Analog to Digital Cognitive Radio

Deborah Cohen, Shahar Tsiper and Yonina C. Eldar

Abstract Enabling cognitive radio (CR) requires revisiting the traditional task of spectrum sensing with specific and demanding requirements in terms of detection performance, real-time processing and robustness to noise. Unfortunately, conventional spectrum sensing methods do not satisfy these demands. In particular, the Nyquist rate of signals typically sensed by a CR is prohibitively high so that sampling at this rate necessitates sophisticated and expensive analog to digital converters, which lead to a torrent of samples. Over the past few years, several sampling methods have been proposed that exploit signals' a priori known structure to sample them below Nyquist. In this chapter, we review some of these techniques and tie them to the task of spectrum sensing for CRs. We then show how other spectrum sensing challenges can be tackled in the sub-Nyquist regime. First, to cope with low signal to noise ratios, spectrum sensing may be based on second-order statistics recovered from the low rate samples. In particular, cyclostationary detection allows to differentiate between communication signals and stationary noise. Next, CR networks, that perform collaborative low rate spectrum sensing, have been proposed to overcome fading and shadowing channel effects. Last, to enhance CR efficiency, we present joint spectrum sensing and direction of arrival estimation methods from sub-Nyquist samples. These allow to map the temporarily vacant bands both in terms of frequency and space. Throughout this chapter, we highlight the relation between theoretical algorithms and results and their practical implementation. We show hardware simulations performed on a prototype built with off-the-shelf devices, demonstrating the feasibility of sub-Nyquist spectrum sensing in the context of CR.

Deborah Cohen
Technion - Israel Institute of Technology, Haifa, Israel, e-mail: debby@technion.ac.il

Shahar Tsiper
Technion - Israel Institute of Technology, Haifa, Israel, e-mail: tsiper@technion.ac.il

Yonina C. Eldar
Technion - Israel Institute of Technology, Haifa, Israel, e-mail: yonina@ee.technion.ac.il

1 Introduction

In order to increase the chance of finding an unoccupied spectral band, Cognitive Radios (CRs) have to sense a wide band of spectrum. Nyquist rates of wideband signals 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. To overcome the rate bottleneck, several sampling methods have been proposed that leverage the a priori known received signal's structure, enabling sampling rate reduction [1, 2]. These include the random demodulator [3], multi-rate sampling [4], multicoset sampling and the modulated wideband converter (MWC) [5, 6, 7, 8].

The CR then performs spectrum sensing on the acquired samples to detect the presence of primary users' (PUs) transmissions. The simplest and most common spectrum sensing approach is energy detection [9], 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. This becomes even more critical in sub-Nyquist regimes since the sensitivity of energy detection is amplified due to aliasing of the noise [10]. Therefore, this scheme fails to meet CR performance requirements in low SNRs. In contrast, matched filter (MF) detection [11, 12], 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. However, this technique requires perfect knowledge of the potential received transmission. When no a priori knowledge is assumed on the received signals' waveform, MF is difficult to implement. A compromise between both methods is cyclostationary detection [13, 14]. This strategy is more robust to noise than energy detection but at the same time only assumes that the signal of interest exhibits cyclostationarity, which is a typical characteristic of communication signals. Consequently, cyclostationary detection is a natural candidate for spectrum sensing from sub-Nyquist samples in low SNRs.

Besides noise, the task of spectrum sensing for CRs is further complicated due to path loss, fading and shadowing [15]. These phenomena are due to the signal's propagation that can be affected by obstacles and multipath, and result in the attenuation of the signal's power. To overcome these practical issues, collaborative CR networks have been considered, where different users share their sensing results and cooperatively decide on the licensed spectrum occupancy [15, 16, 17]. Cooperative spectrum sensing can be classified into three categories based on the way the data is shared by the CRs in the network: centralized, distributed and relay-assisted. In each of these settings, two options of data fusion arise: decision fusion, or hard decision, where the CRs only report their binary local decisions, and measurement fusion, or soft decision, where they share their samples [15]. Cooperation has been shown to improve detection performance and relax sensitivity requirements by exploiting spatial diversity.

Finally, CRs may require, or at least benefit from joint spectrum sensing and direction of arrival (DOA) estimation. DOA recovery can enhance CR performance by allowing exploitation of vacant bands in space in addition to the frequency domain.

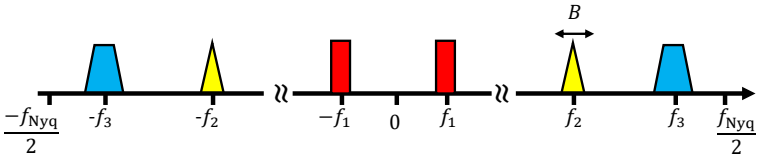


Fig. 1 Multiband model with $K = 6$ bands. Each band does not exceed the bandwidth B and is modulated by an unknown carrier frequency $|f_i| \leq f_{\text{Nyq}}/2$, for $i = 1,2,3$.

For example, a spectral band occupied by a PU situated in a certain direction with respect to the CR, may be used by the latter for transmission to the opposite direction, where receivers do not sense the PU's signal. In order to estimate jointly the carrier frequencies and DOAs of the received transmissions, arrays of sensors have been considered. DOA recovery techniques, such as MUSIC [18, 19], ESPRIT [20] or compressed sensing (CS) [21] techniques, may then be adapted to the joint carrier and DOA estimation problem both in Nyquist and sub-Nyquist regimes

This chapter focuses on the spectrum sensing challenges for CR outlined above. We first review sub-Nyquist sampling methods for multiband signals and then consider different aspects of spectrum sensing performed on low rate samples, including cyclostationary detection, collaborative spectrum sensing and joint carrier frequency and DOA estimation. Our emphasis is on practical low rate acquisition schemes and tailored recovery that can be implemented in real CR settings. The approach adopted here focuses on the analog to digital interface of CRs. In particular, we are concerned with compressive spectrum sensing, including the application of CS to analog signals. Modeling the analog to digital conversion allows demonstrating the realization of the theoretical concepts on hardware prototypes. We focus on the implementation of one sampling scheme reviewed here, the MWC, and show how the same low rate samples can be used in the different extensions of spectrum sensing described above.

2 Sub-Nyquist Sampling for CR

CR receivers sense signals composed of several transmissions with unknown support, spread over a wide spectrum. Such sparse wideband signals belong to the so-called multiband model [6, 7]. An example of a multiband signal $x(t)$ with K bands is illustrated in Fig. 1. The bandwidth of each band is no greater than B, and is centered around unknown carrier frequencies $|f_i| \leq f_{\text{Nyq}}/2$, where f_{Nyq} denotes the signals' Nyquist rate and i indexes the transmissions. Note that, for real-valued signals, K is an even integer due to spectral conjugate symmetry and the number of transmissions is $N_{\text{sig}} = K/2$.

When the frequency support of $x(t)$ is known, classic sampling methods such as demodulation, undersampling ADCs and interleaved ADCs (see [1, 2] and refer-

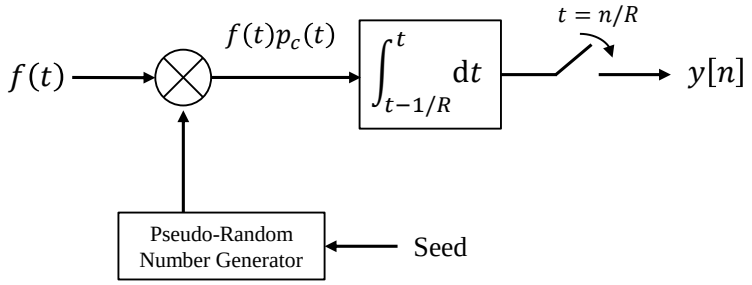


Fig. 2 Block diagram for the random demodulator, including a random number generator, a mixer, an accumulator and a sampler [3].

ences therein) may be used to reduce the sampling rate below Nyquist. Here, since the frequency location of the transmissions are unknown, classic processing first samples $x(t)$ at its Nyquist rate f_{Nyq}, which may be prohibitively high. To overcome the sampling rate bottleneck, several blind sub-Nyquist sampling and recovery schemes have been proposed that exploit the signal's structure and in particular its sparsity in the frequency domain, but do not require knowledge of the carrier frequencies. It has been shown in [6] that the minimal sampling rate for perfect blind recovery in multiband settings is twice the Landau rate [22], that is twice the occupied bandwidth. This rate can be orders of magnitude lower than Nyquist. In the remainder of this section, we survey several sub-Nyquist methods, that theoretically achieve this minimal sampling rate.

2.1 Multitone Model and the Random Demodulator

Tropp et al. [3] consider a discrete multitone model for multiband signals and suggest sampling using the random demodulator, depicted in Fig. 2. Multitone functions are composed of K active tones spread over a bandwidth W, such that

$$f(t) = \sum_{\omega \in \Omega} b_\omega \mathrm{e}^{-2\pi i \omega t}, \quad t \in [0,1)\,. \tag{1}$$

Here, Ω is a set of K normalized frequencies, or tones, that satisfies

$$\Omega \subset \{0, \pm 1, \pm 2, \ldots, \pm(W/2-1), \pm W/2\}, \tag{2}$$

and b_ω, for $\omega \in \Omega$, are a set of complex-valued amplitudes. The number of active tones K is assumed to be much smaller than the bandwidth W. The goal is to recover both the tones ω and the corresponding amplitudes b_ω.

To sample the signal $f(t)$, it is first modulated by a high-rate sequence $p_c(t)$ created by a pseudo-random number generator. It is then integrated and sampled

at a low rate, as shown in Fig. 2. The random sequence used for modulation is a square wave, which alternates between the levels ± 1 with equal probability. The K tones present in $f(t)$ are thus aliased by the pseudo-random sequence. The resulting modulated signal $y(t) = f(t)p_c(t)$ is integrated over a period $1/R$ and sampled at the low rate R. This integrate-and-dump approach results in the following samples

$$y_m = R \int_{m/R}^{(m+1)/R} y(t)\mathrm{d}t, \quad m = 0, 1, \ldots, R-1. \tag{3}$$

The samples y_m acquired by the random demodulator can be written as a linear combination of the $W \times 1$ sparse amplitude vector $\mathbf{b}$ that contains the coefficients b_ω at the corresponding locations ω [3]. In matrix form, we write

$$\mathbf{y} = \mathbf{A}\mathbf{b}, \tag{4}$$

where $\mathbf{y}$ is the vector of size R that contains the samples y_m and $\mathbf{A}$ is the known sampling matrix that describes the overall action of the system on the vector of amplitudes $\mathbf{b}$, namely modulation and filtering (see [3] for more details). Capitalizing on the sparsity of the vector $\mathbf{b}$, the amplitudes b_ω and their respective locations ω can be recovered from the low rate samples $\mathbf{y}$ using CS [21] techniques, in turn allowing for the recovery of $f(t)$. CS provides a framework for simultaneous sensing and compression of finite-dimensional vectors, which relies on linear dimensionality reduction. It provides both recovery conditions and algorithms to reconstruct sparse vectors from low-dimensional measurement vectors, represented as linear combinations of the former. Here, the minimal required number of samples R for perfect recovery of $f(t)$ in noiseless settings is $2K$ [21].

The random demodulator is one of the pioneer attempts to extend the inherently discrete and finite CS theory to analog signals. However, truly analog signals, as those we consider here, require a prohibitively large number of harmonics to approximate them well within the discrete model. When attempting to approximate signals such as those from the multiband model, the number of tones W is on the order of the Nyquist rate and the number of samples R is a multiple of KB. This in turn renders the reconstruction computationally prohibitive and very sensitive to the grid choice (see [1] for a detailed analysis). Furthermore, the time-domain approach precludes processing at a low rate, even for multitone inputs since interpolation to the Nyquist rate is an essential ingredient of signal reconstruction. In terms of hardware and practical implementation, the random demodulator requires accurate modulation by a periodic square mixing sequence and accurate integration, which may be challenging when using analog signal generators, mixers and filters.

In contrast to the random demodulator, which adopts a discrete multitone model, the rest of the approaches we focus on treat the analog multiband model, illustrated in Fig. 1, which is of interest to us in the context of CR.

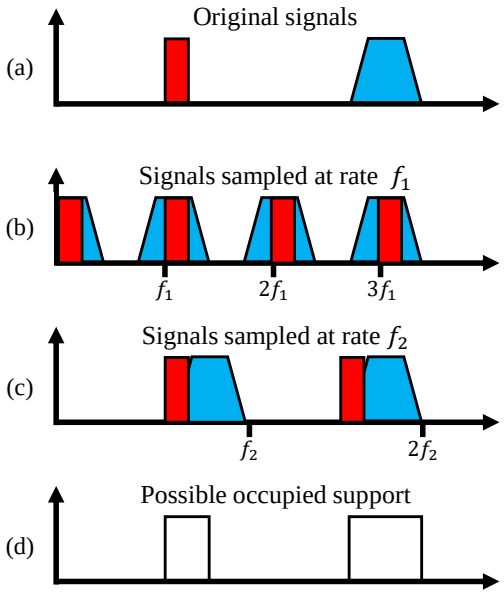


Fig. 3 Action of the SMRS on a multiband signal: (a) the input signal with $K = 2$ bands, (b) signals sampled at rate F_1 in channel 1, (c) signals sampled at rate F_2 in channel 2, and (d) possible support which is the intersection of the supports in channel 1 and 2 [4].

2.2 Multi-Rate Sampling

An alternative sampling approach is based on the synchronous multi-rate sampling (SMRS) [4] scheme, which has been proposed in the context of electro-optical systems to undersample multiband signals. The SMRS samples the input signal at P different sampling rates F_i, each of which is an integer multiple of a basic sampling rate Δf. This procedure aliases the signal with different aliasing intervals, as illustrated in Fig. 3. The Fourier transform of the undersampled signals is then related to the original signal through an underdetermined system of linear equations,

$$\mathbf{z}(f) = \mathbf{Q}\mathbf{x}(f). \tag{5}$$

Here, $\mathbf{x}(f)$ contains frequency slices of size Δf of the original signal $x(t)$ and $\mathbf{z}(f)$ is composed of the Fourier transform of the sampled signal. Each channel contributes $M_i = F_i/\Delta f$ equations to the system (5), which concatenates the observation vector of all the channels. The measurement matrix $\mathbf{Q}$ has exactly P non-zero elements in every column, that correspond to the locations of the spectral replica in each channel baseband $[0, F_i]$.

This approach assumes that either the signal or the sampling time window are finite. The continuous variable f is then discretized to a frequency resolution of Δf. Since $x(t)$ is sparse in the frequency domain, the vector $\mathbf{x}(f)$ is sparse and can be recovered from (5) using CS techniques, for each discrete frequency f. An alternative recovery method, referred to as the reduction procedure, consists of detecting baseband frequencies in which there is no signal, by observing the samples. These

frequencies are assumed to account for the absence of signals of interest in all the frequencies that are down-converted to that baseband frequency. This allows to reduce the number of sampling channels. This assumption does not hold in the case where two or more frequency components cancel each other due to aliasing, which happens with probability zero. The procedure is illustrated in Fig. 3. Once the corresponding components are eliminated from (5), the reduced system can be inverted using the Moore-Penrose pseudo-inverse to recover $\mathbf{x}(f)$.

There are several drawbacks to the SMRS that limit its performance and potential implementation. First, the discretization process affects the SNR since some of the samples are thrown out. Furthermore, spectral components down-converted to off the grid frequencies are missed. In addition, the first recovery approach requires a large number of sampling channels, proportional to the number of active bands K, whereas the reduction procedure does not ensure a unique solution and the inversion problem is ill-posed in many cases. Finally, in practice, synchronization between channels sampling at different rates is challenging. Moreover, this scheme samples wideband signals using low rate samplers. Practical ADCs introduce an inherent bandwidth limitation, modeled by an anti-aliasing low pass filter (LPF) with cut-off frequency determined by the sampling rate, which distorts the samples. To avoid this issue, the multi-rate strategy would require low rate samplers with large analog bandwidth.

2.3 Multicoset sampling

A popular sampling scheme for sampling wideband signals at the Nyquist rate is multicoset or interleaved ADCs [23, 6, 1] in which several channels are used, each operating at a lower rate. We now discuss how such systems can be used in the sub-Nyquist regime.

