Compressed sensing and imaging

Anwei Chai

Computational and Mathematical Engineering
Stanford University

MATH00112230: Selected Topics in Advanced Statistics

Joint with Laurent Demanet (MIT), George Papanicolaou (Stanford)
Part I. Solving linear equations
  - least squares
  - regularization (\(\ell_2\) penalty method)
  - regularization (smoothness)
  - regularization (sparsity \(\ell_1\))

Part II. Imaging localized scatterers
  - Array imaging
  - Kirchhoff migration, least square imaging
  - Multiple signal classification (MUSIC)
  - Singular value decomposition and \(\ell_1\) based method
Part I. Solving linear equations – from least square to compressed sensing
Solving linear equations

- Model:
  \[ y = Ax + n \]  

- \( y \) is data, \( x \) is model image, \( n \) is noise

\[ \begin{align*}
\begin{bmatrix}
\text{h} \\
\text{x1}
\end{bmatrix} &=
\begin{bmatrix}
\mu \\
\text{nxm}
\end{bmatrix} +
\begin{bmatrix}
\lambda \\
\text{nx1}
\end{bmatrix}
\end{align*} \]

\( n > m \)
\( n < m \)
Solving linear equations – least square

- \( n > m \) (over-determined): Find minimizer of \( J(x) = \|Ax - y\|^2 \)
  - Solve \( \nabla J(x) = A^*(Ax - y) = 0 \) for minimizer
    - If \( A^*A \) is invertible, \( \hat{x} = A'^{\dagger}y \) and \( A'^{\dagger} = (A^*A)^{-1}A^* \) (pseudo-inverse)
    - If \( A^*A \) is not invertible (ker \( A \neq 0 \)), \( \hat{x} = A'^{\dagger}y + z, \forall z \in \text{ker } A \)
  - Calculate \( A'^{\dagger} \) from singular value decomposition (SVD)

- \( n < m \) (under-determined): infinitely many solutions if any

  least square gives least norm solution

\[
\hat{x} = A'^{\dagger}y, \quad A'^{\dagger} = A^*(AA^*)^{-1}
\]
Motivation: consider $n = m$ and (1) has solution
\[ \hat{x} = A^{-1}y = x + A^{-1}n \]
bad when $A$ has small singular values

Assume $\|x\|_2$ small
\[
\min \|x\|_2 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \varepsilon
\] (2)

(2) has several equivalent forms
\[
\min \|Ax - y\|_2 \quad \text{s.t.} \quad \|x\|_2 \leq \varepsilon'
\]
\[
\min \|x\|_2^2 + \lambda\|Ax - y\|_2^2
\]
\[
\min \|Ax - y\|_2^2 + \lambda'\|x\|_2^2
\]

Solution to Lagrangian form of (2) is \[ \hat{x} = (A^*A + \lambda I)^{-1}A^*y \]
Assume $\|Bx\|_2$ small: variation of $\ell_2$ penalty
\[
\text{diag}(B, 0) = 1, \quad \text{diag}(B, 1) = -1, \quad \text{diag}(B, k) = 0, \quad k \neq 0, 1
\]

- $Bx$ is finite difference approximation to derivative
- Tikhonov regularization in image processing and Hodrick-Prescott filtering in time series analysis
- Solution is given by $\hat{x} = (A^*A + \lambda B^*B)^{-1}A^*y$
- View point from numerical linear algebra: add extra term ($\lambda I$ or $\lambda B^*B$) to make $A^*A$ well conditioned to take inverse
Solving linear equations – $\ell_1$ penalty

- Assume under-determined $A$ and sparse structure of $x$, i.e., $\|x\|_0 = |\text{supp } x| = |\{ i : x_i \neq 0 \}|$ small
- When $A = I$, choose $x$ s.t.
  \[
  \min \|x\|_0 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \sigma \quad \text{(3)}
  \]
- When $A \neq I$, very difficult to identify nonzeros in $x$
- Relaxation to $\ell_1$ norm of $x$: Basis Pursuit Denoising (BPD) or Least Angle Shrinkage and Selector Operator (LASSO)
  \[
  \min \|x\|_1 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \sigma \quad \text{(4)}
  \]
- Solution is given by soft-threshold
  \[
  \hat{x}_i = S_\tau y_i = \begin{cases} 
  y_i - \tau \text{sgn } y_i, & |y_i| \geq \tau \\
  0, & \text{otherwise}
  \end{cases}
  \]
Solving linear equations – connection with uncertainty principle

