Direction of Arrival Estimation Using Co-Prime Arrays: A Super Resolution Viewpoint

Zhao Tan, *Student Member, IEEE*, Yonina C. Eldar, *Fellow, IEEE*, and Arye Nehorai, *Fellow, IEEE*

***Abstract*—We consider the problem of direction of arrival (DOA) estimation using a recently proposed structure of nonuniform linear arrays, referred to as co-prime arrays. By exploiting the second order statistical information of the received signals, co-prime arrays exhibit $O(MN)$ degrees of freedom with only $M + N$ sensors. A sparsity-based recovery algorithm is proposed to fully utilize these degrees of freedom. The suggested method is based on the developing theory of super resolution, which considers a continuous range of possible sources instead of discretizing this range onto a grid. With this approach, off-grid effects inherent in traditional sparse recovery can be neglected, thus improving the accuracy of DOA estimation. We show that in the noiseless case it is theoretically possible to detect up to $\frac{MN}{2}$ sources with only $2M + N$ sensors. The noise statistics of co-prime arrays are also analyzed to demonstrate the robustness of the proposed optimization scheme. A source number detection method is presented based on the spectrum reconstructed from the sparse method. By extensive numerical examples, we show the superiority of the suggested algorithm in terms of DOA estimation accuracy, degrees of freedom, and resolution ability over previous techniques, such as MUSIC with spatial smoothing and discrete sparse recovery.**

***Index Terms*—Co-prime arrays, continuous sparse recovery, direction of arrival estimation, source number detection, super resolution.**

I. Introduction

In the last few decades, research on direction of arrival (DOA) estimation using array processing has focused primarily on uniform linear arrays (ULA) [1]. It is well known that using a ULA with N sensors, the number of sources that can be resolved by MUSIC-like algorithms is $N - 1$ [2]. New geometries [3], [4] of non-uniform linear arrays have been recently proposed to increase the degrees of freedom of the array by exploiting the covariance matrix of the received signals. Vectorizing the covariance matrix, the system model can be viewed as a virtual array with a wider aperture. In [3], a nested array structure was proposed to increase the degrees of freedom from $O(N)$ to $O(N^2)$, with only $O(N)$ sensors. However, some of the sensors in this structure are closely located, which leads to mutual coupling among these sensors. To overcome this shortcoming, co-prime arrays were proposed in [4]. Such arrays consist of two subarrays with M and N sensors respectively. It was shown that by using $O(M + N)$ number of sensors, this structure can achieve $O(MN)$ degrees of freedom. In this paper we focus on co-prime arrays.

The increased degrees of freedom provided by the co-prime structure can be utilized to improve DOA estimation. To this end, two main methodologies have been proposed to utilize this increased degrees of freedom for co-prime arrays. The first are subspace methods, such as the MUSIC algorithm. In [5], a spatial smoothing technique was implemented prior to the application of MUSIC, and the authors showed that an increased number of sources can be detected. However, the application of spatial smoothing reduces the obtained virtual array aperture [6]. The second approach uses sparsity-based recovery to overcome these disadvantages of subspace methods [6]–[9]. Traditional sparsity techniques discretize the range of interest onto a grid. Off-grid targets can lead to mismatches in the model and deteriorate the performance significantly [10]. In [11], [12] the grid mismatches are estimated simultaneously with the original signal, leading to improved performance over traditional sparse recovery methods. In [13], the joint sparsity between the original signal and the mismatch is exploited during DOA estimation. Due to the first-order approximation used in [13], the estimation performance is still limited by higher-order modeling mismatches.

To overcome grid mismatch of traditional sparsity-based methods, in this paper we apply the recently developed mathematical theory of continuous sparse recovery for super resolution [14]–[16] to DOA estimation with co-prime arrays. The term "super resolution" in this paper is related to the off-grid problem and is different from the traditional definition commonly used in DOA estimation. In [14], [15] it was shown that assuming a signal consists of spikes, the high frequency content of the signal's spectrum can be perfectly recovered in a robust fashion by sampling only the low end of its spectrum, when the minimum distance between spikes satisfies certain requirements. In [16], the author provides performance guarantees on the recovered support set of the sparse signal. One merit of this theory is that it considers all possible locations within the desired range, and thus does not suffer from model mismatches.

Manuscript received December 17, 2013; revised May 27, 2014 and August 25, 2014; accepted August 26, 2014. Date of publication September 04, 2014; date of current version October 02, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chong-Yung Chi. The work of Z. Tan and A. Nehorai was supported by the AFOSR Grant FA9550-11-1-0210, and ONR Grant N000141310050. The work of Y. C. Eldar was supported in part by the Israel Science Foundation under Grant no. 170/10, in part by the Ollendorf Foundation, and in part by a Magneton from the Israel Ministry of Industry and Trade.

Z. Tan and A. Nehorai are with the Preston M. Green Department of Electrical and Systems Engineering Department, Washington University in St. Louis, St. Louis, MO 63130 USA (e-mail: tanz@ese.wustl.edu; nehorai@ese.wustl.edu).

Y. C. Eldar is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: yonina@ee.technion.ac.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2014.2354316

Here we extend the mathematical theory of super resolution to DOA estimation with co-prime arrays under Gaussian noise. The effective noise resulting from the usage of co-prime arrays consists of a term with a known structure and another term containing quadratic combinations of Gaussian noise. Therefore, we modify the reconstruction method to fit these particular noise properties and prove the robustness of our approach by analyzing the noise statistics. We also prove that with $2M+N$ sensors in a co-prime array, it is possible to detect up to $O(MN)$ sources robustly. Previous identifiability research using traditional compressed seining for co-prime arrays [9] was based on the idea of mutual coherence [17]. Using mutual coherence it can be shown that co-prime arrays increase the number of detected sources from $O(M+N)$ to $O(MN)$, but this analysis is valid only for very small values of the number of sources.

Source number detection is another main application of array processing. Various methods have been proposed over the years based on the eigenvalues of the signal space, such as the Akaike information criterion [18], second-order statistic of eigenvalues (SORTE) [19], and the predicted eigen-threshold approach [20]. The authors of [21] showed that among these methods, SORTE often leads to better detection performance. Here we combine the SORTE approach with the spectrum reconstructed from the proposed DOA estimation algorithm to determine the number of sources. Through this source number detection, we identify which reconstructed spikes are true detections and which are false alarms.

The paper is organized as follows. In Section II, we introduce the DOA estimation model and explain how co-prime arrays can increase the degrees of freedom. In Section III, we adapt the theory of super resolution to co-prime arrays, and analyze the robustness of this extension by studying the statistics of the noise pattern in the model. We propose a numerical method to perform DOA estimation for co-prime arrays in Section IV. We then extend this approach to detect the number of sources in Section V. Section VI presents extensive numerical simulations demonstrating the advantages of our method in terms of estimation accuracy, degrees of freedom, and resolution ability.

Throughout the paper, we use capital italic bold letters to represent matrices and operators, and lowercase italic bold letters to represent vectors. For a given matrix $\boldsymbol{A}$, $\boldsymbol{A}^*$ denotes the conjugate transpose matrix, $\boldsymbol{A}^T$ denotes the transpose, $\boldsymbol{A}^{\mathrm{H}}$ represents the conjugate matrix without transpose, and A_{mn} denotes the (m,n)th element of $\boldsymbol{A}$. We use $\otimes$ to denote the Kronecker product of two matrices. For a given operator $\boldsymbol{F}$, $\boldsymbol{F}^*$ denotes the conjugate operator of $\boldsymbol{F}$. Given a vector $\boldsymbol{x}$, we use $\|\boldsymbol{x}\|_1$ and $\|\boldsymbol{x}\|_2$ for the ℓ_1 and ℓ_2 norms; x_i and $x[i]$ are both used to represent the ith element of $\boldsymbol{x}$. Given a function f, $\|f\|_{L_1}, \|f\|_{L_2}, \|f\|_{L_\infty}$ are its $\ell_1, \ell_2, \ell_\infty$ norms.

II. Direction of Arrival Estimation and Co-Prime Arrays

Consider a linear sensor array with L sensors which may be non-uniformly located. Assume that there are K narrow band sources located at $\theta_1, \theta_2, \ldots, \theta_K$ with signal powers $\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2$. The steering vector for the kth source located at θ_k is $\boldsymbol{a}(\theta_k) \in \mathbb{C}^{L\times 1}$ with lth element $e^{\boldsymbol{j}(2\pi/\lambda)d_l\sin(\theta_k)}$, in which d_l is the location of the lth sensor and λ is the wavelength. The data collected by all sensors at time t can be expressed as

$$\boldsymbol{x}(t) = \sum_{k=1}^{K} \boldsymbol{a}(\theta_k)s_k(t) + \boldsymbol{\varepsilon}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{\varepsilon}(t), \tag{1}$$

for $t = 1, \ldots, T$, where $\boldsymbol{\varepsilon}(t) = [\varepsilon_1(t), \varepsilon_2(t), \ldots, \varepsilon_L(t)]^T \in \mathbb{C}^{L\times 1}$ is an i.i.d. white Gaussian noise $\mathcal{CN}(0,\sigma^2)$, $\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2)\ldots, \boldsymbol{a}(\theta_K)] \in \mathbb{C}^{L\times K}$, and $\boldsymbol{s}(t) = [s_1(t), s_2(t), \ldots, s_K(t)]^T$ represents the source signal vector with $s_k(t)$ distributed as $\mathcal{CN}(0, \sigma_k^2)$. We assume that the sources are temporally uncorrelated.

The correlation matrix of the data can be expressed as

$$\begin{aligned}\boldsymbol{R}_{xx} &= E[\boldsymbol{x}(t)\boldsymbol{x}^*(t)]\\ &= \boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{A}^* + \sigma^2\boldsymbol{I}\\ &= \sum_{k=1}^{K}\sigma_k^2\boldsymbol{a}(\theta_k)\boldsymbol{a}^*(\theta_k) + \sigma^2\boldsymbol{I},\end{aligned} \tag{2}$$

in which $\boldsymbol{R}_{ss}$ is a $K\times K$ diagonal matrix with diagonal elements $\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2$. After vectorizing the correlation matrix $\boldsymbol{R}_{xx}$, we have

$$\boldsymbol{z} = \mathrm{vec}(\boldsymbol{R_{xx}}) = \boldsymbol{\Phi}(\theta_1, \theta_2, \ldots, \theta_K)\boldsymbol{s} + \sigma^2\mathbf{1}_n, \tag{3}$$

where

$$\begin{aligned}\boldsymbol{\Phi}(\theta_1, \ldots, \theta_K) &= \boldsymbol{A}^* \odot \boldsymbol{A}\\ &= [\boldsymbol{a}(\theta_1)^{\mathrm{H}}\otimes\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_K)^{\mathrm{H}}\otimes\boldsymbol{a}(\theta_K)],\end{aligned} \tag{4}$$

$\boldsymbol{s} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2]^T$, and $\mathbf{1}_n = [\boldsymbol{e}_1^T, \boldsymbol{e}_2^T, \ldots, \boldsymbol{e}_L^T]^T$ with $\boldsymbol{e}_i$ denoting a vector with all zero elements, except for the ith element, which equals one.