Multicoset sampling may be described as the selection of certain samples from the uniform Nyquist grid, as shown in Fig. 4, where $T_{\text{Nyq}} = 1/f_{\text{Nyq}}$ denotes the Nyquist period. More precisely, the uniform grid is divided into blocks of N consecutive samples, from which only $M < N$ are kept. Mathematically, the ith sampling sequence is defined as

$$x_{c_i}[n] = \begin{cases} x(nT_{\text{Nyq}}), & n = mN + c_i, m \in \mathbb{Z} \\ 0, & \text{otherwise}, \end{cases} \tag{6}$$

where the cosets c_i are ordered integers so that $0 \leq c_1 < c_2 < \cdots < c_M < N$. A possible implementation of the sampling sequences (6) is depicted in Fig. 5. The building blocks are M uniform samplers at rate $1/NT_{\text{Nyq}}$, where the ith sampler is shifted by $c_i T_{\text{Nyq}}$ from the origin. When sampling at the Nyquist rate, $M = N$ and $c_i = (i-1)$.

The samples in the Fourier domain can be written as linear combinations of spectrum slices of $x(t)$, such that [6]

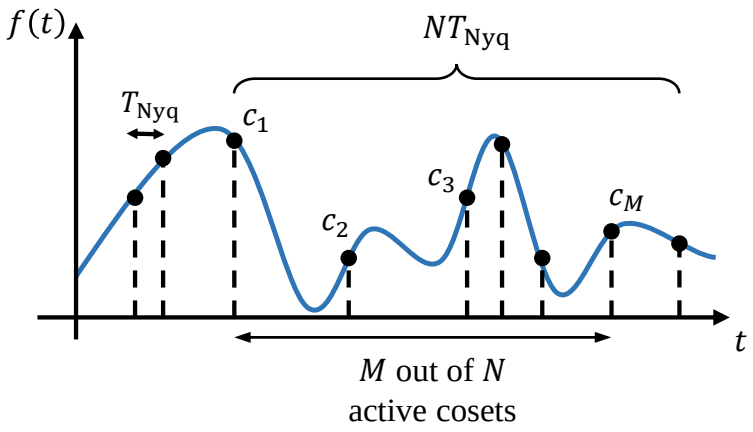


Fig. 4 Illustration of multicoset sampling.

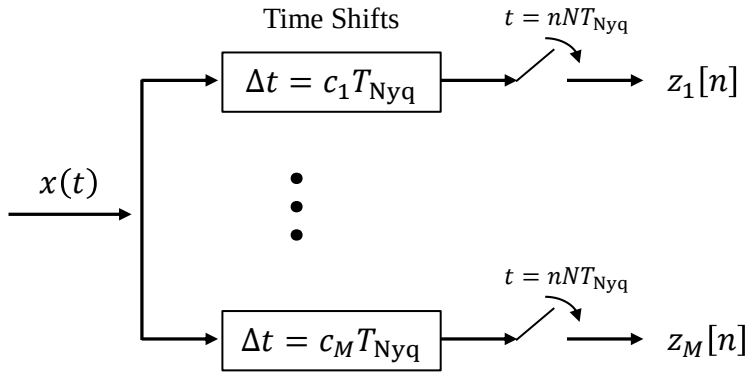


Fig. 5 Schematic implementation of multicoset sampling. The input signal $x(t)$ is inserted into the multicoset sampler that splits the signal into M branches, and delays each one by a fixed coefficient $c_i T_{\text{Nyq}}$. Every branch is sampled at the low rate $1/(NT_{\text{Nyq}})$, and then digitally processed to perform spectrum sensing and signal reconstruction.

$$\mathbf{z}(f) = \mathbf{A}\mathbf{x}(f), \qquad f \in \mathscr{F}_s. \tag{7}$$

Here, $\mathscr{F}_s = [-f_s/2, f_s/2]$ with $f_s = \frac{1}{NT_{\text{Nyq}}} \geq B$ the sampling rate of each channel. The mth row of $\mathbf{z}(f)$ contains the discrete time Fourier transform of the samples $z_m[n]$. The $N \times 1$ vector $\mathbf{x}(f)$ denotes the spectrum slices of $x(t)$, where the ith row of $\mathbf{x}(f)$ is $\mathbf{x}_i(f) = X(f + (i - \lfloor (N+2)/2 \rfloor) f_p)$, and $X(f)$ is the Fourier transform of $x(t)$. Since $x(t)$ is assumed to be sparse, $\mathbf{x}(f)$ is sparse as well, and its support, that is the set that contains the indices corresponding to its non zero rows, is determined

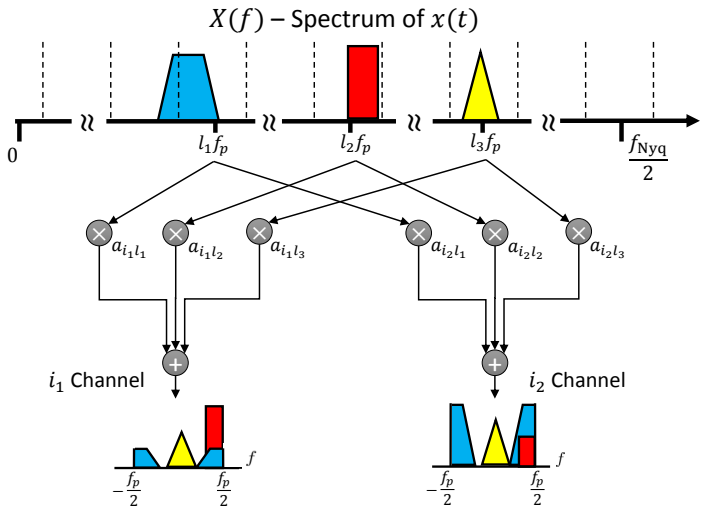


Fig. 6 The spectrum slices of the input signal $\mathbf{x}(f)$ are shown here to be multiplied by the coefficients a_{il} of the sensing matrix $\mathbf{A}$, resulting in the measurements z_i for the ith channel. Note that in multicoset sampling, only the slices' complex phase is modified by the coefficients a_{il}. In the MWC sampling described below, both the phases and amplitudes are affected.

by the frequency locations of the transmissions of $x(t)$. The $M \times N$ sampling matrix $\mathbf{A}$ is a Vandermonde matrix with factors determined by the selected delays or cosets c_i. This relation is illustrated in Fig. 6. In the Nyquist regime, when $M = N$, $\mathbf{A}$ is the Fourier matrix. The recovery processing described below is performed in the time domain, where we have

$$\mathbf{z}[n] = \mathbf{A}\mathbf{x}[n], \qquad n \in \mathbb{Z}. \tag{8}$$

The vector $\mathbf{z}[n]$ collects the measurements at $t = n/f_s$ and $\mathbf{x}[n]$ contains the sample sequences corresponding to the spectrum slices of $x(t)$. Obviously, the sparsity pattern of $\mathbf{x}[n]$ is identical to that of $\mathbf{x}(f)$ and it follows that $\mathbf{x}[n]$ are jointly sparse over time.

Our goal is to recover $\mathbf{x}[n]$ from the samples $\mathbf{z}[n]$. The system (8) is underdetermined due to the sub-Nyquist setup and known as infinite measurement vector (IMV) in the CS literature [21, 2]. The digital reconstruction algorithm consists of the following three stages [6] that we explain in more detail below:

1. The continuous-to-finite (CTF) block constructs a finite frame (or basis) from the samples.
2. The support recovery formulates an optimization problem whose solution's support is identical to the support S of $\mathbf{x}[n]$, that is the active slices.
3. The signal is then digitally recovered by reducing (8) to the support of $\mathbf{x}[n]$.

The recovery of $\mathbf{x}[n]$ for every n independently is inefficient and not robust to noise. Instead, the CTF method, developed in [6], exploits the fact that the bands occupy continuous spectral intervals so that $\mathbf{x}[n]$ are jointly sparse, that is they have the same spectral support S over time. The CTF then produces a finite system of equations, called multiple measurement vectors (MMV) [21, 2] from the infinite number of linear systems described by (8). The samples are first summed as

$$\mathbf{Q} = \sum_n \mathbf{z}[n]\mathbf{z}^H[n], \tag{9}$$

and then decomposed to a frame $\mathbf{V}$ such that $\mathbf{Q} = \mathbf{V}\mathbf{V}^H$. Clearly, there are many possible ways to select $\mathbf{V}$. One option is to construct it by performing an eigen-decomposition of $\mathbf{Q}$ and choosing $\mathbf{V}$ as the matrix of eigenvectors corresponding to the non zero (or large enough) eigenvalues. The finite dimensional MMV system

$$\mathbf{V} = \mathbf{A}\mathbf{U}, \tag{10}$$

is then solved for the sparsest matrix $\mathbf{U}$ with minimal number of non-identically zero rows using CS techniques [21, 2]. The key observation of this recovery strategy is that the indices of the non zero rows of $\mathbf{U}$ coincide with the active spectrum slices of $\mathbf{z}[n]$ [6]. These indices are referred to as the support of $\mathbf{z}[n]$ and are denoted by S.

Once the support S is known, $\mathbf{x}[n]$ is recovered by reducing the system of equations (8) to S. The resulting matrix $\mathbf{A}_S$, that contains the columns of $\mathbf{A}$ corresponding to S, is then inverted

$$\mathbf{x}_S[n] = \mathbf{A}_S^{\dagger}\mathbf{z}[n]. \tag{11}$$

Here, $\mathbf{x}_S[n]$ denotes the vector $\mathbf{x}[n]$ reduced to its support. The remaining entries of $\mathbf{x}[n]$ are equal to zero.

The overall sampling rate of the multicoset system is

$$f_{\text{Total}} = Mf_s = \frac{M}{N} f_{\text{Nyq}}. \tag{12}$$

The minimal number of channels is dictated by CS results [21] which imply that $M \geq 2K$ with $f_s \geq B$ per channel. The sampling rate can thus be as low as $2KB$, which is twice the Landau rate [22].

Although this sampling scheme seems relatively simple and straightforward, it suffers from several practical drawbacks [1]. First, as in the multi-rate approach, multicoset sampling requires low rate ADCs with large analog bandwidth. Another issue arises from the time shift elements, since maintaining accurate time delays between the ADCs on the order of the Nyquist interval T_{Nyq} is difficult. Last, the number of channels M required for recovery of the active bands can be prohibitively high. The MWC, presented in the next section, uses similar recovery techniques while overcoming these practical sampling issues.

2.4 MWC sampling

The MWC [7] exploits the blind recovery ideas developed in [6] and combines them with the advantages of analog RF demodulation. To circumvent the analog bandwidth issue in the ADCs, an RF front-end mixes the input signal $x(t)$ with periodic waveforms. This operation imitates the effect of delayed undersampling used in the multicoset scheme and results in folding the spectrum to baseband with different weights for each frequency interval. The MWC achieves aliasing by mixing the signal, which is filtered prior to sampling. The ADC's input is thus a narrowband signal in contrast with multicoset which samples a wideband signal at a low rate to create aliasing. This characteristic of the MWC enables practical hardware implementation, which will be described in Section 3.

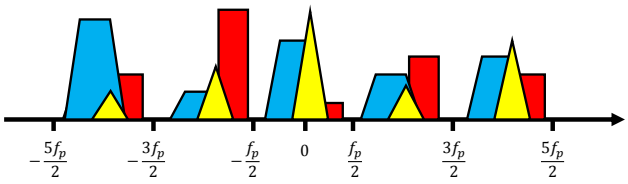
More specifically, the MWC is composed of M parallel channels. In each channel, $x(t)$ is multiplied by a periodic mixing function $p_i(t)$ with period $T_p = 1/f_p$ and Fourier expansion

$$p_i(t) = \sum_{l=-\infty}^{\infty} a_{il} e^{j\frac{2\pi}{T_p}lt}. \tag{13}$$

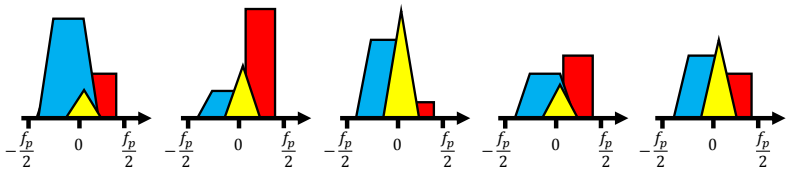
The mixing process aliases the spectrum, such that each band appears in baseband. The signal then goes through a LPF with cut-off frequency $f_s/2$ and is sampled at rate $f_s \geq f_p$. The analog mixture boils down to the same mathematical relation between the samples and the $N = f_{\text{Nyq}}/f_s$ frequency slices of $x(t)$ as in multicoset sampling, namely (7) in frequency and (8) in time, as shown in Fig. 6. Here, the $M \times N$ sampling matrix $\mathbf{A}$ contains the Fourier coefficients a_{il} of the periodic mixing functions. The recovery conditions and algorithm are identical to those described for multicoset sampling.

Choosing the channels' sampling rate f_s to be equal to the mixing rate f_p results in a similar configuration as the multicoset scheme in terms of the number of channels. In this case, the minimal number of channels required for the recovery of K bands is $2K$. The number of branches dictates the total number of hardware devices and thus governs the level of complexity of the practical implementation. Reducing the number of channels is a crucial challenge for practical implementation of a CR receiver. The MWC architecture presents an interesting flexibility property that permits trading channels for sampling rate, allowing to drastically reduce the number of channels, even down to a single channel.

Consider a configuration where $f_s = qf_p$, with odd q. In this case, the ith physical channel provides q equations over $\mathscr{F}_p = [-f_p/2, f_p/2]$, as illustrated in Fig. 7. Conceptually, M physical channels sampled at rate $f_s = qf_p$ are equivalent to Mq channels sampled at $f_s = f_p$. The number of channels is thus reduced at the expense of higher sampling rate f_s in each channel and additional digital processing. The output of each of the M physical channels is digitally demodulated and filtered to produce samples that would result from Mq equivalent virtual branches. This happens in the so-called expander module, directly after the sampling stage and before the digital processing described above, in the context of multicoset sampling. At its



(a) Spectrum of $\tilde{z}_i[n]$ for $f_s = 5f_p$.



(b) Spectrum of $z_{i,j}[n]$ after expansion.

Fig. 7 Illustration of the expander configuration for $q = 5$. (a) Spectrum of the output $\tilde{z}_i[n]$ of the physical ith channel, (b) spectrum of the samples $z_{i,j}[n]$ of the $q = 5$ equivalent virtual channels, for $j = 1,\ldots,5$, after digital expansion.

brink, this strategy allows to collapse a system with M channels to a single branch with sampling rate $f_s = Mf_p$ (further details can be found in [7, 24, 25]).

The MWC sampling and recovery processes are illustrated in Fig. 8. This approach results in a hardware-efficient sub-Nyquist sampling method that does not suffer from the practical limitations described in previous sections, in particular, the analog bandwidth limitation of low rate ADCs. In addition, the number of MWC channels can be drastically reduced below $2K$ to as few as one, using a higher sampling rate f_s in each channel and additional digital processing. This tremendously reduces the burden on hardware implementation. However, the choice of appropriate periodic functions $p_i(t)$ to ensure correct recovery is challenging. Some guidelines are provided in [26, 27, 2].

2.5 Uniform Linear Array based MWC

An alternative sensing configuration, composed of a uniform linear array (ULA) and relying on the sampling paradigm of the MWC, is presented in [28]. The sensing system consists of a ULA composed of M sensors, with two adjacent sensors separated by a distance d, such that $d < c/(|\cos(\theta)|f_{\text{Nyq}})$, where c is the speed of light and θ is the angle representing the DOA of the signal $x(t)$. This system, illustrated in Fig. 9, capitalizes on the different accumulated phases of the input signal between sensors, given by $e^{j2\pi f_i \tau_m}$, where

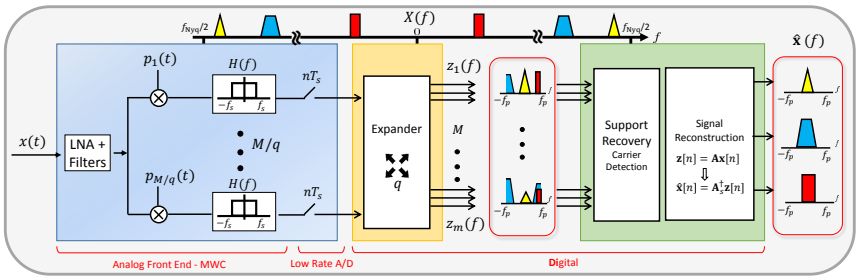


Fig. 8 Schematic implementation of the MWC analog sampling front-end and digital signal recovery from low rate samples.

$$\tau_m = \frac{dm}{c}\cos(\theta) \tag{14}$$

is the delay at the mth sensor with respect to the first one. Each sensor implements one channel of the MWC, that is the input signal is mixed with a periodic function, low-pass filtered and then sampled at a low rate.

