Compressive imaging by wavefield inversion with group sparsity

Felix J. Herrmann

SLIM

Seismic Laboratory for Imaging and Modeling
the University of British Columbia

Drivers

We are *no* longer finding oil...

Seismic imaging & inversion

- from *noisier* and *incomplete* data
- at reduced computational costs
- with improved resolution

Impediments

Seismic-data volumes are *extremely* large

Least-squares migration & *formation* of *image* volumes are *prohibitively* expensive...

[Symes, '08]

SLIM

Impediments

Full-waveform inversion is suffering from

- multimodality, i.e,. a multitude of velocity models explain data
- local minima
- is both over- and undetdetermined

Today's talk

Leverage insights from

- Randomized matrix multiplies and linear regressions, and
- from Compressive Sensing (CS)

for computation of ***image volumes***

Exploit *physical principle* of ***focusing***...

SLIM

Motivation

Faster shot-record depth migrations using phase encoding by Morton & Ober, '98.

Phase encoding of shot records in prestack migration by Romero et. al., '00.

How to choose a subset of frequencies in frequency-domain finite-difference migration by Mulder & Plessix, '04.

Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies by Sirque & Pratt, '04.

Compressed wavefield extrapolation by T. Lin and F.J.H, '07

Efficient Seismic Forward Modeling using Simultaneous Random Sources and Sparsity by N. Neelamani et. al., '08.

Migration Velocity Analysis and Waveform Inversion by Symes, '08

Compressive simultaneous full-waveform simulation by FJH et. al., '09.

[Lailly, '83]
[Tarantola, '84, '86, '87]
[Pratt and co-authors, '96, '98, '99, '03]

SLIM

PDE constrained optimization

$$\min_{\mathbf{U}\in\mathcal{U},\,\mathbf{m}\in\mathcal{M}} \frac{1}{2}\|\mathbf{P}-\mathbf{D}\mathbf{U}\|_2^2 \quad \text{subject to} \quad \mathbf{H}[\mathbf{m}]\mathbf{U}=\mathbf{Q}$$

+ Free surface BC

$\mathbf{P}$ = Total multi-source and multi-frequency data volume

$\mathbf{D}$ = Detection operator

$\mathbf{U}$ = Solution of the Helmholtz equation

$\mathbf{H}$ = Discretized multi-frequency Helmholtz system

$\mathbf{Q}$ = Unknown seismic sources

$\mathbf{m}$ = Unknown model, e.g. $c^{-2}(x)$

[Pratt et. al., '98]
[Plessix '06]

SLIM

Adjoint state

Implicit solves of Helmholtz system for each source $\mathbf{q}$

$$\mathbf{H}[\mathbf{m}]\mathbf{u} = \mathbf{q} \quad \text{and} \quad \mathbf{H}^*[\mathbf{m}]\mathbf{v} = \mathbf{r}$$

with

$$\mathbf{r} = \mathbf{D}^*(\mathbf{p} - \mathcal{F}[\mathbf{m}, \mathbf{q}])$$

and

$$\mathcal{F}[\mathbf{m}, \mathbf{q}] = \mathbf{D}\mathbf{H}^{-1}[\mathbf{m}]\mathbf{q}$$

[Plessix '06]

Gradient

Post-stack migration:

$$\delta\mathbf{m} = \Re\left(\sum_{\omega}\omega^2\sum_{s}(\bar{\mathbf{u}}\odot\mathbf{v})_{s,\omega}\right) = \mathbf{K}^*[\mathbf{m},\mathbf{Q}]\delta\mathbf{d}$$

with

$$\boldsymbol{\delta}\mathbf{d} = \mathrm{vec}(\mathbf{P} - \mathbf{D}\mathbf{H}^{-1}[\mathbf{m}]\mathbf{Q})$$

and

$$\mathcal{F}[\mathrm{m},\mathrm{Q}] = \mathrm{DH}^{-1}[\mathrm{m}]\mathrm{Q}$$

[Erlangga, Oosterlee, Vuik, '06]
[Riyanti et. al., '06]
[Plessix et. al., '07]

SLIM

Helmholtz system

Involves multiple solves of

$$\begin{bmatrix} \mathcal{H}_{\omega_1} & 0 & & \\ 0 & \mathcal{H}_{\omega_2} & \ddots & \\ & \ddots & \ddots & 0 \\ & & 0 & \mathcal{H}_{\omega_{n_f}} \end{bmatrix} \begin{bmatrix} \overbrace{[\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_{n_s}]_{\omega_1}}^{\mathbf{U}_{\omega_1}} \\ \vdots \\ \vdots \\ \underbrace{[\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_{n_s}]_{\omega_{n_f}}}_{\mathbf{U}_{n_f}} \end{bmatrix} = \begin{bmatrix} \overbrace{[\mathbf{q}_1\ \mathbf{q}_2\ \cdots\ \mathbf{q}_{n_s}]_{\omega_1}}^{\mathbf{Q}_{\omega_1}} \\ \vdots \\ \vdots \\ \underbrace{[\mathbf{q}_1\ \mathbf{q}_2\ \cdots\ \mathbf{q}_{n_s}]_{\omega_{n_f}}}_{\mathbf{Q}_{n_f}} \end{bmatrix}$$

and the *adjoint* system

PDE constrained optimization

Design *data-space* Compressive Sampling strategies that

- *reduce* # of frequencies & right-hand-sides
- *commute* with the block-diagonal *Helmholtz* system
- *overhead* **<<** *gain* in computational *speed*

Dimensionality reduction

$$\begin{cases} \mathbf{Q} = \mathbf{D}^* \underbrace{\mathbf{s}}_{\text{single shots}} \\ \mathbf{HU} = \mathbf{Q} \\ \mathbf{y} = \mathbf{RMDU} \end{cases} \quad \Longleftrightarrow \quad \begin{cases} \underline{\mathbf{Q}} = \underline{\mathbf{D}}^* \underbrace{\mathbf{RMs}}_{\text{simul. shots}} \\ \underline{\mathbf{HU}} = \underline{\mathbf{Q}} \\ \underline{\mathbf{y}} = \underline{\mathbf{DU}} \end{cases}$$

Reduced system

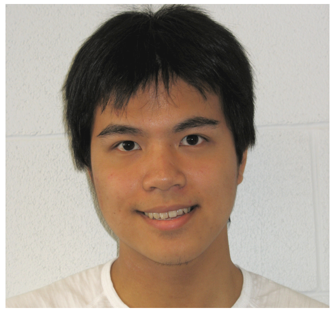
$$\min_{\underline{\mathbf{U}} \in \underline{\mathcal{U}},\, \mathbf{m} \in \mathcal{M}} \frac{1}{2} \|\underline{\mathbf{y}} - \underline{\mathbf{DU}}\|_2^2 \quad \text{subject to} \quad \underline{\mathbf{H}}[\mathbf{m}]\underline{\mathbf{U}} = \underline{\mathbf{Q}}$$

[Neelamani et. al. ‘08]
[FJH et al., ‘07-09]

Recent advances

Numerical modeling costs

- are ***no*** longer determined by the ***size*** of the *discretization*...
- but by *transform-domain* **compressibility** of the ***solution***...

