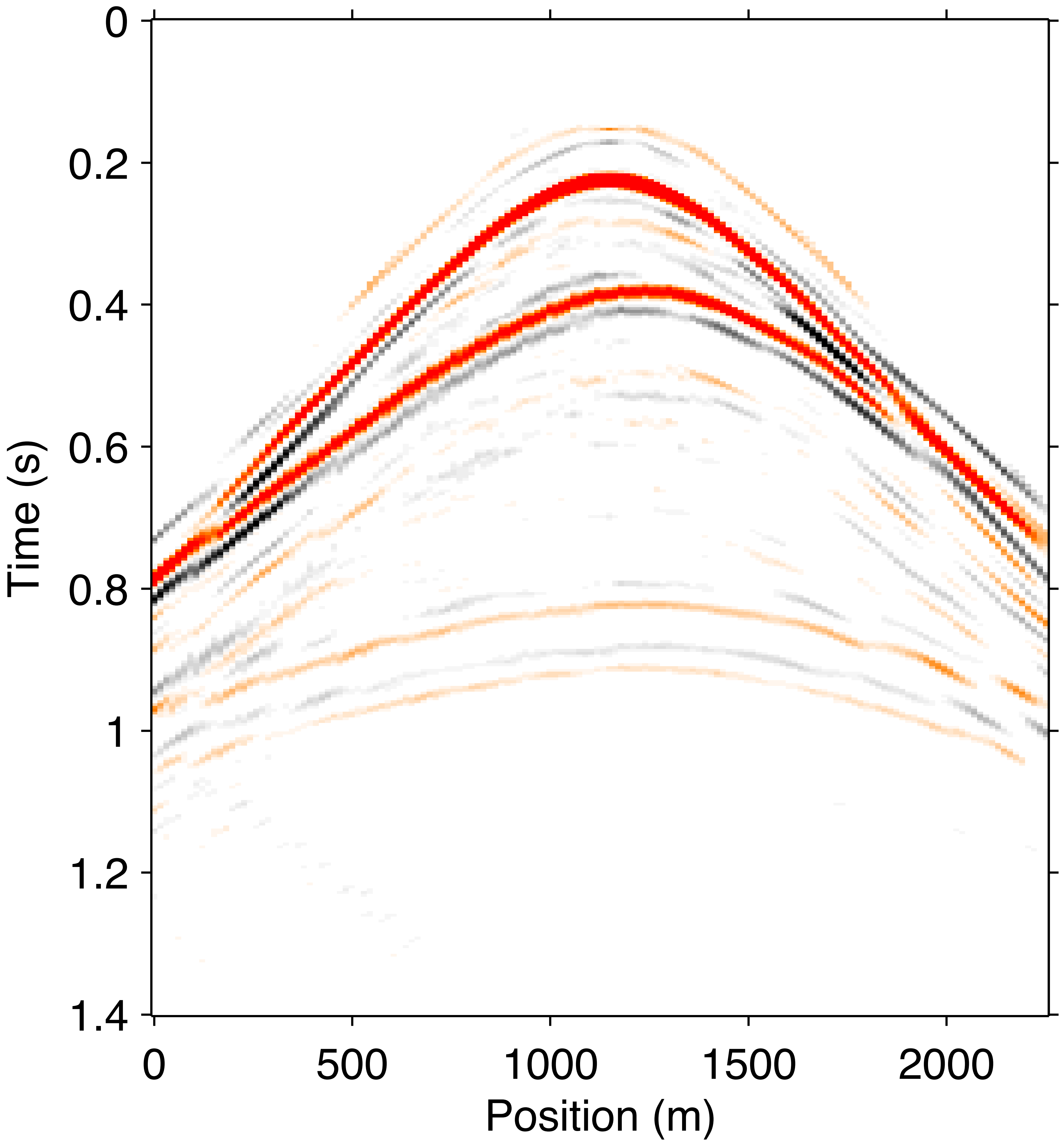
Example

Solution of full data

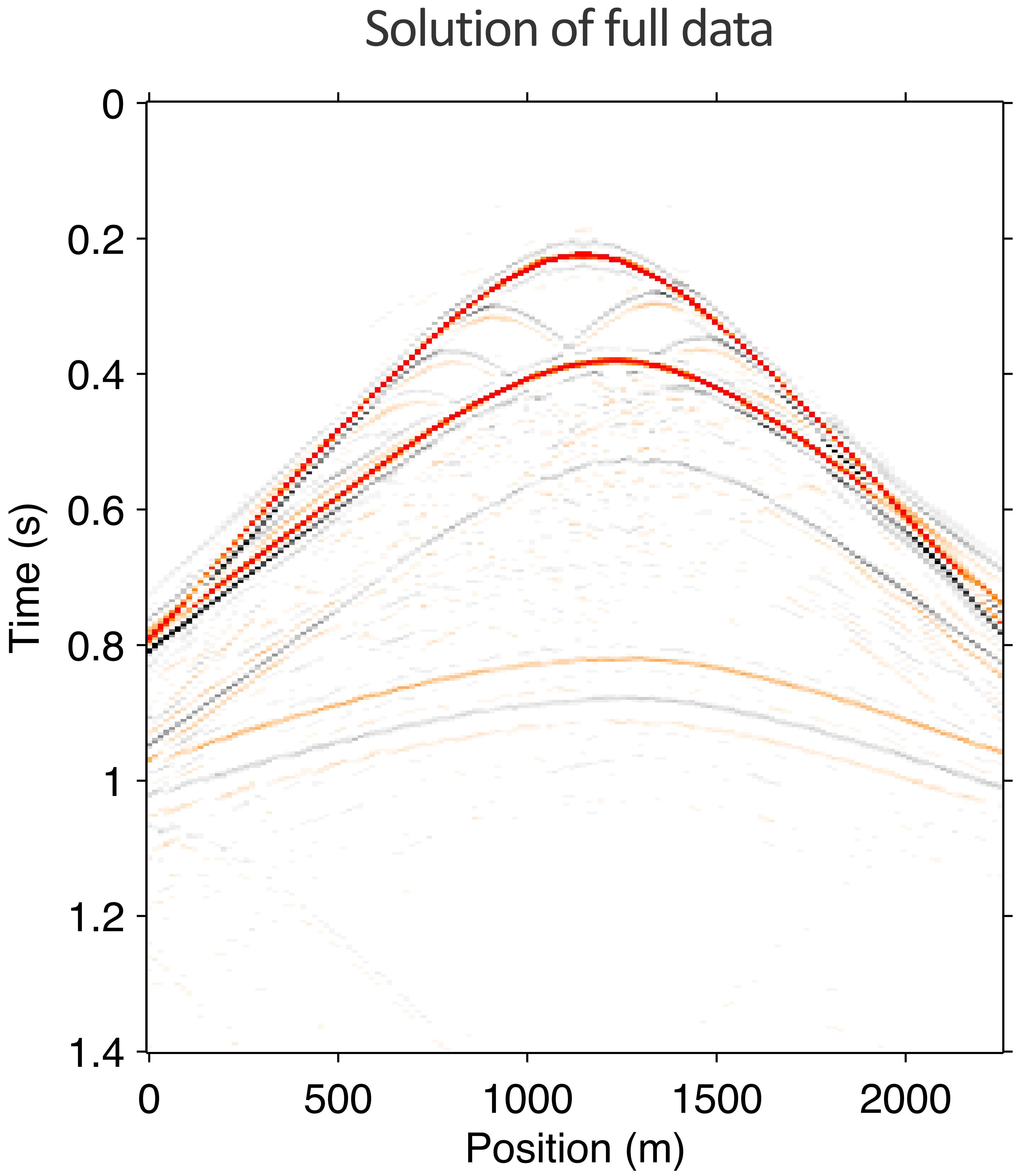


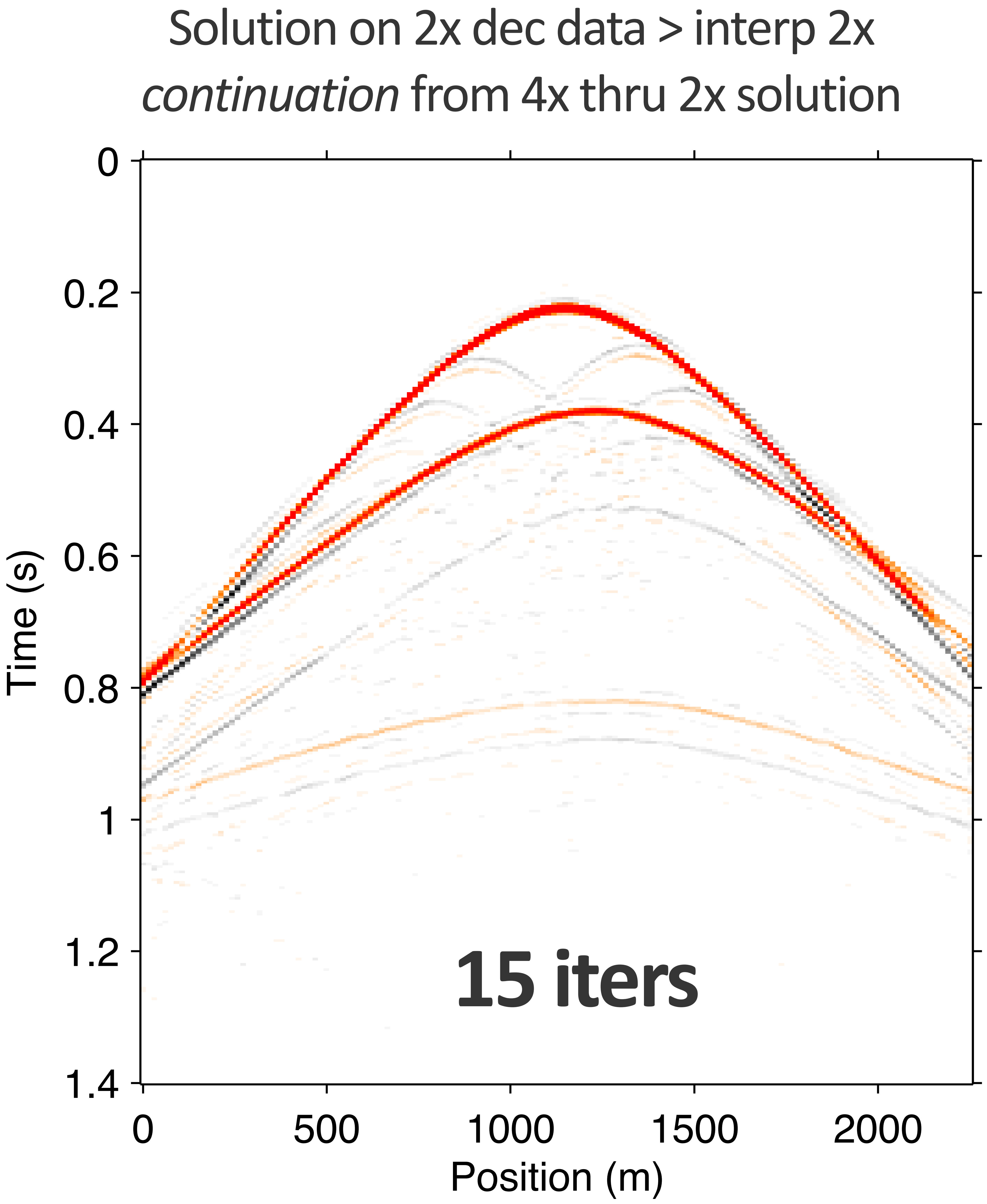
Solution on 2x dec data > interp 2x

continuation from 4x dec solution

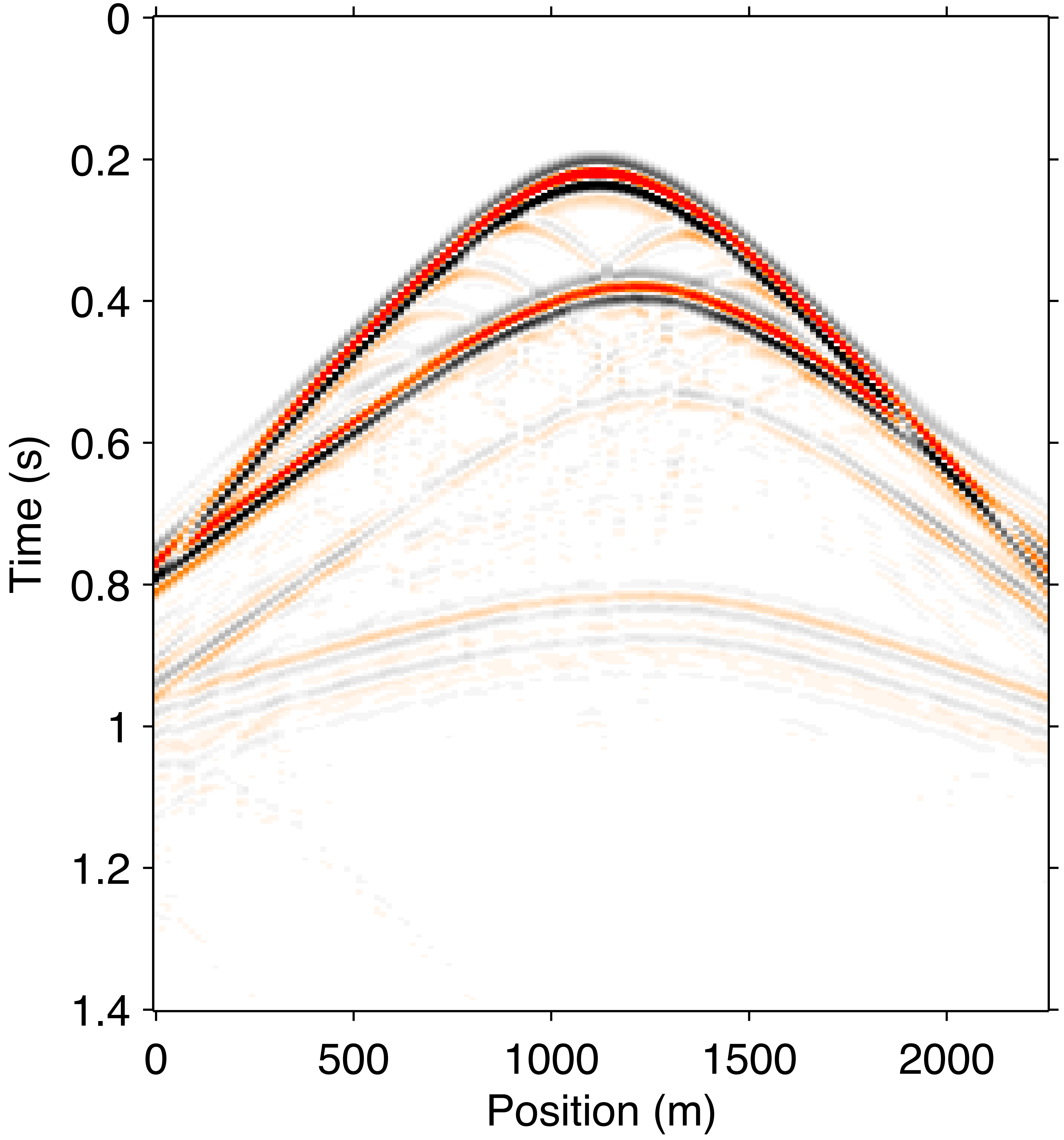


Warm-starting/continuation from coarse solution
Example





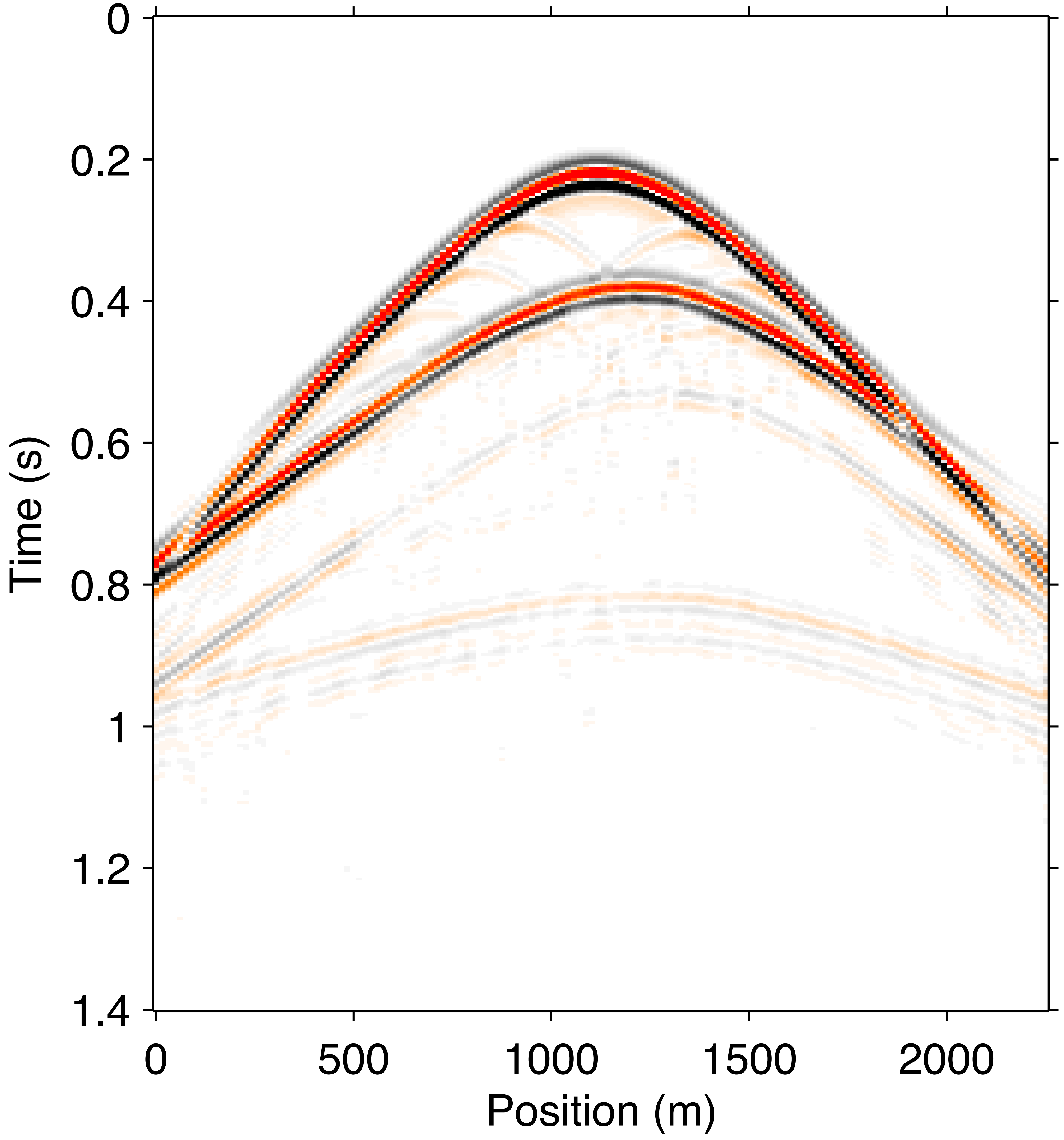
Warm-starting/continuation from coarse solution
Example



Direct Primary

Solved with plain algorithm from finest scale data

Warm-starting/continuation from coarse solution
Example

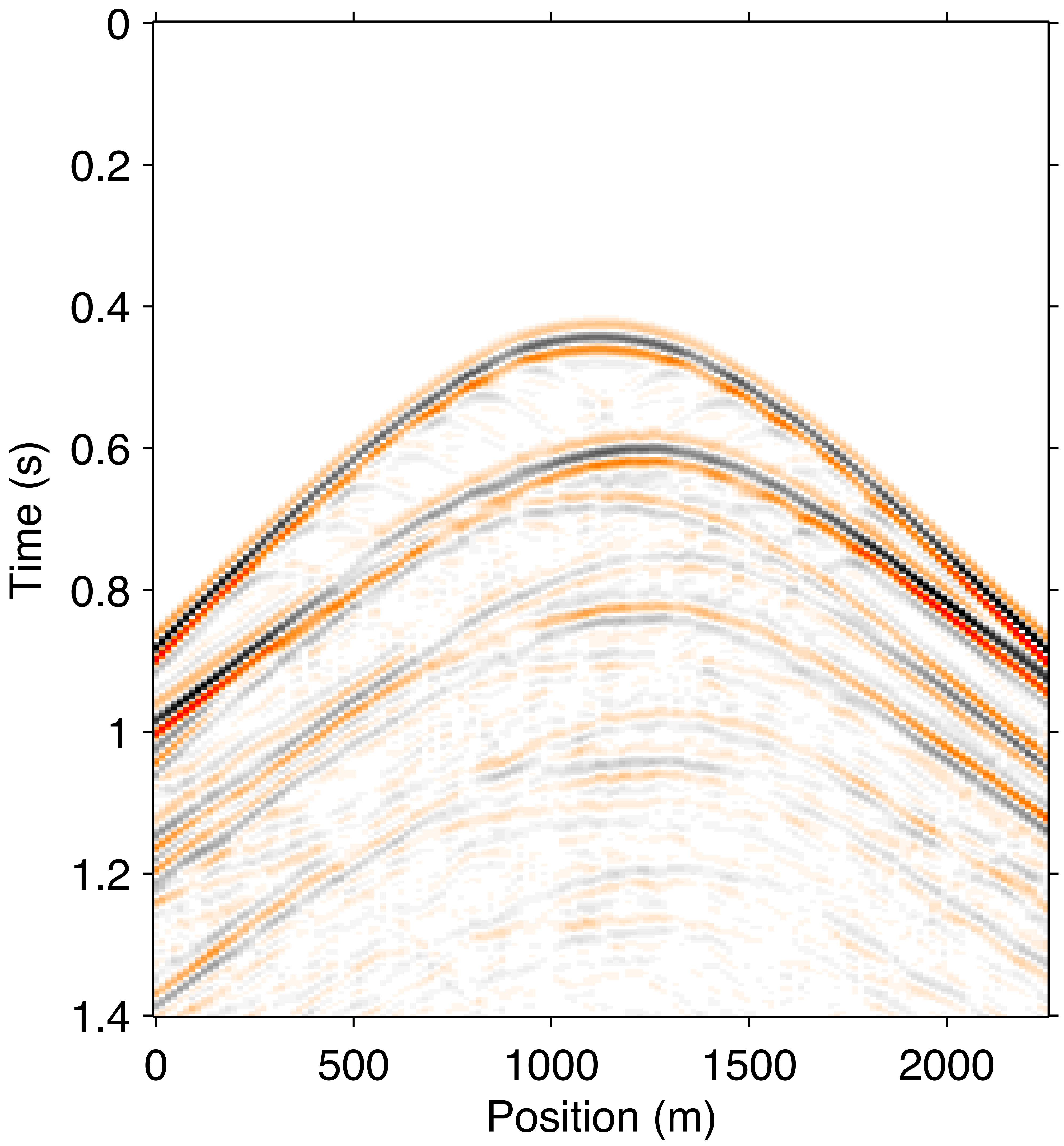


Direct Primary

Solved with spatial sampling continuation

dx = 60m > 30m > 15m

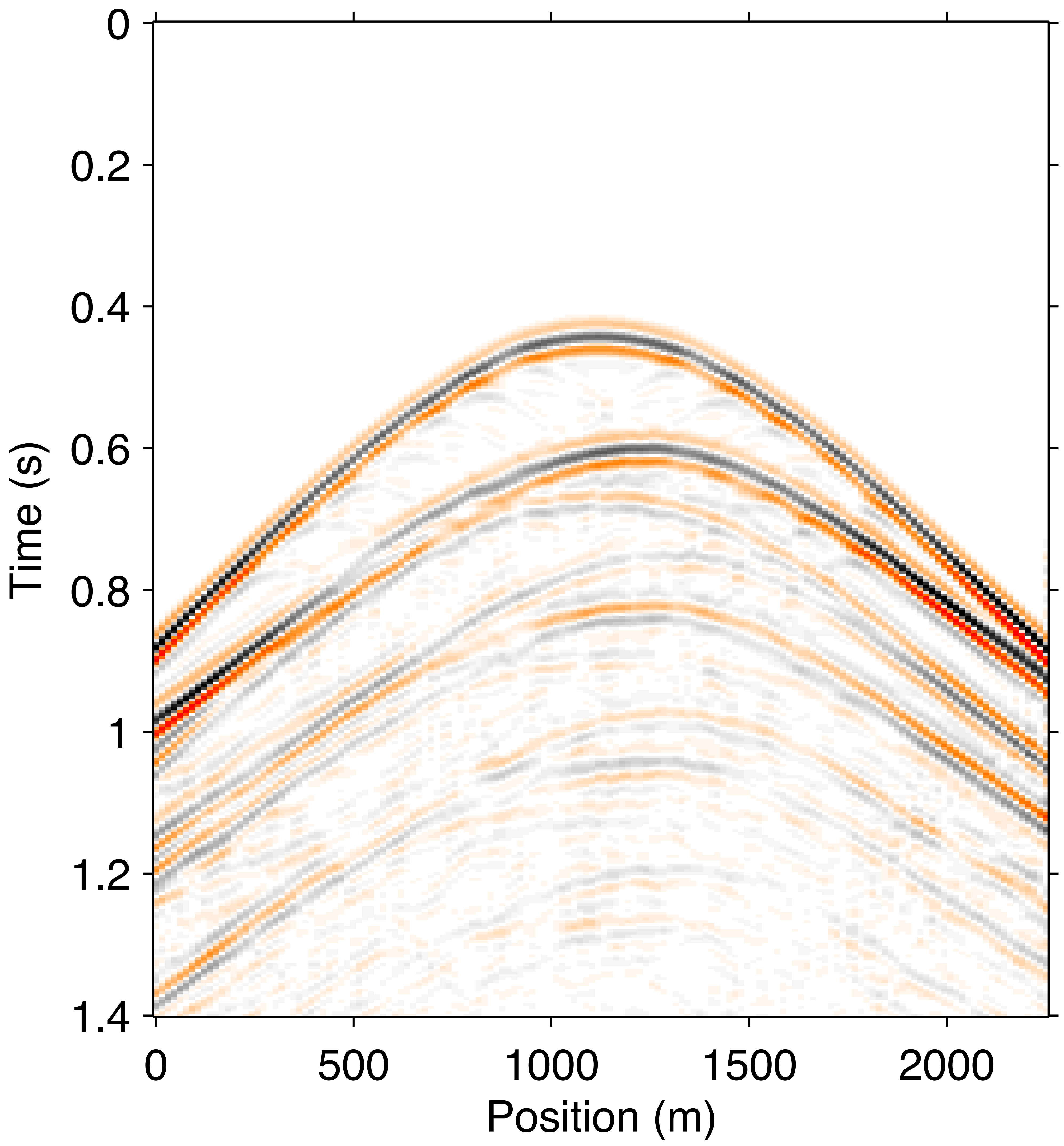
Warm-starting/continuation from coarse solution
Example



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

