Challenges and *opportunities* in sparse wavefield inversion

Felix J. Herrmann

SLIM

Seismic Laboratory for Imaging and Modeling
the University of British Columbia

Motivation

Many successful seismic algorithms are *data*-driven

- Surface-related Multiple Elimination (SRME)
- Estimation of Primaries by Sparse Inversion (EPSI)
- Interferometric deconvolution

Require

- *dense* matrix vector *multiplies*
- *full* (azimuth) sampling
- *memory* and *matvec* make *scaling* to 3-D very *challenging*
- certainly in the light of push for more & more data

Goals

Use *redundant* information residing in *multiple* reflections.

Exploit *data*-space transform-domain *sparsity* & *low rank* to

- *stabilize* wavefield *inversion*
- *reduce system sizes* & *mitigate* cross-talk

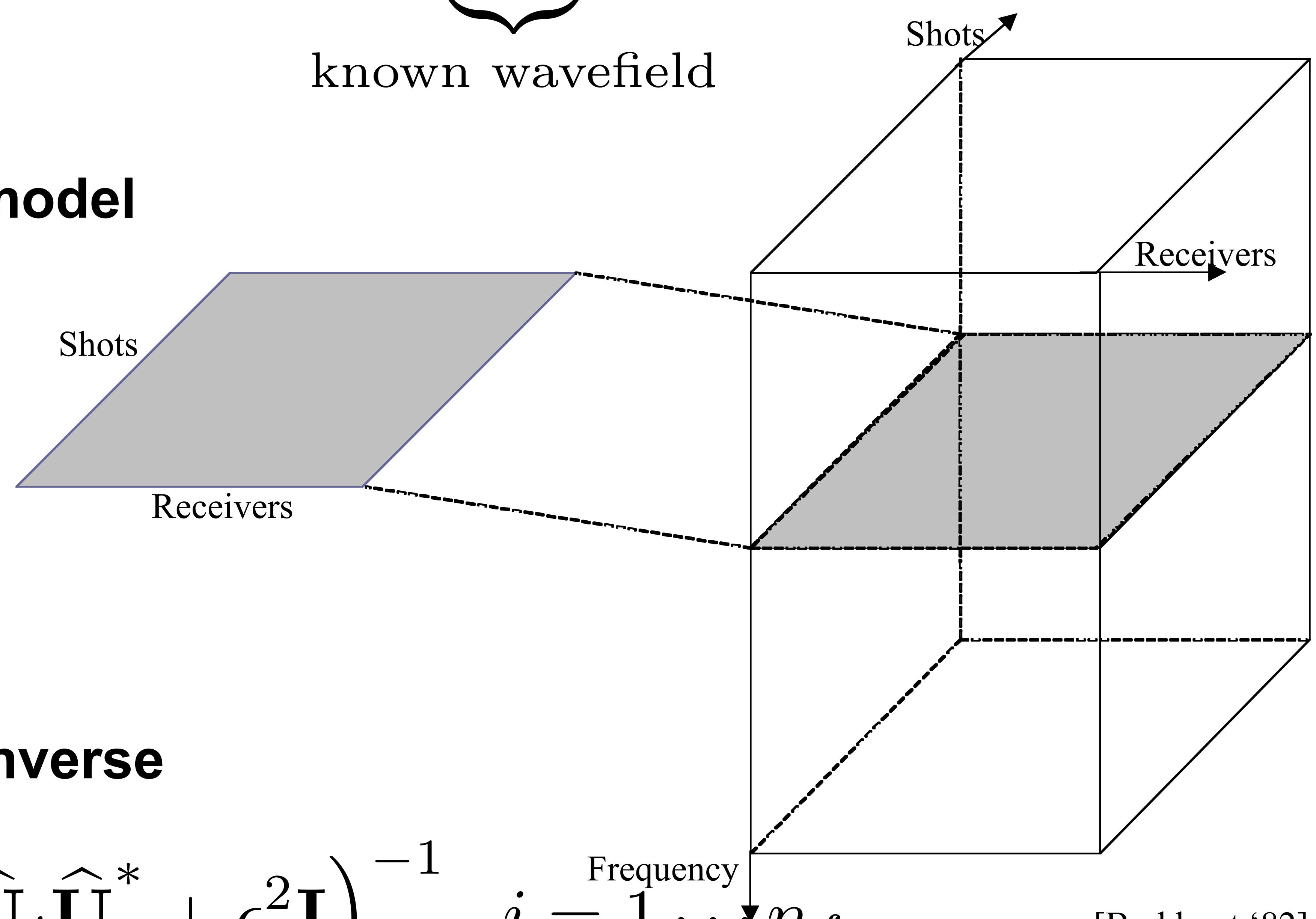
Exploit *adaptive model*-space transform-domain *sparsity* to