Data-space reduction

"Compresive simultaneous full-waveform simulation" by Felix J. Herrmann, Tim T.Y. Lin*, Yogi A. Erlangga. SM 1 Algorithms and Methods. Room 360 A @ 11:25 AM.

"Seismic waveform inversion with Gauss-Newton-Krylov method" by Yogi Erlangga and Felix J Herrmann. SI 3 Methods. Room: 351 F @ 04:25 PM.

Challenges

$$\min_{\mathbf{U}\in\mathcal{U},\,\mathbf{m}\in\mathcal{M}} \frac{1}{2}\|\mathbf{P}-\mathbf{D}\mathbf{U}\|_2^2 \quad \text{subject to} \quad \mathbf{H}[\mathbf{m}]\mathbf{U}=\mathbf{Q}$$

+ Free surface BC

- *there are local minima*--, *i.e.*, many velocity *models* explain *data* within the same *error*
- *continuation methods* only offer partial *solution*
- miss *the ability* to *focus*
- *flexibility* to harvest multiexperiment wavefields U,V for information (e.g., AVO) ...

[Symes, '93,'04, '09]
[Plessix '99]
[Stolk '03]
[de Hoop '03]
[Biondo '04]
[Shen '08]

Migration-velocity analysis

Motivated by *differential semblance*

$$\min_{\mathbf{m}} \|(\mathsf{P}_h \overbrace{\boldsymbol{\delta}\mathbf{I}(\cdot,\ h;\mathbf{m},\delta\mathbf{d})}^{\text{image volume}})\|_2 \quad \text{with} \ \overbrace{\mathsf{P}_h\cdot = \mathbf{h}\cdot}^{\text{annihilator}},$$

redundant coordinate (h)

make the wave equation *nonlocal*-i.e.,

$$\mathbf{m}(x) \mapsto \mathbf{I}(m,h)$$

$$\text{with } m = \frac{1}{2}(x_s + x_r) \text{ and } h = \frac{1}{2}(x_s - x_r)$$

[Symes, '08]

SLIM

Nonlinear MVA

Solve

$$\min_{\mathbf{U}\in\mathcal{U},\mathbf{I}\in\mathcal{I}} \frac{1}{2}\|\mathbf{P}-\mathbf{D}\mathbf{U}\|_2^2 \quad \text{subject to} \quad \begin{cases} \mathbf{H}[\mathbf{I}]\mathbf{U} & = \mathbf{Q} \\ \mathrm{P}_h\mathbf{I} & = \mathbf{0} \end{cases}$$

- increases *degrees of freedom*
- promotes *focusing*
- involves *non-local* modeling

[Symes, '08]

SLIM

Gradient

Pre-stack migration:

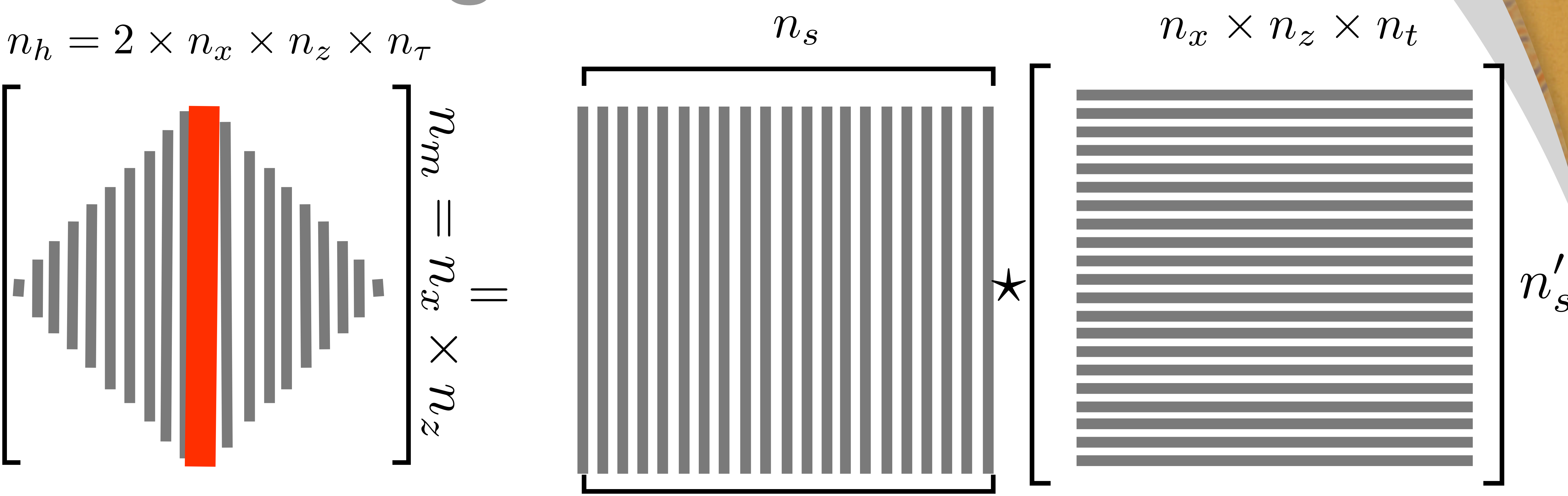
$$\boldsymbol{\delta}\mathbf{I}(m,h,\tau) = \left(\bar{\mathbf{U}} \star \mathbf{V}^T\right) = \underline{\mathbf{K}}^*[\mathbf{m},\mathbf{Q}]\delta\mathbf{d}$$

with for $f = 1 \cdots n_f$

$$\left(\bar{\mathbf{U}} \star \mathbf{V}^T\right) := \mathbf{T}_{(x_s,x_r,\omega)\mapsto(m,h,\tau)} \begin{bmatrix} \bar{\mathbf{U}}_1 & & \\ & \ddots & \\ & & \bar{\mathbf{U}}_{n_f} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \vdots \\ \mathbf{V}_{n_f}^T \end{bmatrix}$$

[Symes, '85]

Image volume



Claerbout's **imaging principle**:

$$
\begin{aligned}
\delta \mathbf{m} &= \boldsymbol{\delta}\mathbf{I}(\cdot, h = 0, \tau = 0) \\
&= \mathbf{K}^*[\mathbf{m}, \mathbf{Q}]\boldsymbol{\delta}\mathbf{d}
\end{aligned}
$$

Image volume

Image *volume large* because of *degrees of freedom*

Wavefields too *expensive* to *store*

Formation *image volume* relies on *full* matrix-matrix *multiplies*

Use *dimensionality reduction* techniques

- beyond *phase-encoded sources*
- "*model-space*" dimension *reduction*

Dimensionality reduction

Look at the Google/Yahoo people for help ...

