Full waveform inversion in the frequency domain

Truncated Newton methods for simultaneous permittivity and conductivity reconstruction

X. Adriaens, C. Geuzaine
Université de Liège

Problem statement

Consider

- a model m
- a wave propagation operator $\mathcal{F}$
- a wavefield u
- a mesurement operator $\mathcal{R}$
- a dataset d

Full wave inversion consists in finding m^* such that

$$\mathcal{R}(u) = d \text{ with } \mathcal{F}(m^*)u = f$$

through the **optimization problem**

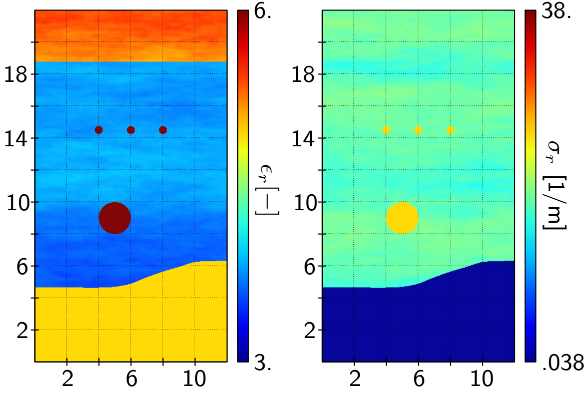
$$m^* = \arg\min_m \mathcal{J}(m) \triangleq \arg\min_m \text{dist}(\mathcal{R}(u(m)), d)$$

Crosshole radar example

Model space

The unknown model is here composed of both **permittivity** and **conductivity** distributions, *i.e.*

$$m \triangleq (\epsilon_r(\boldsymbol{x}), \sigma_r(\boldsymbol{x})).$$



Typical distributions[1] features highly contrasted targets ($\epsilon_{r,\text{pipe}} = 80$) embedded in a layered background with stochastic fluctuations.

[1]Ernst et al., "Full-Waveform Inversion of Crosshole Radar Data Based on 2-D Finite-Difference Time-Domain Solutions of Maxwell's Equations".

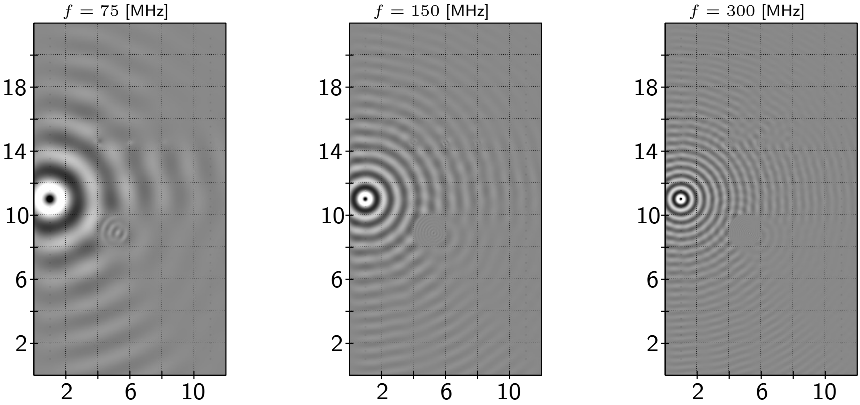
Wavefield space

In electromagnetics (2D TE), the wavefield is an **electric field** whose propagation can be modelled by the **Helmholtz equation**, *i.e.*

$$u \triangleq e(\boldsymbol{x})$$

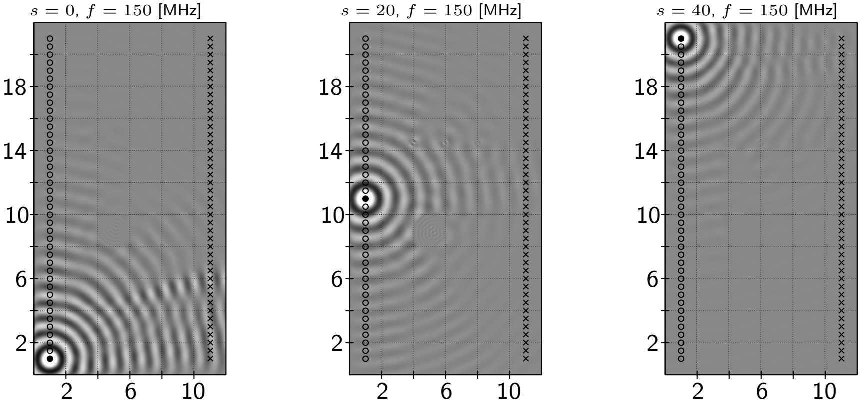
and

$$\mathcal{F}(u, m) \triangleq \Delta e + k^2 \left(\epsilon_r - j \frac{\sigma_r}{k}\right) e$$



Data space I

Several **emitters** (○) are buried in a first hole. Each emitter (•) is successively excited and the response is recorded at all the **receivers** (×), buried in another hole.