- $X$ is transformed $x$ (e.g. in frequency domain), $N_t = |\text{supp } x|$, $N_\omega = |\text{supp } X|$
  - $N_t N_\omega \geq m$, $N_t + N_\omega \geq 2\sqrt{m}$ (Donoho & Stark 1989)
  - $N_t + N_\omega \geq m + 1$ when $m$ is prime (Tao 2003)

Let $A = [I, F]$ where $m = 2n$ and $F$ is normalized Fourier matrix; $T = \text{supp } x \cap \{1, 2, \ldots, n\}$, $W = \text{supp } x \cap \{n + 1, n + 2, \ldots, 2n\}$. Assume there is no noise (Donoho & Huo 1999)
  - $|T| + |W| < \sqrt{n}$, solution of (3) is unique and equals $x$
  - $|T| + |W| < \sqrt{n}/2$, solution of (4) is unique and equals $x$
  - $\exists n$, $\sqrt{n} \in \mathbb{Z}$ and $|T| + |W| = \sqrt{n}$, solutions of (3) and (4) are not unique

if $|T|$ small, no small number of frequencies can recover $x$
Solving linear equations – compressed sensing

- $A = I$: soft-threshold
- $A = [\Phi_1, \Phi_2]$: mutual incoherence
  
  \[ \mu = \max \{|\langle \phi_1, \phi_2 \rangle|, \phi_1 \in \Phi_1, \phi_2 \in \Phi_2\} \quad \text{and} \quad \|x\|_0 < (1 + 1/\mu)/2 \] (also true for sparse solution under general $A$)

- Compressed sensing: sample enough data
  - Counting faces of random polytope (Donoho 2004, Donoho & Tanner 2009)
  - Restricted Isometry Property (RIP) (Candés & Tao 2004-2005)

Matrix $A$ is called to satisfy RIP-S if for any subset $T \in \{1, 2, \ldots, m\}$, $|T| = S$, there exists $0 < \delta_S < 1$ such that:

\[ (1 - \delta_S)\|c\|_2^2 \leq \|A_Tc\|_2^2 \leq (1 + \delta_S)\|c\|_2^2 \]

for any $c \in \mathbb{R}^S$
Solving linear equations – compressed sensing

Three kinds of matrices are proven to satisfy exact reconstruction using (4) with high probability

- uniformly randomly sampled Fourier matrix

Choose $n \geq 22(1 + \delta)\|\mathbf{x}\|_0 \log m$ rows uniformly at random from Fourier matrix $F$. Then (4) recovers $\mathbf{x}$ with probability $p = 1 - \mathcal{O}(m^{-\delta})$

- Gaussian ensemble: $A_{ij}$'s are iid Gaussian random variables up to normalization constant

If entries of $A$ are sampled from $\mathcal{N}(0, 1/n)$ and $S \leq Cn/\log(m/n)$, then reconstruction succeeds with probability $p = 1 - \mathcal{O}(e^{-\gamma m})$ for some $\gamma > 0$

- Bernoulli ensemble: $A_{ij}$'s are iid Bernoulli random variables up to normalization constant

Same as Gaussian ensemble with $A$ sampled from symmetric Bernoulli $\pm 1/\sqrt{n}$
Solving linear equations – compressed sensing

Deterministic exact and stability results are proven under RIP

Assume there is no noise and $x$ is $S$ sparse and $\delta_{2S} + \delta_{3S} < 1$. Then solving (4) gives exact solution.

Consider problem (1). If $\delta_{3S} + 3\delta_{4S} < 2$, then solution $x^*$ to (4) recovers original solution $x$ satisfying $\|x\|_0 \leq S$ up to

$$\|x^* - x\|_2 \leq C_S\sigma$$

If $x_S$ is the approximation to any $x$ by keeping only the first $S$ largest values. Then solution $x^*$ to (4) gives approximation

$$\|x^* - x\|_2 \leq C_S\sigma + C'_S \frac{\|x - x_S\|_1}{\sqrt{S}}$$
Solving linear equations – numerical solution

- **Off-shelf:** any optimization package
  - linear programming for equality constraint
  - second order cone programming for inequality constraint
  - semi-definite programming for higher order problem

  “easy” to implement but extremely inefficient for large scale problem (good for problem of size less than 100)

- **Iterative methods:**
  - Iterative threshold method (IST)
  - Breigman, linearized Breigman
  - Message passing method

  straightforward to implement, fast, and efficient for large scale problem
Solving linear equations – application

- error correction in communication
- mathematical image processing (motion correction)
- medical imaging (MRI to reduce treatment time expense)
- statistics (feature selection)
- a lot more, see list at http://dsp.rice.edu/cs/
Part II. Imaging localized scatterers
Overview – questions and methods

- Compare several imaging methods regarding resolution and stability, i.e. loss of resolution when signal to noise ratio (SNR) is low at imaging array or medium between array and targets is inhomogeneous.

