NON-UNIFORM SAMPLING FOR COMPRESSIVE CYCLIC SPECTRUM RECONSTRUCTION

Dyonisius Dony Ariananda and Geert Leus
Faculty of EEMCS, Delft University of Technology, The Netherlands
{d.a.dyonisius, g.j.t.leus}@tudelft.nl

ABSTRACT

We introduce a new cyclic spectrum estimation method for wide-sense cyclostationary (WSCS) signals sampled at sub-Nyquist rate using non-uniform sampling. We exploit the block Toeplitz structure of the WSCS signal correlation matrix and write the linear relationship between this matrix and the correlations of the sub-Nyquist rate samples as an overdetermined system. We find the condition under which the system matrix has full column rank allowing for least-squares reconstruction of the WSCS signal correlation matrix from the correlations of the compressive measurements. We also evaluate the case when the support of the WSCS signal correlation is limited and look at a special case where each selection matrix is restricted to either an identity matrix or an empty matrix. In the latter case, we can connect the full column rank condition of the system matrix with a circular sparse ruler.

***Index Terms*—** non-uniform sampling, cyclostationary, circular sparse ruler, linear sparse ruler, least-squares

1. INTRODUCTION AND RELATED WORKS

Many researches have been done in the field of compressive sampling a.k.a. sub-Nyquist rate sampling due to the desire to relax the requirements on the analog-to-digital converters while maintaining the possibility for signal reconstruction with no or little information loss. This is possible under the constraint that the signal is sparse in a certain basis [1, 2]. Some applications, such as wideband spectrum sensing for cognitive radio networks, however, require perfect reconstruction of only the power spectrum or cyclic spectrum, instead of the signal itself. Perfect power spectrum reconstruction from sub-Nyquist rate samples has been shown to be possible for a wide-sense stationary (WSS) signal [3, 4] and for a multiband signal with uncorrelated spectra at different bands [5]. This can be performed even without applying a sparsity constraint on the actual power spectrum.

Since a stationary process can be perceived as a special case of a cyclostationary process, which is a process whose statistical characteristics vary periodically with time, the reconstruction of the power spectrum of a WSS signal can be treated as a special case of the reconstruction of the cyclic spectrum of a wide-sense cyclostationary (WSCS) signal. However, while compressively reconstructing the power spectrum of a WSS signal without applying a sparsity constraint on the power spectrum is possible due to the Toeplitz structure in the temporal correlation matrix, it is challenging to find a special structure in the WSCS signal correlation matrix that can be exploited to perform compression. This complicates the cyclic spectrum reconstruction from sub-Nyquist rate samples of a WSCS signal and forces [6] to assume a sparse two-dimensional cyclic spectrum. A similar problem is found in [7], which focuses on only a real multiband signal and arrives at the assumption that the correlation matrix of the entries at different bands has nonzero values only at the diagonal and anti-diagonal elements. A different approach in [8] views the compressive measurements as random linear projections of the original signal and sets the span of the random linear projections equal to an integer multiple of the cyclic period. This allows [8] to exploit the block Toeplitz structure in the correlation matrix and to perform compression yet to present their reconstruction problem as an overdetermined system. The work of [8] does not focus on multi-coset or non-uniform sampling and thus, it does not attempt to find the condition of the system matrix that allows for a least-squares (LS) solution for the overdetermined system. In this paper, we also set the span of the random linear projections equal to an integer multiple of the cyclic period but we focus only on non-uniform sampling. We express the correlations of the sub-Nyquist rate samples as a linear function of the correlation matrix of the corresponding Nyquist-rate samples. We find the condition for the system matrix to have full rank, which enables the LS reconstruction of the correlation matrix of the WSCS signal from the correlations of the compressive measurements. The cyclic spectrum can then be estimated from the reconstructed correlation matrix of the WSCS signal.

This work is supported in part by NWO-STW under the VICI program (project 10382).

2. SYSTEM MODEL AND COMPRESSION

Let us consider a discrete WSCS signal $x[t]$, where the autocorrelation sequence $r_x[t,\tau] = E\{x[t]x^*[t-\tau]\}$ is periodic in t with a period of T. The cyclic autocorrelation sequence and the cyclic power spectrum of $x[t]$ are then respectively given by

$$\tilde{r}_x[f,\tau] = \frac{1}{T}\sum_{t=0}^{T-1} r_x[t,\tau]e^{-j2\pi f(t-\tau/2)/T}, \tag{1a}$$

$$s_x[f,\phi] = \sum_{\tau=-\infty}^{\infty} \tilde{r}_x[f,\tau]e^{-j2\pi\phi\tau}, \tag{1b}$$

