Interferometric waveform inversion

Laurent Demanet

Department of Mathematics, MIT

with Vincent Jugnon (GE) and
Laurent Seppecher (MIT)

Thanks: MITEI seed fund, Earth Resources Lab (ERL), TOTAL SA

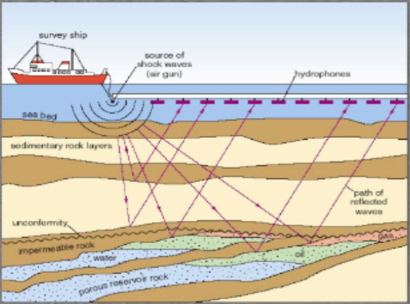
November 2014 – MITEI

The imaging and computing group

Fundamental aspects of inverse wave scattering.

- Scalable computation of high-frequency waves
- Randomized algorithms for HPC

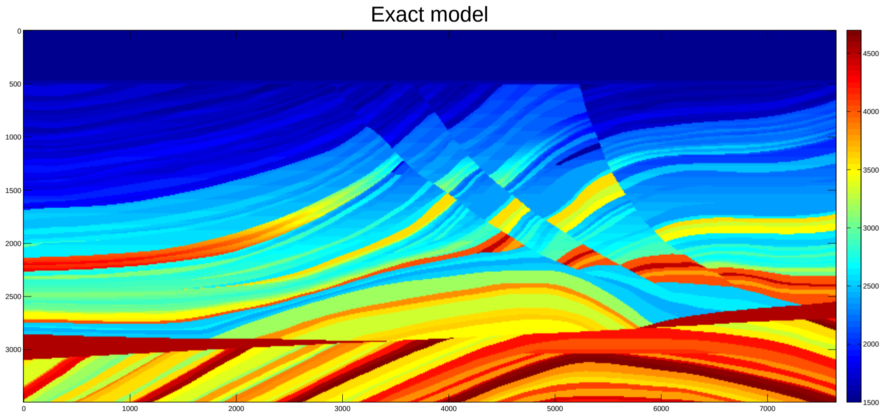
- Nonconvex optimization
- Data processing and learning