- Methods to be compared are:
  - **KM**: travel time migration imaging
  - **LSQ**: least square imaging
  - **MUSIC**: MULTiple SIGNal Classification
  - **CS**: \( \ell_1 \) minimization with or without SVD
Array imaging – schematic

Configuration: range is $100\lambda_0$; array size is between $100\lambda_0$ and $200\lambda_0$ with transducers half of wavelength apart; target area is box of length $20\lambda_0$ with half of wavelength resolution
Array imaging – formulation

Time harmonic Green function in homogeneous medium with speed \( c_0 \) is

\[
\hat{G}_0(x, y, \omega) = \frac{\exp(i \frac{\omega}{c_0}|x-y|)}{4\pi|x-y|}
\]

Response at each sensor \( x_s \) on the array excited by the signal emitted from \( x_s \) and reflected by scatters of reflectivity \( \rho_j, j = 1, \ldots, M \) is given by

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

Let \( \hat{g}_0(y, \omega) = [\hat{G}_0(x_r, y, \omega)] \) and with excitation \( f(\omega) \),

\[
[(\hat{g}_0^T(y, \omega)f(\omega))\hat{g}_0(y, \omega)]\rho = \hat{\Pi}(\omega)f(\omega)
\]
Rewriting by letting $A_f(\omega) = [(\hat{g}_0^T(y, \omega)f(\omega))\hat{g}_0(y, \omega)]$, reflectivity vector $\rho$ satisfies

$$A_f(\omega)\rho = \hat{\Pi}(\omega)f(\omega)$$

(5)

Imaging is inverse problem which needs to solve (5) (with or without subsampling)
Kirchhoff migration and least square imaging

\[ \hat{A}^+ = (A^H A)^{-1} A^* \]

- **LSQ**: \( \min \| \rho \|_2 \) s.t. \( A_f(\omega) \rho = \hat{\Pi}(\omega) f(\omega) \) and solution is given by \( \hat{\rho}_{\text{LSQ}} = A^\dagger(\omega) \hat{\Pi}(\omega) f(\omega) \)

- **KM**: approximate solution \( \rho_{\text{KM}} = A_f^H(\omega) \hat{\Pi}(\omega) f(\omega) \)

**KM** only gives the location of scatterers; **LSQ** computes reflectivity of scatterers due to additional terms in \( A^\dagger \) besides \( A^* \).
Multiple Signal Classification (MUSIC)

MUSIC assumes data are formed by some known basis. In imaging, it means data are formed by Green function vectors $\hat{g}_0(y, \omega)$. Let SVD of response matrix $\hat{\Pi}(\omega)$ be $\hat{\Pi}(\omega) = U \Sigma V^* = \sum_{j=1}^{L} \sigma_j u_j v_j^H$

- Can show $L \approx M$, $u_j = \frac{\hat{g}_0(y_{kj}, \omega)}{\|\hat{g}_0(y_{kj}, \omega)\|_2}$, $v_j = \frac{\bar{u}_j}{\sigma_j}$, $\sigma_j = \rho_j \|\hat{g}_0(y_{kj}, \omega)\|_2^2$

- Assume $\hat{g}_0(y, \omega)$’s are normalized, $u_k^* \hat{g}_0(y^S, \omega)$ has peak when $y^S$ is close to a scatterer

- MUSIC imaging is given by

$$\hat{\rho}_{MUSIC}(y^S) = \frac{\min_{y^S} G_{MUSIC}(y^S)}{G_{MUSIC}(y^S)},$$

where

$$G_{MUSIC}(y^S) = \| P_N \hat{g}_0(y^S, \omega) \|_2^2$$

and

$$P_N \hat{g}_0(y^S, \omega) = \sum_{j=1}^{M} \left( u_j^H \hat{g}_0(y^S, \omega) \right) u_j - \hat{g}_0(y^S, \omega)$$
**MUSIC** only gives location of scatterers in terms of the peak values of the functional. Solving resulting overdetermined least square problem by columns of $A_f(\omega)$ corresponding to large **MUSIC** functional

\[ \text{LSQ } A_f(\omega) \mathbf{p} = \frac{\nabla f}{f} \]
Imaging with $\ell_1$ optimization

