ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/279799937

Sub-Nyquist Sampling and Compressed Sensing in Cognitive Radio Networks

Book · September 2014

DOI: 10.1007/978-3-642-38398-4_6

CITATIONS
0

READS
59

6 authors, including:

Hongjian Sun
Durham University
61 PUBLICATIONS **1,045** CITATIONS

SEE PROFILE

Arumugam Nallanathan
Queen Mary, University of London
381 PUBLICATIONS **5,604** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Molecular Communications View project

Project Simultaneous Wireless InFormation and Energy Transfer View project

All content following this page was uploaded by Arumugam Nallanathan on 22 October 2015.

The user has requested enhancement of the downloaded file.

Sub-Nyquist Sampling and Compressed Sensing in Cognitive Radio Networks

Hongjian Sun, Arumugam Nallanathan, and Jing Jiang

Abstract Cognitive radio has become one of the most promising solutions for addressing the spectral under-utilization problem in wireless communication systems. As a key technology, spectrum sensing enables cognitive radios to find spectrum holes and improve spectral utilization efficiency. To exploit more spectral opportunities, wideband spectrum sensing approaches should be adopted to search multiple frequency bands at a time. However, wideband spectrum sensing systems are difficult to design, due to either high implementation complexity or high financial/energy costs. Sub-Nyquist sampling and compressed sensing play crucial roles in the efficient implementation of wideband spectrum sensing in cognitive radios.

In this chapter, Section 1 presents the fundamentals of cognitive radios. A literature review of spectrum sensing algorithms is given in Section 2. Wideband spectrum sensing algorithms are then discussed in Section 3. Special attention is paid to the use of Sub-Nyquist sampling and compressed sensing techniques for realizing wideband spectrum sensing. Finally, Section 4 shows an adaptive compressed sensing approach for wideband spectrum sensing in cognitive radio networks.

1 Cognitive Radio Networks

Nowadays, radio frequency (RF) spectrum is a scarce and valuable natural resource due to its unique character in wireless communications. Under the current policy, the primary user of a frequency band has exclusive rights of using the licensed band. With the explosive growth of wireless communication applications, the demands

Dr. H. Sun and Dr. A. Nallanathan
Department of Electronic Engineering, King's College London, London, WC2R 2LS, UK. e-mail: mrhjsun@hotmail.com, nallanathan@ieee.org

Dr. J. Jiang
Center for Communication Systems Research, University of Surrey, Guildford, GU2 7XH, UK. e-mail: jing.jiang@surrey.ac.uk

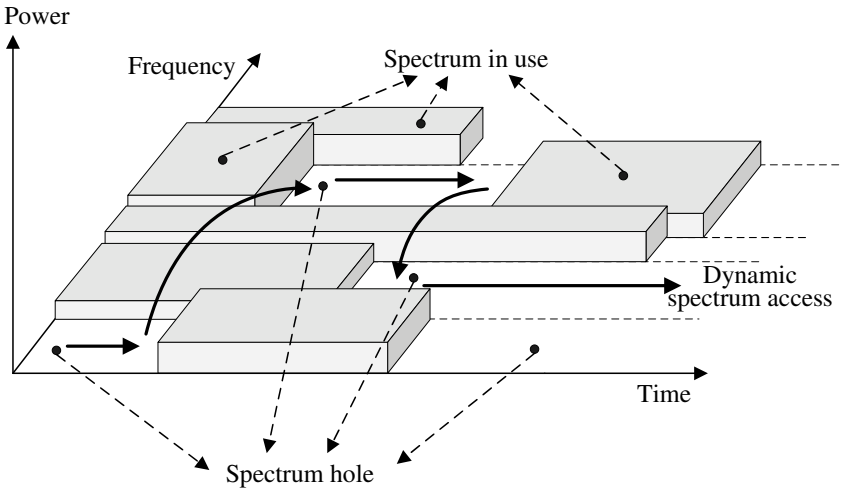

Fig. 1 Dynamic spectrum access and spectrum holes [3].

for the RF spectrum are constantly increasing. It becomes evident that such spectral demands cannot be met under the exclusive spectral allocation policy. On the other hand, it has been reported that the temporal and geographic spectral utilization efficiency is very low. For example, the maximal occupancy of the frequency spectrum between 30 MHz and 3 GHz (in New York City) has been reported to be only 13.1%, with the average occupancy of 5.2% [1]. As depicted by Figure 1, the spectral under-utilization problem can be addressed by allowing secondary users to dynamic access the licensed band when its primary user is absent. *Cognitive radio* is one of the key technologies that could improve the spectral utilization efficiency as suggested by Prof. S. Haykin [2]:

> Cognitive radio is viewed as a novel approach for improving the utilization of a precious natural resource: the radio electromagnetic spectrum.

1.1 Cognitive radio definition and components

The term *cognitive radio*, first coined by Dr. J. Mitola [4], has the following formal definition [2]:

> Cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment (i.e. outside world), and uses the methodology of understanding-by-building to learn from the environment and adapt its internal states to statistical variations in the incoming RF stimuli by making corresponding changes in certain operating parameters (e.g., transmit-power, carrier-frequency, and modulation strategy) in real-time, with two primary objectives in mind:
> • highly reliable communications whenever and wherever needed;
> • efficient utilisation of the radio spectrum.

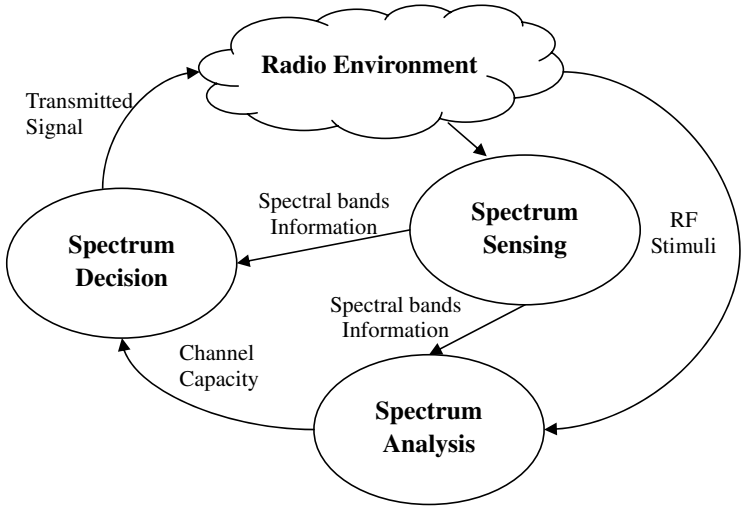

Fig. 2 The cognitive capability of cognitive radio enabled by a basic cognitive cycle [5].

From the definition, the key characteristic of cognitive radio is cognitive capability. It means that cognitive radio should interact with its environment, and intelligently determine appropriate communication parameters based on quality of service (QoS) requirements. These tasks can be implemented by a basic cognitive cycle as illustrated in Figure 2:

- *Spectrum sensing*: To improve the spectral utilization efficiency, cognitive radio should regularly monitor the RF spectral environment. Cognitive radio should not only find spectrum holes, which are not currently used by primary users, by scanning the whole RF spectrum, but also needs to detect the status of primary users for avoiding causing potential interference.
- *Spectrum analysis*: After spectrum sensing, the characteristics of spectrum holes should be estimated. The following parameters need to be known, e.g., channel side information, capacity, delay, and reliability, and will be delivered to the spectrum decision step.
- *Spectrum decision*: Based on the characteristics of spectrum holes, an appropriate spectral band will be chosen for a particular cognitive radio node according to its QoS requirement while considering the whole network fairness. After that, cognitive radio could determine new configuration parameters, e.g., data rate, transmission mode, and bandwidth of the transmission, and then reconfigure itself by using software defined radio techniques.

1.2 Applications of Cognitive Radio Networks

Because cognitive radio is aware of the RF spectral environment and is capable of adapting its transmission parameters to the RF spectral environment, cognitive radio and the concepts of cognitive radio can be applied to a variety of wireless communication environments, especially in commercial and military applications. A few of applications are listed below:

- *Coexistence of wireless technologies:* Cognitive radio techniques were primarily considered for reusing the spectrum that is currently allocated to the TV service. Wireless regional area network (WRAN) users can take advantage of broadband data delivery by the opportunistic usage of the underutilized spectrum. Additionally, the dynamic spectrum access techniques will play an important role in full interoperability and coexistence among diverse technologies for wireless networks. For example, cognitive radio concepts can be used to optimize and manage the spectrum when the wireless local area network (WLAN) and the Bluetooth devices coexist.
- *Military networks:* In military communications, bandwidth is often at a premium. By using cognitive radio concepts, military radios can not only achieve substantial spectral efficiency on a noninterfering basis, but also reduce implementation complexity for defining the spectrum allocation for each user. Furthermore, military radios can obtain benefits from the opportunistic spectrum access function supported by the cognitive radio [6]. For example, the military radios can adapt their transmission parameters to use Global System for Mobile (GSM) bands, or other commercial bands when their original frequencies are jammed. The mechanism of spectrum management can help the military radios achieve information superiority on the battlefield. Furthermore, from the soldiers' perspective, cognitive radio can help the soldiers to reach an objective through its situational awareness.
- *Heterogeneous wireless networks:* From a user's point of view, a cognitive radio device can dynamically discover information about access networks, e.g. WiFi and GSM, and makes decisions on which access network is most suitable for its requirements and preferences. Then the cognitive radio device will reconfigure itself to connect to the best access network. When the environmental conditions change, the cognitive radio device can adapt to these changes. The information as seen by the cognitive radio user is as transparent as possible to changes in the communication environment.

2 Traditional Spectrum Sensing Algorithms

As a key technology in cognitive radio, spectrum sensing should sense spectrum holes and detect the presence/absence of primary users. The most efficient way to sense spectrum holes is to detect active primary transceivers in the vicinity of cogni-

tive radios. However, as some primary receivers are passive, such as TVs, some are difficult to detect in practice. Tractional spectrum sensing techniques can be used to detect the primary transmitters, i.e., matched filtering [7], energy detection [8], cyclostationary detection [9], and wavelet based detection [10]. The implementation of these algorithms requires different conditions, and their detection performance are correspondingly distinguished. The advantages and disadvantages of these algorithms are summarized in Table 1.

Table 1 Summary of advantages and disadvantages of traditional spectrum sensing algorithms.

Spectrum sensing algorithm	Advantages	Disadvantages
Matched filter [7]	Optimal performance Low computational cost	Require prior information of the primary user
Energy detection [8]	Do not require prior information Low computational cost	Poor performance for low SNR Cannot differentiate users
Cyclostationary [9]	Valid in slow SNR region Robust against interference	Require partial prior information High computational cost
Wavelet based detection [10]	Valid for dynamic and wideband spectrum sensing	High sampling rate High computational cost

2.1 Matched filter

A block diagram of a matched filter is shown in Figure 3(a). The matched filter method is an optimal approach for spectrum sensing in the sense that it maximizes the signal-to-noise ratio (SNR) in the presence of additive noise [11]. Another advantage of the matched filter method is that it requires less observation time since the high processing gain can be achieved by coherent detection. For example, to meet a given probability of detection, only $\mathcal{O}(1/\mathrm{SNR})$ samples are required [7]. This advantage is achieved by correlating the received signal with a template to detect the presence of a known signal in the received signal. However, it relies on prior knowledge of the primary user, such as modulation type, and packet format, and requires cognitive radio to be equipped with carrier synchronization and timing devices. With more types of primary users, the implementation complexity grows making the matched filter impractical.

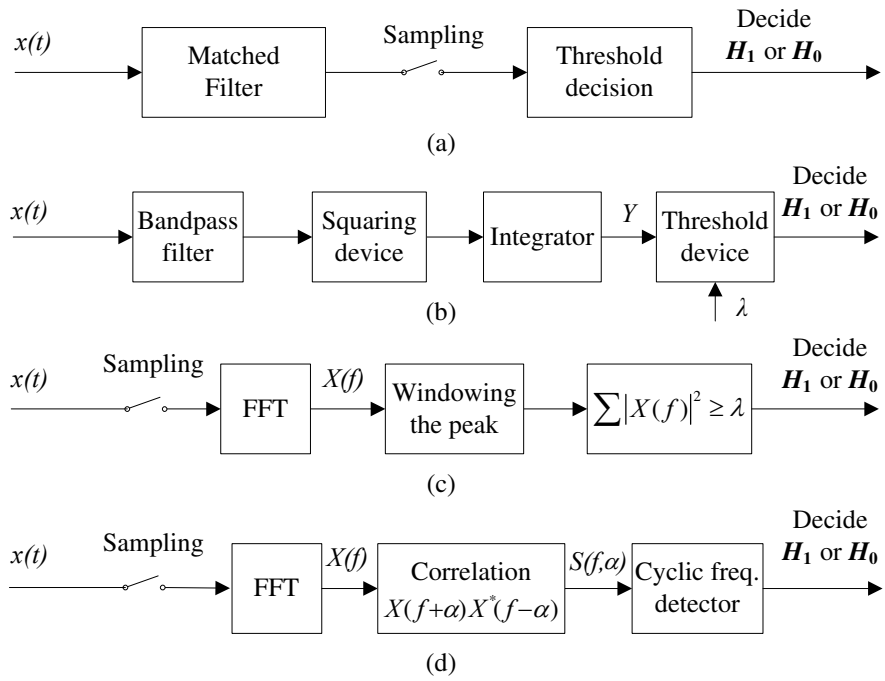

Fig. 3 Block diagrams for traditional spectrum sensing algorithms: (a) matched filter, (b) time domain energy detection, (c) frequency domain energy detection, and (d) cyclostationary detection.

2.2 Energy detection

If the information about the primary user is unknown in cognitive radio, a commonly used method for detecting the primary users is energy detection [8]. Energy detection is a non-coherent detection method that avoids the need for complicated receivers required by a matched filter. An energy detector can be implemented in both the time and the frequency domain. For time domain energy detection as shown in Figure 3(b), a bandpass filter (BPF) is applied to select a center frequency and bandwidth of interest. Then the energy of the received signal is measured by a magnitude squaring device, with an integrator to control the observation time. Finally, the energy of the received signal will be compared with a predetermined threshold to decide whether the primary user is present or not. However, to sense a wide spectrum span, sweeping the BPF will result in a long measurement time. As shown in Figure 3(c), in the frequency domain, the energy detector can be implemented similarly to a spectrum analyzer with a fast Fourier transform (FFT). Specifically, the received signal is sampled at or above the Nyquist rate over a time window. Then the power spectral density (PSD) is computed using an FFT. The FFT is employed to analyze a wide frequency span in a short observation time, rather than sweeping

the BPF in Figure 3(b). Finally, the PSD will be compared with a threshold, λ, to decide whether the corresponding frequency is occupied or not.