- *Fast Monte Carlo Algorithms for Matrices I: Approximating Matrix Multiplication by P. Drineas, R. Kannan, and M.W. Mahoney, '04*
- *Improved Approximation Algorithms for Large Matrices via Random Projections* by Tamás Sarlós, '08

$$\mathbf{AB} \approx \mathbf{A}\left(\mathbf{RM}\right)^{*}\left(\mathbf{RM}\right)\mathbf{B}$$

[Romero et. al., '98.]
[Morton & Ober '00]
[Romberg, '08]

SLIM

Model-space reduction

For each *angular* frequency phase encode

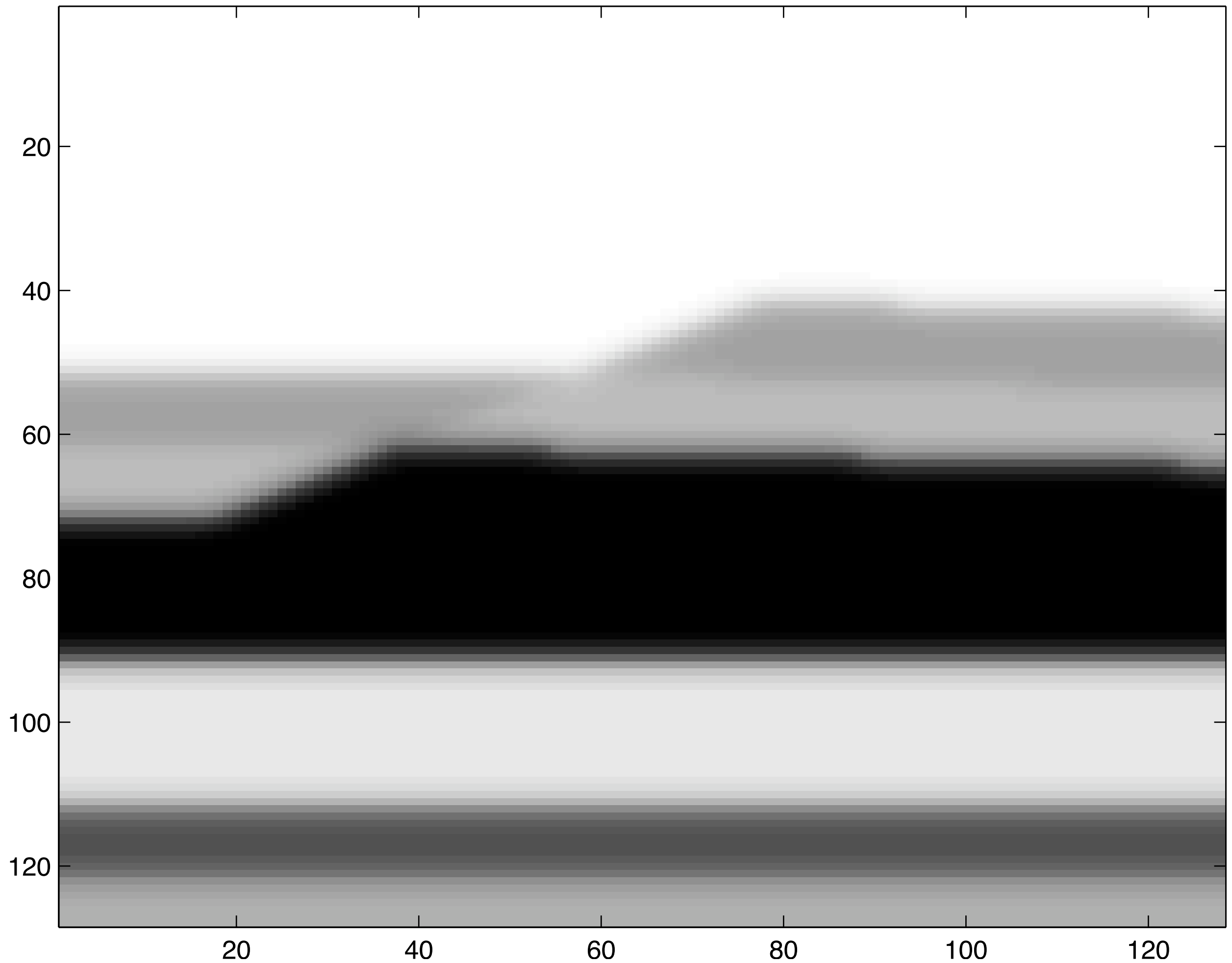
$$\mathbf{RM} := \overbrace{\begin{bmatrix} \mathbf{R}_1^{\sigma} \otimes \mathbf{R}_1^{\rho} \otimes \mathbf{R}_1^{\zeta} \\ \vdots \\ \mathbf{R}_{n'_f}^{\sigma} \otimes \mathbf{R}_{n'_f}^{\rho} \otimes \mathbf{R}_{n'_f}^{\zeta} \end{bmatrix}}^{\text{sub sampler}} \overbrace{\left(\mathbf{F}_3^* \left(e^{\hat{i}\boldsymbol{\theta}}\right)\right) \mathbf{F}_3}^{\text{random phase encoder}},$$

with $n'_f \times n'_\sigma \times n'_\rho \times n'_\zeta \ll n_f \times n_s \times n_r \times n_z$