with $\phi \in [0,1)$ the frequency and $f \in \{0,1,\ldots,T-1\}$ the cyclic frequency. Note that when $x[t]$ is produced by Nyquist-rate sampling at a rate of f_s Hz, f and ϕ correspond to an actual cyclic frequency of $f\frac{f_s}{T}$ Hz and an actual frequency of ϕf_s Hz, respectively. We cascade T consecutive samples $x[t]$ in $\mathbf{x}[n] = [x[nT], x[nT+1],\ldots,x[nT+T-1]]^T$, which is a sequence of stationary vectors with $T\times T$ correlation matrix sequence $\mathbf{R}_x[k] = E\left\{\mathbf{x}[n]\mathbf{x}^H[n-k]\right\} = [r_x[t,kT+t-\tau]]_{t,\tau}$. Note that a one-to-one mapping exists between $r_x[t,\tau]$ and $\mathbf{R}_x[k]$. Next, we stack N consecutive $T\times 1$ stationary vectors $\mathbf{x}[n]$ into an $NT\times 1$ vector $\tilde{\mathbf{x}}[\tilde{n}]$ as $\tilde{\mathbf{x}}[\tilde{n}] = [\mathbf{x}^T[\tilde{n}N], \mathbf{x}^T[\tilde{n}N+1],\ldots,\mathbf{x}^T[\tilde{n}N+N-1]]^T$. The $NT\times NT$ correlation matrix of $\tilde{\mathbf{x}}[\tilde{n}]$ at lag 0 is then given by $\mathbf{R}_{\tilde{x}}[0] = E[\tilde{\mathbf{x}}[\tilde{n}]\tilde{\mathbf{x}}^H[\tilde{n}]]$, whose relationship to $\mathbf{R}_x[k]$ is given by

$$\mathbf{R}_{\tilde{x}}[0] = \begin{bmatrix} \mathbf{R}_x[0] & \mathbf{R}_x[-1] & \ldots & \mathbf{R}_x[-N+1] \\ \mathbf{R}_x[1] & \mathbf{R}_x[0] & \ldots & \mathbf{R}_x[-N+2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_x[N-1] & \mathbf{R}_x[N-2] & \ldots & \mathbf{R}_x[0] \end{bmatrix} \tag{2}$$

which has a block Toeplitz structure. This allows us to perform a temporal compression by introducing an $\tilde{M}\times 1$ vector $\tilde{\mathbf{y}}[\tilde{n}] = [\mathbf{y}^T[\tilde{n}N], \mathbf{y}^T[\tilde{n}N+1],\ldots,\mathbf{y}^T[\tilde{n}N+N-1]]^T$, where $\mathbf{y}[\tilde{n}N+n]$

is an $M_n \times 1$ vector given by

$$\mathbf{y}[\tilde{n}N + n] = \mathbf{C}_n\mathbf{x}[\tilde{n}N + n],\ n = 0, 1, \ldots, N-1, \qquad (3)$$

with $\tilde{M} = \sum_{n=0}^{N-1} M_n$ and $\mathbf{C}_n$ the $M_n \times T$ multi-coset sampling a.k.a. selection matrix whose rows are obtained by selecting the M_n rows of the $T \times T$ identity matrix $\mathbf{I}_T$. We can then write $\tilde{\mathbf{y}}[\tilde{n}]$ as

$$\tilde{\mathbf{y}}[\tilde{n}] = \tilde{\mathbf{C}}\tilde{\mathbf{x}}[\tilde{n}], \qquad (4)$$

where $\tilde{\mathbf{C}}$ is an $\tilde{M} \times NT$ block diagonal matrix given by $\tilde{\mathbf{C}} = \text{diag}\{\mathbf{C}_0, \mathbf{C}_1, \ldots, \mathbf{C}_{N-1}\}$. As we will show in Section 3, for a certain value of n, it is possible that none of the rows of $\mathbf{I}_T$ is selected to form $\mathbf{C}_n$ (i.e., $M_n = 0$). In this case, for that particular n, $\mathbf{y}[\tilde{n}N + n] = \mathbf{C}_n = [\,]$, where $[\,]$ is an empty matrix, as none of the entries of $\mathbf{x}[\tilde{n}N + n]$ is selected. Note from (4) that we take $\tilde{M}$ random linear projections with a total span of NT (N times the period of the autocorrelation sequence $r_x[t, \tau]$) and attain compression by having $\tilde{M} < NT$.

Observe that $\mathbf{y}[\tilde{n}N + n]$ in (3) is generally not a sequence of stationary vectors since $\mathbf{C}_n$ is generally different for different values of n. However, we can obtain a sequence of stationary vectors by collecting $\mathbf{y}[\tilde{n}N+n]$ at different values of $\tilde{n}$ for a given n. We define the $M_n \times M_{n'}$ correlation matrix $\mathbf{R}_{y_{n,n'}} = E[\mathbf{y}[\tilde{n}N+n]\mathbf{y}^H[\tilde{n}N+n']]$, which can be written, by introducing $k = n - n'$, as

$$\mathbf{R}_{y_{n,n-k}} = \mathbf{C}_n\mathbf{R}_x[k]\mathbf{C}_{n-k}^H. \qquad (5)$$

In practice, the computation of $\mathbf{R}_{y_{n,n-k}}$ in (5) must be approximated by taking an average over $\mathbf{y}[\tilde{n}N + n]$ at different indices $\tilde{n}$. We aim to reconstruct $\mathbf{R}_{\tilde{x}}[0]$ in (2), which is equivalent to reconstructing $\mathbf{R}_x[k]$ from $\{\mathbf{R}_{y_{n,n-k}}\}_{n=\max(0,k)}^{N-1+\min(0,k)}$ in (5) for $k = 0, \pm1, \ldots, \pm(N-1)$. Let us now take the realness of $\mathbf{C}_n$ into account, stack the columns of $\mathbf{R}_{y_{n,n-k}}$ in (5) into an $(M_nM_{n-k}) \times 1$ vector vec$(\mathbf{R}_{y_{n,n-k}})$, and rewrite (5) as

$$\text{vec}(\mathbf{R}_{y_{n,n-k}}) = (\mathbf{C}_{n-k} \otimes \mathbf{C}_n)\text{vec}(\mathbf{R}_x[k]), \qquad (6)$$

