Spatial Compressive Sensing for MIMO Radar

Marco Rossi, *Student Member, IEEE*, Alexander M. Haimovich, *Fellow, IEEE*, and Yonina C. Eldar, *Fellow, IEEE*

Abstract—We study compressive sensing in the spatial domain to achieve target localization, specifically direction of arrival (DOA), using multiple-input multiple-output (MIMO) radar. A sparse localization framework is proposed for a MIMO array in which transmit and receive elements are placed at random. This allows for a dramatic reduction in the number of elements needed, while still attaining performance comparable to that of a filled (Nyquist) array. By leveraging properties of structured random matrices, we develop a bound on the coherence of the resulting measurement matrix, and obtain conditions under which the measurement matrix satisfies the so-called *isotropy* property. The coherence and isotropy concepts are used to establish uniform and non-uniform recovery guarantees within the proposed spatial compressive sensing framework. In particular, we show that non-uniform recovery is guaranteed if the product of the number of transmit and receive elements, MN (which is also the number of degrees of freedom), scales with $K(\log G)^2$, where K is the number of targets and G is proportional to the array aperture and determines the angle resolution. In contrast with a filled virtual MIMO array where the product MN scales linearly with G, the logarithmic dependence on G in the proposed framework supports the high-resolution provided by the virtual array aperture while using a small number of MIMO radar elements. In the numerical results we show that, in the proposed framework, compressive sensing recovery algorithms are capable of better performance than classical methods, such as beamforming and MUSIC.

Index Terms—Compressive sensing, direction of arrival estimation, MIMO radar, random arrays.

I. INTRODUCTION

DETECTION, localization, and tracking of targets are basic radar functions. Limited data support and low signal-to-noise ratios (SNR) are among the many challenges frequently faced by localization systems. Another challenge is the presence of nearby targets, whether in terms of location or Doppler, since closely spaced targets are more difficult to discriminate. In multiple-input multiple-output (MIMO) radar, targets are probed with multiple, simultaneous waveforms. Relying on the orthogonality of the waveforms, returns from the targets are jointly processed by multiple receive antennas. MIMO radar is typically used in two antenna configurations, namely distributed [1] and colocated [2]. Depending on the mode of operation and system architecture, MIMO radars have been shown to boost target detection, enhance spatial resolution, and improve interference suppression. These advantages are achieved by providing and exploiting a larger number of degrees of freedom than "conventional" radar.

In this work, we focus on the application of colocated MIMO radar to direction-of-arrival (DOA) estimation. It is well known in array signal processing [3] that DOA resolution improves by increasing the array aperture. However, increasing the aperture without increasing the number of sensors may lead to ambiguities, i.e., measurements explained by erroneous sets of locations. A non-ambiguous uniform linear array (ULA) must have its elements spaced at intervals no larger than $\lambda/2$, where λ is the signal wavelength. For MIMO radar, unambiguous direction finding of targets is possible if N receive elements are spaced $\lambda/2$, and M transmit elements are spaced $N\lambda/2$, a configuration known as *virtual ULA* [2]. In sampling parlance, the $\lambda/2$-spaced array and the MIMO virtual ULA perform spatial sampling at the Nyquist rate. The main disadvantage of this Nyquist setup is that the product of the number of transmit and receive elements, MN, needs to scale linearly with the array aperture, and thus with resolution.

In this paper, we propose the use of a sparse, random array architecture in which a low number of transmit/receive elements are placed at random over a large aperture. This setup is an example of spatial compressive sensing since spatial sampling is applied at sub-Nyquist rates. The goal of spatial compressive sensing is to achieve similar resolution as a filled array, but with significantly fewer elements.

Localizing targets from undersampled array data links random arrays to *compressive sensing* [4]. Random array theory can be traced back to the 1960's. In [5], it is shown that as the number of sensors is increased, the random array pattern, a well known quantity to radar practitioners, converges to its average. This is because the array pattern's variance decreases linearly with the number of elements. This work was extended to MIMO radar in [6]. The main conclusion of the classical random array literature was that the random array pattern can be controlled by using a sufficient number of sensors. However, two fundamental questions were left pending: How many sensors are needed for localization as a function of the number of targets, and which method should be used for localization? Here we suggest that the theory and algorithms of compressed sensing may be used to address these questions.

Early works on compressive sensing radar emphasize that the sparse nature of many radar problems supports the reduction of temporal as well as spatial sampling (an overview is given in [7]). Recent work on compressive sensing for

single-input single-output radar [8]–[11] demonstrates either an increased resolution or a reduction in the temporal sampling rate. Compressive sensing for MIMO radar has been applied both on distributed [12] and colocated [13] setups. Much of the previous literature on compressive sensing for colocated arrays discusses the ULA setup, either within a passive system (with only receive elements) [14] or in a MIMO radar [13], [15] setup. In particular, [15] imposes a MIMO radar virtual ULA and derives bounds on the number of elements to perform range-angle and range-Doppler-angle recovery by using compressive sensing techniques. As discussed above, the (virtual) ULA setup performs Nyquist sampling in the spatial domain. In contrast, we are interested in *spatial* compressive sensing (i.e., reducing the number of antenna elements while fixing the array aperture), and rely on a random array geometry. Links between compressive sensing and random arrays have been explored in [16]. The author shows that spatial compressive sensing can be applied to the passive DOA problem, allowing for a reduction in the number of receiving elements. However, the MIMO radar framework poses a major challenge: contrary to the passive setup, where the rows of the sensing matrix $\mathbf{A}$ are independent, the MIMO radar MN measurements are dependent (they conform to the structure of the MIMO random array steering vector). This lack of independence prevents the application of the vast majority of results in the compressive sensing literature. A MIMO radar random array architecture is studied in [17], but no recovery guarantees are provided.

Low-rate spatial sampling translates into cost savings due to fewer antenna elements involved. It is of practical interest to determine the least amount of elements required to guarantee correct targets recovery. Finding conditions that guarantee recovery has been a main topic of research, and it is one of the underpinnings of compressive sensing theory. Recent work has shown that, for a sufficient number of independent and identically distributed (i.i.d.) compressive sensing measurements, non-uniform recovery can be guaranteed if a specific property of the random sensing matrix, called *isotropy*, holds [18]. While this property plays an important role, this result does not apply to our setup since the MIMO radar MN measurements are not independent. The dependent measurements problem was recently addressed in [19]. There, the authors derived conditions for non-uniform recovery using spatial compressive sensing in a MIMO radar system with N transceivers.

This work expands the literature in several ways. We propose a sparse localization framework for a MIMO random array assuming a general setup of M transmitters and N receivers. We provide a bound on the coherence of the measurement matrix, and determine the conditions under which the isotropy property holds. This allows us to develop both uniform and non-uniform recovery guarantees for target localization in MIMO radar systems. The proposed MIMO random array framework is of practical interest to airborne and other radar applications, where the spacing between antenna elements may vary as a function of aspect angle towards the target, or where exact surveying of element locations is not practical due to natural flexing of the structures involved. Our results show that one can obtain the high-resolution provided by a virtual array aperture while using a reduced number of antenna elements.
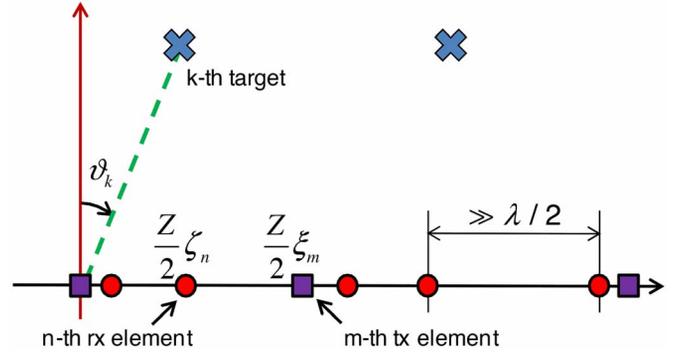
Fig. 1. MIMO radar system model.

The paper is organized as follows: Section II introduces the system model and the proposed sparse localization framework. Section III discusses spatial compressive sensing. Recovery guarantees are derived in Section IV. In Section V, we present numerical results demonstrating the potential of the proposed framework, followed by conclusions in Section VI.

The following notation is used: boldface denotes matrices (uppercase) and vectors (lowercase); for a vector $\mathbf{a}$, the i-th index is $\mathbf{a}_i$, while for a matrix $\mathbf{A}$, the i-th row is denoted by $\mathbf{A}(i,:)$. The complex conjugate operator is $(\cdot)^*$, the transpose operator is $(\cdot)^T$, and the complex conjugate-transpose operator is $(\cdot)^H$. We define $\|\mathbf{X}\|_0$ as the number of non-zero norm rows of $\mathbf{X}$, the support of $\mathbf{X}$ collects the indices of such rows, and a K-sparse matrix satisfies $\|\mathbf{X}\|_0 \leq K$. The operator $\mathbb{E}$ denotes expectation and we define $\psi_x(u) \triangleq \mathbb{E}[\exp(jxu)]$ as the characteristic function of the random variable x. The symbol "$\otimes$" denotes the Kronecker product. The notation $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$ means that the vector $\mathbf{x}$ has a circular symmetric complex normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$. We denote by $K_\alpha(\cdot)$ the modified Bessel function of the second kind.

II. SYSTEM MODEL

A. MIMO Radar Model

We model a MIMO radar system (see Fig. 1) in which N sensors collect a finite train of P pulses sent by M transmitters and returned from K stationary targets. We assume that transmitters and receivers each form a (possibly overlapping) linear array of total aperture Z_{TX} and Z_{RX}, respectively. The quantities Z_{TX} and Z_{RX} are normalized in wavelength units. Defining $Z \triangleq Z_{TX} + Z_{RX}$, the m-th transmitter is at position $Z\xi_m/2$ on the x-axis, while the n-th receiver is at position $Z\zeta_n/2$. Here ξ_m lies in the interval $[-\frac{Z_{TX}}{Z}, \frac{Z_{TX}}{Z}]$, and ζ_n is in $[-\frac{Z_{RX}}{Z}, \frac{Z_{RX}}{Z}]$. This definition ensures that when $Z_{TX} = Z_{RX}$, both ξ_m and ζ_n are confined to the interval $[-\frac{1}{2}, \frac{1}{2}]$, simplifying the notation in the sequel.