- compute *convolutions/correlations via* wave simulators
- *reduce system sizes* & *mitigate* cross-talk

“Typical” approach: damped least-squares

$$\underbrace{\hat{\mathbf{G}}_\mathbf{i}}_{\text{unknown wavefield}} \qquad \overbrace{\hat{\mathbf{U}}_\mathbf{i}}^{\text{known wavefield}} \qquad = \qquad \underbrace{\hat{\mathbf{V}}_\mathbf{i}}_{\text{known wavefield}} \qquad i = 1 \cdots n_f$$

Monochromatic **forward model**

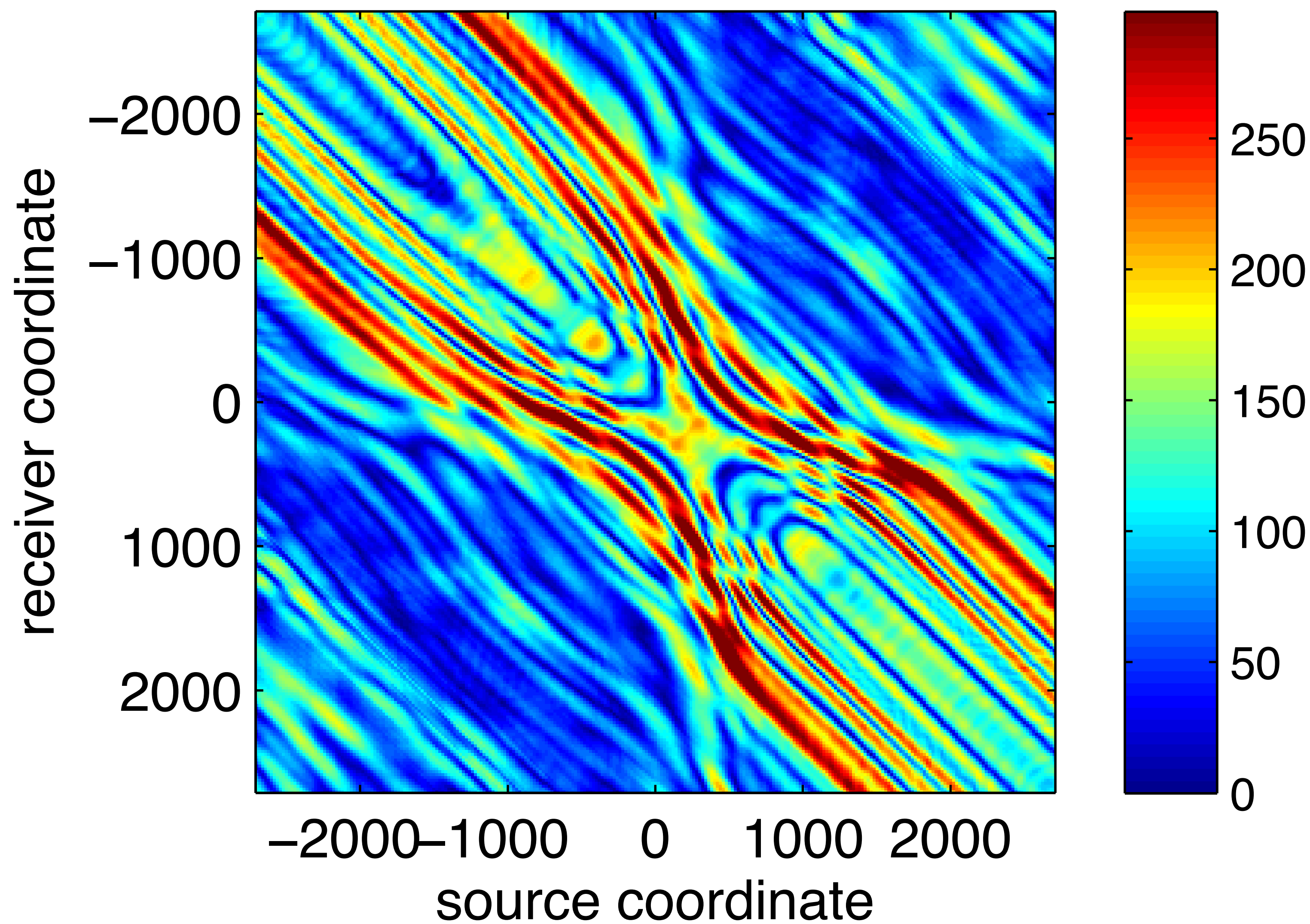


Monochromatic **pseudo-inverse**

$$\widetilde{\widehat{\mathbf{G}}}_i = \widehat{\mathbf{V}}_i \widehat{\mathbf{U}}_i^* \left(\widehat{\mathbf{U}}_i \widehat{\mathbf{U}}_i^* + \epsilon_i^2 \mathbf{I} \right)^{-1}, \; i = 1 \cdots n_f,$$

[Berkhout ‘82]
[F.J.H ’07-’08]
[Wapenaar ‘08]

Data matrix



- Inversion is carried out per frequency slice
- Water level leads to loss in resolution
- Can suffer from instabilities ...