Significant speedup from bootstrapping (in 2D)

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

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

2 times FFT   computing MCG & sum in FX

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

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

Significant speedup from bootstrapping (in 2D)

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

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

2 times FFT   computing MCG & sum in FX

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

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

Significant speedup from bootstrapping (in 3D)

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

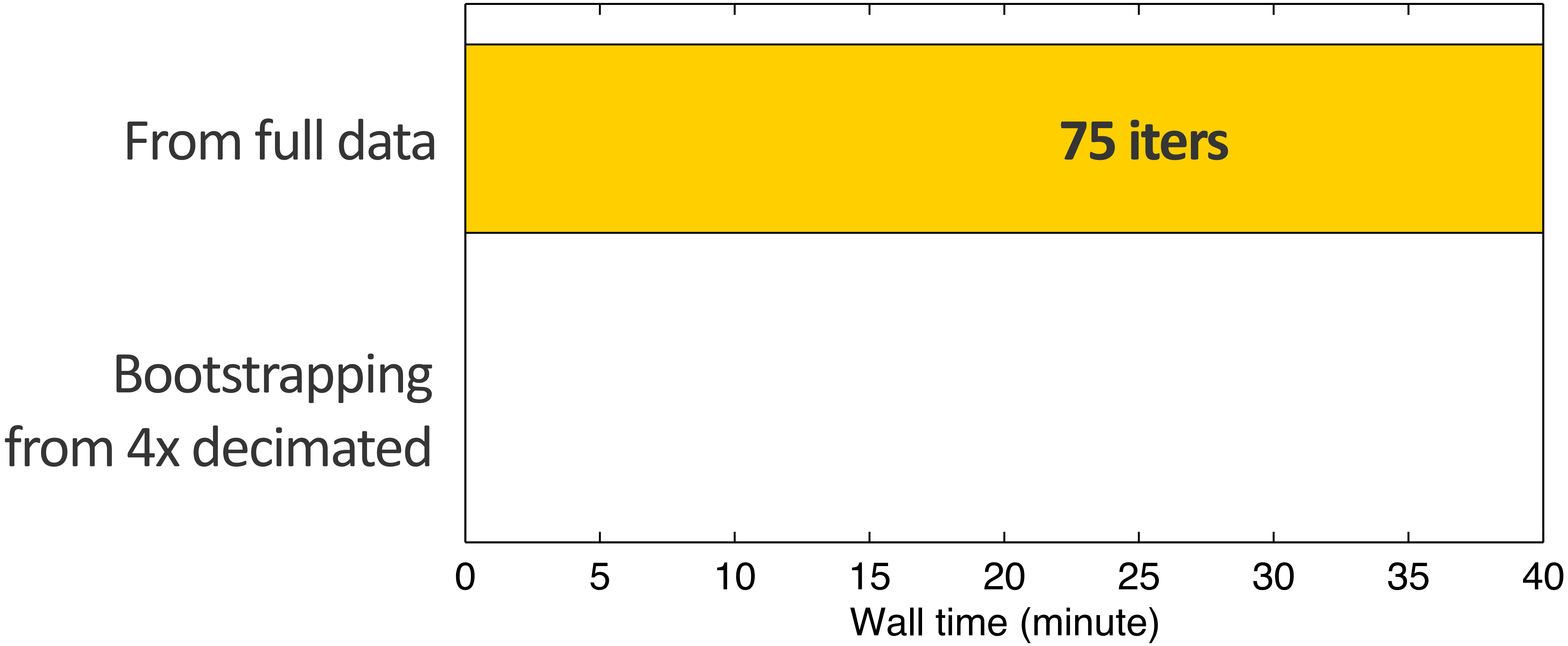
$$\text{Cost}(n) = \mathcal{O}(2n_t n^4 \log n_t) + \mathcal{O}(n_f n^6)$$