Let $s_m(t)$ denote the continuous-time baseband signal transmitted by the m-th transmit antenna and let θ denote the location parameter(s) of a generic target, for example, its azimuth angle. Assume that the propagation is nondispersive and that the transmitted probing signals are narrowband (in the sense that the envelope of the signal does not change appreciably across the antenna array). Then the baseband signal at the target location, considering the p-th transmitted pulse, can be described by (see, e.g., [1])

$$\sum_{m=1}^{M} \exp(j2\pi f_0 \tau_m(\theta)) s_m(t - pT) \triangleq \mathbf{c}^T(\theta)\mathbf{s}(t - pT). \quad (1)$$

Here f_0 is the carrier frequency of the radar, $\tau_m(\theta)$ is the time needed by the signal emitted by the m-th transmit antenna to arrive at the target, $\mathbf{s}(t) \triangleq [s_1(t), \dots, s_M(t)]^T$, T denotes the pulse repetition interval, and

$$\mathbf{c}(\theta) = [\exp(j2\pi f_0 \tau_1(\theta)), \dots, \exp(j2\pi f_0 \tau_M(\theta))]^T \quad (2)$$

is the transmit steering vector. Assuming that the transmit array is calibrated, $\mathbf{c}(\theta)$ is a known function of θ.

To develop an expression for the received signal $r_n(t)$ at the n-th receive antenna, let

$$\mathbf{b}(\theta) = [\exp(j2\pi f_0 \tilde{\tau}_1(\theta)), \dots, \exp(j2\pi f_0 \tilde{\tau}_N(\theta))]^T \quad (3)$$

denote the receive steering vector. Here $\tilde{\tau}_n(\theta)$ is the time needed for the signal reflected by the target located at θ to arrive at the n-th receive antenna. Define the vector of received signals as $\mathbf{r}(t) \triangleq [r_1(t), \dots, r_N(t)]^T$. Under the simplifying assumption of point targets, the received data vector is described by [1]

$$\mathbf{r}(t) = \sum_{k=1}^{K} \sum_{p=1}^{P-1} x_{k,p} \mathbf{b}(\theta_k) \mathbf{c}^T(\theta_k) \mathbf{s}(t - pT) + \mathbf{e}(t) \quad (4)$$

where K is the number of targets that reflect the signals back to the radar receiver, $x_{k,p}$ is the complex amplitude proportional to the radar cross sections of the k-th target relative to pulse p-th, θ_k are locations, and $\mathbf{e}(t)$ denotes the interference plus-noise term. The targets' positions are assumed constant over the observation interval of P pulses. We assume that the target gains $\{x_{k,p}\}$ follow a Swerling Case II model, meaning that they are fixed during the pulse repetition interval T, and vary independently from pulse to pulse [20].

Analyzing how to estimate the number of targets K, or the noise level, without prior information is the topic of current work [21], but outside the scope of this paper. Therefore in the following, we assume that the number of targets K is known and the noise level is available.

B. Problem Formulation

The purpose of the system is to determine the DOA angles to targets of interest. We consider targets associated with a particular range and Doppler bin. Targets in adjacent range-Doppler bins contribute as interferences to the bin of interest. The assumption of a common range bin implies that all waveforms are received with the same time delay after transmission. Since range and Doppler measurements are not of interest, the common time delay and Doppler shift are not explicitly shown in our model. This approach is justified because angle resolution is essentially independent of range-Doppler resolution in antenna arrays [22]. Being capable to handle targets with non-zero Doppler, our approach is applicable to airborne or ground targets. Targets are assumed in the far-field, meaning that a target's DOA parameter $\theta \triangleq \sin \vartheta$ (where ϑ is the DOA angle) is constant across the array. Under these assumptions, the receiver and transmitter steering vectors, $\mathbf{b}(\theta)$ and $\mathbf{c}(\theta)$ respectively, become

$$\mathbf{b}(\theta) = [\exp(j\pi Z\theta \zeta_1), \dots, \exp(j\pi Z\theta \zeta_N)]^T \quad (5)$$

and

$$\mathbf{c}(\theta) = [\exp(j\pi Z\theta \xi_1), \dots, \exp(j\pi Z\theta \xi_M)]^T. \quad (6)$$

By cross-correlating the received signal at each sensor with filters matched to each of the probing waveforms, we obtain

$$\mathbf{y}_p = \text{vec} \left[\int \mathbf{r}(t) \mathbf{s}^H(t - pT) dt \right]$$

$$= \text{vec} \left[\sum_{k=1}^{K} \sum_{p=0}^{P-1} x_{k,p} \mathbf{b}(\theta_k) \mathbf{c}^T(\theta_k) \mathbf{W} + \right.$$

$$\left. + \int \mathbf{e}(t) \mathbf{s}^H(t - pT) dt \right] \quad (7)$$

where the $M \times M$ matrix $\mathbf{W}$ has elements

$$[\mathbf{W}]_{m,j} = \int s_m(t) s_j^*(t) dt. \quad (8)$$

We assume the M probing waveforms to be orthogonal (e.g., pulses modulated by an orthogonal code), therefore $\mathbf{W} = \mathbf{I}$. Defining the $MN \times P$ matrix $\mathbf{Y} \triangleq [\mathbf{y}_1, \dots, \mathbf{y}_P]$, we have from (7)

$$\mathbf{Y} = \tilde{\mathbf{A}}(\boldsymbol{\theta})\tilde{\mathbf{X}} + \mathbf{E}. \quad (9)$$

Here $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_P]$ is a $K \times P$ matrix with $\tilde{\mathbf{x}}_p = [x_{1,p}, \dots, x_{K,p}]^T$,

$$\tilde{\mathbf{A}}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)] \quad (10)$$

is a $MN \times K$ matrix with columns

$$\mathbf{a}(\theta) \triangleq \mathbf{c}(\theta) \otimes \mathbf{b}(\theta) \quad (11)$$

known as the "virtual array" steering vector, and $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_P]$ is $MN \times P$ with $\mathbf{e}_p = \text{vec}[\int \mathbf{e}(t) \mathbf{s}^H(t - pT) dt]$. The term "virtual array" indicates that $\mathbf{a}(\theta)$ can be thought of as a steering vector with MN elements.

Our aim is to recover $\boldsymbol{\theta}$ and $\tilde{\mathbf{X}}$ from $\mathbf{Y}$ using a small number of antenna elements. To do this, we use a sparse localization framework. Neglecting the discretization error, it is assumed that the target possible locations $\boldsymbol{\theta}$ comply with a grid of G points $\phi_{1:G}$ (with $G \gg K$). Since each element of $\boldsymbol{\theta}$ parameterizes one column of $\tilde{\mathbf{A}}(\boldsymbol{\theta})$, it is possible to define an $MN \times G$ dictionary matrix $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_G]$, where $\mathbf{a}_g = \mathbf{a}(\phi_g)$. From (11), the steering vector $\mathbf{a}_g$ is the Kronecker product of the receive steering vector $\mathbf{b}_g = \mathbf{b}(\phi_g)$ and the transmit steering vector $\mathbf{c}_g = \mathbf{c}(\phi_g)$:

$$\mathbf{a}_g = \mathbf{c}_g \otimes \mathbf{b}_g. \quad (12)$$

The received signal is then expressed as

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{E}, \quad (13)$$

where the unknown $G \times P$ matrix $\mathbf{X}$ contains the target locations and gains. Zero rows of $\mathbf{X}$ correspond to grid points without a target. The system model (13) is sparse in the sense that $\mathbf{X}$ has only $K \ll G$ non-zero rows.

Note that in the sparse localization framework, the matrix $\mathbf{A}$ is known, whereas in the array processing model (9), the matrix $\tilde{\mathbf{A}}(\boldsymbol{\theta})$ is unknown. Given the measurements $\mathbf{Y}$ and matrix $\mathbf{A}$, our goal translates into determining the non-zero norm rows' indices of $\mathbf{X}$, i.e., the support of $\mathbf{X}$. The matrix $\mathbf{A}$ is governed by the choice of grid points $\phi_{1:G}$, by the number M of transmitters and their positions, $\xi_{1:M}$, and by the number N of receivers and their positions, $\zeta_{1:N}$. In the following we assume that the transmitter (receiver) elements' positions $\xi_{1:M}$ ($\zeta_{1:N}$) are independent and identically distributed (i.i.d.) random variables governed by a probability density function (pdf) $p(\xi)(p(\zeta))$.

III. SPATIAL COMPRESSIVE SENSING FRAMEWORK

The aim of spatial compressive sensing is to recover the unknown $\mathbf{X}$ from the measurements $\mathbf{Y}$ (see (13)) using a small number of antenna elements, MN, while fixing the array aperture Z. In this section we introduce the proposed spatial compressive sensing framework and overview practical recovery algorithms (the well-known beamforming method as well as compressive sensing based algorithms).

A. Beamforming

Consider the scenario in which the transmitters and receivers locations support the Nyquist array (virtual ULA) geometry. In this setting, the matrix $\mathbf{A}$ in (13) has a Vandermonde structure, and the aperture scales linearly with the number of antenna elements, $Z = (MN - 1)/2$. If we choose the (uniform) grid of possible target locations ϕ_g to match the array resolution, that is $G = 2Z + 1$, then the matrix $\mathbf{A}$ becomes a Fourier matrix. In this case, $\mathbf{Q} \triangleq \mathbf{A}^H \mathbf{A} = MN \cdot \mathbf{I}$. It follows that $\mathbf{X}$ can be estimated as $(1/MN) \cdot \mathbf{A}^H \mathbf{Y}$. In array processing, this method is called beamforming. The support of the unknown $\mathbf{X}$ is recovered by looking for peak values of $\|\mathbf{a}_g^H \mathbf{Y}\|_2$ over the grid points. Beamforming is also applied to estimate the locations of targets not limited to a grid. This is done by finding the peaks of $\|\mathbf{a}^H(\theta)\mathbf{Y}\|_2$, where $\mathbf{a}(\theta)$ is the steering vector (11) swept over the angles of interest. The shortcoming of the Nyquist array setup is that the number of elements MN must scale linearly with the array aperture Z and consequently, with the resolution (i.e., such sampling mode requires $MN = G$).

Spatial compressive sensing implies that a sparse $\mathbf{X}$ can be recovered from a number of spatial measurements significantly lower than the Nyquist array, i.e., $MN \ll G$. The idea is to design the sensing procedure so that the matrix $\mathbf{Q}$ is a scalar multiple of the identity matrix *on average*[1], i.e., $\mathbb{E}[\mathbf{Q}] = MN \cdot \mathbf{I}$, and to control the variance of the non-diagonal elements by using a sufficient number of measurements. Intuitively, the more measurements MN we employ, the closer we get to a diagonal $\mathbf{Q}$. Furthermore, because when $MN < G$, each realization of $\mathbf{Q}$ has non-zero off-diagonal terms, the beamforming metric $\|\mathbf{a}_g^H \mathbf{Y}\|_2$ is affected not only by the g-th row of $\mathbf{X}$ and by the noise, but also by any row of $\mathbf{X}$ that has non-zero norm. This entails that, instead of beamforming, we resort to more sophisticated recovery algorithms, which take advantage of the signal's sparsity to mitigate the mutual interference among non-zero rows of $\mathbf{X}$. A brief overview of compressive sensing recovery methods is provided next.