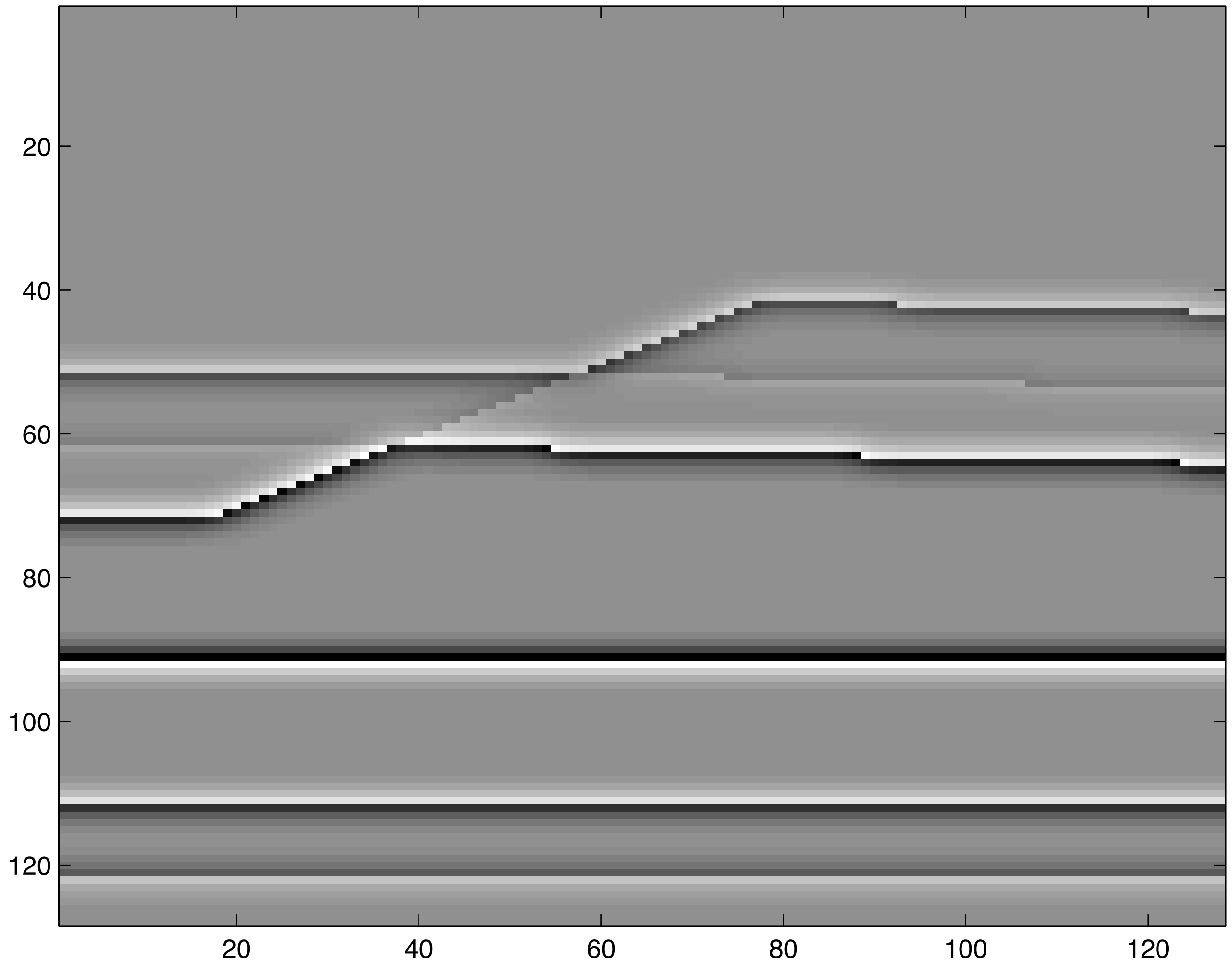
Approximate *image volume* by

$$\boxed{\boldsymbol{\delta}\mathbf{I}(m,h,t) \approx \left(\bar{\mathbf{U}}\left(\mathbf{RM}\right)^* \star \mathbf{RMV}^T\right)}$$