Comparing (1) with (3), we see that $\boldsymbol{s}$ behaves like a coherent source and $\sigma^2\mathbf{1}_n$ becomes a deterministic noise term. The distinct rows in $\boldsymbol{\Phi}$ act as a larger virtual array with sensors located at $d_i - d_j$, with $1 \le i, j \le L$. Traditional DOA estimation algorithms can be implemented to detect more sources when the structure of the sensor array is properly designed. Following this idea, nested arrays [3] and co-prime arrays [4] were introduced, and then shown to increase the degrees of freedom from $O(N)$ to $O(N^2)$, and from $O(M+N)$ to $O(MN)$ respectively. In the following, we focus only on co-prime arrays; the results follow naturally for nested arrays.

Consider a co-prime array structure consisting of two arrays with N and $2M$ sensors respectively. The locations of the N sensors are in the set $\{Mnd, 0 \le n \le N-1\}$, and the locations of the $2M$ sensors are in the set $\{Nmd, 0 \le m \le 2M-1\}$ as illustrated in Fig. 1. The first sensors of these two arrays are collocated. The geometry of such a co-prime array is shown in Fig. 1. The locations of the virtual sensors in $\boldsymbol{\Phi}$ from (3) are given by the cross difference set $\{\pm(Mn - Nm)d, 0 \le n \le N-1, 0 \le m \le 2M-1\}$ and the two self difference sets. In order to implement spatial smoothing of MUSIC, or to use other popular DOA estimation techniques, we are interested in generating a consecutive range of virtual sensors. It was shown

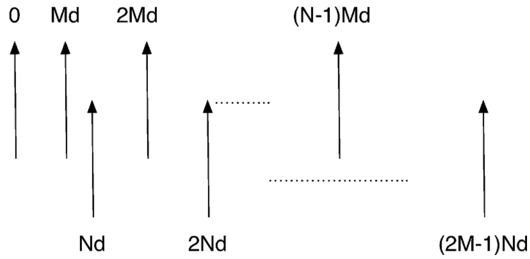

Fig. 1. Geometry of co-prime arrays.

in [5] that when M and N are co-prime numbers, a consecutive range can be created from $-MNd$ to MNd, with

$$\{-MNd, -(MN-1)d, \ldots, -2d, -d, d, 2d, \ldots, (MN-1)d, MNd\}$$

taken from the cross difference set and $\{0d\}$ taken from any one of the self difference sets.

By removing repeated rows of (3) and sorting the remaining rows from $-MNd$ to MNd, we have the linear model rearranged as

$$\tilde{\boldsymbol{z}} = \tilde{\boldsymbol{\Phi}}\boldsymbol{s} + \sigma^2\tilde{\boldsymbol{w}}. \tag{5}$$

It is easy to verify that $\tilde{\boldsymbol{w}} \in \mathbb{R}^{(2MN+1)\times 1}$ is a vector whose elements all equal zero, except the $(MN+1)$th element, which equals one. The matrix $\tilde{\boldsymbol{\Phi}} \in \mathbb{R}^{(2MN+1)\times K}$ is given by

$$\tilde{\boldsymbol{\Phi}} = \begin{bmatrix} e^{-\boldsymbol{j}MNd\frac{2\pi}{\lambda}\sin(\theta_1)} & \ldots & e^{-\boldsymbol{j}MNd\frac{2\pi}{\lambda}\sin(\theta_K)} \\ e^{-\boldsymbol{j}(MN-1)d\frac{2\pi}{\lambda}\sin(\theta_1)} & \ldots & e^{-\boldsymbol{j}(MN-1)d\frac{2\pi}{\lambda}\sin(\theta_K)} \\ \vdots & \ddots & \vdots \\ e^{\boldsymbol{j}MNd\frac{2\pi}{\lambda}\sin(\theta_1)} & \ldots & e^{\boldsymbol{j}MNd\frac{2\pi}{\lambda}\sin(\theta_K)} \end{bmatrix},$$

which is the steering matrix of a ULA with $2MN+1$ sensors. Therefore, (5) can be regarded as a ULA detecting a coherent source $\boldsymbol{s}$ with deterministic noise term $\sigma^2\tilde{\boldsymbol{w}}$. By applying MUSIC with spatial smoothing, the authors in [5] showed that $O(MN)$ sources can be detected, using this approach.

III. Direction of Arrival Estimation With Super Resolution Recovery

In this section we first assume that the signal model (3) is accurate, which means that the number of samples T is infinite, and also that the noise power σ^2 is known a priori. The super resolution theory developed in [14] can then be applied to co-prime arrays to demonstrate that we can detect up to $O(MN)$ sources robustly as long as the distance between any two sources is on the order of $\frac{1}{MN}$. We then consider the case in which the number of time samples T is limited and demonstrate the robustness of super resolution recovery via statistical analysis of the noise structure.

A. Mathematical Theory of Super Resolution

Super resolution seeks to recover high frequency details from the measurement of low frequency components. Mathematically, given a measure $s(\tau)$ with $\tau \in [0,1]$, the Fourier series coefficients are recorded as

$$r(n) = \int_0^1 e^{-\boldsymbol{j}2\pi n\tau}s(\tau)d\tau, \quad n = -f_c, -f_c+1, \ldots, f_c. \tag{6}$$

Using the operator $\boldsymbol{F}$ to denote the low frequency measuring operator which transforms a signal from its continuous time domain into its discrete frequency domain, we can represent (6) as $\boldsymbol{r} = \boldsymbol{F}s$, in which $\boldsymbol{r} = [r(-f_c), \ldots, r(f_c)]^T$ and $s = s(\tau), 0 \le \tau \le 1$.

Suppose that the measure $s(\tau)$ is sparse, i.e., $s(\tau)$ is a weighted sum of several spikes:

$$s(\tau) = \sum_{k=1}^{K} s_k\delta_{\tau_k}, \tag{7}$$

in which s_k can be complex valued and $\tau_k \in [0,1]$ for all k. Then

$$r(n) = \sum_{k=1}^{K} s_k e^{-j2\pi n\tau_k}, \quad n = -f_c, -f_c+1, \ldots, f_c. \tag{8}$$

In order to recover $s(\tau)$ from the measurements $r(n)$, total variation minimization is introduced. This criterion encourages the sparsity in the measure $s(\tau)$, just as ℓ_1 norm minimization produces sparse signals in the discrete space. In the rest of the paper, we will use s to denote the measure $s(\tau)$ for simplicity. The total variation for the complex measure s is defined as

$$\|s\|_{\mathrm{TV}} = \sup\sum_{j=1}^{\infty}|s(B_j)|, \tag{9}$$

the supremum being taken over all partitions of the set $[0,1]$ into countable collections of disjoint measurable sets B_j. When s has the form (7), $\|s\|_{\mathrm{TV}} = \sum_{k=1}^{K} s_k$, which resembles the discrete ℓ_1 norm.

The following convex optimization formula was proposed in [14] to solve the super resolution problem which recovers a sparse measure from $\boldsymbol{r}$:

$$\min_{\tilde{s}} \|\tilde{s}\|_{\mathrm{TV}} \quad \text{s.t.} \quad \boldsymbol{F}\tilde{s} = \boldsymbol{r}. \tag{10}$$

When the distance between any two τ_i and τ_j is larger than $2/f_c$, then the original sparse signal s is the unique solution to the above convex optimization [14]. The continuous optimization (10) can be solved via the dual problem [14]:

$$\begin{aligned} &\max_{\boldsymbol{u},\boldsymbol{Q}} \quad \mathrm{Re}[\boldsymbol{u}^*\boldsymbol{r}] \\ &\text{s.t.} \quad \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{u} \\ \boldsymbol{u}^* & 1 \end{bmatrix} \succeq 0, \\ &\sum_{i=1}^{2MN+1-j} \boldsymbol{Q}_{i,i+j} = \begin{cases} 1 & j = 0, \\ 0 & j = 1, 2, \ldots, 2MN, \end{cases} \end{aligned} \tag{11}$$

where $\boldsymbol{Q} \in \mathbb{C}^{(2MN+1)\times(2MN+1)}$ is a Hermitian matrix and $\boldsymbol{u} \in \mathbb{C}^{2MN+1}$ is the Lagrangian multiplier for the constraint $\boldsymbol{F}\tilde{s} = \boldsymbol{r}$. The primal solution s is obtained through a combined process of rooting finding and least-squares [14].

B. DOA Estimation With TV-Norm Minimization

DOA estimation with co-prime arrays can be related to (8) by a straightforward change of variables. Letting $\tau_k = \frac{d}{\lambda}(1 - \sin(\theta_k))$ for all k, the linear model of (5) can be transformed into

$$r(n) = e^{-\boldsymbol{j}2\pi n \frac{d}{\lambda}}(\tilde{z}_n - \sigma^2 \tilde{w}_n) = e^{-\boldsymbol{j}2\pi n \frac{d}{\lambda}} \sum_{k=1}^{K} s_k e^{\boldsymbol{j}2\pi n \frac{d}{\lambda}\sin(\theta_k)}$$
$$= \sum_{k=1}^{K} s_k e^{-\boldsymbol{j}2\pi n \tau_k} = \int_0^1 e^{-\boldsymbol{j}2\pi n \tau} s(\tau) d\tau, \qquad (12)$$

where $n = -MN, -MN+1, \ldots, MN-1, MN$, and s is a sparse measure given in (7) with $s_k = \sigma_k^2$. Note that the measure s is different from the vector representation $\boldsymbol{s} = [s_1, \ldots, s_K]^T$, and they are related by (7). The change of variables is performed to guarantee that $0 \leq \tau_k \leq 1$. We use $\mathcal{T} = \{\tau_k, 1 \leq k \leq K\}$ to denote the support set.