2 times FFT — computing MCG & sum in FX

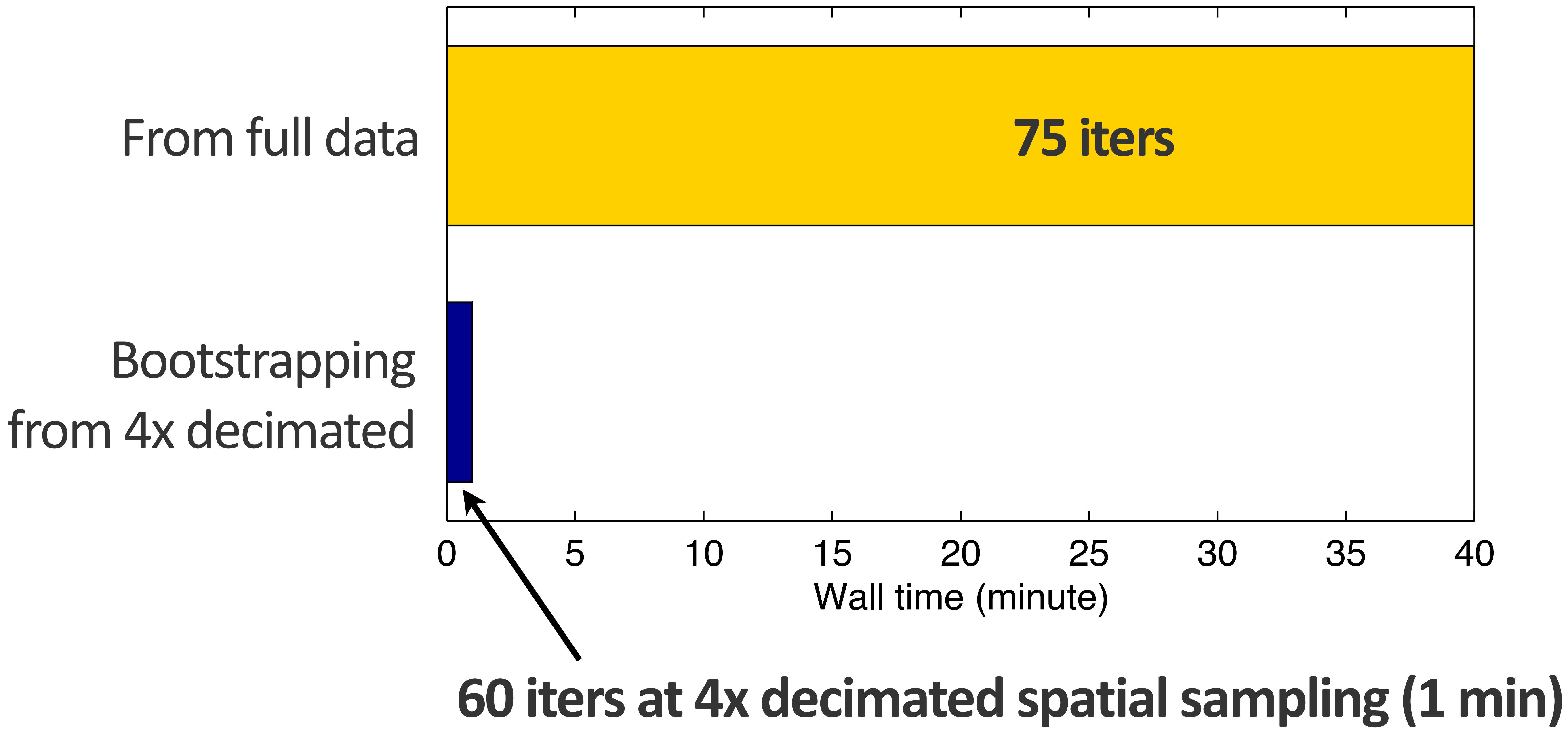
$$\text{Cost}\left(\frac{1}{2}n, \frac{1}{2}n_f\right) = \frac{1}{16}\mathcal{O}(2n_t n^4 \log n_t) + \frac{1}{128}\mathcal{O}(n_f n^6)$$

$$\text{Cost}\left(\frac{1}{4}n, \frac{1}{4}n_f\right) = \frac{1}{256}\mathcal{O}(2n_t n^4 \log n_t) + \frac{1}{8192}\mathcal{O}(n_f n^6)$$

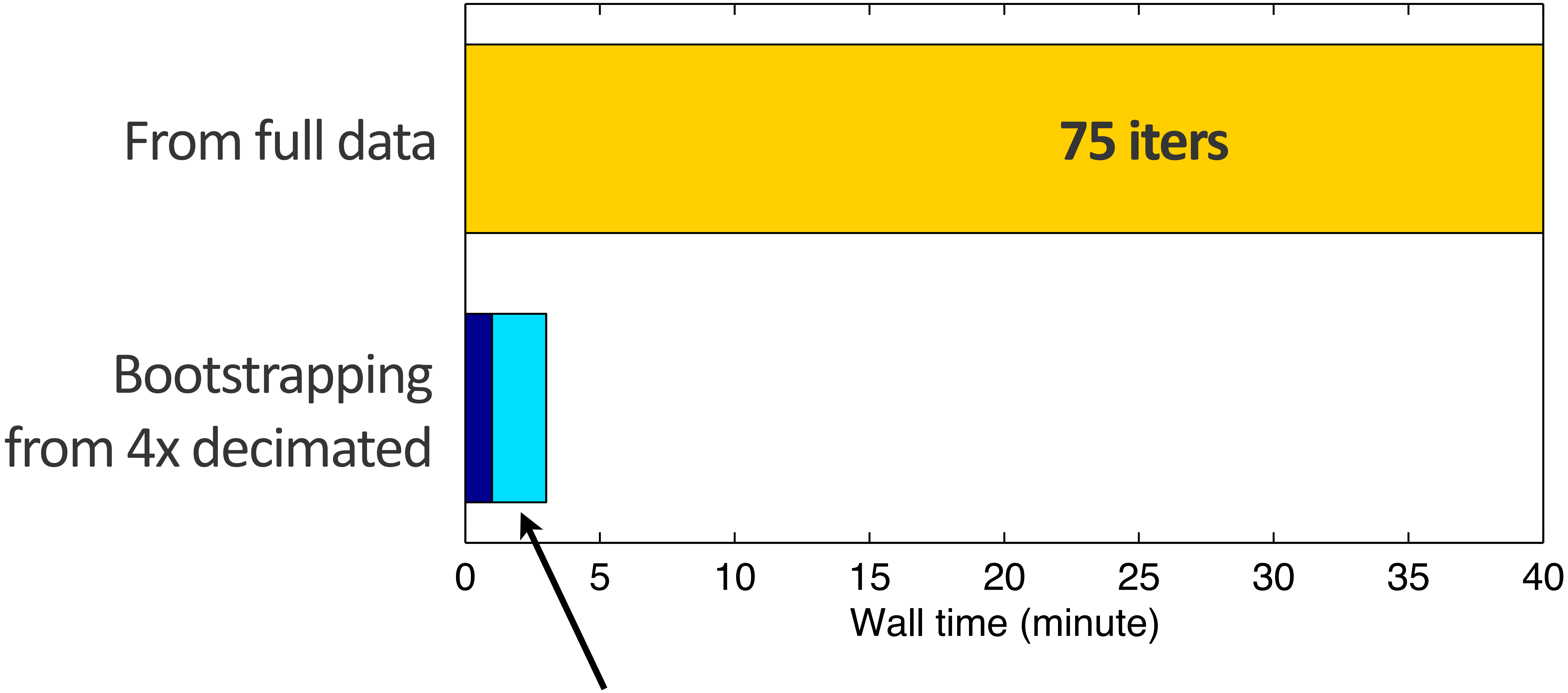
Significant speedup from bootstrapping
Wall times



Significant speedup from bootstrapping
Wall times

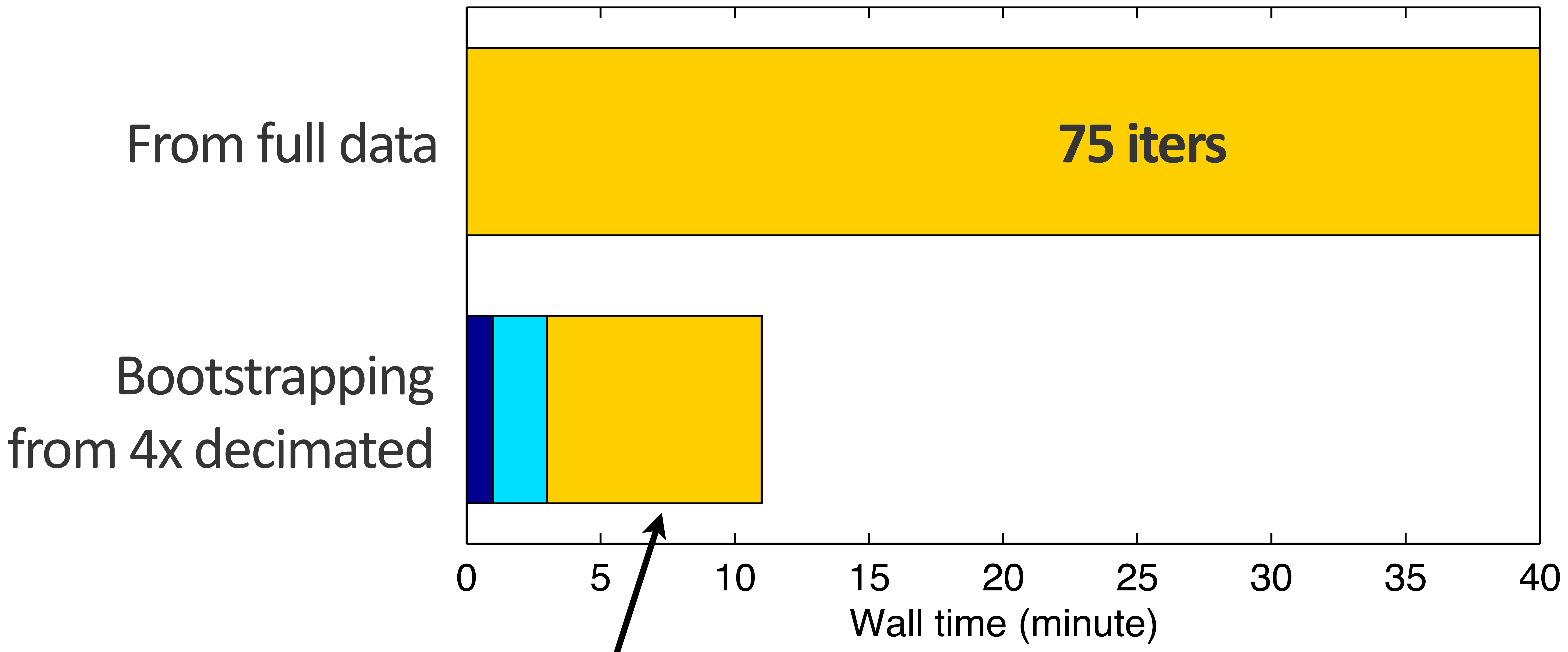


Significant speedup from bootstrapping
Wall times



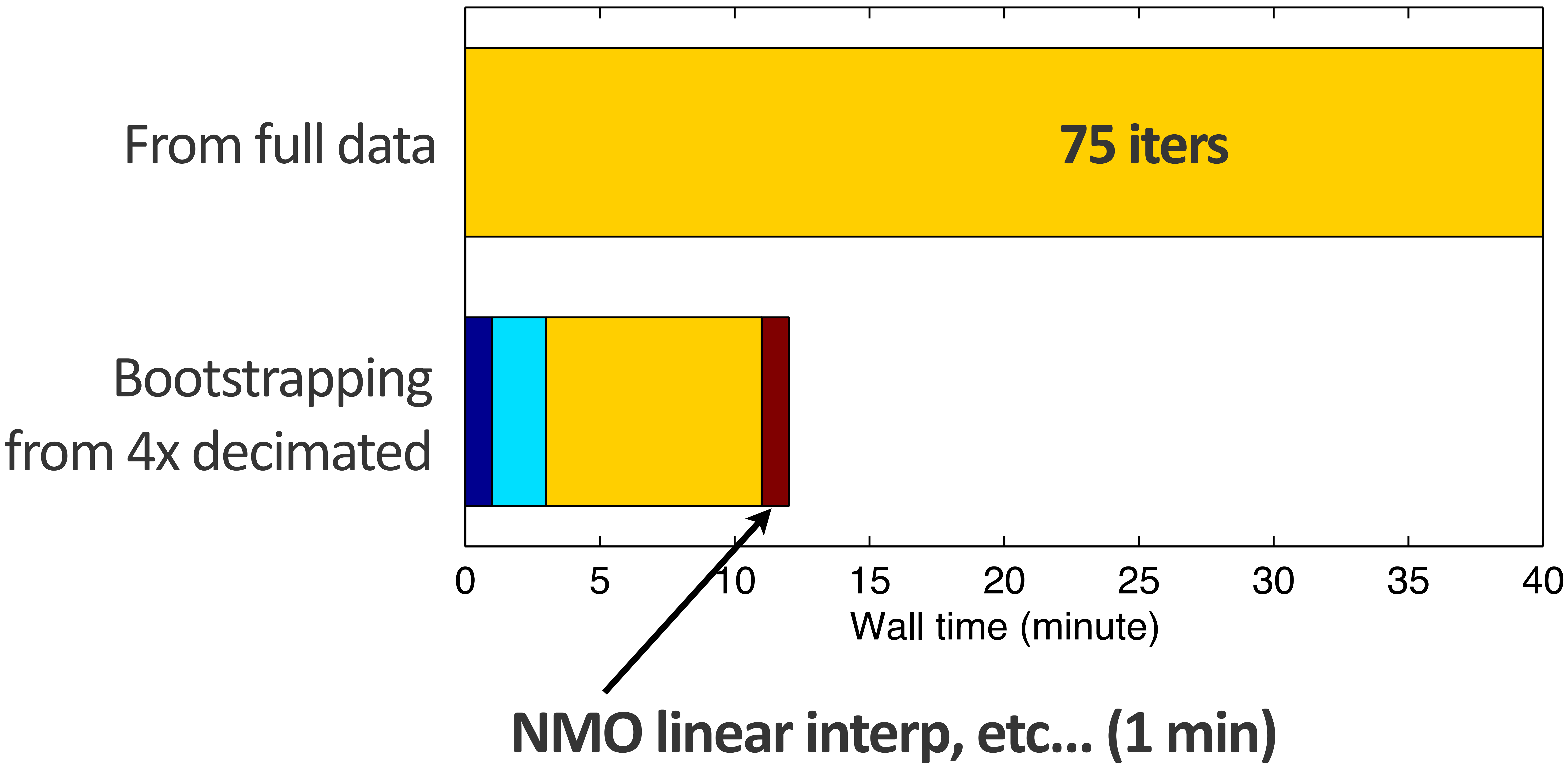
25 iters at 2x decimated spatial sampling (2 min)

