Deborah Cohen, Shahar Tsiper, and Yonina C. Eldar

Analog-to-Digital Cognitive Radio

Sampling, detection, and hardware

The radio spectrum is the radio-frequency (RF) portion of the electromagnetic spectrum. These spectral resources are traditionally allocated to licensed or primary users (PUs) by governmental organizations. As discussed in "Radio-Frequency Spectral Resources," most of the frequency bands are already allocated to one or more PUs. Consequently, new users cannot easily find free frequency bands. Spurred by the ever-increasing demand from new wireless communication applications, this issue has become critical over the past few years.

Various studies [1]–[3] have shown that this overcrowded spectrum is usually significantly underutilized in frequency, time, and space. A 2002 Federal Communications Commission report shows large temporal and geographic variations in

Spectral resources are managed by governmental organizations that allocate them to fixed users. In the United States, regulatory responsibility for the radio spectrum is divided between the Federal Communications Commission, which administers the spectrum for nonfederal use, and the National Telecommunications and Information Administration, which is responsible for federal use.

In Europe, the European Telecommunications Standards Institute and the Electronic Communications Committee cooperate on aspects related to the regulatory environment for the radio spectrum. Figure S1 shows the U.S. frequency allocation chart. As can be seen, all of the bands spanning the very-high-frequency, ultrahigh-frequency, and superhigh-frequency ranges (0.003–30 GHz) are preallocated to one or even several licensed primary users. This poses inherent difficulties in introducing new technologies that require the occasional usage of this spectral range.

FIGURE S1. The U.S. frequency allocation chart. (Image courtesy of the U.S. Department of Commerce, National Telecommunications and Information Administration, Office of Spectrum Management, January 2016.)

terms of the spectrum usage, ranging from 15% to 85% utilization. Figure 1 presents the spectrum occupancy measurements for two major metropolitan areas, New York City and Chicago, Illinois, as conducted by the Shared Spectrum Company (SSC). We observe that the occupancy of most frequency bands does not exceed 50%, and half of them are even below 10% [4]. This shows that spectrum shortage is often due to the inflexibility of the frequency allocation, resulting in low utilization efficiency.

To bridge spectrum scarcity and sparsity and make better use of spectral resources, the cognitive radio (CR) [6]–[8] paradigm has been proposed. The idea of adaptive learning for spectrum sensing can be traced back to Shannon [9]. This technology, which is under development, allows the spectrum to be used more efficiently by granting secondary users opportunistic access to licensed spectral bands when the corresponding PU is not active. A CR transceiver scans for unused bands and changes its transmission and reception parameters to different frequencies depending on the spectrum utilization.

CR faces many issues at various levels, and it challenges traditional analog, digital, and network processing techniques to

FIGURE 1. The spectrum occupancy measurements in New York City from 30 August to 3 September 2004 and Chicago, Illinois, from 16 to 18 November 2005. (Image courtesy of SSC [5].)

meet its specific radio-sensitivity requirements and wide-band frequency agility [10]. The CR cycle includes two major functionalities: spectrum sensing and spectrum access. Through spectrum sensing, CRs collect information about the status of the surrounding spectrum's occupancy, i.e., the PUs' activity, allowing for the adapted exploitation of the vacant spectral bands. To minimize the interference caused to PUs, the spectrum-sensing task performed by a CR should be reliable and fast [11]–[13] and obey the requirements of the IEEE 802.22 protocol, which is described in "IEEE 802.22 Standard for WRAN."

For the past ten years, CR and its challenges have been thoroughly reviewed in the literature. Several works [14]–[19] focus on spectrum-sensing and survey-sensing techniques, along with their performance and limitations. These techniques are essen-

tially energy detection, matched filter, and cyclostationary detection. Collaborative CR networks, where different users share their sensing results and cooperatively decide on the licensed spectrum occupancy, have also been proposed to overcome practical propagation issues, such as path loss, channel fading, and shadowing. Other works [17], [20] deal with spectrum access, which uses the information gathered from spectrum sensing to plan the spectrum exploitation by unlicensed users. This functionality includes spectrum analysis, spectrum access, and spectrum mobility [21]. It ensures coexistence with PUs and other CRs by minimizing interference for the former and sharing the spectrum with the latter.

In this article, we focus on the issue of spectrum sensing from the analog-to-digital interface point of view, which has so far received little attention. Our motivation stems from one of the

During the past decade, the use of analog TV radio bands has been slowly decaying, mainly due to the introduction of digital TV broadcasts based on new transmission infrastructure, such as cable and satellites. The once-crowded, very-high-frequency (VHF) and ultrahigh-frequency (UHF) range (between 54–860 MHz) is less occupied today, particularly in remote rural areas. Concurrently, a new need has emerged in these areas for fast wireless data networks with long-range capabilities of up to 30 km. For these operating range requirements, the VHF/UHF radio spectrum, able to traverse these distances with relatively low transmitting power, is a natural choice. Combining the need for available frequencies in this range with the fact that it is mostly underutilized today has laid the foundation for designing efficient spectrum-sharing techniques, including cognitive radio (CR).

IEEE 802.22 is the first international standard that is especially designed to achieve this goal. It describes a wireless regional area network (WRAN) that relies on new CR technologies and exploits the vacant white space in the VHF/UHF range while still being reserved to licensed TV bands. Managing the spectrum and transmitting broadband communications only in vacant frequency slots enables fast Internet access in hard-to-reach areas with low population density and, therefore, has high potential for extensive worldwide usage.

To exploit analog TV bands' white spaces, an opportunistic approach is selected. Both base stations (BSs) and clients, referred to as *customer-premises equipment* (*CPE*), are spread apart geographically, as seen in Figure S2. To enable a true CR network, the standard defines necessary capabilities from both the BSs and CPEs to enable cognitive sensing and adapt to possible rapid changes incurred by incumbent users.

The physical (PHY) layer is the lowest network layer. It contains the circuitry used to transmit and receive analog communication signals. For implementing CR technologies, the PHY has to rapidly adapt to spectral changes and be both agile and flexible enough to jump between carrier frequencies without losing information. It is required to constantly listen to the operating band during designated
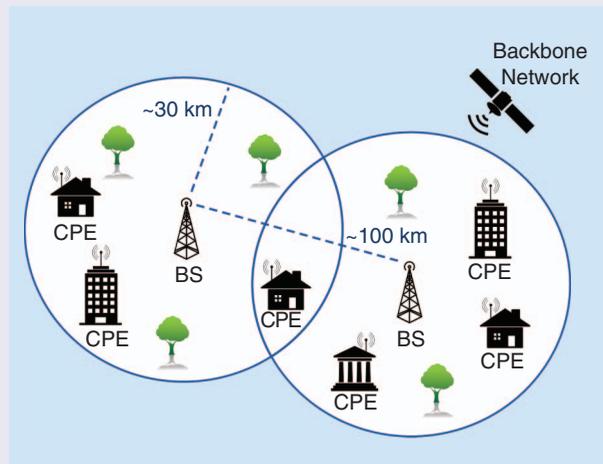
FIGURE S2. The CR network illustration, including BSs and CPE.

quiet times to ensure that the incumbent user does not wake up and require its use. The media access control, situated above the PHY layer in the network hierarchy, is responsible for CR communication management. Its specifications include fast and dynamic adaptation to changes in the environment by constantly sensing the spectrum. The requirements for each of the different network layers are particularly demanding when compared to other radio-frequency standards from the 802.xx family.

To conclude, the 802.22 WRAN standard holds great promise in delivering fast broad-band connections to remote rural areas, with broadcast ranges of up to 30 km, by smartly sharing the already allocated VHF and UHF spectrum. Spectrum coexistence is made possible by CR tools, which require advanced technologies. These introduce tremendous challenges in terms of the necessary hardware to meet such harsh demands while preserving small form factors and low power levels. Sub-Nyquist methods for spectral sensing and reconstruction offer a way to alleviate part of this burden by exploiting the signals' spectral structure, effectively making CR technology more accessible.

main challenges of spectrum sensing in the context of CR, which is the sampling-rate bottleneck. This issue arises because CRs typically deal with wide-band signals with prohibitively high Nyquist rates. Sampling at this rate requires very sophisticated and expensive analog-to-digital converters (ADCs), leading to a torrent of samples. Therefore, the classic spectrum-sensing methods described previously in this section, which are traditionally performed on Nyquist rate samples, are difficult to implement in practice on wide-band signals. Our main goal is to provide an analog-to-digital CR framework, including an analog prepro-

cessing and sub-Nyquist sampling front end and subsequent low-rate digital processing.

We first survey recent methods for spectrum sensing at sub-Nyquist rates, paving the way to efficient sensing with low computational and power requirements. We then review spectrum-sensing strategies that aim at overcoming other diverse challenges in the context of CR, such as coping with low signal-to-noise ratio (SNR) regimes, channel fading, and shadowing effects. Throughout this article, we consider both theoretical and practical aspects and present the hardware implementation

of theoretical concepts, demonstrating real-time wide-band spectrum sensing for CR from low-rate samples.

CR challenges

In this article, we focus on CR spectrum sensing. In practice, the information gathered from spectrum sensing is used to plan spectrum access by the unlicensed users. For completeness, we first briefly review the main components and challenges of this topic.

Spectrum analysis or management, which directly follows spectrum sensing, ensures coexistence with PUs and other CRs. The ambient RF environment is analyzed to characterize the behavior of PUs and the properties of the detected spectrum holes in terms of interference, duration of availability, and more. Then, spectrum access can be optimized to maximize the CR throughput while maintaining interference caused to the licensed users below a target threshold [20]. Several techniques have been proposed to minimize interference to PUs as well as ensure the proper reception of secondary signals, such as waveform design and multicarrier approaches [17], [20]. These are regrouped under the term *spectrum sculpting* [17]. Besides minimizing interference for the PUs, spectrum sharing needs to be coordinated within the CR network. Various power control and resource allocation schemes that deal with this issue are reviewed in [16] and [20]. Spectrum access further requires synchronization between the CR transmitter and receiver [20].

The function of spectrum mobility ensures an adaptation to changes in the spectrum occupancy. When a licensed user starts accessing the channel currently being used by a CR, then the latter has to vacate the band and switch to another free channel. This operation is referred to as a *handoff* [16], [20]. The multicarrier transmission approach described previously in this section maintains uninterrupted communication in such scenarios [17]. Additional issues of security against malicious users and various attacks to the network are discussed in [16], [19], and [20].

We now focus on spectrum sensing, which is the fundamental enabler to spectrum access. To increase the chance of finding an unoccupied spectral band, CRs must sense a wide band of the spectrum. Nyquist rates of wide-band signals are high and can even exceed today's best ADC front-end bandwidths. Apart from that, such high sampling rates generate a large number of samples to process, which consequently affects 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 reduction. These include the random demodulator [22], [23], multirate sampling [24], multicoset sampling, and the modulated wide-band converter (MWC) [13], [25]–[27].

The CR then performs spectrum sensing on the acquired samples to detect PU transmissions. The simplest and most common spectrum-sensing approach is energy detection [28], which does not require any a priori knowledge on the input

> **In practice, the information gathered from spectrum sensing is used to plan spectrum access by the unlicensed users.**

signal. Unfortunately, energy detection is very sensitive to noise and performs poorly in low SNRs. This becomes even more critical in sub-Nyquist regimes because the sensitivity of energy detection is amplified due to noise aliasing [29]. Therefore, this scheme fails to meet CR performance requirements in low SNRs. In contrast, matched-filter (MF) detection [30], [31], 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 potentially received transmission. When no a priori knowledge can be assumed on the received signals' waveform, MF is difficult to implement. A compromise between both methods is cyclostationary detection [32], [33]. This strategy is more robust for 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 as a result of path loss, fading, and shadowing [34]–[36]. These phenomena are due to the signal's propagation, which 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. Cooperative spectrum sensing may be classified into three catagories 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. The first is decision fusion, or hard decision, where the CRs only report their binary local decisions. The second is measurement fusion, or soft decision, where they share their samples [34]. Cooperation has been shown to improve detection performance and relax sensitivity requirements by exploiting spatial diversity [36], [37]. At the medium-access control level, cooperation introduces the need for a tailored communication protocol and a control channel [14], [15], which can be implemented as a dedicated frequency channel or as an underlay ultrawide-band channel [10]. These CR communication challenges are outside the scope of this article.

Finally, CRs may require, or at least benefit from, joint spectrum sensing and direction-of-arrival (DOA) estimation. DOA recovery enhances CR performance by allowing exploitation of vacant bands in space in addition to the frequency domain. 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. To jointly estimate the carrier frequencies and DOAs of the received transmissions, arrays of sensors have been considered, and DOA recovery techniques, such as multiple signal classification (MUSIC) [38], [39], the
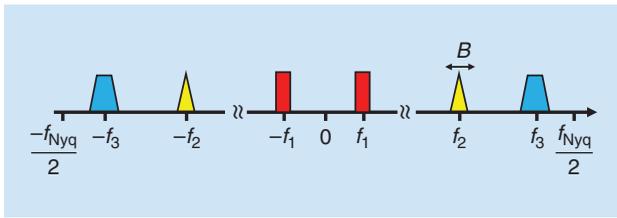
FIGURE 2. A multiband model with $K = 6$ bands. Each band does not exceed the bandwidth B and is modulated by an unknown carrier frequency $|f_i| \le f_{\text{Nyq}}/2$, for $i = 1, 2, 3$.

estimation of signal parameters via rotational invariance techniques (ESPRIT) [40], or compressed-sensing (CS)-based [41] approaches, may then be adapted to the joint carrier and DOA estimation problem. We review several algorithms that jointly estimate the carrier frequencies and DOAs of the received transmissions from the same low-rate samples obtained from an array of sensors.

This article focuses on the previously discussed spectrum-sensing challenges for CR. We first review sub-Nyquist sampling methods for multiband signals and then consider the different aspects of spectrum sensing that are 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. In particular, we examine the analog-to-digital interface of CRs, as opposed to the review in [42], which is concerned with the digital implications of compressive spectrum sensing. This allows for a demonstration of the theoretical concepts on hardware prototypes. In particular, 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 in the first section of this article.

We note that CRs face additional fundamental difficulties, such as the hidden-node problem and potential interference from secondary users to existing primary communication links [17], [20], [21]. These are beyond the scope of this article but need to be addressed as well for CRs to become practical.
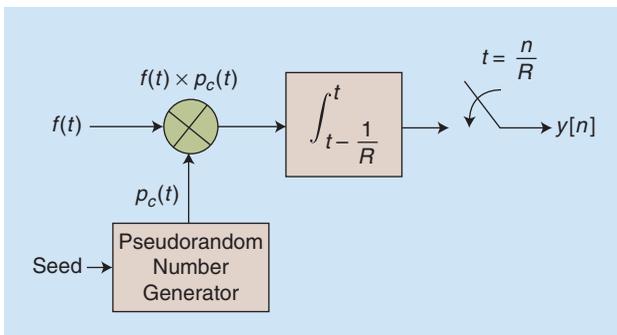
FIGURE 3. A block diagram for the random demodulator, including a random number generator, a mixer, an accumulator, and a sampler [23].

Sub-Nyquist sampling for CR

CR receivers sense signals composed of several transmissions with unknown support that are spread over a wide spectrum. Such sparse wide-band signals belong to the so-called multiband model [25], [26], [43]. An example of a multiband signal $x(t)$ with K bands is illustrated in Figure 2, with individual bandwidths not greater than B, centered around unknown carrier frequencies $|f_i| \le f_{\text{Nyq}}/2$, 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 [43], [46], and references therein), reduce the sampling rate below Nyquist. Here, because the frequency location of the transmissions is unknown, classic processing first samples $x(t)$ at its Nyquist rate f_{Nyq}, which may be prohibitively high. Sensing such wide-band signals may be performed by splitting them into several frequency bands and processing each band separately. This can be done either in series or in parallel. However, this increases the latency in the first case and the hardware complexity in the second.

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 they do not require knowledge of the carrier frequencies. It has been shown in [25] that the minimal sampling rate for perfect blind recovery in multiband settings is twice the Landau rate [47], i.e., 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 the minimal sampling rate. Two main limitations of sub-Nyquist sampling are noise enhancement [29] and the inherent assumption that the channel is sparse and contains spectral holes. The former issue is discussed later in the "Statistics Detection" and "Collaborative-Spectrum Sensing" sections and may be partially resolved by the techniques reviewed next, such as cyclostationary detection and collaborative-spectrum sensing.

Multitone model and the random demodulator

Laska et al. and Tropp et al. [22], [23] consider a discrete multitone model and suggest sampling using the random demodulator, as depicted in Figure 3. Multitone functions are composed of K active tones spread over a bandwidth W, such that

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

Here, Ω is composed of K normalized frequencies, or tones, that are a subset of the integers in the interval $[-W/2, W/2]$, and b_ω, for $\omega \in \Omega$, is 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.

