DOA Estimation Exploiting A Uniform Linear Array with Multiple Co-prime Frequencies

Si Qin, Yimin D. Zhang, Moeness G. Amin, and Braham Himed

Abstract

The co-prime array, which utilizes a co-prime pair of uniform linear sub-arrays, provides a systematical means for sparse array construction. By choosing two co-prime integers M and N, $\mathcal{O}(MN)$ co-array elements can be formed from only $\mathcal{O}(M + N)$ physical sensors. As such, a higher number of degrees-of-freedom (DOFs) is achieved, enabling direction-of-arrival (DOA) estimation of more targets than the number of physical sensors. In this paper, we propose an alternative structure to implement co-prime arrays. A single sparse uniform linear array is used to exploit two or more continuous-wave signals whose frequencies satisfy a co-prime relationship. This extends the co-prime array and filtering to a joint spatio-spectral domain, thereby achieving high flexibility in array structure design to meet system complexity constraints. The DOA estimation is obtained using group sparsity-based compressive sensing techniques. In particular, we use the recently developed complex multitask Bayesian compressive sensing for group sparse signal reconstruction. The achievable number of DOFs is derived for the two-frequency case, and an upper bound of the available DOFs is provided for multi-frequency scenarios. Simulation results demonstrate the effectiveness of the proposed technique and verify the analysis results.

Index Terms

DOA estimation, co-prime array, sparse sampling, group sparsity, sparse Bayesian learning.

S. Qin and M. G. Amin are with the Wireless Communications and Positioning Laboratory, the Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA.

Y. D. Zhang is with the Department of Electrical and Computer Engineering, College of Engineering, Temple University, Philadelphia, PA 19122, USA (e-mail: ydzhang@temple.edu).

B. Himed is with the RF Technology Branch, Air Force Research Lab (AFRL/RYMD), WPAFB, OH 45433, USA.

I. Introduction

An important application of array signal processing is direction-of-arrival (DOA) estimation, which determines the spatial spectrum of the impinging electromagnetic waves. It is well known that an N-element uniform linear array (ULA) has $N-1$ degrees-of-freedom (DOFs), i.e., it resolves up to $N-1$ sources or targets by using conventional DOA estimation methods, such as MUSIC and ESPRIT [3, 4]. On the other hand, a higher number of DOFs can be achieved to resolve more targets by using the same number of array sensors if they are sparsely placed [5, 6]. An increased number of DOFs is usually achieved by exploiting the extended difference co-array whose virtual sensor positions are determined by the lag differences between the physical sensors.

Among a number of techniques that are available for sparse array construction, co-prime array [7] is considered attractive due to its capability of the systematic sparse array design. By choosing two integer numbers M and N to be co-prime, $\mathcal{O}(MN)$ targets can be resolved with $M+N-1$ physical sensors [8]. This co-prime array concept can be generalized by introducing an integer factor that compresses the inter-element spacing of one constituting sub-array, thereby achieving increased DOFs [9, 11]. In addition, by placing the two sub-arrays co-linearly instead of co-located, the number of unique virtual sensors is further increased, which benefits DOA estimation based on sparse signal reconstruction techniques [10, 11].

While the co-prime array concept has been developed using physical uniform linear sub-arrays, we propose in this paper an effective scheme that implements co-prime array configurations using a single sparse ULA with two or more co-prime frequencies. As such, the ULA, whose inter-element spacing is respectively M_1 and M_2 half-wavelengths of the two respective frequencies, with M_1 and M_2 to be mutually co-prime integers, acts as virtual sub-arrays, resulting in an equivalent structure to co-prime arrays. In essence, the proposed approach integrates the concept of co-prime array and co-prime filter to reduce complexity and achieve high system performance. Unlike co-prime arrays, wherein the numbers of sub-array sensors and the inter-element spacings have to satisfy the co-prime relationship, only the frequencies are required to be co-prime in the proposed scheme.

The proposed scheme can be adopted for both passive and active radar systems. The former requires filtering the signal arrivals at the employed co-prime frequencies, whereas the latter requires emitting those frequencies from a single antenna or a phased array and receiving the target backscattering with ULA. The transmitter and receiver can be located or widely separated.

For active sensing, sum co-array of the transmit and receive arrays replaces the difference co-array of the two structures which is associated with receive only operations [12].

In this paper, we derive the analytical expression of the available number of DOFs as a function of the number of physical sensors, L, and the selected co-prime frequencies for the two-frequency case. The results resemble those derived in [9, 11] for a physical co-prime array. The key difference lies in the fact that, unlike the co-prime array where each sub-array uses a different number of sensors, the two virtual sub-arrays in the underlying structure refer to the same physical ULA and thus share the same number of sensors. In addition, the number of physical sensors is not tied to the co-prime frequency multipliers M_1 and M_2. The property enables a higher flexibility in array design and operation. In particular, for a fixed number of physical array sensors, L, we demonstrate that a high number of DOFs, proportional to L^2, can be achieved with large values of M_1 and M_2. When K mutually co-prime frequencies are used, each pair of these frequencies can form a virtual co-prime array as discussed above. Accordingly, $\mathcal{O}(K^2L^2)$ DOFs can be achieved.

It is shown that, in the proposed scheme, the self-lags in the co-array corresponding to each sub-array form a subset of the sub-array cross-lags. As such, the available DOFs are solely determined by the number of cross-lags between the two sub-arrays. Because of the frequency-dependent characteristics of the source, channel and target radar cross section (RCS), the received signal vectors corresponding to the different frequencies have a common spatial support, i.e., DOA, but generally have distinct coefficients. Thus, DOA estimations become a group sparse signal reconstruction problem. In this case, the self-lags obtained for each sub-array can also be exploited for possible performance improvement.

A large number of compressive sensing (CS) techniques have been proposed to deal with this problem. In this paper, we consider the problem under the Bayesian compressive sensing (BCS) or sparse Bayesian learning framework [13–17], which generally achieves a better reconstruction performance over those on the basis of the greedy algorithms and dynamic programming approaches, such as the orthogonal matching pursuit (OMP) [18] and the least absolute shrinkage and selection operator (LASSO) [19] algorithms. In particular, we use the complex multitask Bayesian compressive sensing (CMT-BCS) algorithm [20] to determine the DOAs of group sparse complex signals. This algorithm jointly treats the real and imaginary components of a complex value, in lieu of decomposing them into independent real and imaginary components. As a result, the sparsity of the estimated weight vectors can be improved, yielding better signal

recovery. Group sparsity treatments for real and imaginary entries have been reported in, e.g., [21, 22].

The remainder of the paper is organized as follows. In Section II, we first review the co-prime array concept based on the difference co-array. Then, the array signal model exploiting co-prime frequencies is summarized in Section III. Analytical expressions of array aperture and the number of DOFs are derived in Section IV with respect to two and multiple co-prime frequencies. Sparsity-based DOA estimation exploiting the CMT-BCS is described in Section V. Simulation results are provided in Section VI to compare the performance of DOA estimation for different scenarios and validate the usefulness of the results presented in Section V. Section VII concludes this paper.

Notations: We use lower-case (upper-case) bold characters to denote vectors (matrices). In particular, $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. $(.)^*$ implies complex conjugation, whereas $(.)^T$ and $(.)^H$ respectively denote the transpose and conjugate transpose of a matrix or vector. $\mathrm{vec}(\cdot)$ denotes the vectorization operator that turns a matrix into a vector by stacking all columns on top of each other, and $\mathrm{diag}(\mathbf{x})$ denotes a diagonal matrix that uses the elements of $\mathbf{x}$ as its diagonal elements. $\| \cdot \|_2$ and $\| \cdot \|_1$ respectively denote the Euclidean (l_2) and l_1 norms, and $\mathrm{E}(\cdot)$ is the statistical expectation operator. $\bigotimes$ denotes the Kronecker product, and $\lfloor \cdot \rfloor$ denotes the floor function and returns the largest integer not exceeding the argument. $P_r(\cdot)$ denotes the probability density function (pdf), and $\mathcal{N}(x|a,b)$ denotes that random variable x follows a Gaussian distribution with mean a and variance b. $\mathrm{Re}(x)$ and $\mathrm{Im}(x)$ denote the real and imaginary parts of complex element x, respectively.

II. CO-PRIME ARRAY CONCEPT

In this section, we first review the co-prime array configuration that achieves a higher number of DOFs based on the difference co-array concept. A co-prime array [7] is illustrated in Fig. 1, where M and N are co-prime integers, i.e., their greatest common divisor is one. Without loss of generality, we assume $M < N$. The unit inter-element spacing d is typically set to $\lambda/2$, where λ denotes the wavelength. The array sensors are positioned at

$$\mathbb{P} = \{Mnd| \ 0 \leq n \leq N-1\} \cup \{Nmd| \ 0 \leq m \leq M-1\}. \tag{1}$$

Because the two sub-arrays share the first sensor at the zeroth position, the total number of sensors used in the co-prime array is $M + N - 1$. Note that the minimum inter-element spacing in this co-prime array is $d = \lambda/2$.

