

# Generalized Coprime Sampling of Toeplitz Matrices for Spectrum Estimation

Si Qin, Yimin D. Zhang\*, *Senior Member, IEEE*, Moeness G. Amin, *Fellow, IEEE*,  
and Abdelhak M. Zoubir, *Fellow, IEEE*

**Abstract**—Increased demand on spectrum sensing over a broad frequency band requires a high sampling rate and thus leads to a prohibitive volume of data samples. In some applications, e.g., spectrum estimation, only the second-order statistics are required. In this case, we may use a reduced data sampling rate by exploiting a low-dimensional representation of the original high dimensional signals. In particular, the covariance matrix can be reconstructed from compressed data by utilizing its specific structure, e.g., the Toeplitz property. Among a number of techniques for compressive covariance sampler design, the coprime sampler is considered attractive because it enables a systematic design capability with a significantly reduced sampling rate. In this paper, we propose a general coprime sampling scheme that implements effective compression of Toeplitz covariance matrices. Given a fixed number of data samples, we examine different schemes on covariance matrix acquisition for performance evaluation, comparison and optimal design, based on segmented data sequences.

**Index Terms**—Compressive covariance sampling, structured matrix, coprime sampling, overlapping data segmentation

## I. INTRODUCTION

Various applications require spectrum sensing over a broad frequency band, which demand on the sampling rate and produce a large amount of data. In some cases, the original signal is known to be sparse. This property allows the exploitation of compressive sensing and sparse sampling approaches that enable effective sparse signal reconstruction [3], [4], with no loss of information. The signal reconstruction can be carried out by a number of algorithms, such as orthogonal matching pursuit (OMP), least absolute shrinkage and selection operator (LASSO), and Bayesian compressive sensing [5]–[8].

Spectrum estimation based on the second-order statistics adds to the abovementioned applications for signal reconstruction. In this case, the covariance function and the covariance matrix can be constructed as low-dimensional representations

of the original high-dimensional signals [9], [10]. This fact has motivated the development of an alternate framework, referred to as compressive covariance sampling, in which the signal sparsity is not a requirement [11]–[13].

In this paper, we consider spectrum estimation of wide-sense stationary (WSS) processes utilizing the Toeplitz property of the covariance matrix. Note that, while our focus in this paper is limited to the second-order statistics, extension to techniques based on high-order statistics [14] is straightforward.

Several methods have been developed to tackle similar compressive Toeplitz matrix sampling. For example, a generalized nested sampler [15] was proposed to recover Toeplitz matrices from a compressed covariance matrix. However, this approach assumes an infinite number of data samples and does not consider the achievable reconstruction performance when the number of samples is finite. In addition, it imposes a minimum sampling interval that follows the Nyquist criterion, which makes it ineffective to implement low sampling rate systems for wideband spectrum estimation. In [16], a minimal sparse sampler was proposed through a set of properly designed analog filters and then down-sampling the signals at a reduced rate. A finite number of outputs was divided into multiple blocks without overlapping, and the compressed covariance was estimated by averaging over these blocks. However, the requirement of using the designed analog filters complicates the implementation. In addition, the effect of utilization of overlapping blocks were not considered.

The proposed work is based on the recently developed coprime sampling structure [17], which utilizes only two uniform samplers to sample a WSS process with sampling intervals,  $M$  and  $N$ . The integers  $M$  and  $N$ , which represent the down-sampling rates, are chosen to be coprime. As a result, it generates two sets of uniformly spaced samples with a rate substantially lower than the nested [18] and with fewer samplers than the schemes in [19]–[21].

In this paper, we design a sampling matrix to compress Toeplitz matrices based on a coprime sampling scheme. In particular, our focus is on effective estimations of the Toeplitz covariance matrix and signal spectrum from a finite number of samples of a WSS sequence. Toward this objective, we generalize the coprime sampling approach to achieve a higher number of degrees of freedom (DOFs) and low estimation error. The generalization is carried out in the following two aspects: (a) The first generalization is to use multiple coprime units to obtain a higher number of DOFs and improved power spectrum density (PSD) estimation performance. This

Copyright (c) 2016 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

The work of S. Qin, Y. D. Zhang, and M. G. Amin was supported in part by the Office of Naval Research under Grant No. N00014-13-1-0061. Part of the results was presented at the SPIE Mobile Multimedia/Image Processing, Security, and Applications Conference, Baltimore, MD, May 2015 [1], and the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Shanghai, China, March 2016 [2].

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

Y. D. Zhang is with the Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA, 19122, USA.

A. Zoubir is with Signal Processing Group, Technische Universität Darmstadt, Darmstadt, Germany.

The corresponding author's email address is ydzhang@temple.edu.

is achieved through the use of an integer factor  $p$ , where a coprime unit is defined as a full period of the output sample pattern between  $x[bMN]$  and  $x[(b+1)MN-1]$  for any non-negative integer  $b$ . (b) The second generalization is to exploit overlapping blocks in performing sample averaging, enabling an increased number of blocks to be used for sample averaging, leading to a reduced estimation variance.

The concept of generalized coprime sampling was first developed in [1] where only the abovementioned first generalization is considered, whereas the second generation was introduced in [2]. In this paper, we extend these preliminary results by providing comprehensive theoretical support and performance bound analysis of the developed techniques, and describe the spectrum estimation algorithm based on the cross-covariance between the outputs of the two samplers. A number of simulation results are presented to clearly reveal the relationship between the achieved performance and various parameters related to the sampling strategies and signal conditions.

The rest of the paper is organized as follows. We first introduce the signal model in Section II. Generalized coprime sampling that exploits multiple coprime units is presented in Section III. Section IV describes spectrum estimation based on the generalized coprime sampling scheme, and the corresponding spectrum identifiability, compression factor, and Cramér-Rao bound (CRB) are examined. In Section V, we propose the exploitation of overlapping samples, and show analytically that the overlapping sampling scheme achieves reduced variance in the estimated covariance matrix and signal spectrum. Simulation results are provided in Section VI to numerically verify the effectiveness of the proposed generalization and the analysis. Section VII concludes the 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.  $(\cdot)^*$  implies complex conjugation, whereas  $(\cdot)^T$  and  $(\cdot)^H$  respectively denote the transpose and conjugate transpose of a matrix or a vector.  $\mathbb{E}(\cdot)$  is the statistical expectation operator and  $\otimes$  denotes the Kronecker product.  $\mathbb{R}$  and  $\mathbb{C}$  denote the set of real values and complex values, respectively, while  $\mathbb{N}^+$  denotes the set of positive integers.  $x \sim \mathcal{CN}(a, b)$  denotes that random variable  $x$  follows the complex Gaussian distribution with mean  $a$  and variance  $b$ .  $\lfloor \cdot \rfloor$  denotes the floor function which returns the largest integer not exceeding the argument.  $\text{diag}(\mathbf{x})$  denotes a diagonal matrix that uses the elements of  $\mathbf{x}$  as its diagonal elements, and  $\text{Tr}\{\mathbf{A}\}$  returns the trace of matrix  $\mathbf{A}$ .

## II. SIGNAL MODEL

Assume that a zero-mean WSS process  $X(t), t \in \mathbb{R}$ , which consists of signals corresponding to a number of sparse frequencies, is confined within a bandwidth  $B_s$ . To obtain its PSD, the covariance matrix needs to be provided from a specific realization of  $X(t), t = 0, \dots, T-1$ . It suffices to consider the discrete-time random process,  $X[l]$ , obtained by sampling the analog signal  $X(t)$ , with a Nyquist sampling rate  $f_s = 2B_s$ . Note that the discrete-time process  $X[l]$  remains WSS in the discrete-time sense. Let  $\mathbf{x}_L[l] =$

$[x[l], x[l+1], \dots, x[l+L-1]]^T$  be a realized vector of  $X[l]$ . Then, the resulting semi-positive definite, Hermitian and Toeplitz covariance matrix can be given by

$$\begin{aligned} \mathbf{R}_x &= \mathbb{E}[\mathbf{x}_L[l]\mathbf{x}_L^H[l]] \\ &= \begin{pmatrix} r[0] & r[-1] & \dots & r[-L+1] \\ r[1] & r[0] & \dots & r[-L+2] \\ \vdots & \vdots & \dots & \vdots \\ r[L-2] & r[L-3] & \dots & r[-1] \\ r[L-1] & r[L-2] & \dots & r[0] \end{pmatrix}, \quad (1) \end{aligned}$$

in which the entry  $r[\tau] = \mathbb{E}[x[l]x^*[l-\tau]]$  only depends on the lags  $\tau = -L+1, \dots, L-1$ . It is clear from (1) that  $r[-\tau] = r^*[\tau]$ . In addition, the Toeplitz structure of  $\mathbf{R}_x$  implies that many of its elements are redundant. As a result,  $\mathbf{R}_x$  can be obtained from a sparsely sampled data sequence. This fact motivated compressive covariance sampling [11]–[13].

In this paper, we consider the problem of estimating an  $L \times L$  covariance matrix of  $\mathbf{x}_L[l]$  and the signal PSD from an observation of  $X(t)$  with an available length of  $KT_s$ , where  $K \in \mathbb{N}^+$  and  $K \geq L$ . When sampled at the Nyquist interval  $T_s = 1/f_s$ , it yields  $K$  samples of discrete-time observations  $x[k], k = 0, \dots, K-1$ . A common practice for covariance matrix estimation is to segment the entire discrete-time observation of length  $K$  into multiple length- $L$  blocks, and average the respectively sample covariances [22]. As shown in Fig. 1, the entire observation period is segmented into multiple, possibly overlapping, blocks. In Section III-B, we first consider the non-overlapping segmentation to illustrate the signal model, as shown in Fig. 1(a), whereas the overlapping case depicted in Fig. 1(b) will be discussed in Section III-C. Denote  $B$  as the number of data blocks for the non-overlapping case. We assume for convenience that the  $B$  blocks cover the entire sequence, i.e.,  $BL = K$ .

Denote by  $x_b[l] = x[l + (b-1)L], l = 0, \dots, L-1$ , and  $\mathbf{x}_b = [x_b[0], \dots, x_b[L-1]]^T$  for  $b = 1, \dots, B$ . We sparsely sample each data block using a  $V \times L$  sampling matrix  $\mathbf{A}_s$  to obtain  $\mathbf{y}_b = \mathbf{A}_s \mathbf{x}_b$ , where  $V \ll L$ . The estimated covariance matrix obtained by averaging the available  $B$  blocks and is expressed as

$$\hat{\mathbf{R}}_y = \frac{1}{B} \sum_{b=1}^B \mathbf{y}_b \mathbf{y}_b^H = \mathbf{A}_s \left( \frac{1}{B} \sum_{b=1}^B \mathbf{x}_b \mathbf{x}_b^H \right) \mathbf{A}_s^H = \mathbf{A}_s \hat{\mathbf{R}}_x \mathbf{A}_s^H, \quad (2)$$

where  $\hat{\mathbf{R}}_x$  is an estimated covariance matrix of  $\mathbf{R}_x$ . The compressed covariance matrix  $\hat{\mathbf{R}}_y$  with size  $V \times V$  can be exploited to reconstruct the  $L \times L$  matrix  $\hat{\mathbf{R}}_x$ , provided that it includes all lags  $\tau = 0, \dots, L-1$ . Note that covariances corresponding to negative lags  $\tau = -L+1, \dots, -1$  can be obtained through the Hermitian operation  $r[\tau] = r^*[-\tau]$  and thus does not contain additional information. Reconstruction of full covariance matrix  $\mathbf{R}_x$  from the compressed covariance matrix  $\hat{\mathbf{R}}_y$  can be made possible by designing a proper sampling matrix  $\mathbf{A}_s$ . It is clear that, since there are  $V^2$  entries in  $\hat{\mathbf{R}}_y$ , the number of samples required to enable reconstruction of the Hermitian Toeplitz matrix  $\hat{\mathbf{R}}_x$  is  $\mathcal{O}(\sqrt{L})$ .

