Blind Sub-Nyquist Spectrum Sensing With Modulated Wideband Converter

Peihan Qi, Zan Li, *Senior Member, IEEE*, Hongbin Li, *Senior Member, IEEE*, and Tianyi Xiong, *Student Member, IEEE*

***Abstract*—Rooted in the compressed sensing theory, sub-Nyquist spectrum sensing (SNSS) has been considered as a promising approach to dealing with difficulties and limitations of conventional wideband spectrum sensing in cognitive radio (CR) networks. Most existing SNSS methods require some prior knowledge of the monitored frequency bands, such as the spectrum occupancy/sparsity level and/or the noise power, to determine a termination condition used by an underlying iterative signal recovery process. However, such prior knowledge may be difficult to acquire in practical CR scenarios. To address this problem, we propose a blind SNSS algorithm, referred to as the residual energy ratio based detector (RERD), which bypasses the need for the above-mentioned prior knowledge and performs spectrum sensing in a more autonomous way. The RERD algorithm, which is based on the modulated wideband converter (MWC) sub-Nyquist sampling framework, employs energy ratios of adjacent channels of the MWC as test statistics. We derive closed-form expressions of the decision threshold and the false alarm probability following the Neyman–Pearson criterion. Simulation results show that, without requiring the aforementioned prior knowledge, the RERD algorithm can accurately determine the support of a multiband signal contaminated by background noise in a wide range of signal-to-noise ratio. Moreover, the RERD algorithm is shown to be robust to a range of sparsity orders and different number of sampling channels.**

***Index Terms*—Blind spectrum sensing, cognitive radio, multiband signal, modulated wideband converter, sub-Nyquist sampling.**

I. Introduction

NOWADAYS, cognitive radio (CR) [1], [2] has been widely investigated as a promising solution to deal with the increasing demand for wireless access and the scarcity of available spectrum. Spectrum sensing [3]–[5] which aims to distinguish the presence and absence of primary users (PUs) in certain frequency bands, is one of the fundamental techniques to make CR possible. With the explosive growth of wireless devices and services that require higher data rate in the past few years, wideband spectrum sensing (WSS) has been of interest to seek vacant frequency bands for opportunistic spectrum access [6]–[8]. Conventional WSS methods use signal samples at or above the Nyquist rate to detect and locate spectrum vacancies across a large bandwdith. However, the implementation of such conventional WSS techniques for CR users has encountered some main challenges, namely, the high cost and high power consumption of high-rate analog-to-digital converter (ADC) in the front end, limited memory size, and speed constraint of digital signal processing, etc. [9]. To address these challenges, the recently introduced compressed sensing theory [10], [11], which allows a signal to be sampled at a sub-Nyquist rate and recovered from a significantly reduced number of non-adaptive, linear measurements, can be leveraged for WSS implementation. This has motivated interest in sub-Nyquist spectrum sensing (SNSS) to overcome the the limitations of conventional WSS techniques [12], [13].

Built on several sub-Nyquist sampling frameworks, e.g., the analog-to-information converter (AIC) [14], [15], the multicosets (MC) sampling [16], [17], the multi-rate asynchronous sub-Nyquist sampling (MASS) [18], and the modulated wideband converter (MWC) [19], [20], a number of SNSS methods have been introduced [21]–[33]. Specifically, in [21]–[23], the input signal to AIC is mixed with a high-speed pseudorandom chip sequence, and compressed samples are obtained via integrating and sampling the mixed output at sub-Nyquist sampling rate. AIC-based SNSS methods are usually implemented in two stages. In the first stage, the power spectrum of the original signal is reconstructed from compressed samples using an l_1-norm minimization method. In the second stage, wideband spectrum sensing is performed to determine the locations of the occupied frequency bands. However, l_1-norm minimization methods, e.g., the LASSO algorithm and the GPSR algorithm [24], [25], need to accurately set a termination condition of the underlying iteration process, which is determined from knowledge of the noise power and the sparsity level of the PU signals [23]. In [26], the SNSS algorithm uses the local database to compute the maximum allowable equivalent isotropic radiated power and other revelent prior knowledge, which reduces the required sampling rate and the complexity of signal recovery. This algorithm is applicable in the case where the spectrum of the PU signal changes slowly. If the spectrum is highly dynamic, large computational cost is required to update and maintain the local database. Moreover, it has been found that the AIC framework can be affected by some input signal model mismatch problems [6].

Manuscript received July 28, 2017; revised November 9, 2017 and December 23, 2017; accepted January 7, 2018. Date of publication January 17, 2018; date of current version May 14, 2018. This work was supported in part by the National Natural Science Foundation of China under Grants 61501356 and 61631015, and in part by the Fundamental Research Funds of the Ministry of Education, China, under Grant JB160101. The work of H. Li was supported by the National Science Foundation under Grants ECCS-1408182 and ECCS-1609393. The review of this paper was coordinated by Dr. X. Wang. *(Corresponding author: Zan Li.)*

P. Qi, Z. Li, and T. Xiong are with the State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China (e-mail: phqi@xidian.edu.cn; zanli@xidian.edu.cn; xiongtianyi1989@163.com).

H. Li is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: hli@stevens.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2018.2794779

In [27]–[28], compressed samples for wideband spectrum sensing are acquired by the MC sampling framework, which uses parallel ADC branches and uniformly discretizes the signal at a decimated rate in a time interleaving fashion. Although idea behind the MC framework is intuitive, its application is limited by two issues: (1) the difficulty to synchronize the interleaved ADCs; (2) the signal distortions caused by insufficient bandwidth of commercial ADCs [19]. To relax the strict requirement for time synchronization in the MC sampling framework, the MASS framework which contains multiple ADC branches with different under-sampling rate is designed in [18]. Based on the MASS framework, the SNSS algorithm proposed in [29] is able to determine the presence/absence of the multiple PU signals in a wideband spectrum by only analyzing the statistical properties of the energy of the sub-Nyquist DFT spectrum. The wideband spectrum sensing can be performed without the requirement of spectral recovery in use of this algorithm. However, the decision threshold for the detection process relies on the noise power as a prior knowledge. Meanwhile, the MWC sampling framework [30], which regards the input signal as a sparse union of shift-invariant subspaces, pre-processes the analog signal in multiple channels simultaneously. In each channel, the input signal is multiplied by a periodic random binary sequence, filtered by a lowpass filter, and then sampled by a low-rate ADC. MWC is able to deal with different signal types and can be implemented with off-the-shelf ADCs. With such advantages, MWC-based SNSS methods have drawn increasingly more attention [31]–[33]. In MWC-based SNSS methods, the goal is to find the locations (support set) of the occupied frequency bands by the PU signals, which can be achieved by frame construction and solving a multiple measurement vector (MMV) problem [27], [34] without fully recovering the original signal. The MMV problem can be solved by using either convex relaxation methods or the matching pursuit with a residual-based termination criterion [35]-[38]. In addition, the MMV problem can also be solved by the orthogonal matching pursuit (OMP) with a preset number of iterations, which is estimated either in advance or estimated on the spot [39]–[41].

