Imaging in disordered media

George Papanicolaou
Stanford University

Waves and Imaging in Complex Media

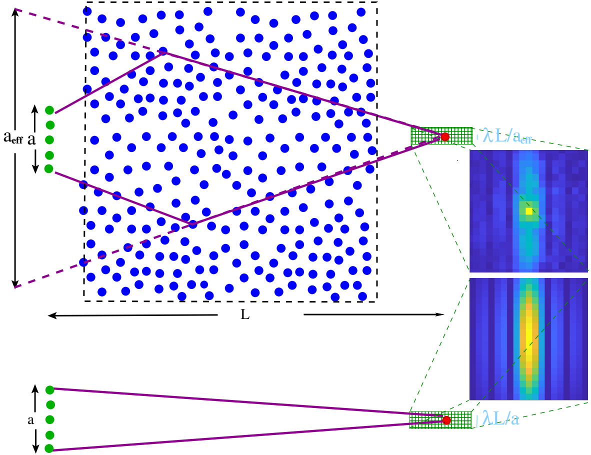
June 11, 2025

with A. Christie, M. Leibovich, M. Moscoso, A. Novikov, C. Tsogka and J. Zheng.

Outline:

Imaging methodology adapted to the availability of an abundance of data

1. When large and diverse data sets are available, such as satellite or conventional SAR surveillance data, imaging methodology changes. For example, data volume is essential for training NNs. But what exactly do we gain with the data-intensive methodology, with or without NN? A lot in fact.
2. We do not only image (statistically speaking: estimate) the location and reflectivity of the object(s) we want. We can and should estimate/assess the ambient medium (as needed) as well.
3. **Main result:** In strongly inhomogeneous media we can **image** with resolution that is better than that of the same imaging system in a homogeneous medium (super-resolution). **Applications**: Imaging through foliage, dusty environments, atmospheric inhomogeneities.
4. Identify advantages and disadvantages in using NN vs more conventional optimization methods. NN gain robustness and relative simplicity but lose some control of accuracy. Overall NN seem to have an edge, but without much of a theory for now.
5. Concurrent issues: Imaging when multiple random media are involved. Application: surveillance SAR.



Imaging schematic and illustration of the effective aperture.[1]

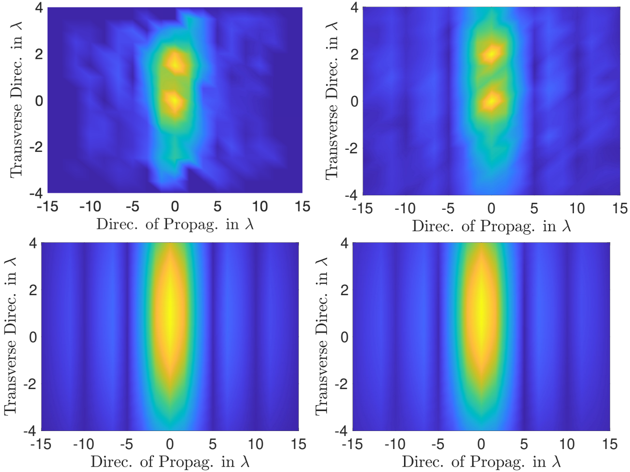
[1] Super-resolution in disordered media using neural networks; https://arxiv.org/pdf/2410.21556

Continued: The simulation setup in the microwave regime

- Array size $a = 30\lambda$ and the bandwidth is 1GHz, with a 5GHz central frequency, C-Band (wavelength $\lambda = 6cm$).
- In the simulations, the cross-range resolution in homogeneous media is $\lambda L/a = 8\lambda$, where a is the physical array aperture and $L = 240\lambda$ is the range.
- In the scattering medium (use the Foldy-Lax equations[2]) the resolution is about 2λ which translates to an effective aperture that is four times the physical aperture of the array, *i.e.,* $a_{eff} \approx 4a = 120\lambda$.
- The range resolution is c_0/B where c_0 is the background propagation speed and B is the bandwidth.
- In this numerical simulation (Foldy-Lax), the range resolution in the scattering medium does not change.

