

Sub-Nyquist Cyclostationary Detection for Cognitive Radio

Deborah Cohen, *Student IEEE* and Yonina C. Eldar, *Fellow IEEE*

Abstract—Cognitive Radio requires efficient and reliable spectrum sensing of wideband signals. In order to cope with the sampling rate bottleneck, new sampling methods have been proposed that sample below the Nyquist rate. However, such techniques decrease the signal to noise ratio (SNR), deteriorating the performance of subsequent energy detection. Cyclostationary detection, which exploits the periodic property of communication signal statistics, absent in stationary noise, is a natural candidate for this setting. In this work, we consider cyclic spectrum recovery from sub-Nyquist samples, in order to achieve both efficiency and robustness to noise. To that end, we propose a structured compressed sensing algorithm, that extends orthogonal matching pursuit to account for the structure imposed by cyclostationarity. Next, we derive a lower bound on the sampling rate required for perfect cyclic spectrum recovery in the presence of stationary noise and show that it can be reconstructed from samples obtained below Nyquist, without any sparsity constraints on the signal. In particular, in non sparse settings, the cyclic spectrum can be recovered at $4/5$ of the Nyquist rate. If the signal of interest is sparse, then the sampling rate may be further reduced to $8/5$ of the Landau rate. Once the cyclic spectrum is recovered, we estimate the number of transmissions that compose the input signal, along with their carrier frequencies and bandwidths. Simulations show that cyclostationary detection outperforms energy detection in low SNRs in the sub-Nyquist regime. This was already known in the Nyquist regime, but is even more pronounced at sub-Nyquist sampling rates.

I. INTRODUCTION

Spectrum sensing has been thoroughly investigated in the signal processing literature. Several sensing schemes have been proposed, with different performance and complexity levels. The simplest approach is energy detection [1], which does not require any a priori knowledge on the input signal. Unfortunately, energy detection is very sensitive to noise and performs poorly in low signal to noise ratio (SNR) regimes. In contrast, matched filter (MF) detection [2], [3], which correlates a known waveform with the input signal to detect the presence of a transmission, is the optimal linear filter for maximizing SNR in the presence of additive stochastic noise. This technique requires perfect knowledge of the potential received transmission. A compromise between both methods is cyclostationary detection [4]–[6]. This approach is more robust to noise than energy detection but at the same time only assumes the signal of interest exhibits cyclostationarity.

Cyclostationary processes have statistical characteristics that vary periodically, arising from the underlying data modulation mechanisms, such as carrier modulation, periodic keying

or pulse modulation. The cyclic spectrum, a characteristic function of such processes, exhibits spectral peaks at certain frequency locations called cyclic frequencies, which are determined by the signal parameters, particularly the carrier frequency and symbol rate [6]. When determining the presence or absence of a signal, cyclostationary detectors exploit one fundamental property of the cyclic spectrum: stationary noise and interference exhibit no spectral correlation. Non-stationary interference can be distinguished from the signal of interest provided that at least one cyclic frequency of the signal is not shared with the interference [6]. This renders such detectors highly robust to noise and interference.

The traditional task of spectrum sensing has recently been facing new challenges due, to a large extent, to cognitive radio (CR) applications [7]. Today, CRs are perceived as a potential solution to the spectrum over-crowdedness issues, bridging between the scarcity of spectral resources and their sparse nature [8]. Even though most of the spectrum is already owned and new users can hardly find free frequency bands, various studies [9]–[11] have shown that it is typically significantly underutilized. CRs allow secondary users to opportunistically use the licensed spectrum when the corresponding primary user (PU) is not active [7]. CR requirements dictate new challenges for its most crucial task, spectrum sensing. On the one hand, detection has to be performed in real time, efficiently and with minimal software and hardware resources. On the other hand, it has to be reliable and able to cope with low SNR regimes.

Nyquist rates of wideband signals, such as those CRs deal with, are high and can even exceed today's best analog-to-digital converters (ADCs) front-end bandwidths. In addition, such high sampling rates generate a large number of samples to process, affecting speed and power consumption. In order to efficiently sample sparse wideband signals, several new sampling methods have recently been proposed [12]–[15] that reduce the sampling rate in multiband settings below the Nyquist rate. This alleviates the burden on both the analog and digital sides by enabling the use of cheaper and lower power reduced rate ADCs and the processing of fewer samples.

The authors of [12] derive the minimal sampling rate allowing for perfect signal reconstruction in noise-free settings and provide specific sampling and recovery techniques. However, when the final goal is spectrum sensing and detection, reconstructing the original signal is unnecessary. Power spectrum reconstruction from sub-Nyquist samples is considered in [16]–[18]. These works seek power spectrum estimates from low rate samples, using multicoset sampling [16]–[18] and the modulated wideband converter (MWC) [18] proposed in [13]. The presence or absence of a signal in a particular

This project is funded by the European Union's Horizon 2020 research and innovation program under grant agreement No. 646804-ERC-COG-BNYQ, and by the Israel Science Foundation under Grant no. 335/14. Deborah Cohen is grateful to the Azrieli Foundation for the award of an Azrieli Fellowship.

frequency band is then assessed with respect to the estimated power within the band. Unfortunately, the sensitivity of energy detection used in the above works is amplified when performed on sub-Nyquist samples due to noise aliasing [19]. Therefore, this scheme fails to meet CR performance requirements in low SNR regimes. On the other hand, little a priori knowledge can be assumed on the received signals, making MF difficult to implement. Consequently, cyclostationary detection is a natural candidate for spectrum sensing from sub-Nyquist samples in low SNRs.

Signal detection using cyclostationarity and its application to spectrum sensing for CR systems in the Nyquist regime, has been thoroughly investigated; see e.g., [5], [8], [20], [21]. Recently, cyclostationary detection from sub-Nyquist samples was treated in [22]–[30]. A general framework is adopted, that exploits a linear relation between the sub-Nyquist and Nyquist samples, over a finite sensing time. In particular, a transformation between the Nyquist cyclic spectrum and the time-varying correlations of the sub-Nyquist samples is derived to retrieve the former from the latter. In [22], the carrier frequencies, symbol periods and modulation types of the transmissions are assumed to be known. In this case, the cyclic spectrum can be reduced to its potential non zero cyclic frequencies, which are recovered using simple least squares (LS). However, this scheme cannot be applied in the context of blind spectrum sensing for CR.

The authors in [23], [24] consider low-pass compressive measurements of the correlation function. The cyclic autocorrelation is then reconstructed at a given lag using compressed sensing (CS) algorithms [31]. Two heuristic detection techniques are developed in order to infer the presence or absence of a transmission. However, it is not clear how the signal itself should be sampled in order to obtain these correlation measurements. In addition, no requirements on the number of samples for recovering the cyclic autocorrelation or guarantees on the detection methods are provided.

In [25], [26], the authors consider a random linear relation between the sub-Nyquist samples autocorrelation and Nyquist cyclic spectrum and formulate a ℓ_1 -norm regularized LS problem, enforcing sparsity on the latter. The same ideas are adopted in [27] but the reconstruction is performed in matrix form, allowing for higher resolution. In [28], correlation lags beyond lag zero are exploited and the span of the random linear projections is extended beyond one period of cyclostationarity of the signal. These two extensions allow for simple LS recovery without any sparsity requirements on the signal.

The main drawback of this digital approach, adopted by all works above, is that it does not deal with the sampling scheme itself, since we do not have access to the Nyquist samples. It simply assumes that the sub-Nyquist samples can be expressed as random linear projections of the Nyquist samples. In addition, due to the inherent finite sensing time, the recovered cyclic frequencies lie on a predefined grid. Therefore, in the above works, the frequencies of interest are assumed to lie on that grid. The theoretical resolution that can be achieved is thus dictated by the sensing time. Moreover, no theoretical guarantees on the minimal sampling rate allowing

for perfect recovery of the cyclic spectrum have been given.

In [29], a concrete sampling scheme is considered, known as multicaset, or non-uniform sampling. The authors derive conditions on the system matrix to have full rank, allowing for perfect cyclic spectrum recovery from the compressive measurements. Random sampling in the form of a successive approximation ADC (SAR-ADC) architecture [32] is used in [30]. The resulting sampling matrix is similar to multicaset, with the distinction that the grid is that of the quantization time rather than the Nyquist grid. While explicit sampling schemes are considered here, the theoretical cyclic spectrum resolution still depends on the sensing time, so that the gridding, or discretization, is part of the theoretical derivations.

In this work, we propose to reconstruct the signal's cyclic spectrum from sub-Nyquist samples obtained using the methods of [12]–[14]. Our theoretical approach does not involve gridding or discretization and the cyclic spectrum can be recovered at any frequency. In addition, the MWC analog front-end presented in [13] is a practical sampling scheme that has been implemented in hardware [33]. We perform cyclostationarity detection on the sub-Nyquist samples, thereby obtaining both an efficient, fast and frugal detector on the one hand and one that is reliable and robust to noise on the other. We derive a sampling rate bound allowing for perfect cyclic spectrum recovery in our settings, for sparse and non sparse signals. We note that the cyclic spectrum can be perfectly recovered in the presence of stationary noise, from compressed samples, except for a limited number of cyclic frequencies that are multiples of the basic low sampling rate. For those, the reconstruction is performed in the presence of bounded noise.

In particular, we show that, in the presence of stationary noise, the cyclic spectrum can be reconstructed from samples obtained at $4/5$ of the Nyquist rate, without any sparsity assumption on the signal. If the signal of interest is sparse, then the sampling rate can be further reduced to $8/5$ of the Landau rate, which is the Lebesgue measure of the occupied bandwidth [34]. Similar results were observed in [18] in the context of power spectrum reconstruction of stationary signals. There, it was shown that the power spectrum of non sparse signals can be retrieved at half the Nyquist rate and that of sparse signals can be perfectly recovered at the Landau rate. Once the cyclic spectrum is reconstructed, we apply our feature extraction algorithm, presented in [35] in the Nyquist regime, that estimates the number of transmissions and their respective carrier frequencies and bandwidths.

The main contributions of this paper are as follows:

- **Low rate sampling and digital processing** - the cyclic spectrum is recovered directly from sub-Nyquist samples. Both sampling and digital processing are performed at a low rate.
- **Structured CS algorithm** - an orthogonal matching pursuit (OMP) based algorithm to reconstruct the cyclic spectrum is proposed that exploits the inherent structure of the correlation matrices between frequency samples of the signal.
- **Robust detection in low SNR** - we show that cyclostationary detection performed by estimating the transmissions carrier frequency and bandwidth is more robust to

noise than energy detection at sub-Nyquist rates.

- **Minimal sampling rate derivation** - a lower bound on the sampling rate required for cyclic spectrum recovery is derived for both sparse and non sparse signals.

This paper is organized as follows. In Section II, we describe the cyclostationary multiband model. Sections III and IV present the sub-Nyquist sampling stage and cyclic spectrum reconstruction algorithm and conditions, respectively. Numerical experiments are presented in Section V.

II. CYCLOSTATIONARY MULTIBAND MODEL

A. Multiband Model

Let $x(t)$ be a real-valued continuous-time signal, supported on $\mathcal{F} = [-1/2T_{\text{Nyq}}, +1/2T_{\text{Nyq}}]$ and composed of up to N_{sig} uncorrelated cyclostationary transmissions corrupted by additive noise, such that

$$x(t) = \sum_{i=1}^{N_{\text{sig}}} s_i(t) + n(t). \quad (1)$$

Here $n(t)$ is a wide-sense stationary bandpass noise and $s_i(t)$ is a zero-mean cyclostationary bandpass process, as defined below, from the class of pulse-amplitude modulation (PAM) signals:

$$s_i(t) = \sqrt{2} \cos(2\pi f_i t) \sum_k a_{ik}^I g_i(t - kT_i) - \sqrt{2} \sin(2\pi f_i t) \sum_k a_{ik}^Q g_i(t - kT_i). \quad (2)$$

The unknown symbols modulating the in-phase and quadrature components are denoted $\{a_{ik}^I\}$ and $\{a_{ik}^Q\}$, respectively, and $g_i(t)$ are the unknown pulse shape functions. The single-sided bandwidth, the carrier frequency and the symbol period are denoted by B_i , f_i and T_i , respectively. Special cases of passband PAM include phase-shift keying (PSK), amplitude and phase modulation (AM-PM) and quadrature amplitude modulation (QAM) [36].

Formally, the Fourier transform of $x(t)$, defined by

$$X(f) = \lim_{T \rightarrow \infty} \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} x(t) e^{-j2\pi f t} dt, \quad (3)$$

is zero for every $f \notin \mathcal{F}$. We denote by $f_{\text{Nyq}} = 1/T_{\text{Nyq}}$ the Nyquist rate of $x(t)$. The number of transmissions N_{sig} , their carrier frequencies, bandwidths, symbol rates and modulations, including the symbols $\{a_{ik}\}$ and the pulse shape functions $g_i(t)$ are unknown, namely the reconstruction of the cyclic spectrum, defined in the next section, is performed in a blind scenario. The single-sided bandwidth of each transmission is only assumed to not exceed a known maximal bandwidth B , namely $B_i \leq B$ for all $1 \leq i \leq N_{\text{sig}}$. If the bandwidth is fully occupied, then $N_{\text{sig}}B$ is on the order of f_{Nyq} .