[1]For instance, this is obtained when using a partial Fourier matrix.

B. Compressive Sensing

One way to classify compressive sensing models is according to the number of pulses P ("snapshots" in array processing problems): single measurement vector (SMV) for $P = 1$ ($\mathbf{Y}$ reduces to a single vector), or multiple measurement vector (MMV) for $P \geq 1$ ($\mathbf{Y}$ is a matrix). The system model in (13) is an example of an MMV setting. For simplicity, in the following we consider an SMV scenario (i.e., $P = 1$, $\mathbf{Y} = \mathbf{y}$, $\mathbf{X} = \mathbf{x}$ and $\mathbf{E} = \mathbf{e}$ in (13)).

In principle, a sparse $\mathbf{x}$ (i.e., it has only $K \ll G$ non-zero rows) can be recovered from the least number of elements MN by solving the non-convex combinatorial ℓ_0-norm problem

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ s.t. } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \sigma, \qquad (14)$$

or one of its equivalent formulations: a cardinality-constrained formulation, $\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \text{ s.t } \|\mathbf{x}\|_0 \leq \gamma$, or a Lagrangian formulation, $\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \nu\|\mathbf{x}\|_0$. These three formulations are equivalent for a proper choice of the parameters γ, ν and σ, which depend on prior information, e.g., the noise level $\|\mathbf{e}\|_2$ or the sparsity K. Unfortunately, the solution to any of these formulations requires an exhaustive search among all combinations of non-zero indices of $\mathbf{x}$, necessitating exponential complexity [4].

A variety of polynomial complexity algorithms have been proposed for obtaining an approximate solution to (14). One family of methods is Matching Pursuit (MP). In its simplest version, an empty provisional support is refined by adding one grid-point index at each iteration. Among the matching pursuit algorithms, the most notable in the SMV setting are Orthogonal Matching Pursuit (OMP) [23], Orthogonal Least Squares (OLS) [24], and CoSaMP [25]. For the general MMV setting, examples are the Rank Aware-Orthogonal Recursive Matching Pursuit (RA-ORMP) algorithm [26] and its generalization, Multi-Branch Matching Pursuit (MBMP) [27]. Another family of methods is known as Basis Pursuit (BP). The BP strategy relaxes the l_0-norm in (14) with the l_1-norm [28]. The reformulation is known as LASSO, defined by

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ s.t. } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \sigma. \qquad (15)$$

Unlike (14), this problem is convex, and a global solution can be found in polynomial time. Since it is a relaxation, the solution obtained could be different from that of (14). Finding conditions that guarantee correct recovery with a specific method (e.g., LASSO) has been a main topic of research and one of the underpinnings of compressive sensing theory [4].

Two kinds of recovery guarantees are defined in compressive sensing: uniform and non-uniform. A uniform recovery guarantee (addressed below by Theorem 2) means that for a fixed instantiation of the random measurement matrix $\mathbf{A}$, all possible K-sparse signals are recovered with high probability. In contrast, a non-uniform recovery result (addressed by Theorem 4) captures the typical recovery behavior for a random measurement matrix $\mathbf{A}$. Specifically, suppose we are given an arbitrary K-sparse vector $\mathbf{x}$, and we then draw $\mathbf{A}$ at random (independent of $\mathbf{x}$). Non-uniform recovery details under what conditions an algorithm will recover $\mathbf{x}$ with high probability. Note that, for a non-uniform guarantee, $\mathbf{A}$ is being asked to recover only a specific $\mathbf{x}$, not *any* K-sparse vectors. Therefore, uniform recovery implies non-uniform recovery, but the converse is not true.

Loosely speaking, a uniform recovery guarantee can be obtained if, with high probability, the matrix $\mathbf{A}$ has small coherence [4]. The coherence is defined as the maximum inner product between the normalized columns of $\mathbf{A}$,

$$\mu \triangleq \max_{i \neq l} \frac{|\mathbf{a}_i^H \mathbf{a}_l|}{\|\mathbf{a}_i\|_2 \|\mathbf{a}_l\|_2}. \tag{16}$$

Alternatively, uniform recovery is guaranteed if $\mathbf{A}$ satisfies the Restricted Isometry Property (RIP) [4] with high probability. Non-uniform recovery follows if a specific property of the random measurement matrix $\mathbf{A}$, called isotropy, holds [18]. The isotropy property states that the components of each row of $\mathbf{A}$ have unit variance and are uncorrelated, i.e.,

$$\mathbb{E}[\mathbf{A}^H(t,:)\mathbf{A}(t,:)] = \mathbf{I} \tag{17}$$

for every t.

Both (16) and (17) suggest that the matrix $\mathbf{Q} \triangleq \mathbf{A}^H \mathbf{A}$ plays a key role in establishing recovery guarantees. Indeed, because in our setting the rows of $\mathbf{A}$ are identically distributed, a simple calculation shows that $\mathbb{E}[\mathbf{A}^H(t,:)\mathbf{A}(t,:)] = \frac{1}{MN}\mathbb{E}[\mathbf{Q}]$, thus the isotropy property requires $\mathbb{E}[\mathbf{Q}] = MN \cdot \mathbf{I}$. Furthermore, as evident by the definition of coherence in (16), μ is the maximum absolute value among normalized off-diagonal elements of $\mathbf{Q}$. In Section IV, by deriving statistics of the matrix $\mathbf{Q}$, we provide conditions on design's quantities ($p(\xi)$, $p(\zeta)$, M, N and $\phi_{1:G}$) to obtain (uniform and non-uniform) recovery guarantees for spatial compressive sensing.

IV. RECOVERY GUARANTEES

In this section, we develop recovery guarantees for sparse localization with MIMO random arrays. In detail, we show how to choose the grid-points $\phi_{1:G}$, the number of elements MN and the distributions governing the element positions $p(\xi)$ and $p(\zeta)$, in order to guarantee target localization by spatial compressive sensing via (15). Due to the role of the matrix $\mathbf{Q}$ in recovery guarantees, we start by studying the statistics of $\mathbf{Q}$.

A. Statistics of $\mathbf{Q} \triangleq \mathbf{A}^H \mathbf{A}$

To study the statistics of $\mathbf{Q}$, we first analyze its relationship to a quantity known to radar practitioners as the *array pattern* [29]. In array processing, the array pattern $\beta(u_{i,l})$ is the system response of an array beamformed in direction ϕ_l to a unit amplitude target located in direction ϕ_i. In other words, $\beta(u_{i,l})$ is the inner product between two normalized columns of the measurement matrix:

$$\beta(u_{i,l}) \triangleq \frac{\mathbf{a}_i^H \mathbf{a}_l}{\|\mathbf{a}_i\|_2 \|\mathbf{a}_l\|_2}$$
$$= \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \exp[ju_{i,l}(\zeta_n + \xi_m)], \tag{18}$$

where we defined

$$u_{i,l} \triangleq \pi Z(\phi_l - \phi_i). \tag{19}$$

The peak of the absolute value of the array pattern for a target colinear with the beamforming direction, $|\beta(0)|$, is called the *mainlobe*. Peaks of $|\beta(u)|$ for $u \neq 0$, are known as *sidelobes,* and the highest among all the sidelobes is called the *peak sidelobe*. Thus the terms $\mathbf{a}_i^H \mathbf{a}_l$ in (18) play the role of sidelobes.

The relation between coherence, isotropy and array pattern is apparent. Indeed, from (16), (18), and the definition of sidelobes, the coherence, in array processing parlance, is the peak sidelobe associated with the matrix $\mathbf{A}$. Similarly, from (17) and (18), the isotropy can be related to the mean array pattern

$$\eta(u_{i,l}) \triangleq \mathbb{E}[\beta(u_{i,l})], \tag{20}$$

where the expectation $\mathbb{E}[\beta(u_{i,l})]$ is taken with respect to the ensemble of element locations. In particular, isotropy requires that $\eta(u_{i,l}) = 0$ for any $i \neq l$.

For a system with randomly placed sensors, the array pattern $\beta(u_{i,l})$ is a stochastic process. Naturally, statistics of the array pattern of a random array depend on the pdf of the sensor locations. In [6], the authors derive the means and the variances of the real and imaginary parts of $\beta(u_{i,l})$. The following proposition formalizes pertinent results from [6]. For the sake of brevity, we drop the dependency on i and l, and denote the array pattern as $\beta(u)$. Define $z \triangleq \zeta + \xi$, and assume that the pdf of z, $p(z)$, is an even function (so that $\text{Im}\eta(u) = 0$). Further, define the variances of the array pattern $\sigma_1^2(u) \triangleq \text{var}[\text{Re}\beta(u)]$, $\sigma_2^2(u) \triangleq \text{var}[\text{Im}\beta(u)]$ and $\sigma_{12}(u) \triangleq \mathbb{E}[(\text{Re}\beta(u) - \eta(u))\text{Im}\beta(u)]$.

Proposition 1: Let the locations ξ of the transmit elements be i.i.d., drawn from a distribution $p(\xi)$, and the locations ζ of the receive elements be i.i.d., drawn from a distribution $p(\zeta)$. Then, for a given u, the following holds:

1) The mean array pattern is the characteristic function of z, i.e.,

$$\eta(u) = \psi_z(u). \tag{21}$$

2) If ξ and ζ are identically distributed, then (22), (23) at the bottom of the page, and $\sigma_{12}(u) = 0$ hold.

   Proof: See Appendix A. $\qquad\square$

Proposition 1 links the probability distributions $p(\xi)$ and $p(\zeta)$ (via $\psi_z(u)$ and $\psi(u)$) to the mean and variances of each element of the matrix $\mathbf{Q}$, i.e., $\beta(u_{i,l}) = \frac{1}{MN}\mathbf{a}_i^H \mathbf{a}_l$. As shown below, this result is used to obtain non-uniform recovery guarantees.

To characterize the statistics of the coherence μ (defined in (16)), we need the distribution of the maximum absolute value among normalized off-diagonal elements of $\mathbf{Q}$.

We now show that, by imposing specific constraints on the grid points $\phi_{1:G}$ and on the probability distributions $p(\xi)$ and $p(\zeta)$, we can characterize the distributions of the elements of $\mathbf{Q}$.

$$\sigma_1^2(u) = \frac{1}{2MN}(1 + \psi_z(2u)) + \psi_z(u)\left[\frac{N+M-2}{2MN}(1 + \psi_\xi(2u)) - \psi_z(u)\frac{N+M-1}{MN}\right] \tag{22}$$

$$\sigma_2^2(u) = \frac{1}{2MN}(1 - \psi_z(2u)) + \psi_z(u)\frac{N+M-2}{2MN}(1 + \psi_\xi(2u)). \tag{23}$$

To do this, we require an intermediate result about the structure of the matrix $\mathbf{Q}$ when $\phi_{1:G}$ is a uniform grid:

Lemma 1: If $\phi_{1:G}$ is a uniform grid, $\mathbf{Q}$ is a Toeplitz matrix.

