Primary estimation with sparsity-promoting bi-convex optimization

Tim Tai-Yi Lin

Ph.D. Final Oral Defence
October 1, 2015
Introduction

Active reflection seismic surveys
Reflectivity model

Green's function

recorded signal

reflectivity series

(+) (-)

geologic section

interfaces

time

depth

Reflectivity model
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**

\[ \begin{align*}
\text{Path}_1 & \quad \ast \quad \text{Path}_2 \quad = \quad \text{Path}_3
\end{align*} \]

**Subtraction of raypath under cross-correlation**

\[ \begin{align*}
\text{Path}_4 & \quad \ast \quad \text{Path}_2 \quad = \quad \text{Path}_5
\end{align*} \]
Modeling all surface multiples by convolution with data

\[ p_m(x, \omega; x_{\text{src}}) = \int_S g(x, \omega; x') r(x, x') p(x', \omega; x_{\text{src}}) \, dx' \]

(frequency domain)
Modeling all surface multiples by convolution with data

- Surface-free Greens function
- Observed pressure data
- True surface multiples
A multidimensional convolution model for seismic data

\[
p_m(x, \omega; x_{src}) = \int_S g(x, \omega; x') r(x, x') p(x', \omega; x_{src}) \, dx' \quad \text{frequency domain}
\]

Primary model

\[
M(g, q; p) := \mathcal{F}_\omega^{-1} \left[ g(x, \omega; x_{src}) q(\omega) + \int_S g(x, \omega; x') r(x, x') p(x', \omega; x_{src}) \, dx' \right]
\]

= \[p_o(x, t; x_{src}) + p_m(x, t; x_{src})
\]

= \[p(x, t; x_{src}) \quad \text{Total data model}\]
\[ G \text{ (Green's function)} \quad \ast \quad P \text{ (Seismic data)} \quad \rightarrow \quad \text{Convolve with wavelet } Q \]

\[ P_m \text{ (Surface multiples)} \quad \rightarrow \quad \text{Po (Primaries)} \]
A multidimensional convolution model for seismic data

\[
p_m(x, \omega; x_{\text{src}}) = \int_S g(x, \omega; x') r(x, x') p(x', \omega; x_{\text{src}}) \, dx' \quad \text{frequency domain}
\]

Primary model

\[
M(g, q; p) := \mathcal{F}_\omega^{-1}[g(x, \omega; x_{\text{src}}) q(\omega)] + \int_S g(x, \omega; x') r(x, x') p(x', \omega; x_{\text{src}}) \, dx'
\]

\[
= p_o(x, t; x_{\text{src}}) + p_m(x, t; x_{\text{src}})
\]

\[
= p(x, t; x_{\text{src}}) \quad \text{Total data model}
\]

Discretizes into

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

\[
M_\omega(G, Q; P) := GQ + GRP
\]

Underlying assumption: there exists \( G \) for all frequencies such that

\[
P \approx GQ - GP
\]

\[
P_o = GQ
\]

"Data matrix" notation

\[
\begin{bmatrix}
G
\end{bmatrix}
\]

\( \omega = \omega_a \) single fixed frequency
Estimation of Primaries by Sparse Inversion
(van Groenestijn and Verschuur, 2009)

\[ P = QG - GP \]

- \( P \): total up-going wavefield
- \( Q \): down-going source signature
- \( G \): primary impulse response

recorded data \hspace{1cm} \text{predicted data from convolution model}
Estimation of Primaries by Sparse Inversion (EPSI)
(van Groenestijn and Verschuur, 2009)

recorded data  predicted data from convolution model

\[ P = QG - GP \]

Inversion objective:

\[ f(G, Q) = \frac{1}{2} \| P - (QG - GP) \|^2 \]

Sparse regularization on \( g \) in physical (time) domain via hard-thresholding on gradients
Primary Estimation instead of Multiple Removal

\[ P = QG - GP \]