In this end,  $\hat{\mathbf{R}}_{\mathbf{x}}$  can be reconstructed as

$$\hat{\mathbf{R}}_{\mathbf{x}} = \begin{pmatrix} \hat{r}[0] & \hat{r}[-1] & \dots & \hat{r}[-L+1] \\ \hat{r}[1] & \hat{r}[0] & \dots & \hat{r}[-L+2] \\ \vdots & \vdots & \dots & \vdots \\ \hat{r}[L-2] & \hat{r}[L-3] & \dots & \hat{r}[-1] \\ \hat{r}[L-1] & \hat{r}[L-2] & \dots & \hat{r}[0] \end{pmatrix}, \quad (3)$$

where  $\hat{r}[\tau]$ ,  $\tau = -L+1, \dots, L-1$  are estimated by averaging all the entries with the same lag  $\tau$  in  $\hat{\mathbf{R}}_{\mathbf{y}}$ .

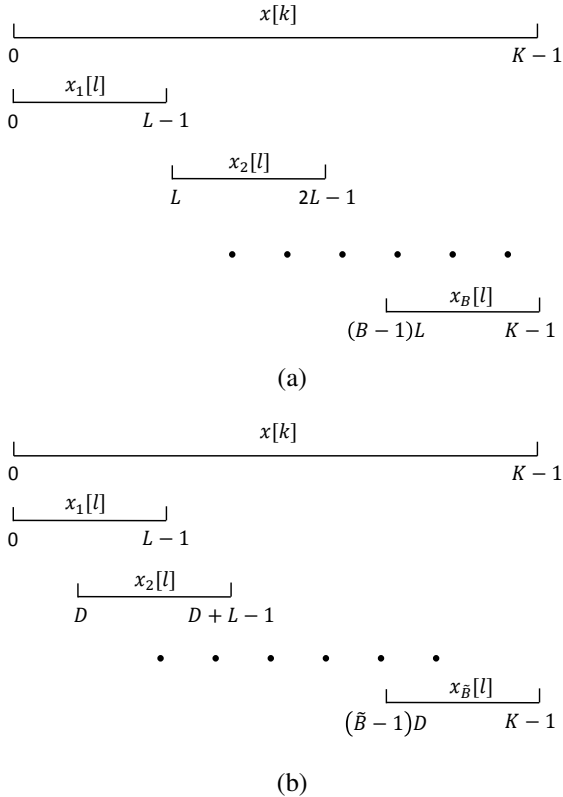


Fig. 1. Illustration of segmentations. (a) Non-overlapping segmentation; (b) Overlapping segmentation.

### III. GENERALIZED COPRIME SAMPLING

Coprime sampling exploits two uniform sub-Nyquist samplers with sampling period being coprime multiples of the Nyquist sampling period [17], [23]. In this section, the generalized coprime sampling scheme is presented in two operations. A multiple coprime unit factor  $p \in \mathbb{N}^+$  [1], aiming to increase the number of lags in the compressed covariance matrix, is first introduced. Then, the utilization of overlapping samples between blocks is pursued to yield a reduced estimation variance through the use of a non-overlapping factor  $q \in \mathbb{N}^+$ .

#### A. The concept of coprime sampling

In coprime sampling, the sampling matrix  $\mathbf{A}_s$  can be denoted as  $\mathbf{A}_s = [\mathbf{A}_{s1}^T \quad \mathbf{A}_{s2}^T]^T$ , where  $\mathbf{A}_{s1}$  and  $\mathbf{A}_{s2}$  are the sub-sampling matrices corresponding to the two coprime samplers.

**Definition 1:** The  $(i, j)$ th entry of the sampling matrices  $\mathbf{A}_{s1}$  and  $\mathbf{A}_{s2}$  can be designed as:

$$[\mathbf{A}_{s1}]_{i,j} = \begin{cases} 1, & j = Mi, \quad i \in \mathbb{N}^+, \\ 0, & \text{elsewhere,} \end{cases}$$

and

$$[\mathbf{A}_{s2}]_{i,j} = \begin{cases} 1, & j = Ni, \quad i \in \mathbb{N}^+, \\ 0, & \text{elsewhere,} \end{cases} \quad (4)$$

where  $M \in \mathbb{N}^+$  and  $N \in \mathbb{N}^+$  are coprime integers.

From a data acquisition perspective, there are two sets of uniformly spaced samples of the input WSS signal  $X(t)$ ,  $t = 0, \dots, T$ , from two samplers with sampling intervals  $MT_s$  and  $NT_s$ , respectively, as illustrated in Fig. 2. Without loss of generality, we assume  $M < N$ . Then, the highest sampling rate of the system is  $1/(MT_s) = f_s/M$  and the two sampled stream outputs can be given as

$$\begin{aligned} y_1[k_1] &= x[Mk_1] = X(Mk_1T_s), \\ y_2[k_2] &= x[Nk_2] = X(Nk_2T_s). \end{aligned} \quad (5)$$

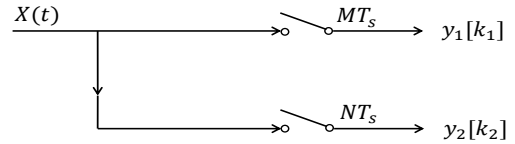


Fig. 2. Coprime sampling structure.

Note that, due to the coprime property of  $M$  and  $N$ , there are no overlapping outputs between the two sets other than  $x[bMN]$  for any non-negative integer  $b$ . The outputs between  $x[(b-1)MN]$  and  $x[bMN-1]$  are referred to as a coprime unit, positioned at

$$\mathbb{P}_b = \{bMN + Mk_1\} \cup \{bMN + Nk_2\}. \quad (6)$$

Over an observation with an available length of  $KT_s$ ,  $K/MN$  coprime units can be obtained, each consists of  $M+N$  physical samples. As such, the total number of physical samples is given by

$$K_s = K \left( \frac{M+N}{MN} \right) = K \left( \frac{1}{M} + \frac{1}{N} \right). \quad (7)$$

For illustration, an example is presented in Fig. 3, where two coprime samplers with  $M = 3$  and  $N = 4$  are considered. The length of  $K = 60$  output streams consist of 5 coprime units, and  $K_s = 35$  physical samples are distributed between  $x[12(b-1)]$  and  $x[12b-1]$ , for  $b = 1, \dots, 5$ , where 5 pairs of samples overlap between the output of the two samplers at positions 0, 12, 24, 36, and 48.

Denote  $\mathbf{y}_{b1} = [y_{b1}[0], \dots, y_{b1}[N-1]]^T$  as an  $N \times 1$  vector, and  $\mathbf{y}_{b2} = [y_{b2}[0], \dots, y_{b2}[M-1]]^T$  as an  $M \times 1$  vector, with  $y_{b1}[k_1] = x[(b-1)MN + Mk_1]$  and  $y_{b2}[k_2] = x[(b-1)MN + Nk_2]$ , where  $0 \leq k_1 \leq N-1$  and  $0 \leq k_2 \leq M-1$ , for  $1 \leq b \leq K/(MN)$ . In addition, let  $\mathbf{y}_b = [\mathbf{y}_{b1}^T \quad \mathbf{y}_{b2}^T]^T$ . As

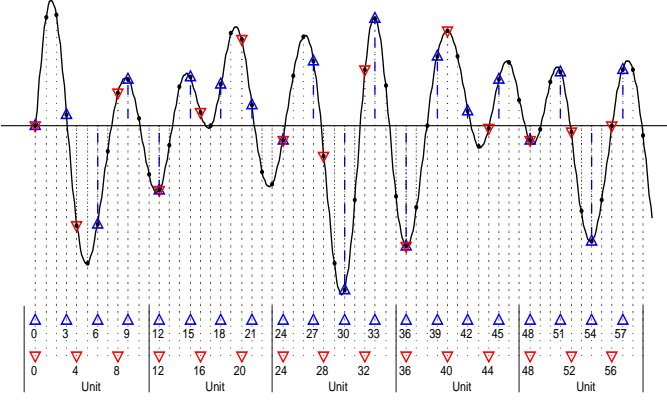


Fig. 3. An example for coprime sampling ( $M = 3$ , and  $N = 4$ ;  $\bullet$ : Nyquist sampler;  $\triangle$ : first sampler outputs;  $\nabla$ : second sampler outputs.)

such, the  $(M + N) \times (M + N)$  covariance matrix  $\mathbf{R}_y$  can be expressed as

$$\mathbf{R}_y = \begin{pmatrix} \mathbf{R}_{y_{11}} & \mathbf{R}_{y_{12}} \\ \mathbf{R}_{y_{21}} & \mathbf{R}_{y_{22}} \end{pmatrix} = \begin{pmatrix} \mathbb{E}[y_{b_1} y_{b_1}^H] & \mathbb{E}[y_{b_1} y_{b_2}^H] \\ \mathbb{E}[y_{b_2} y_{b_1}^H] & \mathbb{E}[y_{b_2} y_{b_2}^H] \end{pmatrix}. \quad (8)$$

In  $\mathbf{R}_y$ , matrices  $\mathbf{R}_{y_{11}}$  and  $\mathbf{R}_{y_{22}}$  contains self-lags of the two sampler output streams, while their cross-lags are included in matrices  $\mathbf{R}_{y_{12}}$  and  $\mathbf{R}_{y_{21}}$ . Note that  $\mathbf{R}_{y_{21}} = \mathbf{R}_{y_{12}}^*$ . In addition, because the two sampler outputs share the first sample in each coprime unit, the self-lags can be taken as cross-lags between every sample from one sampler and the first sample from the other sampler. As such, the self-lags form a subset of the cross-lags. Thus,  $\mathbf{R}_x$  can be reconstructed by using only  $\mathbf{R}_{y_{12}}$ , whose cross-lags (including the negated ones) are given by the following set,

$$\mathbb{L} = \{\tau | \tau = \pm(Mk_1 - Nk_2)\}, \quad (9)$$

where  $0 \leq k_1 \leq N - 1$  and  $0 \leq k_2 \leq M - 1$ .

The prototype scheme uses one coprime unit samples to generate all lags in  $\mathbb{L}$ . However, it should be noticed that they are distributed in the range  $[-M(N - 1), M(N - 1)]$  with some missing integers at  $(aM + bN)$ , where  $a \geq 1$  and  $b \geq 1$ , as shown in Fig. 4(a), for  $M = 3$  and  $N = 4$ . That is, they are not sufficient to reconstruct  $\hat{\mathbf{R}}_x$  with dimension  $L = MN$ . To overcome this limitation, two coprime units from the first sampler and one coprime unit from the second sampler are used to form one block in [17], and the resulting lags are contiguous in the range  $[-MN - N + 1, MN + N - 1]$ , as depicted in Fig. 4(b). This scheme is referred in this paper to as the conventional scheme. In this case, the maximum achievable  $L$  is  $L_{\max} = MN + N$ .