Proof: See Appendix B.   □

Thanks to Lemma 1, whenever $\phi_{1:G}$ is a uniform grid, $\mathbf{Q}$ is described completely by the elements of its first row, $\mathbf{a}_1^H \mathbf{a}_i$ for $i = 1, \ldots, G$. From the definition of $\mathbf{A}$, its columns all have squared-norm equal to MN. Therefore the elements on the main diagonal of $\mathbf{Q}$ are equal to MN. Thus, we need to investigate the remaining random elements, $\mathbf{a}_1^H \mathbf{a}_i$ for $i = 2, \ldots, G$. By exploiting the Kronecker structure of the columns of $\mathbf{A}$ in (12), we can express the elements of $\mathbf{Q}$ as:

$$\mathbf{a}_i^H \mathbf{a}_j \triangleq (\mathbf{c}_i^H \otimes \mathbf{b}_i^H)(\mathbf{c}_j \otimes \mathbf{b}_j)$$
$$= \mathbf{c}_i^H \mathbf{c}_j \mathbf{b}_i^H \mathbf{b}_j, \tag{24}$$

where $\mathbf{b}$ and $\mathbf{c}$ are the steering vectors of the receiver and transmitter arrays, respectively.

From (24), the random variable $\beta(u_{1,i}) \triangleq \frac{1}{MN}\mathbf{a}_1^H \mathbf{a}_i$ is the product between the random variables $\beta_\zeta(u_{1,i}) \triangleq \frac{1}{N}\mathbf{b}_1^H \mathbf{b}_i$ and $\beta_\xi(u_{1,i}) \triangleq \frac{1}{M}\mathbf{c}_1^H \mathbf{c}_i$. As such, the distribution of $\beta(u_{1,i})$ (or equivalently, $\mathbf{a}_1^H \mathbf{a}_i$) can be characterized from the distributions of $\beta_\zeta(u_{1,i})$ and $\beta_\xi(u_{1,i})$. Following the approach in [5], we show in Appendix C that the real and imaginary parts of $\beta_\zeta(u_{1,i})$ (or $\beta_\xi(u_{1,i})$) have an asymptotic joint Gaussian distribution, but, in general, the variances of real and imaginary parts of such variables are not equal. Interestingly, a closed form expression for the cumulative density function (cdf) of the product of $\beta_\zeta(u_{1,i})$ and $\beta_\xi(u_{1,i})$ (i.e., the cdf of $\frac{1}{MN}\mathbf{a}_1^H \mathbf{a}_i$) exists in the special case when $\mathrm{var}[\mathrm{Re}\beta_\zeta(u_{1,i})] = \mathrm{var}[\mathrm{Im}\beta_\zeta(u_{1,i})]$ and $\mathrm{var}[\mathrm{Re}\beta_\xi(u_{1,i})] = \mathrm{var}[\mathrm{Im}\beta_\xi(u_{1,i})]$. By meeting these conditions, in the following theorem we derive an upper bound on the sidelobes' complementary cdf (ccdf), i.e., $\mathrm{Pr}(\frac{1}{MN}|\mathbf{a}_1^H \mathbf{a}_i| > q)$, and show that sidelobes have uniformly distributed phases.

We address two MIMO radar setups: (1) M transmitters and N receivers, where ξ and ζ are independent, and (2) N transceivers, where $\xi_n = \zeta_n$, for all n and $M = N$.

Theorem 1: Let the locations ξ of the transmit elements be drawn i.i.d. from a distribution $p(\xi)$, and the locations ζ of the receive elements be drawn i.i.d. from a distribution $p(\zeta)$. Assume that $p(\xi)$, $p(\zeta)$ and the uniform grid $\phi_{1:G}$ are such that the transmitter and receiver characteristic functions satisfy

$$\psi_\xi(u_{1,i}) = \psi_\xi(2u_{1,i}) = \psi_\zeta(u_{1,i}) = \psi_\zeta(2u_{1,i}) = 0 \tag{25}$$

for $i = 2, \ldots, G$, where $u_{1,i} = \pi Z(\phi_i - \phi_1)$. Then for $i = 2, \ldots, G$:

1) If ξ and ζ are independent:

$$\mathrm{Pr}\left(\frac{1}{MN}|\mathbf{a}_1^H \mathbf{a}_i| > q\right) < x \cdot K_1(x), \tag{26}$$

where $x \triangleq 2\sqrt{MN}q$.

2) If $\xi_n = \zeta_n$ for all n:

$$\mathrm{Pr}\left(\frac{1}{N^2}|\mathbf{a}_1^H \mathbf{a}_i| > q\right) < \exp(-Nq). \tag{27}$$

3) In both scenarios, the phase of $\mathbf{a}_1^H \mathbf{a}_i$ is uniformly distributed on the unit circle, i.e.,

$$\angle \mathbf{a}_1^H \mathbf{a}_i \sim \mathcal{U}[0, 2\pi]. \tag{28}$$

Proof: See Appendix C.   □

This theorem characterizes the distribution of $\frac{1}{MN}\mathbf{a}_1^H \mathbf{a}_i$ for the M transmitters N receivers setup, and for the N transceivers setup. In Section IV-D, we provide a practical setup that satisfies (25). As shown below, this allows to obtain a uniform recovery guarantee for spatial compressive sensing.

B. Uniform Recovery

The following corollary of Theorem 1 bounds the probability that the matrix $\mathbf{A}$ has high coherence, or equivalently, the probability of a peak sidelobe:

Corollary 1: Let the locations of the transmit elements ξ be drawn i.i.d. from a distribution $p(\xi)$, and the locations of the receivers ζ be drawn i.i.d. from a distribution $p(\zeta)$. Assume that the distributions $p(\xi)$ and $p(\zeta)$ and the uniformly spaced grid-points $\phi_{1:G}$ are such that (25) holds for $i = 2, \ldots, G$. Then:

1) If ξ and ζ are independent:

$$\mathrm{Pr}(\mu > q) < 1 - [1 - x \cdot K_1(x)]^{G-1}, \tag{29}$$

where $x \triangleq 2\sqrt{MN}q$.

2) If $\xi_n = \zeta_n$ for all n:

$$\mathrm{Pr}(\mu > q) < 1 - [1 - \exp(-Nq)]^{G-1}. \tag{30}$$

Proof: See Appendix D.   □

Since μ can be interpreted as the peak sidelobe of the array pattern, (29) ((30)) characterizes the probability of having a peak sidelobe higher than q in a system with M transmitters and N receivers (N transceivers). These results are not asymptotic (i.e., they do not need the number of measurements M and N to tend to infinity). To further explore this point, in numerical results we compare these bounds with empirical simulations.

The coherence μ plays a major role in obtaining uniform recovery guarantees for compressive sensing algorithms, as well as guaranteeing the uniqueness of the sparsest solution to (14). For instance, using the coherence μ, it is possible to obtain a bound on the RIP constant, $\delta_K \leq (K - 1)\mu$ [30]. This ensures stable and robust recovery by l_1-minimization (i.e., using (15)) from noisy measurements. By building on Corollary 1, the following theorem establishes the number of elements MN needed to obtain uniform recovery with high probability using (15):

Theorem 2: [Uniform recovery guarantee]. Let the locations ξ of the transmit elements be drawn i.i.d. from a distribution $p(\xi)$, and the locations ζ of the receivers be drawn i.i.d. from a distribution $p(\zeta)$. Let the distributions $p(\xi)$ and $p(\zeta)$, and the uniform grid $\phi_{1:G}$ be such that relations (25) hold for $i = 2, \ldots, G$. Further, let

$$MN \geq C\left(K - \frac{1}{2}\right)^2 \left[\log \gamma + \frac{1}{2}\log(2\log \gamma)\right]^2 \tag{31}$$

where $\gamma \triangleq \sqrt{\pi}G/(2\epsilon)$, and the constant $C = (43 + 12\sqrt{7})/16 \approx 4.6718$. Then, with probability at least $1 - \epsilon$, for any K-sparse signal $\mathbf{x} \in \mathbb{C}^G$ measured from MN MIMO radar measurements $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, with $\|\mathbf{e}\|_2 \leq \sigma$, the solution $\hat{\mathbf{x}}$ of (15) satisfies

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq c\sigma, \tag{32}$$

where c is a constant that depends only on ϵ.

Proof: See Appendix E. □

The significance of (31) is to indicate the number of elements necessary to control the peak sidelobe. This is used to obtain a uniform recovery guarantee for spatial compressive sensing. In addition, the previous theorem ensures exact recovery of any K-sparse signal using (15) in the noise-free case $\sigma = 0$.

It is important to point out that the number of grid points G is not a free variable since $\phi_{1:G}$ must satisfy (25). This point will be explored in Section IV-D, where we show that the resolution G must be linearly proportional to the "virtual" array aperture Z.

Uniform recovery guarantees capture a worst case recovery scenario. Indeed, the average performance is usually much better than that predicted by uniform recovery guarantees. In the following section, we show that if we consider a non-uniform recovery guarantee, then the zero mean conditions (25) can be relaxed, and we obtain recovery guarantees that scale linearly with K.

C. Non-Uniform Recovery

We now investigate non-uniform recovery guarantees. In recent work [18], it has been shown that for a sufficient number of i.i.d. compressive sensing measurements, non-uniform recovery is guaranteed if isotropy holds. However, the result in [18] cannot be directly used in our framework since the MN rows of the matrix $\mathbf{A}$, following (11), are not i.i.d. This scenario is addressed in [19] in which non-uniform recovery is guaranteed for a MIMO radar system with N transceivers if the isotropy property (under the name *aperture condition*) holds. The following theorem derives conditions on grid points $\phi_{1:G}$ and probability distributions $p(\xi)$ and $p(\zeta)$, in order for the random matrix $\mathbf{A}$ to satisfy the isotropy property:

Theorem 3: Let the locations ξ of the transmit elements be drawn i.i.d. from a distribution $p(\xi)$, and the locations ζ of the receivers be drawn i.i.d. from a distribution $p(\zeta)$. For every t, the t-th row of $\mathbf{A}$ in (13) satisfies the isotropy property [18], i.e.,

$$\mathbb{E}[\mathbf{A}^H(t,:)\mathbf{A}(t,:)] = \mathbf{I}, \qquad (33)$$

iff $p(\xi)$, $p(\zeta)$ and $\phi_{1:G}$ are chosen such that, for $i = 2, \ldots, G$,

$$\psi_z(u_{1,i}) = 0, \qquad (34)$$

where $z \triangleq \zeta + \xi$ and $u_{1,i} \triangleq \pi Z(\phi_i - \phi_1)$.