As mentioned above, most existing SNSS methods need the prior knowledge of the background noise power and the sparsity of primary users to ensure the performance of support recovery in noisy environment [42]. However, in practical CR scenarios with PUs continuously entering and leaving the radio environment, such information can hardly be obtained. Hence, we propose a blind sub-Nyquist spectrum sensing algorithm, referred to as the Residual Energy Ratio based Detector (RERD), which aims to determine the support set of the signal without requiring the aforementioned prior knowledge. The proposed RERD algorithm, in each iteration, employs the ratio of the sample energies of adjacent channels of the MWC as the test statistic to determine whether further iteration of signal recovery is needed. Compared with existing SNSS methods, the contributions of the RERD algorithm are as follows:

1) Based on the MWC compressed sampling framework, the SNSS is performed by solving a series of binary hypothesis detection problems. The RERD algorithm is able to control the iterative process of support recovery and find the support set of the PU signals in a more autonomous way relative to the aforementioned SNSS methods.
2) Energy ratios between adjacent channels are employed as the test statistics, whose cumulative distribution function is derived. Closed-form expressions of the false alarm probability and decision threshold are provided.
3) Theoretical and numerical results show that, the conventional SNSS algorithm, when the sparsity order is known, outperforms the proposed RERD algorithm. However, our RERD algorithm can accurately determine the support of a multiband signal even in low SNR and without requiring knowledge of the sparsity level of the PU signals and noise variance. This makes the proposed algorithm effectively a blind sub-Nyquist spectrum sensing method.

In this paper, boldface letters are used to represent a matrices, e.g. $\mathbf{X}$. $\mathbf{X}^T$ and $\mathbf{X}^H$ denote the transpose and the conjugate transpose of $\mathbf{X}$, respectively. $\mathbf{X}^{-1}$ represents the matrix inverse of a square matrix $\mathbf{X}$. $\mathbf{X}^\dagger$ denotes the matrix pseudo-inverse. $\mathbf{X}_k$ corresponds to the kth column vector of $\mathbf{X}$, and $\|\mathbf{X}_k\|_2$ is the l_2-norm of $\mathbf{X}_k$. $E[\cdot]$ and $D[\cdot]$ represent the statistical mean and variance, respectively. $\mathrm{Cov}[\cdot,\cdot]$ denotes the covariance of two random variables, and $\Re[\cdot]$ and $\Im[\cdot]$ take the real and imaginary parts, respectively. The rest of this paper is organized as follows. Section II-A briefly introduces compressed sampling and support recovery using the MWC framework, followed by the system model for the considered SNSS problem in Section II-B. In Section III, the RERD algorithm is presented and closed-form expressions for the false alarm probability and the decision threshold are derived. Section IV provides simulation results to verify the performance of our proposed algorithm. Finally, we conclude our study in Section V.

II. MWC-Based SNSS and Problem Formulation

To facilitate the presentation and analysis of the proposed method in subsequent sections, we first briefly review the conventional MWC-based SNSS framework and then introduce the blind sub-Nyquist spectrum sensing problem.

A. The MWC-Based SNSS

Fig. 1 depicts the flow of the conventional MWC-based SNSS framework [16], [23], [27]. As shown in the Fig. 1(a), the spectrum of a wideband signal $x(t)$, which is bandlimited in the frequency range of $\mathcal{F} = [-f_{\mathrm{NYQ}}/2, f_{\mathrm{NYQ}}/2]$, is divided into a number of subbands with identical bandwidth B. The received signal $x(t)$ at the CR contains a multiband signal $s(t)$ (the PU signal) and a Gaussian white noise $w(t)$ with zero mean and power spectrum density σ_w^2. The PU signal $s(t)$ comprises of M_s disjoint bandpass signals $s_m(t), m = 1, 2, \ldots, M_s$, with bandwidth B_m and central frequency f_{c_m}. Each signal $s_m(t)$ with bandwidth B_m may be located in one or more observation subbands B. Note that, the sparsity order [23], denoted by M, of the input signal in this case is defined as the number of occupied subbands, which is related to the bandwidth of subband B, the signal bandwidth B_m and the central frequency f_{c_m}, where $m = 1, 2, \ldots, M_s$. The set S, whose elements are the indexes of these occupied subbands, is regarded as the frequency support of $x(t)$.

In the compressed sampling stage of the conventional MWC-based SNSS as sketched in Fig. 1(b), the received signal $x(t)$ passed through G channels of the MWC simultaneously, and

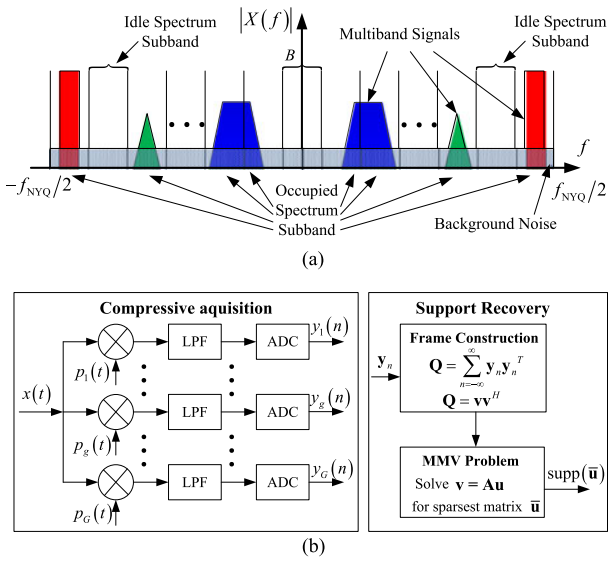


Fig. 1. The conventional MWC-based SNSS. (a) Power spectrum of the received signal $x(t)$. (b) Compressive measurements acquisition and support recovery.

multiplied by a T_p-periodic sequence $p_g(t)$ in the gth channel, which contains J chips alternating between levels ± 1. Afterwards, the mixed signal of each channel is filtered by a low-pass filter (LPF) $q(t)$ with pass band $[-f_s/2, f_s/2]$. Mathematically, the output of the LPF can be expressed as

$$y_g(t) = \int_{-\infty}^{\infty} x(\zeta)\, p_g(\zeta)\, q(t-\zeta)\, d\zeta,\ g = 1, 2, \ldots, G. \quad (1)$$

The compressive measurements $y_g(n)$ can be obtained by discretizing $y_g(t)$ with a sampling interval $T_s = 1/f_s$,

$$y_g(n) = y_g(t)|_{t=nT_s},\ n = 0, 1, \ldots. \quad (2)$$

The discrete-time Fourier transform (DTFT) of $y_g(n)$ is given by

$$\begin{aligned} Y_g\left(e^{j2\pi f/f_s}\right) &= \sum_{n=-\infty}^{\infty} y_g(n) e^{-j2\pi n f/f_s} \\ &= \sum_{\iota=-L_0}^{L_0} c_{g\iota} X(f - \iota f_p),\ f \in \mathcal{F}_s, \end{aligned} \quad (3)$$

where $X(f)$ is the Fourier transform of $x(t)$, $c_{g\iota} = \frac{1}{T_p}\int_0^{T_p} p_g(t) e^{-j\frac{2\pi}{T_p}\iota t} dt$ are the coefficients of the Fourier expansion of $p_g(t)$, and $f_p = 1/T_p$. $L_0 = \lceil \frac{f_{NYQ} + f_s}{2f_p} \rceil - 1$ is chosen as the smallest integer such that the summation in (3) contains all the subbands of $X(f)$ within $\mathcal{F}_s = [-f_s/2, f_s/2]$. Without loss of generality, we set $f_p = f_s$ and $f_p = B$. In this case, if we define $L = 2L_0 + 1$, J equals to L. Since the discrete Fourier transform (DFT) is obtained by uniformly sampling the corresponding DTFT, it follows from (3) that

$$Y_g(k) = \sum_{\iota=-L_0}^{L_0} c_{g\iota} X(k - \iota N), \quad (4)$$

where $g = 1, 2, \ldots, G,\ k = 0, 1, \ldots, N-1$, N is the number of samples in the frequency band $\mathcal{F}$ and $X(k - \iota N)$ is the DFT of the input signal $x(t)$.

