Spatial Compressive Sensing in MIMO Radar with Random Arrays

Marco Rossi[(*)], Alexander M. Haimovich[(*)], and Yonina C. Eldar[(†)]

[(*)]CWCSPR, New Jersey Institute of Technology, Newark, NJ07102-1982, USA

[(†)]Department of EE Technion, Israel Institute of Technology, Haifa 32000, Israel

Email: {marco.rossi, alexander.m.haimovich}@njit.edu, yonina@ee.technion.ac.il

Abstract— We study compressive sensing in the spatial domain for target localization using MIMO radar. By leveraging a joint sparse representation, we extend the single-pulse framework proposed in [1] to a multi-pulse one. For this scenario, we devise a tree-based matching pursuit algorithm to solve the nonconvex localization problem. It is shown that this method can achieve high resolution target localization with a highly undersampled MIMO radar with transmit/receive elements placed at random. Moreover, a lower bound is developed on the number of transmit/receive elements required to ensure accurate target localization with high probability.

I. INTRODUCTION

Detection, estimation, 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, in terms of location or Doppler, since closely spaced targets are more difficult to discriminate. In multiple input multiple output (MIMO) radar [2], targets are probed with multiple, simultaneous waveforms. Relying on the orthogonality of the transmitted waveform, returns from the targets are jointly processed by multiple receive antennas. 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. MIMO radars achieve these advantages by capitalizing on a larger number of degrees of freedom than "conventional" radar. In this work, we focus on the application of MIMO radar to the estimation of direction-of-arrival (DOA). We are particularly interested in a sparse, random array architecture in which a low number of transmit/receive elements are placed at random over a large aperture. We analyze this system from a compressive sensing point of view and propose an algorithm for DOA estimation.

It is well known in array signal processing [3] that resolution improves with the array aperture. A non-ambiguous uniform linear array (ULA) must have its elements spaced at intervals no larger than $\lambda/2$. For a MIMO radar, unambiguous direction finding of targets is possible for $\lambda/2$-spaced receive elements and $N\lambda/2$-spaced transmit elements (a virtual filled array),

where N is the number of receive elements. In compressive sensing parlance, the $\lambda/2$-spaced array and the MIMO virtual filled array perform spatial sampling at Nyquist rate. In this work, we are interested in a random array setup in which the spatial sampling is at sub-Nyquist rates. Recovering targets from undersampled array data, links random arrays to the compressive sensing paradigm (see [1] and references therein).

Whereas compressive sensing literature originated in the Single Measurement Vector (SMV) scenario, array signal processing is set mainly in the Multiple Measurements Vector (MMV) framework. The compressive sensing approach aims to solve the non-convex combinatorial ℓ_0-norm problem (i.e., $\min \|\mathbf{X}\|_0$ subject to $\|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 \le \epsilon$), which can be related to the well known Deterministic Maximum Likelihood (DML) estimator [4]. Both estimators require a multi-dimensional search, infeasible in practical scenarios. The main insight of compressive sensing has been to derive conditions guaranteeing a correct global solution via polynomial complexity (e.g., by solving a convex relaxation of the original problem). These conditions link properties of $\mathbf{X}$ (specifically, the number of active rows/targets K and their linear dependency) with properties of the matrix $\mathbf{A}$ (the smallest number of linearly dependent columns or $\mathrm{spark}(\mathbf{A})$). For a matrix $\mathbf{A}$ with MN rows, it trivially holds that $\mathrm{spark}(\mathbf{A}) \le MN + 1$. Assuming this bound is achieved with equality, a necessary and sufficient lower bound on MN for identifiability of the localization problem is $MN > 2K - \mathrm{rank}(\mathbf{Y})$ [5], [8]. It is well known in array processing that so-called "super-resolution" techniques, e.g., subspace methods such as MUSIC [6] and ESPRIT [7] as well as "rank aware" methods [8], are able to attain this bound with equality in a noiseless setting, when the received signal has a full-rank covariance matrix ($\mathrm{rank}(\mathbf{Y}) = K$). This is a consequence of the fact that subspace methods are "large sample" realizations of the maximum likelihood estimator for $MN > K$ and uncorrelated signals [4]. In a noisy scenario, at high SNR, the Cramér-Rao Bound (CRB) relates the number of elements not only to the number of targets K, but also to the system performance. The CRB for array processing DOA estimation is analyzed in [4], while the CRB for sparse estimation is discussed in [9], [10]. In the current work, we go deeper and uncover relations between compressed sensing DOA estimation and the number of sensors that are necessary for the estimation errors to be local, i.e., such that a local

bound like Cramér-Rao applies.

The work presented in this paper investigates these relations for MIMO random arrays. The paper makes the following specific contributions: (1) it builds on our work in the single pulse (SMV) setting [1] and on the RA-ORMP algorithm [8], to develop a tree-based algorithm, dubbed Multi-Branch Matching Pursuit (MBMP), for the multiple pulse (MMV) setting, typical to radar; (2) it generalizes the lower bound on the number of tx/rx elements originally presented in [1] to account for the complexity of the algorithm. By doing so, we show that it is possible to trade-off sensors for computational complexity. This provides specific insight into links between random arrays and compressive sensing algorithms, and demonstrates that high resolution can be enabled by the proposed algorithm using a relatively low number of randomly placed sensors.