### B. Generalized coprime sampling scheme using non-overlapping blocks

In the sequel, an integer factor  $p \geq 2$ , representing the number of multiple coprime units, is first introduced to achieve a higher value of  $L$ . In each block, outputs from  $p$  coprime units from both samplers, i.e.,  $p(M + N)$  physical samples spawning a time period of  $pMNT_s$ , are used to estimate the

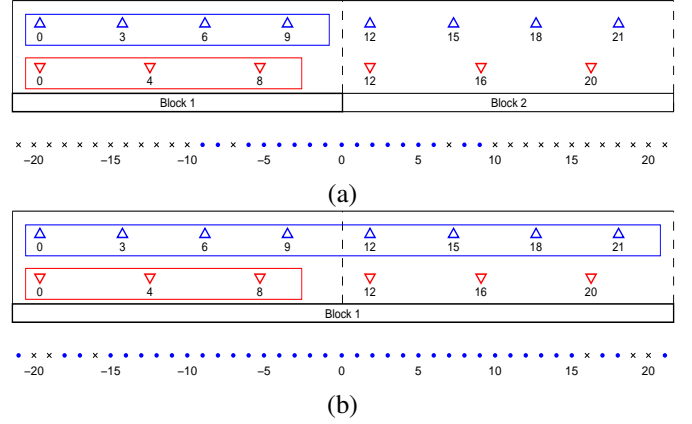


Fig. 4. An example for different schemes ( $M = 3$ , and  $N = 4$ ;  $\triangle$ : first sampler outputs;  $\nabla$ : second sampler outputs;  $\bullet$ : lags;  $\times$ : holes). (a) Prototype; (b) Conventional.)

covariance matrix. In this case, the resulting lags fall into the following set,

$$\tilde{\mathbb{L}} = \{\tau | \tau = \pm(Mk_1 - Nk_2)\}, \quad (10)$$

for  $0 \leq k_1 \leq pN - 1$  and  $0 \leq k_2 \leq pM - 1$ . Note that varying  $p$  changes the set  $\tilde{\mathbb{L}}$ . The following proposition about the set  $\tilde{\mathbb{L}}$  reveals the property of the resulting lag positions.

**Proposition 1:** The set  $\tilde{\mathbb{L}}$  contains all integer lags in the range  $-(p - 1)MN - M - N + 1 \leq \tau \leq (p - 1)MN + M + N - 1$ .

The proof is provided in Appendix A. Note that, all resulting lags using conventional scheme are included in  $\tilde{\mathbb{L}}$  as a special case of  $p = 2$ . For the generalized scheme, the maximum achievable value of  $L$  becomes

$$\tilde{L}_{\max} = (p - 1)MN + M + N, \quad (11)$$

and the number of the corresponding non-overlapping blocks is given by

$$B = \left\lfloor \frac{K}{pMN} \right\rfloor. \quad (12)$$

An example for different values of  $p$  is illustrated in Fig. 5, where  $K = 120$ ,  $M = 3$ , and  $N = 4$  are assumed. For the case of  $p = 2$ , i.e., the conventional scheme, each block forms consecutive lags within  $[-18, 18]$ . That is,  $\hat{\mathbf{R}}_x$  can be reconstructed with a maximum of dimension  $\tilde{L}_{\max} = 19$  by averaging  $B = 5$  blocks. For the case of  $p = 5$ ,  $\tilde{L}_{\max} = 55$  can be obtained by a consecutive lag range of  $[-54, 54]$  in each block, whereas the number of the blocks is reduced to  $B = 2$ .

We examine the compression factor, which is defined as the ratio of the number of entries in  $\hat{\mathbf{R}}_x$  over the corresponding number in  $\hat{\mathbf{R}}_{y_{12}}$ , expressed as

$$\kappa = \frac{L \times L}{pM \times pN} = \frac{L^2}{p^2MN}. \quad (13)$$

Because the maximum value of  $L$  is  $\tilde{L}_{\max} = (p - 1)MN + M + N$ , the maximum achievable value of  $\kappa$  is given by

$$\kappa_{\max} = \frac{[(p - 1)MN + M + N]^2}{p^2MN}. \quad (14)$$

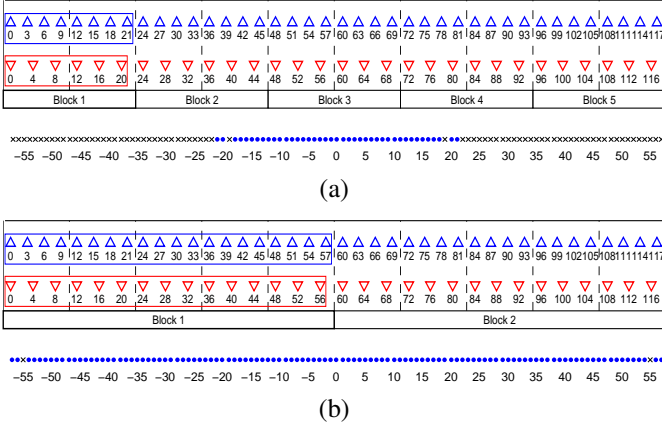


Fig. 5. An example for different values of  $p$  ( $K = 120$ ,  $M = 3$ , and  $N = 4$ ; (a)  $p = 2$ ; (b)  $p = 5$ .)

Fig. 6 shows  $\kappa_{\max}$ , as a function of  $M$ ,  $N$ , and  $p$ . It is clear that  $\kappa_{\max}$  improves as  $M$  and  $N$  increase. Notice that, while the number of entries in  $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$  increases with  $p$ ,  $\kappa_{\max}$  does not significantly change. It asymptotically approaches  $MN$  when  $p \gg 1$ .

For a given number of compression factor, i.e., the constant value of  $MN$ , the optimal coprime pair in terms of total number of physical samples,  $K_s$ , can be derived by solving the optimization problem:

$$\begin{aligned} & \text{Minimize} && K_s = K \left( \frac{1}{M} + \frac{1}{N} \right) \\ & \text{subject to} && MN = \text{constant}, \\ & && 0 < M < N. \end{aligned} \quad (15)$$

It is demonstrated in [23], [31] that the valid optimal coprime pair is the one that has  $M$  and  $N$  as close as possible. This is satisfied by choosing  $N = M + 1$ . This relationship is assumed in the remainder of the paper. In this case,  $K_s$  becomes

$$K_s = K \left( \frac{1}{M} + \frac{1}{M+1} \right), \quad (16)$$

and the corresponding compression factor,  $\kappa_{\max}$ , can be expressed as

$$\kappa_{\max} = \frac{\tilde{L}_{\max}^2}{p^2 M(M+1)} \propto M^2, \quad (17)$$

with  $\tilde{L}_{\max} = (p-1)M^2 + (p+1)M + 1$ .

### C. Utilization of overlapping blocks

The variance of the estimated covariance and spectrum is generally reduced by utilizing a higher number of averaging blocks. In addition to averaging over non-overlapping segments, as discussed earlier, a general and more effective alternative for spectrum estimation is to exploit overlapping segments. In so doing, the number of applicable blocks for sample averaging can be substantially increased. The overlapping samples used are set by non-overlapping factor  $q \in \mathbb{N}^+$ .

As shown in Fig. 1(b), we maintain the same segment length  $pM(M+1)$ , and let the starting points of two adjacent blocks  $D = qM(M+1)$  units apart, where  $1 \leq q \leq p$ . Similarly, we

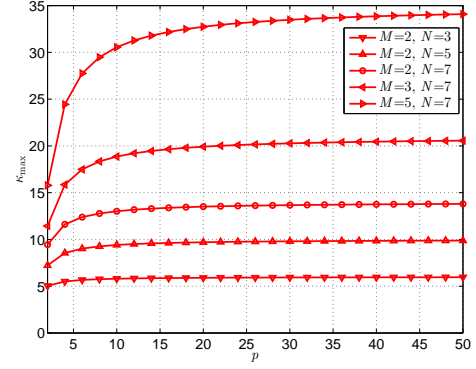


Fig. 6.  $\kappa_{\max}$ , as a function of  $M$ ,  $N$ , and  $p$ .

assume, for convenience, that  $(\tilde{B}-1)D + pM(M+1) = K$  covers the entire recorded sequence.

**Definition 2:** Assume that  $D$  consists of the length of  $q$  coprime units, i.e.,  $D = qM(M+1)$ , where  $1 \leq q \leq p$ . Then, the number of blocks can be expressed as

$$\begin{aligned} \tilde{B} &= \left\lfloor \frac{K - pM(M+1)}{D} \right\rfloor + 1 = \left\lfloor \frac{K}{qM(M+1)} - \frac{p}{q} \right\rfloor + 1 \\ &= \left\lfloor \frac{p}{q} B - \frac{p}{q} \right\rfloor + 1. \end{aligned} \quad (18)$$

It is straightforward to confirm that  $\tilde{B} \geq B$  since  $q \leq p$ . In addition,  $\tilde{B}/B$  approaches  $p/q$  when  $B$  is large. As such,  $p/q$  can be considered as the overlapping ratio that approximately describes the level of additional blocks used for sample averaging. It is clear that  $\tilde{B}$  increases as  $q$  decreases and is maximized when  $q = 1$ . Note that the non-overlapping case can be considered as a special case of  $q = p$  and  $\tilde{B} = B$ .

For illustration, an example of  $p = 5$  and  $q = 1$  is considered in Fig. 7, where  $K$  and  $M$  are assumed to be the same as those in Fig. 5. It is shown that  $\tilde{B} = 6$  blocks can be used in Fig. 7, whereas only  $B = 2$  blocks are obtained in the corresponding non-overlapping scenario, as depicted in Fig. 5(b).

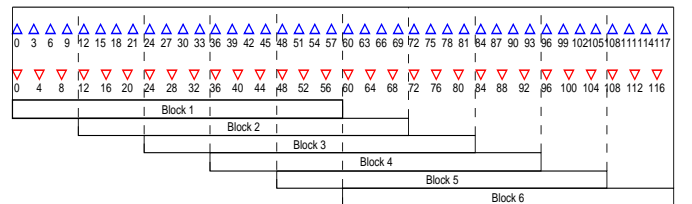


Fig. 7. An example of utilization of overlapping samples ( $K = 120$ ,  $M = 3$ ,  $p = 5$ , and  $q = 1$ .)

Denote  $\tilde{y}_{b_1}[k_1] = x[(b-1) \times qM(M+1) + Mk_1]$  and  $\tilde{y}_{b_2}[k_2] = x[(b-1) \times qM(M+1) + (M+1)k_2]$ , where  $0 \leq k_1 \leq p(M+1) - 1$  and  $0 \leq k_2 \leq pM - 1$ , for  $1 \leq b \leq \tilde{B}$ . In addition, let  $\tilde{\mathbf{y}}_{b_1} = [\tilde{y}_{b_1}[0], \dots, \tilde{y}_{b_1}[p(M+1) - 1]]^T$  and  $\tilde{\mathbf{y}}_{b_2} = [\tilde{y}_{b_2}[0], \dots, \tilde{y}_{b_2}[pM - 1]]^T$ . The covariance matrix  $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$ ,

using the generalized scheme, can be estimated as

$$\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}} = \frac{1}{\tilde{B}} \sum_{\tilde{b}=1}^{\tilde{B}} \tilde{\mathbf{y}}_{b_1} \tilde{\mathbf{y}}_{b_2}^H. \quad (19)$$

Note that, for each  $1 \leq b \leq \tilde{B}$ , the entries  $r_b(k_1, k_2) = \tilde{y}_{b_1}[k_1] \tilde{y}_{b_2}^*[k_2]$  corresponding to the same position  $(k_1, k_2)$  in covariance matrix are still independent. As discussed above, the value of  $\tilde{B}$  is increased from that of  $B$  approximately by a factor of  $p/q$ . Thus, utilizing overlapping blocks for averaging, the variance of the estimated covariance is generally reduced to  $q/p$  of the corresponding non-overlapping case.

