Compressive Sampling for Power Spectrum Estimation*

Dyonisius Dony Ariananda, Geert Leus
Faculty of Electrical Engineering, Mathematics, and Computer Science
Delft University of Technology
Mekelweg 4, 2628 CD, Delft, The Netherlands
{d.a.dyonisius,g.j.t.leus}@tudelft.nl

Abstract

Compressive sampling is a well-known approach to reconstruct sparse signals based on a limited number of measurements. In spectrum sensing applications for cognitive radio though, only reconstruction of the power spectrum of the signal is required, instead of the signal itself. In this paper, we present a new method for power spectrum reconstruction based on samples produced by a sub-Nyquist rate sampling device. The stationary assumption on the received analog signal causes the measurements at the output of the compressive sampling block to be cyclo-stationary, or the measurement vectors to be stationary. We investigate the relationship between the autocorrelation matrix of the measurement vectors and that of the received analog signal, which we represent by its Nyquist rate sampled version. Based on this relationship, we are able to express the autocorrelation sequence of the received wide sense stationary signal as a linear function of the vectorized autocorrelation matrix of the measurement vectors. Depending on the compression rate, we can present the problem as either over-determined or under-determined. Our focus will be mainly on the over-determined case, in which the reconstruction does not require any additional constraints. Two types of sampling matrices are examined, namely complex Gaussian and multi-coset sampling matrices. For both of them, we can derive conditions under which the over-determined system will result in a unique solution for the power spectrum by adopting a simple least squares (LS) algorithm. In the case of multi-coset sampling, further improvement on the quality of the power spectrum estimates can be attained by optimizing the condition of the sampling matrix.

1 Introduction

In cognitive radio networks, the so-called secondary users are allowed to opportunistically rent licensed frequency bands when the licensed users are inactive. In such networks, wideband spectrum sensing is a crucial task for the secondary users that are required to sense the wireless environment over a wide frequency band and to detect the spectrum occupancy. Based on the result of spectrum sensing, the secondary users can subsequently exploit the available frequency holes to set up a communication link. The fact that a broad spectral range has to be sensed in cognitive radio environments has put an additional burden on the existing analog-to-digital converters (ADCs) since the ADCs will need to sample a very large bandwidth signal at the Nyquist rate, which yields a high power consumption [1]. Several works on sub-Nyquist sampling have been conducted in order to alleviate the requirements of the ADCs. Sub-Nyquist sampling based on multi-coset sampling has been evaluated in [2] for the case of multiband

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

signals. Furthermore, [2] has also derived the condition that allows perfect signal reconstruction from the obtained sub-Nyquist samples. In addition, sub-Nyquist sampling for sparse multiband analog signals using a so-called modulated wideband converter (MWC) is introduced in [3]. The MWC basically consists of multiple channels, each of which applies a different mixing function followed by low-pass filtering and low-rate uniform sampling. Similar to [2], [3] also provides the conditions for perfect reconstruction of the original signal based on the output of the MWC.

Note that all the above works focus on perfect reconstruction of the original signal while spectrum sensing applications only require perfect reconstruction of the power spectrum or equivalently the autocorrelation sequence. In [4, 5], power spectrum reconstruction methods based on compressive sampling have been proposed by focusing on the autocorrelation sequence instead of the signal itself. In general, both approaches exploit the sparsity feature of the edge spectrum. The method in [4], however, compresses the autocorrelation function of the Nyquist rate samples, and thus actually still requires Nyquist rate sampling. Hence, [5] attempts to directly perform compressive sampling on the received signal by exploiting the relationship between the autocorrelation sequence of the measurements and the one of the Nyquist rate samples. However, [5] assumes stationarity of the measurements, which is not correct for most compressive sampling matrices.

In this paper, we introduce a new sub-Nyquist sampling based approach for power spectrum reconstruction. Our approach exploits the stationarity of the measurement vectors (or cyclo-stationarity of the measurements) and may not even require the sparsity assumption, which is usually needed for perfect reconstruction of the original signal. We first discuss how to exploit the relationship between the autocorrelation sequence of the Nyquist rate samples and the autocorrelation matrix of the measurement vectors. We subsequently attempt to reformulate the problem by considering the elements of the measurement vector as the outputs of different filters having coefficients given by the rows of the sampling matrix. We then focus on the reconstruction of the power spectrum for over-determined systems by considering two choices for the compressive sampling matrix, namely a complex Gaussian matrix and a multi-coset matrix.

