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

Tim T.Y. Lin and Felix J. Herrmann
Amsterdam, EAGE 2014
From SRME to Robust EPSI

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

\[
\begin{align*}
\text{true primary wavefield} & \quad \text{SRME-produced primary} \\
\mathbf{P}_o &= \mathbf{P} - A(f)\mathbf{P}_o\mathbf{P}
\end{align*}
\]

- \(\mathbf{P}\) 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)

\[ P_o \approx P - A(f)PP \]

- \( P \): total up-going wavefield
- \( P_o \): primary wavefield
- \( A(f) \): “matching” operator
- \( P_{SRMP} \): SRME-produced primary wavefield
From SRME to Robust EPSI

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

\[
\min_A \sum_f \| P - A(f) PP \| \quad \text{SRMP}
\]

\( P \) total up-going wavefield

\( P_o \) primary wavefield

\( A(f) \) “matching” operator
From SRME to Robust EPSI

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

\[
\mathbf{P}_o = \mathbf{P} - A(f)\mathbf{P}_o\mathbf{P}
\]

- \( \mathbf{P} \): 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)

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

- \( P \): total up-going wavefield
- \( P_o \): primary wavefield
- \( A(f) \): “matching” operator

recorded data  predicted data from SRME
From SRME to Robust EPSI

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

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

\[
P_o = QG
\]

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

**Symbols:**
- \(P\): total up-going wavefield
- \(Q\): down-going source signature
- \(G\): primary impulse response
From SRME to Robust EPSI

Based on *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 predicted data from SRME
From SRME to Robust EPSI

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

Inversion objective:

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

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

Inversion objective:

\[ P = QG - GP \]

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

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

**Inversion objective:**

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

![Graph showing recorded data and predicted data from SRME]
From SRME to Robust EPSI

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

Inversion objective:

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

- **P** = predicted data from SRME
- **Q** = recorded data
- **f** = objective function

Position (m) | Time (s)
---|---
0 | 0.2
500 | 0.4
1000 | 0.6
1500 | 0.8
2000 | 1.0
1.2
1.4
From SRME to Robust EPSI

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

Inversion objective:

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

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

Inversion objective:

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

Table:

<table>
<thead>
<tr>
<th>Position (m)</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>500</td>
<td>0.4</td>
</tr>
<tr>
<td>1000</td>
<td>0.6</td>
</tr>
<tr>
<td>1500</td>
<td>0.8</td>
</tr>
<tr>
<td>2000</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Two ways to obtain the final primary wavefield

“Direct” Primary

\[ Q_G = P + G_P \]

“Conservative” Primary

Inversion objective:

\[ f(G, Q) = \frac{1}{2} \| P - (Q_G - G_P) \|_2^2 \]
From SRME to Robust EPSI

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

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

\[
\mathcal{M}(g, q) := F_t^\dagger \text{BlockDiag}_{\omega_1 \ldots \omega_{nf}} [(q(\omega)I - P)^\dagger \otimes I] F_t g
\]

Inversion objective:

\[
f(g, q) = \frac{1}{2} \|p - \mathcal{M}(g, q)\|_2^2
\]
From SRME to Robust EPSI

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

Inversion objective:

\[
\begin{align*}
\text{recorded data} & \quad \text{predicted data from SRME} \\
\mathbf{P} & = \mathbf{QG} - \mathbf{GP}
\end{align*}
\]

\[
f(G, Q) = \frac{1}{2} \| \mathbf{P} - (\mathbf{QG} - \mathbf{GP}) \|_2^2
\]
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 \( \tau_k \) from the Pareto curve

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

\[
\mathbf{q}_{k+1} = \arg \min_{\mathbf{q}} \| \mathbf{p} - \mathcal{M}_{g_{k+1}} \mathbf{q} \|_2
\]
Solving the EPSI problem

Linearizations

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

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

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

In fact it is bilinear:

\[ M_{\tilde{q}} g = \mathcal{M}(g, \tilde{q}) \quad M_{\tilde{g}} q = \mathcal{M}(q, \tilde{g}) \]
Robust EPSI
L1-minimization approach to the EPSI problem