Note that $Y_g(k)$ is a linear weighted combination of the DFT of all the subbands in $\mathcal{F}$. Define vector $\mathbf{Y}_k = [Y_1(k), \ldots, Y_g(k), \ldots, Y_G(k)]^T$ and vector $\mathbf{Z}_k = [Z_1(k), \ldots, Z_l(k), \ldots, Z_L(k)]^T$ with $Z_l(k) = X(k + (l - L_0 - 1)N)$. Then, (4) can be rewritten as $\mathbf{Y}_k = \mathbf{A}\mathbf{Z}_k$, where $\mathbf{A} \in \mathbb{C}^{G \times L}$ is the measurement matrix with elements $\mathbf{A}_{gl} = c_{g,-l+L_0+1}$, $1 \le l \le L$ and $L = 2L_0 + 1$. Collecting the vectors $\mathbf{Y}_k$ and $\mathbf{Z}_k$ in matrices $\mathbf{Y}$ and $\mathbf{Z}$, respectively, a more compact form of the compressed samples can be expressed as

$$\mathbf{Y} = \mathbf{A}\mathbf{Z}. \quad (5)$$

In the stage of support recovery, the autocorrelation matrix $\hat{\mathbf{Q}}$ is firstly computed and decomposed as

$$\hat{\mathbf{Q}} = \sum_{n=0}^{N-1} \mathbf{y}_n \mathbf{y}_n^T \quad (6)$$

$$\hat{\mathbf{Q}} = \mathbf{v}\mathbf{v}^H \quad (7)$$

where compressive measurements vector $\mathbf{y}_n = [y_1(n), \cdots, y_g(n), \cdots, y_G(n)]^T$, $g = 1, 2, \ldots, G$, $n = 0, 1, \cdots, N-1$, and $\mathbf{v}$ is the Hermitian square root of $\hat{\mathbf{Q}}$. As a result, a multiple measurement vectors (MMV) problem can be formed by the measurement matrix $\mathbf{A}$ and the frame $\mathbf{v}$:

$$\mathbf{v} = \mathbf{A}\mathbf{u}. \quad (8)$$

It has been shown that if $x(t)$ is sparse (i.e., $M \ll L$), (8) has a unique solution matrix $\bar{\mathbf{u}}$ with the fewest nonzero rows, the indexes of which are consistent with the support set of $\mathbf{Z}_k$ [34]. Therefore, the support of the vector $\mathbf{Z}_k$ can be obtained as $\hat{S} = \text{supp}(\bar{\mathbf{u}})$, where $\bar{\mathbf{u}}$ can be found by a greedy algorithm [36].

It is necessary to point out that greedy algorithms are implemented in an iterative fashion, where a termination condition of the iterative process should be determined in advance. Most existing greedy algorithms require knowledge of the sparsity order of the received signal or the power level of the background noise to set a termination condition [38], [39]. However, it is often difficult to acquire such prior knowledge in practical CR scenarios due to the time-varying nature of spectrum usage and dynamic fluctuations of the wireless channel. Therefore, accurate termination of the iterative process for these methods is difficult due to lack of such prior knowledge, which may lead to either under-recovery, i.e., the recovered support is only a subset of the true support or over-recovery, i.e., the recovered support is a superset of the true support. The former will increase the missing detection rate while the latter will enlarge the false alarm rate.

B. The Problem

To address the above issue, a blind sub-Nyquist spectrum sensing method referred to as the residual energy ratio based detector (RERD) is proposed in this paper. Specifically, rather than using the aforementioned priori knowledge, the RERD algorithm aims to accurately control the iterative process by

blind residual detection, which performs a decision on the absence/presence of the multiband signal in the residual signal after each iteration and cancellation of a detected subband signal. The decision result is used to determine whether convergence is reached or a further iteration is needed. Mathematically, this process can be formulated as a set of binary hypothesis test problem as

$$\mathbf{Y}^{\tau} = \begin{cases} \mathbf{A}^{\tau}\mathbf{W}, & \mathcal{H}_{0,\tau} \\ \mathbf{A}^{\tau}(\mathbf{S}+\mathbf{W}), & \mathcal{H}_{1,\tau} \end{cases}, \tag{9}$$

In (9), $\mathbf{Y}^{\tau}$ is called the residual matrix, containing the residual of the input signal after the τth iteration, and $\mathbf{A}^{\tau}$ is the equivalent measurement matrix after the τth iteration, both of which are given in the subsequent section. $\mathbf{S}$ and $\mathbf{W} \in \mathbb{C}^{L\times N}$ are composed with elements $\mathbf{S}_{l,k} = S(k+(l-L_0-1)N)$ and $\mathbf{W}_{l,k} = W(k+(l-L_0-1)N)$, which are formed by the DFT of the multiband signal $s(t)$ and background noise $w(t)$, respectively. $\mathcal{H}_{0,\tau}$ denotes the hypothesis that the residual contains only a noise component after the τth iteration, which implies that the iterative process of the greedy algorithm should be terminated. Meanwhile, $\mathcal{H}_{1,\tau}$ indicates that there exists a signal component in the residual after the τth iteration and the iterative process should continue. The problem of interest is to design a suitable decision rule for this sequential residual detection problem. By sequential residual detection, the iterative process is controlled on-line and adaptively rather than by a fixed termination condition that requires additional prior information.

Following the proposed paradigm, the iterative process of support recovery is terminated once $\mathcal{H}_{0,\tau}$ is declared. In turn, an estimated support $\hat{S}$ of the input multiband signal is obtained. To evaluate the performance of the proposed RERD algorithm, several performance metrics are defined as follows

$$P_d = \mathbf{Pr}\{\hat{S} = S\}, \tag{10a}$$

$$P_m = \mathbf{Pr}\{\hat{S} \subset S\}, \tag{10b}$$

$$P_f = \mathbf{Pr}\{\hat{S} \supset S\}, \tag{10c}$$

where S is the true support of the input signal. If the recovered support equals the true one, the associated probability is called the probability of correct detection P_d. Meanwhile, the probability of missed detection P_m and the probability of false alarm P_f refer to the probability associated with under-recovery and over-recovery, respectively.

