A differential equations approach to $l_1$-minimization with applications to array imaging

Miguel Moscoso, Madrid
Alexei Novikov, Penn State
George Papanicolaou, Stanford
Lenya Ryzhik, Stanford

1. The array imaging problem
2. The $l_1$-minimization problem and its variational formulation
3. The ODE formulation
4. The numerical algorithm and its convergence
5. Numerical examples in array imaging
Overall goal: understand behavior of waves in weakly random media – long distances and large times.

Applications: (1) forward problems – when and what to measure, (2) inverse problems – find (small) inclusions

Questions: (i) effective macroscopic models, (ii) correctors to such models to do inverse problems

Today’s talk: inverse problems in array imaging are very underdetermined.
**Array imaging**: see a picture on the board

Array: \(N\) transducers/receivers (one frequency, one transducer).

\(M\) point-like scatterers, unknown reflectivities \(\rho_j > 0\) at unknown \(z_{n_j}\) \((j = 1, \ldots, M)\) — inside the image window.

Image Window:

\(K\) points (pixels) \(z_j, j = 1, \ldots, K\), \(K \gg N, M\).

The linear system: \(Ax = y\); \(A\) – comes from the model (wave equation)

\(x \in \mathbb{R}^K\) — vector of reflectivities \(\rho_j, j = 1, \ldots K\).

\(y \in \mathbb{R}^N\) — vector of measurements at the array
Main aspects of array imaging for today’s talk (noisy linear algebra)

\[ Ax = y, \ x \in \mathbb{R}^K, \ y \in \mathbb{R}^N \]

1. \( K \gg N \) – the system is very underdetermined
2. There is additive measurement noise in \( y \) – ambient noise, multiple scattering between point scatterers, etc.
3. The medium is often unknown – the matrix \( A \) is not known exactly.
4. We are just learning the tools.
The ideal program

0. Formulate the array imaging as a constrained minimization problem.
1. Convert the constrained problem into a PDE (an ODE will do) – because this is what we know how to analyze.
2. Consider the constrained minimization problem with noise.
3. Consider the constrained minimization problem in a random medium.
4. Move on.

Today: a little bit of 1.
What is the right constraint?

Find the "sparsest" solution of an underdetermined system:

\[ Ax = y, \quad x \in \mathbb{R}^n, \quad y \in \mathbb{R}^m. \]

Number of measurements: \( m \ll n \).

Problem\(_0\): recovery of the sparsest solution:

\[
\text{find } \arg\min_{x} \| x \|_{l_0} : \ Ax = y. \quad \text{Hard (Candes, Donoho, ...)}
\]

Problem\(_1\) (the \( l_1 \)-minimization problem) easy:

\[
\text{find } \arg\min_{x} \| x \|_{l_1} \text{ subject to } Ax = y.
\]

Candes, Tao, Donoho, ... under some conditions

Problem\(_0\) \iff\ Problem\(_1\) – this is a mystery in array imaging
The unconstrained relaxed formulation:

\((P_{relax})\) find \(\arg\min \left[ \tau \|x\|_1 + \frac{1}{2} \|Ax - y\|^2 \right].\)

Example: minimize \(f(x) = \tau |x| + \frac{1}{2} (x - 1)^2, \ x \in \mathbb{R}\)

Small \(\tau \in (0, 1)\): \(x_{min} = 1 - \tau\)

Large \(\tau > 1\): \(x_{min} = 0 \) – because \(|x|\) has a corner.

You never get an exact reconstruction of the \(l_1\)-minimization problem and need \(\tau \ll 1\) to be close.

You may not want to find the exact solution but that is a separate issue.
Iterative shrinkage-thresholding algorithm (ISTA) for $P_{relax}$:

$$x_{k+1} = \eta \tau h (x_k - h \nabla f(x_k)),$$

$$f(x) = \frac{1}{2} \| y - Ax \|^2$$

$$\eta_a(x) = \begin{cases} 
    x - a, & \text{if } x > a, \\
    0, & \text{if } -a < x < a, \\
    x + a, & \text{if } x < -a, 
\end{cases}$$

$$\nabla f(x_k) = A^* (Ax_k - y)$$

(1) simple to implement, (2) "time splitting"

$$x_{k+1} = \arg\min_x \{ \tau \| x \|_1 + f(x_k) + \langle x - x_k, \nabla f(x_k) \rangle + \frac{1}{2h} \| x - x_k \|^2 \}.$$  

The Lyapunov function:

$$F(x) = \tau \| x \|_1 + \frac{1}{2} \| y - Ax \|^2: \quad F(x_{k+1}) \leq F(x_k).$$

Converges to minimizer of $(P_{relax})$ if $h < 1/\| A^* A \|$.
Advantage: very easy to implement.