- Solution $\rho$ is sparse
- Application of compressed sensing directly

$$\min \| \rho \|_1 \quad \text{s.t.} \quad \Phi A_f(\omega)\rho = \Phi \hat{\Pi}(\omega)f(\omega)$$

$\Phi$ is uniform random sampling operator

- Accurate results subject to high probability of success and hard to determine the sampling ratio
- Physical experiment by group in Georgia Tech (Gurbuz, McClellan, Scott 2009); theoretical analysis by (Fannjiang 2009)
Imaging with SVD and $\ell_1$ optimization

- Implementation of random sampling, amount of subsampling, flexibility of illumination (i.e. $f(\omega)$)
- SVD of response matrix is powerful to identify number of scatterers and robust to noise
- Green function vectors $\hat{g}_0(y, \omega)$ are asymptotically orthogonal, i.e.

$$\hat{g}_0^H(y_i, \omega)\hat{g}_0(y_j, \omega) = \begin{cases} \|\hat{g}_0(y_i, \omega)\|_2^2, & \text{if } i = j, \\ 0 \ (\text{approximately}), & \text{otherwise,} \end{cases}$$

provided $\lambda \ll \|y - y^S\|_2 \ll L$
Imaging with SVD and $\ell_1$ optimization

Use SVD to reduce dimension (subsample) of data

- Project to space spanned by vectors according to nonzero singular values

$$u_k^*(\omega)A_f(\omega)\rho = u_k^*(\omega)\hat{\Pi}(\omega)f(\omega), \quad k = 1, \ldots, M$$

- Apply illumination vectors $v_k(\omega)$

$$u_k^*(\omega)A_{v_k}(\omega)\rho = u_k^*(\omega)\hat{\Pi}(\omega)v_k(\omega) = \sigma_k, \quad k = 1, \ldots, M$$

- Rewrite new linear system as $B\rho = b$

$$B_{ij} = (\hat{g}_0^T(y_j, \omega)v_i(\omega))\hat{g}_0^*(y_j, \omega)u_i(\omega) = \left(u_i^*\hat{g}_0(y_j, \omega)\right)^2$$

$$b = [\sigma_1, \cdots, \sigma_M]^T$$
Imaging with SVD and $\ell_1$ optimization

- Lower-bound: reduce dimension of data to minimum
- Optimality: use singular vector as illumination
- Guaranteed solution in specific regime ($\lambda \ll \|y - y^S\|_2 \ll L$):
  - $B = [I + E S]$ is full row-rank and contains an identity matrix up to normalization
  - If $\|E\|_1 \leq \varepsilon$ and $\|S\|_1 \leq 1 - \varepsilon$, then solving
    \[
    \min \|\rho\|_1 \quad \text{s.t.} \quad B\rho = b
    \]
    gives the exact solution
  - Proof relies on the fact that the solution to (4) is unique if there exists vector $w$ such that inner product of $w$ and any column vector in $B$ is no greater than 1 (Candés & Tao 2006)
  - No specific numerical method is need since problem size becomes small and $B$ is approximately sparse
Imaging with additive noise

- Use inequality constraint \( \| A_f(\omega) \rho - \hat{\Pi}(\omega) f(\omega) \|_2 \leq \varepsilon \) in LSQ and CS

- Noise is generated on imaging array in homogeneous medium and only affect high frequency data, i.e. space corresponding to zero/small singular values (Borcea, Papanicolaou, Vasquez 2008)

\[
\hat{\Pi}(\omega) + \Delta \hat{\Pi}(\omega) = [U_1 \quad U_2 U_{\Delta \hat{\Pi}}] \begin{bmatrix} \Sigma & \Sigma_{\Delta \hat{\Pi}} \\ \Sigma_{\Delta \hat{\Pi}} & V_{\Delta \hat{\Pi}}^* V_{\Delta \hat{\Pi}} \end{bmatrix} \begin{bmatrix} V_1^* \ V_2^* \end{bmatrix}
\]

- SVD filters out noisy terms in \( \ell_1 \) optimization; no more additional denoising process is needed
Numerical simulation