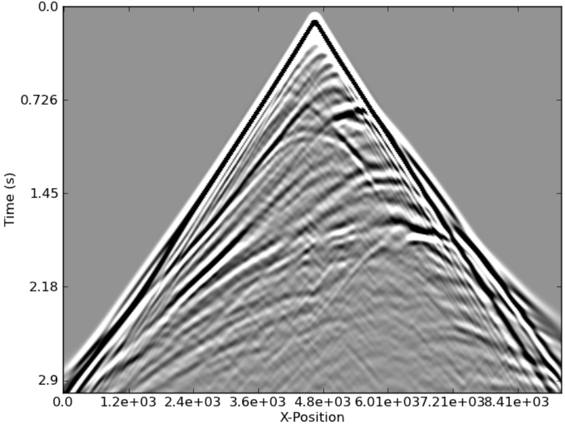
Exploration: surveys / borehole

Monitoring: reservoirs / CO_2 injection sites

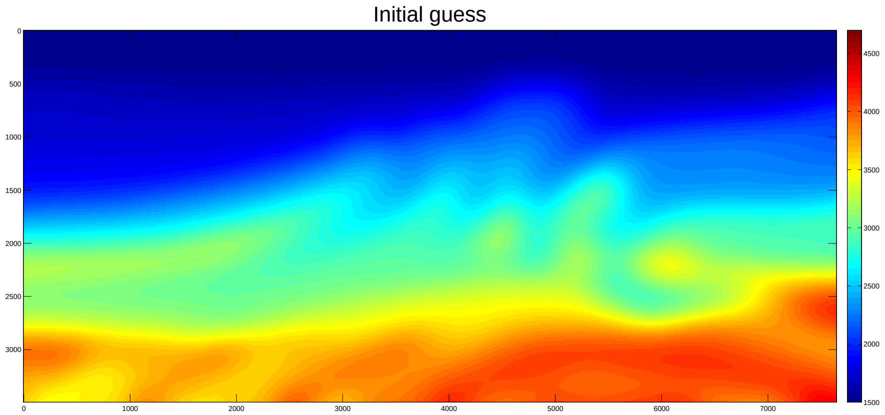
FWI vs. interferometric inversion – m_0



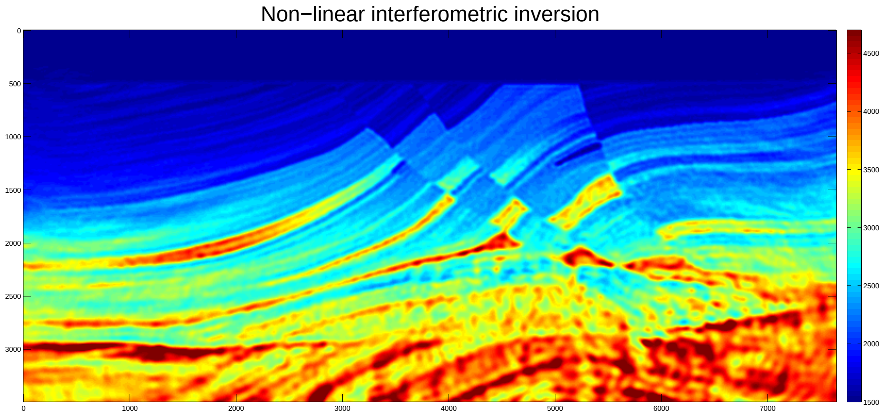
FWI vs. interferometric inversion – d



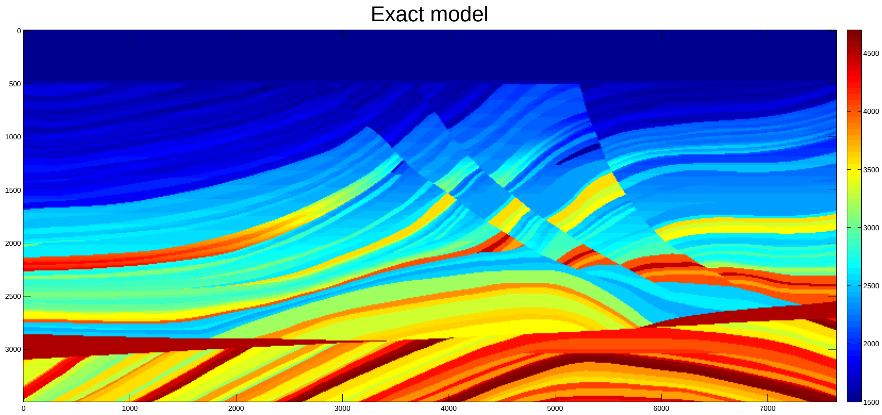
FWI vs. interferometric inversion – initial m



FWI vs. interferometric inversion – *m*



FWI vs. interferometric inversion – m_0



Why seismic imaging in a math dept?

Inverse scattering is challenging:

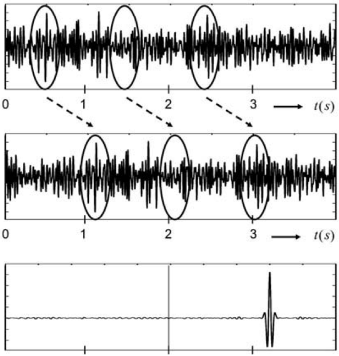
1. The optimization problems are nonconvex.
   Cartoon scenarios: no known reliable minimization method.
2. Current methods not designed for extreme uncertainties.
   Not clear what is data vs. noise.
3. Current algorithms do not scale.
   Small bang for your computer buck.