2 Problem Formulation

Consider the received wide sense stationary analog signal $x(t)$ that is sampled using an analog to information converter (AIC) resulting in a sequence of measurement vectors $\mathbf{y}[k]$. As discussed in [4, 5], the AIC can be interpreted (but not implemented) as a block consisting of a basic Nyquist rate ADC followed by a multiplexing operation collecting N consecutive Nyquist rate samples, and concluded by a multiplication with a compressive sampling matrix, which compresses the number of samples from N to M. Based on the above point of view, we denote the output of the Nyquist rate ADC by $x[n]$ and the output of the multiplexer by the $N \times 1$ vector sequence $\mathbf{x}[k]$ given by:

$$\mathbf{x}[k] = \begin{bmatrix} x[kN] & x[kN+1] & .... & x[kN+N-1] \end{bmatrix}^T. \tag{1}$$

Then, each $N \times 1$ vector $\mathbf{x}[k]$ is compressed by the $M \times N$ compressive sampling matrix $\mathbf{\Phi}$ leading to the $M \times 1$ vector sequence $\mathbf{y}[k]$:

$$\mathbf{y}[k] = \mathbf{\Phi}\mathbf{x}[k]. \tag{2}$$

We denote the autocorrelation sequence of the Nyquist rate samples $x[n]$ by $r_x[l] = E\left[x[n]x^*[n-l]\right]$, where $(.)^*$ represents the complex conjugate operation. The $N \times N$ autocorrelation matrix of $\mathbf{x}[k]$ in (1) can subsequently be constructed as $\mathbf{R}_x = E\left[\mathbf{x}[k]\mathbf{x}^H[k]\right]$, whose elements are given by:

$$[\mathbf{R}_x]_{ij} = r_x[i-j] = r_x^*[j-i], \tag{3}$$

while the $M \times M$ autocorrelation matrix of $\mathbf{y}[k]$ in (2) can be written as:

$$\mathbf{R}_y = E[\mathbf{y}[k]\mathbf{y}^H[k]] = \mathbf{\Phi}\mathbf{R}_x\mathbf{\Phi}^H. \tag{4}$$

Note that the elements of the measurement vectors $\mathbf{y}[k]$ are generally not wide sense stationary because of the nature of the compressive sampling matrix $\mathbf{\Phi}$. Consequently, it is generally not possible to express the elements of $\mathbf{R}_y$ in a similar form as (3). This fact, however, allows us to exploit all columns of $\mathbf{R}_y$ for estimating one of the columns of $\mathbf{R}_x$ since, unlike the columns of $\mathbf{R}_x$, every column of $\mathbf{R}_y$ contains different information. We first stack all columns of $\mathbf{R}_y$ into the $M^2 \times 1$ vector $\text{vec}(\mathbf{R}_y)$, where vec(.) is the operator stacking all columns of a matrix in a large column vector. It is then obvious from (4) that we can write $\text{vec}(\mathbf{R}_y)$ as:

$$\text{vec}(\mathbf{R}_y) = (\mathbf{\Phi}^* \otimes \mathbf{\Phi})\text{vec}(\mathbf{R}_x), \tag{5}$$

where $\otimes$ represents the Kronecker product operation. As mentioned earlier, all columns of $\mathbf{R}_x$ contain the same information, which can be condensed into the $2N-1$ vector $\mathbf{r}_x = [r_x(1-N), \cdots, r_x(-1), r_x(0), r_x(1), \cdots, r_x(N-1)]^T$. We can then express the relationship between $\text{vec}(\mathbf{R}_x)$ and $\mathbf{r}_x$ as:

$$\text{vec}(\mathbf{R}_x) = \mathbf{T}\mathbf{r}_x, \tag{6}$$

where $\mathbf{T}$ is a special $N^2 \times (2N-1)$ repetition matrix. By combining (5) and (6), we obtain:

$$\text{vec}(\mathbf{R}_y) = (\mathbf{\Phi}^* \otimes \mathbf{\Phi})\mathbf{T}\mathbf{r}_x. \tag{7}$$

We introduce the $M^2 \times (2N-1)$ matrix $\mathbf{\Theta} = (\mathbf{\Phi}^* \otimes \mathbf{\Phi})\mathbf{T}$ to simplify the analysis, and rewrite (7) as:

$$\text{vec}(\mathbf{R}_y) = \mathbf{\Theta}\mathbf{r}_x. \tag{8}$$

Given (8), we aim to reconstruct the autocorrelation sequence $\mathbf{r}_x$ from $\text{vec}(\mathbf{R}_y)$ and to use it for calculating the $(2N-1) \times 1$ power spectrum vector $\mathbf{p}_x$ according to:

$$\mathbf{p}_x = \mathbf{F}\mathbf{r}_x \tag{9}$$

where $\mathbf{F}$ is a $(2N-1) \times (2N-1)$ column-permuted version of the discrete Fourier transform (DFT) matrix.

