Decentralized full-waveform inversion

A. Kadu[1], R. Kumar[2]
1Utrecht University, The Netherlands
2Georgia Institute of Technology, Atlanta, USA

2018 EAGE ANNUAL
80TH CONFERENCE + EXHIBITION
COPENHAGEN | DENMARK

Introduction

- Full-waveform inversion

$$\text{minimize} \quad F(\boldsymbol{m}) = \sum_{i=1}^{n_f} K_i(\boldsymbol{m}) + G(\boldsymbol{m})$$

where, $K_i(\boldsymbol{m}) = \|\boldsymbol{b}_i - F_i(\boldsymbol{m}, \boldsymbol{Q}_i)\|_2^2$

- The problem is non-linear and hence, solved using iterative methods

G(m) smooth	G(m) is non-differentiable
$\mathbf{m}^{k+1} = \mathbf{m}^k - \lambda_k \nabla\left(F(\mathbf{m}^k)\right)$	$\mathbf{m}^{k+1} = \prod_G\left(\mathbf{m}^k - \lambda_k \sum_{i=1}^{nf} \nabla\left(K_i(\mathbf{m}^k)\right)\right)$

- For a large-scale problem, the data is divided on local nodes and each update collects the function and gradient values from these nodes.

Challenges

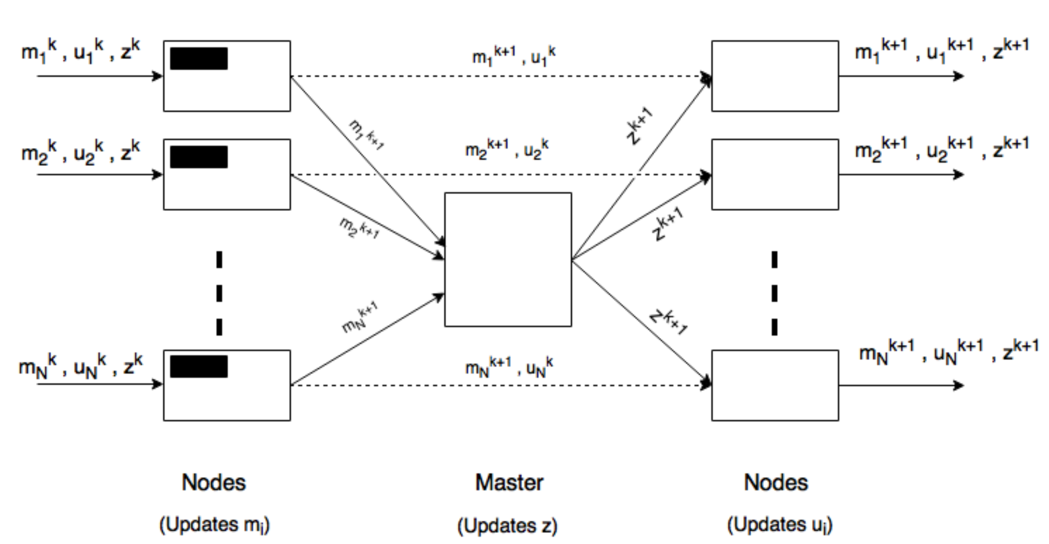
- Substantial network waiting time
- Data sharing concerns when working in multi-client environment
- Requirement of sophisticated optimizer at local nodes

Decentralized FWI

Proposed Optimization problem:

$$\text{minimize} \quad F(\{m_i\}, z) = \sum_{i=1}^{n_f} K_i(m_i) + G(z) \qquad \text{subject to} \quad m_i = z \quad \forall i$$

- Local variables, m_i, are handled by each node separately
- Global variable, z, at central node make sure that inversion is developing a global model
- Constraints $m_i = z$ help to decouple the problem
- The non-smooth regularizer can be handled by central node, hence reducing the requirement of sophisticated solver at local nodes.



Solving Through ADMM

Decentralized FWI Problem:

$$\text{minimize} \quad F(\{m_i\}, z) = \sum_{i=1}^{n_f} K_i(m_i) + G(z) \qquad \text{subject to} \quad m_i = z \quad \forall i$$

Augmented Lagrangian:

$$\mathcal{L}_\rho(\mathbf{m}_i, \mathbf{z}, \mathbf{u}_i) = \sum_{i=1}^{n_f} K_i(\mathbf{m}_i) + G(\mathbf{z}) + \frac{\rho}{2}\sum_{i=1}^{n_f} \|\mathbf{m}_i - \mathbf{z} + \mathbf{u}_i\|^2,$$

Note that the augmented Lagrangian is separable!

