A Dual Formulation of Wavefield Reconstruction Inversion for Large-Scale Seismic Inversion

Gabrio Rizzuti\textsuperscript{1}, Mathias Louboutin\textsuperscript{2}, Rongrong Wang\textsuperscript{3}, and Felix J. Herrmann\textsuperscript{2}

\textsuperscript{1} Georgia Institute of Technology, now Utrecht University
\textsuperscript{2} Georgia Institute of Technology
\textsuperscript{3} Michigan State University
Agenda

- Motivations:
  robustness and complexity

- Theory:
  FWI vs WRI vs WRI*

- Numerical experiments:
  acoustic, TTI, small 3D

- Discussion and conclusions
Full Waveform Inversion

\[
\min_{\mathbf{m}} J_{\text{FWI}}(\mathbf{m}) = \frac{1}{2} ||F(\mathbf{m})\mathbf{q} - \mathbf{d}||^2, \quad F(\mathbf{m}) = RA(\mathbf{m})^{-1}
\]


✔ 3D computations are affordable via large HPC systems (or cloud computing) [Witte et al., 2019]

✗ (Effectively) multimodal problem: it needs a good starting model!
Wavefield Reconstruction Inversion

\[
\min_{\mathbf{m}} J_{\text{WRI}}(\mathbf{m}, \tilde{\mathbf{u}}(\mathbf{m})) = \frac{1}{2} \| R\tilde{\mathbf{u}} - \mathbf{d} \|^2 + \frac{\lambda^2}{2} \| A(\mathbf{m})\tilde{\mathbf{u}} - \mathbf{q} \|^2
\]

[van den Berg, P. M., and Kleinman, R. E., 1997; van Leeuwen, T. and Herrmann, F. J., 2013]

✔ Better conditioning
Wavefield Reconstruction Inversion

\[
\min_{\mathbf{m}} J_{WRI}(\mathbf{m}, \tilde{\mathbf{u}}(\mathbf{m})) = \frac{1}{2} \| \mathbf{d} - \frac{1}{\lambda} \mathbf{A}(\mathbf{m}) \tilde{\mathbf{u}} \|_2^2 + \lambda^2 \| \mathbf{A}(\mathbf{m}) \tilde{\mathbf{u}} - \mathbf{q} \|_2^2
\]

[van den Berg, P. M., and Kleinman, R. E., 1997; van Leeuwen, T., and Herrmann, F. J., 2013]

✔ Better **conditioning**

✗ Augmented solver: hard to scale to 3D
   (explicit time-marching schemes?)
Prior art: extended-source formulation

\[
\min_{\mathbf{m}} \mathcal{J}_{\text{WRI}}(\mathbf{m}, \tilde{\mathbf{u}}) = \frac{1}{2} \| R\tilde{\mathbf{u}} - \mathbf{d} \|^2 + \frac{\lambda^2}{2} \| A(\mathbf{m})\tilde{\mathbf{u}} - \mathbf{q} \|^2
\]

[van den Berg, P. M., and Kleinman, R. E., 1997; van Leeuwen, T. and Herrmann, F. J., 2013]

= 

\[
\min_{\mathbf{m}} \mathcal{J}_{\text{WR1}}(\mathbf{m}, \tilde{\mathbf{q}}) = \frac{1}{2} \| F(\mathbf{m})\tilde{\mathbf{q}} - \mathbf{d} \|^2 + \frac{\lambda^2}{2} \| \tilde{\mathbf{q}} - \mathbf{q} \|^2
\]

[Wang et al, 2016, Huang et al, 2018]
Prior art: augmented state approximation

\[
\min_m \mathcal{J}_{\text{WRI-q}}(m, \bar{q}) = \frac{1}{2} \|F(m)\bar{q} - d\|^2 + \frac{\lambda^2}{2} \|\bar{q} - q\|^2
\]

[Wang et al, 2016]

\[
\begin{bmatrix}
F(m) \\
\lambda I
\end{bmatrix}
\begin{bmatrix}
\bar{q} \\
\lambda q
\end{bmatrix} = \begin{bmatrix}
d \\
\lambda q
\end{bmatrix} \implies \bar{q} \approx q + \frac{F(m)^H(d - F(m)q)}{\lambda^2}
\]

✔ (Approximated) augmented solver: scale to 3D
Prior art: augmented state approximation

\[
\min_{m} J_{WRI-q}(m, \tilde{q}) = \frac{1}{2} \| F(m) \tilde{q} - d \|^2 + \frac{\lambda^2}{2} \| \tilde{q} - q \|^2
\]

[Wang et al, 2016]

\[
\nabla_{m} J_{WRI-q} = \nabla_{m} J_{WRI-q}(m, \tilde{q}) + \frac{\partial \tilde{q}}{\partial m} H \nabla_{q} J_{WRI-q}(m, \tilde{q})
\]

✔ (Approximated) augmented solver: scale to 3D
✘ Gradient computation inconsistency
WRI*: denoising reformulation

\[ \min_{m,u} \frac{1}{2} \| A(m)u - q \|^2 \quad \text{s.t.} \quad \| Ru - d \| \leq \varepsilon \]