where vec(.) is the operator that stacks all columns of a matrix into a column vector and $\otimes$ denotes the Kronecker product operation. In the event where we have either $\mathbf{C}_{n-k} = [\,]$ ($M_{n-k} = 0$) or $\mathbf{C}_n = [\,]$ ($M_n = 0$) in (6) for a particular n, the corresponding vec$(\mathbf{R}_{y_{n,n-k}})$ and $\mathbf{C}_{n-k} \otimes \mathbf{C}_n$ are given by vec$(\mathbf{R}_{y_{n,n-k}}) = \mathbf{C}_{n-k} \otimes \mathbf{C}_n = [\,]$. Stacking vec$(\mathbf{R}_{y_{n,n-k}})$ in (6) in ascending order of n, for all $n = \max(0, k), \ldots, N-1+\min(0,k)$, into a $\gamma_k \times 1$ vector $\mathbf{r}_{y,k}$ with $\gamma_k = \sum_{n=\max(0,k)}^{N-1+\min(0,k)} M_nM_{n-k}$, we can then express $\mathbf{r}_{y,k}$ as

$$\mathbf{r}_{y,k} = \begin{bmatrix} \mathbf{C}_{\max(-k,0)} \otimes \mathbf{C}_{\max(0,k)} \\ \mathbf{C}_{\max(-k,0)+1} \otimes \mathbf{C}_{\max(0,k)+1} \\ \vdots \\ \mathbf{C}_{N-1+\min(-k,0)} \otimes \mathbf{C}_{N-1+\min(0,k)} \end{bmatrix} \text{vec}(\mathbf{R}_x[k])$$
$$= \boldsymbol{\Psi}_k\text{vec}(\mathbf{R}_x[k]), \quad k = 1-N, \ldots, N-2, N-1. \qquad (7)$$

3. PERFECT RECONSTRUCTION

Observe that we can reconstruct vec$(\mathbf{R}_x[k])$ in (7) from $\mathbf{r}_{y,k}$ using LS as long as the $\gamma_k \times T^2$ matrix $\boldsymbol{\Psi}_k$ has full column rank. Hence, $\mathbf{R}_{\tilde{x}}[0]$ in (2) can be reconstructed from $\{\mathbf{r}_{y,k}\}_{k=1-N}^{N-1}$ using LS as long as $\{\boldsymbol{\Psi}_k\}_{k=1-N}^{N-1}$ all have full column rank. In order to simplify the discussion, we consider the following remark.

Remark 1: The full column rank condition of $\boldsymbol{\Psi}_k$ can be achieved only if we have $\gamma_k \geq T^2$. In addition, observe in (7) *that each row of $\boldsymbol{\Psi}_k$ has only a single one in one entry and zeros elsewhere since the rows of $\mathbf{C}_n$ in* (6) *are selected from the rows of $\mathbf{I}_T$. Hence, $\boldsymbol{\Psi}_k$ will have full column rank if and only if each of its columns has at least a single one.*

Let us now introduce the following definition.

Definition 1: Define $\mathcal{C}_n$ as the set containing the indices of the rows of $\mathbf{I}_T$ used in $\mathbf{C}_n$. The set $\mathcal{C}_{n',n}$ is then defined as $\mathcal{C}_{n',n} = \{(i,j)|\forall i \in \mathcal{C}_{n'}, j \in \mathcal{C}_n\}$. Note that we generally have $\mathcal{C}_{n',n} \neq \mathcal{C}_{n,n'}$.

We can then present the following lemma.

Lemma 1: One row of $\mathbf{C}_{n'} \otimes \mathbf{C}_n$ will have a one in the $[(i-1)T + j]$-th entry and zeros elsewhere, if and only if $(i,j) \in \mathcal{C}_{n',n}$, i.e., $\mathbf{C}_{n'}$ contains the i-th row of $\mathbf{I}_T$ and $\mathbf{C}_n$ contains the j-th row of $\mathbf{I}_T$.

Proof: The proof directly follows from the property of the Kronecker product operation. □

Based on Lemma 1, the full column rank condition of $\boldsymbol{\Psi}_k$ in (7) is provided by the following theorem.

Theorem 1: $\boldsymbol{\Psi}_k$ in (7) *has full column rank if and only if*

$$\Gamma_k \equiv \bigcup_{n=\max(0,k)}^{N-1+\min(0,k)} \mathcal{C}_{n-k,n} = \{(1,1), (1,2), \ldots, (1,T),$$
$$(2,1), (2,2), \ldots, (T,T)\}. \qquad (8)$$

Proof: Recall from Lemma 1 that if we have $(i,j) \in \mathcal{C}_{n-k,n}$, $\mathbf{C}_{n-k} \otimes \mathbf{C}_n$ will have a one in the $[(i-1)T + j]$-th column. We can then observe that satisfying (8) is equivalent to ensuring that the $[(i-1)T + j]$-th column of *at least one* of the matrices $\{\mathbf{C}_{n-k} \otimes \mathbf{C}_n\}_{n=\max(0,k)}^{N-1+\min(0,k)}$ contains a one in one entry and zeros elsewhere for all $i, j = 1, 2, \ldots, T$. By taking the structure of $\boldsymbol{\Psi}_k$ in (7) into account, satisfying (8) also guarantees that every column of $\boldsymbol{\Psi}_k$ has at least a single one, which proves the sufficiency part of the theorem due to Remark 1. To prove the necessity part, assume that $\boldsymbol{\Psi}_k$ in (7) has full column rank but there is an (a, b) with $a, b \in \{1, 2, \ldots, T\}$ such that $(a,b) \notin \Gamma_k$. Based on (8) and Lemma 1, this means that none of the matrices $\{\mathbf{C}_{n-k} \otimes \mathbf{C}_n\}_{n=\max(0,k)}^{N-1+\min(0,k)}$ has a one in the $[(a-1)T + b]$-th column. If we take the structure of $\boldsymbol{\Psi}_k$ in (7) into account, this implies that the $[(a-1)T + b]$-th column of $\boldsymbol{\Psi}_k$ only contain zeros. Using Remark 1, this indicates that the full column rank condition of $\boldsymbol{\Psi}_k$ is violated, which contradicts our initial assumption. □

