Some Theoretical Results for Compressed MIMO Radar

Thomas Strohmer*
Dept. Mathematics
UC Davis
Davis CA 95616

Benjamin Friedlander †
Dept. Elec. Eng.
UC Santa Cruz
Santa Cruz, CA 95064

Abstract

We consider a MIMO radar system and derive a theoretical framework for the recoverability of targets in the azimuth-range-Doppler domain via compressive sensing type recovery algorithms. In particular we prove bounds on the achievable resolution and number of detectable targets in the presence of additive noise. Furthermore our theory reveals that even weak targets can be recovered reliably with the proposed approach.

1 Introduction

In recent years there has been considerable interest in a class of radar systems called *MIMO Radar*. The literature on the subject addresses two distinct types of radar systems which we will refer to as *MIMO radar with co-located antennas* [1] and *MIMO radar with widely separated antennas*, [2], which is also referred to as *statistical MIMO*. In this paper we consider only MIMO radars with co-located antennas and all subsequent mentions of MIMO radar refer to this type only.

MIMO radar is characterized by using multiple antennas to simultaneously transmit diverse, usually orthogonal, waveforms in addition to using multiple antennas to receive the reflected signals. MIMO radar has the potential for enhancing spatial resolution and improving interference and jamming suppression. The ability of MIMO radar to shape the transmit beam post facto allows for adapting the transmission based on the received data in a way which is not possible in "conventional" radar.

Most radar scenes are sparse in the sense that only a small fraction of the range-azimuth or range-Doppler-azimuth cells are occupied by objects of interest. In fact in most situations this fraction is very small indeed. This sparsity assumption suggests to approach the MIMO radar problem using the framework of compressed sensing [3]. In this paper we develop some fundamental results about the feasibility of recovering sparse radar scenes using algorithms based on compressed sensing. In particular we prove bounds on the achievable resolution and number of detectable targets in the presence of additive noise. We are able to show that targets can be detected at signal-to-noise ratios comparable to that of matched filter detectors. This is a continuation of our analysis begun in [4], where we investigated the Doppler-free and noiseless case.

*T.S. was supported by the National Science Foundation under grant DTRA-DMS 1042939 and from DARPA via grant N66001-11-1-4090.

†B.F. was supported by the National Science Foundation under grant CCF-0725366.

Notation

Let $\mathbf{v} \in \mathbb{C}^n$. As usual, we define $\|\mathbf{v}\|_1 := \sum_{k=1}^n |\mathbf{v}_k|$ and $\|\mathbf{v}\|_2 := \sqrt{\sum_{k=1}^n |\mathbf{v}_k|^2}$. The operator norm of a matrix $\mathbf{A}$ is the largest singular value of $\mathbf{A}$ and is denoted by $\|\mathbf{A}\|_{\text{op}}$. The coherence of $\mathbf{A}$ is defined as

$$\mu(\mathbf{A}) := \max_{k \neq l} \frac{|\langle \mathbf{A}_k, \mathbf{A}_l \rangle|}{\|\mathbf{A}_k\|_2 \|\mathbf{A}_l\|_2}, \tag{1}$$

where $\mathbf{A}_k$ is the k-th column of $\mathbf{A}$.

2 The Signal Model

Consider a MIMO radar employing N_T antennas at the transmitter and N_R antennas at the receiver. We assume that the element spacing is sufficiently small so that the radar return from a given scatterer is fully correlated across the array. In other words, this is a coherent propagation scenario.

To simplify the presentation we assume that the two arrays are co-located, i.e. this is a mono-static radar. The extension to the bi-static case is straightforward as long as the coherency assumption holds for each array. The arrays are characterized by the array manifolds:

$\mathbf{a}_R(\beta)$ for the receive array and $\mathbf{a}_T(\beta)$ for the transmit array, where $\beta = \sin(\theta)$ is the direction relative to the array. We assume that the arrays and all the scatterers are in the same 2-D plane. The extension to the 3-D case is straightforward and all of the following results hold for that case as well.

For convenience we formulate our theorems and analysis in terms of delay τ instead of range r. This is no loss of generality, as delay and range are related by $\tau = 2r/c$, with c denoting the speed of light.

