Noise robust and time-domain formulations of Wavefield Reconstruction Inversion

Felix J. Herrmann
Noise robust and time-domain formulations of Wavefield Reconstruction Inversion

Mathias Louboutin, Peter Bas, Rongrong Wang, Emmanouil Daskalakis
Motivation

Full-waveform inversion (FWI):
- hampered by poor data & parasitic local minima
- ill-posed ↔ missing frequencies & finite aperture
- high-contrast high-velocity inclusions
- noise in data & errors in modeling

Notoriously difficult inverse problem:
- non-convex
- extremely large scale
Heuristic strategy

Extend & Project
- avoid local minima via extensions
- variable project to fit data

Squeeze
- impose physics & constrain the model

Cycle & Relax
- do multiple warm restarts while relaxing constraints
Heuristic strategy

Extend & Project
- avoid local minima via extensions
- variable project to fit data

Squeeze
- impose physics & constrain the model

Cycle & Relax
- do multiple warm restarts while relaxing constraints

No, this is not a yoga class...
Stylized example

Consider \( F(c)q = \left( \mathcal{F}^{-1} e^{i\omega \| x_1 - x_s \|_2 / c} \mathcal{F} q, \mathcal{F}^{-1} e^{i\omega \| x_2 - x_s \|_2 / c} \mathcal{F} q, \mathcal{F}^{-1} e^{i\omega \| x_3 - x_s \|_2 / c} \mathcal{F} q \right) \)

- seems harmless
- **not so** – oscillatory because of missing low frequencies
Stylized example – extension

Replace
\[ \min_m \frac{1}{2} \sum_{j=1}^{N_s} \| F(m)q_j - d_j \|^2 \]

by
\[ \min_{m, \Delta q} \frac{1}{2} \sum_{j=1}^{N_s} \| F(m)(q_j + \Delta q_j) - d_j \|^2 + \lambda^2 \| \Delta q_j \|^2, \]

with \( \Delta q = [\Delta q_1; \Delta q_2, \ldots, \Delta q_{N_s}] \) slack variables

- coincides with original solution since
  \[ \lambda \uparrow \infty, \quad \| \Delta q \|_2 \downarrow 0 \]

Proxy for extensions of AWI, WRI, and their variants...
Stylized example – extension

Solve by **projecting out** the slack variables

- modify objective for model parameters
- avoids cycle skips
Extensions

Wavefield Reconstruction Inversion:

\[
\min_{m,u} \frac{1}{2} \| Pu - d \|_2^2 + \frac{\lambda^2}{2} \| A(m) u - q \|_2^2
\]

› weak constraints
› “analysis” form

Matched Source Waveform Inversion:

\[
\min_{m,q} \frac{1}{2} \| PA^{-1}(m) q - d \|_2^2 + \frac{\lambda^2}{2} \| W q \|_2^2
\]

› weighted “synthesis” form & weights W focusses


Extensions

Wavefield Reconstruction Inversion:

$$\min_{m,u} \frac{1}{2} \| Pu - d \|^2 + \frac{\lambda^2}{2} \| A(m) u - q \|^2$$

- weak constraints
- “analysis” form

Matched Source Waveform Inversion:

$$\min_{m,q} \frac{1}{2} \| PA^{-1}(m)q - d \|^2 + \frac{\lambda^2}{2} \| W q \|^2$$

- weighted “synthesis” form & weights $W$ focusses

Challenge is to set the Lagrange multiplier!
Tikhonov regularization

Add quadratic penalty terms:

\[
\minimize_m f(m) + \frac{\alpha}{2} \| R_1 m \|^2 + \frac{\beta}{2} \| R_2 m \|^2
\]

- well-known & successful technique
- is differentiable
- not an exact penalty
- gradient & Hessian may become ill-conditioned
- requires non-trivial choices for hyper parameters
- not easily extended to edge-preserving $\ell_1$ - norms & bound constraints
- no guarantees that all model iterates are regularized

\[
\begin{align*}
\text{minimize} \quad & f(m) \\
\text{subject to} \quad & m \in C
\end{align*}
\]
minimize $f(m)$ subject to $m \in C$

