Imaging with intensity-only data

Miguel Moscoso
Universidad Carlos III de Madrid

With Anwei Chai and George Papanicolaou
Overview

- Array imaging with full data
  - Formulation
  - l1 theory
- Imaging with intensity-only data
  - Reformulation
  - Minimum rank problem
- Numerical experiments
- Conclusions
- Papers:
  - Array imaging using intensity-only measurements, IP 27 (2011) 015005.
  - Robust imaging of localized scatterers using the singular value decomposition and l1 minimization, submitted to IP.
In array imaging we want to determine the location and reflectivities of small or distributed reflectors by sending probing signals from the array and recording the backscattered fields.
In array imaging we want to determine the location and reflectivities of small or distributed reflectors by sending probing signals from the array and recording the backscattered fields.

- Narrow / Broad-band
- Homogeneous/Inhomogeneous
- Coherent/Incoherent
- Single/Multiple illum
- Acquisition geometries
- Full data/Intensity-only
- ...

Setup of array imaging
Array imaging

Setup of array imaging

- Point-like scatterers
- Homogeneous medium
- **Coherent imaging**
- Well separated
- Narrow band
- Single illumination
Array imaging

- K pixels in IW
- M scatterers
- N transducers (d units apart)

- $M < N \ll K$
- $\lambda \ll |y_j - y_i| \ll L$
- $a \equiv N \, d \sim L$
Measurements (Full data)

\[ b_r(\omega) = \sum_{j=1}^{M} \rho_j \hat{G}_0(0, y_n, \omega) \hat{G}_0(y_n, x_s, \omega) \]

\[ \hat{G}_0(x, y, \omega) = \frac{\exp(ik|x-y|)}{4\pi|x-y|}, \kappa = \omega/c \]

- \( K \) pixels in IW
- \( M \) scatterers
- \( N \) transducers (d units apart)
- \( M < N \ll K \)
- \( \lambda \ll |y_j - y_i| \ll L \)
- \( \alpha \equiv N d \sim L \)
The coherent inverse problem

\[ b_r(\omega) = \sum_{j=1}^{M} \rho_j \hat{G}_0(x_r, y_n, \omega) \hat{G}_0(y_n, x_s, \omega) \]

\[ \hat{G}_0(x, y, \omega) = \frac{\exp(i\kappa|x - y|)}{4\pi|x - y|}, \kappa = \omega / c \]

Setup of array imaging

- Reflectivity vector \( \mathbf{\rho} = (\rho_1, \rho_2, \rho_3, \ldots, \rho_K)^T \) SPARSE!
- Data vector \( \mathbf{b}(\omega) = (b_1, b_2, b_3, \ldots, b_N)^T \)
- Illumination vector \( \hat{\mathbf{g}}(y_j, \omega) = (\hat{G}_0(x_1, y_j, \omega), \hat{G}_0(x_2, y_j, \omega), \ldots, \hat{G}_0(x_N, y_j, \omega))^T \)
The coherent inverse problem

\[ A_\omega \mathbf{\rho} = \mathbf{b}(\omega) \]

\[
\begin{pmatrix}
\hat{G}_0(y_1, x_s, \omega) \hat{g}(y_1, \omega) \\
\vdots \\
\hat{G}_0(y_K, x_s, \omega) \hat{g}(y_K, \omega)
\end{pmatrix} = 
\begin{pmatrix}
\rho_1 \\
\rho_2 \\
\rho_3 \\
\vdots \\
\rho_K
\end{pmatrix} = 
\begin{pmatrix}
b_1 \\
\vdots \\
b_N
\end{pmatrix}
\]
The linear system

\[
A_\omega \rho = b(\omega)
\]

Array imaging is to solve for \( \rho \)!

Setup of array imaging

\[
\begin{pmatrix}
\hat{G}_0(y_1, x_s, \omega) \hat{g}(y_1, \omega) & \cdots & \hat{G}_0(y_K, x_s, \omega) \hat{g}(y_K, \omega)
\end{pmatrix}
\begin{pmatrix}
\rho_1 \\
\rho_2 \\
\vdots \\
\rho_K
\end{pmatrix}
= 
\begin{pmatrix}
b_1 \\
b_2 \\
\vdots \\
b_N
\end{pmatrix}
\]
The linear system

\[
A_\omega \rho = b(\omega)
\]

Array imaging is to solve for \( \rho \)!

