Interferometric inversion: a robust approach to linear inverse problems

Vincent Jugnon[(*)] *and Laurent Demanet*
Department of Mathematics and Earth Resources Laboratory, Massachusetts Institute of Technology.

SUMMARY

This note describes an interferometric reformulation of linear wave-based inverse problems, such as inverse source and inverse Born scattering. Instead of fitting data directly, we propose to match cross-correlations and other quadratic data combinations that generalize cross-correlations. This modification exhibits surprising robustness to modeling uncertainties of a kinematic nature. Three optimization formulations are proposed: (1) a simple nonconvex problem, (2) a convex relaxation via lifting and a semidefinite constraint, and (3) a fast, relaxed scheme that empirically bypasses the lack of convexity. We describe sufficient conditions for recovery from the lifted scheme (2). We illustrate the robustness of interferometric inversion on two numerical examples, with uncertainty either in the model velocity or the sensor locations.

INTRODUCTION

Linear inverse problems appear in different areas of exploration geophysics. For instance, inverse source problems arise when locating microseismic events, and linear inverse scattering in the Born regime yield model updates for subsurface imaging. These linear problems all take the form

$$Fm = d, \qquad (1)$$

where F is the forward or modeling operator, describing the wave propagation and the acquisition, m is the reflectivity model, and d are the observed data. The classic approach is to use the data (seismograms) directly, to produce an image either

- by migrating the data (Reverse Time Migration),

$$I_{\text{RTM}} = \widetilde{F}^* d$$

  where $\widetilde{F}$ is the **simulation** forward operator and * stands for the adjoint ;

- or by finding a model that best fits the data, in a least-squares sense (Least-Squares Migration),

$$m_{\text{LSM}} = \arg\min_{m} \frac{1}{2}||\widetilde{F}m - d||_2^2 \qquad (2)$$

It is important to note that the physical forward operator F and the one used in simulation $\widetilde{F}$ can be different due to modeling errors or uncertainty. Such errors can happen at different levels:

1. background velocity,
2. sources and receivers positions,
3. sources time profiles.

This list is non-exhaustive and these modeling errors have a very different effect from classic additive white noise, in the sense that they induce a coherent perturbation in the data. As a result, the classic approaches (RTM, LSM) may fail in the presence of such uncertainties.

Robustifying migration

The idea of using interferometry (i.e. products of pairs of data) to make migration robust to modeling uncertainties has already been proposed in the literature (Borcea et al. (2005), Sava and Poliannikov (2008), Schuster et al. (2004)), producing remarkable results.

In their 2005 paper, Borcea et al. developed a comprehensive framework for interferometric migration, in which they proposed the Coherent INTerfermetic imaging functional (CINT), which can be recast in our notation,

$$I_{\text{CINT}} = \text{diag}\{\widetilde{F}^*(E \circ [dd^*])\widetilde{F}\},$$

where dd^* is the matrix of all data pairs products, E is a selector, that is, a sparse matrix with ones for the considered pairs and zeros elsewhere, $\circ$ is the entrywise (Hadamard) product, and diag is the operation of extracting the diagonal of a matrix.

The CINT functional involves $\widetilde{F}^*$, $\widetilde{F}$ and F, F^* (implicitly through dd^*), so cancellation of model errors can still occur, even when F and $\widetilde{F}$ are different. Through a careful analysis of wave propagation in the particular case where the uncertainty consists of random fluctuations of the background velocity, Borcea et al. derive conditions on E under which CINT will be robust.

The purpose of this note is to extend the power of interferometry to inversion.

Inversion from quadratic combinations

We propose to perform inversion using selected data pair products,

$$\text{find } m \text{ s.t. } E \circ (\widetilde{F}mm^*\widetilde{F}^*) = E \circ [dd^*], \qquad (3)$$

i.e., we look for a model m that explains the data pair products $d_i \bar{d}_j$ selected by $(i, j) \in E$. Here, i and j are meta-indices in data space. For example, in an inverse source problem in the frequency domain, $i \equiv (r_i, \omega_i)$ and $j \equiv (r_j, \omega_j)$, and for inverse Born scattering, $i \equiv (r_i, s_i, \omega_i)$ and $j \equiv (r_j, s_j, \omega_j)$.