**Algorithm 1** A Scaled Gradient Projection Algorithm:

1. $n = 0; m^0 \in C; \rho > 0; \epsilon > 0; \sigma \in (0, 1]$;
2. $H$ symmetric with eigenvalues between $\lambda_H^{\text{min}}$ and $\lambda_H^{\text{max}}$;
3. $\xi_1 > 1; \xi_2 > 1; c_0 > \max(0, \rho - \lambda_H^{\text{min}})$;
4. while $n = 0$ or $\frac{\|m^n - m^{n-1}\|}{\|m^n\|} > \epsilon$
   - $\Delta m = \arg\min_{\Delta m \in C-m^n} \Delta m^T \nabla F(m^n) + \frac{1}{2} \Delta m^T (H^n + c_n I) \Delta m$
   - if $F(m^n + \Delta m) - F(m^n) > \sigma (\Delta m^T \nabla F(m^n) + \frac{1}{2} \Delta m^T (H^n + c_n I) \Delta m)$
     - $c_n = \xi_2 c_n$
   - else
     - $m^{n+1} = m^n + \Delta m$
     - $c_{n+1} = \begin{cases} c_n & \text{if } \frac{c_n}{\xi_1} > \max(0, \rho - \lambda_H^{\text{min}}) \\ c_n & \text{otherwise} \end{cases}$
   - Define $H^{n+1}$ to be symmetric Hessian approximation with eigenvalues between $\lambda_H^{\text{min}}$ and $\lambda_H^{\text{max}}$
   - $n = n + 1$
5. end if
6. end while
Regularization w/ constraints

Add multiple constraints:

\[
\text{minimize } f(m) \quad \text{subject to } \quad m \in C_1 \bigcap C_2
\]

- not well-known in our community
- requires understanding of latest optimization techniques
- does not affect gradient & Hessian
- easier parameterization
- able to uniquely project onto intersection of multiple constraint sets
- constraints do not need to be differentiable
- constraints are satisfied at every model iterate

Jean Jacques Moreau
1923–2014


Reduced (2.5 X) BP model – modelling parameters

- number of sources: 132; number of receivers: 311
- receiver spacing: 40m, source spacing: 80m, max offset 11.5 km
- grid size: 20 m
- known Ricker wavelet sources with 15Hz peak frequency
- data available starting at 3 Hz
- 8 simultaneous shots w/ Gaussian weights w/ redraws
- starting model = smoothed true model
- inversion crime but poor data $\frac{\|\text{noise}\|_2}{\|\text{signal}\|_2} = 0.5$
1st cycle cycle

\[ \frac{\|\text{noise}\|_2}{\|\text{signal}\|_2} = 0.5 \]

**FWI**

**WRI**

**bounds only**

**bounds & TV**

Friday, October 6, 2017
2nd cycle

$\|\text{noise}\|_2 / \|\text{signal}\|_2 = 0.5$

bounds only

FWI

WRI

bounds & TV
3rd cycle

\[ \frac{\|\text{noise}\|_2}{\|\text{signal}\|_2} = 0.5 \]

**FWI**

**WRI**

**bounds only**

**bounds & TV**
Today’s agenda

Deal w/ “noise” by

- by handling source-side noise & modeling errors
- automatically select penalty parameter by exploiting duality

Move extensions to 3D

- time-domain WRI
- by exploiting duality
A denoising formulation of Full-Waveform Inversion

Rongrong Wang and Felix J. Herrmann
Motivation

Noises in observed data consist of:

- spatial and spectral discretization errors
- inaccurate PDE modeling, boundary reflections, multiples
- source estimation error, unknown & interfering sources
- trace truncation error
- timing error
- receiver location error
- measurement noise

- interior of the domain
- boundary
Motivation—the Failures of FWI

When measurement noise is "spiky"

(a) Model misfit for FWI
(b) Model misfit for inversion with Student’s t penalty

[Aravkin, A., van Leeuwen, T., & Herrmann, F J, 2012]
Motivation—the Failures of FWI

When water velocity is wrong

True model

FWI inversion

[Peters, A., & Herrmann, F J, 2014]
FWI & its relaxation

FWI requires strict satisfaction of the PDE:

$$\min_{m, u_i, i=1, \ldots, n_s} \sum_{i=1}^{n_s} \| P_{\Omega_i} u_i - d_i \|^2_2$$

subject to $A(m) u_i = q_i, i = 1, \ldots, n_s$

- Implicitly assumes that noise is Gaussian distributed along sources & receivers
- Neglects modeling errors
- Cannot accommodate prior information on noise level
- Becomes problematic when water velocity is wrong