- \( P \): total up-going wavefield
- \( Q \): down-going source signature
- \( G \): primary impulse response

**EPSI**
- Recorded data
- Predicted data from convolution model
Primary Estimation instead of Multiple Removal

\[
P = QG - GP
\]

\[
P_{o} = QG
\]

\[
A(f) = -Q^{-1}
\]

**EPSI** recorded data  predicted data from convolution model

- \(P\) total up-going wavefield
- \(Q\) down-going source signature
- \(G\) primary impulse response
Primary Estimation instead of Multiple Removal

SRME

\[ P_o = P - A(f)P_oP \]

true primary wavefield

SRME-produced primary

\[ P_o = QG \]

\[ A(f) = -Q^{-1} \]

- \( P \): total up-going wavefield
- \( Q \): down-going source signature
- \( G \): primary impulse response
Primary Estimation instead of Multiple Removal

SRME

\[
\begin{align*}
\text{true primary wavefield} & \quad \text{SRME-produced primary} \\
\mathbf{P}_o & \approx \mathbf{P} - A(f) \mathbf{G} \\
\mathbf{P} & \quad \text{total up-going wavefield} \\
\mathbf{Q} & \quad \text{down-going source signature} \\
\mathbf{G} & \quad \text{primary impulse response} \\
\end{align*}
\]

\[
\mathbf{P}_o = \mathbf{Q} \mathbf{G} \\
A(f) = -\mathbf{Q}^{-1}
\]
Primary Estimation instead of Multiple Removal

adaptive subtraction

\[ \min_A \sum_f \| P - A(f) P \| \]

SRMP

\[ P_0 = QG \]
\[ A(f) = -Q^{-1} \]

P  total up-going wavefield
Q  down-going source signature
G  primary impulse response
Primary Estimation instead of Multiple Removal

recorded data predicted data from convolution model

\[ P = QG - GP \]

Inversion objective:

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

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

\[ P = QG - GP \]

Inversion objective:

\[ f(G, Q) = \frac{1}{2} \| P - (QG - GP) \|_2^2 \]
Optimization form

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

- **recorded data:**
  \[ p = M(g, q; p) \]

- **predicted data from SRME:**
  \[ M(g, q; p) = F_\omega^{-1}[M_\omega(G, Q; P)] \]
  \[ M_\omega(G, Q; P) := GQ + GRP \]

**Inversion objective:**

\[ f(g, q) = \frac{1}{2} \| p - M(g, q; p) \|_2^2 \]
Solving the EPSI problem

**Linearizations**

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

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

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

In fact it is bilinear:

\[ M_{\tilde{q}}\mathbf{g} = M(\mathbf{g}, \tilde{\mathbf{q}}) \quad M_{\tilde{g}}\mathbf{q} = M(\mathbf{q}, \tilde{\mathbf{g}}) \]
Solving the EPSI problem

Linearizations

\[ p = M(g, q; p) \]

\[ M_{\tilde{q}} = \left( \frac{\partial M}{\partial g} \right)_{\tilde{q}} \]

\[ M_{\tilde{g}} = \left( \frac{\partial M}{\partial q} \right)_{\tilde{g}} \]

Associated objectives:

\[ f_{\tilde{g}}(g) = \frac{1}{2} \| p - M_{\tilde{q}} g \|_2^2 \quad f_{\tilde{g}}(q) = \frac{1}{2} \| p - M_{\tilde{g}} q \|_2^2 \]
Solving the EPSI problem

Do:

\[ g_{k+1} = g_k + \alpha S(\nabla f_{q_k}(g_k)) \]
\[ q_{k+1} = q_k + \beta \nabla f_{g_{k+1}}(q_k) \]

Gradient sparsity

\( S : \text{pick largest } \rho \text{ elements per trace} \)
Solving the EPSI problem

Data

EPSI Green’s function
Robust EPSI
L1-minimization approach to the EPSI problem

[Lin and Herrmann, 2013 Geophysics]