III. The Proposed RERD Algorithm

In this section, we first present the proposed RERD algorithm, then examine the stochastic properties of the DFT of the compressive samples and channel energy, and finally derive an asymptotical expression of the false alarm probability as well as a closed-form expression of the decision threshold for each iteration in the RERD.

A. The RERD Algorithm

As illustrated in Fig. 2, similar to the SNSS methods based on greedy pursuit, the proposed RERD algorithm also conducts an iterative process consisting of the following steps: select the column of the measurement matrix that is most correlated with the compressive measurement, add the column index to the recovery support, and update the residual matrix by subtracting the estimated signal associated with the current recovered support. However, the RERD algorithm significantly differs from existing greedy pursuit based SNSS methods in the control of the iterative process. Instead of using a preset number of iterations or a stop criterion based on the l_2-norm of the residual matrix [35], the residual detection threshold of our proposed algorithm is adaptively adjusted based on a given false alarm probability. The pseudocode of the proposed RERD is presented in Algorithm 1 which provides more details of Fig. 2.

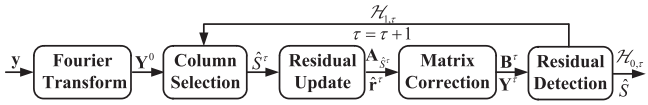


Fig. 2. A schematic of the RERD algorithm.

In line 2, the residual matrix in the τth iteration $\mathbf{Y}^{\tau-1}$ is projected onto each column of the measurement matrix $\mathbf{A}$ and the column that is most correlated with the current residual is selected. In line 3, the signal support $\hat{S}^{\tau}$ is updated by adding column indices I^{τ} and $L+1-I^{\tau}$ (due to spectrum symmetry of real input signal). The operations in line 2 and line 3 correspond to the column selection block in Fig. 2. In line 4 and line 5, the intermediate projector $\mathbf{A}_{\hat{S}^{\tau}}$ is used to compute the occupied subbands $\hat{\mathbf{X}}^{\tau}$, and the residual $\hat{\mathbf{r}}^{\tau}$ is then updated by subtracting the projection of $\hat{\mathbf{X}}^{\tau}$ onto $\mathbf{A}_{\hat{S}^{\tau}}$ from $\mathbf{Y}^0$. These operations correspond to the residual update block in Fig. 2.

In line 6, the residual matrix $\mathbf{Y}^{\tau}$ is calculated by multiplying $\hat{\mathbf{r}}^{\tau}$ with a diagonal correction matrix $\mathbf{M}^{\tau}$, whose elements are determined by matrix $\mathbf{B}^{\tau} = (\mathbf{E} - \mathbf{A}_{\hat{S}^{\tau}}\mathbf{A}^{\dagger}_{\hat{S}^{\tau}})\mathbf{A}$. Matrix $\mathbf{E}$ in this line is a $G \times G$ identity matrix. This matrix correction operation keeps the energy of the row vectors in $\mathbf{A}^{\tau}$ to be equal to 1 in the τth iteration. Under $\mathcal{H}_{0,\tau}$, this operation makes the statistic characteristic of the channel energy T_g^{τ} in different iterations unchanged. $T_g^{\tau},\ g = 1, 2, \ldots, G$ is a channel energy and $\mathbf{A}^{\tau}$ is an equivalent measurement matrix, which are defined in line 7 and line 8, respectively.

In line 7, G branch test statistics are constructed, and the test statistic corresponding to the gth branch detection is given by $o_g^{\tau} = T_g^{\tau}/T_{g+1}^{\tau},\ g = 1, 2, \ldots, G$, where $T_g^{\tau} = \frac{1}{N}\sum_{k=0}^{N-1} |\mathbf{Y}_{g,k}^{\tau}|^2$ and $T_{G+1}^{\tau} = T_1^{\tau}$. In line 8, a pair of decision thresholds $\gamma_g^{\tau,f}$ and $\gamma_g^{\tau,b}$ for the gth branch detection are defined, and their expressions are derived in the subsequent subsection. From line 9 to line 13, each branch detection is performed by comparing the energy ratio $o_g^{\tau}, g = 1, 2, \ldots, G$, with the decision thresholds $\gamma_g^{\tau,f}$ and $\gamma_g^{\tau,b}$. The residual detection in the τth iteration is a collaborative detection scheme which fuses all the G branch detection results with an "OR" rule. The residual detection result is utilized to determine whether a further iteration is needed. In comparison with the conventional MWC-based SNSS [31], [40], [41], the proposed RERD algorithm does not require any prior knowledge of the sparsity level of the PU signal nor the noise power. It can work blindly by employing residual detection to determine the termination condition of the support recovery process.

Algorithm 1: The Proposed RERD.

Input: measurement matrix $\mathbf{A}$; false alarm probability of residual detection P_f;
compressive measurements $\mathbf{y} = [\mathbf{y}_0, \ldots, \mathbf{y}_n, \ldots, \mathbf{y}_{N-1}]$.
Initialize:
signal support $\hat{S}^0 = \emptyset$; iteration index $\tau = 1$;
residual matrix $\mathbf{Y}^0_{G\times N}$ with elements $\mathbf{Y}^0_{g,k} = \sum_{n=0}^{N-1}\mathbf{y}_{g,n}e^{-j2\pi kn/N}$.
1: **Repeat:**
2: $I^\tau = \arg\max_l \|\mathbf{A}_l^T\mathbf{Y}^{\tau-1}\|_2/\|\mathbf{A}_l^T\|_2$;
3: $\hat{S}^\tau = \hat{S}^{\tau-1}\bigcup\{I^\tau,\ L+1-I^\tau\}$;
4: $\hat{\mathbf{X}}^\tau = \mathbf{A}^\dagger_{\hat{S}^\tau}\mathbf{Y}^0$; $\mathbf{A}^\dagger_{S^\tau} = (\mathbf{A}^T_{\hat{S}^\tau}\mathbf{A}_{\hat{S}^\tau})^{-1}\mathbf{A}^T_{\hat{S}^\tau}$;
5: $\hat{\mathbf{r}}^\tau = \mathbf{Y}^0 - \mathbf{A}_{\hat{S}^\tau}\hat{\mathbf{X}}^\tau$;
6: $\mathbf{Y}^\tau = \mathbf{M}^\tau\hat{\mathbf{r}}^\tau$; $\mathbf{M}^\tau = \mathrm{diag}(M_1^\tau,\ldots,M_g^\tau,\ldots,M_G^\tau)$;
$\mathbf{B}^\tau_{G\times L} = (\mathbf{E} - \mathbf{A}_{\hat{S}^\tau}\mathbf{A}^\dagger_{\hat{S}^\tau})\mathbf{A}$;
$M_g^\tau = 1/\sqrt{\sum_{l=1}^{L}|\mathbf{B}^\tau_{g,l}|^2}$;
7: $o_g^\tau = \frac{T_g^\tau}{T_{g+1}^\tau}$; $T_g^\tau = \frac{1}{N}\sum_{k=0}^{N-1}|\mathbf{Y}^\tau_{g,k}|^2$; $g = 1,2,\ldots,G$;
$T^\tau_{G+1} = T^\tau_1$;
8: $\gamma_g^{\tau,f}, \gamma_g^{\tau,b} = \Psi(\mathbf{A}^\tau, P_f^\tau, N)$[§]; $\mathbf{A}^\tau = \mathbf{M}^\tau$
$(\mathbf{E} - \mathbf{A}_{\hat{S}^\tau}\mathbf{A}^\dagger_{\hat{S}^\tau})\mathbf{A}$;
9: **if** $\exists\ o_g^\tau$ s.t. $o_g^\tau \geq \gamma_g^{\tau,f}$ **or** $o_g^\tau \leq \gamma_g^{\tau,b}$ **then**
10: choose $\mathcal{H}_{1,\tau}$; $\tau = \tau + 1$;
11: **else**
12: choose $\mathcal{H}_{0,\tau}$; $\hat{S} = \hat{S}^\tau$;
13: **end if**
14: **until** stopping criterion $\mathcal{H}_{0,\tau}$ is met.
Output: Sparse signal support $\hat{S}$.