Two- vs one-norms

Two-norm inversion:

- tends to *smooth* when *regularizing* the *null* space
- *ineffective* when *dealing* with *cross talk* induced by *randomized* sourcing (e.g., simultaneous)

One-norm *sparsity*-promoting inversion:

- leverages curvelet-domain *sparsity* of data
- highly *effective* for *removal* of source *crosstalk*
- *preserves* frequency *content*

Solutions

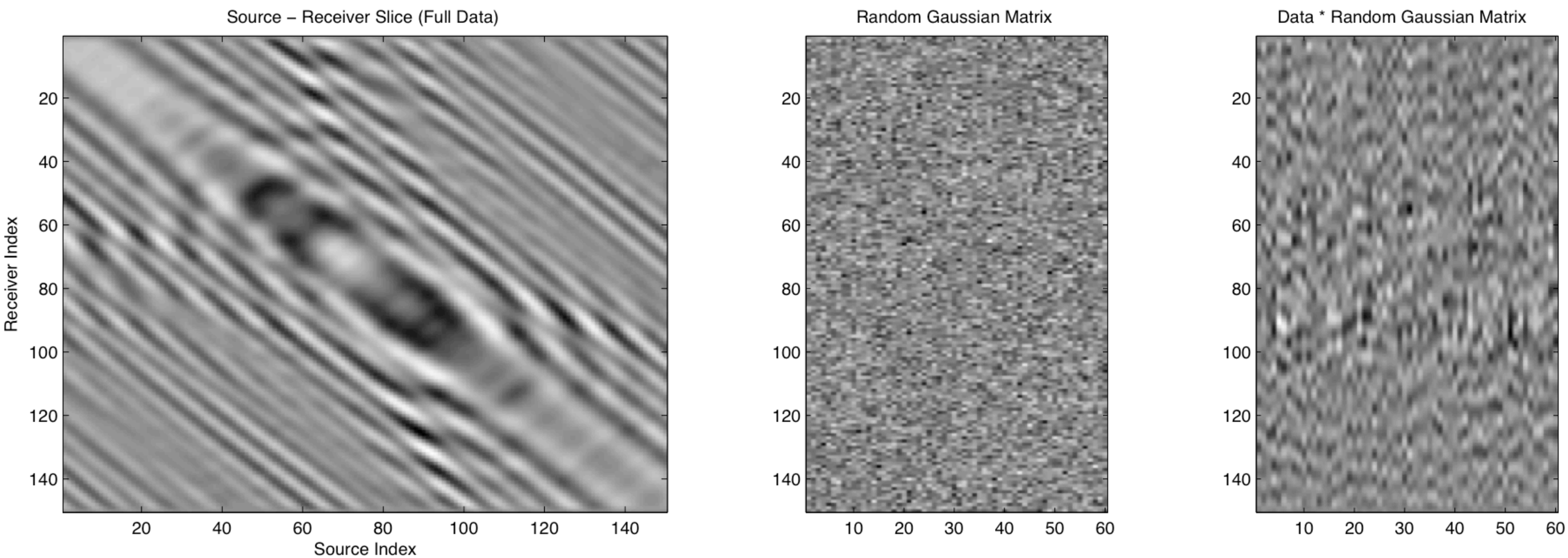
1. *Simultaneous* sourcing in *combination* with *renewals*
   - *reduces* # of shots & leverages *sparsity*-promoting solvers
2. *Wave-equation* based possibly in *combination* with 1.
   - leverages *sparse* wave simulators to carry our multi-D convolutions implicitly
3. Low-rank approximations
   - exploit multi-D structure of seismic wavefields

Randomized source encoding = compressive sensing

$$\left[\boldsymbol{\delta}\mathbf{d}_1, \cdots, \boldsymbol{\delta}\mathbf{d}_{n_s}\right] \quad \mathbf{W} \quad \left[\underline{\boldsymbol{\delta}\mathbf{d}}_1, \cdots, \underline{\boldsymbol{\delta}\mathbf{d}}_{n_s'}\right]$$

or

$$\left(\mathbf{W}^H \otimes \mathbf{I}\right) \mathrm{vec}(\boldsymbol{\delta}\mathbf{D}) = \mathrm{vec}(\underline{\boldsymbol{\delta}\mathbf{D}}) \qquad n_s' \ll n_s$$