Consider the case of $k = N-1$ (we have $\boldsymbol{\Psi}_{N-1} = \mathbf{C}_0 \otimes \mathbf{C}_{N-1}$ in (7)) and $k = -N+1$ (we have $\boldsymbol{\Psi}_{1-N} = \mathbf{C}_{N-1} \otimes \mathbf{C}_0$ in (7)). Observe that the size of both $\boldsymbol{\Psi}_{N-1}$ and $\boldsymbol{\Psi}_{1-N}$ is $M_0M_{N-1} \times T^2$. For this specific case, we have the following theorem.

Theorem 2: $\boldsymbol{\Psi}_{N-1}$ and/or $\boldsymbol{\Psi}_{1-N}$ will have full column rank if and only if $\mathbf{C}_0 = \mathbf{C}_{N-1} = \mathbf{I}_T$.

Proof: The sufficiency part of this theorem is already clear as having $\mathbf{C}_0 = \mathbf{C}_{N-1} = \mathbf{I}_T$ leads to $\boldsymbol{\Psi}_{N-1} = \boldsymbol{\Psi}_{1-N} = \mathbf{I}_{T^2}$. The necessity part is shown for the full column rank condition of $\boldsymbol{\Psi}_{N-1}$ by considering Theorem 1, which requires $\mathcal{C}_{0,N-1} = \{(1,1), (1,2), \ldots, (1,T), (2,1), (2,2), \ldots, (T,T)\}$. This is identical to requiring $\mathcal{C}_0 = \mathcal{C}_{N-1} = \{1, 2, \ldots, T\}$ due to Definition 1. The proof is concluded for $\boldsymbol{\Psi}_{N-1}$. The proof for $\boldsymbol{\Psi}_{1-N}$ follows the same steps. □

Since Theorem 2 requires us to have $\mathbf{C}_0 = \mathbf{C}_{N-1} = \mathbf{I}_T$, our task is to design $\{\mathbf{C}_n\}_{n=1}^{N-2}$ based on Theorem 1. One specific case occurs when we restrict $\mathbf{C}_n$ to either $\mathbf{C}_n = \mathbf{I}_T$ or $\mathbf{C}_n = [\,]$, for each $n = 1, 2, \ldots, N-2$. In this case, $\boldsymbol{\Psi}_k$ in (7) must contain at least one self Kronecker product of $\mathbf{I}_T$ ($\mathbf{I}_T \otimes \mathbf{I}_T$) to preserve its full column rank condition. More precisely, for $k \in \{1-N, \ldots, N-1\}$, we need at least a pair of $n, n' \in \{0, 1, \ldots, N-1\}$ with $n - n' = k$ and $\mathbf{C}_n = \mathbf{C}_{n'} = \mathbf{I}_T$ to ensure the full column rank condition of $\boldsymbol{\Psi}_k$. At this stage, let us review the concept of a linear sparse ruler discussed in [3, 9, 10].

Definition 2: A length-$(N-1)$ linear sparse ruler is defined as a set $\mathcal{P} \subset \{0,1,\ldots,N-1\}$ such that $\{|p-p'||\forall p,p' \in \mathcal{P}\} = \{0,1,\ldots,N-1\}$. It is called minimal if no other linear sparse ruler of length $N-1$ exists with less elements.

Using Definition 2, it is easy to show that the full column rank condition of $\{\boldsymbol{\Psi}_k\}_{k=1-N}^{N-1}$, when we restrict $\mathbf{C}_n$ to either $\mathbf{C}_n = \mathbf{I}_T$ or $\mathbf{C}_n = [\,]$ for each n, follows the following theorem.

Theorem 3: Define W as the number of sampling matrices $\{\mathbf{C}_n\}_{n=0}^{N-1}$ that are set to $\mathbf{I}_T$, i.e., $\mathbf{C}_{n_w} = \mathbf{I}_T$, for $w = 0,1,\ldots,W-1$. The full column rank of all $\{\boldsymbol{\Psi}_k\}_{k=1-N}^{N-1}$ is ensured if and only if the set $\mathcal{W} = \{n_w | n_w \in \{0,1,\ldots,N-1\}, w = 0,1,\ldots,W-1\}$ is a linear sparse ruler.

Under the constraint that we have either $\mathbf{C}_n = \mathbf{I}_T$ or $\mathbf{C}_n = [\,]$ for each $n = 0,1,\ldots,N-1$, it is of interest to obtain the strongest possible compression rate $\tilde{M}/NT$. This is equivalent to minimizing the cardinality of $\mathcal{W}$ in Theorem 3 under the condition that $\mathcal{W}$ is a length-$(N-1)$ linear sparse ruler. This boils down to a length-$(N-1)$ minimal linear sparse ruler problem [3, 9], whose solution minimizes $\tilde{M}/NT$ under the aforementioned constraint while maintaining the identifiability of $\{\mathrm{vec}(\mathbf{R}_x[k])\}_{k=1-N}^{N-1}$ in (7). Once $\mathbf{R}_{\tilde{x}}[0]$ in (2) is reconstructed, we can reconstruct the cyclic autocorrelation sequence $\tilde{r}_x[f,\tau]$ from $\mathbf{R}_{\tilde{x}}[0]$ using (1a) and the cyclic power spectrum $s_x[f,\phi)$ from $\tilde{r}_x[f,\tau]$ using (1b).

