Convex recovery from interferometric measurements

Laurent Demanet* and Vincent Jugnon*

Abstract

This paper discusses some questions that arise when a linear inverse problem involving $Ax = b$ is reformulated in the interferometric framework, where quadratic combinations of b are considered as data in place of b.

First, we show a deterministic recovery result for vectors x from measurements of the form $(Ax)_i\overline{(Ax)_j}$ for some left-invertible A. Recovery is exact, or stable in the noisy case, when the couples (i, j) are chosen as edges of a well-connected graph. One possible way of obtaining the solution is as a feasible point of a simple semidefinite program. Furthermore, we show how the proportionality constant in the error estimate depends on the spectral gap of a data-weighted graph Laplacian.

Second, we present a new application of this formulation to interferometric waveform inversion, where products of the form $(Ax)_i\overline{(Ax)_j}$ in frequency encode generalized cross-correlations in time. In the Born regime, we present numerical evidence that interferometric *inversion* does not suffer from the loss of resolution generally associated with interferometric imaging, and can provide added robustness with respect to a specific kind of kinematic uncertainty in the forward model A. In the non-Born regime, we show that interferometric inversion can successfully be used as a substitute for the classical adjoint-state model updates.

Acknowledgments. The authors would like to thank Amit Singer, George Papanicolaou, and Liliana Borcea for interesting discussions. Some of the results in this paper were reported in the conference proceedings of the 2013 SEG annual meeting [14].

1 Introduction

Throughout this paper, we consider complex quadratic measurements of $x \in \mathbb{C}^n$ of the form

$$B_{ij} = (Ax)_i\overline{(Ax)_j}, \qquad (i, j) \in E, \tag{1}$$

for certain well-chosen couples of indices (i, j), a scenario that we qualify as "interferometric". This combination is special in that it is symmetric in x, and of rank 1 with respect to the indices i and j.

The regime that interests us is when the number m of measurements, i.e., of couples (i, j) in E, is comparable to the number n of unknowns. While phaseless measurements $b_i = |(Ax)_i|^2$ only admit recovery when A has very special structure – such as, being a tall random matrix with Gaussian i.i.d. entries [5, 9] – products $(Ax)_i\overline{(Ax)_j}$ for $i \neq j$ correspond to the idea of phase differences, hence encode much more information. As a consequence, stable recovery occurs under very general conditions: left-invertibility of A and "connectedness" of set E of couples (i, j). These conditions suffices to allow for m to be on the order of n. Various algorithms return x accurately

*Department of Mathematics, MIT. LD is supported by AFOSR, ONR, NSF, Total SA, and the Alfred P. Sloan Foundation. VJ is supported by the Earth Resources Laboratory at MIT.

up to a global phase; we mosty discuss variants of lifting with semidefinite relaxation in this paper. In contrast to other recovery results in matrix completion [19, 6], no randomness is needed in the data model, and our proof technique involves elementary spectral graph theory rather than dual certification or uncertainty principles.

The mathematical content of this paper is the formulation of a quadratic analogue of the well-known relative error bound

$$\frac{\|x - x_0\|}{\|x_0\|} \leq \kappa(A) \, \frac{\|e\|}{\|b\|} \tag{2}$$

for the least-squares solution of the overdetermined linear system $Ax = b$ with $b = Ax_0 + e$, and where $\kappa(A)$ is the condition number of A. Our result is that an inequality of the form (2) still holds in the quadratic case, but with the *square root of the spectral gap of a data-weighted graph Laplacian* in place of $\|b\|$ in the right-hand side. This spectral gap quantifies the connectedness of E, and has the proper homogeneity with respect to b.

The numerical results mostly concern the case when A is a forward model that involves solving a linearized fixed-frequency wave equation, as in seismic or radar imaging. In that case x is a reflectivity, Ax is the wavefield that results from an incoming wave being scattered by x, and $(Ax)_i\overline{(Ax)_j}$ has a special meaning that can be understood in the context of interferometry.

1.1 Interferometry

In optical imaging, an interference fringe of two (possibly complex-valued) wavefields $f(t)$ and $g(t)$, where t is either a time or a space variable, is any combination of the form $|f(t) + g(t + t')|^2$. The sum is a result of the linearity of amplitudes in the fundamental equations of physics (such as Maxwell or Schrödinger), while the modulus squared is simply the result of a detector measuring intensities. The cross term $2\Re(f(t)\overline{g(t + t')})$ in the expansion of the squared modulus manifestly carries the information of destructive vs. constructive interference, hence is a continuous version of what we referred to earlier as an "interferometric measurement".

In particular, when the two signals are sinusoidal at the same frequency, the interferometric combination highlights a phase difference. In astronomical interferometry, the delay t' is for instance chosen so that the two signals interfere constructively, yielding better resolution. Interferometric synthetic aperture radar (InSAR) is a remote sensing technique that uses the fringe from two datasets taken at different times to infer small displacements. In X-ray ptychograhy [20], imaging is done by undoing the interferometric combinations that the diffracted X-rays undergo from encoding masks. These are but three examples in a long list of applications.

Interferometry is also playing an increasingly important role in geophysical imaging, i.e., inversion of the elastic parameters of the portions of the Earth's upper crust from scattered seismic waves. In this context however, the signals are often impulsive rather than monochromatic. As a result, it is more common to perform quadratic combinations of the Fourier transform of seismograms at different receivers, such as $|\widehat{f}(\omega) + \widehat{g}(\omega)|^2$. The cross-term involves $\widehat{f}(\omega)\overline{\widehat{g}(\omega)}$, the Fourier transform of the cross-correlation of f and g. It highlights a time lag in the case when f and g are impulses.

Cross-correlations have been shown to play an important role in imaging, mostly because of their stability to statistical fluctuations of a scattering medium [3] or an incoherent source [16, 11]. Though seismic interferometry is a vast research area [4, 23, 28, 22], explicit inversion of reflectivity parameters from interferometric data has to our knowledge only been considered in [10, 14]. Interferometric inversion offers great promise for model-robust imaging, i.e., recovery of reflectivity maps in a less-sensitive way on specific kinds of errors in the forward model.

Finally, interferometric measurements also play an important role in quantum optical imaging. See [21] for a nice solution to the inverse problem of recovering a scattering dielectric susceptibility from measurements of two-point correlation functions relative to two-photon entangled states.

1.2 Broader context and related work

The setting of this paper is discrete, hence we let i and j in place of either a time of frequency variable. We also specialize to $f = g$, and we let $f = Ax$ to possibly allow an explanation of the signal f by a linear forward model[1] A.

The link between products of the form $f\,\overline{g}$ and squared measurements $|f+g|^2$ goes both ways, as shown by the polarization identity

$$f_i\,\overline{f_j} = \frac{1}{4}\sum_{k=1}^{4} e^{-i\pi k/2}|f_i + e^{i\pi k/2} f_j|^2.$$

Hence any result of robust recovery of f, or x, from couples $f_i\,\overline{f_j}$, implies the same result for recovery from phaseless measurements of the form $|f_i + e^{i\pi k/2} f_j|^2$. This latter setting was precisely considerered by Candès et al. in [5], where an application to X-ray diffraction imaging with a specific choice of masks is discussed. In [1], Alexeev et al. use the same polarization identity to design good measurements for phase retrieval, such that recovery is possible with $m = O(n)$.

Recovery of f_i from $f_i\overline{f_j}$ for some (i,j) when $|f_i| = 1$ (interferometric phase retrieval) can be seen a special case of the problem of angular synchronization considered by Singer [24]. There, rotation matrices R_i are to be recovered (up to a global rotation) from measurements of relative rotations $R_i R_j^{-1}$ for some (i,j). This problem has an important application to cryo-electron microscopy, where the measurements of relative rotations are further corrupted in an a priori unknown fashion (i.e., the set E is to be recovered as well). An impressive recovery result under a Bernoulli model of gross corruption, with a characterization of the critical probability, were recently obtained by Wang and Singer [26]. The spectrum of an adapted graph Laplacian plays an important role in their analysis [2], much as it does in this paper. Singer and Cucuringu also considered the angular synchronization problem from the viewpoint of rigidity theory [25]. For the similar problem of recovery of positions from relative distance, with applications to sensor network localization, see for instance [13].

The algorithmic approach considered in this paper for solving interferometric inversion problems is to formulate them via lifting and semidefinite relaxation. This idea was considered by many groups in recent years [7, 5, 13, 24, 27], and finds its origin in theoretical computer science [12]. As we revise this paper, we also note the recent success of interferometric inversion for passive SAR imaging, under the name low-rank matrix recovery [18].