[Wang, R., and Herrmann, F. J., 2017]
WRI*: Lagrangian

\[
\min_{\mathbf{m}, \mathbf{u}} \frac{1}{2} \| A(\mathbf{m}) \mathbf{u} - \mathbf{q} \|^2 \quad \text{s.t.} \quad \| R \mathbf{u} - \mathbf{d} \| \leq \varepsilon
\]

[Wang, R., and Herrmann, F. J., 2017]

\[
\max_{\mathbf{y}} \min_{\mathbf{m}, \mathbf{u}} \mathcal{L}(\mathbf{m}, \mathbf{u}, \mathbf{y}) = \frac{1}{2} \| A(\mathbf{m}) \mathbf{u} - \mathbf{q} \|^2 + \mathbf{y} \cdot (R \mathbf{u} - \mathbf{d}) - \varepsilon \| \mathbf{y} \|
\]
WRI*: Lagrangian

\[
\min_{m,u} \frac{1}{2} \|A(m)u - q\|^2 \quad \text{s.t.} \quad \|Ru - d\| \leq \varepsilon
\]

[Wang, R., and Herrmann, F. J., 2017]

\[
\max_{y} \min_{m} \tilde{L}(m, y) = \frac{1}{2} \|A(m)\tilde{u} - q\|^2 + y \cdot (R\tilde{u} - d) - \varepsilon \|y\|
\]

\[
A(m)\tilde{u} = q + F(m)^H y
\]
WRI*: augmented state approximation

\[
\min_{m,u} \frac{1}{2} \| A(m)u - q \|^2 \quad \text{s.t.} \quad \| Ru - d \| \leq \varepsilon
\]

[Wang, R., and Herrmann, F. J., 2017]

\[
\min_m \tilde{\mathcal{L}}(m) = \frac{1}{2} \| A(m)\tilde{u} - q \|^2 + \tilde{\mathbf{y}} \cdot (R\tilde{u} - d) - \varepsilon \| \tilde{\mathbf{y}} \|
\]

\[
\tilde{\mathbf{y}} \propto F(m)q - d
\]

can be differentiated through...
WRI*: reduced formulation

\[
\min_{\mathbf{m}} \tilde{\mathcal{L}}(\mathbf{m}) = \frac{1}{2} \| A(\mathbf{m}) \tilde{\mathbf{u}} - \mathbf{q} \|^2 + \tilde{\mathbf{y}} \cdot (R \tilde{\mathbf{u}} - \mathbf{d}) - \varepsilon \| \tilde{\mathbf{y}} \|
\]

- Requires standard wave equation solver: scale to 3D
- No gradient computation inconsistency
- Retains WRI robustness
- Roughly equivalent to 2x the computational cost of FWI
- Roughly equivalent to 2x the computational cost of FWI
- Model resolution generally inferior to WRI
Numerical examples: transmission
Numerical examples: transmission
Numerical examples: transmission
Numerical examples: Marmousi (freq. domain)
Numerical examples: Marmousi (freq. domain)
Numerical examples: Marmousi (freq. domain)
Numerical examples: Marmousi (freq. domain)
Numerical examples: Marmousi (freq. domain)
Numerical examples: Marmousi (freq. domain)
Numerical examples: Marmousi (freq. domain)
Numerical examples: BG Compass (acoustic)
Numerical examples: BG Compass (acoustic)
Numerical examples: BG Compass (acoustic)
Numerical examples: BG Compass (TTI)
Numerical examples: BG Compass (TTI, CIGs)
Numerical examples: small 3D

True

FWI

WRI*
Discussion

• WRI*: 
  - affordable version of WRI
  - retains robustness of WRI
  - 2\textsuperscript{nd} order methods/hybrid schemes to improve resolution

• Source-focusing annihilator might be necessary to avoid local minima [Symes, W. W., 2020]
Discussion

- Better approximation for augmented variable? Need for automatic differentiation! [Ablin et al., 2020]

\[ \nabla_m \tilde{\mathcal{L}} = \nabla_m \tilde{\mathcal{L}} + \frac{\partial \tilde{y}}{\partial m} \nabla_y \tilde{\mathcal{L}} \]

- Some special choice for covariance of data/model error avoid the need for augmented variable approximation
  [van Leeuwen, T., 2019]
Open source implementation

- [github.com/slimgroup/JUDI.jl](https://github.com/slimgroup/JUDI.jl)
- [devitoproject.org](http://devitoproject.org)
- [github.com/slimgroup/ImageGather.jl](https://github.com/slimgroup/ImageGather.jl)
References

- van Leeuwen, T., and Herrmann, F. J., Mitigating local minima in full-waveform inversion by expanding the search space, Geophysical Journal International 195.1 (2013)
- Peters, B., and Herrmann, F. J., and van Leeuwen, T., Wave-equation Based Inversion with the Penalty Method-Adjoint-state Versus Wavefield-Reconstruction Inversion, 76th EAGE Conference (2014)
- Symes, W. W., 2020, “Full Waveform Inversion by Source Extension: Why it works”
- Symes, W. W., 2020, "Wavefield Reconstruction Inversion: an example"
- Van Leeuwen, T., 2019, "A note on extended full waveform inversion"