A particular instance of the interferometric inversion problem is inversion from cross-correlograms. In that case,

$$E_{i,j} = 1 \Leftrightarrow \omega_i = \omega_j.$$

In that case, E considers data pairs from different sources and receivers at the same frequency,

$$d_i \overline{d_j} = d(r_i, s_i, \omega)\overline{d(r_j, s_j, \omega)},$$

where the overline stands for the complex conjugation. This expression is the Fourier transform at frequency ω of the cross-correlogram between trace $(r_i.s_i)$ and trace (r_j, s_j).

Interferometric Inversion

A straightforward approach is to fit the products in a least-squares sense,

$$\widehat{m}_{\text{LS,pairs}} = \arg\min_m ||E \circ (\widetilde{F}mm^*\widetilde{F}^* - dd^*)||_F^2. \qquad (4)$$

While the problem in (3) is quadratic, the least-squares cost in (4) is quartic and non-convex. The introduction of local minima is a highly undesired feature.

The choice of selector E is an important concern. Clearly, there are conditions on E for inversion to be possible. In the extreme case where E is the identity matrix, the problem becomes to estimate the model from intensity-only (phaseless) measurements, which does not in general have a unique solution. On the other hand, it is undesirable to consider too many pairs, both from a computational point of view and for robustness. There is a trade-off between robustness to uncertainties and quantity of information available to ensure invertibility.

In the next section, we provide a way to convexify the interferometric inversion problem. We then state sufficient a posteriori conditions on E for invertibility of this convexified interferometric formulation.

CONVEXIFICATION VIA LIFTING

Lifting was proposed in Chai et al. (2011) to convexify problems such as (3). This idea is natural in the context of recent work on matrix completion of Recht et al. (2010), and Candès and Recht (2009). The concept of lifting was also successfully used for phase retrieval (Candes et al., 2013). The idea is to replace the optimization variable m by the symmetric matrix $M = mm^*$, for which the problem becomes linear (and highly underdetermined). Incorporating the knowledge we have on the solution, the problem becomes

$$\begin{aligned} \text{find} \quad & M \quad \text{s. t.} \\ E \circ [\widetilde{F}M\widetilde{F}^*] &= E \circ [dd^*], \\ M &\succeq 0, \\ \text{rank}(M) &= 1. \end{aligned}$$

The first two constraints (data fit and positive semi-definiteness) are convex, but the rank constraint is not and would lead to a NP hard problem. As we argue below, the rank constraint is surprisingly unnecessary and can often be dropped. We also relax the data pairs fit – an exact fit is ill-advised because of noise and modeling errors – to obtain the following feasibility problem

$$\begin{aligned} \text{find} \quad & M \quad \text{s. t.} , \\ ||\widetilde{F}M\widetilde{F}^* - dd^*||_{\ell_1(E)} &\leq \sigma, \qquad (5) \\ M &\succeq 0. \end{aligned}$$

The approximate fit is expressed in an entry-wise ℓ_1 sense. This feasibility problem is a convex program, for which there exist simple converging iterative methods. Once M is solved for, a model estimate can be obtained by extracting the leading eigenvector of M as

$$\widehat{m} = \sqrt{\eta_1}v_1,$$

where η_1 is the largest eigenvalue of M, and v_1 is the corresponding eigenvector.

In Demanet and Jugnon (2013), we proved a recovery theorem for this estimate. The following definition is needed to explain

this result. The set of selected data pairs E can be viewed as the edges of a graph whose nodes are the data points. If the edges are weighted by the data moduli, the data-weighted graph Laplacian takes the following form

$$\left(L_{|d|}\right)_{ij} = \begin{cases} \sum_{k:(i,k)\in E} |d_k|^2 & \text{if } i = j; \\ -|d_i||d_j| & \text{if } (i,j) \in E; \\ 0 & \text{otherwise.} \end{cases}$$

Its first eigenvalue is zero, and its second eigenvalue λ_2 (a.k.a. spectral gap, or algebraic connectivity of the graph), gives a quantitative indication of the connectedness of the graph.