Approach I

Work with *simultaneous* sources, i.e.,

$$\widehat{\mathbf{V}}\mathbf{W} = \widehat{\mathbf{G}}\widehat{\mathbf{U}}\mathbf{W} \quad \text{with} \quad \mathbf{W} \in \mathbb{C}^{n_s \times n_s'}, n_s' \ll n_s$$

- reduces system size but needs full acquisition for $\widehat{\mathbf{U}}$
- could benefit from *redrawing* simultaneous shots
- $\widehat{\mathbf{G}}$ is '*dense*' & *redundant* in *sparsifying* domain
- still high matvec and storage costs

[Guitton, '02; Berkhout, '05; Whitmore '10; Ning et. al., '10-; Verschuur '11]

Approach II

Wave-equation based:

$$\widehat{\mathbf{G}}\widehat{\mathbf{U}} = \widehat{\mathbf{F}}[\mathbf{m}, \widehat{\mathbf{U}}]$$

with

$$\widehat{\mathbf{F}}[\mathbf{m}, \widehat{\mathbf{Q}}] = \mathbf{R}\widehat{\mathbf{H}}^{-1}[\mathbf{m}]\mathbf{R}^*\widehat{\mathbf{Q}}$$

yielding

$$\widehat{\mathbf{F}}[\mathbf{m}, \mathbf{I}\,]\widehat{\mathbf{U}} = \widehat{\mathbf{F}}[\mathbf{m}, \widehat{\mathbf{U}}]$$

Wave simulator does *heavy* lifting for the multi-D convolutions!

Approach II cont'd

After linearization for EPSI we have

$$\widehat{\mathbf{P}} \approx \widehat{\nabla\mathbf{F}}[\mathbf{m}, \widehat{\mathbf{Q}} - \widehat{\mathbf{P}}]\delta\mathbf{m}$$

and *simultaneous* sourcing

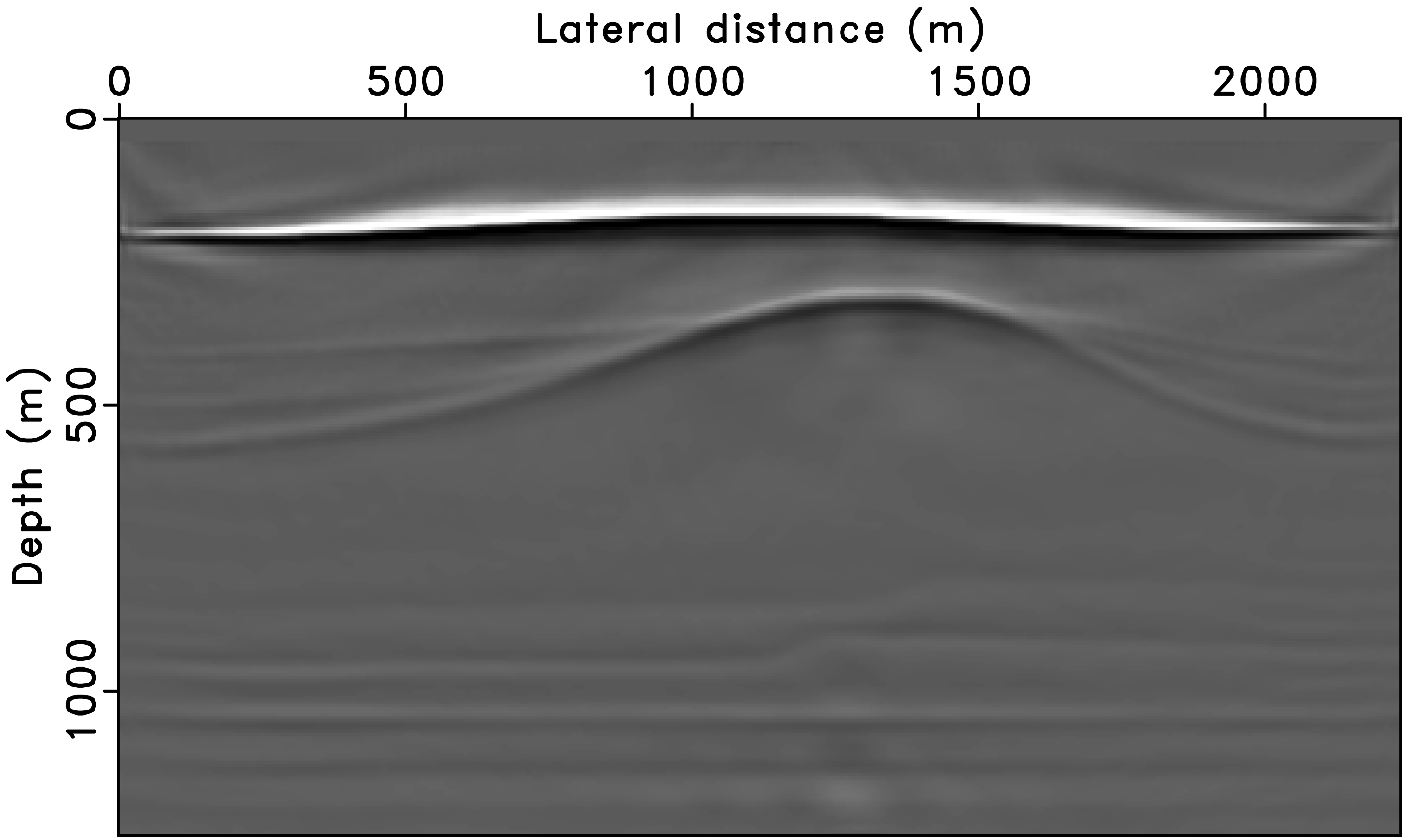
$$\widehat{\mathbf{P}}\mathbf{W} \approx \widehat{\nabla\mathbf{F}}[\mathbf{m}, (\widehat{\mathbf{Q}} - \widehat{\mathbf{P}})\mathbf{W}]\delta\mathbf{m}$$