Significant speedup from bootstrapping
Wall times

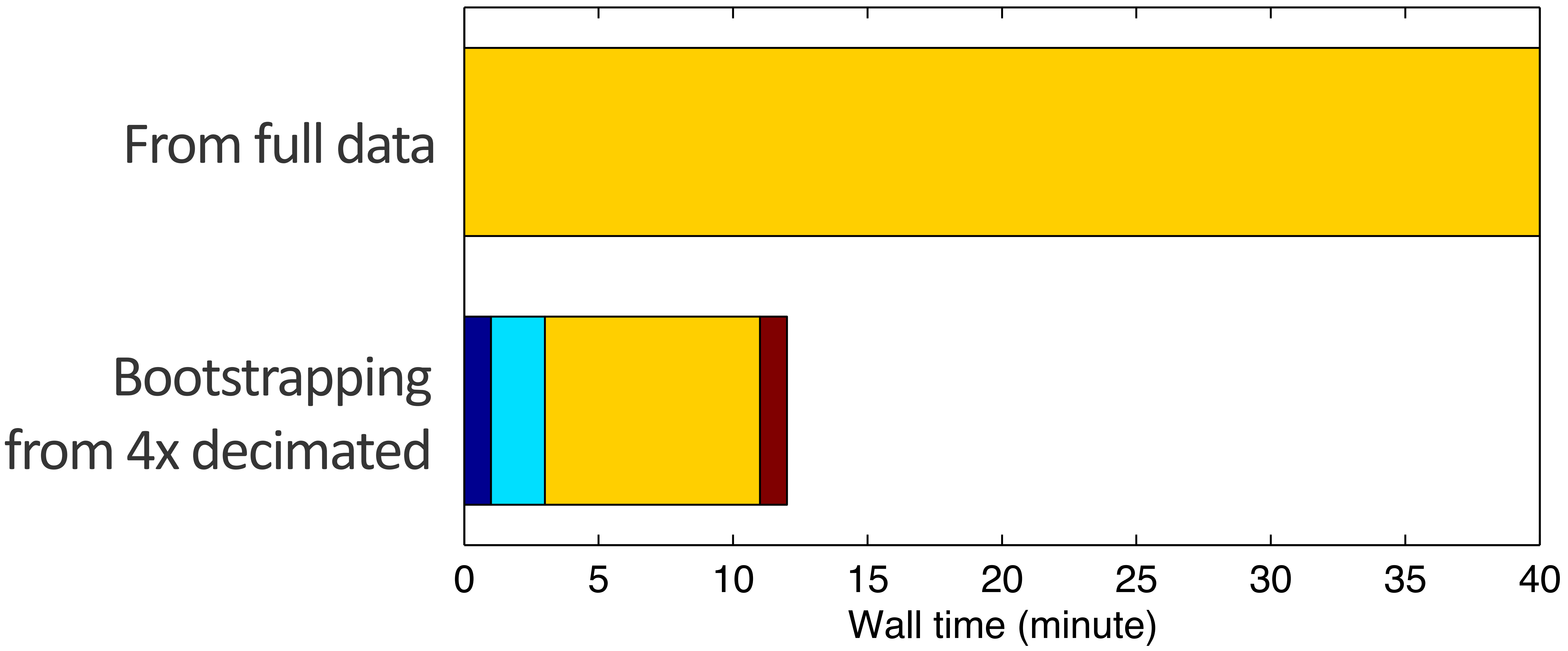


15 iters at full problem size w/ all data (8 min)

