GENERALIZED COPRIME SAMPLING OF TOEPLITZ MATRICES

Si Qin[†]*, Yimin D. Zhang*[‡]*, Moeness G. Amin*[†]*, Abdelhak Zoubir*[♯]

[†] Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA
[‡] Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122, USA
[♯] Signal Processing Group, Technische Universität Darmstadt, Darmstadt, Germany

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. In this paper, we propose a general coprime sampling concept that implements effective compression of Toeplitz covariance matrices. Given a fixed number of data samples, we examine different schemes on covariance matrix acquisition, based on segmented data sequences. The effectiveness of the proposed technique is verified using simulation results.

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

1. INTRODUCTION

Various applications require spectrum sensing over a broad frequency band, which demands a high sampling rate and yields a large volume of data. In some of these applications, the signal is known to be sparse. This property allows the exploitation of compressive sensing and sparse sampling approaches that enable effective sparse signal reconstruction [1, 2], with no loss of information.

In addition to such applications that require signal reconstruction, there is also an important class of applications, e.g., spectrum estimation, where only the second-order statistics are interested. In this case, the covariance function and the covariance matrix can be constructed as low-dimensional representations of high-dimensional signals [3, 4]. This fact motivated the development of an alternative framework, referred to as compressive covariance sampling, in which the signal sparsity is not a requirement [5–8].

In this paper, we focus on spectrum estimation of wide-sense stationary (WSS) processes, utilizing the Toeplitz property of the covariance matrix. Several methods have been developed to tackle similar compressive Toeplitz matrix sampling. For example, a minimal sparse sampler and a generalized nested sampler were respectively proposed in [9] and [10] to recover Toeplitz matrices from a compressed covariance matrix. However, these papers assume an infinite number of data samples, whereas the achievable performance of the reconstruction under a finite number of data samples is not considered. In addition, their minimum sampling interval still needs to be Nyquist. As such, the methods developed therein are infeasible to implement low sampling rate systems for wideband spectrum estimation.

The recent proposed coprime sampling [11] is attractive because it achieves a much lower sampling rate and thus facilitates low-cost implementations. In this paper, our focus is on the effective estimation of the Toeplitz covariance matrix and signal spectrum from a finite number of samples of a WSS sequence. We generalize the coprime sampling approach to achieve a higher number of degrees-of-freedom (DOFs) and lower estimation error. Different schemes for acquisition of covariance matrix entries are analytically derived for quantitative evaluation, comparison, and optimal design.

Notations: We use lower-case (upper-case) bold characters to denote vectors (matrices). In particular, $\mathbf{I}_N$ stands for 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. $\mathbb{R}$ denotes the set of real values, while $\mathbb{N}^+$ represents the set of positive integers. Also, $\mathrm{E}(\cdot)$ is the statistical expectation operator, and $\lfloor \cdot \rfloor$ denotes the floor function which returns the largest integer not exceeding the argument.

2. 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 power spectral density (PSD), the covariance matrix needs to be obtained from a specific realization of $X(t)$, $t = 0, \ldots, T-1$. 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 $\mathbf{x}_L[l] = [x[l], x[l+1], \ldots, x[l+L-1]]^T$. Then, the resulting semi-positive definite and Hermitian covariance matrix can be given by

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

This work was supported in part by the Office of Naval Research (ONR) under grant N00014-13-1-0061.

in which the entry $r[\tau] = \mathrm{E}\left[x[l]x^*[l-\tau]\right]$ only depends on the lags $\tau = -L+1, \ldots, L-1$. 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.

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 $x[k], k = 0, \ldots, K-1$. A common practice for covariance matrix estimation is to segment the entire discrete-time observation of length K into multiple length-L blocks, and the covariance matrix is estimated by averaging the sample covariance over different blocks [12]. As shown in Fig. 1, we segment the entire observation period into multiple, possibly overlapping blocks. In this section, 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 the sequel. Denote B as the number of data blocks obtained for the non-overlapping case, each of length L. We assume for convenience that the B blocks cover the entire recorded sequence, i.e., $BL = K$.