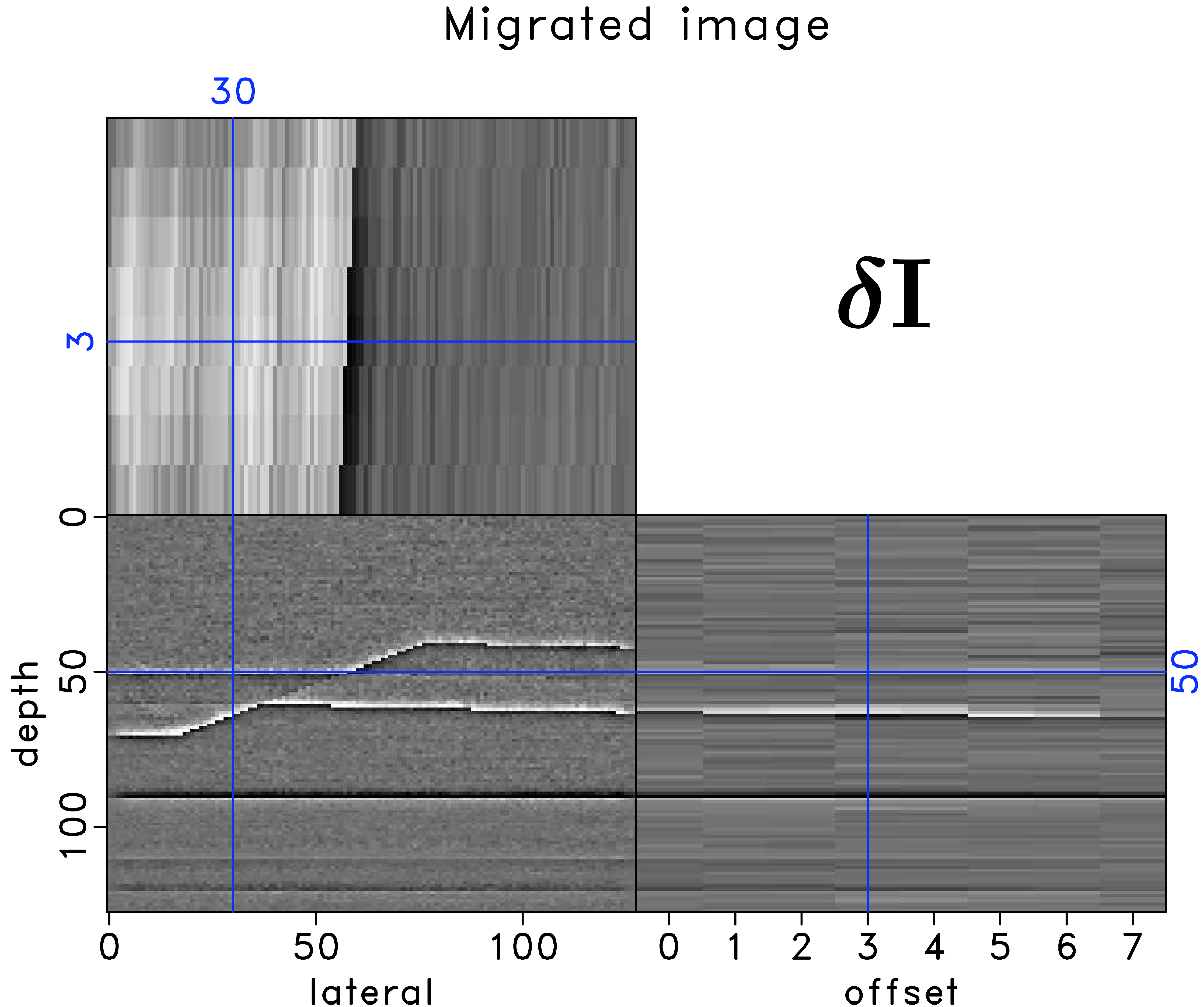
Example



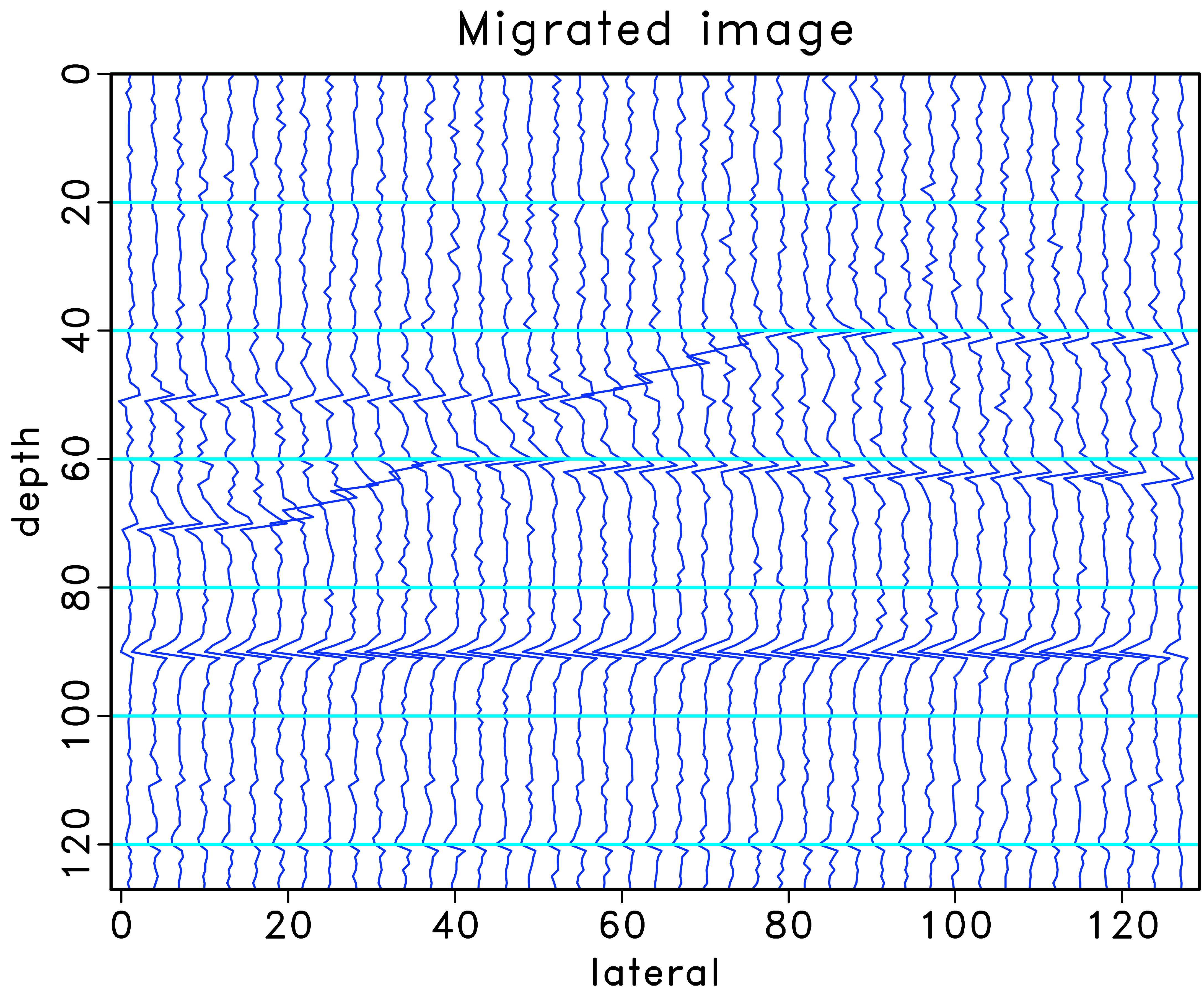
background velocity model



perturbation



Migrated image
30
3
0
50
100
depth
50
δI
0
50
100
lateral
0
1
2
3
4
5
6
7
offset



$$\boldsymbol{\delta}\mathbf{I}(\cdot, h = 0, \tau = 0)$$

Observations

Recovery from 0.1% of the samples ... But is

- noisy
- and *not* focused

Extended linearized inversion

Gauss-Newton updates with *reduced* Hessian

$$\underbrace{\underline{\mathbf{K}}^*\underline{\mathbf{K}}}_{\boldsymbol{\mathcal{H}}^{GN}}\boldsymbol{\delta}\mathbf{I}_{LS} \approx \underbrace{\underline{\mathbf{K}}^*\boldsymbol{\delta}\mathbf{d}}_{\boldsymbol{\delta}\mathbf{I}}$$

Invert with Lanczos (Conjugate gradients)

$$\boldsymbol{\delta}\mathbf{I}_{LS} = \underline{\mathbf{K}}^{\dagger}\boldsymbol{\delta}\mathbf{d} = \min_{\boldsymbol{\delta}\mathbf{I}} \|\boldsymbol{\delta}\mathbf{d} - \underline{\mathbf{K}}\boldsymbol{\delta}\mathbf{I}\|_2$$

Dimensionality reduction

Look again at the Google/Yahoo people for help ...

- *Sampling Algorithms for L2 Regression and Applications* by P. Drineas, M. W. Mahoney, and S. Muthukrishnan, '06
- *Improved Approximation Algorithms for Large Matrices via Random Projections* by Tamás Sarlós, '08

$$\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \approx \min_{\mathbf{x}} \|\mathbf{R}\mathbf{M}\big(\mathbf{b} - \mathbf{A}\mathbf{x}\big)\|_2$$

Sparsity promotion and focusing

Least-squares does not promote *sparsity*.

CS recovers from *dimension reductions* through transform-domain *sparsity* promotion.

Missing fundamental principle of *focusing*...

Use recent results on mixed (1,2)-norm minimization.

