Separating Wheat from Chaff in Array Imaging

Thomas Strohmer
Department of Mathematics
University of California, Davis

BASP
September 5 2011

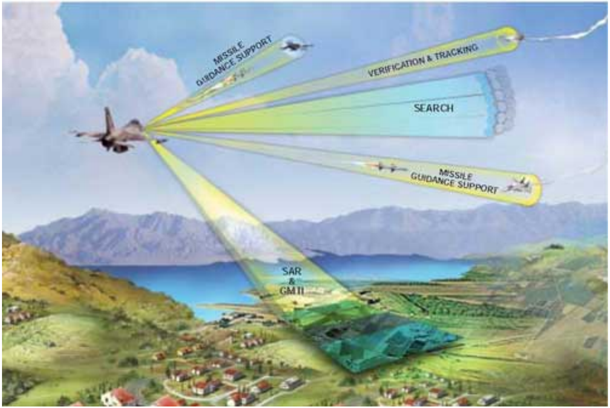
Acknowledgements

Research in collaboration with:
Albert Fannjiang, Benjamin Friedlander, Han Wang, Mike Ya.

This work is sponsored by the Defense Threat Reduction Agency, NSF-DMS, and DARPA.

Array imaging – target detection

Basic problem: Find exact location and velocity of targets by analyzing received (reflected) electromagnetic waves.



Radar imaging – challenges

Problem comes in many different flavors:

- active/passive targets
- clutter/no clutter
- extended targets/point targets (near-field/far-field)
- stationary/moving targets

Radar imaging – challenges

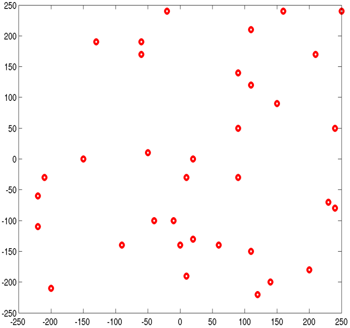
Problem comes in many different flavors:

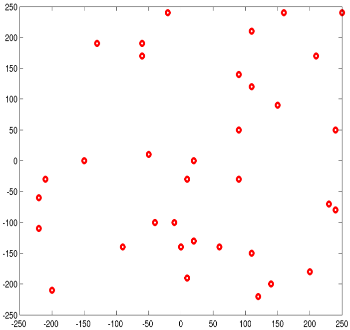
- active/passive targets
- clutter/no clutter
- extended targets/point targets (near-field/far-field)
- stationary/moving targets

Difficulty: Essentially any somewhat realistic setup leads to an underdetermined (linear or nonlinear) system of equations.

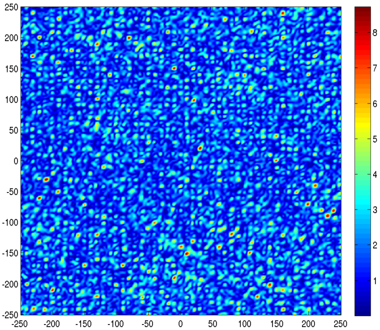
Limitations of existing radar imaging methods

- Methods work well for a “few” targets, but break down if number of targets is “large”
- Limited accuracy and resolution (spatial, delay/Doppler)
- Clutter affects target detection
- Huge computational complexity of more sophisticated (parameter-estimation based) methods
- Theory not well matched to practice (typical assumptions: only one target, number of antennas goes to infinity,...)
- Methods cannot detect “weak” targets (enormous dynamical range across reflection strength of targets)

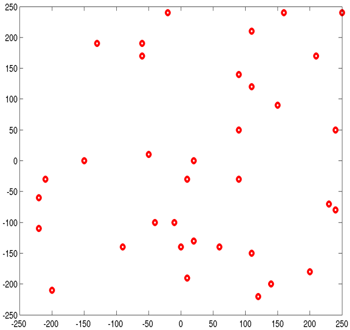

250
200
150
100
50
0
-50
-100
-150
-200
-250
-250 -200 -150 -100 -50 0 50 100 150 200 250



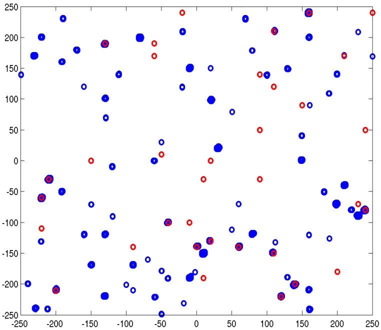
Target scene



Matched field solution



Target scene



Thresholded matched field

Sparsity in array imaging

Conceptual mistake: Existing methods exploit sparsity as as afterthought, an ad-hoc fix, instead of utilizing it in a mathematically rigorous way from the beginning on.

Sparsity in array imaging

Conceptual mistake: Existing methods exploit sparsity as as afterthought, an ad-hoc fix, instead of utilizing it in a mathematically rigorous way from the beginning on.

What is sparse here?
Targets occupy very small area compared to entire domain in which we search for targets (this assumes absence of clutter, otherwise additional linear transform of radar scene may be required). E.g., airborne radar, detection of submarines, ...

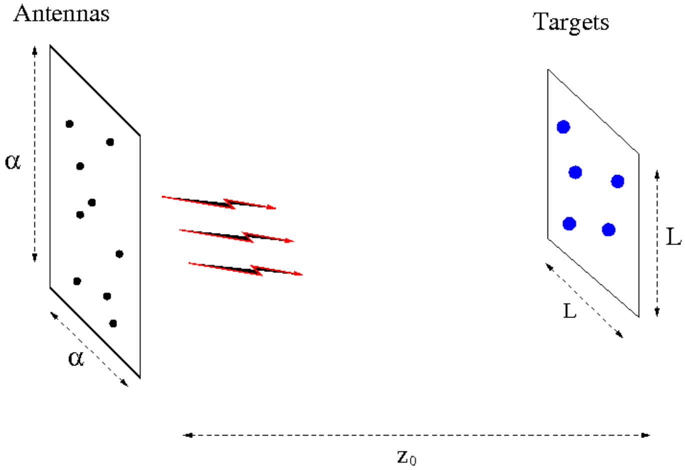
Sparsity in array imaging

Conceptual mistake: Existing methods exploit sparsity as as afterthought, an ad-hoc fix, instead of utilizing it in a mathematically rigorous way from the beginning on.

