MULTI-COSET SAMPLING FOR POWER SPECTRUM BLIND SENSING

Dyonisius Dony Ariananda, Geert Leus*, and Zhi Tian†*

* Faculty of EEMCS, Delft University of Technology, Mekelweg 4, 2628CD, Delft, The Netherlands
† Dept. of Electrical & Computer Engineering Michigan Technological University, Houghton, MI, USA

ABSTRACT

Power spectrum blind sampling (PSBS) consists of a sampling procedure and a reconstruction method that is able to recover the unknown power spectrum of a random signal from the obtained sub-Nyquist-rate samples. It differs from spectrum blind sampling (SBS) that aims to recover the spectrum instead of the power spectrum of the signal. In this paper, a PSBS solution is first presented based on a periodic sampling procedure. Then, a multi-coset implementation for this sampling procedure is developed by solving the so-called minimal sparse ruler problem, and the coprime sampling technique is tailored to fit into the PSBS framework as well. It is shown that the proposed multi-coset implementation based on minimal sparse rulers offers advantages over coprime sampling in terms of reduced sampling rates, increased flexibility and an extended range of estimated auto-correlation lags. These benefits arise without putting any sparsity constraint on the power spectrum. Application to sparse power spectrum recovery is also illustrated.

Index Terms— Multi-coset sampling, sparse ruler, coprime sampling

1. INTRODUCTION

In recent years, spectrum estimation has gained renewed interest because of its significance in the field of cognitive radio networks, where unlicensed secondary users can opportunistically utilize licensed frequency bands when the licensed owners are not active. For such networks, unlicensed users are required to sense the wireless environment over a broad frequency band and identify spectrum holes in the licensed bands that can subsequently be exploited to establish a communication link. In general, a wide frequency range has to be sensed, which requires power-hungry high-rate analog-to-digital converters (ADCs). This issue has spurred researchers to examine specific features of the licensed spectrum that can be exploited to alleviate the requirements of the ADCs. Sparsity of the spectrum or its derivative (the so-called edge spectrum), are often considered [1, 2, 3]. These features enable one to decrease the signal sampling rate below the Nyquist rate while maintaining perfect reconstruction in the noise-free case. Multi-coset sampling investigated in [2, 4] is one popular way to reduce the sampling rate for the case of multiband signals having a frequency support on a union of finite intervals. Similarly, [3] evaluates sub-Nyquist sampling for sparse multiband analog signals by means of a so-called modulated wideband converter (MWC) consisting of multiple branches, each of which employs a different mixing function followed by low-pass filtering and low-rate uniform sampling. Both of the above approaches can be cast into a compressive sampling framework. The reconstruction process can be accomplished by using any sparse recovery methods such as the LASSO algorithm [5], or even using traditional methods such as multiple signal classification (MUSIC) [2] or the minimum variance distortionless response (MVDR) method [6]. The methods discussed in [2, 3] are known as spectrum blind sampling (SBS), where the objective is to sample the signal at minimal rate and reconstruct the unknown spectrum from these samples, given that the spectrum is sparse. It has been found in these works that, for most signals, the minimum average sampling rate is given by the Landau lower bound (as studied in [4]), which is equal to the Nyquist rate multiplied with the frequency occupancy ratio. Nevertheless, in the worst scenario, the minimum average sampling rate increases and it is given by the minimum of twice the Landau lower bound and the Nyquist rate.

While all the above works focus on spectral estimation and intend to obtain perfect reconstruction of the original signal, only the power spectrum (aka the power spectral density), or equivalently, the auto-correlation function, needs to be reconstructed for spectrum sensing applications. In [1, 7, 8], power spectrum estimation approaches have been developed by focusing on the auto-correlation function instead of the original signal itself. The wideband spectrum sensing method in [1] basically exploits the inherent sparsity feature of the edge spectrum. To directly perform compressive sampling on the received signal, [7] exploits the relationship between the auto-correlation function of the Nyquist-rate samples and that of the compressive measurements. However, the measurements are assumed to be wide-sense stationary, which does not hold for most compressive sampling matrices. [8]

This work is supported in part by NWO-STW under the VICI program (project 10382) and US NSF grants #1017887 and #0925881.

employs compressive sampling for power spectrum sensing by introducing K wideband filters in order to identify the occupancy of N available channels with $K < N$. However, [8] only exploits the energy at the output of each filter while more statistical properties could be exploited. Another method that aims to reconstruct the auto-correlation function using sub-Nyquist sampling is provided by [9]. It is based on the use of two uniform sub-Nyquist samplers with sampling periods that are coprime multiples of the Nyquist period.

This paper focuses on the efficient reconstruction of the signal power spectrum. The goal is to develop effective sampling procedures for the power spectrum blind sampling (PSBS) approach introduced in [10]. To this end, a multi-coset implementation is devised by solving the minimal sparse ruler problem. The coprime sampling scheme in [9] is also considered, and tailored to fit into the PSBS framework as a special case of the ruler problem. These two schemes are compared via both analysis and simulations. In general, the developed multi-coset sampling procedures for PSBS can considerably reduce the sampling rate requirements by making proper use of the spectral correlation property, without any sparsity constraints on the power spectrum. Further rate reduction is possible when the signal of interest is sparse, which is illustrated as well.