We will consider the special case of sparse multiband signals as well and show that the sampling rate for perfect cyclic spectrum reconstruction can be further reduced. In this setting, the number of transmissions N_{sig} which dictates the signal sparsity, or at least an upper bound on it, is assumed to be known and $N_{\text{sig}}B \ll f_{\text{Nyq}}$. We denote by $K = 2N_{\text{sig}}$

the upper bound on the number of occupied bands, to account for both the positive and negative frequency bands for each signal.

B. Cyclostationarity

A process $s(t)$ is said to be cyclostationary with period T_0 in the wide sense if its mean $\mathbb{E}[s(t)] = \mu_s(t)$ and autocorrelation $\mathbb{E}[s(t - \tau/2)s(t + \tau/2)] = R_s(t, \tau)$ are both periodic with period T_0 [4]:

$$\mu_s(t + T_0) = \mu_s(t), \quad R_s(t + T_0, \tau) = R_s(t, \tau). \quad (4)$$

Given a wide-sense cyclostationary random process, its autocorrelation $R_s(t, \tau)$ can be expanded in a Fourier series

$$R_s(t, \tau) = \sum_{\alpha} R_s^{\alpha}(\tau) e^{j2\pi\alpha t}, \quad (5)$$

where $\alpha = m/T_0$, $m \in \mathbb{Z}$ and the Fourier coefficients, referred to as cyclic autocorrelation functions, are given by

$$R_s^{\alpha}(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} R_s(t, \tau) e^{-j2\pi\alpha t} dt. \quad (6)$$

The cyclic spectrum is obtained by taking the Fourier transform of (6) with respect to τ , namely

$$S_s^{\alpha}(f) = \int_{-\infty}^{\infty} R_s^{\alpha}(\tau) e^{-j2\pi f \tau} d\tau, \quad (7)$$

where α is referred to as the cyclic frequency and f is the angular frequency [4]. If there is more than one fundamental cyclic frequency $1/T_0$, then the process $s(t)$ is said to be polycyclostationary in the wide sense. In this case, the cyclic spectrum contains harmonics (integer multiples) of each of the fundamental cyclic frequencies [6]. These cyclic frequencies are related to the transmissions carrier frequencies and symbol rates as well as the modulation type.

An alternative interpretation of the cyclic spectrum, which we will exploit, expresses it as the cross-spectral density $S_s^{\alpha}(f) = S_{uv}(f)$ of two frequency-shifted versions of $s(t)$, $u(t)$ and $v(t)$, such that

$$u(t) \triangleq s(t) e^{-j\pi\alpha t}, \quad (8)$$

$$v(t) \triangleq s(t) e^{+j\pi\alpha t}. \quad (9)$$

Then, from [37], it holds that

$$S_s^{\alpha}(f) = S_{uv}(f) = \mathbb{E} \left[S \left(f + \frac{\alpha}{2} \right) S^* \left(f - \frac{\alpha}{2} \right) \right]. \quad (10)$$

The Fourier transform of a wide-sense stationary process $n(t)$ is a white noise process, not necessarily stationary, so that [37]

$$\mathbb{E} [N(f_1)N^*(f_2)] = S_n^0(f_1)\delta(f_1 - f_2). \quad (11)$$

Therefore, stationary noise exhibits no cyclic correlation [6], that is

$$S_n^{\alpha}(f) = 0, \quad \alpha \neq 0. \quad (12)$$

This property is the motivation for cyclostationary detection, in low SNR regimes in particular.

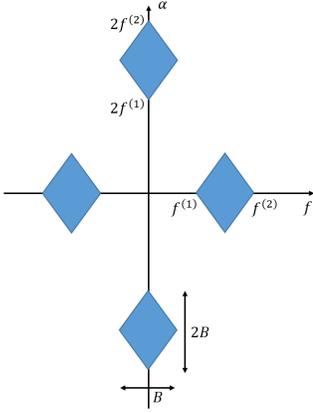


Fig. 1. Support region of the cyclic spectrum of $S_{s_i}^\alpha(f)$.

C. Cyclic Spectrum of Multiband Signals

Denote by $[f_i^{(1)}, f_i^{(2)}]$ the right-side support of the i th transmission $s_i(t)$. Then, $B_i = f_i^{(2)} - f_i^{(1)}$ and $f_i = (f_i^{(1)} + f_i^{(2)})/2$. The support region in the (f, α) plane of the cyclic spectrum $S_{s_i}^\alpha(f)$ of such a bandpass cyclostationary signal is composed of four diamonds, as shown in Fig. 1. More precisely, it holds that [6]

$$S_{s_i}^\alpha(f) = 0, \quad \text{for } \left| |f| - \frac{|\alpha|}{2} \right| \leq f_i^{(1)} \text{ or } |f| + \frac{|\alpha|}{2} \geq f_i^{(2)}. \quad (13)$$

Moreover, since $x(t)$ is bandlimited to \mathcal{F} , it follows that [6]

$$S_x^\alpha(f) = 0, \quad \text{for } |f| + \frac{|\alpha|}{2} \geq \frac{f_{\text{Nyq}}}{2}. \quad (14)$$

Since the transmissions $s_i(t)$ are assumed to be zero-mean and uncorrelated (coming from different sources), the cyclic spectrum of $x(t)$ does not contain any additional harmonics which would result from correlation between different transmissions. It is thus given by

$$S_x^\alpha(f) = \begin{cases} \sum_{i=1}^{N_{\text{sig}}} S_{s_i}^\alpha(f) & \alpha \neq 0 \\ \sum_{i=1}^{N_{\text{sig}}} S_{s_i}^0(f) + S_n^0(f) & \alpha = 0. \end{cases} \quad (15)$$

At the cyclic frequency $\alpha = 0$, the cyclic spectrum reduces to the power spectrum and the occupied bandwidth along the angular frequency axis is $2N_{\text{sig}}B$, to account for both positive and negative frequency bands. This cyclic frequency $\alpha = 0$ contains the noise power spectrum as well. As a consequence, we choose to detect the transmissions and estimate their carriers and bandwidths at cyclic frequencies $\alpha \neq 0$. Note that, for $\alpha \neq 0$, the sum only contains the contributions of the transmissions that exhibit cyclostationarity at the corresponding cyclic frequency α . Therefore, for each $\alpha \neq 0$, the sum typically contains less than N_{sig} non zero elements. It follows from (15) that, besides the noise contribution at the cyclic frequency $\alpha = 0$, the support of $S_x^\alpha(f)$ is composed of $4N_{\text{sig}}$ diamonds, that is four diamonds for each transmission, as those shown in Fig. 1.

The support of $S_x^\alpha(f)$ consists of two types of correlations: the two diamonds lying on the angular frequency f axis and those lying on the cyclic frequency α . The diamonds on the f axis contain self-correlations between a band and its shifted version. These correspond to cyclic peaks at locations (f, α) , with $f \in \mathcal{F}$ and $0 \leq |\alpha| \leq B$. Cyclic features at these locations are the result of the transmissions symbol rate $1/T_i$. The value of T_i can be derived from the bandwidth, as

$$T_i = \frac{1 + \gamma_i}{B_i}, \quad (16)$$

where $0 \leq \gamma_i \leq 1$ is the unknown excess-bandwidth parameter of $g_i(t)$ [36]. Therefore, cyclic peaks that stem from the symbol rate appear at cyclic frequencies $\alpha = \frac{1}{T_i} \leq B_i \leq B$.

The two diamonds lying on the α axis contain cross-correlations between two symmetric bands, belonging to the same transmission. These correspond to cyclic peaks at locations (f, α) , with $0 \leq |f| \leq B/2$ and $B < |\alpha| \leq f_{\text{Nyq}}$. In particular, applying (10) for $\alpha = \pm 2f_i$, namely

$$S_x^{\pm 2f_i}(f) = \mathbb{E}[X(f + f_i)X^*(f - f_i)], \quad (17)$$

computes the correlation between the positive and negative bands of the i th transmission. This generates a peak, or cyclic feature, in the cyclic spectrum at location $(f = 0, \alpha = \pm 2f_i)$. By detecting the peak location, one can estimate the carrier frequency of the corresponding transmission $s_i(t)$. Furthermore, from (13), it is clear that the occupied bandwidth in the cyclic spectrum at $\alpha = \pm 2f_i$ is equal to the bandwidth B_i . These observations are the key to estimating the transmissions carrier frequency and bandwidth. We note that other modulations not included in (2) may not exhibit features at $\alpha = \pm 2f_i$. In such cases, higher order cyclostationary statistics may be used for detection [6], [38].

D. Goal

Our objective is to reconstruct $S_x^\alpha(f)$ from sub-Nyquist samples without any *a priori* knowledge on the support and modulations of $s_i(t)$, $1 \leq i \leq N_{\text{sig}}$. We show that the cyclic spectrum of non sparse signals can be recovered from samples obtained at $4/5$ of the Nyquist rate and if the signal is assumed to be sparse, then the sampling rate can be as low at $8/5$ of the Landau rate.

We then estimate the number of transmissions N_{sig} present in $x(t)$, their carrier frequencies f_i and their bandwidths B_i . Once the signals' carrier frequency and bandwidth are estimated, the occupied support is determined. In [35], we proposed a generic feature extraction algorithm for the estimation of N_{sig} , f_i and B_i , for all $1 \leq i \leq N_{\text{sig}}$, from the cyclic spectrum obtained from Nyquist samples. Here, we apply the same scheme to the reconstructed cyclic spectrum from sub-Nyquist rather than Nyquist samples.

III. SUB-NYQUIST SAMPLING

In this section, we briefly describe the sub-Nyquist sampling schemes we adopt. We consider two different sampling methods: multicoset sampling [12] and the MWC [13] which were previously proposed for sparse multiband signals in conjunction with energy detection. We show that both techniques

lead to identical expressions of the signal cyclic spectrum in terms of correlations between the samples. Therefore, the cyclic spectrum reconstruction stage presented in Section IV can be applied to either of the sampling approaches.

A. Multicoset sampling

Multicoset sampling [39] can be described as the selection of certain samples from the uniform grid. More precisely, the uniform grid is divided into blocks of N consecutive samples, from which only M are kept. The i th sampling sequence is defined as

$$x_{c_i}[n] = \begin{cases} x(nT_{\text{Nyq}}), & n = mN + c_i, m \in \mathbb{Z} \\ 0, & \text{otherwise,} \end{cases} \quad (18)$$

where $0 < c_1 < c_2 < \dots < c_M < N - 1$. Multicoset sampling can be implemented as a multi-channel system with M channels, each composed of a delay unit corresponding to the coset c_i followed by a low rate ADC. Let $f_s = \frac{1}{NT_{\text{Nyq}}} \geq B$ be the sampling rate of each channel and $\mathcal{F}_s = [0, f_s]$. Following the derivations in [12], we obtain

$$\mathbf{z}(\tilde{f}) = \mathbf{A}\mathbf{x}(\tilde{f}), \quad \tilde{f} \in \mathcal{F}_s, \quad (19)$$

where $\mathbf{z}_i(\tilde{f}) = X_{c_i}(e^{j2\pi\tilde{f}T_{\text{Nyq}}})$, $0 \leq i \leq M - 1$, are the discrete-time Fourier transforms (DTFTs) of the multicoset samples and

$$\mathbf{x}_k(\tilde{f}) = X(\tilde{f} + K_k f_s), \quad 1 \leq k \leq N, \quad (20)$$

where $K_k = k - \frac{N+1}{2}$, $1 \leq k \leq N$ for odd N and $K_k = k - \frac{N+2}{2}$, $1 \leq k \leq N$ for even N . Each entry of $\mathbf{x}(f)$ is referred to as a slice of the spectrum of $x(t)$. The ik th element of the $M \times N$ matrix \mathbf{A} is given by

$$\mathbf{A}_{ik} = \frac{1}{NT_{\text{Nyq}}} e^{j\frac{2\pi}{N}c_i K_k}, \quad (21)$$

namely \mathbf{A} is determined by the known sampling pattern $\{c_i\}_{i=1}^M$.

B. MWC sampling

The MWC [13] is composed of M parallel channels. In each channel, an analog mixing front-end, where $x(t)$ is multiplied by a mixing function $p_i(t)$, aliases the spectrum, such that each band appears in baseband, as illustrated in Fig. 2. The mixing functions $p_i(t)$ are required to be periodic with period T_p such that $f_p = 1/T_p \geq B$. The function $p_i(t)$ has a Fourier expansion

$$p_i(t) = \sum_{l=-\infty}^{\infty} c_{il} e^{j\frac{2\pi}{T_p} l t}. \quad (22)$$

In each channel, the signal goes through a lowpass filter with cut-off frequency $f_s/2$ and is sampled at the rate $f_s \geq f_p$, resulting in the samples $y_i[n]$. For the sake of simplicity, we choose $f_s = f_p$. From [13], the relation between the known DTFTs of the samples $y_i[n]$ and the unknown $X(f)$ is given by

$$\mathbf{z}(\tilde{f}) = \mathbf{A}\mathbf{x}(\tilde{f}), \quad \tilde{f} \in \mathcal{F}_s, \quad (23)$$

where $\mathbf{z}(\tilde{f})$ is a vector of length M with i th element $\mathbf{z}_i(\tilde{f}) = Y_i(e^{j2\pi\tilde{f}T_s})$. The entries of the unknown vector $\mathbf{x}(\tilde{f})$ are given

by (20). The $M \times N$ matrix \mathbf{A} contains the known coefficients c_{il} such that

$$\mathbf{A}_{il} = c_{i,-l} = c_{il}^*, \quad (24)$$

where $N = \lceil f_{\text{Nyq}}/f_s \rceil$. This relation is illustrated in Fig. 2.