Past 10-20 years: increasing CS/math/stats components
in addition to geophysics

Interferometry: a new way to view data

Cross-correlations are robust to disordered kinematics

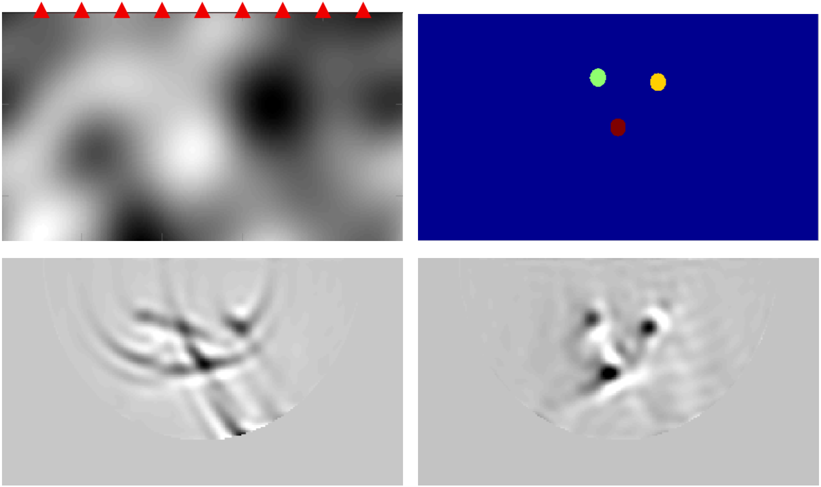
- Time reversal (phys): Fink et al. 1993.
- Time reversal (math) Bal, Papanicolaou, Ryzhik, 2002; Bal, Ryzhik 2003
- CINT imaging: Borcea, Papanicolaou, Tsogka, 2003, 2005
- Seismic interf. (phys): Weaver et al. 2001; Campillo, Paul 2003; Snieder, 2004; Wapenaar et al. 2006
- Seismic interf. (math): Bardos et al. 2008; Colin de Verdiere 2009; Garnier, Papanicolaou, Solna 2009+



Wapenaar et al. 2010

Interferometric inversion: leverage robustness of similar quantities for inverting the underlying physics

Locating microseisms with interferometry



Least-squares inversion   Interferometric inversion

Road ahead

Uncertainties: from quantification to rectification

Toward an understanding of

- How to rethink waveform inversion for robustness
- What is benign vs. unredeemable model ignorance
- New resolution scalings

Development of large-scale optimization methods

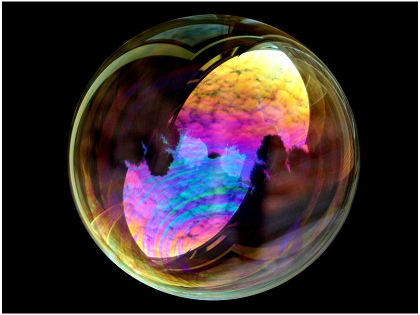
- Computing and data processing in high dimensions
- Close links to recent progress in polynomial optimization and information sciences
- Re-usable software

Bonus slides

What is interferometry?

- Optical interferometry: observe

$$|d_1(x,\omega) + d_2(x,\omega)|^2 = |d_1|^2 + 2\,\mathrm{Re}\, d_1(x,\omega)\overline{d_2(x,\omega)} + |d_2|^2$$



Example:

$$d_1(x,\omega) = e^{i\omega x/c}$$
$$d_2(x,\omega) = e^{i\omega(x+\Delta x)/c}$$
$$\Rightarrow d_1\overline{d_2} = e^{-i\omega(\Delta x)/c}$$

$\Delta x \sim$ micrometer

Interferometry is everywhere

Interferometry in real life:

- X-ray diffraction crystallography: phase retrieval
- Passive SAR imaging from opportunity sources
- Passive seismic imaging from ambient noise

In all cases, $d_1(\omega)\overline{d_2(\omega)}$ with a model $d = Fm$.

Cross-correlation: $d_1(\omega)\overline{d_2(\omega)} = FT$ of $\int d_1(\tau)\overline{d_2(\tau - t)}d\tau$.

More generally, $d(r_1, s_1, \omega_1)\overline{d(r_2, s_2, \omega_2)}$.

Why interferometry in inverse problems?