This configuration has three main advantages over the standard MWC. First, it allows for a simpler design of the mixing functions which can be identical in all sensors. The only requirement on $p(t)$, besides being periodic with period $T_p \leq 1/B$, is that none of its Fourier series coefficients within the signal's Nyquist bandwidth is zero. Second, the ULA based system outperforms the MWC in terms of recovery performance in low SNR regimes. Since all the MWC channels belong to the same sensor, they are all affected by the same additive sensor noise. In the ULA architecture, each channel belongs to a different sensor with uncorrelated sensor noise between channels. The alternative approach benefits from the same flexibility as the standard MWC in terms of collapsing the channels, which translates into reducing the antennas in the alternative configuration. This lead to a trade-off exists between hardware complexity, governed by the number of antennas, and SNR. Finally, as will be shown in Section 6.3, the modified system can be easily extended to enable joint spectrum sensing and DOA estimation.

Similarly to the previous sampling schemes, the samples $\mathbf{z}(f)$ can be expressed as a linear transformation of the unknown vector of slices $\mathbf{x}(f)$, such that

$$\mathbf{z}(f) = \mathbf{A}\mathbf{x}(f), \qquad f \in \mathscr{F}_s. \tag{15}$$

Here, $\mathbf{x}(f)$ is a non sparse vector that contains cyclic shifted, scaled and sampled versions of the active bands, as shown in Fig. 10. In contrast to the previous methods, in this configuration, the matrix $\mathbf{A}$, defined by

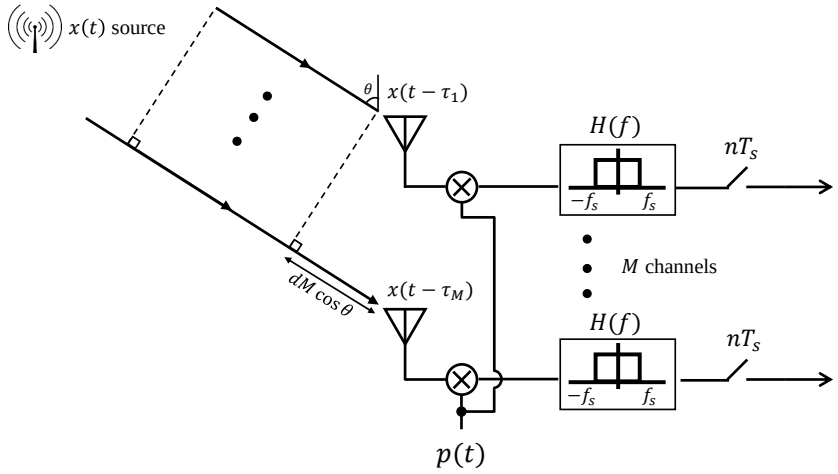


Fig. 9 ULA configuration with M sensors, with distance d between two adjacent sensors. Each sensor includes an analog front-end composed of a mixer with the same periodic function $p(t)$, a LPF and a sampler, at rate f_s.

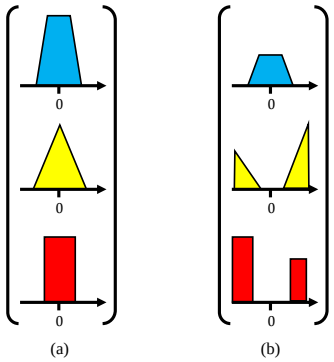


Fig. 10 (a) Original source signals at baseband (before modulation), (b) output signals at baseband $\mathbf{x}(f)$ after modulation, mixing, filtering and sampling.

$$\mathbf{A} = \begin{pmatrix} e^{j2\pi f_1 \tau_1} & \dots & e^{j2\pi f_N \tau_1} \\ \vdots & & \vdots \\ e^{j2\pi f_1 \tau_M} & \dots & e^{j2\pi f_N \tau_M} \end{pmatrix}, \tag{16}$$

depends on the unknown carrier frequencies. As before, in the time domain

$$\mathbf{z}[n] = \mathbf{A}\mathbf{x}[n], \quad n \in \mathbb{Z}. \tag{17}$$

Two approaches are presented in [28] to recover the carrier frequencies of the transmissions composing the input signal. The first is based on CS algorithms and

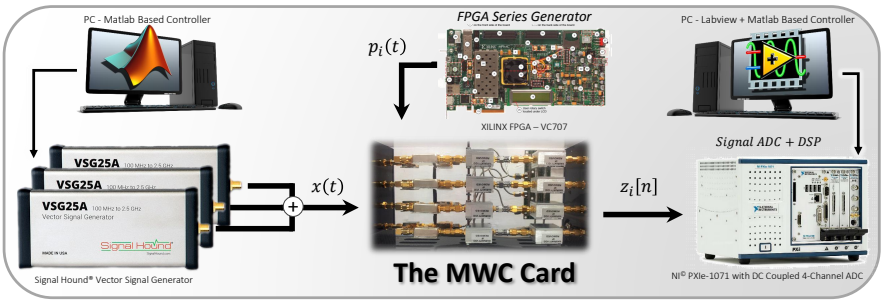


Fig. 11 Hardware implementation of the MWC prototype, including the RF signal generators, analog front-end board, FPGA series generator, ADC and DSP.

assumes that the carriers lie on a predefined grid. In this case, the resulting sensing matrix, which extends $\mathbf{A}$ with respect to the grid, is known and the expanded vector $\mathbf{x}(f)$ is sparse. This leads to a similar system as (7) or (8) which can be solved using the recovery paradigm from [6], described in the context of multicoset sampling.

In the second technique, the grid assumption is dropped and ESPRIT [20] is used to estimate the carrier frequencies. This approach first computes the sample covariance of the measurements

$$\mathbf{R} = \sum_n \mathbf{z}[n]\mathbf{z}^H[n], \tag{18}$$

and performs a singular value decomposition (SVD). The non zero singular values correspond to the signal's subspace and the carrier frequencies are then estimated from these. Once the carriers are recovered, the signal itself is reconstructed by inverting the sampling matrix $\mathbf{A}$ in (17).

The minimal number of sensors required by both reconstruction methods in noiseless settings is $M = 2K$, with each sensor sampling at the minimal rate of $f_s = B$ to allow for perfect signal recovery [28]. The proposed system thus achieves the minimal sampling rate $2KB$ derived in [6]. We note that the expander strategy proposed in the context of the MWC can be applied in this configuration as well.

3 MWC Hardware

3.1 MWC Prototype

One of the main aspects that distinguish the sub-Nyquist MWC from other sampling schemes is its practical implementation [24], proving the feasibility of sub-Nyquist sampling even under distorting effects of analog components and physical phenomena. A hardware prototype, shown in Fig.11, was developed and built according to

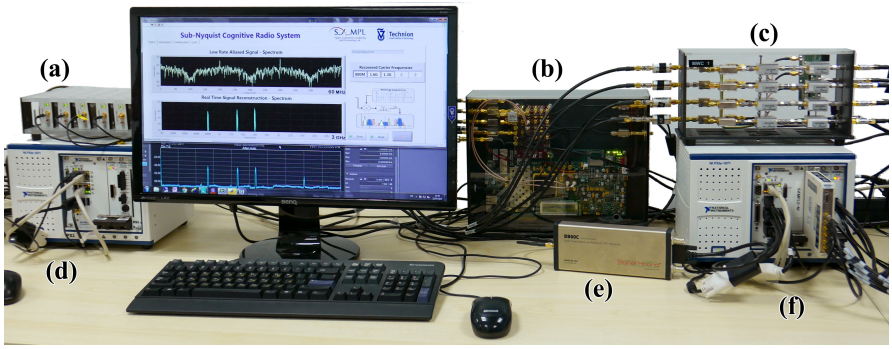


Fig. 12 MWC CR system prototype: (a) vector signal generator (VSG), (b) FPGA mixing sequences generator, (c) MWC analog front-end board, (d) RF combiner, (e) spectrum analyzer, (f) ADC and DSP.

the block diagram in Fig. 8. The main hardware components that were used in the prototype can be seen in Fig. 11. In particular, the system receives an input signal with Nyquist rate of 6 GHz and spectral occupancy of up to 200 MHz, and samples at an effective rate of 480 MHz, that is only 8% of the Nyquist rate and 2.4 times the Landau rate. This rate constitutes a relatively small oversampling factor of 20% with respect to the theoretical lower sampling bound. This section describes the different components of the hardware prototype, shown in Fig. 12, explaining the various considerations that were taken into account when implementing the theoretical concepts on actual analog components.

At the heart of the system lies the proprietary MWC board [24] that implements the sub-Nyquist analog front-end. The card uses a high speed 1-to-4 analog splitter that duplicates the wideband signal to $M = 4$ channels, with an expansion factor of $q = 5$, yielding $Mq = 20$ virtual channels after digital expansion. Then, an analog preprocessing step, composed of preliminary equalization, impedance corrections and gain adjustments, aims at maintaining the dynamic range and fidelity of the input in each channel. Indeed, the signal and mixing sequences must be amplified to specific levels before entering the analog mixers to ensure proper behavior emulating mathematical multiplication with the mixing sequences. The entire analog path of the multiband input signal is described in Fig. 13.

The modulated signal next passes through an analog anti-aliasing LPF. The anti-aliasing filter must be characterized by both an almost linear phase response in the pass band, between 0 to 50 MHz, and an attenuation of more than 20 dB at $f_s/2 = 60$ MHz. A Chebyshev LPF of 7th order with cut-off frequency (-3 dB) of 50 MHz was chosen for the implementation. After impedance and gain corrections, the signal now has a spectral content limited to 50 MHz, that contains a linear combination of the occupied bands with different amplitudes and phases, as seen in Fig. 6. Finally the low rate analog signal is sampled by a National Instruments© ADC operating at 120 MHz, leading to a total sampling rate of 480 MHz.

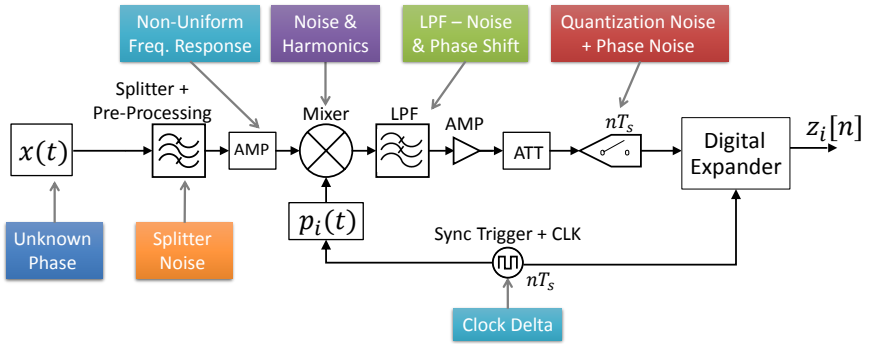


Fig. 13 Hardware RF chain detailed schematics, including amplifiers, attenuators, filters, mixers, samplers and synchronization signals required for precise and accurate operation. The distortions induced by each component are indicated as well.

The mixing sequences that modulate the signal play an essential part in signal recovery. They must have low cross-correlations with each other, while spanning a large bandwidth determined by the Nyquist rate of the input signal, and yet be easy enough to generate with relatively cheap, off-the-shelf hardware. The sequences $p_i(t)$, for $i = 1,\ldots,4$, are chosen as truncated versions of Gold Codes [29], which are commonly used in telecommunication (CDMA) and satellite navigation (GPS). Mixing sequences based on Gold codes were found to give good results in the MWC system [26], primarily due to small bounded cross-correlations within a set.

Since Gold codes are binary, the mixing sequences are restricted to alternating ± 1 values. This fact allows to digitally generate the sequences on a dedicated FPGA. Alternatively, they can be implemented on a small chip with very low power and complexity. The added benefit of producing the mixing sequences on such a platform is that the entire sampling scheme can be synchronized and triggered using the same FPGA with minimally added phase noise and jitter, keeping a closed synchronization loop with the samplers and mixers. A XiLinX VC707 FPGA acts as the central timing unit of the entire sub-Nyquist CR setup by generating the mixing sequences and the synchronization signals required for successful operation. It is crucial that both the mixing period $T_p = 1/f_p$ and the low rate samplers operating at $(q+1)f_p$ (due to intended oversampling) are fully synchronized, in order to ensure correct modeling of the entire system, and consequently guarantee accurate support detection and signal reconstruction.

The digital back-end is implemented using a National Instruments© PXIe-1065 computer with DC coupled ADC. Since the digital processing is performed at the low rate f_s, very low computational load is required in order to achieve real time recovery. MATLAB®and LabVIEW® environments are used for implementing the various digital operations and provide an easy and flexible research platform for further experimentations, as discussed in the next sections. The sampling matrix $\mathbf{A}$ is computed once off-line, using the calibration process outlined in [25].

3.2 Support Recovery

The prototype is fed with RF signals composed of up to 5 carrier transmissions with an unknown total bandwidth occupancy of up to 200 MHz, and Nyquist rate of 6 GHz. An RF input $x(t)$ is generated using vector signal generators (VSG), each producing one modulated data channel with individual bandwidth of up to 20 MHz. The input transmissions then go through an RF combiner, resulting in a dynamic multiband input signal. This allows to test the system's ability to rapidly sense the input spectrum and adapt to changes, as required by modern CR standards (e.g. IEEE 802.22). In addition, the described setup is able to simulate more complex scenarios, including collaborative spectrum sensing [30, 31], joint DOA estimation [28], cyclostationary based detection [32] and various modulation schemes such as PSK, OFDM and more, for verifying sub-Nyquist data reconstruction capabilities.

Support recovery is digitally performed on the low rate samples, as presented above in the context of multicoset sampling. The prototype successfully recovers the support of the transmitted bands transmitted, when SNR levels are above 15 dB, as demonstrated in Fig. 14. Additional simulations presenting different input scenarios can be found in [2]. More sophisticated detection schemes, such as cyclostationary detection, allow to achieve perfect support recovery from the same sub-Nyquist samples in lower SNR regimes of $0-10$ dB, as seen in Figs. 24 and 25, and will be further discussed in Section 4.2.

The main advantage of the MWC is that sensing is performed in real-time for the entire spectral range, even though the operation is performed solely on sub-Nyquist samples, which results in substantial savings in both computational and memory complexity. In additional tests, it is shown that the bandwidth occupied in each band can also be very low without impeding the performance, as seen in Fig. 15, where the support of signals with very low bandwidth (just 10% occupancy within the 20 MHz band) is correctly detected.

3.3 Signal Reconstruction

Once the support is recovered, the data is reconstructed from the sub-Nyquist samples. Reconstruction is performed by inverting the reduced sampling matrix $\mathbf{A}_S$ in the recovered support, applying (11). This step is performed in real-time, reconstructing the signal bands $\mathbf{z}[n]$ one sample at a time, with low complexity due to the small dimensions of the matrix-vector multiplication. We note that reconstruction does not require interpolation to the Nyquist grid. The active transmissions are recovered at the low rate of 20 MHz, corresponding to the bandwidth of the slices $\mathbf{z}(f)$.

The prototype's digital recovery stage is further expanded to support decoding of common communication modulations, including BPSK, QPSK, QAM and OFDM. An example for the decoding of three QPSK modulated bands is given in Fig. 16, where the I/Q constellations are shown after reconstructing the original transmitted

Fig. 14 Screen-shot from the MWC recovery software: low rate samples acquired from one MWC channel at rate 120 MHz (top), digital reconstruction of the entire spectrum, performed from sub-Nyquist samples (middle), true input signal $x(t)$ showed using a fast spectrum analyzer (bottom).

signals $\mathbf{x}_S$ (11), from their low-rate and aliased sampled signals $\mathbf{z}_n$ (8). The I/Q constellations of the baseband signals is displayed, each individually decoded using a general QPSK decoder. In this example, the user broadcasts text strings, that are then deciphered and displayed on screen.

There are no restrictions regarding the modulation type, bandwidth or other parameters, since the baseband information is exactly reconstructed regardless of its respective content. Therefore, any digital modulation method, as well as analog broadcasts, can be transmitted and deciphered without loss of information, by applying any desirable decoding scheme directly on the sub-Nyquist samples.