While \[\|p - M(g_k, q_k)\|_2 > \sigma\]

determine new \(\tau_k\) from the Pareto curve

\[g_{k+1} = \arg\min_g \|p - M_{q_k} g\|_2 \text{ s.t. } \|g\|_1 \leq \tau_k\]

\[q_{k+1} = \arg\min_q \|p - M_{g_{k+1}} q\|_2\]
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 \( \tau_k \) from the Pareto curve

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

\[
\mathbf{q}_{k+1} = \arg \min_{\mathbf{q}} \| \mathbf{p} - \mathbf{M}_{\mathbf{g}_{k+1}} \mathbf{q} \|_2
\]
While $\|p - M(g_k, q_k)\|_2 > \sigma$

determine new $\tau_k$ from the Pareto curve

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

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

$$q_{k+1} = \arg\min_q \|p - M g_{k+1} q\|_2$$

(wavelet matching, solve with LSQR)

[Robust EPSI: L1-minimization approach to the EPSI problem]

[Lin and Herrmann, 2013 Geophysics]
Projected gradient. Our application of the SPG algorithm to solve \((L_\tau)\) follows Birgin, Martínez, and Raydan [5] closely for the minimization of general nonlinear functions over arbitrary convex sets. The method they propose combines projected-gradient search directions with the spectral step length that was introduced by Barzilai and Borwein [1]. A nonmonotone line search is used to accept or reject steps. The key ingredient of Birgin, Martínez, and Raydan’s algorithm is the projection of the gradient direction onto a convex set, which in our case is defined by the constraint in \((L_\tau)\). In their recent report, Figueiredo, Nowak, and Wright [27] describe the remarkable efficiency of an SPG method specialized to \((Q_\lambda P)\). Their approach builds on the earlier report by Dai and Fletcher [18] on the efficiency of a specialized SPG method for general bound-constrained quadratic programs \((Q_\lambda P)\).

2. The Pareto curve. The function \(\phi\) defined by (1.1) yields the optimal value of the constrained problem \((L_\tau)\) for each value of the regularization parameter \(\tau\). Its graph traces the optimal trade-off between the one-norm of the solution \(x\) and the two-norm of the residual \(r\), which defines the Pareto curve. Figure 2.1 shows the graph of \(\phi\) for a typical problem.

The Newton-based root-finding procedure that we propose for locating specific points on the Pareto curve—e.g., finding roots of (1.2)—relies on several important properties of the function \(\phi\). As we show in this section, \(\phi\) is a convex and differentiable function of \(\tau\). The differentiability of \(\phi\) is perhaps unintuitive, given that the one-norm constraint in \((L_\tau)\) is not differentiable. To deal with the nonsmoothness of the one-norm constraint, we appeal to Lagrange duality theory. This approach yields significant insight into the properties of the trade-off curve. We discuss the most important properties below.

2.1. The dual subproblem. The dual of the Lasso problem \((L_\tau)\) plays a prominent role in understanding the Pareto curve. In order to derive the dual of \((L_\tau)\), we first recast \((L_\tau)\) as the equivalent problem

\[
\begin{align*}
\text{minimize} & \quad r, x \quad \|r\|_2 \\
\text{subject to} & \quad Ax + r = b, \quad \|x\|_1 \leq \tau.
\end{align*}
\]

(van den Berg, Friedlander, 2008)
Choosing Tau from the Pareto curve

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

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

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

feasible solution with smallest $\|x\|_1$
Solving the LASSO subproblems

Possible solutions of $Ax=b$

$\ell_1$ ball
($\|x\|_1 = \tau$)

Steepest descent direction
$A^\dagger (Ax - b)$

minimize
subject to
$\|Ax - b\|_2$
$\|x\|_1 \leq \tau$
Solving the LASSO subproblems

\[
\begin{align*}
\text{minimize} & \quad \|Ax - b\|_2 \\
\text{subject to} & \quad \|x\|_1 \leq \tau
\end{align*}
\]