The following notation is used: boldface denotes matrices (uppercase) and vectors (lowercase); $(\cdot)^*$ denotes the complex conjugate operator; $(\cdot)^T$ denotes the transpose operator; $(\cdot)^H$ is the complex conjugate-transpose operator, and $(\cdot)^\dagger$ is the pseudo-inverse. The symbol "$\otimes$" denotes the Kronecker product. Moreover, given a set S of indices, $|S|$ denotes its cardinality, $\mathbf{A}_S$ is the sub-matrix obtained by considering only the columns indexed in S, and we define the projection matrix $\Pi_{\mathbf{A}_S}^\perp \triangleq \mathbf{I} - \mathbf{A}_S \mathbf{A}_S^\dagger$. Finally, $\|\mathbf{X}\|_0$ counts the number of nonzero-norm rows of $\mathbf{X}$.

II. PROBLEM FORMULATION

We model a MIMO radar system where 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 m-th transmitter is at position ξ_m on the x-axis, while the n-th receiver is at position ζ_n (with $\xi_m \in [0, Z_{TX}]$, $\forall m$ and $\zeta_n \in [0, Z_{RX}]$, $\forall n$). Targets are assumed in the far-field, meaning that a target's aspect angle θ_k is constant across the array. The purpose of the system is to determine the DOA angles to the targets. The $N \times 1$ vector representing the sampled received signal for the p-th pulse is

$$\mathbf{R}_p = \sum_{k=1}^{K} x_{k,p} \mathbf{b}(\theta_k) \mathbf{c}^T(\theta_k) \mathbf{S}_p + \mathbf{N}_p \qquad (1)$$

where $x_{k,p}$ is the k-th target's response relative to the p-th pulse, the $N \times 1$ vector $\mathbf{b}(\theta_k) = \frac{1}{\sqrt{N}}\left[\exp\left(j2\pi\frac{\sin\theta_k}{\lambda}\zeta_1\right), \ldots, \exp\left(j2\pi\frac{\sin\theta_k}{\lambda}\zeta_N\right)\right]^T$ accounts for the angular response between the k-th target and each receiver sensor, the $M \times 1$ vector $\mathbf{c}(\theta_k) = \frac{1}{\sqrt{M}}\left[\exp\left(j2\pi\frac{\sin\theta_k}{\lambda}\xi_1\right), \ldots, \exp\left(j2\pi\frac{\sin\theta_k}{\lambda}\xi_M\right)\right]^T$ accounts for the angular response between the k-th target and each transmitter, and the $M \times L$ matrix $\mathbf{S}_p$ contains the L samples of the M signals constituting the p-th pulse transmitted by the MIMO radar. We assume the M transmitted signals to be orthogonal (e.g., pulses modulated by an orthogonal code). The $N \times L$ matrix $\mathbf{N}_p$ models noise (assumed temporally and spatially white). Finally, the DOA angles are assumed constant over the duration of the

P pulses, while the target reflection coefficients $x_{k,p}$ are assumed to remain constant during a pulse period and vary independently from pulse to pulse, following the Swerling Case II target model [11].

Vectorizing the outputs of all the receivers' matched filters, and stacking them column-wise,

$$\mathbf{Y} = \frac{M}{L}\left[(\mathbf{S}_1^* \otimes \mathbf{I}_N)\operatorname{vec}(\mathbf{R}_1), \ldots, (\mathbf{S}_P^* \otimes \mathbf{I}_N)\operatorname{vec}(\mathbf{R}_P)\right], \qquad (2)$$

we obtain

$$\mathbf{Y} = \tilde{\mathbf{A}}(\boldsymbol{\theta})\tilde{\mathbf{X}} + \mathbf{E} \qquad (3)$$

where $\mathbf{Y}$ is a $MN \times P$ matrix, $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_P]$ is a $K \times P$ matrix with $\tilde{\mathbf{x}}_p = [x_{1,p}, \ldots, x_{K,p}]^T$, $\tilde{\mathbf{A}}(\boldsymbol{\theta}) = [\mathbf{c}(\theta_1) \otimes \mathbf{b}(\theta_1), \ldots, \mathbf{c}(\theta_K) \otimes \mathbf{b}(\theta_K)]$ is $MN \times K$, and $\mathbf{E} = [\mathbf{n}_1, \ldots, \mathbf{n}_P]$ is $MN \times P$ with $\mathbf{n}_p \triangleq \frac{M}{L}\left(\mathbf{S}_p^* \otimes \mathbf{I}_N\right)\operatorname{vec}(\mathbf{N}_p)$. To embed the DOA estimation into a sparse localization framework, we discretize the possible targets' locations $\boldsymbol{\theta}$, obtaining a grid of G points $\{\phi_g\}$ (with $G \gg K$). Defining the $MN \times G$ matrix $\mathbf{A} = [\mathbf{a}(\phi_1), \ldots, \mathbf{a}(\phi_G)]$, where $\mathbf{a}(\theta) \triangleq \mathbf{c}(\theta) \otimes \mathbf{b}(\theta)$, the localization problem is expressed as:

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{E} \qquad (4)$$

where the unknown $G \times P$ matrix $\mathbf{X}$ contains the targets locations and gains. Zero rows of $\mathbf{X}$ correspond to grid points without a target. The problem (4) is sparse in the sense that $\mathbf{X}$ has only $K \ll G$ nonzero rows.

