Title

Correlation-informed ordered dictionary learning for imaging in complex media.

Journal

Proceedings of the National Academy of Sciences of the United States of America, 121(11)

ISSN

0027-8424

Authors

Moscoso, Miguel

Novikov, Alexei

Papanicolaou, George

et al.

Publication Date

2024-03-12

DOI

10.1073/pnas.2314697121

Correlation-informed ordered dictionary learning for imaging in complex media

M. Moscoso*, A. Novikov†, G. Papanicolaou‡and C. Tsogka§

March 20, 2024

Abstract

We propose a method for imaging in scattering media when large and diverse data sets are available. It has two steps. Using a dictionary learning algorithm the first step estimates the true Green's function vectors as columns in an unordered sensing matrix. The array data comes from many sparse sets of sources whose location and strength are not known to us. In the second step the columns of the estimated sensing matrix are ordered for imaging using the Multi-Dimensional Scaling algorithm with connectivity information derived from cross correlations of its columns, as in time reversal. For these two steps to work together we need data from large arrays of receivers so the columns of the sensing matrix are incoherent for the first step, as well as from sub-arrays so that they are coherent enough to obtain connectivity needed in the second step. Through simulation experiments, we show that the proposed method is able to provide images in complex media whose resolution is that of a homogeneous medium.

1 Introduction

High-resolution imaging in complex media faces challenges due to wavefront distortion caused by scattering from inhomogeneities. In this paper we introduce a new method for imaging in a randomly inhomogeneous medium involving two basic algorithms. The first is a sparse dictionary learning

*Department of Mathematics, Universidad Carlos III de Madrid, Leganes, Madrid 28911, Spain

†Department of Mathematics, Pennsylvania State University, University Park, PA 16802

‡Department of Mathematics, Stanford University, Stanford, CA 94305

§Department of Applied Mathematics, University of California, Merced, CA 95343

algorithm [17] in order to estimate Green's function vectors between focal or source points in the image window and receiver locations on the array. The second is a Multi-Dimensional Scaling (MDS) algorithm [10] to convert information about correlations of Green's function vectors into positions of the focal points in the image window.

To accomplish the first step, we use a sparsity promoting modification of the Method of Optimal Directions (MOD) [14] to learn an (unordered) dictionary of Green's function vectors that characterize the propagation of signals from a set of focal points, or sources in the image window, to the array. Here unordered means that we do not know which focal points are associated with the estimated column vectors of the dictionary. In this step we assume that an abundance of sensing measurements is available. Specifically, we have access to measurements for multiple signals emanating from many sparse sets of sources, but we do not have prior knowledge of their locations or amplitudes. This dictionary learning algorithm enables us to estimate Green's function vectors with high accuracy under the condition that these vectors are sufficiently incoherent, which means that their normalized inner product is sufficiently small. Given a configuration of sources or focal points in the image window, this implies that the receiver array must be large enough. We present this dictionary learning step in Section 33.1.

The goal of the second step is to associate each Green's function vector with its corresponding focal point in the image window, which means that we want to find the correct order of the columns in the estimated matrix of Green's function vectors. We could back propagate these vectors into the image window using a reference homogeneous medium. This is the Kirchhoff's migration approach that only works well when the fluctuations are weak [8]. For a given set of Green's function vectors we could also try to estimate the position of their focal points using source localization algorithms, commonly used in wireless communications [4, 25, 27, 26]. However, these algorithms use information based on distances between sources and receivers and are therefore very sensitive to noise. They are not suitable for imaging in media with strong fluctuations.

Instead of estimating the distance between a focal point and a receiver, we can obtain a much more accurate estimate of the distance between two nearby focal points. The correlation of the estimated Green's function vectors gives such an estimate but, of course, we do not know where their focal points are located in the image window. By cross correlating each Green's function vector with all the others we can identify its nearest neighbors. This is a key observation that allows us to generate a proxy distance between column or Green's function vectors by counting the smallest number

of neighborhoods that connect them. This provides a connectivity-based proxy distance between all pairs of column vectors that we can use with the Multidimensional Scaling (MDS) algorithm [10] for identifying Green's function vectors with their focal points up to a rotation, translation and scaling. The resulting relative configuration of points can be spatially fixed with a few (two or three) known reference points in the two dimensional image window. The use of a proxy metric based on connectivity is done by the MDS-MAP algorithm [24, 21]. Constructing the connectivity-based proxy distance using cross correlations is described in Section 33.2. We should point out that this use of cross-correlations is different from another use of cross-correlations in imaging, in fields such as geophysics [16, 5], where the gathered data is cross-correlated to estimate the Green's function between two receivers in complex media.

2 Imaging problem setup

Suppose that an array of N receivers records waves generated by sources located over a region of interest, called the image window. The receivers are located at points $\vec{r}_j$, and the sources at unknown locations $\vec{z}_i$. In Figure 1, it is assumed that the array is one-dimensional. The coordinates parallel to this array are the cross-range coordinates, and the ones orthogonal to it the range coordinates. The medium between the array and the unknown sources fluctuates randomly in space as illustrated in Figure 1. The Green's function that characterizes wave propagation of a signal of frequency ω from a point $\vec{z}$ to a point $\vec{r}$ in the random medium satisfies the wave equation

$$\Delta G(\vec{z},\vec{r}) + \kappa^2\, n^2(\vec{z})\, G(\vec{z},\vec{r}) = \delta(\vec{z}-\vec{r}), \tag{1}$$

where $\kappa = \omega/c_0$ is the wavenumber with constant reference wave speed c_0. The random index of refraction is $n(\vec{z}) = c_0/c(\vec{z})$ with local wave speed $c(\vec{z})$. In a homogeneous medium, $c(\vec{z}) \equiv c_0$ for any location $\vec{z}$ and, in this case, $G(\vec{z},\vec{r}) = G_0(\vec{z},\vec{r})$, where

$$G_0(\vec{z},\vec{r}) = \frac{\exp(i\,\kappa\,|\vec{z}-\vec{r}|)}{4\pi|\vec{z}-\vec{r}|}\,. \tag{2}$$

In random media, however, the wave speed $c(\vec{z})$ depends on the position $\vec{z}$. We consider a variable wave speed satisfying

$$\frac{1}{c^2(\vec{z})} = \frac{1}{c_0^2}\left(1+\sigma\mu(\frac{\vec{z}}{l})\right), \tag{3}$$

where l is the correlation length of the inhomogeneities that is characteristic of their size. In (3), σ determines the strength of the fluctuations around the constant speed c_0, and $\mu(\cdot)$ is a stationary random process with zero mean and normalized autocorrelation function $R(|\vec{z}_i - \vec{z}_{i'}|) = \mathbb{E}(\mu(\vec{z}_i)\mu(\vec{z}_{i'}))$, so $R(0) = 1$. .

