Universal Lower Bounds on Sampling Rates for Covariance Estimation

Deborah Cohen
Technion, Haifa, Israel
debby@tx.technion.ac.il

Yonina C. Eldar
Technion, Haifa, Israel
yonina@ee.technion.ac.il

Geert Leus
TU Delft, Delft, The Netherlands
g.j.t.leus@tudelft.nl

***Abstract*—Covariance estimation from compressive samples has become particularly attractive for two main reasons. First, many applications do not require the signal itself, and second-order statistics are oftentimes sufficient. The resulting requirement on the sampling rate of the original signal can therefore be reduced. Second, signal recovery from compressive samples leads to underdetermined systems which require additional constraints, such as the popular sparsity assumption. In contrast, covariance estimation can yield overdetermined problems, even from compressive samples, so that the additional constraints on the signal can be dropped. In this paper, we provide a unified framework for deriving lower bounds on the sampling rate required for covariance estimation of stationary signals, by deriving the lower Beurling density of the difference set associated with the original sampling set. A general sampling scheme is first considered, followed by the analysis of multicoset sampling. We prove that, in both cases, the sampling rate can be arbitrarily low, as was remarked extensively in the literature.**

***Keywords*—Covariance estimation, sub-Nyquist sampling, non-uniform sampling**

I. Introduction

Covariance estimation has been widely considered across different fields of statistical signal processing, such as power spectrum estimation [1], [2], [3], [4], [5], economics and financial time series analysis [6], machine learning [7], [8], and phaseless measurements [3], [9]. In all these examples, second-order statistics suffice for the task at hand and estimation of the signal itself is unnecessary.

In light of new compressed sensing (CS) [10] paradigms, covariance estimation of stationary signals has been recently revisited. This is owing to the fact that, while signal recovery from compressive measurements is an underdetermined problem, compressive covariance recovery has been shown to lead to overdetermined systems in certain settings [11], [1]. In such cases, the sparsity constraint required for signal recovery can be dropped [1]. This is a result of the fact that the covariance of stationary signals is only a function of the time lags. The cardinality of the difference set, namely the set that contains the time lags, is greater than this of its associated original set, and can even be of the order of the square of the cardinality of the original sampling set [12]. We refer to the density of this difference set as the covariance sampling rate. It then follows that the average covariance sampling rate can be above the Nyquist rate, even when the actual sampling rate is below Nyquist by orders of magnitude. Two questions then arise in the context of covariance estimation from non-uniform, low rate samples:

- Let $\tilde{R} = \{t_i\}_{i\in\mathbb{Z}}$ be a sampling set. What can be said about the density of the difference set $R = \{t_i - t_j\}, \forall t_i > t_j \in \tilde{R}$?
- Given a non-uniform finite set of observations $\{x(t_n)\}_{n=1}^{N}$, how good is the corresponding power spectrum estimator?

$$\hat{S}_x(f) = \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{N-k} x(t_i)x(t_{i+n})e^{-j2\pi f(t_{i+n}-t_i)}.$$

Here, the power spectrum is defined by

$$S_x(f) = \int_{-\infty}^{\infty} r_x(\tau)e^{-j2\pi f\tau}\mathrm{d}\tau, \tag{1}$$

where $r_x(\tau) = \mathbb{E}\left[x(t)x(t-\tau)\right]$ is the covariance function of $x(t)$. This work considers the first of these issues. The second one is left for future work.

Many works have considered covariance, or power spectrum, estimation from compressive samples obtained using specific sampling schemes. Masry [13], [14] investigates Poisson sampling for spectral density estimation. The proposed estimator is shown to be consistent for all positive values of the average sampling rate. As opposed to random sampling, a deterministic approach, multicoset, or periodic nonuniform sampling, has been investigated in [15], [1], [11], [16]. In [15], an unbiased power spectrum estimator from multicoset samples is proposed with arbitrarily low average sampling rate. The authors in [1] discuss reconstruction of the covariance or the power spectrum from both underdetermined and overdetermined systems. For the first case, they exploit sparsity properties of the signal and apply CS reconstruction techniques but do not analyze the sampling rate. In the second overdetermined case, they show that the so-called minimal sparse and minimal circular sparse ruler patterns [17] provide optimal solutions for sub-Nyquist sampling, within the class of multicoset samplers, without any prior sparsity assumption. In [11], [16], the authors propose a method to estimate finite resolution approximations to the true power spectrum exploiting multicoset sampling. That is, they estimate the average power within subbands rather than the power spectrum for each frequency. They consider