2.1 The model for the azimuth-delay domain

The i-th transmit antenna repeatedly transmits the signal $s_i(t)$. Let $\mathbf{Z}(t;\beta,\tau)$ be the $N_R \times N_t$ noise free received signal matrix from a unit strength target at direction β and delay τ, where N_t is the number of samples in time. Then

$$\mathbf{Z}(t;\beta,\tau) = \mathbf{a}_R(\beta)\mathbf{a}_T^T(\beta)\mathbf{S}_\tau^*, \tag{2}$$

where $\mathbf{S}_\tau$ is an $N_t \times N_T$ matrix whose columns are the circularly delayed signals $s_i(t-\tau)$, sampled at the discrete time points $t = n\Delta_t, n = 1,\dots,N_t$. If $\tau = 0$, we often write simply $\mathbf{S}$ instead of $\mathbf{S}_0$.

Assuming uniformly spaced linear arrays (ULA), the array manifolds are given by

$$\mathbf{a}_T(\beta) = \begin{bmatrix} 1 \\ e^{j2\pi d_T \beta} \\ \vdots \\ e^{j2\pi d_T \beta (N_T - 1)} \end{bmatrix} \tag{3}$$

and

$$\mathbf{a}_R(\beta) = \begin{bmatrix} 1 \\ e^{j2\pi d_R \beta} \\ \vdots \\ e^{j2\pi d_R \beta (N_R - 1)} \end{bmatrix} \tag{4}$$

where d_T and d_R are the normalized spacings (distance divided by wavelength) between the elements of the transmit and receive arrays, respectively.

The spatial characteristics of a MIMO radar are closely related to that of a virtual array with $N_T N_R$ antennas, whose array manifold is $\mathbf{a}(\beta) = \mathbf{a}_T(\beta) \otimes \mathbf{a}_R(\beta)$. It is known [5] that the following choices for the spacing of the transmit and receive array spacing will yield a uniformly spaced virtual array with half wavelength spacing:

$$d_R = 0.5, d_T = 0.5N_R; \tag{5}$$

$$d_T = 0.5, d_R = 0.5N_T. \tag{6}$$

Both of these choices lead to a virtual array whose aperture is $0.5(N_T N_R - 1)$ wavelengths. This is the largest virtual aperture free of grating lobes. The choices (5) and (6) will also show up in our theoretical analysis, see Theorem 1.

Next let $\mathbf{z}(t;\beta,\tau) = \text{vec}\{\mathbf{Z}(t;\beta,\tau)\}$ be the noise-free vectorized received signal. We set up a discrete delay-azimuth grid $\{(\beta_i,\tau_j)\}, 1 \le i \le N_\beta, 1 \le j \le N_\tau$, where Δ_β and Δ_τ denote the corresponding discretization stepsizes. Using vectors $\mathbf{z}(t;\beta_i,\tau_j)$ for all grid points (β_i,τ_j) we construct a complete response matrix $\mathbf{A}$ whose columns are $\mathbf{z}(t;\beta_i,\tau_j)$ for $1 \le i \le N_\beta$ and $1 \le j \le N_\tau$. In other words, we have N_τ delay values and N_β azimuth values, so that $\mathbf{A}$ is a $N_R N_t \times N_\tau N_\beta$ matrix.

Assume that the radar illuminates a scene consisting of K scatterers located on K points of the (β_i,τ_j) grid. Let $\mathbf{x}$ be a sparse vector whose non-zero elements are the complex amplitudes of the scatterers in the scene. The zero elements corresponds to grid points which are not occupied by scatterers. We can then define the radar signal received from this scene $\mathbf{y}$ by

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v} \tag{7}$$

where $\mathbf{y}$ is a $N_R N_t \times 1$ vector, $\mathbf{x}$ is a $N_\tau N_\beta \times 1$ sparse vector, $\mathbf{v}$ is a $N_R N_t \times 1$ complex Gaussian noise vector, and $\mathbf{A}$ is a $N_R N_t \times N_\tau N_\beta$ matrix.

2.2 The model for the azimuth-delay-Doppler domain

The discussion so far was for the case of a stationary radar scene and a fixed radar, in which case there is no Doppler shift. The extension of this signal model to include the Doppler effect is conceptually straightforward, but leads to a significant increase in the problem dimension.

The signal model for the return from a unit strength scatterer at direction β, delay τ, and Doppler f (corresponding to its radial velocity with respect to the radar) is given by

$$\mathbf{Z}(t;\beta,\tau,f) = \mathbf{a}_R(\beta)\mathbf{a}_T^T(\beta)\mathbf{S}_{\tau,f}^*, \tag{8}$$

where $\mathbf{S}_{\tau,f}$ is a $N_t \times N_T$ matrix whose columns are the circularly delayed and Doppler shifted signals $s_i(t-\tau)e^{j2\pi f t}$.