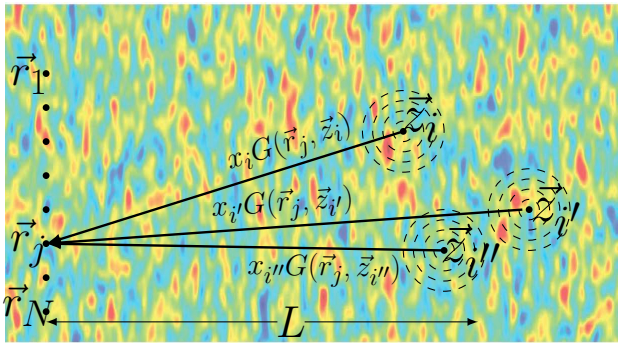

Figure 1: Schematic of the data collected on the array when three sources located at $\vec{z}_i$, $\vec{z}_{i'}$ and $\vec{z}_{i''}$ simultaneously emit signals. It is the superposition of these signals that are recorded on the array of receivers located at points $\vec{r}_j$, $j = 1, \ldots, N$.

We write the data received on the array of N receivers in vector form, with Green's function vector

$$\boldsymbol{g}(\vec{z}) = [G(\vec{r}_1, \vec{z}), G(\vec{r}_2, \vec{z}), \ldots, G(\vec{r}_N, \vec{z})]^T , \tag{4}$$

and we introduce the $N \times K$ sensing matrix

$$\boldsymbol{\mathcal{G}} = [\boldsymbol{g}(\vec{z}_1) \cdots \boldsymbol{g}(\vec{z}_K)] \tag{5}$$

defined on a grid $\{\vec{z}_i\}_{i=1,\ldots,K}$ spanning the image window. The sensing matrix $\boldsymbol{\mathcal{G}}$ maps a distribution of sources in the image window to the (single frequency) data received on the array. The multi-frequency case is described in Section 4. For a given configuration of sources on the grid represented by the vector $\boldsymbol{x} \in \mathbb{C}^K$, the data recorded on the array is given by

$$\boldsymbol{y} = \boldsymbol{\mathcal{G}}\, \boldsymbol{x} , \tag{6}$$

where $\boldsymbol{y} \in \mathbb{C}^N$. Here, $\boldsymbol{x}$ is a vector whose jth component represents the complex amplitude of the source at location $\vec{z}_j$ in the image window, $j = 1, \ldots, K$.

Because the medium is random, the sensing matrix $\boldsymbol{\mathcal{G}}$ in (6) is not known. Our imaging problem is to estimate this matrix from a set of M samples or

observations $\{\boldsymbol{y}_i\}_{i=1,\ldots,M}$, with $\boldsymbol{y}_i = \boldsymbol{\mathcal{G}}\,\boldsymbol{x}_i$. The number of observations is large with respect to the dimension of the vectors $\boldsymbol{x}_i$, i.e., $M \gg K$. Note that $\boldsymbol{x}_i$ is also unknown but we assume that it is sparse, implying that the samples $\boldsymbol{y}_i$ can be represented as a linear combination of a small number of columns of the unknown sensing matrix $\boldsymbol{\mathcal{G}}$. Since we do not know $\boldsymbol{x}_i$, both the locations and the amplitudes of the sources are unknown. We assume that for the few sources that are active for every sample the modulus of their amplitude takes values in a bounded interval away from zero.

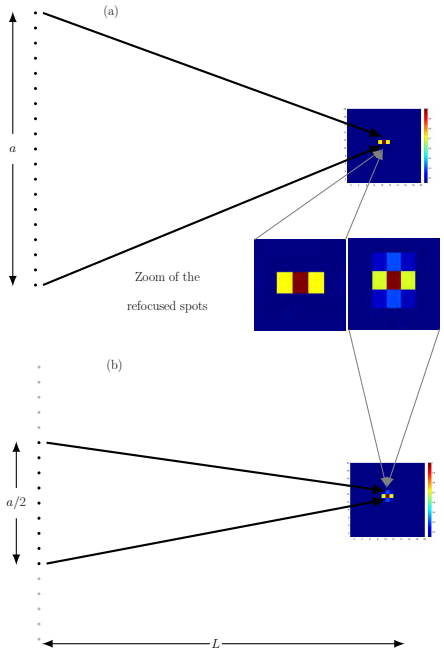

Figure 2: Refocused spots with (a) a large array used in the first step of the algorithm and (b) a small array used in the second step.

We note that, for imaging problems, the coherence between the columns of the sensing matrix $\boldsymbol{\mathcal{G}}$ increases as the grid in the image window becomes finer. This can be challenging for the sparse dictionary learning algorithm, described in the next section, as its convergence is guaranteed under incoherence or restricted isometry property assumptions [1]. Coherence is defined as the maximum of the normalized inner product between different columns of the matrix, *i.e.*,

$$\nu = \max_{\substack{i,j=1 \\ i \neq j}}^{K} \frac{|\boldsymbol{g}(\vec{z}_i)^* \boldsymbol{g}(\vec{z}_j)|}{||\boldsymbol{g}(\vec{z}_i)||\ ||\boldsymbol{g}(\vec{z}_j)||}. \tag{7}$$

Given the size of the recording array and the bandwidth, in order to limit the coherence between the columns of $\boldsymbol{\mathcal{G}}$ we assume that the grid in the image window is not finer than the support of the point spread function in a reference homogeneous medium; see Figure 2-(a). This is essential for the first step of the imaging method, the dictionary learning step. The point spread function of an imaging system is the image one obtains when the signal from a single source is used as input. The proposed two-step imaging method is described next.