The ADMM steps:

$$\mathbf{m}_i^{k+1} = \arg\min_{\mathbf{m}_i}\left\{K_i(\mathbf{m}_i) + \frac{\rho}{2}\|\mathbf{m}_i - (\mathbf{z}^k - \mathbf{u}_i^k)\|^2\right\}, \qquad \text{for } i = 1, \ldots, n_f$$

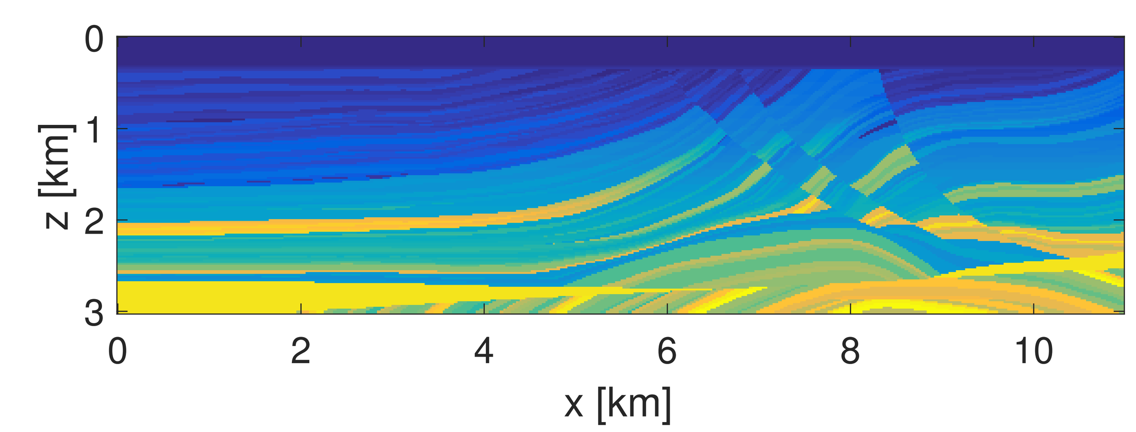
$$\mathbf{z}^{k+1} = \arg\min_{\mathbf{z}}\left\{G(\mathbf{z}) + \frac{n_f \rho}{2}\|\mathbf{z} - (\bar{\mathbf{m}}^{k+1} + \bar{\mathbf{u}}^k)\|^2\right\},$$

$$\mathbf{u}_i^{k+1} = \mathbf{u}_i^k + \mathbf{m}_i^{k+1} - \mathbf{z}^{k+1}, \qquad \text{for } i = 1, \ldots, n_f$$

Proximal Operator:

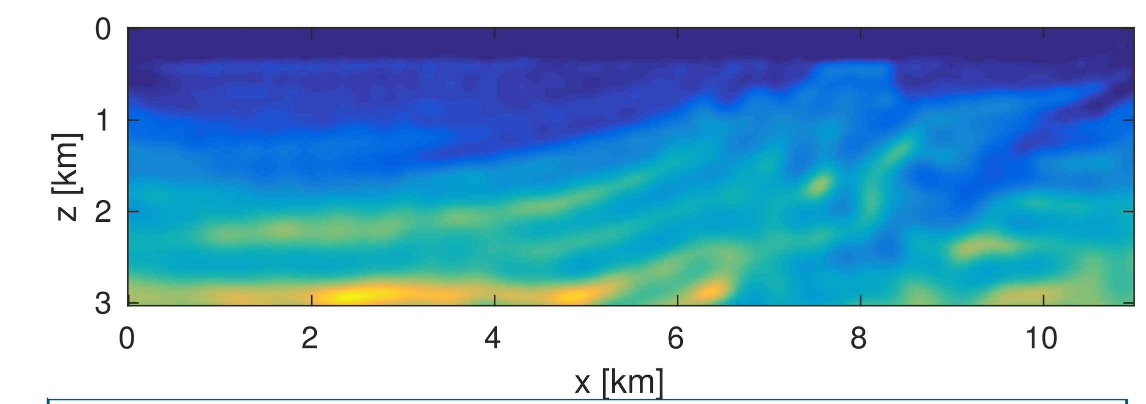
$$\text{prox}_{G,v}(\mathbf{y}) = \arg\min_{\mathbf{z}}\left\{G(\mathbf{z}) + \frac{v}{2}\|\mathbf{z} - \mathbf{y}\|^2\right\}$$