- Setup of simulation

<table>
<thead>
<tr>
<th></th>
<th>two dimension</th>
<th>three dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>length of array</td>
<td>$100 \lambda_0$</td>
<td>$100 \lambda_0$</td>
</tr>
<tr>
<td>distance between detectors $h$</td>
<td>$\lambda_0/2$</td>
<td>$4 \lambda_0$</td>
</tr>
<tr>
<td>array to IW distance $L$</td>
<td>$100 \lambda_0$</td>
<td>$100 \lambda_0$</td>
</tr>
<tr>
<td>size of IW</td>
<td>$20 \lambda_0 \times 20 \lambda_0$</td>
<td>$20 \lambda_0 \times 5 \lambda_0 \times 20 \lambda_0$</td>
</tr>
<tr>
<td>sampling rate within IW $k$</td>
<td>$\lambda_0/2$</td>
<td>$\lambda_0/2$</td>
</tr>
</tbody>
</table>

The size of $\mathcal{A}_f$ is $201 \times 1681$ in two dimension and $676 \times 18491$ in three dimension.

- Perturbation to $\hat{\Pi}(\omega)$ is simulated by generating iid complex $\mathcal{N}(0, \delta p_{avg})$ and signal-to-noise ratio (SNR) is $-10 \log_{10} \delta$
Numerical simulation

A Chai (Stanford)

Compressed sensing and imaging

MATH00112230 29 / 38
### Numerical simulation

<table>
<thead>
<tr>
<th>MUSIC</th>
<th>( \ell_1 )</th>
<th>( \ell_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>loc</td>
<td>val</td>
<td>loc</td>
</tr>
<tr>
<td>(8, 1, 12)</td>
<td>1</td>
<td>(24, 5, 4)</td>
</tr>
<tr>
<td>(24, 5, 4)</td>
<td>0.689</td>
<td>(31, 9, 17)</td>
</tr>
<tr>
<td>(31, 9, 17)</td>
<td>0.4031</td>
<td>(8, 1, 12)</td>
</tr>
<tr>
<td>(25, 8, 30)</td>
<td>0.2139</td>
<td>(21, 11, 13)</td>
</tr>
<tr>
<td>(21, 11, 13)</td>
<td>0.1296</td>
<td>(25, 8, 30)</td>
</tr>
<tr>
<td>(21, 10, 13)</td>
<td>0</td>
<td>(21, 11, 14)</td>
</tr>
<tr>
<td>(31, 10, 17)</td>
<td>0</td>
<td>(7, 1, 12)</td>
</tr>
<tr>
<td>(25, 9, 30)</td>
<td>0</td>
<td>(31, 10, 17)</td>
</tr>
<tr>
<td>(31, 8, 17)</td>
<td>0</td>
<td>(25, 9, 30)</td>
</tr>
<tr>
<td>(25, 7, 30)</td>
<td>0</td>
<td>(24, 4, 4)</td>
</tr>
</tbody>
</table>
Numerical simulation

MUSIC

Kirchhoff migration

Compressed sensing and imaging
### Numerical simulation

#### MUSIC

<table>
<thead>
<tr>
<th>loc</th>
<th>val</th>
<th>$\ell_1$</th>
<th>loc</th>
<th>val</th>
<th>$\ell_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(24, 5, 4)</td>
<td>1</td>
<td>(24, 5, 4)</td>
<td>30.0554</td>
<td>(8, 1, 12)</td>
<td>5.8243</td>
</tr>
<tr>
<td>(31, 9, 17)</td>
<td>0.8352</td>
<td>(31, 9, 17)</td>
<td>28.0076</td>
<td>(21, 11, 13)</td>
<td>4.3916</td>
</tr>
<tr>
<td>(8, 1, 12)</td>
<td>0.5154</td>
<td>(8, 1, 12)</td>
<td>21.0721</td>
<td>(7, 1, 12)</td>
<td>3.5743</td>
</tr>
<tr>
<td>(31, 8, 17)</td>
<td>0.3179</td>
<td>(21, 11, 13)</td>
<td>16.374</td>
<td>(24, 5, 4)</td>
<td>3.508</td>
</tr>
<tr>
<td>(31, 10, 17)</td>
<td>0.3121</td>
<td>(25, 8, 30)</td>
<td>14.5407</td>
<td>(8, 1, 11)</td>
<td>3.3794</td>
</tr>
<tr>
<td>(24, 6, 4)</td>
<td>0.3078</td>
<td>(31, 8, 17)</td>
<td>0.1677</td>
<td>(31, 9, 17)</td>
<td>3.2939</td>
</tr>
<tr>
<td>(24, 4, 4)</td>
<td>0.2941</td>
<td>(21, 10, 13)</td>
<td>0.139</td>
<td>(31, 10, 17)</td>
<td>3.1467</td>
</tr>
<tr>
<td>(21, 11, 13)</td>
<td>0.2711</td>
<td>(25, 7, 30)</td>
<td>0.1294</td>
<td>(24, 4, 4)</td>
<td>3.1138</td>
</tr>
<tr>
<td>(8, 2, 12)</td>
<td>0.2375</td>
<td>(24, 6, 4)</td>
<td>0.0972</td>
<td>(24, 4, 4)</td>
<td>3.0835</td>
</tr>
<tr>
<td>(25, 8, 30)</td>
<td>0.2225</td>
<td>(8, 2, 12)</td>
<td>0.0874</td>
<td>(21, 11, 14)</td>
<td>2.7385</td>
</tr>
</tbody>
</table>
Random medium