What is sparse here?
Targets occupy very small area compared to entire domain in which we search for targets (this assumes absence of clutter, otherwise additional linear transform of radar scene may be required). E.g., airborne radar, detection of submarines, ...

Better approach: Utilize ideas from compressive sensing to exploit sparsity in a rigorous and more systematic way.
Note: We do not want to undersample on purpose, i.e., we don't want to do *compressive sensing*.
We want to use insight from compressive sensing to better deal with ill-posed inverse problem.

Inverse Scattering Problem – Paraxial Regime



Assumption: $\alpha + L \ll z_0$

Helmholtz equation, Green's function, ...

The exact Green's function for Helmholtz's equation which governs wave propagation is

$$G(a, r) = \frac{e^{i\omega\|r-a\|_2}}{4\pi\|r - a\|_2}, \qquad a = (0, \xi, \eta), r = (z_0, x, y)$$

a: point in the sensor domain, r: point in the target domain.
Set phase speed $c = 1$, thus frequency ω = wavenumber.

Helmholtz equation, Green's function, ...

The exact Green's function for Helmholtz's equation which governs wave propagation is

$$G(a,r) = \frac{e^{i\omega\|r-a\|_2}}{4\pi\|r-a\|_2}, \qquad a = (0,\xi,\eta), r = (z_0,x,y)$$

a: point in the sensor domain, r: point in the target domain.
Set phase speed $c = 1$, thus frequency ω = wavenumber.

We assume that distance z_0 between targets and sensors satisfies $z_0 \gg \alpha + L, \quad z_0 \gg \lambda$ (Fresnel diffraction regime)
In this case we can use the paraxial Green's function

$$G(a,r) = \frac{e^{i\omega z_0}}{4\pi z_0} e^{i\omega|x-\xi|^2/(2z_0)} e^{i\omega|y-\eta|^2/(2z_0)}$$

Inverse scattering - Born approximation

Inverse scattering obeys the Lippmann-Schwinger equation and is thus intrinsically nonlinear due to multiple scattering. We use the standard Born approximation: Instead of

$$\tilde{G}(r, a_i) = G(r, a_i) + \sum_{l=1}^{s} \sigma_{j_l} G(r, r_{j_l}) \tilde{G}(r_{j_l}, a_i)$$

we consider

$$\tilde{G}(r, a_i) = G(r, a_i) + \sum_{l=1}^{s} \sigma_{j_l} G(r, r_{j_l}) G(r_{j_l}, a_i)$$

where $\tilde{G}(r, a_i)$ is the exciting field.
We obtain a linearized model ("single-bounce model").

Inverse scattering - Born approximation

Inverse scattering obeys the Lippmann-Schwinger equation and is thus intrinsically nonlinear due to multiple scattering. We use the standard Born approximation: Instead of

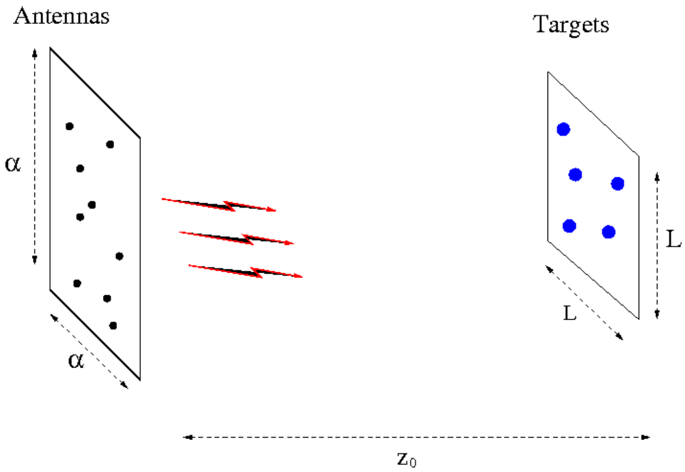
$$\tilde{G}(r, a_i) = G(r, a_i) + \sum_{l=1}^{s} \sigma_{j_l} G(r, r_{j_l}) \tilde{G}(r_{j_l}, a_i)$$

we consider

$$\tilde{G}(r, a_i) = G(r, a_i) + \sum_{l=1}^{s} \sigma_{j_l} G(r, r_{j_l}) G(r_{j_l}, a_i)$$

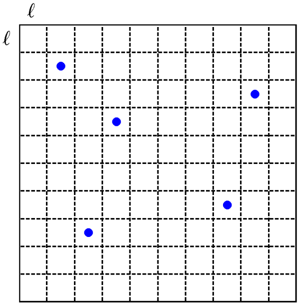
where $\tilde{G}(r, a_i)$ is the exciting field.
We obtain a linearized model (“single-bounce model”).

Inverse Scattering Problem – Paraxial Regime



Assumption: $\alpha + L \ll z_0$

We discretize target domain $[0, L] \times [0, L]$ by a rectangular grid with mesh-size ℓ and denote number of grid points by m.



Let σ_k be the amplitude of the k-th target and let the vector $\boldsymbol{x} \in \mathbb{C}^m$ be defined as

$$\boldsymbol{x}_k = \begin{cases} \sigma_k & \text{if } k\text{-th grid point is location of a target} \\ 0 & \text{else.} \end{cases}$$

Born approximation, Green's function

Due to the Born approximation we have linearized the relation between scatterers and scattered field.
The measurement vector $\boldsymbol{y}$ can be written as $\boldsymbol{Ax} = \boldsymbol{y}$ with $\boldsymbol{y} \in \mathbb{C}^{n^2}$, where n is the number of antennas.