This work is supported in part by the Semiconductor Research Corporation (SRC) through the Texas Analog Center of Excellence (TxACE) at the University of Texas at Dallas (Task ID:1836.114), in part by the Israel Science Foundation under Grant no. 170/10, in part by the Ollendorf foundation, and in part by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI).

both overdetermined and undertermined, or compressive, systems. In the latter case, CS techniques are used, which exploit the signal sparsity, whereas the former setting does not assume any sparsity. In [11], the authors assume that the sampling pattern is such that the system they obtain has a unique solution but no specific sampling pattern or rate satisfying this condition is discussed. In [16], sampling patterns generated uniformly at random and the Golomb ruler are considered in simulations but no analysis of the required rate is performed. Another deterministic sampling scheme, co-prime sampling, is considered in [18]. The covariance is estimated from samples for a co-prime pair of sparse samplers. Due to the co-prime property, a number of consecutive lags can be recovered, yet not as many lags as the sparse ruler for multicoset sampling.

A common conclusion is that the sampling rate for covariance estimation can be arbitrarily low. However, this result is derived for specific sampling schemes only. In this paper, we provide a unified framework for deriving a universal lower bound on the sampling rate for covariance estimation. We first consider a general sampling scheme and show that a sampling set with lower Beurling density zero leads to a difference set of infinite density, under mild conditions. It follows that the minimal sampling rate for covariance estimation is indeed zero. We then turn to multicoset sampling and compute the lower Beurling density of the resulting difference set. This allows us to derive a bound for the sampling rate for this scheme, as a function of the number of sampling channels. We note that we obtain a bound lower than those derived in the literature.

This paper is organized as follows. Section II describes the signal model and formulates the covariance estimation problem. In Section III, we present a lower bound for the sampling rate for covariance estimation. This result is applied to multicoset sampling in Section IV.

II. Signal Model and Problem Formulation

Consider the space $\mathcal{B}$ of stationary wide-sense ergodic bandlimited signals $x(t)$ whose Fourier transform, defined by

$$X(f) = \lim_{T\to\infty} \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} x(t)e^{-j2\pi ft}\mathrm{d}t, \tag{2}$$

is restricted to an unknown support $\mathcal{T}$ which is a subset of $\mathcal{F} = [-f_{\text{Nyq}}/2, f_{\text{Nyq}}/2]$. If $\mathcal{T} \subset \mathcal{F}$, then $x(t)$ is said to be sparse. Denote $\Lambda = \lambda(\text{supp}(X(f)))$, where λ is the Lebesgue measure. Let $\tilde{R} = \{t_i\}_{i\in\mathbb{Z}}$, be a sampling set of $x(t)$.

Our goal is to estimate the covariance function $r_x(\tau)$ of $x(t)$, defined as

$$r_x(\tau) = \mathbb{E}\left[x(t)x(t-\tau)\right], \tag{3}$$

from samples of $x(t)$ at times $t_i \in \tilde{R}$. Since $r_x(\tau)$ is only a function of the time lags τ, we are interested in the difference set $R = \{t_i - t_j\}, \forall t_i > t_j \in \tilde{R}$.