III. THE MULTI-BRANCH MATCHING PURSUIT ALGORITHM

In this section, we detail the proposed MBMP algorithm to address the joint sparse noisy recovery of model (4). Assuming the number of targets K is known, we want to solve the following nonconvex optimization problem:

$$\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 \qquad (5)$$
$$\text{s.t.} \quad \|\mathbf{X}\|_0 \leq K.$$

It is worth noticing that the solution can be parametrized only by its support S, such that the problem can be equivalently posed as $\min_S \left\|\Pi_{\mathbf{A}_S}^\perp \mathbf{Y}\right\|_F^2$ s.t. $|S| \leq K$.

To describe the proposed algorithm, it is instructive to review first the RA-ORMP structure. It starts by determining the row of $\mathbf{X}$ that maximizes the metric given by the inner product between the received signal subspace and each atom in the dictionary. The row's index is added to the active support. At the next iteration, the information associated with the first target is removed, and the metric is updated. The iterations stop when a termination criteria is met, and the current active support is elected as the solution. It is possible to visualize the process as a chain of successive nodes as shown in part (a) of Fig. 1: node (0) is tagged with an empty support. Each node evaluates the prescribed metric based on the support inherited from its parent, and adds an extra row to that support. Note that in the RA-ORMP algorithm, each node has only a *single* child. The proposed MBMP algorithm is a generalization of the
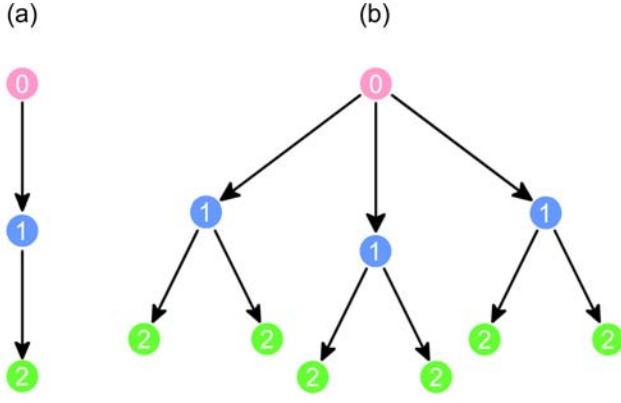
Fig. 1. Graph of: (a) RA-ORMP algorithm; (b) MBMP algorithm with $\mathbf{d} = [3, 2]$. The numbers indicate the nodes' level.

RA-ORMP, obtained by allowing each node to have *multiple* children (the highest peaks of the metric instead of just the maximum peak), such that the chain of nodes becomes a tree (see part (b) of Fig. 1).

With the MBMP algorithm, each node contains a set of indices associated with the active rows. At termination, the support that minimizes (5) is selected as the solution. The structure of the tree depends on the total number of levels K and on the number of allowed branches at each level (assumed constant for nodes within the same level of the tree). The structure can be specified by a length K vector $\mathbf{d}$: element d_i represents the number of branches at level-$(i-1)$. For instance, the tree in Fig. 1(b) has $\mathbf{d} = [3, 2]$ (node 0 has $d_1 = 3$ branches, and each node at level 1 possesses $d_2 = 2$ branches). With the same notation, the RA-ORMP chain in Fig. 1(a) has $\mathbf{d} = [1, 1]$. Notice that the node's level also indicates the cardinal of the associated support. The pseudo-code of the algorithm is detailed in the following table.

Algorithm 1 Multi-branch matching pursuit algorithm

Input: $\mathbf{Y}$, $\mathbf{A}$ and $\mathbf{d}$.

Output: Support of solution to (5).