The response is recorded for several frequencies.

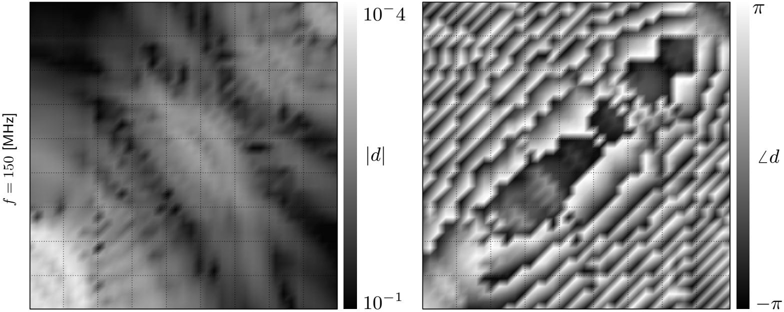
Data space II

A dataset d is thus a $n_s \times n_r \times n_f$ **complex-valued matrix**, *i.e*

$$d \in \mathbb{C}^{n_s \times n_r \times n_f}$$

which is obtained by **pointwise** measurements at the receivers, *i.e.*

$$[\mathcal{R}(u)]_{s,r,f} \triangleq e_{s,f}(\boldsymbol{x}_r).$$



Minimization problem

Descent direction I

Local optimization techniques must be used in practice because the search space is large (*e.g.* $\epsilon_r(\boldsymbol{x})$ and $\sigma_r(\boldsymbol{x})$).

Most algorithms are built on an approximation of the **Newton direction**

$$p \approx -\mathcal{H}^{-1}(j')$$

with j' the gradient and $\mathcal{H}$ the Hessian operator.

Both kernels result from the Taylor expansion of the performance functional

$$\begin{aligned}\mathcal{J}(m+\delta m) &\approx \mathcal{J}(m) + \{D_m\mathcal{J}\}(\delta m) + \frac{1}{2}\{D^2_{mm}\mathcal{J}\}(\delta m, \delta m)\\ &\approx \mathcal{J}(m) + \left\langle j', \delta m\right\rangle_M \quad + \frac{1}{2}\left\langle \mathcal{H}(\delta m), \delta m\right\rangle_M\end{aligned}$$

and thus **strongly depend on the chosen inner product** $\langle , \rangle_M$.

Descent direction II

Gradients and **Hessian operators** are defined by

$$\{D_m \mathcal{J}\}(\delta m) \triangleq \langle j', \delta m \rangle_M , \forall \delta m$$

and

$$\{D^2_{mm} \mathcal{J}\}(\delta m_1, \delta m_2) \triangleq \langle \mathcal{H}(\delta m_1), \delta m_2 \rangle_M , \forall \delta m_1 \forall \delta m_2$$

Changing the inner product is equivalent to **applying a preconditionner** to both kernels[2]. Preconditioning does not change the exact Newton direction but does change approximate solutions.

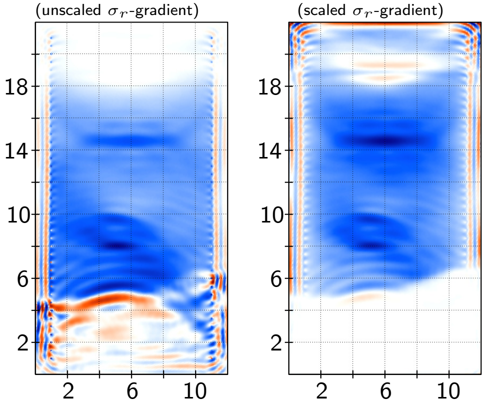
Incorporating **prior knowledge** in this inner product to encourage desired features can thus yield better approximate directions.

[2]Zuberi and Pratt, "Mitigating nonlinearity in full waveform inversion using scaled-Sobolev pre-conditioning".