Consider the space $\mathcal{A}$ of deterministic functions $h(t)$ whose Fourier transform, defined by

$$H(f) = \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}\mathrm{d}t, \tag{4}$$

is restricted to an unknown support $\mathcal{T}$, with Lebesgue measure Λ, which is a subset of $\mathcal{F}$. Similarly to Landau's [19] definition of a sampling set, the authors in [20] define a blind sampling set, namely a sampling set when the signal support is unknown. The set $\tilde{R}$ is a stable sampling set for $\mathcal{A}$ if there exist constants $\alpha > 0$ and $\beta < \infty$ such that

$$\alpha||x-y||^2 \le ||x_{\tilde{R}} - y_{\tilde{R}}||^2 \le \beta||x-y||^2, \quad \forall x, y \in \mathcal{A}, \tag{5}$$

where $x_{\tilde{R}}[i] = x(t_i)$ is the sequence of samples of $x(t)$. The following theorem from [20] derives the conditions of the lower Beurling density of a set $\tilde{R}$ so that it is a sampling set for a space $\mathcal{A}$.

Theorem 1 ([20], Theorem 1)**.** *Let $\mathcal{A}$ be a set of bandlimited signals to $\mathcal{F}$, restricted to an unknown support with Lebesgue measure Λ. If $\tilde{R}$ is a stable sampling set for $\mathcal{A}$, then it must have density*

$$D^-(\tilde{R}) \ge \min\{2\Lambda, f_{Nyq}\}, \tag{6}$$

where

$$D^-(\tilde{R}) = \lim_{r\to\infty} \inf_{y\in\mathbb{R}} \frac{|\tilde{R}\cap[y, y+r]|}{r} \tag{7}$$

is the lower Beurling density of $\tilde{R}$.

Denote $d_{\tilde{R}}(r) = \inf_{y\in\mathbb{R}} |\tilde{R}\cap[y,y+r]|$. Then, $D^-(\tilde{R}) = \lim_{r\to\infty} \frac{d_{\tilde{R}}(r)}{r} = \infty$. Here, we are thus interested in computing the lower Beurling density of the difference set, namely $D^-(R)$, in order to derive a lower bound on the minimal sampling rate required for covariance estimation. By abuse of notation, the sampling times of the covariance are $\tau_n \in R, n \in \mathbb{Z}$, although the covariance is never actually sampled. The covariance average sampling period is therefore $\lim_{n\to\infty} \tau_n/n$ [21].

III. Universal minimal sampling rate

In this section, we consider a general sampling method. We begin by showing that a sampling set with Beurling density zero yields a difference set with infinite Beurling density, under mild conditions. We then conclude that the minimal sampling rate for covariance estimation is zero, as hinted in the literature.

A. Difference set density

Lemma 1. *Let $\tilde{R} = \{t_i\}_{i\in\mathbb{Z}}$, be a sampling set with lower Beurling density $D^-(\tilde{R}) = 0$, so that the set of differences between two sets of size p and q is of the order of $p\cdot q$. Let $R = \{t_i - t_j\}, \forall t_i > t_j \in \tilde{R}$ be the associated difference set. If $\lim_{r\to\infty} \frac{d_{\tilde{R}}(r)}{\sqrt{r}} = \infty$, then, $D^-(R) = \infty$.*

Proof: Let $y \in \mathbb{R}$ and let $r \in \mathbb{R}^+$. Consider $\tilde{R}\cap[y, y+r] = \{t_j\}_{j=[1]}^{[n]}$ and $\tilde{R}\cap[2y+r, 2y+2r] = \{t_i\}_{i=[n+1]}^{[m]}$. It holds that

$$\max_{[1]\le j\le[n],[n+1]\le i\le[m]} t_i - t_j = y + 2r,$$

and

$$\min_{[1]\le j\le[n],[n+1]\le i\le[m]} t_i - t_j = y.$$

Thus, $R\cap[y, y+2r] \supseteq \{t_i - t_j\}, \forall [1] \le j \le [n]$ and $[n+1] \le i \le [m]$. That is, every pair $\{t_j, t_i\}$, where t_j is in the first set, namely the interval $[y, y+r]$ and t_i is in the second set, namely the interval $[2y+r, 2y+2r]$, yields a difference $t_i - t_j$ which is in the interval $[y, y+2r]$. Therefore, from the assumption, the number of such differences is of the order of the product of the number of elements in each set. Since the interval $[2y+r, 2y+2r]$ may contain more differences than those described above, it follows that

