Solution of Time-Harmonic Wave Equation for Full-Waveform Inversion

Rafael Lago, Art Petrenko, Zhilong Fang, Felix Herrmann

EAGE Conference & Exhibition, June 17, 2014
Frequency Domain Full Waveform Inversion Overview
Define the **misfit function** as

\[
\min_{m \in A} \Phi(m) = \sum_{i=1}^{n_s} \| d^i - P_r u^i \|_W.
\]

for some norm $W$ (usually $L^2$ with some regularization), and the **gradient** of the reduced formulation

\[
\nabla \Phi(m) = \sum_{i=1}^{n_s} \mathcal{R} \left\{ (u^i)^T \left[ \frac{\partial A}{\partial m} \right]^T w^i \right\}.
\]

- $d^i$: observed data
- $u^i$: (approximated) computed data
- $m$: Earth parameters; what we are trying to invert!
Forward Modelling - Overview

Forward problem:

\[ u^i = A^{-1}(m)q^i \]

(computation of \( \Phi(m) \))

Backward problem:

\[ w^i = A^{-H}(m)P_r^H(d^i - u^i) \]

(computation of \( \nabla \Phi(m) \))

Both require a PDE solve, computed with an (sufficiently large) accuracy \( \varepsilon \).

- \( A(m) \) operator governing the physics of Earth
- \( d^i, u^i \) observed and computed data
- \( P_r \) restricts the computed data to the receivers
- \( q^i \) source
Frugal FWI Overview
Gradient-Descent with Errors

Let

\[ \nabla \tilde{\Phi}(m_k) = \nabla \Phi(m_k) + e_k \]

for some error \( e_k \). Then, for strongly convex problems:

\[ \Phi(m_k) - \Phi(m_\star) < a_k (\Phi(m_0) - \Phi(m_\star)) \]

\[ a_k = \max \left\{ c^k, \|e_k\|_2^2 \right\}, \quad 0 \leq c \leq 1 \]

where \( c \) is the condition number of the problem.

__________________________________________________________________________

[Friedlander and Schmidt, 2012]
Relaxing the Physics - Approximating $u^i$ and $w^i$

$$\Phi(m) = \sum_{i}^{n_s} \left\| d^i - P_r u^i \right\|_W$$

$$u^i \approx A^{-1}(m)q^i$$

$$\frac{\left\| A(m)u^i - q^i \right\|_2}{\left\| q^i \right\|_2} < \varepsilon$$
Relaxing the Physics - Approximating $u^i$ and $w^i$

\[
\Phi(m) = \sum_{i}^{n_s} \left\| d^i - P_r u^i \right\|_W
\]

\[
u^i \approx A^{-1}(m) q^i
\]

\[
\frac{\left\| A(m) u^i - q^i \right\|_2}{\left\| q^i \right\|_2} < \varepsilon
\]

We hope that “large” $\varepsilon_k$ can “convexify” $\Phi$
Relaxing the Physics - Approximating $u^i$ and $w^i$

\[ \Phi(m) = \sum_{i} \left\| d^i - P_r u^i \right\|_W \]

\[ u^i \approx A^{-1}(m) q^i \]

\[ \frac{\left\| A(m) u^i - q^i \right\|_2}{\left\| q^i \right\|_2} < \varepsilon \]

We hope that "large" $\varepsilon_k$ can "convexify" $\Phi$
and that $\Phi(m, \varepsilon_k)$ "converges" to $\Phi(m, 0)$
Choosing $\varepsilon_k$ - Approximating $u^i$ and $w^i$

$$|\Phi(m, \alpha^k \varepsilon) - \Phi(m, \alpha^{k+1} \varepsilon)| \leq \eta \Phi(m, \alpha^{k+1} \varepsilon)$$

$$u^i \approx A^{-1}(m)q^i$$

$$\frac{\|A(m)u^i - q^i\|_2}{\|q^i\|_2} < \alpha^{k+1} \varepsilon$$

(\textit{use the final tolerance to compute } w^i)

<table>
<thead>
<tr>
<th>Chosen Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$ = 0.5</td>
</tr>
<tr>
<td>$\epsilon$ = $10^{-2}$</td>
</tr>
<tr>
<td>$\eta$ = $5 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

[Herrmann et al., 2013]
Frugal FWI - Source Subsampling

\[ \tilde{\Phi}(m) = \sum_{i \in I_k} \left\| d^i - u^i \right\|_W \]

\[ I_k \subset \{1, 2, \ldots, n_s\}, I_k^# = b_k \]

\( I_k \) is chosen at random without replacement. The expected error is given by

\[ \| e_k \|_2 \propto \sqrt{\frac{1}{b_k} - \frac{1}{n_s}} \]

\[ b_k \sim \min \left\{ \left( \epsilon_k^k + \frac{1}{n_s} \right)^{-1}, n_s \right\}. \]

---

[Friedlander and Schmidt, 2012] and [Herrmann et al., 2013]
CGMN, CRMN & Kaczmarz Sweep
Kaczmarz (Double) Sweep - Overview

\[ u_{i+1} = u_i + \frac{\gamma (q_i - a_i^H u_i) a_i}{\| a_i \|^2} \quad \]

- \( q_i \) \( i \)th element of \( q \)
- \( a_i \) \( i \)th row of \( A \) as a column vector
- \( \gamma \) relaxation parameter \( \in (0, 2) \)