By combining both spectrum sensing and signal reconstruction, the MWC prototype serves as two separate communication devices. The first is a state-of-the-art CR that can perform real time spectrum sensing at sub-Nyquist rates, and the second is a unique receiver able to decode multiple data transmissions simultaneously, regardless of their carrier frequencies, while adapting to temporal spectral changes in real time. In cases where the support of the potential active transmissions is a priori known (e.g. potential cellular carriers), the MWC may be used as an RF demodulator that efficiently acquires several frequency bands simultaneously. Other schemes would require a dedicated demodulation channel for each potentially active band. In this case, the mixing sequences should be designed so that their Fourier coefficients

Fig. 15 The setup is identical to Fig. 14. In this case, the individual transmissions have low bandwidth, highlighting the structure of the signal when folding to baseband.

are non zero only in the bands of interest, increasing SNR, and the support recovery stage is not needed [33].

4 Statistics Detection

In the previous sections, we reviewed recent sub-Nyquist sampling methods that reconstruct a multiband signal, such as a CR signal, from low rate samples. However, the final goal of CRs often only requires detection of the presence or absence of the PUs' transmissions and not necessarily their perfect reconstruction. In this case, several works have proposed performing detection on second-order signal statistics, which share the same frequency support as the original signal. In particular, power and cyclic spectra have been considered for stationary and cyclostationary [13] signals, respectively. Instead of recovering the signal from the low rate samples, its statistics are reconstructed and the support is estimated [34, 35, 36, 37, 38, 39, 40, 32].

Recovering second-order statistics rather than the signal itself benefits from two main advantages. First, it allows to further reduce the sampling rate, as we will dis-

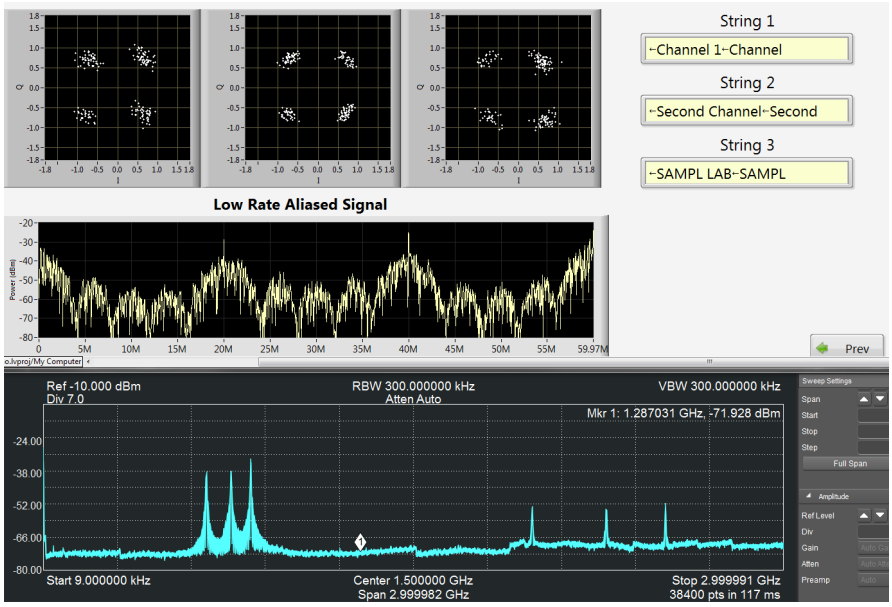


Fig. 16 Demodulation, reconstruction and detection of $N_{\text{sig}} = 3$ inputs from sub-Nyquist samples using the MWC CR prototype. At the bottom, the signal is sampled by an external spectrum analyzer showing the entire bandwidth of 3 GHz. Sub-Nyquist samples from an MWC channel $z_i[n]$ in the Fourier domain are displayed in the middle. The I/Q phase diagrams, showing the modulation pattern of the transmitted bands after reconstruction from the low rate samples, are presented at the top left. In the upper right corner, we see the information that was sent on each carrier, proving successful reconstruction.

cuss in the remainder of this section. Intuitively, statistics have fewer degrees of freedom than the signal itself, requiring less samples for their reconstruction. This follows from the assumption that the signal of interest is either stationary or cyclostationary. Going one step further, the sparsity constraint can even be removed in this case and the power/cyclic spectrum of non sparse signals is recoverable from samples obtained below the Nyquist rate [34, 37, 38, 40, 32]. This is useful for CRs operating in less sparse environments, in which the lower bound of twice the Landau rate may exceed the Nyquist rate. Second, the robustness to noise is increased due to the averaging performed to estimate statistics. This is drastically improved in the case of cyclostationary signals in the presence of stationary noise. Indeed, exploiting cyclostationarity properties exhibited by communication signals allows to separate them from stationary noise, leading to better detection in low SNR regimes [41]. In this section, we first review power spectrum detection techniques in stationary settings and then extend these to cyclic spectrum detection of cyclostationary signals.

4.1 Power Spectrum Based Detection

In the statistical setting, the signal $x(t)$ is modeled as the sum of uncorrelated wide-sense stationary transmissions. The stationarity assumption is key to further reducing the sampling rate. In frequency, stationarity is expressed by the absence of correlation between distinct frequency components. Specifically, as shown in [42], the Fourier transform of a wide-sense stationary signal is a nonstationary white process, such that

$$\mathbb{E}[X(f_1)X^*(f_2)] = S_x(f_1)\delta(f_1 - f_2). \tag{19}$$

Here, the power spectrum $S_x(f)$ of $x(t)$ is the Fourier transform of its autocorrelation $r_x(\tau)$. Thus, obviously, the support of $S_x(f)$ is identical to that of $X(f)$. In addition, due to (19), the autocorrelation matrix of the N spectrum frequency slices of $x(t)$ comprising $\mathbf{x}(f)$ is diagonal, containing only N degrees of freedom, which allows sampling rate reduction.

Another intuitive interpretation to the reduced number of degrees of freedom in statistics recovery is given in the time domain. There, the autocorrelation of stationary signals $r_x(\tau) = \mathbb{E}[x(t)x(t-\tau)]$ is only a function of the time lags τ. The cardinality of the difference set, namely the set that contains the time lags, may be greater than that of its associated original set, up to the order of its square, for an appropriate choice of sampling times [35, 43]. When the sampling scheme is not tailored to power spectrum recovery, the sampling rate can be as low as the Landau rate [38], which constitutes a worst case scenario in terms of sampling rate. With appropriate design, the autocorrelation or power spectrum may be estimated from samples with arbitrarily low average sampling rate [34, 35, 43, 44, 45] at the expense of increased latency.

We first review power spectrum recovery techniques that do not exploit any specific design. We then present methods that further reduce the sampling rate by adapting the sampling scheme to the purpose of autocorrelation or power spectrum estimation. Finally, we extend these results to the cyclostationary model.

4.1.1 Power Spectrum Recovery

In this section, we first focus on sampling with generic MWC or multicoset schemes without specific design of the mixing sequences or cosets, respectively.

To recover $S_x(f)$ from the low rates samples $\mathbf{z}(f)$, consider the correlation matrix of the latter $\mathbf{R}_\mathbf{z}(f) = \mathbb{E}[\mathbf{z}(f)\mathbf{z}^H(f)]$ [38]. Using (7), $\mathbf{R}_\mathbf{z}(f)$ can be related to correlations between the slices $\mathbf{x}(f)$, that is $\mathbf{R}_\mathbf{x}(f) = \mathbb{E}[\mathbf{x}(f)\mathbf{x}^H(f)]$ as follows

$$\mathbf{R}_\mathbf{z}(f) = \mathbf{A}\mathbf{R}_\mathbf{x}(f)\mathbf{A}^H, \qquad f \in \mathscr{F}_s. \tag{20}$$

From (19), the correlation matrix $\mathbf{R}_\mathbf{x}(f)$ is diagonal and contains the power spectrum $S_x(f)$ at the corresponding frequencies, as

$$\mathbf{R}_{\mathbf{x}_{(i,i)}}(f) = S_x(f + if_s - \frac{f_{\text{Nyq}}}{2}), \qquad f \in \mathscr{F}_s. \tag{21}$$

Recovering the power spectrum $S_x(f)$ is thus equivalent to recovering the matrix $\mathbf{R}_\mathbf{x}(f)$. Exploiting the fact that $\mathbf{R}_\mathbf{x}(f)$ is diagonal and denoting by $\mathbf{r}_\mathbf{x}(f)$ its diagonal, (20) can be reduced to

$$\mathbf{r}_\mathbf{z}(f) = (\bar{\mathbf{A}} \odot \mathbf{A})\mathbf{r}_\mathbf{x}(f), \tag{22}$$

where $\mathbf{r}_\mathbf{z}(f) = \text{vec}(\mathbf{R}_\mathbf{z}(f))$ concatenates the columns of $\mathbf{R}_\mathbf{z}(f)$. The matrix $\bar{\mathbf{A}}$ is the conjugate of $\mathbf{A}$ and $\odot$ denotes the Khatri-Rao product [46].

Generic choices of the sampling parameters, either mixing sequences or cosets, which are only required to ensure that $\mathbf{A}$ is full spark, are investigated in [38]. Then, the Khatri-Rao product $(\bar{\mathbf{A}} \odot \mathbf{A})$ is full spark as well if $M > N/2$, that is the number of rows of $\mathbf{A}$ is at least half the number of slices N. The minimal sampling rate to recover $\mathbf{r}_\mathbf{x}(f)$, and consequently $S_x(f)$, from $\mathbf{r}_\mathbf{z}(f)$ in (22) is thus equal to the Landau rate KB, namely half the rate required for signal recovery [38]. The recovery of $\mathbf{r}_\mathbf{x}(f)$ is performed using the procedure presented in the context of signal recovery on (22), that is CTF, support recovery and power spectrum reconstruction (rather than signal reconstruction).

The same result for the minimal sampling rate is valid for non sparse signals, for which KB is in the order of f_{Nyq} [38]. The power spectrum of such signals may be recovered at half their Nyquist rate. This means that even without any sparsity constraints on the signal in crowded environments, a CR can retrieve the power spectrum of the received signal by exploiting the stationarity property of the latter. In this case, the system (22) is overdetermined and $\mathbf{r}_\mathbf{x}(f)$ is obtained by a simple pseudo-inverse operation.

Obviously, in practice, we do not have access to $\mathbf{R}_\mathbf{z}(f)$, which thus needs to be estimated. The overall sensing time is divided into N_f frames of length N_s samples. In [38], different choices of N_f and N_s are examined for a fixed sensing time. In order to estimate the autocorrelation matrix $\mathbf{R}_\mathbf{z}(f)$ in the frequency domain, we first compute the estimates of $\mathbf{z}_i(f), 1 \le i \le M$, denoted by $\hat{\mathbf{z}}_i(f)$, using the fast Fourier transform (FFT) on the samples $z_i[n]$ over a finite time window. We then estimate the elements of $\mathbf{R}_\mathbf{z}(f)$ as

$$\hat{\mathbf{R}}_\mathbf{z}(i,j,f) = \frac{1}{N_f} \sum_{\ell=1}^{N_f} \hat{\mathbf{z}}^\ell(i,f)\hat{\mathbf{z}}^\ell(j,f), \quad f \in \mathscr{F}_s, \tag{23}$$

where $\hat{\mathbf{z}}^\ell(i,f)$ is the value of the FFT of the samples $\mathbf{z}_i[n]$ from the ℓth frame, at frequency f. In practice, the number of samples dictates the number of FFT coefficients in the frequency domain and therefore the resolution of the reconstructed power spectrum.

Once $\hat{\mathbf{r}}_\mathbf{x}(f)$ is reconstructed, the following test statistic,

$$\Gamma_i = \sum_{f \in \mathscr{F}_s} |\hat{\mathbf{r}}_{\mathbf{x}_i}(f)|^2, \qquad 1 \le i \le N, \tag{24}$$

may be adopted in order to detect the occupied support. Here, $\hat{\mathbf{r}}_{\mathbf{x}_i}(f)$ is the ith entry of $\hat{\mathbf{r}}_{\mathbf{x}}(f)$ and the sum is performed over the frequency band of interest to detect the presence of a PU. Alternatively, other detection statistics can be used on the reconstructed power spectrum, such as eigenvalue based test statistics [47].

4.1.2 Power Spectrum Sensing: Tailored Design

Sampling approaches specifically designed for estimating the autocorrelation of stationary signals at much finer lags than the sample spacings have been studied recently in detail [35, 48, 43, 44]. The key observation here is that the autocorrelation is a function of the lags only, namely the differences between pairs of sample times. Thus, it is estimated at all time lags contained in the difference co-array, composed of all the differences between pairs of elements from the original sampling array. Since the size of the difference co-array may be greater than that of the sampling set, it is possible to sample below the Nyquist rate and estimate the correlation at all lags on the Nyquist grid, from the low rate samples. Therefore, the sampling times should be carefully chosen so as to maximize the cardinality of the difference co-array.

The first approach we present adopts multicoset sampling previously reviewed, while specifically designing the cosets to obtain a maximal number of differences. In the previous section, the results were derived for any coset selection. Here, we show that the sampling rate may be lower if the cosets are carefully chosen. When using multicoset sampling, the sampling matrix $\mathbf{A}$ in (20) or (22), is a partial Fourier matrix with (i,k)th element $\mathrm{e}^{j\frac{2\pi}{N}c_i k}$. A typical element of $(\bar{\mathbf{A}} \odot \mathbf{A})$ is then $\mathrm{e}^{j\frac{2\pi}{N}(c_i - c_j)k}$. If all cosets are distinct, then the size of the difference set over one period is greater than or equal to $2M-1$. This bound corresponds to a worst case scenario, as discussed in the previous section and leads to a sampling rate of at least half Nyquist in the non sparse setting and at least Landau for a sparse signal with unknown support. This happens for example is we select the first or last M cosets or if we keep only the even or odd cosets.

To maximize the size of the difference set and increase the rank of $(\bar{\mathbf{A}} \odot \mathbf{A})$, the cosets can be chosen [35, 48] using minimal linear and circular sparse rulers [49]. A linear sparse ruler is a set of integers from the interval $[0,N]$, such that the associated difference sets contains all integers in $[0,N]$. Intuitively, it can be seen as a ruler with some marks erased but still able to measure all integer distances between 0 and its length. For example, consider the minimal sparse ruler of length $N = 10$. This ruler requires $M = 6$ marks, as shown in Fig. 17. Obviously, all the lags $0 \leq \tau \leq 10$ on the integer grid are identifiable. There is no closed form expression for the maximum compression ratio M/N that is achievable using a sparse ruler; however, the following bounds hold

$$\frac{\sqrt{\tau(N-1)}}{N} \leq \frac{M}{N} \leq \frac{\sqrt{3(N-1)}}{N}, \tag{25}$$

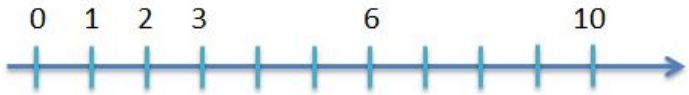


Fig. 17 Minimal sparse ruler of order $M = 6$ and length $N = 10$.

where $\tau \approx 2.4345$ [48]. A circular or modular sparse ruler extends this idea to include periodicity. Such designs that seek minimal sparse rulers, that is rulers with minimal number of marks M, allow to achieve compression ratios M/N on the order of $\sqrt{N}$. As N increases, the compression ratio may be arbitrarily low.

Two additional sampling techniques specifically designed for autocorrelation recovery are nested arrays [43] and co-prime sampling [44], presented in the context of autocorrelation estimation as well as beamforming and DOA estimation applications. In nested and co-prime structures, similarly to multicoset, the corresponding co-arrays have more degrees of freedom than those of the original arrays, leading to a finer grid for the time lags with respect to the sampling times. We now briefly review both sampling structures and their corresponding difference co-arrays and show how the autocorrelation of an arbitrary stationary signal can be recovered on the Nyquist grid from these low rate samples.

In its simplest form, the nested array [43] structure has two levels of sampling density. The first level samples are at the N_1 locations $\{\ell T_{\text{Nyq}}\}_{1\le\ell\le N_1}$ and the second level samples are at the N_2 locations $\{(N_1+1)kT_{\text{Nyq}}\}_{1\le k\le N_2}$. This nonuniform sampling is then repeated with period $(N_1+1)N_2T_{\text{Nyq}}$. Since there are N_1+N_2 samples in intervals of length $(N_1+1)N_2T_{\text{Nyq}}$, the average sampling rate of a nested array sampling set is given by

$$f_s = \frac{N_1+N_2}{(N_1+1)N_2T_{\text{Nyq}}} \equiv \frac{1}{N_1T_{\text{Nyq}}} + \frac{1}{N_2T_{\text{Nyq}}}, \tag{26}$$

which can be arbitrarily low since N_1 and N_2 may be as large as we choose, at the expense of latency.