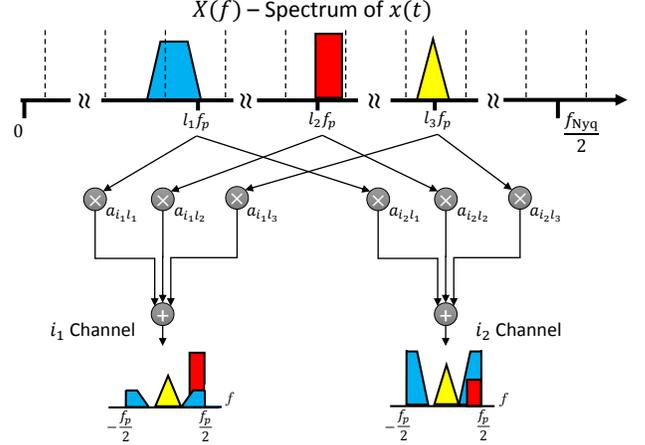


Fig. 2. The spectrum slices $\mathbf{x}(\tilde{f})$ of the input signal are shown here to be multiplied by the coefficients a_{il} of the sensing matrix \mathbf{A} , resulting in the measurements $\mathbf{z}_i(\tilde{f})$ for the i th channel. Note that in multicoset sampling, only the slices' complex phase is modified by the coefficients a_{il} . In the MWC, both the phases and amplitudes are affected in general.

Systems (19) and (23) are identical for both sampling schemes: the only difference is the sampling matrix \mathbf{A} . In the next section, we derive conditions for the reconstruction of the cyclic spectrum from either class of samples. The requirements on the resulting sampling matrix \mathbf{A} are tied to conditions on the sampling pattern for multicoset sampling [12] and on the mixing sequences for the MWC [13]. We then present a method for reconstruction of the analog cyclic spectrum for both sampling schemes jointly. In particular, we will reconstruct $S_x^\alpha(f)$ from correlations between shifted versions of $\mathbf{z}(\tilde{f})$, defined in (19) and (23). We note that for both sampling approaches, the overall sampling rate is

$$f_{tot} = M f_s = \frac{M}{N} f_{\text{Nyq}}. \quad (25)$$

In the simulations, we consider samples obtained using the MWC. However, multicoset samples can be used indifferently.

IV. CYCLIC SPECTRUM RECONSTRUCTION

In this section, we provide a method to reconstruct the cyclic spectrum $S_x^\alpha(f)$ of $x(t)$ from sub-Nyquist samples obtained using either of the sampling schemes described above, namely multicoset and the MWC. We also investigate cyclic spectrum recovery conditions, that is the minimal sampling rate allowing for perfect recovery of $S_x^\alpha(f)$ in the presence of stationary noise. Finally, we consider the special case of power spectrum reconstruction, presented in [18], and compare it to cyclic spectrum in terms of both recovery method and conditions.

A. Relation between Samples and Cyclic Spectrum

From (19) or (23), we have

$$\mathbf{R}_z^a(\tilde{f}) = \mathbf{A}\mathbf{R}_x^a(\tilde{f})\mathbf{A}^H, \quad \tilde{f} \in [0, f_s - a], \quad (26)$$

for all $a \in [0, f_s]$, where

$$\mathbf{R}_x^a(\tilde{f}) = \mathbb{E} \left[\mathbf{x}(\tilde{f})\mathbf{x}^H(\tilde{f} + a) \right], \quad (27)$$

and

$$\mathbf{R}_z^a(\tilde{f}) = \mathbb{E} \left[\mathbf{z}(\tilde{f})\mathbf{z}^H(\tilde{f} + a) \right]. \quad (28)$$

Here $(\cdot)^H$ denotes the Hermitian operation. The entries in the matrix $\mathbf{R}_x^a(\tilde{f})$ are correlations between shifted versions of the slices $\mathbf{x}(\tilde{f})$, namely correlations between frequency-shifted versions of $x(t)$. The variable a controls the shift between the slices, while \tilde{f} , obtaining values in the interval $[0, f_s - a]$, determines the specific frequency location within the slice. Both a and \tilde{f} are continuous and $\mathbf{R}_x^a(\tilde{f})$ can be computed for any combination of $a \in [0, f_s]$ and $\tilde{f} \in [0, f_s - a]$. In practice, with a limited sensing time, a and \tilde{f} are discretized according to the number of samples per channel.

To proceed, we begin by investigating the link between the cyclic spectrum $S_x^\alpha(f)$ and the shifted correlations between the slices $\mathbf{x}(\tilde{f})$, namely the entries of $\mathbf{R}_x^a(\tilde{f})$. We then show how the latter can be recovered from $\mathbf{R}_z^a(\tilde{f})$ using (26).

The alternative definition of the cyclic spectrum (10) implies that the elements in the matrix $\mathbf{R}_x^a(\tilde{f})$ are equal to $S_x^\alpha(f)$ at the corresponding α and f . Indeed, it can easily be shown that

$$\mathbf{R}_x^a(\tilde{f})_{(i,j)} = S_x^\alpha(f), \quad (29)$$

for

$$\begin{aligned} \alpha &= (j - i)f_s + a \\ f &= -\frac{f_{\text{Nyq}}}{2} + \tilde{f} - \frac{f_s}{2} + \frac{(j + i)f_s}{2} + \frac{a}{2}. \end{aligned} \quad (30)$$

Here $\mathbf{R}_x^a(\tilde{f})_{(i,j)}$ denotes the (i, j) th element of $\mathbf{R}_x^a(\tilde{f})$. In particular, from (15), it follows that for $a = 0$, namely with no shift, it holds that

$$\mathbf{R}_x^0(\tilde{f}) = \sum_{k=1}^{N_{\text{sig}}} \mathbf{R}_{\mathbf{s}_k}^0(\tilde{f}) + \mathbf{R}_n^0(\tilde{f}). \quad (31)$$

Here, $\mathbf{R}_n^0(\tilde{f})$ is a diagonal matrix that contains the noise's power spectrum, such that

$$\mathbf{R}_n^0(\tilde{f})_{kk} = S_n^0 \left(-\frac{f_{\text{Nyq}}}{2} + kf_s + \tilde{f} \right), \quad (32)$$

for $0 \leq k \leq N - 1$, and

$$\mathbf{R}_{\mathbf{s}_k}^a(\tilde{f}) = \mathbb{E} \left[\mathbf{s}_k(\tilde{f})\mathbf{s}_k^H(\tilde{f} + a) \right], \quad (33)$$

where $\mathbf{s}_k(\tilde{f})$ is a vector of size N whose non zero elements are the frequency slices from $\mathbf{x}(\tilde{f})$ corresponding to the k th transmission. The diagonal of $\mathbf{R}_{\mathbf{s}_k}^0(\tilde{f})_{kk}$, for $0 \leq k \leq N - 1$, contains the power spectrum of $x(t)$ such that

$$\mathbf{R}_{\mathbf{s}_k}^0(\tilde{f})_{kk} = S_x^0 \left(-\frac{f_{\text{Nyq}}}{2} + kf_s + \tilde{f} \right), \quad (34)$$

is the sum of the transmissions' and noise's power spectrum. Since $n(t)$ is stationary, for $a \neq 0$, we have

$$\mathbf{R}_x^a(\tilde{f}) = \sum_{k=1}^{N_{\text{sig}}} \mathbf{R}_{\mathbf{s}_k}^a(\tilde{f}). \quad (35)$$

Our goal can then be stated as recovery of $\mathbf{R}_x^a(\tilde{f})$, for $a \in [0, f_s]$ and $\tilde{f} \in [0, f_s - a]$, since once $\mathbf{R}_x^a(\tilde{f})$ is known, $S_x^\alpha(f)$ follows for all (α, f) , using (29).

We now consider the structure of the autocorrelation matrices $\mathbf{R}_x^a(\tilde{f})$, which is related to the support of the cyclic spectrum $S_x^\alpha(f)$. In Section II-B, we discussed the support of $S_x^\alpha(f)$, composed of two types of correlations: self-correlations between a band and its shifted version and cross-correlations between shifted versions of symmetric bands belonging to the same transmission. Consider a given frequency location \tilde{f} and shift a . The frequency component $\mathbf{x}_i(\tilde{f})$, for $0 \leq i \leq N - 1$, can be correlated to at most two entries of $\mathbf{x}(\tilde{f} + a)$, one from the same band and one from the symmetric band. The correlated component can be either in the same, respectively symmetric, slice or in one of the adjacent slices. This follows from the fact that each band may split between two slices at most, since we require $f_p \geq B$. Thus, the first correlated entry is either i or $i \pm 1$ and the second is either $N - i$, $N - i + 1$ or $N - 1 + 2$. Since the noise is assumed to be wide-sense stationary, from (11), a noise frequency component is correlated only with itself. Thus, $n(t)$ can contribute non-zero elements only on the diagonal of $\mathbf{R}_x^0(\tilde{f})$.

Figures 4 and 5 illustrate these correlations for $a = 0$ and $a = f_s/2$, respectively. First, in Fig. 3, an illustration of the spectrum of $x(t)$, namely $X(f)$, is presented for the case of a sparse signal buried in stationary bandpass noise. It can be seen that frequency bands of $X(f)$ either appear in one f_p -slice or split between two slices. The resulting vector of spectrum slices $\mathbf{x}(f)$ and the correlations between these slices without any shift, namely $\mathbf{R}_x^0(\tilde{f})$, are shown in Figs. 4(a) and (b), respectively. Define the d -diagonal and d -anti diagonal of an $N \times N$ matrix to be its (i, j) th elements such that $j = i + d$, and $j = N - i + 1 - d$, respectively. In particular, the 0-diagonal stands for the main diagonal and the 0-anti diagonal is the secondary diagonal. In Fig. 4(b), we observe that self-correlations appear only on the main diagonal since every frequency component is correlated with itself. In particular, the main diagonal contains the noise's power spectrum (in green). Cross-correlations between the yellow symmetric triangles appear in the 0-anti diagonal, whereas those of the blue trapeziums are contained in the -1 and 1 -anti diagonals. The red rectangles do not contribute any cross-correlations for a shift of $a = 0$. Figures 5(a) and (b) show the vector $\mathbf{x}(\tilde{f})$ and its shifted version $\mathbf{x}(\tilde{f} + a)$ for $a = f_s/2$, respectively. The resulting correlation matrix $\mathbf{R}_x^a(\tilde{f})$ appears in Fig. 5(c). Here, the self correlations of the blue trapezium appear in the -1 -diagonal. The non zero cross-correlations all appear in the anti-diagonal, for the shift $a = f_s/2$. We note that for this shift, the yellow triangles do not contribute self or cross correlations.

The following four conclusions can be drawn from the observations above on the structure of $\mathbf{R}_x^a(\tilde{f})$, for a given

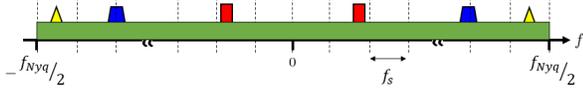


Fig. 3. Original spectrum $X(f)$.

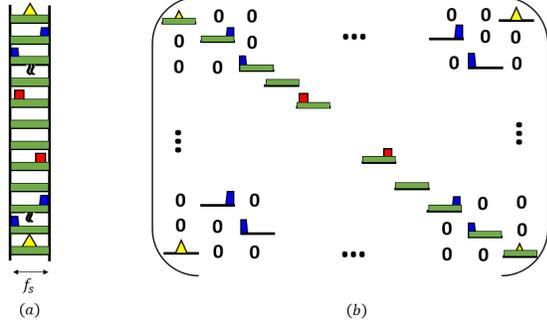


Fig. 4. (a) Spectrum slices vector $\mathbf{x}(\tilde{f})$ (b) correlated slices of $\mathbf{x}(\tilde{f})$ in the matrix $\mathbf{R}_{\mathbf{x}}^a(\tilde{f})$.

$a \in]0, f_s]$ and $\tilde{f} \in [0, f_s - a]$. We will treat the case where $a = 0$ separately since it yields a different structure due to the presence of noise.