Denote by $x_b[l] = x[l + (b-1)L]$, $l = 0, \ldots, L-1$, and $\mathbf{x}_b = [x_b[0], \ldots, x_b[L-1]]^T$ for $b = 1, \ldots, B$. 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 is expressed as

$$\begin{aligned}\hat{\mathbf{R}}_\mathbf{y} &= \frac{1}{B}\sum_{b=1}^{B}\mathbf{y}_b\mathbf{y}_b^H \\ &= \mathbf{A}_s\left(\frac{1}{B}\sum_{b=1}^{B}\mathbf{x}_b\mathbf{x}_b^H\right)\mathbf{A}_s^H = \mathbf{A}_s\hat{\mathbf{R}}_\mathbf{x}\mathbf{A}_s^H, \end{aligned} \quad (2)$$

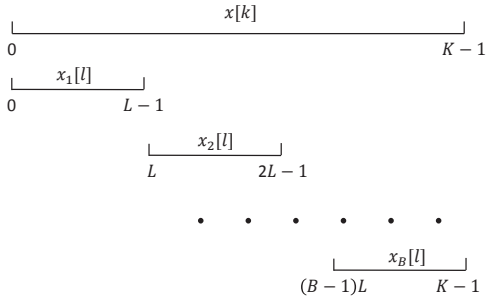
where $\hat{\mathbf{R}}_\mathbf{x}$ is an estimated covariance matrix of $\mathbf{R_x}$. As such, the compressed covariance matrix $\hat{\mathbf{R}}_\mathbf{y}$ with size $V \times V$ can be exploited to reconstruct the $L \times L$ matrix $\hat{\mathbf{R}}_\mathbf{x}$, provided that it includes all lags $\tau = -L+1, \ldots, L-1$ by designing a proper sampling matrix $\mathbf{A}_s$. Note that, since there are V^2 entries in $\hat{\mathbf{R}}_\mathbf{y}$, the sampling size which enables reconstruction of the Toeplitz matrix $\hat{\mathbf{R}}_\mathbf{x}$ is lower bounded by $\sqrt{L}$, i.e., $V \geq \sqrt{L}$. In the end, $\hat{\mathbf{R}}_\mathbf{x}$ can be reconstructed as

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

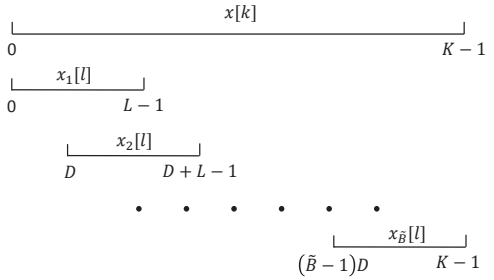
where $\hat{r}[\tau], \tau = -L+1, \ldots, L-1$ are estimated by averaging all the entries with the same lag τ in $\hat{\mathbf{R}}_\mathbf{y}$.

3. GENERALIZED COPRIME SAMPLING

Coprime sampling amounts to exploiting two uniform sub-Nyquist samplers with sampling period being coprime multiples of the Nyquist sampling period [11, 13]. In this section, the generalized coprime sampling scheme is depicted, which is presented in two operations. A multiple coprime unit factor $p \in \mathbb{N}^+$ [14], aiming to enhance the number of lags in the compressed covariance matrix, is first introduced, and then the utilization of overlapping samples between blocks is described to yield a reduced estimation variance through the use of a non-overlapping factor $q \in \mathbb{N}^+$.



(a) Non-overlapping segmentation



(b) Overlapping segmentation

Fig. 1. Illustration of segmentations.

3.1. The concept of coprime sampling

In coprime sampling, the sampling matrix $\mathbf{A}_s$ can be denoted as $\mathbf{A}_s = [\mathbf{A}_{s1}^T \ \ \mathbf{A}_{s2}^T]^T$, where $\mathbf{A}_{s1}$ and $\mathbf{A}_{s2}$ are the sub-sampling matrices corresponding to the 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. As such, from a data acquisition perspective, there are two sets of uniformly spaced samples of the input WSS signal $X(t), t = 0, \ldots, T$ using 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

$$y_1[k_1] = x[Mk_1] = X(Mk_1T_s),$$
$$y_2[k_2] = x[Nk_2] = X(Nk_2T_s). \quad (5)$$

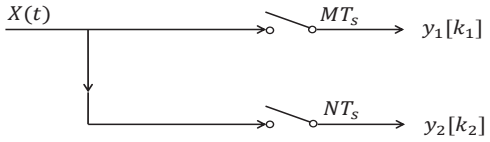


Fig. 2. Coprime sampling structure.