2. PERIODIC SAMPLING

Consider a spectrum sensing application, where the task is to sense the power spectrum of a wide-sense stationary signal $x(t)$. The signal $x(t)$ is assumed to be complex-valued (e.g., the complex envelope of the observed real-valued signal) and bandlimited with bandwidth $1/T$, which also indicates the Nyquist rate. As depicted in Fig. 1, a practical sampling device with M branches is employed, where the ith branch modulates the signal $x(t)$ with a possibly complex-valued periodic waveform $p_i(t)$ of period NT followed by an integrate-and-dump device with period NT (thus with rate equal to $1/N$ times the Nyquist rate). Consequently, the output of the ith branch at the kth sampling index can be written as

$$\begin{aligned} y_i[k] &= \int_{kNT}^{(k+1)NT} p_i(t)x(t)dt \\ &= \int_{kNT}^{(k+1)NT} c_i(t-kNT)x(t)dt \qquad (1) \end{aligned}$$

where $c_i(t)$ yields one period of $p_i(t)$, i.e., $c_i(t) = p_i(t)$ for $0 \le t < NT$ and $c_i(t) = 0$ elsewhere. If $c_i(t)$ is assumed to be a piecewise constant function with constant values in every interval of length T, i.e., $c_i(t) = c_i[-n]$ for $nT \le t < (n+1)T$ with $n = 0, 1, \ldots, N-1$, we can re-express (1) as

$$\begin{aligned} y_i[k] &= \sum_{n=0}^{N-1} c_i[-n] \int_{(kN+n)T}^{(kN+n+1)T} x(t)dt \\ &= \sum_{n=0}^{N-1} c_i[-n]x[kN+n] \\ &= \sum_{n=1-N}^{0} c_i[n]x[kN-n] \qquad (2) \end{aligned}$$

where $x[n]$ can be perceived as the output of an integrate-and-dump process with period T (thus with rate equal to the Nyquist rate) applied to $x(t)$, which is not explicitly carried out due to the high complexity. The average sampling rate of this periodic sampler is equal to the Nyquist rate multiplied by M/N, and we will use $M < N$ to reduce complexity.

The sampler above is similar to the modulated wideband converter of [3], where the values of $c_i[n]$ are randomly selected, e.g., from a random binary set ± 1. It is also possible to implement efficient multi-coset sampling using this sampling device, which is the focus of Section 4.

It is worth noting that (2) can be interpreted as a digital filtering operation of $x[n]$ by the filter $c_i[n]$ of length N followed by an N-fold decimation, i.e., $y_i[k] = z_i[kN]$, where

$$z_i[n] = c_i[n] \star x[n] = \sum_{m=1-N}^{0} c_i[m]x[n-m]$$

with $\star$ representing the convolution operator.

The objective of the PSBS problem is to estimate the power spectrum of $x(t)$ based on the obtained samples $\{y_i[k]\}_{i,k}$, which is equivalent to estimating the corresponding power spectrum of $x[n]$. Clearly, the power spectrum or power spectral density of $x[n]$ is given by

$$P_x(\omega) = \sum_{n=-\infty}^{\infty} r_x[n]e^{-jn\omega}, \quad 0 \le \omega < 2\pi$$

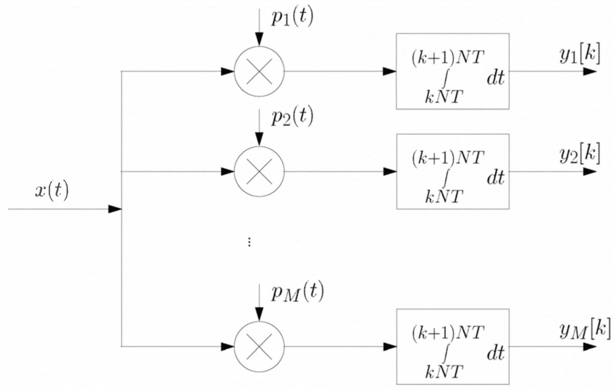


Fig. 1. The adopted sampling device consists of M branches, where each branch modulates the input signal with a periodic waveform followed by an integrate-and-dump process.

where $r_x[n]$ denotes the auto-correlation function of $x[n]$, given by $r_x[n] = E\{x[m]x^*[m-n]\}$. Hence, reconstructing the power spectrum $P_x(\omega)$ amounts to reconstructing the auto-correlation function $r_x[n]$.

The contribution of this work is that we will exploit all the M^2 different cross-spectra of $y_i[k]$ with $y_j[k]$ for $i,j = 0,1,\ldots,M-1$, which will allow rate-compression without imposing any sparsity constraint on $x(t)$. The cross-spectrum or cross spectral density of $y_i[k]$ with $y_j[k]$ is defined as

$$P_{y_i,y_j}(\omega) = \sum_{k=-\infty}^{\infty} r_{y_i,y_j}[k]e^{-jk\omega}, \quad 0 \le \omega < 2\pi$$

where $r_{y_i,y_j}[k] = E\{y_i[l]y_j^*[l-k]\}$ is the cross-correlation function of $y_i[k]$ with $y_j[k]$. These ensemble-mean quantities $\{r_{y_i,y_j}[k]\}_{i,j}$ can be estimated by their sample-averages, which in turn yield the estimates of $\{P_{y_i,y_j}(\omega)\}_{i,j}$.

3. RECONSTRUCTION APPROACH

In this section, we introduce a method to recover $r_x[n]$ given $r_{y_i,y_j}[k]$ for $i,j = 0,1,\ldots,M-1$. Since $y_i[k] = z_i[kN]$, the cross-correlation function of $y_i[k]$ with $y_j[k]$ can be written as the N-fold decimated version of the cross-correlation function of $z_i[n]$ with $z_j[n]$, as follows:

$$\begin{aligned} r_{y_i,y_j}[k] &= E\{y_i[l]y_j^*[l-k]\} \\ &= E\{z_i[lN]z_j^*[(l-k)N]\} = r_{z_i,z_j}[kN]. \end{aligned} \quad (3)$$