Focussed wavefield inversion

Least-squares inversion for image volumes is equivalent to solving multi-D *deconvolution* problem, i.e., by inverting

$$(\mathbf{U}^* * \boldsymbol{\delta}\mathbf{I}) \quad \approx \quad \mathbf{V}^T$$

- reduce dimensionality by *model-space* CS
- include *sparsity promotion* and *focusing*

CS & focus

Compressively sample *augmented* system:

$$\begin{aligned} \mathbf{RM}\left(\mathbf{U}^* * \mathbf{S}^*\mathbf{X}\right) &\approx \mathbf{RMV}^T \\ \mathsf{P}_h\mathbf{X} &\approx \mathbf{0} \end{aligned} \quad \text{or} \quad \mathbf{AX} \approx \mathbf{B}$$

Solve with *mixed* (1,2) norm minimization

Group-sparsity promotion

Back to *optimizers* & CS people

- *Joint* sparsity-promotion with mixed (1,2) norms
- *Joint-sparse recovery from multiple measurements* by E. van den Berg and M. Friedlander, '09

$$\tilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathbf{X}\|_{1,2} \quad \text{subject to} \quad \|\mathbf{A}\mathbf{X} - \mathbf{B}\|_{2,2} \leq \sigma$$

Group-sparsity

Solve by mixed (1,2)-norm minimization:

$$\tilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathbf{X}\|_{1,2} \quad \text{subject to} \quad \|\mathbf{A}\mathbf{X} - \mathbf{B}\|_{2,2} \leq \sigma,$$

with

$$\|\mathbf{X}\|_{1,2} := \sum_{i \in \text{rows}(\mathbf{X})} \|\text{row}_i(\mathbf{X})^*\|_2$$

and

$$\|\mathbf{X}\|_{2,2} := \left(\sum_{i \in \text{rows}(\mathbf{X})} \|\text{row}_i(\mathbf{X})^*\|_2^2 \right)^{\frac{1}{2}}.$$