A theorem about the resolution and degrees of freedom for co-prime arrays can be directly derived using Theorem 1.2 in [14]. Before introducing the theorem, we first define the minimum distance between any two sources as

$$\Delta(\boldsymbol{\theta}) = \min_{\theta_i, \theta_j, \theta_i \neq \theta_j} |\sin(\theta_i) - \sin(\theta_j)|. \qquad (13)$$

Theorem III.1: Consider a co-prime array consisting of two linear arrays with N and $2M$ sensors respectively. The distances between two consecutive sensors are Md for the first array and Nd for the second array, where M and N are co-prime numbers, and $d \leq \frac{\lambda}{2}$. Suppose we have K sources located at $\theta_1, \ldots, \theta_K$. If the minimum distance follows the constraint that

$$\Delta(\boldsymbol{\theta}) \geq \frac{2\lambda}{MNd},$$

then by solving the convex optimization (10) *with the signal model $\boldsymbol{r} = \boldsymbol{F}s$, one can recover the locations θ_k for $k = 1, \ldots, K$ exactly. The number of sources that can be detected is*

$$K_{\max} = \frac{MNd}{\lambda}.$$

With a co-prime array using $2M + N$ sensors, the continuous sparse recovery method can detect up to $\frac{MNd}{\lambda}$ sources when $\Delta(\boldsymbol{\theta}) \geq \frac{2\lambda}{MNd}$. The minimum distance constraint is a sufficient condition. In real applications we can expect a more relaxed distance condition for the sources. We will confirm this point in the numerical results. With the utilization of co-prime arrays, the same number of sensors can detect $O(MN)$ sources as indicated by traditional MUSIC theory [5]. We will show in the numerical examples that implementing the super resolution framework provides more degrees of freedom and finer resolution ability than those of MUSIC. This is because the spatial smoothing in MUSIC reduces the obtained virtual array aperture. For the noiseless case, other methods, such as Prony's method [22] and matrix pencil [23] can be used for exact recovery of $O(MN)$ sources. However, they require prior information about the system order, which we do not require here. Furthermore, these methods are generally sensitive to noise in the model and therefore do not offer robustness guarantees.

C. Noisy Model for Sparse Recovery

In practice, the covariance matrix $\boldsymbol{R}_{xx}$ in (2) is typically unknown, and cannot be estimated exactly unless the number of samples T goes to infinity. Typically the covariance matrix is approximated by the sample covariance:

$$\hat{\boldsymbol{R}}_{xx} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}(t)\boldsymbol{x}^*(t). \qquad (14)$$

Subtracting the noise covariance matrix from both sides, we obtain

$$\hat{\boldsymbol{R}}_{xx} - \sigma^2 \boldsymbol{I} = \boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{A}^* + \boldsymbol{E}. \qquad (15)$$

Here $\boldsymbol{R}_{ss}$ is a diagonal matrix with k-th diagonal element

$$\hat{\sigma}_k^2 = \frac{1}{T}\sum_{t=1}^{T} s_k(t) s_k^*(t), \qquad (16)$$

and the (m,n)th element of $\boldsymbol{E}$ is given by (see (1))

$$E_{mn} = \frac{1}{T}\sum_{t=1}^{T}\sum_{i,j=1, i\neq j}^{K} A_{mi}A_{nj}^* s_i(t) s_j^*(t)$$
$$+ \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{K} A_{mi} s_i(t)\varepsilon_n^*(t) + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{K}\varepsilon_m(t) s_i^*(t) A_{ni}^*$$
$$+ \frac{1}{T}\sum_{t=1}^{T}\varepsilon_m(t)\varepsilon_n^*(t) - \sigma^2 \boldsymbol{I}_{mn}, \quad 1 \leq m, n \leq L. \qquad (17)$$

For simplicity of analysis, we assume that $\varepsilon_j \sim \mathcal{CN}(0, \sigma^2)$ and $s_i(t) \sim \mathcal{CN}(0, \sigma_s^2)$.

Similar to the operation in (3), vectorizing (15) leads to

$$\boldsymbol{z} = \mathrm{vec}(\hat{\boldsymbol{R}}_{\boldsymbol{xx}}) = \boldsymbol{\Phi}(\theta_1, \theta_2, \ldots, \theta_K)\boldsymbol{s} + \sigma^2 \boldsymbol{1}_n + \boldsymbol{\eta}, \qquad (18)$$

where $\boldsymbol{\eta} = \mathrm{vec}(\boldsymbol{E})$ and $\boldsymbol{s} = [\hat{\sigma}_1^2, \ldots, \hat{\sigma}_K^2]^T$. For co-prime arrays, by removing repeated rows in (18), and sorting them as consecutive lags from $-MNd$ to MNd, we get

$$\tilde{\boldsymbol{z}} = \tilde{\boldsymbol{\Phi}}\boldsymbol{s} + \sigma^2 \tilde{\boldsymbol{w}} + \tilde{\boldsymbol{e}}, \qquad (19)$$

in which $\tilde{\boldsymbol{\Phi}}$, and $\tilde{\boldsymbol{w}}$ are defined in (5). The vector $\tilde{\boldsymbol{e}}$ is obtained after rearranging $\boldsymbol{\eta}$, and only one element from $\tilde{\boldsymbol{e}}$ corresponds to the diagonal element from $\boldsymbol{E}$. Applying the transformation technique in (12), we have

$$\boldsymbol{r} = \boldsymbol{F}s + \boldsymbol{e}, \qquad (20)$$

where $e(n) = \tilde{e}(n) e^{-\boldsymbol{j}2\pi n \frac{d}{\lambda}}$, and s is the measure defined in (7) with $s_k = \hat{\sigma}_k^2$. Thus we can formulate the following continuous sparse recovery problem, which considers the noise:

$$\min_s \|s\|_{\mathrm{TV}} \quad \text{s.t.} \quad \|\boldsymbol{F}s - \boldsymbol{r}\|_2 \leq \epsilon. \qquad (21)$$

This optimization can be solved by first solving the dual problem [15]:

$$\max_{\boldsymbol{u}, \boldsymbol{Q}} \quad \mathrm{Re}[\boldsymbol{u}^*\boldsymbol{r}] - \epsilon\|\boldsymbol{u}\|_2$$
$$\text{s.t.} \quad \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{u} \\ \boldsymbol{u}^* & 1 \end{bmatrix} \succeq 0,$$
$$\sum_{i=1}^{2MN+1-j} \boldsymbol{Q}_{i,i+j} = \begin{cases} 1 & j = 0, \\ 0 & j = 1, 2, \ldots, 2MN. \end{cases} \qquad (22)$$

As before, the primal solution is obtained through a combined process of root finding and least-squares [15].

In order to analyze the robustness of the proposed approach for co-prime arrays, we introduce a lemma that shows that the probability of every element in $\boldsymbol{e}$ being larger than a constant is upper bounded. The proof can be found in the appendix.

Lemma III.1: Let E_{mn} be given in (17) *and assume that $\varepsilon_j \sim \mathcal{CN}(0, \sigma^2)$ and $s_i(t) \sim \mathcal{CN}(0, \sigma_s^2)$. Then for $m \neq n$, we have*

$$\Pr(|E_{mn}| \geq \epsilon) \leq 8\exp(-C_1(\epsilon)T) + 16\exp(-C_2(\epsilon)T) + 8\exp(-C_3(\epsilon)T).$$

When $m = n$, and $0 \leq \epsilon \leq 16\sigma^2$, we obtain

$$\Pr(|E_{mn}| \geq \epsilon) \leq 8\exp(-C_1(\epsilon)T) + 16\exp(-C_2(\epsilon)T) + 4\exp(-C_4(\epsilon)T).$$

Here $C_1(\epsilon), C_2(\epsilon), C_3(\epsilon)$ and $C_4(\epsilon)$ are increasing functions of ϵ.

The work of [16] provides an error bound on the support set estimation using (21) with noisy measurements. Combining the result from [16] with Lemma III.1, we have the following theorem.

Theorem III.2: Consider a co-prime array consisting of two linear arrays with N and $2M$ sensors respectively. The distances between two consecutive sensors are Md for the first array and Nd for the second array, where M and N are co-prime numbers, and $d \leq \frac{\lambda}{2}$. Assume T sample points are collected for each receiver. Suppose we have K sources located at $\theta_1, \ldots, \theta_K$. The minimum distance $\Delta(\boldsymbol{\theta}) \geq \frac{2\lambda}{MNd}$. Let $s(\tau) = \sum_{k=1}^{K} s_k \delta_{\tau_k}$ with $\tau_k = \frac{d}{\lambda}(1 - \sin(\theta_k))$ and $s_k = \hat{\sigma}_k^2$. Consider applying the transformation in (12) *and solving the optimization* (21) *with $\epsilon \leq 16\sqrt{2MN+1}\sigma^2$, and denote s_{opt} as the optimal solution, so that*

$$s_{\text{opt}} = \sum_{\tau_{\text{est}}[i] \in \mathcal{T}_{\text{est}}} s_{\text{est}}[i]\delta_{\tau_{\text{est}}[i]}. \tag{23}$$

Then, for every $\tau_k \in \mathcal{T}$

$$\left| s_k - \sum_{|\tau_{\text{est}}[i] - \tau_k| \leq \frac{c}{MN}} s_{\text{est}}[i] \right| \leq C_1\epsilon, \tag{24}$$

$$\sum_{|\tau_{\text{est}}[i] - \tau_k| \leq \frac{c}{MN}} |s_{\text{est}}[i]|(\tau_{\text{est}}[i] - \tau_k)^2 \leq C_2 \frac{\epsilon}{M^2N^2}, \tag{25}$$

and

$$\sum_{\tau_k \in \mathcal{T}} \sum_{|\tau_{\text{est}}[i] - \tau_k| > \frac{c}{MN}} |s_{\text{est}}[i]| \leq C_3\epsilon, \tag{26}$$

with probability at least $1 - \alpha e^{-\gamma(\epsilon)T}$, where $\gamma(\epsilon)$ is an increasing function of ϵ. Here C_1, C_2 and C_3 are positive constants, $c = 0.1649$, and $\mathcal{T} = \{\tau_k : 1 \leq k \leq K\}$ is the support set of the original measure s.

Proof: Since $d \leq \frac{\lambda}{2}$, we have that $\tau_k \in [0, 1]$ for all k after transformation (12). It was shown in [16] that in order to obtain (24)–(26) we only need to show that $\|\boldsymbol{e}\|_2 \leq \epsilon$ with a certain probability in (21). Thus the statistical behavior of $\boldsymbol{e}$ in (20) is analyzed first.