The advantages of energy detection are that prior knowledge of the primary users is not required, and both the implementation and the computational complexity are generally low. In addition, a short observation time is required, for example, $\mathcal{O}(1/\mathrm{SNR}^2)$ samples are required to satisfy a given probability of detection [7]. Although energy detection has a low implementation complexity, it has some drawbacks. A major drawback is that it has poor detection performance under low SNR scenarios as it is a non-coherent detection scheme. Another drawback is that it cannot differentiate between the signal from a primary user and the interference from other cognitive radios, thus, it cannot take advantage of adaptive signal processing, such as interference cancelation. Furthermore, noise level uncertainty can lead to further performance loss. These disadvantages can be overcome by using two-stage spectrum sensing technique, i.e. coarse spectrum sensing and fine spectrum sensing. Coarse spectrum sensing can be implemented by energy detection or wideband spectrum analyzing techniques. The aim of coarse spectrum sensing is to quickly scan the wideband spectrum and identify some possible spectrum holes in a short observation time. By contrast, fine spectrum sensing further investigates and analysis these suspected frequencies. More sophisticated detection techniques can be used at this stage, such as cyclostationary detection described below.

2.3 Cyclostationary detection

A block diagram of cyclostationary detection is shown in Figure 3(d). Cyclostationary detection is a method for detecting the primary users by exploiting the cyclostationary features in the modulated signals. In most cases, the received signals in cognitive radios are modulated signals, which in general exhibit built-in-periodicity within the training sequence or cyclic prefixes. This periodicity is generated by the primary transmitter so that the primary receiver can use it for parameter estimation, such as channel estimation, and pulse timing [12]. The cyclic correlation function, also called cyclic spectrum function (CSF), is used for detecting signals with a particular modulation type in the presence of noise. This is because noise is usually wide-sense stationary (WSS) without correlation, by contrast, modulated signals are cyclostationary with spectral correlation. Furthermore, since different modulated signals will exhibit different characteristics, cyclostationary detection can be used for distinguishing between different types of transmitted signals, noise, and interference in low SNR environments. One of the drawbacks of cyclostationary detection is that it still requires partial information of the primary user. Another drawback is that the computational cost is high as the CSF is a two-dimensional function dependent on frequency and cyclic frequency [9].

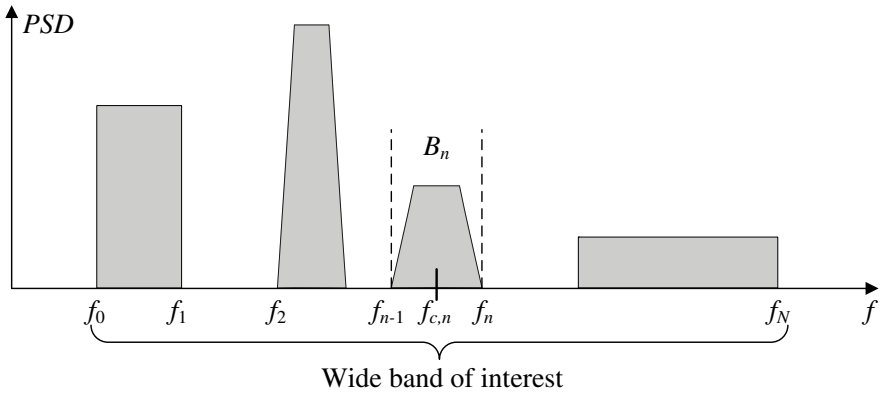

Fig. 4 Demonstration of the Fourier spectrum of interest. The PSD is smooth within each subband, and exhibits discontinuities and irregularities with the adjacent subbands [10, 13].

2.4 Wavelet based spectrum sensing

In [10], Tian and Giannakis proposed a wavelet-based spectrum sensing approach. It provides an advantage of flexibility in adapting to a dynamic spectrum. In this approach, the PSD of the Fourier spectrum is modeled as a train of consecutive frequency subbands, where the PSD is smooth within each subband but exhibits discontinuities and irregularities on the border of two neighboring subbands as shown in Figure 4. The wavelet transform of the wideband PSD is used to locate the singularities of the PSD.

Let $\varphi(f)$ be a wavelet smoothing function, the dilation of $\varphi(f)$ is given by

$$\varphi_d(f) = \frac{1}{d}\varphi\left(\frac{f}{d}\right) \tag{1}$$

where d is a dyadic scale that can take values that are powers of 2, i.e. $d = 2^j$. The continuous wavelet transform (CWT) of the PSD is given by [10]

$$\mathrm{CWT}\{S(f)\} = S(f) * \varphi_d(f) \tag{2}$$

where "$*$" denotes the convolution and $S(f)$ is the PSD of the received signal.

Then the first and second derivative of the $\mathrm{CWT}\{S(f)\}$ are used to locate the irregularities and discontinuities in the PSD. Specifically, the boundaries of each subbands are located by using the local maxima of the first derivative of $\mathrm{CWT}\{S(f)\}$, and locations of the subbands are finally tracked by finding zero crossings in the second derivative of $\mathrm{CWT}\{S(f)\}$. By controlling the wavelet smoothing function, the wavelet-based spectrum sensing approach has flexibility in adapting to the dynamic spectrum.

3 Wideband Spectrum Sensing Algorithms

As the discussions in previous section, spectrum sensing is composed of data acquisition (sampling) process and decision-making process. For implementing wideband data acquisition, cognitive radio needs some essential components, i.e., wideband antenna, wideband RF front end, and high speed analog-to-digital converter (ADC). Considering the Nyquist sampling theory, the sampling rate of ADC is required to exceed $2W$ samples per second (known as Nyquist rate), if W denotes the bandwidth of the received signal (e.g., bandwidth $W = 10$ GHz). In [14], Yoon *et al.* have shown that the -10 dB bandwidth of the newly designed antenna can be 14.2 GHz. Hao and Hong [15] have designed a compact highly selective wideband bandpass filter with a bandwidth of 13.2 GHz. By contrast, the development of ADC technology is relatively behind. When we require an ADC to have a high resolution and a reasonable power consumption, the achievable sampling rate of the state-of-the-art ADC is 3.6 Gsps [16]. Thus, ADC becomes a bottleneck in such a wideband data acquisition system. Even if there exists ADC with more than 20 Gsps sampling rate, the real-time digital signal processing of 20 Gb/s of data could be very expensive. This dilemma motivates researchers to look for technologies to reduce the sampling rate while retaining W by using sub-Nyquist sampling techniques.

Sub-Nyquist sampling refers to the problem of recovering signals from partial measurements that are obtained by using sampling rate lower than the Nyquist rate [17]. Three important sub-Nyquist sampling techniques are: multi-coset sub-Nyquist sampling, multi-rate sub-Nyquist sampling, and compressed sensing based sub-Nyquist sampling.

3.1 Multi-coset Sub-Nyquist Sampling

Multi-coset sampling is a selection of some samples from a uniform grid, which can be obtained when uniformly sampling signal at a rate of f_N greater than the Nyquist rate. The uniform grid is then divided into blocks of L consecutive samples, and in each block $v(v < L)$ samples are retained while the rest of samples, i.e. $L - v$ samples, are skipped. A constant set C that describes the indexes of these v samples in each block is called a sampling pattern as

$$C = \{t^i\}_{i=1}^{v}, \quad 0 \le t^1 < t^2 < \cdots < t^v \le L - 1. \tag{3}$$

As shown in Figure 5, the multi-coset sampling can be implemented by using v sampling channels with sampling rate of $\frac{f_N}{L}$, where the i-th sampling channel is offset by $\frac{t^i}{f_N}$ from the origin as below

$$x^i[n] = \begin{cases} x(\frac{n}{f_N}), & n = mL + t^i,\ m \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

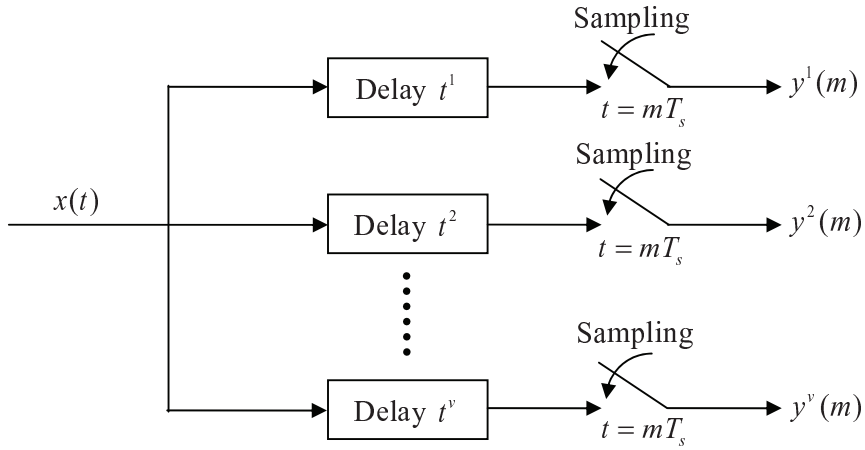

Fig. 5 Block diagram of multi-coset sub-Nyquist sampling.

where $x(t)$ denotes the received signal to be sampled.

The discrete-time Fourier transform (DTFT) of the samples can be linked to the unknown Fourier transform of signal $x(t)$ by

$$\mathbf{Y}(f) = \Phi\mathbf{X}(f) \tag{5}$$

where $\mathbf{Y}(f)$ denotes a vector of DTFT of these measurements from v channels, $\mathbf{X}(f)$ is a vector of the Fourier transform of $x(t)$, and Φ is the measurement matrix whose elements are determined by the sampling pattern C. The problem of wideband spectrum sensing is thus equivalent to recovering $\mathbf{X}(f)$ from $\mathbf{Y}(f)$. In order to get a unique solution from (5), every set of v columns of Φ should be linearly independent. However, searching for this sampling pattern is a combinatorial problem.

In [18, 19], some sampling patterns are proved to be valid for reconstruction. The advantage of multi-coset sampling is that the sampling rate in each channel is L times lower than the Nyquist rate. Moreover, the number of measurements is $\frac{v}{L}$ lower than the Nyquist sampling case. One drawback of the multi-coset sampling is that accurate time offsets between sampling channels are required to satisfy a specific sampling pattern. Another one is that the number of sampling channels should be sufficiently high [20].

3.2 Multi-rate Sub-Nyquist Sampling

An alternative model for compressing the wideband spectrum in the analog domain is a multirate sampling system as shown in Figure 6. Asynchronous multirate sampling (MRS) and synchronous multirate sampling (SMRS) were used for recon-

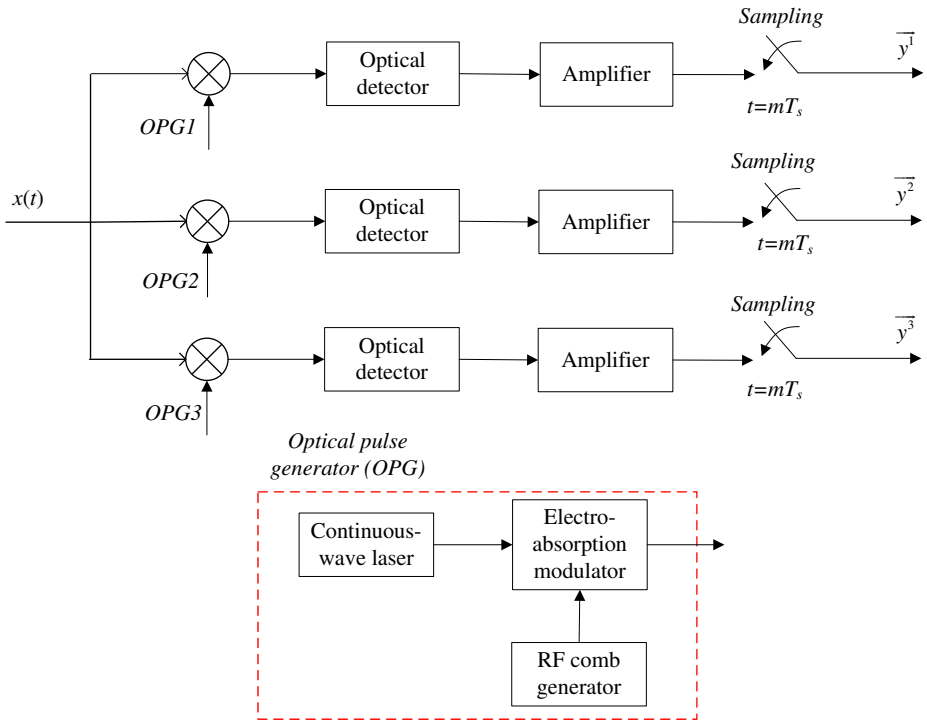

Fig. 6 Multirate sampling system implemented by electro-optical devices [21]. In each channel, the received signal is modulated by a train of short optical pulses. The modulated signal is then detected by an optical detector, amplified, and sampled by a low-rate ADC.

structing sparse multiband signals in [22] and [23], respectively. In addition, MRS has been successfully implemented in experiments using an electro-optical system with three sampling channels as described in [21]. Both systems employ three optical pulsed sources that operate at different rates and at different wavelengths. The received signal is modulated with optical pulses, which provided by an optical pulse generator (OPG), in each channel. In order to reconstruct a wideband signal with an 18 GHz bandwidth, the modulated pulses are amplified, and sampled by an ADC at a rate of 4 GHz in each channel.