$$|R \cap [y, y+2r]| \geq C|\tilde{R} \cap [y, y+r]| \cdot |\tilde{R} \cap [2y+r, 2y+2r]|, \quad (8)$$

for some $C > 0$. Therefore,

$$\begin{aligned} D^{-}(R) &= \lim_{r\to\infty} \inf_{y\in\mathbb{R}} C \frac{|R \cap [y, y+2r]|}{2r} \\ &\geq \lim_{r\to\infty} \inf_{y\in\mathbb{R}} \frac{C}{2} \frac{|\tilde{R} \cap [y, y+r]|}{\sqrt{r}} \cdot \frac{|\tilde{R} \cap [2y+r, 2y+2r]|}{\sqrt{r}} \\ &= \infty, \end{aligned} \quad (9)$$

where the last equation follows from the fact that $\lim_{r\to\infty} \frac{d(r)}{\sqrt{r}} = \infty$. ■

We note that the conditions of Lemma 1 do not hold for uniform sampling, but are respected by multicoset sampling.

B. Sampling rate bound

We now apply this result to the covariance function $r_x(\tau)$.

Theorem 2. *Let $x(t) \in \mathcal{B}$ and R be as in Lemma 1. Then, the minimal rate for perfect recovery of $r_x(\tau)$ is 0.*

Proof: From [22], $S_x(f) = \mathbb{E}|X(f)|^2$, where $S_x(f)$ is defined in (1). Thus, obviously, the frequency support of $r_x(\tau)$ is identical to that of $x(t)$, and $r_x(\tau)$ is bandlimited as well.

Let $\mathcal{A}$ be as in Theorem 1. Define

$$h_T(\tau) = \frac{1}{2T} \int_{-T}^{T} x(t)x(t-\tau)\mathrm{d}t. \quad (10)$$

It holds that $h_T(t) \in \mathcal{A}$. Now, let $\tau_n \in R, n \in \mathbb{Z}$. From Lemma 1, $D^{-}(R) = \infty$. Thus, from Theorem 1, R is a stable sampling set for $\mathcal{A}$.

From the wide-sense ergodicity of $x(t)$, it holds that

$$\lim_{T\to\infty} h_T(\tau) = r_x(\tau) \quad (11)$$

where the convergence is in the mean square sense. ■

Therefore, the covariance of $x(t)$ can by recovered from samples of $x(t)$ from a sampling set $\tilde{R}$ as defined in Lemma 1, for any f_{Nyq} and any Λ. The result of Theorem 2 has widely been hinted at in the literature, at least since the 70s and the work of Masry on Poisson sampling, which states that "it is shown that the estimate [of the spectral density function] is mean-square consistent for all positive values of the average sampling rate" [13]. More recently, Tarczynski [15], using multicoset (or periodic non uniform) sampling shows that "in the case of PSD [power spectral density] estimation, the average sampling rate can be arbitrarily low". In [11], where multicoset sampling is considered as well, "the noncompressive estimates can theoretically be computed at arbitrarily low sampling rates". In the context of co-prime sampling [18], the authors show that "the sampling rate can be made arbitrarily small". Here, we demonstrate that the universal sampling rate lower bound for covariance estimation is zero, regardless of the specific sampling schemes.

In the next section, we turn to study multicoset sampling, and derive a lower bound on the minimal sampling rate for covariance estimation.

IV. Minimal rate for multicoset sampling

Consider now multicoset sampling [23]. Multicoset sampling can be described as the selection of certain samples from the uniform grid. More precisely, the uniform grid is divided into blocks of n consecutive samples, from which only m are kept. The ith sampling sequence is defined as

$$x_{c_i}[l] = \begin{cases} x(lT), & l = n(k+c_i), k \in \mathbb{Z} \\ 0, & \text{otherwise}, \end{cases} \quad (12)$$

where $0 < c_1 < c_2 < \cdots < c_m < 1$ and $T = 1/f_{\text{Nyq}}$.