Highly *efficient* formulation that

- *reduces* # of PDE solves
- is *conducive* to image-domain *sparsity*-promotion

Migration from marine '*simultaneous*' data

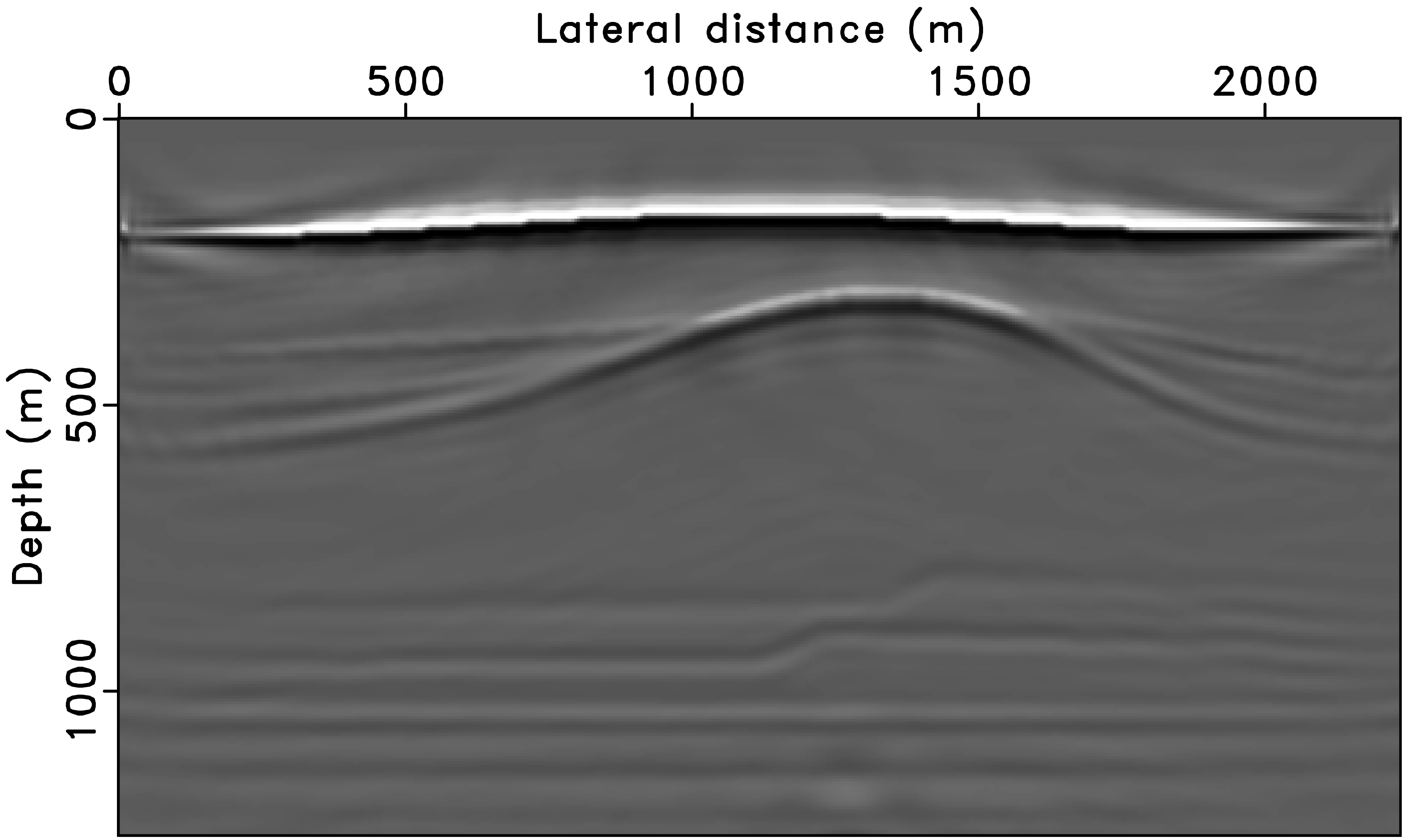
inversion from *EPSI inverted* Green's function