As before we let $\mathbf{z}(t;\beta,\tau,f) = \text{vec}\{\mathbf{Z}(t;\beta,\tau,f)\}$ be the noise-free vectorized received signal. We extend the discrete delay-azimuth grid by adding a discretized Doppler component (with stepsize Δ_f and corresponding Doppler values $f = k\Delta_f, k = 1,\dots,N_f$) and obtain a uniform delay-azimuth-Doppler grid $\{(\beta_i,\tau_j,f_k)\}$. Using vectors $\mathbf{z}(t;\beta_i,\tau_j,f_k)$ for all discrete (β_i,τ_j,f_k) we construct a complete response matrix $\mathbf{A}$ whose columns are $\mathbf{z}(t;\beta_i,\tau_j,f_k)$ for $1 \le i \le N_\beta$, $1 \le j \le N_\tau$, $1 \le k \le N_f$.

Assume that the radar illuminates a scene consisting of K scatterers located on K points of the (β_i,τ_j,f_k) grid. Let $\mathbf{x}$ be a sparse vector whose non-zero elements are the complex amplitudes of the scatterers in the scene. The zero elements corresponds to grid points which are not occupied by scatterers. We can then define the radar signal received from this scene $\mathbf{y}$ by $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$, where $\mathbf{y}$ is a $N_R N_t \times 1$ vector, $\mathbf{x}$ is a $N_\tau N_\beta N_f \times 1$ sparse vector, $\mathbf{v}$ is a $N_R N_t \times 1$ complex Gaussian noise vector, and $\mathbf{A}$ is a $N_R N_t \times N_\tau N_\beta N_f$ matrix.

2.3 The scattering model

We introduce the following *generic K-sparse scatterer model*:

- The support $I \subset \{1,\dots,N_\tau N_\beta\}$ (or $I \subset \{1,\dots,N_\tau N_f N_\beta\}$ for the Doppler case) of the K nonzero coefficients of $\mathbf{x}$ is selected uniformly at random.
- The phases of $\mathbf{x}$ form a Steinhaus sequence, i.e., they are random and uniformly distributed on the unit circle.

We do not impose any condition on the amplitudes of the non-zero entries of $\mathbf{x}$. We do assume however that the targets are exactly located at the discretized grid points. This is certainly an idealized assumption, that is not satisfied in this strict sense in practice, resulting in a "gridding error". We refer the reader to [6, 7] for an analysis of the associated perturbation error.

2.4 The recovery algorithm – Debiased Lasso

A standard approach to finding a sparse (and under appropriate conditions *the sparsest*) solution to a noisy system $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$ is via

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{x}\|_1, \qquad (9)$$

which is also known as lasso [8]. Here $\lambda > 0$ is a regularization parameter.

In this paper we adopt the following two-step version of lasso. In the first step we compute an estimate $\tilde{K}$ for the support of $\mathbf{x}$ by solving (9). In the second step we estimate the amplitudes of $\mathbf{x}$ by solving the reduced-size least squares problem $\min \|\mathbf{A}_{\tilde{K}}\mathbf{x}_{\tilde{K}} - \mathbf{y}\|_2$, where $\mathbf{A}_{\tilde{K}}$ is the submatrix of $\mathbf{A}$ consisting of the columns corresponding to the index set $\tilde{K}$, and similarly for $\mathbf{x}_{\tilde{K}}$. This is a standard way to "debias" the solution, we thus will call this approach in the sequel *debiased lasso*.

3 Performance bounds for sparse MIMO radar

We assume that $s_i(t)$ is a periodic, complex-valued, continuous-time white Gaussian noise signal of period-duration T seconds and bandwidth B. The transmit waveforms are normalized so that the total transmit power is fixed, independent of the number of transmit antennas. Thus, we assume that the entries of $s_i(t)$ have variance $\frac{1}{N_T}$. It is convenient to introduce the finite-length vector $\mathbf{s}_i$ associated with s_i, via $\mathbf{s}_i(l) := s_i(l\Delta_t), l = 1,\dots,N_t$, where $\Delta_t = \frac{1}{2B}$ and $N_t = T/\Delta_t$.