1.3 Recovery of unknown phases

Let us start by describing the simpler problem of interferometric phase recovery, when $A = I$ and we furthermore assume $|x_i| = 1$. Given a vector $x_0 \in \mathbb{C}^n$ such that $|(x_0)_i| = 1$, a set E of pairs (i,j), and noisy interferometric data $B_{ij} = (x_0)_i\overline{(x_0)_j} + \varepsilon_{ij}$, find a vector x such that

$$|x_i| = 1, \qquad \sum_{(i,j)\in E} |x_i\overline{x_j} - B_{ij}| \leq \sigma, \tag{3}$$

[1] Such as scattering from a reflectivity profile x in the Born approximation, for which A is a wave equation Green's function.

for some $\sigma > 0$. Here and below, if no heuristic is provided for σ, we may cast the problem as a minimization problem for the misfit and obtain σ a posteriori.

The choice of the elementwise ℓ_1 norm over E is arbitrary, but convenient for the analysis in the sequel[2]. We aim to find situations in which this problem has a solution x close to x_0, up to a global phase. Notice that x_0 is feasible for (3), hence a solution exists, as soon as $\sigma \geq \sum_{i,j\in E} |\varepsilon_{ij}|$.

The relaxation by *lifting* of this problem is to find X (a proxy for xx^*) such that

$$X_{ii} = 1, \qquad \sum_{(i,j)\in E} |X_{ij} - B_{ij}| \leq \sigma, \qquad X \succeq 0,$$
$$\text{then let } x \text{ be the top eigenvector of } X \text{ with } \|x\|^2 = n. \tag{4}$$

The notation $X \succeq 0$ means that X is symmetric and positive semi-definite. Again, the feasibility problem (4) has at least one solution ($X_0 = x_0 x_0^*$) as soon as $\sigma \geq \sum_{i,j\in E} |\varepsilon_{ij}|$.

The set E generates edges of a graph $G = (V, E)$, where the nodes in V are indexed by i. Without loss of generality, we consider E to be symmetric. By convention, G does not contain loops, i.e., the diagonal $j = i$ is not part of E. (Measurements on the diagonal are not informative for the phase recovery problem, since $|(x_0)_i|^2 = 1$.)

The graph Laplacian on G is

$$L_{ij} = \begin{cases} d_i & \text{if } i = j; \\ -1 & \text{if } (i,j) \in E; \\ 0 & \text{otherwise,} \end{cases}$$

where d_i is the node degree $d_i = \sum_{j:(i,j)\in E} 1$. Observe that L is symmetric and $L \succeq 0$ by Gershgorin. Denote by $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ the eigenvalues of L sorted in increasing order. Then $\lambda_1 = 0$ with the constant eigenvector $v_1 = \frac{1}{\sqrt{n}}$. The second eigenvalue is zero if and only if G has two or more disconnected components. When $\lambda_2 > 0$, its value is a measure of connectedness of the graph. Note that $\lambda_n \leq 2d$ by Gershgorin again, where $d = \max_i d_i$ is the maximum degree.

Since $\lambda_1 = 0$, the second eigenvalue λ_2 is called the spectral gap. It is a central quantity in the study of expander graphs: it relates to

- the edge expansion (Cheeger constant, large if λ_2 is large);
- the degree of separation between any two nodes (small if λ_2 is large); and
- the speed of mixing of a random walk on the graph (fast if λ_2 is large).

More information about spectral graph theory can be found, e.g., in the lecture notes by Lovasz [17]. It is easy to show with interlacing theorems that adding an edge to E, or removing a node from V, both increase λ_2. The spectral gap plays an important role in the following stability result.

In the sequel, we denote the componentwise ℓ_1 norm on the set E by $\|\cdot\|_1$.

Theorem 1. *Assume $\|\varepsilon\|_1 + \sigma \leq n\lambda_2$, where λ_2 is the second eigenvalue of the graph Laplacian L on G. Any solution x of (3) or (4) obeys*

$$\|x - e^{i\alpha} x_0\| \leq 4\sqrt{\frac{\|\varepsilon\|_1 + \sigma}{\lambda_2}},$$

for some $\alpha \in [0, 2\pi)$.

[2]The choice of ℓ_1 norm as "least unsquared deviation" is central in [26] for the type of outlier-robust recovery behavior documented there.

Manifestly, recovery is exact (up to the global phase ambiguity) as soon as $\varepsilon = 0$ and $\sigma = 0$, provided $\lambda_2 \neq 0$, i.e., the graph G is connected. The easiest way to construct expander graphs (graphs with large λ_2) is to set up a probabilistic model with a Bernoulli distribution for each edge in an i.i.d. fashion, a model known as the Erdős-Rényi random graph. It can be shown that such graphs have a spectral gap bounded away from zero independently of n with $m = O(n \log n)$ edges.

A stronger result is available when the noise ε originates at the level of x_0, i.e., $B = x_0 x_0^* + \varepsilon$ has the form $(x_0 + e)(x_0 + e)^*$.

Corollary 2. *Assume $\varepsilon = (x_0+e)(x_0+e)^* - x_0 x_0^*$ and $\sigma \leq n\lambda_2$, where λ_2 is the second eigenvalue of the graph Laplacian L on G. Any solution x of (3) or (4) obeys*

$$\|x - e^{i\alpha} x_0\| \leq 4 \sqrt{\frac{\sigma}{\lambda_2}} + \|e\|,$$

for some $\alpha \in [0, 2\pi)$.

Proof. Apply theorem 1 with $\varepsilon = 0$, $x_0 + e$ in place of x_0, then use the triangle inequality. □

In the setting of the corollary, problem (3) always has $x = x_0 + e$ as a solution, hence is feasible even when $\sigma = 0$.

Let us briefly review the eigenvector method for interferometric recovery. In [24], Singer proposed to use the first eigenvector of the (noisy) data-weighted graph Laplacian as an estimator of the vector of phases. A similar idea appears in the work of Montanari et al. as the first step of their OptSpace algorithm [15], and in the work of Chatterjee on universal thresholding [8]. In our setting, this means defining

$$(\widetilde{\mathcal{L}})_{ij} = \begin{cases} d_i & \text{if } i = j; \\ -B_{ij} & \text{if } (i,j) \in E; \\ 0 & \text{otherwise,} \end{cases}$$

and letting $x = \widetilde{v}_1 \sqrt{n}$ where v_1 is the unit-norm eigenvector of $\widetilde{\mathcal{L}}$ with smallest eigenvalue. Denote by $\widetilde{\lambda}_1 \leq \widetilde{\lambda}_2 \leq \ldots$ the eigenvalues of $\widetilde{\mathcal{L}}$. The following result is known from [2], but we provide an elementary proof for completeness of the exposition.

Theorem 3. *Assume $\|\varepsilon\| \leq \widetilde{\lambda}_2/2$. Then the result x of the eigenvector method obeys*

$$\|x - e^{i\alpha} x_0\| \leq \sqrt{2n}\, \frac{\|\varepsilon\|}{\widetilde{\lambda}_2},$$

for some $\alpha \in [0, 2\pi)$.

Alternatively, we may express the inequality in terms of λ_2, the spectral gap of the noise-free Laplacian L defined earlier, by noticing[3] that $\widetilde{\lambda}_2 \geq \lambda_2 - \|\varepsilon\|$. Both λ_2 and $\widetilde{\lambda}_2$ are computationally accessible. In the case when $|B_{ij}| = 1$, we have $\widetilde{\lambda}_1 \geq 0$, hence $\widetilde{\lambda}_2$ is (slightly) greater than the spectral gap $\widetilde{\lambda}_2 - \widetilde{\lambda}_1$ of $\widetilde{\mathcal{L}}$. Note that the $1/\widetilde{\lambda}_2$ scaling appears to be sharp in view of the numerical experiments reported in section 3. The inverse square root scaling of theorem 1 is stronger in the presence of small spectral gaps, but the noise scaling is weaker in theorem 1 than in theorem 3.

[3] This owes to $\|\mathcal{L} - \widetilde{\mathcal{L}}\| \leq \|\varepsilon\|$, with $\mathcal{L} = \Lambda L \Lambda^*$ the noise-free Laplacian with phases introduced at the beginning of section 2.1.

1.4 Interferometric recovery

The more general version of the interferometric recovery problem is to consider a left-invertible tall matrix A, linear measurements $b = Ax_0$ for some vector x_0 (without condition on the modulus of either b_i or $(x_0)_i$), noisy interferometric measurements $B_{ij} = b_i\overline{b_j} + \varepsilon_{ij}$ for (i,j) in some set E, and find x subject to

$$\sum_{(i,j)\in E\cup D} |(Ax)_i\overline{(Ax)_j} - B_{ij}| \leq \sigma. \tag{5}$$

Notice that we now take the union of the diagonal $D = \{(i,i)\}$ with E. Without loss of generality we assume that $\varepsilon_{ij} = \overline{\varepsilon_{ji}}$, which can be achieved by symmetrizing the measurements, i.e., substituting $\frac{B_{ij}+B_{ji}}{2}$ for B_{ij}.