- **Conclusion 1:** The non zeros entries of $\mathbf{R}_{\mathbf{x}}^a(\tilde{f})$ are contained in its $-1, 0$ and 1 -diagonals and $-1, 0$ and 1 -anti diagonals.
- **Conclusion 2:** The i th row of $\mathbf{R}_{\mathbf{x}}^a(\tilde{f})$ contains at most two non zero elements at locations $(i, g(i))$ and $(i, g(N - i + 1))$, where

$$g(i) = \begin{cases} i \text{ or } i \pm 1 & 2 \leq i \leq N - 1, \\ i \text{ or } i + 1 & i = 1, \\ i \text{ or } i - 1 & i = N. \end{cases} \quad (36)$$

- **Conclusion 3:** The j th column of $\mathbf{R}_{\mathbf{x}}^a(\tilde{f})$ contains at most two non zero elements at locations $(g(j), j)$ and $(g(N - j + 1), j)$, for $1 \leq j \leq N$.
- **Conclusion 4:** For each specific frequency \tilde{f} , a transmission contributes to at most two slices, one in the negative and one in the positive frequencies. This is due to the assumption that $f_p \geq B$. Therefore, $\mathbf{R}_{\mathbf{x}}^a(\tilde{f})$ contains

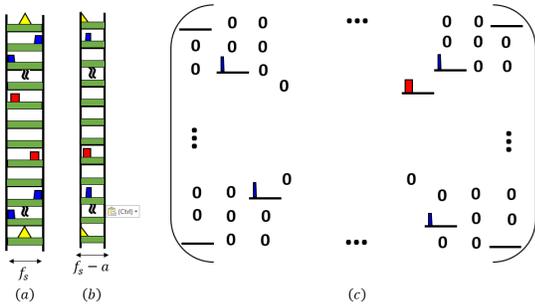


Fig. 5. (a) Spectrum slices vector $\mathbf{x}(\tilde{f})$ (b) spectrum slices shifted vector $\mathbf{x}(\tilde{f} + a)$ for $a = f_s/2$ (c) correlated slices of $\mathbf{x}(\tilde{f})$ and $\mathbf{x}(\tilde{f} + a)$ in the matrix $\mathbf{R}_{\mathbf{x}}^a(\tilde{f})$, with $a = f_s/2$.

at most $K = 2N_{\text{sig}}$ rows/columns that have non zero elements. Without any sparsity assumption, it is obvious that $K = N$ even if the number of transmissions is greater than $N/2$.

From **Conclusions 2** and **4**, it follows that $\mathbf{R}_{\mathbf{x}}^a(\tilde{f})$ is $2K$ -sparse and has additional structure described in **Conclusions 1-3**.

Since the non zero elements of $\mathbf{R}_{\mathbf{x}}^a(\tilde{f})$ only lie on the 3 main and anti-diagonals, (26) can be further reduced to

$$\mathbf{r}_{\mathbf{z}}^a(\tilde{f}) = (\bar{\mathbf{A}} \otimes \mathbf{A}) \text{vec}(\mathbf{R}_{\mathbf{x}}^a(\tilde{f})) = (\bar{\mathbf{A}} \otimes \mathbf{A}) \mathbf{B} \mathbf{r}_{\mathbf{x}}^a(\tilde{f}) \triangleq \Phi \mathbf{r}_{\mathbf{x}}^a(\tilde{f}), \quad (37)$$

where $\bar{\mathbf{A}}$ denotes the conjugate matrix of \mathbf{A} and

$$\Phi = (\bar{\mathbf{A}} \otimes \mathbf{A}) \mathbf{B}. \quad (38)$$

Here \otimes is the Kronecker product, $\mathbf{r}_{\mathbf{z}}^a(\tilde{f}) = \text{vec}(\mathbf{R}_{\mathbf{z}}^a(\tilde{f}))$, where $\text{vec}(\cdot)$ denotes the column stack concatenation operation, and \mathbf{B} is a $N^2 \times (6N - 4)$ selection matrix that selects the elements of the $-1, 0$ and 1 -diagonals and anti-diagonals of $\mathbf{R}_{\mathbf{x}}^a(\tilde{f})$ from the vector $\text{vec}(\mathbf{R}_{\mathbf{x}}^a(\tilde{f}))$. The resulting $(6N - 4) \times 1$ vector composed of these selected elements is denoted by $\mathbf{r}_{\mathbf{x}}^a(\tilde{f})$.

From **Conclusions 2-4**, the vector $\mathbf{r}_{\mathbf{x}}^a(\tilde{f})$ is $2K$ -sparse and its support presents additional structure. Denote by Σ_k the set of k -sparse vectors that belong to a linear subspace Σ . In our case, $\mathbf{r}_{\mathbf{x}}^a(\tilde{f})$ belongs to $\Sigma_k = \Sigma_{2K}$, defined as

$$\Sigma_{2K} = \{\mathbf{x} \in \mathbb{R}^{(6N-4) \times 1} \mid |S(\mathbf{x})| \leq 2K \wedge S(\mathbf{x}) \in \mathcal{I}\}, \quad (39)$$

where $S(\mathbf{x})$ denotes the support of \mathbf{x} and the set \mathcal{I} is determined by the following properties:

- 1) If the group indexed by j is in \mathcal{I} , then the group indexed by $(N + 1 - j)$ is in \mathcal{I} as well. The groups are defined in the following items.
- 2) If the group indexed by $2 \leq j \leq N - 1$ is in \mathcal{I} , then it means that at most one of the entries $\{6j - 1, 6j, 6j + 1\}$ and at most one of the entries $\{6j + 2, 6j + 3, 6j + 4\}$ are in \mathcal{I} .
- 3) If the group indexed by $j = 1$ is in \mathcal{I} , then at most one of the entries $\{1, 2\}$ and at most one of the entries $\{3, 4\}$ are in \mathcal{I} .
- 4) If the group indexed by $j = N$ is in \mathcal{I} , then at most one of the entries $\{6N - 7, 6N - 6\}$ and at most one of the entries $\{6N - 5, 6N - 4\}$ are in \mathcal{I} .

Item 1) follows from **Conclusion 4** and items 2)-3) and 4) follow from **Conclusions 2-3**. We note that the indexation above is valid if the selection matrix in (37) selects the elements of $\text{vec}(\mathbf{R}_{\mathbf{x}}^a(\tilde{f}))$ row by row. Any other selection order would yield a different indexation.

Consider now the case in which $a = 0$. The matrix $\mathbf{R}_{\mathbf{x}}^0(\tilde{f})$ contains the power spectrum of the transmissions and the noise on its main diagonal and cyclic components of the transmissions on its $-1, 0$ and 1 -anti diagonals. The remaining elements are zero. Denote by $\mathbf{r}_{\mathbf{x}_1}^0$ the $N \times 1$ non-sparse vector composed of the diagonal of $\mathbf{R}_{\mathbf{x}}^0(\tilde{f})$ and by $\mathbf{r}_{\mathbf{x}_2}^0$ the $(3N - 2) \times 1$ sparse vector composed of its $-1, 0$ and 1 -anti diagonals. Let $\Phi_1 = \bar{\mathbf{A}} \odot \mathbf{A}$ where \odot denotes the Khatri-Rao product and $\Phi_2 = (\bar{\mathbf{A}} \otimes \mathbf{A}) \bar{\mathbf{B}}$ with $\bar{\mathbf{B}}$ a $N^2 \times (3N - 2)$ selection matrix that selects the elements of the $-1, 0$ and 1 -anti diagonals of $\mathbf{R}_{\mathbf{x}}^0(\tilde{f})$ from the vector

$\text{vec}(\mathbf{R}_x^0(\tilde{f}))$. Combining (31) and (37), we can write $\mathbf{r}_z^0(\tilde{f})$ as the sum of two components as follows,

$$\mathbf{r}_z^0(\tilde{f}) = \Phi_2 \mathbf{r}_{x_2}^0(\tilde{f}) + \mathbf{w}(\tilde{f}), \quad (40)$$

where $\mathbf{w}(\tilde{f}) = \Phi_1 \mathbf{r}_{x_1}^0(\tilde{f})$ is the noise component. We note that the signal's power spectrum is buried in the noise $\mathbf{w}(\tilde{f})$. Therefore, we do not recover it and signal detection will be performed only on cyclic frequencies $\alpha \neq 0$.

From (14) and (29), by recovering $\mathbf{r}_x^a(\tilde{f}) \in \Sigma_{2K}$ for all $a \in [0, f_s]$, $\tilde{f} \in [0, f_s - a]$, we recover the entire cyclic spectrum of $x(t)$. We consider only $a \geq 0$ and consequently $\alpha \geq 0$. We thus only reconstruct half of the cyclic spectrum, known to be symmetric [6]. In (37), there is no noise component, even if the signal $x(t)$ is corrupted by additional stationary noise. For the corresponding cyclic frequencies, we can therefore achieve perfect recovery. In contrast, in (40), namely for $a = 0$, there is an additional noise component. From (30), this case corresponds to cyclic frequencies which are multiples of the channels' sampling frequency f_s . For these frequencies, the recovery of the sparse vector $\mathbf{r}_{x_2}^0(\tilde{f}) \in \Sigma_{2K}$ is not perfect and is performed in the presence of bounded noise. In the simulations, we observe that for detection purposes, this noisy recovery is satisfactory. To achieve perfect recovery for these cyclic frequencies as well, we may sample the signal using a different sampling frequency f_{s_2} .

B. Cyclic Spectrum Recovery Conditions

We now consider conditions for perfect recovery of the cyclic spectrum $S_x^\alpha(f)$ from sub-Nyquist samples. Corollary 1 below derives sufficient conditions on the minimal number of channels M for perfect recovery of $\mathbf{R}_x^a(\tilde{f})$, for any $a \in]0, f_s]$ and $\tilde{f} \in [0, f_s - a]$ in the presence of additive stationary noise. As stated above, for $a = 0$, the recovery is noisy. From **Conclusion 1**, we only need to recover $\mathbf{r}_x^a(\tilde{f})$ from (37) or (40) for $a \neq 0$ or $a = 0$, respectively. Theorem 1 first states sufficient conditions for reconstruction of the vector $\mathbf{r}_x^a(\tilde{f}) \in \Sigma_{2K}$. To that end, we rely on the following Lemma which is well known in the CS literature [15], [31].

Lemma 1. *For any vector $\mathbf{y} \in \mathbb{R}^m$, there exists at most one signal $\mathbf{x} \in \Sigma_k$ such that $\mathbf{y} = \mathbf{A}\mathbf{x}$ if and only if all sets of $2k$ columns of \mathbf{A} belonging to W , such that*

$$W = \{S(\mathbf{x}_1) \cup S(\mathbf{x}_2) | \mathbf{x}_1, \mathbf{x}_2 \in \Sigma_k\}, \quad (41)$$

are linearly independent. In particular, for uniqueness we must have that $m \geq 2k$.

Using Lemma 1, it follows that in order to perfectly recover $\mathbf{r}_x^a(\tilde{f}) \in \Sigma_{2K}$ from $\mathbf{r}_z^a(\tilde{f})$, we need to ensure that all sets of $4K$ columns of Φ belonging to W are linearly independent. In our case, W is the set of unions of the supports of two vectors from Σ_{2K} , as defined in (41). The following theorem relates spark properties of the sampling matrix \mathbf{A} to rank properties of $M^2 \times 4K$ sub-matrices of Φ , whose columns belong to W , which we denote by Φ^W . Since Φ_2 in (40) is a submatrix of Φ , we only need to consider recovery conditions for the case $a \neq 0$. For $a = 0$, under these conditions, we do not achieve perfect recovery due to the presence of noise in (40).

Theorem 1. *Let Φ be defined in (38), with \mathbf{A} of size $M \times N$ ($M \leq N$) such that $\text{spark}(\mathbf{A}) = M + 1$. The $M^2 \times 4K$ matrix Φ^W , whose columns belong to W defined in (41) with $\Sigma_k = \Sigma_{2K}$ from (39), is full column rank if $M > \frac{8}{5}K$.*

We note that if $K \geq 2$, namely there is at least one transmission, then $M > \frac{8}{5}K$ implies $M^2 \geq 4K$ and Φ^W is a tall matrix.

Proof: Recall that the matrix Φ is expressed as

$$\Phi = \begin{bmatrix} \bar{\mathbf{a}}_1 \otimes \mathbf{a}_1 & \bar{\mathbf{a}}_2 \otimes \mathbf{a}_2 & \dots & \bar{\mathbf{a}}_N \otimes \mathbf{a}_N & \dots \\ \dots & \bar{\mathbf{a}}_1 \otimes \mathbf{a}_2 & \bar{\mathbf{a}}_2 \otimes \mathbf{a}_3 & \dots & \bar{\mathbf{a}}_{N-1} \otimes \mathbf{a}_N \\ \dots & \bar{\mathbf{a}}_2 \otimes \mathbf{a}_1 & \bar{\mathbf{a}}_3 \otimes \mathbf{a}_2 & \dots & \bar{\mathbf{a}}_N \otimes \mathbf{a}_{N-1} \\ \dots & \bar{\mathbf{a}}_1 \otimes \mathbf{a}_N & \bar{\mathbf{a}}_2 \otimes \mathbf{a}_{N-1} & \dots & \bar{\mathbf{a}}_N \otimes \mathbf{a}_1 & \dots \\ \dots & \bar{\mathbf{a}}_1 \otimes \mathbf{a}_{N-1} & \bar{\mathbf{a}}_2 \otimes \mathbf{a}_{N-2} & \dots & \bar{\mathbf{a}}_{N-1} \otimes \mathbf{a}_1 \\ \dots & \bar{\mathbf{a}}_2 \otimes \mathbf{a}_N & \bar{\mathbf{a}}_3 \otimes \mathbf{a}_{N-1} & \dots & \bar{\mathbf{a}}_N \otimes \mathbf{a}_2 \end{bmatrix}, \quad (42)$$

where \mathbf{a}_i and $\bar{\mathbf{a}}_i$ denote the i th column of \mathbf{A} and its conjugate, respectively. Assume by contradiction that the columns of Φ^W are linearly dependent. Then, there exist $\beta_1, \dots, \beta_{4K}$, not all zeros, such that

$$\sum_{j=1}^{4K} \beta_j \phi_j^W = \mathbf{0}, \quad (43)$$

where ϕ_j^W denotes the j th column of Φ^W and is of the form

$$\phi_j^W = \bar{\mathbf{a}}_{[j]} \otimes \mathbf{a}_{g([j])} \text{ or } \bar{\mathbf{a}}_{[j]} \otimes \mathbf{a}_{g([N+1-j])}. \quad (44)$$

Here $\mathbf{a}_{[j]}$ denotes the column of \mathbf{A} that corresponds to the j th selected column of Φ and $g(\cdot)$ is defined in (36). Denote by k_0 the number of indices $[j]$, for $1 \leq j \leq K$, that appear twice in Φ^W . Obviously, $0 \leq k_0 \leq K$ and k_0 is even.