3 Problem Reformulation

Each row of $\mathbf{\Phi}$ in (7) can also be considered as a unique discrete waveform or filter and we can thus express $\mathbf{\Phi}$ in terms of its row vectors:

$$\mathbf{\Phi} = \begin{bmatrix} \boldsymbol{\varphi}_0 & \boldsymbol{\varphi}_1 & \boldsymbol{\varphi}_2 & \dots & \boldsymbol{\varphi}_{M-1} \end{bmatrix}^T, \tag{10}$$

with $\boldsymbol{\varphi}_i = [\varphi_i[0], \varphi_i[-1], \dots, \varphi_i[1-N]]^T$. If the vector sequence $\mathbf{y}[k]$ in (2) is defined as a collection of M parallel scalar sequences $y_i[k]$, i.e., $\mathbf{y}[k] = [y_0[k], y_1[k], \dots, y_{M-1}[k]]^T$, we can consider $y_i[k]$ as the N-fold decimated version of the sequence obtained by filtering $x[n]$ with $\varphi_i[n]$, $y_i[k] = \sum_{n=1-N}^{0} \varphi_i[n]x[kN-n]$. In this case, we can express the auto- and cross-correlations between the elements of the measurement vector $\mathbf{y}[k]$ in $\text{vec}(\mathbf{R}_y)$ in terms of the auto- and cross-correlation sequences between the different sequences $y_i[k]$ evaluated at lag zero. By denoting the correlation sequence between

$y_i[k]$ and $y_j[k]$ at lag l by $r_{y_i,y_j}[l]$ and taking the deterministic nature of the filter coefficients into account, $r_{y_i,y_j}[l]$ can be written as:

$$r_{y_i,y_j}[l] = E\left\{y_i[k]y_j^*[k-l]\right\} = \sum_{n=1-N}^{0} \varphi_i[n] \sum_{p=1-N}^{0} \varphi_j^*[p] r_x[lN+p-n]. \quad (11)$$

By using variable substitution, (11) can be rewritten as:

$$r_{y_i,y_j}[l] = \sum_{s=1-N}^{N-1} r_{\varphi_i,\varphi_j}[s] r_x[lN-s], \quad (12)$$

where $r_{\varphi_i,\varphi_j}[l] = \sum_{n=1-N}^{0} \varphi_i[n]\varphi_j^*[n-l]$ is the deterministic correlation sequence between $\varphi_i[n]$ and $\varphi_j[n]$. By taking (12) into account, we can express $\text{vec}(\mathbf{R}_y)$ in (8) as:

$$\text{vec}(\mathbf{R}_y) = \left[\mathbf{r}_{y_0}^T[0], \mathbf{r}_{y_1}^T[0], \ldots, \mathbf{r}_{y_{M-1}}^T[0]\right]^T, \quad (13)$$

where $\mathbf{r}_{y_i}[0] = \left[r_{y_0,y_i}[0], r_{y_1,y_i}[0], \ldots, r_{y_{M-1},y_i}[0]\right]^T$. From some elementary mathematical computations on (7) and (8) as well as from (12), we can observe that $\mathbf{\Theta}$ in (8) is composed of the auto- and cross-correlations between the rows of $\mathbf{\Phi}$. By specifying $\mathbf{r}_{\varphi_i,\varphi_j}$ as $\mathbf{r}_{\varphi_i,\varphi_j} = \left[r_{\varphi_i,\varphi_j}[N-1], \cdots, r_{\varphi_i,\varphi_j}[0], \cdots, r_{\varphi_i,\varphi_j}[1-N]\right]^T$, $\mathbf{\Theta}$ can be written as:

$$\mathbf{\Theta} = \left[\mathbf{r}_{\varphi_0,\varphi_0}, \ldots, \mathbf{r}_{\varphi_{M-1},\varphi_0}, \ldots, \mathbf{r}_{\varphi_0,\varphi_{M-1}}, \ldots, \mathbf{r}_{\varphi_{M-1},\varphi_{M-1}}\right]^T. \quad (14)$$

From (13) and (14), we can find that $\text{vec}(\mathbf{R}_y)$ is obtained by simply multiplying the autocorrelation sequence $\mathbf{r}_x$ with $\mathbf{\Theta}$, whose elements are given by the deterministic auto- and cross-correlations between the rows of $\mathbf{\Phi}$.