Now, consider the difference co-array which has contribution from the cross-differences and the self-differences. The non negative cross-differences, normalized by T_{Nyq} for clarity, are given by

$$n = (N_1+1)k-\ell, \quad 1\le k\le N_2, 1\le \ell\le N_1. \tag{27}$$

All differences in the range $1 \le n \le (N_1+1)N_2-1$ are covered, except for the multiples of N_1+1. These are are precisely the self differences among the second array. As a result, the difference co-array is a filled array represented by the set of all integers $-[(N_1+1)N_2-1] \le n \le [(N_1+1)N_2-1]$. Going back to our autocorrelation or power spectrum estimation problem, this result shows that by proper averaging,

we can estimate $R(\tau)$ at any lag τ on the Nyquist grid for any stationary signal from the nested array samples, with arbitrarily low sampling rate.

Co-prime sampling involves two uniform sampling sets with spacing $N_1 T_{\text{Nyq}}$ and $N_2 T_{\text{Nyq}}$ respectively, where N_1 and N_2 are co-prime integers. Therefore, the average sampling rate of such a sampling set, given by

$$f_s = \frac{1}{N_1 T_{\text{Nyq}}} + \frac{1}{N_2 T_{\text{Nyq}}}, \tag{28}$$

can be made arbitrarily small compared to the Nyquist rate $1/T_{\text{Nyq}}$.

The associated difference set normalized by T_{Nyq} is composed of elements of the form $n = N_1 k - N_2 \ell$. Since N_1 and N_2 are co-prime, there exist integers k and ℓ such that the above difference achieves any integer value n. Therefore, the autocorrelation can be estimated by proper averaging, as

$$\hat{R}[n] = \frac{1}{Q} \sum_{q=0}^{Q-1} x(N_1(k+N_2 q)) x^*(N_2(\ell + N_1 q)), \tag{29}$$

where Q is the number of snapshots used for averaging. Again, the autocorrelation of any stationary signal may be estimated over the Nyquist grid from samples with arbitrarily low rate, and without any sparsity constraint.

The main drawback of both techniques, besides the practical issue of analog bandwidth and channel synchronization similarly to multicoset sampling, is the added latency required for sufficient averaging. In addition, nested array sampling still requires one sampler operating at the Nyquist rate. Thus, there is no saving in terms of hardware, but only in memory and computation.

4.2 Cyclostationary Detection

Communication signals typically exhibit statistical periodicity, due to modulation schemes such as carrier modulation or periodic keying [50]. Therefore, such signals are better modeled as cyclostationary rather than stationary processes. A characteristic function of such processes, the cyclic spectrum $S_x^{\alpha}(f)$, extends the traditional power spectrum to a two dimensional map, with respect to two frequency variables, angular and cyclic. The cyclic spectrum exhibits spectral peaks at certain frequency locations, the cyclic frequencies, which are determined by the signal's parameters, particularly the carrier frequency and symbol rate [41]. This constitutes the main advantage of cyclostationary detection. Stationary noise and interference exhibit no spectral correlation [41], as shown in (19), rendering such detectors highly robust to noise. Compressive power spectrum recovery techniques have been extended to reconstruction of the cyclic spectrum from the same compressive measurements. In this section, we first provide some general background on cyclostationarity and then review sub-Nyquist cyclostationary detection approaches.

4.2.1 Cyclostationarity

A process $x(t)$ is said to be wide-sense cyclostationary with period T_0 if its mean $\mu_x(t) = \mathbb{E}[x(t)]$ and autocorrelation $R_s(t,\tau) = \mathbb{E}[x(t)x(t+\tau)]$ are both periodic with period T_0 [13], that is

$$\mu_x(t+T_0) = \mu_x(t), \qquad R_x(t+T_0,\tau) = R_x(t,\tau), \tag{30}$$

for all $t \in \mathbb{R}$. Given a wide-sense cyclostationary random process, its autocorrelation $R_x(t,\tau)$ can be expanded in a Fourier series

$$R_x(t,\tau) = \sum_{\alpha} R_x^{\alpha}(\tau) \mathrm{e}^{j2\pi\alpha t}, \tag{31}$$

where the sum is over integer multiples of the fundamental frequency $1/T_0$ and the Fourier coefficients, referred to as cyclic autocorrelation functions, are given by

$$R_x^{\alpha}(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} R_x(t,\tau) \mathrm{e}^{-j2\pi\alpha t} \mathrm{d}t. \tag{32}$$

The cyclic spectrum is the Fourier transform of (32) with respect to τ, namely

$$S_x^{\alpha}(f) = \int_{-\infty}^{\infty} R_x^{\alpha}(\tau) \mathrm{e}^{-j2\pi f\tau} \mathrm{d}\tau, \tag{33}$$

where α is referred to as the cyclic frequency and f is the angular frequency [13]. If there is more than one fundamental frequency $1/T_0$, then the process $x(t)$ is said to be poly-cyclostationary in the wide sense. In this case, the cyclic spectrum contains harmonics (integer multiples) of each of the fundamental cyclic frequencies [41]. These cyclic frequencies are governed by the transmissions' carrier frequencies and symbol rates as well as modulation types.

An alternative and more intuitive interpretation of the cyclic spectrum expresses it as the cross-spectral density $S_x^{\alpha}(f) = S_{uv}(f)$ of two frequency-shifted versions of $x(t)$, $u(t)$ and $v(t)$, such that

$$u(t) \triangleq x(t)\mathrm{e}^{-j\pi\alpha t}, \quad v(t) \triangleq x(t)\mathrm{e}^{+j\pi\alpha t}. \tag{34}$$

Then, from [42], it holds that

$$S_x^{\alpha}(f) = S_{uv}(f) = \mathbb{E}\left[X\left(f+\frac{\alpha}{2}\right)X^*\left(f-\frac{\alpha}{2}\right)\right]. \tag{35}$$

Thus, the cyclic spectrum $S_x^{\alpha}(f)$ measures correlations between different spectral components of $x(t)$. Stationary signals, which do not exhibit spectral correlation between distinct frequency components, appear only at $\alpha = 0$. This property is the key to robust detection of cyclostationary signals in the presence of stationary noise, in low SNR regimes.

The support region in the (f, α) plane of the cyclic spectrum of a bandpass cyclostationary signal is composed of four diamonds, as shown in Fig. 18. Therefore, the cyclic spectrum $S_x^\alpha(f)$ of a multiband signal with K uncorrelated transmissions is supported over $4K$ diamond-shaped areas. Figure 19 illustrates the cyclic spectrum of two modulation types, AM and BPSK.

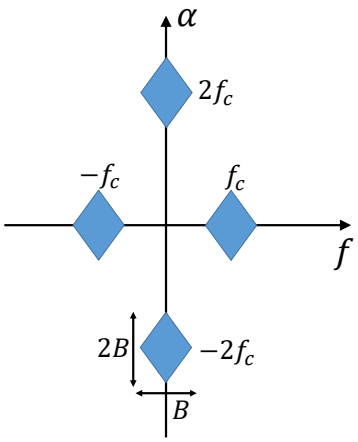


Fig. 18 Support region of the cyclic spectrum of a bandpass cyclostationary signal with carrier frequency f_c and bandwidth B.

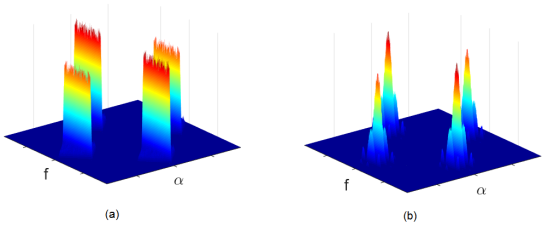


Fig. 19 Cyclic spectrum magnitude of signals with (a) AM and (b) BPSK modulations.

4.2.2 Cyclic Spectrum Recovery

In the previous section, we showed how the power spectrum $S_x(f)$ can be reconstructed from correlations $\mathbf{R_z}(f)$ between the samples obtained using the MWC or multicoset sampling. To that end, we first related $S_x(f)$ to the slices' correla-

tion matrix $\mathbf{R}_{\mathbf{x}}(f)$ and then recovered the latter from $\mathbf{R}_{\mathbf{z}}(f)$. Here, this approach is extended to the cyclic spectrum $S_x^{\alpha}(f)$. We first show how it is related to shifted correlations between the slices, namely $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f}) = \mathbb{E}\left[\mathbf{x}(\tilde{f})\mathbf{x}^{H}(\tilde{f}+a)\right]$, for $a \in [0, f_s]$ and $\tilde{f} \in [0, f_s - a]$. Next, similarly to power spectrum recovery, $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})$ is reconstructed from shifted correlations of the samples $\mathbf{R}_{\mathbf{z}}^{a}(\tilde{f}) = \mathbb{E}\left[\mathbf{z}(\tilde{f})\mathbf{z}^{H}(\tilde{f}+a)\right]$. Once the cyclic spectrum $S_x^{\alpha}(f)$ is recovered, we estimate the transmissions' carriers and bandwidth by locating its peaks. Since the cyclic spectrum of stationary noise $n(t)$ is zero for $\alpha \neq 0$, cyclostationary detection is more robust to noise than stationary detection.

The alternative definition of the cyclic spectrum (35), implies that the elements in the matrix $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})$ are equal to $S_x^{\alpha}(f)$ at the corresponding α and f. Indeed, it is easily shown [32] that

$$\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})_{(i,j)} = S_x^{\alpha}(f), \tag{36}$$

for

$$\begin{aligned} \alpha &= (j-i)f_s + a \\ f &= -\frac{f_{\mathrm{Nyq}}}{2} + \tilde{f} - \frac{f_s}{2} + \frac{(j+i)f_s}{2} + \frac{a}{2}. \end{aligned} \tag{37}$$

Here $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})_{(i,j)}$ denotes the (i,j)th element of $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})$. This means that each entry of the cyclic spectrum $S_x^{\alpha}(f)$ can be mapped to an element from one of the correlation matrices $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})$, and vice versa. Using (7) and similarly to (20) in the context of power spectrum recovery, we relate the shifted correlations matrices of $\mathbf{x}(f)$ and $\mathbf{z}(f)$ as

$$\mathbf{R}_{\mathbf{z}}^{a}(\tilde{f}) = \mathbf{A}\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})\mathbf{A}^{H}, \quad \tilde{f} \in [0, f_s - a], \tag{38}$$

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

Recall that, in the context of stationary signals, $\mathbf{R}_{\mathbf{x}}(f)$ is diagonal. Here, understanding the structure of $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})$ is more involved. It was shown [32] that $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})$ contains non zero elements on its 0, 1 and -1 diagonals and anti-diagonals. Besides the non zero entries being contained only in the three main and anti-diagonals, additional structure is exhibited, limiting to two the number of non zero elements per row and column of the matrix $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})$. The above pattern follows from two main considerations. First, each frequency component, namely each entry of $\mathbf{x}(f)$, is correlated to at most two frequencies from the shifted vector of slices $\mathbf{x}(\tilde{f}+a)$, one from the same frequency band and one from the symmetric band. Second, the correlated component can be either in the same/symmetric slice or in one of the adjacent slices.

Figures 21 and 22 illustrate these correlations for $a = 0$ and $a = f_s/2$, respectively. First, in Fig. 20, an illustration of the spectrum of $x(t)$, namely $X(f)$, is presented for the case of a sparse signal buried in stationary noise. It can be seen that frequency bands of $X(f)$ either appear in one f_p-slice or are split between two slices at most since $f_p \geq B$. The resulting vector of spectrum slices $\mathbf{x}(f)$ and the correlations between these slices without any shift, namely $\mathbf{R}_{\mathbf{x}}^{0}(\tilde{f})$, are shown in Figs. 21(a) and (b), respectively. In Fig. 21(b), we observe that self-correlations appear only

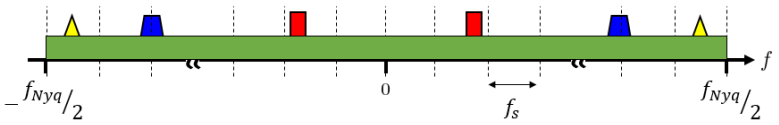


Fig. 20 Original spectrum $X(f)$. The cyclostationary transmissions are shown as a triangle, trapezoid and rectangle shaped spectral components, buried in a flat stationary noise.

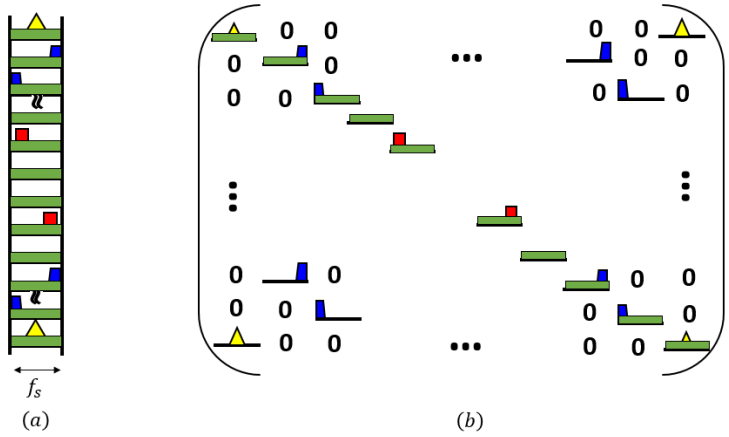


Fig. 21 (a) Spectrum slices vector $\mathbf{x}(\tilde{f})$, (b) correlated slices of $\mathbf{x}(\tilde{f})$ in the matrix $\mathbf{R}_{\mathbf{x}}^{0}(\tilde{f})$.

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 trapezes are contained in the -1 and $+1$ anti diagonals. The red rectangles do not contribute any cross-correlations for $a = 0$.

Figures 22(a) and (b) show the vector $x(\tilde{f})$ and its shifted version $x(\tilde{f}+a)$ for $a = f_s/2$, respectively. The resulting correlation matrix $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})$ appears in Fig. 22(c). Here, the self correlations of the triangle shaped frequency component appear in the main diagonal and that of the trapezoid shaped component in the -1 diagonal. The cross-correlations all appear in the anti-diagonal for the shift $a = f_s/2$. Note that since the noise is assumed to be wide-sense stationary, from (19), a noise frequency component is correlated only with itself. Thus, $n(t)$ contributes non-zero elements only on the diagonal of $\mathbf{R}_{\mathbf{x}}^{0}(\tilde{f})$.

To recover $\mathbf{R}_{\mathbf{x}}^{a}(\tilde{f})$ from $\mathbf{R}_{\mathbf{z}}^{a}(\tilde{f})$, structured CS techniques are used in [32] that aim at reconstructing a sparse matrix while taking into account its specific structure, as described above. Once the cyclic spectrum is reconstructed, the number of transmissions and their respective carrier frequencies and bandwidths are estimated, as discussed in the next section. The detection performed on the cyclic spectrum is more robust to stationary noise than power spectrum based detection, at the expense

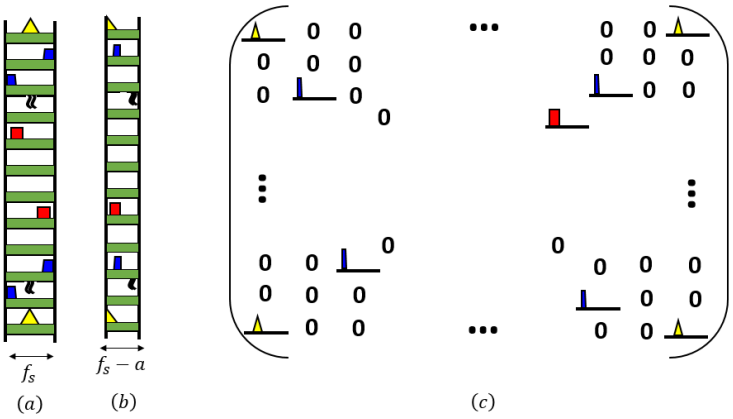


Fig. 22 (a) Spectrum slices vector $\mathbf{x}(\tilde{f})$, (b) spectrum slices shifted vector $\mathbf{x}(\tilde{f}+a)$, for $a = f_s/2$, (c) correlated slices of $\mathbf{x}(\tilde{f})$ and $\mathbf{x}(\tilde{f}+a)$ in the matrix $\mathbf{R}^a_{\mathbf{x}}(\tilde{f})$, with $a = f_s/2$.

of a slightly higher sampling rate, as shown in [32]. More precisely, in the presence of stationary noise, the cyclic spectrum may 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 minimal sampling rate is further reduced to $8/5$ of the Landau rate [32].