From 1)-4) in the definition of Σ_{2K} , we can express Φ^W as

$$\Phi^W = \begin{bmatrix} \bar{\mathbf{a}}_{[1]} \otimes \mathbf{a}_{g([1])} & \dots & \bar{\mathbf{a}}_{[K]} \otimes \mathbf{a}_{g([K])} & \dots \\ \dots & \bar{\mathbf{a}}_{[1]} \otimes \mathbf{a}_{h([1])} & \dots & \bar{\mathbf{a}}_{[k_0/2]} \otimes \mathbf{a}_{h([k_0/2])} & \dots \\ \dots & \bar{\mathbf{a}}_{[K-k_0/2]} \otimes \mathbf{a}_{h([K-k_0/2])} & \dots & \bar{\mathbf{a}}_{[K]} \otimes \mathbf{a}_{h([K])} & \dots \\ \dots & \bar{\mathbf{a}}_{[K+1]} \otimes \mathbf{a}_{g([K+1])} & \dots & \bar{\mathbf{a}}_{[2K-k_0]} \otimes \mathbf{a}_{g([2K-k_0])} & \dots \\ \dots & \bar{\mathbf{a}}_{[1]} \otimes \mathbf{a}_{g([K])} & \dots & \bar{\mathbf{a}}_{[K]} \otimes \mathbf{a}_{g([1])} & \dots \\ \dots & \bar{\mathbf{a}}_{[1]} \otimes \mathbf{a}_{h([K])} & \dots & \bar{\mathbf{a}}_{[k_0/2]} \otimes \mathbf{a}_{h([K-k_0/2])} & \dots \\ \dots & \bar{\mathbf{a}}_{[k_0/2]} \otimes \mathbf{a}_{h([K-k_0/2])} & \dots & \bar{\mathbf{a}}_{[K]} \otimes \mathbf{a}_{h([1])} & \dots \\ \dots & \bar{\mathbf{a}}_{[K+1]} \otimes \mathbf{a}_{[2K-k_0]} & \dots & \bar{\mathbf{a}}_{g([2K-k_0])} \otimes \mathbf{a}_{g([K+1])} \end{bmatrix}. \quad (45)$$

Here, $h(i)$ is defined as $g(i)$ and we can have either $h(i) = g(i)$ or $h(i) \neq g(i)$. By rearranging the columns of Φ^W with respect to the left entry index of each Kronecker product, the system of equations (43) can be written as

$$\mathbf{C}\bar{\mathbf{A}}^T = \mathbf{0}, \quad (46)$$

where the $M \times (2K - k_0)$ matrix \mathbf{C} is defined by

$$\mathbf{C} = [\beta_1 \mathbf{a}_{g([1])} + \beta_{K+1} \mathbf{a}_{h([1])} + \beta_{2K+1} \mathbf{a}_{g([K])} + \beta_{3K+1} \mathbf{a}_{h([K])} \\ \dots \beta_K \mathbf{a}_{g([K])} + \beta_{K+k_0} \mathbf{a}_{h([K])} + \beta_{3K} \mathbf{a}_{g([1])} + \beta_{3K+k_0} \mathbf{a}_{h([1])} \\ \dots \beta_{K+k_0+1} \mathbf{a}_{g([K+1])} + \beta_{3K+k_0+1} \mathbf{a}_{g([2K-k_0])} \\ \dots \beta_{2K} \mathbf{a}_{g([2K-k_0])} + \beta_{4K} \mathbf{a}_{g([K+1])}]. \quad (47)$$

We next show that $\text{rank}(\mathbf{C}\bar{\mathbf{A}}^T) > 0$ in order to contradict the assumption that there exist $\beta_1, \dots, \beta_{4K}$, not all zeros, such that (43) holds. Denote by n_0 the number of quadruplets $(\beta_j, \beta_{K+j}, \beta_{2K+j}, \beta_{3K+j})$, for $1 \leq j \leq K$, or pairs $(\beta_{k_0+j}, \beta_{2K+k_0+j})$, for $K+1 \leq j \leq 2K-k_0$, with one non zero element at least. Obviously, $1 \leq n_0 \leq 2K-k_0$. Let $\tilde{\mathbf{C}}$ be the $M \times n_0$ matrix composed of the n_0 columns of \mathbf{C} corresponding to these non zero quadruplets/pairs and let $\tilde{\mathbf{A}}^T$ be constructed out of the corresponding n_0 rows of $\bar{\mathbf{A}}^T$. Then, from the Sylvester rank inequality,

$$\text{rank}(\mathbf{C}\bar{\mathbf{A}}^T) = \text{rank}(\tilde{\mathbf{C}}\tilde{\mathbf{A}}^T) \geq \text{rank}(\tilde{\mathbf{C}}) + \text{rank}(\tilde{\mathbf{A}}^T) - n_0. \quad (48)$$

In addition,

$$\text{rank}(\tilde{\mathbf{A}}^T) \geq \min(n_0, \text{spark}(\mathbf{A}) - 1), \quad (49)$$

and

$$\text{rank}(\tilde{\mathbf{C}}) \geq \min\left(\left\lceil \frac{n_0}{4} \right\rceil, \frac{\text{spark}(\mathbf{A}) - 1}{4}\right). \quad (50)$$

The last inequality follows from the fact that any linear combination of n columns of \mathbf{C} is a linear combination of at least $\lceil n/4 \rceil$ distinct columns of \mathbf{A} . Therefore,

$$\text{rank}(\mathbf{C}\bar{\mathbf{A}}^T) \geq \min\left(\left\lceil \frac{n_0}{4} \right\rceil, \frac{5}{4}M - n_0\right). \quad (51)$$

Since $M > \frac{8}{5}K$, it holds that $\text{rank}(\mathbf{C}\bar{\mathbf{A}}^T) \geq 1$, for all $1 \leq n_0 \leq 2K - k_0$, $0 \leq k_0 \leq K$, contradicting (46). \square

Corollary 1, which directly follows from Theorem 1, provides sufficient conditions for perfect recovery of $\mathbf{r}_x^a(\tilde{f})$.

Corollary 1. *If*

- 1) $\text{spark}(\mathbf{A}) = M + 1$,
- 2) $M > \frac{8}{5}K$,

then the system (37) has a unique solution in Σ_{2K} .

Proof: Conditions 1)-2) ensure that all sets of $2K$ columns belonging to W are linearly independent, from Theorem 1. Thus, the proof follows from Lemma 1. \square

Under the conditions of Corollary 1, the cyclic spectrum $S_x(f)$ can be perfectly recovered for cyclic frequencies α which are not multiples of f_s . For α that are multiples of f_s , the recovery is performed in the presence of bounded noise, yielding a bounded reconstruction error. For detection purposes, this has proven satisfactory in the simulations.

Without any sparsity assumption, we can repeat the proof of Theorem 1 with $K = N$ and $k_0 = N$, leading to $1 \leq n_0 \leq N$. We thus obtain that, if $M > \frac{4}{5}N$, then we can perfectly recover the cyclic spectrum of $x(t)$. The minimal sampling rate is then

$$f_{\min_0} = Mf_s = \frac{4}{5}Nf_s = \frac{4}{5}f_{\text{Nyq}}. \quad (52)$$

This means that even without any sparsity constraints on the signal, we can retrieve its cyclic spectrum from samples below the Nyquist rate, by exploiting its cyclostationary properties. A similar result was already observed in [18] in the context of power spectrum reconstruction of wide-sense stationary signals in noiseless settings. In that case, the power spectrum slices appear only on the diagonal of the matrix $\mathbf{R}_x^0(\tilde{f})$ and it follows that $\Phi = \bar{\mathbf{A}} \odot \mathbf{A}$. Power spectrum recovery is therefore

a special case of cyclic spectrum reconstruction, treated here. There, it was shown that the power spectrum can be retrieved at half the Nyquist rate without any sparsity constraints. Here, we extend this result to cyclic spectrum reconstruction, which requires a higher rate.

If $x(t)$ is assumed to be sparse in the frequency domain, with $K = 2N_{\text{sig}} \ll N$, then the minimal sampling rate for perfect reconstruction of its cyclic spectrum is

$$f_{\min} = Mf_s = \frac{16}{5}N_{\text{sig}}B = \frac{8}{5}f_{\text{Landau}}. \quad (53)$$

It was shown in [18], that the power spectrum of a stationary sparse signal can be perfectly recovered at its Landau rate. Again, the minimal sampling rate for cyclic spectrum recovery is slightly higher than that required for power spectrum reconstruction.

C. Cyclic Spectrum Recovery

So far, we only discussed conditions for perfect recovery of the cyclic spectrum, namely for (37) and (40) to have unique solutions. We now provide an algorithm for cyclic spectrum reconstruction. To account for the structure of $\mathbf{r}_x^a(\tilde{f})$ for $a \neq 0$, we extend orthogonal matching pursuit (OMP) [15], [31]. In each iteration, we add an internal loop that, for a selected element originally from the diagonals of $\mathbf{R}_x^a(\tilde{f})$, checks for a corresponding non-zero element from the anti-diagonals, and vice versa, as defined by the set \mathcal{I} . For $a = 0$, we use the standard OMP [15], [31]. We note that for all $a \geq 0$ we can exploit the additional symmetric structure of $\mathbf{r}_x^a(\tilde{f})$ as defined by Property 1 of \mathcal{I} . Our structured OMP method (assuming that the columns of Φ are normalized) is formally defined by Algorithm 1.

In Algorithm 1, $\mathbf{v} = \mathbf{r}_x^a(\tilde{f})$, Λ_i denotes the set of complementary indices with respect to i according to Property 2 of \mathcal{I} , namely for $2 \leq i \leq N-1$,

$$\Lambda_i = \begin{cases} \{6d_i + 5, 6d_i + 6, 6d_i + 7\} & \text{if } 0 < \text{mod}(i-4, 6) \leq 3 \\ \{6d_i + 8, 6d_i + 9, 6d_i + 10\} & \text{else,} \end{cases} \quad (54)$$

where $d_i = \lfloor \frac{i-4}{6} \rfloor$. For $i = 1$ and $i = N$, Λ_i is similarly defined according to Properties 3 and 4, respectively. The vector \mathbf{w}^S is the reduction of \mathbf{w} to the support S , Φ_S contains the corresponding columns of Φ , S^c denotes the complementary set of S and \dagger is the Moore-Penrose pseudo-inverse. The halting criterion in Algorithm 1, as for standard OMP, can be sparsity-based if the true sparsity is known, or at least an upper bound for it, or residual-based.

Similarly to [12], the set (37) consists of an infinite number of linear systems since \tilde{f} is a continuous variable. Since the support S is common to $\mathbf{r}_x^a(\tilde{f})$ for all $\tilde{f} \in \mathcal{F}_s$, we propose to recover it jointly instead of solving (37) for each \tilde{f} individually, thus increasing efficiency and robustness to noise. To that end, we use the support recovery paradigm from [12] that produces a finite system of equations, called multiple measurement vectors (MMV) from an infinite number of linear systems. This reduction is performed by what is referred to as the continuous to finite (CTF) block. The cyclic spectrum reconstruction of both sparse and non sparse signals can then

Algorithm 1 Structured OMP

Input: observation vector \mathbf{v} of size m , measurement matrix Φ of size $m \times n$, threshold $\epsilon > 0$

Output: index set S containing the locations of the non zero indices of \mathbf{u} , estimate for sparse vector $\hat{\mathbf{u}}$

Initialize: residual $\mathbf{r}_0 = \mathbf{v}$, index set $S_0 = \emptyset$, possible index set $\Gamma_0 = \{1, \dots, n\}$, estimate $\hat{\mathbf{u}} = \mathbf{0}$, $\ell = 0$

while halting criterion false **do**

$\ell \leftarrow \ell + 1$

$\mathbf{b} \leftarrow \Phi^* \mathbf{r}_\ell$

$S_\ell \leftarrow S_\ell \cup \arg \max_{i \in \Gamma_\ell} \mathbf{b}_i$

$\Gamma_\ell \leftarrow \Gamma_\ell \setminus \Lambda_i$

$(\hat{\mathbf{u}}_\ell)^{S_\ell} \leftarrow \Phi_{S_\ell}^\dagger \mathbf{v}$, $(\hat{\mathbf{u}}_\ell)^{S_\ell^c} \leftarrow \mathbf{0}$

$\delta_0 \leftarrow \|\mathbf{v} - \Phi \hat{\mathbf{u}}_\ell\|^2$

for $j \in \Lambda_i$ **do**

$\mathbf{w}_j^{S_\ell} \leftarrow \Phi_{S_\ell \cup j}^\dagger \mathbf{v}$, $\mathbf{w}_j^{S_\ell^c} \leftarrow \mathbf{0}$