Theorem 1. *Assume $||\varepsilon||_1 + \sigma \leq \lambda_2/2$, where ε is the noise over the products of data pairs and λ_2 is the second eigenvalue of the data-weighted graph Laplacian $L_{|d|}$ formed from E. Any estimate $\widehat{m}$ obtained from (5) obeys*

$$\frac{||\widehat{m} \pm m||}{||m||} \leq 15\,\kappa(F)^2\sqrt{\frac{||\varepsilon||_1 + \sigma}{\lambda_2}},$$

where $\kappa(F)$ is the condition number of F.

Up to a sign, the relative error estimate obtained from (5) will be bounded by the norm of the error on the data pairs product ε, the threshold imposed on the approximate fit σ, and $\frac{1}{\sqrt{\lambda_2}}$.

The theorem gives sufficient conditions for invertibility:

- for λ_2 to be non-zero (and the bound to be finite), the graph has to be connected,
- the better connected the graph is, the larger λ_2, and the more stable the recovery.

It is important to stress that this estimate gives sufficient conditions on E for recovery to be possible and stable to additive noise ε, but not to modeling error (in the theorem, $F = \widetilde{F}$). It does not provide an explanation of the robust behavior of our correlations-based approach.

NUMERICAL ILLUSTRATION

A practical algorithm

The convexified formulation in (5) is too costly to solve at the scale of even toy problems. Let N be the total number of degrees of freedom of your unknown model m ; then the variable M of (5) is a $N \times N$ matrix, on which we want to impose positive semi-definiteness and approximate fit. As of 2013 and to our knowledge, there is no time-efficient and memory-efficient algorithm to solve this type of semi-definite program when N ranges from 10^4 to 10^6.

We consider instead a non-convex relaxation of the feasibility problem (5), in which we limit the rank of M to K, as in (Burer and Monteiro (2003)). We may then write $M = RR^*$ where R is $N \times K$ and $K \ll N$. We replace the approximate ℓ_1 fit by Frobenius minimization. Regularization is also added to handle noise and uncertainty, yielding

$$\widehat{R} = \arg\min_R ||E \circ (\widetilde{F}RR^*\widetilde{F} * -dd^*)||_F^2 + \lambda ||R||_F^2. \qquad (6)$$

Interferometric Inversion

An estimate of m is obtained from $\widehat{R}$ extracting the leading eigenvector of $\widehat{R}\widehat{R}^*$. Note that the Frobenius regularization on R is equivalent to a trace regularization on $M = RR^*$, which is known to promote low-rank character of the matrix (Candes et al., 2013).

The rank-K relaxation (6) can be seen as a generalization of the straightforward least-squares formulation (4). The two formulations coincide in the limit case $K = 1$. The strength of (6) is that the optimization variable is in a slightly bigger space than formulation (4).

The rank-K relaxation (6) is still non-convex, but in practice, no local minimum has been observed even for $K = 2$, whereas the issue often arises for the least-squares approach (4).

In the following two subsections we compare reconstructions from classic Least-Squares Migration (2) and from interferometric inversion using (6).

Example 1: an inverse source problem with background velocity uncertainty

Here we consider a constant density acoustics inverse source problem, which reads in the Fourier domain

$$-(\Delta + \omega^2 m_0(x))\widehat{u}_s(x,\omega) = \widehat{w}(\omega)m(x)$$

$$Fm = d(x_r,\omega) = \widehat{u}_s(x_r,\omega)$$

$$m_0(x) = \frac{1}{c_0(x)^2}.$$

Waves are propagating from a source term with known time signature w. The problem is to reconstruct the spatial distribution m.

The source distribution is the Shepp-Logan phantom. The waves are measured on receivers surrounding the domain (Fig.1 (top)). Equispaced frequencies are considered on the bandwidth of w.

A significant modeling error is assumed to have been made on the background velocity. In the experiment, the waves propagated with unit speed $c_0(x) \equiv 1$, but in the simulation, the waves propagate more slowly, $\widetilde{c}_0(x) \equiv 0.95$.

As shown in Fig.1 (middle), Least-Squares Migration of the data cannot handle this type of uncertainty and produces a strongly defocused image.

In contrast, interferometric inversion, shown in Fig.1 (bottom), enjoys a better resolution. In this case, the price to pay for focusing is positioning: the interferometric reconstruction is a shrunk version of the true source distribution.