4 Reconstruction

We try to recover the power spectrum by first reconstructing the autocorrelation sequence $\mathbf{r}_x$ from (8) for $\mathbf{\Theta}$ given by (14). The reconstruction problem can be classified into two different cases, the under- and over-determined cases (the determined case is viewed as part of the over-determined case). It is important to note how our method may boil down to an over-determined system (even with significant compression, i.e., $M \ll N$) while common compressive sampling problems generally result in an under-determined system. This is due to the fact that we concentrate on reconstructing statistical parameters (namely auto- and cross-correlation sequences), which enables us to gain much more system equations. While additional constraints (such as sparsity assumptions, as discussed in [6]) are obviously required in the under-determined case, this is not the case for the over-determined case which will be the focus of this paper. More details on the under-determined case can be found in [4, 5, 6].

In general, the system will be over-determined when $M^2 \geq 2N-1$. If $\mathbf{\Theta}$ has full column rank, we can calculate the autocorrelation vector $\mathbf{r}_x$ from (8) as the LS solution of (8):

$$\hat{\mathbf{r}}_x = (\mathbf{\Theta}^H\mathbf{\Theta})^{-1}\mathbf{\Theta}^H\text{vec}(\mathbf{R}_y). \quad (15)$$

The power spectrum estimate $\hat{\mathbf{p}}_x$ can subsequently be calculated from (9). It is interesting to note that we can reconstruct the statistics of the signal without requiring any sparsity assumptions. This is in contrast with the general compressive sampling framework where signal sparsity is required to ensure perfect reconstruction. In the following two sub-sections, we discuss the design of the sampling matrix in order to guarantee the uniqueness of the LS solution ($\mathbf{\Theta}$ having full column rank). We consider the complex Gaussian matrix and multi-coset matrix cases.

4.1 Complex Gaussian Sampling Matrix

When every element of $\mathbf{\Phi}$ is randomly generated following a Gaussian distribution, the probability that $\mathbf{\Theta}$ will have full column rank is very high once $M^2 \geq (2N-1)$ and this can happen for $M \ll N$. Hence, we propose to use a complex Gaussian matrix as a possible realization of the sampling matrix $\mathbf{\Phi}$ to ensure the full column rank property of $\mathbf{\Theta}$.

4.2 Multi-Coset Sampling Matrix

Multi-coset sampling can also be tailored to fit into our framework. From (10), a multi-coset sampling matrix can be constructed by selecting M different rows from the identity matrix $\mathbf{I}_N$ leading to an $M \times N$ multi-coset sampling matrix $\mathbf{\Phi}$. Note that we cannot choose the rows of $\mathbf{I}_N$ in a random way for a given M since some conditions have to be fulfilled to guarantee the full column rank property of $\mathbf{\Theta}$ in (14). When we employ a multi-coset sampling matrix, each row of $\mathbf{\Theta}$ will only contain a single one and will have zeros elsewhere. Therefore, to obtain a full column rank $\mathbf{\Theta}$, we must select a proper combination of rows of $\mathbf{I}_N$ that leads to $\mathbf{\Theta}$ having at least a single one in every column. Moreover, our aim is to keep the number of selected rows minimal in order to minimize the compression rate M/N. By assuming that $\varphi_j[n] = \delta[-n-n_j]$ for $j = 0, 1, 2, \dots, N-1$, it is obvious from (12) that the correlation $r_{\varphi_i,\varphi_j}[l]$ is given by $r_{\varphi_i,\varphi_j}[l] = \delta[l - n_i + n_j]$. The next step is to form $\mathbf{\Phi}$ by selecting M out of N rows of $\mathbf{I}_N$ subject to the constraints on $\mathbf{\Theta}$ stated earlier. By introducing S as a set of M indices chosen from $\{0, 1, \dots, N-1\}$, which represents the rows from $\mathbf{I}_N$ that are going to be selected and Ω as a set given by:

$$\Omega = \{|n_i - n_j| : \forall n_i, n_j \in S\}, \tag{16}$$

we can express the multi-coset sampling matrix construction problem as:

$$\min_S |S| \text{ s.t. } \Omega = \{0, 1, \dots, N-1\}, \tag{17}$$

where $|S|$ represents the cardinality of the set S. In fact, this problem is a so-called minimal $(N-1)$-length sparse ruler problem. An $(N-1)$-length sparse ruler can be considered as a ruler having $k < N$ distance marks $0 = n_0 < n_1 < \ldots < n_{k-1} = N-1$ and that is still able to measure all integer distances from 0 up to $N-1$. Observe that Ω in (17) is the set of integer distances that can be measured by the $(N-1)$-length sparse ruler with all marks $n_i \in S$. The $(N-1)$-length sparse ruler with k distance marks is called minimal if there is no $(N-1)$-length sparse ruler with $k-1$ marks. Solving this minimal sparse ruler problem basically means minimizing the compression rate M/N, while maintaining uniqueness of the solution of the LS reconstruction problem. The minimal sparse ruler problem has for instance been studied in [7].

Note that in practice, the received signal $x(t)$ will be of finite length. Consequently, $\mathbf{R}_x$ in (3) will not be a perfect Toeplitz matrix and this contributes to additional errors in the LS estimate $\hat{\mathbf{r}}_x$ in (15). In order to get a better error averaging, it is desirable to have a sufficient number of ones in each column of $\mathbf{\Theta}$. Given M, the best way to get the optimum error averaging is by ensuring that the number of ones in each column of $\mathbf{\Theta}$ is as equal as possible. With respect to this issue, we propose two possible approaches for the construction of a multi-coset sampling matrix. In scheme I, we start by generating a multi-coset sampling matrix based on a minimal sparse ruler leading to minimum M/N. The larger M/N cases are then realized by selecting additional rows of $\mathbf{I}_N$ that minimize $\text{Var}[\sum_{n_i \in S} \sum_{n_j \in S} r_{\varphi_i,\varphi_j}[l]]$, namely the variance of the multiplicity of ones in the columns of $\mathbf{\Theta}$. In scheme II, we do not start from a minimal sparse ruler. Instead, we directly minimize $\text{Var}[\sum_{n_i \in S} \sum_{n_j \in S} r_{\varphi_i,\varphi_j}[l]]$ while minimizing M/N. Note that our second approach might not be able to reach the smallest possible value of M/N as provided by a minimal sparse ruler design. The procedures for scheme I and II are described below.

1. Denote $S^{(k)}$ as the set of indices selected from $\{0, 1, \ldots, N-1\}$ after the k-th iteration and $\Omega^{(k)} = \left\{|n_i - n_j| : \forall n_i, n_j \in S^{(k)}\right\}$. Initialize $S^{(0)} = \{0, N-1\}$ for scheme II. For scheme I, the initial value for $S^{(0)}$ is obtained by solving the minimal $(N-1)$-length sparse ruler problem.

2. Denote $\Xi^{(k)} = \{0, 1, \ldots, N-1\} \setminus S^{(k-1)}$. For each $n_{q'} \in \Xi^{(k)}$ define the set $C_{q'} = \left(\bigcup_{n_i \in S^{(k-1)}} \{|n_{q'} - n_i|\}\right) \cup \Omega^{(k-1)}$.

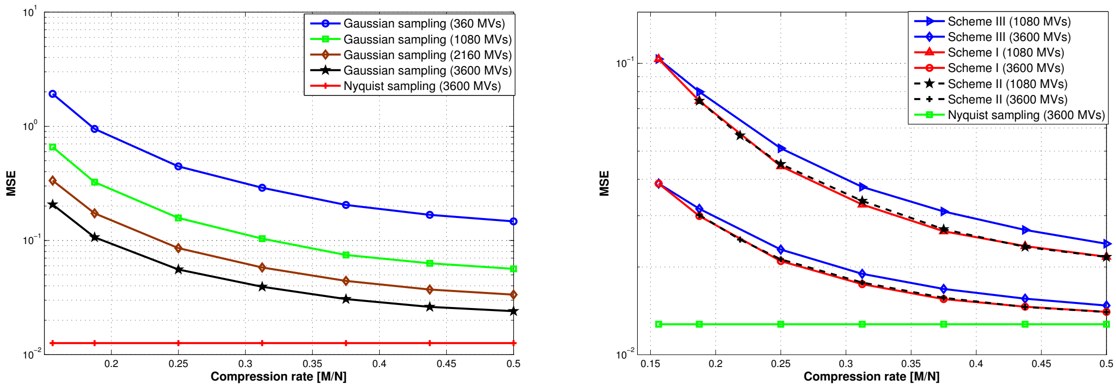
3. Define the set $\Gamma^{(k)} = \{n_{q''} | n_{q''} = \arg \max_{n_{q'} \in \Xi^{(k)}} |C_{q'}|\}$, where $|C_{q'}|$ denotes the cardinality of set $C_{q'}$.