Denote $\mathbf{p} = [p_1, ..., p_{M+N-1}]^T$ as the positions of the array sensors, where $p_i \in \mathbb{P}$, $i = 1, ..., M + N - 1$, and the first sensor, located at $p_1 = 0$, is assumed as the reference. Assume that Q uncorrelated signals impinging on the array from angles $\Theta = [\theta_1, ..., \theta_Q]^T$, and their discretized baseband waveforms are expressed as $s_q(t), t = 1, ..., T$, for $q = 1, ..., Q$. Then, the data vector received at the co-prime array is expressed as,

$$\mathbf{x}(t) = \sum_{q=1}^{Q} \mathbf{a}(\theta_q) s_q(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{2}$$

where

$$\mathbf{a}(\theta_q) = \left[1, e^{j\frac{2\pi p_2}{\lambda}\sin(\theta_q)}, ..., e^{j\frac{2\pi p_{M+N-1}}{\lambda}\sin(\theta_q)}\right]^T \tag{3}$$

is the steering vector of the array corresponding to θ_q, $\mathbf{A} = [\mathbf{a}(\theta_1), ..., \mathbf{a}(\theta_Q)]$, and $\mathbf{s}(t) = [s_1(t), ..., s_Q(t)]^T$. The elements of the noise vector $\mathbf{n}(t)$ are assumed to be independent and identically distributed (i.i.d.) random variables following the complex Gaussian distribution $\mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{M+N-1})$.

The covariance matrix of the data vector $\mathbf{x}(t)$ is obtained as

$$
\begin{aligned}
\mathbf{R_{xx}} &= \mathrm{E}[\mathbf{x}(t)\mathbf{x}^H(t)] = \mathbf{A}\mathbf{R_{ss}}\mathbf{A}^H + \sigma_n^2 \mathbf{I}_{M+N-1} \\
&= \sum_{q=1}^{Q} \sigma_q^2 \mathbf{a}(\theta_q)\mathbf{a}^H(\theta_q) + \sigma_n^2 \mathbf{I}_{M+N-1},
\end{aligned} \tag{4}
$$

where $\mathbf{R_{ss}} = \mathrm{E}[\mathbf{s}(t)\mathbf{s}^H(t)] = \mathrm{diag}([\sigma_1^2, ..., \sigma_Q^2])$ with σ_q^2 denoting the input signal power of the qth target, $q = 1, ..., Q$. In practice, the covariance matrix is estimated using the T available samples, i.e.,

$$\hat{\mathbf{R}}_{\mathbf{xx}} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}(t)\mathbf{x}^H(t). \tag{5}$$

By vectorizing the matrix $\hat{\mathbf{R}}_{\mathbf{xx}}$, we obtain the following measurement vector:

$$\mathbf{z} = \mathrm{vec}(\hat{\mathbf{R}}_{\mathbf{xx}}) = \tilde{\mathbf{A}}\mathbf{b} + \sigma_n^2 \mathbf{i}, \tag{6}$$

where $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}(\theta_1), \ldots, \tilde{\mathbf{a}}(\theta_Q)]$, $\tilde{\mathbf{a}}(\theta_q) = \mathbf{a}^*(\theta_q) \bigotimes \mathbf{a}(\theta_q)$, for $1 \leq q \leq Q$. In addition, $\mathbf{b} = [\sigma_1^2, \ldots, \sigma_Q^2]^T$ and $\mathbf{i} = \mathrm{vec}(\mathbf{I}_{M+N-1})$. Benefiting from the Vandermonde vector $\mathbf{a}(\theta_q)$, we can regard $\mathbf{z}$ as a received signal from a single snapshot $\mathbf{b}$ and the matrix $\tilde{\mathbf{A}}$ behaves as the manifold matrix of a larger virtual array which has sensors located at the lags between two sub-arrays. From a pair of antennas located at the ith and kth positions in $\mathbf{p}$, the correlation $\mathrm{E}[x_i(t)x_k^*(t)]$ yields the (i, k)th entry in $\mathbf{R_{xx}}$ with lag $p_i - p_k$. As such, all the available values of i and k,

where $0 \leq i \leq M + N - 1$ and $0 \leq k \leq M + N - 1$, yield virtual sensors of the following difference co-array:

$$\mathbb{C}_P = \{\mathbf{z} \mid \mathbf{z} = \mathbf{u} - \mathbf{v}, \mathbf{u} \in \mathbb{P}, \mathbf{v} \in \mathbb{P}\}. \tag{7}$$

The significance of the difference co-array is that the correlation of the received signal can be calculated at all lags in the set $\mathbb{C}_P$. Any application which depends only on such correlation (e.g., DOA estimation) can exploit all the DOFs offered by the resulting co-array structure. Using a part or the entire set of the distinct lag entries in the set $\mathbb{C}_P$, instead of the original physical array, to perform DOA estimation, we can increase the parameter identifiability. The maximum number of the DOFs is determined by the number of unique elements in the following set

$$\mathbb{L}_P = \{l_P \mid l_P d \in \mathbb{C}_P\}. \tag{8}$$

III. SYSTEM MODEL

As described in the previous section, a higher number of DOFs is achieved using a co-prime array. Such a co-prime array structure was originally developed using two physical uniform linear sub-arrays with co-prime inter-element spacing [7]. In this paper, we extend that concept to a sparse ULA with two or multiple co-prime frequencies, offering improved capabilities and flexibilities to achieve better performance using a single ULA.

Assume K continuous-wave (CW) signals with co-prime frequencies are received at an L-element ULA with inter-element spacing D. By co-prime frequencies, we mean that the ratio between carrier frequencies equals the ratio between co-prime integers. For a CW waveform with frequency $f_k, k = 1, \ldots, K$, the return signal from the Q far-field targets, located at DOAs $\theta_q, q = 1, 2, \cdots, Q$, are expressed in a vector form as

$$\check{\mathbf{x}}_k(t) = \exp(j2\pi f_k t) \sum_{q=1}^{Q} \rho_{kq}(t) \mathbf{a}_k(\theta_q) + \check{\mathbf{n}}_k(t), \quad k = 1, \ldots, K, \tag{9}$$

where $\rho_{kq}(t)$ is the complex envelop of the signal q corresponding to f_k, which does not vary with the receive antennas, but is in general frequency-dependent due to the different propagation phase delays. We assume $\rho_{kq}(t)$ to be uncorrelated for different targets over one scan due to target motion or RCS fluctuations (Swerling II). In addition, $\mathbf{a}_k(\theta_q)$ is the steering vector corresponding to θ_q for frequency f_k, expressed as

$$\mathbf{a}_k(\theta_q) = \left[1, e^{-j\frac{2\pi D}{\lambda_k} \sin(\theta_q)}, ..., e^{-j\frac{2\pi (L-1)D}{\lambda_k} \sin(\theta_q)}\right]^T, \tag{10}$$

where $\lambda_k = c/f_k$ denotes the wavelength corresponding to f_k, and c is the velocity of wave propagation. Furthermore, $\breve{\mathbf{n}}_k(t)$ is the additive noise vector whose elements are assumed to be spatially and temporally white, and are independent of the target signals.

After converting the received signal vector to baseband using the respective frequencies, followed by low-pass filtering, we obtain

$$\mathbf{x}_k(t) = \sum_{q=1}^{Q} \rho_{kq}(t)\mathbf{a}_k(\theta_q) + \mathbf{n}_k(t) = \mathbf{A}_k\mathbf{s}_k(t) + \mathbf{n}_k(t), \quad k = 1, \ldots, K, \tag{11}$$

where $\mathbf{A}_k = [\mathbf{a}_k(\theta_1), \cdots, \mathbf{a}_k(\theta_Q)]$ and $\mathbf{s}_k(t) = [\rho_{k1}(t), \cdots, \rho_{kQ}(t)]^T$. We denote the noise variance at the filter output as $\sigma_{n_k}^2$.

For convenience, $M_k, k = 1, \ldots, K$, are denoted as mutually co-prime integers. Without loss of generality, we assume that they are sorted in a descending order, i.e., $M_1 < M_2 < \ldots < M_K$. In addition, we assume that D is integer multiples of the half-wavelengths of all frequencies, such that $M_k = 2D/\lambda_k, k = 1, \ldots, K$. As such, the ULA is sparse (spatially undersampled) at each frequency by a factor of M_k. In this case and for clarity, we can rewrite the steering vectors in a frequency-independent form, expressed as

$$\mathbf{a}_k(\theta_q) = \left[1, e^{-jM_k\pi \sin(\theta_q)}, ..., e^{-jM_k(L-1)\pi \sin(\theta_q)}\right]^T. \tag{12}$$

It is clear that the DOA estimation problem is similar to the co-prime arrays considered in [7, 11]. There are K uniform linear sub-arrays with a respective co-prime inter-element spacing. It is noted, however, that unlike a co-prime array, in which the numbers of sub-array sensors are different, all sub-arrays in the underlying virtual co-prime array structure share the same number of sensors, L. In addition, the DOA estimation method needs to account for the fact that signals corresponding to different virtual arrays have distinct phases. In the next two sections, we respectively analyze the achievable DOFs and describe group sparse CS-based DOA estimation technique.

IV. ANALYSIS OF ACHIEVABLE DOFS

Similar to the co-prime array, the parameter identifiability can be improved using correlation-aware techniques. In this section, we consider the virtual array constructed by exploiting multiple co-prime frequencies and derive the analytical expressions of the number of DOFs.