Compressed sensing (CS) [41], [43] is a framework for the simultaneous sensing and compression of finite-dimensional vectors, which relies on linear dimensionality reduction. In particular, the field of CS focuses on the following recovery problem:

$$\mathbf{z} = \mathbf{A}\mathbf{x}, \tag{S1}$$

where $\mathbf{x}$ is an $N \times 1$ sparse vector, i.e., with few nonzero entries, and $\mathbf{z}$ is a vector of the measurements of size $M < N$. CS provides recovery conditions and algorithms to reconstruct $\mathbf{x}$ from the low-dimensional vector $\mathbf{z}$.

The different spectrum-sensing applications described in this article primarily deal with analog signals and sampling techniques, while CS inherently defines a digital framework. We will discuss how the analog approaches of low-rate sampling use CS as a tool for recovery and adapt it to the analog setting. To that end, we describe here two CS greedy recovery algorithms that solve the optimization problem

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t. } \mathbf{z} = \mathbf{A}\mathbf{x}, \tag{S2}$$

where $\|\cdot\|_0$ denotes the ℓ_0-norm. The first algorithm we consider belongs to the family of matching pursuit (MP) methods [44]. The orthogonal MP algorithm iteratively proceeds by finding the column of $\mathbf{A}$ that is most correlated to the signal residual $\mathbf{r}$,

$$i = \arg\max|\mathbf{A}^H \mathbf{r}|, \tag{S3}$$

where the absolute value is computed elementwise. The residual is obtained by subtracting the contribution of a partial estimate $\hat{\mathbf{x}}_\ell$ of the signal at the ℓth iteration, from $\mathbf{z}$, as follows:

$$\mathbf{r} = \mathbf{z} - \mathbf{A}\hat{\mathbf{x}}_\ell. \tag{S4}$$

Once the support set is updated by adding the index i, the coefficients of $\hat{\mathbf{x}}_\ell$ over the support set are updated, so as to minimize the residual error.

Other greedy techniques include thresholding algorithms. Here, we focus on the iterative hard thresholding method proposed in [45]. Starting from an initial estimate $\hat{\mathbf{x}}_0 = 0$, the algorithm iterates a gradient descent step with step size μ followed by hard thresholding, i.e.,

$$\hat{x}_\ell = \mathcal{T}(\hat{\mathbf{x}}_{\ell-1} + \mu \mathbf{A}^H(z - \mathbf{A}\hat{\mathbf{x}}_{\ell-1}), k), \tag{S5}$$

until a convergence criterion is met. Here, $\mathcal{T}(\mathbf{x}, k)$ denotes a thresholding operator on $\mathbf{x}$ that sets all but the k entries of $\mathbf{x}$ with the largest magnitudes to zero, and k is the sparsity level of $\mathbf{x}$, assumed to be known.

These two greedy algorithms and other CS recovery techniques can be adapted to further settings, such as multiple measurement vectors, where the measurements $\mathbf{z}$ and sparse objective $\mathbf{x}$ become matrices, infinite measurement vectors, block sparsity, and more, as we will partially discuss in the article. Further details on CS recovery conditions and techniques can be found in [41] and [43].

To sample the signal $f(t)$, it is first modulated by a high-rate sequence $p_c(t)$ created by a pseudorandom number generator. It is then integrated and sampled at a low rate, as shown in Figure 3. The random sequence used for demodulation 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 pseudorandom sequence. The resulting demodulated signal $y(t) = f(t)p_c(t)$ is then integrated over a period $1/R$ and sampled at the low rate R. This integrate-and-dump approach results in the following samples:

$$y_m = R \int_{m/R}^{(m+1)/R} y(t)\,dt, \quad m = 0, 1, \ldots, R - 1. \tag{2}$$

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 [22], [23]. In matrix form, we write

$$\mathbf{y} = \mathbf{\Phi}\mathbf{b}, \tag{3}$$

where $\mathbf{y}$ is the vector of size R that contains the samples y_m and $\mathbf{\Phi}$ is the known sampling matrix that describes the overall action of the system on the vector of amplitudes $\mathbf{b}$, i.e., demodulation and filtering (see [23] 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 [41] techniques like those discussed in "Compressed Sensing Recovery," in turn, allowing for the recovery of $f(t)$. The minimal required number of samples R for perfect recovery of $f(t)$ in a noiseless setting is $2K$ [41].

The random demodulator is one of the pioneering and innovative attempts to extend the inherently discrete and finite CS theory to analog signals. However, truly analog signals, such 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 like those from the multiband model, the number of tones W is of 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
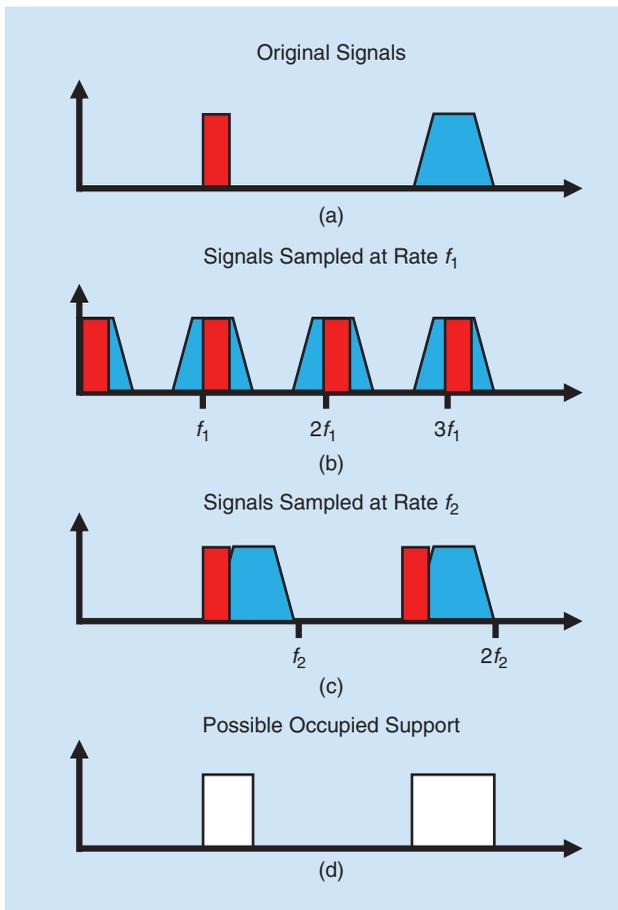
FIGURE 4. The action of the SMRS on a multiband signal. (a) The input signal with $K = 2$ bands, (b) signals sampled at rate F_1 in channel 1, (c) signals sampled at rate F_2 in channel 2, and (d) possible support that is the intersection of the supports in channels 1 and 2 [24].

to the grid choice (see [46] for a detailed analysis). Furthermore, the time-domain approach precludes processing at a low rate, even for multitone inputs, because 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 implemented with 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 handle the analog multiband model (Figure 2), which is of interest in the context of CR.

Multirate sampling

An alternative sampling approach is based on the synchronous multirate sampling (SMRS) [24] 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 intervals, as illustrated in Figure 4. The Fourier transform of the undersampled signals is then related

to the original signal through an underdetermined system of linear equations,

$$\mathbf{z}(f) = \mathbf{Q}\mathbf{x}(f). \tag{4}$$

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 (9), which concatenates the observation vector of all the channels. The measurement matrix $\mathbf{Q}$ has exactly P nonzero elements in each column, which 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 is finite. The continuous variable f is then discretized to a frequency resolution of Δf. Because $x(t)$ is sparse in the frequency domain, the vector $\mathbf{x}(f)$ is sparse and can be recovered from (9) using CS techniques for each discrete frequency f. An alternative recovery method, referred to as the *reduction procedure* and illustrated in Figure 4, 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 downconverted to that baseband frequency, which allows a reduction of the number of sampling channels. This assumption does not hold in the case where two or more frequency components cancel each other out due to aliasing, which happens with probability zero. Once the corresponding components are eliminated from (4), the reduced system is 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 because some of the samples are thrown out. Furthermore, spectral components downconverted 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. Another difficulty is that, in practice, synchronization between channels sampling at different rates is challenging. Finally, this scheme samples wide-band signals using low-rate samplers. Practical ADCs introduce an inherent bandwidth limitation, modeled by an antialiasing low-pass filter (LPF) with a cutoff frequency determined by the sampling rate, which distorts the samples. Thus, the SMRS implementation requires low-rate samplers with large analog bandwidth.

Multicoset sampling

A popular scheme for sampling wide-band signals at the Nyquist rate involves multicoset or interleaved ADCs [25], [46], [48], in which several channels are used, each operating at a lower rate. Multicoset sampling may be described as the selection of certain samples from the uniform Nyquist grid, as shown in Figure 5, where $T_{Nyq} = 1/f_{Nyq}$ 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

$$x_{c_i}[n] = \begin{cases} x(nT_{Nyq}), & n = mN + c_i, m \in \mathbb{Z} \\ 0, & \text{otherwise}, \end{cases} \quad (5)$$

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

The samples in the Fourier domain can then be written as linear combinations of spectrum slices of $x(t)$, such that [25]

$$\mathbf{z}(f) = \mathbf{A}\mathbf{x}(f), \quad f \in \mathcal{F}_s. \quad (6)$$

Here, $\mathbf{x}(f)$ denotes the spectrum slices of $x(t)$, $f_s = (1/NT_{Nyq}) \ge B$ is the sampling rate of each channel, and $\mathcal{F}_s = [-f_s/2, f_s/2]$. This relation is illustrated in Figure 7. The $M \times N$ sampling matrix $\mathbf{A}$ is determined by the selected cosets c_i. The recovery processing described later in this section is performed in the time domain, where we can write

$$\mathbf{z}[n] = \mathbf{A}\mathbf{x}[n], \quad n \in \mathbb{Z}, \quad (7)$$

The vector $\mathbf{z}[n]$ collects the measurements at $t = n/f_s$, and $\mathbf{x}[n]$ contains the samples sequences corresponding to the spectrum slices of $x(t)$.

The goal can then be stated as the recovery of $\mathbf{x}[n]$ from the samples $\mathbf{z}[n]$. The system (7) is underdetermined due to the sub-Nyquist setup and known as *infinite measurement vectors* (*IMVs*) in the CS literature [41], [43]. With respect to these two properties, the digital reconstruction algorithm encompasses the following three stages [25] that we explain in more detail next:

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 that of S of $\mathbf{x}[n]$, i.e., the active slices.
3) The signal is digitally recovered by reducing (7) 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 [25], exploits the fact that the bands occupy continuous spectral intervals so that all $\mathbf{x}[n]$ are jointly sparse, i.e., they have the same spectral support S over time. The CTF produces a finite system of equations, called *multiple measurement vectors* (*MMVs*) [41], [43], from the infinite number of linear systems described by (7). The samples are first summed as

$$\mathbf{Q} = \sum_n \mathbf{z}[n]\mathbf{z}^H[n] \quad (8)$$

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
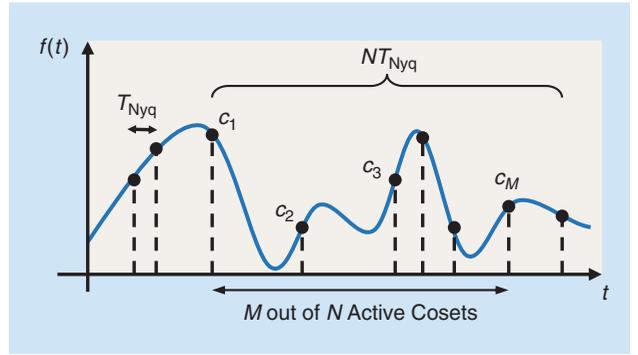
FIGURE 5. An illustration of multicoset sampling.
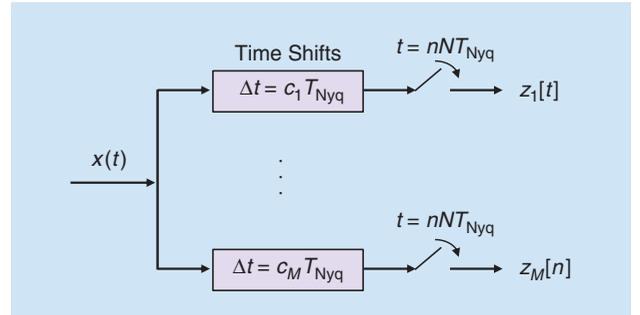
FIGURE 6. A schematic implementation of multicoset sampling. The input signal $x(t)$ is inserted into the multicoset sampler, which splits the signal into M branches and delays each one by a fixed coefficient $c_i T_{Nyq}$. Every branch is then sampled at the low rate $1/(NT_{Nyq})$ and then digitally processed to perform spectrum sensing and signal reconstruction.
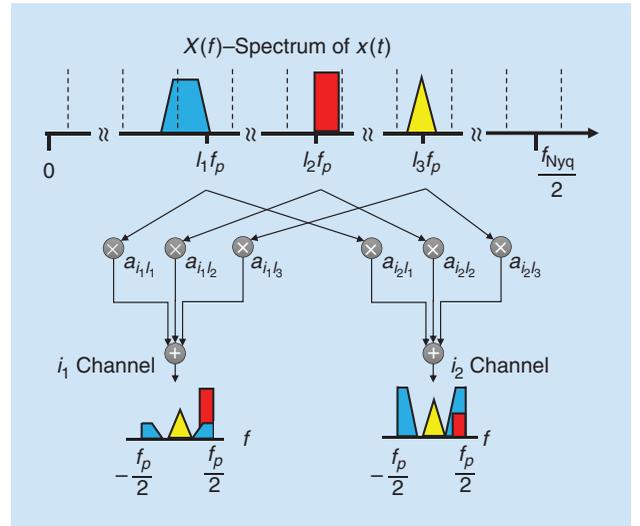
FIGURE 7. The spectrum slices of the input signal $x(f)$ are shown here to be multiplied by the coefficients a_{il} of the sensing matrix $\mathbf{A}$, resulting in the measurements z_i for the ith channel. Note that, in multicoset sampling, only the slices' complex phase is modified by the coefficients a_{il}. In the MWC sampling described later, in the "MWC Sampling" section, both the phases and amplitudes are affected.

option is to construct it by performing an eigendecomposition of $\mathbf{Q}$ and choosing $\mathbf{V}$ as the matrix of eigenvectors corresponding to the nonzero eigenvalues. The finite-dimensional MMV system

$$\mathbf{V} = \mathbf{AU} \tag{9}$$

is then solved for the sparsest matrix $\mathbf{U}$ with minimal number of nonidentically zero rows using CS techniques [41], [43]. The key observation of this recovery strategy is that the indices of the nonzero rows of $\mathbf{U}$ coincide with the active spectrum slices of $\mathbf{z}[n]$ [25]. 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 (7) to S. The resulting matrix $\mathbf{A}_S$, which contains the columns of $\mathbf{A}$ corresponding to S, is then inverted as follows:

$$\mathbf{x}_S[n] = \mathbf{A}_S^\dagger \mathbf{z}[n]. \tag{10}$$

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

$$f_{\text{tot}} = Mf_s = \frac{M}{N}f_{\text{Nyq}}. \tag{11}$$

The minimal number of channels is dictated by CS results [41] such 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 [47].

Although this sampling scheme seems relatively simple and straightforward, it suffers from several practical drawbacks [46]. First, as in the SMRS approach, multicoset sampling requires low-rate ADCs with large analog bandwidth. Another issue arises from the time-shift elements, because maintaining accurate time delays between the ADCs to the order of the Nyquist interval T_{Nyq} is difficult. Finally, 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.
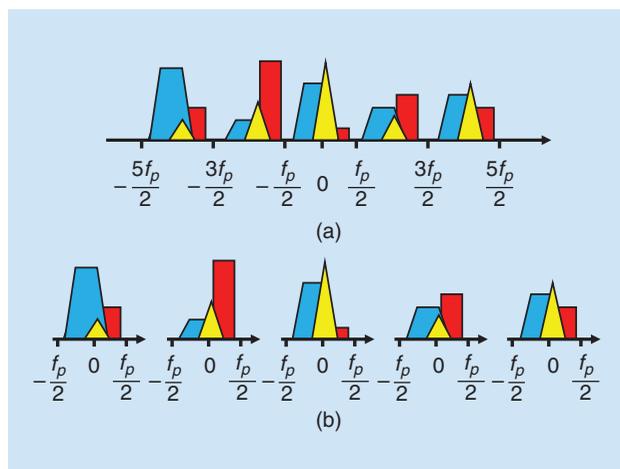
MWC sampling

The MWC [26] exploits the blind recovery ideas developed in [25] 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 the baseband with different weights for each frequency interval. These characteristics of the MWC enable practical hardware implementation, which will be described in the "MWC Hardware" section.

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

$$p_i(t) = \sum_{l=-\infty}^{\infty} a_{il} e^{j\frac{2\pi}{T_p}lt}. \tag{12}$$

The mixing process aliases the spectrum such that each band appears at a baseband. The signal then goes through an LPF with cutoff 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, i.e., (6) in frequency and (7) in time, as shown in Figure 7. 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 implementation complexity. Reducing the number of channels is a crucial challenge for practical implementation of a CR receiver. The MWC architecture presents an interesting flexibility that permits trading between physical channels and sampling rate, allowing a drastic reduction in the number of channels, even down to a single channel.