Born approximation, Green's function

Due to the Born approximation we have linearized the relation between scatterers and scattered field.
The measurement vector $\boldsymbol{y}$ can be written as $\boldsymbol{Ax} = \boldsymbol{y}$ with $\boldsymbol{y} \in \mathbb{C}^{n^2}$, where n is the number of antennas.

The corresponding $n^2 \times m$ sensing matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A}_{lj} = G(a_i, r_j)G(r_j, a_k), \qquad l = i(n-1) + k,$$

where $G(a, r)$ is the paraxial Green's function

$$G(a, r) = \frac{e^{i\omega z_0}}{4\pi z_0} e^{i\omega|x-\xi|^2/(2z_0)} e^{i\omega|y-\eta|^2/(2z_0)}$$

Note: Unless we use crude discretization we have $n^2 \ll m$.
Hence the system is underdetermined!

Consider the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$, where $\boldsymbol{A}$ is a $k \times m$ matrix with full row-rank and $k \leq m$: System is underdetermined!
$\Rightarrow$ Infinitely many solutions, which one is the right one?

Sparse Recovery/Compressive Sensing

Consider the system $\boldsymbol{Ax} = \boldsymbol{y}$, where $\boldsymbol{A}$ is a $k \times m$ matrix with full row-rank and $k \leq m$: System is underdetermined!
$\Rightarrow$ Infinitely many solutions, which one is the right one?
Assume that $\boldsymbol{x} \in \mathbb{C}^m$ satisfies $\|\boldsymbol{x}\|_0 := |\operatorname{supp} x| \ll m$.
If $s := \|x\|_0 \leq k/2$, then in theory we can recover $\boldsymbol{x}$ by solving

$$\min \|\boldsymbol{x}\|_0 \quad \text{s.t.} \quad \boldsymbol{Ax} = \boldsymbol{y} \tag{L0}$$

Sparse Recovery/Compressive Sensing

Consider the system $\boldsymbol{Ax} = \boldsymbol{y}$, where $\boldsymbol{A}$ is a $k \times m$ matrix with full row-rank and $k \leq m$: System is underdetermined!
$\Rightarrow$ Infinitely many solutions, which one is the right one?
Assume that $\boldsymbol{x} \in \mathbb{C}^m$ satisfies $\|\boldsymbol{x}\|_0 := |\operatorname{supp} x| \ll m$.
If $s := \|x\|_0 \leq k/2$, then in theory we can recover $\boldsymbol{x}$ by solving

$$\min \|\boldsymbol{x}\|_0 \quad \text{s.t.} \quad \boldsymbol{Ax} = \boldsymbol{y} \tag{L0}$$

Instead of trying to solve this NP-hard problem we consider its convex relaxation (known as Basis Pursuit)

$$\min \|\boldsymbol{x}\|_1 \quad \text{s.t.} \quad \boldsymbol{Ax} = \boldsymbol{y} \tag{L1}$$

If x is sparse then the solution of (L1) is identical to the solution of (L0) under certain conditions on A: RIP, incoherence, ...

Conditions on matrix $\boldsymbol{A}$ for (L0)-(L1) equivalence

Coherence of $\boldsymbol{A}$: [Donoho-Huo, Tropp, ...]

$$\mu(\boldsymbol{A}) = \max_{k \neq \ell} \frac{|\langle \boldsymbol{A}_k, \boldsymbol{A}_\ell \rangle|}{\|\boldsymbol{A}_k\|_2 \|\boldsymbol{A}_\ell\|_2}, \qquad \text{where } \boldsymbol{A}_i \text{ is } i\text{-th column of } \boldsymbol{A}.$$

Restricted Isometry Property: [Candes-Tao] Any submatrix formed by r arbitrary columns of A is almost an isometry.

Conditions on matrix $\boldsymbol{A}$ for (L0)-(L1) equivalence

Coherence of $\boldsymbol{A}$: [Donoho-Huo, Tropp, ...]

$$\mu(\boldsymbol{A}) = \max_{k \neq \ell} \frac{|\langle \boldsymbol{A}_k, \boldsymbol{A}_\ell \rangle|}{\|\boldsymbol{A}_k\|_2 \|\boldsymbol{A}_\ell\|_2}, \qquad \text{where } \boldsymbol{A}_i \text{ is } i\text{-th column of } \boldsymbol{A}.$$

Restricted Isometry Property: [Candes-Tao] Any submatrix formed by r arbitrary columns of A is almost an isometry.

Examples for good matrices: Gaussian matrices, partial random Fourier matrices, equiangular tight frames, ...

Typical result for those matrices: if $s < \mathcal{O}(k/\ln m)$, then with high probability BP gives the sparsest solution.

Compressive sensing matrices in practice

Known good CS matrices: ... are either random matrices (Gaussian matrices, Bernoulli matrices, partial random Fourier matrices) or deterministic matrices (equiangular tight frames, mutually unbiased bases, ...).
Well developed theory for these cases.

Compressive sensing matrices in practice

Known good CS matrices: ... are either random matrices (Gaussian matrices, Bernoulli matrices, partial random Fourier matrices) or deterministic matrices (equiangular tight frames, mutually unbiased bases, ...).
Well developed theory for these cases.

Challenge: In many applications we cannot simply choose the matrix $\boldsymbol{A}$ as we please. Structure of $\boldsymbol{A}$ is governed (at least in part) by underlying physical process, such as the properties of wave propagation.

Is our A a good compressive sensing matrix?

What is under our control? Number of antennas, antenna locations, mesh size for discretization

Discretization, Regularization, Coherence

Questions of discretization are usually ignored in compressive sensing, but play a key role in many real world applications.

Two contradicting desires:

- Want as high resolution as possible: choose very small meshsize ℓ.
- Want matrix to have small coherence or to satisfy RIP with good bounds: choose large meshsize ℓ.

Discretization, Regularization, Coherence

Questions of discretization are usually ignored in compressive sensing, but play a key role in many real world applications.