It is straightforward to show that $r_{z_i,z_j}[n]$ can be written as

$$r_{z_i,z_j}[n] = r_{c_i,c_j}[n] \star r_x[n] = \sum_{m=-N+1}^{N-1} r_{c_i,c_j}[m]r_x[n-m] \quad (4)$$

where $r_{c_i,c_j}[n]$ is the "deterministic" cross-correlation function between $c_i[n]$ and $c_j[n]$:

$$r_{c_i,c_j}[n] = c_i[n] \star c_j^*[-n] = \sum_{m=1-N}^{0} c_i[m]c_j^*[m-n]. \quad (5)$$

From (3) and (4), it holds that

$$\begin{aligned} r_{y_i,y_j}[k] &= r_{z_i,z_j}[kN] = \sum_{m=-N+1}^{N-1} r_{c_i,c_j}[m]r_x[kN-m] \\ &= \sum_{l=0}^{1} \mathbf{r}_{c_i,c_j}^T[l]\mathbf{r}_x[k-l] \end{aligned} \quad (6)$$

which uses the following definitions:

$$\begin{aligned} \mathbf{r}_{c_i,c_j}[k] = [&r_{c_i,c_j}[kN], r_{c_i,c_j}[kN-1], \ldots, \\ &r_{c_i,c_j}[(k-1)N+1]]^T \\ \mathbf{r}_x[k] = [&r_x[kN], r_x[kN+1], \ldots, r_x[(k+1)N-1]]^T. \end{aligned} \quad (7)$$

By cascading the M^2 different cross-correlation functions $r_{y_i,y_j}[k]$, we form the vector $\mathbf{r}_y[k] = [\ldots, r_{y_i,y_j}[k], \ldots]^T$ of length M^2 for $i,j = 0,1,\ldots,M-1$, which can be derived from (6) as

$$\mathbf{r}_y[k] = \sum_{l=0}^{1} \mathbf{R}_c[l]\mathbf{r}_x[k-l] \quad (8)$$

where $\mathbf{R}_c[k]$ is the $M^2 \times N$ matrix given by $\mathbf{R}_c[k] = [\ldots, \mathbf{r}_{c_i,c_j}[k], \ldots]^T$, for $i,j = 0,1,\ldots,M-1$.

Suppose that $\mathbf{r}_y[k]$ has a support limited to $-L \le k \le L$. Accordingly, the support of $r_x[n]$ is limited to $-LN \le n \le LN$, which means that the support of $\mathbf{r}_x[k]$ is also limited to $-L \le k \le L$. All these quantities are gathered into vectors, as follows:

$$\mathbf{r}_y = [\mathbf{r}_y^T[0], \mathbf{r}_y^T[1], \ldots, \mathbf{r}_y^T[L], \mathbf{r}_y^T[-L], \ldots, \mathbf{r}_y^T[-1]]^T \quad (9)$$

$$\mathbf{r}_x = [\mathbf{r}_x^T[0], \mathbf{r}_x^T[1], \ldots, \mathbf{r}_x^T[L], \mathbf{r}_x^T[-L], \ldots, \mathbf{r}_x^T[-1]]^T. \quad (10)$$

Based on (8), and the fact that the first column of $\mathbf{R}_c[1]$ as well as the last $N-1$ entries of $\mathbf{r}_x[L]$ are zero, the relation between $\mathbf{r}_y$ and $\mathbf{r}_x$ can finally be expressed as

$$\mathbf{r}_y = \mathbf{R}_c\mathbf{r}_x \quad (11)$$

where $\mathbf{R}_c$ is the $(2L+1)M^2 \times (2L+1)N$ matrix given by

$$\mathbf{R}_c = \begin{bmatrix} \mathbf{R}_c[0] & & & & \mathbf{R}_c[1] \\ \mathbf{R}_c[1] & \mathbf{R}_c[0] & & & \\ & \mathbf{R}_c[1] & \mathbf{R}_c[0] & & \\ & & \ddots & \ddots & \\ & & & \mathbf{R}_c[1] & \mathbf{R}_c[0] \end{bmatrix}. \quad (12)$$

If $\mathbf{R}_c$ has full column rank, it is possible to solve (11) using least-squares (LS). The necessary condition is $M^2 \ge N$.

It is possible to simplify the inverse problem for (11). Note that $\mathbf{R}_c$ is a block circulant matrix with blocks of size $M^2 \times N$, and hence can easily be converted into a block diagonal matrix $\mathbf{Q}_c$ with blocks of size $M^2 \times N$. This can be performed simply by using the $(2L+1)$-point (inverse) discrete Fourier transform ((I)DFT), that is,

$$\mathbf{R}_c = (\mathbf{F}_{2L+1}^{-1} \otimes \mathbf{I}_{M^2})\mathbf{Q}_c(\mathbf{F}_{2L+1} \otimes \mathbf{I}_N)$$

where $\mathbf{F}_{2L+1}$ is the $(2L+1) \times (2L+1)$ DFT matrix, and $\mathbf{Q}_c = \text{diag}\{\mathbf{Q}_c(0), \mathbf{Q}_c(2\pi\frac{1}{2L+1}), \ldots, \mathbf{Q}_c(2\pi\frac{2L}{2L+1})\}$ with $\mathbf{Q}_c(\omega)$ being the $M^2 \times N$ matrix spectrum of the $M^2 \times N$ matrix filter $\mathbf{R}_c[k]$:

$$\mathbf{Q}_c(\omega) = \sum_{k=0}^{1} \mathbf{R}_c[k]e^{-jk\omega}. \quad (13)$$