\(\ell_1\) ball 
\((\|x\|_1 = \tau)\)

Subtract away shortest L2 distance to the ball
Solving the LASSO subproblems

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

$\ell_1$ ball
($\|x\|_1 = \tau$)

“projected gradient”
L1 projection and sparsity

variable $g$ at beginning of LASSO

$$g_{k+1} = \arg \min_g \| p - M_{q_k} g \|_2 \text{ s.t. } \| g \|_1 \leq \tau_k$$
L1 projection and sparsity

**variable g at end of LASSO**

\[ g_{k+1} = \arg \min_{g} \| p - M_q g \|_2 \text{ s.t. } \| g \|_1 \leq \tau_k \]

Emits “deconvolved” solution
Robust EPSI
L1-minimization approach to the EPSI problem

[Lin and Herrmann, 2013 Geophysics]

\[
\text{While } \|p - \mathcal{M}(g_k, q_k)\|_2 > \sigma
\]

determine new \( \tau_k \) from the Pareto curve

\[
g_{k+1} = \arg \min_g \|p - M_{q_k} g\|_2 \text{ s.t. } \|g\|_1 \leq \tau_k
\]

\[
q_{k+1} = \arg \min_q \|p - M_{g_{k+1}} q\|_2
\]
Overall REPSI solution path

\[
\begin{align*}
\minimize \quad & \|x\|_1 \\
\text{subject to} \quad & \|Ax - b\|_2 \leq \sigma
\end{align*}
\]

- Solution of Lasso problem (used to improve \(q\))
- Intermediate Lasso solutions
- Gradient update on \(g\)
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

Time (s) 0 5 10 15 20 25 30 35 40 45
CMP coordinate (km)
Pluto1.5  REPSI primary
80 total gradient iters

CMP coordinate (km)
Pluto1.5  REPSI multiple model
80 total gradient iters
Pluto1.5 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)
Gulf of Suez data
NMO-corrected stack
Gulf of Suez data
NMO-corrected stack

REPSI primary wavefield
Gulf of Suez data
NMO-corrected stack

REPSI + transform domain primary wavefield
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*
So what happens...

...when there are gaps in your data?
Chapter 4

Mitigating data gaps with scattering terms
Data with missing traces $P'$  

Exact multiples  

EPSI Predicted -GP'  
(with perfect G)
Missing contributions

Exact multiples

EPSI Predicted -GP'
(with perfect G)
Main idea

Augment the convolutional model

\[ P = GQ + GRP \]

to account for the missing contribution
Trace Mask

Masking operator $K$
Trace Mask

Bisects wavefield data to unknown/uncertain traces

(e.g., near-offset)
Trace Mask

Masking operator $K$

Time domain: $Kp$

Frequency slices: $K \circ P$
Trace Mask

Complement of Masking operator $K_c$

Time domain: $K_c p$

Frequency slices: $K_c \circ P$
Bisected data variables $P' + P'' = P$

Known data traces: $P' := K \circ P$

Unknown data traces: $P'' := K_c \circ P$
Bisected data variables \[ P' + P'' = P \]

Known data traces: \[ P' := K \circ P \]

Unknown data traces: \[ P'' := K_c \circ P \]
\[ = K_c \circ (GQ + RGP' + RGP'') \]
Bisected data variables \( P' + P'' = P \)

Known data traces: \( P' := K \circ P \)

Unknown data traces: \( P'' := K_c \circ P \\ = K_c \circ (GQ + RGP' + RGP'') \)
Bisected data variables \( P' + P'' = P \)

Known data traces: \( P' := K \circ P \)

Unknown data traces: \( P'' = K_c \circ P \)
\[ = K_c \circ (GQ + RGP' + RGP'') \]