Migration from marine '*simultaneous*' data

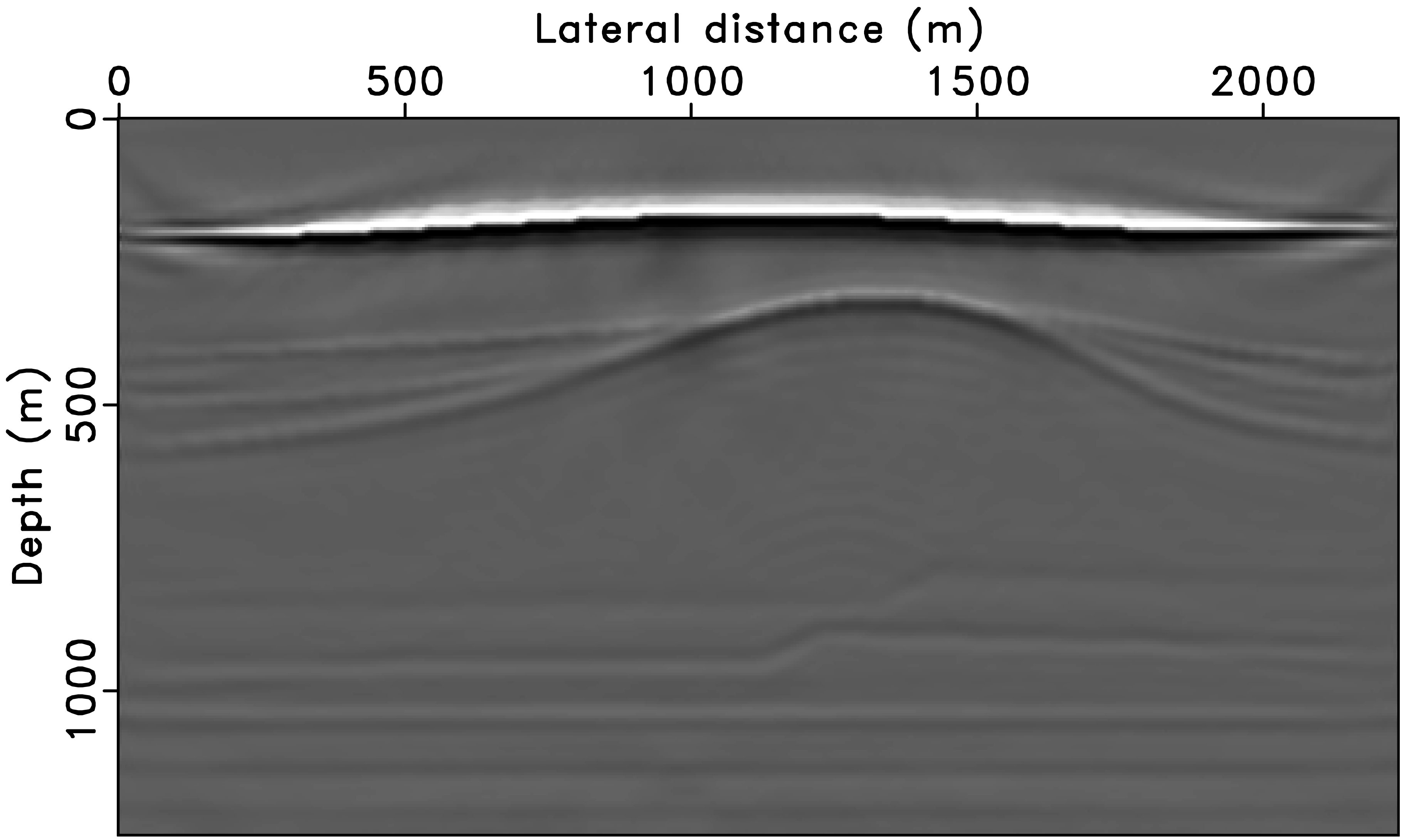
inversion from *total* data