Since we no longer have a unit-modulus condition, the relevant notion of graph Laplacian is now data-dependent. It reads

$$\big(L_{|b|}\big)_{ij} = \begin{cases} \sum_{k:(i,k)\in E} |b_k|^2 & \text{if } i = j; \\ -|b_i||b_j| & \text{if } (i,j)\in E; \\ 0 & \text{otherwise.} \end{cases}$$

The connectedness properties of the underlying graph now depend on the size of $|b_i|$: the edge (i,j) carries valuable information if and only if both $|b_i|$ and $|b_j|$ are large.

A few different recovery formulations arise naturally in the context of lifting and semidefinite relaxation.

- The basic lifted formulation is to find some X such that

$$\sum_{(i,j)\in E\cup D} |(AXA^*)_{ij} - B_{ij}| \leq \sigma, \qquad X \succeq 0,$$
$$\text{then let } x = x_1\sqrt{\eta_1}, \text{ where } (\eta_1, x_1) \text{ is the top eigen-pair of } X. \tag{6}$$

  Our main result is as follows.

  Theorem 4. *Assume $\|\varepsilon\|_1 + \sigma \leq \lambda_2/2$, where λ_2 is the second eigenvalue of the data-weighted graph Laplacian $L_{|b|}$. Any solution x of (6) obeys*

$$\frac{\|x - e^{i\alpha}x_0\|}{\|x_0\|} \leq 15\ \kappa(A)^2\ \sqrt{\frac{\|\varepsilon\|_1 + \sigma}{\lambda_2}},$$

  for some $\alpha \in [0, 2\pi)$, and where $\kappa(A)$ is the condition number of A.

  The quadratic dependence on $\kappa(A)$ is necessary[4]. In section 3, we numerically verify the inverse square root scaling in terms of λ_2.

  If the noise originates from $b + e$ rather than $bb^* + \varepsilon$, the error bound is again improved to

$$\frac{\|x - e^{i\alpha}x_0\|}{\|x_0\|} \leq 15\ \kappa(A)^2\ \sqrt{\frac{\sigma}{\lambda_2}} + \kappa(A)\ \frac{\|e\|}{\|b\|},$$

  for the same reason as earlier.

[4]The following example shows why that is the case. For any X_0 and invertible A, the solution to $AXA^* = AX_0A^* + \varepsilon$ is $X = X_0 + A^+\varepsilon(A^*)^+$. Let $X_0 = e_1e_1^T$, $\varepsilon = \delta e_1e_1^*$ for some small δ, and $A^+ = I + Ne_1e_1^T$. Then $X = (1+\delta N^2)e_1e_1^T$, and the square root of its leading eigenvalue is $\sqrt{\eta_1} \simeq 1 + \frac{1}{2}\delta N^2$. As a result, x is perturbation of x_0 by a quantity of magnitude $O(\delta\|A^+\|^2)$.

- An alternative, two-step lifting formulation is to find x through Y such that

$$\sum_{(i,j)\in E\cup D} |Y_{ij} - B_{ij}| \le \sigma, \qquad Y \succeq 0,$$
$$\text{then let } x = A^+ y_1 \sqrt{\eta_1}, \text{ where } (\eta_1, y_1) \text{ is the top eigen-pair of } Y. \tag{7}$$

  The dependence on the condition number of A is more favorable than for the basic lifting formulation.

  Theorem 5. *Assume $\|\varepsilon\|_1 + \sigma \le \lambda_2/2$, where λ_2 is the second eigenvalue of the data-weighted graph Laplacian $L_{|b|}$. Any solution x of (5) or (7) obeys*

$$\frac{\|x - e^{i\alpha} x_0\|}{\|x_0\|} \le 15\ \kappa(A) \sqrt{\frac{\|\varepsilon\|_1 + \sigma}{\lambda_2}},$$

  for some $\alpha \in [0, 2\pi)$.

  However, this formulation may be more computationally expensive than the one-step variant if data (b) space is much larger than model (x) space.

The quantity λ_2 is not computationally accessible in general, but it can be related to the second eigenvalue $\widetilde{\lambda}_2$ of the noisy data-weighted Laplacian,

$$\left(\widetilde{\mathcal{L}}_B\right)_{ij} = \begin{cases} \sum_{k:(i,k)\in E} B_{kk} & \text{if } i = j; \\ -B_{ij} & \text{if } (i,j) \in E; \\ 0 & \text{otherwise.} \end{cases}$$

It is straightforward to show that $\lambda_2 \ge \widetilde{\lambda}_2 - [\,(d+1)\|\varepsilon\|_\infty + \|\varepsilon\|\,]$, where $\|\cdot\|_\infty$ is the elementwise maximum on $E \cup D$, $\|\cdot\|$ is the spectral norm, and d is the maximum node degree.

2 Proofs

2.1 Proof of theorem 1.

Observe that if x is feasible for (3), then xx^* is feasible for (4), and has $e^{i\alpha}x$ as leading eigenvector. Hence we focus without loss of generality on (4).

As in [24], consider the Laplacian matrix weighted with the unknown phases,

$$\mathcal{L} = \Lambda L \Lambda^*,$$

with $\Lambda = \mathrm{diag}(x_0)$. In other words $\mathcal{L}_{ij} = (X_0)_{ij} L_{ij}$ with $X_0 = x_0 x_0^*$. We still have $\mathcal{L} \succeq 0$ and $\lambda_1 = 0$, but now $v_1 = \frac{1}{\sqrt{n}} x_0$. Here and below, λ and v refer to $\mathcal{L}$, and v has unit ℓ_2 norm.

The idea of the proof is to compare X with the rank-1 spectral projectors $v_j v_j^*$ of $\mathcal{L}$. Let $\langle A, B\rangle = \mathrm{tr}(AB^*)$ be the Frobenius inner product. Any X obeying (3) can be written as $X = X_0 + \widetilde{\varepsilon}$ with $\|\widetilde{\varepsilon}\|_1 \le \|\varepsilon\|_1 + \sigma$. We have

$$\langle X, \mathcal{L}\rangle = \langle X_0, \mathcal{L}\rangle + \langle \widetilde{\varepsilon}, \mathcal{L}\rangle$$

A short computation shows that

$$\begin{aligned}\langle X_0, \mathcal{L}\rangle &= \sum_i (X_0)_{ii}\overline{\mathcal{L}_{ii}} + \sum_{(i,j)\in E} (X_0)_{ij}\overline{\mathcal{L}_{ij}} \\ &= -\sum_i d_i + \sum_{(i,j)\in E} (x_0)_i\overline{(x_0)_j}\ \overline{(x_0)_i}(x_0)_j \\ &= \sum_i \left[-d_i + \sum_{j:(i,j)\in E} 1\right] \\ &= 0.\end{aligned}$$

Since $|L_{ij}| = 1$ on E, the error term is simply bounded as

$$|\langle \widetilde{\varepsilon}, \mathcal{L}\rangle| \le \|\widetilde{\varepsilon}\|_1$$

On the other hand the Laplacian expands as

$$\mathcal{L} = \sum_j v_j \lambda_j v_j^*,$$

so we can introduce a convenient normalization factor $1/n$ and write

$$\langle \frac{X}{n}, \mathcal{L}\rangle = \sum_j c_j \lambda_j, \tag{8}$$

with

$$c_j = \langle \frac{X}{n}, v_j v_j^*\rangle = \frac{v_j^* X v_j}{n}.$$

Notice that $c_j \ge 0$ since we require $X \succeq 0$. Their sum is

$$\sum_j c_j = \langle \frac{X}{n}, \sum_j v_j v_j^*\rangle = \langle \frac{X}{n}, I\rangle = \frac{\mathrm{tr}(X)}{n} = 1.$$

Hence (8) is a convex combination of the eigenvalues of $\mathcal{L}$, bounded by $\|\widetilde{\varepsilon}\|_1/n$. The smaller this bound, the more lopsided the convex combination toward λ_1, i.e., the larger c_1. The following lemma makes this observation precise.