4. Search among the candidate set $\Gamma^{(k)}$ for the item n_q that satisfies:

$$n_q = \arg \min_{n_{q''} \in \Gamma^{(k)}} \mathrm{Var}\left[\sum_{n_i \in S^{(k-1)} \cup \{n_{q''}\}} \sum_{n_j \in S^{(k-1)} \cup \{n_{q''}\}} r_{\varphi_i, \varphi_j}[l]\right]$$

5. We set $S^{(k)} = S^{(k-1)} \cup \{n_q\}$ and $\Omega^{(k)} = \left(\bigcup_{n_i \in S^{(k-1)}} \{|n_q - n_i|\}\right) \cup \Omega^{(k-1)}$. As long as $|S^{(k)}|$ is less than the desired M (or $|\Omega^{(k)}| < N$ for scheme II), go to step 2.

6. Construct $\mathbf{\Phi}$ by selecting M out of N rows of $\mathbf{I}_N$ where the M indices are given by $S^{(k)}$ with $k = M - 2$ for scheme II and $k = M - |S^{(0)}|$ for scheme I.

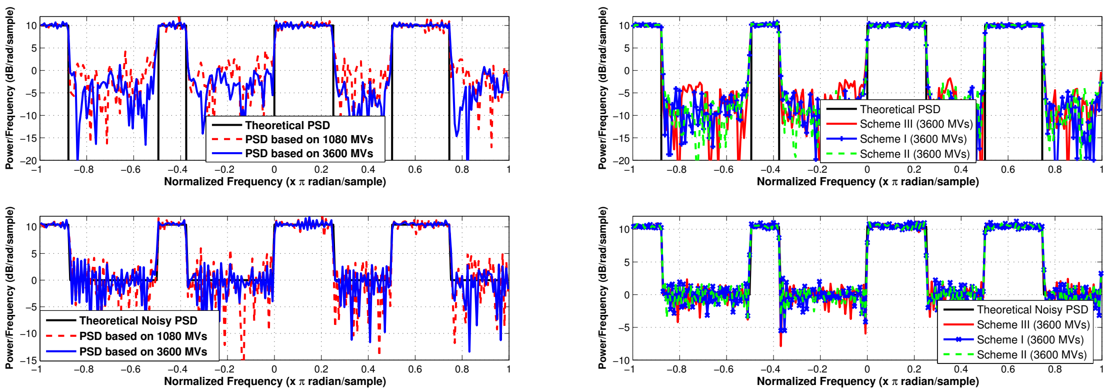


(a) Gaussian sampling case (b) Multi-coset sampling case

Figure 1: The MSE between the estimated power spectrum and the theoretical one for a noiseless signal.