$P_{\Omega_i}$: restriction operator
$q_i$: $i$th source
$A$: Time stepping or Helmholtz operator
$d_i$: Observed data for the $i$th source
$u_i$: wavefield associated to the $i$th source
FWI & its relaxation

Direct relaxation of PDE constraint:

\[
\begin{align*}
\min_{m, u_i, i=1, \ldots, n_s} & \sum_i \| P_{\Omega_i} u_i - d_i \|_2 \\
\text{subject to } & \| A(m) u_i - q_i \|_2 \leq \epsilon, i = 1, \ldots, n_s
\end{align*}
\]
FWI & its relaxation

Direct relaxation of PDE constraint:

$$\min_{m,u_i,i=1,...,n_s} \sum_i \|P_{\Omega_i} u_i - d_i\|_2$$

subject to $\|A(m)u_i - q_i\|_2 \leq \varepsilon, i = 1, \ldots, n_s$

Hard to choose!
FWI & its relaxation

Direct relaxation of PDE constraint

\[
\begin{align*}
\min_{m, u_i, i=1,\ldots,n_s} \sum_i \| P_{\Omega_i} u_i - d_i \|_2 \\
\text{subject to } \| A(m) u_i - q_i \|_2 \leq \epsilon, i = 1, \ldots, n_s
\end{align*}
\]

Hard to choose!

Flip the objective and the constraint

\[
\begin{align*}
\min_{m, u_i, i=1,\ldots,n_s} \| A(m) u_i - q_i \|_2^2 \\
\text{subject to } \| P_{\Omega_i} u_i - d_i \|_2 \leq \epsilon_i, i = 1, \ldots, n_s
\end{align*}
\]
FWI & its relaxation

Direct relaxation of PDE constraint

\[
\min_{m, u, i=1, \ldots, n_s} \sum_{i} \| P_{\Omega_i} u_i - d_i \|^2_2 \\
\text{subject to } \| A(m) u_i - q_i \|^2_2 \leq \varepsilon, i = 1, \ldots, n_s
\]

Flip the objective and the constraint

\[
\min_{m, u, i=1, \ldots, n_s} \| A(m) u_i - q_i \|^2_2 \\
\text{subject to } \| P_{\Omega_i} u_i - d_i \|^2_2 \leq \varepsilon_i, i = 1, \ldots, n_s
\]
FWI & its relaxation

Direct relaxation of PDE constraint
\[
\min_{m,u_i,i=1,\ldots,n_s} \sum_i \| P_{\Omega_i} u_i - d_i \|_2
\]
subject to \( \| A(m) u_i - q_i \|_2 \leq \epsilon_i, i = 1, \ldots, n_s \)

Flip the objective and the constraint
\[
\min_{m,u_i,i=1,\ldots,n_s} \| A(m) u_i - q_i \|_2^2
\]
subject to \( \| P_{\Omega_i} u_i - d_i \|_2 \leq \epsilon_i, i = 1, \ldots, n_s \)

Decompose wavefield variables
\[
u_i = P_{\Omega_i}^T P_{\Omega_i}^c u_i + P_{\Omega_i}^T P_{\Omega_i} u_i
\]
Boundary part

Interior part

Hard to choose!
Noise level

Decompose wavefield variables

Boundary part

Interior part
FWI & its relaxation

Direct relaxation of PDE constraint

\[
\min_{m,u_i,i=1,\ldots,n_s} \sum_i \| P_{\Omega_i} u_i - d_i \|_2
\]

subject to \( \| A(m) u_i - q_i \|_2 \leq \varepsilon_i, i = 1, \ldots, n_s \)

Flip the objective and the constraint

\[
\min_{m,u_i,i=1,\ldots,n_s} \| A(m) u_i - q_i \|_2^2
\]

subject to \( \| P_{\Omega_i} u_i - d_i \|_2 \leq \varepsilon_i, i = 1, \ldots, n_s \)

Decompose wavefield variables

\[
u_i = P_{\Omega_i}^T P_{\Omega_i} u_i + P_{\Omega_i}^T P_{\Omega_i}^c u_i + b_i
\]

- Boundary part
- Interior part

Hard to choose!
Noise level
FWI & its relaxation

Direct relaxation of PDE constraint

$$\min_{m,u_i,i=1,...,n_s} \sum_i \| P_{\Omega_i} u_i - d_i \|_2$$
subject to $\| A(m) u_i - q_i \|_2 \leq \varepsilon, i = 1, ..., n_s$