B. False Alarm Probability

In Line 8 of the RERD algorithm, the false alarm probability of residual detection in the τth iteration can be expressed as

$$P_f^\tau = 1 - \prod_{g=1}^{G}(1 - P_{f,g}^\tau), \tag{11}$$

where $P_{f,g}^\tau$ is the false alarm probability of the gth branch detection. Note that, since energy ratio is employed as the test statistic, it is reasonable to compare o_g^τ with a lower threshold $\gamma_g^{\tau,f}$ and an upper threshold $\gamma_g^{\tau,b}$ to achieve better performance. The branch detection is thus a bidirectional decision process, whereby $P_{f,g}^\tau$ is the sum of the forward false alarm probability $P_{f,g}^{\tau,f}$ and backward false alarm probability $P_{f,g}^{\tau,b}$, which are defined as

$$P_{f,g}^{\tau,f} = \Pr\{o_g^\tau \geq \gamma_g^{\tau,f}|\mathcal{H}_{0,\tau}\}, \tag{12a}$$

$$P_{f,g}^{\tau,b} = \Pr\{o_g^\tau \leq \gamma_g^{\tau,b}|\mathcal{H}_{0,\tau}\}. \tag{12b}$$

Consequently, the expression for $P_{f,g}^{\tau,f}$ and $P_{f,g}^{\tau,b}$ should be deduced first to obtain the P_f^τ in the τth iteration. Under $\mathcal{H}_{0,\tau}$, the contribution from the occupied spectrum subbands has been removed from the original signal $\mathbf{Y}^0$, and the entries of the residual matrix $\mathbf{Y}^\tau$ is the DFT of the idle spectrum subbands which consists of white noise only. In this case, based on the statistical analysis on the entries of the residual matrix $\mathbf{Y}^\tau$, and the channel energy T_g^τ, the distribution of the test statistics o_g^τ and expressions for $P_{f,g}^{\tau,f}$ and $P_{f,g}^{\tau,b}$ can be derived progressively. The final result is given in the following theorem.

Theorem 1: Under $\mathcal{H}_{0,\tau}$, when the sample number N is sufficiently large, the false alarm probability $P_{f,g}^{\tau,f}$ and $P_{f,g}^{\tau,b}$ corresponding to preset decision thresholds $\gamma_g^{\tau,f}$ and $\gamma_g^{\tau,b}$ are

$$P_{f,g}^{\tau,f} = \Pr\left(\frac{T_g^\tau}{T_{g+1}^\tau} \geq \gamma_g^{\tau,f}\middle|\mathcal{H}_{0,\tau}\right) = 1 - \Phi\left(\frac{\sqrt{N}\gamma_g^{\tau,f} - \sqrt{N}}{\sqrt{1 - 2\gamma_g^{\tau,f}\rho_g^\tau + (\gamma_g^{\tau,f})^2}}\right), \tag{13a}$$

$$P_{f,g}^{\tau,b} = \Pr\left(\frac{T_g^\tau}{T_{g+1}^\tau} \leq \gamma_g^{\tau,b}\middle|\mathcal{H}_{0,\tau}\right) = \Phi\left(\frac{\sqrt{N}\gamma_g^{\tau,b} - \sqrt{N}}{\sqrt{1 - 2\gamma_g^{\tau,b}\rho_g^\tau + (\gamma_g^{\tau,b})^2}}\right), \tag{13b}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{\eta^2}{2}}d\eta$ is the Gaussian cumulative distribution function (CDF) and the correlation coefficient is given by

$$\rho_g^\tau \triangleq \frac{(\rho_g^{\tau,0})^2 + (\rho_g^{\tau,1})^2 + (\rho_g^{\tau,2})^2 + (\rho_g^{\tau,3})^2}{2}, \tag{14}$$

where

$$\rho_g^{\tau,0} = \sum_{l=1}^{L}[\Re(\mathbf{A}^\tau_{g,l})\Im(\mathbf{A}^\tau_{g+1,l}) - \Re(\mathbf{A}^\tau_{g+1,l})\Im(\mathbf{A}^\tau_{g,l})], \tag{15a}$$

$$\rho_g^{\tau,1} = \sum_{l=1}^{L}[\Im(\mathbf{A}^\tau_{g,l})\Re(\mathbf{A}^\tau_{g+1,l}) - \Re(\mathbf{A}^\tau_{g,l})\Im(\mathbf{A}^\tau_{g+1,l})], \tag{15b}$$

$$\rho_g^{\tau,2} = \sum_{l=1}^{L}[\Re(\mathbf{A}^\tau_{g,l})\Re(\mathbf{A}^\tau_{g+1,l}) + \Im(\mathbf{A}^\tau_{g+1,l})\Im(\mathbf{A}^\tau_{g,l})], \tag{15c}$$

$$\rho_g^{\tau,3} = \sum_{l=1}^{L}[\Re(\mathbf{A}^\tau_{g,l})\Re(\mathbf{A}^\tau_{g+1,l}) + \Im(\mathbf{A}^\tau_{g+1,l})\Im(\mathbf{A}^\tau_{g,l})], \tag{15d}$$

and $\mathbf{A}^\tau = \mathbf{M}^\tau(\mathbf{E} - \mathbf{A}_{\hat{S}^\tau}\mathbf{A}^\dagger_{\hat{S}^\tau})\mathbf{A}$ is the equivalent measurement matrix in the τth iteration.

