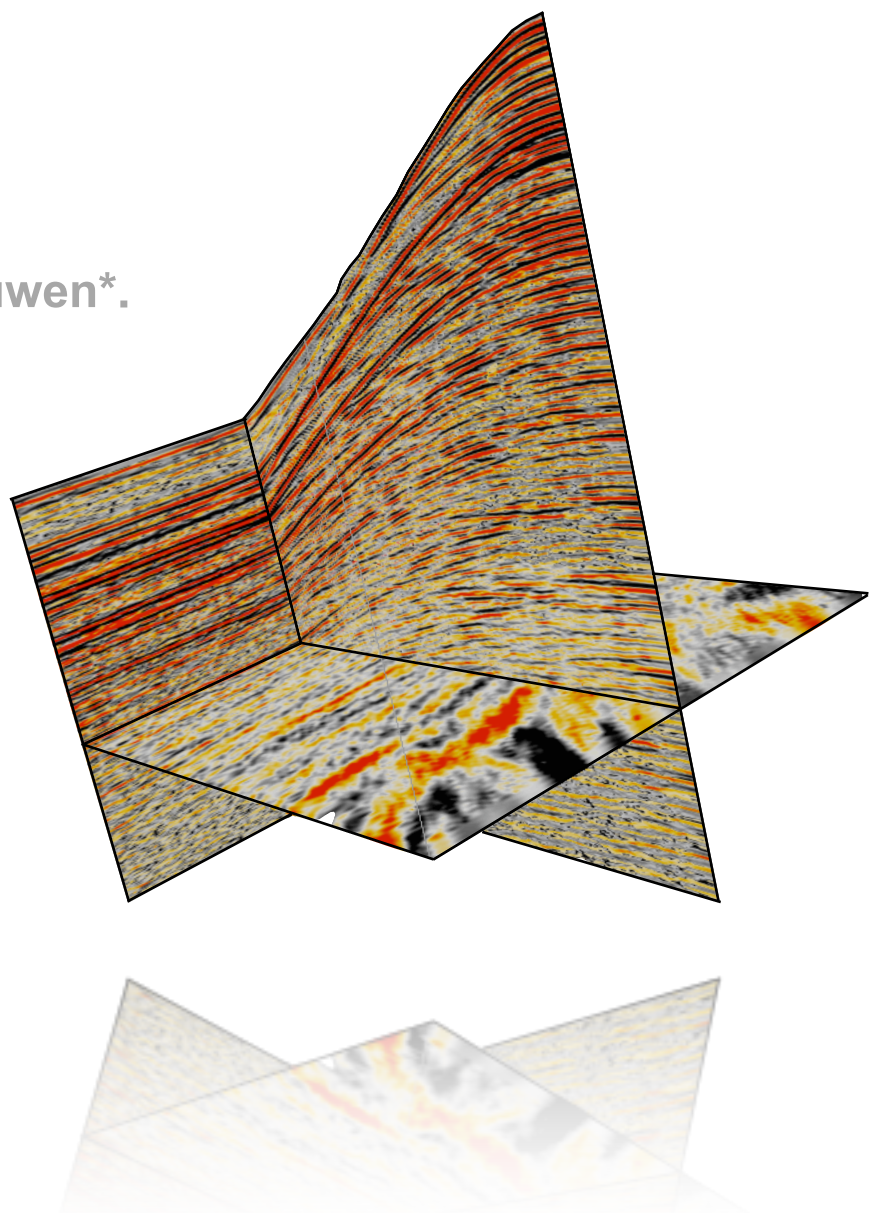
Sparsity-promoting GN method

Aleksandr Aravkin*
saravkin@eos.ubc.ca

Joint work with

Xiang Li*, Felix J. Herrmann*, Tristan van Leeuwen*.



***Seismic Laboratory for Imaging & Modeling**
Department of Earth & Ocean Sciences
The University of British Columbia

Sindad Consortium
Whistler, 2011

Outline

- FWI Mathematical Formulation
  - Convex-composite structure of inverse problem
- Quick Review of Gauss-Newton (GN) Method
- Motivations for extensions and modifications
  - Overwhelming data volume
  - Ill-posed problem and need for regularization
- Overview of Modified Gauss-Newton Method

Nonlinear Least Squares Formulation

- We consider inverse problems of the form

$$\mathbf{D} = \boldsymbol{\mathcal{F}}(\mathbf{m}; \mathbf{Q}) + \boldsymbol{\epsilon}$$

$\mathbf{D}$	$n \times m$ matrix of observations
$\mathbf{Q}$	$l \times m$ array of source parameters
$\mathbf{m}$	parameters to be recovered
$\boldsymbol{\mathcal{F}}(\mathbf{m}; \mathbf{Q})$	Forward model (calculated data)
$\boldsymbol{\epsilon}$	Model for error, typically Gaussian i.i.d.