Theorem 1 *Consider $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$, where $\mathbf{A}$ is as defined in Subsection 2.1 and $\mathbf{v}_i \in \mathcal{CN}(0,\sigma^2)$. Choose the discretization stepsizes to be $\Delta_\beta = \frac{2}{N_R N_T}$ and $\Delta_\tau = \frac{1}{2B}$. Let $d_T = 1/2, d_R = N_T/2$ or $d_T = N_R/2, d_R = 1/2$, and suppose that*

$$N_t \ge 128 \qquad \text{and} \qquad \big(\log(N_\tau N_\beta)\big)^3 \le N_t, \qquad (10)$$

If $\mathbf{x}$ is drawn from the generic K-sparse scatterer model with

$$K \le \frac{c_0 N_\tau N_R}{3\log(N_\tau N_\beta)} \qquad (11)$$

for some constant $c_0 > 0$, and if

$$\min_{k\in S} |\mathbf{x}_k| > \frac{10\sigma}{\sqrt{N_R N_t}}\sqrt{2\log N_\tau N_\beta}, \qquad (12)$$

then the solution $\tilde{\mathbf{x}}$ of the debiased lasso computed with $\lambda = 2\sigma\sqrt{2\log(N_\tau N_\beta)}$ obeys with high probability

$$\operatorname{supp}(\tilde{\mathbf{x}}) = \operatorname{supp}(\mathbf{x}), \qquad (13)$$

and

$$\frac{\|\tilde{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le \frac{6\sigma\sqrt{N_t N_R}}{\|\mathbf{y}\|_2}. \qquad (14)$$

Remarks:

- The notion "with high probability" in the above theorem (as well as in the other theorems and lemmata below) can be quantified explicitly, but the resulting explicit probabilities are somewhat lengthy, therefore we leave their precise form for the journal version of this paper.
- The assumptions in (10) are fairly mild and easy to satisfy in practice.
- We emphasize that at least in theory there is no constraint on the dynamical range of the target amplitudes. The lasso estimate will recover all target locations correctly as long as they exceed the noise level (12), regardless of the dynamical range between the targets.

- Note that equation (12) can be written as $\mathrm{SNR}_k > 100\frac{2\log N_\tau N_\beta}{N_R N_t}$ where $\mathrm{SNR}_k = |x_k|^2/\sigma^2$ is the input signal-to-noise ratio for the k-th scatterer.
- As noted in [9], one can replace the factor 10 in (12) by a factor $(1+\varepsilon)$ for some $\varepsilon > 0$, at the cost of a somewhat reduced probability of success and slightly stronger conditions on the coherence and sparsity. This means the factor 100 in the remark above can be replaced by a much smaller value.

The proof of the above theorem is rather involved and too long to be included in this brief paper. The full proof of this theorem, as well as other results presented in this paper can be found in the journal version of this paper. Here, we can only sketch the key steps. We rely on the following theorem which is a slight generalization of a theorem by Candès and Plan [9].

Theorem 2 *Given $\mathbf{y} = \mathbf{\Psi}\mathbf{x}+\mathbf{v}$, where $\mathbf{\Psi} \in \mathbb{C}^{n\times m}$ has all unit-ℓ_2-norm columns, $\mathbf{x}$ is drawn from the generic K-sparse model and $\mathbf{v}_i \sim \mathcal{CN}(0,\sigma^2)$. Assume that*

$$\mu(\mathbf{\Psi}) \le \frac{C_0}{\log m}, \tag{15}$$

where $C_0 > 0$ is a constant independent of n, m. Furthermore, suppose

$$K \le \frac{c_0 m}{\|\mathbf{\Psi}\|_{op}^2 \log m} \tag{16}$$

for some constant $c_0 > 0$ and that

$$\min_{k\in S} |\mathbf{x}_k| > 8\sigma\sqrt{2\log m}. \tag{17}$$

Then the solution $\hat{\mathbf{x}}$ to the debiased lasso computed with $\lambda = 2\sigma\sqrt{2\log m}$ obeys

$$\mathrm{supp}(\hat{\mathbf{x}}) = \mathrm{supp}(\mathbf{x}), \tag{18}$$

and

$$\frac{\|\hat{\mathbf{x}}-\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le \frac{3\sigma\sqrt{n}}{\|\mathbf{y}\|_2} \tag{19}$$

with high probability.

In order to apply Theorem 2 we need two key lemmata, one concerns a bound for the operator norm of $\mathbf{A}$, the other one concerns a bound for the coherence of $\mathbf{A}$. In many cases (such as for Gaussian random matrices or tight frames) it is fairly simple to bound $\|\mathbf{A}\|_{\mathrm{op}}$. In our case this is less straightforward. Indeed, the proof of the lemma below relies on a careful exploitation of the structure of the matrix $\mathbf{A}^*\mathbf{A}$ and various large deviation inequalities for random matrices.