Lemma 1. *Let $\mu = \sum_j c_j\lambda_j$ with $c_j \ge 0$, $\sum_j c_j = 1$, and $\lambda_1 = 0$. If $\mu \le \lambda_2$, then*

$$c_1 \ge 1 - \frac{\mu}{\lambda_2}.$$

Proof of lemma 1.

$$\mu = \sum_{i\ge 2} c_j\lambda_j \ge \lambda_2 \sum_{j\ge 2} c_j = \lambda_2(1 - c_1),$$

then isolate c_1. □

Assuming $\|\widetilde{\varepsilon}\|_1 \le n\lambda_2$, we now have

$$\langle \frac{X}{n}, v_1 v_1^*\rangle \ge 1 - \frac{\|\widetilde{\varepsilon}\|_1}{n\lambda_2}.$$

We can further bound

$$\|\frac{X}{n} - v_1 v_1^*\|_F^2 = \mathrm{tr}\left[\left(\frac{X}{n} - v_1 v_1^*\right)^2\right]$$
$$= \mathrm{tr}((v_1 v_1^*)^2) + \frac{\mathrm{tr}(X^2)}{n^2} - 2\ \mathrm{tr}\left(\frac{X}{n} v_1 v_1^*\right).$$

The first term is 1. The second term is less than 1, since $\mathrm{tr}(X^2) \leq \mathrm{tr}(X)^2$ for positive semidefinite matrices. Therefore,

$$\|\frac{X}{n} - v_1 v_1^*\|_F^2 \leq 2 - 2\ \mathrm{tr}\left(\frac{X}{n} v_1 v_1^*\right) \leq \frac{2\|\widetilde{\varepsilon}\|_1}{n\lambda_2}.$$

We can now control the departure of the top eigenvector of X/n from v_1 by the following lemma. It is analogous to the sin theta theorem of Davis-Kahan, except for the choice of normalization of the vectors. (It is also a generalization of a lemma used by one of us in [9] (section 4.2).) The proof is only given for completeness.

Lemma 2. *Consider any Hermitian $X \in \mathbb{C}^{n\times n}$, and any $v \in \mathbb{C}^n$, such that $\|X - vv^*\| < \frac{\|v\|^2}{2}$. Let η_1 be the leading eigenvalue of X, and x_1 the corresponding unit-norm eigenvector. Let x be defined either as (a) $x_1\|v\|$, or as (b) $x_1\sqrt{\eta_1}$. Then*

$$\|\ x\|x\| - e^{i\alpha} v\|v\|\ \| \leq 2\sqrt{2}\ \|X - vv^*\|,$$

for some $\alpha \in [0, 2\pi)$.

Proof of Lemma 2. Let $\delta = \|X - vv^*\|$. Notice that $\|vv^*\| = \|v\|^2$. Decompose $X = \sum_{j=1}^n x_j \eta_j x_j^*$ with eigenvalues η_j sorted in decreasing order. By perturbation theory for symmetric matrices (Weyl's inequalities),

$$\max\{|\|v\|^2 - \eta_1|, |\eta_2|, \ldots, |\eta_n|\} \leq \delta, \tag{9}$$

so it is clear that $\eta_1 > 0$, and that the eigenspace of η_1 is one-dimensional, as soon as $\delta < \frac{\|v\|^2}{2}$.

Let us deal first with the case (a) when $x = x_1\|v\|$. Consider

$$vv^* - xx^* = vv^* - X + Y,$$

where

$$Y = x_1(\|v\|^2 - \eta_1)x_1^* + \sum_{j=2}^n x_j \eta_j x_j^*.$$

From (9), it is clear that $\|Y\| \leq \delta$. Let $v_1 = v/\|v\|$. We get

$$\|vv^* - xx^*\| \leq \|vv^* - X\| + \|Y\| \leq 2\delta.$$

Pick α so that $|v^*x| = e^{-i\alpha} v^* x$. Then

$$\begin{aligned}
\|\ v\|v\| - e^{-i\alpha} x\|x\|\ \|^2 &= \|v\|^4 + \|x\|^4 - 2\,\|v\|\,\|x\|\,\Re\, e^{-i\alpha} v^* x \\
&= \|v\|^4 + \|x\|^4 - 2\,\|v\|\,\|x\|\,|v^* x| \qquad \text{by definition of } \alpha \\
&\leq \|v\|^4 + \|x\|^4 - 2\,|v^* x|^2 \qquad \text{by Cauchy-Schwarz} \\
&= \|vv^* - xx^*\|_F^2 \\
&\leq 2\|vv^* - xx^*\|^2 \qquad \text{because } vv^* - xx^* \text{ has rank 2} \\
&\leq 8\delta^2.
\end{aligned}$$

The case (b) when $x = x_1\sqrt{\eta_1}$ is treated analogously. The only difference is now that

$$Y = \sum_{j=2}^{n} x_j \eta_j x_j^*.$$

A fortiori, $\|Y\| \leq \delta$ as well.

□

Part (a) of lemma 2 is applied with X/n in place of X, and v_1 in place of v. In that case, $\|v_1\| = 1$. We conclude the proof by noticing that $v_1 = \frac{x_0}{\sqrt{n}}$, and that the output x of the lifting method is normalized so that $x_1 = \frac{x}{\sqrt{n}}$.

2.2 Proof of theorem 3

The proof is a simple argument of perturbation of eigenvectors. We either assume $\varepsilon_{ij} = \overline{\varepsilon_{ji}}$ or enforce it by symmetrizing the measurements. Define $\mathcal{L}$ as previously, and notice that $\|\mathcal{L} - \widetilde{\mathcal{L}}\| \leq \|\varepsilon\|$. Consider the eigen-decompositions

$$\mathcal{L}v_j = \lambda_j v_j, \qquad \widetilde{\mathcal{L}}\widetilde{v}_j = \widetilde{\lambda}_j \widetilde{v}_j,$$

with $\lambda_1 = 0$. Form

$$\widetilde{\mathcal{L}}v_j = \lambda_j v_j + r_j,$$

with $\|r_j\| \leq \|\varepsilon\|$. Take the dot product of the equation above with $\widetilde{v}_k$ to obtain

$$\langle \widetilde{v}_k, r_j \rangle = (\widetilde{\lambda}_k - \lambda_j)\langle \widetilde{v}_k, v_j \rangle.$$

Let $j = 1$, and use $\lambda_1 = 0$. We get

$$\sum_{k\geq 2} |\langle \widetilde{v}_k, v_1 \rangle|^2 \leq \frac{\sum_{k\geq 2} |\langle \widetilde{v}_k, r_1 \rangle|^2}{\max_{k\geq 2} |\widetilde{\lambda}_k|^2} \leq \frac{\|\varepsilon\|^2}{\widetilde{\lambda}_2^2}.$$

As a result,

$$|\langle \widetilde{v}_1, v_1 \rangle|^2 \geq 1 - \frac{\|\varepsilon\|^2}{\widetilde{\lambda}_2^2}.$$

Choose α so that $\langle e^{i\alpha}\widetilde{v}_1, v_1 \rangle = |\langle \widetilde{v}_1, v_1 \rangle|$. Then

$$\begin{aligned}
\|v_1 - e^{i\alpha}\widetilde{v}_1\|^2 &= 2 - 2\Re\langle e^{i\alpha}\widetilde{v}_1, v_1 \rangle \\
&= 2 - 2|\langle \widetilde{v}_1, v_1 \rangle| \\
&\leq 2 - 2|\langle \widetilde{v}_1, v_1 \rangle|^2 \\
&\leq 2\frac{\|\varepsilon\|^2}{\widetilde{\lambda}_2^2}.
\end{aligned}$$

Conclude by multiplying through by n and taking a square root.

2.3 Proof of theorem 4.

The proof follows the argument in section 2.1; we mostly highlight the modifications.