4.2.3 Carrier frequency and bandwidth estimation

Many detection and classification algorithms based on cyclostationarity have been proposed (see reviews [13, 14]). To assess the presence or absence of a signal, a first group of techniques requires a priori knowledge of its parameters and particularly of the carrier frequency, which is the information that CRs should uncover in the first place. A second strategy focuses on a single transmission, which does not fit the multiband model. Alternative approaches apply machine learning tools to the modulation classification of a single signal with unknown carrier frequency and symbol rate. Besides being only suitable for a single transmission, these methods require a training phase, which would be a main drawback for CR purposes. In particular, these techniques can only cope with PUs whose modulation type and parameters were part of the training set.

For CR purposes, we need a detector designed to comply with certain requirements: (1) carrier frequency and bandwidth estimation rather than simple detection of the presence or absence of a signal; (2) blind detection, namely without knowledge of the carrier frequencies, bandwidths and symbol rates of the transmissions; (3) simultaneous detection of several transmissions; (4) non-learning approach, i.e. with no training phase. The parameter estimation algorithm, presented in [51] is a

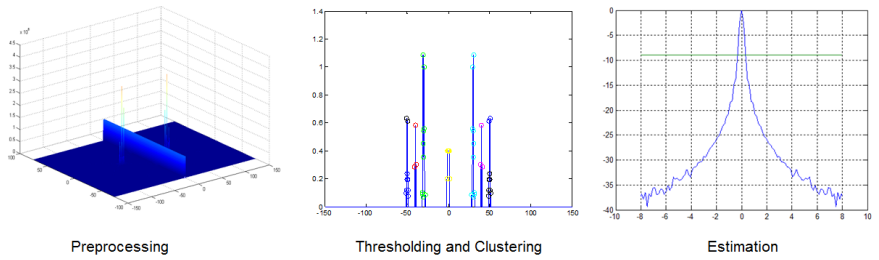


Fig. 23 Carrier frequency and bandwidth estimation from the cyclic spectrum: preprocessing (left), thresholding and clustering (middle), parameter estimation (right).

simple parameter extraction method from the cyclic spectrum of multiband signals, that answers these requirements. It allows the estimation of several carriers and several bandwidths simultaneously, as well as the number of transmissions, namely half the number of occupied bands $K/2$ for real-valued signals. The proposed parameter estimation algorithm can be decomposed into the following four main steps: preprocessing, thresholding, clustering and parameter estimation, as illustrated in Fig. 23.

The preprocessing simply aims at compensating for the presence of stationary noise in the cyclic spectrum at the cyclic frequency $\alpha = 0$, by attenuating the energy of the cyclic spectrum at this frequency. Thresholding is then applied to the resulting cyclic spectrum in order to find its peaks. The locations and values of the selected peaks are then clustered using k-means to find the corresponding cyclic feature, after estimating the number of clusters by applying the elbow method [52]. It follows that, apart from the cluster present in DC, the number of real signals, namely $N_{\text{sig}} = K/2$, is equal to half the number of clusters. Next, the carrier frequency f_i, which corresponds to the highest peak [41], is estimated for each transmission. The bandwidth B_i is found by locating the edge of the support of the angular frequencies at the corresponding cyclic frequency $\alpha_i = 2f_i$.

Results presented in [32] demonstrate that cyclostationary based detection, as described in this section, outperforms energy detection carried on the signal's spectrum or power spectrum, at the expense of increased complexity. We now show similar results obtained from hardware simulations, performed using the prototype from Fig. 12.

4.3 Hardware Simulations: Robustness to Noise

Cyclostationary detection has been implemented in the MWC CR prototype. The analog front-end is identical to that of the original prototype and only the digital recovery part is modified since the cyclic spectrum is recovered directly from the MWC low rate samples. Preliminary testing suggests that sensing success is achievable at SNRs lower by 10 dB than those allowed by energy detection performed on

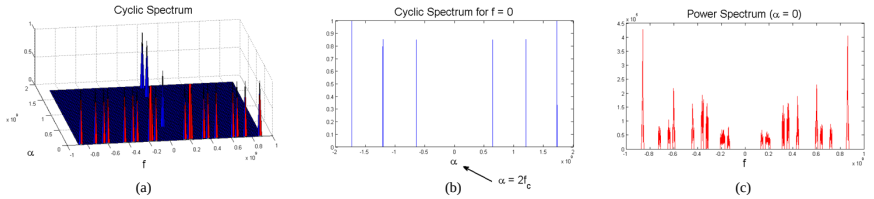


Fig. 24 Screen shot from the MWC with cyclostationary detection. The input signal is composed of $N_{\text{sig}} = 3$ transmissions (or $K = 6$ bands) with carriers $f_1 = 320$MHz, $f_2 = 600$MHz and $f_3 = 860$MHz. (a) The recovered cyclic spectrum from low rate samples. (b) The cyclic spectrum profile at the angular frequency $f = 0$; the cyclic peaks are clearly visible at twice the carrier frequencies. (c) The power spectrum recovery is displayed and shown to fail in the presence of noise.

the recovered spectrum or power spectrum. Representative results shown in Figs. 24 and 25 demonstrate the advantage of cyclostationary detection over energy detection in the presence of noise. The figures show the reconstructed cyclic spectrum from samples of the MWC prototype, as well as cross-sections at $f = 0$ and $\alpha = 0$, which corresponds to the power spectrum. This increased robustness to noise comes at the expense of more complex digital processing on the low rate samples, stemming from the higher dimensionality involved, since we reconstruct the 2-dimensional cyclic spectrum rather than the 1-dimensional (power) spectrum.

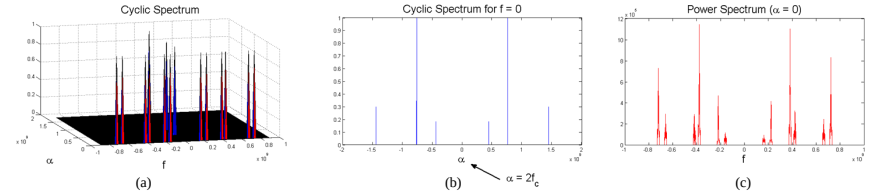


Fig. 25 The setup is identical to Fig. 24, with carrier frequencies $f_1 = 220$MHz, $f_2 = 380$MHz and $f_3 = 720$MHz.

5 Collaborative Spectrum Sensing

5.1 Collaborative Model

Until now, we assumed direct observation of the spectrum. In practice, the task of spectrum sensing for CR is further complicated due to physical channel effects such as path loss, fading and shadowing [15]. To overcome these practical issues, collaborative CR networks have been considered, where different users share their sensing results and cooperatively decide on the licensed spectrum occupancy.

The different collaborative approaches can be distinguished according to several criteria [15]. First, cooperation can be either centralized or distributed. In centralized settings, the data is sent to a fusion center which combines the shared data to jointly estimate the spectrum or determine its occupancy. In the distributed approach, the CRs communicate among themselves and iteratively converge to a common estimate or decision. While centralized cooperation does not require iterations and can reach the optimal estimate based on the shared data, convergence to this estimate is not always guaranteed in its distributed counterpart. On the other hand, the latter is less power hungry and more robust to node and link failure, increasing the network survivability. An additional criterion concerns the shared data type; the CRs may share local binary decisions on the spectrum occupation (hard decision) or a portion of their samples (soft decision).

We consider the following collaborative model. A network of N_{rec} CRs receives the N_{sig} transmissions, such that the received signal at the jth CR is given by

$$x^{(j)}(t) = \sum_{i=1}^{N_{\text{sig}}} r_{ij}(t) = \sum_{i=1}^{N_{\text{sig}}} s_i(t) * h_{ij}(t). \tag{39}$$

The channel response $h_{ij}(t)$ is determined by fading and shadowing effects. Typical models are Rayleigh fading, or small-scale fading, and log-normal shadowing, or large-scale fading [16, 53, 54]. In the frequency domain, the Fourier transform of the jth received signal is given by

$$X^{(j)}(f) = \sum_{i=1}^{N_{\text{sig}}} S_i(f) H_{ij}(f). \tag{40}$$

Therefore, the support of $x^{(j)}(t)$ is included in the support of the original signal $x(t)$. Since the transmissions are affected differently by fading and shadowing from each transmitter to each CR, we can assume that the union of their respective supports is equivalent to the frequency support of $x(t)$. The goal here is therefore to assess the support of the transmitted signal $x(t)$ from sub-Nyquist samples of the received $x^{(j)}(t), 1 \le j \le N_{\text{rec}}$, by exploiting their joint frequency sparsity.

A simple and naive approach is to perform support recovery at each CR from its low rate samples and combine the local binary decisions, either in a fusion center for centralized collaboration or in a distributed manner. In this hard decision strategy, the combination can be performed using several fusion rules such as AND, OR or majority rule. Although this method is attractive due to its simplicity and low communication overhead, it typically achieves lower performance than its soft decision counterpart. To mitigate the communication overhead, soft decision based methods can rely on sharing observations based on the low rate samples with smaller dimensions, rather than the samples themselves. In the next section, we review such techniques both in centralized and distributed contexts.

5.2 Centralized Collaborative Support Recovery

One approach [55, 56] to centralized spectrum sensing considers a digital model based upon a linear relation between the M sub-Nyquist samples $\mathbf{z}^{(j)}$ at CR j and N Nyquist samples $\mathbf{x}^{(j)}$ obtained for a given sensing time frame, namely

$$\mathbf{z}^{(j)} = \mathbf{A}\mathbf{x}^{(j)}, \tag{41}$$

where $\mathbf{A}$ is the sampling matrix. In particular, the authors consider multicoset sampling where $\mathbf{z}$ selects certain samples from the Nyquist grid $\mathbf{x}$ and $\mathbf{A}$ is the corresponding selection matrix. The goal is to recover the power spectrum of the true signal $\mathbf{x}$, assumed to be stationary. To that end, the covariance matrices of sub-Nyquist and Nyquist samples are related by the following quadratic equation

$$\mathbf{R_z}^{(j)} = \mathbf{A}\mathbf{R_x}^{(j)}\mathbf{A}^H, \tag{42}$$

where $\mathbf{R_x}^{(j)}$ is diagonal. Each CR sends its autocorrelation matrix $\mathbf{R_z}^{(j)}$ to the fusion center. The common sparsity of the diagonal of $\mathbf{R_x}^{(j)}$ is then exploited in the frequency domain across all CRs to jointly reconstruct them at the fusion center, using a modified simultaneous orthogonal matching pursuit (SOMP) [21] algorithm.

In [55] only the autocorrelation $\mathbf{R_z}^{(j)}$ between the samples of each CR j, are considered. This approach is extended in [56] to include cross-correlations between measurements from different CRs,

$$\mathbf{R}_{\mathbf{z}^{(j)}\mathbf{z}^{(k)}} = \mathbb{E}[\mathbf{z}^{(j)}(\mathbf{z}^{(k)})^H], \tag{43}$$

where j and k are the indices of two CRs. Here, each CR sends its measurement vector $\mathbf{z}^{(j)}$ to the fusion center and the cross-correlations are then computed. The cross-correlations are related to the common power spectrum $\mathbf{s_x} = \mathbf{F}\mathbf{r_x}$ by

$$\mathbf{R}_{\mathbf{z}^{(j)}\mathbf{z}^{(j)}} = \mathbf{C}^{(j)}\mathbf{S_x}(\mathbf{C}^{(j)})^H, \tag{44}$$

where $\mathbf{S_x}$ is the diagonal matrix that contains the power spectrum vector $\mathbf{s_x}$ and $\mathbf{C}^{(j)} = \mathbf{A}^{(j)}\mathbf{F}^H\mathbf{H}^{(j)}$. The sampling matrix $\mathbf{A}^{(j)}$ can be different for each CR j, $\mathbf{F}$ is the $N \times N$ Fourier matrix and $\mathbf{H}^{(j)}$ is a diagonal matrix that contains the frequency channel state information (CSI). After vectorization, similarly to (42), the N^2 auto and cross-correlation measurements are concatenated and the goal is to recover $\mathbf{s_x}$, that is the diagonal of $\mathbf{S_x}$ from these. It is shown that if the total number of samples $N_{\text{rec}}M$ is greater than N and these are suitably chosen to account for enough measurement diversity, then the power spectrum $\mathbf{s_x}$ of a non sparse signal can be recovered from compressed samples from a sufficient number of CRs. This shows that the minimal rate per CR is lower by a factor of N_{rec} with respect to that required for an individual CR and the number of receivers may be traded for the number of samples per CR. However, increasing the number of samples per CR does not increase spatial diversity, as does increasing the number of receivers. A drawback

of this technique is that CSI is traditionally unknown by the CRs and should be estimated prior to detection.

An alternative approach [31] relies on the analog model from (7) and does not assume any a priori knowledge on the CSI. This method considers collaborative spectrum sensing from samples acquired via multicoset sampling or the MWC at each CR. In this approach, the jth CR shares its observation matrix $\mathbf{V}^{(j)}$, as defined in (10), rather than the sub-Nyquist samples themselves, and its measurement matrix $\mathbf{A}^{(j)}$, with a fusion center. The sampling matrices are considered to be different from one another in order to allow for more measurement diversity. However, the same known matrix may be used to reduce the communication overhead. The underlying matrices $\mathbf{U}^{(j)}$ are jointly sparse since fading and shadowing do not affect the original signal's support. Capitalizing on the joint support of $\mathbf{U}^{(j)}$, the support of the transmitted signal $x(t)$ can be recovered at the fusion center by solving

$$\begin{aligned} \arg\min_{\mathbf{U}^{(j)}} \quad & \bigcup_{i=1}^{N_{\text{rec}}} ||\mathbf{U}^{(j)}||_0 \\ \text{s.t.} \quad & \mathbf{V}^{(j)} = \mathbf{A}^{(j)}\mathbf{U}^{(j)}, \text{ for all } 1 \leq j \leq N_{\text{rec}}. \end{aligned} \tag{45}$$

To recover the joint support of $\mathbf{U}^{(j)}$ from the observation matrices $\mathbf{V}^{(j)}$, both the orthogonal matching pursuit (OMP) and iterative hard thresholding (IHT) algorithms, two popular CS techniques, are extended to the collaborative setting [31]. Previously we considered support recovery from an individual CR, which boils down to an MMV system of equations (10). CS algorithms have been extended to this case, such as SOMP from [57] and simultaneous IHT (SIHT) presented in [58]. These now need to account for the joint sparsity across the CRs.

The distributed CS-SOMP (DCS-SOMP) algorithm [59], which extends the original SOMP to allow for different sampling matrices $\mathbf{A}^{(j)}$ for each receiver is adapted to the CR collaborative setting [31]. The main modification appears in the computation of the index that accounts for the greatest amount of residual energy. Here, the selected index is the one that maximizes the sum of residual projections over all the receivers. Once the shared support is updated, the residual matrices can be computed for each CR separately. The resulting modified algorithm is referred to as block sparse OMP (BSOMP) [31]. The sparse IHT algorithm can also be extended to this setting by selecting the indices of the common support though averaging over all the estimated $\mathbf{U}^{(j)}$ in each iteration. Once the support is selected, the update calculations are performed separately for each receiver [31]. Both methods are suitable for centralized cooperation, in the presence of a fusion center. As in the previous approach, if the CSI is known, then the sampling rate per CR can be reduced by a factor of N_{rec} with respect to the rate required from an individual CR.