Proof: To prove Theorem 1, we start from Line 6

$$\begin{aligned}\mathbf{Y}^\tau &= \mathbf{M}^\tau\hat{\mathbf{r}}^\tau = \mathbf{M}^\tau(\mathbf{Y}^0 - \mathbf{A}_{\hat{S}^\tau}\hat{\mathbf{X}}^\tau)\\ &= \mathbf{M}^\tau(\mathbf{Y}^0 - \mathbf{A}_{\hat{S}^\tau}\mathbf{A}^\dagger_{\hat{S}^\tau}\mathbf{Y}^0)\\ &= \mathbf{M}^\tau(\mathbf{E} - \mathbf{A}_{\hat{S}^\tau}\mathbf{A}^\dagger_{\hat{S}^\tau})\mathbf{A}\mathbf{Z},\end{aligned} \tag{16}$$

where $\mathbf{E} \in \mathbb{C}^{G\times G}$ is an identity matrix. If we define $\mathbf{A}^\tau = \mathbf{M}^\tau(\mathbf{E} - \mathbf{A}_{\hat{S}^\tau}\mathbf{A}^\dagger_{\hat{S}^\tau})\mathbf{A}$ as an equivalent measurement matrix,

[§]The function Ψ is given in the following section. Note that it is related to the measurement matrix, the false alarm probability of residual detection and the number of compressive measurement vectors.

$$\gamma_g^{\tau,f} = \frac{N - \rho_g^\tau \left[\Phi^{-1}\left(1 - P_{f,g}^{\tau,f}\right)\right]^2 + \Phi^{-1}\left(1 - P_{f,g}^{\tau,f}\right)\sqrt{(\rho_g^{\tau\,2} - 1)\left[\Phi^{-1}\left(1 - P_{f,g}^{\tau,f}\right)\right]^2 + 2N(\rho_g^\tau - 1)}}{N - \left[\Phi^{-1}\left(1 - P_{f,g}^{\tau,f}\right)\right]^2}, \tag{22a}$$

$$\gamma_g^{\tau,b} = \frac{N - \rho_g^\tau \left[\Phi^{-1}\left(P_{f,g}^{\tau,b}\right)\right]^2 - \Phi^{-1}\left(P_{f,g}^{\tau,b}\right)\sqrt{((\rho_g^\tau)^2 - 1)\left[\Phi^{-1}\left(P_{f,g}^{\tau,b}\right)\right]^2 + 2N(\rho_g^\tau - 1)}}{N - \left[\Phi^{-1}\left(P_{f,g}^{\tau,b}\right)\right]^2}, \tag{22b}$$

$\mathbf{Y}^\tau = \mathbf{A}^\tau \mathbf{Z}$ can be regarded as a compressive sampling on the matrix $\mathbf{Z}$ with a new measurement matrix $\mathbf{A}^\tau$. Under $\mathcal{H}_{0,\tau}$, using the associative and distributive laws of matrix, it can be shown that the columns with indexes taken from support $\hat{S}^\tau$ are identical in matrix $\mathbf{A}$ and $\mathbf{A}_{\hat{S}^\tau}$ ($\mathbf{A}_{\hat{S}^\tau}^\dagger \mathbf{A}$), while the columns vectors $\mathbf{A}_l^\tau$, $l \in \hat{S}^\tau$ are 0, where $\hat{S}^\tau$ is the complete support of multiband signal. Then $\mathbf{A}^\tau$ acquires nothing from the occupied subbands in this case. Hence, even if we replace the occupied subbands with idle subbands, this will cause no changes in the statistical properties of $\mathbf{Y}^\tau$, and we can infer the statistical characteristic of $\mathbf{Y}_{g,k}^\tau$ form $\mathbf{Y}^\tau = \mathbf{A}^\tau \mathbf{W}$ instead of $\mathbf{Y}^\tau = \mathbf{A}^\tau \mathbf{Z}$ for simplicity.

Considering the linearity of the DFT and the assumption that the noise $w(t)$ has zero mean and power spectrum density σ_w^2, $\mathrm{W}_{l,k}$ is a complex Gaussian random variable with zero mean and variance $LN\sigma_w^2$. Since the sum of the modular square of the row vector in $\mathbf{A}^\tau$ is 1, $\mathbf{Y}_{g,k}^\tau$, which can be expressed as $\mathbf{Y}_{g,k}^\tau = \sum_{l=1}^{L} \mathbf{A}_{g,l}\mathbf{W}_{l,k}$ is also a complex Gaussian random variable with zero mean and variance $LN\sigma_w^2$. Based on the statistical characteristic of $\mathbf{Y}_{g,k}^\tau$, Lemma 1 in Appendix indicates that T_g^τ is a sum of N independent random variables $|\mathbf{Y}_{g,k}^\tau|^2$, $k = 0, 1, \ldots, N-1$ with identical mean and variance under $\mathcal{H}_{0,\tau}$. Thus, for large N and following the central limit theorem, T_g^τ is approximately a Gaussian variable with mean and variance given by

$$E\left[T_g^\tau | \mathcal{H}_{0,\tau}\right] = \frac{1}{N}\sum_{k=0}^{N-1} E\left[\left|\mathbf{Y}_{g,k}^\tau\right|^2\right] = LN\sigma_w^2, \tag{17a}$$

$$D\left[T_g^\tau | \mathcal{H}_{0,\tau}\right] = \frac{1}{N}\sum_{k=0}^{N-1} D\left[\left|\mathbf{Y}_{g,k}^\tau\right|^2\right] = L^2N^2\sigma_w^4. \tag{17b}$$

The covariance of different channel energies are

$$\begin{aligned}
&\mathrm{Cov}\left[T_g^\tau, T_{g+1}^\tau\right] \\
&= \mathrm{Cov}\left[\frac{1}{N}\sum_{k=0}^{N-1}\left|Y_g^\tau(k)\right|^2, \frac{1}{N}\sum_{k=0}^{N-1}\left|Y_{g+1}^\tau(k)\right|^2\right] \\
&= \frac{1}{N^2}\sum_{k=0}^{N-1}\mathrm{Cov}\left[\left|Y_g^\tau(k)\right|^2, \left|Y_{g+1}^\tau(k)\right|^2\right] \\
&= \frac{(\rho_g^{\tau,0})^2 + (\rho_g^{\tau,1})^2 + (\rho_g^{\tau,2})^2 + (\rho_g^{\tau,3})^2}{2} L^2N^2\sigma_w^4,
\end{aligned} \tag{18}$$

where $\rho_g^{\tau,0}$, $\rho_g^{\tau,1}$, $\rho_g^{\tau,2}$ and $\rho_g^{\tau,3}$ are defined in Theorem 1. The last equality of (18) follows from (24) of Lemma 1. Therefore, the correlation coefficient ρ_g^τ between T_g^τ and T_{g+1}^τ can be calculated as

$$\begin{aligned}
\rho_g^\tau &= \frac{\mathrm{Cov}\left[T_g^\tau, T_{g+1}^\tau\right]}{\sqrt{D\left[T_g^\tau | \mathcal{H}_{0,\tau}\right] D\left[T_{g+1}^\tau | \mathcal{H}_{0,\tau}\right]}} \\
&= \frac{(\rho_g^{\tau,0})^2 + (\rho_g^{\tau,1})^2 + (\rho_g^{\tau,2})^2 + (\rho_g^{\tau,3})^2}{2}.
\end{aligned} \tag{19}$$

According to the derivation above, the test statistics o_g^τ under $\mathcal{H}_{0,\tau}$ is the ratio of two statistically correlated Gaussian variables whose statistics are given by (17), (18), and (19). To find the CDF of o_g^τ, we employ the following property.

Property 1: If the correlation coefficient between two Gaussian random variable I_1 and I_2 is ρ, the CDF of $R = I_1/I_2$ is [45]

$$F_R(r) = \Pr(R < r) = \Phi\left(\frac{r\mu_2 - \mu_1}{\sqrt{\sigma_1^2 - 2r\rho\sigma_1\sigma_2 + r^2\sigma_2^2}}\right), \tag{20}$$

where $\Phi(x)$ the Gaussian CDF defined in Theorem 1, μ_1, μ_2, σ_1^2, and σ_2^2 are the mean and variance of I_1 and I_2, respectively and they are all positive.