Two contradicting desires:

- Want as high resolution as possible: choose very small meshsize ℓ.
- Want matrix to have small coherence or to satisfy RIP with good bounds: choose large meshsize ℓ.

Making meshsize small increases coherence and increases underdeterminedness of $\boldsymbol{Ax} = \boldsymbol{y}$.
Thus finding optimal discretization is crucial to make sparse recovery possible!

Compressive array imaging - theory

Theorem: [A.Fannjiang, M.Yan, T.S. '09].

Assume that the n sensors are randomly distributed and that

$$\frac{\ell\alpha}{\lambda z_0} \geq 1 \tag{1}$$

Let $s \leq \mathcal{O}(n^2/(\ln m)^2)$. Then with high probability s randomly distributed targets can be located exactly by Basis Pursuit.

The relation (1) indicates the existence of a threshold, an optimal aperture size $\alpha = \lambda z_0/\ell$, or equivalently, an optimal mesh size ℓ for the discretization of the target domain

$$\ell \geq \frac{\lambda z_0}{\alpha}$$

Proof-sketch: Based on coherence of matrix $\boldsymbol{A}$ and a theorem by J. Tropp. We need to show that

- $\boldsymbol{A}$ has full rank
- derive a good bound for $\|\boldsymbol{A}\|_2$
- obtain an estimate for the coherence μ of $\boldsymbol{A}$.

Coherence estimate for $\boldsymbol{A}$: (see also Rauhut and Kunis, 2008)
Assume $\frac{\ell\alpha}{\lambda z_0} \geq 1$ and $m \leq \delta e^{K^2/2}$ for some $\delta, K > 0$. Then

$$\mu(\boldsymbol{A}) \leq 2K^2/n$$

with probability greater than $(1-\delta)^2$.
Remarks:

- Coherence estimate is optimal w.r.t. n.
- $\mu(\boldsymbol{A})$ is roughly constant for $\alpha \geq \lambda z_0/\ell$
- Threshold for mesh-size $\ell = z_0\lambda/\alpha$. Choose ℓ too small, and BP will fail.

Resolution then and now ...

The resolution limit

$$\ell \geq z_0 \lambda / \alpha$$

is the classical Rayleigh limit

Does Sparse Remote Sensing have the same resolution limit as standard methods?

What have we gained, if anything?

Resolution then and now ...

The resolution limit

$$\ell \geq z_0\lambda/\alpha$$

is the classical Rayleigh limit

Does Sparse Remote Sensing have the same resolution limit as standard methods?

What have we gained, if anything?

A resolution limit does not mean that this resolution is actually achievable!

With compressive sensing approach we obtain guaranteed recovery (with high probability) of a substantial number of targets with comparably small number of antennas.

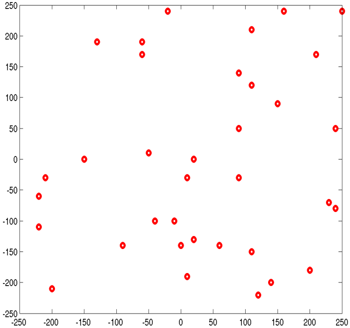
Inverse scattering, numerical example

20 antennas, $\alpha = 100, z_0 = 10000, \lambda = 0.1$, this implies $\ell = 10$ as optimal mesh size for discretization of target domain.
35 randomly distributed targets. Matrix $\boldsymbol{A}$ is of size 400×2500.
We use exact Green's function for computing actual wave propagation (i.e., for computing $\boldsymbol{y}$), but Born approximation and paraxial approximation for setting up the matrix $\boldsymbol{A}$.
That means we have twofold model mismatch:

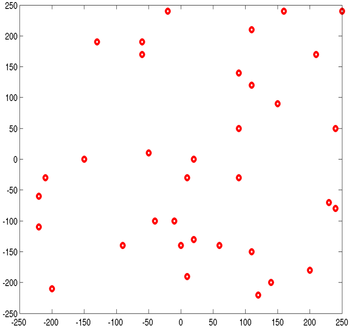
- We use Born approximation, but actual system of equations is nonlinear
- We use paraxial approximation, i.e., we have a linear perturbation of matrix $\boldsymbol{A}$

[M.Herman, T.S., '08] BP is robust in presence of linear matrix perturbation if matrix satisfies RIP (extends result by Candes).