A. Difference set density

In the following lemma, we derive the lower Beurling density of the difference set of a multicoset sampling set.

Lemma 2. *Let $\tilde{R}_p = \{nT(k+c_1), nT(k+c_2), \ldots, nT(k+c_m)\}_{k\in\mathbb{Z}}$, be a periodic sampling set with period nT, where $0 \leq c_i < 1$ for $i \in 1, 2, \ldots, m$ and $c_i \neq c_j$ for $i \neq j$. Assume that $c_i - c_j \neq c_k - c_l, \forall i \neq k, j \neq l$, namely the differences between two distinct cosets are unique. Let $R = \{t_i > t_j\}, \forall t_i, t_j \in \tilde{R}$. The lower Beurling density of the sampling set R is given by*

$$D^{-}(R) = \frac{m(m-1)+1}{nT}. \quad (13)$$

Proof: The sampling set R can be expressed as the union of $m(m-1)+1$ uniform sampling sets

$$R = R_u \bigcup_{1\leq i\neq j\leq m} R_{u_{ij}}, \quad (14)$$

where $R_u = \{nTk\}_{k\in\mathbb{Z}}$ and $R_{u_{ij}} = \{nT(k+c_i-c_j)\}_{k\in\mathbb{Z}}$ are uniform sampling sets and $\bigcup$ denotes union between sets. Since $c_i - c_j \neq c_k - c_l, \forall i \neq k, j \neq l$, then R_u and $R_{u_{ij}}, 1 \leq i \neq j \leq m$ are disjoint sets. It follows that

$$\begin{aligned} |R \cap [y, y+r]| &= |R_u| + \sum_{1\leq i\neq j\leq m} |R_{u_{ij}}| \\ &= \left\lfloor \frac{y+r}{nT} \right\rfloor - \left\lceil \frac{y}{nT} \right\rceil + 1 \\ &+ \sum_{1\leq i\neq j\leq m} \left(\left\lfloor \frac{y-(c_i-c_j)nT+r}{nT} \right\rfloor - \left\lceil \frac{y-(c_i-c_j)nT}{nT} \right\rceil + 1 \right), \end{aligned}$$

where the last equality comes from the derivation of the Beurling density of a uniform sampling set [24]. Therefore,

$$\begin{aligned} (m(m-1)+1) \left\lfloor \frac{r}{nT} \right\rfloor &\leq \inf_{y\in\mathbb{R}} |R \cap [y, y+r]| \\ &\leq (m(m-1)+1) \left(\left\lfloor \frac{r}{nT} \right\rfloor + 1 \right). \end{aligned} \quad (15)$$

Dividing the equations by r and taking the limit as $r \to \infty$ leads to

$$D^{-}(R) = \lim_{r\to\infty} \inf_{y\in\mathbb{R}} \frac{|R \cap [y, y+r]|}{r} = \frac{m(m-1)+1}{nT}, \quad (16)$$

from the sandwich theorem. ■

B. Sampling rate bound

Theorem 3 presents a lower bound for the sampling rate using multicoset sampling.

Theorem 3. *Let $x(t) \in \mathcal{B}$. Let $C = \{c_i\}_{i=1}^{m}$ be a sampling pattern in multicoset sampling as in Lemma 2. The minimal rate for perfect recovery of $r_x(\tau)$ when using multicoset sampling with n channels is given by*

$$\frac{m}{nT} \geq \frac{1}{T} \min \left\{ \frac{1+\sqrt{8\Lambda T n - 3}}{2n}, \frac{1+\sqrt{4n-3}}{2n} \right\}. \tag{17}$$