A. Analysis of DOFs with two co-prime frequencies

We first consider the problem when two frequencies M_1 and M_2, $M_1 < M_2$, are used. As shown in Fig. 2, the sensors of the two equivalent sub-arrays are located at

$$\tilde{\mathbb{P}} = \{M_1 l_1 d_0 | 0 \leq l_1 \leq L - 1\} \bigcup \{M_2 l_2 d_0 | 0 \leq l_2 \leq L - 1\}, \tag{13}$$

where d_0 denotes a half-wavelength unit inter-element spacing in a normalized frequency sense (i.e., no specific frequency is referred to), and l_1 and l_2 are the respective indices of the sensor positions of the two equivalent sub-arrays. As such, the aperture of this equivalent co-prime array structure is $M_2(L-1)d_0$. In addition, the two uniform linear sub-arrays in the underlying problem have the same L sensors, which align in the zeroth position and whenever l_2/M_2 is an integer. Therefore, there are $2L - 1 - \lfloor (L-1)/M_2 \rfloor$ equivalent sensors. It is noted that, when $M_2 < L$, there are overlaps among the equivalent sensors, resulting in a reduced number of DOFs. Therefore, we only consider the $M_2 \geq L$ case in the remainder of this paper.

Because each sub-array is linear and uniformly spaced and the two sub-arrays share the first sensor at the zeroth position, a self-lag position of a sub-array can always be taken as the cross-lag position between a sensor of this sub-array and the first sensor of the other sub-array. In other words, the self-lag positions form a subset of the cross-lag positions [11]. Therefore, we only consider the cross-lags when determining the number of DOFs. In this array configuration, the cross-lags of the two equivalent sub-arrays are given by the following set,

$$\tilde{\mathbb{L}}_c = \{\tilde{l}_c | \tilde{l}_c = M_1 l_1 - M_2 l_2\}, \tag{14}$$

and the corresponding mirrored set,

$$\tilde{\mathbb{L}}_c^- = \{\tilde{l}_c | \tilde{l}_c = M_2 l_2 - M_1 l_1\} = \{-\tilde{l}_c | \tilde{l}_c \in \tilde{\mathbb{L}}_c\}, \tag{15}$$

where $0 \leq l_1 \leq L - 1$ and $0 \leq l_2 \leq L - 1$. The achievable DOFs from the difference co-array is determined by the unique elements in the following set

$$\tilde{\mathbb{L}}_P = \tilde{\mathbb{L}}_c \bigcup \tilde{\mathbb{L}}_c^-. \tag{16}$$

Overall, there are $2L^2$ lags in the set $\tilde{\mathbb{L}}_P$, which contains both non-overlapping and overlapping lags. To obtain a higher number of DOFs, which is determined by the number of unique lags in the set $\tilde{\mathbb{L}}_P$, we can choose different pairs of M_1 and M_2 to reduce the redundancies in both $\tilde{\mathbb{L}}_c$ and $\tilde{\mathbb{L}}_c^-$, as well as the overlapping lags between $\tilde{\mathbb{L}}_c$ and $\tilde{\mathbb{L}}_c^-$.

Denote η as the number of unique lags in the set $\tilde{\mathbb{L}}_P$. The following proposition reveals the analytical relationship between η and different choice of M_1 and M_2.

Proposition 1: For a virtual array constructed from a ULA with inter-element spacing D using two co-prime frequencies with $D = \frac{1}{2}M_1\lambda_1 = \frac{1}{2}M_2\lambda_2$, the number of unique lags is given by

$$\eta = 2L^2 - 1 - \max\{0, 2L - 1 - M_2\}\min\{M_1 + 1, 2L - 1 - M_1\}. \tag{17}$$

It can be expressed for three different cases:

(a) For $M_2 \geq 2L - 1$, $\eta = 2L^2 - 1$;

(b) For $L \leq M_2 < 2L - 1$ and $L \leq M_1 < M_2$, $\eta = 2L^2 - 1 - (2L - 1 - M_2)(2L - 1 - M_1)$;

(c) For $L \leq M_2 < 2L - 1$ and $1 \leq M_1 < L$, $\eta = 2L^2 - 1 - (2L - 1 - M_2)(M_1 + 1)$.

The proof is provided in Appendix A.

The number of DOFs in the co-array can be obtained as $(\eta + 1)/2$ [23]. It indicates that η achieves the maximum value of $2L^2 - 1$ in case (a), irrespective of M_1, provided that $M_1 < M_2$ is satisfied. In practice, however, a large value of M_2 would increase the number of missing positions, i.e., holes in the difference co-array. For cases (b) and (c), η depends on the values of both M_1 and M_2 and is maximized when $M_1 = 1$ or $M_1 = M_2 - 1$. The latter case yields a smaller frequency separation between f_1 and f_2, whereas the former configuration represents a nested structure [24]. A nested array is usually designed such that the virtual sensors in the resulting co-array are all contiguous and is considered as a special case of the generalized co-prime array in [11].

For an illustrative purpose, examples for different pairs of M_1 and M_2 are presented in Fig. 3 and Fig. 4, where the physical ULA has 4 sensors in all cases. The equivalent sensor positions are illustrated in Fig. 3, whereas the respective co-arrays are presented in Fig. 4. Note that the holes are indicated by " $\times$ ". It is clear that the difference co-arrays for all cases have more virtual sensors than the number of physical sensors in the original ULA. Compared to the other examples, there are more duplications in the $M_1 = 2$ and $M_2 = 3 < L$ case depicted in Fig. 3(a), leading to a reduction of the DOFs in the co-array, as shown in Fig. 4(a). Also, there are 19 unique lags for the $M_1 = 3$ and $L < M_2 = 4 < 2L - 1$ case in Fig. 4(b), whereas it increases to 31 in Fig. 4 (c) for $M_1 = 6$ and $M_2 = 7 \geq 2L - 1$, due to fewer overlapping lags between $\tilde{\mathbb{L}}_c$ and $\tilde{\mathbb{L}}_c^-$. The nested structure with $M_1 = 1$ and $M_2 = L = 4$ is depicted in Fig. 4(d) as a special case of $L \leq M_2 < 2L - 1$. It is evident that, in this case, all 25 lags are contiguous.

B. Analysis of DOFs with multiple co-prime frequencies

When more than two mutually co-prime frequencies are used, each co-prime frequency pair forms a virtual co-prime array corresponding to the two frequencies. Therefore, for K mutually co-prime frequencies, there are $\binom{K}{2} = \frac{K(K-1)}{2}$ co-prime frequency pairs. As a consequence, the number of DOFs in the resulting co-array is determined by the cardinality of the unique sum set of lags obtained in each co-prime frequency pair, which generally increases with the number of frequencies being used. However, a general expression of the DOF for different choices of the co-prime frequencies is rather complicated and does not necessarily provide meaningful insights. Instead, we provide the maximum number of achievable DOFs in the following proposition, which corresponds to the case where each pair achieves the maximum number of DOFs with minimum overlapping between different frequency pairs.

Proposition 2: The maximum number of achievable unique lags of the co-array generated from the equivalent sub-arrays is given by

$$\eta = (L^2 - 1)(K^2 - K) - 2(L - 1)(K^2 - 2K) + 1. \tag{18}$$

The proof is provided in Appendix B.

It is clear that $\eta \propto O(K^2 L^2)$, since there are $O(K^2)$ frequency pairs and $O(L^2)$ unique lags for each pair. To achieve the upper bound of DOFs, however, it requires a large separation between different multipliers $M_k, k = 1, \ldots, K$, so that the number of overlapping lags between different frequency pairs is minimized.

V. COMPRESSIVE SENSING BASED DOA ESTIMATION

While the DOA estimation problem considered here appears to be similar to that discussed in [7, 25], the CS method exploited therein cannot be readily applied to the underlying problem. A major distinction is that the target reflection coefficients $\rho_{kq}, q = 1, \ldots, Q$, differ at different frequencies $k = 1, \ldots, K$, due to differences in their propagation phase delays and target reflectivities. As such, the phase term of the cross-correlation between the received data vectors for different frequencies depends not only on the spatial angle, but also on the unknown phase difference in the reflection coefficients and propagation delays. In this section, we formulate the DOA estimation problem as a group sparsity based signal recovery problem.

A. DOA estimation using only cross-lags

As discussed earlier, a full number of unique lags is achieved in the resulting co-array by using the cross-lags between the sub-arrays. As such, the spatial spectra can be estimated based only on the cross-lag correlations without loss of DOFs.