3 The two-step imaging method

The first step of our imaging method uses a dictionary learning algorithm and allows us to recover the columns of the sensing matrix up to a permutation. Although in most of the applications of dictionary learning the order of the columns is not an issue, it is essential in the imaging case. That is because even though we recover the Green's function vectors, the imaging problem is still not solved as we do not know their correspondence with the grid points in the image window. To create an image we need to associate each column of the estimated matrix $\widehat{\boldsymbol{\mathcal{G}}}$ with the corresponding focal point in the image window. This is challenging since the propagation medium is unknown, so we cannot just back-propagate the recovered Green's function vectors, as it is done in time reversal.

In the second step of our method, we deal with the *focal spot localization problem* where we associate each recovered Green's function vector $\widehat{\boldsymbol{g}}_i$ with its corresponding focal or source point in the image window. The key idea is the use of cross-correlations between the columns of the estimated matrix $\widehat{\boldsymbol{\mathcal{G}}}$ to identify the nearest neighbors of their focal or source points in the image window and then infer their associated focal points on the grid. Coherence becomes crucial in this step, as we cannot identify the nearest neighbors of each focal point if $\widehat{\boldsymbol{\mathcal{G}}}$ is column-incoherent. We can do this, however, by using for the cross-correlations a suitable subset of the elements of the estimated columns $\widehat{\boldsymbol{g}}_i$ corresponding to a fraction of the size of the recording array. An example is illustrated in Figure 2-(b) where a subarray of half the size is used. The grid reconstruction problem is then solved using the MDS algorithm with a proxy distance, as in sensor network localization problems [21]. No Euclidian distance information is known but a proxy distance is obtained from connectivity information at limited range that is recovered from cross-correlations.

The cross-correlations formed can be interpreted as time reversal exper-

iments. That is, the signals recorded on the array are re-emitted in the same medium and given the time-reversibility of the wave equation, those signals will focus at the location from which the original signal was emitted. Our connectivity reconstruction relies on this fundamental property of the wave equation and therefore is robust to the complexity of the medium. Our numerical simulations confirm that this is the case.

3.1 Step one: Dictionary learning

We discuss here an algorithm aimed at learning from the gathered data a dictionary $\mathbf{A} \in \mathbb{C}^{N\times K}$ that represents the normalized sensing matrix (5). We assume that the recorded signals $\boldsymbol{y}_i \in \mathbb{C}^N$, $i = 1, \ldots, M$, the data, come from only a few sources and can therefore be represented as a linear combination of a small number of columns in the dictionary $\mathbf{A}$ we want to determine. This means that $\boldsymbol{y}_i = \mathbf{A}\,\boldsymbol{x}_i$, where $\boldsymbol{x}_i \in \mathbb{C}^K$ are sparse vectors that represent unknown collections of sources firing at the same time. For imaging applications, we can assume that the columns of $\mathbf{A}$ have unit lengths. We also assume that we know the dimension K (with usually $K \gg N$), which is the number of points in the image window and therefore specifies the resolution of the image. An estimate of K can be based on the resolution of the imaging setup expected in a homogeneous medium. In a random medium resolution in time reversal, but not in imaging, will improve [15, 7] and therefore K could be larger. We assume here that K is chosen based on a homogeneous medium.

To find the dictionary $\mathbf{A}$ and the layout of sources, we define the matrix $\mathbf{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M] \in \mathbb{C}^{K\times M}$ and the data matrix $\mathbf{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M] \in \mathbb{C}^{N\times M}$, and solve the problem

$$\begin{aligned} \min_{\mathbf{A},\mathbf{X}} \quad & \|\mathbf{A}\mathbf{X} - \mathbf{Y}\|_F^2 \\ \text{s.t.} \quad & \|\boldsymbol{x}_i\|_0 \le s,\ i = 1, \ldots, M, \end{aligned} \tag{8}$$

where $\|\cdot\|_0$ counts the number of non-zero elements and s is the expected sparsity level. The decomposition $\mathbf{Y} = \mathbf{A}\mathbf{X}$ is unique up to permutations of the columns of $\mathbf{A}$ and rows of $\mathbf{X}$ provided that the data $\mathbf{Y}$ is rich enough [3].

How much data do we need, that is, how big should M be? As already noted, in imaging we usually have $K \gg N$. If the sparsity s is fixed independent of N then the condition

$$M > K \log K \tag{9}$$

is sufficient for a suitable probabilistic model of $\mathbf{A}, \mathbf{X}, \mathbf{Y}$ [2, 20].

Problem (8) is non-convex, as the constraint is not convex and both $\mathbf{A}$ and $\mathbf{X}$ are unknown. However, its solution can be found efficiently by means of an alternating optimization procedure that uses the ℓ_1-norm instead of the sparsity count, provided the initialization is close enough to the true solution and the columns of $\mathbf{X}$ are sparse enough. We refer the reader to [1] for local convergence guarantees.