Consider a configuration where $f_s = qf_p$ with an odd q. In this case, the ith physical channel provides q equations over $\mathcal{F}_p = [-f_p/2, f_p/2]$, as illustrated in Figure 8. Conceptually, M physical channels sampling at rate $f_s = qf_p$ are then equivalent to Mq channels sampling at $f_s = f_p$. The number of channels is thus reduced at the expense of the 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 previously in the "Multicoset Sampling" section, in the context of multicoset sampling. At its brink, this strategy allows collapsing a system with M channels to a single branch with
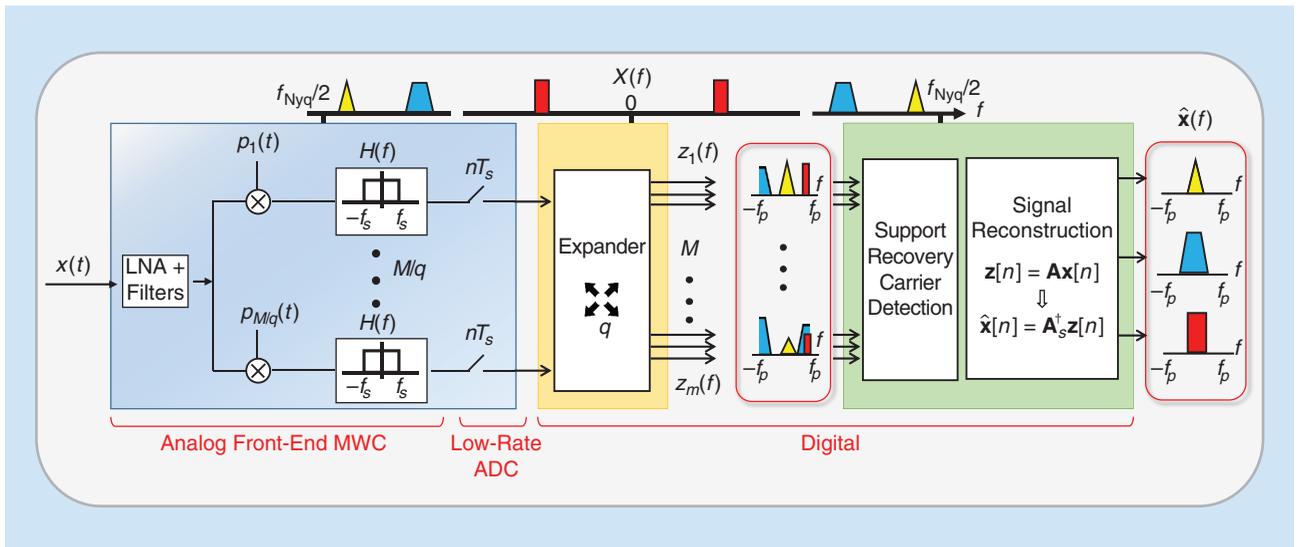
FIGURE 9. A schematic implementation of the MWC analog sampling front end and digital signal recovery from low-rate samples. LNA: low-noise amplifier.

the sampling rate $f_s = Mf_p$ (further details can be found in [26], [49], and [50]).

The MWC sampling and recovery processes are illustrated in Figure 9. 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. Additionally, 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 complexity of the hardware implementation. However, the choice of appropriate periodic functions $p_i(t)$ to ensure correct recovery is challenging. Some guidelines are provided in [43], [51], and [52].

Uniform-linear-array-based MWC

An alternative sensing configuration that is composed of a uniform linear array (ULA) and relies on the sampling paradigm of the MWC is presented in [53]. 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 of arrival (AOA) of the signal $x(t)$. This system, illustrated in Figure 10, capitalizes on the different accumulated phases of the input signal between sensors, given by $e^{j2\pi f_i \tau_m}$, where

$$\tau_m = \frac{dm}{c} \cos(\theta) \qquad (13)$$

is the delay at the mth sensor with respect to the first one. Each sensor implements one channel of the MWC, i.e., 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 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, is
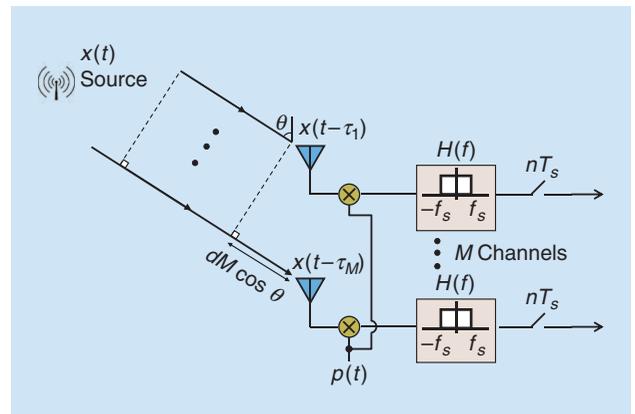
FIGURE 10. A uniform linear array configuration with M sensors, with the distance d between two adjacent sensors. Each sensor includes an analog front end composed of a mixer with the same periodic function $p(t)$, an LPF, and a sampler at the rate f_s.

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. Because all of 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. This alternative architecture benefits from the same flexibility as the standard MWC in terms of collapsing the channels, which translates here into reducing the antennas. This leads to a tradeoff between hardware complexity, governed by the number of antennas, and SNR. Finally, as will be discussed later in the "Joint-Carrier Frequency and Direction Estimation" section, the modified system can be easily extended to enable joint spectrum sensing and DOA estimation.

Similar to the previous schemes, the samples $\mathbf{z}(f)$ can be expressed as a linear transformation of the unknown vector of slices $\mathbf{x}(f)$, such that
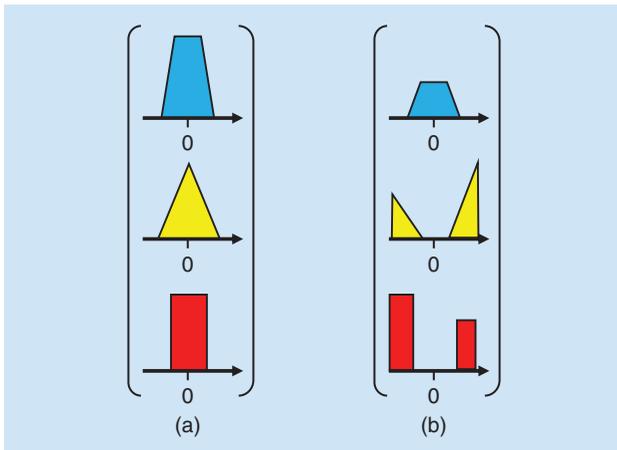
FIGURE 11. (a) The original source signals at baseband (before modulation) and (b) the output signals at baseband $x(f)$ after modulation, mixing, filtering, and sampling.

$$\mathbf{z}(f) = \mathbf{A}\mathbf{x}(f), \quad f \in \mathcal{F}_s. \tag{14}$$

Here, $\mathbf{x}(f)$ is a nonsparse vector that contains cyclic shifted, scaled, and sampled versions of the active bands, as shown in Figure 11. In contrast to the previous schemes, in this configuration, the matrix $\mathbf{A}$, defined by

$$A = \begin{pmatrix} e^{j2\pi f_1 \tau_1} & \cdots & e^{j2\pi f_N \tau_1} \\ \vdots & & \vdots \\ e^{j2\pi f_1 \tau_M} & \cdots & e^{j2\pi f_N \tau_M} \end{pmatrix}, \tag{15}$$

depends on the unknown carrier frequencies.

Two approaches are presented in [53] to recover the carrier frequencies of the transmissions composing the input signal. The first is based on CS algorithms and 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 (6) or (7), which can be solved using the recovery paradigm from [25], described in the context of multicoset sampling. In the second technique, the grid assumption is dropped, and ESPRIT [40] is used to estimate the carrier frequencies. This approach first computes the sample covariance of the measurements $\mathbf{z}[n]$ as

$$\mathbf{R} = \sum_n \mathbf{z}[n]\mathbf{z}^H[n] \tag{16}$$

and performs a singular value decomposition. The nonzero 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 (14).

The minimal number of sensors required by both reconstruction methods in a noiseless setting is $M = 2K$, with each sensor sampling at the minimal rate of $f_s = B$ to allow for perfect signal recovery [53]. The proposed system thus achieves the minimal sampling rate $2KB$ derived in [25]. We note that the expander strategy proposed in the context of the MWC can be applied in this configuration as well.

MWC hardware

MWC prototype

One of the main aspects that distinguishes the sub-Nyquist MWC from other sampling schemes is its practical implementation [49], proving the feasibility of sub-Nyquist sampling even under distorting effects of analog components and physical phenomena. A hardware prototype, shown in Figure 12, was developed and built according to the block diagram in Figure 9. In particular, the system receives an input signal with a Nyquist rate of 6 GHz, a spectral occupancy of up to 200 MHz, and samples at an effective rate of 480 MHz, i.e., only 8% of the Nyquist rate and 2.4 times the Landau rate.
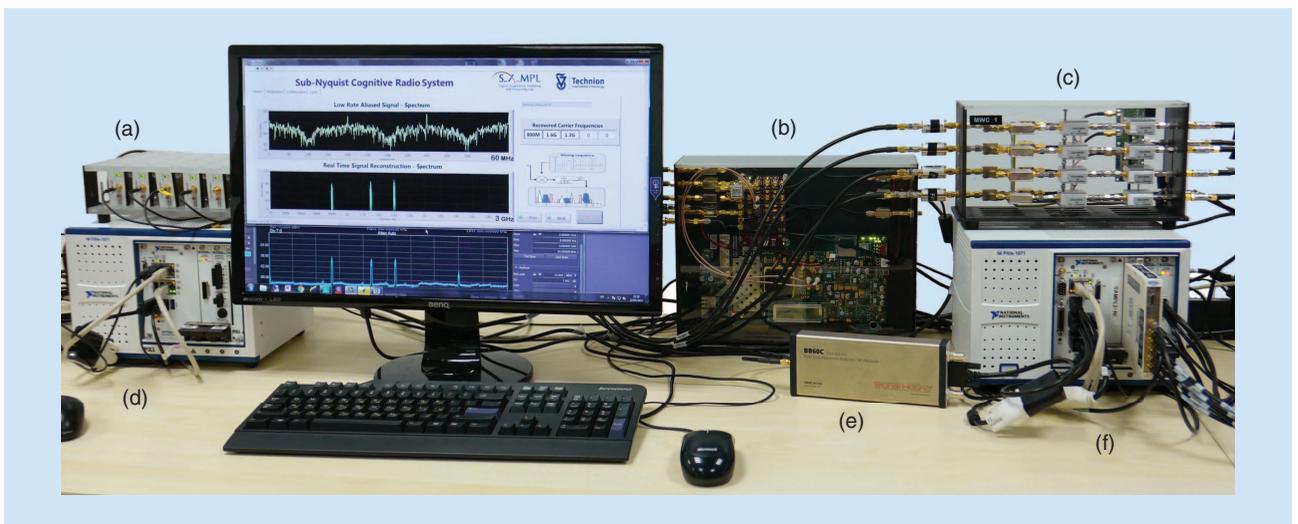
FIGURE 12. The MWC CR system prototype: (a) vector signal generators, (b) an FPGA mixing sequences generator, (c) an MWC analog front-end board, (d) an RF combiner, (e) a spectrum analyzer, and (f) an ADC and digital signal processor digital signal processor.

This rate constitutes a relatively small oversampling factor of 20% with respect to the theoretical lower sampling bound. This section describes the different elements of the hardware prototype, which are shown in Figure 12, and explains 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 [49] that implements the sub-Nyquist analog front end. The card uses a high-speed one-to-four analog splitter that duplicates the wide-band signal to $M = 4$ channels, with an expansion factor of $q = 5$, yielding $M_q = 20$ 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 can be seen in Figure 13.

The modulated signal next passes through an analog anti-aliasing LPF. The antialiasing filter must be characterized by both an almost linear phase response in the passband from 0 to 50 MHz and an attenuation of more than 20 db at $f_s/2 = 60$ MHz. A seventh-order type I Chebyshev LPF with a cutoff frequency (-3 dB) of 50 MHz was chosen for the implementation. After impedance and gain corrections, the spectral content of the signal is limited to 50 MHz and contains a linear combination of the occupied bands with different amplitudes and phases, as seen in Figure 7. 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.

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 [54], which are commonly used in telecommunication (code division multiple access) and satellite navigation (global positioning system). Mixing sequences based on Gold codes were found to give good results in the MWC system [55], primarily due to their small cross correlations.

Because Gold codes are binary, the mixing sequences are restricted to alternating ± 1 values. This fact allows the digital generation of the sequences on a dedicated field-programmable gate array (FPGA). Alternatively, they can be implemented on a small microchip with very low power and complexity. The added benefit of producing the mixing sequences on such a platform is that the entire sampling scheme may 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. An XiLinX VC707 FPGA evaluation kit 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 and the low-rate samplers operating at $(q + 1)f_p$ (due to intended oversampling) are fully synchronized 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 a dc-coupled ADC. Because the digital processing is performed at the low rate f_s, very low computational load is required 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 **A** is computed once offline, using the calibration process outlined in [50] and described in "Hardware Calibration of the MWC."
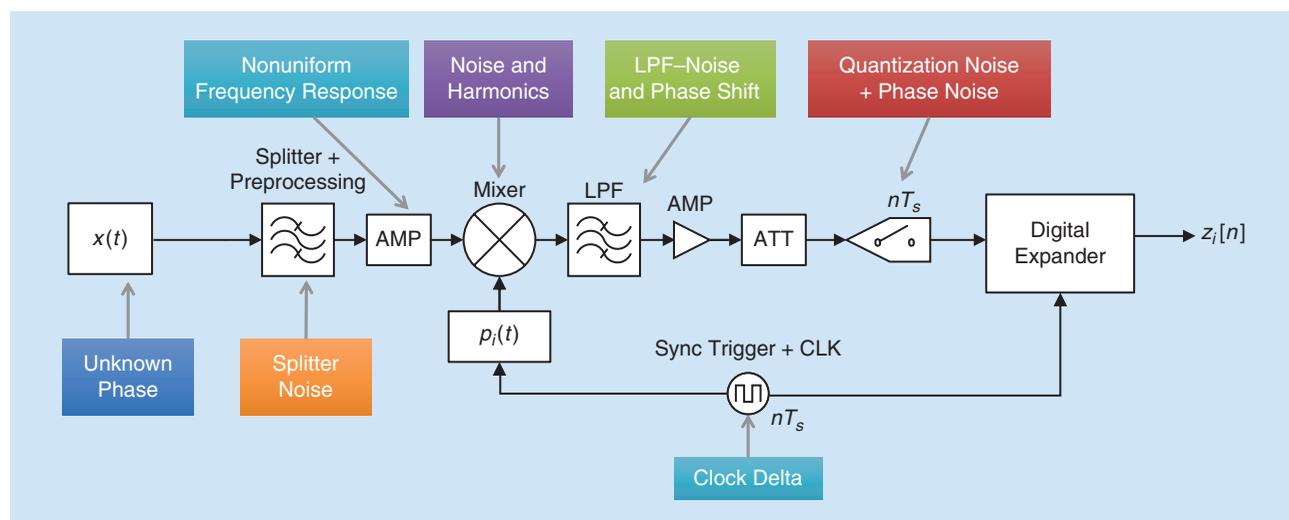
FIGURE 13. The hardware RF chain detailed schematics, including the amplifiers, attenuators, filters, mixers, samplers, and synchronization signals required for precise and accurate operation. The distortions induced by each component are indicated as well. Freq.: frequency; ATT: attenuator; CLK: clock.

In the sampling system described previously in the "MWC Sampling" section, the system matrix **A** is theoretically known and contains the Fourier series coefficients of the mixing sequences, such that

$$\mathbf{A}_{il} = c_{il} = \frac{1}{T_p} \int_0^{T_p} p_i(t) \, e^{-i \frac{2\pi}{T_p} l t} \, dt. \tag{S6}$$

When calculating the matrix coefficients using (S6), perfect support recovery and signal reconstruction are guaranteed both theoretically and by numerical verification performed in software simulations. However, in practice, analog and physical distortions and imperfections affect the mixing and sampling process, and some modeling assumptions that describe the system matrix in theory no longer hold. The main effects that distort the transfer function are described as follows.

- The mixing procedure introduces nonlinearities. In general, mixers are intended to modulate narrow-band signals with one sine carrier, as opposed to our mixing sequences that effectively contain over 100 different harmonics.
- The analog filters have a nonflat frequency response, both in amplitude and phase.
- The actual design uses amplifiers and attenuators. These components exhibit a nonlinear frequency response.
- The phase noise and jitter, due to variations in components, cables, and clock deltas, lead to unknown and varying time shifts between the mixing and sampling stages.

An accurate method for estimating the effective **A** is crucial for successful support recovery and signal reconstruction. An adaptive calibration scheme is proposed in [50]. The calibration procedure estimates the elements of **A** with no prior knowledge on the mixing series $p_i(t)$, except for their period length T_p.

FIGURE S3. The hardware reconstruction success rate of the calibrated matrix versus the theoretical matrix.