Lemma 3 *Let $\mathbf{A}$ be as defined in Theorem 1. Then with high probability there holds*

$$\|\mathbf{A}\|_{op}^2 \ge N_t N_R N_T (1+\log N_t), \tag{20}$$

where $C > 0$ is some numerical constant.

Next we estimate the coherence of $\mathbf{A}$. Since the columns of $\mathbf{A}$ do not all have the same norm, we will proceed in two steps. First, in Lemma 4 we bound the modulus of the inner product of any two columns of $\mathbf{A}$ and then use this result to bound the coherence of a properly normalized version of $\mathbf{A}$ in Lemma 5. Since the columns of $\mathbf{A}$ depend on azimuth and delay, we index them via the double-index (τ,β). Thus the (τ,β)-th column of $\mathbf{A}$ is $\mathbf{A}_{\tau,\beta}$.

Lemma 4 *Let $\mathbf{A}$ be as defined in Theorem 1. Assume that*

$$\log(N_\tau N_\beta) \le \frac{N_t}{23}, \tag{21}$$

then with high probability there holds

$$\max_{(\tau,\beta)\neq(\tau',\beta')} \left|\langle \mathbf{A}_{\tau,\beta}, \mathbf{A}_{\tau',\beta'}\rangle\right| \le 3N_R\sqrt{N_t \log(N_\tau N_\beta)}. \tag{22}$$

The key to proving Theorem 1 is now to combine Lemma 3 and Lemma 4 with Theorem 2. However, the latter theorem requires the matrix to have columns of unit-norm, whereas the columns of our matrix $\mathbf{A}$ have all different norms. Thus instead of $\mathbf{A}\mathbf{x} = \mathbf{y}$ we now consider

$$\tilde{\mathbf{A}}\mathbf{z} = \mathbf{y}, \qquad \text{where } \tilde{\mathbf{A}} := \mathbf{A}\mathbf{D}^{-1} \text{ and } \mathbf{z} := \mathbf{D}\mathbf{x}. \tag{23}$$

Here $\mathbf{D}$ is the $N_\tau N_\beta \times N_\tau N_\beta$ diagonal matrix defined by

$$\mathbf{D}_{(\tau,\beta),(\tau,\beta)} = \|\mathbf{A}_{\tau,\beta}\|_2. \tag{24}$$

When proving Theorem 1, we first establish the claims for the system $\tilde{\mathbf{A}}\mathbf{z} = \mathbf{y}$ in (23) where $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{D}^{-1}$, $\mathbf{z} = \mathbf{D}\mathbf{x}$, and then we switch back to $\mathbf{A}\mathbf{x} = \mathbf{y}$. Note that $\mathbf{x}$ and $\mathbf{z}$ have the same sparsity and we can easily recover $\mathbf{x}$ from $\mathbf{z}$ via $\mathbf{x} = \mathbf{D}^{-1}\mathbf{z}$.

The following lemma gives a bound for $\mu(\tilde{\mathbf{A}})$ and $\|\tilde{\mathbf{A}}\|_{\mathrm{op}}$ in terms of the corresponding bounds for $\mathbf{A}$.

Lemma 5 *Let $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{D}^{-1}$, where the $\mathbf{D}$ the diagonal matrix is defined by (24). Under the conditions of Theorem 1, there holds with high probability*

$$\|\tilde{\mathbf{A}}\|_{op}^2 < 3(1+\log N_t) \tag{25}$$

and

$$\mu(\tilde{\mathbf{A}}) \le 6\sqrt{\frac{1}{N_t}\log(N_\tau N_R N_T)}. \tag{26}$$

If we consider a Gaussian random matrix of the same dimensions as $\mathbf{A}$, then its coherence would scale like $\sqrt{\frac{1}{N_t N_R}\log(N_\tau N_R N_T)}$, thus we would gain a factor of $\sqrt{1/N_R}$ compared to (26). However the coherence proved in Lemma 5 is of the optimal order for the matrix at hand. The reason why the factor $\sqrt{1/N_R}$ cannot appear in (26) is due to the "decoupling" of the receive antennas and the waveforms in the matrix $\mathbf{A}$, which becomes apparent when we recall that we can express the columns of $\mathbf{A}$ as $\mathbf{A}_{\tau,\beta} = \mathbf{a}_R(\beta) \otimes (\mathbf{S}_\tau \mathbf{a}_T(\beta))$.