- \( N<<K \Rightarrow \) underdetermined
- but, ...

\[
\begin{pmatrix}
\tilde{G}_0(y_1, x_s, \omega) \tilde{g}(y_1, \omega) \\
\vdots \\
\tilde{G}_0(y_2, x_s, \omega) \tilde{g}(y_2, \omega) \\
\vdots \\
\tilde{G}_0(y_K, x_s, \omega) \tilde{g}(y_K, \omega)
\end{pmatrix}
\begin{pmatrix}
\rho_1 \\
\rho_2 \\
\rho_3 \\
\vdots \\
\rho_K
\end{pmatrix}
= 
\begin{pmatrix}
b_1 \\
\vdots \\
b_N
\end{pmatrix}
\]
The linear system

\[ A_\omega \rho = b(\omega) \]

Array imaging is to solve for \( \rho \)!

- \( N < K \Rightarrow \) underdetermined
- but, the solution is sparse!!

Setup of array imaging

\[
\begin{pmatrix}
\hat{g}_0(y_1, x_s, \omega) \hat{g}(y_1, \omega) \\
\vdots \\
\hat{g}_0(y_K, x_s, \omega) \hat{g}(y_K, \omega)
\end{pmatrix}
\begin{pmatrix}
\hat{G}_0(y_1, x_s, \omega) \\
\vdots \\
\hat{G}_0(y_K, x_s, \omega)
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}
= 
\begin{pmatrix}
b_1 \\
\vdots \\
b_N
\end{pmatrix}
\]
Imaging methods

\[
A_\omega \rho = b(\omega)
\]

\[A \in \mathbb{C}^{N \times K} \text{ is fat (} N << K \text{)}\]

- **l2-norm minimization.** Since \( N << K \), we have fewer constraints than unknowns, and therefore there are infinite solutions. We can pick one by finding the ‘smallest’ one.

\[
\min \| \rho \|_2 \quad \text{subject to} \quad A_\omega \rho = b(\omega)
\]

The solution to this problem is given by

\[
\rho_{l2} = A_\omega^* (A_\omega A_\omega^*)^{-1} b(\omega)
\]
Imaging methods

\[
A_\omega \rho = b(\omega) \quad A \in \mathbb{C}^{N \times K} \text{ is fat (} N << K \text{)}
\]

- **l2-norm minimization.** Since \( N << K \), we have fewer constraints than unknowns, and therefore there are infinite solutions. We can pick one by finding the smallest one.

\[
\min \| \rho \|_2 \quad \text{subject to} \quad A_\omega \rho = b(\omega)
\]

The solution to this problem is given by

\[
\rho_{l2} = A_\omega^* (A_\omega A_\omega^*)^{-1} b(\omega)
\]

- **Kirchhoff migration.** For large scale matrices, \((A_\omega A_\omega^*)^{-1}\) is computationally expensive, and often ill-conditioned! Hence it is often replace by the identity.

\[
\rho_{KM} = A_\omega^* b(\omega)
\]
Imaging methods

\[ A_\omega \rho = b(\omega) \]

\[ A \in \mathbb{C}^{N \times K} \text{ is fat } (N \ll K) \]

- **l2-norm minimization.** When the data is **contaminated by noise** we need to relax the constraint

\[
\min \| \rho \|_2 \quad \text{subject to} \quad \| A_\omega \rho - b(\omega) \| < \varepsilon
\]

The solution to this problem is given by

\[
\rho_{l2}(\tau) = (A_\omega^* A_\omega + \tau I)^{-1} A_\omega^* b(\omega)
\]

Related to the **Tikhonov regularized** method for ill-posed problems that gives preference to a particular solution including a regularization term

\[
\rho_{l2}(\tau) = \arg\min_{\rho} \left\{ \frac{\tau}{2} \| \rho \|^2 + \frac{1}{2} \| A_\omega \rho - b(\omega) \|^2 \right\}
\]
Imaging methods

- Both l2-nom minimization and KM are usually applied to broadband imaging, where they perform quite well when the medium is homogeneous or weakly scattering, leading to range resolutions proportional to the reciprocal of the bandwidth.

- Cross range resolution is proportional to $\lambda L/a$, for both broad and narrow band imaging (assuming $\lambda \ll a \ll L$).