Significant speedup from bootstrapping
Wall times

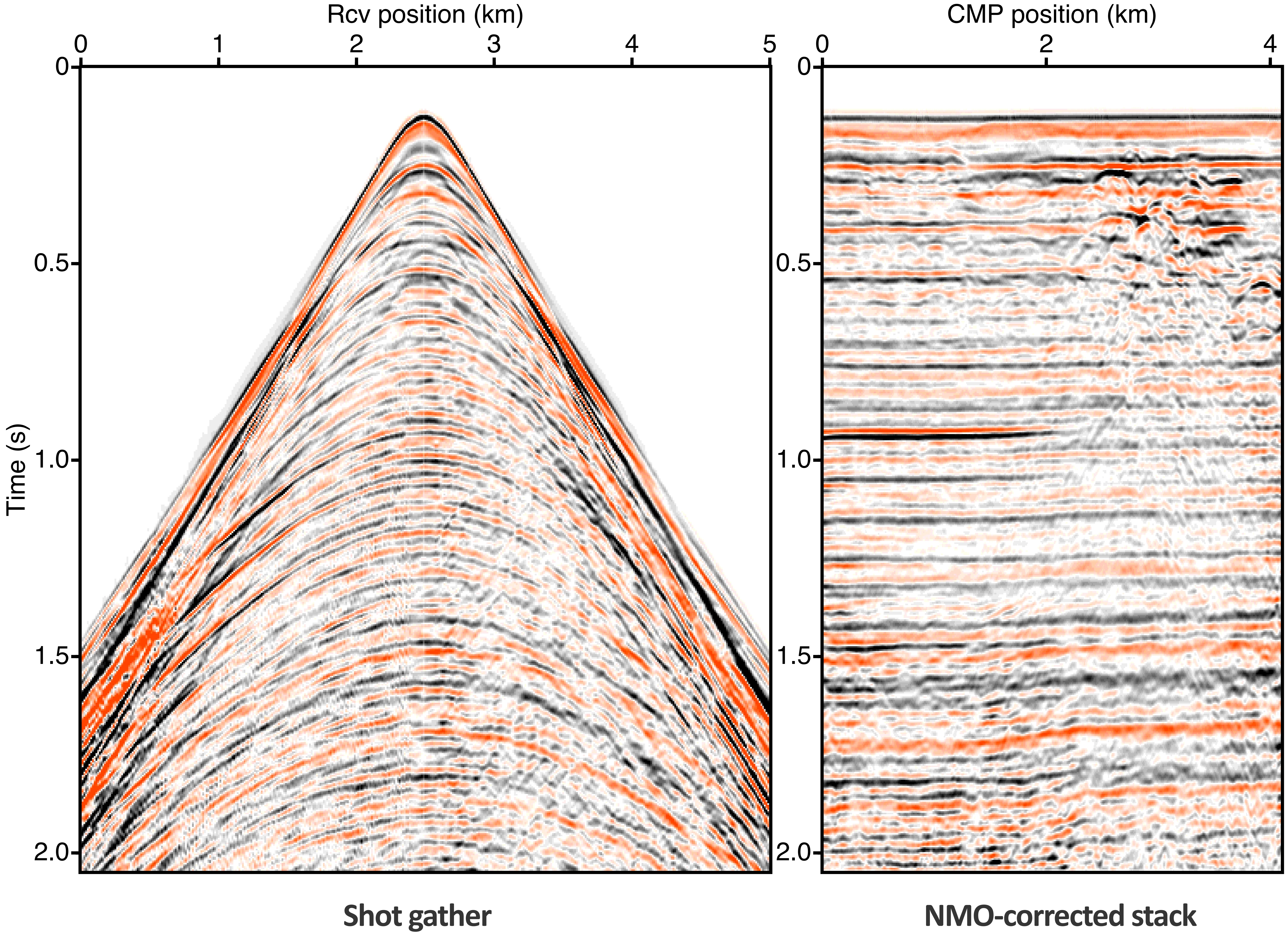


Significant speedup from bootstrapping
Wall times



Field data example

North Sea dataset



North sea data

Shot gather and stack

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

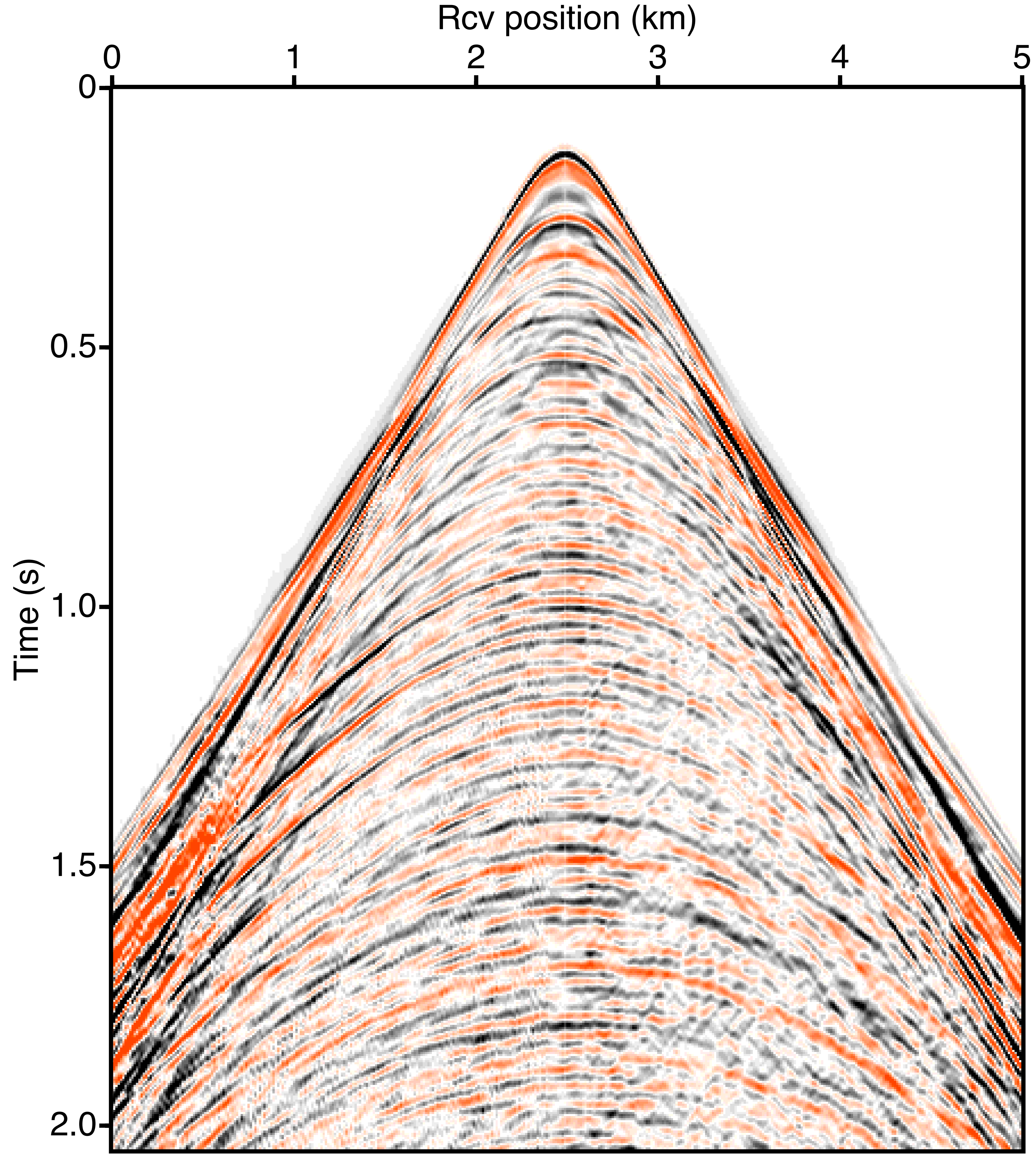
401 source and reciever

12.5 m spatial grid

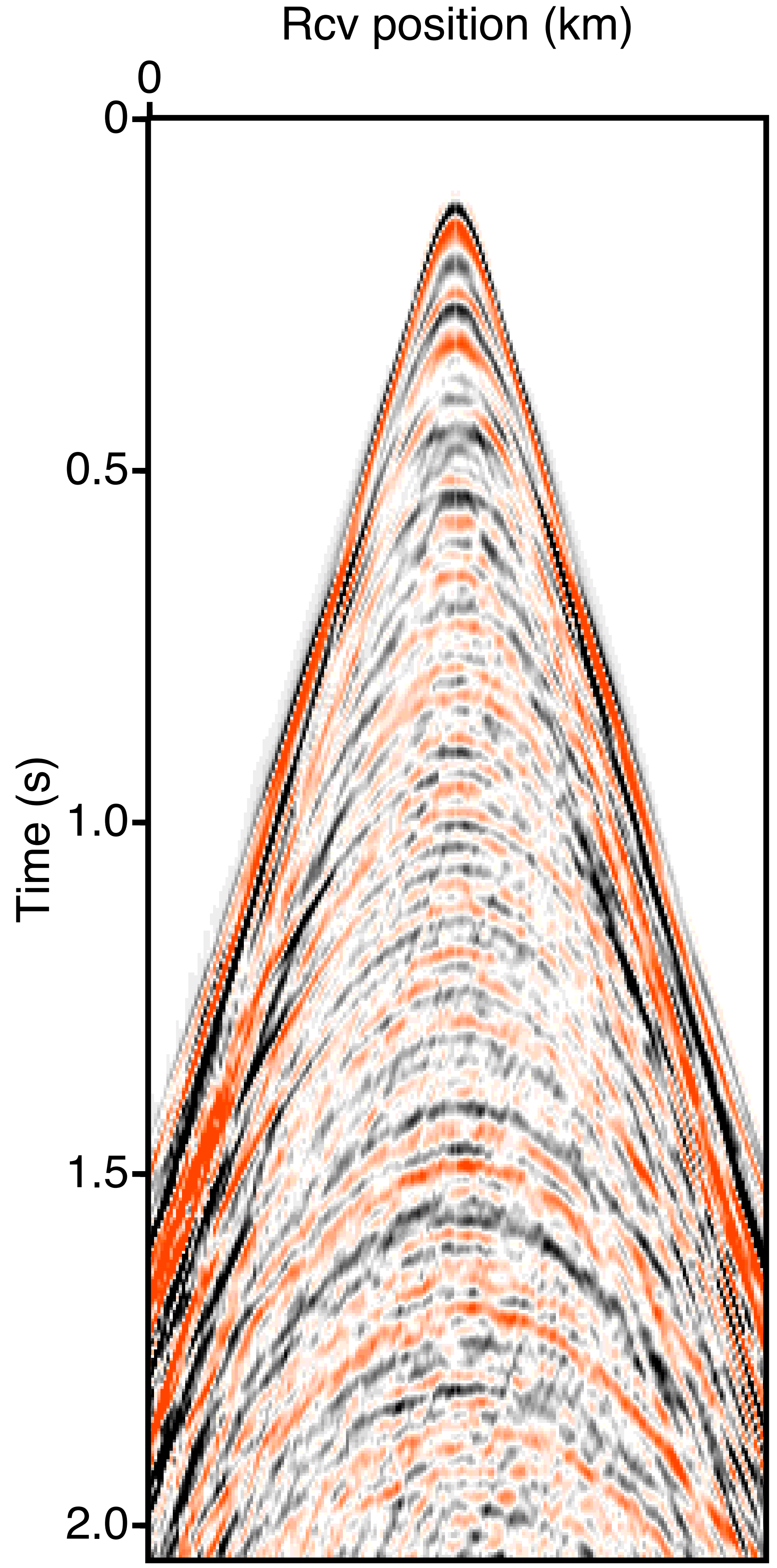
4 ms time sampling

Decimated wavefields

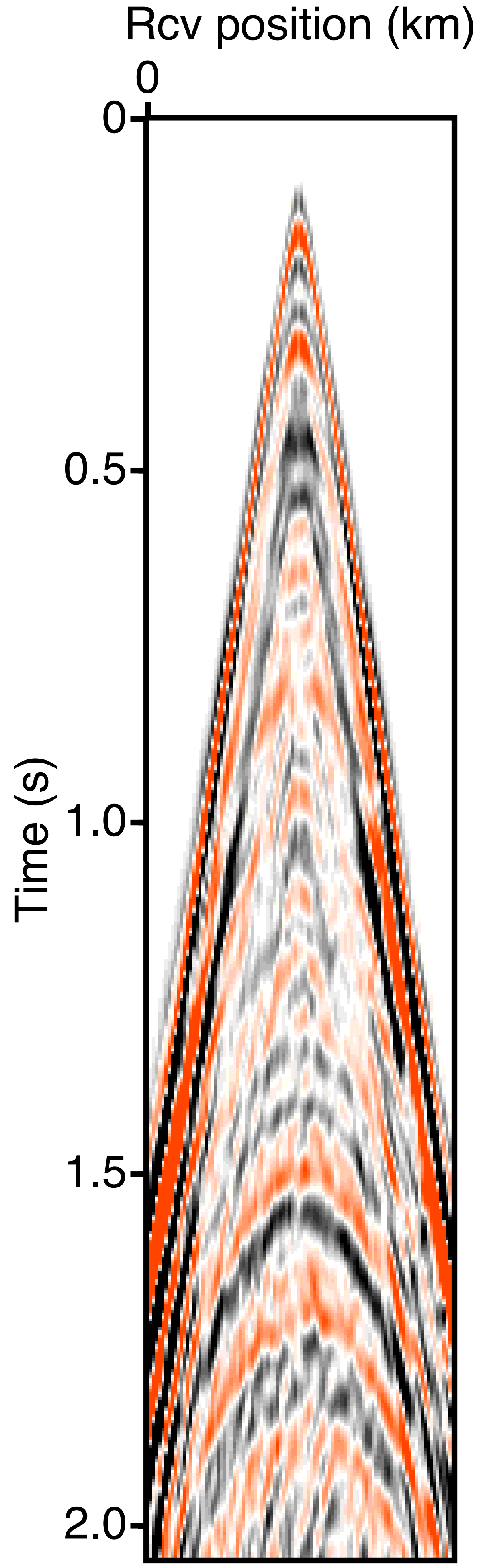
Original (dx = 12.5m)