Either:

- Observe $d_1(\omega)\overline{d_2(\omega)}$:
  measurements come in this form.
- Decide to form $d_1(\omega)\overline{d_2(\omega)}$:
  leverage stability of quadratic quantities.

Agenda

1. How to reliably solve such quadratic systems

Stability to data errors

Measurements and inverse problems

Linear measurements from an imaging device: $d_i = (Fm_0)_i + e_i$, $i = (r, s, \omega)$. Usual proposal:

$$\min_m \sum_i |d_i - (Fm)_i|^2$$

- -

Invert for m, but leverage the model robustness of $d_i \overline{d_j}$. Form interferometric measurements: $D_{ij} = d_i \overline{d_j}$ for some pairs (i, j). Simplest expression:

$$\min_m \sum_{(i,j) \in E} |D_{ij} - (Fm)_i \overline{(Fm)_j}|^2$$

for some incomplete set E of data pairs. Compensates for *some* uncertainties in F.

Fitting linear vs. interferometric measurements

$$\min_{m} \sum_{(i,j)\in E} |D_{ij} - (Fm)_i \overline{(Fm)_j}|^2$$

Problem 1: quartic objective with spurious local minima

Solution: relaxation

- Lifting: new unknown M for mm^*, subject to $\text{rank}(M) = 1$.
- Semidefinite relaxation: constrain instead to $M \succeq 0$.
- When done, compute leading eigenvector of M, get $e^{i\alpha} m$.

Turned the problem into one of rank-1 matrix recovery

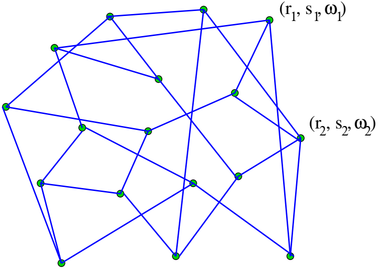
(links to Lasserre 2001; Chai et al. 2010; and Candès et al, 2011)

Fitting linear vs. interferometric measurements

$$\min_m \sum_{(i,j)\in E} |D_{ij} - (Fm)_i \overline{(Fm)_j}|^2$$

Problem 2: size of E needs to be kept in check: N^2 vs N.

Solution: view E as the set of edges of a graph G, and pose the problem as one of graph design.



Interferometric waveform inversion

Well-posedness: G should be well connected.

- Adjacency matrix: $A_{ij} = 1$ if $(i, j) \in E$, zero otherwise
- Laplacian matrix: $N - A$, with N a diagonal of node degrees
- Data-weighted Laplacian matrix: $L_{ij} = |d_i||d_j|$, $i \neq j$.

Note $\lambda_1(L) = 0$.

Definition

G is an expander graph if L has a large spectral gap $\lambda_2(L)$, i.e., when $\lambda_2(L)$ is a nonnegligible fraction of $\lambda_N(L)$.

Interferometric waveform inversion

Let $0 = \lambda_1(L) < \lambda_2(L) < \ldots$ be the egval of the data-weighted L.

Theorem (D, Jugnon, 2013)

Consider noisy data $D_{ij} = (Fm_0)_i \overline{(Fm_0)_j} + \epsilon_{ij}$. *Assume* F *is invertible. Assume* G *is connected with loops. Consider any method that imposes* $\sum_{(i,j)\in E} |D_{ij} - (FMF^*)_{ij}| \leq \sigma$ *with* $M \succeq 0$, *then defines* m *as the leading eigenvector of* M. *Assume* $\|\epsilon\|_1 + \sigma \leq \lambda_2(L)/2$. *Then*

$$\frac{\|m - e^{i\alpha} m_0\|}{\|m_0\|} \leq 15\, \kappa(F)^2 \sqrt{\frac{\|\epsilon\|_1 + \sigma}{\lambda_2(L)}}$$

Compare with least-squares: $\frac{\|m - m_0\|}{\|m_0\|} \leq \kappa(F)\, \frac{\|e\|}{\|d\|}$.