Proof: Let $\tilde{R} = \{nT(k+c_i)\}$ and R be as in Lemma 2. From Lemma 2, $D^{-}(R) = \frac{m(m-1)+1}{nT}$. Now, from [20], if R is a blind sampling set for $r_x(\tau)$, then

$$D^{-}(R) \geq \min \{2\Lambda, f_{\text{Nyq}}\}. \tag{18}$$

Therefore, here, if R is a blind sampling set for $S_x(f)$, then

$$\frac{m(m-1)+1}{nT} \geq \min \left\{ 2\Lambda, \frac{1}{T} \right\}, \tag{19}$$

which leads to

$$\frac{m}{n} \geq \min \left\{ \frac{1+\sqrt{8\lambda n - 3}}{2n}, \frac{1+\sqrt{4n-3}}{2n} \right\}. \tag{20}$$

■

Since Theorem 3 holds for every n, the minimal sampling rate for perfect reconstruction of $r_x(\tau)$ using multicoset sampling is 0.

C. Comparison with previous work and discussion

Note that for the case where $x(t)$ is not sparse, namely $\Lambda = 1/T$, (19) reduces to

$$\frac{m(m-1)+1}{n} \geq 1, \tag{21}$$

which is exactly the same condition as equation (28) in [17].

In [17], the minimal sparse ruler has been proposed as a solution to recover all n lags $t_i - t_j$. In that case, the difference set is a uniform sampling set with spacing T and the covariance can be estimated at all lags τ. However, the minimal sparse ruler does not achieve the minimal rate (17). Figure 1 shows the minimal sampling rate derived in (17) and the lower bound of the minimal sampling rate that can be achieved using the minimal sparse ruler ([17], equation (34)), namely

$$\frac{m}{n} \geq \frac{\sqrt{\tau(n-1)}}{n}, \quad \tau \approx 2.4345. \tag{22}$$

The authors suggest that this result is due to the fact that when the selection of samples is limited to the Nyquist grid, the bound (17) cannot be achieved. This bound requires all coset differences to be unique, which only holds for specific finite m and n and are referred to as perfect Golomb ruler [25]. Therefore, we propose selecting the samples non-uniformly and not on a grid. The following toy example illustrates this statement.

Consider the minimal sparse ruler of length 10. This ruler requires 6 marks, as shown in Fig. 2. Obviously, all the lags $1 \leq \tau \leq 10$ are identifiable. Now, consider the non-uniform ruler in Fig. 3. This ruler has only 5 marks and yields the following 10 lags or differences $\{0.9, 2.1, 3.2, 3.8, 4.1, 5.9, 7, 7.9, 9.1, 10\}$. The number of lags that are obtained from both rulers is identical. While the minimal sparse ruler yields uniform lags, the non-uniform ruler gives non-uniform differences. In this example, the average covariance sampling rate is identical in both schemes, and the covariance can be reconstructed from either set of samples. Obviously, reconstruction from non-uniform samples is more complex than reconstruction from uniform samples [26]. However, the minimal sampling rate when we allow for non-uniform lags is lower than the minimal sampling rate required to reconstruct all the differences on a grid, as shown here.

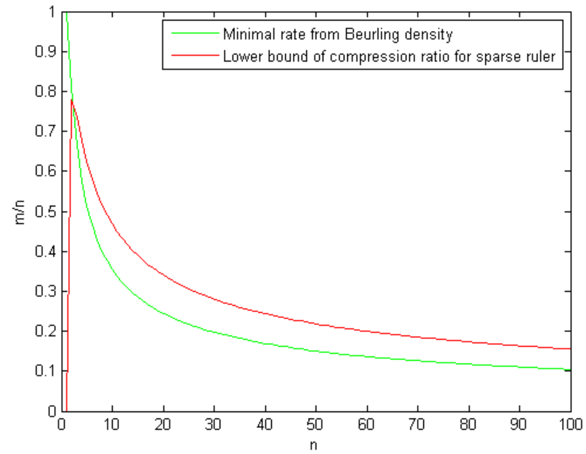

Fig. 1. Minimal compression ratio for multicoset sampling

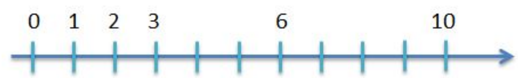

Fig. 2. Minimal sparse ruler.

Fig. 3. Non-uniform "ruler".