Disadvantages:
(1) for a fixed $\tau$ the solution to $P_{relax}$ is not the same as the solution to Problem$_1$ – exact recovery from noiseless data is not achieved, unless $\tau$ is sent to zero.
(2) Convergence is slow for small $\tau$. Must use a very carefully chosen decreasing sequence of $\tau$.

A tiny sample of methods (it is a jungle out there):
Beck-Taboulle’08: FISTA – a much better convergence rate.
Osher school – Bregman iteration, linearized Bergman iteration, augmented Lagrangian methods.
Main results

We consider Problem 1: \( \min ||x||_{l_1}, \text{subject to } Ax = y. \)

Only assumptions – a unique minimizer \( \bar{x}, \) and \( A \) has full rank: the matrix \( AA^* \) is invertible.

The min-max variational principle:

\[
\bar{F} = \max_z \min_x \left\{ F(x, z) \equiv \tau ||x||_{l_1} + \frac{1}{2} ||Ax - y||^2 + \langle z, y - Ax \rangle \right\}.
\]

Theorem. (MNPR’11; Rockafellar’73) The max-min is attained, and \( \tau ||\bar{x}||_{l_1} = \max_x \min_z F(x, z) \)

Moreover, (1) \( \tau ||\bar{x}||_{l_1} = F(\bar{x}, z) \) for any \( z, \)

(2) if \( \min_x F(x, z) = \tau ||\bar{x}||_{l_1} \) for some \( z, \)

then \( \text{argmin}_x F(x, z) = \bar{x}. \)
How this works: solve equation $x = 1$

Example: $\max_{z \in \mathbb{R}} \min_{x \in \mathbb{R}} \left\{ \tau |x| + \frac{1}{2} (x - 1)^2 + z(1 - x) \right\}$.

$$l(z) = \min_x \left\{ \tau |x| + \frac{1}{2} (x - 1)^2 + z(1 - x) \right\} = \begin{cases} 
\tau |1 + z - \tau| + 1/2(\tau^2 - z^2), & z > \tau - 1, \\
1/2 + z, & -\tau - 1 < z < \tau - 1, \\
\tau |1 + z + \tau| + 1/2(\tau^2 - z^2), & z < -\tau - 1.
\end{cases}$$

$l(z)$ attains its maximum at $z = \tau$ for all $\tau > 0$, and

$$\arg\min_{x \in \mathbb{R}} \left\{ \tau |x| + \frac{1}{2} (x - 1)^2 + \tau(1 - x) \right\} = 1.$$ 

Exact recovery of $x = 1$ for all $\tau > 0$.

Choose $\tau$ large or small, depending on the application.
Why the quadratic term?

A Lagrange multiplier? \( \tau \|x\|_{l_1} + \langle z, y - Ax \rangle \),

The Euler-Lagrange equations give the sub-differential:

\[
[A^*z]_i = \begin{cases} 
\tau, & \text{if } \bar{x}_i > 0, \\
-\tau, & \text{if } \bar{x}_i < 0,
\end{cases}
\]

and \( |[A^*z]_i| \leq \tau \).

But: look at \( \tau |x| + z(1 - x) \). Then \( z = \tau \) satisfies the sub-differential condition, and

\[
\tau |x| + \tau (1 - x) = \begin{cases} 
\tau, & \text{if } x > 0, \\
\tau (1 - 2x), & \text{if } x < 0,
\end{cases}
\]

which has no unique minimum. Hence, the quadratic term.
The ODE approach

An obvious degeneracy: \( F(\bar{x}, z) = \tau \|\bar{x}\|_{l_1} \) for all \( z \in \mathbb{R}^m \).

We can only hope to recover \( \bar{x} \) — there is no “optimal” \( z \).

The ODE point of view:

\[
\frac{dx}{dt} = -\nabla_x F(x, z), \quad \frac{dz}{dt} = \nabla_z F(x, z),
\]

hope that \( x(t) \to \bar{x} \) as \( t \to +\infty \).

A (not only technical) difficulty: \( F(x, z) \) is not differentiable in \( x \) at the points where \( x_j = 0 \) for some \( j = 1, \ldots, n \).
Crandall-Liggett theory (primitivized)

\[ \frac{du}{dt} = Au, \quad u(0) = u_0. \]

1. If \( A \) is an accretive (contractive) operator then a strong solution exists that is differentiable for a.e. \( t \), no need for Lipschitzianity.