Specifically, if $\mathbf{A}$ is known, then $\mathbf{X}$ can be obtained from the ℓ_1-norm minimization problem

$$\min \|\mathbf{X}\|_1 \;\text{ subject to }\; \mathbf{A}\mathbf{X} = \mathbf{Y}\,, \tag{10}$$

that can be easily solved by several different algorithms [12, 22, 13, 6]. Here we solve (10) using a Generalized Lagrangian Multiplier Algorithm (GeLMA) [18]. In the next step, if $\mathbf{X}$ is known, the minimization problem for $\mathbf{A}$

$$\min_{\mathbf{A}} \quad \|\mathbf{A}\mathbf{X} - \mathbf{Y}\|_F^2 \quad , \tag{11}$$

can be easily solved. The exact solution is $\mathbf{A} = \mathbf{Y}\mathbf{X}^T(\mathbf{X}\mathbf{X}^T)^{-1}$, provided $\mathbf{X}\mathbf{X}^T$ is invertible (actually stably invertible). Once $\mathbf{A}$ has been computed we normalize its columns to one.

To summarize, in order to solve (8), we alternate between problems (10) and (11) to update $\mathbf{X}$ and $\mathbf{A}$ sequentially one after the other. Each iteration has two steps, starting from an initial guess $\mathbf{X}_0$ and $\mathbf{A}_0$. At the beginning of iteration $l \geq 1$ we have $\mathbf{A}_{l-1}$ and $\mathbf{X}_{l-1}$, and we solve (10) with $\mathbf{A} = \mathbf{A}_{l-1}$ to obtain $\mathbf{X}_l$ using GeLMA. Then, we solve (11) with $\mathbf{X} = \mathbf{X}_l$ fixed to find

$$\mathbf{A}_l = \mathbf{Y}\mathbf{X}_l^T(\mathbf{X}_l\mathbf{X}_l^T)^{-1} \tag{12}$$

in the second step. This is very much like the Method of Optimal Directions (MOD) algorithm proposed in [14] for signal compression that uses Matching Pursuit for finding $\mathbf{X}_l$ instead of GeLMA in the first step.

Our numerical experiments indicate that, given a suitable initialization, this dictionary learning algorithm can construct the matrix of Green's functions for wave propagation in random media. However, the columns of this matrix, the dictionary, are unordered. They cannot be used for imaging because we do not know the points where they focus in the image window. The next Section addresses this problem.

3.2 Step two: Grid reconstruction

We describe now an algorithm for finding the focal spots in the image window from the *estimated* Green's function vectors $\{\widehat{\boldsymbol{g}}_i\}_{i=1}^K$. It is the range-free or

connectivity based sensor localization algorithm [24], analyzed in [21]. Our main contribution here is to determine connectivity from cross-correlations of the Green's function vectors $\{\widehat{\boldsymbol{g}}_i\}_{i=1}^K$, which now must retain some coherence. However, the dictionary learning algorithm of the previous section requires incoherence, which means that the ν in (7) for the estimated $\{\widehat{\boldsymbol{g}}_i\}_{i=1}^K$ is small. By using a subset of the components of the $\{\widehat{\boldsymbol{g}}_i\}_{i=1}^K$, corresponding to a subarray as illustrated in Figure 2, we increase their coherence. In this Section we use the same notation $\{\widehat{\boldsymbol{g}}_i\}_{i=1}^K$ for their subsampled components.

This increased coherence is essential for the connectivity-based localization algorithm because it allows us to introduce a graph $G = (V, E)$ where the vertex set $V = \{1, 2, \ldots, K\}$ is associated with the estimated Green's function vectors $\{\widehat{\boldsymbol{g}}_i\}_{i=1}^K$. A pair of vertices (i, j) is then connected by an edge in E, and assigned the value one in the adjacency matrix of the graph, if the cross-correlation $\widehat{\boldsymbol{g}}_i^* \widehat{\boldsymbol{g}}_j$ is sufficiently close to one in absolute value. Otherwise, the pair (i, j) is not connected, and zero is entered in the adjacency matrix. The size of the subarray of receivers is adjusted so that each vertex has up to $k = 2r$ edges, where $r = 2, 3$ is the ambient dimension of the image window.

The proxy distance between two Green's function vectors is now the geodesic graph distance between their corresponding vertices. That is, the proxy distance between $\widehat{\boldsymbol{g}}_i$ and $\widehat{\boldsymbol{g}}_j$, denoted by $\widehat{d}_{ij}$, is the number of edges in the shortest path connecting i and j. We use this proxy distance as a replacement of the Euclidean distance between pairs of focal points in the image window associated with Green's function vectors in the MDS algorithm. The resulting configuration of focal points $\widehat{Z} = [\widehat{\vec{z}}_1, \widehat{\vec{z}}_2, \ldots, \widehat{\vec{z}}_K]^T$ in the image window provides an estimate for the true configuration of focal points $Z = [\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_K]^T$, up to rotation, translation and scaling. This is the MDS-MAP algorithm [24] with our correlation-based proxy distance as in Algorithm 1.

When the true Euclidean distance $D = (d_{ij})$ is used instead of $\widehat{D} = (\widehat{d}_{il})$ the classical metric MDS algorithm [10] recovers the configuration of focal points $Z = [\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_K]^T$ up to rotation and translation. In this case the input is a $K \times K$ squared distance matrix D with entries d_{ij}, $d_{ij} = (\vec{z}_i - \vec{z}_j)^T(\vec{z}_i - \vec{z}_j)$, and the output the $K \times r$ configuration matrix of focal points Z. We have that [10]

$$-\frac{1}{2} LDL = LZZ^T L\,, \tag{13}$$

where $L = \mathbf{I}_K - \mathbb{1}_K \mathbb{1}_K^T / K$ is a centering matrix, with $\mathbf{I}_K$ the $K \times K$ identity matrix, and $\mathbb{1}_K$ the column vector of all ones. This means that the matrices

Algorithm 1 Reconstruction of focal points in image window

INPUT: $N \times K$ matrix $\widehat{\boldsymbol{G}}$ with columns $\widehat{\boldsymbol{g}}_i$, space dimensions $r = 2, 3$.
OUTPUT: Matrix $\widehat{Z}$ whose column vectors are the estimated coordinates of the grid points $\widehat{\vec{z}}_i$, $i = 1, \ldots, K$.