Let $b_i = |b_i|e^{i\phi_i}$. The Laplacian with phases is $\mathcal{L}_b = \Lambda_\phi L_{|b|} \Lambda_\phi^*$, with $\Lambda_\phi = \mathrm{diag}(e^{i\phi_i})$. Explicitly,

$$(\mathcal{L}_b)_{ij} = \begin{cases} \sum_{k:(i,k)\in E} |b_k|^2 & \text{if } i = j; \\ -b_i\overline{b_j} & \text{if } (i,j) \in E; \\ 0 & \text{otherwise,} \end{cases}$$

The matrix $Y = AXA^*$ is compared to the rank-1 spectral projectors of $\mathcal{L}_b$. We can write it as $Y = bb^* + \widetilde{\varepsilon}$ with $\|\widetilde{\varepsilon}\|_1 \le \|\varepsilon\|_1 + \sigma$. The computation of $\langle bb^*, \mathcal{L}_b\rangle$ is now

$$\begin{aligned} \langle bb^*, \mathcal{L}_b\rangle &= \sum_i b_i\overline{b_i\mathcal{L}_{ii}} + \sum_{(i,j)\in E} b_i\overline{b_j}\,\overline{\mathcal{L}_{ij}} \\ &= -\sum_i |b_i|^2 \sum_{k:(i,k)\in E} |b_k|^2 + \sum_{(i,j)\in E} b_i\overline{b_j}\,\overline{b_i}b_j \\ &= \sum_i |b_i|^2 \left[-\sum_{j:(i,j)\in E} |b_j|^2 + \sum_{j:(i,j)\in E} |b_j|^2 \right] \\ &= 0. \end{aligned}$$

The error term is now bounded (in a rather crude fashion) as

$$|\langle \widetilde{\varepsilon}, \mathcal{L}_b\rangle| \le \|\mathcal{L}_b\|_\infty \|\widetilde{\varepsilon}\|_1 \le \left[\max_i \sum_{j:(i,j)\in E} |b_j|^2 \right] \|\widetilde{\varepsilon}\|_1 \le \|b\|^2 \|\widetilde{\varepsilon}\|_1.$$

Upon normalizing Y to unit trace, we get

$$|\langle \frac{Y}{\mathrm{tr}(Y)}, \mathcal{L}_b\rangle| \le \frac{\|b\|^2\|\widetilde{\varepsilon}\|_1}{\|b\|^2 + \mathrm{tr}(\widetilde{\varepsilon})} \le 2\|\widetilde{\varepsilon}\|_1,$$

where the last inequality follows from

$$\begin{aligned} |\mathrm{tr}(\widetilde{\varepsilon})| &\le \|\widetilde{\varepsilon}\|_1 \\ &\le \|\varepsilon\|_1 + \sigma \\ &\le \lambda_2/2 \qquad \text{(assumption of the theorem)} \\ &\le \|b\|^2/2 \qquad \text{(by Gershgorin).} \end{aligned}$$

On the other hand, we expand

$$\langle \frac{Y}{\mathrm{tr}(Y)}, \mathcal{L}_b\rangle = \sum_j c_j\lambda_j,$$

and use $X \succeq 0 \Rightarrow Y \succeq 0$ to get $c_j \ge 0$, $\sum_j c_j = 1$. Since $2\|\widetilde{\varepsilon}\|_1 \le \lambda_2$, we conclude as in section 2.1 that

$$\langle \frac{Y}{\mathrm{tr}(Y)}, v_1v_1^*\rangle \ge 1 - \frac{2\|\widetilde{\varepsilon}\|_1}{\lambda_2},$$

hence

$$\|\frac{Y}{\mathrm{tr}(Y)} - v_1v_1^*\|_F^2 \le 4\frac{\|\widetilde{\varepsilon}\|_1}{\lambda_2}. \tag{10}$$

For $X = A^+Y(A^*)^+$, we get

$$\|\frac{X}{\mathrm{tr}(Y)} - (A^+v_1)(A^+v_1)^*\|_F^2 \leq 4\|A^+\|^4 \frac{\|\widetilde{\varepsilon}\|_1}{\lambda_2}.$$

Call the right-hand side δ^2. Recall that $v_1 = b/\|b\|$ hence $A^+v_1 = x_0/\|b\|$. Using $\mathrm{tr}(Y) = \|b\|^2 + \mathrm{tr}(\widetilde{\varepsilon})$, we get

$$\|X - x_0x_0^*\| \leq \delta\,\mathrm{tr}(Y) + \frac{\|x_0\|^2}{\|b\|^2}\,|\mathrm{tr}(\widetilde{\varepsilon})|. \tag{11}$$

Elementary calculations based on the bound $\|\widetilde{\varepsilon}\|_1 \leq \lambda_2/2 \leq \|b\|^2/2$ allow to further bound the above quantity by $\frac{(6+\sqrt{2})}{4}\delta\|b\|^2$. We can now call upon lemma 2, part (b), to obtain

$$\|x\|x\| - e^{i\alpha}x_0\|x_0\|\| \leq 2\sqrt{2}\frac{(6+\sqrt{2})}{4}\delta\|b\|^2,$$

where $x = x_1\sqrt{\lambda_1(X)}$ is the leading eigenvector of X normalized so that $\|x\|^2 = \lambda_1(X)$ is the leading eigenvalue of X. We use (11) one more time to bound

$$|\lambda_1(X) - \|x_0\|^2| \leq \frac{(6+\sqrt{2})}{4}\delta\|b\|^2,$$

hence

$$\begin{aligned}
\|x_0\|\,\|x - e^{i\alpha}x_0\| &\leq \|x\|x\| - e^{i\alpha}x_0\|x_0\| + \|x\|\,|\|x\| - \|x_0\||\\
&\leq 2\sqrt{2}\frac{(6+\sqrt{2})}{4}\delta\|b\|^2 + \frac{\|x\|}{\|x\| + \|x_0\|}|\|x\|^2 - \|x_0\|^2|\\
&\leq (2\sqrt{2}+1)\frac{(6+\sqrt{2})}{4}\delta\|b\|^2.
\end{aligned}$$

Use $\|b\| \leq \|A\|\,\|x_0\|$ and the formula for δ to conclude that

$$\|x - e^{i\alpha}x_0\| \leq C\,\|x_0\|\,\kappa(A)^2\sqrt{\frac{\|\widetilde{\varepsilon}\|_1}{\lambda_2}},$$

with $C = 2(2\sqrt{2}+1)\frac{(6+\sqrt{2})}{4} \leq 15$.

2.4 Proof of theorem 5.

The proof proceeds as in the previous section, up to equation (10). The rest of the reasoning is a close mirror of the one in the previous section, with Y in place of X, y in place of x, b in place of x_0, and δ re-set to $2\sqrt{\|\widetilde{\varepsilon}\|_1/\lambda_2}$. We obtain

$$\|y - e^{i\alpha}b\| \leq 15\,\|b\|\sqrt{\frac{\|\widetilde{\varepsilon}\|_1}{\lambda_2}}.$$

We conclude by letting $x = A^+y$, $x_0 = A^+b$, and using $\|b\| \leq \|A\|\|x_0\|$.

3 Numerical illustrations

3.1 Influence of the spectral gap

We investigate the scalings of the bounds for phase recovery given by theorems 1 and 3 on toy examples ($n = 2^7$), with respect to the spectral gap. This is achieved by considering three types of graphs :

- the path P_n which is proven to be the connected graph with the smallest spectral gap[5];
- graphs obtained by adding randomly K edges to P_n with K ranging from 1 to 50;
- Erdős-Rényi random graphs with probability ranging from 0.03 to 0.05, conditioned on connectedness (positive specrtal gap).

A realization of the two latter types of graphs is given in figure 1.

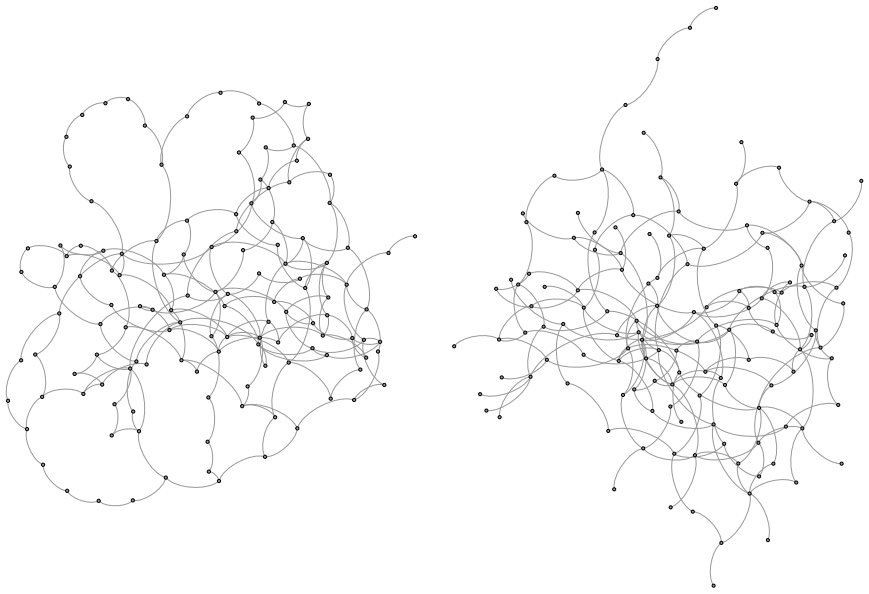

Figure 1: P_n + random edges (left), Erdős-Rényi random graph (right)

To study the eigenvector method, we draw one realization of a symmetric error matrix ε with $\varepsilon_{ij} \sim \mathbb{C}\mathcal{N}(0, \eta^2)$, with $\eta = 10^{-8}$. The spectral norm of the noise (used in theorem 3) is $||\varepsilon|| \sim 2 \times 10^{-7}$.
For different realizations of the aforementioned graphs, we estimate the solution with the eigenvector method and plot the ℓ_2 recovery error as a function of $\widetilde{\lambda}_2$. See figure 2.