3.1 Some theoretical results for the Doppler case

We use the same transmission waveforms of bandwidth B as in the previous section. However, due to the Doppler effect the received signal will have a somewhat larger bandwidth $B_1 > B$. But in practice this increase in bandwidth is small, therefore we assume for simplicity $B_1 \approx B$ in the sequel.

Theorem 6 *Consider $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$, where $\mathbf{A}$ is as defined in Subsection 2.2 and $\mathbf{v}_i \in \mathcal{CN}(0, \sigma^2)$. Choose the discretization stepsizes to be $\Delta_\beta = \frac{2}{N_R N_T}$, $\Delta_\tau = \frac{1}{2B}$ and $\Delta_f = \frac{1}{T}$. Let $d_T = 1/2, d_R = N_T/2$ or $d_T = N_R/2, d_R = 1/2$, and suppose that*

$$N_t \geq 128 \qquad \text{and} \qquad \big(\log(N_\tau N_\beta)\big)^3 \leq N_t,$$

If $\mathbf{x}$ is drawn from the generic K-sparse scatterer model with

$$K \leq \frac{c_0 N_\tau N_f N_R}{6\log(N_\tau N_f N_\beta)}$$

for some constant $c_0 > 0$, and if

$$\min_{k\in S} |\mathbf{x}_k| > \frac{10\sigma}{\sqrt{N_R N_t}}\sqrt{2\log N_\tau N_f N_\beta}, \qquad (27)$$

then the solution $\tilde{\mathbf{x}}$ of the debiased lasso computed with $\lambda = 2\sigma\sqrt{2\log(N_\tau N_f N_\beta)}$ obeys with high probability

$$\operatorname{supp}(\tilde{\mathbf{x}}) = \operatorname{supp}(\mathbf{x}),$$

and

$$\frac{\|\tilde{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \frac{6\sigma\sqrt{N_t N_R}}{\|\mathbf{y}\|_2}.$$

The proof is very similar to that of Theorem 1. We need to establish the analogs of the key steps, Lemma 3 and Lemma 4.

Lemma 7 *Let $\mathbf{A}$ be as defined in Theorem 6. Then with high probability*

$$\|\mathbf{A}\|_{op}^2 \leq 2N_t N_f N_R N_T.$$

Next we establish a coherence bound for $\mathbf{A}$.

Lemma 8 *Let $\mathbf{A}$ be as defined in Theorem 6. Assume that*

$$\log(N N_\beta) < \frac{N_t}{23},$$

where N can be chosen to be either N_τ, N_f or $\sqrt{N_\tau N_f}$, then with high probability

$$\max\left|\langle \mathbf{A}_{\tau,f,\beta}, \mathbf{A}_{\tau',f',\beta'}\rangle\right| \leq 3N_R\sqrt{N_t\log(N_\tau N_f N_\beta)}.$$

Remark: Note that equation (27) can be written as $\mathrm{SNR}_k > 100\frac{2\log N_\tau N_f N_\beta}{N_R N_t}$ where $\mathrm{SNR}_k = |x_k|^2/\sigma^2$ is the input signal-to-noise ratio for the k-th scatterer.

References

[1] J. Li and P. Stoica, "MIMO Radar with Colocated Antennas: Review of Some Recent Work," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 106–114, 2007.

[2] A. Haimovich, R. Blum, and L. Cimini, "MIMO radar with widely separated antennas," *IEEE Signal Processing Magazine*, pp. 116–129, 2008.

[3] M. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. on Signal Processing*, vol. 57, no. 6, pp. 2275–2284, 2009.

[4] T. Strohmer and B. Friedlander, "Compressed sensing for MIMO radar - algorithms and performance," in *Asilomar Conference Signals, Systems, Computers*, (Asilomar), 2009.

[5] B. Friedlander, "On the relationship between MIMO and SIMO radars," *IEEE Trans. Signal Processing*, vol. 57, pp. 394–398, January 2009.

[6] M. Herman and T. Strohmer, "General deviants: an analysis of perturbations in compressed sensing," *IEEE Journal of Selected Topics in Signal Processing: Special Issue on Compressive Sensing*, vol. 4, no. 2, pp. 342–349, 2010.

[7] Y. Chi, A. Pezeshki, L. Scharf, and R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, (Dallas, Texas), 2010.

[8] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.

[9] E. Candès and Y. Plan, "Near-ideal model selection by ℓ_1 minimization," *Annals of Statistics*, vol. 37, no. 5A, pp. 2145–2177, 2009.