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

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

Denote $\mathbf{y}_{b_1} = [y_{b_1}[0], \ldots, y_{b_1}[N-1]]^T$ as an $N \times 1$ vector, and $\mathbf{y}_{b_2} = [y_{b_2}[0], \ldots, y_{b_2}[M-1]]^T$ as an $M \times 1$ vector, with $y_{b_1}[k_1] = x[(b-1)MN + Mk_1]$ and $y_{b_2}[k_2] = x[(b-1)MN + Nk_2]$, where $0 \le k_1 \le N-1$ and $0 \le k_2 \le M-1$, for $1 \le b \le K/(MN)$. In addition, let $\mathbf{y}_b = [\mathbf{y}_{b_1}^T \ \ \mathbf{y}_{b_2}^T]^T$. As such, the $(M+N) \times (M+N)$ covariance matrix $\mathbf{R}_\mathbf{y}$ can be expressed as

$$\mathbf{R}_\mathbf{y} = \begin{pmatrix} \mathbf{R}_{\mathbf{y}_{11}} & \mathbf{R}_{\mathbf{y}_{12}} \\ \mathbf{R}_{\mathbf{y}_{21}} & \mathbf{R}_{\mathbf{y}_{22}} \end{pmatrix} = \begin{pmatrix} \mathrm{E}[\mathbf{y}_{b_1}\mathbf{y}_{b_1}^H] & \mathrm{E}[\mathbf{y}_{b_1}\mathbf{y}_{b_2}^H] \\ \mathrm{E}[\mathbf{y}_{b_2}\mathbf{y}_{b_1}^H] & \mathrm{E}[\mathbf{y}_{b_2}\mathbf{y}_{b_2}^H] \end{pmatrix}. \quad (7)$$

In $\mathbf{R}_\mathbf{y}$, matrices $\mathbf{R}_{\mathbf{y}_{11}}$ and $\mathbf{R}_{\mathbf{y}_{22}}$ contains self-lags of the two sampler output streams, while their cross-lags are included in matrices $\mathbf{R}_{\mathbf{y}_{12}}$ and $\mathbf{R}_{\mathbf{y}_{21}}$. Note that, $\mathbf{R}_{\mathbf{y}_{21}} = \mathbf{R}_{\mathbf{y}_{12}}^*$. In addition, because the two sampled 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, $\mathbf{R}_\mathbf{x}$ can be reconstructed by using only $\mathbf{R}_{\mathbf{y}_{12}}$, whose cross-lags (including the negated ones) are given by the following set,

$$\mathbb{L} = \{\tau | \tau = Mk_1 - Nk_2\} \cup \{\tau | \tau = Nk_2 - Mk_1\}, \quad (8)$$

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

The prototype scheme uses one coprime unit samples to generate all lags in $\mathbb{L}$. However, it should be noticed that there are missing integers in the range $[-MN, MN]$, that is, they are not sufficient to reconstruct $\hat{\mathbf{R}}_\mathbf{x}$ with dimension $L = MN$. To achieve this, two coprime units from the first sampler and one coprime unit from the second sampler are used to form one block in [11], and the resulting lags are continuous in the range $[-MN - N + 1, MN + N - 1]$. This scheme is referred to as the conventional scheme in this paper. In this case, the maximum achievable L is $L_{\max} = MN + N$.

3.2. Generalized coprime sampling scheme

In the sequel, a multiple coprime unit factor $p \in \mathbb{N}^+$ is first introduced to achieve a larger achievable value of L. In each block, p coprime unit outputs from both samplers, i.e., $p(M+N)$ physical samples in the length of pMN, are used to estimate the covariance matrix. In this case, the resulting lags fall into the following set,

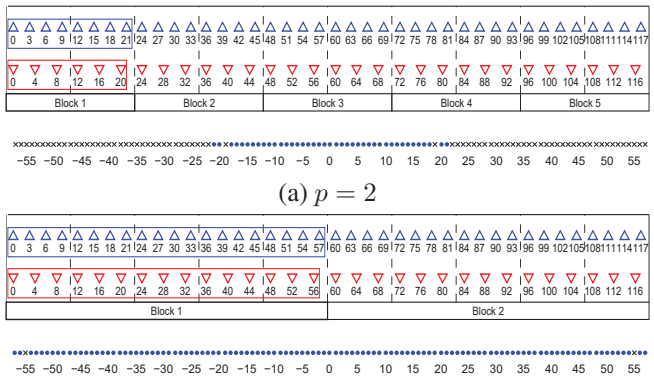
$$\tilde{\mathbb{L}} = \{\tau | \tau = Mk_1 - Nk_2\} \cup \{\tau | \tau = Nk_2 - Mk_1\}, \quad (9)$$

for $0 \le k_1 \le pN-1$ and $0 \le k_2 \le pM-1$. 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 $-(p-1)MN - M - N + 1 \le \tau \le (p-1)MN + M + N - 1$.