Proof: See Appendix F. □

Theorem 3 links grid points $\phi_{1:G}$ and probability distributions $p(\xi)$ and $p(\zeta)$ (through the characteristic function of z) with the isotropy property of $\mathbf{A}$. When (34) holds, it can be shown that the aperture condition used in [19] holds too. Therefore, using the same approach as in [19], non-uniform recovery of K targets via (15) is guaranteed in the proposed spatial compressive sensing framework. The following Theorem customizes Theorem 2.1 in [19] to our framework:

Theorem 4: [Non-uniform recovery guarantee] Consider a K-sparse $\mathbf{x} \in \mathbb{C}^G$ measured from MN MIMO radar measurements $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where $\|\mathbf{e}\|_2 \leq \sigma$. Let $\varepsilon > 0$ be an arbitrary scalar, and suppose that the random matrix $\mathbf{A}$ satisfies the isotropy property, $\mathbb{E}[\mathbf{A}^H(t,:)\mathbf{A}(t,:)] = \mathbf{I} \, \forall t$. Then with probability at least $1 - \varepsilon$, the solution $\hat{\mathbf{x}}$ to (15) obeys

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq C_1 \sigma \sqrt{\frac{K}{MN}}, \qquad (35)$$

provided that the number of rows of $\mathbf{A}$ meets

$$MN \geq CK \log^2\left(\frac{cG}{\varepsilon}\right), \qquad (36)$$

where C_1, C and c are constants.

Proof: The theorem results from Theorem 2.1 in [19] by performing the following substitutions: K for s (sparsity), MN for n^2 (number of rows of $\mathbf{A}$), and G for N (number of columns of $\mathbf{A}$). Since in this work we consider K-sparse signals, in (35) we discarded the term that accounts for nearly-sparse signals present in [19]. □

Theorem 4 shows that, when the isotropy property is satisfied, the proposed framework enables us to localize K targets using about $MN = K(\log G)^2$ MIMO radar measurements.

Some comments are in order. First, it is important to stress that in (36), the number of elements scales linearly with the sparsity K. This is in contrast with uniform recovery bounds based on coherence (e.g., (31)), which scale quadratically with K. Moreover, the significance of the logarithmic dependence on G is that the proposed framework enables high resolution with a small number of MIMO radar elements. This is in contrast with a filled virtual MIMO array where the product MN scales linearly with G. Again, it is crucial to point out that the number of grid points G is not a free variable, because the grid points $\phi_{1:G}$ must satisfy (34). Second, differently from (35), in (32) the error did not depend on K, M and N. Third, (35) shows that reconstruction is stable even when the measurements are noisy. Additionally, we see from (35) that when $\sigma = 0$, Theorem 4 guarantees exact reconstruction with high probability, when (36) holds. Both results above can be extended to approximately sparse vectors, in which case an extra term appears in the right hand-side of (32) and (35). This situation may emerge when targets are not exactly on a grid, however, the analysis of such scenario is outside the scope of this paper. Finally, to suggest some intuition into the above conditions, notice that recovery can be guaranteed by requiring the matrix $\mathbf{A}$ to satisfy the isotropy property, $\mathbb{E}[\mathbf{Q}] = MN \cdot \mathbf{I}$, and by controlling the variances of the non-diagonal elements of $\mathbf{Q}$ (which, according to (22) and (23), scale with $1/MN$) through the use of a sufficient number of measurements MN.

D. Element Locations and Grid-points

We now provide an example of $p(\xi)$, $p(\zeta)$ and $\phi_{1:G}$ that meet the requirements of Theorem 1 and Theorem 3.

The conditions needed by each theorem constrain the characteristic function of the random variables ξ, ζ. Let, $Z_{TX} = Z_{RX} = Z/2$, such that the random variables ξ and ζ are both confined to the interval $[-\frac{1}{2}, \frac{1}{2}]$. The characteristic function of a uniform random variable $\zeta \sim \mathcal{U}[-\frac{1}{2}, \frac{1}{2}]$ is the sinc function, i.e.,

$$\psi_\zeta(u) = \frac{\sin(u/2)}{u/2}. \qquad (37)$$

Therefore, when ζ is uniformly distributed, by choosing $\phi_{1:G}$ as a uniform grid of $2/Z$-spaced points in the range $[-1, 1]$, we have that $\psi_\zeta(u_{i,l}) = \psi_\zeta(2u_{i,l}) = 0$ for any $i \neq l$ (since $u_{i,l} \triangleq \pi Z(\phi_l - \phi_i) = 2\pi|i - l|$). The number of grid points G is not a free variable, because the grid points $\phi_{1:G}$ must satisfy (25) or (34). For instance, in the example above, $\phi_{1:G}$ must be a uniform

grid of $2/Z$-spaced points between $[-1, 1]$, and, assuming that Z is an integer, the number of grid points is $G = Z + 1$.

The dependence between the number of grid points G and the virtual array aperture Z can be understood by noticing that both (25) and (34) impose that grid points are placed at the zeros of the characteristic function of the relative random variable (i.e., the sinc function). The spacing of the zeros is dictated by the virtual array aperture Z. The bigger the aperture the more grid points fit in the range $[-1, 1]$.

Summarizing, choose $\phi_{1:G}$ as a uniform grid of $2/Z$-spaced points in the range $[-1, 1]$. Then:

1) If *both* ζ and ξ are uniformly distributed, relations (25) hold, and we can invoke Theorem 1 (for uniform recovery);

2) If *either* ζ or ξ are uniformly distributed, relation (34) holds, and we can invoke Theorem 3 (for non-uniform recovery).

Note that non-uniform recovery, i.e., (34), requires only one density function, say $p(\zeta)$, to be uniform, while the other distribution, $p(\xi)$, can be arbitrarily chosen, e.g., it can be even deterministically dependent on ζ. For instance, (34) is satisfied in a MIMO radar system with N transceivers, i.e., when $\zeta_{1:N}$ are i.i.d. uniform distributed and we deterministically set $\xi_n = \zeta_n$.

Finally, we remark that the analysis provided in this section regarding the statistics of the matrix $\mathbf{A}$ may be used with block-sparsity results in the compressive sensing literature [4] to obtain guarantees for the general MMV scenario.

V. NUMERICAL RESULTS

In this section, we present numerical results illustrating the proposed spatial compressive sensing framework.

We design an example to follow Theorem 1, in which $p(\xi)$ and $p(\zeta)$ are both uniform distributions, and $\phi_{1:G}$ represents a uniform grid of $2/Z$-spaced points in the interval $[-1, 1]$, which implies that the number of grid points is $G = Z + 1$. The system transmits a total of P pulses. When expressed in discrete form, each pulse consists of M orthogonal codes composed by M symbols. In particular, we select the codes to be the rows of the $M \times M$ Fourier matrix. Equal length apertures were assumed for the transmit and receive arrays, i.e., $Z_{TX} = Z_{RX} = Z/2$. The target gains were given by $x_k = \exp(-j\varphi_k)$, with φ_k drawn i.i.d., uniform over $[0, 2\pi)$, for all $k = 1, \ldots, K$ (where K is the number of targets). The noise (see (13)) was assumed to be distributed as $\mathrm{vec}(\mathbf{E}) \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$ and the SNR is defined as $-10\log_{10}\sigma^2$. From the definition of the measurement matrix $\mathbf{A}$, its columns all have squared-norm equal to MN. Throughout the numerical results, we normalize the columns of $\mathbf{A}$ to have unit norm.

We first investigate the statistics of the matrix $\mathbf{Q}$ discussed in Section III. In particular, we analyze the coherence μ of the measurement matrix $\mathbf{A}$ compared to the result given in Theorem 1. The virtual aperture was $Z = 250$ (thus $G = 251$). In Fig. 2, we plot the ccdf of the coherence μ, i.e., $\Pr(\mu > q)$, as a function of the number of elements for (a) the M transmitters and N receivers setup and (b) the N transceivers setup. As a reference, we also plot the upper bound given in (29) and (30), respectively. It can be seen how the upper bound becomes tighter and tighter as the number of elements increases. In addition, it is interesting to notice that the coherence of the matrix $\mathbf{A}$ for the N' transceivers setup is very close to the coherence of
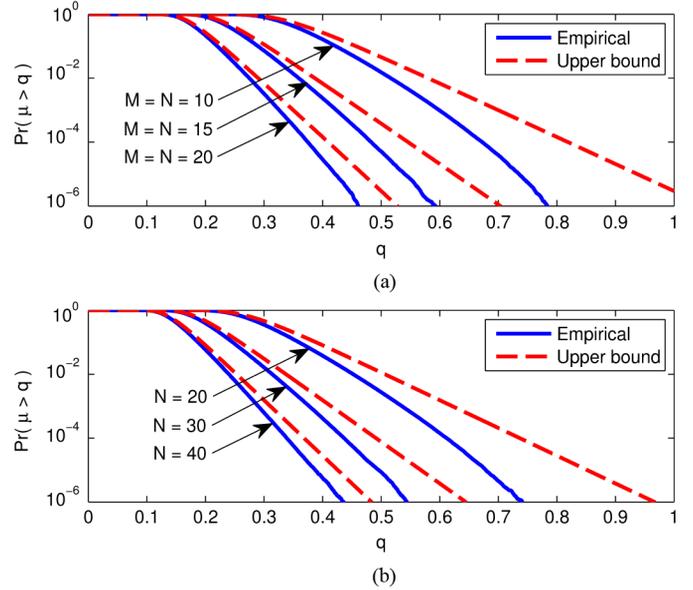
Fig. 2. Empirical ccdf of the coherence of the measurement matrix $\mathbf{A}$ and its upper bound as a function of the number of elements. (a) considers the M transmitters and N receivers setup and the upper bound is given in (29); (b) considers the N transceivers setup and the upper bound is given in (30).

the matrix $\mathbf{A}$ for the the M transmitters and N receivers setup when $M = N = N'/2$.

We next present localization performance using practical algorithms. We implemented target localization using LASSO following the algorithm proposed in [19] to solve problem (15). In addition, we implement Beamforming, OLS, OMP, CoSaMP, FOCUSS [31] and MBMP. In the MMV setup we also compare MBMP, RA-ORMP [26], M-FOCUSS [31], and MUSIC [32]. Concerning MBMP, it requires as input a K length branch vector $\mathbf{d}$, which define the algorithm's complexity (see [27] for details on setting parameters for MBMP). The output of MBMP is the estimated support. Notice that, when $\mathbf{d} = [1, \ldots, 1]$, MBMP reduces to OLS in the SMV scenario, and to RA-ORMP in the MMV scenario. We define the support recovery error when the estimated support does not coincide with the true one. For algorithms that return an estimate $\hat{\mathbf{x}}$ of the sparse vector $\mathbf{x}$ (e.g., LASSO, FOCUSS and MUSIC), the support was then identified as the K largest modulo entries of the signal $\hat{\mathbf{x}}$.