Inner product examples

Example 1: spatial scaling

$$\langle \cdot, \cdot \rangle_M = \left\langle \cdot \sqrt{h_a(\boldsymbol{x})}, \sqrt{h_a(\boldsymbol{x})} \cdot \right\rangle_{L_2}$$



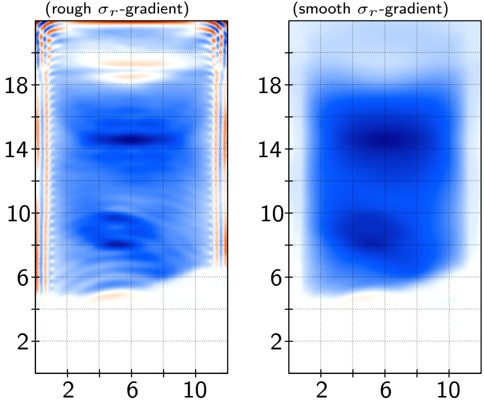
Consider an appropriate **spatial weight**, *e.g.* the diagonal of the Gauss-Newton Hessian[3] $h_a(\boldsymbol{x})$.

This inner product **restores balance** between gradient contributions everywhere.

[3] Pan, Innanen, and Liao, "Accelerating Hessian-free Gauss-Newton full-waveform inversion via l-BFGS preconditioned conjugate-gradient algorithm".

Example 2: spatial smoothing

$$\langle \cdot, \cdot \rangle_M = \left\langle \cdot \sqrt{h_a(\boldsymbol{x})}, \sqrt{h_a(\boldsymbol{x})} \cdot \right\rangle_{L_2} + \mu \left\langle \boldsymbol{\nabla} \cdot, \boldsymbol{\nabla} \cdot \right\rangle_{L_2}$$



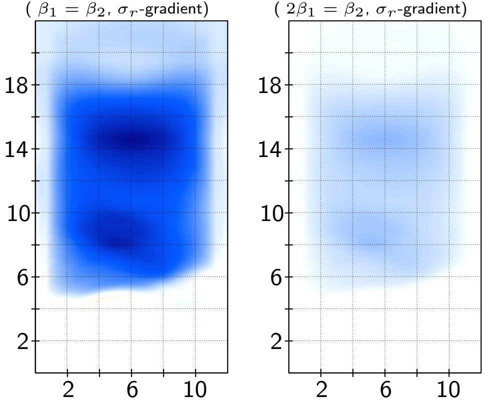
Adding such a **regularization term** penalizes rough models. Kernels w.r.t this inner product are therefore smoother.

Encouraging **smooth updates** early in inversion processes is a strategy to avoid local minima trapping[4].

[4] Zuberi and Pratt, "Mitigating nonlinearity in full waveform inversion using scaled-Sobolev pre-conditioning".

Example 3: parameter weighting

$$\langle m_a, m_b \rangle_M = \beta_1^2 \langle \epsilon_a, \epsilon_b \rangle + \beta_2^2 \langle \sigma_a, \sigma_b \rangle$$



The weight β governs the **relative importance** given to permittivities and conductivities.

By changing its value, it is possible to **reconstruct one parameter first** and again to avoid local minima trapping[5].

[5] Lavoué et al., "Two-dimensional permittivity and conductivity imaging by full waveform inversion of multioffset GPR data: a frequency-domain quasi-Newton approach".

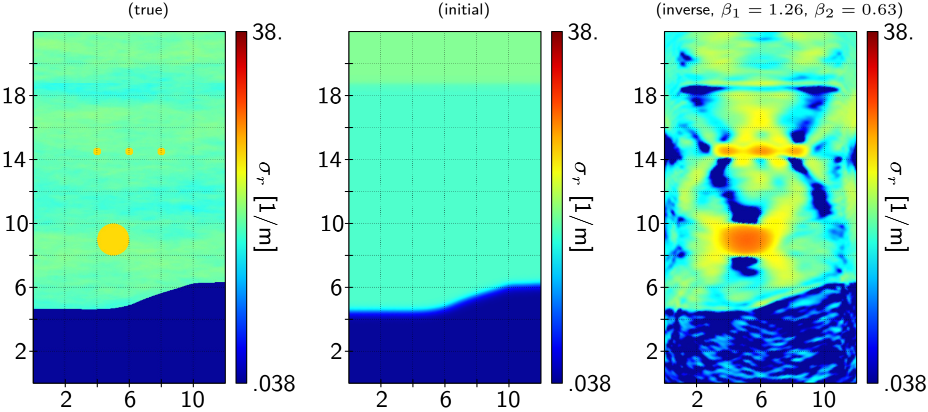
Crosshole radar application (synthetic)