Then,  $\hat{\mathbf{R}}_{\mathbf{x}}$  with dimension  $L \times L$ , where  $L \leq \tilde{L}_{\max}$ , can be reconstructed as

$$\hat{\mathbf{R}}_{\mathbf{x}} = \begin{pmatrix} \hat{r}[0] & \hat{r}[-1] & \dots & \hat{r}[-L+1] \\ \hat{r}[1] & \hat{r}[0] & \dots & \hat{r}[-L+2] \\ \vdots & \vdots & \dots & \vdots \\ \hat{r}[L-2] & \hat{r}[L-3] & \dots & \hat{r}[-1] \\ \hat{r}[L-1] & \hat{r}[L-2] & \dots & \hat{r}[0] \end{pmatrix}, \quad (20)$$

where  $\hat{r}[\tau], \tau = -L+1, \dots, L-1$ , are estimated by averaging all the entries with the same lag  $\tau$  in  $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$ .

We make the following two remarks:

1. In the generalized coprime sampling scheme,  $(\tilde{B} - 1)qM(M+1) + pM(M+1) = K$ , where  $p, q \in \mathbb{N}^+$ , is assumed to cover the entire recorded sequence. When  $\tilde{B} = 1$ , factor  $q$  does not have a physical meaning. Thus,  $\tilde{B} \geq 2$  needs to be guaranteed, which is equivalent to

$$p + q \leq \frac{K}{M(M+1)}. \quad (21)$$

As such, the range of the pair of  $(p, q)$  falls into the following set,

$$\Pi_{p,q} = \left\{ (p, q) \mid p + q \leq \frac{K}{M(M+1)}, 1 \leq q \leq p, p, q \in \mathbb{N}^+ \right\}. \quad (22)$$

2. The covariance matrix  $\mathbf{R}_{\tilde{\mathbf{y}}_{12}}$  is estimated using the  $\tilde{B}$  available samples. In practice,  $p$  and  $q$  are generally chosen to yield the large number of blocks  $\tilde{B}$  to achieve to rigid estimation of  $\mathbf{R}_{\tilde{\mathbf{y}}_{12}}$ .

3. As  $p$  increases, a higher number of DOFs in the compressed covariance matrix  $\mathbf{R}_{\tilde{\mathbf{y}}_{12}}$  can be achieved. As a result, we can reconstruct covariance matrix  $\hat{\mathbf{R}}_{\mathbf{x}}$  with a higher dimension, yielding an improved spectrum resolution and estimation accuracy. When  $q$  increases, the estimation accuracy can be improved because a higher number of blocks are used in the averaging. However, such higher dimension and higher number of blocks result in a higher computational complexity.

#### IV. SPECTRUM ESTIMATION AND THE CRB

Spectrum estimation deals with the problem of estimating the PSD of a random process, and finds applications in the context of dynamic spectrum sharing [24]. In this case, a broad frequency band should be sensed in order to locate the unoccupied spectrum before establishing a communication link. Sub-Nyquist sampling for cognitive radios is a widely studied topic, e.g., in [25]–[30].

Generally, power spectrum sensing can be classified into two major categories. The first category reconstructs the signal waveforms and then estimate the power spectrum, whereas the second category estimates the power spectrum from the signal covariance, i.e., the second-order statistics. The approach discussed in [25]–[27] belongs to the former where the signals are assumed to be sparse in some domain and sub-Nyquist sampling is implemented to recover the signal waveforms through compressive sensing. The approach adopted in this paper, along with [28]–[30] and several other references [11]–[13], [15], [16] belong to the second category. Note, however, that this paper makes significant difference to the papers in its category, as our major contribution is the generalization of the coprime sampling, where the multiple unit factor  $p$  is used to improve the degrees-of-freedom (DOFs) and spectrum resolution, and the non-overlapping factor  $q$  is used to improve the estimation accuracy. Such generalization and the related analyses are novel.

##### A. Spectrum estimation

The well-known Wiener-Khinchin theorem proves that the PSD of a signal and the covariance function form a Fourier transform pair, expressed as

$$P[f] = \sum_{\tau=-\infty}^{\infty} r[\tau] e^{-j2\pi\tau f/f_s}. \quad (23)$$

Therefore, once  $\hat{\mathbf{R}}_{\mathbf{x}}$  is reconstructed, then  $P[f]$  can be estimated by employing the discrete Fourier transform which does not require the assumptions of signal sparsity in the frequency domain. The applicability to continuous spectrum signals will be demonstrated using a simulation example in Section V.

For signals with sparse and discrete spectrum, however, we can further achieve high-resolution spectrum estimation by exploiting subspace-based spectrum estimation techniques, in lieu of the Fourier transform. As such, in the following, we focus on the spectrum estimation of sparse spectrum signals which consist of a sum of multiple sinusoids, and the corresponding CRB analysis is provided.

Assume that  $x[k]$ , for  $k = 0, 1, \dots, K-1$ , are samples of the analog signal  $X(t)$ , which can be presented as a sum of  $I$  independent frequency components

$$x[k] = \sum_{i=0}^{I-1} \sigma_i e^{-\frac{j2\pi k f_i}{f_s}} + n[k], \quad (24)$$

of frequency  $f_i$  and complex magnitudes  $\sigma_i, i = 0, \dots, I-1$ . The additive noise  $n[k]$  is assumed to be an independent and identically distributed (i.i.d.) random variable following the zero-mean complex Gaussian distribution with a variance  $\sigma_n^2$ , i.e.,  $n[k] \sim \mathcal{CN}(0, \sigma_n^2)$ .

Using the generalized coprime sampling scheme, in the  $b$ th block,  $0 \leq b \leq \tilde{B}-1$ , the received outputs at the two coprime

samplers can be respectively written as

$$\begin{aligned}\tilde{y}_{b_1}[k_1] &= x[(b-1) \times qM(M+1) + Mk_1] \\ &= \sum_{i=0}^{I-1} \sigma_i e^{\frac{-j2\pi[(b-1) \times qM(M+1) + Mk_1]f_i}{f_s}} + n_{b_1}[k_1], \quad (25) \\ \tilde{y}_{b_2}[k_2] &= x[(b-1) \times qM(M+1) + (M+1)k_2] \\ &= \sum_{i=0}^{I-1} \sigma_i e^{\frac{-j2\pi[(b-1) \times qM(M+1) + (M+1)k_2]f_i}{f_s}} + n_{b_2}[k_2], \quad (26)\end{aligned}$$

where  $0 \leq k_1 \leq p(M+1) - 1$ ,  $0 \leq k_2 \leq pM - 1$ , and the range of the pair  $(p, q)$  is given in  $\Pi_{p,q}$ . Stacking  $\tilde{y}_{b_1}[k_1]$  for  $0 \leq k_1 \leq p(M+1) - 1$  and  $\tilde{y}_{b_2}[k_2]$  for  $0 \leq k_2 \leq pM - 1$ , yields the following received vector data

$$\begin{aligned}\tilde{\mathbf{y}}_{b_1} &= \sum_{i=0}^{I-1} \mathbf{a}_{b_1}(f_i) e^{\frac{-j2\pi[(b-1) \times qM(M+1)]f_i}{f_s}} \sigma_i = \mathbf{A}_{b_1} \mathbf{s} \Phi + \mathbf{n}_{b_1}, \\ \tilde{\mathbf{y}}_{b_2} &= \sum_{i=0}^{I-1} \mathbf{a}_{b_2}(f_i) e^{\frac{-j2\pi[(b-1) \times qM(M+1)]f_i}{f_s}} \sigma_i = \mathbf{A}_{b_2} \mathbf{s} \Phi + \mathbf{n}_{b_2}, \quad (27)\end{aligned}$$

where  $\mathbf{s} = [\sigma_1, \dots, \sigma_I]^T$ ,  $\mathbf{A}_{b_1} = [\mathbf{a}_{b_1}(f_1), \dots, \mathbf{a}_{b_1}(f_I)]$ , and  $\mathbf{A}_{b_2} = [\mathbf{a}_{b_2}(f_1), \dots, \mathbf{a}_{b_2}(f_I)]$  with

$$\mathbf{a}_{b_1}(f_i) = \left[ 1, e^{\frac{-j2\pi M f_i}{f_s}}, \dots, e^{\frac{-j2\pi[p(M+1)-1]M f_i}{f_s}} \right]^T, \quad (28)$$

$$\mathbf{a}_{b_2}(f_i) = \left[ 1, e^{\frac{-j2\pi(M+1) f_i}{f_s}}, \dots, e^{\frac{-j2\pi(pM-1)(M+1) f_i}{f_s}} \right]^T. \quad (29)$$

In addition,  $\Phi$  is a diagonal matrix given by

$$\Phi = \text{diag}\left( \left[ e^{\frac{-j2\pi[(b-1) \times qM(M+1)]f_1}{f_s}}, \dots, e^{\frac{-j2\pi[(b-1) \times qM(M+1)]f_I}{f_s}} \right] \right). \quad (30)$$

Note that the noise vectors  $\mathbf{n}_{b_1}$  and  $\mathbf{n}_{b_2}$  follow the complex Gaussian distribution  $\mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{pM})$  and  $\mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{p(M+1)})$ , respectively. Then, the compressed covariance matrix  $\mathbf{R}_{\tilde{\mathbf{y}}_{12}}$  is obtained as

$$\begin{aligned}\mathbf{R}_{\tilde{\mathbf{y}}_{12}} &= \mathbb{E}[\tilde{\mathbf{y}}_{b_1} \tilde{\mathbf{y}}_{b_2}^H] = \mathbf{A}_{b_1} \mathbf{R}_{ss} \mathbf{A}_{b_2}^H + \sigma_n^2 \mathbf{i}_{\tilde{\mathbf{y}}_{12}} \\ &= \sum_{i=0}^{I-1} \sigma_i^2 \mathbf{a}_{b_1}(f_i) \mathbf{a}_{b_2}^H(f_i) + \sigma_n^2 \mathbf{i}_{\tilde{\mathbf{y}}_{12}}, \quad (31)\end{aligned}$$

where  $\mathbf{i}_{\tilde{\mathbf{y}}_{12}}$  returns a  $pM \times p(M+1)$  matrix with ones on the main diagonal and zeros elsewhere. Note that, the following vector with elements corresponding to different lags,  $\mathbf{a}(f_i) = [1, e^{\frac{-j2\pi f_i}{f_s}}, e^{\frac{-j4\pi f_i}{f_s}}, \dots, e^{\frac{-j2(L-1)\pi f_i}{f_s}}]^T$ , can be extracted based on  $\mathbf{a}_{b_1}(f_i) \otimes \mathbf{a}_{b_2}^*(f_i)$  for  $1 \leq i \leq I$ . Thus,  $\mathbf{R}_{\mathbf{x}} \in \mathbb{C}^{L \times L}$ , where  $L \leq \tilde{L}_{\max}$ , can be reconstructed and expressed as

$$\mathbf{R}_{\mathbf{x}} = \sum_{i=0}^{I-1} \sigma_i^2 \mathbf{a}(f_i) \mathbf{a}^H(f_i) + \sigma_n^2 \mathbf{I}_L. \quad (32)$$

In practice,  $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$  is estimated by averaging the available  $\tilde{B}$  blocks as in (19), and  $\hat{\mathbf{R}}_{\mathbf{x}}$  is reconstructed as in (20). The spectrum can be estimated using a variety of methods (e.g., [33]), with respect to  $\hat{\mathbf{R}}_{\mathbf{x}}$ . It is well known that subspace-based methods are popular candidates to achieve a high spectrum

resolution with a moderate computational complexity. The multiple signal classification (MUSIC) algorithm [34] is used to evaluate the performance of our approach. Note that the extension of other spectrum estimation techniques [35], [36] is straightforward.