For choosing E, we have followed an idea from Borcea et al. (2005): if the modeling error has a smooth effect on the data, then we should consider data pairs that are close to one another both in receiver position and in frequency.

Example 2: a inverse scattering problem with receivers position uncertainty

We now turn to an active setting. Constant density acoustic wave are generated by sources surrounding the domain. The incident wavefields are scattered by the reflectivity m (which is again the Shepp-Logan phantom). Receivers all around the domain measure the Born (linear) scattered wavefield

$$-(\Delta + \omega^2 m_0(x))\widehat{u}_{0,s}(x,\omega) = \widehat{w}(\omega)\delta(x - x_s)$$
$$-(\Delta + \omega^2 m_0(x))\widehat{u}_{1,s}(x,\omega) = \omega^2\widehat{u}_{0,s}(x,\omega)m(x)$$

$$Fm = d(x_r,x_s,\omega) = \widehat{u}_{1,s}(x_r,\omega).$$

The isolation of the Born scattered wavefield (primary reflections) from the full scattered field, although a very difficult task in practice, is assumed to be performed perfectly in this paper. In this example, the modeling error is assumed to be on the receiver positions. In the experiment, they have a smooth random deviation from the circle as shown in Fig.1 (top).

Once again, Least-Squares Migration produces a poor result (Fig.2 (middle)), where the features of the phantom do not appear, and the strong outer layer is reconstructed with a significant oscillatory error.

Interferometric inversion produces a better reconstruction (Fig.2 (bottom)), where some features of the phantom are recovered and the outer layer is well-resolved. The same selector as in the previous case has been used.

CONCLUSION

We have developed an new approach to classic linear inverse problems based on wave-propagation. By shifting the focus from the data itself to products between data pairs (a generalization of correlations), we have derived a formulation robust to modeling errors.

The interferometric inversion becomes convex upon semidefinite relaxation. Developing an algorithm to solve this lifted problem at interesting scales is a difficult problem that we circumvent via an ad-hoc rank-2 relaxation scheme.

The choice of the selector E, crucial for robustness, needs to be investigated more thoroughly. It is possible to derive heuristics based on the type of considered uncertainty, but a more detailed study of what an optimal selector would be in terms of robustness and algebraic connectivity is desired.

Even though briefly mentioned, approximate fit of the data pairs (or Frobenius regularization of misfit minimization) is central to the method we present here. Parameters of this regularization and their effect on the behavior of the algorithm need to be better understood.

Finally, we are interested in applying interferometric inversion to non-linear problems like full waveform inversion. In this case, the lift idea does not apply anymore (the problem is non-linear to begin with), but the naive least-squares fit of data pairs products (4) can still be implemented. The introduction of non-convexity is not as important an issue because the FWI cost is strongly non-convex in the first place. Preliminary results show that inversion from data pairs is possible in this case. Whether interferometric inversion will retain its robustness in the non-linear setting is a matter of current investigation.