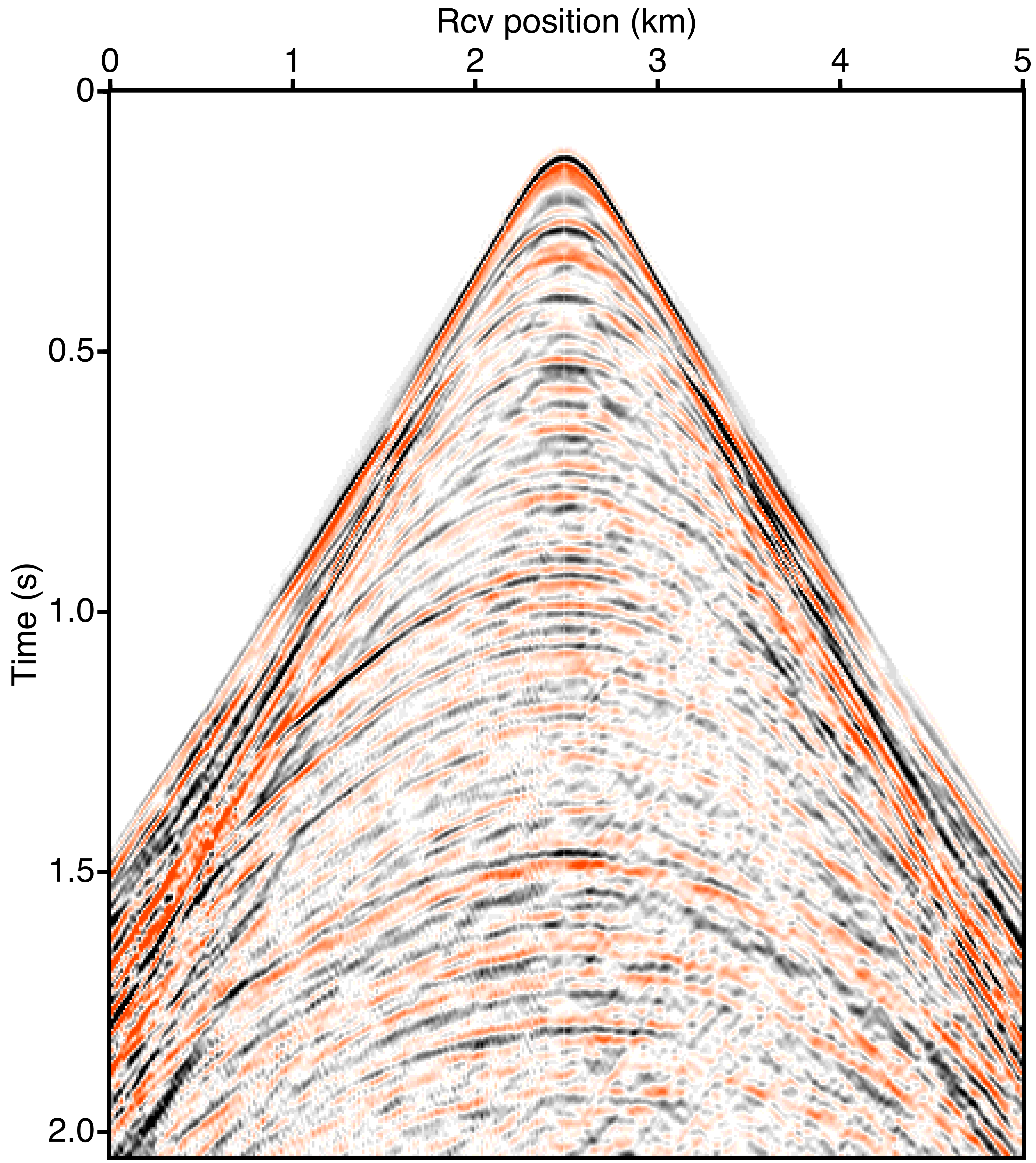
2x decimated
lowpass 40Hz



4x decimated
lowpass 20Hz



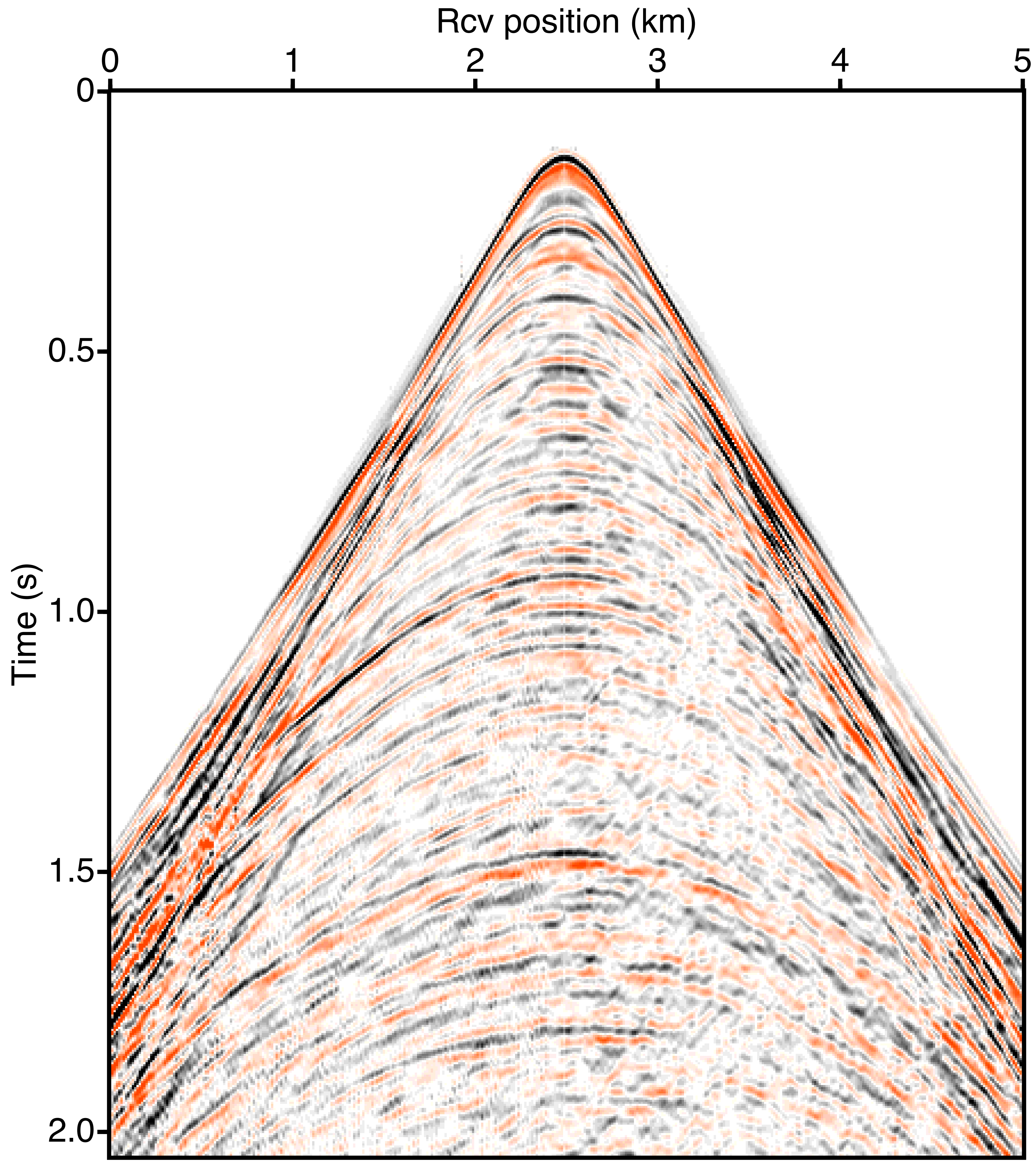
Solution wavefield comparison



Direct Primary

Solved with plain algorithm from finest scale data

Solution wavefield comparison

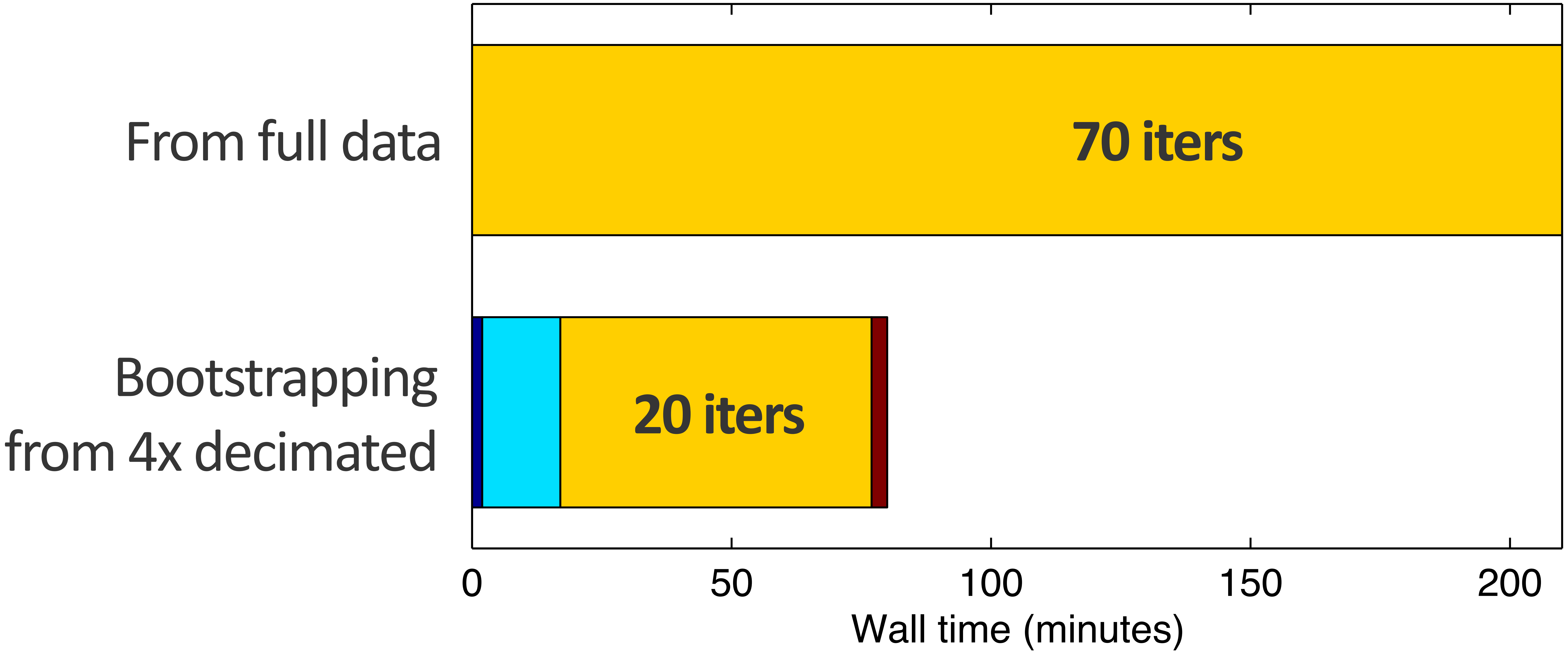


Direct Primary

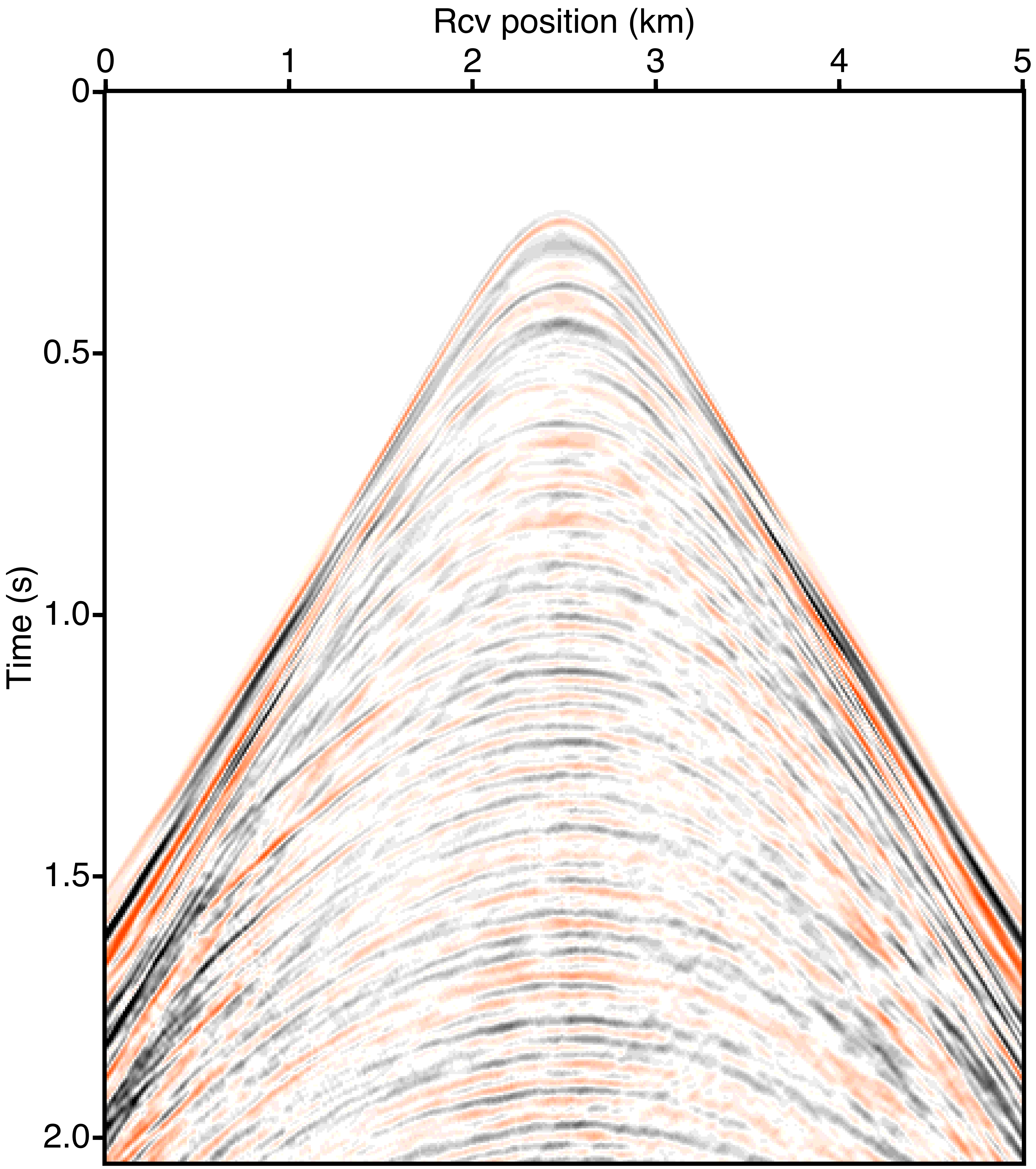
Solved with spatial sampling continuation

dx = 50m > 25m > 12.5m

Runtime breakdown
(wall time)