The cross-lag covariance matrix $\mathbf{R}_{\mathbf{xx}}^{(i,k)}$ between the $L \times 1$ received data vectors $\mathbf{x}_i(t)$ and $\mathbf{x}_k(t)$, for $1 \leq i \neq k \leq K$, is obtained as

$$\mathbf{R}_{\mathbf{xx}}^{(i,k)} = \mathrm{E}\left[\mathbf{x}_i(t)\mathbf{x}_k^H(t)\right] = \mathbf{A}_i \mathbf{R}_{\mathbf{ss}}^{(i,k)} \mathbf{A}_k^H = \sum_{q=1}^{Q} \sigma_q^{(i,k)} \mathbf{a}_i(\theta_q)\mathbf{a}_k^H(\theta_q), \qquad (19)$$

where $\mathbf{R}_{\mathbf{ss}}^{(i,k)} = \mathrm{E}[\mathbf{s}_i(t)\mathbf{s}_k^H(t)] = \mathrm{diag}([\sigma_1^{(i,k)}, \ldots \sigma_Q^{(i,k)}])$ is the cross-correlation matrix between the received signals at the ith and kth frequencies. Note that $\sigma_q^{(i,k)}$, $q = 1, \ldots, Q$, in general, takes a complex value. Vectorizing $\mathbf{R}_{\mathbf{xx}}^{(i,k)}$ in (19), we obtain

$$\mathbf{z}_{ik} = \mathrm{vec}\left(\mathbf{R}_{\mathbf{xx}}^{(i,k)}\right) = \tilde{\mathbf{A}}_{ik}\mathbf{b}_{ik}, \quad i \neq k \in [1, \ldots, K], \qquad (20)$$

where $\tilde{\mathbf{A}}_{ik} = [\tilde{\mathbf{a}}_{ik}(\theta_1), ..., \tilde{\mathbf{a}}_{ik}(\theta_Q)]$ with $\tilde{\mathbf{a}}_{ik}(\theta_q) = \mathbf{a}_i^*(\theta_q) \bigotimes \mathbf{a}_k(\theta_q)$, and $\mathbf{b}_{ik} = [\sigma_1^{(i,k)}, ..., \sigma_Q^{(i,k)}]^T$. It is noted that the $L^2 \times 1$ vector $\mathbf{z}_{ik}$ can be sparsely represented in the spatial domain over the entire angular grids as

$$\mathbf{z}_{ik} = \tilde{\mathbf{A}}_{ik}^o \mathbf{b}_{ik}^o, \quad i \neq k \in [1, \ldots, K], \qquad (21)$$

where $\tilde{\mathbf{A}}_{ik}^o$ is defined as the collection of steering vectors $\tilde{\mathbf{a}}_{ik}$ over the entire possible grids θ_g for $g = 1, \ldots, G$, with $G \gg Q$. It is important to note that the angle positions of the signal arrivals $\theta_q, q = 1, \ldots, Q$, are indicated by the non-zero entries in vector $\mathbf{b}_{ik}^o$, whose values describe the corresponding coefficients. Generally, the non-zero entries take different values with respect to different frequency pairs but share the same positions because they correspond to the DOAs of the same Q targets. Therefore, $\mathbf{b}_{ik}^o$ exhibits a group sparsity across the K frequencies and, as such, the DOA estimation problem can be solved in the context of group sparse reconstruction.

B. DOA estimation using both self- and cross-lags

While CS-based DOA estimation can be performed based only on the cross-lag correlations without losing the available co-array DOFs, the utilization of both self- and cross-lags makes full use of the observed data and may yield performance improvement.

The self-lag covariance matrix for the data vector $\mathbf{x}_k(t)$, corresponding to the kth frequency for $1 \leq k \leq K$, can be obtained as

$$\mathbf{R}_{\mathbf{xx}}^{(k,k)} = \mathrm{E}\left[\mathbf{x}_k(t)\mathbf{x}_k^H(t)\right] = \mathbf{A}_k \mathbf{R}_{\mathbf{ss}}^{(k,k)} \mathbf{A}_k^H + \sigma_{n_k}^2 \mathbf{I}_L = \sum_{q=1}^{Q} \sigma_{kq}^2 \mathbf{a}_k(\theta_q) \mathbf{a}_k^H(\theta_q) + \sigma_{n_k}^2 \mathbf{I}_L, \quad (22)$$

where $\mathbf{R}_{\mathbf{ss}}^{(k,k)} = \mathrm{E}\left[\mathbf{s}_k(t)\mathbf{s}_k^H(t)\right] = \mathrm{diag}([\sigma_{k1}^2, \ldots \sigma_{kQ}^2])$ is the auto-covariance matrix corresponding to the kth frequency, and the signal power σ_{kq}^2, $q = 1, \ldots, Q$, is real and positive. Similarly, vectorizing $\mathbf{R}_{\mathbf{xx}}^{(k,k)}$ in (22) yields an $L^2 \times 1$ vector

$$\mathbf{z}_{kk} = \mathrm{vec}\left(\mathbf{R}_{\mathbf{xx}}^{(k,k)}\right) = \tilde{\mathbf{A}}_k \mathbf{b}_k + \sigma_{n_k}^2 \mathbf{i}, \quad k \in [1, \ldots, K], \quad (23)$$

where $\tilde{\mathbf{A}}_k = [\tilde{\mathbf{a}}_k(\theta_1), \ldots, \tilde{\mathbf{a}}_k(\theta_Q)]$, $\tilde{\mathbf{a}}_k(\theta_q) = \mathbf{a}_k^*(\theta_q) \bigotimes \mathbf{a}_k(\theta_q)$, $\mathbf{b}_k = [\sigma_{k1}^2, \ldots, \sigma_{kQ}^2]^T$, and $\mathbf{i} = \mathrm{vec}(\mathbf{I}_L)$. Similarly, $\mathbf{z}_{kk}$ can be sparsely represented as

$$\mathbf{z}_{kk} = \tilde{\mathbf{B}}_{kk}^o \tilde{\mathbf{b}}_{kk}^o, \quad k \in [1, \ldots, K], \quad (24)$$

where $\tilde{\mathbf{B}}_{kk}^o = \left[\tilde{\mathbf{A}}_k^o, \mathbf{i}\right]$ and $\tilde{\mathbf{b}}_{kk}^o = \left[\mathbf{b}_{kk}^{o\,T}, \sigma_{n_k}^2\right]^T$. Herein, $\tilde{\mathbf{A}}_k^o$ is the collection of steering vectors $\tilde{\mathbf{a}}_k(\theta_g)$ with $g = 1, \ldots, G$, and $\mathbf{b}_{kk}^o$ is the sparse vector whose non-zero entry positions correspond to the DOAs of the signals. Similar to $\mathbf{b}_{ik}^o$ in Eqn. (21), $\tilde{\mathbf{b}}_{kk}^o$ also exhibits a group sparsity across the K frequencies and shares the same sparsity pattern with $\mathbf{b}_{ik}^o$. Thus, by combing the results of $\mathbf{z}_{ik}$ and $\mathbf{z}_{kk}$, both self- and cross-lag covariances can be fully utilized for possible performance improvement based on group sparsity.

By using $\tilde{\mathbf{z}}_{ik}, i, k \in [1, \ldots, K]$, to denote both cross-lag vector $\mathbf{z}_{ik}$, $i \neq k$, and self-lag vector $\mathbf{z}_{kk}$, the DOA estimation problem using both self- and cross-lag covariances can be reformulated as:

$$\tilde{\mathbf{z}}_{ik} = \tilde{\mathbf{B}}_{ik}^o \tilde{\mathbf{b}}_{ik}^o + \boldsymbol{\epsilon}_{ik}, \quad i, k \in [1, \ldots, K], \quad (25)$$

where each vector $\tilde{\mathbf{z}}_{ik}$ employs its respective $L^2 \times (G+1)$ dictionary matrix,

$$\tilde{\mathbf{B}}_{ik}^o = \begin{cases} \left[\tilde{\mathbf{A}}_{kk}^o, \mathbf{i}\right], & i = k, \\ \left[\tilde{\mathbf{A}}_{ik}^o, \mathbf{0}\right], & i \neq k, \end{cases} \quad (26)$$

and $\mathbf{0}$ denotes the all zero vector of $L^2 \times 1$. An $L^2 \times 1$ error vector $\boldsymbol{\epsilon}_{ik}$ is included in (25) to account for the discrepancies between the statistical expectation and the sample average in computing the covariance matrices $\mathbf{R}_{\mathbf{xx}}^{(i,k)}, i, k = 1, \ldots, K$. The discrepancies are modelled as i.i.d. complex Gaussian as a result of a sufficiently large number of samples employed in the averaging.

Note that exploiting the self-lag covariances, together with the cross-lags, requires expanding the dimension of the unknown sparse vector $\tilde{\mathbf{b}}_{ik}^o$ by an additional element of the noise power $\sigma_{n_k}^2$. In this case, the first G elements of the obtained estimates of $\tilde{\mathbf{b}}_{ik}^o$ are used to determine the DOAs, whereas the last element of $\tilde{\mathbf{b}}_{ik}^o$ is discarded.

A number of effective algorithms within the convex optimization and Bayesian sparse learning frameworks are available to solve the complex-valued group sparse reconstruction problem. In this paper, the CMT-BCS algorithm proposed in [20] and summarized in Section V-C is used due to its superior performance and robustness to dictionary coherence.

C. CMT-BCS algorithm

We use the CMT-BCS to determine the DOAs of the targets which are treated as group sparse complex observations. In this subsection, we briefly review the CMT-BCS approach based on [20]. Assume that the entries in the sparse vectors $\mathbf{r}_{ik}$ are drawn from the product of the following zero-mean Gaussian distributions:

$$\tilde{\mathbf{b}}_{ik}^{o^g} \sim \mathcal{N}(\tilde{\mathbf{b}}_{ik}^{o^g}|\mathbf{0}, \alpha_g \mathbf{I}_2), \quad g \in [1, \ldots, G], \tag{27}$$

where $\tilde{\mathbf{b}}_{ik}^{o^g}$ is a vector consisting of the real part coefficient, $\tilde{\mathbf{b}}_{ik}^{o^{g R}}$, and the imagery part coefficient, $\tilde{\mathbf{b}}_{ik}^{o^{g I}}$, with respect to the gth grid in $\tilde{\mathbf{b}}_{ik}^o$. In addition, $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_G]^T$ is a vector that contains variances of $\tilde{\mathbf{b}}_{ik}^{o^g}$, $g = 1, ..., G$. Note that the vector $\boldsymbol{\alpha}$ is shared by all groups to enforce the group sparsity. It is easy to confirm that $\tilde{\mathbf{b}}_{ik}^{o^g}$ trends to be zero when α_g is set to zero [14].