\( \dot{u} = \text{sgn}(u) \) vs. \( \dot{u} = -\text{sgn}(u) \).

2. This solution can be found by an implicit discretization in time.
Crandall-Liggett for our problem: the sub-differential $\partial \| x \|_{l_1}$

$$\partial \| x \|_{l_1} = \text{sgn}(x_1) \times \cdots \times \text{sgn}(x_n) \subset \mathbb{R}^n$$

$\text{sgn}(s)$ – a subset of $\mathbb{R}$: $\text{sgn}(s) = \begin{cases} 
{1} & \text{if } s > 0, \\
{-1} & \text{if } s < 0, \\
[-1, 1] & \text{if } s = 0.
\end{cases}$

Replace the discontinuous ODEs by Crandall-Ligget equations

$$\frac{dx}{dt} - A^*(z - Ax + y) \in -\tau\partial \| x \|_{l_1},$$
$$\frac{dz}{dt} = y - Ax.$$
The implicit scheme (also in Chambolle, Pock’11?):

\[
(1) \quad \frac{x_{k+1} - x_k}{\Delta t} = -\tau \xi_{k+1} + A^*(z_{k+1} + y - Ax_{k+1}),
\]

\[
(2) \quad \frac{z_{k+1} - z_k}{\Delta t} = y - Ax_{k+1},
\]

\[
(3) \quad \xi_{k+1} \text{ is a vector in the set } \partial \|x_{k+1}\|_{l_1}.
\]

Initial data: \( x_0 = x, \ z_0 = 0 \).

**Proposition.** Solution converges as \( \Delta t \to 0 \), uniformly on finite time intervals, to the unique strong solution of ODEs.

Because the right side is accretive and Crandall-Liggett applies.
How it works

The toy problem $\dot{r} = -\text{sgn}(r)$.

An explicit discretization: $\frac{r_{k+1} - r_k}{\Delta t} = -\xi_k$, $\xi_k \in \text{sgn}(r_k)$, oscillates around $r = 0$ if $r_k \in [-\Delta t, \Delta t]$, and will never converge to $x = 0$ for $\Delta t > 0$.

The implicit scheme: $\frac{r_{k+1} - r_k}{\Delta t} = -\xi_{k+1}$, $\xi_{k+1} \in \text{sgn}(r_{k+1})$ is different.

Forces $r_{k+1} = 0$ if $r_k \in [-\Delta t, \Delta t]$ equivalent to soft thresholding:

$r_{k+1} = \eta \Delta t(r_k)$.

The importance of being earnest and using an implicit scheme.
Convergence of the numerical algorithm

**Theorem (MNPR’12).** Given any $\delta > 0$ there exists $h > 0$ an $T > 0$ so that for all $0 < \Delta t < h$ and all $k > \lceil T/\Delta t \rceil$ we have $|x_k - \bar{x}| < \delta$. The time $T$ depends on $\delta$, the initial data $x \in \mathbb{R}^n$, and the norm $\|AA^*\|$ but not on dimension $n$.

Proved via convergence of ODEs:

**Theorem.** For any $\delta > 0$ there exists $T = T(\delta)$ such that the solution of ODEs satisfies $\|x(t) - \bar{x}\| < \delta$, for all $t > T$. The time $T(\delta)$ depends only on $\delta$, the initial data $x_0$, and $\|AA^*\|$ but not on the dimension $n$. 
**GeLMA**: an Euler quazi-explicit modification that is easier to implement, generalized Lagrange multiplier algorithm:

\[
x_{k+1} = x_k - \xi_{k+1} + \Delta t(A^*(z_k + y - Ax_k)),
\]

\[
z_{k+1} = z_k + \Delta t(y - Ax_k),
\]

where \( \xi_{k+1} \in \tau \Delta t \partial \|x_{k+1}\|_1 \) is a vector in the subdifferential of \( \tau \Delta t \|x_{k+1}\|_1 \). Equivalent to soft thresholding:

\[
x_{k+1} = \eta \tau \Delta t \left( x_k + A^*(z_k + y - Ax_k) \Delta t \right),
\]

\[
z_{k+1} = z_k + \Delta t(y - Ax_k).
\]

Converges if \( \Delta t < 1/||A|| \). GeLMA is very easy to implement.
The sub-differential condition: what happens to $z$

A by-product of the proof (another rediscovery of something VERY old):