V. Conclusion

We presented lower bounds for the sampling rate required for covariance estimation in the context of a general sampling scheme, and for multicoset sampling in particular. A similar analysis of other sampling schemes, based on the derivation of the lower Beurling density of the resulting difference set, is left to future work.

References

[1] D. D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. on Signal Processing*, vol. 60, pp. 4775–4789, Sept. 2012.

[2] Y. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wideband spectrum sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'2009*, 2009, pp. 2337–2340.

[3] Y. Chen, Y. Chi, and A. Goldsmith, "Exact and stable covariance estimation from quadratic sampling via convex programming," *arXiv preprint arXiv:1310.0807*, 2013.

[4] C. P. Yen, Y. Tsai, and X. Wang, "Wideband spectrum sensing based on sub-Nyquist sampling," *IEEE Trans. on Image Processing*, vol. 61, pp. 3028–3040, 2013.

[5] D. Cohen and Y. C. Eldar, "Sub-Nyquist sampling for power spectrum sensing in cognitive radios: A unified approach," *CoRR*, vol. abs/1308.5149, 2013.

[6] J. J. Bai and S. Shi, "Estimating high dimensional covariance matrices and its applications," *Annals of Economics and Finance*, vol. 12, no. 2, pp. 199–215, 2011.

[7] Y. Zhang and J. Schneider, "Learning multiple tasks with a sparse matrix-normal penalty," in *Advances in Neural Information Processing Systems*, vol. 23, 2010, pp. 2550–2558.

[8] A. Hero and B. Rajaratnam, "Hub discovery in partial correlation graphs," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 6064–6078, 2012.

[9] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," 2012, to appear in SIAM J. on Imaging Sciences.

[10] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge University Press, 2015.

[11] M. A. Lexa, M. E. Davies, J. S. Thompson, and J. Nikolic, "Compressive power spectral density estimation," *IEEE ICASSP*, pp. 22 – 27, May 2011.

[12] P. Pal and P. P. Vaidyanathan, "Nested array: A novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. on Signal Processing*, vol. 58, pp. 4167–4181, Aug. 2010.

[13] E. Masry and M. C. Lui, "Discrete-time spectral estimation of continuous-parameter processes - a new consistent estimate," *IEEE Trans. Information Theory*, pp. 298–312, 1976.

[14] E. Masry, "Poisson sampling and spectral estimation of continuous-time processes," *IEEE Trans. Information Theory*, pp. 173–183, Mar. 1978.

[15] D. Qu and A. Tarczynski, "A novel spectral estimation method by using periodic nonuniform sampling," *Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)*, pp. 1134–1138, 2007.

[16] M. A. Lexa, M. E. Davies, and J. S. Thompson, "Compressive and noncompressive power spectral density estimation from periodic nonuniform samples," *CoRR*, vol. abs/1110.2722, 2011.

[17] D. Romero, R. Lopez-Valcarce, and G. Leus, "Compression limits for random vectors with linearly parameterized second-order statistics," *arXiv:1311.0737 [math.ST]*, Nov. 2013.

[18] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. on Signal Processing*, vol. 59, pp. 573–586, 2011.

[19] H. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math*, vol. 117, pp. 37–52, Jul. 1967.

[20] M. Mishali and Y. C. Eldar, "Blind multi-band signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. on Signal Processing*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.

[21] Y. C. Eldar and A. V. Oppenheim, "Filter bank reconstruction of bandlimited signals from nonuniform and generalized samples," *IEEE Trans. Signal Processing*, pp. 2864–2875, Oct. 2000.

[22] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 1991.

[23] R. Venkataramani and Y. Bresler, "Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2173 – 2183, Sept. 2000.

[24] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[25] K. Drakakis, "A review of the available construction methods for Golomb rulers," *Advances in Mathematics of Communications*, vol. 3, pp. 235–250, Aug. 2009.

[26] K. Yao and J. B. Thomas, "On nonuniform sampling of bandwidth-limited signals," *IEEE Trans. Circuit Theory*, pp. 251–257, Dec. 1956.