In [22], the sampling channels of the MRS can be implemented separately without synchronisation. However, reconstruction of the spectrum requires that each frequency of the signal must be non-aliased in at least one of the sampling channels. In [23] SMRS reconstructs the spectrum from linear equations, which relate the Fourier transform of the signal to the Fourier transform of its samples. Using compressed sensing theory, sufficient conditions for perfectly reconstructing the spectrum are obtained; $v \geq 2k$ (the Fourier transform of the signal is k-sparse) sampling channels are required. In order to reconstruct the spectrum using MRS with fewer sampling channels, the spectrum to be recovered should possess certain properties,

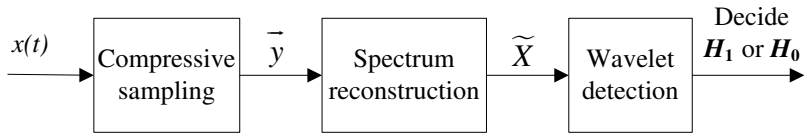

Fig. 7 Block diagram of the compressed sensing based wideband spectrum sensing algorithm.

e.g., minimal bands, and uniqueness. Nonetheless, the spectral components from primary users may not possess these properties. Obviously, even though the multi-rate sampling system has broad application, there is a long way to go to implement it in a cognitive radio network because of its stringent requirements on both optical devices and the number of sampling channels.

3.3 Compressed sensing based sub-Nyquist sampling

In the classic work [13], Tian and Giannakis introduced compressed sensing theory to realize wideband spectrum sensing by exploiting the sparsity of radio signals. The technique takes advantage of using fewer samples closer to the information rate, rather than the inverse of the bandwidth, to perform wideband spectrum sensing. After reconstruction of the wideband spectrum, wavelet-based edge detection was used to detect the wideband spectrum as shown in Figure 7.

Let $x(t)$ represent a wideband signal received at cognitive radio. If $x(t)$ is sampled at the Nyquist sampling rate, the sequence vector, i.e. $\mathbf{x}$ ($\mathbf{x} \in \mathbb{C}^N$), will be obtained. The Fourier transform of the sequence, $\mathbf{X} = \mathbf{F}\mathbf{x}$, will therefore be alias-free, where $\mathbf{F}$ denotes the Fourier matrix. When the spectrum, $\mathbf{X}$, is k-sparse ($k \ll N$), which means k out of N values in $\mathbf{X}$ are not neglectable, $x(t)$ can be sampled at a sub-Nyquist rate while its spectrum can be reconstructed with a high probability. The sub-sampled/compressed signal, $\mathbf{y} \in \mathbb{C}^M$ ($k < M \ll N$), is linked to the Nyquist sequence $\mathbf{x}$ by [13],

$$\mathbf{y} = \Phi\mathbf{x} \tag{6}$$

where $\Phi \in \mathbb{C}^{M\times N}$ is the measurement matrix, which is a selection matrix that randomly chooses M columns of the size-N identity matrix. Namely, $N - M$ samples out of N samples are skipped. The relationship between the spectrum $\mathbf{X}$ and the compressed sequence $\mathbf{y}$ is given by [13]

$$\mathbf{y} = \Phi\mathbf{F}^{-1}\mathbf{X} \tag{7}$$

where $\mathbf{F}^{-1}$ denotes the inverse Fourier matrix.

Approximating $\mathbf{X}$ from $\mathbf{y}$ in (7) is a linear inverse problem and is NP-hard. The basis pursuit (BP) [24] algorithm can be used to solve $\mathbf{X}$ by linear programming [13]:

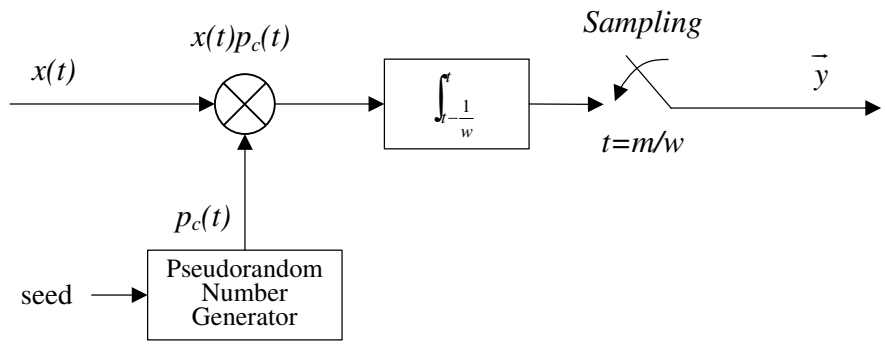

Fig. 8 Block diagram for the analog-to-information converter [26]. The received signal, $x(t)$, is randomly demodulated by a pseudorandom chipping sequence, integrated by an accumulator, and sampled at a sub-Nyquist rate.

$$\widetilde{X} = \arg\min \|\mathbf{X}\|_1, \ \text{ s. t. } \ \mathbf{y} = \boldsymbol{\Phi}\mathbf{F}^{-1}\mathbf{X}. \tag{8}$$

After reconstructing the full spectrum $\mathbf{X}$, the PSD is calculated using $\widetilde{X}$. Then the wavelet detection approach can be used to analyze the edges in the PSD. Although less measurements are used for characterizing the wideband spectrum, the requirement of high sampling rate on ADC is not relaxed. By contrast, in [25], Polo *et al.* suggested using an analog-to-information converter (AIC) model (also known as random demodulator, [26]) for compressing the wideband signal in the analog domain. The block diagram of AIC is given in Figure 8.

A pseudorandom number generator is used to produce a discrete-time sequence $\varepsilon_0, \varepsilon_1, \cdots$, called a chipping sequence, the number of which takes values of ± 1 with equal probability. The waveform should randomly alternate at or above the Nyquist rate, i.e. $\varpi \geq 2W$, where W is the bandwidth of signal. The output of the pseudorandom number generator, i.e. $p_c(t)$, is employed to demodulate a continuous-time input $x(t)$ by a mixer. Then an accumulator sums the demodulated signal for $1/w$ seconds, and the filtered signal is sampled at a sub-Nyquist rate of w. This sampling approach is called integrate-and-dump sampling since the accumulator is reset after each sample is taken. The samples acquired by the AIC, $\mathbf{y} \in \mathbb{C}^w$, can be related to the received signal, $\mathbf{x} \in \mathbb{C}^{\varpi}$, by

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{x} \tag{9}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{w \times \varpi}$ is the measurement matrix describing the overall action of the AIC system on the input signal $\mathbf{x}$. The signal $\mathbf{x}$ can be identified by solving the convex optimization problem,

$$\widetilde{x} = \arg\min \|\mathbf{x}\|_1, \ \text{ s. t. } \ \mathbf{y} = \boldsymbol{\Phi}\mathbf{x}, \tag{10}$$

by BP or other greedy pursuit algorithms. The PSD of the wideband spectrum can be estimated using the recovered signal $\widetilde{x}$, followed by a hypothesis test on the PSD. Alternatively, the PSD can be directly recovered from the measurements using com-

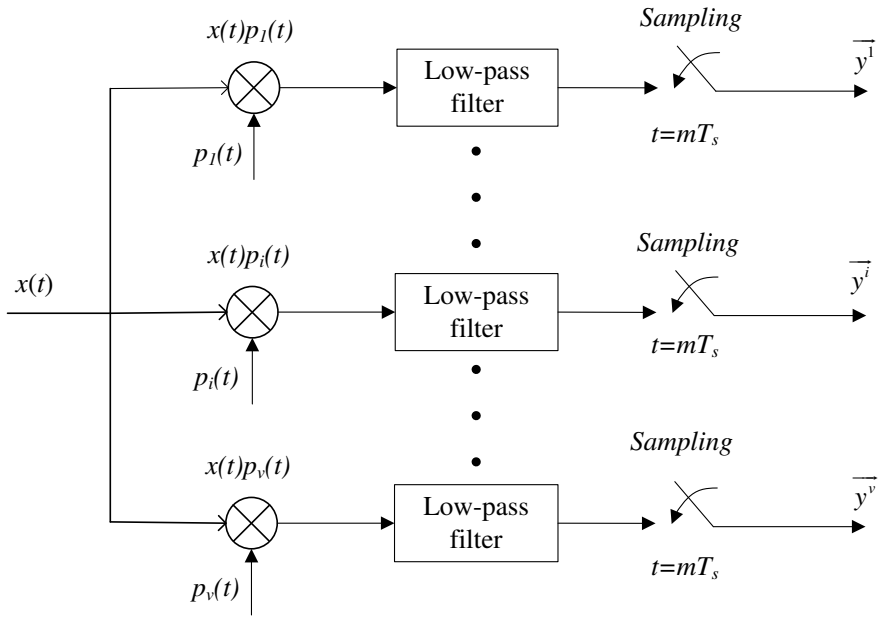

Fig. 9 Block diagram for the modulated wideband converter [27]. In each channel, the received signal is demodulated by a pseudorandom sequence, filtered by a low-pass filter, and sampled at a sub-Nyquist rate $\frac{1}{T_s}$.

pressed sensing algorithms [25]. Although the AIC bypasses the requirement for a high sampling rate ADC, it leads to a high computational complexity as the huge-scale of the measurement matrix. Furthermore, it has been identified that the AIC model can easily be influenced by design imperfections or model mismatches [27].

In [27], Mishali and Eldar proposed a parallel implementation of the AIC model, called modulated wideband converter (MWC), as shown in Figure 9. The key difference is that in each channel the accumulator for integrate-and-dump sampling is replaced by a general low-pass filter. One of the benefits of introducing parallel structure is that the dimension of the measurement matrix is reduced making the reconstruction easier. Another benefit is that it provides robustness to noise and model mismatch. On the other hand, the implementation complexity increases as multiple sampling channels are involved. An implementation issue of using MWC is that the storage and transmission of the measurement matrix must be considered when it is used in a distributed cognitive radio network under a data fusion collaborative scheme.

4 Adaptive Compressed Sensing Framework for Wideband Spectrum Sensing

The compressed sensing technologies require that the signal to be sampled should be sparse in a suitable basis. If it is sparse, the signal can be reconstructed from partial measurements by using some recovery algorithms, e.g., orthogonal matching pursuit (OMP) or compressive sampling matching pursuit (CoSaMP)[28]. Given the low spectral occupancy, the wideband signal that is received by cognitive radios can be assumed to be sparse in the frequency domain [13]. If this sparsity level (denoted by k) is known, we can choose an appropriate number of measurements M to secure the quality of spectral recovery, e.g., $M = C_0 k \log(N/k)$, where C_0 denotes a constant and N denotes the number of measurements when using the Nyquist rate [13]. However, in order to avoid incorrect spectral recovery in the cognitive radio system, traditional compressed sensing approaches must pessimistically choose the parameter C_0, which results in excessive number of measurements. As shown in Fig. 10, considering $k = 10$, traditional compressed sensing approaches tend to choose $M = 37\%N$ measurements for achieving a high successful recovery rate. We note that, with $20\%N$ measurements, we can still achieve 50% successful recovery rate. If these 50% successful recovery cases can be identified, we could save the number of measurements. In addition, in a practical cognitive radio system, the sparsity level of the instantaneous spectrum is often unknown or difficult to estimate because of either the dynamic activities of primary users or the time-varying fading channels between the primary users and cognitive radios. Due to this sparsity level uncertainty, traditional compressed sensing approaches should further increase the number of measurements. For example, in Fig. 10, if k is known to be $10 \leq k \leq 20$, traditional compressed sensing approaches would select $M = 50\%N$, which does not fully exploit the advantages of using compressed sensing technologies for wideband spectrum sensing. Further, the sparsity level uncertainty could also result in early or late termination of greedy recovery algorithms. Due to the effects of under-fitting or over-fitting caused by the early or late iteration termination, traditional compressed sensing recovery algorithms will lead to unfavorable spectral recovery quality.

To address these challenges, adaptive compressed sensing approach should be adopted for reconstructing the wideband spectrum by using an appropriate number of compressive measurements without prior knowledge of the instantaneous spectral sparsity level. Specifically, the adaptive framework divides the spectrum sensing interval into several equal-length time slots, and performs compressive measurements in each time slot. The measurements are then partitioned into two complementary subsets, performing the spectral recovery on the training subset, and validating the recovery result on the testing subset. Both the signal acquisition and the spectral estimation will be terminated if the designed ℓ_1 norm validation parameter meets certain requirements. In the next section, we will introduce the adaptive compressed sensing approach in detail for addressing wideband spectrum sensing issues in cognitive radios.

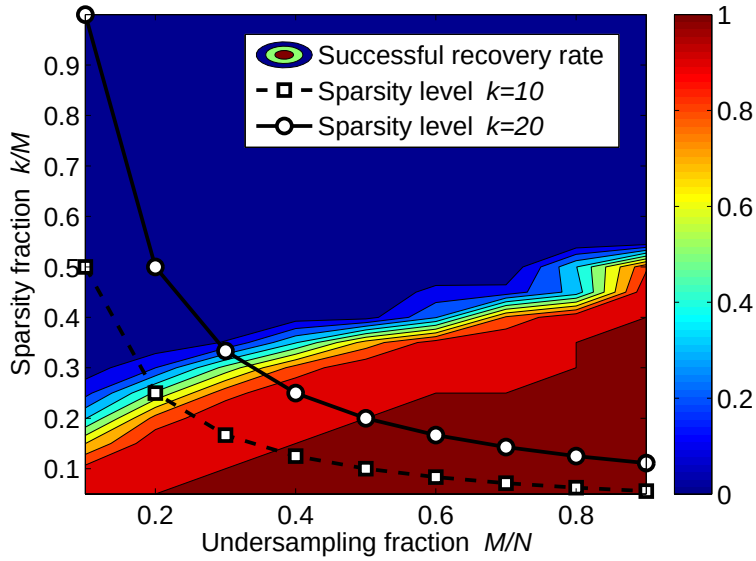

Fig. 10 An example of a traditional compressed sensing system, where the successful recovery rate varies when the number of measurements and the sparsity level vary. In simulations, considering $N = 200$, we varied the number of measurements M from 20 to 180 in eight equal-length steps. The sparsity level k was set to between 1 and M. The measurement matrix was assumed to be Gaussian. The figure was obtained with 5000 trials of each parameter setting.

4.1 Problem Statement

Suppose that an analog primary signal $x(t)$ is received at a cognitive radio, and the frequency range of $x(t)$ is $0 \sim W$ (Hz). If the signal $x(t)$ were sampled at the sampling rate f (Hz) in the observation time τ (seconds), a signal vector $\mathbf{x} \in \mathbb{C}^{N\times 1}$ would be obtained, where N denotes the number of samples and can be written as $N = f\tau$. Without loss of generality, we assume that N is an integer number. However, here we consider that the signal is sampled at sub-Nyquist rate as enhanced by compressed sensing.