- Objective below is **Convex-Composite:**

$$\min_{\mathbf{m}} \Phi(\mathbf{x}) = \|\mathbf{D} - \boldsymbol{\mathcal{F}}(\mathbf{m}; \mathbf{Q})\|_F^2 = \sum_{i=1}^{m} \|\mathbf{d}_i - \boldsymbol{\mathcal{F}}(\mathbf{m})\mathbf{q}_i\|_2^2$$

Overview of Gauss-Newton method

- Gauss-Newton method respects the convex-composite structure:

- Objective: $$\min_{\mathbf{m}} \Phi(\mathbf{m}) = \|\mathbf{D} - \boldsymbol{\mathcal{F}}(\mathbf{m};\mathbf{Q})\|_F^2$$

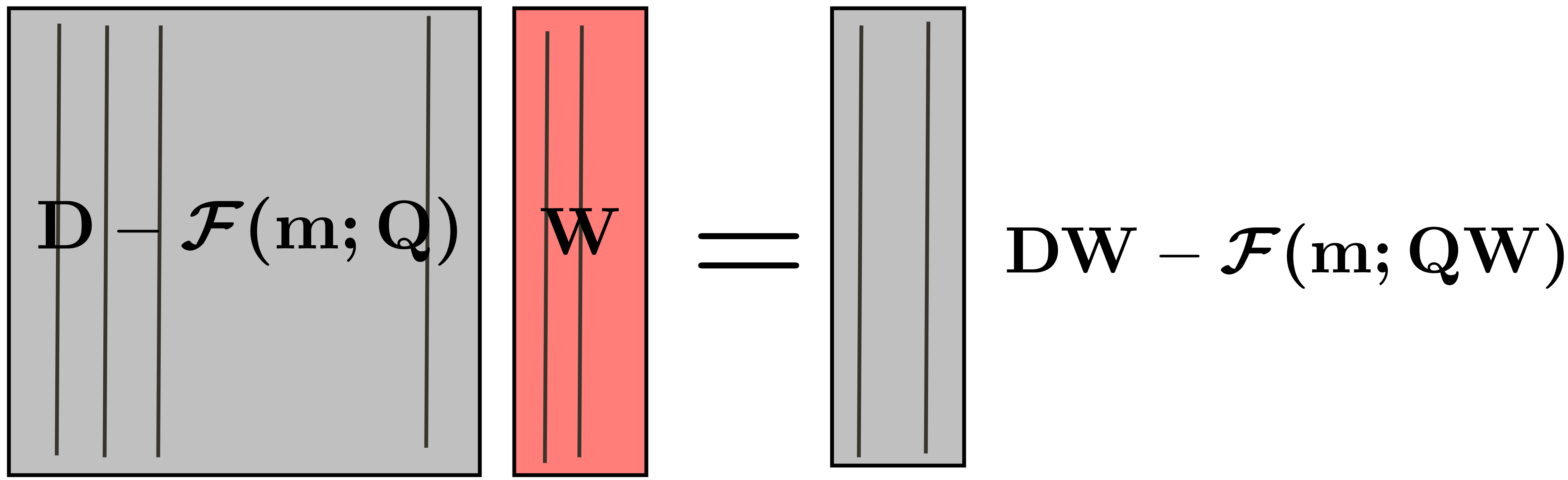
- Iterated algorithm: $$\mathbf{m}^{k+1} = \mathbf{m}^k + \alpha_k \overline{\boldsymbol{\delta}\mathbf{m}}$$

- Update $\overline{\boldsymbol{\delta}\mathbf{m}}$ solves $$\min_{\boldsymbol{\delta}\mathbf{m}} \|\mathbf{D} - \boldsymbol{\mathcal{F}}[\mathbf{m}^k;\mathbf{Q}] - \nabla\boldsymbol{\mathcal{F}}[\mathbf{m}^k;\mathbf{Q}]\boldsymbol{\delta}\mathbf{m}\|_F^2$$