1: Tag root node with $S = \emptyset$ and $\bar{S} = \{1, \ldots, G\}$.
2: **while** $\exists$ leaf node with $|S| < length(\mathbf{d})$
3:      Select a leaf node with $|S| < length(\mathbf{d})$.
4:      Set $\Gamma = S$, $\bar{\Gamma} = \bar{S}$ and $i = |\Gamma|$.
5:      Evaluate the signal subspace $\mathbf{U}$ of the residual $\Pi_{\mathbf{A}_\Gamma}^{\perp} \mathbf{Y}$.
6:      $\hat{g}_{1:d_{i+1}}$ = indexes of the d_{i+1} highest *peaks* of $\frac{\left\| \mathbf{U}^H \Pi_{\mathbf{A}_\Gamma}^{\perp} \mathbf{a}_g \right\|_2}{\left\| \Pi_{\mathbf{A}_\Gamma}^{\perp} \mathbf{a}_g \right\|_2}$ with $g \in \bar{\Gamma}$.
7:      **for** $d = 1, \ldots, d_{i+1}$
8:          Tag a child node with:
             $S = \Gamma \cup \hat{g}_d$.
             $\bar{S} = \bar{\Gamma} \setminus \{\hat{g}_{1:d}\}$.
             $f = \left\| \Pi_{\mathbf{A}_S}^{\perp} \mathbf{Y} \right\|_F^2$.
9:      **end**
10: **end**
11: Return S of a node tagged with the minimum f.

Computational Complexity

Given the $MN \times G$ matrix $\mathbf{A}$, the MBMP algorithm with branch vector $\mathbf{d} = [d_1, \ldots, d_K]$ has complexity $O\left(\min\left(P, MN\right) MNGD\right)$ where $D \triangleq \sum_{i=1}^{K} \prod_{j=1}^{i} d_j$. This should be compared with the exhaustive search strategy, which requires $O\left(\min\left(P, MN\right) MNG^K K^2\right)$ multiplications, i.e., is exponential in K. By properly selecting $\mathbf{d}$, we can effectively trade-off complexity with performance. For instance, when all $d_i = 1$, the algorithm becomes equivalent to the RA-ORMP, whose complexity is linear in K. Moreover, we are interested in solutions that require a relatively small number of transmitters and receivers. Lower bounds on the number of array elements are discussed in the next section.

IV. LOWER BOUNDS ON THE NUMBER OF SENSORS

In [1], we established a lower bound on the number of elements of a random array MIMO radar system for a single measurement vector (SMV) of sub-Nyquist spatial samples. The bound was customized to the OLS algorithm [12] to guarantee correct recovery with high probability. In this section, we propose a new bound tailored to the MBMP algorithm. This bound is developed for the SMV (i.e., $\mathbf{X}$ in (4) is $G \times 1$), with its extension to MMV, in future work. For simplicity, it is assumed that targets are located on grid points, and measurements are noiseless.

To guarantee correct recovery of the targets by the MBMP algorithm, the correct support for vector $\mathbf{X}$ in (4) has to be contained in one of the tree's paths. This translates into requiring that at the i-th level of the tree, one of the d_{i+1} branches selects an index belonging to the true support. Thus, at the level-0, it is required that the d_1-ranked highest peak response at grid points ϕ_g without targets is smaller than the highest response of a target,

$$\frac{d_1\text{-ranked } peak \text{ of } \left| \mathbf{y}^H \mathbf{a}\left(\phi_g\right) \right| \text{ for } \phi_g \notin \boldsymbol{\theta}}{\text{largest value of } \left| \mathbf{y}^H \mathbf{a}\left(\phi_g\right) \right| \text{ for } \phi_g \in \boldsymbol{\theta}} < 1 \quad (6)$$

where ϕ_g, $g = 1, \ldots, G$ are the directions associated with the grid, and $\boldsymbol{\theta} = \{\theta_k\}_{k=1}^{K}$ is the set of target locations. In phased array parlance, the numerator is the d_1-ranked highest *sidelobe*. Since the noiseless snapshot across the array is given by $\mathbf{y} = \sum_{k=1}^{K} \mathbf{a}\left(\theta_k\right) x_k$, correct target estimation at the first level is guaranteed if the d_1-ranked sidelobe is smaller than the response of at least one of the targets. It is worth pointing out that this is a much weaker condition with respect to the bound proposed in [1], which required that the 1-ranked sidelobe (largest) is smaller than the response of at least one of the targets.

Following steps similar to [1], it can be shown that a lower bound on the number of MIMO radar elements required to guarantee that the sidelobe condition (6) is met with probability greater than $1 - \gamma$ is

$$MN \geq K \ln \frac{n}{d_1 \gamma} \quad (7)$$

where the number of sidelobes n is approximately two times the "virtual array" aperture, i.e., $n \approx 2\left(Z_{TX} + Z_{RX}\right)/\lambda$. A

proof outline is provided in the appendix.