Compute $G = (V, E)$, with $V = \{1, 2, \ldots, K\}$ and E so that each node is connected to $2r$ neighbors; those corresponding to the 2r-largest values of $|\widehat{\boldsymbol{g}}_i^* \widehat{\boldsymbol{g}}_j|$.
Compute the proxy for distance matrix $\widehat{D}$:

if $(i, j) \in E$ **then**
$\widehat{d}_{ij} = 1$.
else
$\widehat{d}_{ij}$ = shortest path along G
end if

Compute $P = -\frac{1}{2} L \widehat{D} L^T$, where $L = \mathbf{I}_K - \mathbb{1}_K \mathbb{1}_K^T / K$.
Diagonalize P: $P = U \Sigma U^T$.
Compute $\widehat{Z} = U_r \Sigma_r^{1/2}$.

ZZ^T and $-D/2$ are equal when the center of mass of the configuration is moved to zero. For the Euclidean distance matrix D, Algorithm 1 determines the Euclidean coordinates of the focal points.

The rank of the matrix $P = -\frac{1}{2} LDL$ equals the ambient dimension r of the image window when the Euclidian distance matrix D is used in the MDS algorithm. When we use the geodesic distance $\widehat{D}$ on the graph then the rank of P is not equal to r any more. However, the first r singular vectors of P are close to the true coordinates Z (up to centering and rotation) [21]. This is illustrated in Figure 3. The absolute location of the focal points in the image window can be determined using the true location of a few of them, the anchors. These anchors allow us to find the proper rigid transformation and scaling to superimpose the given configuration over them. The anchors can be known a priori or their location can be estimated using coherent interferometric imaging [9]. The number of anchors needed is small, typically $r + 1$.

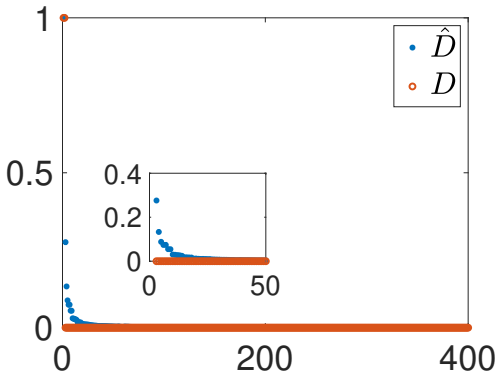

Figure 3: The singular values of the doubly centered distance matrix P normalized by the maximal one are plotted with red circles when D is used (rank is exactly 2) and with blue stars when $\widehat{D}$ is used. There are exactly 2 top singular values in the second case as well, plotted with blue stars, and the lower eigenvalues drop to zero fast but are not immediately zero as with the red circles.

4 Numerical experiments

To simulate wave propagation in random media we use the random travel time model ([9, 19] and references therein) which provides an analytical approximation for the Green's function in (1) in the high-frequency regime in random media with weak fluctuations and large correlation lengths $\ell = 100\lambda$ compared to the central wavelength λ, given by

$$G(\vec{z},\vec{r}) = G_0(\vec{z},\vec{r}) \exp\left[i\sigma\kappa|\vec{z}-\vec{r}| \int_0^1 \mu(\frac{\vec{z}}{l} + \frac{s}{l}(\vec{r}-\vec{z}))\, ds\right]. \qquad (14)$$

Comparing Eqs. (14) and (2) for an homogeneous medium we see that, in this regime, only the phases are perturbed by the random medium while the magnitudes remain unchanged. This model is widely used in adaptive optics as it captures well waveform distortions in heterogeneous media [23]. It does not capture, however, the delay spread (coda) due to multiple scattering and does not provide cross-range diversity.

In our numerical experiments the distance between the array and the image window $L = 100\ell$ is large, so the small distortions produced by each inhomogeneity build up over the propagation distance and are significant at the receivers. The strength of the fluctuations σ is scaled by the dimensionless parameter $\lambda/\sqrt{lL}$, for which the standard deviation of the random phase

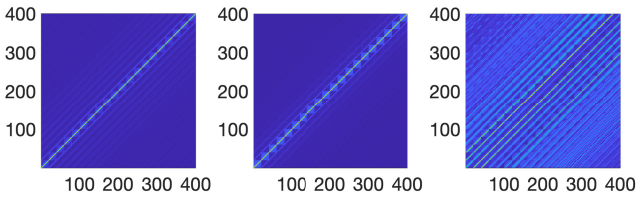

Figure 4: Cross-correlations of the sensing matrix. Left: cross-correlations of the sensing matrix in the random medium. Center: cross-correlations of the sensing matrix in the homogeneous medium. Right cross-correlations between the sensing matrix in the random and homogeneous media. Strength of the fluctuations of the random medium $\tilde{\sigma} = 0.8$.

fluctuations in the Green's function is $\mathcal{O}(1)$. The strength of the fluctuations $\tilde{\sigma} = \sigma/(\lambda/\sqrt{lL})$ in the simulations is $\tilde{\sigma} = 0.6$ or $\tilde{\sigma} = 0.8$.

The challenge is to obtain statistically stable results in these media. Some previously proposed methods, such as coherent interferometry, provide stable images but at the expense of resolution, causing some blurring [9]. Our approach produces stable images whose resolution is comparable to that of a homogeneous medium.

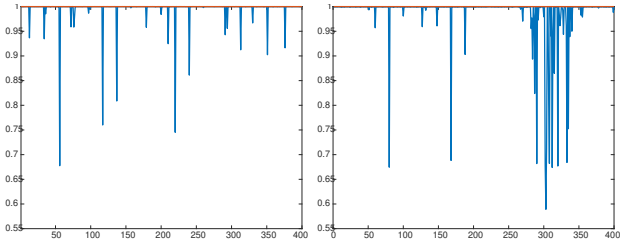

Figure 5: Maximum correlation between each estimated column and the true ones in the sensing matrix, as in (15). In red the results when the columns of the matrix are more incoherent ($a = 48\ell$). In blue the results when the columns of the matrix are more coherent ($a = 24\ell$). Sparsity $s = 4$ on the left and $s = 8$ on the right.