Note that

$$\begin{aligned} \Pr(\|\boldsymbol{e}\|_2 \leq \epsilon) &\geq \Pr\left(\cap_{n=-MN}^{MN} |e(n)| \leq \frac{\epsilon}{\sqrt{2MN+1}}\right) \\ &= 1 - \Pr\left(\cup_{n=-MN}^{MN} |e(n)| \geq \frac{\epsilon}{\sqrt{2MN+1}}\right) \\ &\geq 1 - \sum_{n=-MN}^{MN} \Pr\left(|e(n)| \geq \frac{\epsilon}{\sqrt{2MN+1}}\right), \end{aligned} \tag{27}$$

which leads to the inequality

$$\begin{aligned} \Pr(\|\boldsymbol{e}\|_2 \geq \epsilon) &\leq \sum_{n=-MN}^{MN} \Pr\left(|e(n)| \geq \frac{\epsilon}{\sqrt{2MN+1}}\right) \\ &= \sum_{n=-MN}^{MN} \Pr\left(|\tilde{e}(n)| \geq \frac{\epsilon}{\sqrt{2MN+1}}\right). \end{aligned} \tag{28}$$

The equality follows from the fact that $|e(n)| = |\tilde{e}(n)|$ according to (19) and (20). Recall that $2MN$ elements of $\tilde{\boldsymbol{e}}$ are taken from E_{mn} when $m \neq n$, and one element of $\tilde{\boldsymbol{e}}$ is taken from E_{mn} when $m = n$. Therefore, by applying the results of Lemma III.1, we can show that $\|\boldsymbol{F}\boldsymbol{s} - \boldsymbol{r}\|_2 = \|\boldsymbol{e}\|_2 \leq \epsilon$ with probability at least $1 - \alpha e^{-\gamma(\epsilon)T}$, and $\gamma(\epsilon)$ is a increasing function of ϵ. □

Equations (24) and (25) show that the estimated support set clusters tightly around the true support, while (26) indicates that the false peaks in the estimated set $\mathcal{T}_{\text{est}}$ have small amplitudes. A numerical method is proposed in the next section to further refine the estimation, using a discrete sparse recovery method after obtaining $\mathcal{T}_{\text{est}}$.

When both DOAs and signal powers are of interest, we combine the statistical analysis of the noise structure in co-prime arrays with the super resolution results in [15] to give a performance guarantee on the reconstruction of the sparse measure s. Since s is a sparse measure, there is no point in bounding $s - s_{\text{opt}}$ directly. Instead K_h, which is a low pass filter with cut-off frequency $f_h > MN$, is introduced. This kernel is referred to as the Fejér kernel, and is given by

$$\begin{aligned} K_h(t) &= \frac{1}{f_h} \sum_{k=-f_h}^{f_h} (f_h + 1 - |k|) e^{\boldsymbol{j}2\pi kt} \\ &= \frac{1}{f_h + 1} \left(\frac{\sin(\pi(f_h+1)t)}{\sin(\pi t)} \right). \end{aligned} \tag{29}$$

The cut-off frequency f_h can be much higher than MN. Thus using $K_h(t)$, we can show that by solving the convex optimization problem in (21) the high resolution details of the original measure $s(\tau) = \sum_{k=1}^{K} s_k \delta_{\tau_k}$ can be recovered with high probability, even though the sample size T is finite.

Theorem III.3: Let the co-prime arrays and the locations of the sources have the same setup as in Theorem III.2. The solution of the convex optimization (21) *satisfies*

$$\|K_h * (s_{\text{opt}} - s)\|_{L_1} \leq C_0 \frac{f_h^2}{M^2N^2}\epsilon, \tag{30}$$

with probability at least $1 - \alpha e^{-\gamma(\epsilon)T}$ when $\epsilon \leq 16\sqrt{2MN+1}\sigma^2$, where $\gamma(\epsilon)$ is a increasing function of ϵ. Here C_0 is a positive constant number.

The proof can be obtained by combining Lemma III.1 with the techniques in [15]. Theorem III.3 allows to choose the cut-off frequency f_h as large as one wants in order to bound the reconstruction error up to a certain resolution. However, this will entail an increase in the reconstruction error which is proportional to f_h^2. This theorem also shows that the reconstruction of s is stable in the presence of noise. The probability of successful reconstruction goes to one exponentially fast as the number of samples T goes to ∞. When fixing the probability of a stable reconstruction, by increasing the number of samples T we can allow for a decreased ϵ since $\gamma(\epsilon)$ is an increasing function. Therefore we can have a larger f_h without increasing the error bound in (30). By collecting more samples, one can stably reconstruct the measure s as if we had an even wider aperture f_h.

IV. DOA Estimation via Semidefinite Programming and Root Finding

We now derive an optimization framework to reconstruct s for co-prime arrays. Since the optimization is performed on a continuous domain, we will refer to the proposed algorithm as the continuous sparse recovery in the rest of this paper.

For DOA estimation the noise power σ^2 is often unknown. Therefore, the optimization must be modified to include this effect. A more realistic optimization is reformulated as

$$\min_{s,\sigma^2\geq 0} \|s\|_{\mathrm{TV}} \quad \text{s.t.} \quad \|\boldsymbol{r} - \boldsymbol{F}s - \sigma^2\boldsymbol{w}\|_2 \leq \epsilon, \tag{31}$$

in which $w_n = \tilde{w}_n e^{-\boldsymbol{j}2\pi n\frac{d}{\lambda}}$, and $\tilde{\boldsymbol{w}}$ is defined in (5). The dual problem takes on the form

$$\begin{aligned} \max_{\boldsymbol{u}\in\mathbb{C}^{2MN+1}} \quad & \mathrm{Re}[\boldsymbol{u}^*\boldsymbol{r}] - \epsilon\|\boldsymbol{u}\|_2 \\ & \text{s.t.} \quad \|\boldsymbol{F}^*\boldsymbol{u}\|_{L_\infty} \leq 1,\ \mathrm{Re}[\boldsymbol{u}^*\boldsymbol{w}] \leq 0. \end{aligned} \tag{32}$$

The derivation of (32) is given in the Appendix. Since $\boldsymbol{u} = \boldsymbol{0}$ is a feasible solution, strong duality holds according to the general Slater condition [24].

Due to the first constraint in (32), the problem itself is still an infinite dimensional optimization. It was shown in [14] that the first constraint can be recast as a semidefinite matrix constraint. Thus the infinite dimensional dual problem is equivalent to the following semidefinite program (SDP):

$$\begin{aligned} \max_{\boldsymbol{u},\boldsymbol{Q}} \quad & \mathrm{Re}[\boldsymbol{u}^*\boldsymbol{r}] - \epsilon\|\boldsymbol{u}\|_2 \\ \text{s.t.} \quad & \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{u} \\ \boldsymbol{u}^* & 1 \end{bmatrix} \succeq 0, \quad \mathrm{Re}[\boldsymbol{u}^*\boldsymbol{w}] \leq 0, \\ & \sum_{i=1}^{2MN+1-j} \boldsymbol{Q}_{i,i+j} = \begin{cases} 1 & j = 0, \\ 0 & j = 1, 2, \ldots, 2MN. \end{cases} \end{aligned} \tag{33}$$

Here $\boldsymbol{Q} \in \mathbb{C}^{(2MN+1)\times(2MN+1)}$ is a Hermitian matrix. This optimization problem can be easily solved, for example by using the CVX package [24], to yield the optimal dual solution.

The following lemma is introduced to link the solutions of the primal and dual problems.

Lemma IV.1: Let s_{opt} and $\boldsymbol{u}_{\mathrm{opt}} \in \mathbb{C}^{2MN+1}$ be the optimal solutions of the primal problem (31) *and dual problem* (33) *respectively. Then*

$$(\boldsymbol{F}^*\boldsymbol{u}_{\mathrm{opt}})(\tau) = \mathrm{sgn}(s_{\mathrm{opt}}(\tau)) \tag{34}$$

for all τ such that $s_{\mathrm{opt}}(\tau) \neq 0$. Here $\boldsymbol{F}^*$ is the adjoint operator of $\boldsymbol{F}$, and it transforms a vector into a continuous signal by taking the inverse Fourier transform.

Proof: Let σ_{opt}^2 be the noise power estimated in the primal problem. Since strong duality holds, we have

$$\begin{aligned} \|s_{\mathrm{opt}}\|_{\mathrm{TV}} &= \mathrm{Re}\langle \boldsymbol{r}, \boldsymbol{u}_{\mathrm{opt}}\rangle - \epsilon\|\boldsymbol{u}_{\mathrm{opt}}\|_2 \\ &= \mathrm{Re}\langle \boldsymbol{r} - \boldsymbol{F}s_{\mathrm{opt}} - \sigma_{\mathrm{opt}}^2\boldsymbol{w}, \boldsymbol{u}_{\mathrm{opt}}\rangle - \epsilon\|\boldsymbol{u}_{\mathrm{opt}}\|_2 \\ &\quad + \mathrm{Re}\langle \boldsymbol{F}s_{\mathrm{opt}} + \sigma_{\mathrm{opt}}^2\boldsymbol{w}, \boldsymbol{u}_{\mathrm{opt}}\rangle \\ &\leq \mathrm{Re}\langle \boldsymbol{F}s_{\mathrm{opt}} + \sigma_{\mathrm{opt}}^2\boldsymbol{w}, \boldsymbol{u}_{\mathrm{opt}}\rangle \leq \mathrm{Re}\langle \boldsymbol{F}s_{\mathrm{opt}}, \boldsymbol{u}_{\mathrm{opt}}\rangle. \end{aligned} \tag{35}$$

The first inequality follows from the Cauchy-Schwarz inequality and the fact that $\|\boldsymbol{r} - \boldsymbol{F}s_{\mathrm{opt}} - \sigma_{\mathrm{opt}}^2\boldsymbol{w}\|_2 \leq \epsilon$. The second inequality results from $\mathrm{Re}[\boldsymbol{u}_{\mathrm{opt}}^*\boldsymbol{w}] \leq 0$. In addition, we also have

$$\mathrm{Re}\langle \boldsymbol{F}s_{\mathrm{opt}}, \boldsymbol{u}_{\mathrm{opt}}\rangle \leq \|\boldsymbol{F}^*\boldsymbol{u}_{\mathrm{opt}}\|_{L_\infty}\|s_{\mathrm{opt}}\|_{\mathrm{TV}} \leq \|s_{\mathrm{opt}}\|_{\mathrm{TV}}, \tag{36}$$

where we used the fact that $\|\boldsymbol{F}^*\boldsymbol{u}_{\mathrm{opt}}\|_{L_\infty} \leq 1$. Combining (35) and (36) leads to $\|s_{\mathrm{opt}}\|_{\mathrm{TV}} = \mathrm{Re}\langle s_{\mathrm{opt}}, \boldsymbol{F}^*\boldsymbol{u}_{\mathrm{opt}}\rangle$, which implies (34). ☐