As a result, by defining the $(2L+1)N \times 1$ vector $\mathbf{q}_x$ and the $(2L+1)M^2 \times 1$ vector $\mathbf{q}_y$ as

$$\mathbf{q}_x = (\mathbf{F}_{2L+1} \otimes \mathbf{I}_N)\mathbf{r}_x \quad (14)$$

$$\mathbf{q}_y = (\mathbf{F}_{2L+1} \otimes \mathbf{I}_{M^2})\mathbf{r}_y \quad (15)$$

we can re-write (11) as

$$\mathbf{q}_y = \mathbf{Q}_c \mathbf{q}_x. \tag{16}$$

From (9), (10), (14), and (15), it is also possible to interpret $\mathbf{q}_x$ and $\mathbf{q}_y$ as

$$\mathbf{q}_x = [\mathbf{q}_x^T(0), \mathbf{q}_x^T(2\pi\tfrac{1}{2L+1}), \ldots, \mathbf{q}_x^T(2\pi\tfrac{2L}{2L+1})]^T$$
$$\mathbf{q}_y = [\mathbf{q}_y^T(0), \mathbf{q}_y^T(2\pi\tfrac{1}{2L+1}), \ldots, \mathbf{q}_y^T(2\pi\tfrac{2L}{2L+1})]^T$$

where $\mathbf{q}_x(\omega)$ of length N and $\mathbf{q}_y(\omega)$ of length M^2 are the vector-form spectra of the cross-correlation sequences $\mathbf{r}_x[k]$ and $\mathbf{r}_y[k]$ respectively, as follows:

$$\begin{aligned} \mathbf{q}_y(\omega) &= \sum_{k=-L}^{L} \mathbf{r}_y[k] e^{-jk\omega} \\ \mathbf{q}_x(\omega) &= \sum_{k=-L}^{L} \mathbf{r}_x[k] e^{-jk\omega}. \end{aligned} \tag{17}$$

Here, $\mathbf{q}_y$ is available from the sample-averaging versions of $\mathbf{r}_y[k]$, while $\mathbf{q}_x$, and hence $\mathbf{r}_x[k]$, are to be estimated. With (17), (16) can be written as a set of $2L+1$ matrix equations:

$$\mathbf{q}_y(2\pi\tfrac{l}{2L+1}) = \mathbf{Q}_c(2\pi\tfrac{l}{2L+1})\mathbf{q}_x(2\pi\tfrac{l}{2L+1}),\ l = 0, 1, \ldots, 2L. \tag{18}$$

By assuming that $\mathbf{Q}_c(2\pi\frac{l}{2L+1})$ has full column rank for $l = 0, 1, \ldots, 2L$, $\mathbf{q}_x$ can be simply solved using LS from (18) for $l = 0, 1, \ldots, 2L$.

Having estimated $\mathbf{q}_x$, $\mathbf{r}_x$ can be reconstructed using (14), and the $(2L+1)N \times 1$ power spectrum vector $\mathbf{p}_x$ can be computed as

$$\mathbf{p}_x = \mathbf{F}\mathbf{r}_x \tag{19}$$

where $\mathbf{F}$ is the $(2L+1)N \times (2L+1)N$ DFT matrix.

When the power spectrum $\mathbf{p}_x$ in (19) is sparse, it is possible to utilize the sparsity knowledge in our approach to further reduce the sampling rate requirements, that is, allow for $M^2 < N$. To do so, we combine (14), (16), and (19) to yield

$$\mathbf{q}_y = \mathbf{\Phi}\mathbf{p}_x \tag{20}$$

where $\mathbf{\Phi} = \mathbf{Q}_c(\mathbf{F}_{2L+1} \otimes \mathbf{I}_N)\mathbf{F}^{-1}$ is of size $(2L+1)M^2 \times (2L+1)N$. Then, the LASSO technique in [5] can be applied, which regularizes the LS with an additional ℓ_1-norm penalty term to induce a sparse solution to the power spectrum $\mathbf{p}_x$, as follows:

$$\hat{\mathbf{p}}_x = \arg\min_{\mathbf{p}_x} \|\mathbf{q}_y - \mathbf{\Phi}\mathbf{p}_x\|_2^2 + \lambda \|\mathbf{p}_x\|_1 \tag{21}$$

where the weight $\lambda \geq 0$ balances the sparsity-bias tradeoff.

4. MULTI-COSET SAMPLING IMPLEMENTATION

This section presents some specific implementations for the sampling procedure of the developed PSBS approach. For the sampling device in Section 2, a novel multi-coset implementation is developed based on the so-called minimal sparse ruler problem, and the coprime sampling approach in [9] is presented as a special case of the ruler problem.

4.1. Proposed Multi-Coset Sampling Approach

As indicated in (2), multi-coset sampling can be implemented by simply setting for every branch i, one different entry of $c_i[n]$ to one and the others to zero, i.e., $c_i[n] = 1$ if $-n = n_i$ and $c_i[n] = 0$ if $-n \neq n_i$, where $n_i \neq n_j$ whenever $i \neq j$. Concisely, $c_i[n] = \delta[-n - n_i]$, where $n_i \neq n_j$, $\forall i \neq j$. Note that this is equivalent to selecting M different rows from the identity matrix $\mathbf{I}_N$. However, this row selection cannot be random, since we need to deterministically ensure full column rank of $\{\mathbf{Q}_c(2\pi\frac{l}{2L+1})\}_{l=0}^{2L}$ in (18) or equivalently of $\mathbf{R}_c$ in (11). Since every row of $\mathbf{R}_c$ only contains a single one, it is clear that the full rank conditions can be satisfied by ensuring that $\mathbf{R}_c$ has at least a single one in each of its columns. We can find from (5) and (7) that when $\mathbf{R}_c[0]$ has a one at the column corresponding to lag $-n$, $\mathbf{R}_c[1]$ has a one at the column corresponding to lag n. This means that if the first $\lfloor \frac{N}{2} \rfloor + 1$ columns of $\mathbf{R}_c[0]$ all have at least a single one, we know that also the last $\lfloor \frac{N}{2} \rfloor$ columns of $\mathbf{R}_c[1]$ all have at least a single one, where $\lfloor x \rfloor$ denotes the largest integer not greater than x. Our task to ensure that all columns of $\mathbf{R}_c$ have at least a single one can thus be achieved by ensuring that the first $\lfloor \frac{N}{2} \rfloor + 1$ columns of $\mathbf{R}_c[0]$ all have at least a single one. Now, the problem becomes how to select an appropriate combination of rows of $\mathbf{I}_N$ to generate the coefficients of $c_i[n]$ for $i = 0, 1, \ldots, M-1$, such that $\mathbf{R}_c[0]$ has at least a single one in each of its first $\lfloor \frac{N}{2} \rfloor + 1$ columns. Further, it is desired that the number of rows selected is minimal, in order to minimize the number of branches M and thus minimize the compression rate M/N.