To encourage the sparsity of $\tilde{\mathbf{b}}_{ik}^o$, a Gamma prior is placed on α_g^{-1}, which is conjugate to the Gaussian distribution,

$$\alpha_g^{-1} \sim \mathrm{Gamma}(\alpha_g^{-1}|a, b), \quad g \in [1, \ldots, G], \tag{28}$$

where $\mathrm{Gamma}(x^{-1}|a, b) = \Gamma(a)^{-1} b^a x^{-(a-1)} e^{-\frac{b}{x}}$, with $\Gamma(\cdot)$ denoting the Gamma function, and a and b are hyper-parameters.

As the covariance matrix is estimated from the received data samples, a Gaussian prior $\mathcal{N}(\mathbf{0}, \beta_0 \mathbf{I}_2)$ is also placed on the $\boldsymbol{\epsilon}_{ik}$. Similarly, the Gamma prior is placed on β_0^{-1} with hyper-parameters c and d.

The CMT-BCS algorithm carries out a Bayesian inference by the Gibbs samplers [20]. Once the parameters $\boldsymbol{\alpha}$ and β_0 are estimated by maximizing the marginal likelihood, the joint pos-

terior density function of $\tilde{\mathbf{b}}_{ik}^o$ can be obtained analytically using Bayes' rule. Define $\tilde{\mathbf{b}}_{ik}^{o_{RI}} = \left[(\tilde{\mathbf{b}}_{ik}^{o_R})^T, (\tilde{\mathbf{b}}_{ik}^{o_I})^T \right]^T$, with $\tilde{\mathbf{b}}_{ik}^{o_R} = [b_{ik}^{o_1R}, \ldots, b_{ik}^{o_GR}]^T$ and $\tilde{\mathbf{b}}_{ik}^{o_I} = [b_{ik}^{o_1I}, \ldots, b_{ik}^{o_GI}]^T$. Then,

$$\Pr(\tilde{\mathbf{b}}_{ik}^{o_{RI}} | \tilde{\mathbf{z}}_{ik}, \tilde{\mathbf{B}}_{ik}^o, \boldsymbol{\alpha}, \beta_0) = \mathcal{N}(\tilde{\mathbf{b}}_{ik}^{o_{RI}} | \boldsymbol{\mu}_{ik}, \boldsymbol{\Sigma}_{ik}),$$

where

$$\tilde{\mathbf{z}}_{ik}^{RI} = \left[\mathrm{Re}(\tilde{\mathbf{z}}_{ik})^T, \mathrm{Im}(\tilde{\mathbf{z}}_{ik})^T \right]^T \tag{29}$$

$$\boldsymbol{\mu}_{ik} = \beta_0^{-1} \boldsymbol{\Sigma}_{ik} \boldsymbol{\Psi}_{ik}^T \tilde{\mathbf{z}}_{ik}^{RI}, \tag{30}$$

$$\boldsymbol{\Sigma}_{ik} = \left[\beta_0^{-1} \boldsymbol{\Psi}_{ik}^T \boldsymbol{\Psi}_{ik} + \mathbf{F}^{-1} \right]^{-1}, \tag{31}$$

$$\boldsymbol{\Psi} = \begin{bmatrix} \mathrm{Re}(\tilde{\mathbf{B}}_{ik}^o) & -\mathrm{Im}(\tilde{\mathbf{B}}_{ik}^o) \\ \mathrm{Im}(\tilde{\mathbf{B}}_{ik}^o) & \mathrm{Re}(\tilde{\mathbf{B}}_{ik}^o) \end{bmatrix}, \tag{32}$$

$$\mathbf{F} = \mathrm{diag}(\alpha_1, \ldots, \alpha_G, \alpha_1, \ldots, \alpha_G). \tag{33}$$

Note that the mean and variance of each scattering coefficients can be derived using Eqns. (30) and (31) when $\boldsymbol{\alpha}$ and β_0 are given. On the other hand, the values of $\boldsymbol{\alpha}$ and β_0 are determined by maximizing the logarithm of the marginal likelihood, i.e.,

$$\{\boldsymbol{\alpha}, \beta_0\} = \arg \max_{\boldsymbol{\alpha}, \beta_0} \mathcal{L}(\boldsymbol{\alpha}, \beta_0), \tag{34}$$

where

$$\mathcal{L}(\boldsymbol{\alpha}, \beta_0) = \sum_{i,k=1}^{K} \log \Pr(\tilde{\mathbf{b}}_{ik}^{o_{RI}} | \boldsymbol{\alpha}, \beta_0)$$

$$= \mathrm{const} - \frac{1}{2} \sum_{i,k=1}^{K} \log |\mathbf{C}_{ik}| + (\tilde{\mathbf{z}}_{ik}^{RI})^T \mathbf{C}_{ik}^{-1} \tilde{\mathbf{z}}_{ik}^{RI}, \tag{35}$$

and $\mathbf{C}_{ik} = \beta_0 \mathbf{I} + \boldsymbol{\Psi}_{ik} \mathbf{F} \boldsymbol{\Psi}_{ik}^T$. A type-II maximum likelihood (ML) approximation [26] employs the point estimates for $\boldsymbol{\alpha}$ and β_0 to maximize Eqn. (35), which can be implemented via the expectation maximization (EM) algorithm to yield

$$\alpha_g^{(\mathrm{new})} = \frac{1}{K^2} \sum_{i,k=1}^{K} (\mu_{ik,g}^2 + \mu_{ik,g+G}^2 + \Sigma_{ik,gg} + \Sigma_{ik,(g+G)(g+G)}), \quad g \in [1, \ldots, G], \tag{36}$$

$$\beta_0^{(\mathrm{new})} = \frac{1}{2GK^2} \sum_{i,k=1}^{K} \left(\mathrm{Tr}[\boldsymbol{\Sigma}_{ik} \boldsymbol{\Psi}_{ik}^T \boldsymbol{\Psi}_{ik}] + ||\tilde{\mathbf{z}}_{ik}^{RI} - \boldsymbol{\Psi}_{ik} \boldsymbol{\mu}_{ik}||_2^2 \right), \tag{37}$$

where $\mu_{ik,g}^2$ and $\mu_{ik,g+G}^2$ are the gth and $(g+G)$th elements in vector $\boldsymbol{\mu}_{ik}$, and $\Sigma_{ik,gg}$ and $\Sigma_{ik,(g+G)(g+G)}$ are the (g,g) and $(g+G, g+G)$ entries in matrix $\boldsymbol{\Sigma}_{ik}$. Because $\boldsymbol{\alpha}$ and β_0 depend

on $\boldsymbol{\mu}_{ik}$ and $\boldsymbol{\Sigma}_{ik}$, the CMT-BCS algorithm is iterative and iterates between Eqns. (30)-(31) and Eqns. (36)-(37), until a convergence criterion is satisfied or the maximum number of iterations is reached.

VI. SIMULATION RESULTS

In the simulations, the CMT-BCS algorithm is used to estimate the DOAs of the signal arrivals with hyper-parameters $a = b = c = d = 0$. The maximum number of iterations in the Gibbs sampling is set to 200, and the sampler with the maximum marginal likelihood in the last 20 samples is chosen as the estimate of $\tilde{\mathbf{b}}_{ik}^{o}$.

We present four examples to demonstrate the effectiveness of the proposed technique. For all examples, Q targets, which are uniformly distributed between $-60°$ and $60°$ are assumed to imping a ULA with $L = 4$. The grid interval in the angular space is set to $0.25°$. In addition, the noise power at each frequency is assumed to be identical and the phase difference between the received signal corresponding to each frequency pair is independent and uniformly distributed over $[0, 2\pi)$. We evaluate the performance through Monte Carlo simulations. The root mean-square error (RMSE) of the estimated DOA of the signal arrivals, expressed as

$$\text{RMSE} = \sqrt{\frac{1}{IQ} \sum_{i=1}^{I} \sum_{q=1}^{Q} (\hat{\theta}_q(i) - \theta_q)^2},$$

is used as the metric for performance evaluation with respect to the input SNR, where $\hat{\theta}_q(i)$ is the estimate of θ_q for the ith Monte Carlo trial, $i = 1, \dots, I$. We use $I = 500$ independent trials in all simulations.