- Range resolution:
  - For narrowband is $\lambda L^2/a^2$
  - For broadband is $c/B$ ($c$ is the speed of propagation and $B$ is the bandwidth)

- Can we devise an imaging method whose resolution is dictated by the sparsity of the image instead of the signal’s bandwidth?
Imaging methods

\[
A_\omega \rho = b(\omega)
\]

\(A \in \mathbb{C}^{N \times K}\) is fat \((N << K)\)

- l1-norm minimization. Among the infinite solutions we seek the one that is sparsest. This changes the problem substantially because we exploit the sparsity of the solution! The sparsest solution is given by

\[
\min \|\rho\|_0 \quad subject \ to \quad A_\omega \rho = b(\omega) \tag{10}
\]
Imaging methods

\[ A_\omega \rho = b(\omega) \quad A \in \mathbb{C}^{N \times K} \text{ is fat } (N \ll K) \]

- \textbf{l1-norm minimization.} Among the infinite solutions we seek the one that is sparsest. This changes the problem substantially because we exploit the sparsity of the solution! The sparsest solution is given by

\[
\min \| \rho \|_{l_0} \quad \text{subject to} \quad A_\omega \rho = b(\omega) \quad (l0)
\]

However this problem is NP-hard. Hence we use the convex relaxation

\[
\min \| \rho \|_{l_1} \quad \text{subject to} \quad A_\omega \rho = b(\omega) \quad (l1)
\]

Compressed Sensing (Donoho, Candès and Tao) gives conditions under which (l0) and (l1) have the same unique solution*.

Imaging methods

\[
A_\omega \rho = b(\omega) \quad A \in \mathbb{C}^{N \times K} \text{ is fat (} N << K \text{)}
\]

- **l1-norm minimization.** Among the infinite solutions we seek the one that is sparsest. This changes the problem substantially because we exploit the sparsity of the solution! The sparsest solution is given by

\[
\min \| \rho \|_0 \text{ subject to } A_\omega \rho = b(\omega) \quad (l0)
\]

However this problem is NP-hard. Hence we use the convex relaxation

\[
\min \| \rho \|_1 \text{ subject to } A_\omega \rho = b(\omega) \quad (l1)
\]

**Definition:** Restricted isometry property (RIP)

\[
(1 - \delta_s) \| \rho \|_2 \leq \| A_\omega \rho \|_2 \leq (1 + \delta_s) \| \rho \|_2
\]
Imaging methods

\[ A_\omega \rho = b(\omega) \]

\[ A \in \mathbb{C}^{N \times K} \text{ is fat (} N << K \)\]

- **l1-norm minimization.** Among the infinite solutions we seek the one that is sparsest. This changes the problem substantially because we exploit the sparsity of the solution! The sparsest solution is given by

\[
\min \| \rho \|_{l_0} \text{ subject to } A_\omega \rho = b(\omega) \quad (l0)
\]

However this problem is NP-hard. Hence we use the convex relaxation

\[
\min \| \rho \|_{l_1} \text{ subject to } A_\omega \rho = b(\omega) \quad (l1)
\]

(l0) and (l1) problems are equivalent in the sense:

- if \( \delta_{2s} < 1 \), the (l0) problem has a unique s-sparse solution
- if \( \delta_{2s} < \sqrt{2} - 1 \), the solution to (l1) problem is that of the (l0)
Key property in array imaging for RIP

\[ A_\omega \rho = b(\omega) \]

\( A \in \mathbb{C}^{N \times K} \) is fat \((N \ll K)\)

The vectors \( \hat{g}(y_j, \omega) \), for \( y_j \) in IW must be nearly orthogonal

\[ \hat{g}(y_i, \omega)^* \hat{g}(y_j, \omega) \approx 0 \text{, } y_i \neq y_j \]

which holds if \( \lambda \ll \|y_i - y_j\| \ll L \) (see, Borcea, Papanicolaou, Tsogka).
Numerical examples
### Numerical examples

<table>
<thead>
<tr>
<th>Set-up</th>
<th>Units normalized by $\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of the array, a</td>
<td>100</td>
</tr>
<tr>
<td>Distance between detectors, $h$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>Distance from the array to IW, $L$</td>
<td>100</td>
</tr>
<tr>
<td>Size of IW</td>
<td>41x 41</td>
</tr>
<tr>
<td>Sampling rate of IW</td>
<td>1</td>
</tr>
</tbody>
</table>

The size of $A_w$ is 201x1681

- 88% of the equations are missing

Phase information is available in these experiments!!
Numerical examples

<table>
<thead>
<tr>
<th>Set-up</th>
<th>Units normalized by $\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of the array, $a$</td>
<td>100</td>
</tr>
<tr>
<td>Distance between detectors, $h$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>Distance from the array to IW, $L$</td>
<td>100</td>
</tr>
<tr>
<td>Size of IW</td>
<td>$41 \times 41$</td>
</tr>
<tr>
<td>Sampling rate of IW</td>
<td>1</td>
</tr>
</tbody>
</table>