The compressed sensing theory relies on the fact that we can represent many signals using only a few non-zero coefficients in a suitable basis or dictionary. Such signals may therefore be acquired by sub-Nyquist sampling, which leads to fewer samples than predicted on the basis of Nyquist sampling theory. The sub-Nyquist sampler, e.g., the random demodulator [26, 29, 30], will generate a vector of compressive measurements $\mathbf{y} \in \mathbb{C}^{M\times 1}$ ($M \ll N$) via random projections of the signal vector $\mathbf{x}$. Mathematically, the compressive measurement vector $\mathbf{y}$ can be written as

$$\mathbf{y} = \Phi\mathbf{x} \tag{11}$$

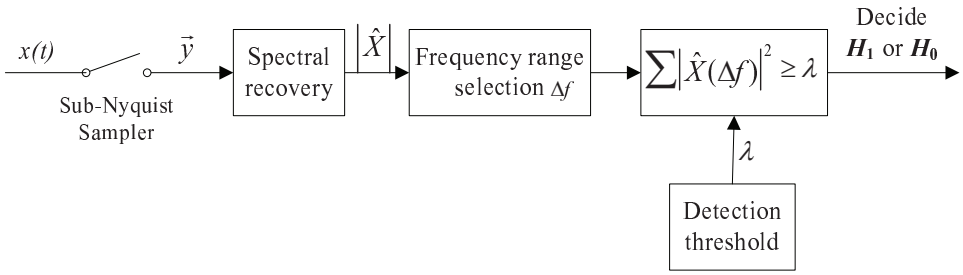

Fig. 11 Diagram of compressed sensing based spectrum sensing approach when using the spectral domain energy detection approach.

where $\mathbf{x}$ denotes the signal vector obtained by using sampling rate higher than or equal to the Nyquist rate (i.e., $f \geq 2W$), and Φ denotes an $M \times N$ measurement matrix. Of course, there is no hope to reconstruct an arbitrary N-dimensional signal $\mathbf{x}$ from partial measurements $\mathbf{y}$. However, if the signal $\mathbf{x}$ is k-sparse ($k < M \ll N$) in some basis, there do exist measurement matrices that allow us to recover $\mathbf{x}$ from $\mathbf{y}$ using some recovery algorithms.

Based on the fact of spectral sparseness in a cognitive radio system [13], the compressed sensing technologies can be applied for signal acquisition at cognitive radios. A block diagram of a typical compressed sensing based spectrum sensing infrastructure is shown in Fig. 11. The goal is to reconstruct the Fourier spectrum $\mathbf{X} = \mathbf{Fx}$ from partial measurements $\mathbf{y}$, and to perform spectrum sensing based on the reconstructed spectrum $\hat{X}$. Due to the advantages of short running time and good sampling efficiency, greedy recovery algorithms are often used in some practical scenarios where the signal processing should be performed on a near real-time basis in addition to computational capability constraints.

After the spectral recovery, spectrum sensing approaches can be performed by using the reconstructed spectrum $\hat{X}$. A typical spectrum sensing approach is spectral domain energy detection as the discussions in Section 2. As depicted in Fig. 11, this approach extracts the reconstructed spectrum in the frequency range of interest, e.g., Δf, and then calculates the signal energy in the spectral domain. The output energy will be compared with a detection threshold (denoted by λ) to decide whether the corresponding frequency band is occupied or not, i.e., choosing between hypotheses $\mathscr{H}_1$ (presence of primary users) and $\mathscr{H}_0$ (absence of primary users).

It can be easily understood that the performance of such an infrastructure will highly depend on the recovery quality of the Fourier spectrum $\mathbf{X}$. From the compressed sensing theory, we know that the recovery quality is determined by: the sparsity level, the choice of measurement matrix, the recovery algorithm, and the number of measurements. The spectral sparsity level in a cognitive radio system is mainly determined by the activities of primary users within a specific frequency range and the medium access control (MAC) of the cognitive radios. One elegant metric for evaluating the suitability of a chosen measurement matrix is the restricted isometry property (RIP) [31]. For a comprehensive understanding of RIP and mea-

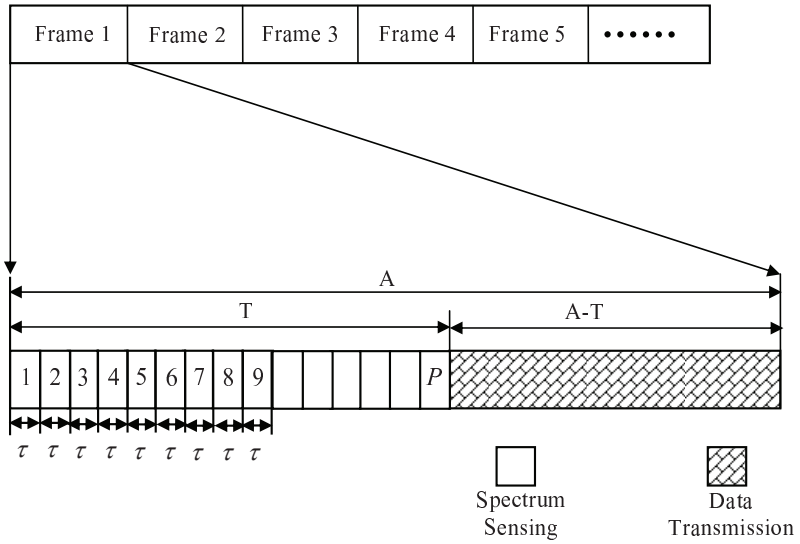

Fig. 12 Frame of periodic spectrum sensing in cognitive radio networks.

surement matrix design, we refer the reader to [32] and references therein. In the following, we will concentrate on addressing: the choice of the number of measurements and the design of the recovery algorithm. We will discuss an adaptive sensing framework enabling us to gradually acquire spectral measurements. Both the signal acquisition and the spectral estimation will be terminated when certain halting criterions are met, thereby avoiding the problems of excessive or insufficient numbers of compressive measurements.

4.2 System Description

Consider a cognitive radio system using a periodic spectrum sensing infrastructure in which each frame is comprised of a spectrum sensing time slot and a data transmission time slot, as shown in Fig. 12. The length of each frame is A (seconds), and the duration of spectrum sensing is T $(0 < T < A)$. The remaining time $A - T$ is used for data transmission. Further, we assume that the spectrum sensing duration T is carefully chosen so that the symbols from primary users, and the channels between the primary users and cognitive radios are quasi-stationary. We propose to divide the spectrum sensing duration T into P equal-length mini-time slots, each of which has length $\tau = \frac{T}{P}$, as depicted in Fig. 12. As enforced by protocols, e.g., at the MAC layer [33], all cognitive radios can keep quiet during the spectrum sensing interval. Therefore, the spectral components of the Fourier spectrum $\mathbf{X} = \mathbf{F}\mathbf{x}$ arise only from primary users and background noise. Due to the low spectral occupancy [13], the

Table 2 Compressed Adaptive Sensing (CASe) Framework

Input: Sensing duration T, N, noise variance δ^2, threshold ϖ accuracy ε in the noiseless case, accuracy ε in the noisy case.
1. *Initialize:* Divide T into P time slots, each has length $\tau = \frac{T}{P}$, index $p = 0$. 2. *While* the halting criterion is false and $p < P$, *do* a). Increment p by 1. b). Perform compressive sampling in the time slot p using rate f_s. c). If $p = 1$, partition the measurement vector into: the training set $\mathbf{y}_1$ and testing set $\mathbf{V}$ as in (12)-(13). d). Concatenate the training sets from the time slots $1, \cdots, p$ to form $\mathbf{Y}_p$ as in (13). e). Estimate the spectrum from $\mathbf{Y}_p$ using spectral recovery algorithm resulting in the spectral estimate $\hat{X}_p$. f). Calculate the validation parameter using $\mathbf{V}$ and Ψ: $\rho_p = \frac{\lVert \mathbf{V} - \Psi \mathbf{F}^{-1} \hat{X}_p \rVert_1}{r}$. 3. *Check and make decision:* *If* the halting criterion is true a). Terminate the signal acquisition. b). Perform spectrum sensing using the reconstructed spectrum $\hat{X}_p$. c). Choose un-occupied bands, and start the data transmission. *Else if* $p = P$ a). Terminate the signal acquisition. b). Report its reconstruction is not trustworthy. c). Increase f_s and wait for next spectrum sensing frame. *end*
Halting Criterion: $\frac{\sqrt{\frac{\pi N}{2}}\rho_p}{1-\varepsilon} \leq \varpi$, in the noiseless measurement case. $\lvert \rho_p - \sqrt{\frac{\pi}{2}}\delta \rvert \leq \varepsilon$, in the noisy measurement case.

Fourier spectrum $\mathbf{X}$ can be assumed to be k-sparse, which means it consists only of k largest values that are not negligible. The spectral sparsity level k is unknown except that $k \leq k_{\max}$, where $k_{\max}$ is a known parameter. This assumption is reasonable because the maximal occupancy of the spectrum can be estimated by long-term spectral usage measurements.

For simplicity, we name the adaptive compressed sensing-based wideband spectrum sensing approach as: compressed adaptive sensing (CASe). The aim of CASe is to gradually acquire compressive measurements, reconstruct the wideband spectrum $\mathbf{X}$, and terminate the signal acquisition if and only if the current spectral recovery performance is satisfactory. The work procedure of CASe is shown in Table 2. We assume that cognitive radio performs compressive measurements using the same sub-Nyquist sampling rate f_s ($f_s < 2W$) in all P mini time slots. In each time slot, an m-length measurement vector would be obtained, where $m = f_s \tau = \frac{f_s T}{P}$ is assumed to be an integer. Without loss of generality, the measurement matrices of P time slots are assumed to follow the same distribution, e.g., the standard normal distribution, or the Bernoulli distribution with equal probability of ± 1. We partition the measurement set of the first time slot into two complementary subsets, i.e., validating the spectral recovery result using the testing subset $\mathbf{V}$ ($\mathbf{V} \in \mathbb{C}^{r \times 1}$, $0 < r < m$) which

is given by

$$\mathbf{V} = \Psi \mathbf{F}^{-1} \mathbf{X} \tag{12}$$

and performing the spectral recovery using the training subset $\mathbf{y}_1$ ($\mathbf{y}_1 \in \mathbb{C}^{(m-r)\times 1}$), where $\Psi \in \mathbb{C}^{r\times N}$ denotes the testing matrix. The measurements of other time slots, i.e., $\mathbf{y}_i$, $\forall i \in [2,P]$, are used only as the training subsets for spectral recovery. We concatenate the training subsets of all p time slots as

$$\mathbf{Y}_p \triangleq \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_p \end{pmatrix} = \Phi_p \mathbf{F}^{-1} \mathbf{X}_p \tag{13}$$

where $\mathbf{Y}_p \in \mathbb{C}^{(pm-r)\times 1}$ denotes the concatenated measurement vector, Φ_p denotes the measurement matrix after p time slots, and $\mathbf{X}_p$ denotes the signal spectrum. It should be noted that Φ_p and the testing matrix Ψ are chosen to be different but have the same distribution, and the signal spectrum $\mathbf{X}_p$ is always noisy, e.g., due to the receiver noise. We then gradually estimate the spectrum from $\mathbf{Y}_1, \mathbf{Y}_2, \cdots, \mathbf{Y}_p$ using a certain compressed sensing recovery algorithm, leading to a sequence of spectral estimates $\hat{X}_1, \hat{X}_2, \cdots, \hat{X}_p$.

4.3 Acquisition Termination Metric

We hope that the signal acquisition procedure can be terminated if we find a good spectral approximation $\hat{X}_p$ that makes the spectral recovery error $\|\mathbf{X} - \hat{X}_p\|_2$ sufficiently small. The remaining spectrum sensing time slots, i.e., $p+1, \cdots, P$, can be used for data transmission. If this target can be achieved, we could not only improve the cognitive radio system throughput (due to the longer data transmission time), but could also obtain measurement savings, leading to both energy and computational savings. However, the spectral recovery error $\|\mathbf{X} - \hat{X}_p\|_2$ is typically not known as $\mathbf{X}$ is unknown under the sub-Nyquist sampling rate. Hence, when using traditional compressed sensing approaches, we do not know when we should terminate the signal acquisition procedure. In this chapter, we propose to use the following validation parameter as a proxy for $\|\mathbf{X} - \hat{X}_p\|_2$:

$$\rho_p \triangleq \frac{\|\mathbf{V} - \Psi \mathbf{F}^{-1} \hat{X}_p\|_1}{r} \tag{14}$$

and terminate the signal acquisition if the validation parameter ρ_p is smaller than a predetermined threshold. This is based on the following observation:

Theorem 1: Assume that $\Phi_1, \cdots, \Phi_P$ and Ψ follow the same distribution, i.e., either the standard normal distribution or the Bernoulli distribution with equal probability of ± 1. Let $\varepsilon \in (0, \frac{1}{2})$, $\xi \in (0,1)$, and $r = C\varepsilon^{-2} \log \frac{4}{\xi}$ (C is a constant). Then

using $\mathbf{V}$ for testing the spectral estimate $\hat{X}_p$, the validation parameter ρ_p satisfies:

$$\Pr\left[(1-\varepsilon)\|\mathbf{X}-\hat{X}_p\|_2 \leq \sqrt{\frac{\pi N}{2}}\rho_p \leq (1+\varepsilon)\|\mathbf{X}-\hat{X}_p\|_2\right] \geq 1-\xi \tag{15}$$

where ξ can also be written as $\xi = 4\exp(-\frac{r\varepsilon^2}{C})$.

The proof of Theorem 1 is given in Appendix A.

Remark 1: In Theorem 1, we can see that, with either higher ε or greater r, we have higher confidence for estimating the actual spectral recovery error $\|\mathbf{X}-\hat{X}_p\|_2$. Fig. 13(a) shows the influence of using different number of measurements for testing the spectral estimate when the number of time slots increases. The spectral occupancy is assumed to be 6%, which means the spectral sparsity level $k = 6\%N = 120$ where $N = 2000$. It can be seen that with more testing data, the validation result is more trustworthy. Furthermore, we can find that even with $r = 5$ measurements for testing, the validation result is still very close to the actual recovery error. The choice of parameter C in Theorem 1 depends on the concentration property of random variables in the measurement matrix $\mathbf{\Psi}$. For a good $\mathbf{\Psi}$, e.g., the measurement matrix with random variables following either the Gaussian or Bernoulli distribution, C could be a small number.