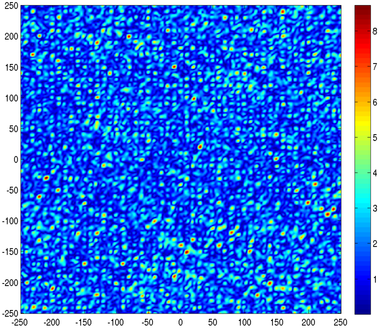
Inverse Scattering - numerical example



Inverse Scattering - numerical example

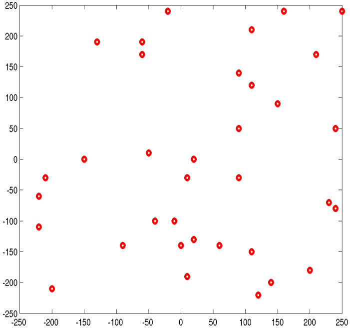


Target scene

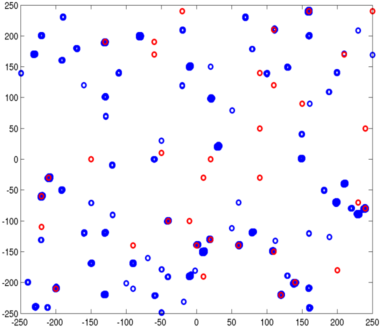


Matched field solution

Inverse Scattering - numerical example

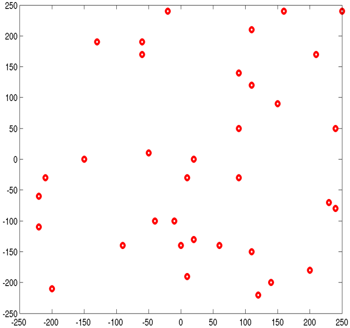


Target scene

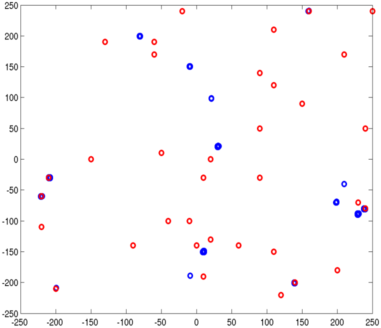


Thresholded matched field

Inverse Scattering - numerical example

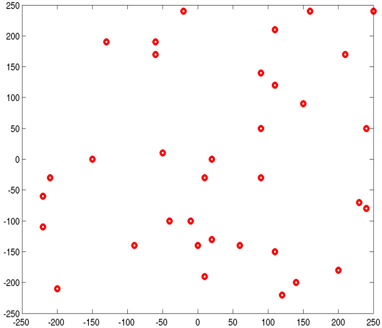


Target scene

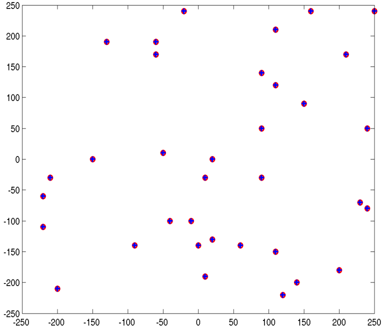


Thresholded matched field

Inverse Scattering - numerical example

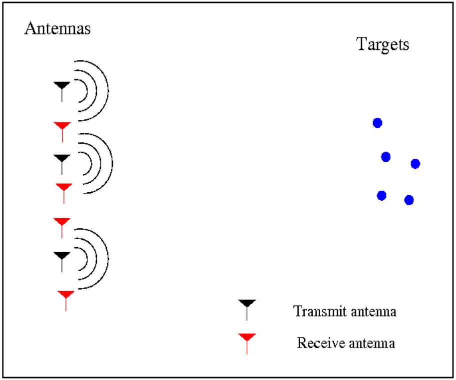


Target scene



Basis Pursuit

MIMO Radar - Signal model



- N_T transmit antennas, N_R receive antennas
- Co-located antennas (monostatic setup)
- Coherent propagation scenario
- k-th antenna sends signal s_k of bandwidth B and period T

Assume we take N_s samples of the received radar signal.
Let $\boldsymbol{Z}(t;\theta,r)$ denote the received $N_R \times N_s$ signal matrix from a unit-strength target at direction θ and range r. Then

$$\boldsymbol{Z}(t;\theta,r) = \boldsymbol{a}_R(\theta)\boldsymbol{a}_T^T(\theta)\boldsymbol{S}(t-\tau),$$

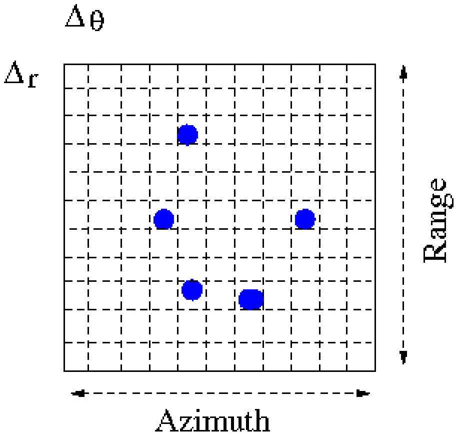
where $\boldsymbol{S}$ is an $N_T \times N$ matrix whose rows contain the circularly delayed signals $s_k(t-\tau), t = 1,\ldots,N$; and $\tau = 2r/c$ with c denoting the speed of light.
$\boldsymbol{a}_T(\theta)$ and $\boldsymbol{a}_R(\theta)$ are the transmit- and receive array manifolds, which for uniformly spaced linear arrays can be written as

$$\boldsymbol{a}_R(\theta) = \begin{bmatrix} 1 \\ e^{j2\pi d_R \sin\theta} \\ \vdots \\ e^{j2\pi d_R (N_R-1)\sin\theta} \end{bmatrix}, \quad \boldsymbol{a}_T(\theta) = \begin{bmatrix} 1 \\ e^{j2\pi d_T \sin\theta} \\ \vdots \\ e^{j2\pi d_T (N_T-1)\sin\theta} \end{bmatrix}$$

where d_R and d_T are the normalized spacings (distance divided by wavelength) between antenna elements.

From signal model to linear system of equations

We discretize range/azimuth domain with step-sizes Δ_r, Δ_θ and obtain a range/azimuth grid (θ_i, r_j), $1 \leq i \leq N_\theta, 1 \leq j \leq N_r$. Here, N_r, N_θ denote the number of grid points in each axis.



From signal model to linear system of equations

- We construct the response matrix $\boldsymbol{A}$, whose columns are the vectors $\boldsymbol{z}(t; \theta_i, r_j) := \text{vec}\{\boldsymbol{Z}(t; \theta_i, r_j)\}$. Each $\boldsymbol{z}$ has length $N_R N$, hence $\boldsymbol{A}$ is an $N_R N_s \times N_\theta N_r$ matrix.
- Assume the radar scene consists of s scatterers located on s points of the (θ_i, r_j)-grid. Let $\boldsymbol{x}$ be the $N_\theta N_r \times 1$ vector, whose non-zero elements are the amplitudes of the scatterers. That means $\boldsymbol{x}$ has s non-zero elements (but we do not know their location!).
- The received radar signal $\boldsymbol{y}$ is now given by

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{v},$$

  where $\boldsymbol{v}$ is Gaussian noise with variance σ.
- Note: Unless we use crude discretization we have $N_R N_s < N_\theta N_r$. Hence the system is underdetermined

Non-stationary radar scene – Doppler effect

In presence of Doppler shift f_d, we need to replace $\boldsymbol{Z}(t;\theta,r)$ by

$$\boldsymbol{Z}(t;\theta,r,f_d) = \boldsymbol{a}_R(\theta)\boldsymbol{a}_T^T(\theta)\boldsymbol{S}(t-\tau,f_d),$$

where the entries of $\boldsymbol{S}$ are the circularly delayed and Doppler shifted signals $s_k(t-\tau)e^{j2\pi f_d t}$.
Discretizing the “Doppler domain” with N_f grid points and setting up the response matrix $\boldsymbol{A}$ analogously to before, we obtain the system of equations

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{v},$$

where $\boldsymbol{A}$ is now an $N_R N_s \times N_\theta N_r N_f$ matrix.
Thus the system is even more underdetermined than before.