Condition (6), and consequently (7), derives from level-0 of the MBMP algorithm. Given a correct decision at this level (say θ_1), the algorithm projects the received vector and the dictionary on the null-space of the steering vector $\mathbf{a}(\theta_1)$. The updated received vector becomes $\sum_{k=2}^{K} \mathbf{a}(\theta_k) x_k$. From MIMO random array theory [1], the new sidelobes $\left| \sum_{k=2}^{K} x_k \mathbf{a}^H(\theta_k) \mathbf{a}(\phi_g) \right|$, $\phi_g \neq \{\theta_k\}_{k=2}^{K}$, are Rayleigh distributed with a *lower* variance $\left(\sum_{k=2}^{K} |x_k|^2 \right) / MN$ than for the earlier case (i.e., $\left(\sum_{k=1}^{K} |x_k|^2 \right) / MN$). Intuitively, we have removed the interference to the other targets from the correct estimated target. Thus, as long as the new denominator $\max_{\phi_g \in \boldsymbol{\theta}} \left| \sum_{k=2}^{K} x_k \mathbf{a}^H(\theta_k) \mathbf{a}(\phi_g) \right| \approx \max_{k \in \{2,...,K\}} |x_k|$, and for a proper $d_2 \leq d_1$, the correct target recovery at the second level imposes a looser bound, i.e., $MN \geq (K-1) \ln \frac{n}{d_2 \gamma}$, guaranteed by (7). A similar argument can be invoked to guarantee a correct decision for successive levels.

V. NUMERICAL RESULTS

In this section, we present numerical results to demonstrate the potential of the MBMP algorithm for the target localization problem, and to investigate the fitness of the proposed lower bound. Consider a MIMO radar that transmits P pulses of orthogonal spread spectrum waveforms of length M each. The waveforms were chosen as the $M \times M$ Fourier matrix. Equal length apertures were assumed for the transmit and receive arrays, i.e., $Z_{TX} = Z_{RX} = Z$ and, unless otherwise stated we set $Z = 50\lambda$. To ensure the aperture length, elements anchor the ends at locations 0 and Z of both the transmit and receive arrays. The locations of the remaining sensors are drawn uniformly at random. We assume that the number of targets K is known. In addition to MBMP, we implement target localization using other compressive sensing techniques (RA-ORMP and Reg. M-FOCUSS [13]) as well as the MUSIC estimator, famous in array processing. In addition, we compare the random array geometry with a ULA array geometry. Specifically, a "virtual ULA" is obtained by N receive elements spaced by $\lambda/2$, and M transmit elements spaced by $N\lambda/2$. The "virtual ULA" array (labeled "ULA" on the figures) results in a significantly smaller aperture then the random array with the same number of tx/rx sensors. To gain further insight, we include for the "virtual ULA", results obtained with the grid-free ESPRIT algorithm. In each figure, we perform independent Monte Carlo realizations varying the target's responses ($x_{k,p} = \exp(-j\varphi_k)$ with $\varphi_k \sim \mathcal{U}(0, 2\pi)$ $\forall k,p$), the noise (vec$(\mathbf{N}_p) \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$ $\forall p$), and the random array sensors' positions. Moreover, the SNR is defined as $10 \log_{10} \sigma^2$ and, since targets' locations are held fixed, grid points are spaced linearly in the range $-80° + \varepsilon$ to $80° + \varepsilon$, where $\varepsilon \sim \mathcal{U}(-80/G, 80/G)$ is randomized throughout the Monte Carlo realizations. The choice of G is related to the Rayleigh resolution limit for the random array. For instance, when $Z = 50\lambda$, the resolution is $\lambda/(2Z) \approx 0.6°$, and we place 4 grid-points per lobe, i.e., $G \approx 4 \times 160/0.6 = 1001$.
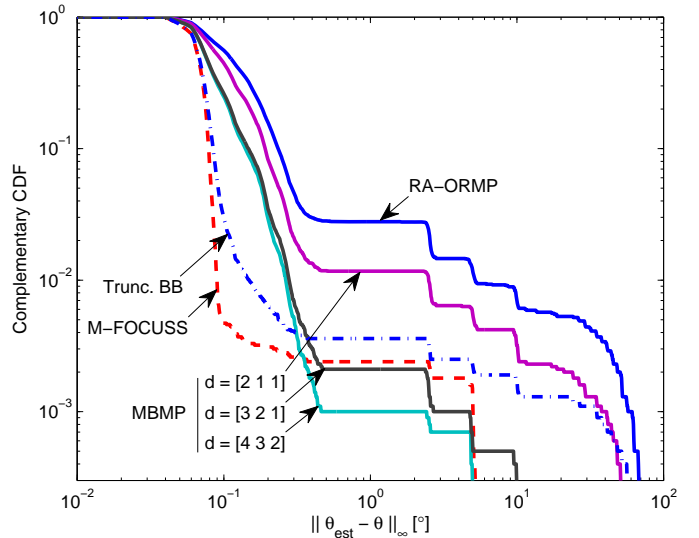

Fig. 2. CCDFs of the MBMP algorithm with different tree's spans. Noiseless SMV setting.

The same grid is used for all the recovery methods. As a measure of estimation accuracy, for each realization, we collect the *largest modulo* of the targets' estimation error, i.e., $\left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \right\|_{\infty} \triangleq \max_k \left| \hat{\theta}_k - \theta_k \right|$. We then plot the complementary cumulative distribution function (CCDF), defined as $C(x) \triangleq \Pr(X \geq x) = 1 - F(x)$, where F is the cumulative distribution function. The function C is the probability of having an error *greater* then the abscissa, such that a good technique shifts the CCDF towards the bottom-left of the figure. This choice highlights both the resolution and the probability of ambiguities (sidelobes' detections).