Eqn. (13) follows immediately by applying Property 1 to the test statistic of branch detection $o_g^\tau = T_g^\tau / T_{g+1}^\tau$. ■

C. Decision Threshold of Branch Detection

Suppose all the false alarm probabilities of branch detection are identical, (11) can be simplified as

$$P_f^\tau = 1 - (1 - P_{f_g}^\tau)^G. \tag{21}$$

Following the Neyman-Pearson criterion, if the target false alarm probability of residual detection P_f^τ is given, we can calculate $P_{f_g}^\tau$ from (21), and then the decision thresholds $\gamma_g^{\tau,f}$ and $\gamma_g^{\tau,b}$ for branch detection in the τth iteration can be obtained by inverting (13), where the expression at the right side of the equal sign in (22), shown at the top of the page, is defined as the function $\Psi(\cdot)$ and $\Phi^{-1}(x)$ is the inverse function of $\Phi(x)$. From the process of the RERD and the threshold expression, we know that the threshold $\gamma_g^{\tau,f}$ and $\gamma_g^{\tau,b}$ are determined by the compressive measurements number N, false alarm probability P_f^τ, and the correlation coefficient ρ_g^τ which only depends on the equivalent measurement matrix $\mathbf{A}^\tau$. Therefore, the RERD threshold $\gamma_g^{\tau,f}$ and $\gamma_g^{\tau,b}$ can be accurately computed from these parameters. Notably, the RERD threshold $\gamma_g^{\tau,f}$ and $\gamma_g^{\tau,b}$ are not related to the noise variance σ_w^2, and does not require any prior knowledge about the noise $w(t)$. Therefore, the RERD is inherently a blind sub-Nyqusit spectrum sensing scheme.

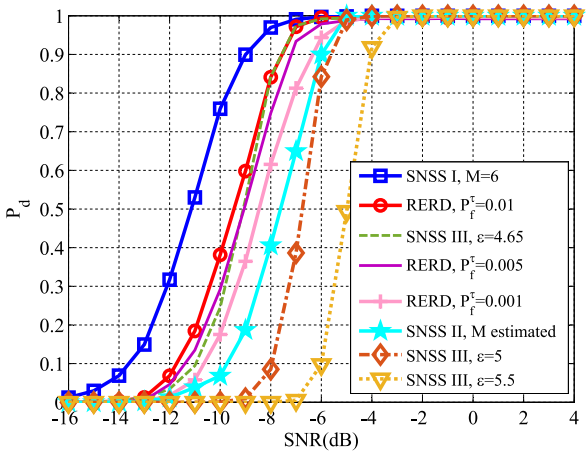


Fig. 3. Comparison between the correct detection probability of conventional SNSS method and the RERD algorithm.

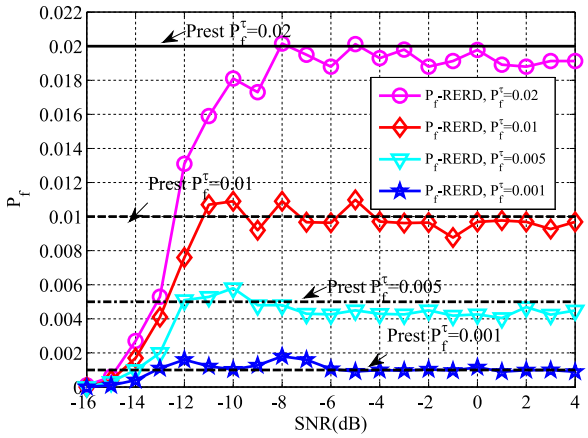


Fig. 4. False alarm probability of the RERD algorithm with different P_f^τ.

IV. Performance Simulation and Analysis

In this section, we first discuss the performance simulation setting, followed by several performance evaluation results.

A. Simulation Setup

In the following simulations, the system simulation parameters of the MWC are: the equivalent sampling rate $f_{\mathrm{NYQ}} =$ 6GHz, the number of sampling channels $G = 40$, the period of each random sequence $T_p = 19.17$ns with $L = 115$ chips of random ± 1 in each period, the sampling rate of single channel $f_s = B = f_p$, and the cutoff frequency of LPF in each channel equals to $f_s/2$, where $f_p = 1/T_p = 52.2$ MHz. The transmission signal of PU consists of M_s digital modulation signals. The symbol rate s_r of each signal is 20 MBaud ($s_r < f_p$), and the carriers frequency of each PU signal is $l' f_p$, where l' is an integer and can be randomly selected in the set $[1, 2, \ldots, L_0]$ with equal probability. It means that each PU signal will be located at the center of two subband (due to spectrum symmetry of real input signal). Thus $M = 2M_s$ subbands have been occupied in the frequency domain. The signal-to-noise ratio (SNR) used in simulations is defined as $\mathrm{SNR} = (P/\sigma_w^2)$, where P is the total power of M_s PU signals with identical power. According to the rule of constant false alarm, P_f^τ is fixed, and false alarm probability of branch detection $P_{f,g}^\tau$ is computed from (21). We set $P_{f,g}^{\tau,f} = P_{f,g}^{\tau,b} = P_{f,g}^\tau/2$ in the simulations. The results as below are all averaged over 10^4 Monte Carlo experiments.

B. Simulation Results

Fig. 3 compares the detection probability P_d of the RERD algorithm with that of the conventional SNSS methods. In the simulation, $M = 6$ bands are active, $N = 1000$ compressive measurement vectors are used, and SNR ranges from -16dB to 4dB. Three SNSS methods with different conditions of termination iteration are chosen as benchmarks. We named the SNSS method in references [19], [39] as "SNSS I", which terminates the iteration of the support recovery according to a known sparsity order. The SNSS method in reference [40] is defined as "SNSS II", which terminates the iteration of the support recovery according to an estimated sparsity order. The SNSS method in reference [34] is termed as "SNSS III", which decides whether the next iteration is necessary according to the l_2-norm value of the residual matrix. In the RERD algorithm, $P_f^\tau = [0.001, 0.005, 0.01]$ for each iteration are chosen, respectively. As shown in the figure, the detection performance of the RERD algorithm slightly improves with the increasing false alarm probability of the RERD algorithm. Benefitting from the priori knowledge [38], the detection performance of the SNSS I method with known sparsity order is better than the that of the RERD blind sub-Nyquist spectrum sensing method. Nevertheless, when the sparsity order is unknown and difficult to accurately estimate in practical scenarios [41], SNSS II method suffers a significant performance loss in the detection probability. Meanwhile, the detection performance of the SNSS III method with the stop criterion $\|\hat{\mathbf{r}}^\tau\| \le \varepsilon$ experiences notable fluctuations. When $\varepsilon = 4.65$, its performance is slightly superior to that of the RERD method. When $\varepsilon = 5$ and $\varepsilon = 5.5$, their performances decrease significantly and become inferior to that of the RERD algorithm. The RERD holds an advantage over the SNSS III method when it is difficult to set an appropriate value ε due to uncertainty in the noise level. Our results show that, even in the absence of a priori knowledge and at low SNR, the proposed method can obtain good detection performance.