We consider the following setup for our numerical simulations. In the first step of dictionary learning for the sensing matrix $\boldsymbol{\mathcal{G}}$ we use the multi-frequency data recorded with a large array aperture $a = 48\ell$ with $N_r =$

145 equally spaced receivers (see Fig. 2). In the second step for the grid reconstruction, we use the data corresponding to half the array aperture, so $a = 24\ell$. The bandwidth $[0.5f, f]$, with $f = c_0/\lambda$, is discretized with $N_f = 10$ equally spaced frequencies, and is the same for both steps of the algorithm. We organize the multiple frequency data column-wise, so

$$\mathbf{Y} = [\mathbf{Y}(f_1)^\mathsf{T}, \mathbf{Y}(f_2)^\mathsf{T}, \ldots, \mathbf{Y}(f_{N_f})^\mathsf{T}]^\mathsf{T}.$$

The multi-frequency sensing matrix is now

$$\boldsymbol{g}_i = [\boldsymbol{g}(\vec{z}_i, f_1)^\mathsf{T}, \boldsymbol{g}(\vec{z}_i, f_2)^\mathsf{T}, \ldots, \boldsymbol{g}(\vec{z}_i, f_{N_f})^\mathsf{T}]^\mathsf{T},$$

$i = 1, \ldots, K$. Thus, the sensing matrix $\boldsymbol{\mathcal{G}} = [\boldsymbol{g}_1 \cdots \boldsymbol{g}_K]$ has dimensions $N \times K$ with $N = N_r \cdot N_f$. The sampling of the 20×20 points in the image window is based on the homogeneous medium array resolution, $O(\lambda L/a)$ in cross-range and $O(c_0/B)$ in range [11, 7]. This means that we implicitly assume that the sources are far part in the sense that the random Green's function vectors corresponding to this discretization are non-coherent, *i.e.,* the coherence ν (7) is small.

In Figure 4, we assume that the sensing matrices corresponding to a random medium $\boldsymbol{\mathcal{G}}$ and to the homogeneous medium $\boldsymbol{\mathcal{G}}_0$ are known, and we show the cross-correlation matrices of $\boldsymbol{\mathcal{G}}^*\boldsymbol{\mathcal{G}}$ (left), $\boldsymbol{\mathcal{G}}_0^*\boldsymbol{\mathcal{G}}_0$ (center), and $\boldsymbol{\mathcal{G}}^*\boldsymbol{\mathcal{G}}_0$ (right). For the homogeneous medium the Green's function used is given by (2). Each row i in these images corresponds to a time reversal experiment where a source located at $\vec{z}_i$ emits a pulse, and the recorded signals are time reversed and emitted back into the medium. When the waves are re-emitted into the same medium in which the measurements were obtained, as in the left and center images of Figure 4, they retrace the original scattering process and arrive back approximately at the point at which they were emitted, that is, the focal point. However, when the back-propagation is done in a different medium, as in the right image of this figure, there is no re-focusing. In Figure 4, the large values (lighter blue color) correspond to re-focusing points. This figure shows that (a) time reversal of waves into random and homogeneous media are similar, and (b) that we cannot use the homogeneous medium to recover this structure if there is scattering.

In our numerical experiments, we assume that we have a diverse set of data $\mathbf{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_M]$, with $\boldsymbol{y}_i = \boldsymbol{\mathcal{G}}\,\boldsymbol{x}_i$. Both the sensing matrix $\boldsymbol{\mathcal{G}} \in C^{N \times K}$ and the sparse vectors $\boldsymbol{x}_i \in \mathbb{C}^K$ are unknown. We assume data corresponding to a large number of experiments, so $M \gg K$. Given this set of data, we want to recover the columns of the sensing matrix $\boldsymbol{\mathcal{G}} = [\boldsymbol{g}_1 \cdots \boldsymbol{g}_K]$, whose rank is approximately $200 < K = 400$ and whose

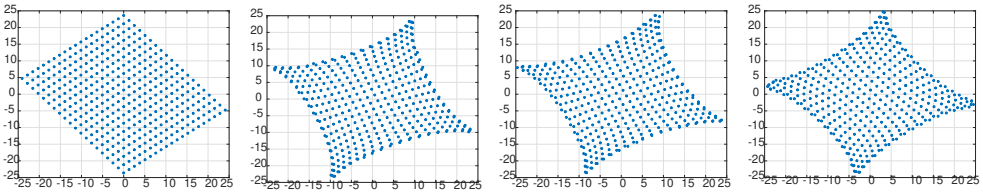

Figure 6: From left to right: Grid reconstruction from true Euclidean distances using the MDS algorithm when all the pairwise distances are assumed known; from true Euclidean distances when only distances corresponding to the four nearest neighbors are assumed known; using the MDS-MAP algorithm with geodesic graph distances for $\tilde{\sigma} = 0.6$; and using the MDS-MAP algorithm with geodesic graph distances for $\tilde{\sigma} = 0.8$. Sparsity $s = 8$ in all cases.

coherence is $\nu = 0.7$. The sensing matrix is rank-deficient because the resolution of the image window is high, with pixel sizes $\lambda L/a$ in cross-range and c_0/B in range.

The results of the first step of the proposed strategy are depicted in Figure 5 for large (red lines) and small (blue lines) arrays. We solve problem (8) as described in Section 33.1 and measure the success of this first step as follows. We form, for every (normalized) recovered column $\widehat{\boldsymbol{g}}_i$, the cross-correlations with all the columns of the true sensing matrix $\boldsymbol{\mathcal{G}}$ and represent in Figure 5 the maximum value

$$C_{max}(i) = \max_j |\widehat{\boldsymbol{g}}_i^T \boldsymbol{g}_j| \tag{15}$$

for a sparsity level $s = 4$ (left) and $s = 8$ (right). We observe values very close to 1 in both cases when the columns of the sensing matrices are for large array apertures (red lines). This means that the true Green's function vectors are recovered when large arrays are used because they are incoherent. However, when smaller arrays are used (blue lines) the Green's function vectors are coherent and some of them are not recovered. It is important to recover accurately all, or almost all, Green's function vectors because, otherwise, we cannot establish their connectivity properly and, therefore, we cannot reconstruct the grid in the image window in the second step.