Analyzing how to set the noise parameter σ in (14), or the sparsity K, without prior information is the topic of current work [21], but outside the scope of this paper. Therefore, we assume that the noise level is available and that the number of targets K is known (notice that this information is needed by all the algorithms including MBMP). The virtual aperture was $Z = 250$ (thus $G = 251$), and tests were carried out for $K = 5$ targets. The SNR was 20 dB throughout.

The main focus of the paper is to reduce the number of antenna elements while avoiding sidelobes errors and while preserving the high-resolution provided by the virtual array aperture Z (i.e., recovering $2/Z$-spaced targets). Therefore, to account for errors due to sidelobes (an erroneous target is estimated at a sidelobe location) and unresolved targets (the responses of two targets in consecutive grid-points is merged in only one grid-point), we consider as performance metric the support recovery error probability, defined as the error event when at least one target is estimated erroneously.
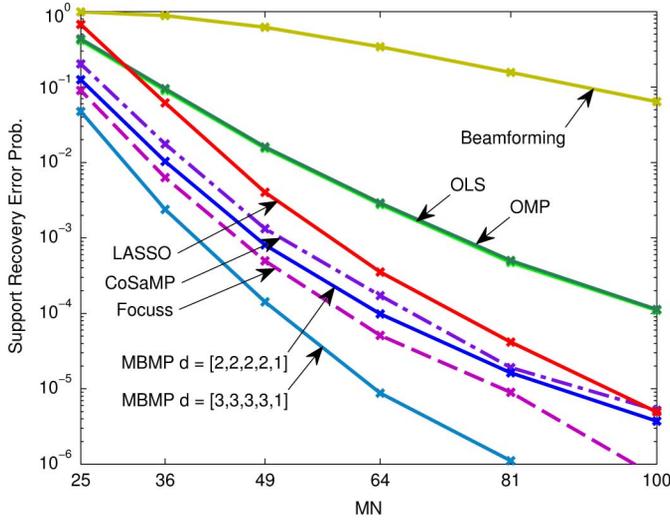
Fig. 3. Probability of support recovery error as a function of the number of rows MN of $\mathbf{A}$. Non-uniform SMV setup. The system settings are $Z = 250$, $G = 251$, $P = 1$ and $K = 5$ targets with $|x_k| = 1$ for all k. The SNR is 20 db.

Fig. 4. Probability of support recovery error as a function of the number of rows MN of $\mathbf{A}$. Uniform SMV setup. The system settings are $Z = 250$, $G = 251$, $P = 1$ and $K = 5$ targets with $|x_k| = 1$ for all k. The SNR is 20 db.

We first treat the non-uniform guarantee setting. Monte Carlo simulations were carried out using independent realizations of target gains, targets locations, noise and element positions. Fig. 3 illustrates the probability of support recovery error as a function of the number of measurements MN. From the figure, it can be seen that compressive sensing algorithms enable better performance (smaller probability of sidelobe error and better resolution) than beamforming, which is not well-suited for the sparse recovery framework. Among compressive sensing algorithms, two main groups appear: on one side, OLS and OMP, which both have practically the same performance; on the other side, LASSO, CoSaMP, FOCUSS and MBMP. Among the latter group, it is important to point out that, although the recovery guarantee established in Theorem 4 requires the solution of (15), and thus using LASSO, MBMP provides a viable and competitive way to perform target localization.

We next consider uniform guarantees. In this setup, we first generate a realization of the matrix $\mathbf{A}$ by drawing at random the element positions. Maintaining the matrix $\mathbf{A}$ fixed, we perform 500 Monte Carlo simulations using independent realizations of target gains, targets locations and noise. For each recovery method, we defined a support recovery error if an error occur in any of the 500 simulations. We then average throughout element positions realizations. Fig. 4 illustrates the probability of support recovery error as a function of the number of measurements MN. The difference among OLS/OMP and the more sophisticated methods (i.e., LASSO, CoSaMP, FOCUSS and MBMP) is even more evident in this setup (e.g., at $MN = 81$, the probability of OLS/OMP error is greater than 0.1), confirming the theoretical finding [33] of OLS/OMP unfitness to deliver uniform recovery. On the other hand, MBMP, an extension of OLS, still provides competitive performance. In particular, MBMP with $\mathbf{d} = [3, 3, 3, 3, 1]$ outperforms the other methods.

The theoretical results presented in this work focus on the SMV setting. However, in practice several snapshots can be available. To explore the benefits of the proposed MIMO random array framework in such case, in Fig. 5 we consider an MMV setting ($P = 5$) and we compare sparse recovery
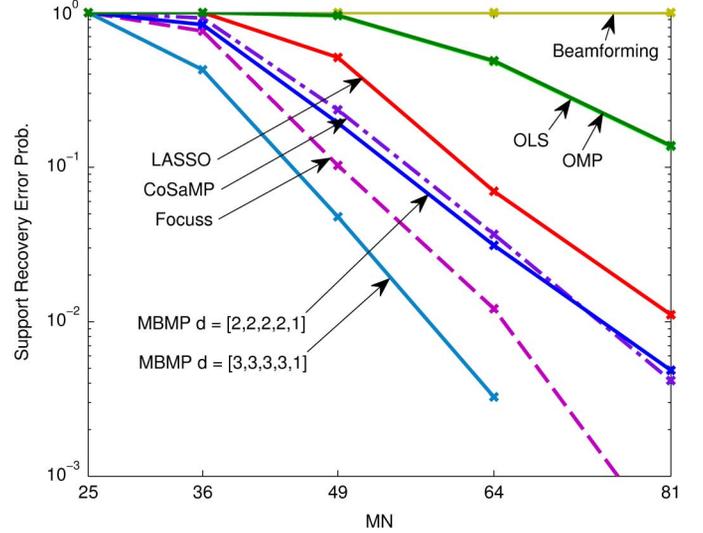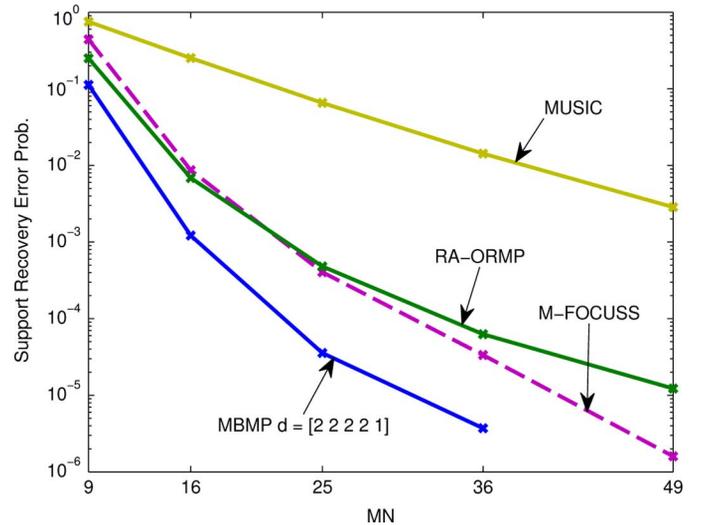
Fig. 5. Probability of support recovery error as a function of the number of rows MN of $\mathbf{A}$. Non-uniform MMV setup. The system settings are $Z = 250$, $G = 251$, $P = 5$ and $K = 5$ targets with $|x_{k,p}| = 1$ for all k, p. The SNR is 20 db.

methods with the well-known MUSIC algorithm. We evaluate five different elements configurations: $[M, N] = [3, 3], [4, 4], [5, 5], [6, 6]$ and $[7, 7]$. The figure illustrates the probability of support recovery error as a function of the number of measurements MN (nonuniform setup). Sparse recovery algorithms have better performances than MUSIC, and the availability of multiple snapshots allows to considerably reduce the number of antenna elements. Moreover, in the MMV setting, algorithms which are able to exploit the signal subspace information (e.g., MBMP and RA-ORMP) posses a clear advantage over those algorithms that are unable (e.g., M-FOCUSS). For instance, this can be appreciated by the difference in performance of FOCUSS and MBMP with $\mathbf{d} = [2, 2, 2, 2, 1]$ when comparing the SMV (Fig. 3) and MMV (Fig. 5) settings. The numerical simulations presented in this paper considered a medium SNR level and show a superior performance of sparse recovery methods over classical methods (e.g., beamforming or MUSIC)

in the proposed framework. Since the sparsity property, upon which sparse recovery methods rely, is independent from the SNR, we expect a similar behavior also at low SNR (e.g., SNR = 0 dB or lower).

VI. CONCLUSIONS

We propose a sparse framework to address the source localization problem for a random array MIMO radar system. We link system design quantities, i.e., the probability distributions $p(\xi)$ and $p(\zeta)$ of the tx/rx sensors location and the sparse localization grid points $\phi_{1:G}$, with the statistics of the Gram matrix $\mathbf{Q}$ and the related coherence of the matrix $\mathbf{A}$. Based on this result, we were able to develop uniform and non-uniform recovery guarantees for spatial compressive sensing. We show that within the proposed framework, it is possible to localize K targets using about $MN = K(\log G)^2$ MIMO radar noisy measurements, where G is proportional to the array aperture and determines the angle resolution. In other words, the proposed framework supports the high-resolution provided by the virtual array aperture while using a reduced number of MIMO radar elements. This is in contrast with a filled virtual MIMO array for which the product MN scales linearly with G. Moreover, since the results characterize the product of the number of transmit and receive elements, MIMO random array implementation further reduces the total number of antenna elements needed. From numerical simulations it emerges that, in the proposed framework, compressive sensing recovery algorithms (e.g., MBMP) are capable of better performance (i.e., smaller probability of sidelobe errors and better resolution) than classical methods, such as beamforming and MUSIC.

APPENDIX

Proof of Proposition 1:

Mean: The mean $\eta(u)$ is by definition the expectation of the random array pattern, i.e., $\mathbb{E}[\beta(u)]$, over $z_{mn} = \xi_m + \zeta_n$. The expectation and the summations can be interchanged obtaining

$$\eta(u) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \mathbb{E}[\exp(juz_{mn})]. \tag{38}$$

Moreover, the average of $\exp[ju_{i,l}(\zeta_n + \xi_m)]$ does not depend on the index n and m, since $\zeta_{1:N}$ are identically distributed, and so are $\xi_{1:M}$. By dropping the indexes of ζ_n and ξ_m and using $z = \xi + \zeta$, we have the sum of MN identical terms, divided by MN. Thus $\eta(u)$ equals $\mathbb{E}[\exp(juz)]$, the characteristic function of the random variable z.