4. LIMITED CORRELATION SUPPORT

Let us now assume that the support of $\mathbf{R}_x[k]$ in (5) is limited to $1-N \le k \le N-1$ and that the elements of $\mathbf{R}_x[k]$ are all close to 0 for $|k| \ge N$. In addition, unlike in the previous sections, we also exploit the correlation between $\tilde{\mathbf{y}}[\tilde{n}]$ in (4) and its neighboring blocks $\tilde{\mathbf{y}}[\tilde{n}+1]$ and $\tilde{\mathbf{y}}[\tilde{n}-1]$. While we have written $\mathbf{R}_{y_{n,n-k}}$ in (5) as $\mathbf{R}_{y_{n,n-k}} = E[\mathbf{y}[\tilde{n}N+n]\mathbf{y}^H[\tilde{n}N+n-k]]$ we can now also write it, for example, as $\mathbf{R}_{y_{n,n-k}} = E[\mathbf{y}[\tilde{n}N+n]\mathbf{y}^H[(\tilde{n}-1)N+n-k+N]]$ for $k>0$ or as $\mathbf{R}_{y_{n,n-k}} = E[\mathbf{y}[\tilde{n}N+n]\mathbf{y}^H[(\tilde{n}+1)N+n-k-N]]$ for $k<0$. By considering (3), we can now also rewrite (5) as

$$\mathbf{R}_{y_{n,n-k}} = \mathbf{C}_n \mathbf{R}_x[k] \mathbf{C}^H_{(n-k) \bmod N}, \tag{9}$$

with $n \bmod N$ the remainder of the integer division n/N. We stack the columns of $\mathbf{R}_{y_{n,n-k}}$ into $\mathrm{vec}(\mathbf{R}_{y_{n,n-k}})$ as in Section 2 but we now cascade $\mathrm{vec}(\mathbf{R}_{y_{n,n-k}})$, for all $n = 0,1,\ldots,N-1$, in increasing order of n into a $\tilde{\gamma}_k \times 1$ vector $\tilde{\mathbf{r}}_{y,k}$ with $\tilde{\gamma}_k = \sum_{n=0}^{N-1} M_n M_{(n-k) \bmod N}$. We can then express $\tilde{\mathbf{r}}_{y,k}$ as

$$\begin{aligned}\tilde{\mathbf{r}}_{y,k} &= \begin{bmatrix} \mathbf{C}_{(-k) \bmod N} \otimes \mathbf{C}_0 \\ \mathbf{C}_{(1-k) \bmod N} \otimes \mathbf{C}_1 \\ \vdots \\ \mathbf{C}_{(N-1-k) \bmod N} \otimes \mathbf{C}_{N-1} \end{bmatrix} \mathrm{vec}(\mathbf{R}_x[k]) \\ &= \tilde{\boldsymbol{\Psi}}_k \mathrm{vec}(\mathbf{R}_x[k]), \quad k = 1-N,\ldots,N-2,N-1. \end{aligned}\tag{10}$$

We can again reconstruct $\mathrm{vec}(\mathbf{R}_x[k])$ from $\tilde{\mathbf{r}}_{y,k}$ in (10) using LS where the condition that ensures a full column rank $\tilde{\boldsymbol{\Psi}}_k$ in (10) can be found by following the procedure in Section 3. By applying Remark 1 to $\tilde{\boldsymbol{\Psi}}_k$, using Definition 1 as well as Lemma 1, and following an analysis similar to the proof of Theorem 1, we can find that $\tilde{\boldsymbol{\Psi}}_k$ in (10) has full column rank if and only if

$$\begin{aligned}\tilde{\Gamma}_k \equiv \bigcup_{n=0}^{N-1} \mathcal{C}_{(n-k) \bmod N, n} &= \{(1,1),(1,2),\ldots,(1,T), \\ &(2,1),(2,2),\ldots,(T,T)\}. \end{aligned}\tag{11}$$

It is interesting to observe that, for this limited correlation support case, we do not have any condition similar to Theorem 2.

We again focus on the case where $\mathbf{C}_n$ is restricted to either $\mathbf{C}_n = \mathbf{I}_T$ or $\mathbf{C}_n = [\,]$, for each $n = 0,1,\ldots,N-1$. In order to maintain the full column rank condition of $\tilde{\boldsymbol{\Psi}}_k$ in (10), $\tilde{\boldsymbol{\Psi}}_k$ must contain at least one self Kronecker product of $\mathbf{I}_T$, which is equivalent, for each $k \in \{1-N,\ldots,N-1\}$, to having at least a pair of $n,n' \in \{0,1,\ldots,N-1\}$ with $(n-n') \bmod N = k$ and $\mathbf{C}_n = \mathbf{C}_{n'} = \mathbf{I}_T$. We now look at the concept of a circular sparse ruler discussed in [9].