**Theorem.** As $t \to +\infty$, $z(t)$ converges to a vector $z$ such that $[A^*z]_i = \text{sgn}(\bar{x})$ if $\bar{x} \neq 0$, and $|[A^*z]_i| \leq 1$ if $\bar{x}_i = 0$. 
The ODE approximation

An approximation of \( \text{sgn} x \): \( G_\varepsilon(s) = \begin{cases} 1, & \text{if } s > \varepsilon, \\ s/\varepsilon, & \text{if } |s| < \varepsilon, \\ -1, & \text{if } s < -\varepsilon. \end{cases} \)

The regularized ODE system:

\[
\frac{dx_\varepsilon}{dt} = -\tau G_\varepsilon(x_\varepsilon) + A^* (z + y - Ax_\varepsilon), \\
\frac{dz_\varepsilon}{dt} = y - Ax_\varepsilon.
\]

\( F(x, z) \) is replaced by \( F_\varepsilon(x, z) = \tau \sum_{j=1}^{n} r_\varepsilon(x_j) + f(x) + \langle z, y-Ax \rangle. \)

Here \( r_\varepsilon(s) = \begin{cases} |s|, & \text{if } s > \varepsilon, \\ s^2/(2\varepsilon) + \varepsilon/2, & \text{if } |s| < \varepsilon, \\ |s|, & \text{if } s < -\varepsilon, \end{cases} \)

is an approximation of \( |s| \) (the Huber function).
Theorem. For any $\delta > 0$ there exists $\varepsilon_0 = \varepsilon_0$ and $T = T(\delta)$ such that for any $\varepsilon$, $0 < \varepsilon < \varepsilon_0$ the solution of $\text{ODE}_\varepsilon$ satisfies $\|x_\varepsilon(t) - \bar{x}\| < \delta$, for all $t > T$. The time $T(\delta)$ depends only on $\delta$, the initial data $x_0$, and $\|AA^*\|$ but not on the dimension $n$.

This is the key ingredient in the proof of convergence: solution $x(t)$ of $\text{ODE}$ is the limit of $x_\varepsilon(t)$ as $\varepsilon \to 0$, and thus it obeys the same bounds as $x_\varepsilon(t)$, finishing the proof.

The main ingredient in the proof: energy estimates for $\text{ODE}_\varepsilon$. Elementary but not quite straightforward.
Basics of the proof: \( \bar{x} \) — the unique minimizer, set \( x_\varepsilon = \bar{x} + q_\varepsilon \):

\[
\frac{dq_\varepsilon}{dt} = -\tau G_\varepsilon (\bar{x} + q_\varepsilon) + A^* (z_\varepsilon - Aq_\varepsilon), \quad \frac{dz_\varepsilon}{dt} = -Aq_\varepsilon.
\]

Lemma 1. (The Lyapunov function)

\[
\|\dot{q}_\varepsilon(t)\|^2 + \|Aq_\varepsilon(t)\|^2 + \int_0^\infty \|A\dot{q}_\varepsilon(s)\|^2 ds < C_0,
\]

Why: \( \ddot{q}_\varepsilon + A^* A(\dot{q}_\varepsilon + q_\varepsilon) = -\tau g^\varepsilon (\bar{x} + q_\varepsilon) \dot{q}_\varepsilon \),

\( g^\varepsilon(x) \) a diagonal matrix \( g^\varepsilon_{ii}(x) = \begin{cases} 0, & \text{if } |x_i| > \varepsilon, \\ 1/\varepsilon, & \text{if } |x_i| < \varepsilon. \end{cases} \)

Singular pendulum, friction, forcing, expect \( Aq_\varepsilon(t) \to 0 \) as \( t \to +\infty \). Forcing and friction drive \( q_\varepsilon(t) \to 0 \).
Strategy of the proof

0. Need to use the fact that $\bar{x}$ is a minimizer – otherwise $q_\varepsilon(t) \to 0$ is nonsense!

1. Get a priori bounds on $q_\varepsilon(t)$ and $z_\varepsilon(t)$.
2. Show that any limit point of $q_\varepsilon(t)$ as $t \to +\infty$ is close to zero.
Lemma 2. \( \|z_\varepsilon(t)\| \leq C \) for all \( t > 0 \).