According to Lemma IV.1, the supports of $s_{\mathrm{opt}}(\tau)$ satisfy (34), and thus can be retrieved by root-finding based on the trigonometric polynomial $1-|\boldsymbol{F}^*\boldsymbol{u}_{\mathrm{opt}}(\tau)|^2 = 0$. Let $\mathcal{T}_{\mathrm{est}}$ denote the recovered set of roots of this polynomial with cardinality K_{est}, and let $\tau_{\mathrm{est}}[i]$ denote elements in $\mathcal{T}_{\mathrm{est}}$ with $1 \leq i \leq K_{\mathrm{est}}$. A matrix $\boldsymbol{F}_{\mathrm{est}} \in \mathbb{C}^{(2MN+1)\times K_{\mathrm{set}}}$ can then be formulated, with measurement $\boldsymbol{r}$ expressed as

$$\boldsymbol{r} = \boldsymbol{F}_{\mathrm{est}}\boldsymbol{s}_0 + \sigma^2\boldsymbol{w} + \boldsymbol{e}, \tag{37}$$

in which $\boldsymbol{s}_0 \in \mathbb{R}^{K_{\mathrm{est}}}$ and

$$\boldsymbol{F}_{\mathrm{est}} = \begin{bmatrix} e^{-\boldsymbol{j}MNd2\pi\tau_{\mathrm{est}}[1]} & \cdots & e^{-\boldsymbol{j}MNd2\pi\tau_{\mathrm{est}}[K_{\mathrm{est}}]} \\ e^{-\boldsymbol{j}(MN-1)d2\pi\tau_{\mathrm{est}}[1]} & \cdots & e^{-\boldsymbol{j}(MN-1)d2\pi\tau_{\mathrm{est}}[K_{\mathrm{est}}]} \\ \vdots & \ddots & \vdots \\ e^{\boldsymbol{j}MNd2\pi\tau_{\mathrm{est}}[1]} & \cdots & e^{\boldsymbol{j}MNd2\pi\tau_{\mathrm{est}}[K_{\mathrm{est}}]} \end{bmatrix}.$$

Due to numerical issues in the root finding process, the cardinality of $\mathcal{T}_{\mathrm{est}}$ is normally larger than the cardinality of $\mathcal{T}$, i.e., $K_{\mathrm{est}} \geq K$. It is possible in some cases that $K_{\mathrm{est}} \geq 2MN+1$, which would lead to an ill-conditioned linear system (37). Sparsity can then be exploited on $\boldsymbol{s}_0$. A convex optimization in the discrete domain can be formulated as

$$\min_{\boldsymbol{s}_0,\sigma^2\geq 0} \|\boldsymbol{s}_0\|_1 \quad \text{s.t.} \quad \|\boldsymbol{r} - \boldsymbol{F}_{\mathrm{est}}\boldsymbol{s}_0 - \sigma^2\boldsymbol{w}\|_2 \leq \epsilon_d. \tag{38}$$

We choose ϵ_d in (38) to be larger than ϵ in (31) since the noise level is expected to be higher in (37) due to inevitable error introduced in the root finding process. Assuming that the solution of (38) is $\boldsymbol{s}_{\text{est}} \in \mathbb{R}^{K_{\text{est}}}$, the estimation of the measure $s(\tau)$ in the continuous domain can be represented as

$$s_{\text{opt}}(\tau) = \sum_{i=1}^{K_{\text{est}}} s_{\text{est}}[i]\delta_{\tau_{\text{est}}[i]}. \tag{39}$$

V. Extension: Source Number Detection

Conventional source number detection for array processing is typically performed by exploiting eigenvalues from the sample covariance matrix. For co-prime arrays, this covariance matrix can be obtained by performing spatial smoothing on $\tilde{\boldsymbol{z}}$ in (5). The same idea can also be implemented on the sparse signal $\boldsymbol{s}_{\text{est}}$ recovered from the previous section. Ideally, after sorting its elements in a descending order, the signal $\boldsymbol{s}_{\text{est}}$ reconstructed from (38) should follow

$$\begin{aligned} s^2_{\text{est}}[1] \geq s^2_{\text{est}}[2] \geq \cdots s^2_{\text{est}}[K] \\ \geq s^2_{\text{est}}[K+1] = \cdots = s^2_{\text{est}}[K_{\text{est}}] = 0. \end{aligned} \tag{40}$$

The SORTE algorithm can be applied to this series. The difference of the elements from $\boldsymbol{s}_{\text{est}}$ is

$$\nabla s_{\text{est}}[i] = s^2_{\text{est}}[i] - s^2_{\text{est}}[i+1],\ i = 1, \ldots, K_{\text{est}} - 1. \tag{41}$$

The gap measure in SORTE is given as

$$\text{SORTE}(i) = \begin{cases} \frac{\text{var}[i+1]}{\text{var}[i]} & \text{var}[i] \neq 0, \\ +\infty & \text{var}[i] = 0, \end{cases} \quad i = 1, \ldots, K_{\text{est}}-2, \tag{42}$$

where

$$\text{var}[i] = \frac{1}{K_{\text{est}} - i} \sum_{m=i}^{K_{\text{est}}-1} \left(\nabla s_{\text{est}}[m] - \frac{1}{K_{\text{est}} - i} \sum_{n=i}^{K_{\text{est}}-1} \nabla s_{\text{est}}[n] \right)^2. \tag{43}$$

The number of sources can be estimated as

$$\hat{K} = \text{argmin}_i \quad \text{SORTE}(i). \tag{44}$$

This approach requires $K_{\text{est}} > 2$ due to the definition of $\text{SORTE}(i)$ in (42). When $K_{\text{est}} \leq 2$, since $\mathcal{T}_{\text{est}}$ is obtained from the rooting finding process based on the continuous sparse recovery, we simply let $\hat{K} = K_{\text{est}}$. We will refer to this continuous sparse recovery based SORTE as CSORTE.

VI. Numerical Results

In this section, we present several numerical examples to show the merits of implementing our continuous sparse recovery techniques to co-prime arrays.

We consider a co-prime array with 10 sensors. One set of sensors is located at positions $[0, 3, 6, 9, 12]d$, and the second set is located at $[0, 5, 10, 15, 20, 25]d$, where d is taken as half of the wavelength. The first sensors from both sets are collocated. It is easy to show that the correlation matrix generates a virtual array with lags from $-17d$ to $17d$. We compare continuous sparse recovery (CSR) techniques with MUSIC and also with discrete sparse recovery method (DSR) considering grid mismatches [13]. In [13], a LASSO formulation is used to perform the DOA estimation. Here we implement an equivalent form of LASSO, i.e., Basis Pursuit, to perform the comparison. The MUSIC method in this simulation follows the spatial smoothing technique in [5]. For the discrete sparse recovery method, we take the grid from -1 to 1, with step size 0.005 for $\sin(\theta)$. The noise levels ϵ in the optimization formulas are chosen by cross validation. We consider 15 narrow band signals located at

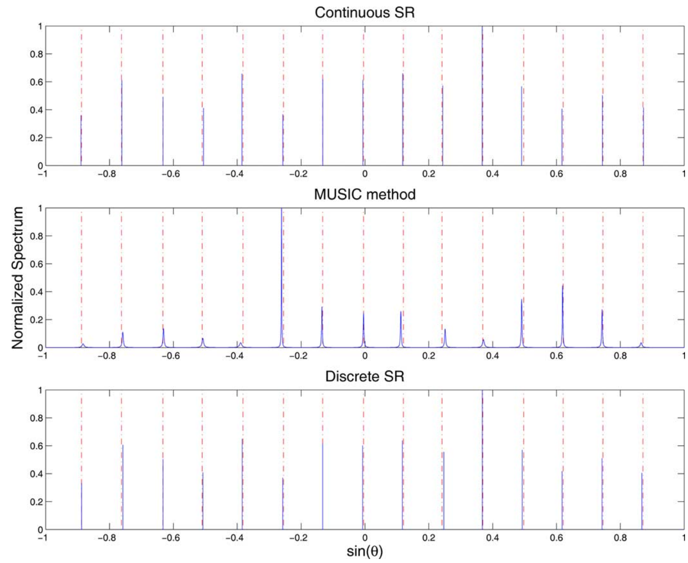

Fig. 2. Normalized spectra for CSR, MUSIC, and DSR, with $T = 500$ and $\text{SNR} = -10$ dB.

$$\begin{aligned} \sin(\boldsymbol{\theta}) = [&-0.8876, -0.7624, -0.6326, -0.5096, \\ &-0.3818, -0.2552, -0.1324, -0.0046, 0.1206, \\ &0.2414, 0.3692, 0.4972, 0.6208, 0.7454, 0.8704]. \end{aligned}$$

We show that continuous sparse recovery yields better results in terms of detection ability, resolution, and estimation accuracy.

A. Degrees of Freedom

In this first numerical example, we verify that the proposed continuous sparse recovery increases the degrees of freedom to $O(MN)$ by implementing the co-prime arrays' structure. The number of time samples is 500 and the SNR is chosen to be -10 dB. The ϵ for CSR is taken as 5, and ϵ_d is taken as 10 while DSR uses $\epsilon = 10$. In Fig. 2, we use a dashed line to represent the true directions of arrival. The CPU time for running CSR was 7.30 seconds. DSR took 7.82 seconds, whereas MUSIC algorithm took only 0.81 seconds. In MUSIC, we implement a root MUSIC algorithm to estimate the location of each source, where the number of sources is assumed to be given. The average estimation errors for CSR, DSR, and root MUSIC are 0.23%, 0.26%, and 0.42% respectively. We can see that all three methods achieve $O(MN)$. In the following subsection, we test the estimation accuracy of these three methods via Monte Carlo simulations.