Setup

Waveforms: s_k is a periodic, continuous-time white-noise signal of duration T seconds, filtered by an ideal lowpass filter with cutoff frequency B Hertz.

Antennas: Let $d_T = \frac{N_R}{2}, d_R = \frac{1}{2}$ (or $d_T = \frac{1}{2} d_R = \frac{N_T}{2}$).

Discretization:

Azimuth is discretized as $\beta = n\Delta_\beta$ where $\Delta_\beta = \frac{2}{N_R N_T}$, $n = -\frac{N_R N_T}{2}, \ldots, \frac{N_R N_T - 1}{2}$ and $\beta = \sin\theta$.

Range is discretized as $\tau = m\Delta_\tau$ where $\Delta_\tau = \frac{1}{2B}$, $m = 0, \ldots, N_s - 1$.

Generic sparse scatterer model: Location of the S scatterers is selected uniformly at random, amplitudes of scatterers have random phases

LASSO:

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1 .$$

Theorem (no Doppler): [B.Friedlander, T.S, '11].

Assume that **x** is drawn from the generic S-sparse scatterer model with

$$S \leq \frac{c_0 N_r N_R}{4 \log(N_r N_R N_T)} \tag{1}$$

for some constant $c_0 > 0$. Furthermore, suppose that

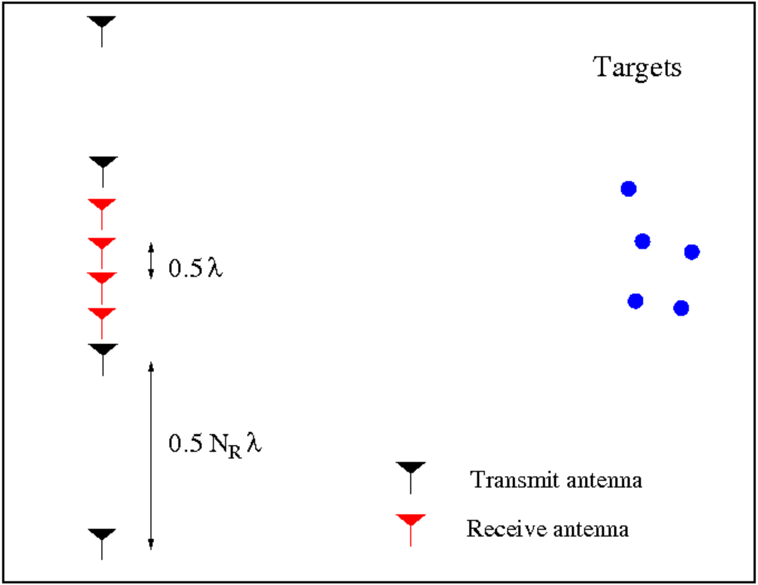
$$\log^3(N_r N_R N_T) \leq N_s. \tag{2}$$

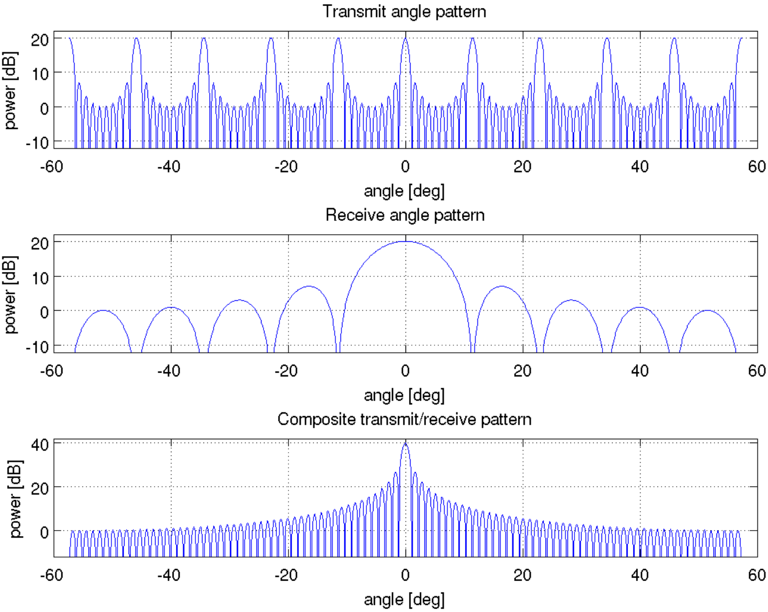
If

$$\min_k |\mathbf{x}_k| > (1+\varepsilon)\sigma\sqrt{2\log N_r N_R N_T}, \tag{3}$$

then with high probability the Lasso estimate computed with $\lambda = 2\sqrt{2\log(N_r N_R N_T)}$ obeys

$$\operatorname{supp}(\hat{\mathbf{x}}) = \operatorname{supp}(\mathbf{x}), \quad \text{and} \quad \frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \frac{3\sigma\sqrt{N_r N_R N_T}}{\|\mathbf{y}\|_2}$$


Targets
0.5 λ
0.5 N_R λ
Transmit antenna
Receive antenna


Transmit angle pattern
power [dB]
20
10
0
-10
-60
-40
-20
0
20
40
60
angle [deg]
Receive angle pattern
power [dB]
20
10
0
-10
-60
-40
-20
0
20
40
60
angle [deg]
Composite transmit/receive pattern
power [dB]
40
20
0
-60
-40
-20
0
20
40
60
angle [deg]

Proof-sketch:

Proof is based on careful analysis of structure of $\boldsymbol{A}$ and a theorem by Candes-Plan

Key steps:

- Need bound on $\|\boldsymbol{A}\|_{\mathrm{op}}$.
- Need bound on coherence $\mu(\boldsymbol{A})$.