Definition 3: A circular sparse ruler of length $N-1$ is defined as a set $\mathcal{Q} \subset \{0,1,\ldots,N-1\}$ such that $\{(q-q') \bmod N | \forall q,q' \in \mathcal{Q}\} = \{0,1,\ldots,N-1\}$. It is called minimal if no other circular sparse ruler of length $N-1$ exists with less elements.

Using Definition 3, it is obvious that the full column rank condition of $\{\tilde{\boldsymbol{\Psi}}_k\}_{k=1-N}^{N-1}$, when we restrict $\{\mathbf{C}_n\}_{n=0}^{N-1}$ to either $\mathbf{C}_n = \mathbf{I}_T$ or $\mathbf{C}_n = [\,]$, follows the following theorem.

Theorem 4: Recall from Theorem 3 that $\mathcal{W} = \{n_w | n_w \in \{0,1,\ldots,N-1\}, w = 0,1,\ldots,W-1\}$ with $\{\mathbf{C}_{n_w}\}_{w=0}^{W-1} = \mathbf{I}_T$. The full column rank of all $\{\tilde{\boldsymbol{\Psi}}_k\}_{k=1-N}^{N-1}$ in (10) is guaranteed if the set $\mathcal{W}$ is a circular sparse ruler.

Again, if we restrict $\mathbf{C}_n$ to either $\mathbf{C}_n = \mathbf{I}_T$ or $\mathbf{C}_n = [\,]$, for each $n = 0,1,\ldots,N-1$, the best compression $\tilde{M}/NT$ is obtained by minimizing the cardinality of $\mathcal{W}$ in Theorem 4 under the condition that $\mathcal{W}$ is a length-$(N-1)$ circular sparse ruler, which boils down to a length-$(N-1)$ minimal circular sparse ruler problem [9].

5. SELECTION MATRIX CONSTRUCTION

5.1. General Case

We focus on the formation of $\{\mathbf{C}_n\}_{n=0}^{N-1}$ for the general case discussed in Sections 2 and 3. Recall that Theorem 2 requires us to have $\mathbf{C}_0 = \mathbf{C}_{N-1} = \mathbf{I}_T$ to obtain full column rank $\boldsymbol{\Psi}_{N-1}$ and $\boldsymbol{\Psi}_{1-N}$. Also note from (7) that, for a larger n, less number of matrices $\{\mathbf{C}_n\}_{n=0}^{N-1}$ are contained in $\boldsymbol{\Psi}_n$. It is thus reasonable to construct $\{\mathbf{C}_n\}_{n=0}^{N-1}$ by evaluating the rank condition of $\boldsymbol{\Psi}_n$ starting from $n = N-1$ to $n = 0$. Note from (7) that the full column rank of $\boldsymbol{\Psi}_{-n}$ is ensured once $\boldsymbol{\Psi}_n$ has full column rank. Let us start by evaluating $\boldsymbol{\Psi}_{N-2} = [(\mathbf{C}_0 \otimes \mathbf{C}_{N-2})^T, (\mathbf{C}_1 \otimes \mathbf{C}_{N-1})^T]^T = [(\mathbf{I}_T \otimes \mathbf{C}_{N-2})^T, (\mathbf{C}_1 \otimes \mathbf{I}_T)^T]^T$. Assuming $\mathcal{C}_{N-2} = \{m_1, m_2, \ldots, m_{M_{N-2}}\}$ and $\mathcal{C}_1 = \{\tilde{m}_1, \tilde{m}_2, \ldots, \tilde{m}_{M_1}\}$, we have from Definition 1 $\mathcal{C}_{0,N-2} = \{(1,m_1),(1,m_2),\ldots,(1,m_{M_{N-2}}),(2,m_1),(2,m_2),\ldots,(T,m_{M_{N-2}})\}$ and $\mathcal{C}_{1,N-1} = \{(\tilde{m}_1,1),(\tilde{m}_1,2),\ \ldots,(\tilde{m}_1,T),(\tilde{m}_2,1),(\tilde{m}_2,2),\ldots,(\tilde{m}_{M_1},T)\}$. Note from Theorem 1 that $\boldsymbol{\Psi}_{N-2}$ has full column rank if and only if $\Gamma_{N-2} = \mathcal{C}_{0,N-2} \cup \mathcal{C}_{1,N-1}$ satisfies (8). Observe that this is possible only if we have at least one of $\mathbf{C}_{N-2}$ and $\mathbf{C}_1$ equal to $\mathbf{I}_T$. Once we set either $\mathbf{C}_{N-2}$ or $\mathbf{C}_1$ to $\mathbf{I}_T$, it is reasonable in this step to set the other one to [] as we want to minimize the compression rate. We proceed to evaluate $\boldsymbol{\Psi}_{N-3} = [(\mathbf{C}_0 \otimes \mathbf{C}_{N-3})^T, (\mathbf{C}_1 \otimes \mathbf{C}_{N-2})^T, (\mathbf{C}_2 \otimes \mathbf{C}_{N-1})^T]^T = [(\mathbf{I}_T \otimes \mathbf{C}_{N-3})^T, (\mathbf{C}_1 \otimes \mathbf{C}_{N-2})^T, (\mathbf{C}_2 \otimes \mathbf{I}_T)^T]^T$. As we have set either $\mathbf{C}_{N-2}$ or $\mathbf{C}_1$ to $\mathbf{I}_T$ and the other one to [], we have $\mathbf{C}_1 \otimes \mathbf{C}_{N-2} = [\,]$ and $\boldsymbol{\Psi}_{N-3} = [(\mathbf{I}_T \otimes \mathbf{C}_{N-3})^T, (\mathbf{C}_2 \otimes \mathbf{I}_T)^T]^T$. Using the same analysis used when we considered $\boldsymbol{\Psi}_{N-2}$, we clearly have to set either $\mathbf{C}_{N-3}$ or $\mathbf{C}_2$ to $\mathbf{I}_T$. Again, it is reasonable to set the other matrix to [ ].