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 $\tilde{L}_{\max} = (p-1)MN + M + N$, and the number of the corresponding non-overlapping blocks is given by $B = \lfloor \frac{K}{pMN} \rfloor$.

An example for different values of p is illustrated in Fig. 3, 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, the $\hat{\mathbf{R}}_\mathbf{x}$ 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, as a trade-off of a reduced $B = 2$.



(a) $p = 2$

(b) $p = 5$

Fig. 3. Example for different values of p ($K = 120$, $M = 3$, and $N = 4$; △: first sampler outputs; ▽: second sampler outputs; •: lags; ×: holes.)

The variance of the estimated covariance and spectrum is generally reduced by utilizing a higher number of blocks used for averaging. In addition to averaging over non-overlapping segments, as discussed earlier, a more general and 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. As a result, the variance of the estimated covariance and spectrum can be reduced. In the second operation of the generalization, overlapping samples are used to increase the number of available blocks, yielding a reduced estimation variance through a use of non-overlapping factor $q \in \mathbb{N}^+$.

As shown in Fig 1(b), we maintain the same segment length pMN, and take $\tilde{B}$ ($\tilde{B} \geq B$) blocks with the starting points of these blocks D ($D \leq pMN$) units apart. Similarly, we assume, for convenience, that $(\tilde{B}-1)D + pMN = K$ covers the entire recorded sequence.

Definition 2: Assume that D consists of the length of q coprime units, i.e., $D = qMN$, where $1 \leq q \leq p$. Then, the number of blocks can be expressed as $\tilde{B} = \lfloor \frac{p}{q}B - \frac{p}{q} \rfloor + 1$.

It is straightforward to confirm that $\tilde{B} \geq B$ since $q \leq p$, as shown in Fig. 4 ($\tilde{B} = 6$ in Fig. 4 versus $B = 2$ in Fig 1(b)). Note that in the non-overlapping case, B can be considered as a special case of $\tilde{B}$, i.e., when $q = p$. In addition, $\tilde{B}$ increases as q decreases and is maximized when $q = 1$.

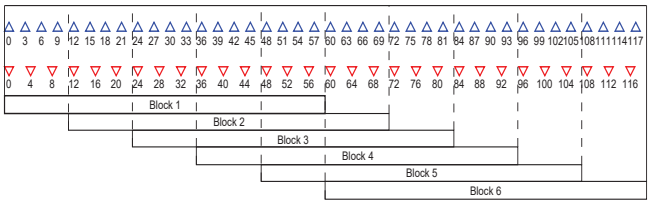


Fig. 4. Example of utilization of overlapping samples ($K = 120$, $M = 3$, $N = 4$, $p = 5$, and $q = 1$.)

Denote $\tilde{y}_{b_1}[k_1] = x[(b-1) \times qMN + Mk_1]$ and $\tilde{y}_{b_2}[k_2] = x[(b-1) \times qMN + Nk_2]$, where $0 \leq k_1 \leq pN-1$ and $0 \leq k_2 \leq pM-1$, for $1 \leq b \leq \tilde{B}$. In addition, let $\tilde{\mathbf{y}}_{b_1} = [\tilde{y}_{b_1}[0], \ldots, \tilde{y}_{b_1}[pN-1]]^T$ and $\tilde{\mathbf{y}}_{b_2} = [\tilde{y}_{b_2}[0], \ldots, \tilde{y}_{b_2}[pM-1]]^T$. The covariance matrix $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$, using the generalized scheme, can be estimated as $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}} = \frac{1}{\tilde{B}} \sum_{\tilde{b}=1}^{\tilde{B}} \tilde{\mathbf{y}}_{b_1} \tilde{\mathbf{y}}_{b_2}^H$. Then, $\hat{\mathbf{R}}_{\mathbf{x}}$ with dimension $L \times L$, where $L \leq \tilde{L}_{\max}$, can be reconstructed.

3.3. Compression factor

We examine the compression factor, which is defined as the ratio of the number of entries in $\hat{\mathbf{R}}_{\mathbf{x}}$ over the corresponding number in $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}_{12}}$, expressed as $\kappa = L^2/(pM \times pN)$. Because the maximum value of L is $\tilde{L}_{\max} = (p-1)MN + M + N$, the maximum achievable value of κ is given by

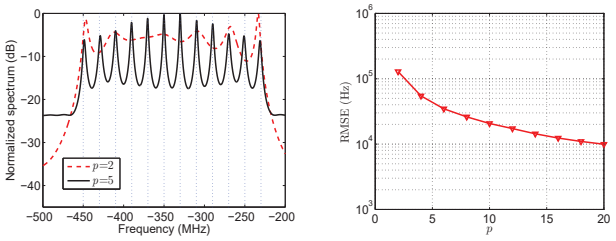
$$\kappa_{\max} = \frac{[(p-1)MN + M + N]^2}{pM \times pN}. \qquad (10)$$

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 $p \gg 1$.