Difficulty: $\boldsymbol{A}$ is a both a random and a deterministic matrix.

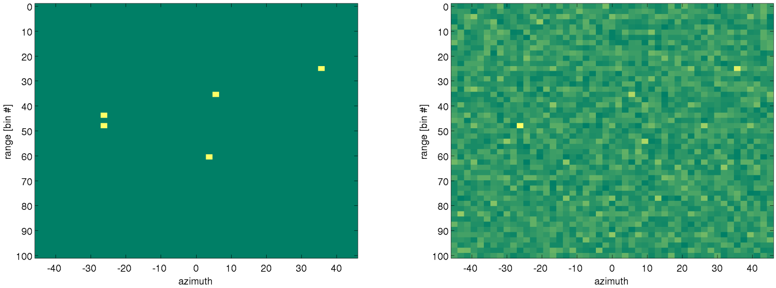
Key tools:

- Under the right conditions $\boldsymbol{A}\boldsymbol{A}^*$ is a block-Toeplitz matrix with circulant blocks (but $\boldsymbol{A}$ is not!)
- Incoherence of $\boldsymbol{S}$ comes into play
- Use bounds for quadratic forms (a'la Wright-Hanson)
- Concentration of measure
- Exploit specific choice for transmit/receive antenna spacing

Optimality of estimates

- Bounds on norm and coherence are optimal (up to small constants and probability)
- Coherence: $\mu(\boldsymbol{A}) \leq 2\sqrt{\frac{1}{N_s}\log(N_r N_R N_T)}$. Why does $\mu(\boldsymbol{A})$ only scale with N_s and not with the number of rows, $N_R N_s$? Comes from "decoupling": $\boldsymbol{A}_{\tau,\beta} = \boldsymbol{a}_R \otimes (\boldsymbol{S}_\tau \boldsymbol{a}_T)$
- Randomness on target locations can likely be removed (a'la Rauhut ...)

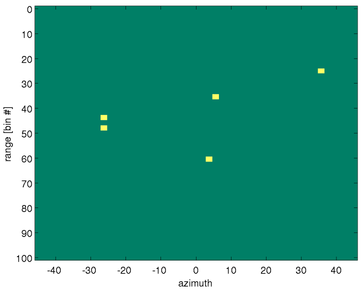
MIMO radar - simulations

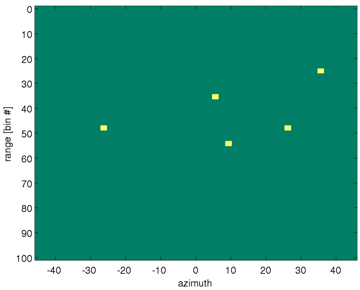


5 targets, target strengths range from 0.1 to 10

“Reconstruction” via matched filter (spectrogram) misses targets

MIMO radar - simulations

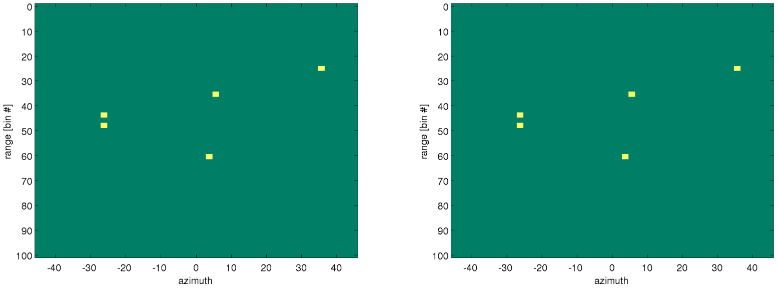




5 targets, target strengths range from 0.1 to 10

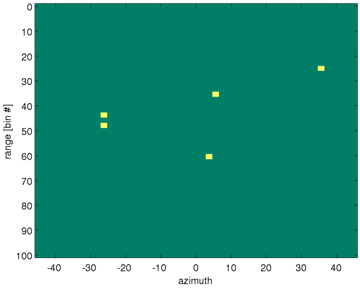
“Reconstruction” via matched filter (spectrogram) misses targets

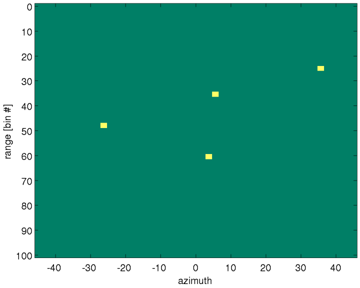
MIMO radar - simulations



Reconstruction via Lasso: Noise level is such that noise condition of theorem is satisfied by all targets

MIMO radar - simulations



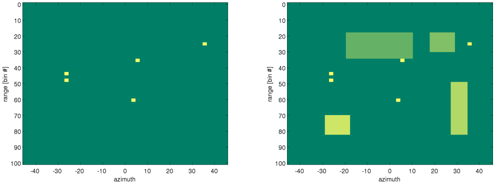


Reconstruction via Lasso: Noise level is such that noise condition of theorem is satisfied by four strongest targets

Clutter - separating wheat from chaff

- In many imaging scenarios the object of interest (target) is surrounded by objects we don't care about (clutter).
- Can we automatically separate clutter from targets?
- We need some model for clutter. E.g. clutter is fairly stationary compared to moving targets

Clutter and radar



- Assuming some model for the clutter can we improve target detection by automatically separating targets from clutter?
- We measure $y = Ax$, where x contains targets as well as clutter. Clutter has the effect that x is no longer sparse, thus standard sparse recovery techniques will fail!
- Want framework that can deal with rather vague prior information about clutter.