Because our system is not time invariant (e.g., samplers) or linear (e.g., mixers), one cannot find the entries of the system matrix by simply measuring its response to an impulse. To circumvent this difficulty, the system's response is investigated for every frequency band of the spectrum by injecting known sinusoidal inputs sequentially. In each iteration, the following input

$$x_l(t) = \alpha_l \sin[2\pi(lf_p + f_0)t], l \in [0, 1, \dots, L_0] \tag{S7}$$

is fed to the system. Here, $f_0 = 0.1 f_p$ was heuristically chosen, and the amplitudes α_l increase with l to compensate for the attenuation of the Fourier coefficients of the mixing sequences at high frequencies. Every sine wave corresponds to a specific spectral band and translates to a relevant column of the matrix **A**. The sub-Nyquist samples are then digitally processed to estimate the system matrix coefficients column by column. Performance of the calibrated system is illustrated in Figure S3.

Support recovery

The prototype is fed with RF signals composed of up to five carrier transmissions with an unknown total bandwidth occupancy of up to 200 MHz and a Nyquist rate of 6 GHz. An RF input $x(t)$ is generated using vector signal generators, 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 for testing of the system's ability to rapidly sense the input spectrum and adapt to changes, as required by modern CR standards (see "IEEE 802.22 Standard for WRAN" for details). In addition, the described setup can simulate more complex scenarios, including collaborative-spectrum sensing [56], [57], joint DOA estimation [53], cyclostationary-based detection [58], and various modulation schemes, such as phase-shift keying (PSK), orthogonal fre-

quency-division multiplexing (OFDM), and more, for verifying sub-Nyquist data reconstruction capabilities.

Support recovery is digitally performed on the low-rate samples, as previously presented in the "Sub-Nyquist Sampling for CR" section in the context of multicoset sampling. The prototype successfully recovers the support of the transmitted bands when SNR levels are above 15 dB, as demonstrated in Figure 14 (see Table 1 for Figure 14 parameters). More sophisticated detection schemes, such as cyclostationary detection, allow perfect support recovery from the same sub-Nyquist samples in lower SNR regimes of 0–15 dB, as seen in the "Cyclostationary Detection" section, and will be further discussed next.

The main advantage of the MWC is that sensing is done 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 small without impeding the performance, as shown in Figure 15, where the support of signals with very low bandwidth (just 10% occupancy within the 20-MHz band) is correctly detected.

Signal reconstruction

Once the support is recovered, the data are reconstructed from the sub-Nyquist samples by applying (10). 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 inherent bandwidth of the slices $\mathbf{z}(f)$.

The prototype's digital recovery stage is further expanded to support decoding of common communication modulations, including binary PSK (BPSK), quarternary PSK (QPSK), quadrature amplitude modulation, and OFDM. An example for the decoding of three QPSK modulated bands is presented in Figure 16. The I-axis (in phase) and Q-axis (quadrature) (I/Q) constellations are shown after reconstructing the original transmitted signals $\mathbf{x}_S$ using (10), from their low-rate and aliased sampled signals $\mathbf{z}_n$. The I/Q constellations of the baseband signals are displayed, each individually translated using a general QPSK decoder. In this example, the user broadcasts text strings, which are then deciphered and displayed on screen. There are no restrictions regarding the modulation type, bandwidth, or other parameters because the baseband information is precisely reconstructed regardless of its respective content. Therefore, any digital modulation method, as well as analog broadcasts, can be transmitted and deciphered without the loss of information by applying any desirable decoding scheme directly to 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 performs real-time spectrum sensing at sub-Nyquist rates, and the second is a receiver that is able to simultaneously decode multiple data transmissions, regardless of their carrier frequencies, while adapting to 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 and deciphers 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 are nonzero only in the bands of inter-

> The main advantage of the MWC is that sensing is done 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.
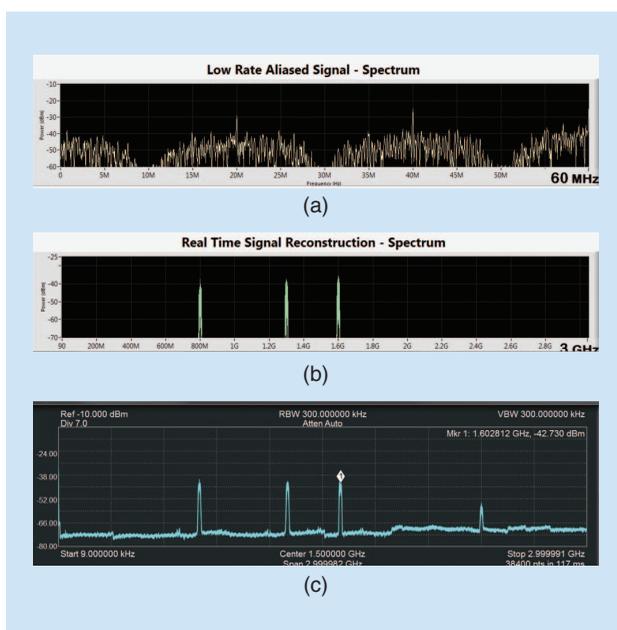
FIGURE 14. A screenshot from the MWC recovery software: (a) low-rate samples acquired from one MWC channel at rate 120 MHz; (b) digital reconstruction of the entire spectrum, performed from sub-Nyquist samples; and (c) true input signal $x(t)$ showed using a fast spectrum analyzer.

est, increasing SNR, and the support recovery stage is not needed [59].

Statistics detection

In the previous section, 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 and not necessarily a perfect reconstruction of the PUs' transmissions. In this regard, 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 [32] (see "Cyclostationarity") signals, respectively. Instead of recovering the signal from the low-rate samples, its statistics are reconstructed, and its support is estimated [58], [60]–[71].

Recovering second-order statistics rather than the signal itself benefits from two main advantages. First, it allows a further reduction of the sampling rate, as we will discuss later in this section. Intuitively, statistics have less degrees of freedom than the signal itself, requiring fewer 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 nonsparse signals can be recovered

Cyclostationary processes have statistical characteristics that vary periodically with time. Examples of periodic phenomena that give rise to random data abound in engineering and science. In particular, in communications, periodicity arises from the underlying data modulation mechanisms, such as carrier modulation, periodic keying, or pulse modulation [81]. A characteristic function of such processes, referred to as the *cyclic spectrum*, extends the traditional power spectrum to a two-dimensional map, with respect to angular and cyclic frequency variables. The cyclic spectrum exhibits spectral peaks at certain frequency locations, called the *cyclic frequencies*, which are determined by the signal's parameters, particularly the carrier frequency and symbol rate [72].

A process $x(t)$ is said to be wide-sense cyclostationary with the 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 the period T_0 [32], i.e.,

$$\mu_x(t + T_0) = \mu_x(t), R_x(t + T_0, \tau) = R_x(t, \tau), \quad \text{(S8)}$$

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:

$$R_x(t, \tau) = \sum_\alpha R_x^\alpha(\tau) e^{j2\pi\alpha t}, \quad \text{(S9)}$$

where the sum is over integer multiples of the fundamental frequency $1/T_0$ and the Fourier coefficients $R_x^\alpha(\tau)$ are referred to as *cyclic autocorrelation functions*. The cyclic spectrum is obtained by taking the Fourier transform of the cyclic autocorrelation functions with respect to τ, i.e.,

$$S_x^\alpha(f) = \int_{-\infty}^{\infty} R_x^\alpha(\tau) e^{-j2\pi f\tau} d\tau, \quad \text{(S10)}$$

where α is referred to as the *cyclic frequency* and f is the *angular frequency* [32]. If there is more than one fundamental frequency $1/T_0$, then the process $x(t)$ is said to be polycyclostationary in the wide sense. In this case, the cyclic spectrum contains harmonics (integer multiples) of each of the fundamental cyclic frequencies [72]. 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

$$u(t) \triangleq x(t) e^{-j\pi\alpha t}, \quad v(t) \triangleq x(t) e^{+j\pi\alpha t}. \quad \text{(S11)}$$

Then, from [73], it holds that
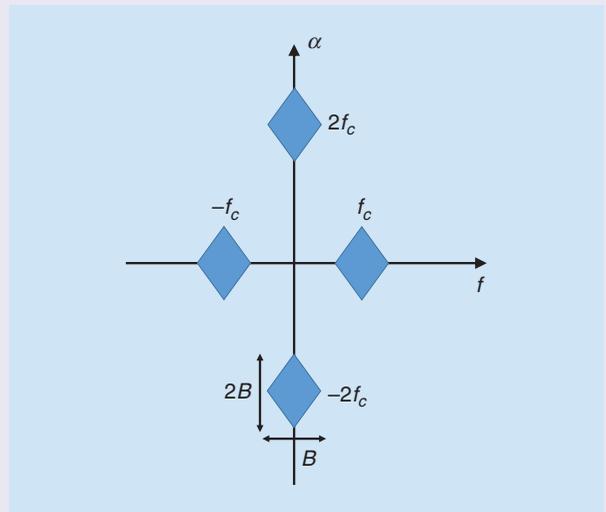
FIGURE S4. The support region of the cyclic spectrum of a bandpass cyclostationary signal with the carrier frequency f_c and bandwidth B.
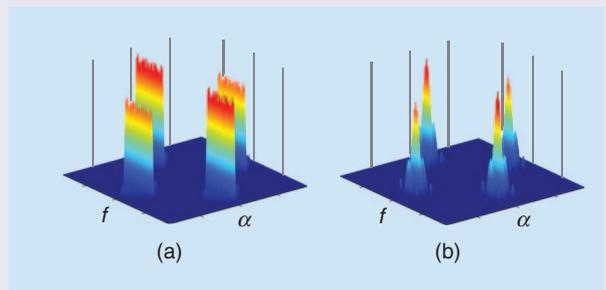
FIGURE S5. The cyclic-spectrum magnitude of bandpass signals with (a) amplitude modulation and (b) BPSK modulation.

$$S_x^\alpha(f) = S_{uv}(f) = \mathbb{E}\left[X\left(f + \frac{\alpha}{2}\right) X^*\left(f - \frac{\alpha}{2}\right) \right]. \quad \text{(S12)}$$

As expressed in (S12), the cyclic spectrum $S_x^\alpha(f)$ measures correlations between different spectral components of $x(t)$. Stationary signals, which do not exhibit a spectral correlation between distinct frequency components, appear only at $\alpha = 0$. This property is the key to the robust detection of cyclostationary signals in the presence of stationary noise, in low signal-to-noise ratio 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 Figure S4. Thus, the cyclic spectrum $S_x^\alpha(f)$ of a multiband signal with K uncorrelated transmissions is supported over $4K$ diamond-shaped areas. Figure S5 illustrates the cyclic spectrum of two modulation types, amplitude modulation and binary phase-shift keying (BPSK).

Table 1. The MWC parameters used in the setup of Figure 14.

	Value	Notes
f_s	120 MHz	$(q+1)f_P$ —Sampling rate
f_P	20 MHz	$1/T_P$
q	5	The expansion factor
M/q	4	The number of hardware channels
f_{max}	3 GHz	$f_{max} = f_{Nyq}/2$
B	18.5 MHz	The bandwidth on each carrier
M_P	305	The number of ±1 intervals in each period of $p_i(t)$

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

from samples obtained below the Nyquist rate [58], [60], [61], [65], [66], [70]. 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, robustness to noise is increased due to the averaging performed when estimating 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 them to separate from stationary noise, leading to better detection in low SNR regimes [72]. In this section, we first review power-spectrum detection techniques in stationary settings and then extend these to the cyclic-spectrum detection of cyclostationary signals.

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 [73], the Fourier transform of a wide-sense stationary signal is a nonstationary white process, such that

$$\mathbb{E}[X(f_1)X^*(f_2)] = S_x(f_1)\delta(f_1 - f_2). \tag{17}$$

Here, the power spectrum $S_x(f)$ of $x(t)$ is the Fourier transform of its autocorrelation $r_x(\tau)$. Thus, it is obvious that the support of $S_x(f)$ is identical to $X(f)$. In this case, we will see that the autocorrelation matrix of the N spectrum frequency slices of $x(t)$ contained in $\mathbf{x}(f)$ is diagonal and presents only N degrees of freedom, which is key to the sampling-rate reduction.

Another intuitive interpretation can be given in the time domain. The autocorrelation of stationary signals $r_x(\tau) = \mathbb{E}[x(t)x(t-\tau)]$ is a function only of the time lags τ and not of the sampling times t. Therefore, the autocorrelation recovery capabilities of a sampling set are determined by the density of the associated difference set—the set that contains the time lags. The cardinality of the difference set, which depends on the choice of sampling times, is, in the worst case, equal to twice that of the original set. This corresponds, e.g., to a
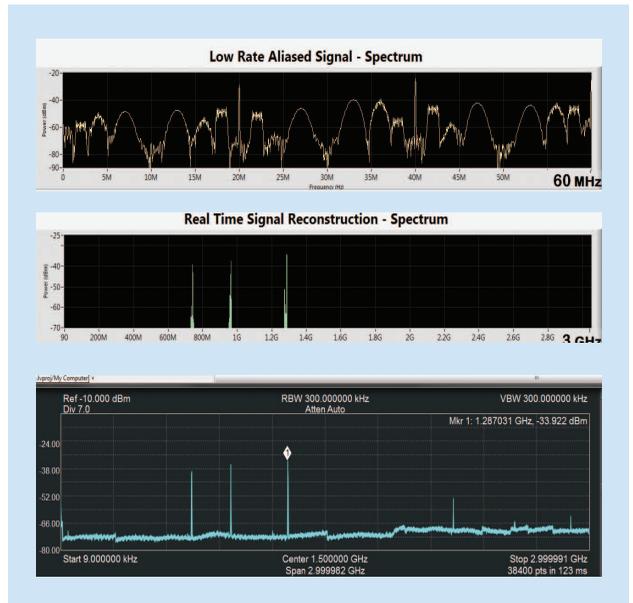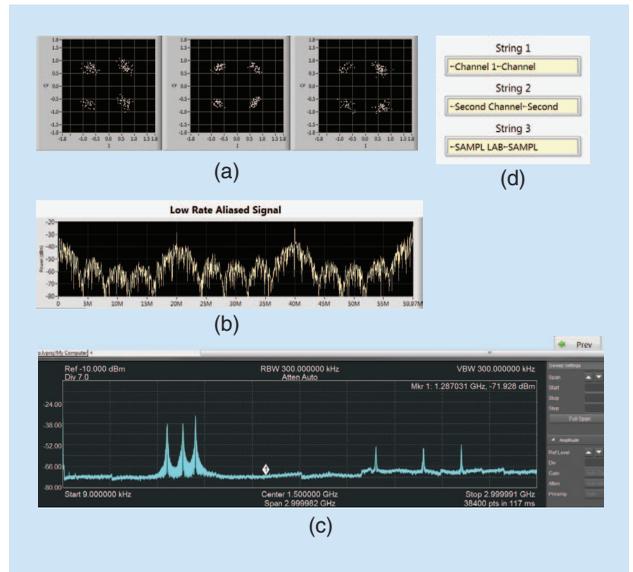
FIGURE 16. The demodulation, reconstruction, and detection of $N_{sig} = 3$ inputs from sub-Nyquist samples using the MWC CR prototype: (a) I/Q phase diagrams, showing the modulation pattern of the transmitted bands after reconstruction from the low-rate samples, (b) sub-Nyquist samples from an MWC channel $z_i[n]$ in the Fourier domain, (c) signal sampled by an external spectrum analyzer showing the entire bandwidth of 3 GHz, and (d) information sent on each carrier, proving successful reconstruction.

uniform sampling set. As a consequence, when the sampling scheme is not tailored to a power-spectrum recovery, the minimal required sampling rate in the blind setting is the Landau rate [66]. However, for an appropriate choice of sampling times [62], [74], the cardinality of the difference set can be the square of the original sampling set. In this case, under certain

conditions, the density of the different set can be arbitrarily high even if the density of the sampling set reaches zero. With an appropriate design, the autocorrelation or power spectrum may thus be estimated from samples with arbitrarily low average sampling rates [60]–[62], [74]–[76] 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.

Power-spectrum recovery

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

To recover $S_x(f)$ from the low-rates samples $\mathbf{z}(f)$ obtained via multicoset sampling or the MWC [66], consider the correlation matrix of the latter $\mathbf{R}_\mathbf{z}(f) = \mathbb{E}[\mathbf{z}(f)\mathbf{z}^H(f)]$. From (6), $\mathbf{R}_\mathbf{z}(f)$ can be related to correlations between the slices $\mathbf{x}(f)$, i.e., $\mathbf{R}_\mathbf{x}(f) = \mathbb{E}[\mathbf{x}(f)\mathbf{x}^H(f)]$, as follows:

$$\mathbf{R}_\mathbf{z}(f) = \mathbf{A}\mathbf{R}_\mathbf{x}(f)\mathbf{A}^H, \quad f \in \mathcal{F}_s. \quad (18)$$

From (17), the correlation matrix $\mathbf{R}_\mathbf{x}(f)$ is diagonal and contains the power spectrum $S_x(f)$ at the corresponding frequencies, as

$$\mathbf{R}_{\mathbf{x}(i,i)}(f) = S_x\left(f + if_s - \frac{f_{\mathrm{Nyq}}}{2}\right), \quad f \in \mathcal{F}_s. \quad (19)$$

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 its diagonal by using $\mathbf{r}_\mathbf{x}(f)$, (18) can be reduced to

$$\mathbf{r}_\mathbf{z}(f) = (\bar{\mathbf{A}} \odot \mathbf{A})\mathbf{r}_\mathbf{x}(f), \quad (20)$$

where $\mathbf{r}_\mathbf{z}(f) = \mathrm{vec}(\mathbf{R}_\mathbf{z}(f))$ 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 [77].