Challenges and Solutions

- Overwhelming data volume motivates Dimensionality Reduction
  - Simultaneous shots (random mixtures of right hand sides)
  - Random sequential shots (for marine acquisition)
- FWI is an Ill-Posed problem
  - Many potential solutions explain the same data
  - Different optimization algorithms may go to different solutions from the same initial model.
- Challenge is to design methods that
  - take advantage of prior knowledge to go to the RIGHT solution!
  - work well in a dimensionality reduction regime

Dimensionality Reduction Revisited



- Taking **W** to be Gaussian corresponds to simultaneous shots
- Taking **W** to be random columns of **I** gives sequential shots
- To reduce dimensionality, we can work with:

$$\underline{\Phi}(\mathbf{x}) = \|\underline{\mathbf{D}} - \mathcal{F}(\mathbf{m}; \underline{\mathbf{Q}})\|_F^2 , \quad \underline{\mathbf{D}} = \mathbf{DW} , \quad \underline{\mathbf{Q}} = \mathbf{QW}$$

Gauss-Newton Method Revisited

- Focus on a dimensionality reduced problem:
- Objective: $$\min_{\mathbf{m}} \underline{\Phi}(\mathbf{m}) = \|\underline{\mathbf{D}} - \boldsymbol{\mathcal{F}}(\mathbf{m}; \underline{\mathbf{Q}})\|_F^2$$
- Iterated algorithm: $$\mathbf{m}^{k+1} = \mathbf{m}^k + \alpha_k \overline{\boldsymbol{\delta}\mathbf{m}}$$
- We can modify the Gauss-Newton subproblem in a number ways, and still prove convergence of the algorithm:

$$\overline{\boldsymbol{\delta}\mathbf{m}} \quad \text{solves} \quad \min_{\boldsymbol{\delta}\mathbf{m}} \|\underline{\mathbf{D}} - \boldsymbol{\mathcal{F}}[\mathbf{m}^k; \underline{\mathbf{Q}}] - \nabla\boldsymbol{\mathcal{F}}[\mathbf{m}^k; \underline{\mathbf{Q}}]\boldsymbol{\delta}\mathbf{m}\|_F^2$$

$$\text{subject to } \|\boldsymbol{\delta}\mathbf{m}\| \leq \tau^k$$

Promoting Curvelet Sparsity of UPDATES

- **Idea:** replace usual GN subproblem with a LASSO problem:

$$\min_{\mathbf{y}} \quad \|\underline{\mathbf{D}} - \mathcal{F}[\mathbf{m}^k;\underline{\mathbf{Q}}] - \nabla\mathcal{F}[\mathbf{m}^k;\underline{\mathbf{Q}}]\underbrace{\mathcal{C}^*\mathbf{y}}_{\boldsymbol{\delta}\mathbf{m}}\|_F^2$$

$$\text{s.t.} \qquad \|\mathbf{y}\|_1 \leq \tau$$

- **CS-Perspective (Dim. Red.):** we are trying to recover the TRUE Gauss-Newton update from a subsampled data volume.

- **Regularization perspective (Ill-Posed):** We force the Gauss-Newton subproblem to stop early, and customize it to promote sparsity of updates.

- Both perspectives exploit curvelet sparsity of updates.

Why Curvelet Sparsity of Updates?

- Expression for monochromatic FWI gradient:

$$\partial_{\mathbf{m}_i}\Phi(\mathbf{m}) = \bar{\mathbf{v}}^* \frac{\partial \mathbf{H}[\mathbf{m}]}{\partial \mathbf{m}_i}\bar{\mathbf{u}}$$

  – This gradient can be interpreted as a ‘correlation’ of forward an adjoint wavefields, and hence is still compressible in curvelets, even when the current model estimate is far from the truth.

- Closed form expression for Gauss-Newton update:

$$\boldsymbol{\delta}\mathbf{m} = -\left(\nabla\boldsymbol{\mathcal{F}}[\mathbf{m};\mathbf{Q}]^T\boldsymbol{\mathcal{F}}[\mathbf{m};\mathbf{Q}]^T\right)^{-1}\nabla\Phi(\mathbf{m})$$

  – Action of Gauss-Newton Hessian or its inverse is diagonal in phase space, so preserves the sparsity of the gradient in curvelets.

SLIM

Acknowledgements

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BP, Chevron, BGP, ConocoPhillips, PGS, Petrobras, Total SA, and WesternGeco.