5.3 Distributed Collaborative Support Recovery

In the distributed approach, there is no fusion center and the CRs are restricted to communicate only with their neighbors. Both the digital and analog centralized approaches have been extended to the distributed settings. First, in [60, 61], the digital model (41) is used, and the low rate samples $\mathbf{z}^{(j)}$ of the jth CR are expressed with respect to the spectrum $\mathbf{w}^{(j)} = \mathbf{F}\mathbf{x}^{(j)}$, such that

$$\mathbf{z}^{(j)} = \mathbf{A}^{(j)}\mathbf{F}^H\mathbf{H}^{(j)}\mathbf{w}^{(j)}. \tag{46}$$

Both unknown and known CSI cases are considered. In the first case, each CR computes its local binary decision $\mathbf{c}^{(j)}$ for each spectral band by recovering the sparse spectrum using CS techniques and comparing the local spectrum estimate $\mathbf{w}^{(j)}$ to a chosen threshold. Then, an average consensus approach is adopted, with respect to the shared hard decision. Specifically, each node j broadcasts its current decision $\mathbf{c}^{(j)}(t)$ to its neighbors $\mathscr{N}^{(j)}$ and updates itself by adding a weighted sum of the local discrepancies, that is

$$\mathbf{c}^{(j)}(t+1) = \mathbf{c}^{(j)}(t) + \sum_{k\in\mathscr{N}^{(j)}} \alpha_{jk}(\mathbf{c}^{(k)}(t) - \mathbf{c}^{(j)}(t)), \tag{47}$$

where α_{jk} is a weight associated with the edge (j,k). If the CSI is known, then the joint spectrum itself and not only its support can be collaboratively recovered. Each CR iteratively solves an ℓ_1 optimization problem for the sparse spectrum $\mathbf{w}^{(j)}$ constrained to consent with one-hop neighbors. In [60], the proposed algorithm iterates through the following steps: local spectrum reconstruction given the support and consensus averaging on the spectrum estimate. In [61], a distributed augmented Lagrangian algorithm is adopted.

Another approach extends the the method presented in [31], based on the analog model (7), to comply with distributed settings [30]. The ith CR contacts a random neighbor j, chosen with some probability P_{ij}, according to the Metropolis-Hastings scheme for random transition probabilities,

$$P_{ij} = \begin{cases} \min\{\frac{1}{d_i}, \frac{1}{d_j}\} & (i,j) \in E, \\ \sum_{(i,k)\in E} \max\{0, \frac{1}{d_i} - \frac{1}{d_k}\} & i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{48}$$

Here d_i denotes the cardinality of the neighbor set of the ith CR and E is the set of communication links between CRs in the network.

A single vector, computed from the low rate samples (and that will be defined below for each recovery algorithm), is passed between the CR nodes in the network, rather than the samples themselves, effectively reducing communication overhead. When a CR receives this vector, it performs local computation to update both the shared vector and its own estimate of the signal support accordingly. Finally, the vector is sent to a neighbor CR, chosen according to the random walk with prob-

ability (48). Two distributed algorithms are presented in [30]. The first, distributed one-step greedy algorithm (DOSGA), extends the OSGA from [59] to distributed settings. The second method, referred to as randomized distributed IHT (RDSIHT), adapts the centralized BSIHT [31] to the distributed case.

To describe the DOSGA algorithm, we first present its centralized counterpart OSGA. Each CR computes the ℓ_2-norm of the projections of the observation matrix $\mathbf{V}^{(j)}$ onto the columns of the measurement matrix $\mathbf{A}^{(j)}$, stored in the vector $\mathbf{w}^{(j)}$ of size N. The fusion center then averages over all receivers' vectors, such that

$$\hat{\mathbf{w}} = \frac{1}{N_{\text{rec}}} \sum_{j=1}^{N_{\text{rec}}} \mathbf{w}^{(j)}, \tag{49}$$

and retains the highest values of $\hat{\mathbf{w}}$ whose indices constitute the support of interest. In the absence of a fusion center, finding this average is a standard distributed average consensus problem, also referred to as distributed averaging or distributed consensus. DOSGA [30] then uses a randomized gossip algorithm [62] for this purpose, where the vectors $\mathbf{w}^{(j)}$ are exchanged, with the Metropolis-Hastings transition probabilities.

Next, we turn to the RDSIHT algorithm, which adapts the centralized BSIHT [31] to the distributed scenario. The distributed approach from [30] was inspired by the randomized incremental subgradient method proposed in [63], and recent work on a stochastic version of IHT [64]. A vector $\mathbf{w}$ of size N, that sums the ℓ_2-norms of the rows of the estimates of $\mathbf{U}^{(j)}$ before thresholding, is shared in the network through a random walk. The indices of its k largest values correspond to the current estimated support. When the ith CR receives $\mathbf{w}$, it locally updates it by performing a gradient step using its own objective function that is then added to $\mathbf{w}$. Next, it selects a neighbor j to send the vector to with probability P_{ij} (48). The joint sparsity across the CRs is exploited by sharing one common vector $\mathbf{w}$ by the network. It is shown numerically in [30] that both DOSGA and RDSIHT converge to their centralized counterparts.

5.4 Hardware Simulations: Collaborative vs. Individual Spectrum Sensing

Here, we would like to confirm that the collaborative algorithms for spectrum sensing perform better than their individual counterparts. We demonstrate a collaborative setting simulated on the MWC CR prototype, as can be seen in Fig. 26. During the simulations conducted, $N_{\text{rec}} = 5$ CR receivers, spread across different locations are emulated, denoted by white circles on the transmitter/receiver map. The transmitters are also positioned in various locations depicted by green x-marks. The transmitter positions and broadcasts are mimicking the effects of physical channel phenomena, i.e. fading and shadowing. The frequency support recovered by each of the CRs is

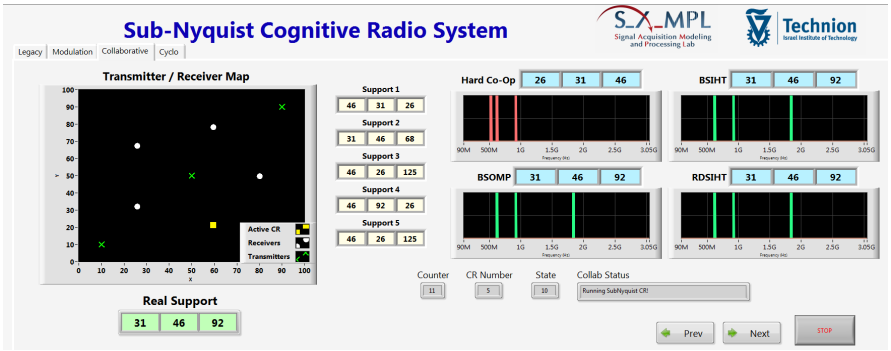


Fig. 26 Screen shot from the MWC CR collaborative hardware prototype. On the upper left side, we see the spatial map of the receivers in white, and transmitters in green. On the bottom left, the occupied band indices of the real spectral support are shown, while to the right of the transmitter/receiver map, the estimated indices by each CR individually are presented. On the right, we see the spectrum sensing results of 4 different algorithms: Hard Co-Op (hard decision collaboration that selects the most popular frequency band indices), BSIHT, BSOMP and RDSIHT. These results show both the superiority of collaborative spectrum sensing methods over individual detection and that of soft decision methods compared to the plain union of all the CR results.

false, since they individually receive only a partial spectral image of their surroundings, as expected in a real-world scenario.

In all simulated scenarios, collaborative spectrum sensing outperforms detection realized by individual CRs. This result is expected, since the soft collaborative methods take advantage of the spatial deployment of the receivers to reproduce the exact spectral map of the environment. Moreover, the centralized and distributed algorithms BSOMP, BSIHT and RDSIHT, based on soft decisions, showed superior results in comparison with a hard decision method. The same result can be seen in Fig.26, where the hard decision support algorithm (Hard Co-Op) fails to recover the entire active frequency support (depicted by red bins).

6 Joint Carrier Frequency and Direction Estimation

The final extension we consider is joint spectrum sensing and DOA estimation. In order for CRs to map vacant bands more efficiently, spatial information about the PUs' locations can be of great interest. Consider the network of CRs presented in Fig. 27 and focus on CR1. Now, picture a scenario where PU2, with DOA θ_2 with respect to CR1, is transmitting in a certain frequency band with carrier f_2. Assuming that CR2 does not receive PU2's transmission, CR1 may transmit in the same frequency band in the opposite direction of PU2 towards CR2. DOA estimation can thus enhance CR performance by allowing exploitation of vacant bands in space in addition to the frequency domain.

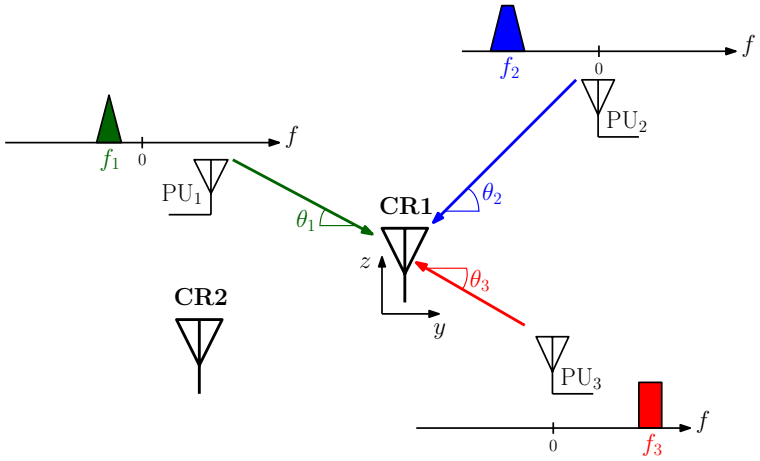


Fig. 27 Illustration of $N_{\text{sig}} = 3$ source signals in the yz plane. Each transmission is associated with a carrier frequency f_i and DOA θ_i.

6.1 Model and System Description

To formulate the joint spectrum sensing and DOA estimation problem mathematically, assume that the input signal $x(t)$ is composed of N_{sig} source signals $s_i(t)$ with both unknown and different carrier frequencies f_i and DOAs θ_i. The main difference between this scenario and the one that has been discussed in the previous sections is the additional unknown DOAs θ_i. Figure 27 illustrates this signal model. To recover the unknown DOAs, an array of sensors is required. A similar problem thoroughly treated in the literature, is the 2D-DOA recovery problem, where two angles are traditionally recovered and paired. In our case, the second variable is the signal's carrier frequency instead of an additional angle.

6.2 Multicoset Approach

A few works have recently considered joint DOA and spectrum sensing of multiband signals from sub-Nyquist samples. In [65] and [66], the low rate samples are obtained using multicoset sampling. In [65], which considers the digital model (41), both time and spatial compression are applied by selecting samples from the Nyquist grid and receivers from a uniform linear array (ULA), such that

$$\mathbf{Z}[n] = \mathbf{C}_s \mathbf{X}[n] \mathbf{C}_t. \tag{50}$$

Here, $\mathbf{X}[n]$ is the matrix of Nyquist samples from all receivers in the ULA, the selection matrices $\mathbf{C}_s$ and $\mathbf{C}_t$ operate on the spatial and time domain, respectively to form the matrix of compressed samples $\mathbf{Z}[n]$. The 2D power spectrum matrix of the

underlying signal is then reconstructed from the samples, where every row gives the power spectrum in the frequency domain for a given DOA and every column provides the power spectrum information in the angular domain for a given frequency.

In [66], an L-shaped array with two interleaved, or multicoset, channels, with a fixed delay between the two, τT_{Nyq} with $\tau \in [0,1]$, samples the signal below the Nyquist rate. The delayed path signal received at the mth sensor is approximated by

$$x_m^d(t) = \sum_{i=1}^{K} s_i(t) \mathrm{e}^{j2\pi f_i \left(t - \tau T_{\text{Nyq}} + \tau_m(\theta_i)\right)}, \tag{51}$$

using the narrowband assumption on the envelope $s_i(t)$. Here, K denotes the number of transmissions, f_i and θ_i are the carrier and DOAs of transmission i and $\tau_m(\theta_i)$ is the time difference between the mth element to the reference point for a plane wave arriving from the source i in direction θ_i. The samples of both paths is then written in frequency as

$$\mathbf{Z} = \mathbf{AX}, \quad \mathbf{Z}^d = \mathbf{ADX}, \tag{52}$$

where $\mathbf{Z}$ and $\mathbf{Z}^d$ concatenate all samples for each sensor of the direct and delayed path, respectively. Here, $\mathbf{A}$ is the unknown steering matrix that depends on f_i and θ_i, $\mathbf{X}$ is the unknown matrix that contains frequency slices of the signal and

$$\mathbf{D} \triangleq \operatorname{diag}\left[\mathrm{e}^{j2\pi f_1 \tau T_{\text{Nyq}}} \ \cdots \ \mathrm{e}^{j2\pi f_K \tau T_{\text{Nyq}}}\right]. \tag{53}$$

Exploiting correlations between samples from the direct path and its delayed version, the frequencies and their corresponding DOAs are estimated using MUSIC [18, 19]. However, the pairing issue between the two, that is matching each frequency with its corresponding angle, is not discussed.

In the next section, we describe the compressed carrier and DOA estimation (CaSCADE) system, presented in [28], that utilizes the sampling principles of the MWC. This technique addresses the pairing problem and avoids the hardware issues involved in multicoset sampling.

6.3 The CaSCADE System

The CaSCADE system implements the modified, or ULA based, MWC over an L-shaped array with $2M-1$ sensors (M sensors along the y axis and M sensors along the z axis including a common sensor at the origin) in the yz plane. Each transmission $s_i(t)$ impinges on the array with its corresponding DOA θ_i, as shown in Fig. 28. The array sensors have the same sampling pattern as the alternative MWC. Each sensor is composed of an analog mixing front-end, implementing one physical branch of the MWC, that includes a mixer, a LPF and a sampler.

By treating the L-shaped array as two orthogonal ULAs, one along the y axis and the other along the z axis, two systems of equations similar to (15) can be derived. For the ULA along the y axis, we obtain

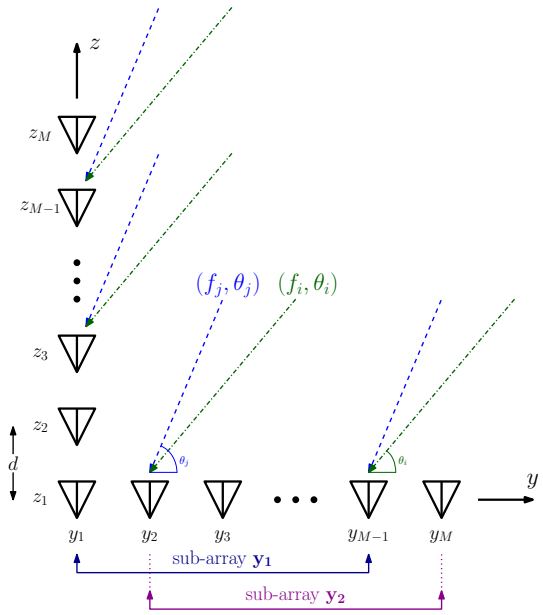


Fig. 28 CaSCADE system: L-shaped array with M sensors along the y axis and M sensors along the z axis including a common sensor at the origin. The sub-arrays $\mathbf{y}_1$ and $\mathbf{y}_2$ (and similarly $\mathbf{z}_1$ and $\mathbf{z}_2$) are defined in the derivation of the 2D-ESPRIT algorithm.

$$\mathbf{y}(f) = \mathbf{A}_y \mathbf{x}(f), \quad f \in \mathscr{F}_s, \tag{54}$$

where

$$\mathbf{A}_y = \begin{bmatrix} e^{j2\pi f_1 \tau_1^y(\theta_1)} & \dots & e^{j2\pi f_K \tau_1^y(\theta_K)} \\ \vdots & & \vdots \\ e^{j2\pi f_1 \tau_M^y(\theta_1)} & \dots & e^{j2\pi f_K \tau_M^y(\theta_K)} \end{bmatrix}. \tag{55}$$

Similarly, along the z axis, we get

$$\mathbf{z}(f) = \mathbf{A}_z \mathbf{x}(f), \quad f \in \mathscr{F}_s, \tag{56}$$

where $\mathbf{A}_z$ is defined accordingly. Here,

$$\tau_m^y(\theta) = \frac{dm}{c}\cos(\theta), \quad \tau_m^z(\theta) = \frac{dm}{c}\sin(\theta) \tag{57}$$

denote the delays at the mth sensors in the y and z axis respectively, with respect to the first sensor. The matrices $\mathbf{A}_y$ and $\mathbf{A}_z$ thus depend on both the unknown carrier frequencies and DOAs. In the time domain,

$$\mathbf{y}[n] = \mathbf{A}_y \mathbf{x}[n], \quad n \in \mathbb{Z} \tag{58}$$

$$\mathbf{z}[n] = \mathbf{A}_z \mathbf{x}[n], \quad n \in \mathbb{Z}. \tag{59}$$

Two joint recovery approaches for the carrier frequencies and DOAs of the transmissions are proposed in [28]. Note that once the carriers and DOAs are estimated, the signals can be reconstructed, as shown for the alternative MWC. For the sake of simplicity, a statistical model where $x(t)$ is wide-sense stationary is considered. The first recovery approach is based on CS techniques and allows recovery of both parameters assuming the electronic angles $f_i \cos\theta_i$ and $f_i \sin\theta_i$ lie on a predefined grid. The CS problem is formulated in such a way that no pairing issue arises between the carrier frequencies and their corresponding DOAs. To that end, the samples from both ULAs are concatenated into one vector $\mathbf{v}[n] = \left[\mathbf{y}^T[n]\,\mathbf{z}^T[n]\right]^T$, whose correlation matrix,

$$\mathbf{R} = \mathbb{E}\left[\mathbf{v}[n]\mathbf{v}^H[n]\right] = \mathbf{A}\mathbf{R}_x\mathbf{A}^H, \tag{60}$$

is computed. Here, $\mathbf{A} = [\mathbf{A}_y^T\,\mathbf{A}_z^T]^T$ and the autocorrelation matrix $\mathbf{R}_x = \mathbb{E}\left[\mathbf{x}[k]\mathbf{x}^H[k]\right]$ is sparse and diagonal, from the stationarity of $x(t)$. From the grid assumption, (60) may be discretized with respect to the possible values taken by the electronic angles. The resulting sparse matrix derived from $\mathbf{R}_x$ is diagonal as well, and its sparse diagonal is recovered using traditional CS techniques, similarly to (22).