Implications

Proof: concentration of a convex combination.
Alternatively, L is a dual certificate $\quad(L1 = 0, \quad L \succ 0 \text{ on } 1^{\perp})$

Corollary

Phase retrieval from $|(Fm_0)_i + e^{ik2\pi/3}(Fm_0)_j|^2 + \epsilon_{ij}$, $k = 1, 2, 3$.
Depending on how G is chosen, the recovery is

- *minimally robust with $3N$ measurements ($1/\sqrt{\lambda_2} \sim N$);*
- *robust with $3\frac{d_G}{2}N$ measurements (d_G = node degree)*

(see PhaseLift, Candès et al. 2011)

Erdos-Renyi graph: expected $\lambda_2 > c > 0$ with degree $O(\log N)$.

Related: Singer et al. 2011

Use a nonsmooth iterative method

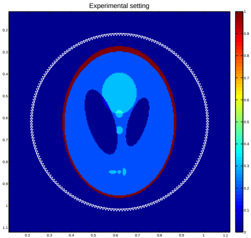
Practical numerical aspects:

- Keep M separated throughout: let $M = RR^T$ with $\text{rank}(R) = 2$ (Burer-Monteiro 2003). Quasi-Newton: LBFGS.
- Convergence: slowdown vs. LS of a factor 10-100.

Agenda

2. *Why is interferometric inversion interesting*

New way of handling model errors (errors in F)



Radar / seismic / ultrasound forward scattering map $\mathcal{F} : m \mapsto d$.

- m: reflectivity map, $m = 1/c^2$.
- d: wavefield data $d_{r,s,\omega}$.
- $\mathcal{F}$: solve a wave equation and sample at receivers. Linearized as $F = \delta\mathcal{F}/\delta m$. Explicitly:

$$d = Fm$$

$$\Leftrightarrow$$

$$(m_0 \partial_{tt} - \Delta)\, u_s = -m\, \partial_{tt} u_{0,s}, \qquad d_{r,s,\omega} = \hat{u}_s(x_r, \omega)$$

Interferometric inversion

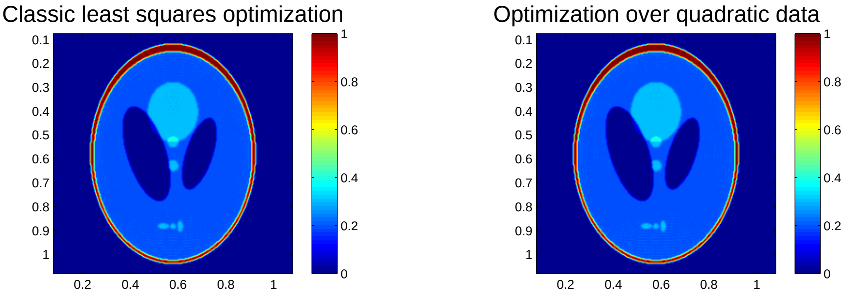
$$\min_M \operatorname{tr}(M) \quad \text{s.t.} \quad M \succeq 0, \quad \sum_{(i,j)\in E} |d_i \overline{d_j} - (FMF^*)_{ij}|^2 \leq \sigma^2.$$

Choice of E: random, sparse pairs of moderately close r, s, ω.

(E should not be the identity $i = j$. It should not be the complete graph either.)

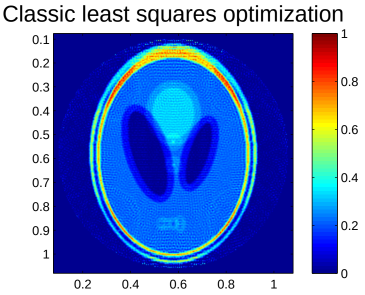
Choice of σ: *above some threshold* in order to get robustness in F.