Remark 2: Theorem 1 can be used to provide tight upper and lower bounds on the unknown recovery error $\|\mathbf{X}-\hat{X}_p\|_2$ by using (15) such that

$$\frac{\sqrt{\frac{\pi N}{2}}\rho_p}{1+\varepsilon} \leq \|\mathbf{X}-\hat{X}_p\|_2 \leq \frac{\sqrt{\frac{\pi N}{2}}\rho_p}{1-\varepsilon}. \tag{16}$$

Fig. 13(b) compares the actual recovery error $\|\mathbf{X}-\hat{X}_p\|_2$ and the validation parameter $\sqrt{\frac{\pi N}{2}}\rho_p$ when the spectral sparsity level varies. It is evident that the validation parameter can closely fit the unknown actual recovery error. The upper and lower bounds on the actual recovery error that we obtained in (16) can correctly predict the trend of the actual recovery error even if either p or k vary. Fig. 13(b) also illustrates that the lower the sparsity level, the fewer time slots (thereby the fewer compressive measurements) are required for reconstructing the spectrum. When the spectral occupancy is 12% (i.e., $k = 12\%N = 240$), the CASe framework requires $p = 7$ mini-time slots, i.e., $M = pm = 1400$ measurements in total. On the other hand, when $k = 100$, only $p = 3$ time slots and $M = pm = 600$ measurements are required. The remaining time slots can be used for data transmission, which can therefore lead to higher throughput than the cognitive radio system using traditional compressed sensing approaches. If we require $\|\mathbf{X}-\hat{X}_p\|_2$ (unknown) to be less than a tolerable recovery error threshold ϖ, we can let the upper bound on (16) to be a proxy for $\|\mathbf{X}-\hat{X}_p\|_2$. As shown in Table 2, we choose the upper bound on (16) as the signal acquisition termination metric in the noiseless case. If it is less than or equal to the threshold ϖ, i.e., $\|\mathbf{X}-\hat{X}_p\|_2 \leq \frac{\sqrt{\frac{\pi N}{2}}\rho_p}{1-\varepsilon} \leq \varpi$, the signal acquisition can be ter-

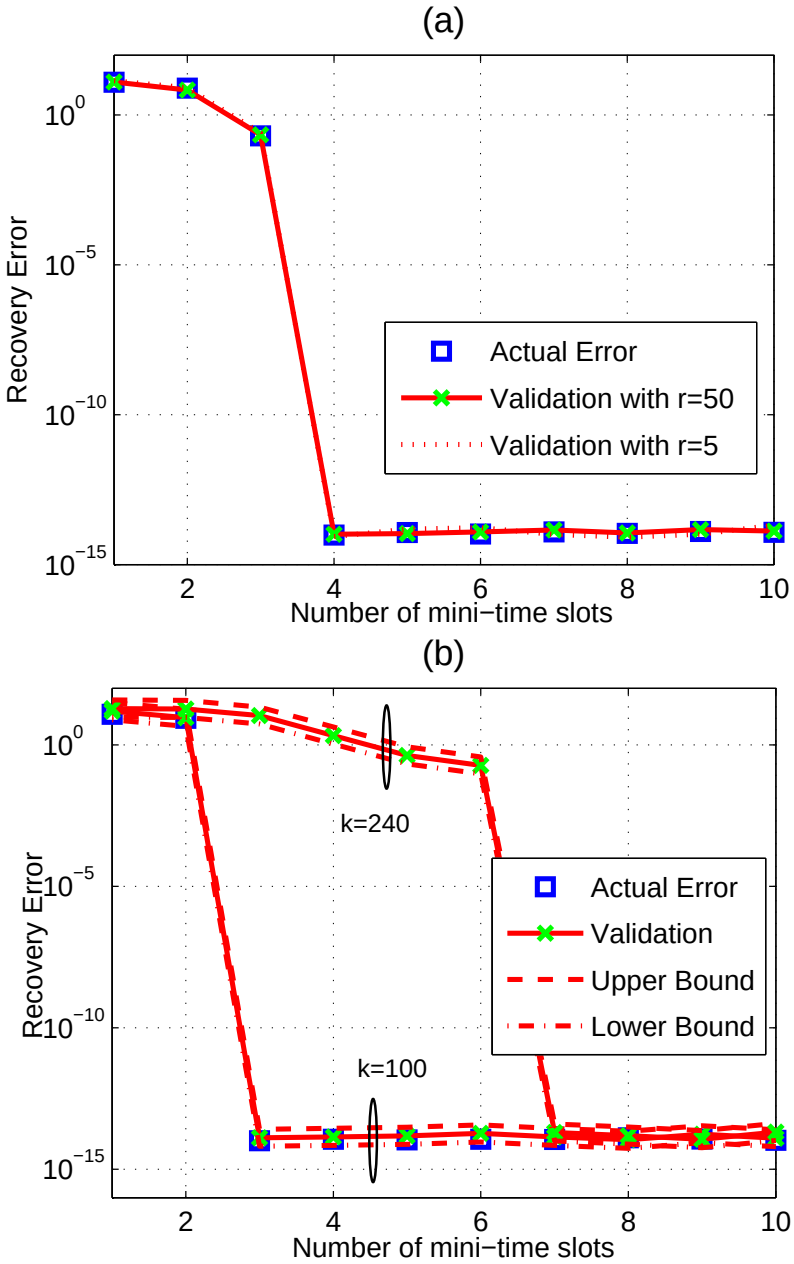

Fig. 13 Comparison of the actual recovery error and the proposed validation parameter when the number of mini time slots increases: (a) different number of measurements for validation when the spectral sparsity level $k = 120$, and (b) different spectral sparsity levels when $r = 50$. It was assumed that there is no measurement noise in the compressive measurements. The upper and lower bounds on the actual recovery error are given in (16).

minated. This approach, to some extent, decreases the probabilities of excessive or insufficient numbers of measurements.

4.4 Noisy compressed adaptive sensing

Due to either the quantization error of ADC or the imperfect design of sub-Nyquist sampler, the measurement noise may exist when performing compressive measurements. In this section, the ℓ_1 norm validation approach is further studied to fit the CASe framework in the noisy case. After that, we present a sparsity-aware recovery algorithm that can correctly terminate greedy iterations when the spectral sparsity level is unknown and the effects of measurement noise are not negligible.

In the noisy signal measurement case, the concatenated training set $\mathbf{Y}_p$ and the testing subset $\mathbf{V}$ can be written as

$$\mathbf{Y}_p = \boldsymbol{\Phi}_p \mathbf{F}^{-1} \mathbf{X}_p + \mathbf{n} \tag{17}$$

and

$$\mathbf{V} = \boldsymbol{\Psi} \mathbf{F}^{-1} \mathbf{X} + \mathbf{n} \tag{18}$$

respectively, where the measurement noise $\mathbf{n}$ is additive noise (added to the real compressed signal after the random projection) generated by the signal measurement procedure, i.e., signal quantization. The measurement noise can be modeled by circular complex additive white Gaussian noise (AWGN). Without loss of generality, we assume that $\mathbf{n}$ has an upper bound $\bar{n}$, and has zero mean and known variance δ^2, i.e., $\mathbf{n} \sim \mathscr{CN}(0, \delta^2)$. For example, if the measurement noise $\mathbf{n}$ is generated by the quantization noise of a uniform quantizer, the noise variance δ^2 can be estimated by $\Delta^2/12$ and $\mathbf{n} \leq \bar{n} = \Delta$, where Δ denotes the cell width.

If ρ_p is close enough to $\sqrt{\frac{\pi}{2}}\delta$, the signal acquisition procedure can be safely terminated. This observation is due to the following theorem:

Theorem 2: Let $\varepsilon > 0$, $\delta > 0$, $\rho \in (0,1)$, $\nu \geq \frac{\sqrt{2/\pi}}{\delta}\bar{n} - 1$, and $r = \ln\left(\frac{2}{\rho}\right)\frac{3(4-\pi)\delta^2 + \sqrt{2\pi}\varepsilon\delta\nu}{3\varepsilon^2}$. If the best spectral approximation exists within the sequence of spectral estimates $\hat{X}_1, \cdots, \hat{X}_P$, then there exists a validation parameter ρ_p that satisfies

$$\Pr\left[\sqrt{\frac{\pi}{2}}\delta - \varepsilon \leq \rho_p \leq \sqrt{\frac{\pi}{2}}\delta + \varepsilon\right] > 1 - \rho, \tag{19}$$

where ρ is given by $\rho = 2\exp\left(-\frac{3r\varepsilon^2}{3(4-\pi)\delta^2 + \sqrt{2\pi}\varepsilon\delta\nu}\right)$.

The proof of Theorem 2 is given in Appendix B.

Remark 3: It is worthwhile to note that Theorem 2 addresses the problem of finding the best spectral approximation, i.e., $\hat{X}_p = \mathbf{X}^\star$, that minimizes $\|\mathbf{X} - \hat{X}_p\|_2$ among all possible spectral estimates in the noisy case. This is different from Theorem 1, which focuses on finding a satisfactory spectral estimate $\hat{X}_p$ that makes $\|\mathbf{X} - \hat{X}_p\|_2 \leq \varpi$ in the noiseless case. Using Theorem 1, we should carefully choose

the tolerable recovery error threshold ϖ in order to avoid excessive or insufficient numbers of measurements. In addition, in Theorem 1, the relation between the tolerable recovery error threshold ϖ and the probability of finding the best spectral approximation is unknown. By contrast, Theorem 2 shows that if there exists a best spectral approximation, the corresponding validation parameter should be within a certain small range with a probability greater than $1-\rho$. Thus, if the result of Theorem 2 is used as the signal acquisition termination metric, the issues of excessive or insufficient numbers of measurements can be solved.

Remark 4: If the best spectral approximation exists, the probability of finding it exponentially increases as the size of testing set (i.e., r) increases. It means that if we monitor ρ_p, we have a higher probability of finding the best spectral approximation when using more measurements for validation. However, we should note that there is a trade off between the size of the training set and the size of the testing set for a fixed sub-Nyquist sampling rate. On the one hand, a smaller r (i.e., larger training set for a fixed m) could result in better spectral recovery, while on the other hand, the probability of finding the best spectral approximation decreases as r becomes small. In addition, for a fixed degree of confidence $1-\rho$, we face a trade off between the accuracy ε and the size of the testing set r, as shown in Theorem 2. At the expense of the accuracy ε (i.e., larger ε), r can be small. We should also emphasize that, as we can see in (34), linear increase of the standard deviation δ will lead to quadratic growth in the size of the testing set. This is the reason why we should carefully consider the effects of measurements noise in the validation approach.

4.5 Sparsity-Aware Recovery Algorithm

As the above discussions indicated, Theorem 2 can be used for identifying the best spectral approximation to $\mathbf{X}$ from the spectral estimate sequence $\hat{X}_1, \hat{X}_2, \cdots, \hat{X}_p$, which is calculated by increasing the number of measurements in the proposed CASe framework. We note that Theorem 2 can also be used for preventing over-fitting or under-fitting in greedy recovery algorithms. Greedy recovery algorithms iteratively generate a sequence of estimates $\hat{X}_p^1, \hat{X}_p^2, \cdots, \hat{X}_p^t$, where the best spectral estimate may exist under certain system parameter choices. For example, the OMP algorithm chooses one column from the measurement matrix at a time for reconstructing $\mathbf{X}$ from $\mathbf{y}$. After $t=k$ iterations, the k-sparse vector $\hat{X}^k$ will be returned as an approximation to $\mathbf{X}$. Note that OMP requires the sparsity level k as an input, and such an input is commonly needed by most greedy recovery algorithms. However, the sparsity level k of the spectrum in the cognitive radio system is often unknown, and therefore traditional greedy compressed sensing algorithms will result in either early or late termination of greedy algorithms. Then the problems of under-fitting and over-fitting arise, leading to inferior spectral recovery performance. In order to reconstruct the full spectrum in the case of unknown k, we propose to use the testing set for validating the spectral estimate sequence $\hat{X}_p^1, \hat{X}_p^2, \cdots, \hat{X}_p^t$, and termi-

Table 3 Sparsity-Aware OMP Algorithm

Input: training set $\mathbf{Y}_p$, testing set $\mathbf{V}$, measurement matrix $\boldsymbol{\Phi}_p$, testing matrix $\boldsymbol{\Psi}$, noise variance δ^2, accuracy ε, $k_{\max}$.
1. *Initialize:* Index set $\Lambda^0 = \emptyset$, residual $\mathbf{R}_p^0 = \mathbf{Y}_p$, and iteration index $t = 0$. Let $\rho_p^t = C_1$ ($\forall t \in [0, k_{\max}]$), where C_1 is a large constant. 2. *While* $\lvert\rho_p^t - \sqrt{\frac{\pi}{2}}\delta\rvert > \varepsilon$ and $t < k_{\max}$, *do* a). Increment t by 1. b). Find the index λ^t that solves the optimization problem: $\lambda^t = \arg\max_{j=1,\cdots,N} \lvert < \mathbf{R}_p^{t-1}, \boldsymbol{\Phi}_p^j > \rvert$. c). Augment the index set $\Lambda^t = \Lambda^{t-1} \cup \{\lambda^t\}$, and revise the matrix $\Theta_p^t = \boldsymbol{\Phi}_p(\Lambda^t)$ by only selecting the column index belongs to Λ^t, other columns are all zeros. d). Solve a least squares problem: $\hat{X}_p^t = \arg\min_{\mathbf{X}} \lVert\mathbf{Y}_p - \Theta_p^t \mathbf{F}^{-1}\mathbf{X}\rVert_2$. e). Calculate the validation parameter using $\mathbf{V}$ and $\boldsymbol{\Psi}$: $\rho_p^t = \frac{\lVert\mathbf{V} - \boldsymbol{\Psi}\mathbf{F}^{-1}\hat{X}_p^t\rVert_1}{r}$. f). Update residual: $\mathbf{R}_p^t = \mathbf{Y}_p - \boldsymbol{\Phi}_p \mathbf{F}^{-1}\hat{X}_p^t$.
Output: $\hat{X}_p = \arg\min_{\hat{X}_p^t} \lvert\rho_p^t - \sqrt{\frac{\pi}{2}}\delta\rvert, \ \forall t \in [1, k_{\max}]$

nate the iterations if the current validation parameter satisfies the conditions given in Theorem 2.