The second recovery approach, inspired by [67], extends the ESPRIT algorithm to the joint estimation of carriers and DOAs, while overcoming the pairing issue. The 2D-ESPRIT algorithm presented in [28] is directly applied to the sub-Nyquist samples, by considering cross-correlation matrices between the sub-arrays of both axis. Dropping the time variable n for clarity, the samples from the sub-arrays can be written as

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{A}_{y_1}\mathbf{x}, \qquad \mathbf{y}_2 = \mathbf{A}_{y_2}\mathbf{x} \\ \mathbf{z}_1 &= \mathbf{A}_{z_1}\mathbf{x}, \qquad \mathbf{z}_2 = \mathbf{A}_{z_2}\mathbf{x}, \end{aligned} \tag{61}$$

where $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{z}_1$, $\mathbf{z}_2$ are samples from the sub-arrays shown in Fig. 28. The matrices $\mathbf{A}_{y_1}$ and $\mathbf{A}_{y_2}$ are the first and last $M-1$ rows of $\mathbf{A}_y$, respectively and $\mathbf{A}_{z_1}$ and $\mathbf{A}_{z_2}$ are similarly defined. Each couple of sub-array matrices along the same axis are related by

$$\begin{aligned} \mathbf{A}_{y_2} &= \mathbf{A}_{y_1}\mathbf{D}_\phi \\ \mathbf{A}_{z_2} &= \mathbf{A}_{z_1}\mathbf{D}_\psi, \end{aligned} \tag{62}$$

where

$$\begin{aligned} \mathbf{D}_\phi &\triangleq \operatorname{diag}\left[\mathrm{e}^{j2\pi f_1 \tau_1^y(\theta_1)} \;\cdots\; \mathrm{e}^{j2\pi f_K \tau_1^y(\theta_K)}\right] \\ \mathbf{D}_\psi &\triangleq \operatorname{diag}\left[\mathrm{e}^{j2\pi f_1 \tau_1^z(\theta_1)} \;\cdots\; \mathrm{e}^{j2\pi f_K \tau_1^z(\theta_K)}\right]. \end{aligned} \tag{63}$$

We can see from (63) that the carrier frequencies f_i and DOAs θ_i are embedded in the diagonal matrices $\mathbf{D}_\phi$ and $\mathbf{D}_\psi$. Applying the ESPRIT framework on cross-correlations matrices between the subarrays of both axis, allows to jointly recover $\mathbf{D}_\phi$ and $\mathbf{D}_\psi$ [28]. This leads to proper pairing of the corresponding elements $f_i\tau_1^y(\theta_i)$ and $f_i\tau_1^z(\theta_i)$. The DOAs θ_i and carrier frequencies f_i are then given by

$$\theta_i = \tan^{-1}\left(\frac{\angle(D_\psi)_{ii}}{\angle(D_\phi)_{ii}}\right) \quad f_i = \frac{\angle(D_\phi)_{ii}}{2\pi\frac{d}{c}\cos(\theta_i)}. \tag{64}$$

It is proven in [28] that the minimal number of sensors required for perfect recovery is $2K+1$. This leads to a minimal sampling rate of $(2K+1)B$, which is slightly higher than the minimal rate $2KB$ required for spectrum sensing in the absence of DOA recovery. These ideas can also be extended to jointly recover the transmissions' carrier frequencies, azimuth and elevation angles in a 3D framework.

7 Summary

In this chapter, we reviewed several challenges imposed on the traditional task of spectrum sensing by the new application of CR. We first investigated sub-Nyquist sampling schemes, enabling sampling and processing of wideband signals at low rate, by exploiting their a priori known structure. A possible extension of these works is to include adaptive updating of the detected support, triggered by a change in a PU's activity, either starting a transmission in a previously vacant band or withdrawing from an active band. To increase efficiency, this should be performed by taking the current detected support as a prior and updating it with respect to the newly acquired samples, without going through the entire support recovery process from scratch. Additional preliminary assumptions on the structure or statistical behavior of the potentially active signals, such as statistics on channel occupancy, can be exploited as well.

We then considered detection challenges in the presence of noise, where second-order statistics recovery, and in particular cyclostationary detection, are shown to perform better than simple energy detection. Next, fading and shadowing channel effects were overcome by collaborative CR networks. We then addressed the joint spectrum sensing and DOA estimation problem, allowing for better exploitation of frequency vacant bands by exploiting spatial sparsity as well. All these methods should next be combined in order to map the occupied spectrum, in frequency, time and space, thus maximizing the CR network's throughput. This would require an adequate spectrum access protocol as well, that translates the data acquired by spectrum sensing into transmission opportunities for the CRs.

An essential part of the approach adopted in this survey is the relation between the theoretical algorithms and practical implementation, demonstrating real-time spectrum sensing from low rate samples using off-the-shelf hardware components. Indeed, we believe that prototype development is an important component to enabling sub-Nyquist sampling as a solution to the task of spectrum sensing in CR platforms.

References

1. M. Mishali and Y. C. Eldar, "Sub-Nyquist sampling: Bridging theory and practice," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 98–124, Nov. 2011.
2. Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge University Press, 2015.
3. J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Inf. Theory*, vol. 56, pp. 520–544, Jan. 2010.
4. M. Fleyer, A. Linden, M. Horowitz, and A. Rosenthal, "Multirate synchronous sampling of sparse multiband signals," *IEEE Trans. Signal Process.*, vol. 58, pp. 1144–1156, Mar. 2010.
5. M. Mishali and Y. C. Eldar, "Wideband spectrum sensing at sub-Nyquist rates," *IEEE Signal Process. Mag.*, vol. 28, pp. 102–135, Jul. 2011.
6. ——, "Blind multi-band signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.
7. ——, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Select. Topics Signal Process.*, vol. 4, no. 2, pp. 375–391, Apr. 2010.
8. M. Mishali, Y. C. Eldar, and A. J. Elron, "Xampling: Signal acquisition and processing in union of subspaces," *IEEE Trans. Signal Process.*, vol. 59, pp. 4719–4734, Oct. 2011.
9. H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, pp. 523–531, Apr. 1967.
10. E. Arias-Castro and Y. C. Eldar, "Noise folding in compressed sensing," *IEEE Signal Process. Lett.*, vol. 18, no. 8, pp. 478–481, Aug. 2011.
11. 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.
12. G. L. Turin, "An introduction to matched filters," *IRE Trans. Inf. Theory*, vol. 6, pp. 311–329, Jun. 1960.
13. W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, pp. 639–697, 2006.
14. A. Napolitano, "Cyclostationarity: New trends and applications," *Signal Process.*, vol. 120, pp. 385–408, Mar. 2016.
15. I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Phys. Commun.*, vol. 4, pp. 40–62, Mar. 2011.
16. S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative sensing among cognitive radios," *IEEE Int. Conf. Commun.*, pp. 1658–1663, 2006.
17. K. B. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 878–893, 2009.
18. V. F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophysics, J. Roy. Astron. Soc.*, vol. 33, pp. 347–366, 1973.
19. R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagation*, vol. 34, pp. 276–280, Mar. 1986.
20. R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Signal Process.*, vol. 37, pp. 984–995, Jul. 1989.
21. Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.
22. H. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math*, vol. 117, pp. 37–52, Jul. 1967.
23. R. Venkataramani and Y. Bresler, "Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2173–2183, Sept. 2000.
24. M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-Nyquist rates," *IET Circuits, Devices and Syst.*, vol. 5, pp. 8–20, Jan. 2011.
25. E. Israeli, S. Tsiper, D. Cohen, A. Reysenson, R. Hilgendorf, E. Shoshan, and Y. C. Eldar, "Hardware calibration of the modulated wideband converter," *IEEE Global Commun. Conf.*, pp. 948–953, Dec. 2014.

26. M. Mishali and Y. C. Eldar, "Expected RIP: Conditioning of the modulated wideband converter," *IEEE Inf. Theory Workshop*, pp. 343–347, 2009.
27. H. W. L. Gan and H. Wang, "Deterministic binary sequences for modulated wideband converter," in *Int. Conf. Sampling Theory and Applications*, 2013.
28. S. Stein, O. Yair, D. Cohen, and Y. C. Eldar, "CaSCADE: Compressed carrier and DOA estimation," *arXiv:1604.02723 [cs.IT]*, 2016.
29. R. Gold, "Optimal binary sequences for spread spectrum multiplexing (corresp.)," *IEEE Trans. Inf. Theory*, vol. 13, no. 4, pp. 619–621, 1967.
30. D. Cohen, A. Akiva, B. Avraham, and Y. C. Eldar, "Distributed cooperative spectrum sensing from sub-Nyquist samples for cognitive radios," *IEEE Workshop Signal Proc. Advances Wireless Commun.*, pp. 336–340, Jun. 2015.
31. ——, "Centralized cooperative spectrum sensing from sub-Nyquist samples for cognitive radios," *IEEE Int. Conf. Commun.*, pp. 7487–7491, Jun. 2015.
32. D. Cohen and Y. C. Eldar, "Sub-Nyquist cyclostationary detection for cognitive radio," *arXiv:1604.02659 [cs.IT]*, 2016.
33. D. Adams, Y. Eldar, and B. Murmann, "A mixer frontend for a four-channel modulated wideband converter with 62 db blocker rejection," in *2016 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, May 2016, pp. 286–289.
34. M. A. Lexa, M. E. Davies, and J. S. Thompson, "Compressive and noncompressive power spectral density estimation from periodic nonuniform samples," *CoRR*, vol. abs/1110.2722, 2011.
35. D. D. Arianandа and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. Signal Process.*, vol. 60, pp. 4775–4789, Sept. 2012.
36. D. Romero and G. Leus, "Compressive covariance sampling," *Proc. Inf. Theory and Applications Workshop*, pp. 1–8, Feb. 2013.
37. C. P. Yen, Y. Tsai, and X. Wang, "Wideband spectrum sensing based on sub-Nyquist sampling," *IEEE Trans. Signal Process.*, vol. 61, pp. 3028–3040, Jun. 2013.
38. 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.
39. Z. Tian, Y. Tafesse, and B. M. Sadler, "Cyclic feature detection with sub-nyquist sampling for wideband spectrum sensing," *IEEE J. Select. Topics Signal Process.*, vol. 6, no. 1, pp. 58–69, Feb. 2012.
40. G. Leus and Z. Tian, "Recovering second-order statistics from compressive measurements," *IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pp. 337–340, Dec. 2011.
41. W. Gardner, *Statistical spectral analysis: a non probabilistic theory*. Prentice Hall, 1988.
42. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 1991.
43. P. Pal and P. P. Vaidyanathan, "Nested array: A novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, pp. 4167–4181, Aug. 2010.
44. P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, pp. 573–586, 2011.
45. D. Qu and A. Tarczynski, "A novel spectral estimation method by using periodic nonuniform sampling," *Asilomar Conf. Signals, Syst. and Computers*, pp. 1134–1138, 2007.
46. C. G. Khatri and C. R. Rao, "Solutions to some functional equations and their applications to characterization of probability distributions," *Sankhyā: The Indian J. Statistics, Series A*, pp. 167–180, 1968.
47. M. Arts, A. Bollig, and R. Mathar, "Analytical test statistic distributions of the mmme eigenvalue-based detector for spectrum sensing," in *2015 International Symposium on Wireless Communication Systems (ISWCS)*. IEEE, 2015, pp. 496–500.
48. D. Romero, R. López-Valcarce, and G. Leus, "Compression limits for random vectors with linearly parameterized second-order statistics," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1410–1425, 2015.
49. J. Leech, "On the representation of 1, 2,, n by differences," *J. London Mathematical Soc.*, vol. 1, no. 2, pp. 160–169, 1956.

50. W. A. Gardner, "The spectral correlation theory of cyclostationary time-series," *Signal Process.*, vol. 11, pp. 13–36, Jul. 1986.
51. D. Cohen, L. Pollak, and Y. C. Eldar, "Carrier frequency and bandwidth estimation of cyclostationary multiband signals," *IEEE ICASSP*, Mar. 2016.
52. R. L. Thorndike, "Who belong in the family?" *Psychometrika*, vol. 18, pp. 267–276, Dec. 1953.
53. A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," *IEEE Int. Symp. New Frontiers in Dynamic Spectrum Access Networks*, pp. 131–136, Nov. 2005.
54. B. Sklar, "Rayleigh fading channels in mobile digital communication systems part I: Characterization," *IEEE Commun. Mag.*, vol. 35, pp. 90–100, Jul. 1997.
55. Y. Wang, A. Pandharipande, Y. L. Polo, , and G. Leus, "Distributed compressive wide-band spectrum sensing," *IEEE Inf. Theory and Applications Workshop*, pp. 178–183, Feb. 2009.
56. D. D. Arianandа and G. Leus, "A study on cooperative compressive wideband power spectrum sensing," *Joint WIC/IEEE Symp. Inf. Theory and Signal Process.*, pp. 102–109, 2012.
57. J. Tropp, A. C. Gilbert, and M. J. Strauss, "Simultaneous sparse approximation via greedy pursuit," *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, vol. 5, pp. 721–724, Mar. 2005.
58. A. Makhzani and S. Valaee, "Reconstruction of jointly sparse signals using iterative hard thresholding," *IEEE Int. Conf. Commun.*, pp. 3564–3568, 2012.
59. M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," *IEEE Asilomar Conf. Signals, Syst. and Computers*, pp. 1537–1541, 2005.
60. Z. Tian, "Compressed wideband sensing in cooperative cognitive radio networks," *IEEE Global Commun. Conf.*, pp. 1–5, Dec. 2008.
61. F. Zeng, C. Li, and Z. Tian, "Distributed compressive spectrum sensing in cooperative multihop cognitive networks," *J. Select. Topics Signal Process.*, vol. 5, pp. 37–48, Feb. 2011.
62. S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, pp. 2508–2530, 2006.
63. B. J. M. Rabi and M. Johansson, "A randomized incremental subgradient method for distributed optimization in networked systems," *SIAM J. Optimization*, vol. 20, no. 3, pp. 1157–1170, 2009.
64. N. Nguyen, D. Needell, and T. Woolf, "Linear convergence of stochastic iterative greedy algorithms with sparse constraints," *CoRR*, vol. abs/1407.0088, 2014. [Online]. Available: http://arxiv.org/abs/1407.0088
65. D. D. Arianandа and G. Leus, "Compressive joint angular-frequency power spectrum estimation," *Proc. European Signal Process. Conf.*, pp. 1–5, Sep. 2013.
66. A. A. Kumar, S. G. Razul, and C. S. See, "An efficient sub-Nyquist receiver architecture for spectrum blind reconstruction and direction of arrival estimation," *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, pp. 6781–6785, May 2014.
67. J.-F. Gu, W.-P. Zhu, and M. N. S. Swamy, "Joint 2-D DOA estimation via sparse L-shaped array," *IEEE Trans. Signal Process.*, vol. 63, Mar. 2015.