A. Example I: Achievable number of DOFs with two co-prime frequencies

The number of achievable DOFs from the 4-element ULA with two co-prime frequencies is first illustrated in Fig. 5. $Q = 9$ targets are considered, which are much larger than the number of physical sensors. Two co-prime frequencies with $M_1 = 3$ and $M_2 = 4$ are exploited. Because the virtual sensor lags are obtained from the estimated covariance matrix based on the received data samples, as in Eqn. (5), the virtual steering matrix is sensitive to the noise contamination. To clearly demonstrate the number of achievable DOFs, therefore, we use 10000 noise-free snapshots to obtain a relatively clean covariance matrix. Fig. 5(a) shows the estimated spatial spectrum from the proposed co-array, which yields a co-array with $\eta = 19$ virtual sensors, and

the result of the conventional non-co-array scenario is depicted in Fig. 5(b). It is clear that the co-array provides a sufficient number of DOFs to correctly identify the DOAs of all 9 targets, whereas the non-co-array approach fails.

B. Example II: DOA estimation using only cross-lags vs. both self- and cross-lags

In Figs. 6 and 7, the results obtained by using both self- and cross-lags are compared to those using only the cross-lags. $Q = 6$ targets are considered and two co-prime frequencies with $M_1 = 3$ and $M_2 = 4$ are exploited. The RMSE with respect to the input SNR is depicted in Fig. 6, where 2000 snapshots are used. At a moderate or high SNR, the utilization of both self- and cross-lag covariances benefits from additional measurement offered by the self-lags, resulting in the improved performance than the cross-lag only scenario. In Fig. 7, such improvement is demonstrated with fewer false peaks in the estimated spectra, where the input SNR is 10 dB. On the other hand, in the low SNR region, as shown in Fig. 6, the performance of the algorithm using cross-lag covariances only is better than the results using both self- and cross-lag covariances. In this case, both vectors $\mathbf{z}_{ik}$ and $\mathbf{z}_{kk}$ are highly perturbed by the noise. The inclusion of self-lag covariance matrices causes additional errors in the noise power estimation in (25), whereas this term does not exist in the cross-lag covariances.

C. Example III: DOA estimation using different frequency pairs

This example compares the DOA estimation performance when different frequency pairs are used. In the first frequency pair, $M_1 = 3$ and $M_2 = 4$ are assumed, yielding $\eta = 19$ elements in the virtual co-array. In the second frequency pair, we assume $M_1 = 6$ and $M_2 = 7$, resulting in $\eta = 31$ virtual co-array lags. In Fig. 8, the RMSE performance is presented as a function of the number of targets, Q, where SNR is assumed to be 10 dB and 2000 snapshots are exploited. The result shows that the second frequency pair outperforms the first one due to its higher number of DOFs and the larger aperture.

D. Example IV: Performance of multiple co-prime frequency cases

To demonstrate the merits of exploiting multiple co-prime frequencies, we first consider a three frequency case with $M_1 = 5$, $M_2 = 6$ and $M_3 = 7$. Similarly, 10000 noise-free snapshots are used to obtain a relatively clean covariance matrix. Fig. 9 shows the estimated spectrum for

$Q = 13$ targets. Note in this case that all targets are resolved correctly due to a high number of DOFs and a small number of missing positions in the co-array.

Then, the RMSE performance of the three frequency case is presented in Fig. 10 with respect to the input SNR, where $Q = 13$ and 2000 snapshots are assumed. For comparison purposes, a four frequency scenario with $M_1 = 5$, $M_2 = 6$, $M_3 = 7$, and $M_4 = 11$ is also considered. It is clearly shown that the performance is significantly improved as the number of frequencies is increased.

VII. CONCLUSIONS

In this paper, we developed a co-prime array implementation using a sparse uniform linear array with multiple co-prime frequencies. We derived the analytical expression for the number of unique lags of the yielding difference co-array to determine the number of detectable targets. The complex multitask Bayesian compressive sensing algorithm was used to exploit the group sparse direction-of-arrivals (DOAs) across different frequencies for effective spatial spectrum estimation. The number of detectable targets and the DOA performance are improved as the number of frequencies increases. The effectiveness of the proposed technique and analysis is verified using simulation results.

VIII. APPENDIX

A. Proof of Proposition 1

Denote η_t and η_o as the total number of lags in $\tilde{\mathbb{L}}_P$ and the number of overlaps between the set $\tilde{\mathbb{L}}_c$ and $\tilde{\mathbb{L}}_c^-$, respectively. Then, the number of distinct lags in $\tilde{\mathbb{L}}_P$ can be expressed as

$$\eta = \eta_t - \eta_o. \tag{38}$$

Both $\tilde{\mathbb{L}}_c$ and $\tilde{\mathbb{L}}_c^-$ have L^2 distinct lags due to the co-primality of M_1 and M_2. It is easy to confirm that

$$\eta_t = 2L^2. \tag{39}$$

Given arbitrary lags $\tilde{l}_{c_m} = M_1 l_{1_m} - M_2 l_{2_m}$ and $\tilde{l}_{c_n} = M_2 l_{2_n} - M_1 l_{1_n}$ in set $\tilde{\mathbb{L}}_c$ and $\tilde{\mathbb{L}}_c^-$, respectively, where the indexes $0 \le l_{1_m} \le L - 1$, $0 \le l_{2_m} \le L - 1$, $0 \le l_{1_n} \le L - 1$ and $0 \le l_{2_n} \le L - 1$. Had $\tilde{l}_{c_m} = \tilde{l}_{c_n}$ been held, we would have $M_1(l_{1_m} + l_{1_n}) = M_2(l_{2_m} + l_{2_n})$.

It is evident that they overlap at 0 position provided $l_{1_m} = l_{1_n} = l_{2_m} = l_{2_n} = 0$. When $l_{1_m} + l_{1_n} \neq 0$, the requirement is equivalent to

$$\frac{M_1}{M_2} = \frac{l_{2_m} + l_{2_n}}{l_{1_m} + l_{1_n}}. \tag{40}$$

(a) When $M_2 \geq 2L - 1$, the maximum value of $l_{1_m} + l_{1_n}$ is less than M_2. Since M_1 and M_2 are co-prime, it is indicated that M_1/M_2 cannot be reduced to a ratio of smaller integers. As a result, (40) cannot be hold. In other word, $\tilde{\mathbb{L}}_c$ and $\tilde{\mathbb{L}}_c^-$ only coincide at 0 position, i.e.,

$$\eta_o = 1. \tag{41}$$

Substituting (39) and (41) into (38), we can obtain

$$\eta = 2L^2 - 1. \tag{42}$$

(b) When $L \leq M_2 < 2L - 1$, the relationship $0 \leq l_{1_m} + l_{1_n} \leq 2L - 2 < 2M_2$ is guaranteed. Due to the co-primality of M_1 and M_2, (40) is valid if and only if

$$l_{2_m} + l_{2_n} = M_1,$$
$$l_{1_m} + l_{1_n} = M_2. \tag{43}$$

Since $0 \leq l_{1_m}, l_{2_m} \leq L - 1$, the requirement is equivalent to

$$M_1 - (L - 1) \leq l_{2_n} \leq M_1,$$
$$M_2 - (L - 1) \leq l_{1_n} \leq M_2. \tag{44}$$

Because $0 \leq l_{1_n}, l_{2_n} \leq L - 1$, we obtain the following relationship

$$\max\{M_1 - (L - 1), 0\} \leq l_{2_n} \leq \min\{M_1, L - 1\},$$
$$M_2 - (L - 1) \leq l_{1_n} \leq L - 1, \tag{45}$$

where $\max\{a, b\}$ and $\min\{a, b\}$ are operators, returning maximum and minimum values between a and b, respectively. Since $L \leq M_1 < M_2$, Eqn. (45) becomes

$$M_1 - (L - 1) \leq l_{2_n} \leq L - 1,$$
$$M_2 - (L - 1) \leq l_{1_n} \leq L - 1. \tag{46}$$

It is indicated that $2L - 1 - M_1$ and $2L - 1 - M_2$ integers are in the respective range of l_{2_n} and l_{1_n}. In addition to 0 position, there are $(2L - 1 - M_1)(2L - 1 - M_2)$ combination to satisfy (40), i.e,

$$\eta_o = (2L - 1 - M_1)(2L - 1 - M_2) + 1. \tag{47}$$

Substituting (39) and (47) into (38), we can obtain

$$\eta = 2L^2 - 1 - (2L - 1 - M_1)(2L - 1 - M_2). \tag{48}$$

(c) When $L \leq M_2 < 2L - 1$ and $1 \leq M_1 < L$, (45) is equivalent to

$$0 \leq l_{2_n} \leq M_1,$$

$$M_2 - (L - 1) \leq l_{1_n} \leq L - 1. \tag{49}$$

As such, there are $(M_1 + 1)(2L - 1 - M_2)$ integers satisfying (40). Therefore,

$$\eta_o = (M_1 + 1)(2L - 1 - M_2) + 1. \tag{50}$$

Substituting (39) and (50) into (38), we can obtain

$$\eta = 2L^2 - 1 - (M_1 + 1)(2L - 1 - M_2). \tag{51}$$

B. Proof of Proposition 2

(a) When K multiple frequencies are exploited, there are $K(K-1)/2$ pairs of frequencies. As such, the total number of lags, η_t, which includes both unique and overlapping lags, is

$$\eta_t = K(K-1)L^2, \tag{52}$$

as each pair has $2L^2$ lags. To obtain the maximum number of achievable unique lags of the co-array, we consider the case that each pair achieves its respective maximum number of unique lags, as described in Section IV-A, and the number of overlapping lags between different pairs is minimum. In this case, redundancy between different co-prime pairs happens at the following two cases: (a) The zeroth entry is shared by all $K(K - 1)/2$ pairs of co-prime frequencies with a total number of $K(K - 1)$ overlapping lags, whereas the unique lag in this position is 1; (b) At all self-lag positions because the array sensors corresponding to each frequency are used to generate $K - 1$ co-prime frequency pairs. As each frequency yields $2(L-1)$ non-zero self-lags in $\tilde{\mathbb{L}}_c \bigcup \tilde{\mathbb{L}}_c^-$, there are $K(K-1) \times 2(L-1)$ total lag entries with $2K(L-1)$ unique lags, yielding $2K(K-2)(L-1)$ redundancies to be discounted in computing the available unique lags. As a result, we can obtain the maximum number of the achievable unique lags of the co-array as

$$\eta = \eta_t - \eta_o = (L^2 - 1)(K^2 - K) - 2(L - 1)(K^2 - 2K) + 1. \tag{53}$$

REFERENCES

[1] Y. D. Zhang, M. G. Amin, F. Ahmad, and B. Himed, "DOA estimation using a sparse uniform linear array with two CW signals of co-prime frequencies," in *Proc. IEEE Int. Workshop on Comp. Adv. in Multi-Sensor Adaptive Proc.*, Saint Martin, Dec. 2013, pp. 404–407.