As shown in Table 3, we present a sparsity-aware OMP algorithm. One important advantage of the proposed algorithm is that it does not require the instantaneous spectral sparsity level k, but requires instead its upper bound $k_{\max}$ which can be easily known. In each iteration, the column index $\lambda^t \in [1, N]$ that has the maximum correlation between the residual and the measurement matrix will be found, and be merged with the previously computed spectral support to form a new spectral support Λ^t. After that, the full spectrum is recovered by solving a least squares problem as shown in the step 2-*d*) of Table 3. Note that $\Theta_p^t \triangleq \boldsymbol{\Phi}_p(\Lambda^t)$ is the sub-matrix obtained by only selecting the columns whose indices are within Λ^t in the matrix $\boldsymbol{\Phi}_p$, while other columns are set to all zeros. For a spectral estimate $\hat{X}_p^t$, we validate it by using the validation parameter ρ_p^t, which can be calculated by using the testing set $\mathbf{V}$ and the spectral estimate $\hat{X}_p^t$ as shown in the step 2-*e*) of Table 3. The residual is then updated. We emphasize that the proposed algorithm monitors the validation parameter ρ_p^t, instead of the residual $\|R_p^t\|_2 \leq \varpi$ as used in the traditional greedy recovery algorithms. Based on Theorem 2, if the best spectral estimate is included in the spectral estimate sequence $\hat{X}_p^1, \hat{X}_p^2, \cdots, \hat{X}_p^t$, the probability of finding it will be greater than $1 - 2\exp\left(-\frac{3r\varepsilon^2}{3(4-\pi)\delta^2 + \sqrt{2\pi}\varepsilon\delta\nu}\right)$. In other words, the probability of under-/over-fitting is less than or equal to $2\exp\left(-\frac{3r\varepsilon^2}{3(4-\pi)\delta^2 + \sqrt{2\pi}\varepsilon\delta\nu}\right)$, and becomes smaller as r increases.

For the proposed spectral recovery algorithm, there is a key parameter we need to know, i.e., ε. The following quadratic equation regarding ε holds by using (33):

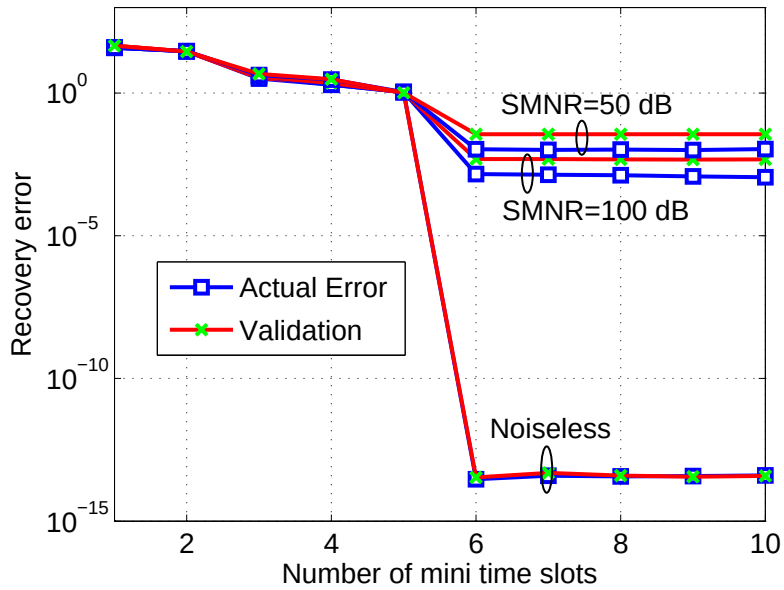

Fig. 14 The effects of measurement noise on both the actual recovery error and the proposed validation parameter when the SMNR varies. The spectral sparsity level was set to $k = 120$.

$$r \cdot \varepsilon^2 - \frac{\sqrt{2\pi}}{3} \ln\left(\frac{2}{\rho}\right) \delta v \cdot \varepsilon - (4-\pi) \ln\left(\frac{2}{\rho}\right) \delta^2 = 0. \tag{20}$$

It can be easily determined that the discriminant of the above quadratic equation is positive, so there are two distinct real roots. The following positive root can be used to determine ε:

$$\varepsilon = \left[\frac{\sqrt{2\pi} \ln\left(\frac{2}{\rho}\right) \delta v \pm \delta \sqrt{2\pi \ln^2\left(\frac{2}{\rho}\right) v^2 + 36(4-\pi) \ln\left(\frac{2}{\rho}\right) r}}{6r} \right]^+ \tag{21}$$

where $[x]^+$ denotes $\max(x, 0)$.

4.6 Numerical Results

In our simulations, we adopt the wideband analog signal model in [27] and let the received signal $x(t)$ at a cognitive radio to be of the form

$$x(t) = \sum_{l=1}^{N_b} \sqrt{E_l B_l} \cdot \mathrm{sinc}\,(B_l(t-\alpha)) \cdot \cos\,(2\pi f_l(t-\alpha)) + z(t) \tag{22}$$

where $\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$, α denotes a random time offset smaller than $T/2$, $z(t)$ is AWGN (i.e., $z(t) \sim \mathcal{N}(0,1)$), and E_l is the received power for the subband l at cognitive radio. The received signal $x(t)$ consists of $N_b = 8$ non-overlapping subbands. The l-th subband is in the frequency range of $[f_l - \frac{B_l}{2}, f_l + \frac{B_l}{2}]$, where the bandwidth $B_l = 10 \sim 30$ MHz and f_l denotes the center frequency. The center frequency of the subband l is randomly located within $[\frac{B_l}{2}, W - \frac{B_l}{2}]$ (i.e., $f_l \in [\frac{B_l}{2}, W - \frac{B_l}{2}]$), where the overall signal bandwidth $W = 2$ GHz. Therefore, the Nyquist rate is $f = 2W = 4$ GHz, and the spectral occupancy (i.e., $\frac{\sum_{l=1}^{8} B_l}{W}$) is a random number between 4% and 12%. We emphasize that the spectral occupancy of 4% $\sim$ 12% in our simulations is very close to the spectral measurements in New York City as noted above. The received signal-to-noise ratios (SNRs) of these 8 active subbands are random natural numbers between 5 dB and 25 dB. The spectrum sensing duration is chosen to be $T = 5\ \mu$s, during which the symbols from primary users and the channels between the primary users and cognitive radios are assumed to be quasi-stationary. We then divide T into $P = 10$ mini time slots, each of which has $\tau = \frac{T}{P} = 0.5\ \mu$s. If the received signal $x(t)$ were sampled at the Nyquist rate, the number of Nyquist samples in each time slot would be $N = 2W\tau = 2,000$. It can be calculated that the spectral sparsity level k is in the range of $4\% \times N = 80 \leq k \leq 12\% \times N = 240$. In the proposed framework, rather than using the Nyquist sampling rate, we adopt the sub-Nyquist sampling rate $f_s = 400$ MHz; thus, the number of measurements in each time slot is $m = f_s\tau = 200$. In other words, the undersampling fraction in each time slot is $m/N = 10\%$. For the purpose of testing/validation, $r = 50$ measurements in the first time slot are reserved, while the remaining measurements are used for reconstructing the spectrum. The measurement matrices, i.e., $\boldsymbol{\Phi}_p$ and $\boldsymbol{\Psi}$, follow the standard normal distribution with zero mean and unit variance. Due to the imperfect design of signal measurement devices, the measurement noise may exist. In the noisy case, the measurement noise is assumed to be circular complex AWGN, i.e., $\mathbf{n} \sim \mathcal{CN}(0, \delta^2)$. As the measurement noise in this chapter is mainly due to the signal quantization in the ADCs, we set the signal-to-measurement-noise ratios (SMNR) to be 50 dB and 100 dB. This is because the SMNR of the uniform quantization increases 6 dB for each one-bit; thus, the SMNR of 8-bit quantization is 48 dB and the SMNR of 16-bit quantization is 96 dB, which are approximately 50 dB and 100 dB.

Firstly, we consider the effects of measurement noise to both the spectral recovery quality and the validation parameter. In Fig. 14, the spectral sparsity level is set to $k = 120$. We can see that, in either the noiseless measurement case or the noisy measurement case, the proposed CASe framework can reconstruct the spectrum using 6 time slots. The spectral recovery quality becomes worse when the measurement noise level increases. In the noiseless case, the proposed validation parameter can closely fit the actual recovery error. By contrast, there is a gap between the actual recovery error and the validation result when the measurement noise exists. This is

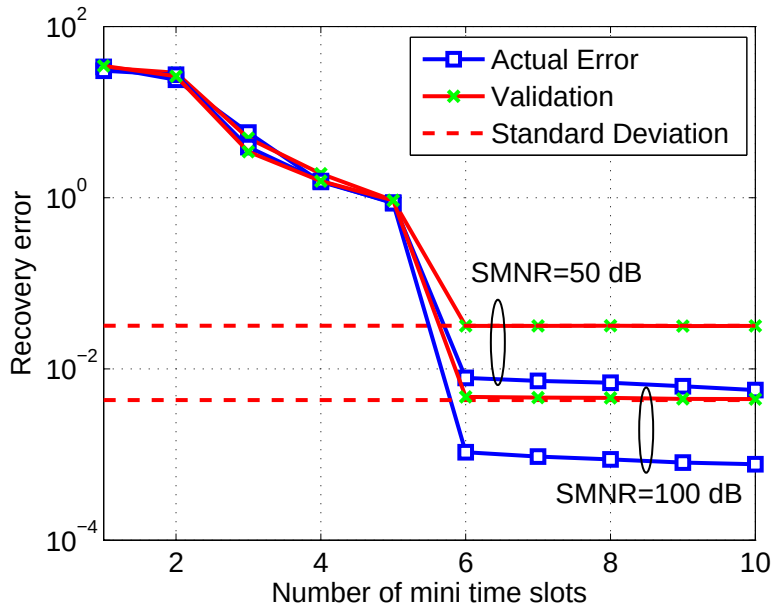

Fig. 15 Comparison of the validation parameter and the actual recovery error when the best spectral approximation occurs. The dash line denotes the predicted validation value, i.e., $\sqrt{\frac{\pi}{2}}\delta$ (scaled standard deviation), as used in Theorem 2.

because, on the one hand, the actual recovery error $\|\mathbf{X}-\hat{X}_p\|_2$ can be very small, e.g., 10^{-14} in the case of best spectral approximation, on the other hand, the validation parameter is mainly determined by the noise level as shown in Theorem 2. This implies that the effects of measurement noise should be carefully considered even if $\hat{X}_p$ is the best spectral approximation. In Fig. 15, it is seen that when the best spectral approximation occurs (i.e., the actual recovery error is small enough), the validation parameter is very close to the scaled noise standard deviation, i.e., $\sqrt{\frac{\pi}{2}}\delta$. This observation validates the results of Theorem 2. If the validation method is used for designing the termination metric of the signal acquisition, such as in the algorithm given in Table 2, the problems of insufficient or excessive numbers of measurements can be solved.

Secondly, Fig. 16 analyzes the spectral recovery performance when using different compressed sensing approaches. In these simulations, in order to find the best spectral approximation with high confidence, the accuracy parameter ε in (19) is set to $\delta/2$ and the number of testing measurements is $r = 50$. As depicted in Fig. 16(a), the proposed CASe framework can adaptively adapt its number of measurements to the unknown spectral sparsity level k. The corresponding spectral recovery performance is shown in Fig. 16(b), where the spectral recovery mean square error (MSE) of different compressed sensing approaches is given. We can see that, even with the total number of measurements $M = 1300$, the performance of the traditional com-

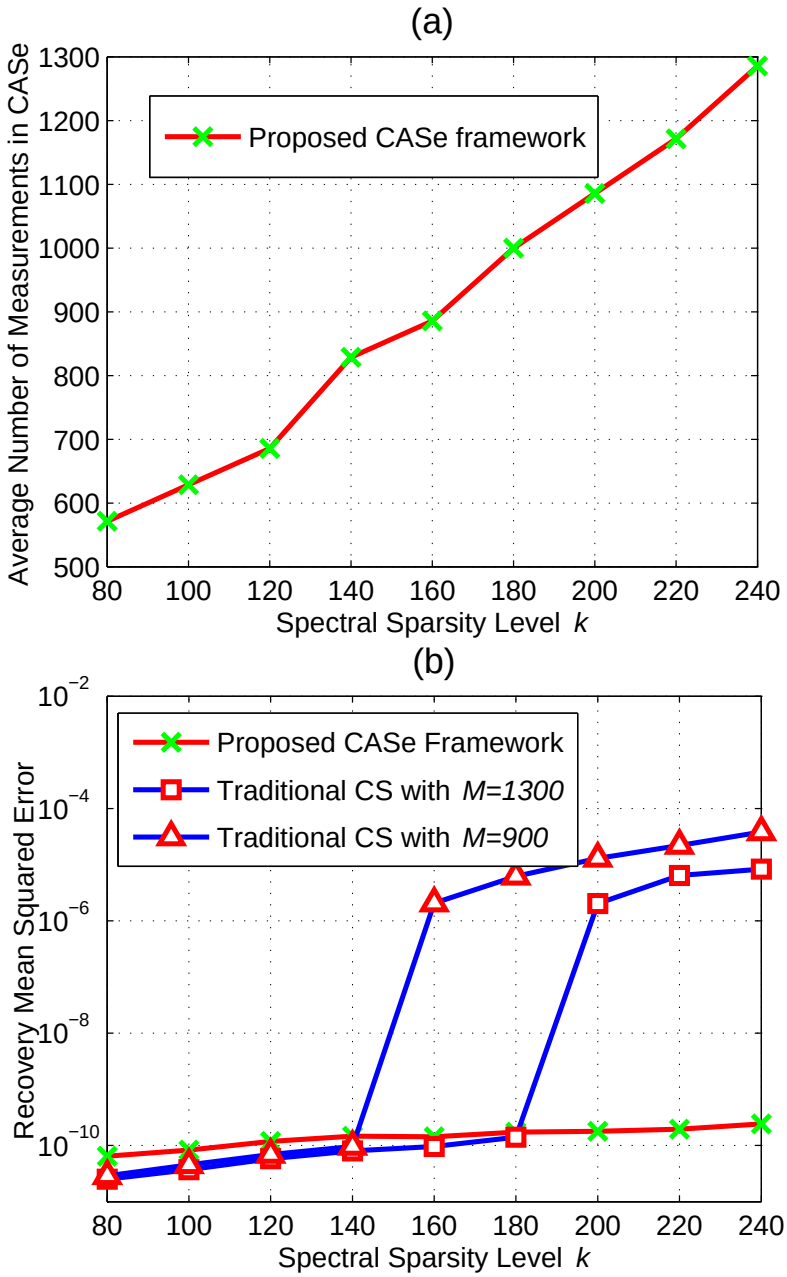

Fig. 16 Performance analysis of spectral recovery when using different compressed sensing approaches: (a) the average number of measurements required by CASe, and (b) the spectral recovery mean square error. The SMNR was set to 100 dB.