Figure 6 shows the results for the grid reconstructions accomplished with Algorithm 1 described in Section 33.2 using $k = 4$ neighbors. This algorithm provides the correspondence between the Green's function vectors found in the first step and their focal points in the image window. From left to right we show the results when (left) all the pairwise Euclidean distances

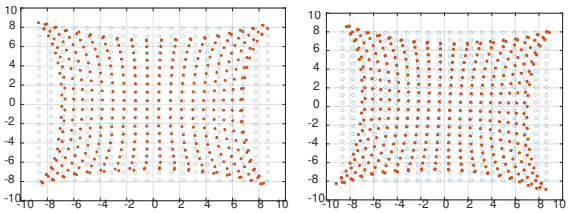

Figure 7: Using three points as anchors, *i.e.* assuming the location of those three points is known, we can estimate the scaling and the rotation needed to recover the absolute grid positions. We compare the recovered locations (red stars) with the true ones (blue circles) where $\tilde{\sigma} = 0.6$ (left) and $\tilde{\sigma} = 0.8$ (right). Sparsity $s = 8$ in both cases.

between the focal points are known in a homogeneous medium, (second from the left) when only Euclidean distances between the four nearest neighbors are known in a homogeneous medium, (second from the right) using only connectivity information in a random medium with $\tilde{\sigma} = 0.6$, and (right) using only connectivity information in a random medium with $\tilde{\sigma} = 0.8$. In all the cases, the sparsity is $s = 8$.

Algorithm 1 provides grid positions up to a rigid transformation and scaling. In Figure 7, we post-process the results shown in the second from the right and right images in Figure 6 to transform them to absolute positions using three anchors. We observe that the grids are quite well reconstructed near the center but bent towards the edges. This occurs because our geodesic graph distance is the scaled l_1 distance on the grid, and an embedding of such distances into Euclidean spaces leads to such distortions. Naturally, there is no grid deformation shown in the left image of Figure 6 since Euclidean distances are used for all focal points.

After the two steps of the proposed strategy, we recover the ordered sensing matrix $\widehat{\boldsymbol{\mathcal{G}}}$, so we can image any signal measured at the array into the image window. In Figure 8, we back-propagate a signal $\boldsymbol{g}(\vec{z}_j)$ from a source located at $\vec{z}_j$ using the recovered Green's function vectors. Thus, the image formed at points $\vec{z}_i,\ i = 1, \ldots, K$, is

$$\mathcal{I}(\vec{z}_i; \vec{z}_j) = |\widehat{\boldsymbol{g}}(\vec{z}_i)^* \boldsymbol{g}(\vec{z}_j)| \,. \tag{16}$$

As before, the hat in (16) denotes the recovered Green's function vectors of the sensing matrix using the two step method introduced here. As illustrated in Figure 8, the produced image (right) is similar to the one obtained using the true Green's functions (left) and significantly better than the one obtained using the homogeneous Green's function (center). This figure shows

the need for recovering accurate estimates of the Green's function vectors for imaging in random media since the ones corresponding to a reference homogeneous medium provide very noisy, useless images with (16). The top and bottom rows are images obtained in different realizations of the random media with $\tilde{\sigma} = 0.6$ and $\tilde{\sigma} = 0.8$, respectively.

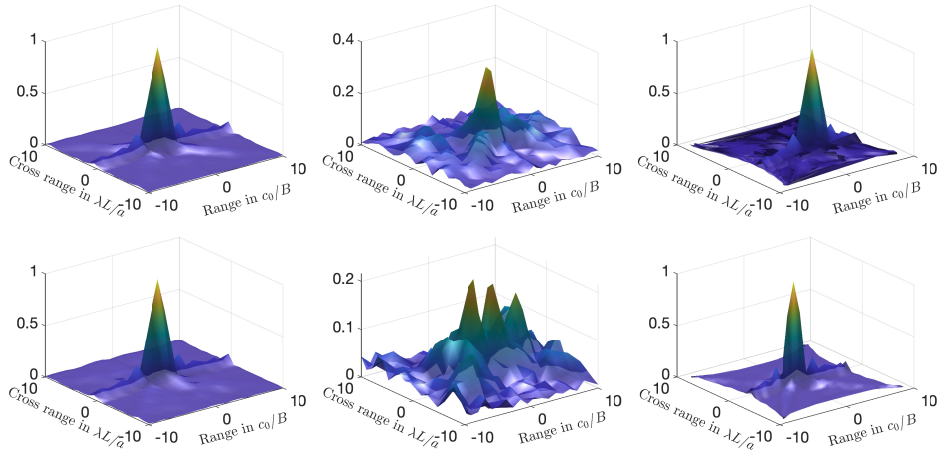

Figure 8: From left to right, image formed with (16) using the true random Green's functions, the homogeneous Green's functions and the recovered ones with the proposed method. Here the sparsity is $s = 8$. The strength of the fluctuations is $\tilde{\sigma} = 0.6$ for the top row and $\tilde{\sigma} = 0.8$ for the bottom row.

5 Discussion

We propose here a data-driven approach for imaging in random media. The data are multiple signals recorded by an array of receivers and coming from many sets of sparse sources whose location and strengths are unknown. Imaging is done with a two step method. In the first step, we show that dictionary learning algorithms can estimate the random Green's function vectors as columns of an unordered matrix. Ordering is not an issue in other applications of dictionary learning but it is essential for imaging as we need to know the correspondence between them and the grid points in the image window. This is done in the second step using a multidimensional scaling algorithm that uses local information at each grid point to compute the global layout of the grid points. A key element of the proposed approach is that the local information is obtained from the cross-correlations of the

estimated Green's function vectors. These cross-correlations determine the connectivity between the grid points.