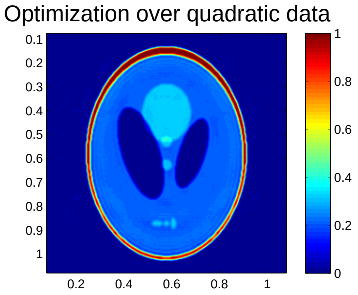
Linear vs. interferometric inversion



Exact $m_0 = 1$ (exact F)

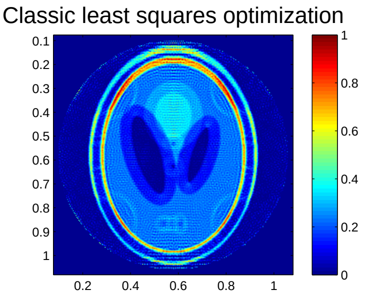
Linear vs. interferometric inversion

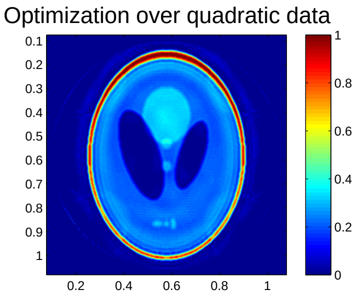




Inexact $m_0 = 1.06$ (approximate F)

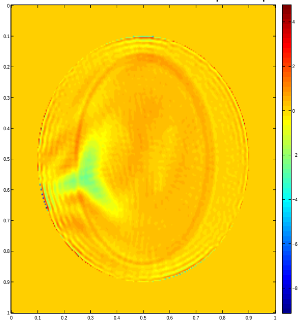
Linear vs. interferometric inversion





Inexact $m_0 = 1.11$ (approximate F)

Local minima in unrelaxed quartic formulation



Exact $m_0 = 1$ (exact F)

Robustness to m_0: explanation

With $m_0 = 1/c^2$, let

- $F, \tilde{F}$: full-aperture Born modeling operators in velocity $c, \tilde{c}$, frequency ω, receiver radius R.

Theorem (D., Jugnon, 2014)

Let

$$\tilde{m}(x) = \frac{c}{\tilde{c}}\, m\left(\frac{c}{\tilde{c}}\, x\right), \qquad \tilde{M} = \tilde{m} \otimes \tilde{m}$$

Then

$$\|Fm - \tilde{F}\tilde{m}\|_2 \gtrsim \|Fm\|_2$$

and

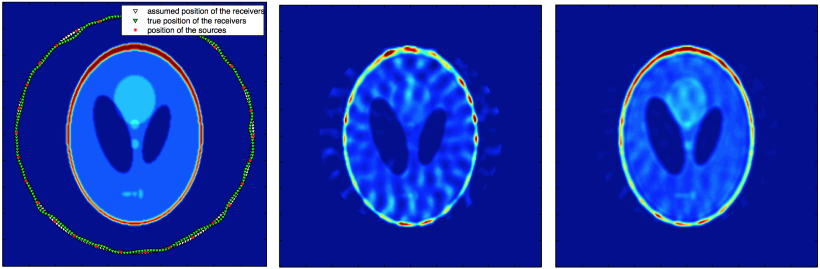
$$\|FMF^* - \tilde{F}\tilde{M}\tilde{F}^*\|_1 \lesssim \frac{\omega\, diam^2(m)}{c\, R}\, \|FMF^*\|_1$$

$\Rightarrow$ $\tilde{m}$ compatible with interferometric measurements.

Robustness to sensor location

Full aperture, MIMO, uniform medium $m_0 = 1$,

but wrong sensor locations.



E: spacing between correlated r or s $<$ decoherence angle.

Agenda

3. *Not-so-special quadratic systems*

Toward the convex relaxation of inverse wave scattering

Inverse scattering

Inverse scattering: determine m from

$$\begin{aligned} m\,\partial_{tt}u - \Delta u &= f, \\ Su &= d. \end{aligned}$$

Hard because $m \to d = \mathcal{F}(m)$ is very nonlinear (high frequencies).

No convincing algorithmic progress in 30+ years

Basis for two major seismic imaging conferences each year.

Relaxation of inverse scattering

Inverse scattering: determine m from

$$\begin{aligned} m\,\partial_{tt}u - \Delta u &= f, \\ Su &= d \end{aligned}$$

Quadratic system in the entries of $v = (1, m, u)^T$.