Because $c_i[n] = \delta[-n - n_i]$, it is clear from (5) that

$$r_{c_i,c_j}[n] = \delta[n - n_i + n_j] \tag{22}$$

which depends on the differences $n_i - n_j$. Let S denote a set of M indices selected from $\{0, 1, \ldots, N-1\}$, representing the rows of $\mathbf{I}_N$ that are to be selected by the multi-coset sampler. Let Ω denote the set of index-differences, given by

$$\Omega = \{|n_i - n_j|\, |\forall n_i, n_j \in S\}. \tag{23}$$

Then, the problem of constructing the sampler coefficients $\{c_i[n]\}_{i=0}^{M-1}$ becomes:

$$\min_S |S| \quad \text{s.t.} \quad \Omega = \left\{0, 1, \ldots, \left\lfloor \frac{N}{2} \right\rfloor \right\} \tag{24}$$

where $|S|$ denotes the cardinality of the set S. While the best solution that minimizes $|S|$ in (24) is still under investigation, one possible way to approximate the desired solution of (24) is by reformulating the problem as a so-called minimal length-$\lfloor \frac{N}{2} \rfloor$ sparse ruler problem and solving it. This is done by defining S' as a set of M indices selected from $\{0, 1, \ldots, \lfloor \frac{N}{2} \rfloor\}$ and solving:

$$\min_{S'} |S'| \quad \text{s.t.} \quad \Omega = \left\{0, 1, \ldots, \left\lfloor \frac{N}{2} \right\rfloor\right\} \tag{25}$$

A sparse ruler with length $\lfloor \frac{N}{2} \rfloor$ can be considered as a ruler having $k < \lfloor \frac{N}{2} \rfloor + 1$ distance marks $0 = n_0 < n_1 < \ldots < n_{k-1} = \lfloor \frac{N}{2} \rfloor$, but is still able to measure all integer distances from 0 up to $\lfloor \frac{N}{2} \rfloor$. Note that Ω in (25) represents the set of integer distances that can be measured by the length-$\lfloor \frac{N}{2} \rfloor$ sparse ruler with all marks $n_i \in S'$. The length-$\lfloor \frac{N}{2} \rfloor$ sparse ruler with k distance marks is called minimal if there is no length-$\lfloor \frac{N}{2} \rfloor$ sparse ruler with $k-1$ marks. The minimal sparse ruler problem has been studied in literature, see for example, [11]. Many exact and approximate solutions for this problem have been pre-calculated and tabulated. By making the connection between our multi-coset design problem and the sparse ruler problem, we are able to construct the sampler coefficients $\{c_i[n]\}_i$ using any known sparse ruler, which ensures the full rank property of $\mathbf{R}_c$ and hence the uniqueness of the simple LS solution to power spectrum reconstruction.

4.2. Coprime Sampling

The coprime sampling technique in [9] can also be fit into the PSBS framework. Specifically, it utilizes two uniform sub-Nyquist samplers of different rates $1/(AT)$ and $1/(BT)$ respectively, where $1/T$ is the Nyquist rate, and A and B are coprime numbers with $A < B$. The idea is to collect B consecutive samples from the first sampler, at the positions bAT with $0 \leq b \leq B-1$, and $2A-1$ consecutive samples from the second sampler, at positions aBT with $1 \leq a \leq 2A-1$. When these two sets of samples are combined, it is possible to obtain all the auto-correlation lags from $-ABT$ to ABT.

In our framework, this approach can be viewed as using the sampling device in Section 2 with $M = B + 2A - 1$ branches. The period of the modulation process as well as that of the integrate-and-dump device for each branch are then given by NT with $N = 2AB$. The coefficients $c_i[n]$ for every branch i of this coprime sampler can be written as $c_i[n] = \delta[-n - n_i]$, for $i = 0, 1, \ldots, M-1$, where

$$\begin{cases} n_i = iA, & \text{for } 0 \leq i \leq B-1 \\ n_i = (i - B + 1)B, & \text{for } B \leq i \leq B + 2A - 2 \end{cases}.$$

The resulting $\mathbf{R}_c$ in (12) under this coprime sampling scheme has full column rank according to the discussion from the previous section. However, we decide to follow the procedure of [9] here by cascading all M^2 vectors $\mathbf{r}_{c_i,c_j} = \left[r_{c_i,c_j}[0], \ldots, r_{c_i,c_j}[-AB], r_{c_i,c_j}[AB], \ldots, r_{c_i,c_j}[1]\right]^T$, i.e., $\mathbf{R}_{c,II} = [\ldots, \mathbf{r}_{c_i,c_j}, \ldots]^T$ for $0 \leq i, j \leq M-1$. The reconstruction problem for the coprime sampling case can then be expressed in terms of $\mathbf{r}_y[0]$ in (8) as

$$\mathbf{r}_y[0] = \mathbf{R}_{c,II} \mathbf{r}_{x,II}, \tag{26}$$

where $\mathbf{r}_{x,II} = [r_x[0], \ldots, r_x[AB], r_x[-AB], \ldots, r_x[-1]]^T$. Since the $M^2 \times (2AB+1)$ matrix $\mathbf{R}_{c,II}$ has at least a single one in all of its columns, it has full column rank, allowing us to reconstruct $\mathbf{r}_{x,II}$ using LS. Finally, the power spectrum vector $\mathbf{p}_x$ can be computed based on (19) by simply replacing $\mathbf{r}_x$ with a zero-padded version of $\mathbf{r}_{x,II}$.