The numerical experiments show that using these two basic algorithms, dictionary learning and multidimensional scaling, we can form stable images in random media without loss of resolution. This demonstration of principle can be extended to imaging in more complex ambient media but this may require additional steps in the imaging method.

6 Materials and methods

All the imaging data used in the numerical experiments are generated numerically as described in detail in Section 4. The implementation of the two-step imaging method that we are proposing is also described in detail in this section.

7 Acknowledgments

Miguel Moscoso's work was supported by the Spanish AEI grant PID2020-115088RB-I00. Alexei Novikov's work was partially supported by AFOSR FA9550-23-1-0352 and FA9550-23-1-0523. The work of George Papanicolaou was partially supported by AFOSR FA9550-23-1-0352. The work of Chrysoula Tsogka was partially supported by AFOSR FA9550-23-1-0352 and FA9550-21-1-0196.

References

[1] A. Agarwal, A. Anandkumar, P. Jain, and P. Netrapalli, *Learning sparsely used overcomplete dictionaries via alternating minimization,* SIAM Journal on Optimization **26**, 2775–2799 (2016).

[2] A. Agarwal, A. Anandkumar, and P. Netrapalli, *A clustering approach to learning sparsely used overcomplete dictionaries,* IEEE Transactions on Information Theory **63**, 575–592 (2017).

[3] M. E. M. Aharon, M. Elad and A. Bruckstein, *K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,* IEEE Transactions on Signal Processing **54**, 4311–4322 (2006).

[4] J. Aspnes, T. Eren, D.K Goldenberg, W. Whiteley, Y.R. Yang, and B.D. O. Anderson, A Theory of Network Localization, Mobile Computing, IEEE Transactions **5**, 1663–1678 (2007).

[5] A. Bakulin and R. Calvert, *The virtual source method: Theory and case study,* Geophysics, **71**, SI139-SI150 (2006).

[6] A. Beck and M. Teboulle, *A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems,* SIAM J. Img. Sci. **2**, 183–202 (2009),

[7] L. Borcea, G. Papanicolaou, C. Tsogka, and James Berryman, *Imaging and time reversal in random media*, Inverse Problems **18**, 1247-1279 (2002).

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

[9] L. Borcea, J. Garnier, G. Papanicolaou and C. Tsogka, *Enhanced Statistical Stability in Coherent Interferometric Imaging*, Inverse Problems **27**, 085004 (2011)

[10] I. Borg and P. Groenen, *Modern Multidimensional Scaling: Theory and Applications*, Springer Series in Statistics, (2005)

[11] M. Born and E. Wolf, *Principles of Optics*, Academic Press, (1970)

[12] G. Davis, S. Mallat, and M. Avellaneda, *Adaptive greedy approximations,* Journal of Constructive Approximation **13**, 57–98 (1997).

[13] I. Daubechies, M. Defrise, and C. De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint,* Comm. Pure Appl. Math. **57**, 1413–1457 (2004).

[14] K. Engan, S.O. Aase and J. Hakon Husoy, *Method of optimal directions for frame design,* in 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258), **5**, 2443–2446 (1999).

[15] M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, M. Tanter, J.-L. Thomas and F. Wu, *Time-reversed acoustics*, Rep. Prog. Phys. **63**, 1933-1995 (2000).

[16] J. Garnier and G. Papanicolaou, *Passive Imaging with Ambient Noise*, Cambridge University Press, ISBN 978-1-107-13563-5, New York, NY, United States, 2016.

[17] K. Kreutz-Delgado, J.F.Murray, D.B. Rao, K. Engan, T.W. Lee, T.J. Sejnowski, *Dictionary learning algorithms for sparse representation,* Neural Comput. **15** , 349–396 (2003).

[18] M. Moscoso, A. Novikov, G. Papanicolaou, and L. Ryzhik, *A differential equations approach to l1-minimization with applications to array imaging,* Inverse Problems **28**, 105001 (2012).

[19] M. Moscoso, A. Novikov, G. Papanicolaou, and C. Tsogka, *Multifrequency Interferometric Imaging with Intensity-Only Measurements,* SIAM J. Img. Sci. **10**, 1005–1032 (2017).

[20] A. Novikov, S. White, *Spectral subspace dictionary learning,* Proceedings of Machine Learning Research, 34th International Conference on Algorithmic Learning Theory, 1–36 (2023).

[21] S. Oh, A. Montanari and A. Karbasi, *Sensor network localization from local connectivity: Performance analysis for the MDS-MAP algorithm,* 2010 IEEE Information Theory Workshop on Information Theory (ITW 2010, Cairo), Cairo, Egypt, pp. 1–5 (2010).

[22] M. R. Osborne, B. Presnell, and B. A. Turlach, *On the lasso and its dual,* Journal of Computational and Graphical Statistics *9*, 319–337 (2000).

[23] S.M. Rytov, Y.A. Kravtsov, and V.I. Tatarskii, Principles of statistical radiophysics. 4. Wave Propagation through random media, Springer Verlag, Berlin, 1989.

[24] Y. Shang, W. Ruml, Y. Zhang, and M. P. J. Fromherz, *Localization from mere connectivity,* in MobiHoc '03: Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing. New York, NY, USA: ACM, pp. 201–212 (2003).

[25] R.G. Stansfield, Statistical theory of d.f. fixing, Journal of the Institution of Electrical Engineers **94**, 762–770 (1947).

[26] P. Wu, S. Su, Z. Zuo, X. Guo, B. Sun and X. Wen, Time Difference of Arrival (TDoA) Localization Combining Weighted Least Squares and Firefly Algorithm, Sensors **19**, 2554 (2019).

[27] A. Yeredor and E. Angel, Joint TDOA and FDOA Estimation: A Conditional Bound and Its Use for Optimally Weighted Localization,IEEE Transactions on Signal Processing **59**, 1612–1623 (2011).