SLIM



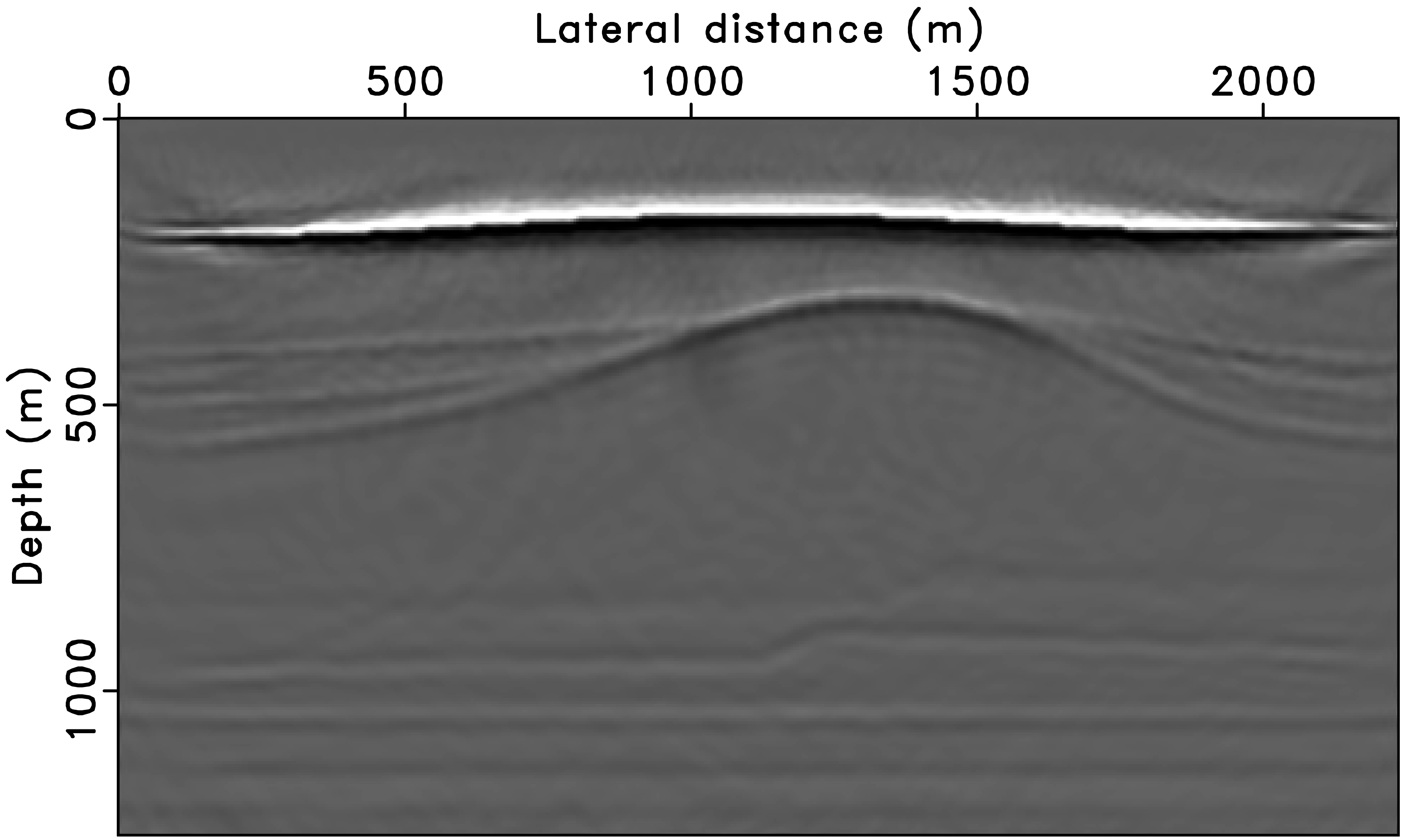
Migration from complete data with source-encoding