The MUSIC approach is based on eigen-decomposition of the reconstructed covariance matrix  $\hat{\mathbf{R}}_{\mathbf{x}}$ , given by

$$\hat{\mathbf{R}}_{\mathbf{x}} = \hat{\mathbf{U}} \hat{\mathbf{\Lambda}} \hat{\mathbf{U}}^H, \quad (33)$$

where  $\hat{\mathbf{\Lambda}} = \text{diag}\{\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_L\}$  is the diagonal matrix of the eigenvalues in a descending order, and the  $L \times L$  matrix  $\hat{\mathbf{U}}$  contains the corresponding eigenvectors. The MUSIC algorithm requires the information of the rank of the signal subspace, i.e., the number of carrier frequencies of the signal arrivals. Various mathematical criteria, such as Akaike information criterion (AIC) [37], minimum description length (MDL) [38], and Bayesian information criterion (BIC) [39], can be employed to achieve the rank estimation. In this paper, we apply the BIC on  $\hat{\mathbf{R}}_{\mathbf{x}}$  to obtain the value of  $I$ . It was shown that that BIC based methods [40]–[42] generally outperform other methods, such as those developed based on AIC and MDL [43]–[45] due to the stronger consistency, particularly when the number of array sensors is large and the number of samples is small. Then, Eqn. (33) can be decomposed as

$$\hat{\mathbf{R}}_{\mathbf{x}} = \hat{\mathbf{U}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{U}}_s^H + \hat{\mathbf{U}}_n \hat{\mathbf{\Lambda}}_n \hat{\mathbf{U}}_n^H, \quad (34)$$

where  $\hat{\mathbf{U}}_s \in \mathbb{C}^{L \times I}$  and  $\hat{\mathbf{U}}_n \in \mathbb{C}^{L \times (L-I)}$  contain the signal and noise subspace sample eigenvectors, respectively, and the corresponding sample eigenvalues are included in the diagonal matrices  $\hat{\mathbf{\Lambda}}_s = \text{diag}\{\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_I\}$  and  $\hat{\mathbf{\Lambda}}_n = \text{diag}\{\hat{\lambda}_{I+1}, \hat{\lambda}_{I+2}, \dots, \hat{\lambda}_L\}$ . Then, the spectrum can be estimated as

$$\hat{P}(f) = \frac{1}{\mathbf{a}^H(f) \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H \mathbf{a}(f)}, \quad (35)$$

where  $f$  is defined as the collection over all possible grids in the spectrum and the values of  $f$  that produce peaks in the estimator  $\hat{P}(f)$  are taken as estimates of the frequencies  $f_i, i = 1, \dots, I$ . The spectrum identifiability and resolution are improved as  $L$  increases, and they are optimized when  $L = \tilde{L}_{\max}$ . This relationship is assumed in the remainder of the paper.

## B. The Cramér-Rao Bound (CRB)

The CRB offers a lower bound on the variances of unbiased estimates of the parameters. The specific CRB expressions given in [46]–[48] are valid only when the number of frequencies is less than the number of physical samples ( $I < K_s$ ). This is because the expressions are based on the inverse of the matrix  $\mathbf{A}^H \mathbf{A}$ , where  $\mathbf{A}$  is the so-called array or frequency manifold matrix. However, the assumption  $I < K_s$  is not requirement for the existence of CRB, because even when  $I \geq K_s$ , with proper prior information, the Fisher information matrix (FIM) can remain nonsingular (invertible) under a much broader range of conditions. Thus, we use the inverse of FIM as the CRB expression. After we have submitted the previous version of the manuscript, several papers have been published on the CRB analysis of the directions of arrival estimation

when more sources than the number of sensors are handled in the context of coarrays. We have cited these papers as references [49]–[51]. However, none of these papers provide revealing solutions in a compact matrix form.

For a set of vectors  $\tilde{\mathbf{y}}_b = [\tilde{\mathbf{y}}_{b_1}^T \tilde{\mathbf{y}}_{b_2}^T]^T, b = 1, \dots, \tilde{B}$ , the CRB is calculated by the well-known expression [47] involving the FIM elements

$$F_{\alpha_i \alpha_j} = \tilde{B} \text{Tr} \left\{ \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \alpha_i} \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \alpha_j} \right\}, \quad (36)$$

for unknown variables  $\alpha_i$  and  $\alpha_j$ , where  $\mathbf{R}_{\tilde{\mathbf{y}}}$  is expressed as

$$\mathbf{R}_{\tilde{\mathbf{y}}} = \text{E}[\tilde{\mathbf{y}}_b \tilde{\mathbf{y}}_b^H] = \sum_{i=0}^{I-1} \sigma_i^2 \mathbf{a}_b(f_i) \mathbf{a}_b^H(f_i) + \sigma_n^2 \mathbf{I}_{p(2M+1)}, \quad (37)$$

and  $\mathbf{a}_b(f_i) = [\mathbf{a}_{b_1}^T(f_i) \mathbf{a}_{b_2}^T(f_i)]^T$ .

In the underlying case, the unknown parameters are the  $I$  signal frequencies  $f_i$  and powers  $\sigma_i^2$  for  $i = 1, \dots, I$ , as well as the noise power  $\sigma_n^2$ . Therefore, the elements in the  $(2I + 1) \times (2I + 1)$  Fisher matrix  $\mathbf{F}$  can be written in terms of the block matrices, for  $i, j = 1, \dots, I$ , given by

$$\begin{aligned} F_{i,j} &= \tilde{B} \text{Tr} \left\{ \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial f_i} \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial f_j} \right\}, \\ F_{i,j+I} &= \tilde{B} \text{Tr} \left\{ \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial f_i} \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_j^2} \right\}, \\ F_{i,2I+1} &= \tilde{B} \text{Tr} \left\{ \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial f_i} \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_n^2} \right\}, \\ F_{i+I,j} &= \tilde{B} \text{Tr} \left\{ \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_i^2} \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial f_j} \right\}, \\ F_{i+I,j+I} &= \tilde{B} \text{Tr} \left\{ \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_i^2} \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_j^2} \right\}, \\ F_{i+I,2I+1} &= \tilde{B} \text{Tr} \left\{ \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_i^2} \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_n^2} \right\}, \\ F_{2I+1,2I+1} &= \tilde{B} \text{Tr} \left\{ \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_n^2} \mathbf{R}_{\tilde{\mathbf{y}}}^{-1} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_n^2} \right\}, \end{aligned} \quad (38)$$

where

$$\begin{aligned} \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial f_i} &= \sigma_i^2 \left[ \frac{\partial \mathbf{a}_b(f_i)}{\partial f_i} \mathbf{a}_b^H(f_i) + \mathbf{a}_b(f_i) \frac{\partial \mathbf{a}_b^H(f_i)}{\partial f_i} \right], \\ \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_i^2} &= \mathbf{a}_b(f_i) \mathbf{a}_b^H(f_i), \\ \frac{\partial \mathbf{R}_{\tilde{\mathbf{y}}}}{\partial \sigma_n^2} &= \mathbf{I}_{p(2M+1)}. \end{aligned} \quad (39)$$

Then, the CRB of estimated frequencies is obtained as

$$\text{CRB}(f_i) = [\mathbf{F}^{-1}]_{i,i}. \quad (40)$$

## V. SIMULATION RESULTS

For illustrative purposes, we demonstrate the spectrum estimation performance under different choices of the arguments within the generalized coprime sampling scheme. Assume that  $I$  frequency components with identical powers are distributed in the frequency band  $[-500, 500]$  MHz. Assume that  $K = 50000$  samples are generated with a Nyquist sampling rate  $f_s = 1$  GHz. In addition, the noise power is assumed to be

identical across the entire spectrum. The MUSIC method is used to estimate the power spectrum. Our benchmarks are the spectrum DOFs and their statistical performance. The latter is evaluated in terms of average relative root mean square error (RMSE) of the estimated frequencies, defined as

$$\text{Relative RMSE}(f_i) = \frac{1}{f_s} \sqrt{\frac{1}{500I} \sum_{n=1}^{500} \sum_{i=1}^I (\hat{f}_i(n) - f_i)^2}, \quad (41)$$

where  $\hat{f}_i(n)$  is the estimate of  $f_i$  from the  $n$ th Monte Carlo trial,  $n = 1, \dots, 500$ .

### A. The performance of coprime sampling

We first illustrate the performance of coprime sampling. Herein, the conventional coprime sampling scheme is considered, i.e.,  $p = 2$ . In addition,  $M = 3$  is assumed. As such, the  $L \times L = 19 \times 19$  covariance matrix  $\hat{\mathbf{R}}_{\mathbf{x}}$  can be reconstructed from  $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$  with dimension  $pM \times p(M+1) = 6 \times 8$ . Thus, the resulting compression factor is  $\kappa_{\max} \approx 7.52$  and up to  $L - 1 = 18$  frequencies can be estimated.

In Fig. 8, we consider  $I = 18$  frequencies with  $\delta_f = 50$  MHz separation in the presence of noise with a 0 dB SNR. It is evident that all 18 frequencies can be identified correctly. In Fig. 9, the RMSE results are shown as a function of the input SNR, where  $I = 1$  is assumed. As expected, it displays a strong inverse semi-logarithmic dependence on the input SNR. It is also observed that there is a gap between the RMSE and CRB even in the high SNR region, due to estimation bias. The errors are mainly generated in two aspects. On one hand,  $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$  is used to reconstruct  $\hat{\mathbf{R}}_{\mathbf{x}}$ . On the other hand, only consecutive lag entries in  $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$  are exploited. It is observed that the bias errors increase with  $I$  due to a higher frequency components, as shown in Fig. 10, where the input SNR is set to 0 dB.

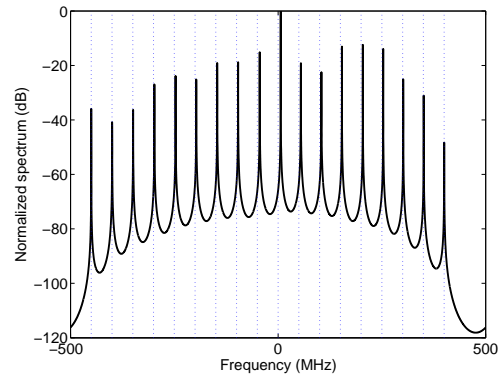


Fig. 8. Estimated spectrum ( $I = 18$  and input SNR=0 dB).