At this stage, it can be found that we might face two possibilities with respect to the next considered $\boldsymbol{\Psi}_n$. First, we might be again required to set one of two selection matrices to $\mathbf{I}_N$ in order to ensure the full column rank of the next considered $\boldsymbol{\Psi}_n$. Second, we might have an option to set two selection matrices to [ ] while maintaining the full column rank of the next considered $\boldsymbol{\Psi}_n$. These two possibilities are repeatedly faced for every considered $\boldsymbol{\Psi}_n$ as long as we set one selection matrix to $\mathbf{I}_N$ and the other one to [ ], when the first one occurs, and we set both selection matrices to [], when the second one occurs. If this procedure is followed, we can find that each of $\{\mathbf{C}_n\}_{n=0}^{N-1}$ is equal to either $\mathbf{I}_T$ or [ ], which is the constraint that we considered when we formulated Theorem 3. Considering that this procedure is a reasonable way to minimize the compression rate, we suggest to design $\{\mathbf{C}_n\}_{n=0}^{N-1}$ by following Theorem 3 and minimizing the cardinality of $\mathcal{W}$ in Theorem 3, i.e., solving the minimal linear sparse ruler problem. Many solutions for the minimal linear sparse ruler problem have been tabulated.

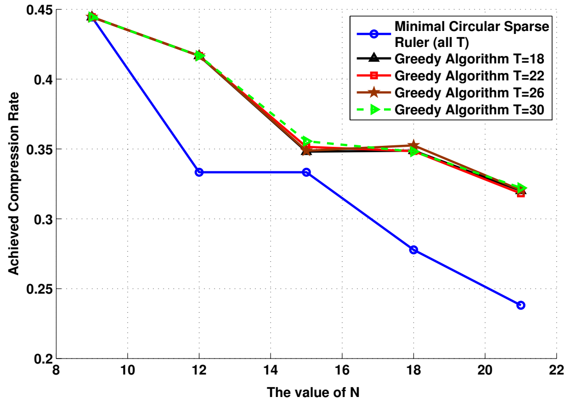

Fig. 1. The achievable compression rate for the selection matrices designed using the greedy algorithm in Table 1 and those designed based on the minimal circular sparse ruler.

5.2. Limited Correlation Support Case

Recall that we do not have any condition similar to Theorem 2 for the limited correlation support case. Hence, unlike in Section 5.1, it is reasonable for this case to expect that there might be better options than setting $\mathbf{C}_n$ either to $\mathbf{I}_T$ or [ ] for all n, which leads to a circular sparse ruler based matrices design in Theorem 4. This motivates us to propose an algorithm, provided in Table 1, for designing $\{\mathbf{C}_n\}_{n=0}^{N-1}$ for the limited correlation support case in Section 4. The main for loop in step 4 of Table 1 decides on some rows of $\{\mathbf{C}_n\}_{n=0}^{N-1}$ while focusing on the full column rank condition of $\mathbf{\Psi}_k$. Only non-negative k is considered since $\mathbf{\Psi}_{-k}$ has full column rank once $\mathbf{\Psi}_k$ has full column rank. Each element of the indicator matrix $\mathbf{Z}^{(k)}$ is equal to either 0 or 1. $[\mathbf{Z}^{(k)}]_{i,j} = 1$ indicates that the $((i-1)T+j)$-th column of $\mathbf{\Psi}_k$ already has at least a single one. The while loop in step 5 checks if all columns of the considered $\mathbf{\Psi}_k$ have at a least single one and, as long as this is not the case, it will iterate and add one row to one of the matrices $\{\mathbf{C}_n\}_{n=0}^{N-1}$. The inner for loop in steps 6-9 determines the indices of the candidate rows of $\mathbf{I}_N$ to be added to each of the matrices $\{\mathbf{C}_n\}_{n=0}^{N-1}$. However, steps 10-11 will decide that only one of the matrices $\{\mathbf{C}_n\}_{n=0}^{N-1}$ is going to be updated in each iteration of the while loop. Fig. 1 describes the achievable compression rate for the selection matrices designed using the greedy algorithm in Table 1. We run the algorithm 1000 times and pick the matrix offering the best compression for each N and T. Here, T is varied from $T = 18$ to $T = 30$ and N is varied from $N = 9$ to $N = 21$. We also plot the achievable compression rate for the selection matrices designed based on the minimal circular sparse ruler, which is independent of T since the minimal circular sparse ruler based $\mathbf{C}_n$ is set to either $\mathbf{I}_T$ or []. Observe that the minimal circular sparse ruler based selection matrices offer a stronger compression than the ones produced by the greedy algorithm. However, the minimal circular sparse ruler problem is a combinatorial problem whose solution has to be found using a brute force, which might be computationally infeasible for a large N. Moreover, the circular sparse ruler based selection matrices lead to many Nyquist-spaced samples and thus, the greedy algorithm might be more attractive for some applications.

Table 1. A greedy algorithm to find a sub-optimal solution for $\{\mathbf{C}_n\}_{n=0}^{N-1}$ for limited correlation support case.