[2] S. Qin, Y. D. Zhang, and M. G. Amin, "DOA estimation exploiting coprime frequencies," in *Proc. SPIE Wireless Sens., Localization, Process. Conf.*, Baltimore, MD, May 2014, vol. 9103, pp. 91030E1–91030E7.

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

[4] R. Roy and T. Kailath, "ESPRIT – Estimation of signal parameters via rotation invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 17, no. 7, pp. 984–995, July 1989.

[5] R. T. Hoctor and S. A. Kassam, "The unifying role of the co-array in aperture synthesis for coherent and incoherent imaging," *Proc. IEEE*, vol. 78, no. 4, pp. 735–752, April 1990.

[6] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas Propagat.*, vol. 16, no. 2, pp. 172–175, March 1968.

[7] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 573–586, Feb. 2011.

[8] P. Pal and P. P. Vaidyanathan, "Coprime sampling and the MUSIC algorithm," in *Proc. IEEE Digital Signal Process. Workshop and IEEE Signal Process. Education Workshop*, Sedona, AZ, Jan. 2011, pp. 289–294.

[9] S. Qin, Y. D. Zhang, and M. G. Amin, "Generalized coprime array configurations," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, A Coruña, Spain, June 2014, pp. 529–532.

[10] Y. D. Zhang, S. Qin, and M. G. Amin, "DOA estimation exploiting coprime arrays with sparse sensor spacing," in *Proc. IEEE ICASSP*, Florence, Italy, May 2014, pp. 2267–2271.

[11] S. Qin, Y. D. Zhang, and M. G. Amin, "Generalized coprime array configurations for direction-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1377–1390, March 2015.

[12] S. Qin, Y. D. Zhang, and M. G. Amin, "DOA estimation of mixed coherent and uncorrelated signals exploiting a nested MIMO system," in *Proc. IEEE Benjamin Franklin Symposium on Microwave and Antenna Sub-systems*, Philadelphia, PA, Sep. 2014.

[13] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Machine Learning Research*, vol. 1, no. 9, pp. 211–244, 2001.

[14] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, 2008.

[15] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 92–106, 2009.

[16] Z. Zhang and D. B. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayyesian learning," *IEEE J. Sel. Topics Signal Process.*, vol. 57, no. 1, pp. 92–106, 2009.

[17] Z. Zhang and D. B. Rao, "Extension of SBL algorithms for the recovery of block sparse signals with intrablock

correlation," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 2009–2015, April 2013.

[18] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Info. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[19] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal Statistical Society*, Series B, vol. 58, no. 1, pp. 267–288, 1996.

[20] Q. Wu, Y. D. Zhang, and M. G. Amin, "Complex multitask Bayesian compressive sensing," in *Proc. IEEE ICASSP*, Florence, Italy, May 2014, pp. 3375–3379.

[21] D. Wipf and S. Nagarajan, "Beamforming using the relevance vector machine," in *Proc. Int. Conf. on Machine Learning*, Corvallis, OR, pp. 1023–1030, 2007.

[22] L. Poli, G. Oliveri, F. Viani, and A. Massa, "MT-BCS-based microwave imaging approach through minimum-norm current expansion," *IEEE Trans. Antennas Propagat.*, vol. 61, no. 9, pp. 4722–4732, Sept. 2013.

[23] P. Pal and P. P. Vaidyanathan, "Pushing the limits of sparse support recovery using correlation information," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 711–726, Feb. 2015.

[24] P. Pal and P. P. Vaidyanathan, "Nested Arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4167–4181, Aug. 2010.

[25] Y. D. Zhang, M. G. Amin, and B. Himed, "Sparsity-based DOA estimation using co-prime arrays," in *Proc. IEEE ICASSP*, Vancouver, Canada, May 2013, pp. 3967–3971.

[26] D. J. C. MacKay, "Bayesian interpolation," *Neural Comput.*, vol. 4, no. 3, pp. 415–447, 1992.

Fig. 1. The coprime array configuration.

(a) A sparse ULA
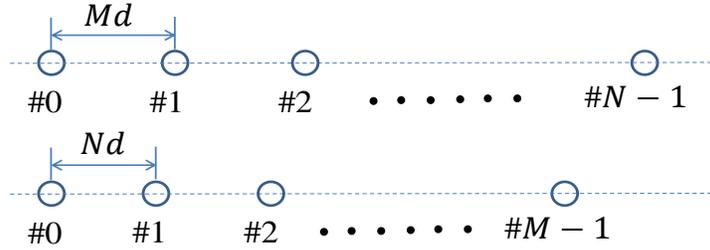
(b) Equivalent structure with two coprime frequencies
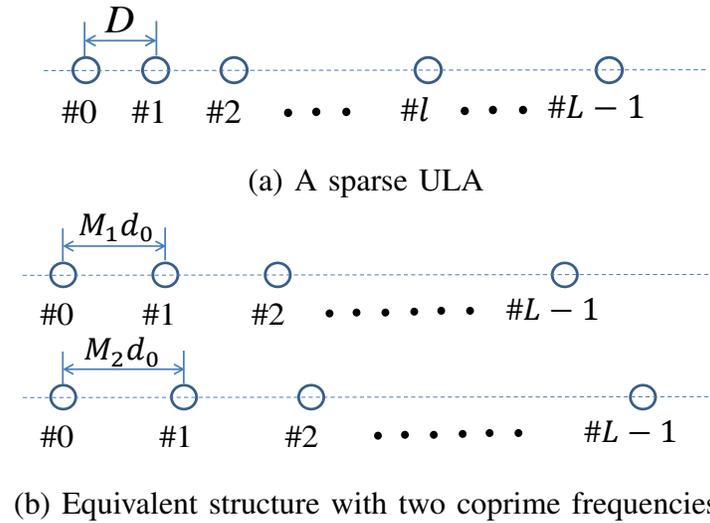
Fig. 2. A sparse ULA with two coprime frequencies configuration.

(a) $M_1=2$ and $M_2=3$

(b) $M_1=3$ and $M_2=4$

(c) $M_1=6$ and $M_2=7$
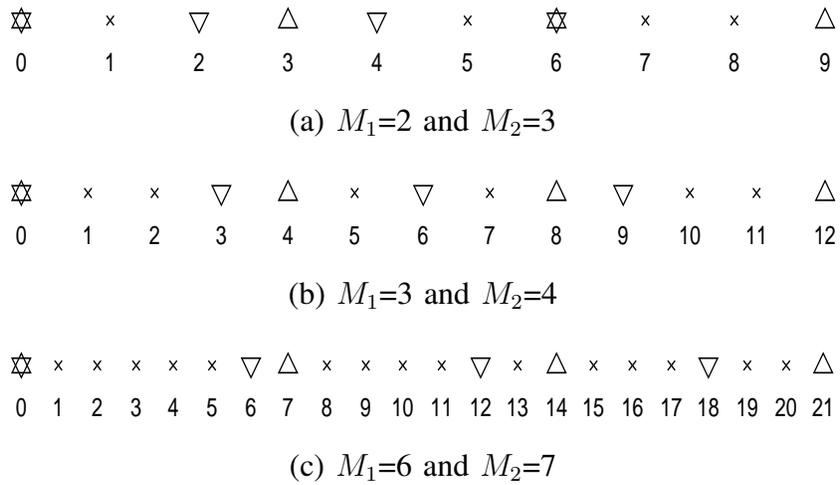
Fig. 3. Equivalent sensor positions for different M_1 and M_2 with $L = 4$ elements ULA (∇: Sub-array with M_1; $\triangle$: Sub-array with M_2).