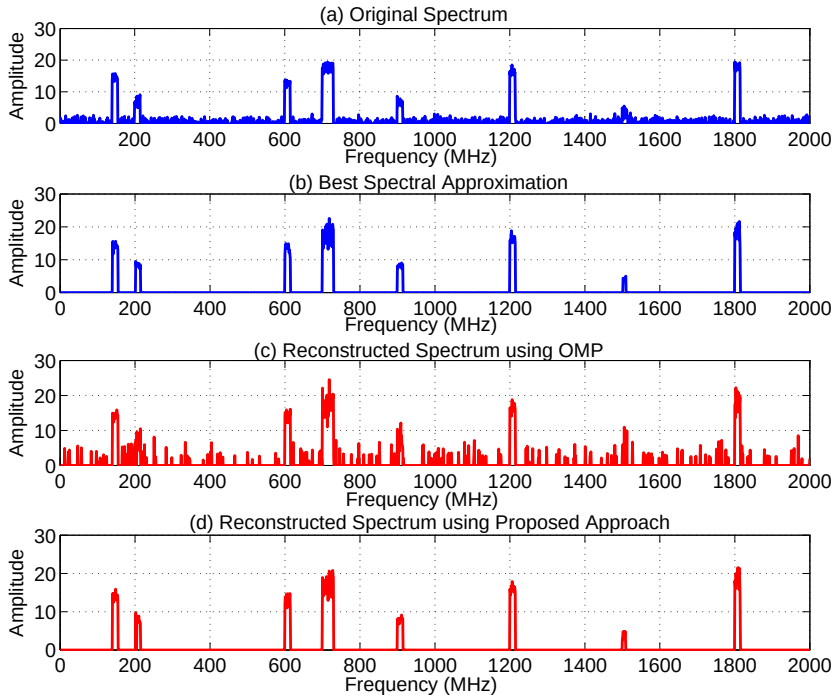

Fig. 17 Examples of the reconstructed spectrum when using different recovery algorithms. The spectral sparsity level was assumed to be $k = 150$, with the total number of measurements $M = 800$. The received SNRs of these 8 active subbands were set to random natural numbers between 5 dB and 25 dB. The SMNR was set to 50 dB.

pressed sensing system is inferior to that of the proposed CASe framework as the traditional compressed sensing system cannot deal with the case of $k \geq 200$. Note that, if we assume that the spectral sparsity level k has a uniform distribution between 80 and 240, the average number of measurements required by CASe is 900. Compared to the traditional compressed sensing system with $M = 900$, it is obvious that the CASe framework has much lower MSE for most of $k \in [80, 240]$.

Thirdly, Fig. 17 shows examples of the original spectrum and the reconstructed spectrum when using different spectral recovery algorithms, i.e., OMP and the proposed algorithm. We can see that the recovery performance of the proposed algorithm is superior to that of the traditional OMP algorithm. As the sparsity level is unknown and has the range of $80 \leq k \leq 240$, if the OMP algorithm is used, the problems of either under-fitting (i.e., iteration is terminated earlier as k is underestimated) or over-fitting exist. As the problem of under-fitting could lead to the missed detection of primary users which may cause harmful interference to primary users, the traditional OMP algorithm should prevent the under-fitting from occurring, and tends to choose more number of iterations. In the case of over-fitting, the traditional OMP algorithm will result in a “noisy” reconstructed spectrum as de-

picted in Fig. 17(c). With the aid of the testing set, the proposed approach has an improved recovery performance as shown in Fig. 17(d). Compared with the OMP algorithm, the proposed algorithm provides better spectral estimate, and is much more similar to the best spectral approximation in Fig. 17(b). It is worthwhile to emphasize that the proposed algorithm will have more noticeable improvement over the OMP algorithm when there is larger uncertainty in the spectral sparsity level k.

Finally, Fig. 18 further explores the performance of different recovery algorithms. In order to illustrate the performance of CASe when using different recovery algorithms, the MSE of the reconstructed spectrum is given in Fig. 18(a). It can be seen that the gain of using the proposed algorithm over OMP is approximately one order of magnitude in MSE. This is because the proposed algorithm can terminate the iteration at the right iteration index; by contrast, when using OMP, the problems of either under-fitting or over-fitting exist, leading to either incomplete spectral recovery or noisy spectral recovery. As a consequence, we can see from Fig. 18(b) that, for a fixed SMNR=50 dB, the proposed algorithm has much lower recovered error rate than the OMP algorithm. We note that the recovered error rate is defined as the probability of simulated mean MSE larger than the target MSE.

4.7 Discussions and Conclusions

4.7.1 Discussions

The CASe framework shares its goals with some recent efforts that have looked at testing the actual error directly from compressed data. The ℓ_2 norm cross validation approach for compressed sensing has been studied by Ward [34], and Boufounos *et al.* [35]. These results are very remarkable as they allow us to verify the actual decoding error with almost no effort (i.e., a very few measurements are reserved for testing). We note that the results here are different from those in these papers. In particular, we have studied a different validation approach, i.e., the ℓ_1 norm is used for validating the recovery result, rather than the ℓ_2 norm. In addition, the effects of measurement noise were carefully considered in our analysis. By contrast, Ward's validation approach did not model the effects of measurement noise. When the proposed ℓ_1 norm validation approach is used in compressed sensing technologies, it could be a useful complement to the work in [34, 35]. It should also be emphasized that, compared to the ℓ_2 norm validation approach, the proposed ℓ_1 norm validation approach is less sensitive to outliers. As shown in Fig. 19(a), when outliers exist in the testing set, the validation parameter of using the ℓ_1 norm is one order in magnitude lower than that of using the ℓ_2 norm. Moreover, we note that using compressed sensing technologies for wideband spectrum sensing in a cognitive radio system, we cannot avoid outliers. This is because the ADC is not a noise-free device, and the non-linearity of ADC could be a source of generating outliers. Furthermore, in a real-time compressed sensing device such as the random demodulator

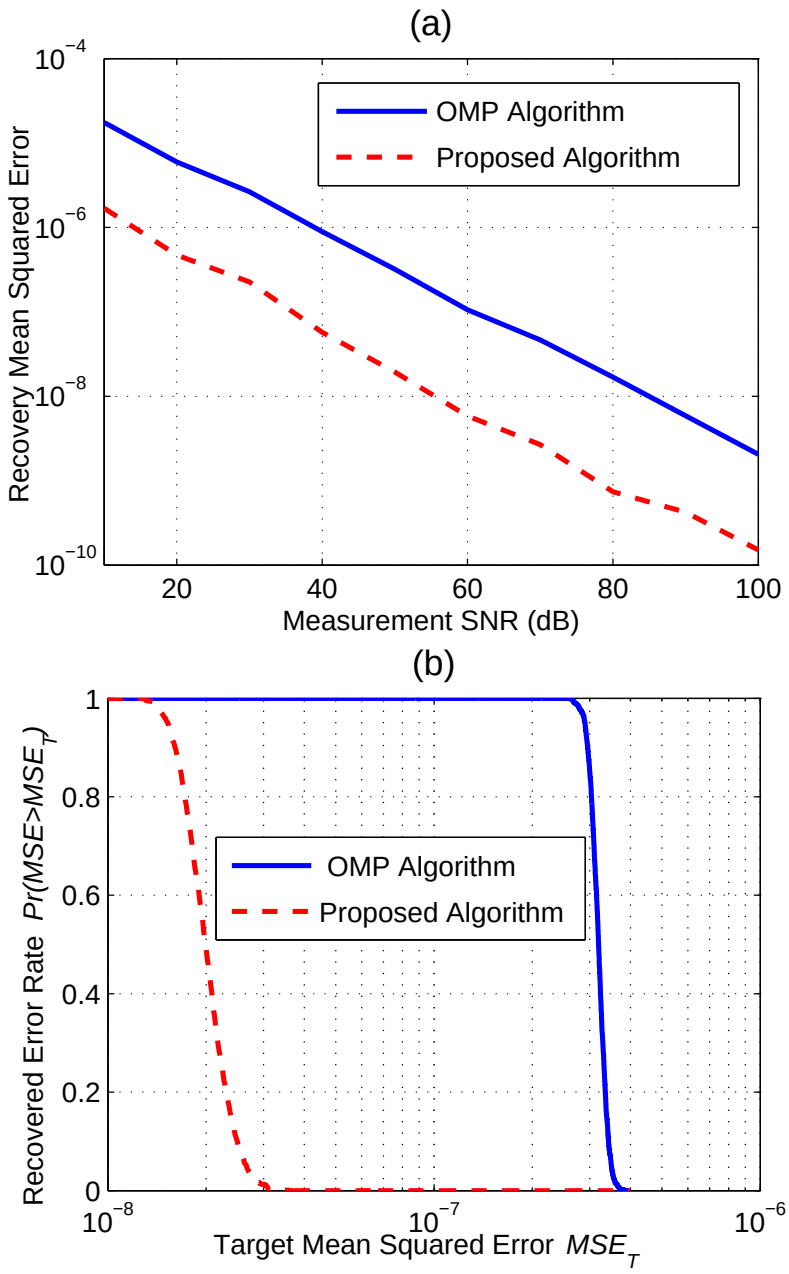

Fig. 18 Performance comparison of different recovery algorithms: (a) the spectral recovery mean square error when the SMNR increases, and (b) the recovered error rate $\Pr(MSE > MSE_T)$ when SMNR=50 dB. The spectral sparsity level was assumed to be $k = 120$, with the average number of measurements $M = 800$.

in [26, 29, 30], imperfect synchronization of the pseudo-random sequence generator and the low-rate ADC could result in outliers.

A natural technique for choosing the stopping time of the measurement would be sequential detection [36], in which we collect one sample at a time until we have enough observations to generate a final decision. However, we note that, in the compressed sensing-based spectrum sensing system, the sequential measurements cannot be directly used for performing sequential test. This is because, due to the sub-Nyquist sampling, there exists spectral aliasing phenomenon, which makes frequencies become indistinguishable. Thus, in order to apply sequential detection, the wideband spectrum should be reconstructed before each sequential test for avoiding spectral aliasing. In such a scenario, sequential detection could lead to high computational costs. Malioutov *et al.* [37] have studied a typical compressed sensing-based sequential measurement system, where the decoder can receive compressed samples sequentially. It has been shown that such a system can successfully estimate the current decoding error by using some additional samples. Nevertheless, it is not proper to apply the compressed sensing-based sequential measurement setup in cognitive radio systems. Because, in this scheme, the wideband spectrum should be repeatedly reconstructed for each additional measurement that could lead to high computational costs and large spectrum sensing overhead in cognitive radios. For example, using the CoSaMP algorithm [28], the running time in each reconstruction is $\mathscr{O}(\beta N)$, where β denotes the current number of measurements. Thus, the total running time for the sequential measurement setup is $\mathscr{O}(\frac{M(M+1)N}{2})$, where M denotes the number of measurements till the termination of measurement. By contrast, in our proposed system, the spectrum sensing time slot is divided into P equal-length mini-time slots, and the wideband spectrum is reconstructed after each mini-time slot. The total running time of the proposed system is therefore $\mathscr{O}(\frac{M(P+1)N}{2})$, where $P \ll M$. Fig. 19(b) shows that the spectrum sensing overhead (due to the spectral reconstruction) of the sequential compressed sensing system is several times higher than that of the proposed system. Furthermore, another advantage of the proposed system is that, by changing the length of mini-time slot (thus the value of P because $P = \frac{M}{m}$), we can control the trade-off between the cost of computation and the cost of acquiring additional measurements.

4.7.2 Conclusions

We have presented a novel framework, i.e., CASe, for wideband spectrum sensing in cognitive radio systems. It has been shown that CASe can considerably improve the spectral recovery performance when the sparsity level of the spectrum is unknown, thanks to the ℓ_1 norm validation approach. We have shown that the proposed validation parameter can be a very good proxy for the actual spectral recovery error in the noiseless measurement case even if the testing set is small. The proper use of the validation approach could solve the problems of excessive or insufficient numbers of measurements, thereby improving not only the energy-efficiency of cognitive radio, but also the throughput of cognitive radio networks. In addition, we have shown