Trace stencil

Recursively defined
Modify the modeling operator

\[ M(G, Q; P') = GQ + RGP' \\
+ RGP'' \]
Modify the modeling operator

\[ \tilde{\mathcal{M}}(G, Q; P') = GQ + RGP' \\
+ RGK_c \circ (GQ + RGP') \\
+ O(G^3) \]

Tilde: modified with higher-order terms
Modify the modeling operator

\[ \tilde{M}(G, Q; P') = GQ + RGP' + RGK_c \circ (GQ + RGP') + \mathcal{O}(G^3) \]
Modify the modeling operator

\[ \tilde{M}(G, Q; P') = GQ + RGP' \]
\[ + \mathrm{RGK}_c \odot (GQ + RGP') \]
\[ + \mathcal{O}(G^3) \]
Modify the modeling operator

\[ \tilde{\mathcal{M}}(G, Q; P') = GQ + RGP' \]

2nd Order autoconvolution term (of G)

\[ + RGK_c \circ (GQ + RGP') 
\]

\[ + O(G^3) \]
Modify the modeling operator

Trace mask over all modeled wavefield

\[ \tilde{M}(G, Q; P') = K \circ \left[ GQ + RGP' \right] \]

2nd Order autoconvolution term

\[ + RGK_c \circ (GQ + RGP') \]

\[ + \mathcal{O}(G^3) \]
Modify the modeling operator

\[ \tilde{M}(G, Q; P') = \sum_{n=0}^{\infty} (RGK_c \circ) \circ (GQ + RGP') \]

2nd Order autoconvolution term + \[ RGK_c \circ (GQ + RGP') \]

3rd Order autoconvolution term + \[ RGK_c \circ (RGK_c \circ (GQ + RGP')) \]

+ \[ \mathcal{O}(G^4) \]

:= \sum_{n=0}^{\infty} (RGK_c \circ)^n (GQ + 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)
Position (km)  |  Time (s)
---|---
0 | 0
0.5 | 0.5
1 | 1
1.5 | 1.5
2 | 2

0 | 0
0.2 | 0.2
0.4 | 0.4
0.6 | 0.6
0.8 | 0.8
1 | 1
1.2 | 1.2
1.4 | 1.4

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

\[ \tilde{M}(G, Q; P') = K \circ [GQ + RGP'] + RGK_c \circ (GQ + RGP') + RGK_c \circ (RGK_c \circ (GQ + RGP')) + \mathcal{O}(G^4) \]

\[ := K \circ \sum_{n=0}^{\infty} (RGK_c \circ)^n (GQ + RGP') \]
Solution strategy
(dealing with non-linear modeling operator)
Autoconvolving Robust EPSI
Accounting for unknown data with $G$

While $\|p - \tilde{M}(g_k, q_k)\|_2 > \sigma$

determine new $\tau_k$ from the Pareto curve

$g_{k+1} = \arg \min_g \|p - \tilde{M}(g, q_k)\|_2 \text{ s.t. } \|g\|_1 \leq \tau_k$

$q_{k+1} = \arg \min_q \|p - \tilde{M}_{g_{k+1}} q\|_2$
Strategy 1: Re-linearization
Using $G$ from previous iter in higher-order terms

While $\|p - \mathcal{M}(g_k, q_k)\|_2 > \sigma$

determine new $\tau_k$ from the Pareto curve

$$g_{k+1} = \arg \min_g \|p - \mathcal{M}_{g_k} g\|_2 \text{ s.t. } \|g\|_1 \leq \tau_k$$

fix at $g_k$ for autoconv terms

$$q_{k+1} = \arg \min_q \|p - \mathcal{M}_{g_{k+1}} q\|_2$$
Strategy 2: Modified Gauss-Newton
Obtain Jacobian using G from previous iter