Application choices

- 41 emitters and 41 receivers (spacing $\lessapprox$ background wavelength)
- 7 frequencies ($75, 90, 120, 150, 180, 225, 300$ [MHz])
- Individual/Sequential inversion from low to high frequency
- Minimization by a trust-region Newton method[6]
- Scaled, smoothed and weighted inner products
- Distance between data measured using least square norm
- Forward problems solved using finite elements
- Measured data generated synthetically (*i.e* with a numerical model)

[6] Adriaens, Métivier, and Geuzaine, "A trust-region Newton method for frequency-domain full waveform inversion Trust-Region Newton Method".

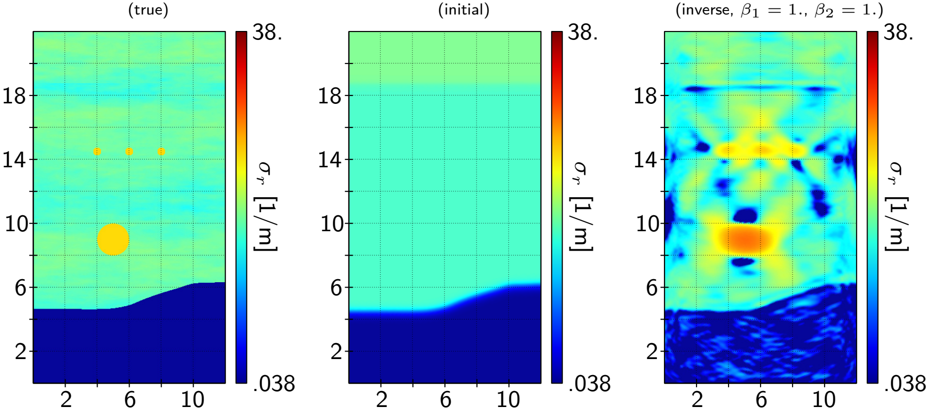
Results: conductivity



- Locations are correct
- Amplitudes are wrong
- Background is not resolved
- Background features artifacts

Reconstruction quality depends on β

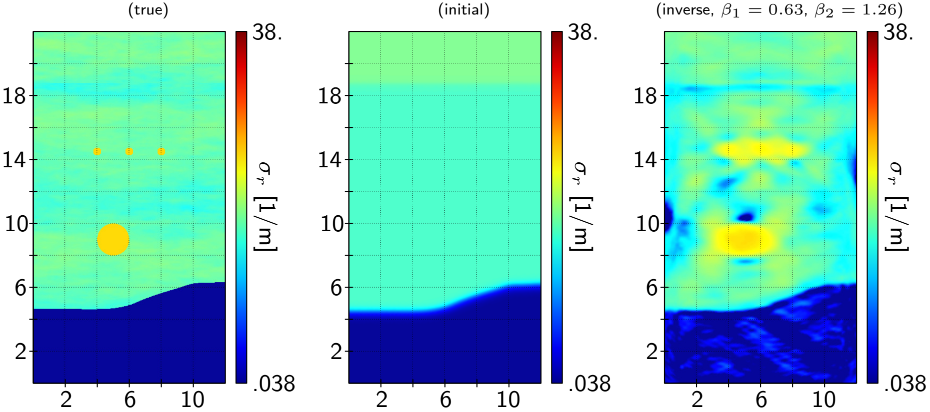
Results: conductivity



- Locations are correct
- Amplitudes are wrong
- Background is not resolved
- Background features artifacts