B. Estimation Accuracy

In this section, we compare CSR, DSR and MUSIC via Monte Carlo simulations. Since traditional MUSIC does not yield the DOA of each source directly, we consider the Root MUSIC algorithm instead. For simplicity, we will still refer to it as MUSIC in this section. The number of sources is assumed to be known for the MUSIC algorithm in this simulation, whereas sparse

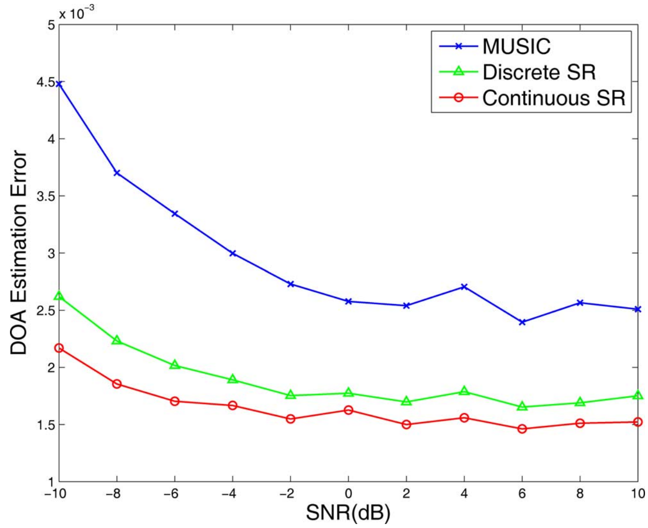

Fig. 3. DOA estimation errors for CSR, MUSIC, and DSR, with $T = 500$.

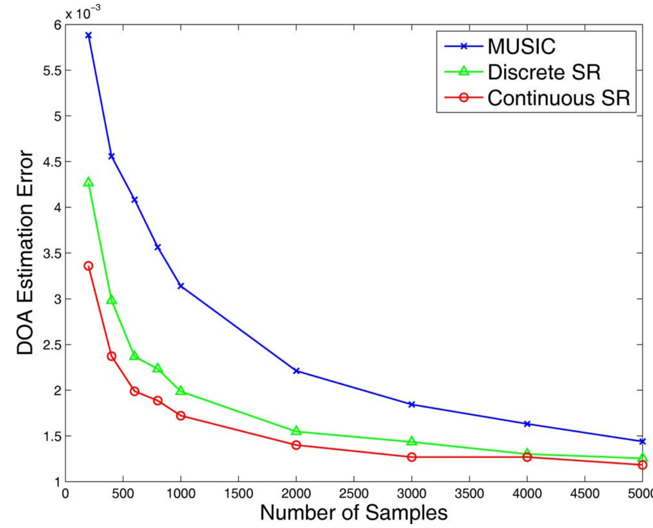

Fig. 4. DOA estimation error for CSR, MUSIC, and DSR, with $\mathrm{SNR} = -10$ dB.

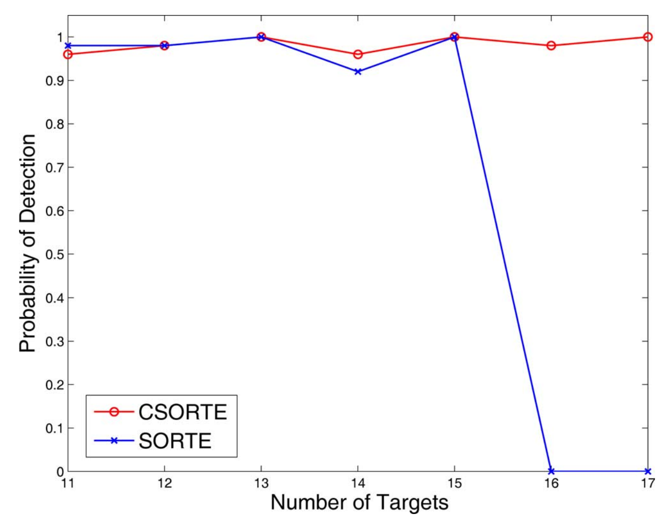

Fig. 5. Source number detection using CSORTE and SORTE, with $\mathrm{SNR} = 0$ dB, $T = 3000$.

methods do not assume this a priori. The values of ϵ and ϵ_d are chosen to be 5 and 10, while discrete SR uses $\epsilon = 10$.

Fig. 3 shows the DOA estimation error as a function of SNR after 50 Monte Carlo simulations. The estimation error is calculated based on the sine function of the DOAs. The average CPU times for running CSR, DSR and MUSIC are 6.93 s, 9.30 s, and 1.46 s respectively. We can see that CSR performs better than DSR uniformly with less computing time. Both sparse recovery methods achieve better DOA estimation accuracy than MUSIC. The accuracy of DSR can be further improved by taking a finer grid with a smaller step-size. However, this will slow down DSR further.

In Fig. 4 we show that with a varying number of snapshots the proposed CSR also exhibits better estimation accuracy than either DSR or MUSIC. The average CPU times for running CSR, DSR and MUSIC are 6.50 s, 7.91 s, and 1.43 s respectively. The performance of MUSIC and DSR approaches the performance of CSR when the number of snapshots is close to 5000. We can see that implementing CSR can save sampling time by taking a small number of snapshots to achieve the same estimation accuracy as the MUSIC algorithm. The parameters ϵ and ϵ_d are equal to 5 and 10 in this simulation.

C. Source Number Detection Performance Comparison

We now compare the source number detection performance of the proposed CSORTE with that of traditional SORTE applied to the covariance matrix after spatial smoothing. The SNR is set to 0 dB while the number of snapshots is 3000. We vary the number of sources from 11 to 17. Since this co-prime array structure yields consecutive lags from $-17d$ to $17d$, 17 is the maximum number of sources that can be detected theoretically via techniques based on the covariance matrix.

Fig. 7 shows the probability of detection with respect to the number of sources after 50 Monte Carlo simulations. In CSR, ϵ is chosen to be 5σ, and ϵ_d is set to be 2ϵ. When the number of sources is less than 15, CSORTE and SORTE yield comparable result. However, SORTE fails after the number of sources is larger than 15, while CSORTE provides stable performance and also exhibits perfect detection even when the number of sources reaches the theoretical limit of 17. DSR can also be combined with SORTE to perform source number detection. However, the detection accuracy is jeopardized by the spurious signal from the reconstructed signals using DSR. Therefore SORTE based on DSR is not included here.

D. Resolution Ability

Finally we compare the resolution abilities of CSR and MUSIC, and show that CSR is capable of resolving very closely located signals. In the first simulation, two sources are closely located at -32° and -30°. The value of ϵ is chosen to be 0.7σ and ϵ_d is set to be 2ϵ in CSR, where σ is the noise power.

Fig. 6 shows a numerical example in which the SNR is 0 dB and the number of snapshots is 500. Normalized spectra are plotted for three methods. MUSIC method A is the MUSIC algorithm with the assumption that the number of sources is known while the MUSIC method B is MUSIC relying on traditional SORTE to provide the estimated number of sources. We can see that MUSIC method B fails to resolve these two

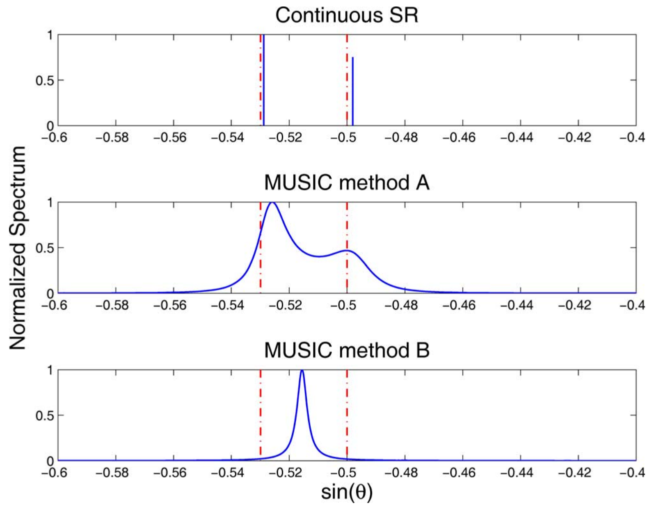

Fig. 6. Source number detection using CSR and the MUSIC algorithm, with $\mathrm{SNR} = 0\ \mathrm{dB}, T = 500$.

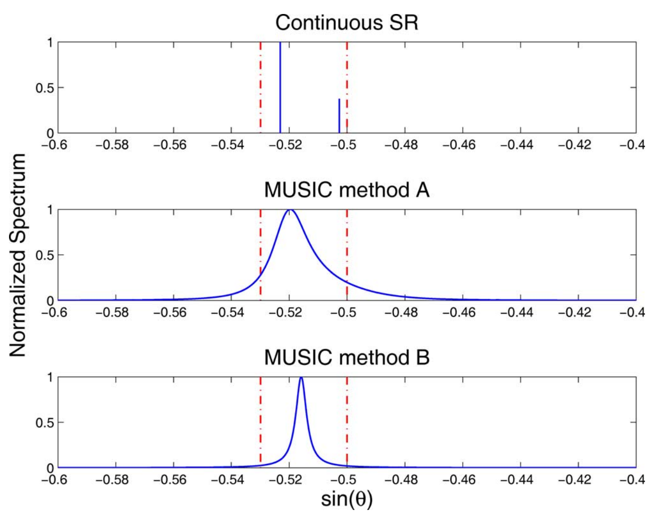

Fig. 7. Source number detection using CSR and MUSIC algorithm, with $\mathrm{SNR} = -5\ \mathrm{dB}, T = 500$.

targets because traditional SORTE fails to estimate the number of sources correctly. CSR resolves the two sources successfully even though a priori information about the number of sources is not assumed to be given. In Fig. 7, we lower the SNR to $-5\ \mathrm{dB}$, and we notice that even given the number of sources, MUSIC fails to resolve the two closely located sources while CSR resolves them successfully.

Note, that while a separation of $\frac{2\lambda}{MNd}$ is sufficient for Theorem III.1 to hold, in real applications, we expect to observe a better result. Thus we intentionally chose two sources which are more closely located, to show that the proposed method still works even with a stronger constraint.

Finally we conduct a simulation based on Monte Carlo runs to compare the resolution ability of CSORTE and the traditional SORTE algorithm. Fig. 8 shows the resolution performance in detecting two sources located at -32° and -30°, using CSORTE and SORTE after 50 Monte Carlo runs. The parameter ϵ is chosen to be 0.7σ, and ϵ_d is set to be 2ϵ in CSR. We can see that CSORTE outperforms traditional SORTE when detecting the two closely located sources.