Flip the objective and the constraint

$$\min_{m,u_i,i=1,...,n_s} \| A(m) u_i - q_i \|_2$$
subject to $\| P_{\Omega_i} u_i - d_i \|_2 \leq \varepsilon, i = 1, ..., n_s$

Decompose wavefield variables

$$u_i = P_{\Omega_i}^T P_{\Omega_i} u_i + P_{\Omega_i}^T b_i$$

$$\min_{m,b_i,v_i,i=1,...,n_s} \| A(m) (P_{\Omega_i}^T v_i + P_{\Omega_i}^T b_i) - q_i \|_2$$
subject to $\| b_i - d_i \|_2 \leq \varepsilon, i = 1, ..., n_s$
The denoising formulation (FWIDN)

**Denoising formulation of FWI:**

$$\min_{m, b_i, v_i, i=1, \ldots, n_s} \|A(m)(P_{\Omega_i}^T v_i + P_{\Omega_i}^T b_i) - q_i\|^2_2$$

subject to $\|b_i - d_i\|_2 \leq \epsilon_i, i = 1, \ldots, n_s$

**Pros:**
- allows noise levels $\epsilon_i$ to vary with sources, and allows $\epsilon_i = 0$
- ensures reasonable PDE fidelity while preventing overfit
- all pros of WRI

**Cons:** algorithmically & computationally more demanding
FWI-DN – a more general form

Weighted/preconditioned least-squares objective:

\[
\min_{m,b_i,v_i,i=1,...,n_s} \left\| D_z (A(m)(P^T_{\Omega_i}v_i + P^T_{\Omega_i}b_i) - q_i) \right\|^2_2
\]

subject to \( \|b_i - d_i\|_2 \leq \epsilon_i, i = 1, ..., n_s \)

- \( D_z \) resharpe PDE misfit distribution
- Imposes looser PDE constraint at shallow part where the model is “noisier”
- Examples of \( D_z \): linear depth weighting, two-level depth weighting

\[
D_z f(x, z) = z f(x, z) \quad D_z f(x, z) = \chi_{z < z_0} f(x, z) + 2 \chi_{z \geq z_0} f(x, z)
\]
Solving FWI-DN

**Strategy:** alternatively update \( m \) and \( b_i, i = 1, \ldots, n_s \)

At iteration \( k \),

1. fix \( m^k \), solve for \( b_i^{k+1}, i = 1, \ldots, n_s \) from

\[
(b_i^{k+1}, v_i^{k+1}) = \arg \min_{b_i, v_i, i=1,\ldots,n_s} \| D_z(A(m^k)(P^T_{\Omega_i} v_i + P^T_{\Omega_i} b_i) - q_i) \|_2^2 \quad (P_d)
\]

subject to \( \| b_i - d_i \|_2 \leq \epsilon_i, i = 1, \ldots, n_s \)

2. for fixed \( b_i^{k+1}, i = 1, \ldots, n_s \), update \( m^k \) by solving \( T \) steps of

\[
\min_{m,v_i,i=1,\ldots,n_s} \sum_{i=1}^{n_s} \| D_z(A(m)(P^T_{\Omega_i} v_i + P^T_{\Omega_i} b_i^{k+1}) - q_i) \|_2^2 \quad (P_m)
\]
Solving for \((P_d)\) — a denoising step

\[
(P_d) \quad \text{Fix } m^k, \text{ solve for } b_i^{k+1} \text{ from}
\]

\[
(b_i^{k+1}, v_i^{k+1}) = \arg \min_{b_i, v_i} \|D_z(A(m^k)(P_{\Omega_i}^{T} v_i + P_{\Omega_i}^{T} b_i) - q_i)\|_2^2
\]

subject to \(\|b_i - d_i\|_2 \leq \epsilon_i\),
Solving for \((P_d)\) — a denoising step

\[
(P_d) \quad \text{Fix } m^k, \text{ solve for } b_i^{k+1} \text{ from }
\]
\[
(b_i^{k+1}, v_i^{k+1}) = \arg \min_{b_i,v_i} \| D_z(A(m^k)(P_{1i}^T v_i + P_{2i}^T b_i) - q_i) \|_2^2 \\
\text{subject to } \| b_i - d_i \|_2 \leq \epsilon_i,
\]