	Algorithm
1:	Introduce $\mathbf{Z}^{(k)}$ as a $T \times T$ indicator matrix with respect to $\mathbf{\Psi}_k$ and denote its element at the i-th row and the j-th column by $[\mathbf{Z}^{(k)}]_{i,j}$.
2:	For all $k = 0, \dots, N-1$, initialize $\mathbf{Z}^{(k)} = \mathbf{0}_{T\times T}$ with $\mathbf{0}_{T\times T}$ a $T \times T$ matrix containing only zeros.
3:	For $k = N-1$, randomly select $i, j \in \{1, 2, \dots, T\}$, and set $\mathcal{C}_0 = i$, $\mathcal{C}_{N-1} = j$, and $[\mathbf{Z}^{(N-1)}]_{i,j} = 1$.
4:	for $k = N-1$ to 0 in decreasing order do
5:	while $\mathbf{Z}^{(k)}$ has at least one zero entry do
6:	for $n = 0$ to $N-1$ do
7:	Define a set $\Xi = \{1, 2, \dots, T\} \setminus \mathcal{C}_n$ and a function $f(g'_n, n)$ as $f(g'_n, n) = \sum_{i' \in \mathcal{C}_{(n-k) \bmod N}} (1 - [\mathbf{Z}^{(k)}]_{i',g'_n}) + \sum_{i'' \in \mathcal{C}_{(n+k) \bmod N}} (1 - [\mathbf{Z}^{(k)}]_{g'_n,i''})$.
8:	Search in Ξ for the element g_n that satisfies: $g_n = \arg\max_{g'_n \in \Xi} f(g'_n, n)$, randomly pick one if we have multiple g_n, and set $h_n = f(g_n, n)$.
9:	end for
10:	Find $\bar{n}$ such that $h_{\bar{n}}$ is the maximum of $\{h_n\}_{n=0}^{N-1}$, randomly pick one if have multiple maxima $h_{\bar{n}}$, and update $\mathcal{C}_{\bar{n}}$ to $\mathcal{C}_{\bar{n}} = \mathcal{C}_{\bar{n}} \cup \{g_{\bar{n}}\}$.
11:	For all $i' \in \mathcal{C}_{(n-k) \bmod N}$ and $i'' \in \mathcal{C}_{(n+k) \bmod N}$, set $[\mathbf{Z}^{(k)}]_{i',g_{\bar{n}}}$ and $[\mathbf{Z}^{(k)}]_{g_{\bar{n}},i''}$ to 1, respectively.
12:	end while
13:	if $k > 0$ do
14:	for $n = 0$ to $N-1$ do
15:	For all $i' \in \mathcal{C}_{(n-k+1) \bmod N}$, $i'' \in \mathcal{C}_{(n+k-1) \bmod N}$ and $j' \in \mathcal{C}_n$, set $[\mathbf{Z}^{(k-1)}]_{i',j'}$ and $[\mathbf{Z}^{(k-1)}]_{j',i''}$ to 1.
16:	end for
17:	end if
18:	end for

6. CONCLUSION

The knowledge of the cyclic period allows us to set the span of the random linear projections such that the block Toeplitz structure emerges in the WSCS signal correlation matrix. We have shown how to exploit this block Toeplitz structure, have introduced compression using non-uniform sampling, and have presented the reconstruction problem as an overdetermined system. We have presented the condition for the system matrix to have full column rank, which allows for LS reconstruction of WSCS signal correlation matrix. We considered two cases, the general case and the limited correlation support case. For the general case, we proposed the minimal linear sparse ruler based sampling matrices design as a reasonable way to approximately minimize the compression rate. For the limited correlation support case, we proposed a greedy algorithm to find a suboptimal solution for the sampling matrices, which might be more attractive in particular situations than the minimal circular sparse ruler based solution, though the latter appears to offer a stronger compression.

7. REFERENCES

[1] E.J. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, February 2006.

[2] M. Mishali and Y.C. Eldar, “Blind multiband signal reconstruction: Compressed sensing for analog signals,” *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 993–1009, March 2009.

[3] D.D. Ariananda and G. Leus, “Compressive wideband power spectrum estimation,” *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4775–4789, September 2012.

[4] M.A Lexa, M.E Davies, J.S. Thompson, and J. Nikolic, “Compressive power spectral density estimation,” in *Proc. IEEE International Conference Acoustics, Speech and Signal Processing (ICASSP 2011)*. Prague, May 2011, pp. 3884–3887.

[5] C.P. Yen, Y. Tsai, and X. Wang, “Wideband spectrum sensing based on sub-nyquist sampling,” *IEEE Transactions on Signal Processing*, vol. 61, no. 12, pp. 3028–3040, June 2013.

[6] Z. Tian, Y. Tafesse, and B.M. Sadler, “Cyclic feature detection with sub-nyquist sampling for wideband spectrum sensing,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 1, pp. 58–69, February 2012.

[7] D. Cohen, E. Rebeiz, Y.C. Eldar, and D. Cabric, “Cyclic spectrum reconstruction and cyclostationary detection from subnyquist samples,” in *Proc. IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. Darmstadt, June 2013, pp. 420–424.

[8] G. Leus and Z. Tian, “Recovering second-order statistics from compressive measurements,” in *Proc. IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*. San Juan, December 2011, pp. 337–340.

[9] D. Romero and G. Leus, “Compressive covariance sampling,” in *Proc. Information Theory and Applications Workshop (ITA 2013)*. San Diego, February 2013.

[10] J. Leech, “On the representation of 1,2,...,n by differences,” *Journal of the London Mathematical Society*, vol. 31, pp. 160–169, April 1956.