N	Length $\lfloor \frac{N}{2} \rfloor$	Nr. of marks M	M/N
2	1	2	1
11	5	4	0.3636
18	9	5	0.2778
39	19	8	0.2051
78	39	11	0.1410
84	42	11	0.1310
127	63	14	0.1102

Table 1. Examples of minimal sparse rulers

5. DISCUSSION

To compare the developed sampler implementations, we can cast the coprime sampling solution of Section 4.2 as a ruler problem as well. Specifically, the ruler has $2A + B - 1$ distance marks $0 = n_0 < n_1 < \ldots < n_{2A+B-2} = 2AB - B$. However, this ruler should be considered as an incomplete ruler instead of a sparse ruler, since it can only measure all integer distances from 0 up to AB instead of from 0 up to $2AB - B$. Based on this ruler problem perspective, it is trivial to show that our proposed multi-coset sampler offers a better compression rate M/N than that offered by coprime sampling for the same value of N. For example, Table 1 shows that for $N = 84$ at least $M = 11$ marks are required to obtain a length-42 sparse ruler, leading to a compression rate of $M/N = 0.1310$. The comparable coprime sampling scenario with $N = 84$ is obtained by setting either $A = 6$ and $B = 7$, or $A = 3$ and $B = 14$, or $A = 2$ and $B = 21$. The best choice is given by $A = 6$ and $B = 7$ leading to $M/N = 0.2143$, which is higher than the optimal compression rate produced by our scheme. It is also interesting to note that coprime sampling in this current scheme can only estimate a small range of correlation lags if N is fixed, whereas the proposed multi-coset sampler can estimate any number of lags since we have the option to set L in (9) and (10) to any number.

Another drawback of the coprime sampling approach is that it cannot be implemented for arbitrary values of N without invoking Nyquist-rate sampling. As an example, consider the case where $N = 122$ leading to $AB = 61$. Since 61 is a prime number, we have to select A and B equal to 1 and 61, respectively, since A is assumed to be smaller than B. However, this means we have to sample the signal at the Nyquist rate, which is exactly what we want to avoid. On the other hand, our proposed multi-coset approach based on the minimal sparse ruler design can be used for arbitrary values of N without having to sample the signal at the Nyquist rate.

6. SIMULATIONS

This section presents some simulation results that illustrate the effectiveness of our proposed approach. First, we compare our proposed multi-coset sampling based on the minimal sparse ruler with coprime sampling, both using an LS approach. Next, we assume a sparse spectrum and compare the sparsity-agnostic approach (e.g., LS) with the sparsity-aware approach (e.g., LASSO).

6.1. Minimal Sparse Ruler versus Coprime Sampling

In this part, we consider a complex baseband representation of an OFDM signal with 16-QAM data symbols, 8192 frequency tones that span a frequency band from $-\pi$ to π, and a cyclic prefix length of 1024. We only activate 3072 frequency tones in the bands $[-\pi, -0.75\pi]$, $[0, 0.25\pi]$, and $[0.5\pi, 0.75\pi]$. The transmitted signal $x(t)$ has a power of 10 dB. We set N to $N = 84$ and vary the compression rate M/N. Two different approaches are examined, namely our proposed multi-coset sampling based on the minimal sparse ruler and the one based on coprime sampling, which are respectively discussed in Sections 4.1 and 4.2. Both estimates are computed using LS.

In the first approach, we set L in (9) equal to $L = 1$ and generate the coefficients of $c_i[n]$ for $i = 0, 1, \dots, M-1$ based on the length-42 minimum sparse ruler having $M = 11$ distance marks. This is equivalent to selecting the corresponding $M = 11$ rows from the first 43 rows of the identity matrix $\mathbf{I}_{84}$ leading to $M = 11$ branches in our sampling device. This leads to matrices $\mathbf{R}_c[0]$ and $\mathbf{R}_c[1]$ of size 121×84 in (8). The larger M/N cases are then realized by randomly adding additional rows of $\mathbf{I}_{84}$ to the already selected 11 rows.

In the second approach, the coefficients of $c_i[n]$ are generated based on coprime sampling. We select two coprime numbers, namely $A = 6$ and $B = 7$ leading to $M = 18$ branches. This leads to a matrix $\mathbf{R}_{c,II}$ of size 324×85 in (26). Again, the larger M/N cases are realized by randomly adding additional rows of $\mathbf{I}_{84}$ to the already selected 18 filters $c_i[n]$.

In Figs. 2 and 3, the normalized mean squared error (MSE) between the estimated power spectrum and the theoretical one is computed for both the sparse ruler based multi-coset sampling and coprime sampling. While no noise is considered in Fig. 2, random Gaussian noise with unit variance is introduced in Fig. 3. The normalized MSE is calculated according to:

$$MSE = \frac{E\left\{\|\hat{\mathbf{p}}_x - \mathbf{p}_x\|_2^2\right\}}{\|\mathbf{p}_x\|_2^2} \tag{27}$$

where $\mathbf{p}_x$ represents the theoretical power spectrum vector. The normalized MSE is computed for different numbers of measurement vectors (MVs) as an attempt to represent different sensing times. The normalized MSE between the estimated power spectrum produced by Nyquist-rate sampling and the theoretical one for different sensing times is also plotted as a reference. Note that the Nyquist-rate based estimated power spectrum is obtained from our proposed LS approach by setting $M = N$. It is clear from the figures that the quality of the estimation improves as M/N increases and it slowly converges towards that of the Nyquist rate. We can also notice that the normalized MSE improves as the sensing time increases, which is to be expected. It is also found that the proposed sparse ruler based multi-coset sampling generally performs better than coprime sampling. The way $\mathbf{r}_{c_i,c_j}[k]$ in (7) is arranged in $\mathbf{R}_c$ (see (11)) allows us to get a better error averaging when we reconstruct $\mathbf{r}_x$. This error appears due to the finite length effect of the received signal $x(t)$. The same trend is also observed in Fig. 4 depicting the estimated power spectrum (for both the sparse ruler based multi-coset sampling and coprime sampling) as well as the theoretical one for $M/N = 0.5$.