The size of $A_w$ is $201 \times 1681$
Numerical examples

\[
\ell_2\| \text{noise} \%
\]
Numerical examples

**Noise**

- **0 %**
- **10 %**
- **100 %**

**Graphs**

- **\( \epsilon_2 \)**
- **KM**
Numerical examples

- **e2**
  - Noise: 0%
  - Noise: 10%
  - Noise: 100%

- **KM**

- **e1**
Intensity-only data

Phase information **IS NOT** available
While in coherent imaging both amplitude and phase are available, in many applications ONLY the intensity of the data is available.

\[ \mathbf{b}_1(\omega) = (|b_1|^2, |b_2|^2, |b_3|^2, ..., |b_N|^2)^T \equiv \text{diag}(\mathbf{b}(\omega) \mathbf{b}(\omega)^*) \]

Hence the equation for \( \rho \) in intensity-based imaging is given by

\[ \text{diag}(A_\omega \rho \rho^* A_\omega^*) = \mathbf{b}_1(\omega) \]
While in coherent imaging both amplitude and phase are available, in many applications only the intensity of the data is available.

\[ \mathbf{b}_I(\omega) = (|b_1|^2, |b_2|^2, |b_3|^2, ..., |b_N|^2)^T \equiv \text{diag}(\mathbf{b}(\omega) \mathbf{b}(\omega)^*) \]

Hence the equation for \( \mathbf{\rho} \) in intensity-based imaging is given by

\[ \text{diag}(A_\omega \mathbf{\rho} \mathbf{\rho}^* A_\omega^*) = \mathbf{b}_I(\omega) \]

Straightforward formulation:

\[ \min \| \mathbf{\rho} \|_{l_1} \text{ subject to } \text{diag}(A_\omega \mathbf{\rho} \mathbf{\rho}^* A_\omega^*) = \mathbf{b}_I(\omega) \]
Intensity-only data

We introduce the matrix variable \( X := \mathbf{\rho} \mathbf{\rho}^* \in \mathbb{R}^{K \times K} \) and the linear map \( \mathcal{L}_\omega: \mathbb{R}^{K \times K} \rightarrow \mathbb{R}^N \) such that \( \mathcal{L}_\omega (X) := \text{diag}(A_\omega X A_\omega^*) \) for any matrix \( X \in \mathbb{R}^{K \times K} \). Hence, we have

\[
\mathcal{L}_\omega (X) := b_1(\omega)
\]

Once \( X \) is found we obtain \( \mathbf{\rho} \) by taking \( \mathbf{\rho} = \sqrt{\text{diag}(X)} \).
We introduce the matrix variable $X := \rho \rho^* \in \mathbb{R}^{K \times K}$ and the linear map $\mathcal{L}_\omega : \mathbb{R}^{K \times K} \to \mathbb{R}^N$ such that

$$\mathcal{L}_\omega (X) := \text{diag}(A_\omega X A_\omega^*)$$

for any matrix $X \in \mathbb{R}^{K \times K}$. Hence, we have

$$\mathcal{L}_\omega (X) := b_1(\omega)$$

$X$ is:

- Low rank (rank 1)
- Sparse (due to the sparsity of $\rho$)
- Real symmetric and positive semidefinite (outer product of $\rho$ and $\rho^*$)

**Rank minimization problem:**

$$\min \text{rank}(X) \text{ subject to } \mathcal{L}_\omega (X) := b_1(\omega), \quad X \geq 0$$

where $X \geq 0$ is an entry-wise inequality
Minimizing the rank of a matrix subject to constraints is a general problem that arises in many applications including control, statistics, system identification, etc. But it is a non-convex optimization problem for which the best exact algorithms require a exponential time in the dimensions of the matrix $X$.

We use convex heuristics to replace rank($X$) by the nuclear norm $\|X\| = \text{sum of singular values of } X$ as the objective function.

**Rank minimization problem:**

$$\min \text{ rank}(X) \text{ subject to } \mathcal{L}_\omega (X) := b_I(\omega), \quad X \geq 0$$

where $X \geq 0$ is an entry-wise inequality
Minimizing the rank of a matrix subject to constraints is a general problem that arises in many applications including control, statistics, system identification, etc. But it is a non-convex optimization problem for which the best exact algorithms require a exponential time in the dimensions of the matrix $X$.

We use convex heuristics to replace $\text{rank}(X)$ by the nuclear norm $\|X\| = \text{sum of singular values of } X$ as the objective function.