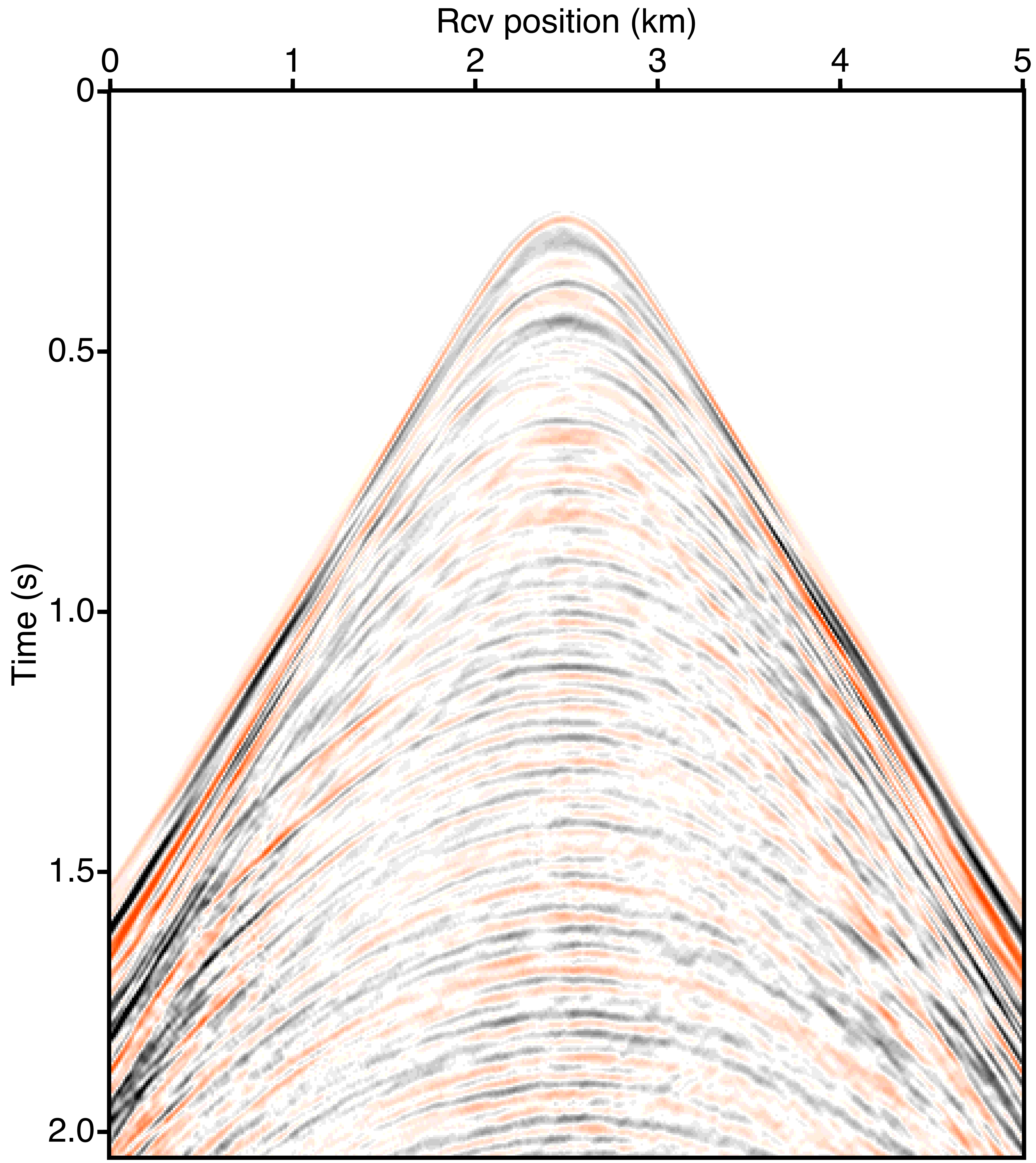
Solution multiple comparison



Predicted Surface Multiple

Solved with plain algorithm from finest scale data

Solution multiple comparison



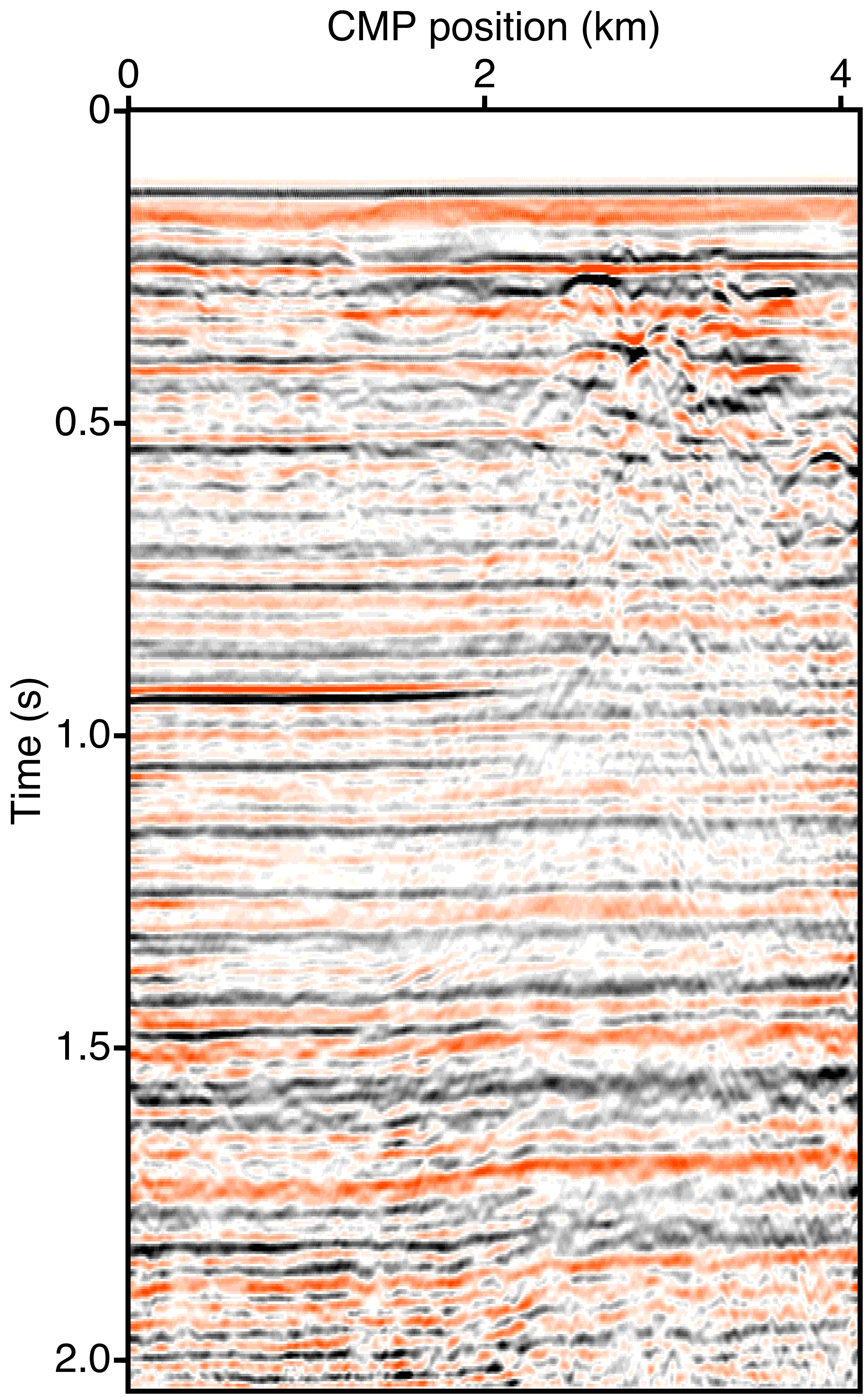
Predicted Surface Multiple

Solved with spatial sampling continuation

dx = 50m > 25m > 12.5m

Solution stack comparison

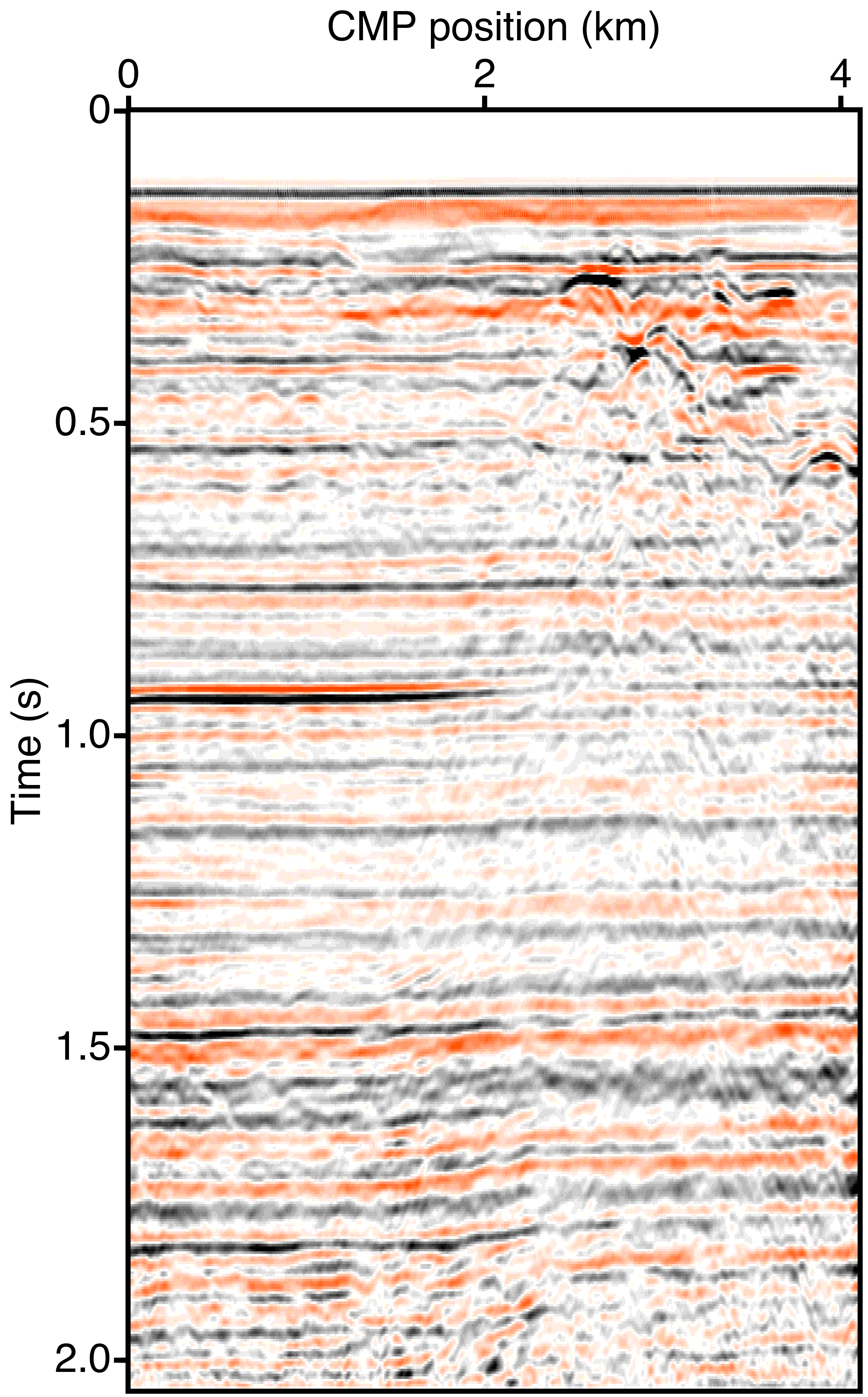
NMO Stack
original data



Solution stack comparison

REPSI Primaries NMO Stack

Solved with plain algorithm from finest scale data

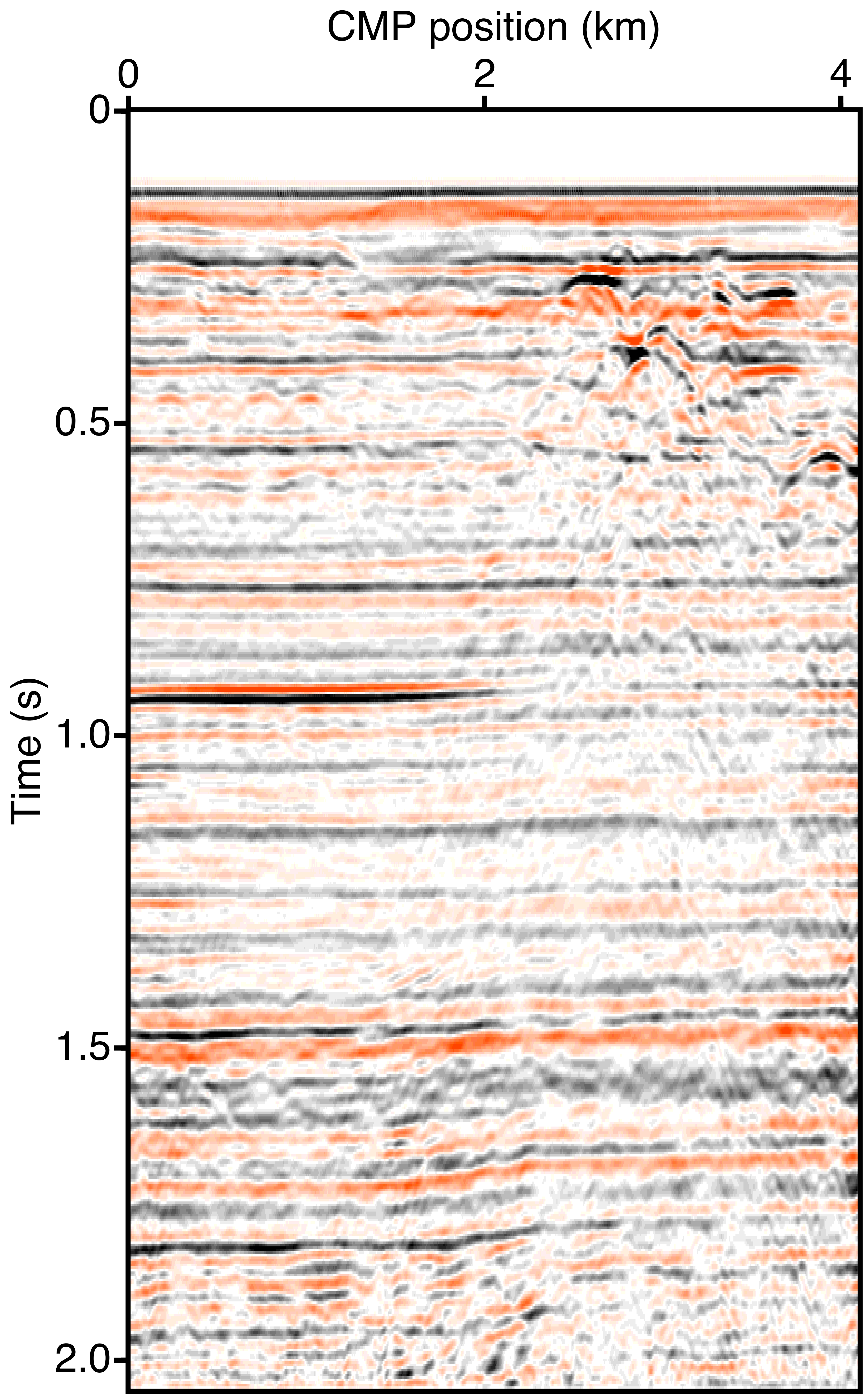


Solution stack comparison

REPSI Primaries NMO Stack

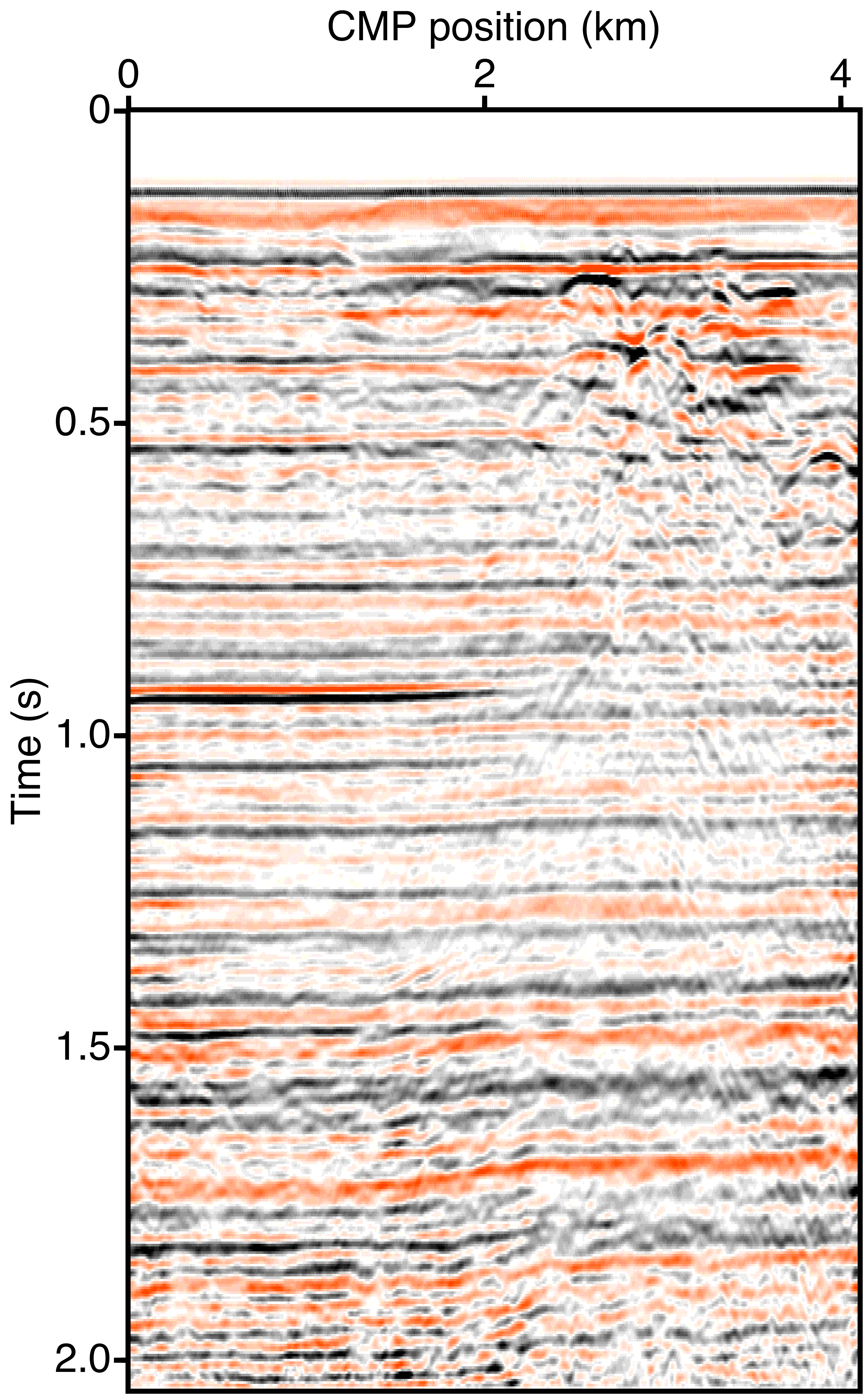
Solved with spatial sampling continuation