[2] P.-D. Letourneau, Y. Wu, G. Papanicolaou, J. Garnier, and E. Darve, A numerical study of super-resolution through fast 3D wideband algorithm for scattering in highly-heterogeneous media, Wave Motion, 70, 113-134 (2017)

The results: Super-resolution



TL: Physical time reversal; TR: The new imaging method;
BL: Homogeneous medium; BR: Kirchhoff migration

Continued: Super-resolution vs. time reversal

- **Top left**: physical time reversal focusing on the known grid using the true Green's functions.
- **Top right**: the proposed imaging algorithm, migration image using the estimated dictionary elements on the reconstructed grid from the ordered sensing matrix. We can see that we achieve similar resolution in both cases, which is able to resolve to closely located sources.
- **Bottom left**: migration image in a homogeneous medium.
- **Bottom right**: Homogeneous medium Green's function migration applied to random medium data. The resolution is inferior to that obtained by the estimated sensing matrix.

The imaging method without NN

The imaging problem comes down to solving a very large linear system[3]

$$\mathbf{y} = \mathcal{G}\,\mathbf{x}\,, \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^N$ is the recorded array data vector. We consider $\mathbf{x}$ as the image, that is, a vector whose k-th component represents the complex amplitude of the source at location $\vec{z}_k$ in the image window, $k = 1, \ldots, K$, and $\mathcal{G}$ is the $N \times K$ sensing matrix, that is, the matrix whose columns are the Green's functions for wave propagation in the medium between the array and the image window. Typically $K \gg N$.

New Imaging Problem: Given $\mathbf{y}_i = \mathcal{G}\,\mathbf{x}_i$ for $i = 1, 2, \ldots, M$ data, estimate, in addition, the sensing matrix $\mathcal{G}$. Here $M \gg K$.

Once $\mathcal{G}$ has been estimated imaging with any data vector $\mathbf{y}$ is routine inversion of a linear system for the image vector $\mathbf{x}$, which is "conventional imaging".

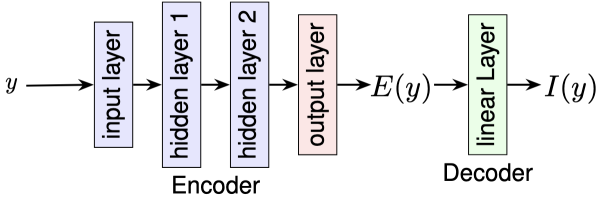
Essential assumption: The image vector $\mathbf{x}$ is sparse.

[3] We sssume single frequency in this presentation for simplicity but multiple frequencies (broadband) are used in theory and simulations

How to estimate the sensing matrix $\mathcal{G}$

- Step 1. Estimate the columns of columns of $\mathcal{G}$ as an **unordered** set (dictionary learning) either with an encoder-decoder NN or with alternating optimization (MOD): Minimize over $\hat{\mathcal{G}}$ ($N \times K$) and $\mathbf{X}$ ($K \times M$) the error $\|\hat{\mathcal{G}}\mathbf{X} - \mathbf{Y}\|_F^2$ where $\mathbf{Y}$ is the $N \times M$ matrix of data vectors, the expected sparsity level of each column of $\mathbf{X}$ is fixed, and F denotes the Frobenius norm. This **requires** a good initialization as it is non-convex! It is a stable (because of sparsity) form of blind deconvolution. With NN we do **not** need an initialization but we need multiple versions of the output $\hat{\mathcal{G}}$ followed by a **clustering** method to improve accuracy. The NN approach is simpler and more robust but may be less accurate. It can also be fully replicated by the DL algorithm, with random initialization and followed by clustering.
- Step 2. **Order** the columns of the estimated sensing matrix $\hat{\mathcal{G}}$. That is, identify their source point in the image window.