[5]As justified by the decreasing property of λ_2 under edge removal, mentioned earlier.

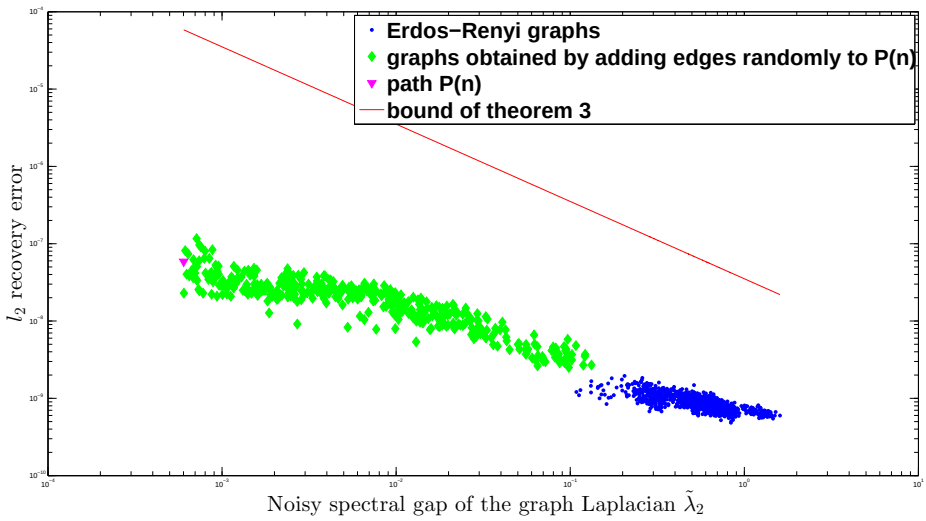

Figure 2: Recovery error for the eigenvector method as a function of $\widetilde{\lambda}_2$

To study the feasibility method, we consider the case of an approximate fit ($\sigma = 10^{-4}$) in the noiseless case ($\varepsilon = 0$). The feasibility problem (4) is solved using the Matlab toolbox cvx which calls the toolbox SeDuMi. An interior point algorithm (centering predictor-corrector) is used. The recovery error as a function of the spectral gap λ_2 is illustrated in figure 3.

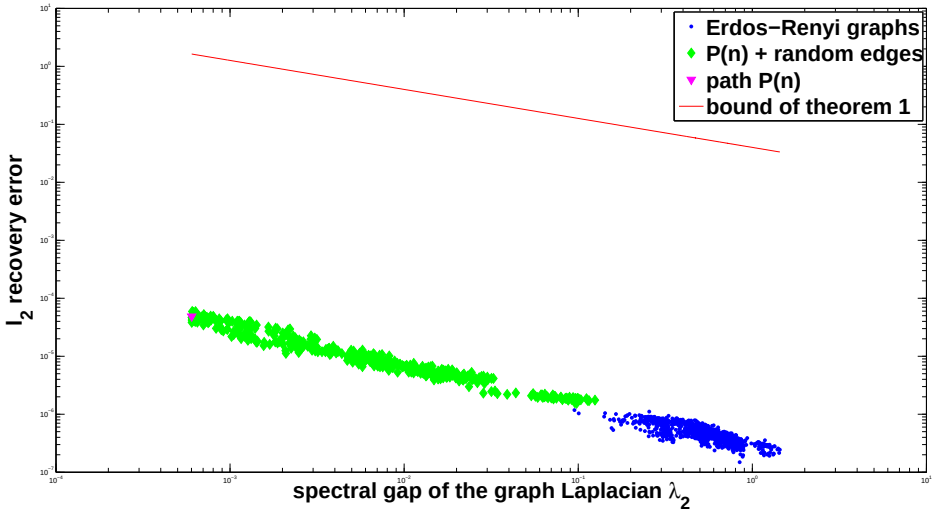

Figure 3: Recovery error for the feasibility method as a function of λ_2

3.2 Interferometric inverse scattering

An important application of interferometric ideas is to the inversion of a medium's index of refraction from recordings of waves scattered by that medium, as in seismic imaging.

In a first numerical experiment, we let $b = Ax$ where x is a reflectivity profile in a rectangle (perturbation of the index of refraction), and A is an acoustic time-harmonic wave equation that

maps this reflectivity profile in a linearized way to the solution wavefield b sampled at receivers placed at the top edge of the rectangle (surface). The sources in the wave equation are pointwise, and also placed at the surface. The wave equation is discretized via finite differences, with schemes of different orders for the data modeling step and for the inversion step. The data index i runs over receiver locations x_r, frequencies ω, and source positions x_s (which define different wave equations with different righ-hand sides.) Figure 4 shows robust recovery of x from noisy b, both by least-squares and by interferometric inversion. Here the noise model is taken to be Gaussian,

$$\widetilde{b}_i = b_i + \eta_i \quad \eta_i \sim \mathbb{C}\mathcal{N}(0, \sigma^2)$$

where $\sigma = 0.1\dfrac{||b||_2}{\sqrt{2n}}$, so that $\dfrac{||\eta||_2}{||b||_2} = 0.1$ (10% additive noise).

In this case the graph E is taken to be an Erdős-Rényi random graph with $p = 1.5\dfrac{log(N)}{N}$ to ensure connectedness. The computational method used for handling this example is a scalable rank-2 relaxation scheme, which is not new and is explained in the companion note [14]. The details concerning the time-harmonic wave equation can also be found in this note.

In a second numerical experiment, we show that interferometric inversion is still accurate and stable, even when the forward model $b = \mathcal{A}(x)$ is the full wave equation that maps the index of refraction x to the wavefield b nonlinearly (no Born approximation.) Again, 10% Gaussian noise is added. Figure 5 shows the result of nonlinear least-squares inversion, and the corresponding interferometric inversion result. To produce this latter result, we simply replace the classical adjoint-state (preconditioned) gradient step by interferometric inversion applied to the data residual.

Perhaps surprisingly, there is no apparent loss of resolution in either of Figure 4 (bottom) or Figure 5 (bottom), compared to their least-squares counterparts. This behavior is due to the precaution to *invert* from interferometric measurements, rather than just image with a backprojection operator as in [4]. This observation was also a key conclusion in the recent paper by Mason et al. [18]. Note that "backprojecting the cross-correlations" can be shown to be related to the first iteration of gradient descent in the *lifted semidefinite formulation* of interferometric inversion.

So far, our numerical experiments merely show that interferometric inversion can be accurate and stable under minimal assumptions on the graph E of data pair products. In the next section, we show that there is also an important *rationale* for switching to the interferometric formulation: its results display robustness vis-a-vis some uncertainties in the forward model A.