Proof. Lemma 1+ integration by parts:
\[
\ddot{z}_\varepsilon + AA^*(\dot{z}_\varepsilon + z_\varepsilon) = \tau AG_\varepsilon (\bar{x} + q_\varepsilon(t)).
\]

Lemma 3. (\( Aq_\varepsilon(t) \) is small sometimes). For any \( k \in \mathbb{N} \) there exists \( t_k < C_1k^3 \) such that for all \( t \in (t_k, t_k + C_2k) \) we have \( \|Aq_\varepsilon(t)\| \leq C_1/k \).

Proof. Lemmas 1+2, and \( \dot{z}_\varepsilon = -Aq_\varepsilon \).

Lemma 4. (A bound on \( x_\varepsilon(t) \).)

\[
\|\bar{x} + q_\varepsilon(t)\| \leq C.
\]

Here the precise form of the forcing in the pendulum is used.
\[ \left| \int_0^t \langle g^\varepsilon (\bar{x} + q^\varepsilon) \dot{q}^\varepsilon, q^\varepsilon \rangle ds \right| \leq C. \]

Fix \( 1 \leq j \leq n \) look at

\[
I = \int_0^t g^\varepsilon_{jj} (\bar{x} + q^\varepsilon(s)) q^\varepsilon,j(s) \dot{q}^\varepsilon,j(s) ds
= \frac{1}{\varepsilon} \sum_{k=1}^{Q} \int_{s_k}^{s'_k} q^\varepsilon,j(s) \dot{q}^\varepsilon,j(s) ds
= \frac{1}{2\varepsilon} \sum_{k=1}^{Q} (\|q^\varepsilon,j(s'_k)\|^2 - \|q^\varepsilon,j(s_k)\|^2)
\]

\((s_k, s'_k)\) — the time \( q_j(s) \) spends in \((-\bar{x}_j - \varepsilon, -\bar{x}_j + \varepsilon)\), \( q_j(s_k) = \bar{x}_j \pm \varepsilon, q_j(s'_k) = q_j(s_{k+1}) \), whence a telescoping sum, giving

\[
I = \frac{1}{2\varepsilon} (\|q^\varepsilon,j(s'_Q)\|^2 - \|q^\varepsilon,j(s_1)\|^2) \leq C
\]
Lemma 5. There exists a constant $C > 0$ so that for any $k \in \mathbb{N}$ there exists a time $s_k < C k^3$ such that $\| A q_\varepsilon (s_k) \|^2 + \| \dot{q}_\varepsilon (s_k) \|^2 \leq C / k$ for all $\varepsilon < \varepsilon_0$.

Corollary. For any $n \in \mathbb{N}$ there exists a time $s_n = s_n(\varepsilon) < C n^3$ such that $\| A q_\varepsilon (s) \|^2 + \| \dot{q}_\varepsilon (s) \|^2 \leq C / n$ for all $s > s_n$.

Proof. This is simply because $\| A q_\varepsilon (t) \|^2 + \| \dot{q}_\varepsilon (t) \|^2$ is decreasing in $t$ and Lemma 5.
Why the correct limit: for any $\delta_0 > 0$ for all $t > T_0$ and $\varepsilon < \varepsilon_0$

$$\|A^*z_\varepsilon(t) - \tau G_\varepsilon(\bar{x} + q_\varepsilon(t))\| \leq \delta_0, \|Aq_\varepsilon(t)\| \leq \delta_0$$

Hence: $$|(A^*z_\varepsilon(t) - \tau G_\varepsilon(\bar{x} + q_\varepsilon(t)) \cdot (\bar{x} + q_\varepsilon(t))| \leq \delta_0\|\bar{x} + q_\varepsilon(t)\|.$$

But: $$G_\varepsilon(\bar{x} + q_\varepsilon(t)) \cdot (\bar{x} + q_\varepsilon(t)) \approx \|\bar{x} + q_\varepsilon(t)\|_{l_1}$$

and $$|A^*z_\varepsilon(t) \cdot (\bar{x} + q_\varepsilon(t))| \approx |A^*z_\varepsilon(t) \cdot \bar{x}| \leq (\tau + \delta_0)\|\bar{x}\|_{l_1}$$

Therefore: $$\|\bar{x} + q_\varepsilon(t)\|_{l_1} - \|\bar{x}\|_{l_1} \leq \text{small}$$

Hence $(Aq_\varepsilon(t) \approx 0) q_\varepsilon(t)$ is close to 0.
Application to array imaging in a homogeneous medium

The array: $N$ receivers/transducers at $x_p$, $p = 1, \ldots, N$. $M$ point scatterers — unknown reflectivities $\rho_j > 0$, at unknown positions $y_{nj} \ (j = 1, \ldots, M)$. The image window with a uniform mesh $y_j, \ j = 1, \ldots, K$. Free space Green function $\tilde{G}_0(y, x, \omega) = \frac{\exp(-i\omega|x-y|/c)}{4\pi|x-y|}$. Born approximation: neglect multiple scattering.