### B. The generalized coprime sampling scheme versus other schemes

Next, we compare the generalized coprime sampling scheme with the nested sampler and the sparse ruler based sampler, where the same number of physical samples is assumed. For the coprime sampler, we set  $M = 3$ , and thus there are



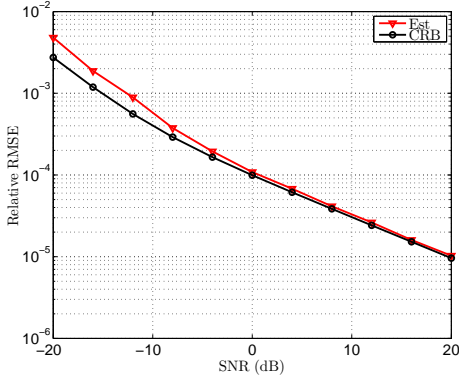


Fig. 9. Relative RMSE versus SNR ( $I = 1$ ).

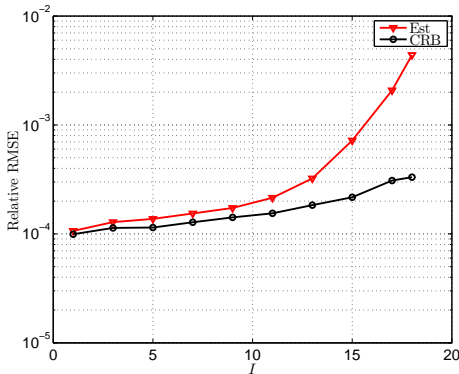


Fig. 10. Relative RMSE versus  $I$  (SNR=0 dB).

$2M + 1 = 7$  physical samples in each coprime unit. The sampling patterns corresponding to the nested sampler and the sparse ruler based sampler that yield the same 7-sample unit are  $[0 \ 1 \ 2 \ 3 \ 7 \ 11 \ 15]$  and  $[0 \ 1 \ 4 \ 10 \ 12 \ 15 \ 17]$ , respectively. In this simulation,  $p = 3$  coprime units are used to form the covariance matrix for the generalized coprime scheme, whereas the nested sampler and minimal sparse ruler based sampler each uses one unit as in [15] and [12]. Their relative RMSEs are depicted as a function of input SNR in Fig. 11, where  $I = 5$  frequencies are considered. It is clear that the generalized coprime scheme outperform the other two sampling schemes due to the higher number of DOFs and improved resolution.

### C. Relative RMSE for various $p$

In Figs. 12–14, we compare the performance corresponding to different choices of  $p$  under different criteria, where non-overlapping segmentation is used.

Figs. 12 and Fig. 13 examine the performance for different choices of  $p$ , based on the same compression factor, where  $M = 3$  is assumed. In Fig. 12, the distinction on spectrum identifiability is depicted for the cases of  $p = 10$  and  $p = 45$ . We consider  $I = 100$  frequencies with  $\delta_f = 2$  MHz separation in the presence of noise with a 0 dB SNR. It is evident that only the scenario of  $p = 45$  can resolve all frequencies correctly, although in the case of  $p = 10$  the number of

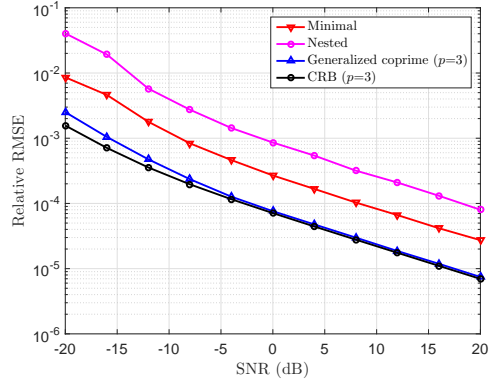


Fig. 11. Relative RMSE versus SNR for different sampling schemes ( $I=5$ ).

DOFs  $L - 1 = 114$ , is slightly higher than the number of frequency components. Fig. 13 presents the RMSE and CRB with respect to  $p$ , where  $I = 5$  is assumed. It is observed that the estimation performance is improved as  $p$  increases. In addition, the bias error between the estimated frequency and the CRB becomes smaller, since the ratio between the number of consecutive lags and the number of total lags in  $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$  increases with  $p$ . In summary, a higher value of  $p$  can improve DOFs and spectrum estimation performance under the same compression factor. However, the requirement of storage space and the computational load become higher, due to the resulting higher value of  $L$ .

In Fig. 14, we present the relative RMSE as a function of the input SNR for different values of  $(p, M)$  pairs, where the dimension of the covariance matrix is  $L = (p - 1)M^2 + (p + 1)M + 1 = 161$ , and the number of frequencies is  $I = 5$ . It is clear that, as the value of  $M$  decreases (and so does the compression factor  $\kappa_{\max}$  because  $\kappa_{\max} \propto M^2$ ), the estimated relative RMSE is reduced since a higher number of physical samples ( $K_s = K(1/M + 1/(M + 1))$ ) can be used.

### D. Relative RMSE for various $q$

Finally, the advantage of utilization of overlapping blocks is demonstrated in Fig. 15, where  $M = 3$  and  $p = 12$  are assumed and  $I = 5$  frequency is considered with a 0 dB SNR. In addition,  $q$  is chosen within the range of  $\{1, 2, 3, 4, 6, 12\}$ . It is evident that the estimation performance can be improved as  $q$  decreases, compared to the non-overlapping case, i.e.,  $q = p = 12$ .

### E. Relative RMSE versus $K$

In Fig. 16, we present the relative RMSE performance with respect to  $K$ , where  $M=3$  is assumed, and  $I=5$  frequencies with a 0 dB input SNR are considered. It is evident that the estimated relative RMSE performance is improved as  $K$  increases because a higher number of blocks is used to reduce the noise effect. Asymptotically, when  $K$  is large, the relative RMSE asymptotically decreases with a factor of  $1/\sqrt{K}$ . In addition, various cases with different values of  $p$  and  $q$  are compared in this figure. By assuming a large value of  $p$  and a small value of  $q$ , the generalized coprime sampling

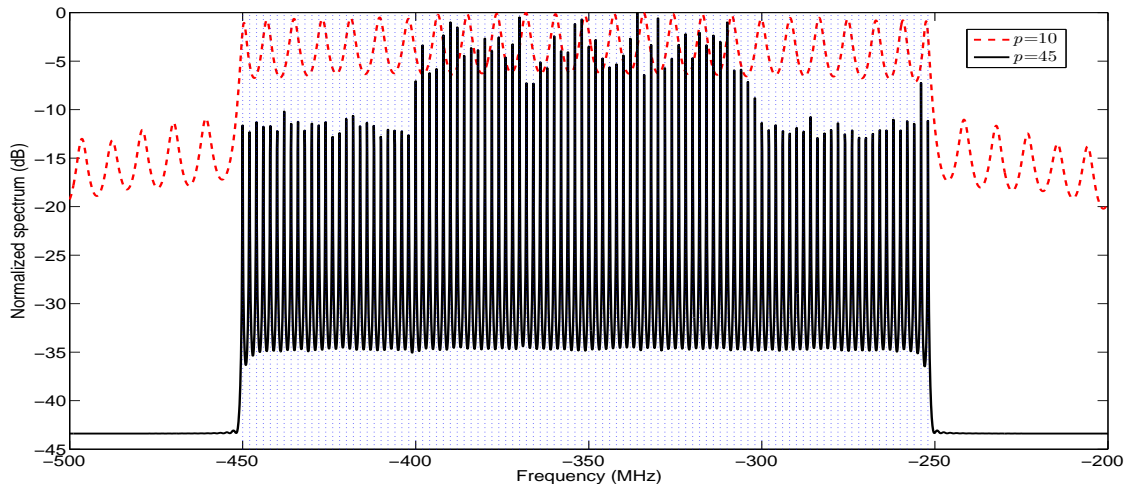


Fig. 12. Estimated spectra for the cases of  $p = 10$  and  $p = 45$  ( $M = 3$  and input SNR=0 dB).

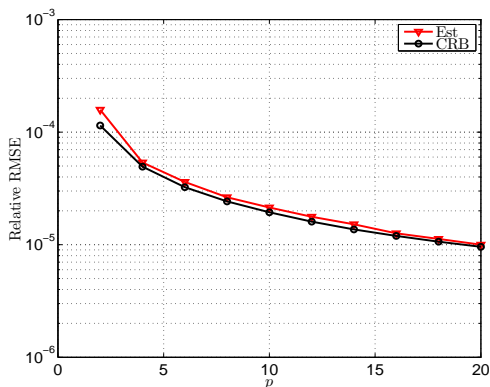


Fig. 13. Relative RMSE versus  $p$ , based on the same  $M$  ( $I = 5$  and  $M = 3$ ).

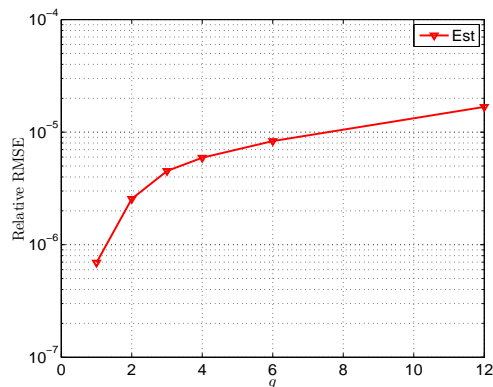


Fig. 15. Relative RMSE versus  $q$  ( $M = 3$ ,  $p = 12$ , and  $I = 5$ ).

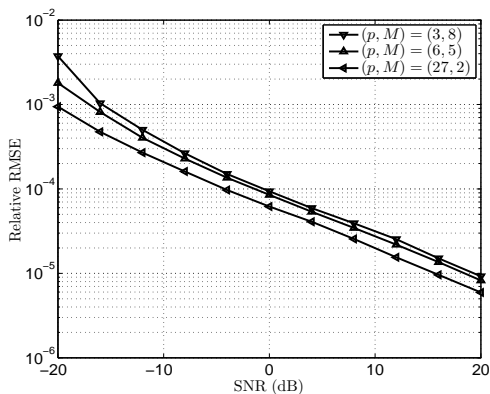


Fig. 14. Relative RMSE versus SNR, based on the same  $L$  ( $I = 5$  and  $L = 161$ ).

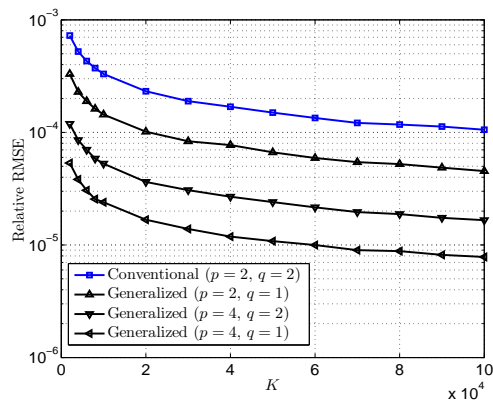


Fig. 16. Relative RMSE versus  $K$  ( $M=3$  and  $I=5$ ).

scheme improves the RMSE performance as it benefits from the high dimension of the reconstructed covariance matrix and the utilization of overlapping blocks, respectively.