5 Simulation Study

In this section, the numerical results from a simulation study for both multi-coset and complex Gaussian sampling are presented. Consider a complex baseband representation of an OFDM signal having 8192 sub-carriers spanning a frequency band from $-\pi$ to π, 16 QAM data symbols and a cyclic prefix length of 1024. We activate 3072 sub-carriers in the bands $[-\pi, -0.875\pi]$, $[-0.5\pi, -0.375\pi]$, $[0, 0.25\pi]$ and $[0.5\pi, 0.75\pi]$. The power of the transmitted signal $x(t)$ is set to 10dB. The value of N is fixed to $N = 128$ and the compression rate M/N is varied while ensuring the full column rank property of $\mathbf{\Theta}$. We randomly generate the entries of the complex Gaussian sampling matrix with zero mean and variance $1/M$ and we keep it fixed over the different simulation runs. Fig. 1a depicts the mean squared error (MSE) between the estimated power spectrum and the theoretical one when a complex Gaussian sampling matrix is used. No noise is assumed in this figure. The MSE is computed based on $\text{MSE} = E\left\{(\|\hat{\mathbf{p}}_x - \mathbf{p}_x\|_2^2)/\|\mathbf{p}_x\|_2^2\right\}$ where $\mathbf{p}_x$ denotes the theoretical power spectrum vector. In order to simulate different sensing times, we calculate the MSE for different numbers of collected measurement vectors (MVs) $\mathbf{y}[k]$ in (2). The MSE between the estimated power spectrum produced by Nyquist rate sampling and the theoretical one is also plotted as a line for the sake of reference. From the figure, it is obvious that the quality of the estimation produced by complex Gaussian sampling improves with M/N, although its performance converges very slowly towards that of Nyquist rate sampling. We can also observe that the MSE improves as the sensing time increases due to the fact that our estimated autocorrelation value $r_{y_i,y_j}[0]$ in (13) approaches the actual value. Fig. 2a describes the estimated power spectrum together with the theoretical one for $M/N = 0.5$ and different sensing times. A complex Gaussian sampling matrix is also considered here. Observe that for longer sensing times, the presence of the active bands can be better located.



(a) Gaussian sampling case (b) Multi-coset sampling case

Figure 2: Estimated power spectrum for $M/N = 0.5$ together with the theoretical one; top: noise-free; bottom: noisy.

We investigate three different schemes for the construction of the multi-coset sampling matrix. For scheme I, it turns out that the minimum number of distance marks for a 127-length sparse ruler is 20 and thus we choose the corresponding 20 rows from the 128 rows of the identity matrix $\mathbf{I}_{128}$ to construct a 20×128 matrix $\mathbf{\Phi}$. The larger M/N cases are then realized by selecting additional rows of $\mathbf{I}_N$ that minimize the variance of the multiplicity of ones in the columns of $\mathbf{\Theta}$ (as mentioned in Section 4.2). In scheme II, the sampling matrix for every M/N case is constructed according to the procedures explained in Section 4.2. However, the minimum M/N achievable by this scheme is higher than that by scheme I, namely 24/128 instead of 20/128. Similar to scheme I, scheme III also starts from the sampling matrix generated based on the minimal 127-length sparse ruler. However, instead of focusing on the multiplicity of ones in the columns of $\mathbf{\Theta}$, the larger M/N cases are realized by randomly adding additional rows of $\mathbf{I}_{128}$ into the already selected 20 rows. From Fig. 1b, it is clear that schemes I and II have outperformed scheme III, meaning that ensuring the number of ones in each column of $\mathbf{\Theta}$ to be as equal as possible leads to better error averaging. A similar trend is also found in Fig. 2b. Note that from Figs. 1 and 2, multi-coset sampling appears to offer a better performance than complex Gaussian sampling.

6 Conclusion

In this work, a new method for power spectrum estimation based on samples produced by a sub-Nyquist rate sampling device has been proposed. In general, the cyclostationarity of the measurements is exploited to obtain more linear equations for the reconstruction problem. We concentrate on the over-determined case and investigate the full column rank property of the reconstruction matrix $\mathbf{\Theta}$ such that a simple LS algorithm can be employed to recover the power spectrum of stationary signals even without any sparsity assumption. We investigate two realizations of the sampling matrix, namely a multi-coset and a complex Gaussian sampling matrix. For the case of multi-coset sampling, we introduce new methods to optimize the performance by minimizing the variance of the multiplicity of ones in each column of $\mathbf{\Theta}$. The simulation study for both sampling matrix realizations has indicated the satisfactory performance of our method, which is able to correctly locate the occupied bands, thus making it a suitable candidate for power spectrum sensing in a cognitive radio network.

References

[1] B. Le, T. Rondeau, J. Reed, and C. Bostian "Analog-to-digital converters," *IEEE Signal Processsing Magazine*, vol. 22, no. 6, Nov. 2005.

[2] 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, Oct. 2001.

[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] Z. Tian and G.B. Giannakis, "Compressed sensing for wideband cognitive radios," *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 4, pp. IV/1357-IV/1360, April 2007.

[5] 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*, pp. 2337-2340, April 2009.

[6] E.J. Candès, J. Romberg and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transaction on Information Theroy*, Vol. 52, No. 2, pp. 489-509, Feb. 2006.

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