While \( \|p - \tilde{\mathcal{M}}(g_k, q_k)\|_2 > \sigma \)

determine new \( \tau_k \) from the Pareto curve

\[
\begin{align*}
g_{k+1} &= g_k + \arg\min_{\Delta g} \| r_k - \partial_{(g_k, q_k)} \tilde{\mathcal{M}} \Delta g \|_2 \text{ s.t. } \| \Delta g \|_1 \leq \tau_k \\
q_{k+1} &= \arg\min_{q} \| p - \tilde{\mathcal{M}} g_{k+1} q \|_2
\end{align*}
\]
Field data example
North Sea dataset
North Sea dataset
100m near-offset
regularized to 12.5m dx
and 4km fixed-spread
from streamer
4ms sampling
NMO stack
Parabolic Radon Interp
NMO stack
Re-linearization
Using 3rd Order terms
NMO stack
Re-linearization Using 3rd Order terms
NMO stack
Re-linearization
Using 2nd Order terms
Radon interp - Re-linearization 2nd order

Re-linearization 3rd - 2nd order

NMO stack
Difference plots
NMO stack
Modified Gauss-Newton
Using 3rd Order terms
Diff: GN vs Relinearize

GN 3rd Order Multiple

Re-lin. 3rd Order Multiple
Summary: Near-offset mitigation using scattering

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

\[
\tilde{\mathbf{M}}(\mathbf{G}, \mathbf{Q}; \mathbf{P}') = \mathbf{K} \circ [\mathbf{GQ} + \mathbf{RGP}']
\]

2nd Order Scattering

\[ + \mathbf{RGK}_c \circ (\mathbf{GQ} + \mathbf{RGP}') \]

3rd Order Scattering

\[ + \mathbf{RGK}_c \circ (\mathbf{RGK}_c \circ (\mathbf{GQ} + \mathbf{RGP}')) \]

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

\[ := \mathbf{K} \circ \sum_{n=0}^{\infty} (\mathbf{RGK}_c \circ)^n (\mathbf{GQ} + \mathbf{RGP}') \]
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
Impulse response solutions

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

- 40 min
- 6 min
- 1.5 min
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):\( n = n_{rcv} = n_{src} \)

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

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):

\[ n = n_{rcv} = n_{src} \]

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

- 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

75 iters

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

1600m/s NMO, linear interp 2x

Solution of 4x decimated data

Solve
Warm-starting/continuation from coarse solution

Example

Solution of 2x decimated data

Solution on 2x dec data

continuation from 4x dec solution

25 iters
Warm-starting/continuation from coarse solution

Example

Solution of full data

Solution on 2x dec data

continuation from 4x dec solution
Warm-starting/continuation from coarse solution

Example

Solution of full data

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

Prolongation
Warm-starting/continuation from coarse solution

Example

Solution of full data

Solution on 2x dec data > interp 2x
continuation from 4x thru 2x solution

Solve

15 iters
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
\(dx = 60m > 30m > 15m\)
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

- From full data
- Bootstrapping from 4x decimated

Wall time (minute)

- 75 iters
Significant speedup from bootstrapping
Wall times

From full data

Bootstrapping from 4x decimated

60 iters at 4x decimated spatial sampling (1 min)
Significant speedup from bootstrapping

Wall times

From full data

75 iters

Bootstrapping from 4x decimated

25 iters at 2x decimated spatial sampling (2 min)
Significant speedup from bootstrapping

Wall times

- From full data: 75 iters
- Bootstrapping from 4x decimated:
  - 15 iters at full problem size w/ all data (8 min)
Significant speedup from bootstrapping

Wall times

From full data

Bootstrapping from 4x decimated

NMO linear interp, etc... (1 min)
Significant speedup from bootstrapping

Wall times

From full data

Bootstrapping from 4x decimated
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)
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


- 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
- 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
- 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
- Designing Simultaneous Acquisitions with Compressive Sensing, Tim T. Y. Lin and Felix J. Herrmann, 71st EAGE Conference & Exhibition, 2009
- 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