Generic choices of the sampling parameters, either mixing sequences or cosets, which are only required to ensure that $\mathbf{A}$ is a full spark, are investigated in [66]. The Khatri–Rao product $(\bar{\mathbf{A}} \odot \mathbf{A})$ is a full spark as well if $M > N/2$, i.e., 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 (20) is thus equal to the Landau rate KB, i.e., half the rate required for signal recovery [66]. The recovery of $\mathbf{r}_\mathbf{x}(f)$ is performed using the procedure presented in the context of signal recovery on (20), i.e., CTF, support recovery, and power-spectrum reconstruction (rather than signal reconstruction).

The same result for the minimal sampling rate is valid for nonsparse signals, for which KB is f_{Nyq} [66]. The power spectrum of such signals can 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 its power spectrum by exploiting stationarity. In this case, the system (20) is overdetermined, and $\mathbf{r}_\mathbf{x}(f)$ is obtained by a simple pseudoinverse operation.

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 or windows of length N_s samples. In [66], different choices of N_f and N_s are examined for a fixed sensing time. To estimate the autocorrelation matrix $\mathbf{R}_\mathbf{z}(f)$ in the frequency domain, estimates of $\mathbf{z}_i(f), 1 \leq i \leq M$, denoted by $\hat{\mathbf{z}}_i(f)$, are first computed using the fast Fourier transform on the samples $z_i[n]$ over a finite time window. The elements of $\mathbf{R}_\mathbf{z}(f)$ are then obtained by averaging the correlations between $\hat{\mathbf{z}}_i(f)$ over the frames. In practice, the number of samples dictates the number of discrete Fourier transform coefficients in the frequency domain and, therefore, the resolution of the reconstructed power spectrum. Once $\hat{\mathbf{r}}_\mathbf{x}(f)$ is reconstructed, several detection statistics can be computed, such as power or eigenvalue-based test statistics [78].

Sampling scheme design

Sampling approaches specifically designed for estimating the autocorrelation of stationary signals at much finer lags than the sample spacings have recently been studied in detail [62], [74], [75], [79]. These rely on the observation that the autocorrelation is a function of the lags only, i.e., the differences between the pairs of sampling times. Thus, the correlation may be estimated at time lags contained within the difference set, also referred to as the *difference coarray*, which is the set composed of all differences between the pairs of sampling times. Because the size of the difference set may be greater than that of the original sampling set, depending on the choice of sampling times, we may need fewer sampling times for autocorrelation recovery than for signal recovery. Therefore, the sampling times should be carefully chosen so as to maximize the cardinality of the difference set.

The first approach we present, for autocorrelation recovery at sub-Nyquist rates, adopts multicoset sampling and designs the cosets in (5) 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 (18) or (20) is a partial Fourier matrix with the (i,k)th element $e^{j(2\pi/N)c_i k}$. A typical element of $(\bar{\mathbf{A}} \odot \mathbf{A})$ is then $e^{j(2\pi/N)(c_i-c_j)k}$. 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 nonsparse setting and at least Landau for a sparse signal with unknown support. This happens, e.g., if we select the first or last M cosets or if we only keep 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 [62], [79] using minimal linear and circular sparse rulers [80]. A linear sparse ruler is a set of integers from the interval $[0, N]$, such that the

associated difference set contains all integers in $[0, N]$. Intuitively, it may be viewed as a ruler with some marks erased but still able to measure all integer distances between zero and its length. For example, consider the minimal sparse ruler of length $N = 10$. This ruler requires $M = 6$ marks, as shown in Figure 17. All of 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:

$$\frac{\sqrt{\tau(N-1)}}{N} \leq \frac{M}{N} \leq \frac{\sqrt{3(N-1)}}{N}, \tag{21}$$

where $\tau \approx 2.4345$ [79]. A circular or modular sparse ruler extends this idea to include periodicity. Such designs that seek minimal sparse rulers, i.e., rulers with a minimal number of marks M, can achieve compression ratios M/N on the order of $\sqrt{N}$. As N increases, the compression ratio may become arbitrarily low.

Two additional sampling techniques specifically designed for autocorrelation recovery are nested arrays [74] and coprime sampling [75], presented in the context of autocorrelation estimation as well as beamforming and DOA estimation applications. In nested and coprime structures, similar to multicosets, the corresponding coarrays 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 quickly review both sampling structures and their corresponding difference coarrays and show how the autocorrelation of an arbitrary stationary signal can be recovered on the Nyquist grid from low-rate samples.

In its simplest form, the nested-array [74] structure has two levels of sampling density. The first-level samples are at the N_1 locations $\{\ell T_{\mathrm{Nyq}}\}_{1 \leq \ell \leq N_1}$, and the second-level samples are at the N_2 locations $\{(N_1 + 1)k T_{\mathrm{Nyq}}\}_{1 \leq k \leq N_2}$. This nonuniform sampling is then repeated with the period $(N_1 + 1)N_2 T_{\mathrm{Nyq}}$. Because there are $N_1 + N_2$ samples in intervals of length $(N_1 + 1)N_2 T_{\mathrm{Nyq}}$, the average sampling rate of a nested-array sampling set is given by

$$f_s = \frac{N_1 + N_2}{(N_1 + 1)N_2 T_{\mathrm{Nyq}}} \approx \frac{1}{N_1 T_{\mathrm{Nyq}}} + \frac{1}{N_2 T_{\mathrm{Nyq}}}. \tag{22}$$

This rate can be arbitrarily low because N_1 and N_2 may be as large as we choose at the expense of latency.

Now, consider the difference coarray, which has a contribution from the cross differences and the self-differences. The nonnegative cross differences, normalized by T_{Nyq} for clarity, are given by

$$n = (N_1 + 1)k - \ell, \quad 1 \leq k \leq N_2, 1 \leq \ell \leq N_1. \tag{23}$$

All differences in the range $1 \leq n \leq (N_1 + 1)N_2 - 1$ are covered except for multiples of $N_1 + 1$. These are precisely the self-differences among the second array. As a result, the difference coarray is a filled array composed of all integers
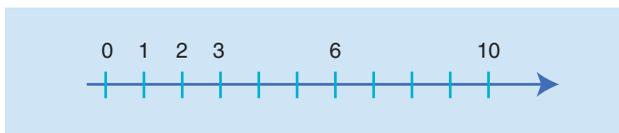
FIGURE 17. A minimal sparse ruler of order $M = 6$ and length $N = 10$.

$-[(N_1 + 1)N_2 - 1] \leq n \leq [(N_1 + 1)N_2 - 1]$. Going back to our autocorrelation estimation problem, this result shows that, by proper averaging, we can estimate $R(\tau)$ at any lag τ on the Nyquist grid from nested-array samples with arbitrarily low sampling rate as

$$\hat{R}[n] = \begin{cases} \frac{1}{Q} \sum_{q=0}^{Q-1} x(\tilde{N}(k + q)) x^*(\ell + \tilde{N}q), & n = \tilde{N}k - \ell, \\ \frac{1}{Q} \sum_{q=0}^{Q-1} x(\tilde{N}(k + q)) x^*(\tilde{N}(\ell + q)), & n = \tilde{N}(k - \ell), \end{cases} \tag{24}$$

where Q is the number of snapshots used for averaging and $\tilde{N} = N_1 + 1$. Here, k and ℓ are such that (23) holds.

Another sampling technique designed for autocorrelation recovery is coprime sampling. It involves two uniform sampling sets with spacing $N_1 T_{\mathrm{Nyq}}$ and $N_2 T_{\mathrm{Nyq}}$, respectively, where N_1 and N_2 are coprime integers. The average sampling rate of such a sampling set, given by

$$f_s = \frac{1}{N_1 T_{\mathrm{Nyq}}} + \frac{1}{N_2 T_{\mathrm{Nyq}}}, \tag{25}$$

can, again, be made arbitrarily small compared to the Nyquist rate $1/T_{\mathrm{Nyq}}$ by choosing arbitrarily large coprime numbers N_1 and N_2. The associated difference set, normalized by T_{Nyq}, is composed of elements of the form $n = N_1 k - N_2 \ell$. Because N_1 and N_2 are coprime, there exist integers k and ℓ such that the previous difference achieves any integer value n. Therefore, the autocorrelation may be estimated by proper averaging, as

$$\hat{R}[n] = \frac{1}{Q} \sum_{q=0}^{Q-1} x(N_1(k + N_2 q)) x^*(N_2(\ell + N_1 q)), \tag{26}$$

where k and ℓ are such that $n = N_1 k - N_2 \ell$.

The main drawback of both techniques, besides the practical issue of analog bandwidth that is similar to multicoset sampling, is the added latency required for averaging. Furthermore, in practice, synchronizing ADCs with different sampling rates can be challenging. Finally, 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.

Cyclostationary detection
Communication signals often exhibit statistical periodicity, due to schemes like carrier modulation or periodic keying [81]. Therefore, such signals are cyclostationary processes

[72]. A characteristic function of these processes, the cyclic spectrum $S_x^\alpha(f)$, exhibits spectral peaks at the cyclic frequencies, which are determined by the signal's parameters of periodicity, such as the carrier frequency and symbol rate [72]. The formal definition of the cyclic spectrum is presented in "Cyclostationarity." Cyclostationary-based detection distinguishes between the signals of interest, assumed to be cyclostationary, and stationary noise and interference by measuring spectral correlation [72]. Because stationary noise and interference exhibit no spectral correlation, as shown in (17), such detectors are highly robust to noise. With noise enhancement being one of the main limitations of sub-Nyquist sampling-based techniques, cyclostationary detection performed on the reconstructed cyclic spectrum from compressive measurements is a natural candidate for improving sub-Nyquist detection.

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
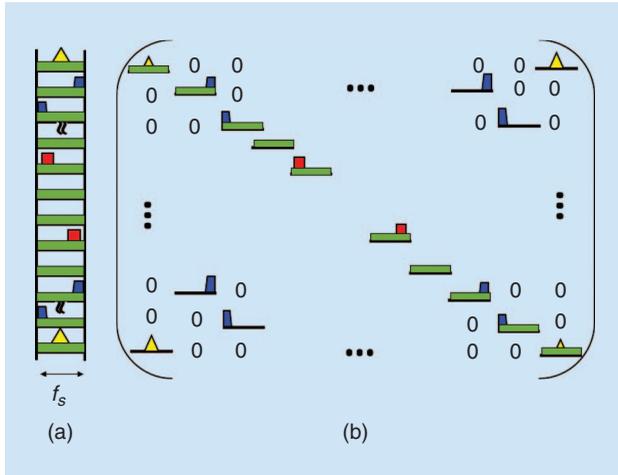
FIGURE 18. (a) The spectrum slices vector $x(\tilde{f})$ and (b) correlated slices of $x(\tilde{f})$ in the matrix $R_x^0(\tilde{f})$.
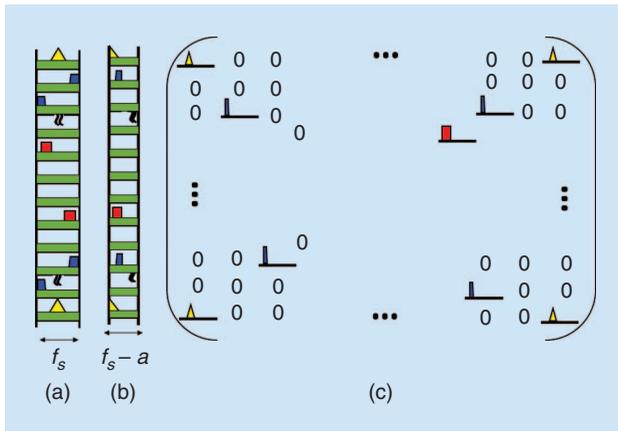
FIGURE 19. (a) The spectrum slices vector $x(\tilde{f})$; (b) spectrum slices shifted vector $x(\tilde{f}+a)$, for $a = f_s/2$; and (c) correlated slices of $x(\tilde{f})$ and $x(\tilde{f}+a)$ in the matrix $R_x^a(\tilde{f})$, with $a = f_s/2$.

the samples obtained using the MWC or multicoset sampling. To that end, we first related $S_x(f)$ to the slices' correlation matrix $\mathbf{R_x}(f)$ and then recovered the latter from $\mathbf{R_z}(f)$. 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, $\mathbf{R_x^a}(\tilde{f}) = \mathbb{E}[\mathbf{x}(\tilde{f})\mathbf{x}^H(\tilde{f}+a)]$ for $a \in [0, f_s]$ and $\tilde{f} \in [0, f_s - a]$. Next, similar to power-spectrum recovery, $\mathbf{R_x^a}(\tilde{f})$ is reconstructed from shifted correlations of the samples $\mathbf{R_z^a}(\tilde{f}) = \mathbb{E}[\mathbf{z}(\tilde{f})\mathbf{z}^H(\tilde{f}+a)]$. Once the cyclic spectrum $S_x^\alpha(f)$ is recovered, the transmissions' carriers and bandwidth may be estimated by locating its peaks. Because 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 (S12), presented in "Cyclostationary," implies that the elements in the matrix $\mathbf{R_x^a}(\tilde{f})$ are equal to $S_x^\alpha(f)$ at the corresponding α and f. Indeed, it can easily be shown [58] that

$$\mathbf{R_x^a}(\tilde{f})_{(i,j)} = S_x^\alpha(f), \qquad (27)$$

for

$$\alpha = (j - i)f_s + a$$
$$f = -\frac{f_{\text{Nyq}}}{2} + \tilde{f} - \frac{f_s}{2} + \frac{(j+i)f_s}{2} + \frac{a}{2}. \qquad (28)$$

Here, $\mathbf{R_x^a}(\tilde{f})_{(i,j)}$ denotes the (i, j)th element of $\mathbf{R_x^a}(\tilde{f})$. 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 $\mathbf{R_x^a}(\tilde{f})$, and vice versa. From (6) and similar to (18) in the context of power-spectrum recovery, we relate the shifted correlation matrices of $\mathbf{x}(f)$ and $\mathbf{z}(f)$ as

$$\mathbf{R_z^a}(\tilde{f}) = \mathbf{A}\mathbf{R_x^a}(\tilde{f})\mathbf{A}^H, \quad \tilde{f} \in [0, f_s - a] \qquad (29)$$

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

Recall that, in the context of stationary signals, $\mathbf{R_x}(f)$ is diagonal. Here, understanding the structure of $\mathbf{R_x^a}(\tilde{f})$ is more involved. In [58], it is shown that $\mathbf{R_x^a}(\tilde{f})$ contains nonzero elements on its 0, 1, and −1 diagonals and antidiagonals. Besides the nonzero entries being contained in only the three main and antidiagonals, the additional structure is exhibited, limiting the number of nonzero elements to two per row and column of the matrix $\mathbf{R_x^a}(\tilde{f})$. This pattern follows from two main considerations. First, each frequency component, i.e., 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 18 and 19 illustrate these correlations for $a = 0$ and $a = f_s/2$, respectively. First, in Figure 20, a sketch of the spectrum of $x(t)$, i.e., $X(f)$, is presented for the case of a sparse signal buried in stationary noise. It can be seen that frequency bands of $X(f)$ may either appear in one f_p slice or split between two slices at most because $f_p \geq B$. The

resulting vector of spectrum slices $\mathbf{x}(f)$ and the correlations between these slices without any shift, i.e., $\mathbf{R}_x^0(\tilde{f})$, are shown in Figure 18(a) and (b), respectively. In Figure 18(b), we observe that self-correlations appear on only the main diagonal because 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 zero antidiagonal, whereas those of the blue trapezes are contained in the –1 and +1 antidiagonals. The red rectangles do not contribute any cross correlations for $a = 0$.

Figure 19(a) and (b) shows the vector $x(\tilde{f})$ and its shifted version $x(\tilde{f}+a)$ for $a = f_s/2$, respectively. The resulting correlation matrix $\mathbf{R}_x^a(\tilde{f})$ appears in Figure 19(c). Here, the self-correlations of the yellow triangle appear in the main diagonal and the blue trapeze in the -1 diagonal. The cross correlations all appear in the antidiagonal for the shift $a = f_s/2$. Note that, because the noise is assumed to be wide-sense stationary, from (17), a noise-frequency component is correlated only with itself. Thus, $n(t)$ only contributes nonzero elements on the diagonal of $\mathbf{R}_x^0(\tilde{f})$.