Nuclear norm minimization problem:

$$\|X\|_* \text{ subject to } \mathcal{L}_\omega (X) := b_1(\omega), \quad X \geq 0$$

where $X \geq 0$ is an entry-wise inequality
Intensity-only data

Minimizing the rank of a matrix subject to constraints is a general problem that arises in many applications including control, statistics, system identification, etc. But it is a non-convex optimization problem for which the best exact algorithms require a exponential time in the dimensions of the matrix $X$.

We use convex heuristics to replace rank$(X)$ by the nuclear norm $\|X\| = \text{sum of singular values of } X$ as the objective function.

Moreover, because $X$ is positive semidefinite, the nuclear norm is the same as the trace. Therefore, the relaxed formulation is also equivalent to the trace heuristic

Trace minimization problem:

$$\min \text{trace}(X) \text{ subject to } \mathcal{L}_\omega (X) := b_1(\omega), \quad X \succeq 0, \quad X \geq 0$$

where $X \geq 0$ is an entry-wise inequality and the symbol $\succeq 0$ stands for p.s.d.
Main result

In general these two problems are not equivalent. Moreover, their solutions may not be unique.

We prove that under certain configurations of the imaging problem the rank minimization problem has a unique solution. However, we have not been able to prove that they have the same solution. Nevertheless, our numerical simulations show that the trace minimization does yield the solution of minimum complexity.

\[
\begin{align*}
\min \text{rank}(X) \quad \text{subject to} \quad L_\omega (X) := b_1(\omega), & \quad X \geq 0 \\
\min \text{trace}(X) \quad \text{subject to} \quad L_\omega (X) := b_1(\omega), & \quad X \succeq 0, \quad X \geq 0
\end{align*}
\]
Main result

In general these two problems are not equivalent. Moreover, their solutions may not be unique.

**We prove that under certain configurations of the imaging problem the rank minimization problem has a unique solution.** However, we have not been able to prove that they have the same solution. Nevertheless, our numerical simulations show that the trace minimization does yield the solution of minimum complexity.

We use the restricted isometry property, generalized to matrix variable problems, and the special structure of the imaging problem, to prove this result (analysis of sums of products of four and eight Green's functions in an asymptotic regime).

\[
\begin{align*}
\min \text{rank}(X) & \text{ subject to } L_\omega (X) := b_I(\omega), \quad X \geq 0 \\
\min \text{trace}(X) & \text{ subject to } L_\omega (X) := b_I(\omega), \quad X \geq 0
\end{align*}
\]
Numerical examples
### Numerical examples

<table>
<thead>
<tr>
<th>Set-up</th>
<th>Units normalized by $\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of the array, $a$</td>
<td>100</td>
</tr>
<tr>
<td>Distance between detectors, $h$</td>
<td>5</td>
</tr>
<tr>
<td>Distance from the array to IW, $L$</td>
<td>100</td>
</tr>
<tr>
<td>Size of IW</td>
<td>11x 11</td>
</tr>
<tr>
<td>Sampling rate of IW</td>
<td>1</td>
</tr>
</tbody>
</table>

The size of $A_w$ is 21x121
- 82% of the equations are missing

Phase information **IS NOT available** in this experiments!!
## Numerical examples

<table>
<thead>
<tr>
<th>Set-up</th>
<th>Units normalized by λ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of the array, a</td>
<td>100</td>
</tr>
<tr>
<td>Distance between detectors, h</td>
<td>5</td>
</tr>
<tr>
<td>Distance from the array to IW, L</td>
<td>100</td>
</tr>
<tr>
<td>Size of IW</td>
<td>11x 11</td>
</tr>
<tr>
<td>Sampling rate of IW</td>
<td>1</td>
</tr>
</tbody>
</table>

**SINGLE ILLUM.**

![original config.](image1.png)

![0 % noise](image2.png)

![0.5 % noise](image3.png)
Numerical examples

**MULTIPLE ILLUM.**

- 5
- 10

**SINGLE ILLUM.**

- 0.5 % noise
- 0 %
- 0.5 %

original config.
Conclusions

• Effectiveness of the formulation of the intensity-only problem as finding the minimum rank solution of a linear matrix equation associated with the locations and reflectivities of the scatterers.
• We seek a matrix solution of minimal complexity.
• Images are very sensitive to noise (much more than with full data) but it can reduce using multiple illumination.
• Need to optimize the illumination which, from previous work (Borcea Papanicolaou and Tsogka), makes a big difference.
• Medical and other imaging applications can benefit substantially from intensity-only imaging as the cost of the devices is much lower.