Computational details

Compute image volume for varying

- horizontal offset, i.e.,

$$h = h_x = \frac{1}{2}(x_s - x_r) \text{ for } z = z_s = z_r \text{ fixed}$$

- ‘vertical’ time offset

$$\tau = \frac{1}{2}(t_s - t_r)$$

- combine

$$\mathbf{h} = (h, \tau) \text{ and } \mathsf{P}_{\mathbf{h}} = h \otimes \tau$$

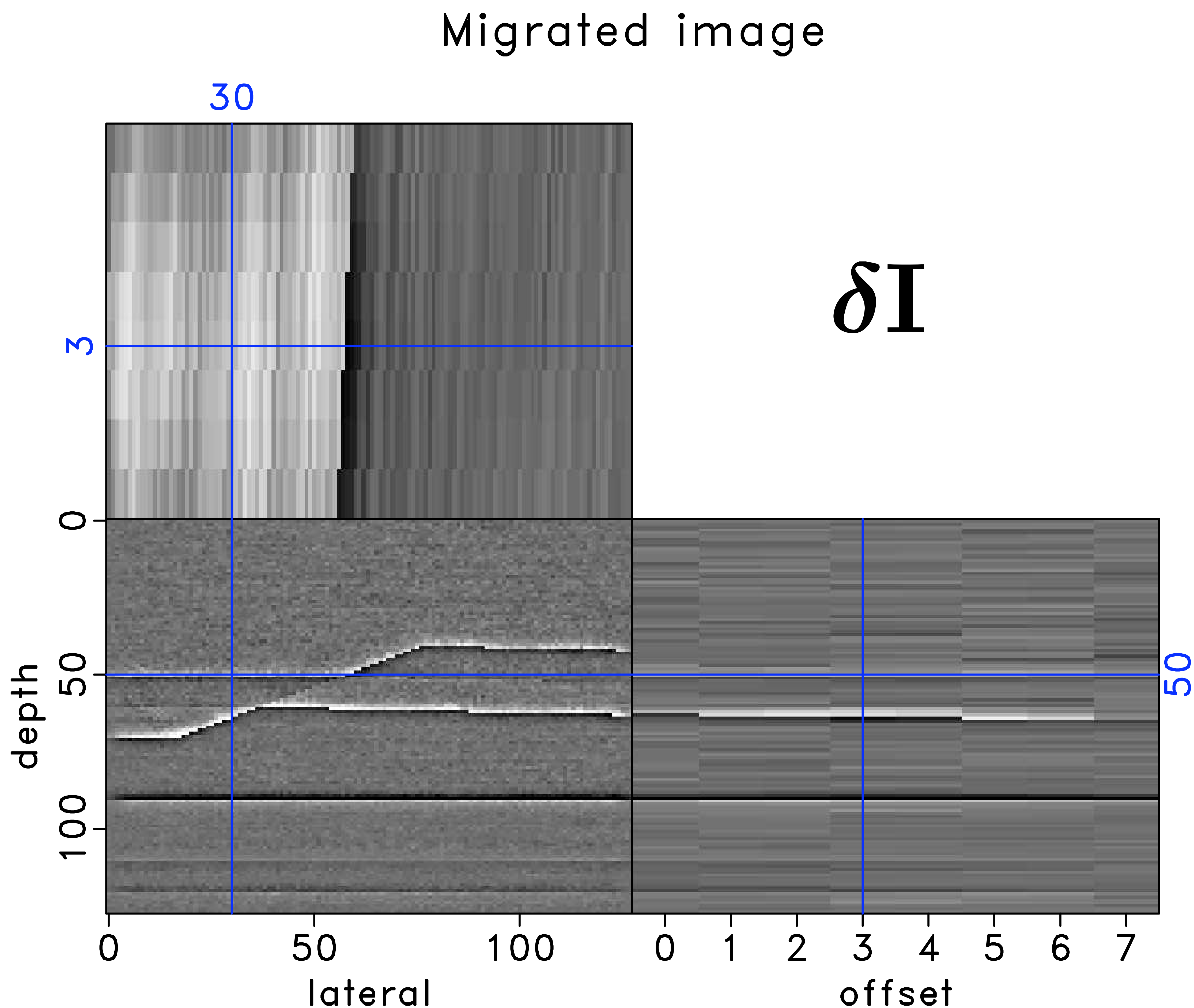
Recovery

Solve

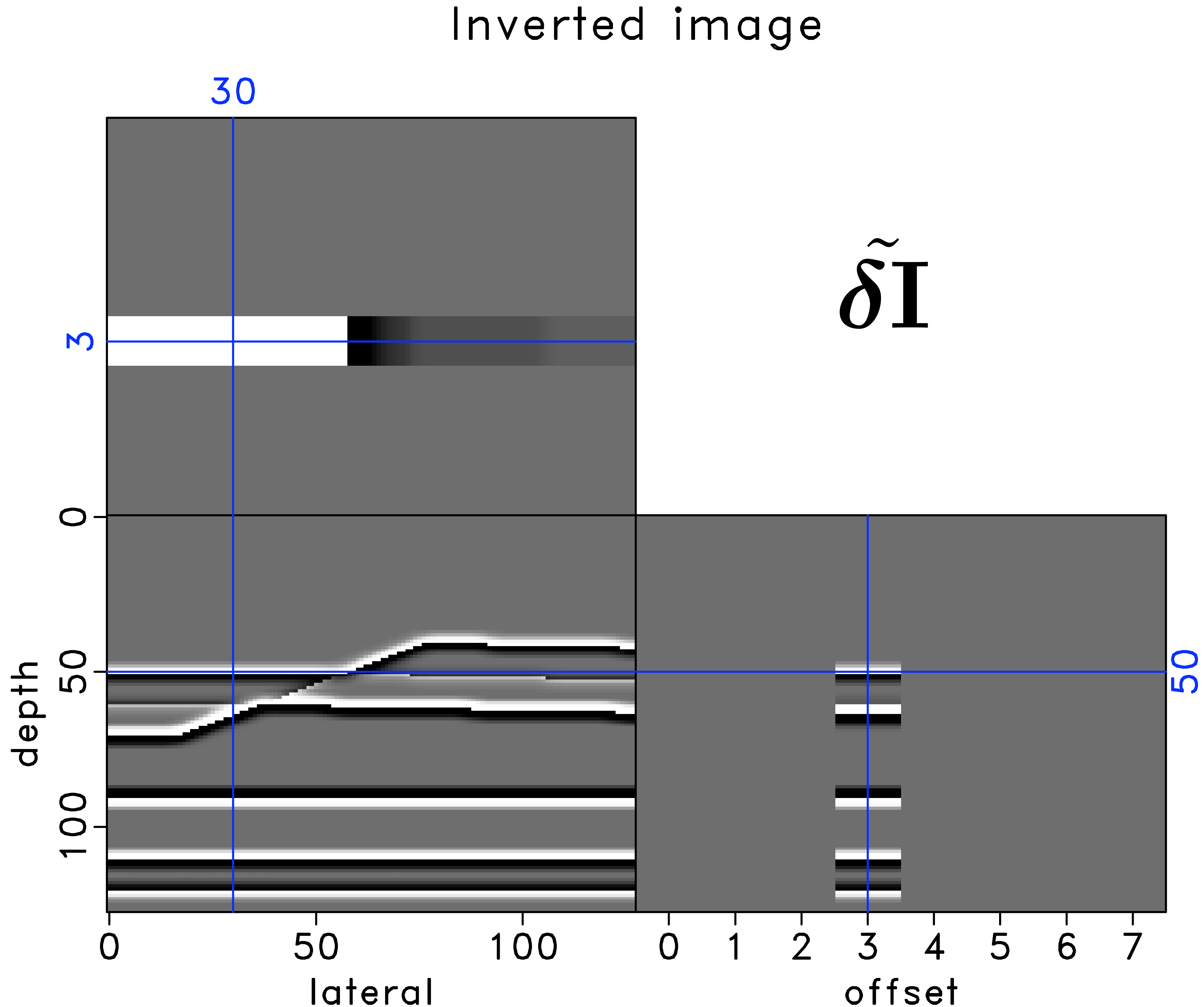
$$\begin{cases} \tilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathbf{X}\|_{1,2} \quad \text{subject to} \quad \|\mathbf{A}\mathbf{X} - \mathbf{B}\|_{2,2} \leq \sigma \\ \tilde{\delta\mathbf{I}} = \mathbf{S}^*\tilde{\mathbf{X}} \end{cases}$$

with

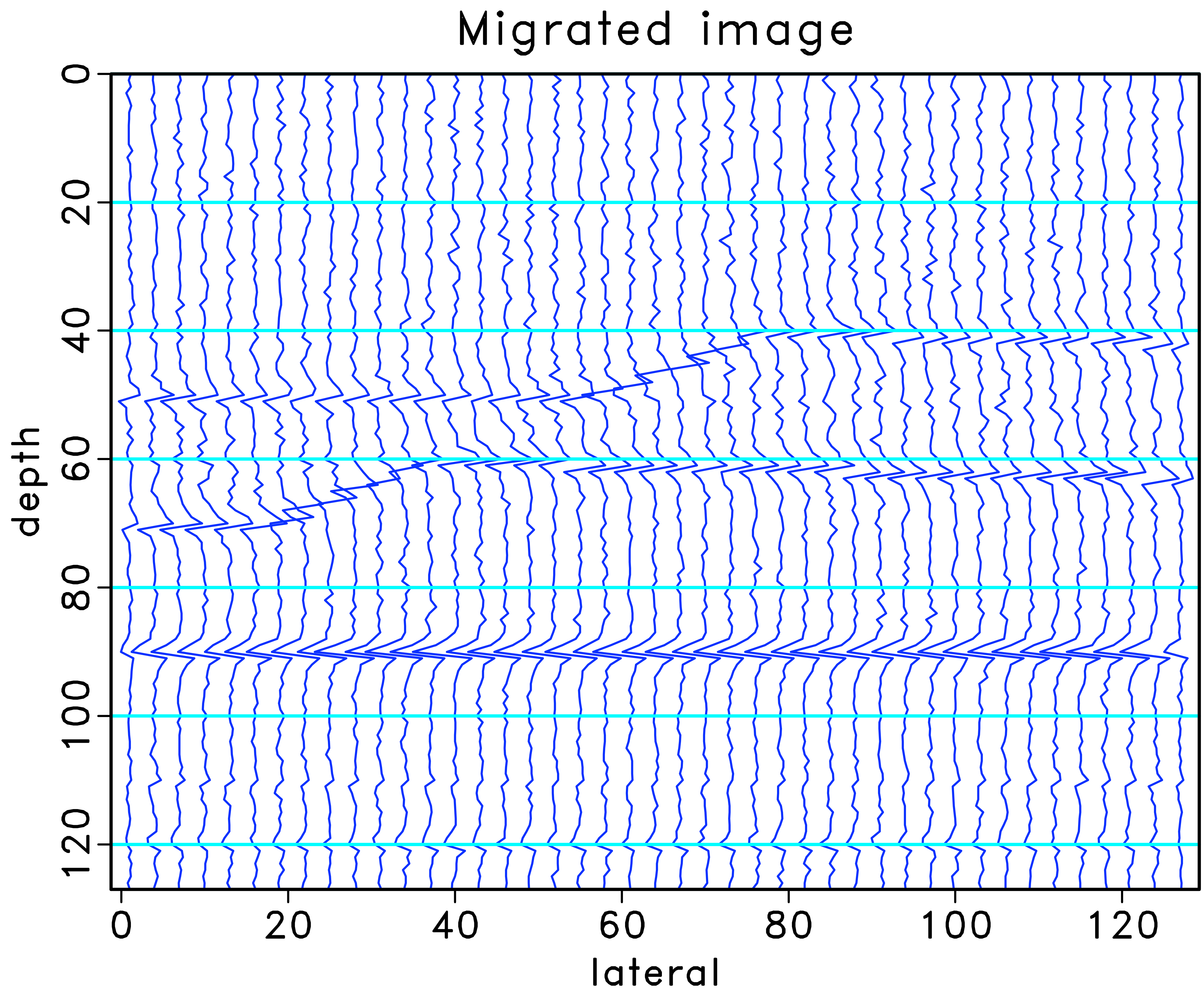
$$\mathbf{S} = \mathbf{W} \otimes \mathbf{W}$$