$\delta_j \leftarrow \|\mathbf{v} - \Phi \mathbf{w}_j\|^2$

end for

if $\delta_0 - \min_{j \in \Lambda_i} \delta_j > \epsilon$ **then**

$S_\ell \leftarrow S_\ell \cup \arg \min_{j \in \Lambda_i} \delta_j$

$(\hat{\mathbf{u}}_\ell)^{S_\ell} \leftarrow \Phi_{S_\ell}^\dagger \mathbf{v}$, $(\hat{\mathbf{u}}_\ell)^{S_\ell^c} \leftarrow \mathbf{0}$

end if

$\mathbf{r} \leftarrow \mathbf{v} - \Phi \hat{\mathbf{u}}_\ell$

end while

return S_ℓ and $\hat{\mathbf{u}}_\ell$

be divided into two stages: support recovery, performed by the CTF, and cyclic spectrum recovery. From (37), for $a \in]0, f_s]$, we have

$$\mathbf{Q}^a = \Phi \mathbf{Z}^a \Phi^H \quad (55)$$

where

$$\mathbf{Q}^a = \int_{\tilde{f} \in \mathcal{F}_s} \mathbf{r}_z^a(\tilde{f}) \mathbf{r}_z^{aH}(\tilde{f}) d\tilde{f} \quad (56)$$

is an $M \times M$ matrix and

$$\mathbf{Z}^a = \int_{\tilde{f} \in \mathcal{F}_s} \mathbf{r}_x^a(\tilde{f}) \mathbf{r}_x^{aH}(\tilde{f}) d\tilde{f} \quad (57)$$

is an $N \times N$ matrix. Then, any matrix \mathbf{V}^a for which $\mathbf{Q}^a = \mathbf{V}^a (\mathbf{V}^a)^H$ is a frame for $\mathbf{r}_z^a(\mathcal{F}_s) = \{\mathbf{r}_z^a(\tilde{f}) | \tilde{f} \in \mathcal{F}_s\}$ [12], [40]. Clearly, there are many possible ways to select \mathbf{V}^a . We construct it by performing an eigendecomposition of \mathbf{Q}^a and choosing \mathbf{V}^a as the matrix of eigenvectors corresponding to the non zero eigenvalues. We then define the following linear system

$$\mathbf{V}^a = \Phi \mathbf{U}^a. \quad (58)$$

For $a = 0$, identical derivations can be carried out by replacing Φ by Φ_2 . From [12] (Propositions 2-3), the support of the unique sparsest solution of (58) is the same as the support of $\mathbf{r}_x^a(\tilde{f})$ in our original set of equations (37) (or (40)). For simplicity, Algorithm 1 presents the single measurement vector (SMV) version of the recovery algorithm, which can be adapted to the MMV settings, similarly to the simultaneous OMP [15], [31], to solve (58).

As discussed above, $\mathbf{r}_x^a(\tilde{f})$ is $2K$ -sparse for each specific $\tilde{f} \in [0, f_s - a]$, for all $a \in [0, f_s]$. However, after combining

the frequencies, the matrix \mathbf{U}^a is $4K$ -sparse (at most), since the spectrum of each transmission may split between two slices. Therefore, the above algorithm, referred to as SBR4 in [12] (for signal reconstruction as opposed to cyclic spectrum reconstruction), requires a minimal sampling rate of $2f_{\min}$ for sparse signals or $2f_{\min_0}$ for non sparse signals. In order to achieve the minimal rate f_{\min} or f_{\min_0} , the SBR2 algorithm regains the factor of two in the sampling rate at the expense of increased complexity [12]. In a nutshell, SBR2 is a recursive algorithm that alternates between the CTF described above and a bi-section process. The bi-section splits the original frequency interval into two equal width intervals on which the CTF is applied, until the level of sparsity of \mathbf{U}^a is less or equal to $2K$. We refer the reader to [12] for more details.

Once the support S is known, perfect reconstruction of the cyclic spectrum is obtained by

$$\begin{aligned} (\hat{\mathbf{r}}_x^a)^S(\tilde{f}) &= \Phi_S^\dagger \mathbf{r}_z^a(\tilde{f}) \\ \hat{\mathbf{r}}_{x_i}^a(\tilde{f}) &= 0 \quad \forall i \notin S, \end{aligned} \quad (59)$$

for all $a \in]0, f_s]$ and

$$\begin{aligned} (\hat{\mathbf{r}}_x^0)^S(\tilde{f}) &= (\Phi_2)_S^\dagger \mathbf{r}_z^0(\tilde{f}) \\ \hat{\mathbf{r}}_{x_i}^0(\tilde{f}) &= 0 \quad \forall i \notin S. \end{aligned} \quad (60)$$

The cyclic spectrum $S_x^\alpha(f)$ is then assembled using (29) for (f, α) defined in (30).

D. Carrier Frequency and Bandwidth Estimation

Once the cyclic spectrum $S_x^\alpha(f)$ is reconstructed from the sub-Nyquist samples, we apply our carrier frequency and bandwidth estimation algorithm from [35]. Our approach is a simple parameter extraction method from the cyclic spectrum of multiband signals. It allows the estimation of several carriers and several bandwidths simultaneously, as well as that of the number of transmissions N_{sig} . The proposed algorithm consists of the following five steps: preprocessing, thresholding, clustering, parameter estimation, corrections. Here, we briefly describe the algorithm steps. The reader is referred to [35] for more details.

The preprocessing aims to compensate for the presence of stationary noise in the cyclic spectrum at the cyclic frequency $\alpha = 0$, by attenuating the cyclic spectrum energy around this frequency. Thresholding is then applied to the resulting cyclic spectrum in order to find its peaks. For each cyclic frequency α , we retain the value of the cyclic spectrum at $f = 0$. The locations and values of the selected peaks are then clustered to find the corresponding cyclic feature. Before separating the clusters, we start by estimating their number using the elbow method, which can be traced to speculation by Thorndike [41]. Clustering is then performed using the k-means method. At the end of the process, each cluster represents a cyclic feature. It follows that, apart from the cluster present at DC which we remove, the number of signals N_{sig} is equal to half the number of clusters. Next, we estimate the carrier frequency f_i and bandwidth B_i of each transmission. The carrier frequency yields the highest correlation [6] and thus the highest peak, at the cyclic frequency equal to twice its value, namely $\alpha = 2f_i$. It is therefore estimated as half the cyclic frequency of the

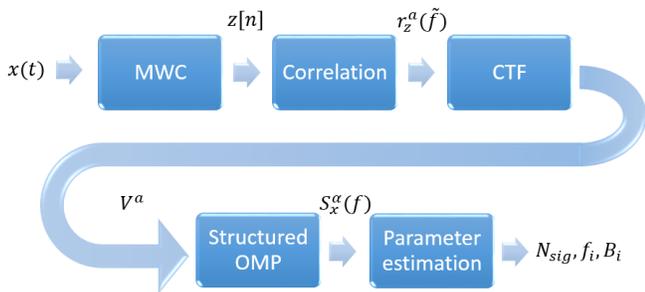


Fig. 6. Processing flow diagram. The input signal $x(t)$ is first fed to the MWC analog front-end which generates the low rate samples $z[n]$ shown in (23) in the frequency domain. The correlations $\mathbf{r}_z^a(\tilde{f})$ between frequency shifted versions of the samples are then computed by vectorizing (28). The CTF next produces a finite set of equations, which are solved using Algorithm 1, exploiting the known structure of the sparse vectors that compose the cyclic spectrum $S_x^\alpha(f)$. The number of transmissions N_{sig} , their respective carrier frequencies f_i and bandwidths B_i are finally estimated from $S_x^\alpha(f)$. The entire processing is performed at a low rate.

highest peak within the clusters belonging to the same signal. The bandwidth is found by locating the edge of the support of the angular frequencies.

The processing flow of our low rate sampling and cyclic spectrum recovery algorithm is summarized in Fig. 6.

V. SIMULATION RESULTS

We now demonstrate via simulations cyclic spectrum reconstruction from sub-Nyquist samples and investigate the performance of our carrier frequency and bandwidth estimation algorithm. We compare our approach to energy detection and investigate the impact of the sampling rate on the detection performance. Throughout the simulations we use the MWC analog front-end [33] for the sampling stage.

A. Preliminaries

We begin by explaining how we estimate the elements of $\mathbf{R}_z^a(\tilde{f})$ in (28). The overall sensing time is divided into P time windows of length Q samples. We first compute the estimates of $\mathbf{z}_i(\tilde{f})$, $1 \leq i \leq M$ using the fast Fourier transform (FFT) on the samples $\mathbf{z}_i[n]$ over a finite time window. We then estimate the elements of $\mathbf{R}_z^a(\tilde{f})$ as

$$\hat{\mathbf{R}}_z^a(\tilde{f})_{(i,j)} = \frac{1}{P} \sum_{p=1}^P \hat{\mathbf{z}}_i^p(\tilde{f}) \hat{\mathbf{z}}_j^p(\tilde{f} + a), \quad (61)$$

for $a \in [0, f_s]$ and $f \in [0, f_s - a]$. Here, $\hat{\mathbf{z}}_i^p(\tilde{f})$ is the estimate of $\mathbf{z}_i(\tilde{f})$ from the p th time window.

The cyclic spectrum recovery processing is presented here in the frequency domain. The reconstruction can be equivalently performed in the time domain by modulating the slices to replace the frequency shift $\tilde{f} + a$. Then, $\hat{\mathbf{r}}_x^a[n]$ can be recovered using the time equivalent of (59)-(60). However, the carrier frequencies f_i and bandwidths B_i estimation is performed on the cyclic spectrum, in the frequency domain. Thus, the Fourier transform of $\hat{\mathbf{r}}_x^a[n]$ needs to be computed, and $S_x^\alpha(f)$ is then mapped using (29) for (f, α) defined in (30). Therefore,

we choose to perform the entire processing in the frequency domain. Another reason to do so is that SBR2 can obviously be performed in the frequency domain only, as opposed to SBR4 which can be carried out both in time and frequency.

We note that in theory, our approach does not require any discretization, neither in the angular frequency f nor in the cyclic frequency α . Indeed, $\mathbf{R}_x^a(\tilde{f})$ can be computed for any $a \in [0, f_s]$ and $\tilde{f} \in [0, f_s - a]$. This distinguishes our scheme from those based on a transformation between Nyquist and sub-Nyquist samples, where the resolution is theoretically inherent to the problem dimension and dictated by the length of the Nyquist samples vector. In practice, the resolution both in f and α obviously depends on the sensing time and is determined by the number of samples, namely the number of discrete Fourier transform (DFT) coefficients of $\hat{\mathbf{z}}(\tilde{f})$.

We compare our cyclostationary detection to energy detection based on power spectrum recovery, presented in [18]. There, it was shown that power spectrum sensing outperforms spectrum sensing, namely energy detection performed on the recovered signal itself. This power spectrum reconstruction approach is a special case of ours for $a = 0$, when only the matrix $\mathbf{R}_x^0(\tilde{f})$ is considered and only its diagonal is reconstructed. Therefore, we compare our detection approach performed on $S_x^\alpha(f)$ for $\alpha \neq 0$ to energy detection carried out on $S_x^\alpha(f)$, for $\alpha = 0$, corresponding to the diagonal elements of $\mathbf{R}_x^0(\tilde{f})$.

Throughout the simulations we consider additive white Gaussian noise (AWGN) $n(t)$. The SNR is defined as the ratio between the power of the wideband signal and that of the wideband noise as follows

$$\text{SNR} = \frac{\sum_{i=1}^{N_{sig}} \|s_i(t)\|^2}{\|n(t)\|^2}. \quad (62)$$

B. Cyclic Spectrum Recovery

We first illustrate cyclic spectrum reconstruction from sub-Nyquist samples. We consider $x(t)$ composed of $N_{sig} = 3$ AM transmissions. Each transmission has bandwidth $B = 80\text{MHz}$ and the carrier frequencies are drawn uniformly at random in $[0, \frac{f_{Nyq}}{2}]$, with $f_{Nyq} = 6.4\text{GHz}$. The SNR is set to -5dB . In the sampling stage, we use the MWC with $M = 11$ channels, each sampling at $f_s = 95\text{MHz}$. The overall sampling rate is therefore 1.05GHz , that is 2.2 times the Landau rate and 16% of the Nyquist rate. Here, the theoretical minimal sampling rate is $f_{min} = 768\text{MHz}$. Figure 7 shows the original and reconstructed power spectrum of $x(t)$, namely $S_x^0(f)$. In this experiment, the carriers are $f_1 = 97\text{MHz}$, $f_2 = 573\text{MHz}$ and $f_3 = 1.4\text{GHz}$. We observe that the recovery of the power spectrum failed; some occupied bands were not reconstructed and others that only contained noise were identified as active. This is due to the poor performance of energy detection in low SNR regimes. In can be seen in Fig. 7 that the signal spectrum is not buried in noise and the transmissions would have been perfectly detected using energy detection on Nyquist samples. However, in sub-Nyquist regimes, the aliasing decreases the SNR [19], as can be seen in Fig. 8, which shows the DFT of the samples of one channel. As a consequence, energy detection failed in this sub-Nyquist regime.