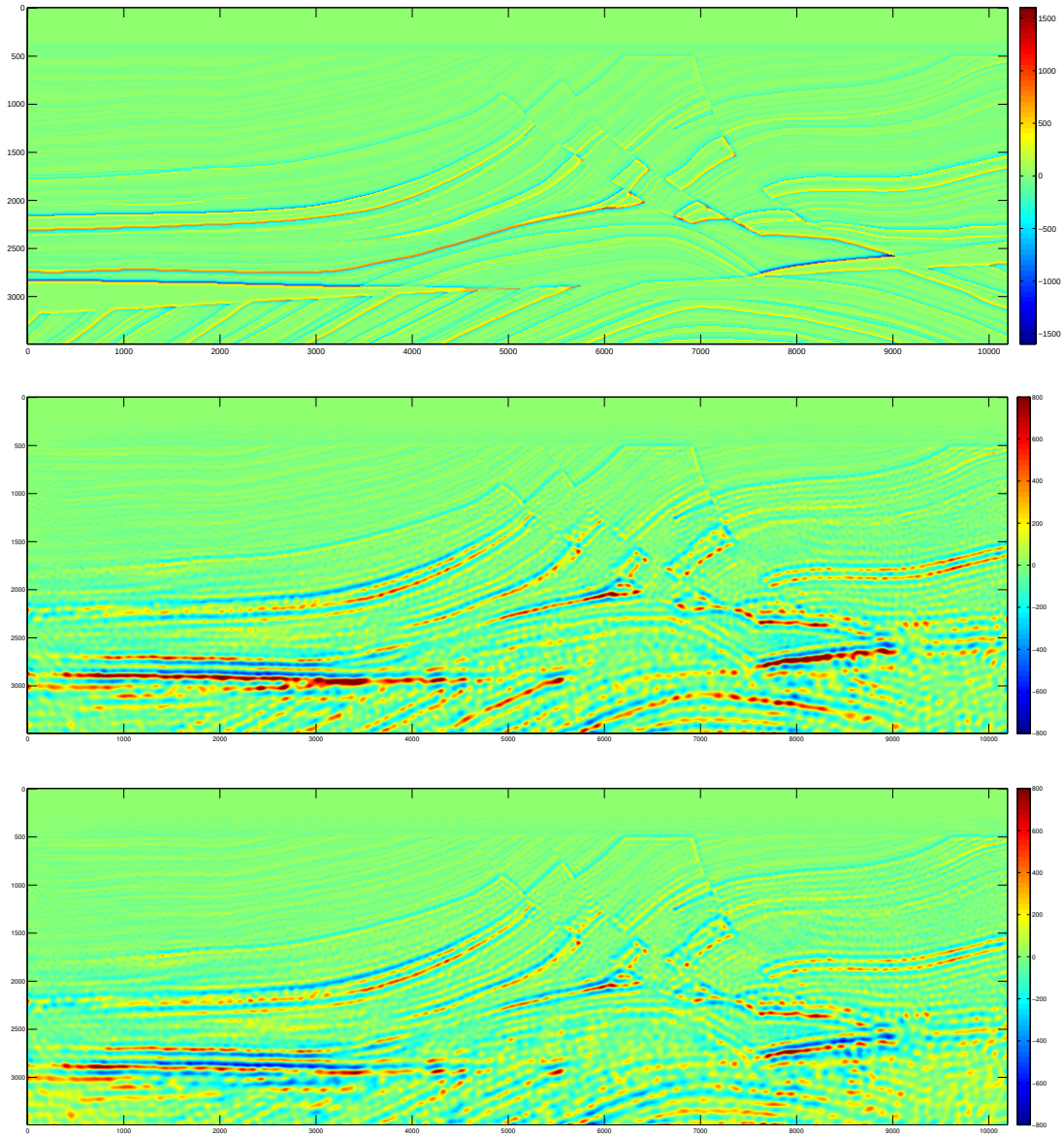

Figure 4: Top: true, unknown reflectivity profile x used to generate data. Middle: least-squares solution. Bottom: result of interferometric inversion.

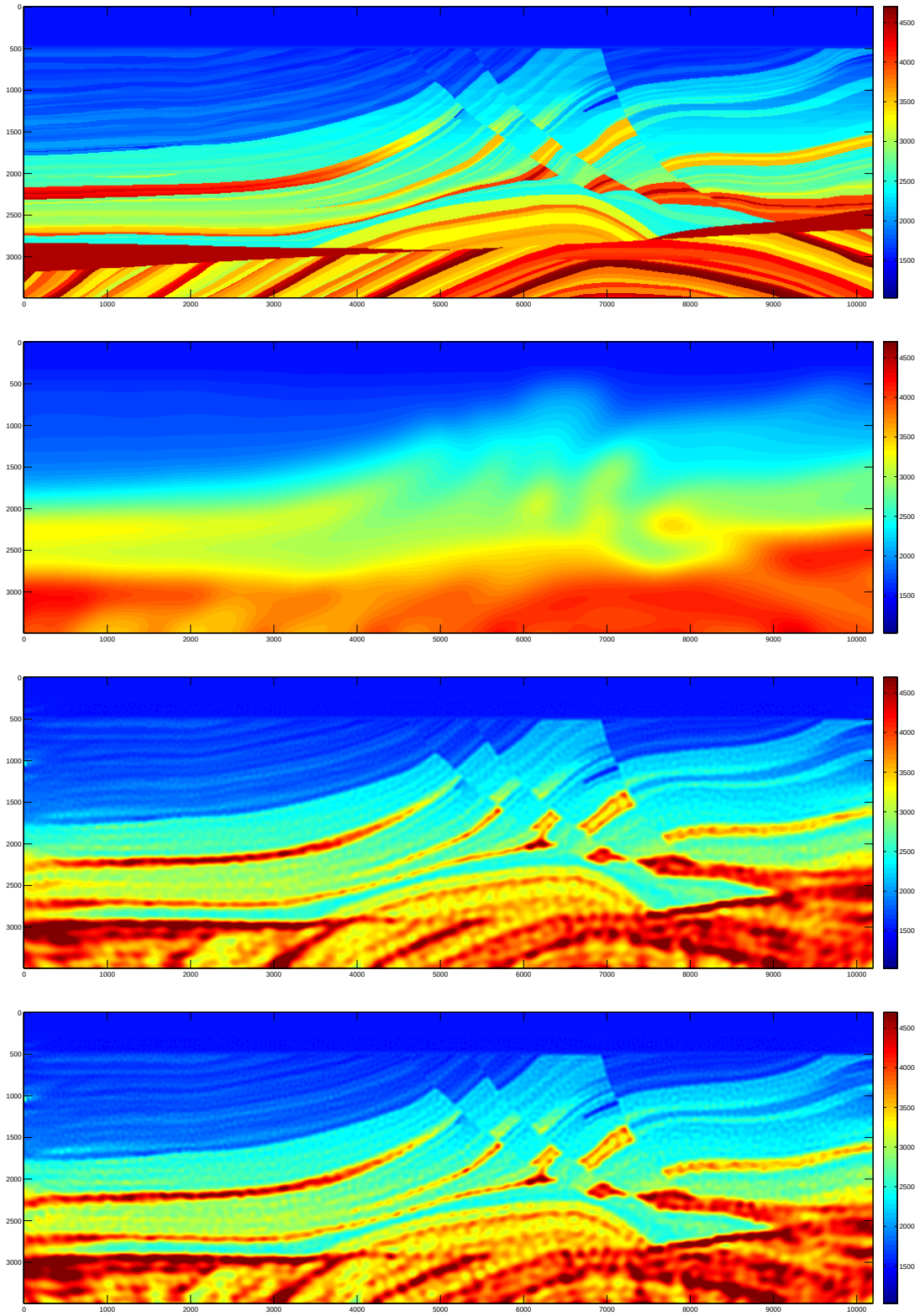

Figure 5: Top: true, unknown map of the index of refraction x used to generate data. Second: initial guess for either inversion scheme. Third: nonlinear least-squares solution, a.k.a. full waveform inversion. Bottom: result of interferometric inversion.

3.3 Model robustness

In this section we continue to consider wave-based imaging scenarios, but more constrained in the sense that the sources and receivers completely surround the object to be imaged, i.e., provide a full aperture. However, we now change some of the parameters in the forward map, so that the A map used for data modeling is not the same as the $\widetilde{A}$ map used for inversion. The interferometrix mask E is also different, and more physically meaningful than in the previous section: it indicates a small band around the diagonal $i = j$, i.e., it selects nearby receiver positions and frequencies, as in [4]. In all our experiments with model errors, it is crucial that the data misfit parameter σ be positive nonzero – it would be a mistake to try and explain the data perfectly with a wrong model.

Figure 6 illustrates the behavior of interferometric vs. classical least-squares inversion, on two different imaging examples with the Shepp-Logan phantom. We only report on the high-level conclusions; all the details about optimization schemes and wave equations can be found in the companion note [14].

- In a first (passive) scenario, called inverse source, A maps the right-hand-side of a time-harmonic wave equation to its solution. The wave speed profile is uniform, but *different* for the data generation and inversion stages. In the former case, $c = 1$, while in the latter case, $\widetilde{c} = 0.95$.

  As shown in Figure 6 (middle left), least-squares inversion does not properly handle this type of uncertainty and produces a defocused image. In contrast, interferometric inversion, shown in Figure 6 (bottom left), enjoys a better resolution. The price to pay for focusing is positioning: although we do not have a mathematical proof of this fact, the interferometric reconstruction is near a shrunk version of the true source distribution.

- In a second (active) scenario, called inverse scattering, A maps a reflectivity profile, which appears multiplied by an incident field in the righ-hand-side of a time-harmonic wave equation, to the solution of this latter equation. The wave speed profile is uniform ($c = 1$). In this example, the modeling error is assumed to be on the receiver positions: they have a smooth random deviation from a circle, as shown in Figure 6 (top right).

  Again, least-squares inversion produces a poor result (Figure 6 (middle right)), where the features of the phantom are not clearly reconstructed, and the strong outer layer is affected by a significant oscillatory error. Interferometric inversion produces a somewhat more focused reconstruction (Figure 6 (bottom right)), where more features of the phantom are recovered and the outer layer is well-resolved.

We leave numerous important questions unanswered, such as how to optimize the choice of the selector E in the presence of model uncertainties; how to select the data misfit parameter $\sigma > 0$ and whether there is a phase transition in σ for model robustness vs. lack thereof; how to justify the gain in focusing in some reconstructions; and whether this gain in focusing is part of a tradeoff with the geometric faithfulness of the recovered image.