dx = 50m > 25m > 12.5m



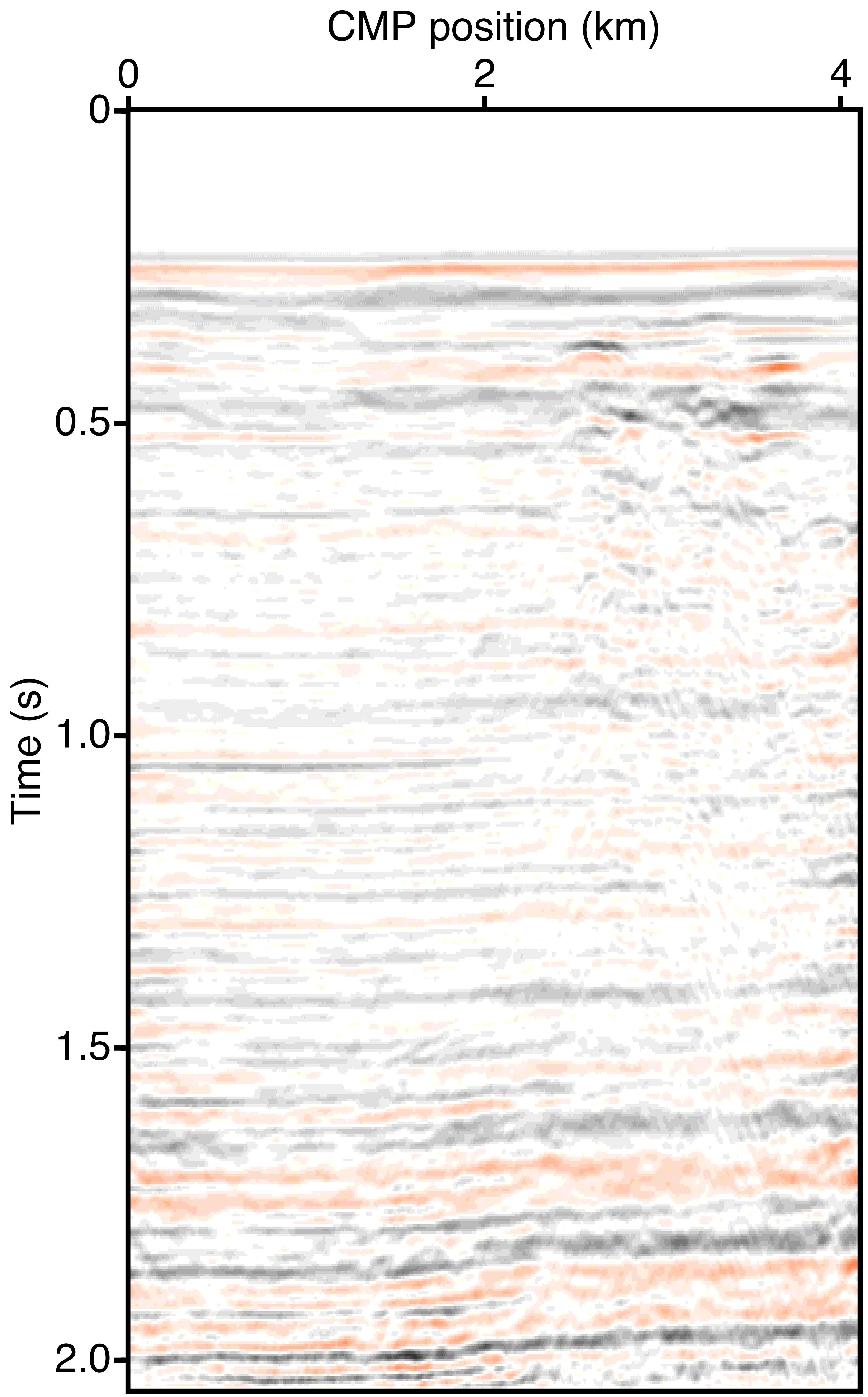
Solution stack comparison

NMO Stack
original data



Solution stack comparison

REPSI Multiples NMO Stack
Solved with plain algorithm from finest scale data

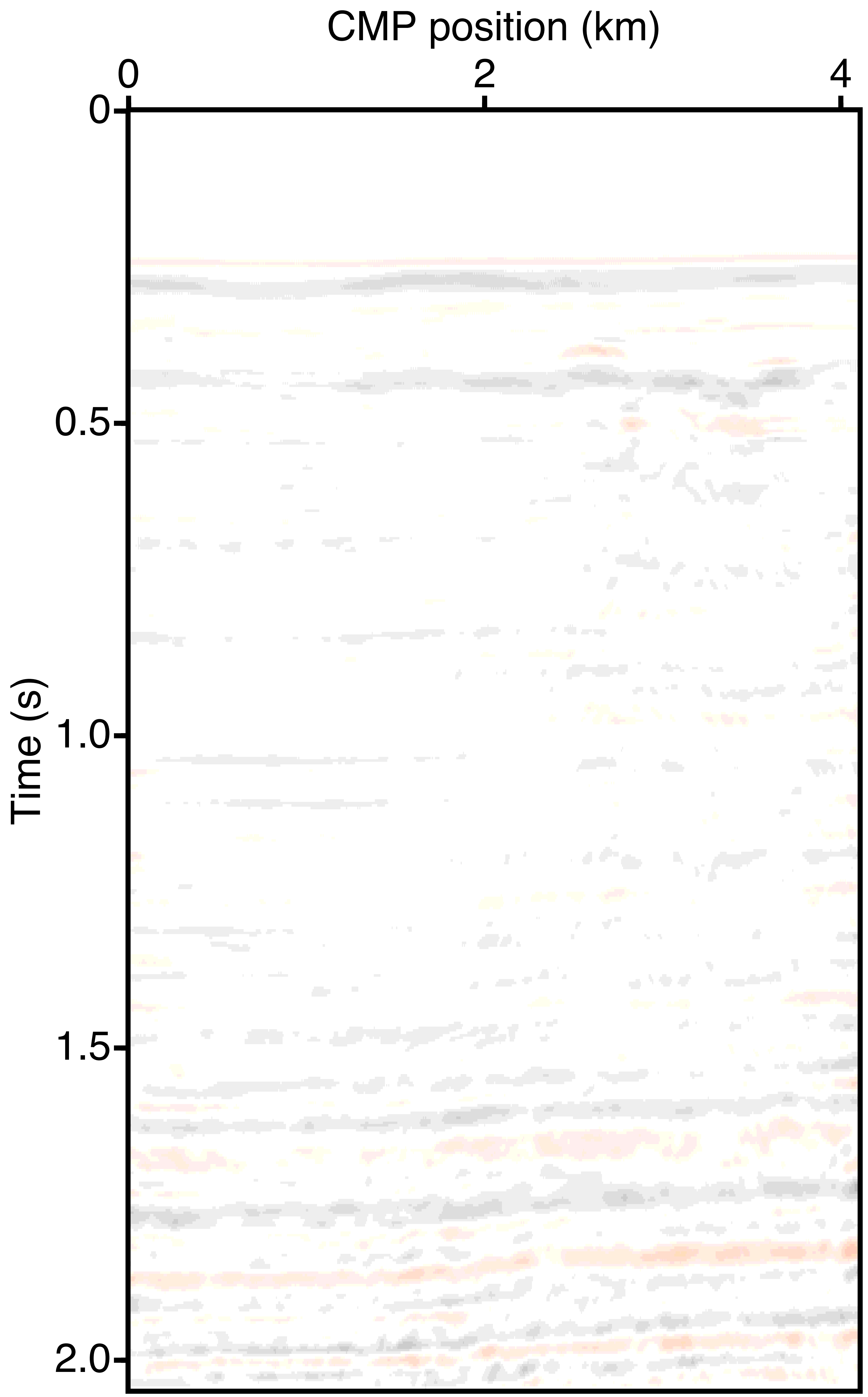


SLIM

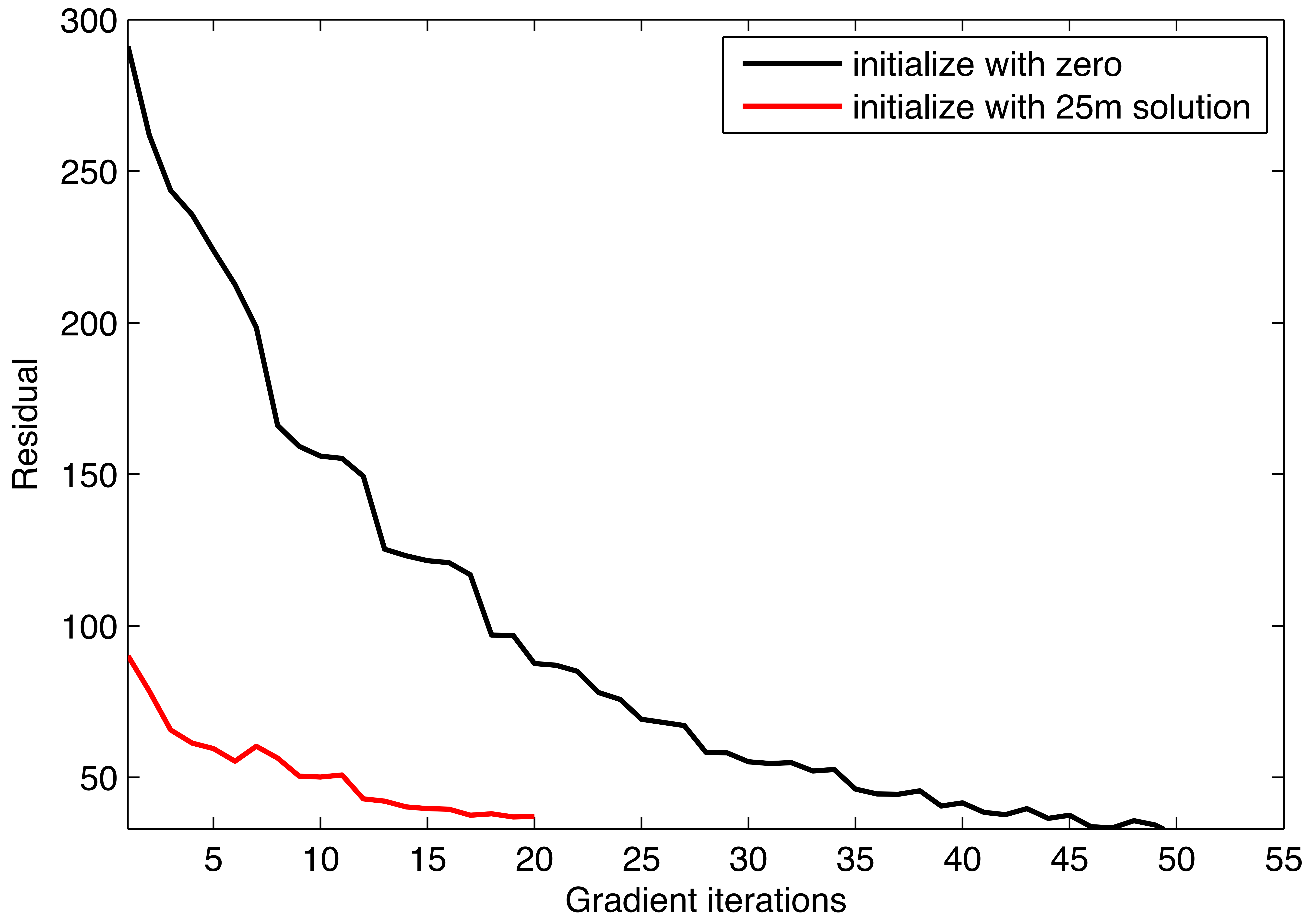
Solution stack comparison

REPSI Multiples NMO Stack

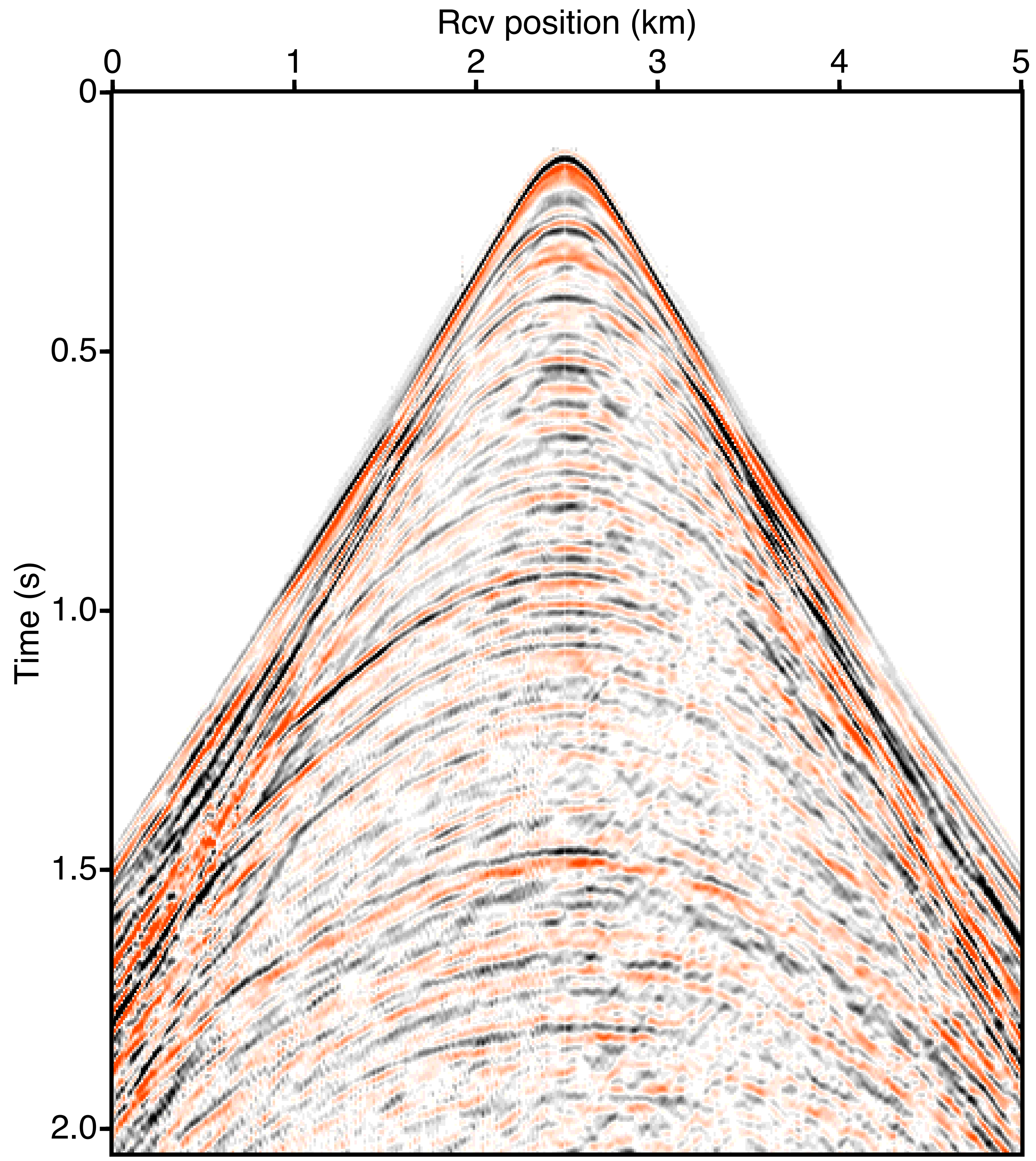
Difference:
- plain algorithm
- accelerated algorithm

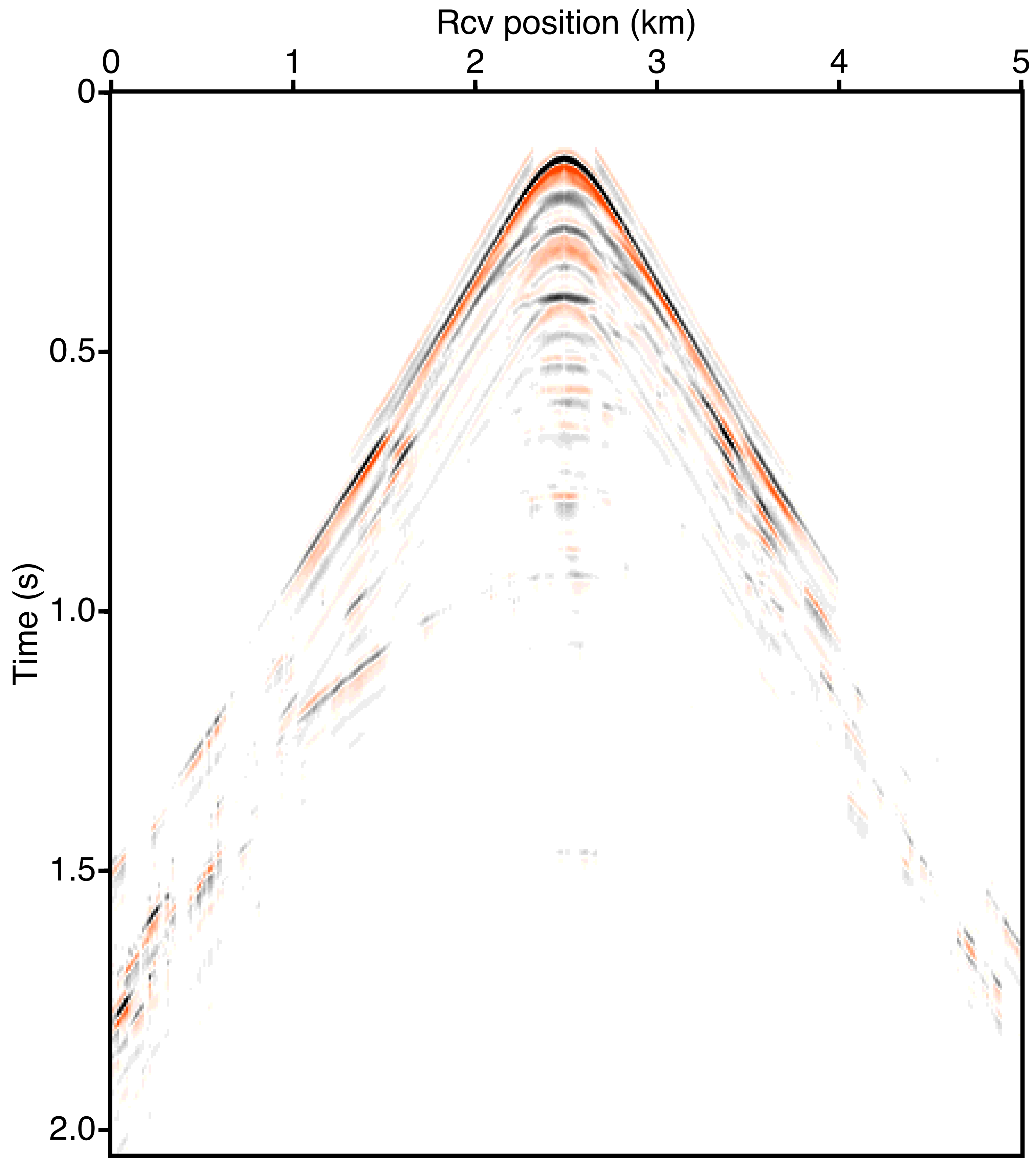


Warm-start vs from zero residual graph
(for full scale problem)



Warm-start vs from zero 'G' shot gathers





Acceleration strategy summary

Start REPSI with *decimated* data, *lowpass* to avoid spatial aliasing

Once "enough" progress is made, continue with fine-scale data

Significant savings in computation cost, 100x to 200x SRMP becomes more like 20x to 30x

How low can we go? Depends on the ability of sparsity-regularized inversion to resolve wavefronts under reduced bandwidth.

Acknowledgements

- Eric Verschuur and the DELPHI team
- PGS for permission to use the field dataset

SINBAD

NSERC CRSNG

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