To recover $\mathbf{R}_x^a(\tilde{f})$ from $\mathbf{R}_z^a(\tilde{f})$, structured CS techniques are used in [58] that aim at reconstructing a sparse matrix while taking into account its specific structure, as previously described. 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 of a slightly higher sampling rate, as shown in [58]. More precisely, in the presence of stationary noise, the cyclic spectrum may be reconstructed from samples obtained at four fifths 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 eight fifths of the Landau rate [58].

Carrier frequency and bandwidth estimation

Many detection and classification algorithms based on cyclostationarity have been proposed (see the reviews [32] and [33]). We first survey several detection and classification approaches and then explain why they do not adhere to CR requirements. 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 [82]–[87], which is the information that CRs need to determine in the first place. A second strategy focuses on a single transmission [84], [88], which does not fit the multiband model. Alternative approaches apply machine-learning tools to the modulation classification of a single signal with an unknown carrier frequency and symbol rate [89]–[92]. Besides being suitable for only a single transmission, these methods require a training phase. Thus, these techniques can only cope with

> **The detection performed on the cyclic spectrum is more robust to stationary noise than power-spectrum-based detection, at the expense of a slightly higher sampling rate.**
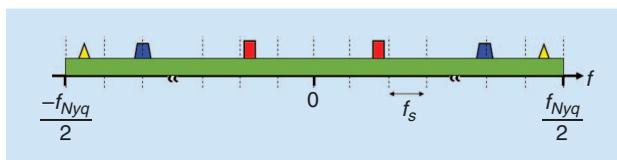
FIGURE 20. The original spectrum $X(f)$. The cyclostationary transmissions are shown in yellow, blue, and red, buried in stationary noise shown in green.

PUs whose modulation type and parameters were part of the training set.

For CR purposes, we seek a detector designed to comply with the following requirements: 1) carrier frequency and bandwidth estimation rather than simple detection of the presence or absence of a signal; 2) blind detection, i.e., without knowledge of the carrier frequencies, bandwidths, and symbol rates of the transmissions; 3) simultaneous detection of several transmissions; and 4) preferably a nonlearning approach, i.e., with no training phase. The parameter estimation algorithm, presented in [93], is a simple parameter extraction method from the cyclic spectrum of multiband signals that satisfies these requirements. It allows the estimation of several carriers and bandwidths simultaneously as well as the number of transmissions. The proposed parameter estimation algorithm can be decomposed into four main steps: preprocessing, thresholding, clustering, and parameter estimation, as illustrated in Figure 21.

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 to find its peaks. The locations and values of the selected peaks are then clustered using the k-means to find the corresponding cyclic feature after estimating the number of clusters by applying the elbow method [94]. It follows that, apart from the cluster present in dc, the number of real signals, i.e., $N_{\text{sig}} = K/2$, is equal to half the number of clusters. Next, the carrier frequency f_i, which corresponds to the highest peak [72], 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 [58] 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 shown in Figure 12.

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 because the cyclic spectrum is recovered directly from the MWC low-rate samples. Preliminary testing suggests that
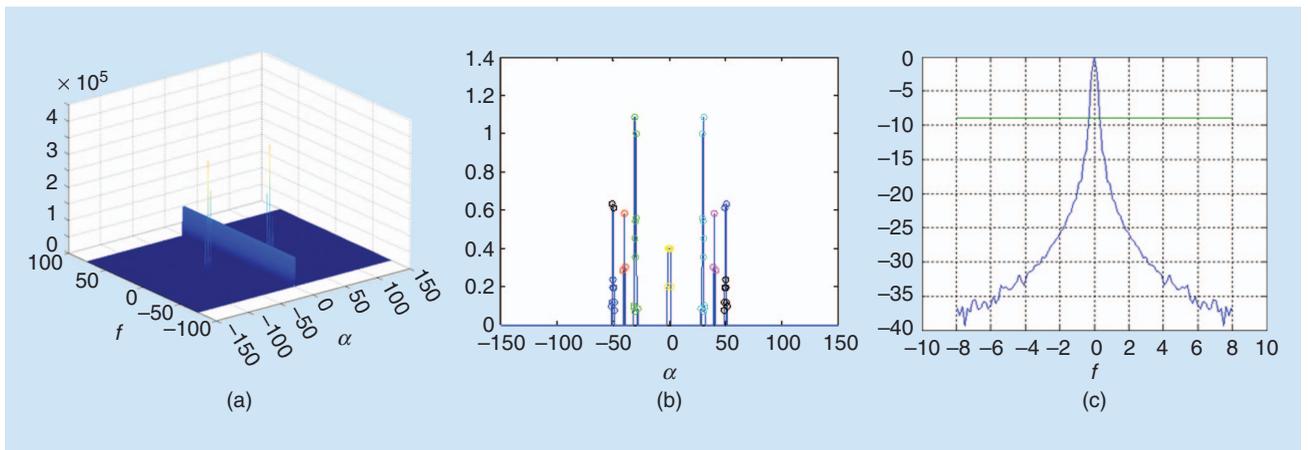
FIGURE 21. The carrier frequency and bandwidth estimation from the cyclic spectrum: (a) preprocessing, (b) thresholding and clustering, and (c) parameter estimation.
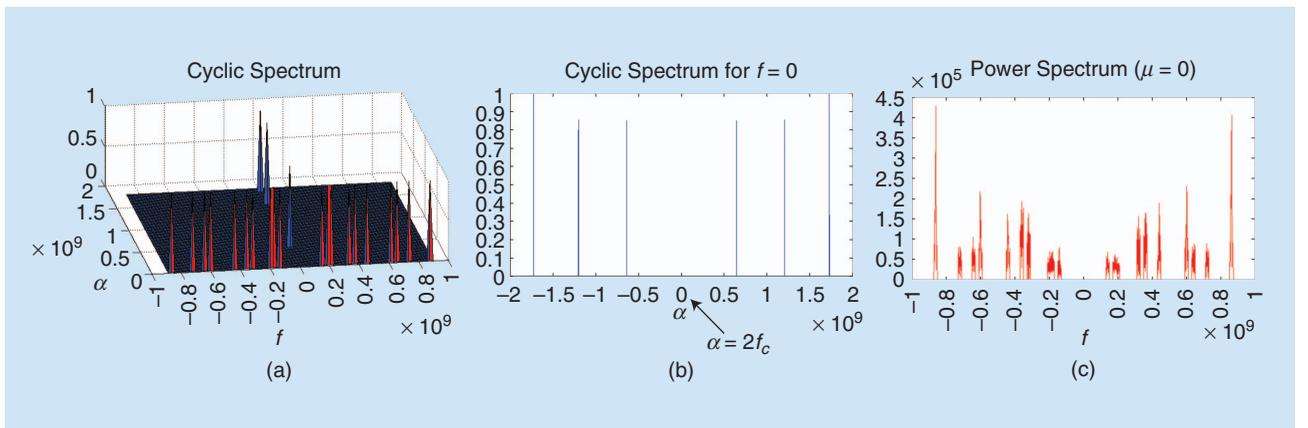
FIGURE 22. A screenshot from the MWC with cyclostationary detection. The input signal is composed of $N_{sig} = 3$ transmissions (or $K = 6$ bands) with carriers $f_1 = 320$, $f_2 = 600$, and $f_3 = 860$ MHz. (a) The recovered cyclic spectrum from low-rate samples; (b) the cyclic-spectrum profile at the angular frequency $f = 0$, where the cyclic peaks are clearly visible at twice the carrier frequencies; and (c) the power-spectrum recovery failing in the presence of noise.

sensing success is still achievable at SNRs that are 15-dB lower than those allowed energy detection performed on the recovered spectrum or power spectrum. Representative results shown in Figure 22 demonstrate the advantage of cyclostationary detection over energy detection in the presence of noise. The figure presents the reconstructed cyclic spectrum from samples obtained using the MWC prototype as well as cross sections at $f = 0$ and $\alpha = 0$, which correspond to the power spectrum. This increased robustness to noise comes at the expense of more involved digital processing on the low-rate samples. The additional complexity stems from the higher dimensionality involved, because we reconstruct the two-dimensional (2-D) cyclic spectrum rather than the one-dimensional (power) spectrum.

Collaborative-spectrum sensing

Collaborative model
Until now, we assumed a 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 [34]–[36], as described in "Channel Fading and Shadowing." 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 [34]. First, cooperation can be either centralized or distributed. In centralized settings, the data streams are sent to a fusion center that combines them 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. However, the latter is less power hungry and more robust for node and link failures, 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).

Analog signals transmitted over physical channels are affected by two main phenomena: Rayleigh fading, or small-scale fading, and log-normal shadowing, or large-scale fading [35], [36], [95]. The received signal is generally described in terms of the transmitted signal $s_i(t)$ convolved with the impulse response of the channel $h_{ij}(t)$, i.e.,

$$r_{ij}(t) = s_i(t) * h_{ij}(t), \tag{S13}$$

where $r_{ij}(t)$ is the received signal corresponding to the ith transmission at the jth cognitive radio and * denotes the convolution. Fading and shadowing affect the channel response $h_{ij}(t)$.

Rayleigh fading

For most practical channels, the free-space propagation model, which only accounts for path loss, is inadequate to describe the channel. A signal typically travels from transmitter to receiver over multiple reflective paths, which is traditionally modeled as Rayleigh fading. This implies that the amplitude and phase of the channel response $h_{ij}(t) = R(t) e^{j\phi(t)}$ are stochastically independent and identically distributed processes. The amplitude $R(t)$, for $t \in \mathbb{R}$, follows the Rayleigh distribution, given by

$$p_R(r) = \begin{cases} \dfrac{r}{\sigma^2} e^{-r^2/2\sigma^2} & r \geq 0 \\ 0 & \text{otherwise}, \end{cases} \tag{S14}$$

where $2\sigma^2$ is the mean power [95]. The phase $\phi(t)$, for $t \in \mathbb{R}$, is uniformly distributed over the interval $[0, 2\pi)$.

Log-normal shadowing

Large-scale fading represents the average signal power attenuation or path loss due to motion over large areas. The resulting channel frequency response is, therefore, a constant. This phenomenon is affected by prominent terrain contours between the transmitter and receiver. Empirical measurements suggest that this type of fading, or shadowing, follows a normal distribution in decibel units [96], or, alternatively, the linear channel gain may be modeled as a log-normal random variable [36]. Therefore, the path loss (PL) measured in decibels is expressed as

$$PL = PL_0 + 10\gamma \log \frac{d}{d_0} + X_\sigma. \tag{S15}$$

Here, the reference distance d_0 corresponds to a point located in the far field of the antenna (typically, 1 km for large cells). The PL to the reference point PL_0 is usually found through field measurements or calculated using free-space PL. The value of the PL exponent γ depends on the frequency, antenna heights, and propagation environment. Finally, X_σ denotes a Gaussian random variable (in decibels) with the variance σ^2 determined heuristically as well [95]. The shadowed received signal is thus given by

$$r_{ij}(t) = 10^{-PL_{ij}/20} \cdot s_i(t), \tag{S16}$$

where PL_{ij} denotes the PL between the ith transmitter and the jth receiver and the channel response $h_{ij}(t) = 10^{-PL_{ij}/20} \delta(t)$.

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

$$x^{(j)}(t) = \sum_{i=1}^{N_{sig}} r_{ij}(t) = \sum_{i=1}^{N_{sig}} s_i(t) * h_{ij}(t). \tag{30}$$

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 [35], [36], [95], as described in "Channel Fading and Shadowing." In the frequency domain, the Fourier transform of the jth received signal is given by

$$X^{(j)}(f) = \sum_{i=1}^{N_{sig}} S_i(f) H_{ij}(f). \tag{31}$$

Therefore, the support of $x^{(j)}(t)$ is included in the support of the original signal $x(t)$. Because 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 to assess the support of the transmitted signal $x(t)$ from sub-Nyquist samples of the received $x^{(j)}(t), 1 \leq j \leq N_{rec}$, 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 a 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 may rely on sharing observations based on the low-rate samples, rather than the samples themselves. In the next section, we review such techniques both in centralized and distributed contexts.

Centralized collaborative support recovery

One approach [97], [98] to centralized spectrum sensing considers a digital model based on 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, i.e.,

$$\mathbf{z}^{(j)} = \mathbf{A}\mathbf{x}^{(j)}, \tag{32}$$

where $\mathbf{A}$ is the sampling matrix. This model has been extensively studied by Leus et al. [97]–[99]. In all these works, the channel state information (CSI) is assumed to be known, and the joint power spectrum of x is reconstructed, where $\mathbf{x}$ denotes the $N \times 1$ vector of Nyquist samples. CSI is traditionally unknown by the CRs and should be estimated prior to detection to enable the use of this method. The autocorrelation of the Nyquist samples is first related to that of the sub-Nyquist samples. Then, in [97], the common sparsity of $\mathbf{r}_\mathbf{x}^{(j)}$ is exploited in the frequency domain across all CRs to jointly reconstruct them at the fusion center, using a modified simultaneous orthogonal matching pursuit (SOMP) [41] algorithm. In [98] and [99], besides exploiting joint sparsity, cross correlations between measurements from different CRs are used, i.e., $\mathbf{R}_{\mathbf{z}_j\mathbf{z}_k}$, where j and k are the indices of two CRs. These cross correlations are related to the common power spectrum $\mathbf{s}_\mathbf{x} = \mathbf{F}\mathbf{r}_\mathbf{x}$, where F is the discrete Fourier transform matrix. It is shown that if the total number of samples $N_{rec} M$ is greater than N, and these are suitably chosen to account for enough measurement diversity, then the power spectrum $\mathbf{s}_\mathbf{x}$ of a nonsparse signal can be recovered from compressed samples from a sufficient number of CRs. This shows that 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.

An alternative approach [57] relies on the analog model from (6) 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 (9), 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 to allow for more measurement diversity. However, the same known matrix can be used to reduce the communication overhead. The underlying matrices $\mathbf{U}^{(j)}$ are jointly sparse because fading (S13) and (S14) and shadowing (S16) do not affect the original signal's support. Capitalizing on the joint support of $\mathbf{U}^{(j)}$, the support of the transmitted signal $x(t)$ is recovered at the fusion center by solving

$$\underset{\mathbf{U}^{(j)}}{\arg\min} \quad \bigcup_{i=1}^{N_{rec}} \left\| \mathbf{U}^{(j)} \right\|_0$$
$$\text{s.t.} \quad \mathbf{V}^{(j)} = \mathbf{A}^{(j)}\mathbf{U}^{(j)}, \text{ for all } 1 \le j \le N_{rec}. \tag{33}$$

To recover the joint support of $\mathbf{U}^{(j)}$ from the observation matrices $\mathbf{V}^{(j)}$, both orthogonal matching pursuit (OMP) and iterative hard thresholding (IHT), described in "Compressed

Sensing Recovery," are extended to the collaborative setting [57]. Previously, in the "Sub-Nyquist Sampling for CR" section, we considered support recovery from an individual CR, which boils down to an MMV system of equations (9). CS algorithms have been extended to this case, such as SOMP from [100] and simultaneous IHT (SIHT), as presented in [101]. These now need to account for the joint sparsity across the CRs.

The distributed CS-SOMP (DCS-SOMP) algorithm [102], which extends the original SOMP to allow for different sampling matrices $\mathbf{A}^{(j)}$ for each receiver, is adapted to the CR collaborative setting in [57]. 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*) [57]. Next, the block sparse IHT (BSIHT) algorithm [57] extends SIHT by selecting the indices of the common support though averaging over all of the estimated $\mathbf{U}^{(j)}$ in each iteration. Once the support is selected, the update calculations are performed separately for each receiver. 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.

Distributed collaborative support recovery

In the distributed approach, there is no fusion center, and the CRs can only communicate with their neighbors. Both the digital and analog centralized approaches have been extended to the distributed settings. In [103] and [104], a digital model (32) is used where the spectrum is divided into N known slots. Both unknown and known CSI cases are considered. In the first case, each CR computes its local binary decision for every spectral band by recovering the sparse spectrum using CS techniques. Then, an average consensus approach is adopted with respect to the shared hard decision. If the CSI is known, similar to [98] and [99], then the spectrum is recovered in a distributed fashion. In [103], the proposed algorithm iterates through the following two steps: local spectrum reconstruction given an estimated support and consensus averaging on the support of the nodes. In [104], a distributed augmented Lagrangian is adopted.