Lift to X, a proxy for $v \otimes v^*$, and get the equivalent rank-1 matrix recovery problem

$$\begin{aligned} \partial_t^2 \mathrm{diag} X_{23} - \Delta X_{13} &= f, \\ SX_{13} &= d, \\ X_{11} &= 1, \\ \mathrm{rank}(X) &= 1. \end{aligned}$$

Relaxation of inverse scattering

- First-order relaxation L1: drop $\mathrm{rank}(X) = 1$, add $X \succeq 0$. (add min Tr.)

  Exact for some problems (e.g. phase retrieval) but not for inverse scattering. Manifestation: ill-posedness.

- Second-order relaxation L2: drop $\mathrm{rank}(X) = 1$, but encode determinant conditions $X_{ij}X_{k\ell} - X_{ik}X_{j\ell} = 0$ via a lifting proxy $\mathcal{X}$ for $X \otimes X$. Add $X \succeq 0$ and $\mathcal{X} \succeq 0$.

- Increasingly tighter Lk: hierarchy of moment / sum of squares relaxation of Lasserre (2001) for polynomial optimization.

No large-scale implementation to date for Lk, $k \geq 2$.

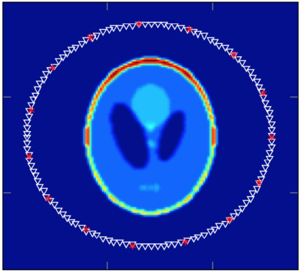
Relaxation of inverse scattering

- First-order relaxation L1-Rk: relax $\text{rank}(X) = 1$ into $\text{rank}(X) \leq k$ (nonconvex!), with $X \succeq 0$, and use your favorite descent method.

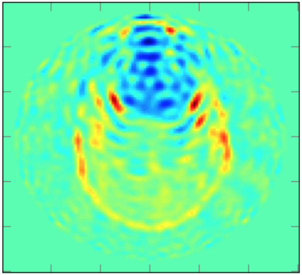
  Still ill-posed for inverse scattering.

- Adjoint-state constraints: pick an underdetermined subset of constraints $A(X) = b$. Solve it in the ULS sense: new constraint: $X = A^* \lambda$. Project out the resulting variable λ by the adjoint-state method, for speed.

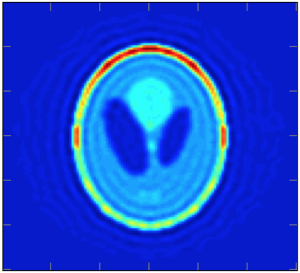
  Well-posed, enlarged attraction basin.



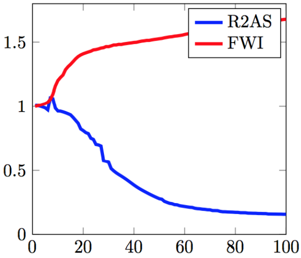
Setup



Classical FWI



R2+AS



Model error

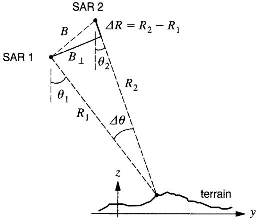
Relaxation methods in imaging

Conclusions:

- A systematic way of reducing quadratic/algebraic systems to rank-1 matrix recovery problems.
- Implications for model-robust imaging from interferometric data: errors in F, not in d.
- Toward convex SDP relaxation of inverse scattering: Lasserre hierarchy / sum of squares.

What is InSAR?

- Interferometric SAR: two recorded waveforms $d_1(\omega)$, $d_2(\omega)$



Form

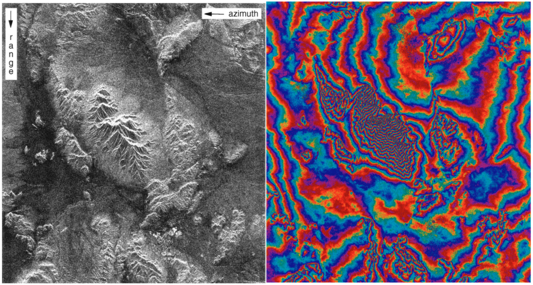
$$\color{red}{d_1(\omega)\overline{d_2(\omega)}} \sim e^{2i\omega R_1/c} e^{-2i\omega R_2/c}$$
$$\sim e^{i\phi}$$

Then relate $\phi = 2k(R_1 - R_2)/c$ to h.