[Lin and Herrmann, 2013 Geophysics]

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

determine new $\tau_k$ from the Pareto curve

$g_{k+1} = \arg \min_g \| p - M_{q_k} g \|_2$ 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

[Lin and Herrmann, 2013 *Geophysics*]

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

determine new $\mathbf{\tau}_k$ from the Pareto curve

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

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

Emits sparse, or “deconvolved” solution
L1 projection and sparsity

variable g at beginning of LASSO

\[ g_{k+1} = \arg \min_g \| p - M_q g \|_2 \text{ s.t. } \| g \|_1 \leq \tau_k \]
variable $g$ at **end** of LASSO

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

Emits “deconvolved” solution
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

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

Zero–offset trace, 1140m

- IR Reference
- IR 1:2 trace, lowpass 30Hz
- IR 1:4 trace, lowpass 15Hz
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 \textbf{70-100 iterations} to converge (each iteration is doing 2 SRME multiple prediction), can we make it \textbf{FASTER}?

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

Previous Q is \textit{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

Original (dx = 15m)

<table>
<thead>
<tr>
<th>Position (m)</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>500</td>
</tr>
<tr>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td>2000</td>
<td>1500</td>
</tr>
<tr>
<td>3000</td>
<td>2000</td>
</tr>
</tbody>
</table>

2x decimated lowpass 30Hz

<table>
<thead>
<tr>
<th>Position (m)</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>400</td>
<td>400</td>
</tr>
</tbody>
</table>

4x decimated lowpass 15Hz

<table>
<thead>
<tr>
<th>Position (m)</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>600</td>
<td>600</td>
</tr>
</tbody>
</table>
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
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

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
Warm-starting/continuation from coarse solution

Example

Solution of full data

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

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
\(dx = 60m > 30m > 15m\)
Significant speedup from bootstrapping (in 2D)

Per-iteration FLOPs cost (one forward/adjoint):

\[
\begin{align*}
\text{Cost}(n) &= O(2n_t n^2 \log n_t) + O(n_f n^3) \\
&= 2 \text{ times FFT computing MCG \& sum in FX} \\
\text{Cost}\left(\frac{1}{2}n\right) &= \frac{1}{4} O(2n_t n^2 \log n_t) + \frac{1}{8} O(n_f n^3) \\
\text{Cost}\left(\frac{1}{4}n\right) &= \frac{1}{16} O(2n_t n^2 \log n_t) + \frac{1}{64} O(n_f n^3)
\end{align*}
\]
Significant speedup from bootstrapping (in 2D)

Per-iteration FLOPs cost (one forward/adjoint):

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

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

- 2 times FFT
- computing MCG & sum in FX

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

\[
\text{Cost} \left( \frac{1}{4} n, \frac{1}{4} n_f \right) = \frac{1}{16} O(2n_t n^2 \log n_t) + \frac{1}{128} O(n_f n^3)
\]
Significant speedup from bootstrapping (in 3D)

Per-iteration FLOPs cost (one forward/adjoint):

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

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

2 times FFT computing MCG & sum in FX

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

\[
\text{Cost} \left( \frac{1}{4} n, \frac{1}{4} n_f \right) = \frac{1}{256} O(2n_t n^4 \log n_t) + \frac{1}{8192} O(n_f n^6)
\]
Significant speedup from bootstrapping

Wall times

From full data

Bootstrapping from 4x decimated

75 iters
Significant speedup from bootstrapping

Wall times

- From full data: 75 iters
- 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

<table>
<thead>
<tr>
<th>Wall time (minute)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>35</td>
</tr>
<tr>
<td>40</td>
</tr>
</tbody>
</table>

75 iters

Bootstrapping from 4x decimated

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

Wall times

From full data

Bootstrapping from 4x decimated
Field data example
North Sea dataset
North sea data

**Shot gather and stack**

Stream data (regularized to fixed-spread data)

401 source and receiver

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)

- From full data: 70 iters
- Bootstrapping from 4x decimated: 20 iters
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
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