4. SIMULATION RESULTS

For illustrative purposes, we demonstrate the key distinctions on the performance of spectrum estimation between different choices of arguments in the generalized coprime sampling scheme, based on the same sampling rates and compression factor, i.e., the same values of $M = 3$ and $N = 4$. Assume that I sinusoidal signals with identical powers are distributed in the frequency band $[-500, 500]$ MHz, and the ith signal is located at $[-450+(i-1)\delta_f]$ MHz for $i = 1, \ldots, I$, where δ_f is the frequency separation. 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 [15] is used to estimate the power spectrum. Our benchmarks are the spectrum identifiability and their statistical performance. The latter is evaluated in terms of average root mean square error (RMSE) of the estimated frequencies, based on 500 independent runs.

Fig. 5 examines the performance for different choices of p. In Fig. 5(a), the distinction on spectrum identifiability is depicted for the cases of $p = 2$ and $p = 5$. We consider $I = 12$ frequencies with $\delta_f = 20$ MHz separation in the presence of noise with a 0 dB SNR. It is evident that only the scenario of $p = 5$ can resolve all frequencies correctly, although the case of $p = 2$ also has a sufficient number of DOFs. Fig. 5(b) presents the statistical performance with respect to p, where $I = 5$ is assumed. It is observed that the estimation performance is improved, as p increases. As a summary, a larger p can improve the identifiability and estimation performance in the spectrum. However, the requirement of storage space and the computational load become higher, due to the resulting larger L.



(a) (b)

Fig. 5. The performance for different choices of p. (a) Estimated spectra for the cases of $p = 2$ and $p = 5$ ($I = 12$); (b) RMSE versus p ($I = 5$).

The advantage of utilization of the overlapping blocks is demonstrated in Fig. 6, where $p = 12$ is 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., $q = p = 12$.

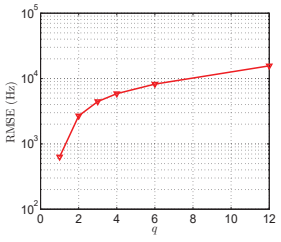


Fig. 6. RMSE versus q ($p = 12$ and $I = 5$).

5. CONCLUSIONS

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

6. REFERENCES

[1] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Info. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] G. Dasarathy, P. Shah, B. N. Bhaskar and R. Nowak, "Sketching sparse matrices," *arXiv preprint arXiv:1303. 6544,* 2003.

[4] Y. Chen, Y. Chi, and A. Goldsmith, "Exact and stable covariance estimation from quadratic sampling via convex programming," *arXiv preprint arXiv:1310. 0807v4,* 2013.

[5] 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.*, San Juan, Puerto Rico, Dec. 2011, pp. 337–340.

[6] D. D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4775–4789, Sept. 2012.

[7] D. Romero and G. Leus, "Compressive covariance sampling," in *Proc. Info. Theory Appl. Workshop*, San Diego, CA, Feb. 2013, pp. 1–8.

[8] 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.*, Prague, Czech Republic, May 2011, pp. 3884–3887.

[9] Z. Tian, Y. Tafesse, and B. M. Sadler, "Cyclic feature detection with sub-Nyquist sampling for wideband spectrum sensing," *IEEE J. Selected Topics in Signal Process.*, vol. 6, no. 1, pp. 58–69, 2012.

[10] H. Qiao and P. Pal, "Generalized nested sampling for compression and exact recovery of sysmmetric Teoplitz matrices," in *Proc. IEEE GlobalSIP Sympo. on Info. Process. for Big Data.*, Atlanta, GA, Dec. 2014, pp. 337–340.

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

[12] 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 Electroacoustics*, vol. AU-15, no. 2, pp. 70–73, 1967.

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

[14] S. Qin, Y. D. Zhang, and M. G. Amin, "High-resolution frequency estimation using generalized coprime sampling," in *Proc. SPIE Wireless Sens., Localization, Process. Conf.*, Baltimore, MD, May 2015, vol. 9497, pp. 94970K1–94970K7.

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