Credit:

Bamler, Hartl

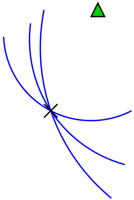
ESA

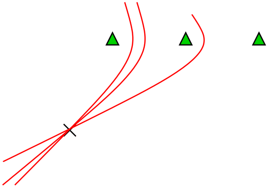


Physical content of interferometry

Example: correlograms

	Waveform	Traveltime
Absolute	$d(r,s,\omega)$	$\tau(r,x)$
Relative	$d(r_1,s,\omega)\overline{d(r_2,s,\omega)}$	$\tau(r_1,x)-\tau(r_2,x)$

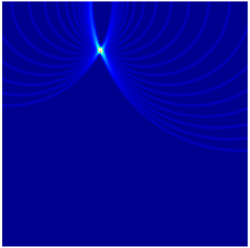




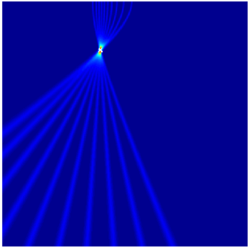
Example: focal spot / diffraction point

Correlations are also robust to background velocity $m_0 = 1/c_0^2$.

Image of a point scatterer: exact $c_0 = 1$.



Backprojection
$\widetilde{m}_{\text{mig}}(x)$

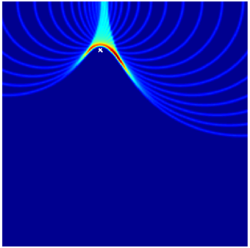


Interferometric imaging
$\widetilde{m}_{\text{int}}(x)$

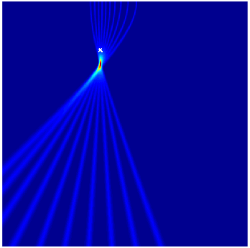
Example: focal spot / diffraction point

Correlations are also robust to background velocity $m_0 = 1/c_0^2$.

Image of a point scatterer: inexact $c_0 = 0.9$.



Backprojection
$\widetilde{m}_{\text{mig}}(x)$



Interferometric imaging
$\widetilde{m}_{\text{int}}(x)$

Robust imaging with correlations

Usual imaging: backprojection/migration $(F^* d)(x) \simeq$

$$\widetilde{m}_{\text{mig}}(x) = \sum_{r,s} \int e^{i\omega(\tau(s,x)+\tau(x,r))} \, d(r,s,\omega) \, d\omega$$

Interferometric imaging (Borcea et al., 2002, '03, '05):

$$\tau(s,x) + \tau(x,r_1) - \tau(s,x) - \tau(x,r_2)$$

and "backproject the correlations"

$$\widetilde{m}_{\text{int}}(x) = \sum_{r_1,r_2,s} \int e^{i\omega(\tau(x,r_1)-\tau(x,r_2))} d(r_1,s,\omega)\overline{d(r_2,s,\omega)} \, d\omega$$

(+ windowing tricks)

Use a nonsmooth iterative method

Practical numerical aspects:

- Douglas-Rachford splitting (or other nonsmooth primal-dual method)

$$\min f(M)+g(M), \qquad J_f = (I+\gamma\partial f)^{-1}, \qquad J_g = (I+\gamma\partial g)^{-1}$$

$$R_f = 2J_f - I, \qquad R_g = 2J_g - I$$

$$y^{k+1} = \frac{1}{2}\left[R_f R_g + I\right] y^k$$
$$M^k = J_g(y^k)$$

- Keep M separated throughout. Low-rank heuristics for $P_{M\succeq 0}$.
- Convergence: slowdown vs. LS of a factor 10-100.