Variance: Let $\sigma_1^2(u) = \mathrm{Re}\beta(u)$ and $\sigma_2^2(u) = \mathrm{Im}\beta(u)$. For brevity of notation, we drop the dependency on u. First notice that since $p(z)$ is even, its characteristic function is real, thus so is the mean value of the array pattern η. We also have that $\mathbb{E}[(\mathrm{Re}\beta - \eta)\mathrm{Im}\beta] = \mathbb{E}[\mathrm{Re}\beta\,\mathrm{Im}\beta] - \eta\mathbb{E}[\mathrm{Im}\beta] = 0$, since the real and imaginary parts are uncorrelated and because $\mathbb{E}[\mathrm{Im}\beta] = 0$. Next, we need to evaluate $\sigma_1^2 \triangleq \mathbb{E}[(\mathrm{Re}\beta - \eta)^2]$ and $\sigma_2^2 \triangleq \mathbb{E}[(\mathrm{Im}\beta)^2]$. In order to derive these quantities, we consider the expectations given by $\mathbb{E}[(\beta - \eta)^2]$ and $\mathbb{E}[|\beta - \eta|^2]$. It can be shown that,

$$\mathbb{E}[(\beta - \eta)^2] = \sigma_1^2 - \sigma_2^2 + j2\sigma_{12} \tag{39}$$

and

$$\mathbb{E}[|\beta - \eta|^2] = \sigma_1^2 + \sigma_2^2. \tag{40}$$

Substituting the definition of the random array pattern $\beta(u)$ (18) and (21) in (39) and (40), we obtain (22) and (23).

Proof of Lemma 1: From (18) we have that $\mathbf{a}_i^H \mathbf{a}_l = MN \cdot \beta(u_{i,l})$, where $u_{i,l} \triangleq \pi Z(\phi_l - \phi_i)$. When $\phi_{1:G}$ is a uniform grid, $\phi_l - \phi_i$ is constant whenever $i - l$ is constant, i.e., along every diagonal of the matrix $\mathbf{Q}$. Since $\beta(u_{i,l})$ depends only on the term $\phi_l - \phi_i$ (not on the actual ϕ_i and ϕ_j), $\mathbf{Q}$ is a Toeplitz matrix.

Proof of Theorem 1: We define the array pattern associated with the transmitter as

$$\beta_\zeta(u_{i,l}) \triangleq \frac{1}{N} \sum_{n=1}^{N} \exp[ju_{i,l}\zeta_n] = \frac{1}{N}\mathbf{b}_i^H\mathbf{b}_l \tag{41}$$

and with the receiver arrays as:

$$\beta_\xi(u_{i,l}) \triangleq \frac{1}{M} \sum_{m=1}^{M} \exp[ju_{i,l}\xi_m] = \frac{1}{M}\mathbf{c}_i^H\mathbf{c}_l. \tag{42}$$

Statistical properties of random arrays were analyzed in [5] in the case of passive localization (i.e., an array with only receiving elements). The following lemma customizes useful results from [5]:

Lemma 2: Let the locations $\zeta_{1:N}$ of the receiving array be i.i.d., drawn from an even distribution $p(\zeta)$ and consider a given u. Then $\beta_\zeta(u)$ is asymptotically jointly Gaussian distributed (we neglect the dependency on u):

$$\begin{bmatrix} \mathrm{Re}\beta_\zeta \\ \mathrm{Im}\beta_\zeta \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathrm{Re}\psi_\zeta \\ \mathrm{Im}\psi_\zeta \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right) \tag{43}$$

where $\sigma_1^2(u) = \frac{1}{2N}[1 + \psi_\zeta(2u)] - \frac{1}{N}\psi_\zeta^2(u)$ and $\sigma_2^2(u) = \frac{1}{2N}[1 - \psi_\zeta(2u)]$.

Proof: See [5]. $\square$

The joint distribution of $\mathrm{Re}\beta_\xi(u)$ and $\mathrm{Im}\beta_\xi(u)$ can be obtained similarly.

For a given $i \in \{2, \ldots, G\}$, using Lemma 1 and the assumption that the mean patterns of both the transmitter and receiver arrays satisfy (25), i.e., $\psi_\xi(u_{1,i}) = \psi_\xi(2u_{1,i}) = \psi_\zeta(u_{1,i}) = \psi_\zeta(2u_{1,i}) = 0$, we have that, for both transmitter and receiver arrays, the array pattern evaluated at any grid point is being drawn from an asymptotically complex normal distribution with variance defined by the number of transmit and receive elements, i.e., $\beta_\xi(u_{1,i}) \sim \mathcal{CN}(0, \frac{1}{M})$ and $\beta_\zeta(u_{1,i}) \sim \mathcal{CN}(0, \frac{1}{N})$. It follows that the random variable $q = \frac{1}{N}|\mathbf{b}_1^H\mathbf{b}_i|$ can be approximated as belonging to Rayleigh distribution, i.e., $p(q) = (q/\sigma^2)\exp(-q^2/2\sigma^2)$, where $\sigma^2 = 1/(2N)$, and similarly the random variable $\frac{1}{N}|\mathbf{c}_1^H\mathbf{c}_i|$ is governed by a Rayleigh distribution with variance $\sigma^2 = 1/(2M)$.

If ξ and ζ are independent (part 1), the two random variables $\frac{1}{N}|\mathbf{b}_1^H\mathbf{b}_i|$ and $\frac{1}{N}|\mathbf{c}_1^H\mathbf{c}_i|$ are independent. Using (24), we have that the distribution of $\frac{1}{MN}|\mathbf{a}_1^H\mathbf{a}_i|$ is the product of two independent Rayleigh distributed variables. The cumulative density function of such a variable is given in [34]. It follows that the ccdf of $\frac{1}{MN}|\mathbf{a}_1^H\mathbf{a}_i|$ satisfies

$$\Pr\left(\frac{1}{MN}|\mathbf{a}_1^H\mathbf{a}_i| > q \right) < x \cdot K_1(x), \tag{44}$$

where $x \triangleq 2\sqrt{MN}q$.

If $\xi_n = \zeta_n$ for all n (part 2), by using (24), we have that $\frac{1}{N^2}|\mathbf{a}_1^H \mathbf{a}_i| = \left(\frac{1}{N}|\mathbf{b}_1^H \mathbf{b}_i|\right)^2$. Since the random variable $\frac{1}{N}|\mathbf{b}_1^H \mathbf{b}_i|$ has a Rayleigh distribution, $\frac{1}{N^2}|\mathbf{a}_1^H \mathbf{a}_i|$ is distributed as the square of a Rayleigh distribution, which has cdf $1 - \exp(-Nq)$. As such, its ccdf satisfies

$$\Pr\left(\frac{1}{N^2}|\mathbf{a}_1^H \mathbf{a}_i| > q\right) < \exp(-Nq). \qquad (45)$$

Part 3 follows because, from (24), the phase of $\mathbf{a}_1^H \mathbf{a}_i$ is the sum of the phases of $\mathbf{b}_1^H \mathbf{b}_i$ and $\mathbf{c}_1^H \mathbf{c}_i$. In the case of transceivers the phase of $\mathbf{a}_1^H \mathbf{a}_i$ is evidently uniform since it is the same phase of $\mathbf{b}_1^H \mathbf{b}_i$. In the case of M transmitter and N receivers, since both $\mathbf{b}_1^H \mathbf{b}_i$ and $\mathbf{c}_1^H \mathbf{c}_i$ are two independent circular symmetric complex normal variables, the sum of the phases is itself uniformly distributed over $[0, 2\pi)$.

Proof of Corollary 1: We take the conservative assumption of independence between the $G-1$ random variables $|\frac{1}{MN}\mathbf{a}_1^H \mathbf{a}_i|$, for $i = 2, \ldots, G$. If ξ and ζ are independent (part 1), from (26), the ccdf of the maximum among $G-1$ such variables (which gives the coherence), is upper bounded by

$$\Pr\left(\max_{i>1}\left|\frac{1}{MN}\mathbf{a}_1^H \mathbf{a}_i\right| > q\right) < 1 - [1 - x \cdot K_1(x)]^{G-1}, \qquad (46)$$

where $x \triangleq 2\sqrt{MN}q$.

If $\xi_n = \zeta_n$ for all n (part 2), by using (27), the ccdf of the maximum among $G-1$ such variables, is upper bounded by

$$\Pr\left(\max_{i>1}\left|\frac{1}{N^2}\mathbf{a}_1^H \mathbf{a}_i\right| > q\right) < 1 - [1 - \exp(Nq)]^{G-1}. \qquad (47)$$

This concludes the proof.

Proof of Theorem 2: The theorem follows by combining the claims of Theorem 2.7 in [30] and Corollary 1. Theorem 2.7 in [30] provides stable recovery guarantees for any K-sparse signal if the measurement matrix $\mathbf{A}$ has RIP $\delta_{2K} < 2/(3 + \sqrt{7/4}) \triangleq \alpha$. The goal is therefore to bound the RIP δ_{2K} of the spatial compressive sensing measurement matrix $\mathbf{A}$ with probability higher than $1 - \epsilon$. In other words, we want to find how many measurements MN we need to satisfy $\Pr(\delta_{2K} \leq \alpha) \geq 1 - \epsilon$. By using $\delta_{2K} \leq (2K-1)\mu$ [30], we have that $\delta_{2K} \leq \alpha$ if $(2K-1)\mu \leq \alpha$. Moreover, the condition $\Pr((2K-1)\mu \leq \alpha) \geq 1 - \epsilon$ is equivalent to $\Pr(\mu > \alpha/(2K-1)) < \epsilon$. Therefore by invoking (29) in Corollary 1, we can write

$$\Pr(\delta_{2K} > \alpha) < 1 - [1 - x \cdot K_1(x)]^{G-1} \qquad (48)$$

where, by combining $x \triangleq 2\sqrt{MN}q$ and $q = \alpha/(2K-1)$, we have $x = 2\sqrt{MN}\alpha/(2K-1)$. We thus look for the value MN that makes the right hand-side equal to ϵ.

We first approximate the modified Bessel function of the second kind $K_1(x)$ for a large absolute value and small phase of the argument (in our setting, the argument x is real) [35]: $K_1(x) \approx \sqrt{\frac{\pi}{2x}}\exp(-x)$. We thus would like to enforce $1 - [1 - \sqrt{\frac{\pi x}{2}}\exp(-x)]^{G-1} = \epsilon$. Defining $t \triangleq \sqrt{\frac{\pi x}{2}}\exp(-x)$ and linearizing the function $(1-t)^{G-1}$ around $t = 0$, we obtain $G\sqrt{\frac{\pi x}{2}}\exp(-x) = \epsilon$, where, for simplicity, we used G in place of $G-1$. This equation can be rewritten in the form $-2x\exp(-2x) = -\gamma^{-2}$, where $\gamma \triangleq \frac{\sqrt{\pi}G}{2\epsilon}$. The inverse function of such equation is called the Lambert W function [36]. For real arguments, it is not injective, therefore it is divided in two branches: $x > 1/2$ or $x \leq 1/2$. Since in our setup $x \leq 1/2$, the lower branch, denoted W_{-1}, is considered,

and our solution satisfies $-2x = W_{-1}(-\gamma^{-2})$. By using the asymptotic expansion $W_{-1}(-\gamma^{-2}) \approx -2\ln\gamma - \ln(2\ln\gamma)$ and solving for MN we obtain (31). The claim of the theorem follows from Theorem 2.7 in [30]. Finally, since in this work we consider K-sparse signals, in the error term (32), we discarded the term for nearly-sparse signals present in [30].