Adapt Robust PCA ideas to clutter separation

Compressive Completion

- Assume we have a sequence of observations $y_1, y_2, \ldots, y_n$, where $y_k = Ax_k$. Write x_k as k-th column of the matrix X and let $y := [y_1; \ldots y_n]$. Then $AX = y$.
- In radar this system will be highly underdetermined, but X will not be sparse in presence of clutter!
  Example: STAP radar.
- Assume that $X = L + S$, where the matrix L represents the static background (clutter) and S encodes the innovation (moving targets). L will be a low-rank matrix and S a sparse matrix.

Compressive Completion

- Assume we have a sequence of observations $y_1, y_2, \ldots, y_n$, where $y_k = Ax_k$. Write x_k as k-th column of the matrix X and let $y := [y_1; \ldots y_n]$. Then $AX = y$.
- In radar this system will be highly underdetermined, but X will not be sparse in presence of clutter!
  Example: STAP radar.
- Assume that $X = L + S$, where the matrix L represents the static background (clutter) and S encodes the innovation (moving targets). L will be a low-rank matrix and S a sparse matrix.

To recover S from y we propose to solve

$$\min \ \|L\|_* + \lambda\|S\|_1 \quad \text{s.t. } \|A(L+S) - y\|_2 \leq \varepsilon$$

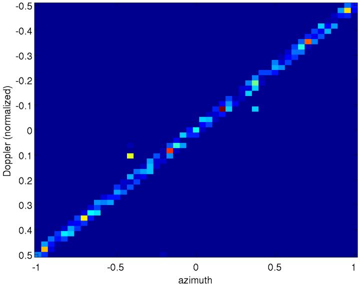
Nuclear norm $\|\cdot\|_*$ serves as proxy for minimizing rank of L
1-norm $\|\cdot\|_1$ serves as proxy for sparsity of S.

Compressive Completion

- Approach combines tools from compressive sensing with tools from matrix completion $\rightarrow$ Compressive Completion
- Approach inspired by work of Chandrasekaran et al. and Candes et al..
- Key difference: They observe entire matrix X, but we observe only small number of linear measurements of X, which makes it way more challenging.
- We cannot first recover X from y and then separate X into L and S, since $AX = y$ is underdetermined and X is not sparse!
- Initial theoretical results: Can prove that under certain conditions compressive completion gives exact result.

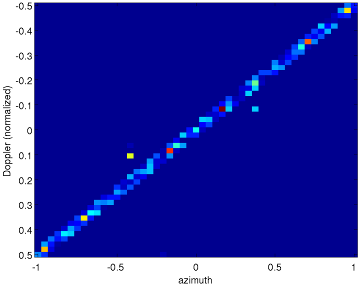
Initial numerical simulations

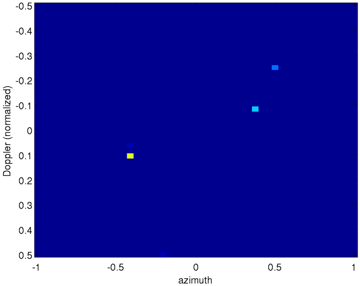
Stylized STAP example:



Initial numerical simulations

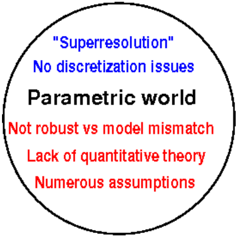
Stylized STAP example:





Standard compressive sensing recovery methods would fail miserably, only carefully exploiting a priori information about clutter makes success possible!
Would need a very large number of antennas to achieve similar results with standard techniques!

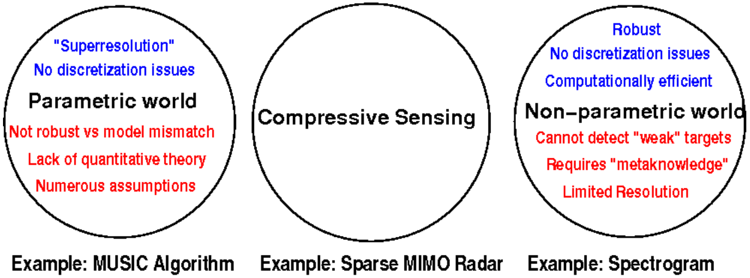
Parametric vs Nonparametric World



Example: MUSIC Algorithm

Example: Spectrogram

Parametric vs Nonparametric World



Caught between two worlds

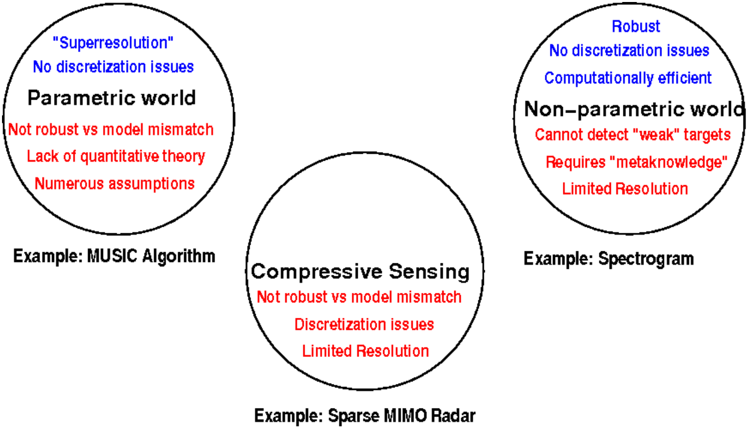
Parametric world:
Maximum Likelihood

$$\min f(y; x_1, \ldots, x_s)$$

Non-Parametric world:
Spectrogram + “Thresholding”

$$A^* y$$

Parametric vs Nonparametric World



Parametric vs Nonparametric World