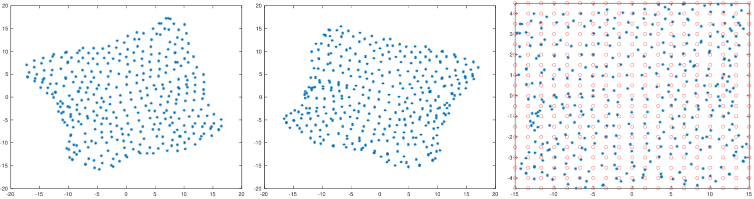
The Neural Network



Neural network architecture.
We use a two hidden layer network for the encoder. The output of the encoder $E(y)$ is fed into the linear decoder network with output $I(y)$. Modulus thresholding activation is used in the intermediate layers and LeakyReLU activation is used for the other layers. Before activation functions are applied, the input is standardized using batch normalization. We train the networks to minimize the loss function
$\mathcal{L} = \|I(y) - y\|_2^2 + \mu \| E(y)\|_1$.
Note the sparsity penalty in the loss function, which is essential.

Ordering the columns of the estimated sensing matrix $\hat{\mathcal{G}}$

It amounts to reconstructing the grid in the image window using MultiDimensional Scaling (MDS)



Left: Grid reconstructed with the true Green's functions using MDS;
Middle: Grid obtained from the estimated Green's functions using MDS;
Right: Grid obtained from MDS after rotation and scaling assuming the location of 3 "anchor" points is known, superimposed over true grid, plotted with red circles.

Ordering the columns of $\hat{\mathcal{G}}$ continued

- This is Step 2: We want to find the focal spots in the image window from the *estimated* columns of the sensing matrix $\{\hat{g}_i\}_{i=1}^K$.
- We use MDS. **What is it?** In its simplest form it is a least squares algorithm that finds the location of points $\{z_i\}_{i=1}^K$ in the image window when their pairwise distance $\|z_i - z_j\|$ is given, and this is up to an overall translation, rotation and scaling. Most important application to date (20-25 years old): **sensor location**, but with a metric that is not Euclidean.
- We use MDS with a **proxy metric** that is formed using the inner products of the estimated columns $\{\hat{g}_i\}_{i=1}^K$.
- From these inner products, which contain all possible imaging information, we do a proxy MDS and find the ordering.
- MDS can also be done with NN. But when we have a rectangular grid in the image window there is no gain. If, however, the image window has an irregular/unknown gridding then the NN improves performance by some 10-20 percent (this is a separate work). **Application:** SAR with an irregular surveillance pattern that is not known accurately.

Proxy MDS

- We can use MDS in a non-metric context where we do not have Euclidean pairwise distances. Instead, we associate with each column of the sensing matrix the vertex of a graph, From the inner products that are close to one (ranked inner products) we form neighborhoods and link them with edges. The pairwise distance between columns is then their geodesic graph distance (minimum number of edges needed to connect two vertices in a simply connected graph).
- This works quite well in the Foldy-Lax array imaging data **provided** the pixelization of the image window is calibrated correctly. This is how "classical" imaging resolution connects with MDS methodology, and we need some minimum coherence for the DL estimated columns.
- Important trade-off: DL needs incoherence but MDS needs some coherence.

Multiple random media

- If we have full-diversity data for all the different random media then **pooled** DL works well and we end up with a large sensing matrix (unordered at first) which is the concatenation of the sensing matrices of each random medium. MDS and clustering works fine in this context.
- If we do not have full data sets but the random media do not change much then the concatenation will work for DL. MDS works as well.
- The incremental multiple random media case is important in many surveillance applications.

Concluding remarks

- When we have large and diverse data sets we can image with exceptional resolution and stability
- The methodology is very different from previous ones. We now set out to estimate "everything", even "noise"
- Going forward: Multiple random media, when the ambient medium changes (weakly or strongly). Recent results tell us that this can actually lead to better resolution by some 10-20 percent by improving the MDS step, when the ambient medium changes slowly. In preparation: Surveillance SAR simulations, including satellite LEO SAR.