The Lagrangian dual of \((P_d)\) is

\[
\max_{\lambda \geq 0} \phi(\lambda)
\]

where

\[
\phi(\lambda) = \min_{u_i} \| D_z(A(m)u_i - q_i) \|_2^2 + \lambda \| P_{\Omega_i} u_i - d_i \|_2^2 - \lambda \epsilon_i
\]

The Lagrangian dual of \((P_d)\) is

\[
\max_{\lambda \geq 0} \phi(\lambda)
\]

where

\[
\phi(\lambda) = \min_{u_i} \|D_z(A(m)u_i - q_i)\|_2^2 + \lambda \|P_{\Omega_i} u_i - d_i\|_2^2 - \lambda \epsilon_i
\]

Solving for \((P_d)\) — a denoising step

Fix \(m^k\), solve for \(b_i^{k+1}\) from

\[
(b_i^{k+1}, v_i^{k+1}) = \arg \min_{b_i, v_i} \|D_z(A(m^k)(P_{\Omega_i}^T v_i + P_{\Omega_i}^T b_i) - q_i)\|_2^2
\]

subject to \(\|b_i - d_i\|_2 \leq \epsilon_i\).

Solving for \( (P_d) \) — a denoising step

\( \phi(\lambda) \) has closed-form gradient & Hessian

\[
\phi'(\lambda) = \|P_{\Omega_i} \bar{u}_i(\lambda) - d_i \|^2 - \epsilon_i \\
\phi''(\lambda) = -2(P_{\Omega_i} \bar{u}_i(\lambda) - d_i)^T P_{\Omega_i} C^{-1} P_{\Omega_i}^T (P_{\Omega_i} \bar{u}_i - d_i)
\]

where

\[
C = A(m)^T D_z^T D_z A(m) + \lambda P_{\Omega_i}^T P_{\Omega_i} \bar{u}_i(\lambda) = \begin{bmatrix} D_z(A(m^k)) \\ \sqrt{\lambda} P_{\Omega_i} \end{bmatrix}^\dagger \begin{bmatrix} D_z(q_i) \\ \sqrt{\lambda} d_i \end{bmatrix}
\]

Newton steps for \( \lambda \)

\[
\lambda^{k+1} = \lambda^k - \phi'(\lambda) / \phi''(\lambda)
\]

After finding the minimizer \( \lambda^* \), the primal optimizers are

\[
v_{i}^{k+1} = P_{\Omega_i} \bar{u}_i(\lambda^*), \quad v_{i}^{k+1} = P_{\Omega_i} \bar{u}_i(\lambda^*)
\]
Solving for \((P_m)\)

For fixed \(b^k_{i+1}, i = 1, \ldots, n_s\), update \(m^k\) by solving T steps of

\[
(P_m) \\ \min_{m, v_i, i=1, \ldots, n_s} \sum_{i=1}^{n_s} \|D_z(A(m)(P^T_{\Omega_i} v_i + P^T_{\Omega_i} b^k_{i+1}) - q_i)\|^2_2
\]
Solving for \((P_m)\)

For fixed \(b_{i}^{k+1}, i = 1, ..., n_s\), update \(m^k\) by solving \(T\) steps of

\[
(P_m) \min_{m,v_i, i = 1, ..., n_s} \sum_{i=1}^{n_s} \|D_z(A(m)(P_{\Omega_i}^Tv_i + P_{\Omega_i}^Tb_i^{k+1}) - q_i)\|_2^2
\]

Solve by variable projection

[Aravkin and van Leeuwen, 2012]
Solving for \((P_m)\)

For fixed \(b_i^{k+1}, i = 1, \ldots, n_s\), update \(m^k\) by solving \(T\) steps of

\[
(P_m) \quad \min_{m, v_i, i=1,\ldots,n_s} \sum_{i=1}^{n_s} \|D_z(A(m)(P_{\Omega_i}^T v_i + P_{\Omega_i}^T b_i^{k+1}) - q_i)\|^2_2
\]

Solve by variable projection \([\text{Aravkin and van Leeuwen, 2012}]\)

\[
\min_{m, v_i, i=1,\ldots,n_s} \sum_{i=1}^{n_s} \|D_z(A(m)(P_{\Omega_i}^T v_i + P_{\Omega_i}^T b_i^{k+1}) - q_i)\|^2_2 \equiv f(m, v_1, \ldots v_{n_s})
\]

\[
\iff \min_{m, v_i, i=1,\ldots,n_s} f(m, v_1, \ldots, v_{n_s}) = \min_m f(m, \bar{v}_1, \ldots, \bar{v}_{n_s}) \quad (\bar{v}_1, \ldots, \bar{v}_{n_s}) = \arg \min_{v_1, \ldots, v_{n_s}} f(m, v_1, \ldots, v_{n_s})
\]
Algorithm and Complexity