Reconstruction quality depends on β

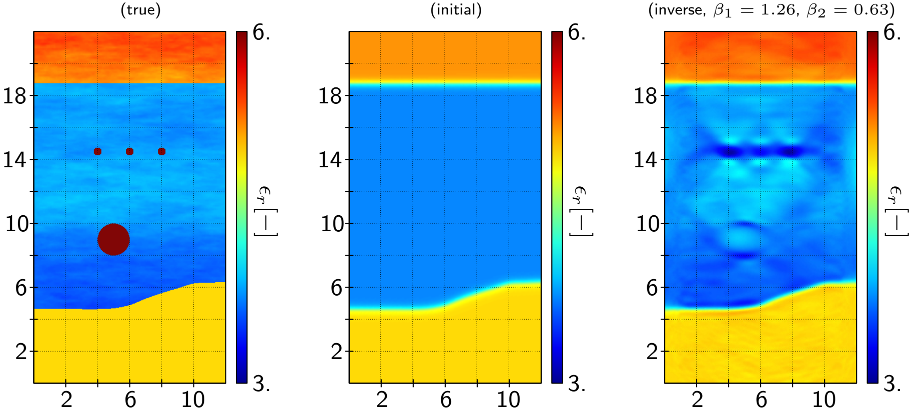
Results: conductivity



- Locations are correct
- Amplitudes are wrong
- Background is not resolved
- Background features artifacts

Reconstruction quality depends on β

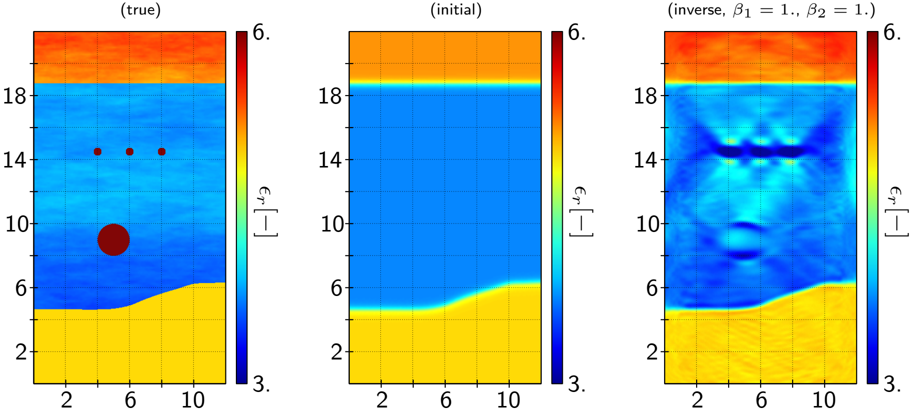
Results: permittivity



- Locations are correct
- Amplitudes are wrong
- Background is not resolved
- Background features artifacts

Reconstruction quality depends on β

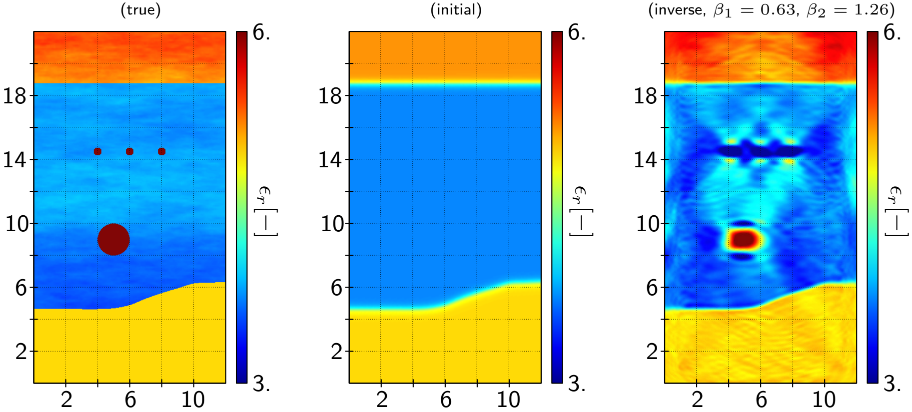
Results: permittivity



- Locations are correct
- Amplitudes are wrong
- Background is not resolved
- Background features artifacts

Reconstruction quality depends on β

Results: permittivity



- Locations are correct
- Amplitudes are wrong
- Background is not resolved
- Background features artifacts

Reconstruction quality depends on β

Conclusion

1. Pure Newton methods are attractive but prohibitively expensive
2. Truncated Newton methods must be used instead
3. Truncated Newton methods however require a preconditioning strategy
4. Modifying the inner product is an elegant preconditioning strategy
5. Modifying the inner product should be investigated further, especially to retrieve highly contrasted media

Conclusion

1. Pure Newton methods are attractive but prohibitively expensive
2. Truncated Newton methods must be used instead
3. Truncated Newton methods however require a preconditioning strategy
4. Modifying the inner product is an elegant preconditioning strategy
5. Modifying the inner product should be investigated further, especially to retrieve highly contrasted media

Thank you for your attention