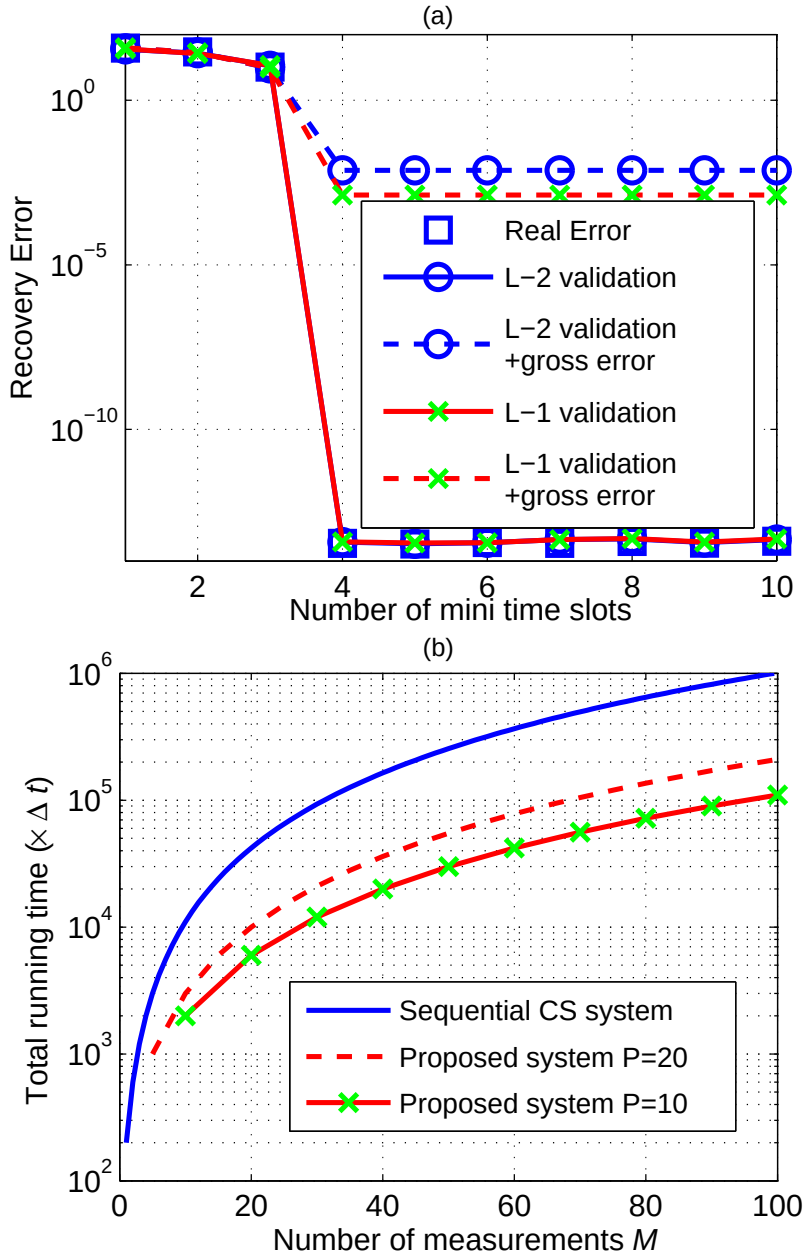

Fig. 19 Comparison between the proposed system and the existing systems: (a) Sensitivity test of both the ℓ_1 norm validation and the ℓ_2 norm validation approaches against outliers. In simulations, the measurement error was added to a single sample of the testing set, and the magnitude of the measurement error was set to 100 dB lower than that of the sample. (b) Total running time of reconstructing the spectrum for both the sequential compressed sensing measurement setup and the proposed system when using the CoSaMP algorithm. In simulations, $N = 200$, and $M = Pm = 100$ where m denotes the number of measurements in each mini-time slot and P is the number of mini-time slots.

that, in the case of noisy compressive measurements, if the best spectral approximation exists, then the corresponding validation parameter has a very large probability of being within a certain small range. Based on this property, we have proposed a sparsity-aware recovery algorithm for reconstructing the wideband spectrum without the knowledge of the spectral sparsity level. In the proposed algorithm, if the best spectral approximation exists, then the correct iteration termination index can be found with high probability; therefore, the issues of under-/over-fitting are addressed.

Simulation results have shown that the proposed framework can correctly terminate the signal acquisition that saves both spectrum sensing time slots and signal acquisition energy, while providing better spectral recovery performance than traditional compressed sensing approaches. Compared with the existing greedy recovery algorithm, the proposed sparsity-aware algorithm can achieve lower MSE for reconstructing the spectrum and better spectrum sensing performance. As the RF spectrum is the lifeblood of wireless communication systems and the wideband techniques could potentially offer greater capacity, we expect that the proposed framework has a broad range of applications, e.g., broadband spectral analyzers, signals-intelligence receivers, and ultra wideband radars. Moreover, the proposed ℓ_1 norm validation approach can be used in other compressed sensing applications, e.g., a compressed sensing based communication system where we need to terminate the decoding algorithm with high confidence and small predictable decoding error.

Acknowledgements H. Sun and A. Nallanathan acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) with Grant No. EP/I000054/1.

Appendix

Proof of Theorem 1

Using a variant of the Johnson-Lindenstrauss lemma as shown in Theorem 5.1 of [38], we have

$$\Pr\left[(1-\varepsilon)\|\mathbf{x}\|_2 \leq \frac{\|\boldsymbol{\Psi}\mathbf{x}\|_1}{\sqrt{2/\pi}\, r} \leq (1+\varepsilon)\|\mathbf{x}\|_2\right] \geq 1-\xi. \tag{23}$$

Defining $\mathbf{x} \triangleq \mathbf{F}^{-1}(\mathbf{X}-\hat{X}_p)$ in (23), we obtain

$$\Pr\left[(1-\varepsilon)\|\mathbf{F}^{-1}(\mathbf{X}-\hat{X}_p)\|_2 \leq \frac{\|\boldsymbol{\Psi}\mathbf{F}^{-1}(\mathbf{X}-\hat{X}_p)\|_1}{\sqrt{2/\pi}\, r} \leq (1+\varepsilon)\|\mathbf{F}^{-1}(\mathbf{X}-\hat{X}_p)\|_2\right] \geq 1-\xi. \tag{24}$$

The above inequality can be rewritten by using (12) and (14)

$$\Pr\left[(1-\varepsilon)\|\mathbf{F}^{-1}(\mathbf{X}-\hat{X}_p)\|_2 \le \sqrt{\frac{\pi}{2}}\rho_p \le (1+\varepsilon)\|\mathbf{F}^{-1}(\mathbf{X}-\hat{X}_p)\|_2\right] \ge 1-\xi. \tag{25}$$

Applying Parseval's relation to (25), we have

$$\Pr\left[(1-\varepsilon)\|\mathbf{X}-\hat{X}_p\|_2 \le \sqrt{\frac{\pi N}{2}}\rho_p \le (1+\varepsilon)\|\mathbf{X}-\hat{X}_p\|_2\right] \ge 1-\xi. \tag{26}$$

Thus, Theorem 1 follows.

Proof of Theorem 2

The best spectral approximation $\mathbf{X}^\star$ means that $\|\mathbf{X}^\star-\mathbf{X}\|_2$ is sufficiently small. Without loss of generality, we approximate $\mathbf{X}^\star$ by $\mathbf{X}$. Thus, if $\hat{X}_p$ is the best spectral approximation, the validation parameter can be rewritten by using (18)

$$\rho_p = \frac{\|\mathbf{V}-\mathbf{\Psi}\mathbf{F}^{-1}\hat{X}_p\|_1}{r} = \frac{\|\mathbf{n}\|_1}{r} = \frac{\sum_{i=1}^r |n^i|}{r}. \tag{27}$$

As the measurement noise $n^i \sim \mathscr{CN}(0,\delta^2)$, its absolute value $|n^i|$ follows the Rayleigh distribution with mean $\sqrt{\frac{\pi}{2}}\delta$ and variance $\frac{4-\pi}{2}\delta^2$. Using the cumulative distribution function of the Rayleigh distribution, we have $\Pr(|n^i| \le x) = 1-\exp(-\frac{x^2}{2\delta^2})$. Further, as the measurement noise level has an upper-bound $\bar{n}$ in practice, there exists a sufficiently large parameter ν that makes $|n^i| \le \bar{n} \le (\nu+1)\sqrt{\frac{\pi}{2}}\delta$ almost surely. If we define a new variable $D_i = |n^i| - \sqrt{\frac{\pi}{2}}\delta$, we obtain $\mathbb{E}[D_i]=0$, $\mathbb{E}[D_i^2] = \frac{4-\pi}{2}\delta^2$, and $|D_i| \le \sqrt{\frac{\pi}{2}}\delta\nu$. Based on the Bernstein's inequality [39], the following inequality holds

$$\Pr\left[\left|\sum_{i=1}^r D_i\right| > \varepsilon\right] = \Pr\left[\left|\sum_{i=1}^r |n^i| - r\sqrt{\frac{\pi}{2}}\delta\right| > \varepsilon\right] \tag{28}$$

$$\le 2\exp\left(-\frac{\varepsilon^2/2}{\sum_{i=1}^r \mathbb{E}[D_i^2] + \overline{D}\varepsilon/3}\right) \tag{29}$$

$$\le 2\exp\left(-\frac{3\varepsilon^2}{3(4-\pi)r\delta^2 + \sqrt{2\pi}\varepsilon\delta\nu}\right) \tag{30}$$

where $\overline{D} = \sqrt{\frac{\pi}{2}}\delta\nu$ denotes the upper-bound on $|D_i|$.

Simply replacing ε by $r\varepsilon$ in (30) while using (27), we can rewrite (30) as

$$\Pr\left[\left|\rho_p - \sqrt{\frac{\pi}{2}}\delta\right| > \varepsilon\right] \le 2\exp\left(-\frac{3r\varepsilon^2}{3(4-\pi)\delta^2 + \sqrt{2\pi}\varepsilon\delta\nu}\right). \tag{31}$$

Using (31), we end up with

$$\Pr\left[\left|\rho_p - \sqrt{\frac{\pi}{2}}\delta\right| \le \varepsilon\right] > 1 - 2\exp\left(-\frac{3r\varepsilon^2}{3(4-\pi)\delta^2 + \sqrt{2\pi}\varepsilon\delta\nu}\right). \tag{32}$$

To derive the required r, we set the lower probability bound in (32) as

$$1 - 2\exp\left(-\frac{3r\varepsilon^2}{3(4-\pi)\delta^2 + \sqrt{2\pi}\varepsilon\delta\nu}\right) = 1 - \rho. \tag{33}$$

Solving the above equation, we obtain

$$r = \ln\left(\frac{2}{\rho}\right)\frac{3(4-\pi)\delta^2 + \sqrt{2\pi}\varepsilon\delta\nu}{3\varepsilon^2}. \tag{34}$$

This completes the proof of Theorem 2.

References

1. M. A. McHenry, "NSF spectrum occupancy measurements project summary," Shared Spectrum Company, Tech. Rep., Aug. 2005.
2. S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201 – 220, Feb. 2005.
3. I. Akyildiz, W.-Y. Lee, M. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 40 –48, Apr. 2008.
4. J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Dept. of Teleinformatics, Royal Institute of Technology Stockholm, Sweden, 8 May, 2000.
5. I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, 2006.
6. H. Ekram and B. V. K, *Cognitive Wireless Communications Networks*, B. V. K, Ed. Springer Publication, 2007.
7. D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, vol. 1, 2004, pp. 772–776.
8. H. Sun, D. Laurenson, and C.-X. Wang, "Computationally tractable model of energy detection performance over slow fading channels," *IEEE Communications Letters*, vol. 14, no. 10, pp. 924–926, Oct. 2010.
9. E. Hossain, D. Niyato, and Z. Han, *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Cambridge University Press, July 2009.
10. Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Proc. IEEE Cognitive Radio Oriented Wireless Networks and Communications*, Mykonos Island, Greece, June 2006, pp. 1–5.
11. J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2001.
12. T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116 –130, 2009.
13. Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Hawaii, USA, April 2007, pp. 1357–1360.

14. M.-H. Yoon, Y. Shin, H.-K. Ryu, and J.-M. Woo, "Ultra-wideband loop antenna," *Electronics Letters*, vol. 46, no. 18, pp. 1249–1251, Sept. 2010.
15. Z.-C. Hao and J.-S. Hong, "Highly selective ultra wideband bandpass filters with quasi-elliptic function response," *IET Microwaves, Antennas Propagation*, vol. 5, no. 9, pp. 1103–1108, 2011.
16. [Online]. Available: www.national.com/pf/DC/ADC12D1800.html
17. H. Sun, W.-Y. Chiu, J. Jiang, A. Nallanathan, and H. V. Poor, "Wideband spectrum sensing with sub-Nyquist sampling in cognitive radios," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 6068–6073, nov. 2012.
18. R. Venkataramani and Y. Bresler, "Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals," *IEEE Trans. Information Theory*, vol. 46, no. 6, pp. 2173 –2183, Sep. 2000.
19. T. Tao, "An uncertainty principle for cyclic groups of prime order," *Math. Res. Lett*, vol. 12, pp. 121–127, 2005.
20. M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 993–1009, March 2009.
21. A. Feldster, Y. Shapira, M. Horowitz, A. Rosenthal, S. Zach, and L. Singer, "Optical undersampling and reconstruction of several bandwidth-limited signals," *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 1027 –1033, April 2009.
22. A. Rosenthal, A. Linden, and M. Horowitz, "Multi-rate asynchronous sampling of sparse multi-band signals," 2008, arXiv.org:0807.1222.
23. M. Fleyer, A. Rosenthal, A. Linden, and M. Horowitz, "Multirate synchronous sampling of sparse multiband signals," 2008, arXiv.org:0806.0579.
24. S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001. [Online]. Available: http://www.jstor.org/stable/3649687
25. Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Taipei, April 2009, pp. 2337–2340.
26. J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Information Theory*, vol. 56, no. 1, pp. 520–544, Jan. 2010.
27. M. Mishali and Y. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375 –391, april 2010.
28. D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301 – 321, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/B6WB3-4T1Y404-1/2/a3a764ae1efc1bd0569dcde301f0c6f1
29. J. Laska, S. Kirolos, Y. Massoud, R. Baraniuk, A. Gilbert, M. Iwen, and M. Strauss, "Random sampling for analog-to-information conversion of wideband signals," in *Proc. IEEE DCAS*, Oct. 2006, pp. 119–122.
30. 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 International Symposium on Circuits and Systems ISCAS 2007*, 27–30 May 2007, pp. 1959–1962.
31. E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Tr. IT*, vol. 51(12), pp. 4203–4215, 2005.
32. J. Haupt, W. U. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," *IEEE SP Mag.*, vol. 25(2), pp. 92–101, 2008.
33. Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1128–1140, Mar. 2009.
34. R. Ward, "Compressed sensing with cross validation," *IEEE Trans. on Information Theory*, vol. 55, no. 12, pp. 5773–5782, Dec. 2009.

35. P. Boufounos, M. Duarte, and R. Baraniuk, "Sparse signal reconstruction from noisy compressive measurements using cross validation," in *Proc. IEEE/SP 14th Workshop on Statistical Signal Processing*, Madison, WI, USA, Aug. 2007, pp. 299–303.
36. S. Chaudhari, V. Koivunen, and H. V. Poor, "Autocorrelation-based decentralized sequential detection of OFDM signals in cognitive radios," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2690–2700, July 2009.
37. D. Malioutov, S. Sanghavi, and A. Willsky, "Sequential compressed sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 435–444, April 2010.
38. J. Matousek, "On variants of the Johnson-Lindenstrauss lemma," *Random Structures and Algorithms*, vol. 33, pp. 142–156, 2008.
39. M. Hazewinkel, Ed., *Encyclopaedia of Mathematics*. New York: Springer, Nov. 1987, vol. 1.

View publication stats