Inputs: $m_0, d_i, q_i, \ i = 1, \ldots, n_s, T, K$

For $\omega = \omega_1, \ldots, \omega_n$ do

- solve $(P_d)$ using $T$ iterations of Newton updates on $\lambda$
- perform $K$ gradient or L-BFGS updates on $m$ towards the minimizer of $(P_m)$

Endfor

On average, 1 update of $m$ requires: 2 PDE solves for FWI

- 2 PDE solves for WRI
- 3-5 PDE solves for FWI-DN
Test 1: robustness under non-uniform noise along sources
Test 1

- Frequency continuation using batches [3,3.5][3.5,4]....[14.5,15]Hz
- Source spacing: 240m
- Receiver spacing: 48m
- SNR=0 for low frequency data 3-10Hz
- SNR=25dB for high frequency data 10-15Hz
- Linear depth weighting
- Noise level $\epsilon_{sj} = 3\epsilon_{si}$ where $i = 1, \ldots, \lfloor n_s/2 \rfloor, j = \lfloor n_s/2 \rfloor + 1, \ldots, n_s$
Method for comparison: weighted FWI

\[
\min_m \sum_{i \in N_1} 9 \| P_{\Omega_i} A^{-1}(m) q_i - d_i \|_2^2 + \min_m \sum_{i \in N_2} \| P_{\Omega_i} A^{-1}(m) q_i - d_i \|_2^2
\]

where \( N_1 = \{1, \ldots, \lfloor \frac{n_s}{2} \rfloor \}, N_2 = \{ \lceil \frac{n_s}{2} \rceil + 1, \ldots, n_s \} \)
Test 1

Inverted model w/ weighted FWI

Inverted model w/ FWI-DN
Test 2: robustness under modeling error
Test 2

- Frequency continuation using batches [3,3.5][3.5,4]....[14.5,15]Hz
- Source spacing : 240m
- Receiver spacing : 48m
- Source depth: 12m
- True source q: ricker wavelet at 10Hz
- Source used for inversion: 0.8q
- Linear depth weighting
Test 2

Inverted model w/ FWI

Inverted model w/ FWI-DN with $\epsilon = 0$
Conclusion

- We proposed a denoising version of FWI
- We observed weighted/preconditioned PDE misfits dramatically increase robustness to modeling error
- The formulation makes incorporating prior knowledge of noise level convenient w/o increasing too much of complexity
Cost saving, time-domain dual formulation of WRI

Felix J. Herrmann, Mathias Louboutin, Peter Bas, Rongrong Wang, Emmanouil Daskalakis
Time domain implementation through dual formulation

Primal formulation
\[
\min_{U,m} \| A(m)U - Q \|_2 \\
\text{s.t.} \quad \| P_\Omega U - D \|_2 \leq \epsilon
\]

Dual formulation
\[
\min_m \left\{ \max_y \frac{1}{2} \| F^T(y) \|_2^2 + \langle y, D - FQ \rangle + \epsilon \| y \|_2 \right\}
\]

where \( F = P_\Omega A^{-1}(m) \)

Variable Size: \( O(N_xN_zN_TN_s) \) \( O(N_rN_sN_T + N_xN_z) \)

At optimal points: \( A(m)U^* = Q - F^T y^* \)

Algorithm
\[
G_y = FF^Ty + D - FQ + \epsilon \frac{y}{\| y \|_2} \\
G_m = \text{Jacobian}(m, Q - F^T y)
\]

Step 1: L-BFGS on \( y \)
Step 2: Gradient descent on \( m \)
Algorithm

Using the following gradients:

\[
\frac{\partial p(m, y)}{\partial y} = FF^T y + D - FQ + \epsilon \frac{y}{\|y\|_2}
\]

\[
\frac{\partial p(m, y)}{\partial m} = J^T (m, \tilde{Q}) y
\]

1. Solve with L-BFGS or gradient descent on the variable pair \((m, y)\)

2. Alternating updates of the two variables
Preliminary results

Original time-harmonic WRI

New time-domain WRI