(a) M_1=2 and M_2=3

(b) M_1=3 and M_2=4

(c) M_1=6 and M_2=7

(d) M_1=1 and M_2=4
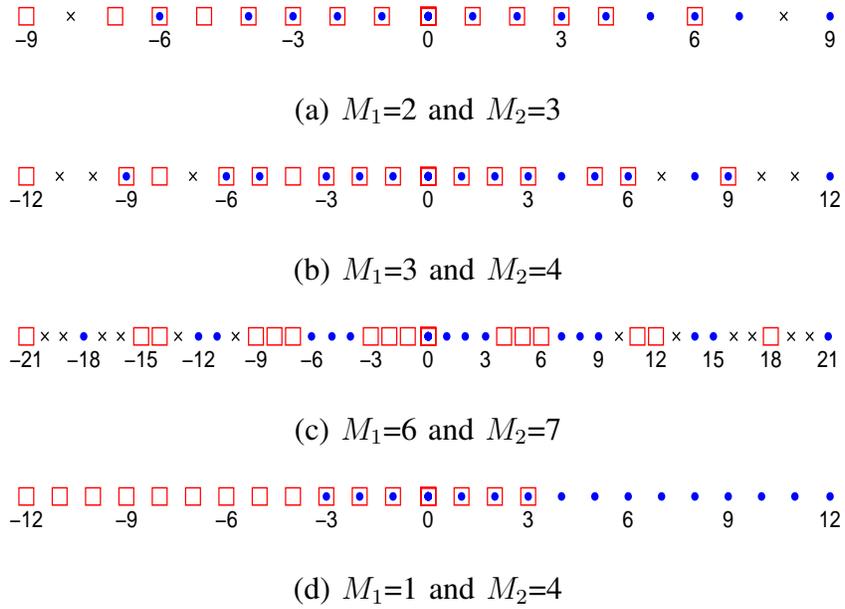
Fig. 4. The co-array lag positions in the set $\tilde{\mathbb{L}}_c \bigcup \tilde{\mathbb{L}}_c^-$ with $L = 4$ element ULA ($\bullet$: Positions in $\tilde{\mathbb{L}}_c$; $\square$: Positions in $\tilde{\mathbb{L}}_c^-$).

(a) Co-array scenario

(b) Non-co-array scenario
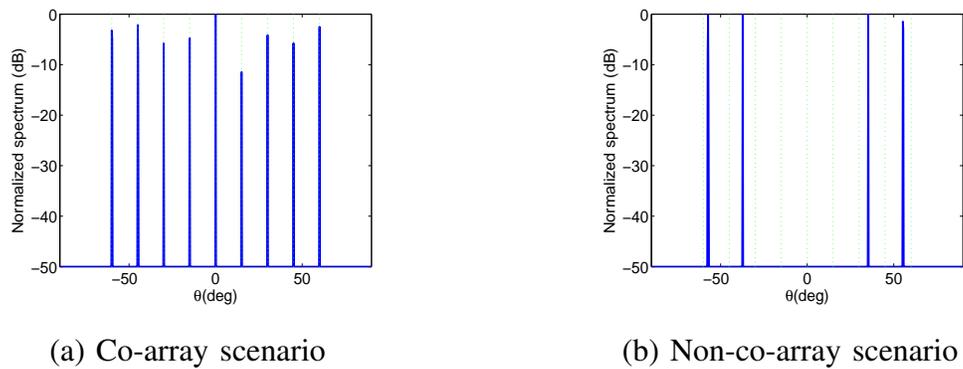
Fig. 5. Estimated spectrum using co-array and non-co-array scenarios ($M_1 = 3$, $M_2 = 4$, $Q = 9$, and 10000 noise-free snapshots).
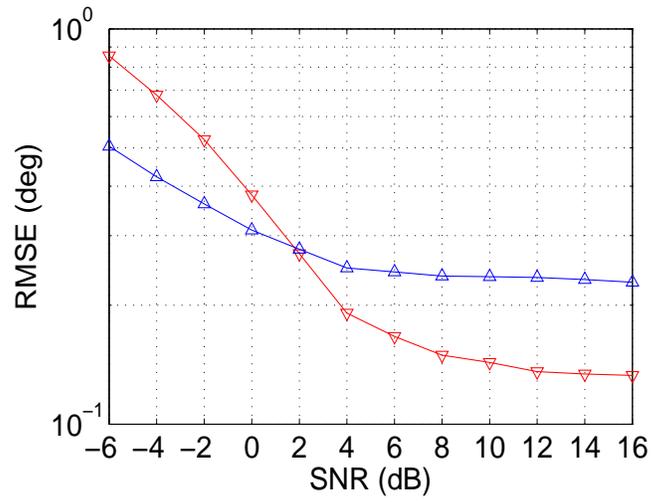
Fig. 6. RMSE versus input SNR ($M_1 = 3$, $M_2 = 4$, $Q = 6$, and 2000 snapshots; ∇: Use self- and cross-lags; $\triangle$: Use cross-lags only).

(a) Using self- and cross-lags

(b) Using cross-lags only
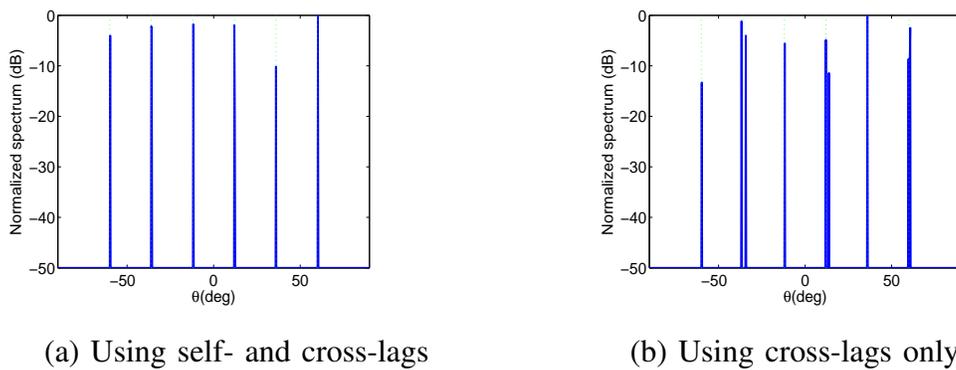
Fig. 7. Spatial spectra estimated using different lags ($M_1 = 3$, $M_2 = 4$, $Q = 6$, SNR=10 dB, and 2000 snapshots).
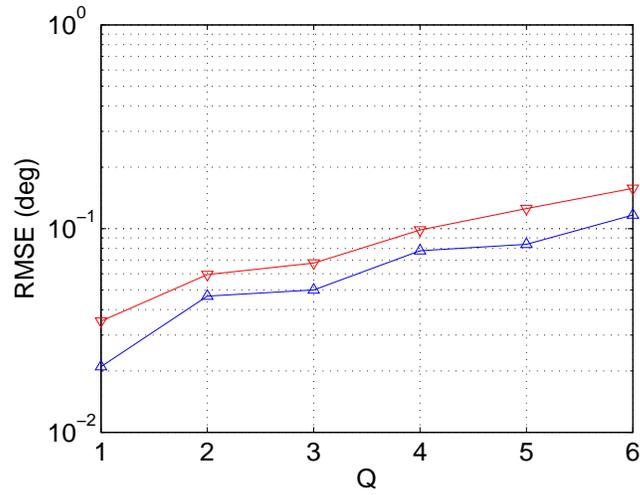
Fig. 8. RMSE versus signal number for different frequency pairs (SNR=10 dB and 2000 snapshots; ∇: first pair ($M_1 = 3$ and $M_2 = 4$); $\triangle$: second pair ($M_1 = 6$ and $M_2 = 7$)).
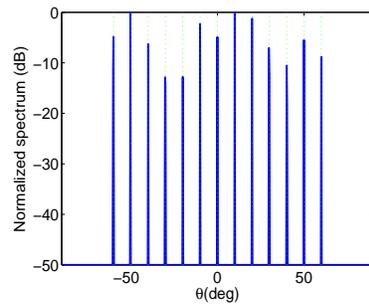
$M_1 = 5$, $M_1 = 6$ and $M_2 = 7$

Fig. 9. Spatial spectra estimated for three frequency case ($M_1 = 5$, $M_2 = 6$, $M_3 = 7$, $Q = 13$ and 10000 noise-free snapshots).
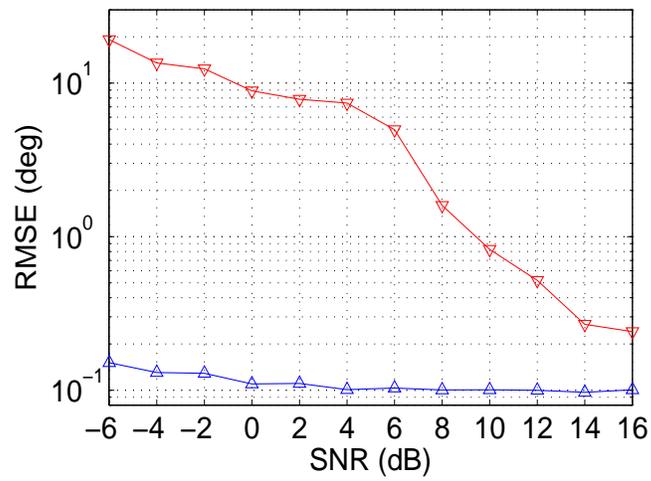
Fig. 10. RMSE versus SNR for different number of frequencies ($Q = 13$ and 2000 snapshots; $\triangledown$: $M_1 = 5$, $M_2 = 6$ and $M_3 = 7$; $\triangle$: $M_1 = 5$, $M_2 = 6$, $M_3 = 7$ and $M_4 = 11$.