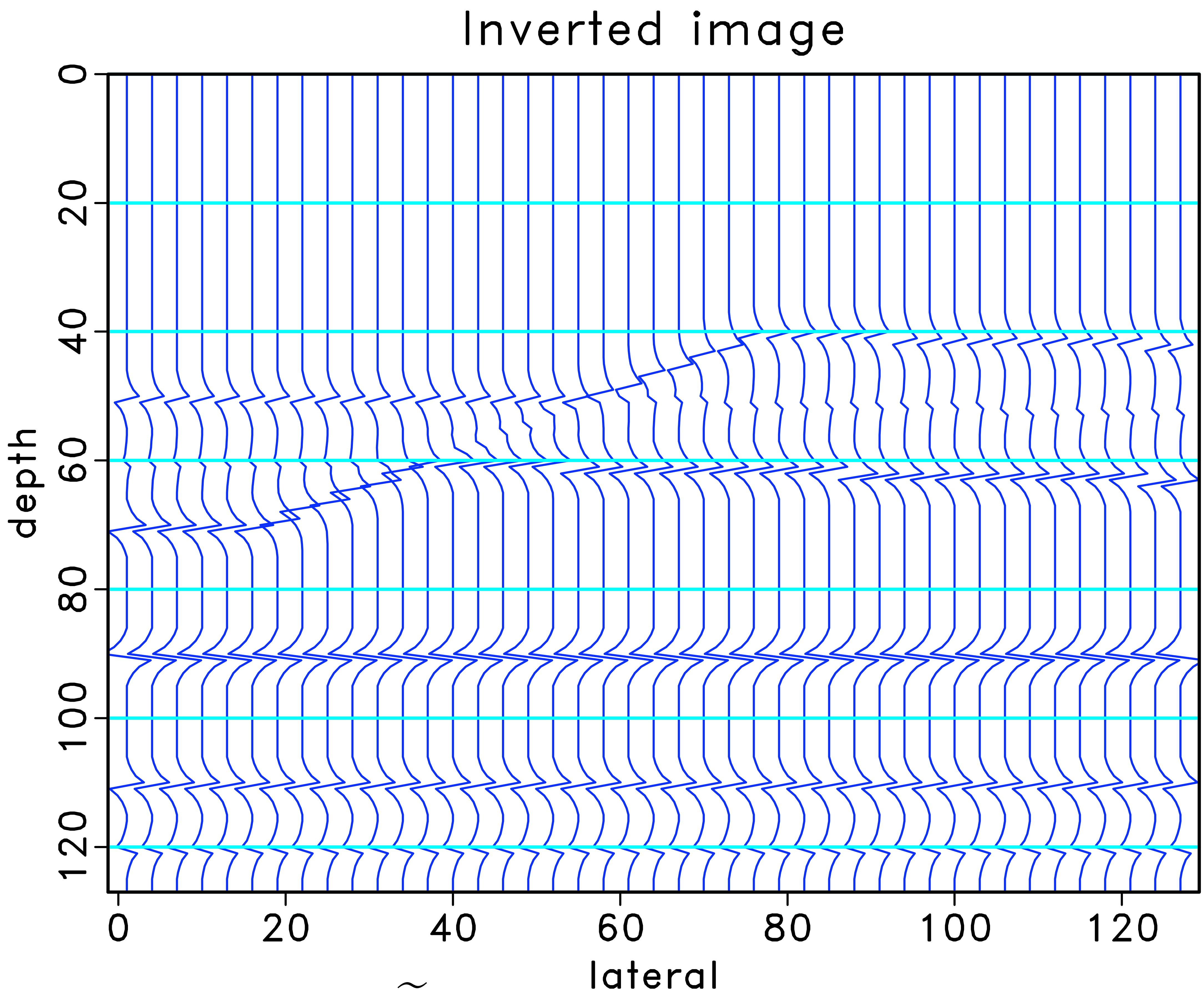
Migrated image
30
3
0
50
100
depth
50
δI
0
50
100
lateral
0
1
2
3
4
5
6
7
offset



Inverted image
δĨ
30
3
0
50
100
depth
0
50
100
lateral
0
1
2
3
4
5
6
7
offset
50



$$\boldsymbol{\delta}\mathbf{I}(m, h = 0, \tau = 0)$$



$$\widetilde{\boldsymbol{\delta}\mathbf{I}}(m, h = 0, \tau = 0)$$

SLIM

Observations

Model ***reductions*** by CS are ***essential***

Group *sparsity* promotion allows for

- *focusing* amongst *offsets* via two-norm minimization
- *sparsity*-promotion amongst *images* via one-norm minimization

Perspective of solving the *extended* formulation...

CS alternative

Computational *complexity* **no** longer depends on the *size of the discretization*...

...but on *transform-domain compressibility* of the *solution*...

Conclusions & outlook

CS allows for model-size reduction

- image volumes computable
- recoverable with group-sparsity promotion
- promotes sparsity & focusing

Opens perspectives towards

- harvesting image volumes for information
- Symes' extension

SLIM

Acknowledgments

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 (334810-05) of Felix J. Herrmann.

This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BP, Petrobras, and Schlumberger.

Thanks to the DELPHI consortium for their hospitality

Thanks to Eric Verschuur and Martin Lacasse for fruitful discussions.

SLIM

Thank you for your attention!

more information

slim.eos.ubc.ca