The centralized approach, based on the analog model (6), presented in [57], is modified in [56] to comply with distributed settings. The ith CR contacts a random neighbor j, chosen with some probability P_{ij}, according to the Metropolis–Hastings scheme for random transition probabilities:

$$P_{ij} = \begin{cases} \min\left\{\dfrac{1}{d_i}, \dfrac{1}{d_j}\right\} & (i,j) \in E, \\ \displaystyle\sum_{(i,k) \in E} \max\left\{0, \dfrac{1}{d_i} - \dfrac{1}{d_k}\right\} & i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$
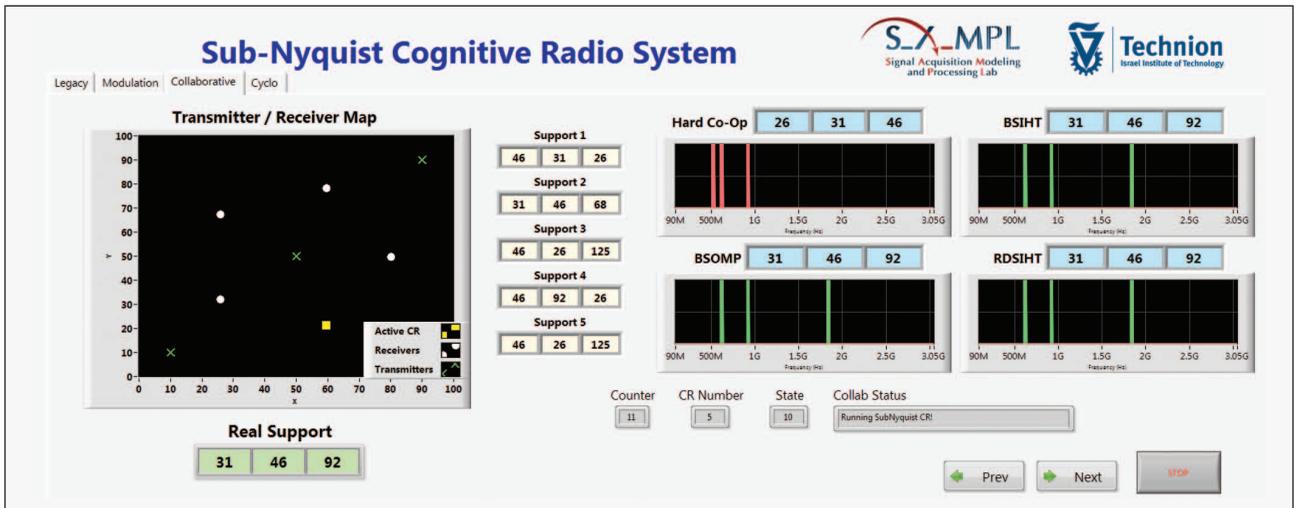
FIGURE 23. A screenshot from the MWC CR collaborative hardware prototype. On the upper left side, we see the spatial map of the receivers in white and the transmitters in green. On the bottom left, the occupied band indices of the real spectral support are shown, while, to the right of the transmitter/receiver map, the estimated indices by each CR individually are presented. On the right, we see the spectrum-sensing results of four different algorithms: hard co-op (hard-decision collaboration, which 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 CR results.

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 later in this section 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 a 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 probability (34). Two distributed algorithms are presented in [56]. The first is the distributed one-step greedy algorithm (DOSGA), which extends the OSGA from [102] to distributed settings. The second method, referred to as *randomized distributed IHT* (*RDSIHT*), adapts the centralized BSIHT [57] 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

$$\hat{\mathbf{w}} = \frac{1}{N_{\text{rec}}} \sum_{j=1}^{N_{\text{rec}}} \mathbf{w}^{(j)} \tag{35}$$

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 [56] then uses a randomized gossip algorithm [105] for this purpose, with the Metropolis–Hastings transition probabilities.

Next, we turn to the RDSIHT algorithm, which adapts the centralized BSIHT algorithm [57] to the distributed scenario. The distributed approach from [56] was inspired by the randomized incremental subgradient method proposed in [106] and recent work on a stochastic version of IHT [107]. A vector $\mathbf{w}$ of size N, which sums the ℓ_2 norms of the rows of the estimates of $\mathbf{U}^{(j)}$ before thresholding, is shared in the network through 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} (34). The joint sparsity across the CRs is exploited by sharing one common vector $\mathbf{w}$ by the network. It is shown numerically in [56] that both DOSGA and RDSIHT converge to their centralized counterparts.

Hardware simulations: Collaborative versus individual spectrum sensing

We now confirm that the collaborative algorithms for spectrum sensing perform better than their individual counterparts by demonstrating a collaborative setting implemented using the MWC CR prototype, as can be seen in Figure 23. During the conducted experiments, $N_{\text{rec}} = 5$ 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 positions and broadcasts of the transmitter mimic the effects of physical channel phenomena, i.e., fading and shadowing. The frequency support recovered by each of the CRs is false because 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, because 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 Figure 23, where the hard-decision support algorithm (hard co-op) fails to recover the entire active frequency support (depicted by red bins).

Joint-carrier frequency and direction estimation

Many signal processing applications, including CR, may require or at least benefit from joint spectrum sensing and DOA estimation. For CRs to map vacant bands more efficiently, spatial information about the PUs' locations are of great interest. Consider the network of CRs presented in Figure 24 and focus on CR1. Now, imagine a scenario where PU2, with AOA 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 could transmit in the same frequency band in the opposite direction of PU2 toward CR2. The DOA estimation can thus enhance CR performance by allowing the exploitation of vacant bands in space in addition to the frequency domain.

Model and system description

To mathematically formulate our problem, 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 AOAs θ_i. The main difference between this scenario and the one that has been discussed in the previous sections is the additional unknown AOAs θ_i. Figure 24 illustrates this signal model. To recover the unknown AOAs, an array of

FIGURE 24. An illustration of $N_{sig} = 3$ source signals in the yz plane. Each transmission is associated with a carrier frequency f_i and AOA θ_i.

> 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.

sensors is required. A similar problem thoroughly treated in the literature is the 2-D-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.

Multicoset approach

A few works have recently considered joint DOA and spectrum sensing of multiband signals from sub-Nyquist samples. In [108] and [109], low-rate samples are obtained using the multicoset sampling scheme. In [108], which considers the digital model (32), both time and spatial compression are applied by selecting samples from the Nyquist grid and receivers from a ULA, such that

$$\mathbf{Z}[n] = \mathbf{C}_s \mathbf{X}[n] \mathbf{C}_t. \tag{36}$$

Here, $\mathbf{X}[n]$ is the matrix of Nyquist samples from all receivers in the ULA, and 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 2-D power-spectrum matrix of the underlying signal is then reconstructed from the samples, where every row describes the power spectrum in the frequency domain for a given AOA and every column corresponds to the power-spectrum information in the angular domain for a given frequency.

In [109], an L-shaped array with two interleaved, or multicoset, channels, with a fixed delay between the two, samples the signal below the Nyquist rate. Then, exploiting correlations between samples from the direct path and its delayed version, the frequencies and their corresponding AOAs are estimated using MUSIC [38], [39]. However, the pairing issue between the two, i.e., 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 [53], that utilizes the sampling principles of the MWC. This technique addresses the pairing issue and avoids the hardware issues involved in multicoset sampling.

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 AOA θ_i, as shown in Figure 25. 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, an 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
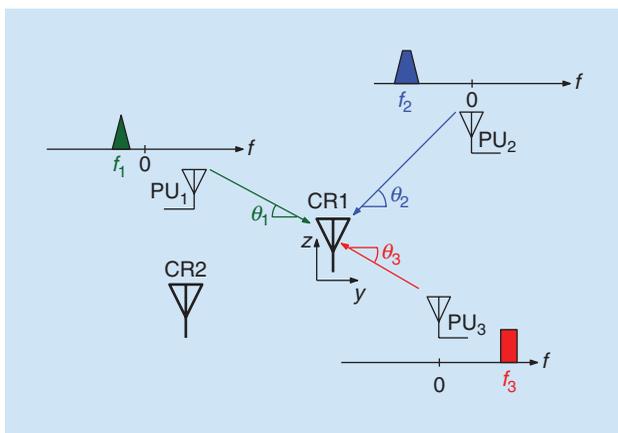
of equations similar to (14) can be derived. For the ULA along the y axis, we obtain

$$\mathbf{y}(f) = \mathbf{A}_y \mathbf{x}(f), \quad f \in \mathcal{F}_s, \tag{37}$$

where

$$\mathbf{A}_y = \begin{bmatrix} e^{j2\pi f_1 \tau_1^y(\theta_1)} & \cdots & e^{j2\pi f_K \tau_1^y(\theta_K)} \\ \vdots & & \vdots \\ e^{j2\pi f_1 \tau_M^y(\theta_1)} & \cdots & e^{j2\pi f_K \tau_M^y(\theta_K)} \end{bmatrix} \tag{38}$$

and $\tau_m^y(\theta) = (dm/c)\cos(\theta)$ denote the delays at the mth sensors with respect to the first sensor. Along the z axis, the samples $\mathbf{z}(f)$, sampling matrix $\mathbf{A}_z$, and delays τ_m^z are similarly defined.

Two joint recovery approaches for the carrier frequencies and DOAs of the transmissions are proposed in [53]. Note that, once the carriers and AOAs 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 time-domain samples from both ULAs are concatenated into one vector $\mathbf{v}[n] = [\mathbf{y}^T[n]\,\mathbf{z}^T[n]]^T$, whose correlation matrix,

$$\mathbf{R} = \mathbb{E}[\mathbf{v}[n]\mathbf{v}^H[n]] = \mathbf{A}\mathbf{R}_x \mathbf{A}^H, \tag{39}$$

is computed. Here, $\mathbf{A} = [\mathbf{A}_y^T, \mathbf{A}_z^T]^T$ and the autocorrelation matrix $\mathbf{R}_x = \mathbb{E}[\mathbf{x}[k]\mathbf{x}^H[k]]$ are sparse and diagonal from the stationarity of $x(t)$. From the grid assumption, (39) can 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 can be recovered using traditional CS techniques, similar to (15).

The second recovery approach, inspired by [110] and [111], extends the ESPRIT algorithm to the joint estimation of carriers and DOAs while overcoming the pairing issue. The 2-D-ESPRIT algorithm presented in [53] is directly applied to the sub-Nyquist samples by considering cross-correlation matrices between the subarrays of both axes. Dropping the time variable n for clarity, the samples from the subarrays can be written as

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{A}_{y_1}\mathbf{x}, \quad \mathbf{y}_2 = \mathbf{A}_{y_2}\mathbf{x} \\ \mathbf{z}_1 &= \mathbf{A}_{z_1}\mathbf{x}, \quad \mathbf{z}_2 = \mathbf{A}_{z_2}\mathbf{x}, \end{aligned} \tag{40}$$

where $\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1$, and $\mathbf{z}_2$ are samples from the subarrays shown in Figure 25. 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 subarray matrices along the same axis are related as
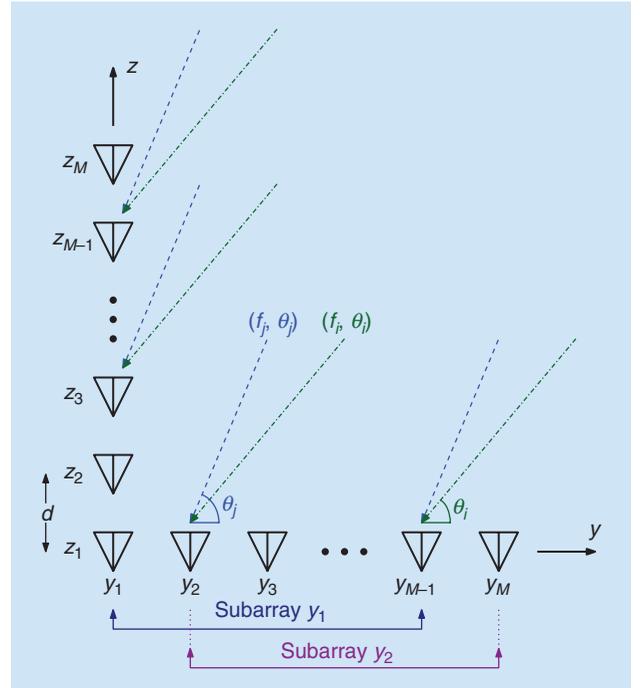
FIGURE 25. The CaSCADE system: an L-shaped array with M sensors along the y axis and M sensors along the z axis, including a common sensor at the origin. The subarrays $\mathbf{y}_1$ and $\mathbf{y}_2$ (and similarly $\mathbf{z}_1$ and $\mathbf{z}_2$) are defined in the derivation of the 2-D-ESPRIT algorithm.

$$\begin{aligned} \mathbf{A}_{y_2} &= \mathbf{A}_{y_1}\mathbf{\Phi} \\ \mathbf{A}_{z_2} &= \mathbf{A}_{z_1}\mathbf{\Psi}, \end{aligned} \tag{41}$$

where

$$\begin{aligned} \mathbf{\Phi} &\triangleq \mathrm{diag}\left[e^{j2\pi f_1 \tau_1^y(\theta_1)} \quad \cdots \quad e^{j2\pi f_K \tau_1^y(\theta_K)}\right] \\ \mathbf{\Psi} &\triangleq \mathrm{diag}\left[e^{j2\pi f_1 \tau_1^z(\theta_1)} \quad \cdots \quad e^{j2\pi f_K \tau_1^z(\theta_K)}\right]. \end{aligned} \tag{42}$$

We see from (42) that the carrier frequencies f_i and AOAs θ_i are embedded in the diagonal matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$. Applying the ESPRIT framework on cross-correlation matrices between the subarrays of both axes allows it to jointly recover $\mathbf{\Phi}$ and $\mathbf{\Psi}$ [53]. This leads to a proper pairing of the corresponding elements $f_i \tau_1^y(\theta_i)$ and $f_i \tau_1^z(\theta_i)$. The AOAs θ_i and carrier frequencies f_i are then given by

$$\theta_i = \tan^{-1}\left(\frac{\angle\mathbf{\Psi}_{ii}}{\angle\mathbf{\Phi}_{ii}}\right), \quad f_i = \frac{\angle\mathbf{\Phi}_{ii}}{2\pi\dfrac{d}{c}\cos(\theta_i)}. \tag{43}$$

It is proven in [53] 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 be extended to jointly recover the transmissions' carrier frequencies, azimuth, and elevation angles in a three-dimensional framework.

Conclusions and future challenges

In this article, 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 the sampling and processing of wide-band signals at low rates by exploiting their a priori known structure. A possible extension of these works is to include an adaptive detected support update, which is triggered by a change in a PU's activity, either by 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. Additional preliminary assumptions on the structure or statistical behavior of the potentially active signals, such as statistics on channel occupancy, can also be exploited.

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. Finally, we addressed the joint spectrum sensing and DOA estimation problem, allowing for better exploitation of frequency vacant bands by utilizing spatial sparsity as well. All of these methods should be combined to map the occupied spectrum in frequency, time, and space, thus maximizing the CR network's throughput. This requires an adequate spectrum access protocol as well, one that translates the data acquired by spectrum sensing into transmission opportunities for the CRs. Spectrum access challenges and algorithms were outside this article's scope.

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. We believe that prototype development is the key to enabling sub-Nyquist sampling as a solution to the task of spectrum sensing in CR platforms. A natural next step in that direction is the implementation of a complete CR network, collaboratively performing joint spectrum sensing and DOA estimation followed by spectrum access. This prototype should then be tested on real data to assess its true capabilities. We believe that a sub-Nyquist digital-to-analog interface can alleviate many of the bottlenecks currently hindering the development of CR systems, consequently allowing the fast deployment of low-rate, simple, and efficient CR devices that will be using currently available hardware.

For CRs to become a viable solution to the spectrum shortage, other main challenges need to be addressed, as discussed previously in the "CR Challenges" section. The issue of coexistence with existing communication links from PUs is crucial and very particular to the CR scenario. Here, coexistence is not a symmetric requirement, as CRs are prohibited from interfering with PUs. Another challenge is the establishment of a communication channel for CRs to be able to exchange the locations of their current transmission bands. Finally, the issue of security against attacks to the CR networks still has numerous unresolved questions.

Acknowledgment

Authors

Deborah Cohen (debby@technion.ac.il) received her B.S. degree in electrical engineering (summa cum laude) in 2010 and her Ph.D. degree in electrical engineering in 2016, both from the Technion–Israel Institute of Technology, Haifa. She has been granted several awards, including the Meyer Foundation Excellence Award, the David and Tova Freud and Ruth Freud-Brendel Memorial Scholarship, the Sandor Szego Award, the Vivian Konigsberg Award, and the Muriel and David Jacknow Award for Excellence in Teaching. Since 2014, she has been an Azrieli fellow. She is currently a research scientist with Google Israel. Her research interests include theoretical aspects of signal processing, compressed sensing, reinforcement learning, and machine learning for dialogues.

Shahar Tsiper (tsiper@technion.ac.il) received his B.S. degree in electrical engineering (cum laude) and his B.S. degree in physics (cum laude), both in 2015, from the Technion–Israel Institute of Technology, Haifa. He is a graduate of the EMET excellent students program and has been granted several awards, including the Meyer Foundation Excellence Award; the METRO 450 Consortium Excellence in Research Award; the Kasher Award; the Best Show & Tell Demo Award at the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing; and the Ollendorf Fellowship Award for Excellent M.S. Students. He is currently pursuing an M.S. degree in electrical engineering at the Technion.