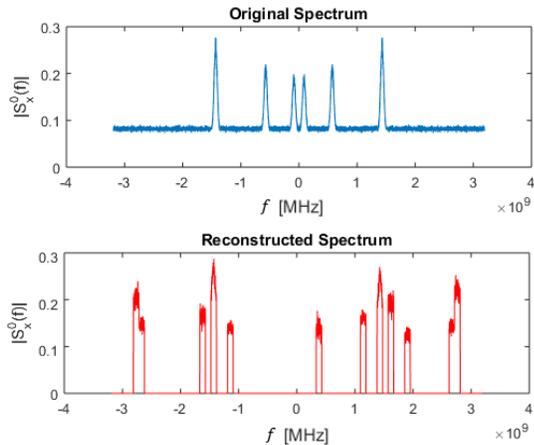


Fig. 7. Original (top) and reconstructed (bottom) power spectrum ($\alpha = 0$).

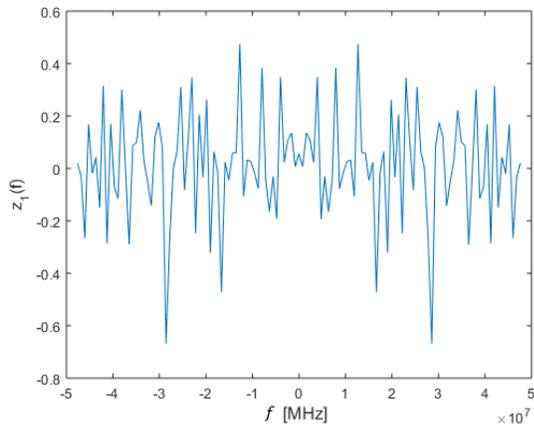


Fig. 8. Sub-Nyquist samples in the first channel in the frequency domain.

The reconstructed cyclic spectrum of $x(t)$ is presented in Fig. 9 (the reader is referred to the colored version for a clearer figure), where we observe that the noise contribution is concentrated at $\alpha = 0$ while it is significantly lower at the non zero cyclic frequencies. For clarity, we focus on the one-dimensional section of $S_x^\alpha(f)$ for $f = 0$, shown in Fig. 10. It can be clearly seen that the highest peaks (at least 6dB above the lower peaks) are located at $\alpha_i = 2f_i$ for all 3 active transmissions. This illustrates the advantage of cyclostationary detection in comparison with energy detection.

Next, we perform cyclostationary detection on the reconstructed cyclic spectrum. We compare the performance of cyclostationary and energy detection performed on the reconstructed cyclic and power spectrum, respectively. For cyclostationary detection, we use a single-cycle detector which computes the energy at several frequencies around $f = 0$ and at a single cyclic frequency α . In the simulations, we consider AM modulated signals. We address a blind scenario where the carrier frequencies of the signals occupying the wideband channel are unknown and we have $N_{\text{sig}} = 3$ potentially active transmissions, with single-sided bandwidth $B = 100\text{MHz}$. For each iteration, the alternative and null hypotheses define the

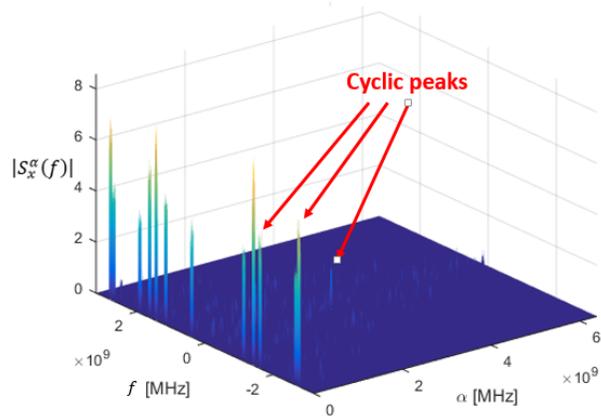


Fig. 9. Reconstructed cyclic spectrum.

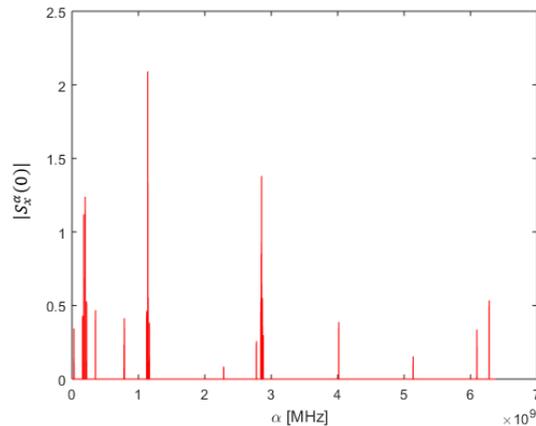


Fig. 10. Reconstructed cyclic spectrum for $f = 0$, $S_x^\alpha(0)$, as a function of the cyclic frequency α .

presence or absence of one out of the N_{sig} transmissions. We refer to that transmission as the signal of interest. The Nyquist rate of $x(t)$ is $f_{\text{Nyq}} = 10\text{GHz}$. We consider $N = 64$ spectral bands and $M = 7$ analog channels, each sampling at $f_s = 156\text{MHz}$. The overall sampling rate is $Mf_s = 1.09\text{GHz}$ which is 182% of the Landau rate and 10.9% of the Nyquist rate. Here, the theoretical minimal sampling rate is $f_{\text{min}} = 960\text{MHz}$. The receiver operating characteristic (ROC) curve is shown in Fig. 11 for different SNR regimes (the averages were performed over $P = 15$ time windows). Detection occurs if the presence of the signal of interest is correctly detected while false alarm is declared if a detection is claimed while the signal of interest is absent. It can be seen that cyclostationary detection outperforms energy detection in low SNR regimes, as expected. This is already known in the Nyquist regime and is now shown on samples obtained at a sub-Nyquist rate.

C. Carrier Frequencies and Bandwidths Recovery

We now demonstrate carrier frequency and bandwidth estimation from sub-Nyquist samples. We first illustrate our algorithm process on a specific experiment. We consider $x(t)$

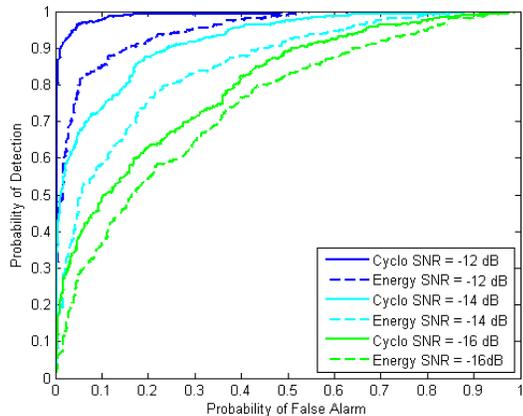


Fig. 11. ROC for both energy and cyclostationary detection.

composed of $N_{\text{sig}} = 3$ BPSK transmissions, which have cyclic features at the locations $(f, \alpha) = (0, \pm 2f_i), (\pm f_i, \pm \frac{1}{T_i})$, where f_i is the carrier frequency and T_i is the symbol period of the i th transmission [6]. Each transmission has bandwidth $B_i = 18\text{MHz}$ and symbol rate $T_i = 1/B_i = 0.56\mu\text{s}$, and the carrier frequencies are drawn uniformly at random in $[0, \frac{f_{\text{Nyq}}}{2}]$, with $f_{\text{Nyq}} = 1\text{GHz}$. In this experiment, the selected carriers are $f_1 = 163.18\text{MHz}$, $f_2 = 209.69\text{MHz}$ and $f_3 = 396.12\text{MHz}$. The SNR is set to -5dB . In the sampling stage, we use the MWC with $M = 9$ channels, each sampling at $f_s = 23.26\text{MHz}$. The overall sampling rate is therefore 210MHz , that is a little below twice the Landau rate and 21% of the Nyquist rate. Here, the theoretical minimal sampling rate is $f_{\text{min}} = 172.8\text{MHz}$.

Figure 12 presents the original and reconstructed power spectrum using $P = 100$ time windows. We observe that the signal's spectrum was not perfectly recovered due to the noise. The reconstructed cyclic spectrum, including the power spectrum, estimated over $P = 100$ time windows as well, is shown in Fig. 13 and the section corresponding to $f = 0$ can be seen in Fig. 14. The cyclic peaks at the locations $(f, \alpha) = (0, \pm 2f_i)$, for $i = 1, 2, 3$ can be observed in both figures. In Fig. 15, we illustrate the clustering stage of our algorithm as a function of α for $f = 0$. The estimated number of clusters is 6, yielding a correctly estimated number of signals $\hat{N}_{\text{sig}} = 3$. The estimated carrier frequencies using cyclostationary based estimation are $\hat{f}_1 = 162.66\text{MHz}$, $\hat{f}_2 = 209.19\text{MHz}$ and $\hat{f}_3 = 395.11\text{MHz}$, and the corresponding estimated bandwidths are $\hat{B}_1 = 17.4\text{MHz}$, $\hat{B}_2 = 17.4\text{MHz}$ and $\hat{B}_3 = 17.0\text{MHz}$. Using energy based estimation, we obtain $\hat{N}_{\text{sig}} = 5$ signals, with estimated carrier frequencies $\hat{f}_1 = 93.04\text{MHz}$, $\hat{f}_2 = 162.82\text{MHz}$, $\hat{f}_3 = 255.86\text{MHz}$ and $\hat{f}_4 = 383.89\text{MHz}$, $\hat{f}_5 = 465.21\text{MHz}$ and estimated bandwidths $\hat{B}_1 = \hat{B}_2 = \hat{B}_3 = \hat{B}_5 = 23.1\text{MHz}$, $\hat{B}_4 = 46.3\text{MHz}$. Clearly, cyclostationary detection succeeded where energy detection failed.

We now investigate the performance of our carrier frequency and bandwidth estimation algorithm from sub-Nyquist samples with respect to SNR and compare it to energy detection. We

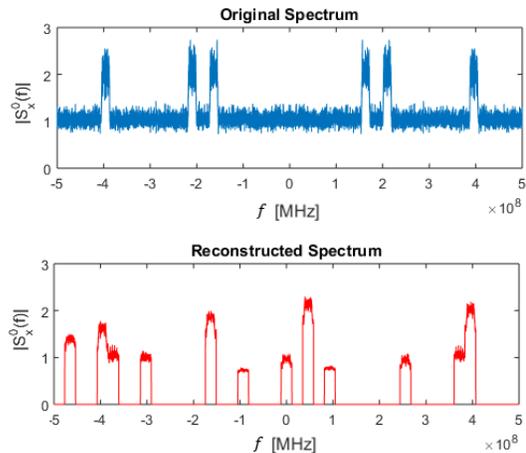


Fig. 12. Original and reconstructed power spectrum.

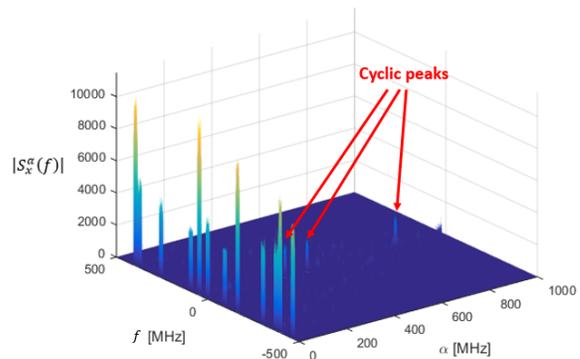


Fig. 13. Reconstructed cyclic spectrum.

consider $x(t)$ composed of $N_{\text{sig}} = 3$ BPSK transmissions with identical parameters as in the previous section. The sampling parameters remain the same as well. In each experiment, we draw the carrier frequencies uniformly at random and generate the transmissions. The results are averaged over 1000 realizations.

Figure 16 shows the probability of detection of both cyclostationary (blue) and energy (red) detection. A detection is reported if the distance between the true and recovered carrier frequencies is below 10 times the frequency resolution, which is equal to 0.388MHz . The average number of false alarms, namely unoccupied bands that are labeled as detection, is shown in Fig. 17. Clearly, cyclostationarity outperforms the energy approach in terms of probability of detection. Cyclostationary detection also yields fewer false alarms. For high SNRs, the gap between the performance of both schemes is small, since energy detection still succeeds in these regimes. This gap widens with SNR decrease, where the advantage of cyclostationary detection is clearly marked. The curves for both cyclostationary and energy detection show a rapid decrease of performance below a certain SNR level. We note that this level is lower for cyclostationary detection. This behavior is common to CS based recovery algorithms, which

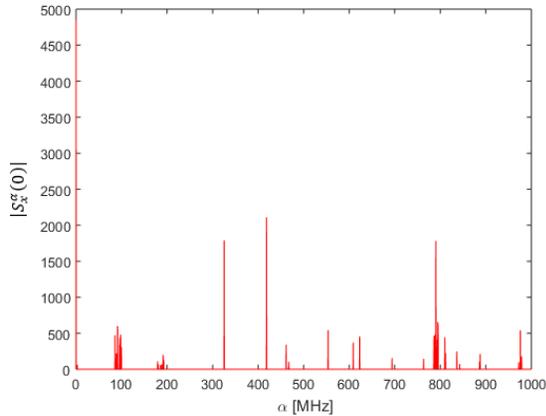


Fig. 14. Reconstructed cyclic spectrum for $f = 0$, $S_x^\alpha(0)$, as a function of the cyclic frequency α .