Results

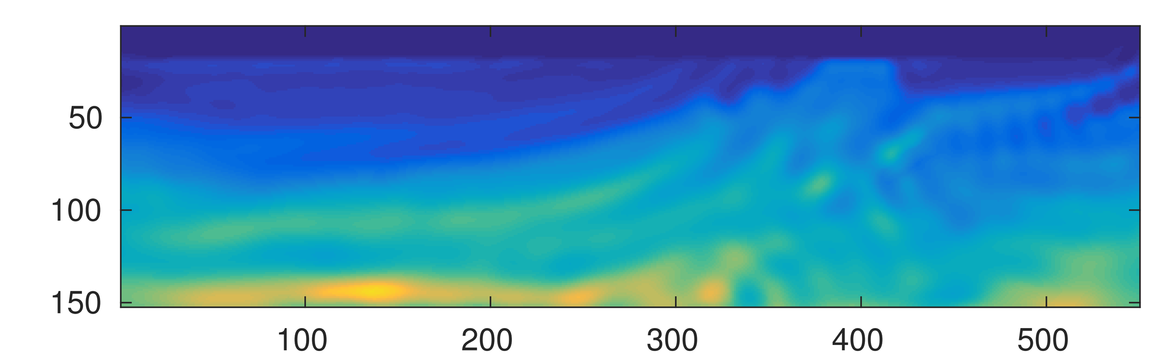


True Marmousi Model discretized on a 20 m x 20 m grid

Experimental Setup	
Number of Sources	110
Source grid spacing	100 m
Source depth	8 m
Number of receivers	220
Receiver grid spacing	50 m
Receiver depth	10 m
Ricker wavelet peak frequency	10 Hz
Phase shift	0^0
Frequencies	20 in [2,3] Hz



General FWI Results after 20 iterations



Decentralized FWI results after 5 global iterations with max 5 local iterations

	General FWI	Decentralized FWI
Model misfit	0.85	0.78
Data misfit	0.58	0.60
Total communications	400	80
Average waiting time	50 sec	4 sec

Conclusions

- Introduced a decentralized full-waveform inversion to handle *large-scale* data on distributed platform.
- The method separates decision variables into local and global components.
  - Local components are inverted on the local machines
  - Global updates are performed once local updates are done
- We propose ADMM method to solve the resulting problem
- The proposed method decouples the regularization from data fitting and reduces network time drastically.

Future Work

- Scalability: Experiments on 3D FWI
- Robustness: Is the proposed method robust to node failures?
- Multi-modality: Can we handle more parameters than just velocity? For example, density, anisotropy, etc.

References

Bertsekas, D.P. and Tsitsiklis, J.N. [1989] Parallel and distributed computation: numerical methods, 23

Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. [2011] Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning, 3(1), 1–122.

Combettes, P.L. and Pesquet, J.C. [2011] Proximal splitting methods in signal processing. In: Fixed- point algorithms for inverse problems in science and engineering, Springer, 185–212.

Moreau, J.J. [1965] Proximité et dualité dans un espace hilbertien. Bull. Soc. Math. France, 93(2), 273–299.

Nedic, A. and Ozdaglar, A. [2009] Distributed subgradient methods for multi-agent optimization. IEEE Transactions on Automatic Control, 54(1), 48–61.

Pratt, R.G. [1999] Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. Geophysics, 64(3), 888–901.

Schmidt, M., Berg, E., Friedlander, M. and Murphy, K. [2009] Optimizing costly functions with simple constraints: A limited-memory projected quasi-newton algorithm. In: Artificial Intelligence and Statistics. 456–463.

Virieux, J. and Operto, S. [2009] An overview of full-waveform inversion in exploration geophysics. Geophysics, 74(6), WCC1–WCC26.

Acknowledgements

The research is part of 'Computational Sciences for Energy Research'

FOM
NWO Netherlands Organisation for Scientific Research

RK would like to thank SINBAD II project members

SINBAD

We thank Tristan van Leeuwen for helpful suggestions.

Contact Information

Ajinkya Kadu – a.a.kadu@uu.nl - ajinkyakadu125.github.io
Rajiv Kumar – rkumar@eoas.ubc.ca -