Phase only perturbation model of random medium.

Realization of random medium with strength of fluctuation $\sigma_0 = 3\%$ and correlation length half of wavelength.
Wave speed in a random medium is modeled by

\[ c^{-2} = c_0^{-2}(1 + \sigma_0 \mu) \]

where \( \mu(\cdot) \) is zero mean, isotropic, weakly stationary random fluctuation. Green function in random medium is modeled by

\[ \hat{G}(x, y, \omega) = \hat{G}_0(x, y, \omega) e^{i \frac{\omega \sigma_0}{c_0} |x-y| \int_0^1 \mu \left( \frac{s}{\ell} + \frac{s}{\ell} (y-x) \right) ds} \]

and response matrix is obtained by substituting \( \hat{G}_0 \) by \( \hat{G} \). As in the case of additive noise, inequality constraints are used in \( \text{LSQ} \) and \( \text{CS} \), but the noise level \( \varepsilon \) is usually impossible to estimate accurately, which limits the use of minimization approaches in random medium.
The moment formula holds for random medium model:

\[
\mathbb{E} \left( \hat{G}(x, y, \omega) \hat{G}(x, y', \omega) \right) \approx \hat{G}_0(x, y, \omega) \hat{G}_0(x, y', \omega) e^{-\frac{(\frac{\omega}{c_0})^2 |y-y'|^2}{2L^2}} a_e^2,
\]

where

\[
a_e = \sigma_0 L \left( -R(0) - \frac{2}{3} \frac{L}{\ell} \int_0^{\infty} dt \frac{R'(t)}{t} \right)^{1/2}
\]

is the effective aperture of random medium.
Analysis: statistical stability

As size of array becomes larger, the random medium model satisfies

\[
\mathbb{E} \left| \hat{G}(x, y, \omega) \hat{G}(x, y', \omega) - \mathbb{E} \left( \hat{G}(x, y, \omega) \hat{G}(x, y', \omega) \right) \right|^2 \\
\approx \frac{1}{(4\pi)^4 |x - y|^2 |x - y'|^2} \left( 1 - e^{-\left( \frac{\omega}{c_0} \right)^2 \frac{a^2}{L^2} |y - y'|^2} \right),
\]

\[
\mathbb{E} \left| \hat{g}^H(y, \omega) \hat{g}(y', \omega) - \mathbb{E} \left( \hat{g}^H(y, \omega) \hat{g}(y', \omega) \right) \right|^2 \\
\mathbb{E} \left| \hat{g}(y, \omega) \right|^2 
\rightarrow 0, \quad a \rightarrow \infty.
\]
Array length: $100\lambda_0$ (top), $150\lambda_0$ (middle), $200\lambda_0$ (bottom). Fluctuation of random medium is 0.3%. Resolution improvement is clear for KM, MUSIC, SVD which implies statistical stability.
**Conclusion: Performance of methods**

**KM:** Simple to implement for large arrays, low resolution, wide area imaging, robust to additive noise and random media.

**LSQ:** More costly to implement for large arrays, improves resolution, localized or wide area imaging, not so robust to noise and random media.

**MUSIC:** Simple to implement for small (sparse) targets using the SVD, has good resolution, relatively robust to noise and random media; the modified method can be the basis of more elaborate imaging algorithms to achieve various effects.

**CS:** Very effective for sparse targets in homogeneous media, fast algorithm, excellent accuracy even with substantial subsampling. Not robust to noise or random media (physical experiment by group from Duke).