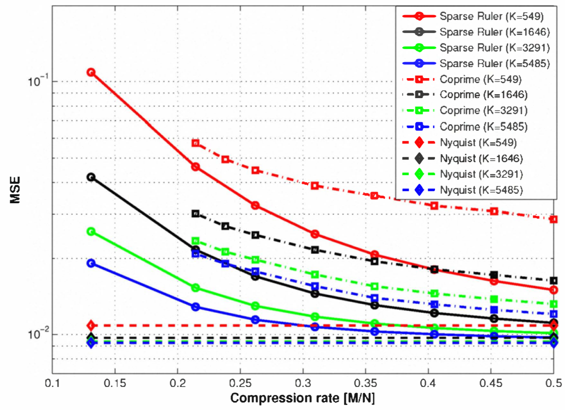


Fig. 2. The normalized MSE between the estimated power spectrum (coprime and sparse ruler based multi-coset sampling) and the theoretical one for a noiseless signal and various numbers of MVs (K).

6.2. Sparsity-Agnostic versus Sparsity-Aware

In this section, we generate a sparse discrete-time real-valued signal $x[n]$ having a frequency support $[0.2\pi, 0.26\pi]$, $[0.5\pi, 0.56\pi]$, and $[0.7\pi, 0.76\pi]$. We set N to $N = 78$ and limit L to $L = 1$. The number of branches M is varied between $M = 15$ and $M = 39$. The coefficients of $c_i[n]$ are generated as in Section 6.1. We then test a sparsity-aware power spectrum estimator that solves (21) using coordinate-descent LASSO, and compare it with the sparsity-agnostic

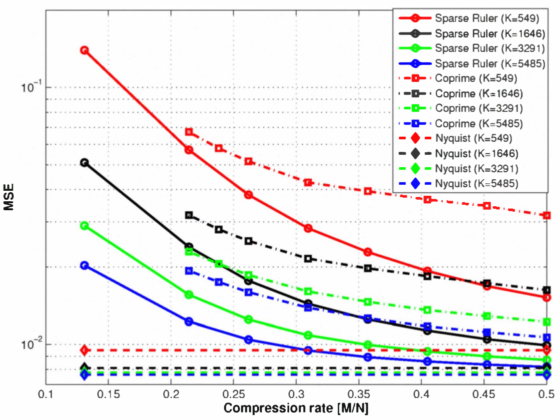


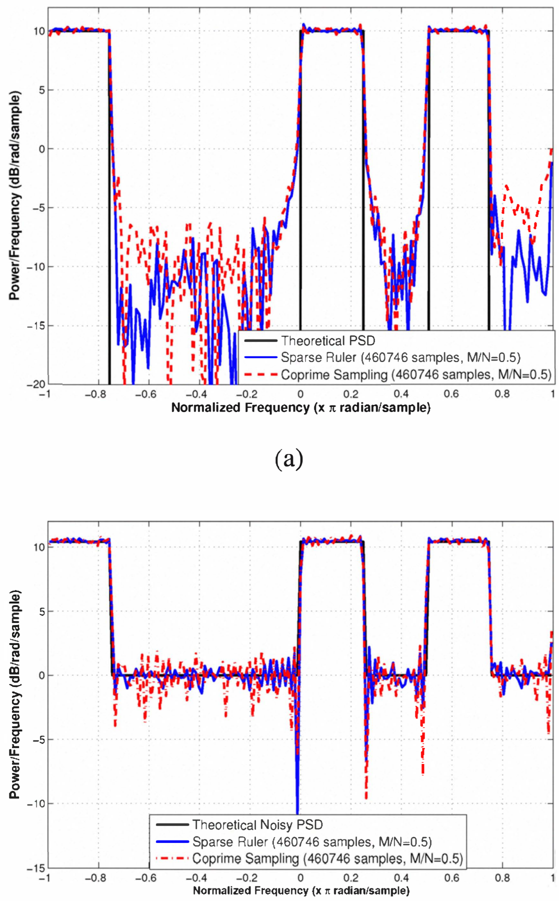
Fig. 3. The normalized MSE between the estimated power spectrum (coprime and sparse ruler based multi-coset sampling) and the theoretical one for a noisy signal and various numbers of MVs (K). The signal to noise ratio (SNR) is set to 10 dB.

LS estimator which amounts to solving (21) in closed form with $\lambda = 0$.

In Fig. 5, the normalized MSE between the estimated power spectrum and the theoretical one is computed for a noise-free scenario. The normalized MSE is calculated according to (27) for different numbers of MVs. It is clear that the coordinate-descent LASSO generally performs better than the LS approach due to the additional ℓ_1-norm penalty, which induces a sparse solution. This is true even if the LS approach is computed based on samples that are obtained from Nyquist-rate sampling ($M = N$ case). A similar trend can be observed in Fig. 6 depicting a realization of the estimated power spectrum (for both coordinate-descent LASSO and LS) and the theoretical one for $M/N = 0.5$ and different sensing times. Note that our approach using any of the two reconstruction methods is generally able to locate the presence of the active bands.