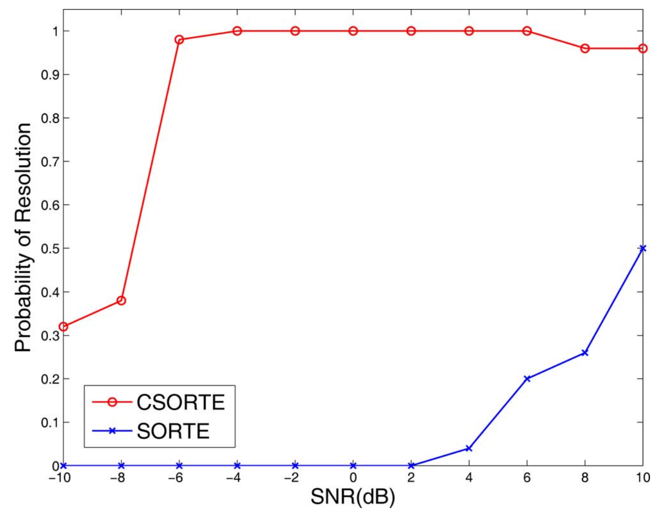

Fig. 8. Comparison of resolution performance of CSORTE and SORTE, with $T = 2000$.

VII. Conclusions and Future Work

In this work, we extended the mathematical theory of super resolution to DOA estimation using co-prime arrays. A primal-dual approach was utilized to transform the original infinite dimensional optimization problem to a solvable semidefinite program. After estimating the candidate support sets by root finding, a small scale sparse recovery problem is solved. The robustness of the proposed super resolution approach was verified by performing statistical analysis of the noise inherent in co-prime array processing. A source number detection algorithm was then proposed by combining the existing SORTE algorithm with the reconstructed spectrum from convex optimization. Via numerical examples, we showed that the proposed method achieves a more accurate DOA estimation while providing more degrees of freedom, and also exhibits improved resolution ability over traditional MUSIC with spatial smoothing.

Although implementing the continuous sparse recovery method saves sampling time in obtaining a certain estimation accuracy compared with MUSIC, one shortcoming of this approach is that solving the semidefinite program is more time consuming than MUSIC. Fast algorithm development could be an interesting topic for future work. It is also of interest to develop a systematic way to choose ϵ and ϵ_d in the optimization formulas. One major assumption made in current research on co-prime arrays research is that the sources are uncorrelated. Incorporating correlations among sources is another important topic for future work.

Appendix

To derive the statistical behavior of each element in $\boldsymbol{E}$ in Lemma III.1 we rely on two lemmas regarding the concentration behavior of complex Gaussian random variables. Their proofs are based on results from [25].

Lemma A.1: Let $x(t)$ *and* $y(t), t = 1, \ldots, T$ *be sequences of i.i.d., circularly-symmetric complex normal variables with zero* *mean and variances equal to σ_x^2 and σ_y^2 respectively. That is $x(t) \sim \mathcal{CN}(0,\sigma_x^2)$ and $y(t) \sim \mathcal{CN}(0,\sigma_y^2)$. Then*

$$\Pr\left(\left|\sum_{t=1}^{T} x(t)y^*(t)\right| \geq \epsilon\right) \leq 8\exp\left(-\frac{\epsilon^2}{16\sigma_x\sigma_y(T\sigma_x\sigma_y+\frac{\epsilon}{4})}\right).$$

Proof: First we have

$$\begin{aligned}\sum_{t=1}^{T} x(t)y^*(t) &= \sum_{t=1}^{T}\mathrm{Re}[x(t)]\mathrm{Re}[y(t)] + \sum_{t=1}^{T}\mathrm{Im}[x(t)]\mathrm{Im}[y(t)]\\ &\quad -\boldsymbol{j}\sum_{t=1}^{T}\mathrm{Re}[x(t)]\mathrm{Im}[y(t)] + \boldsymbol{j}\sum_{t=1}^{T}\mathrm{Im}[x(t)]\mathrm{Re}[y(t)].\end{aligned}$$

Following the same procedure used in the proof of Lemma III.1, we have

$$\begin{aligned}&\Pr\left(\left|\sum_{t=1}^{T} x(t)y^*(t)\right| \geq \epsilon\right)\\ &\leq \Pr\left(\left|\sum_{t=1}^{T}\mathrm{Re}[x(t)]\mathrm{Re}[y(t)]\right| \geq \frac{\epsilon}{4}\right)\\ &\quad + \Pr\left(\left|\sum_{t=1}^{T}\mathrm{Im}[x(t)]\mathrm{Im}[y(t)]\right| \geq \frac{\epsilon}{4}\right)\\ &\quad + \Pr\left(\left|\sum_{t=1}^{T}\mathrm{Re}[x(t)]\mathrm{Im}[y(t)]\right| \geq \frac{\epsilon}{4}\right)\\ &\quad + \Pr\left(\left|\sum_{t=1}^{T}\mathrm{Im}[x(t)]\mathrm{Re}[y(t)]\right| \geq \frac{\epsilon}{4}\right).\end{aligned}$$

Applying Lemma 6 from [25] concludes the proof. □

Before introducing the next lemma, we need to show that the square sums of i.i.d Gaussian random variables concentrate around the sum of their variances. The results below rely on Lemma 7 from [25].

Lemma A.2: Let $x(t), t = 1,\ldots,T$ be a sequence of i.i.d. Gaussian random variables with zero mean and variance equal to σ^2, i.e., $x(t) \sim \mathcal{N}(0,\sigma^2)$. Then

$$\Pr\left(\left|\sum_{t=1}^{T} x^2(t) - T\sigma^2\right| \geq \epsilon\right) \leq 2\exp\left(-\frac{\epsilon^2}{16\sigma^4 T}\right)$$

for $0 \leq \epsilon \leq 4\sigma^2 T$.

Proof: From the results in [25], for any positive c, we have the asymmetric bounds

$$\begin{aligned}\Pr\left(\sum_{t=1}^{T} x^2(t) - T\sigma^2 \geq 2\sigma^2\sqrt{Tc} + 2\sigma^2 c\right) &\leq \exp(-c),\\ \Pr\left(\sum_{t=1}^{T} x^2(t) - T\sigma^2 \leq -2\sigma^2\sqrt{Tc}\right) &\leq \exp(-c).\end{aligned}$$

When $0 \leq c \leq T$, we obtain

$$\begin{aligned}\Pr\left(\sum_{t=1}^{T} x^2(t) - T\sigma^2 \geq 4\sigma^2\sqrt{Tc}\right) &\leq \exp(-c),\\ \Pr\left(\sum_{t=1}^{T} x^2(t) - T\sigma^2 \leq -4\sigma^2\sqrt{Tc}\right) &\leq \exp(-c).\end{aligned}$$

Combing the above two inequalities leads to

$$\Pr\left(\left|\sum_{t=1}^{T} x^2(t) - T\sigma^2\right| \geq 4\sigma^2\sqrt{Tc}\right) \leq 2\exp(-c),$$

which yields the result by replacing $4\sigma^2\sqrt{Tc}$ with ϵ while maintaining $0 \leq c \leq T$. □

Lemma A.3: Let $x(t) \sim \mathcal{CN}(0,\sigma_x^2), t = 1,\ldots,T$ be a sequence of i.i.d., circularly-symmetric complex normal random variable. When $0 \leq \epsilon \leq 4\sigma_x^2 T$, we have

$$\Pr\left(\left|\sum_{t=1}^{T} |x(t)|^2 - T\sigma_x^2\right| \geq \epsilon\right) \leq 4\exp\left(-\frac{\epsilon^2}{16T\sigma_x^4}\right).$$

Proof: We begin by noting that

$$\sum_{t=1}^{T} |x(t)|^2 - T\sigma_x^2 = \sum_{t=1}^{T}[\mathrm{Re}\, x(t)]^2 + \sum_{t=1}^{T}[\mathrm{Im}\, x(t)]^2 - T\sigma_x^2.$$

Therefore

$$\begin{aligned}&\Pr\left(\left|\sum_{t=1}^{T} |x(t)|^2 - T\sigma_x^2\right| \geq \epsilon\right)\\ &\leq \Pr\left(\left|\sum_{t=1}^{T}[\mathrm{Re}\, x(t)]^2 - \frac{T\sigma_x^2}{2}\right| \geq \frac{\epsilon}{2}\right)\\ &\quad + \Pr\left(\left|\sum_{t=1}^{T}[\mathrm{Im}\, x(t)]^2 - \frac{T\sigma_x^2}{2}\right| \geq \frac{\epsilon}{2}\right).\end{aligned}$$

Applying Lemma A.2, we establish the result. □

Proof of Lemma III.1: We use T_1, T_2, and T_3 to denote the first three terms in (17). The last two terms are denoted by T_4. Then

$$\begin{aligned}\Pr(|E_{mn}| \leq \epsilon) &\geq \Pr\left(\cap_{i=1}^4 |T_i| \leq \frac{\epsilon}{4}\right) = s1 - \Pr\left(\cup_{i=1}^4 |T_i| \geq \frac{\epsilon}{4}\right)\\ &\geq 1 - \sum_{i=1}^{4}\Pr\left(|T_i| \geq \frac{\epsilon}{4}\right),\end{aligned}$$

which leads to the inequality

$$\Pr(|E_{mn}| \geq \epsilon) \leq \sum_{i=1}^{4}\Pr\left(|T_i| \geq \frac{\epsilon}{4}\right). \tag{45}$$

We also have

$$|T_1| = \frac{1}{T}\sum_{t=1}^{T}\sum_{i,j=1,i\neq j}^{K} A_{mi}A_{nj}^* s_i(t)s_j^*(t)$$
$$\leq \frac{1}{T}\sum_{i,j=1,i\neq j}^{K}|A_{mi}A_{ni}^*|\left|\sum_{t=1}^{T}s_i(t)s_j^*(t)\right|$$
$$\leq \frac{1}{T}\sum_{i,j=1,i\neq j}^{K}\left|\sum_{t=1}^{T}s_i(t)s_j^*(t)\right|. \quad (46)$$

The last inequality follows from the fact that $|A_{mn}| \leq 1$ for all m, n. Thus

$$\Pr\left(|T_1| \geq \frac{\epsilon}{4}\right) \leq \Pr\left(\sum_{i,j=1,i\neq j}^{K}\left|\sum_{t=1}^{T}s_i(t)s_j^*(t)\right| \geq \frac{\epsilon T}{4}\right).$$

Clearly

$$\Pr\left(|T_1| \geq \frac{\epsilon}{4}\right) \leq \Pr\left(\left|\sum_{t=1}^{T}s_{i_0}(t)s_{j_0}^*(t)\right| \geq \frac{\epsilon T}{4K(K-1)}\right),$$

for some i_0, j_0 with $i_0 \neq j_0$. Using Lemma A.1

$$\Pr\left(|T_1| \geq \frac{\epsilon}{4}\right) \leq 8\exp(-C_1(\epsilon)T), \quad (47)$$

with $C_1(\epsilon) = \frac{\epsilon^2}{16\sigma_s^2 K(K-1)(16\sigma_s^2 K(K-1)+\epsilon)}$.