We analyze first the behavior of the proposed algorithm in a noiseless SMV setting. Fig. 2 plots the CCDFs of the recovery error for a variety of tree spans of the MBMP algorithm (i.e., different branch vectors $\mathbf{d}$), and compares it with the performance of certain compressive sensing techniques, as well as with the Truncated BB algorithm proposed in [1] truncated at 100 iterations. The system settings are $M = N = 5$ elements, noiseless $\sigma^2 = 0$, and $K = 3$ targets at $\boldsymbol{\theta} = [-5°, 0°, 5°]$. It can be seen how the probability of errors greater then $0.6°$ (i.e., the random array resolution) diminishes as the tree's span is increased. Moreover, the proposed MBMP algorithm outperforms the Truncated BB with less then half of the latter complexity (i.e., from the discussion of computational complexity, for $\mathbf{d} = [4, 3, 2]$, we have $D = 40$). This favorable outcome stems from the rules used to build the algorithm's tree.

Next, we focus on a scenario with high noise and multiple pulses (MMV). The system settings are $M = N = 5$, $P = 200$, $\sigma^2 = 1$, and $K = 4$ targets at $\boldsymbol{\theta} = [-7.5°, -2.5°, 2.5°, 7.5°]$. Fig. 3 plots the CCDFs of the recovery error for various versions of the MBMP algorithm (i.e., for different branch vector $\mathbf{d}$), and compares it with the performance of compressive sensing techniques and with
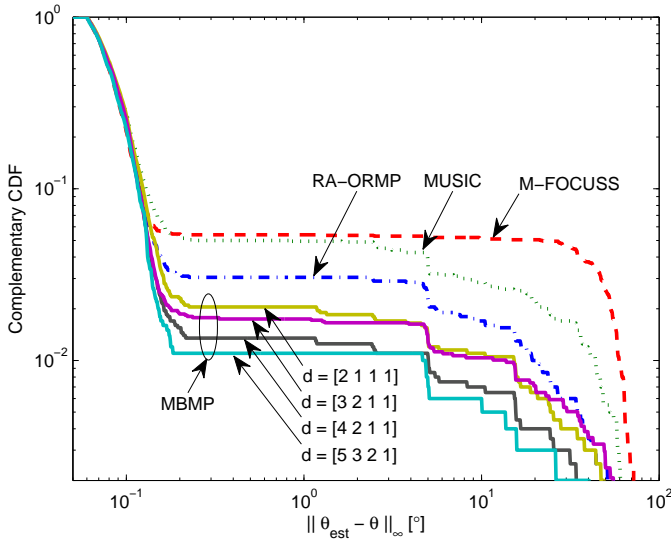
Fig. 3. CCDFs of the MBMP algorithm with different tree's spans. Noisy MMV setting.

Fig. 4. Sidelobe probability of the MBMP algorithm (a) varying the noise level; (b) varying the number of pulses.

the MUSIC estimator. The main difference from the noiseless SMV scenario is that M-FOCUSS is now performing much worse than MBMP. It is well known that for a large sample (high P) the estimated covariance matrix becomes accurate. Subspace methods exploits that through the separation of signal and noise subspaces. This gain is not exploited by the M-FOCUSS algorithm, resulting in poorer performance. Nonetheless, MUSIC (a subspace method) does not perform that well either. Now, this can be explained by the non-iterative decision strategy of the algorithm: the K targets are estimated *simultaneously* as the K highest peaks of the MUSIC spectrum. In contrast, MBMP refines its estimates iteratively based on a successive-decision strategy (following the tree's levels). As a result, this strategy enables to reduce the errors without compromising the computational complexity.

We then investigate the behavior of the proposed algorithm ($\mathbf{d} = [5, 3, 2, 1]$) as a function of SNR and sample support P. We compare it with the performance of compressive sensing techniques and with the MUSIC estimator. The system settings are $M = N = 5$ and $K = 4$ targets at $\boldsymbol{\theta} = [-7.5°, -2.5°, 2.5°, 7.5°]$. Fig. 4(a) plots the probability of having a recovery error greater then $1°$ when the number of samples is $P = 5$. Fig. 4(b) plots the probability of having a recovery error greater then $1°$ in a scenario with high noise ($SNR = 5\ dB$). It can be seen the superior performances of the MBMP algorithm with respect to other methods.