Yonina C. Eldar (yonina@ee.technion.ac.il) is a professor in the Department of Electrical Engineering at the Technion–Israel Institute of Technology, Haifa, where she holds the Edwards Chair in engineering. She is also an adjunct professor at Duke University, Durham, North Carolina, and a research affiliate with the Research Laboratory of Electronics at the Massachusetts Institute of Technology, Cambridge, and she was a visiting professor at Stanford University, California. She is a member of the Israel Academy of Sciences and Humanities and a Fellow of the IEEE and the European Association for Signal Processing. She has received many awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award, the IEEE/Aerospace and Electronic Systems Society Fred Nathanson Memorial Radar Award, the IEEE Kiyo Tomiyasu Award, the Michael Bruno Memorial Award from the Rothschild Foundation, the Weizmann Prize for Exact Sciences, and the Wolf Foundation Krill Prize for Excellence in Scientific Research. She is the editor-in-chief of *Foundations and Trends in Signal Processing* and serves the IEEE on several technical and award committees.

References

[1] Federal Communications Commission. (2002, Nov.). Spectrum policy task force report. Tech. Rep. 02-135. [Online]. Available: http://www.gov/edocs_public/attachmatch/DOC228542A1.pdf

[2] M. McHenry. (2005, Aug.). NSF spectrum occupancy measurements project summary. Shared Spectrum Co. [Online]. Available: http://www.sharedspectrum.com

[3] R. I. C. Chiang, G. B. Rowe, and K. W. Sowerby, "A quantitative analysis of spectral occupancy measurements for cognitive radio," in *Proc. IEEE Vehicular Technology Conf.*, 2007, pp. 3016–3020.

[4] Y. Liu, "Beyond white space: Robust spectrum sensing and channel statistics based spectrum accessing strategies for cognitive radio network," Ph.D. dissertation, Univ. Texas, Austin, 2013.

[5] T. Erpek, M. Lofquist, and K. Patton. (2007, Sept. 18–20). Spectrum occupancy measurements: Loring Commerce Centre, Limestone, Maine. Shared Spectrum Company. Vienna, VA. [Online]. Available: http://www.sharedspectrum.com/wp-content/uploads/Loring_Spectrum_Occupancy_Measurements_v2_3.pdf

[6] J. Mitola, "Software radios: Survey, critical evaluation and future directions," *IEEE Aerospace Electron. Syst. Mag.*, vol. 8, pp. 25–36, Apr. 1993.

[7] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 201–220, Feb. 2005.

[8] J. Mitola and C. Q. Maguire, Jr., "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, pp. 13–18, Aug. 1999.

[9] C. E. Shannon, "Computers and automata," *Proc. IRE*, vol. 41, pp. 1234–1241, Oct. 1953.

[10] D. Cabric, S. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proc. Asilomar Conf. Signals, Systems and Computers*, 2004, pp. 772–776.

[11] E. G. Larsson and M. Skoglund, "Cognitive radio in a frequency-planned environment: Some basic limits," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4800–4806, Dec. 2008.

[12] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Proc. Annu. Allerton Conf. Communication, Control, and Computing*, 2004, pp. 1662–1671.

[13] M. Mishali and Y. C. Eldar, "Wideband spectrum sensing at sub-Nyquist rates," *IEEE Signal Process. Mag.*, vol. 28, pp. 102–135, July 2011.

[14] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: Requirements, challenges and design trade-offs," *IEEE Commun. Mag.*, vol. 46, pp. 32–39, Apr. 2008.

[15] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radios: State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, pp. 101–116, May 2012.

[16] B. Wang and K. Liu, "Advances in cognitive radio networks: A survey," *IEEE J. Sel. Topics Signal Process.*, vol. 5, pp. 5–23, 2011.

[17] J. Ma, G. Y. Li, and B. H. F. Juang, "Signal processing in cognitive radio," *Proc. IEEE*, vol. 97, pp. 805–823, 2009.

[18] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: Challenges and solutions," *EURASIP J. Advances Signal Process.*, vol. 2010, pp. 1–15, 2010.

[19] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, pp. 116–130, 2009.

[20] E. Hossain, D. Niyato, and D. I. Kim, "Evolution and future trends of research in cognitive radio: A contemporary survey," *Wireless Commun. Mobile Computing*, vol. 15, pp. 1530–1564, 2015.

[21] E. Biglieri, A. J. Goldsmith, L. J. Greenstein, N. B. Mandayam, and H. V. Poor, *Principles of Cognitive Radio*. Cambridge. U.K.: Cambridge Univ. Press, 2012.

[22] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *Proc. IEEE Int. Symp. Circuits and Systems*, 2007, pp. 1959–1962.

[23] 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.

[24] 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.

[25] M. Mishali and Y. C. Eldar, "Blind multi-band signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.

[26] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 375–391, Apr. 2010.

[27] 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.

[28] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, pp. 523–531, Apr. 1967.

[29] 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.

[30] D. O. North, "An analysis of the factors which determine signal/noise discrimination in pulsed carrier systems," *Proc. IEEE*, vol. 51, pp. 1016–1027, July 1963.

[31] G. L. Turin, "An introduction to matched filters," *IRE Trans. Inf. Theory*, vol. 6, pp. 311–329, June 1960.

[32] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, pp. 639–697, 2006.

[33] A. Napolitano, "Cyclostationarity: New trends and applications," *Signal Process.*, vol. 120, pp. 385–408, Mar. 2016.

[34] 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.

[35] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative sensing among cognitive radios," in *Proc. IEEE Int. Conf. Communications*, 2006, pp. 1658–1663.

[36] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. IEEE Int. Symp. New Frontiers in Dynamic Spectrum Access Networks*, 2005, pp. 131–136.

[37] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, pp. 1128–1140, 2009.

[38] V. F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophys. J. R. Astron. Soc.*, vol. 33, pp. 347–366, 1973.

[39] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, pp. 276–280, Mar. 1986.

[40] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Signal Process.*, vol. 37, pp. 984–995, July 1989.

[41] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[42] S. K. Sharma, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Application of compressive sensing in cognitive radio communications: A survey," *Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1838–1860, 2016.

[43] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[44] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, pp. 3397–3415, Dec. 1993.

[45] T. Blumensath and M. Davies, "Gradient pursuits," *IEEE Trans. Signal Process.*, vol. 56, pp. 2370–2382, June 2008.

[46] 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.

[47] H. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math.*, vol. 117, pp. 37–52, July 1967.

[48] 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.

[49] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-Nyquist rates," *IET Circuits, Devices Syst.*, vol. 5, pp. 8–20, Jan. 2011.

[50] E. Israeli, S. Tsiper, D. Cohen, A. Reysenson, R. Hilgendorf, E. Shoshan, and Y. C. Eldar, "Hardware calibration of the modulated wideband converter," in *Proc. IEEE Global Communications Conf.*, 2014, pp. 948–953.

[51] M. Mishali and Y. C. Eldar, "Expected RIP: Conditioning of the modulated wideband converter," in *Proc. IEEE Information Theory Workshop*, 2009, pp. 343–347.

[52] H. W. L. Gan and H. Wang, "Deterministic binary sequences for modulated wideband converter," in *Proc. Int. Conf. Sampling Theory and Applications*, 2013.

[53] S. Stein, O. Yair, D. Cohen, and Y. C. Eldar. (2016, Apr. 10). CaSCADE: Compressed carrier and DOA estimation. arXiv. [Online]. Available: https://arxiv.org/abs/1604.02723

[54] R. Gold, "Optimal binary sequences for spread spectrum multiplexing (corresp.)," *IEEE Trans. Inf. Theory*, vol. 13, no. 4, pp. 619–621, 1967.

[55] M. Mishali and Y. C. Eldar, "Expected RIP: Conditioning of the modulated wideband converter," in *Proc. IEEE Information Theory Workshop*, 2009, pp. 343–347.

[56] D. Cohen, A. Akiva, B. Avraham, and Y. C. Eldar, "Distributed cooperative spectrum sensing from sub-Nyquist samples for cognitive radios," in *Proc. IEEE Workshop Signal Processing Advances Wireless Communications*, 2015, pp. 336–340.

[57] D. Cohen, A. Akiva, B. Avraham, and Y. C. Eldar, "Centralized cooperative spectrum sensing from sub-Nyquist samples for cognitive radios," in Proc. *IEEE Int. Conf. Communications*, 2015, pp. 7487–7491.

[58] D. Cohen and Y. C. Eldar. (2016, Apr. 10). Sub-Nyquist cyclostationary detection for cognitive radio. arXiv. [Online]. Available: https://arxiv.org/abs/1604.02659

[59] D. Adams, Y. C. Eldar, and B. Murmann, "A mixer frontend for a four-channel modulated wideband converter with 62 db blocker rejection," in *Proc. 2016 IEEE Radio Frequency Integrated Circuits Symp. (RFIC)*, 2016, pp. 286–289.

[60] M. A. Lexa, M. E. Davies, J. S. Thompson, and J. Nikolic, "Compressive power spectral density estimation," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2011, pp. 22–27.

[61] M. A. Lexa, M. E. Davies, and J. S. Thompson. (2011, Oct. 12). Compressive and noncompressive power spectral density estimation from periodic nonuniform samples. arXiv. [Online]. Available: https://arxiv.org/abs/1110.2722

[62] D. D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. Signal Process.*, vol. 60, pp. 4775–4789, Sept. 2012.

[63] D. D. Ariananda, G. Leus, and Z. Tian, "Multi-coset sampling for power spectrum blind sensing," in *Proc. Int. Conf. Digital Signal Processing*, 2011, pp. 1–8.

[64] D. Romero and G. Leus, "Compressive covariance sampling," in *Proc. Information Theory and Applications Workshop*, 2013, pp. 1–8.

[65] 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, June 2013.

[66] 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.

[67] Z. Tian, "Cyclic feature based wideband spectrum sensing using compressive sampling," in *Proc. IEEE Int. Conf. Communications*, 2011, pp. 1–5.

[68] Z. Tian, Y. Tafesse, and B. M. Sadler, "Cyclic feature detection with sub-Nyquist sampling for wideband spectrum sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 1, pp. 58–69, Feb. 2012.

[69] S. A. Razavi, M. Valkama, and D. Cabric, "High-resolution cyclic spectrum reconstruction from sub-Nyquist samples," in *Proc. IEEE Workshop Signal Processing Advances Wireless Communications*, 2013, pp. 250–254.

[70] G. Leus and Z. Tian, "Recovering second-order statistics from compressive measurements," in *Proc. IEEE Int. Workshop Computational Advances Multi-Sensor Adaptive Processing*, 2011, pp. 337–340.

[71] D. D. Ariananda and G. Leus, "Non-uniform sampling for compressive cyclic spectrum reconstruction," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2014, pp. 41–45.

[72] W. Gardner, *Statistical Spectral Analysis: A Non Probabilistic Theory*. Englewood Cliffs, NJ: Prentice Hall, 1988.

[73] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw Hill, 1991.

[74] 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.

[75] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, pp. 573–586, 2011.

[76] D. Qu and A. Tarczynski, "A novel spectral estimation method by using periodic nonuniform sampling," in *Proc. Asilomar Conf. Signals, Systems and Computers*, 2007, pp. 1134–1138.

[77] C. G. Khatri and C. R. Rao, "Solutions to some functional equations and their applications to characterization of probability distributions," *Sankhya: Indian J. Statistics, Series A*, pp. 167–180, 1968.

[78] M. Arts, A. Bollig, and R. Mathar, "Analytical test statistic distributions of the MMME eigenvalue-based detector for spectrum sensing," in *Proc. IEEE 2015 Int. Symp. Wireless Communication Systems (ISWCS)*, pp. 496–500.

[79] 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.

[80] J. Leech, "On the representation of 1, 2, n by differences," *J. London Mathematical Soc.*, vol. 1, no. 2, pp. 160–169, 1956.

[81] W. A. Gardner, "The spectral correlation theory of cyclostationary time-series," *Signal Process.*, vol. 11, pp. 13–36, July 1986.

[82] P. S. Aparna and M. Jayasheela, "Cyclostationary feature detection in cognitive radio using different modulation schemes," *Int. J. Comput. Applicat.*, vol. 47, pp. 12–16, June 2012.

[83] S. Y. Kumar and M. Z. A. Khan, "Robust, block based cyclostationary detection," in *Proc. Int. Conf. Telecommunications*, 2013, pp. 1–5.

[84] A. V. Dandawate and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Trans. Signal Process.*, pp. 2355–2369, 1994.

[85] M. Ghozzi, F. Marx, M. Dohler, and J. Palicot, "Cyclostatilonarilty-based test for detection of vacant frequency bands," in *Proc. Int. ICST Conf. Cognitive Radio Oriented Wireless Networks and Communications*, 2006, pp. 1–5.

[86] Y. T. Chan, J. W. Plews, and K. C. Ho, "Symbol rate estimation by the wavelet transform," in *Proc. IEEE Int. Symp. Circuits and Systems*, 1997, pp. 177–180.

[87] E. Rebeiz and D. Cabric, "Blind modulation classification based on spectral correlation and its robustness to timing mismatch," in *Proc. Military Communications Conf.*, 2011, pp. 277–282.

[88] R. J. Mammone, R. J. Rothaker, C. I. Podilchuk, S. Davidovici, and D. L. Schilling, "Estimation of carrier frequency, modulation type and bit rate of an unknown modulated signal," in *Proc. IEEE Int. Conf. Communications*, 1987, pp. 1006–1012.

[89] A. Badawy and T. Khattab, "Robust, block based cyclostationary detection," in *Proc. IEEE Wireless Communication and Networking Conf.*, 2014, pp. 253–258.

[90] A. Fehske, J. Gaeddert, and J. H. Reed, "A new approach to signal classification using spectral correlation and neural networks," in *Proc. IEEE Int. Symp. New Frontiers in Dynamic Spectrum Access Networks*, 2005, pp. 144–150.

[91] P. S. Aparna and M. Jayasheela, "Cyclostationary feature detection in cognitive radio using different modulation schemes," *Int. J. Comput. Applicat.*, vol. 47, pp. 12–16, June 2012.

[92] X. He, Z. Zeng, and C. Guo, "Signal classification based on cyclostationary spectral analysis and HMM/SVM in cognitive radio," in *Proc. Int. Conf. Measuring Technology and Mechatronics Automation*, 2009, pp. 309–312.

[93] D. Cohen, L. Pollak, and Y. C. Eldar, "Carrier frequency and bandwidth estimation of cyclostationary multiband signals," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, 2016.

[94] R. L. Thorndike, "Who belong in the family?" *Psychometrika*, vol. 18, pp. 267–276, Dec. 1953.

[95] B. Sklar, "Rayleigh fading channels in mobile digital communication systems part I: Characterization," *IEEE Commun. Mag.*, vol. 35, pp. 90–100, July 1997.

[96] V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, and R. Bianchi, "An empirically based path loss model for wireless channels in suburban environments," *IEEE J. Sel. Areas Commun.*, vol. 17, pp. 1205–1211, Sept. 2006.

[97] Y. Wang, A. Pandharipande, Y. L. Polo, and G. Leus, "Distributed compressive wide-band spectrum sensing," in *Proc. IEEE Information Theory and Applications Workshop*, 2009, pp. 178–183.

[98] D. D. Ariananda and G. Leus, "A study on cooperative compressive wideband power spectrum sensing," in *Proc. Joint WIC/IEEE Symp. Information Theory and Signal Processing*, 2012, pp. 102–109.

[99] D. D. Ariananda and G. Leus, "Cooperative compressive wideband power spectrum sensing," in *Proc. IEEE Asilomar Conf. Signals, Systems and Computers*, 2012, pp. 303–307.

[100] J. Tropp, A. C. Gilbert, , and M. J. Strauss, "Simultaneous sparse approximation via greedy pursuit," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2005, pp. 721–724.

[101] A. Makhzani and S. Valaee, "Reconstruction of jointly sparse signals using iterative hard thresholding," in *Proc. IEEE Int. Conf. Communications*, 2012, pp. 3564–3568.

[102] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Proc. IEEE Asilomar Conf. Signals, Systems and Computers*, 2005, pp. 1537–1541.

[103] Z. Tian, "Compressed wideband sensing in cooperative cognitive radio networks," in *Proc. IEEE Global Communications Conf.*, 2008, pp. 1–5.

[104] F. Zeng, C. Li, and Z. Tian, "Distributed compressive spectrum sensing in cooperative multihop cognitive networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, pp. 37–48, Feb. 2011.

[105] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, pp. 2508–2530, 2006.

[106] B. J. M. Rabi and M. Johansson, "A randomized incremental subgradient method for distributed optimization in networked systems," *SIAM J. Optim.*, vol. 20, no. 3, pp. 1157–1170, 2009.

[107] N. Nguyen, D. Needell, and T. Woolf. (2014). Linear convergence of stochastic iterative greedy algorithms with sparse constraints. CoRR. [Online]. abs/1407.0088. Available: http://arxiv.org/abs/1407.0088

[108] D. D. Ariananda and G. Leus, "Compressive joint angular-frequency power spectrum estimation," in *Proc. European Signal Processing Conf.*, 2013, pp. 1–5.

[109] 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," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2014, pp. 6781–6785.

[110] J.-F. Gu and P. Wei, "Joint SVD of two cross-correlation matrices to achieve automatic pairing in 2-D angle estimation problems," *IEEE Antennas Wireless Propag. Lett.*, vol. 6, pp. 553–556, 2007.

[111] 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.

SP