7. CONCLUSIONS

This paper has developed a multi-coset sampling version of power spectrum blind sampling (PSBS), which is used to estimate the power spectrum of wide-sense stationary signals based on samples obtained from a sub-Nyquist sampling device without any sparsity constraint. The design of the multi-coset sampling device is cast as a minimal sparse ruler problem. It is shown that any sparse ruler can produce a multi-coset sampling design that guarantees the full rank condition of the formulated sampling problem, and hence ensures the uniqueness of the power spectrum estimates as solutions to a set of simple least-squares problems. Further, when minimal sparse rulers are used, the resulting samplers approximate the



(a)

(b)

Fig. 4. Estimated power spectrum for $M/N = 0.5$ (coprime and sparse ruler based multi-coset sampling) and the theoretical power spectrum; (a) noise-free; (b) noisy (SNR=10 dB)

minimum sampling rates and hence the strong compression under the PSBS framework. The coprime sampler turns out to be an efficient ruler as well, but it is an incomplete ruler. It has been shown that the proposed minimal sparse ruler solution to multi-coset sampler design offers better compression rates and a larger range of estimated correlation lags than coprime sampling, for the same signal size N; further, it is more flexible and can be employed for arbitrary values of signal size N. These advantages are corroborated by simulations. When the power spectrum of interest is sparse, the PSBS framework can also incorporate the sparsity knowledge straightforwardly, resulting in even more compression.

8. REFERENCES

[1] Z. Tian and G.B. Giannakis, "Compressed sensing for wideband cognitive radios," *Proc. of IEEE International*

Conference on Acoustics, Speech and Signal Processing (ICASSP 2007), Vol. 4, Honolulu, HI, pp. IV/1357-IV/1360, April 2007.

[2] Y. Bresler, "Spectrum-blind sampling and compressive sensing for continuous-index signals," *Proc. of Information Theory and Applications Workshop (ITA 2008)*, pp. 547-554, January 2008.

[3] M. Mishali and Y. Eldar, "From theory to practice: sub-Nyquist sampling of sparse wideband analog signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375-391, April 2010.

[4] R. Venkataramani and Y. Bresler, "Optimal sub-Nyquist nonuniform sampling and reconstruction for multiband signals," *IEEE Transaction on Signal Processing*, vol. 49, no. 10, pp. 2301-2313, October 2001.

[5] R. Tibshirani, "Regression shrinkage and selection via thelasso," *Journal of the Royal Statistical Society*, Series B, vol.58, pp. 267288, 1996.

[6] Y. Wang, A. Pandharipande and G. Leus, "Compressive sampling based MVDR spectrum sensing," *Proc. of the International Workshop on Cognitive Information Processing (CIP 2010)*, Elba Island, Italy, pp. 333-337, June 2010.

[7] Y.L. Polo, Y. Wang, A. Pandharipande and G. Leus, "Compressive wide-band spectrum sensing," *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009)*, Taipei, pp. 2337-2340, April 2009.

[8] V. Havary-Nassab, S. Hassan and S. Valaee, "Compressive detection for wide-band spectrum sensing," *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010)*, Dallas, Texas, pp. 3094-3097, March 2010.

[9] P. Pal and P.P. Vaidyanathan, "Coprime sampling and the MUSIC algorithm," *Proc. of IEEE Digital Signal Processing and Signal Processing Education Workshop*, Sedona, Arizona, pp. 289-294, January 2011.

[10] G. Leus and D.D. Ariananda, "Power Spectrum Blind Sampling," submitted to *IEEE Signal Processing Letters.*

[11] J. Leech, "On the representation of 1,2,...,n by differences," *Journal of the London Mathematical Society*, Vol. 31, pp. 160-169, April 1956.

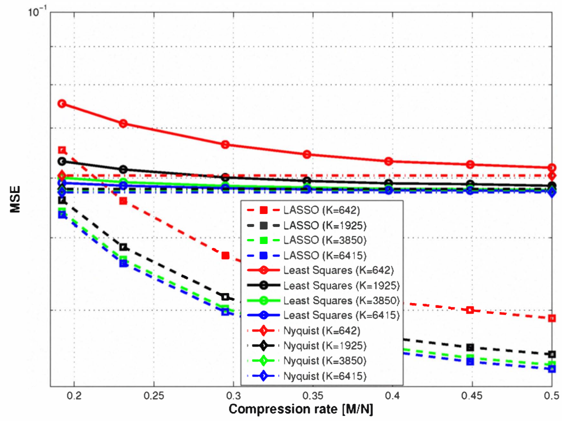


Fig. 5. The normalized MSE between the estimated power spectrum (coordinate-descent LASSO and LS) and the theoretical one for a noiseless sparse real signal and various numbers of MVs (K).

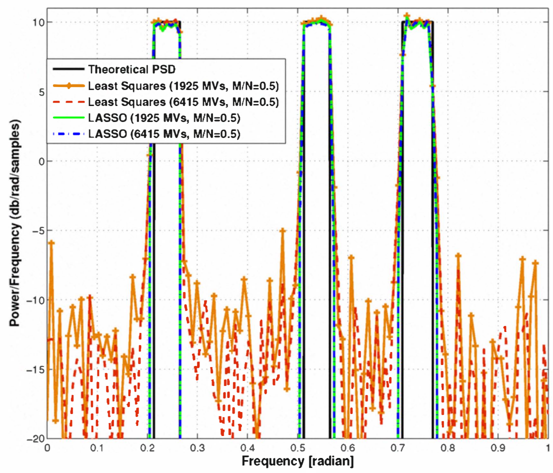


Fig. 6. Estimated power spectrum for $M/N = 0.5$ (coordinate-descent LASSO and LS) and the theoretical power spectrum for a noiseless sparse real signal and various numbers of MVs (K)