Proof of Theorem 3: Because the variables $\zeta_{1:N}$ are identically distributed, and so are $\xi_{1:M}$, the average $\mathbb{E}[\mathbf{A}^H(t, :)\mathbf{A}(t, :)]$ does not depend on the index $t = N(m-1) + n$, where the last relation follows from the definition of $\mathbf{a}_g$. Therefore, we have $\mathbb{E}[\mathbf{A}^H(t, :)\mathbf{A}(t, :)] = \frac{1}{MN}\sum_{t=1}^{MN}\mathbb{E}[\mathbf{A}^H(t, :)\mathbf{A}(t, :)] = \frac{1}{MN}\mathbb{E}[\mathbf{Q}]$. Thanks to Lemma 1, we can focus only on the first row of $\frac{1}{MN}\mathbb{E}[\mathbf{Q}]$. Using (18), the elements of the first row of such matrix are $\eta(u_{1,i})$ for $i = 1, \ldots, G$. From (21) in Proposition 1, we know that $\eta(u_{1,i}) = \psi_z(u_{1,i})$. Thus, requiring (34), i.e., $\psi_z(u_{1,i}) = 0$ for $i = 2, \ldots, G$, together with the fact that $\exp(jzu_{1,1}) = 1$ (because $u_{1,1} \triangleq \pi Z(\phi_1 - \phi_1) = 0$), gives the "if" direction of the claim.

The "only if" direction follows by noticing that when (34) is not satisfied there will be at least one i such that $\eta(u_{1,i}) \neq 0$. Therefore, the matrix $\mathbf{A}$ does not satisfy the isotropy property, showing that (34) is also a necessary condition.

References

[1] A. M. Haimovich, R. Blum, and L. Cimini, "MIMO radar with widely separated antennas," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 116–129, 2008.

[2] P. Stoica and J. Li, "MIMO radar with colocated antennas," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 106–114, Sep. 2007.

[3] H. L. VanTrees, *Detection, Estimation and Modulation Theory: Optimum Array Processing*. New York, NY, USA: Wiley, 2002, vol. 4.

[4] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[5] Y. Lo, "A mathematical theory of antenna arrays with randomly spaced elements," *IEEE Trans. Antennas Propag.*, vol. 12, no. 3, pp. 257–268, May 1964.

[6] M. A. Haleem and A. M. Haimovich, "On the distribution of ambiguity levels in MIMO radar," presented at the 42nd Asilomar Conf. Signals, Syst. Comput., Pacific Grove, CA, USA, Oct. 2008.

[7] J. H. G. Ender, "On compressive sensing applied to radar," *Elsevier J. Signal Process.*, vol. 90, pp. 1402–1414, Nov. 2009.

[8] K. Gedalyahu and Y. C. Eldar, "Time-delay estimation from low-rate samples: A union of subspaces approach," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3017–3031, Jun. 2010.

[9] W. U. Bajwa, K. Gedalyahu, and Y. C. Eldar, "Identification of parametric underspread linear systems with application to super-resolution radar," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2548–2561, Aug. 2011.

[10] O. Bar-Ilan and Y. C. Eldar, "Sub-Nyquist radar via Doppler focusing," in *IEEE Trans. Signal Process.*, Nov. 2012 [Online]. Available: http://arxiv.org/abs/1211.0722,, ArXiv preprint, submitted for publication

[11] E. Baransky, G. Itzhak, I. Shmuel, N. Wagner, E. Shoshan, and Y. C. Eldar, "A sub-Nyquist radar prototype: Hardware and algorithms," *IEEE Trans. Aerosp. Electron. Syst., Special Issue on Compressed Sensing for Radar*, Aug. 2012 [Online]. Available: http://arxiv.org/abs/1208.2515, ArXiv preprint, submitted for publication

[12] S. Gogineni and A. Nehorai, "Target estimation using sparse modeling for distributed MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, pp. 5315–5325, Nov. 2011.

[13] T. Strohmer and B. Friedlander, "Compressed sensing for MIMO radar — algorithms and performance," presented at the 43rd Asilomar Conf. Signals, Syst., Comput., Pacific Grove, CA, USA, Nov. 2009.

[14] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, 2005.

[15] T. Strohmer and B. Friedlander, "Analysis of sparse MIMO radar," 2012 [Online]. Available: http://arxiv.org/abs/1203.2690, ArXiv preprint.

[16] L. Carin, "On the relationship between compressive sensing and random sensor arrays," *IEEE Antennas Propag. Mag.*, vol. 5, no. 5, pp. 72–81, Oct. 2009.

[17] Y. Yu, A. P. Petropulu, and H. V. Poor, "MIMO radar using compressive sampling," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 146–163, Feb. 2010.

[18] E. J. Candes and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7235–7254, 2011.

[19] M. Hügel, H. Rauhut, and T. Strohmer, "Remote sensing via l_1 minimization," , 2012 [Online]. Available: ArXiv preprint arXiv:1205.1366

[20] M. Skolnik, *Introduction to Radar Systems*, 3rd ed. New York, NY, USA: McGraw-Hill, 2002.

[21] M. Rossi, A. M. Haimovich, and Y. C. Eldar, "Compressive sensing with unknown parameters," in *Proc. 46th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 4-7, 2012.

[22] J. V. DiFranco and W. L. Rubin, "Spatial ambiguity and resolution for array antenna systems," *IEEE Trans. Military Electron.*, vol. 9, no. 3, pp. 229–237, Jul. 1965.

[23] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," presented at the 27th Asilomar Conf. Signals, Syst., Comput., Pacific Grove, CA, USA, Nov. 1993.

[24] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *Int. J. Control*, vol. 50, no. 5, pp. 1873–1896, 1989.

[25] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Computat. Harmon. Anal.*, vol. 26, pp. 301–321, 2008.

[26] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1135–1146, Jun. 2012.

[27] M. Rossi, A. M. Haimovich, and Y. C. Eldar, "Spatial compressive sensing in MIMO radar with random arrays," presented at the CISS , Princeton, NJ, USA, Mar. 21–23, 2012.

[28] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[29] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing Concepts and Techniques*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[30] H. Rauhut, "Compressive sensing and structured random matrices," *Theoret. Found. Numer. Methods Sparse Recov.*, vol. 9, pp. 1–92, 2010.

[31] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.

[32] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[33] H. Rauhut, "On the impossibility of uniform sparse reconstruction using greedy methods," *Sampl. Theory Signal Image Process.*, vol. 7, no. 2, pp. 197–215, 2008.

[34] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook For Engineers and Scientists*. Dordrecht, The Netherlands: Springer Netherlands, 2002.

[35] M. A. S. AbramowitzI, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards*. Washington, DC, USA: U.S. Dept. Commerce, 1972.

[36] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Adv. Comput. Math.*, vol. 5, pp. 329–359, 1996.

Marco Rossi (S'10) received the B.Sc. and M.Sc. degrees in information engineering from Politecnico di Milano, Milan, Italy, in 2004 and 2007, respectively.

From November 2006 to November 2007, as part of his M.Sc. thesis, he worked in the Radio System Technology Research Group at Nokia-Siemens Network, Milan, Italy. He is currently a Ph.D. candidate in electrical and computer engineering at the Center for Wireless Communications and Signal Processing Research (CWCSPR), New Jersey Institute of Technology (NJIT), Newark. His research interests are in optimization methods, with an emphasis on non-convex and discrete problems, and their applications to signal processing and wireless communications.

Mr. Rossi was a recipient of the Ross Fellowship Scholarship 2010–2012.

Alexander M. Haimovich (S'82–M'87–SM'97–F'12) received the B.Sc. degree in electrical engineering from the Technion-Israel Institute of Technology, Haifa, in 1977, the M.Sc. degree in electrical engineering from Drexel University, in 1983, and the Ph.D. degree in systems from the University of Pennsylvania, Philadelphia, in 1989.

From 1983 to 1990, he was a design engineer and staff consultant at AEL Industries. He served as Chief Scientist of JJM Systems from 1990 until 1992. He is the Ying Wu endowed Chair and a Professor of Electrical and Computer Engineering at the New Jersey Institute of Technology (NJIT), Newark, where he has been on the faculty since 1992. His research interests include MIMO radar, source localization, communication systems, and wireless networks.

Yonina C. Eldar (S'98–M'02–SM'07–F'12) received the B.Sc. degree in physics and the B.Sc. degree in electrical engineering both from Tel-Aviv University (TAU), Tel-Aviv, Israel, in 1995 and 1996, respectively, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 2002.

From January 2002 to July 2002, she was a Postdoctoral Fellow at the Digital Signal Processing Group, MIT. She is currently a Professor in the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa, and holds The Edwards Chair in Engineering. She is also a Research Affiliate with the Research Laboratory of Electronics at MIT and a Visiting Professor at Stanford University, Stanford, CA. Her research interests are in the broad areas of statistical signal processing, sampling theory and compressed sensing, optimization methods, and their applications to biology and optics.

Dr. Eldar was in the program for outstanding students at TAU from 1992 to 1996. In 1998, she held the Rosenblith Fellowship for study in electrical engineering at MIT, and in 2000, she held an IBM Research Fellowship. From 2002 to 2005, she was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. In 2004, she was awarded the Wolf Foundation Krill Prize for Excellence in Scientific Research, in 2005 the Andre and Bella Meyer Lectureship, in 2007 the Henry Taub Prize for Excellence in Research, in 2008 the Hershel Rich Innovation Award, the Award for Women with Distinguished Contributions, the Muriel & David Jacknow Award for Excellence in Teaching, and the Technion Outstanding Lecture Award, in 2009 the Technion's Award for Excellence in Teaching, in 2010 the Michael Bruno Memorial Award from the Rothschild Foundation, and in 2011 the Weizmann Prize for Exact Sciences. In 2012 she was elected to the Young Israel Academy of Science and to the Israel Committee for Higher Education. In 2013 she received the Technion's Award for Excellence in Teaching, and the Hershel Rich Innovation Award. She received several best paper awards together with her research students and colleagues. She is a Signal Processing Society Distinguished Lecturer, and Editor in Chief of Foundations and Trends in Signal Processing. In the past, she was a member of the IEEE Signal Processing Theory and Methods and Bio Imaging Signal Processing technical committees, and served as an associate editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the EURASIP *Journal of Signal Processing*, the SIAM *Journal on Matrix Analysis and Applications*, and the SIAM *Journal on Imaging Sciences*.