Kaczmarz sweeps guarantees convergence in a finite (possibly large) number of steps.
Kaczmarz (Double) Sweep - Overview

\[
Q = Q_1Q_2...Q_NQ_NQ_{N-1}...Q_1 \quad Q_i = I - \frac{\gamma}{\|a_i\|^2_2}a_ia_i^H
\]

\[
u := Qu + Rq \quad \implies \quad (I - Q)u = Rq
\]

- \(Q\) is symmetric positive definite
- We can use \textbf{CG} to solve this system
- Neither \(Q\) nor \(R\) need to be computed in practice
- Equivalent to using \textbf{CG on the normal equations}, preconditioned by SSOR
1 (CGMN).

\[ p_0 = r_0 = \text{dk}swp(A, u_0, b, \gamma) - u_0; \]

while not converged do

\[ q_k = p_k - \text{dk}swp(A, p_k, 0, \gamma); \]
\[ \alpha_k = \langle r_k, r_k \rangle / \langle p_k, q_k \rangle; \]
\[ u_{k+1} = u_k + \alpha_k r_k; \]
\[ r_{k+1} = r_k - \alpha_k q_k; \]
\[ \beta_k = \langle r_{k+1}, r_{k+1} \rangle / \langle r_k, r_k \rangle; \]
\[ p_{k+1} = r_k + \beta_k p_k; \]
\[ k = k + 1; \]

end while

---

Very low memory cost

Very simple implementation

Suitable for any matrix \( A \) (even nonsquare)

On CG:[Hestenes and Stiefel, 1952], on CGMN: [Björck and Elfving, 1979]
while not converged do

\[ Ar_k := r_k - d \cdot \text{kswp}(A, r_k, 0, \gamma); \]
\[ \beta_k = \frac{\langle r_k, Ar_k \rangle}{\langle r_{k-1}, Ar_{k-1} \rangle}; \]
\[ p_k = r_k + \beta_k p_{k-1}; \]
\[ Ap_k = Ar_k + \beta_k Ap_{k-1}; \]
\[ \alpha_k = \frac{\langle r_k, Ar_k \rangle}{\langle Ap_k, Ap_k \rangle}; \]
\[ u_{k+1} = u_k + \alpha_k r_k; \]
\[ r_{k+1} = r_k - \alpha_k q_k; \]
\[ k = k + 1; \]

end while

---

Very low memory cost

One extra vector storage

One extra inner product

Minimal residual properties

Numerical Experiment
Forward Modeling - SEG/EAGE Overthrust

Solution of Time-Harmonic Wave Equation for Full-Waveform Inversion

- 20.1×20.1×4.7 km$^3$
- 100m grid spacing
- $O(1.9 \times 10^6)$ points
- 3Hz, $n_\lambda = 7.2$
- $v_{\text{min}} = 2179m/s$
- $v_{\text{max}} = 6000m/s$
- PML: 15 points
Solution of Time-Harmonic Wave Equation for Full-Waveform Inversion

- CGMN prec. res.
- CRMN prec. res.
- CRMN error
- CGMN error

- $4 \times 4 \times 1.2 \ km^3$
- $20m$ grid spacing
- $\mathcal{O}(2.5 \times 10^6)$ points
- $3\text{Hz}$, $n_\lambda = 22.7$
- $v_{min} = 1365 m/s$
- $v_{max} = 4991 m/s$
- PML: 15 points
The True Stopping Criterion

CGMN/CRMN stops when

$$\frac{\|r_j\|_2}{\|r_0\|_2} < \alpha^{k+1} \epsilon$$

For $k$ satisfying

$$\left| \Phi(m, \alpha^k \epsilon) - \Phi(m, \alpha^{k+1} \epsilon) \right| \leq \eta \Phi(m, \alpha^{k+1} \epsilon)$$

Chosen Parameters

- $\alpha = 0.5$
- $\epsilon = 10^{-2}$
- $\eta = 5 \times 10^{-2}$
FFWI-CGMN vs. FFWI-CRMN - Overthrust

Table: Total number of iterations of CGMN and CRMN during the inversion for each frequency slice

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>CGMN Iterations</th>
<th>CRMN Iterations</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>23,403</td>
<td>19,846</td>
<td>18%</td>
</tr>
<tr>
<td>6</td>
<td>30,189</td>
<td>24,387</td>
<td>24%</td>
</tr>
<tr>
<td>8</td>
<td>34,724</td>
<td>26,265</td>
<td>32%</td>
</tr>
</tbody>
</table>
Conclusions & Future Work
Conclusions & Future Work

- Smaller error computed by CGMN does not bring any improvement to Frugal FWI
- CRMN seems to be a feasible option
- Does the performance gain of CRMN grow with the frequency?
- Does the same result hold for other kind of PDEs?
- Does this behaviour holds for other models?
- Does this behaviour holds for other heuristics?
Questions?
Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations.

Geometric aspects in the theory of Krylov subspace methods.

CG versus MINRES: An empirical comparison.

Hybrid deterministic-stochastic methods for data fitting.


Rafael Lago
Component-averaged row projections: A robust, block-parallel scheme for sparse linear systems.


CARP-CG: A robust and efficient parallel solver for linear systems, applied to strongly convection dominated pdes.

Parallel solution of high frequency Helmholtz equations using high order finite difference schemes.

Frugal full-waveform inversion: from theory to a practical algorithm.

Methods of conjugate gradients for solving linear systems.

3d finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study.

Relaxationsmethoden bester strategie zur losung linearer gleichungssystem.