#### F. Estimation for continuous spectrum

Finally, we consider an example of continuous spectrum signals in Fig. 17, where  $x(t)$  is assumed to have continuous rectangular spectrum supports in  $[-350, -230]$  MHz and  $[150, 280]$  MHz. Multiple coprime unit factors of  $p = 2, 3, 7$

are considered. As  $p$  increases, it is clear that the mainlobe becomes closer to the ideal signal bandwidth due to the larger dimension of the reconstructed matrix  $\hat{\mathbf{R}}_{\mathbf{x}}$ . For comparison, the case of  $p = 7$  and  $q = 1$  generally outperforms the case of  $p = 7$  and  $q = 7$  because a higher number of blocks, achieved by using overlapping segmentation, become available for averaging.

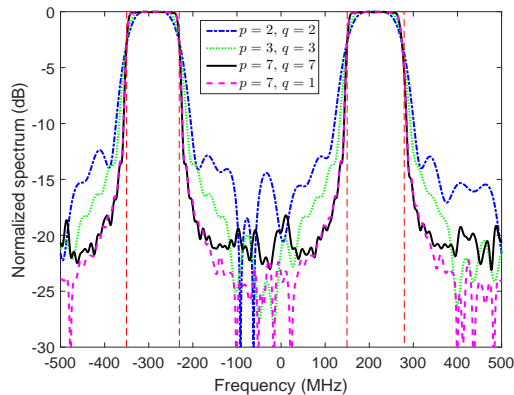


Fig. 17. Estimated spectrum.

## VI. CONCLUSIONS

We proposed an effective approach to compressively sample wide-sense stationary processes. The coprime sampling matrix was used to obtain a compressed representation for their second-order statistics. Using a fixed number of data, different schemes for the acquisition of a covariance matrix were presented, based on segmenting the data sequence. The performance of these schemes was compared and numerically evaluated. The effectiveness of the proposed technique was evidently verified using simulation results.

## VII. APPENDIX

### Proof of Proposition 1

For the convenience of presentation, we define the function  $\Gamma([k_{1\min}, k_{1\max}], [k_{2\min}, k_{2\max}])$  as the operation  $\pm(Mk_1 - Nk_2)$  with  $k_1 \in [k_{1\min}, k_{1\max}]$  and  $k_2 \in [k_{2\min}, k_{2\max}]$ . Denote

$$\tilde{\mathbb{L}}_1 = \{\tau_1 | \Gamma([0, pN - 1], [0, M - 1])\}, \quad (42)$$

$$\tilde{\mathbb{L}}_2 = \{\tau_2 | \Gamma([0, N - 1], [0, pM - 1])\}, \quad (43)$$

and the proposition 1 can be obtained by proving the following propositions:

- 1(a)  $\tilde{\mathbb{L}} = \tilde{\mathbb{L}}_1 \cup \tilde{\mathbb{L}}_2$ .
- 1(b) For the set  $\tilde{\mathbb{L}}_1$ , it contains all integer lags in the range  $-(p-1)MN - N + 1 \leq \tau_1 \leq (p-1)MN + N - 1$ , and the ‘‘holes’’ are located at  $\pm[(p-1)MN + aM + bN]$ , where  $a \geq 0$  and  $b > 0$  are integers.
- 1(c) For the set  $\tilde{\mathbb{L}}_2$ , it contains all integer lags in the range  $-(p-1)MN - M + 1 \leq \tau_2 \leq (p-1)MN + M - 1$ , and the ‘‘holes’’ are located at  $\pm[(p-1)MN + aM + bN]$ , where  $a > 0$  and  $b \geq 0$  are integers.
- 1(d) The first pair of holes  $\pm[(p-1)MN + bN]$  in  $\tilde{\mathbb{L}}_1$ , where  $b \geq 0$ , can be aligned by the non-consecutive element in  $\tilde{\mathbb{L}}_2$ .

*Proof of proposition 1(a):* The lag set

$$\begin{aligned} \tilde{\mathbb{L}} &= \{\Gamma([0, pN - 1], [0, pM - 1])\}, \\ &= \bigcup_{p_1=1}^p \{\Gamma([0, pN - 1], [(p_1 - 1)M, p_1M - 1])\} \cup \\ &\quad \bigcup_{p_2=1}^p \{\Gamma([(p_2 - 1)N, p_2N - 1], [0, pM - 1])\} \\ &= \tilde{\mathbb{L}}_1 \cup \left( \bigcup_{p_1=2}^p \{\Gamma([0, pN - 1], [(p_1 - 1)M, p_1M - 1])\} \right) \cup \\ &\quad \tilde{\mathbb{L}}_2 \cup \left( \bigcup_{p_2=2}^p \{\Gamma([(p_2 - 1)N, p_2N - 1], [0, pM - 1])\} \right). \end{aligned} \quad (44)$$

Note that the union of the sets  $\bigcup_{p_1=2}^p \{\Gamma([0, pN - 1], [(p_1 - 1)M, p_1M - 1])\}$  and  $\bigcup_{p_2=2}^p \{\Gamma([(p_2 - 1)N, p_2N - 1], [0, pM - 1])\}$  is the subset of  $\tilde{\mathbb{L}}_1 \cup \tilde{\mathbb{L}}_2$ . Therefore, (44) can be simplified as

$$\tilde{\mathbb{L}} = \tilde{\mathbb{L}}_1 \cup \tilde{\mathbb{L}}_2. \quad (45)$$

*Proof of proposition 1(b):* Given any integer  $\tau_1$  satisfying

$$0 \leq \tau_1 \leq (p-1)MN + N - 1, \quad (46)$$

we need to prove that there exist integers  $k_1 \in [0, pN - 1]$  and  $k_2 \in [0, M - 1]$  such that  $\tau_1 = Mk_1 - Nk_2$  holds. The requirement  $k_2 \in [0, M - 1]$  is equivalent to

$$0 \leq Nk_2 \leq MN - N. \quad (47)$$

Because  $Mk_1 = \tau_1 + Nk_2$ , we obtain the following relationship by combining (46) and (47),

$$0 \leq Mk_1 \leq pMN - 1 < pMN. \quad (48)$$

This result can be equivalently expressed as  $0 \leq k_1 < pN$ . Because  $k_1$  is an integer, this requirement is equivalent to

$$0 \leq k_1 \leq pN - 1, \quad (49)$$

which is satisfied in the underlying coprime array configuration.

Next, we prove the hole positions by contradiction. We suppose  $Mk_1 - Nk_2 = (p-1)MN + aM + bN$  holds for some integers  $k_1 \in [0, pN - 1]$  and  $k_2 \in [0, M - 1]$ , where  $a \geq 0$  and  $b > 0$  are integers, then relationship

$$\frac{M}{N} = \frac{k_2 - M + b}{k_1 - pN - a} \quad (50)$$

must be valid. From  $k_1 \in [0, pN - 1]$  and  $a \geq 0$ , we find  $Mk_1 - Nk_2 = (p-1)MN + aM + bN < pMN$ , and then  $b < M$ . As a result,  $|k_2 - M + b| < M$ . Due to the coprimality between  $M$  and  $N$ , we cannot find a  $k_1$  to satisfy (50). Therefore,  $Mk_1 - Nk_2 \neq (p-1)MN + aM + bN$ , i.e., there are holes at  $(p-1)MN + aM + bN$  in  $\tilde{\mathbb{L}}_1$ .

Due to the symmetry of  $\tilde{\mathbb{L}}_1$ , we can draw the conclusions that  $\tilde{\mathbb{L}}_1$  all integer lags in the range  $-(p-1)MN - N + 1 \leq \tau_1 \leq (p-1)MN + N - 1$ , and the ‘‘holes’’ are located at  $\pm[(p-1)MN + aM + bN]$ , where  $a \geq 0$  and  $b > 0$  are integers.

*Proof of proposition 1(c):* We omit the proof of proposition 1(c), which can be proved by using the same method as in the proof of proposition 1(b).

*Proof of proposition 1(d):* Based on the proposition 1(b), there are holes  $(p-1)MN + aM + bN$  in  $\tilde{\mathbb{L}}_1$ , where  $a \geq 0$  and  $b > 0$  are integers. If the holes are aligned by the elements in  $\tilde{\mathbb{L}}_2$ , the following relationship

$$(p-1)MN + aM + bN = \pm(Mk_1 - Nk_2) \quad (51)$$

must be valid for  $k_1 \in [0, N-1]$  and  $k_2 \in [0, pM-1]$ . The requirement is equivalent to

$$(p-1)MN + aM + (b+k_2)N = Mk_1,$$

or

$$(p-1)MN + (a+k_1)M + bN = Nk_2,$$

i.e.,

$$b = -k_2, \quad \text{or} \quad a = -k_1. \quad (52)$$

It is only possible for  $a = k_1 = 0$  when  $k_1 \in [0, N-1]$ ,  $k_2 \in [0, pM-1]$ ,  $a \in [0, \infty)$ , and  $b \in (0, \infty)$ . Then, the requirement further becomes

$$(p-1)M + b = k_2. \quad (53)$$

In the proof of proposition 1(b), it is shown that  $b < M$ , i.e.,  $b \leq M-1$ . As such,  $k_2 \in ((p-1)M, pM-1] \subseteq [0, pM-1]$ . Therefore, the holes  $(p-1)MN + bN$  ( $a = 0$ ) in  $\tilde{\mathbb{L}}_1$  are aligned by the element in  $\tilde{\mathbb{L}}_2$  for some integers  $k_2 \in [0, pM-1]$ . As a result, the first hole outside the consecutive range of  $\tilde{\mathbb{L}}$  becomes  $(p-1)MN + M + N$ . Then, the set  $\tilde{\mathbb{L}}$  contains all integer lags in the range