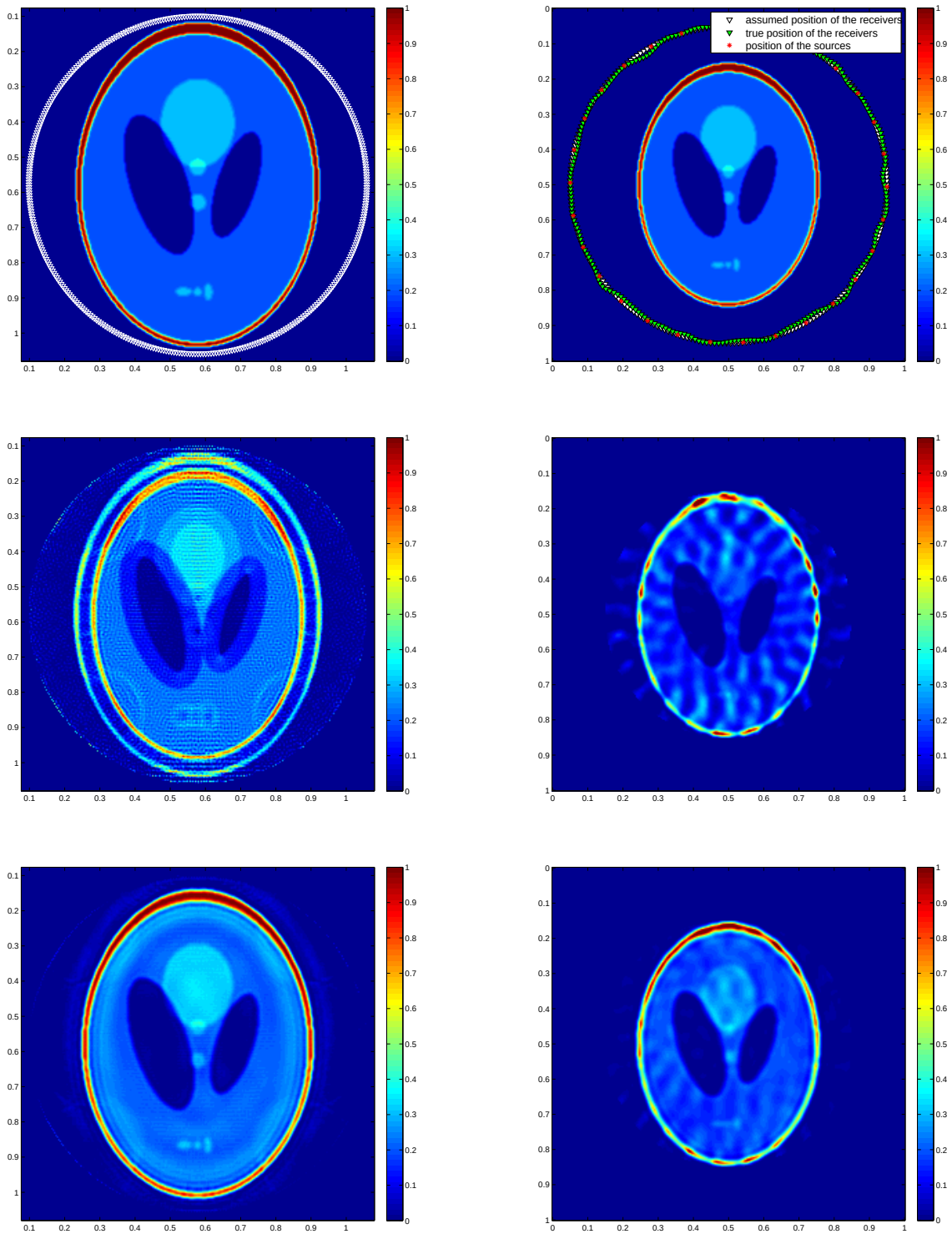

Figure 6: Top: Setting of the inverse source experiment (left), and setting of the inverse scattering experiment (right). Middle: least-squares reconstruction for inverse source (left) and inverse scattering (right). Bottom: interferometric reconstruction for inverse source (left) and inverse scattering (right).

References

[1] B. Alexeev, A. S. Bandeira, M. Fickus and D. G. Mixon. Phase retrieval with polarization, *SIAM J. Imaging Sci.*, 7(1), 3566

[2] A. S. Bandeira, A. Singer, and D. A. Spielman. A Cheeger inequality for the graph connection Laplacian, *SIAM. J. Matrix Anal. Appl.*, 34(4), 16111630

[3] P. Blomgren, G. Papanicolaou, and H. Zhao, Super-resolution in time-reversal acoustics, *J. Acoust. Soc. Am.* 111(1), 230-248, 2002

[4] L. Borcea, G. Papanicolaou, and C. Tsogka, Interferometric array imaging in clutter, *Inv. Prob.* 21(4), 1419-1460, 2005

[5] E. J. Candes, Y. C. Eldar, T. Strohmer and V. Voroninski. Phase retrieval via matrix completion, *SIAM J. Imaging Sci.*, 6(1), 199-225, 2013

[6] E. J. Candes, B. Recht, Exact Matrix Completion via Convex Optimization, *Found. Comput. Math.*, 9-6, 717-772, 2009

[7] A. Chai, M. Moscoso, and G. Papanicolaou. Array imaging using intensity-only measurements, *Inverse Problems 27.1 015005*, 2011

[8] S. Chatterjee, Matrix estimation by Universal Singular Value Thresholding, *Ann. Statist.* Volume 43, Number 1 (2015), 177-214

[9] L. Demanet and P. Hand. Stable optimizationless recovery from phaseless linear measurements, *Journal of Fourier Analysis and Applications*, Volume 20, Issue 1, (2014) pp 199221

[10] E. Dussaud, *Velocity analysis in the presence of uncertainty*, Ph.D. thesis, Computational and Applied Mathematics, Rice University, 2005

[11] J. Garnier, Imaging in randomly layered media by cross-correlating noisy signals, *SIAM Multiscale Model. Simul.* 4, 610-640, 2005

[12] M. Goemans and D. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *Journal of the ACM*, 42(6),1115-1145, 1995

[13] A. Javanmard, A. Montanari, Localization from incomplete noisy distance measurements, *Found. Comput. Math.* 13, 297-345, 2013

[14] V. Jugnon and L. Demanet, Interferometric inversion: a robust approach to linear inverse problems, in *Proc. SEG Annual Meeting, 2013*

[15] R. Keshavan, A. Montanari, S. Oh, Matrix Completion from Noisy Entries, *Journal of Machine Learning Research* 11, 2057-2078, 2010

[16] O. I. Lobkis and R. L. Weaver, On the emergence of the Greens function in the correlations of a diffuse field, *J. Acoustic. Soc. Am.*, 110, 3011-3017, 2001

[17] L. Lovasz, *Eigenvalues of graphs*, Lecture notes, 2007

[18] E. Mason, I. Son, B. Yazici, Passive Synthetic Aperture Radar Imaging Using Low-Rank Matrix Recovery Methods, *IEEE Journal of Selected Topics in Signal Processing*, 9-8, 2015

[19] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization, *SIAM Review* 52-3, 471-501, 2010

[20] J. M. Rodenburg, A. C. Hurst, A. G. Cullis, B. R. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson. Hard-x-ray lensless imaging of extended objects, *Physical review letters 98, no. 3 034801*, 2007

[21] J. Schotland, Quantum imaging and inverse scattering, *Optics letters*, 35(20), 3309-3311, 2010

[22] G. T. Schuster, *Seismic Interferometry* Cambridge University Press, 2009

[23] G. T. Schuster, J. Yu, J. Sheng, and J. Rickett, Interferometric/daylight seismic imaging *Geophysics* 157(2), 838-852, 2004

[24] A. Singer. Angular synchronization by eigenvectors and semidefinite programming, *Applied and computational harmonic analysis 30.1 20-36*, 2011

[25] A. Singer, M. Cucuringu, Uniqueness of low-rank matrix completion by rigidity theory, *SIAM. J. Matrix Anal. Appl.* 31(4), 16211641, 2010

[26] L. Wang and A. Singer. Exact and Stable Recovery of Rotations for Robust Synchronization, *Information and Inference: A Journal of the IMA* (2013) 2, 145193

[27] I. Waldspurger, A. d'Aspremont, S. Mallat, Phase recovery, maxcut and complex semidefinite programming, *Mathematical Programming*, Volume 149, Issue 1 (2015), pp 4781

[28] K. Wapenaar and J. Fokkema, Greens function representations for seismic interferometry, *Geophysics*, 71, SI33-SI46, 2006