For the second term T_2, we have

$$|T_2| = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{K}A_{mi}s_i(t)\varepsilon_n^*(t)$$
$$\leq \frac{1}{T}\sum_{i=1}^{K}|A_{mi}|\left|\sum_{t=1}^{T}s_i(t)\varepsilon_n(t)^*\right|$$
$$\leq \frac{1}{T}\sum_{i=1}^{K}\left|\sum_{t=1}^{T}s_i(t)\varepsilon_n(t)^*\right|. \quad (48)$$

Following similar arguments as for T_1, we obtain that

$$\Pr\left(|T_2| \geq \frac{\epsilon}{4}\right) \leq \Pr\left(\left|\sum_{t=1}^{T}s_{i_0}(t)\varepsilon_n^*(t)\right| \geq \frac{\epsilon T}{4K}\right).$$

Applying Lemma A.1, we have

$$\Pr\left(|T_2| \geq \frac{\epsilon}{4}\right) \leq 8\exp(-C_2(\epsilon)T), \quad (49)$$

with $C_2(\epsilon) = \frac{\epsilon^2}{16\sigma_s\sigma K(16\sigma_s\sigma K+\epsilon)}$.

For the third term, we have the same results as the second one, given as

$$\Pr\left(|T_3| \geq \frac{\epsilon}{4}\right) \leq 8\exp(-C_2(\epsilon)T). \quad (50)$$

When $m \neq n$, the last term $T_4 = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_m(t)\varepsilon_n^*(t)$, and by Lemma A.1,

$$\Pr\left(|T_4| \geq \frac{\epsilon}{4}\right) \leq 8\exp(-C_3(\epsilon)T), \quad (51)$$

with $C_3(\epsilon) = \frac{\epsilon^2}{16\sigma^2(16\sigma^2+\epsilon)}$. When $m = n$, the last term is given as $T_4 = \frac{1}{T}\sum_{t=1}^{T}|\varepsilon_m(t)|^2 - \sigma^2$, thus the probability is bounded by

$$\Pr\left(|T_4| \geq \frac{\epsilon}{4}\right) \leq 4\exp(-C_4(\epsilon)T), \quad (52)$$

where $C_4(\epsilon) = \frac{\epsilon^2}{256\sigma_\varepsilon^2}$ and $\epsilon \leq 16\sigma^2$ according to Lemma A.3. Applying the results from (47), (49), (50), (51) and (52) to inequality (45), we obtain the desired result. □

A. Derivation of the Dual Problem in Section IV

By introducing the variable $\boldsymbol{z} \in \mathbb{C}^{2MN+1}$, the original primal problem is equivalent to the following optimization:

$$\min_{\boldsymbol{s},\sigma^2\geq 0,\boldsymbol{z}} \quad \|\boldsymbol{s}\|_{\mathrm{TV}}$$
$$\text{s.t.} \quad \|\boldsymbol{z}\|_2 \leq \epsilon, \quad \boldsymbol{z} = \boldsymbol{r} - \boldsymbol{F}\boldsymbol{s} - \sigma^2\boldsymbol{w}.$$

With the Lagrangian multiplier $v \geq 0$ and $\boldsymbol{u} \in \mathbb{C}^{2MN+1}$, the Lagrangian function is given as

$$L(\boldsymbol{s},\boldsymbol{z},\sigma^2,\boldsymbol{u},v) = \|\boldsymbol{s}\|_{\mathrm{TV}} + v(\|\boldsymbol{z}\|_2 - \epsilon) + \mathrm{Re}[\boldsymbol{u}^*(\boldsymbol{r} - \boldsymbol{F}\boldsymbol{s} - \sigma^2\boldsymbol{w} - \boldsymbol{z})].$$

The dual function is given as

$$g(\boldsymbol{u},v) = \mathrm{Re}[\boldsymbol{u}^*\boldsymbol{r}] - v\epsilon + \inf_{\boldsymbol{s},\boldsymbol{z},\sigma^2\geq 0}\{\|\boldsymbol{s}\|_{\mathrm{TV}} - \mathrm{Re}[\boldsymbol{u}^*\boldsymbol{F}\boldsymbol{s}] - \sigma^2\mathrm{Re}[\boldsymbol{u}^*\boldsymbol{w}] + v\|\boldsymbol{z}\|_2 - \boldsymbol{u}^*\boldsymbol{z}\}.$$

The Lagrangian multipliers $\boldsymbol{u}$ and v in the domain of the dual function have to satisfy the following three constraints:

$$\|\boldsymbol{F}^*\boldsymbol{u}\|_{L_\infty} \leq 1, \mathrm{Re}[\boldsymbol{u}^*\boldsymbol{w}] \leq 0, v\frac{\boldsymbol{z}}{\|\boldsymbol{z}\|_2} = \boldsymbol{u}.$$

From the third constraint, we have $v = \|\boldsymbol{u}\|_2$, resulting in the dual problem stated in (33).

REFERENCES

[1] H. L. V. Trees, *Optimum Array Processing: Part IV of Detection, Estimation and Modulation Theory*. New York, NY, USA: Wiley Intersci., 2002.

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

[3] 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, 2010.

[4] 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, 2011.

[5] P. Pal and P. P. Vaidyanathan, “Coprime sampling and the music algorithm,” in *Proc. IEEE Digit. Signal Process. Workshop and IEEE Signal Process. Educ. Workshop (DSP/SPE)*, 2011, pp. 289–294.

[6] Y. D. Zhang, M. G. Amin, and B. Himed, “Sparsity-based DOA estimation using co-prime arrays,” in *Proc. IEEE Int. Conf. Acoust, Speech, Signal Process.*, Vancouver, Canada, May 2013, pp. 3967–3971.

[7] P. Pal and P. P. Vaidyanathan, “Correlation-aware techniques for sparse support recovery,” in *Proc. IEEE Statist. Signal Process. Workshop (SSP)*, 2012, pp. 53–56.

[8] P. Pal and P. P. Vaidyanathan, “On application of LASSO for sparse support recovery with imperfect correlation awareness,” in *Proc. Conf. Rec. 46th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, 2012, pp. 958–962.

[9] P. Pal and P. P. Vaidyanathan, "Correlation-aware sparse support recovery: Gaussian sources," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2013, pp. 5880–5884.

[10] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2182–2195, May 2011.

[11] Z. Tan, P. Yang, and A. Nehorai, "Joint-sparse recovery in compressed sensing with dictionary mismatch," in *Proc. IEEE 5th IEEE Int. Workshop Computat. Adv. Multi-Sens. Adapt. Process. (CAMSAP)*, 2013, pp. 248–251.

[12] Z. Yang, C. Zhang, and L. Xie, "Robustly stable signal recovery in compressed sensing with structured matrix perturbation," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4658–4671, 2012.

[13] Z. Tan and A. Nehorai, "Sparse direction of arrival estimation using co-prime arrays with off-grid targets," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 26–29, 2014.

[14] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.*, vol. 67, no. 6, pp. 906–956.

[15] E. J. Candès and C. Fernandez-Granda, "Super-resolution from noisy data," *J. Fourier Anal. Appl.*, vol. 19, no. 6, pp. 1229–1254, Aug. 2013.

[16] C. Fernandez-Granda, "Support detection in super-resolution," in *Proc. 10th Int. Conf. Sampling Theory Appl.*, 2013, pp. 145–148.

[17] J. A. Tropp, "Just relax: Convex programming methods for identifying sparse signals in noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1030–1051, 2006.

[18] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, 1974.

[19] Z. He, A. Cichocki, S. Xie, and K. Choi, "Detecting the number of clusters in n-way probabilistic clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 2006–2021, 2010.

[20] W. Chen, K. M. Wong, and J. P. Reilly, "Detection of the number of signals: A predicted eigen-threshold approach," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1088–1098, 1991.

[21] K. Han and A. Nehorai, "Improved source number detection and direction estimation with nested arrays and ulas using jackknifing," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 6118–6128, 2013.

[22] R. Prony, "Essai expérimental et analytique: Sur les lois de la dilatabilité de fluidesélastique et sur celles de la force expansive de la vapeur de l'alkool,à diérentes températures," *J. de l'Ecole Polytechn.*, vol. 1, no. 2, pp. 24–76.

[23] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, no. 5, pp. 814–824, May 1990.

[24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[25] J. Haupt, W. U. Bajwa, G. Raz, and R. Nowak, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5862–5875, 2010.

Zhao Tan (S'12) received the B.Sc. degree in electronic information science and technology from Fudan University of China in 2010. He received M.Sc. degree in electrical engineering from Washington University in St. Louis in 2013.

He is currently working towards a Ph.D. degree with the Preston M. Green Department of Electrical and Systems Engineering at Washington University under the guidance of Dr. A. Nehorai. His research interests are mainly in the areas of optimization algorithms, compressed sensing, sensor arrays, radar signal processing, and state estimation and scheduling in smart grid.

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

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

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

Arye Nehorai (S'80–M'83–SM'90–F'94) received the B.Sc. and M.Sc. degrees from the Technion, Israel, and the Ph.D. from Stanford University, CA.

He is the Eugene and Martha Lohman Professor and Chair of the Preston M. Green Department of Electrical and Systems Engineering (ESE), Professor in the Department of Biomedical Engineering (by courtesy) and in the Division of Biology and Biomedical Studies (DBBS) at Washington University, St. Louis (WUSTL), MO. He serves as Director of the Center for Sensor Signal and Information Processing at WUSTL. Under his leadership as department chair, the undergraduate enrollment has more than tripled in the last four years. Earlier, he was a faculty member at Yale University and the University of Illinois at Chicago.

Dr. Nehorai served as Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2000 to 2002. From 2003 to 2005, he was the Vice President (Publications) of the IEEE Signal Processing Society (SPS), the Chair of the Publications Board, and a member of the Executive Committee of this Society. He was the founding editor of the special columns on Leadership Reflections in IEEE SIGNAL PROCESSING MAGAZINE from 2003 to 2006. He received the 2006 IEEE SPS Technical Achievement Award and the 2010 IEEE SPS Meritorious Service Award. He was elected Distinguished Lecturer of the IEEE SPS for a term lasting from 2004 to 2005. He received several best paper awards in IEEE journals and conferences. In 2001, he was named University Scholar of the University of Illinois. He has been a Fellow of the Royal Statistical Society since 1996 and a Fellow of AAAS since 2012.