$$-(p-1)MN - M - N + 1 \leq \tau \leq (p-1)MN + M + N - 1. \quad (54)$$

## REFERENCES

- [1] S. Qin, Y. D. Zhang, and M. G. Amin, "High-resolution frequency estimation using generalized coprime sampling," in *Proc. SPIE Mobile, Multimedia/Image Process., Secur. Appl. Conf. (SPIE)*, Baltimore, MD, 2015, vol. 9497, pp. 94970K1–94970K7.
- [2] S. Qin, Y. D. Zhang, M. G. Amin, and A. M. Zoubir, "Generalized coprime sampling of Toeplitz matrices," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Shanghai, China, 2016, pp. 4468–4472.
- [3] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [4] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [5] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [6] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. R. Stat. Soc., Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.
- [7] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 92–106, 2009.
- [8] Q. Wu, Y. D. Zhang, M. G. Amin, and B. Himed, "Multi-task Bayesian compressive sensing exploiting intra-task dependency," *IEEE Signal Process. Lett.*, vol. 22, no. 4, pp. 430–434, 2015.
- [9] G. Dasarathy, P. Shah, B. N. Bhaskar, and R. Nowak, "Sketching sparse matrices, covariances, and graphs via tensor products," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1373–1388, 2015.
- [10] Y. Chen, Y. Chi, and A. Goldsmith, "Exact and stable covariance estimation from quadratic sampling via convex programming," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 4034–4059, 2015.
- [11] G. Leus and Z. Tian, "Recovering second-order statistics from compressive measurements," in *Proc. IEEE Int. Workshop on Comp. Adv. in Multi-Sensor Adaptive Process. (CAMSAP)*, San Juan, Puerto Rico, 2011, pp. 337–340.
- [12] D. D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4775–4789, 2012.
- [13] D. Romero and G. Leus, "Compressive covariance sampling," in *Proc. Inf. Theory Appl. Workshop (ITA)*, San Diego, CA, 2013, pp. 1–8.
- [14] C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Process. Mag.*, vol. 10, no. 3, pp. 10–37, 1993.
- [15] H. Qiao and P. Pal, "Generalized nested sampling for compression and exact recovery of symmetric Toeplitz matrices," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Atlanta, GA, 2014, pp. 443–447.
- [16] Z. Tian, Y. Tafesse, and B. M. Sadler, "Cyclic feature detection with sub-Nyquist sampling for wideband spectrum sensing," *IEEE J. Sel. Top. Signal Process.*, vol. 6, no. 1, pp. 58–69, 2012.
- [17] 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.
- [18] 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.
- [19] M. A. Lexa, M. E. Davis, J. S. Thompson, and J. Nikolic, "Compressive power spectral density estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Prague, Czech Republic, 2011, pp. 3884–3887.
- [20] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Taipei, Taiwan, 2009, pp. 2337–2340.
- [21] D. D. Ariananda, G. Leus, and Z. Tian, "Multi-coset sampling for power spectrum blind sensing," in *Proc. Int. Conf. Digit. Signal Process. (DSP)*, Corfu, Greece, 2011, pp. 1–8.
- [22] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: a method based on time-averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, 1967.
- [23] 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, 2015.
- [24] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, 2007.
- [25] M. Mishali and Y. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Top. Signal Process.*, vol. 4, no. 2, pp. 375–391, 2010.
- [26] R. Venkataramani and Y. Bresler, "Perfect reconstruction formulas and bound on aliasing error in sub-Nyquist nonuniform sampling of multiband signals," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2173–2183, 2000.
- [27] M. Mishali and Y. Eldar, "Blind multiband signal reconstruc-

- tion: Compressed sensing for analog signals,” *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, 2009.
- [28] H. Sun, W. Y. Chiu, J. Jiang, A. Nallanathan and H. V. Poor, “Wideband spectrum sensing with sub-Nyquist sampling in cognitive radios,” *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 6068–6073, 2012.
- [29] D. Cohen and Y. C. Eldar, “Sub-Nyquist sampling for power spectrum sensing in cognitive radios: A unified approach,” *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3897–3910, 2014.
- [30] M. Shaghghi and S. A. Vorobyov, “Finite-length and asymptotic analysis of averaged correlogram for undersampled data,” *Appl. Comput. Harmon. Anal.*, <http://dx.doi.org/10.1016/j.acha.2016.02.001>.
- [31] K. Adhikari, J. R. Buck and K. E. Wage, “Extending coprime sensor arrays to achieve the peak side lobe height of a full uniform linear array,” *EURASIP J. Wireless Commun. Netw.*, doi:10.1186/1687–6180–2014–148, 2014.
- [32] P. Stoica and A. Nehorai, “MUSIC, maximum likelihood, and Cramér-Rao bound,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 5, pp. 720–741, 1989.
- [33] P. Stoica and R. L. Moses, *Spectrum Analysis of Signals*, Upper Saddle River, NJ: Prentice-Hall, 2005.
- [34] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.
- [35] 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, 1989.
- [36] 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, 1990.
- [37] H. Akaike, “A new look at the statistical model identification,” *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [38] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 33, no. 2, pp. 387–392, 1985.
- [39] G. Schwarz, “Estimating the dimension of a model,” *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, 1978.
- [40] Z. Lu and A. M. Zoubir, “Generalized Bayesian information criterion for source enumeration in array processing,” *IEEE Trans. Signal Process.*, vol. 61, no. 6, pp. 1470–1480, 2013.
- [41] Z. Lu and A. M. Zoubir, “Source enumeration in array processing using a two-step test,” *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2718–2727, 2015.
- [42] L. Huang, Y. Xiao, K. Liu, H. C. So, and J.-K. Zhang, “Bayesian information criterion for source enumeration in large-scale adaptive antenna array,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3018–3032, 2016.
- [43] 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.
- [44] C. D. Giurcaneanu, S. A. Razavi, and A. Liski, “Variable selection in linear regression: Several approaches based on normalized maximum likelihood,” *Signal Process.*, vol. 91, no. 8, pp. 1671–1692, 2011.
- [45] L. Huang and H. C. So, “Source enumeration via MDL criterion based on linear shrinkage estimation of noise subspace covariance matrix,” *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4806–4821, 2013.
- [46] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York: Wiley, 2002.
- [47] P. Stoica and A. Nehorai, “Performance study of conditional and unconditional direction-of-arrival estimation,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 10, pp. 1783–1795, 1990.
- [48] M. Shaghghi and S. A. Vorobyov, “Cramer-Rao bound for sparse signals fitting the low-rank model with small number of parameters,” *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1497–1501, 2015.
- [49] C.-L. Liu and P. P. Vaidyanathan, “Cramer-Rao bounds for coprime and other sparse arrays, which find more sources than sensors,” *Digital Signal Process.*, doi: 10.1016/j.dsp.2016.04.011, 2016.
- [50] M. Wang and A. Nehorai, “Coarrays, MUSIC, and the Cramer-Rao bound,” arXiv:1605.03620, 2016.
- [51] A. Koochakzadeh and P. Pal, “Cramer-Rao bounds for underdetermined source localization,” *IEEE Signal Process. Lett.*, vol. 23, no. 7, pp. 919–923, 2016.



**Si Qin** received the B.S. and M.S. degrees in Electrical Engineering from Nanjing University of Science and Technology, Nanjing, China, in 2010 and 2013, respectively. Currently, he is a Research Assistant at the Center for Advanced Communications, Villanova University, Villanova, PA, working toward his Ph.D. degree in Electrical Engineering. His research interests include direction-of-arrival estimation, sparse array and signal processing, and radar signal processing.

Mr. Si Qin received the First Prize of Student Paper Competition at 2014 IEEE Benjamin Franklin Symposium on Microwave and Antenna Sub-systems (BenMAS) and Best Student Paper Award at 2012 IEEE International Conference on Microwave and Millimeter Wave Technology (ICMMT).



**Yimin D. Zhang** (SM'01) received his Ph.D. degree from the University of Tsukuba, Tsukuba, Japan, in 1988.

Dr. Zhang joined the faculty of the Department of Radio Engineering, Southeast University, Nanjing, China, in 1988. He served as a Director and Technical Manager at the Oriental Science Laboratory, Yokohama, Japan, from 1989 to 1995, a Senior Technical Manager at the Communication Laboratory Japan, Kawasaki, Japan, from 1995 to 1997, and a Visiting Researcher at the ATR Adaptive Communications Research Laboratories, Kyoto, Japan, from 1997 to 1998. He was with the Villanova University, Villanova, PA, USA, from 1998 to 2015, where he was a Research Professor with the Center for Advanced Communications. Since 2015, he has been with the Department of Electrical and Computer Engineering, College of Engineering, Temple University, Philadelphia, PA, USA, where he is currently an Associate Professor. His general research interests lie in the areas of statistical signal and array processing applied for radar, communications, and navigation, including compressive sensing, convex optimization, time-frequency analysis, MIMO system, radar imaging, target localization and tracking, wireless networks, and jammer suppression. He has published more than 300 journal articles and conference papers and 12 book chapters. Dr. Zhang received the 2016 Premium Award from the Institution of Engineering and Technology (IET) for Best Paper in IET Radar, Sonar & Navigation.

Dr. Zhang is a member of the Sensor Array and Multichannel Technical Committee of the IEEE Signal Processing Society. He is an Associate Editor for the *IEEE Transactions on Signal Processing*, and serves on the Editorial Board of the *Signal Processing* journal. He was an Associate Editor for the *IEEE Signal Processing Letters* during 2006–2010, and an Associate Editor for the *Journal of the Franklin Institute* during 2007–2013.



**Moeness G. Amin** (F'01) received the Ph.D. degree in electrical engineering from the University of Colorado, Boulder, CO, USA, in 1984. Since 1985, he has been with the Faculty of the Department of Electrical and Computer Engineering, Villanova University, Villanova, PA, USA, where he became the Director of the Center for Advanced Communications, College of Engineering, in 2002. Dr. Amin is a Fellow of the Institute of Electrical and Electronics Engineers; Fellow of the International Society of Optical Engineering; Fellow of

the Institute of Engineering and Technology (IET), and a Fellow of the European Association for Signal Processing (EURASIP). Dr. Amin is the Recipient of the 2016 Alexander von Humboldt Research Award, the 2016 IET Achievement Medal, the 2014 IEEE Signal Processing Society Technical Achievement Award, the 2009 Individual Technical Achievement Award from the European Association for Signal Processing, the 2015 IEEE Aerospace and Electronic Systems Society Warren D. White Award for Excellence in Radar Engineering, the IEEE Third Millennium Medal, the 2010 NATO Scientific Achievement Award, and the 2010 Chief of Naval Research Challenge Award. He is the recipient of the 1997 Villanova University Outstanding Faculty Research Award and the 1997 IEEE Philadelphia Section Award. He was a Distinguished Lecturer of the IEEE Signal Processing Society, 2003-2004, and is presently the Chair of the Electrical Cluster of the Franklin Institute Committee on Science and the Arts. Dr. Amin has over 700 journal and conference publications in signal processing theory and applications. He co-authored 20 book chapters and is the Editor of the three books *Through the Wall Radar Imaging*, *Compressive Sensing for Urban Radar*, and *Radar for Indoor Monitoring* published by CRC Press in 2011, 2014, and 2017, respectively.



**Abdelhak M. Zoubir** (F'08) is a Fellow of the IEEE and IEEE Distinguished Lecturer (Class 2010–2011). He received his Dr.-Ing. from Ruhr-Universität Bochum, Germany, in 1992. He was with Queensland University of Technology, Australia from 1992–1998 where he was Associate Professor. In 1999, he joined Curtin University of Technology, Australia as a Professor of Telecommunications. In 2003, he moved to Technische Universität Darmstadt, Germany as Professor of Signal Processing and Head of the Signal Processing Group. His

research interest lies in statistical methods for signal processing with emphasis on bootstrap techniques, robust detection and estimation and array processing applied to telecommunications, radar, sonar, automotive monitoring and safety, and biomedicine. He published over 400 journal and conference papers on these areas. Professor Zoubir served as General Chair and Technical Chair of numerous international IEEE conferences, most recently he was the Technical Co-Chair of ICASSP-14 held in Florence, Italy. Dr. Zoubir also served on publication boards of various journals and he served a three-year term as Editor-In-Chief of the IEEE Signal Processing Magazine (2012–2014). He was Chair (2010–2011), Vice-Chair (2008–2009) and Member (2002–2007) of the IEEE SPS Technical Committee Signal Processing Theory and Methods (SPTM). He was a Member of the IEEE SPS Technical Committee Sensor Array and Multi-channel Signal Processing (SAM) from 2007 until 2012. He also serves on the Board of Directors of the European Association of Signal Processing (EURASIP) (2009–2016) and on the Board of Governors of the IEEE SPS (2015–2017). He is the President-Elect of EURASIP, starting his term in January 2017.