inversion from *total* data, 10 super-shots



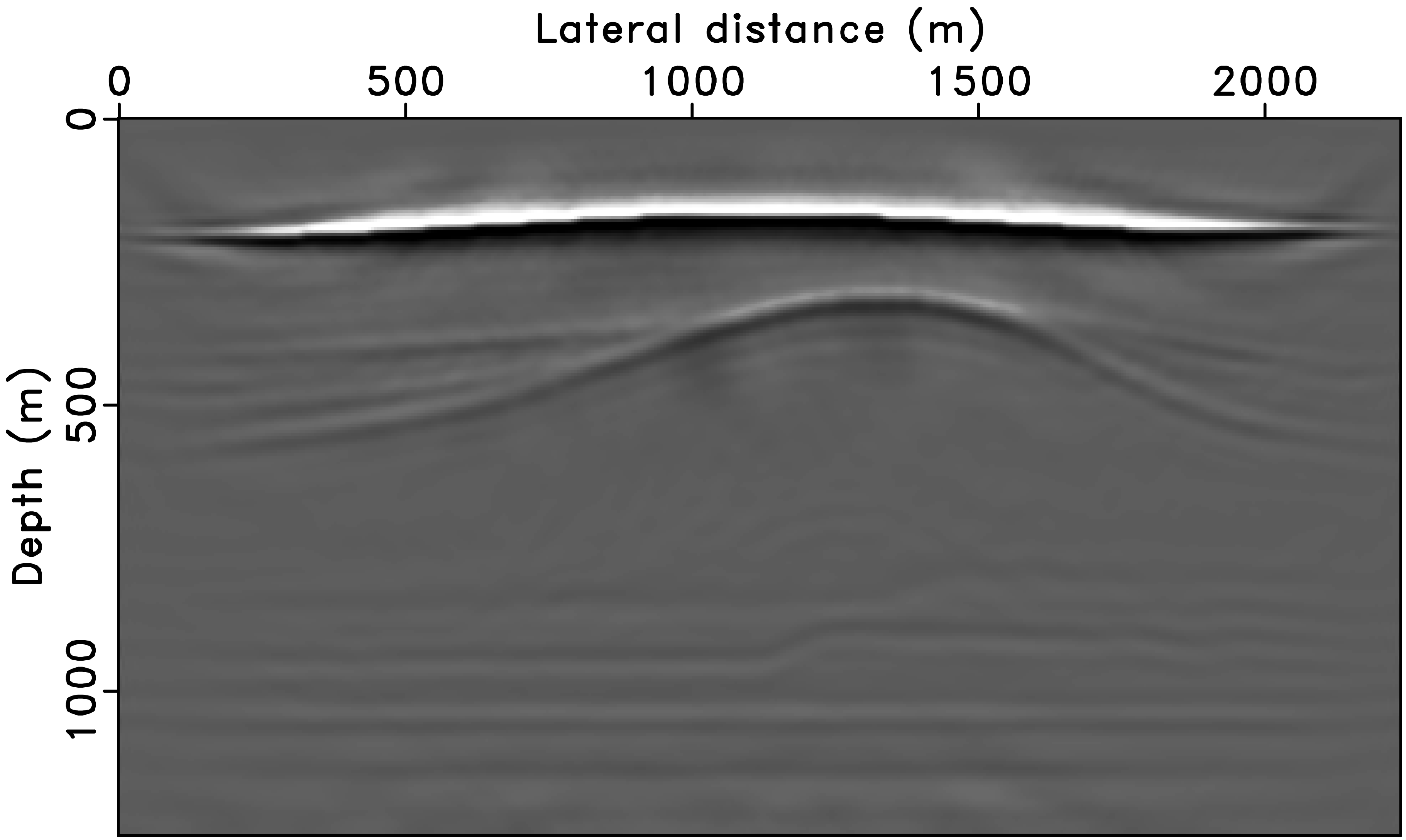
Migration from complete data with source-encoding

inversion from *total* data, 2 super-shots, no renewal



Migration from complete data with source-encoding

inversion from *total* data, 2 super-shots, renewal



Observations

Combination of *EPSI* or *interferometric* deconvolution

with *imaging* via *areal* sources allows us to

- exploit *image*-domain *sparsity* & *information* from *multiples*
- do multi-D convolutions/correlations with wave solver

Costs and *reliance* on *full* sampling can be brought down by

- simultaneous sourcing, random time dithering, or a combination thereof
- but *adaptive* method *requires velocity* information

Approach III

With 'black-box' access to matvecs
(SRME multiple prediction including on-the-fly interpolation)

- use *randomized* SVDs allow us to do a low-rank approximation to factorize

$$\widehat{\mathbf{P}} \approx \widehat{\mathbf{L}}\widehat{\mathbf{R}}^T$$

- reduces memory imprint and matvec costs
- allows us to conduct *velocity* analysis
- but requires 'full' data

Opportunity

Adapt *recent* matrix-*completion* techniques with *maxnorms*

Allows us to estimate low-rank approximations from incomplete directly data by solving

$$\underset{\mathbf{L},\mathbf{R}}{\text{minimize}} \, \|\mathbf{b} - \boldsymbol{\mathcal{A}}(\mathbf{L}\mathbf{R}^*)\|_2^2 + \mu \|\mathbf{L}\mathbf{R}^*\|_*$$

- *nuclear* norm is *approximated* by *maxnorm*
- opens *possibility* to *scale* to 3D
- *challenge* is to find *appropriate* low-rank 'domain' (e.g., midpoint/offset)

SLIM

Observations

‘Multiples *facilitate recovery* from *severe* undersamplings.

Large data volumes *impede* data-space recovery and *require* exploration of other types of structure.

Image-domain wave simulators can *carry* the weight of ‘data-driven’ *approaches*

- and *really* shine with *simultaneous* sources & renewals
- but require velocity-model information

Q: relationship free surface BC & EPSI-like techniques?