The back-scattered field:

$$b_r(\omega) = \sum_{j=1}^{M} \rho_j \tilde{G}_0(x_r, y_{nj}, \omega) \tilde{G}_0(y_{nj}, x_s, \omega).$$
The linear system: relate the reflectivity $\rho_{0j}$ at each grid point $y_j$ of the IW ($j = 1, \ldots, K$) and the data $b_r(\omega)$ measured at the array ($r = 1, \ldots, N$).

Reflectivity $\rho_0 = (\rho_{01}, \rho_{02}, \ldots, \rho_{0K})^T \in \mathbb{R}^K$ — sparse, $M \ll K$.

The data vector $b(\omega) = (b_1, b_2, \ldots, b_N)^T \in \mathbb{R}^N$.

The linear system $A_\omega \rho_0 = b(\omega)$,

$A_\omega$ — $N \times K$ matrix

$N \ll K$ — an underdetermined linear system.

Array imaging is to find $\rho_0$. 
Numerical Examples

Array: 21 transducers, 5\(\lambda\) apart. Imaging window: 41 \(\times\) 41 grid points, distance 120\(\lambda\) away.

Left: true configuration, right: GeLMA reconstruction
Comparison between the exact solutions (blue circles) and the solutions with GeLMA (green crosses)
Left: true configuration, right: GeLMA reconstruction
Comparison between the exact solutions (blue circles) and the solutions with GeLMA (green crosses)
Left: true configuration, right: GeLMA reconstruction
Comparison between the exact solutions (blue circles) and the solutions with GeLMA (green crosses)
Rate of convergence as $\tau$ varies

$\|\rho - \rho_0\|_2$ vs. the iteration number, $\tau = \alpha \|A^T_\omega b(\omega)\|_\infty$: $\alpha = 2$ (solid line), $\alpha = 5$ (dashed line), $\alpha = 10$ (dot-dashed line), and $\alpha = 20$ (dotted line). Large $\tau$ speeds up convergence.
Comparison to FISTA

For ISTA or FISTA need $\tau < \| A_T^T b(\omega) \|_\infty$ – otherwise, they converge to $\rho = 0$. FISTA:

$$
\rho^{(k)} = \eta \tau \alpha_k (\xi^{(k)} - \alpha_k \nabla f(\xi^{(k)})),
$$

$$
\alpha_{k+1} = \frac{1 + \sqrt{1 + 4 \alpha_k^2}}{2},
$$

$$
\xi^{(k+1)} = \rho^{(k)} + \frac{\alpha_k - 1}{\alpha_{k+1}} (\rho^{(k)} - \rho^{(k-1)}).
$$
FISTA with $\tau = 0.01\|A^T_\omega b(\omega)\|_\infty$. Convergence rate is much slower than for GeLMA with $\tau = 20\|A^T_\omega b(\omega)\|_\infty$. FISTA does not obtain the exact solution. To achieve the exact solution, we would have to let $\tau \to 0$. 

39
Robustness to noise

Noisy data $b(\omega) + e(\omega)$. 

$e(\omega)$ – independent Gaussian random variables with zero mean, standard deviation $\beta \|b(\omega)\|_{\ell_2}/\sqrt{N}$. 

Four scatterers 

$\beta = 0.05$, $\beta = 0.1$, and $\beta = 0.3$. 

$\tau = \alpha \|A_\omega^T b(\omega)\|_{\infty}$, $\alpha = 2, 20, 100$. 
$\alpha = 2$. Noise; 5% (left), 10% (middle), 30% (right).
$\alpha = 20$. Noise; 5\% (left), 10\% (middle), 30\% (right).
\( \alpha = 200 \). Noise; 5\% (left), 10\% (middle), 30\% (right).

A bit better than \( \alpha = 20 \).
Imaging with correlations (with O. Pinaud)
Conclusions

1. The min-max variational principle gives an exact reconstruction.
2. GeLMA converges fast for various $\tau > 0$ – no serious limitation on the size of $\tau$.
3. ODE techniques provide an interesting (good for us) framework for the analysis
4. Empirically: fast convergence and robustness to noise
Open problems

1. Rigorous study of the convergence rates
2. Rigorous study of the robustness to noise
3. Effect of randomness in $A$ – connection to correlations and CINT-type methods