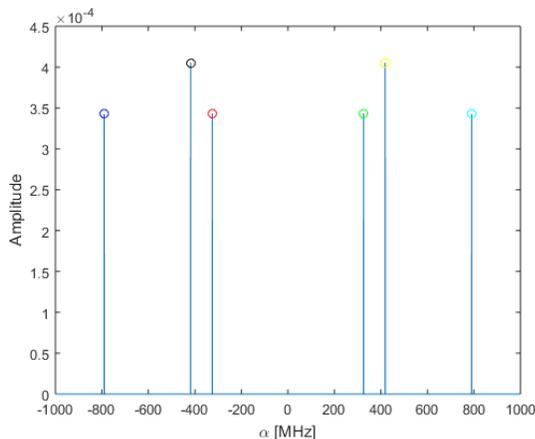


Fig. 15. Clustering with $k = 6$.

fail in the presence of large noise and yield wrong signal support, leading to misdetections and false alarms. When the SNR becomes too low, cyclostationary detection fails as well, due to the finite sensing time and averaging.

Next, we compare cyclic spectrum reconstruction from Nyquist and sub-Nyquist samples. We consider the same parameters as above for the signal generation and the sampling front-end. From Fig. 18, which shows the detection performance in both regimes, it can be seen that the gap between them is not large. The loss in performance due to the reduced number of samples is small since it is compensated by cyclostationary detection, which is robust to noise.

In Fig. 19, we wish to validate the derived theoretical minimal sampling rate. In the settings described above, the lower bound is $f_{\min} = 172.8\text{MHz}$, which corresponds to a minimal number of channels $M_{\min} = 10$ for perfect cyclic spectrum recovery. It can be seen in the figure that beyond 10 channels, the probability of detection is close to 1 in the noiseless regime. Detection errors are due to the finite sensing time and averaging. In the presence of noise, the probability of detection is slightly lower and the number of channels required to reach its maximal value is higher. Below 10 channels,

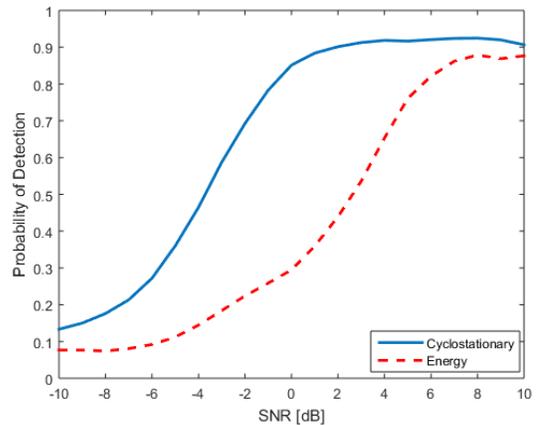


Fig. 16. Probability of detection - cyclostationary vs. energy detection.

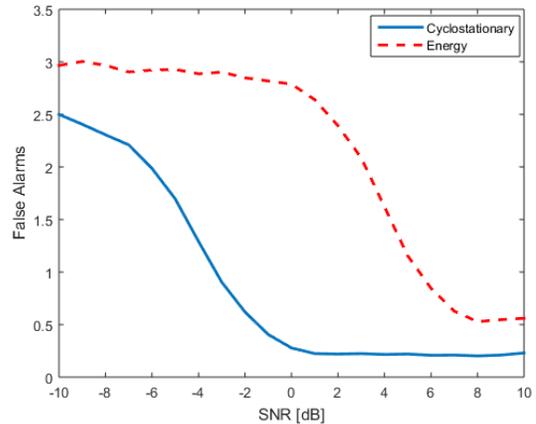


Fig. 17. False alarm - cyclostationary vs. energy detection.

the cyclic spectrum cannot be perfectly recovered and the detection performance decreases with the number of channels.

Finally, we compare the recovery performance of our structured OMP presented in Algorithm 1 with the traditional OMP. Here, we consider $N_{\text{sig}} = 4$ transmissions and $M = 14$ sampling channels. The remaining parameters are identical to those in the previous experiments. The added performance of exploiting the structure of the correlation matrices can be observed in Fig. 20 above a certain SNR value.

VI. CONCLUSION

In this paper, we considered cyclostationary detection in a sub-Nyquist regime, to cope with efficiency and robustness requirements for spectrum sensing in the context of CR. We presented a cyclic spectrum reconstruction algorithm from sub-Nyquist samples along with recovery conditions for both sparse and non sparse signals. We showed that even if the signal is not sparse, its cyclic spectrum can be recovered from samples obtained below the Nyquist rate. The minimal rates obtained for both the sparse and non sparse cases are found to be higher than those required for power spectrum recovery and

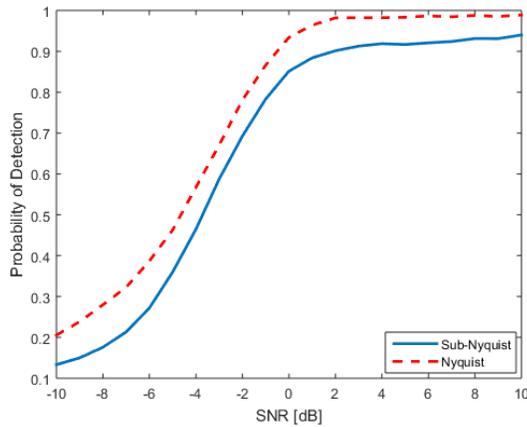


Fig. 18. Probability of detection - sub-Nyquist vs. Nyquist sampling.

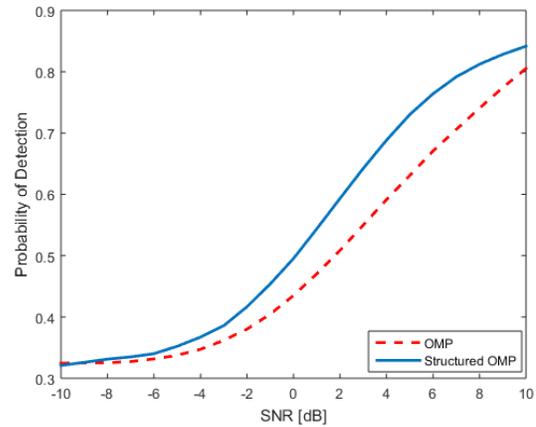


Fig. 20. Probability of detection - structured OMP vs. traditional OMP.

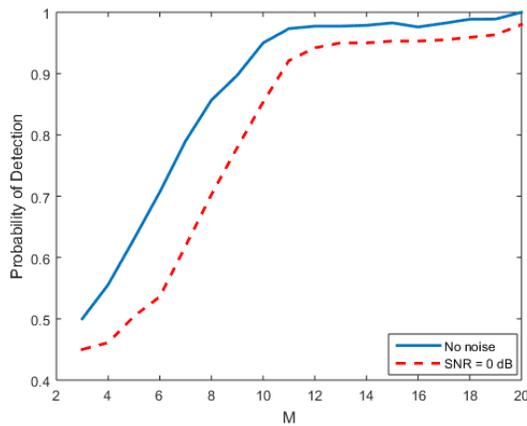


Fig. 19. Probability of detection - sampling rate.

lower than the rates required for signal reconstruction. Once the cyclic spectrum is recovered, we applied our feature extraction algorithm that estimates the number of transmissions and their respective carrier frequency and bandwidth. Simulations performed at low SNRs validate that cyclostationary detection outperforms energy detection in the sub-Nyquist regime, as well as the theoretical lower sampling bound.

REFERENCES

- [1] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, pp. 523–531, Apr. 1967.
- [2] D. O. North, "An analysis of the factors which determine signal/noise discrimination in pulsed carrier systems," *Proc. IEEE*, vol. 51, pp. 1016–1027, Jul. 1963.
- [3] G. L. Turin, "An introduction to matched filters," *IRE Trans. Inf. Theory*, vol. 6, pp. 311–329, Jun. 1960.
- [4] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, pp. 639–697, Apr. 2006.
- [5] A. Napolitano, "Cyclostationarity: New trends and applications," *Signal Process.*, vol. 120, pp. 385–408, Jan. 2016.
- [6] W. Gardner, *Statistical spectral analysis: a non probabilistic theory*. Prentice Hall, 1988.
- [7] J. Mitola, "Software radios: Survey, critical evaluation and future directions," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 8, pp. 25–36, Apr. 1993.
- [8] E. Axell, G. Leus, E. Larsson, and H. Poor, "Spectrum sensing for cognitive radios: State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, pp. 101–116, May 2012.
- [9] FCC, "Spectrum policy task force report: Federal communications commission, tech. rep. 02-135. [online]," http://www.gov/edocs_public/attachmatch/DOC228542A1.pdf, Nov. 2002.
- [10] M. McHenry, "NSF spectrum occupancy measurements project summary. shared spectrum co., tech. rep. [online]," <http://www.sharedspectrum.com>, Aug. 2005.
- [11] R. I. C. Chiang, G. B. Rowe, and K. W. Sowerby, "A quantitative analysis of spectral occupancy measurements for cognitive radio," *IEEE Vehicular Technology Conf.*, pp. 3016–3020, Apr. 2007.
- [12] M. Mishali and Y. C. Eldar, "Blind multi-band signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, pp. 993–1009, Mar. 2009.
- [13] —, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Topics Signal Process.*, vol. 4, pp. 375–391, Apr. 2010.
- [14] —, "Sub-Nyquist sampling: Bridging theory and practice," *IEEE Signal Process. Mag.*, vol. 28, pp. 98–124, Nov. 2011.
- [15] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge University Press, 2015.
- [16] M. A. Lexa, M. E. Davies, J. S. Thompson, and J. Nikolic, "Compressive power spectral density estimation," *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, vol. 57, pp. 22–27, May 2011.
- [17] D. D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. Signal Process.*, vol. 60, pp. 4775–4789, Sept. 2012.
- [18] D. Cohen and Y. C. Eldar, "Sub-Nyquist sampling for power spectrum sensing in cognitive radios: A unified approach," *IEEE Trans. Signal Process.*, vol. 62, pp. 3897–3910, Aug. 2014.
- [19] E. Arias-Castro and Y. C. Eldar, "Noise folding in compressed sensing," *IEEE Signal Process. Lett.*, vol. 18, pp. 478–481, Aug. 2011.
- [20] C. M. Spooner and R. B. Nicholls, *Spectrum sensing based on spectral correlation*, in: B. Fette (Ed.), *Cognitive Radio Technology, 2nd ed. (Chapter 18)*. Elsevier, 2009.
- [21] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proc. IEEE*, vol. 97, pp. 849–877, May 2009.
- [22] E. Rebeiz, V. Jain, and D. Cabric, "Cyclostationarity-based low complexity wideband spectrum sensing using compressive sampling," *IEEE Int. Conf. Commun.*, pp. 1619–1623, Jun. 2012.
- [23] Z. Khalaf, A. Nafkha, and J. Palicot, "Blind cyclostationary feature detector based on sparsity hypotheses for cognitive radio equipment," *IEEE Int. Midwest Symposium Circuits and Syst.*, pp. 1–4, Aug. 2011.
- [24] —, "Blind spectrum detector for cognitive radio using compressive sensing," *IEEE Global Comm. Conf.*, pp. 1–5, Dec. 2011.
- [25] Z. Tian, "Cyclic feature based wideband spectrum sensing using compressive sampling," *IEEE Int. Conf. Commun.*, pp. 1–5, Jun. 2011.
- [26] Z. Tian, Y. Tafesse, and B. M. Sadler, "Cyclic feature detection with sub-Nyquist sampling for wideband spectrum sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 6, pp. 58–69, Feb. 2012.

- [27] S. A. Razavi, M. Valkama, and D. Cabric, "High-resolution cyclic spectrum reconstruction from sub-Nyquist samples," *IEEE Int. Workshop Signal Process. Advances Wireless Comm.*, pp. 250–254, Jun. 2013.
- [28] G. Leus and Z. Tian, "Recovering second-order statistics from compressive measurements," *IEEE Int. Workshop Computational Advances in Multi-Sensor Adaptive Process.*, pp. 337–340, Dec. 2011.
- [29] D. D. Ariananda and G. Leus, "Non-uniform sampling for compressive cyclic spectrum reconstruction," *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, pp. 41–45, May 2014.
- [30] L. Zhu, C. E. Luo, and J. H. McClellan, "Cyclostationary-based wide-band spectrum sensing using random sampling," *IEEE Global Conf. Signal and Inf. Process.*, pp. 1202–1205, Dec. 2013.
- [31] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.
- [32] C. E. Luo and J. H. McClellan, "Compressive sampling with a successive approximation ADC architecture," *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, pp. 3920–3923, May 2011.
- [33] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-nyquist rates," *IET Circuits, Devices & Syst.*, vol. 5, pp. 8–20, Jan. 2011.
- [34] H. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math*, vol. 117, pp. 37–52, Jul. 1967.
- [35] D. Cohen, L. Pollak, and Y. C. Eldar, "Carrier frequency and bandwidth estimation of cyclostationary multiband signals," *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, pp. 3716–3720, Mar. 2016.
- [36] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*. Springer, 2003.
- [37] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 1991.
- [38] C. W. Lim and M. B. Wakin, "Compressive temporal higher order cyclostationary statistics," *IEEE Trans. Signal Process.*, vol. 63, pp. 2942–2956, Jun. 2015.
- [39] R. Venkataramani and Y. Bresler, "Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2173 – 2183, Sept. 2000.
- [40] M. Mishali and Y. C. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Process.*, vol. 56, pp. 4692–4702, Oct. 2008.
- [41] R. L. Thorndike, "Who belong in the family?" *Psychometrika*, vol. 18, pp. 267–276, Dec. 1953.