In the next experiment, we investigate the resolution capabilities with respect to SNR and sample support. We are asking how high must SNR be to deliver "super-resolution" (i.e., resolve targets less then one beamwidth apart)? We compare a MIMO random array architecture and a "virtual ULA" configuration. The system settings are $M = N = 3$ sensors and $K = 3$ targets at $\boldsymbol{\theta} = [-5°, 0°, 5°]$. The random array has a virtual aperture of 14λ (resulting in a beamwidth of $\approx 4°$),

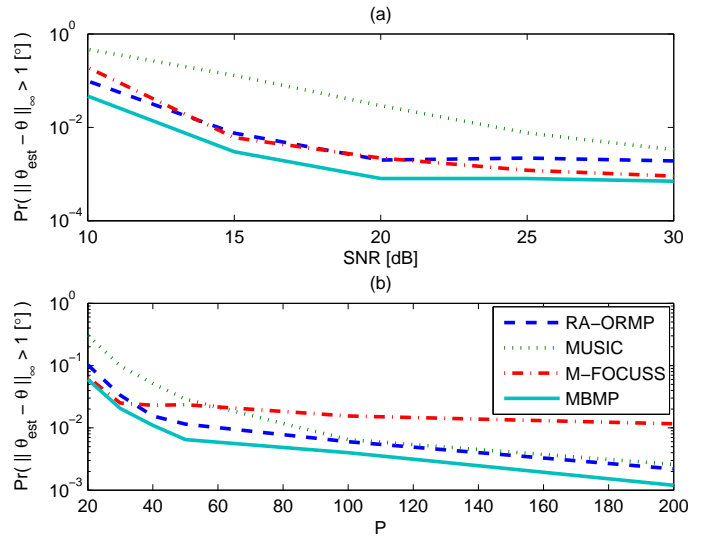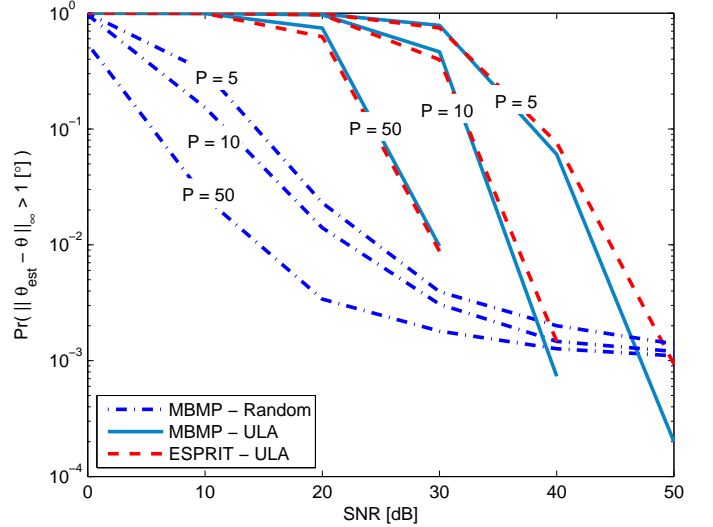Fig. 5. Sidelobe error probability for random array and ULA with the MBMP and ESPRIT algorithms.

while the "virtual ULA" has an aperture of 4λ (beamwidth of $\approx 14°$). The number of sensors and the array aperture have been selected such that the random array beamwidth is less then the minimum target separation $\Delta\theta = 5°$, while the "virtual ULA" beamwidth is greater then $\Delta\theta$. Once again, this is to investigate how an ambiguous but large aperture MIMO random array compares with an ambiguity-free "virtual ULA" where resolution is constrained to the array aperture. In a noiseless scenario, we expect super-resolution methods to recover all the targets with probability 1, irrespective of "virtual ULA" aperture. Fig. 5 plots the probability of having a recovery error greater then $1°$ varying the noise level and for different number of pulses ($P = 5$, 10 and 50). The MBMP algorithm ($\mathbf{d} = [5, 3, 2, 1]$) is employed in the MIMO

TABLE I
LOWER BOUND ON THE NUMBER OF SENSORS

				d_1				
Z	K	M	N	5	4	3	2	1
100	5	7	7	51.9	52.5	49.8	50.3	45.1
100	4	6	6	35.6	36.1	33.9	39.2	32.9
100	3	6	6	35.9	33.3	32.1	35.4	29.3
50	5	7	7	49.5	50.1	47.3	48.7	44.3
50	4	6	6	38.6	38.4	35.2	36.0	32.4
50	3	6	6	34.7	35.3	32.7	35.3	28.1

random array and in the "virtual ULA" setting. In the latter, the ESPRIT algorithm is shown for comparison. The superiority of the random array, which has higher array resolution, is evident. Moreover, the proposed algorithm is found to be more robust then ESPRIT to limited data support.

Finally, we address the goodness of the proposed lower bound (7). To be consistent with the analysis, we choose the noiseless SMV scenario (i.e., $\sigma^2 = 0$ and $P = 1$). The MBMP algorithm is run with different branches' vectors $\mathbf{d}$, in particular we select d_1 from 1 to 5. Table I reports the value of $K \ln \frac{n}{d_1 \gamma}$ for different scenarios varying the random array aperture Z, the number of targets K and the number of sensors $M = N$. The γ used in evaluating the formula is the empirical sidelobe probability (i.e., $\Pr \left(\left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \right\|_\infty > 1° \right)$) experienced by the MBMP algorithm in the different scenarios throughout Monte-Carlo simulations. A surprisingly good fit between the predicted value and the true MN emerges, particularly in the way (7) captures the role of d_1.