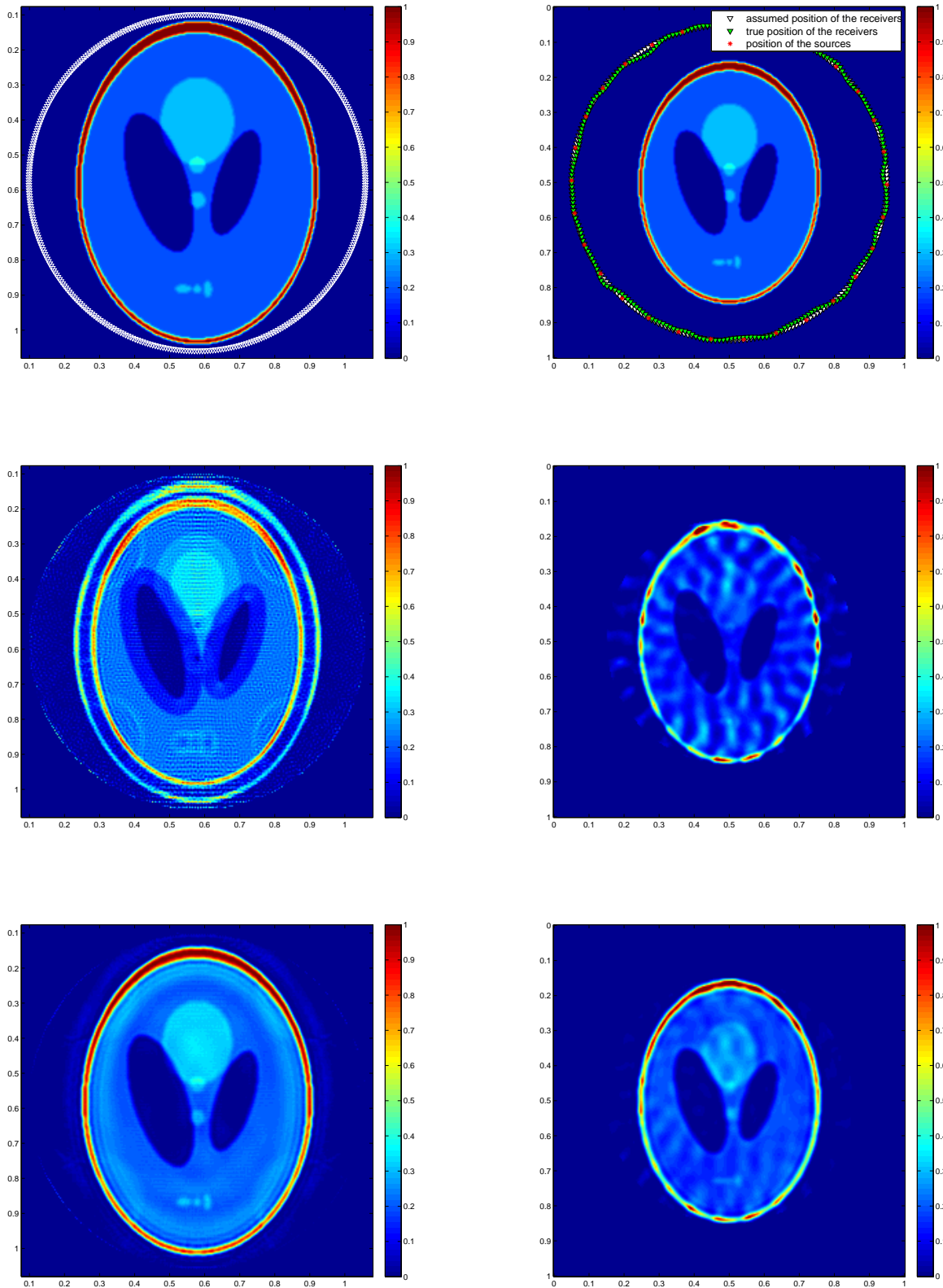## ACKNOWLEDGEMENTS

Interferometric Inversion

Figure 1: (top) Setting of the inverse source experiment (middle) Least-Squares Migration reconstruction, and (bottom) interferometric reconstruction.

Figure 2: (top) Setting of the inverse scattering experiment (middle) Least-Squares Migration reconstruction, and (bottom) interferometric reconstruction.

REFERENCES

Borcea, L., G. Papanicolaou, and C. Tsogka, 2005, Interferometric array imaging in clutter: Inverse Problems, **21**, 1419.

Burer, S., and R. Monteiro, 2003, A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization: Mathematical Programming (Series B), **95**, 329–357.

Candes, E. J., Y. C. Eldar, T. Strohmer, and V. Voroninski, 2013, Phase retrieval via matrix completion: SIAM Journal on Imaging Sciences, **6**, 199–225.

Candès, E. J., and B. Recht, 2009, Exact matrix completion via convex optimization: Foundations of Computational mathematics, **9**, 717–772.

Chai, A., M. Moscoso, and G. Papanicolaou, 2011, Array imaging using intensity-only measurements: Inverse Problems, **27**, 015005.

Demanet, L., and V. Jugnon, 2013, Convex recovery from interferometric measurements: to appear.

Recht, B., M. Fazel, and P. A. Parrilo, 2010, Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization: SIAM review, **52**, 471–501.

Sava, P., and O. Poliannikov, 2008, Interferometric imaging condition for wave-equation migration: Geophysics, **73**, S47–S61.

Schuster, G., J. Yu, J. Sheng, and J. Rickett, 2004, Interferometric/daylight seismic imaging: Geophysical Journal International, **157**, 838–852.