VI. CONCLUSIONS

We address the source localization problem in MIMO radar by using a sparse representation framework. We develop a global search algorithm for the sparse recovery problem, and derive an explicit lower bound on the number of random array elements needed to achieve a prescribed probability of correct DOA estimation. The lower bound provides specific insight into links between random arrays and compressive sensing algorithms, and demonstrates that a high resolution can be obtained with a relatively low number of randomly placed sensors.

VII. APPENDIX

Here we provide an outline of the proof to obtain (7), i.e., a lower bound on MN such that the first level of the algorithm is successful with probability $P_s \geq 1 - \gamma$ (for small γ, e.g., $\gamma \approx 10^{-3}$).
Defining $q \triangleq \max_k \left| x_k + \sum_{j \neq k} x_j \mathbf{a}^H (\theta_j) \mathbf{a} (\theta_k) \right|$, the probability of correct estimation, i.e., when (6) holds true, is

$$P_s = \sum_{i=0}^{d_1-1} \Pr \left[\text{exactly } i \text{ sidelobes} > q \right]. \quad (8)$$

As in [1], we approximate the sidelobes using independent Rayleigh variables with variance $\left(\sum_{k=1}^{K} |x_k|^2 \right) / MN$. For the

sake of notation brevity, we define $\beta \triangleq \exp \left(-\frac{MN}{\sum_{k=1}^{K} |x_k|^2} q^2 \right)$. Using the binomial distribution, we have

$$P_s = \sum_{i=0}^{d_1-1} \binom{n}{i} \beta^i [1 - \beta]^{n-i}. \quad (9)$$

The condition $P_s \geq 1 - \gamma$ is thus equivalent to

$$\gamma \geq 1 - \sum_{i=0}^{d_1-1} \binom{n}{i} \beta^i [1 - \beta]^{n-i}. \quad (10)$$

The quantity on the right-hand side can be upper-bounded by $\frac{n}{d_1} \beta$. Using this bound and solving for MN we have

$$MN \geq \frac{\sum_{k=1}^{K} |x_k|^2}{q^2} \ln \left(\frac{n}{d_1 \gamma} \right). \quad (11)$$

We next focus on the term q. Assuming that the targets are separated by at least one lobe, using the same argument in [1], the set $\left\{ \left| x_k + \sum_{j \neq k} x_j \mathbf{a}^H (\theta_j) \mathbf{a} (\theta_k) \right| \quad \text{for } k = 1, \ldots, K \right\}$ can be approximated by K independent Ricean variables with mean parameter $|x_k|$ and variance $\frac{1}{MN} \sum_{j \neq k} |x_j|^2$. The maximum over k is a random variable which depends on the targets' modulo. When MN is sufficiently high, $q \approx \max_k |x_k|$. Assuming equal modulo targets and using $q = |x_k|$, the quantity $\frac{1}{q^2} \sum_{k=1}^{K} |x_k|^2$ tends to K and (11) becomes (7). While several assumptions are used in this outline, the proposed bound (7) is reinforced by numerical results.

REFERENCES

[1] M. Rossi, A. M. Haimovich, and Y. C. Eldar, "Global Methods for Compressive Sensing in MIMO Radar with Distributed Sensors," in Proc. Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, Nov. 6-9, 2011. Available at: *http://web.njit.edu/~mr227/*
[2] A. M. Haimovich, R. Blum, and L. Cimini, "MIMO radar with widely separated antennas," IEEE Sig. Proc. Mag., vol.25, no.1, pp.116–129, 2008.
[3] H. L. VanTrees, *Detection, Estimation and Modulation Theory: Optimum Array Processing Vol. 4*. New York: Wiley, 2002.
[4] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound," IEEE Trans. Acoust., Speech, & Sig. Proc., vol. 37, pp 720–741, May 1989.
[5] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays ," Acoustics, Speech and Signal Processing, IEEE Transactions on , vol.37, no.7, pp.996-1000, Jul 1989.
[6] R. Schmidt, "Multiple emitter location and signal parameter estimation," Antennas and Propagation, IEEE Transactions on, vol. 34, no. 3, pp. 276-280, 1986.
[7] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," IEEE Trans. Acoust., Speech, Sig. Proc., vol.37, no.7, pp.984–995, Jul. 1989.
[8] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," to appear in IEEE Trans. on Info. Theory.
[9] Z. Ben-Haim and Y. C. Eldar, "The Cramér-Rao bound for estimating a sparse parameter vector," IEEE Trans. Signal Process., vol. 58, no. 6, pp. 3384–3389, June 2010.
[10] Z. Ben-Haim, T. Michaeli and Y. C. Eldar, "Performance bounds and design criteria for estimating finite rate of innovation signals," submitted to IEEE Trans. Information Theory.
[11] M. Skolnik, *Introduction to Radar Systems*. McGraw-Hill, 3rd ed., 2002.
[12] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," Int. Journal of Control, vol.50, no.5, pp.1873–1896, 1989.
[13] 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, July 2005.