From Fig. 3, we can see that the RERD algorithm with a preset P_f^τ is able to recover more elements of the true support than the SNSS III with a certain ε, and the number of recovered support elements of the RERD algorithm is less than or equal to that of the SNSS I method. In reference [39], the convergence of SNSS I and SNSS III have been analyzed, and the error bound corresponding to these SNSS methods have been listed. Hence, the error bound of the RERD algorithm lies between that of SNSS I and SNSS III, and the convergence of the RERD algorithm can be ensured.

The false alarm probability of the RERD algorithm obtained by simulation is compared with its preset value in Fig. 4. It is shown that the false alarm probability is around the preset value with slight fluctuation in relatively large SNR region. From (10c), the false alarm probability P_f of the RERD algorithm is defined as the probability of over-recovery. The over-recovery means excessive support recovery beyond the true signal

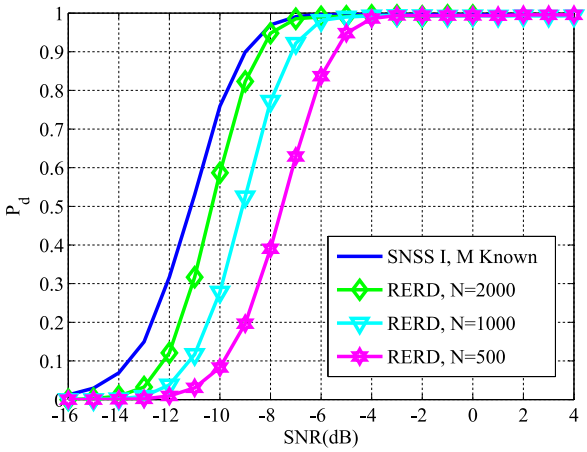


Fig. 5. Correct detection probability of the RERD algorithm with different sample number N.

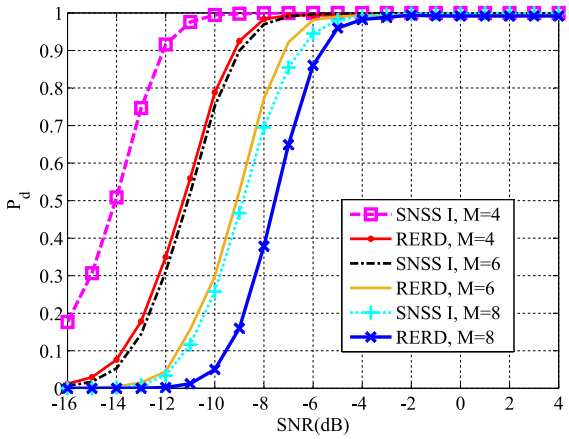


Fig. 6. Correct detection probability of the RERD algorithm with different numbers of occupied subbands M.

support set S. Hence, false alarm probability P_f is related to the correct detection probability P_d and the false alarm probability of excessive support recovery. When SNR is sufficiently large, the correct detection probability P_d almost reaches 1, and P_f of the RERD algorithm is determined by the false alarm probability P_f^{τ} of excessive iteration. These results offer an intuition for the definition of the false alarm probability.

Fig. 5 gives the correct detection probability of the RERD when the number of compressive measurement vectors N is set as 500, 1000, and 2000, respectively. It is seen that the correct detection probability of the RERD algorithm is close to 99% with just 500 compressive samples at SNR $= -4$ dB, and the detection performance can be improved by increasing the number of samples.

Fig. 6 gives the correct detection probability of the RERD when the number of occupied bands M is 4, 6, and 8, respectively. In view of the simulation setting, there exists $M_s = M/2$ nonadjacent signals which possesses an identical signal power P/M_s in the full band. As shown in the figure, the RERD algorithm can provide stable detection performance with different numbers of occupied bands. The fewer bands being occupied, the better the detection performance. This is because when M_s is larger, the power P/M_s of each signal decreases, the detection

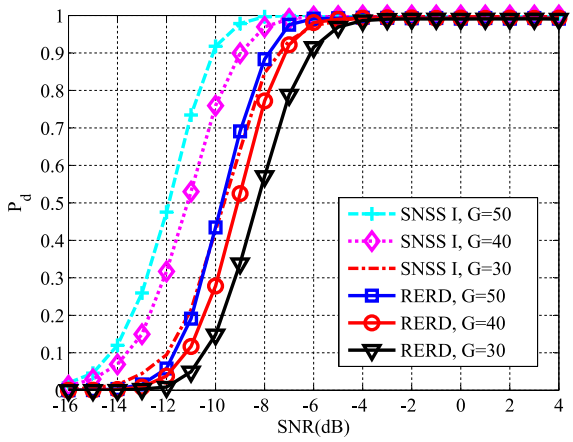


Fig. 7. Correct detection probability of the RERD algorithm with different numbers of channels G.

TABLE I
COMPARISON OF COMPUTATIONAL COMPLEXITY BETWEEN THE RERD AND THE SNSS I

Complexity / Algorithm	CMC	FFT	CMD	OMP	RD
SNSS I	NG^2	—	$\mathcal{O}(G^3)$	$\mathcal{O}(M_sLG^2)$	—
Proposed RERD	—	$NG\log_2(N)$	—	$\mathcal{O}(M_sNG^2)$	M_sGN

probability of each signal in different iterations also decreases because the signal is weaker, and so is the overall detection performance. If the power of each signal remains unchanged, the correct detection probability of the RERD will remain so as well.

Fig. 7 shows the correct detection probability of the RERD when the number of sampling channels G is set to be 30, 40, and 50, respectively. The number of occupied bands and compressive measurement vectors are fixed as $M = 6$ and $N = 1000$, respectively. It is shown that, when the condition $G \geq 2M$ is fulfilled, the RERD algorithm has the ability to achieve good correct detection performance. Note that the condition $G \geq 2M$ is the sufficient condition for successful support recovery based on the MWC in this simulation [19], [34]. The P_d of the RERD algorithm is seen to improve with more sampling channels. Besides, this simulation result offers some guideline in selecting the number of sampling channels for the modulated wideband converter.

Finally, we briefly compare the computational complexity of the SNSS I method and the RERD algorithm in Table I. The SNSS I method [19] involves the following steps: correlation matrix construction (CMC), correlation matrix decomposition (CMD) and orthogonal matching pursuit (OMP). CMC and CMD have computational complexity $\mathcal{O}(NG^2)$ and $\mathcal{O}(G^3)$ separately. In each iteration of the OMP, LG^2 multiplications are needed. If the sparsity order is known in advance, the computational cost of the OMP can be estimated as M_sLG^2. The RERD algorithm employs the FFT, OMP and residual detection (RD) based iteration control to achieve blind spectrum sensing. When performing $N-$point FFT on compressive measurements from G channels simultaneously, the associated complexity is $\mathcal{O}(GN\log_2 N)$. As shown in line 2 of the RERD algorithm, the computational complexity of the OMP in RERD is dominated by the index selection, and $\mathcal{O}(M_sNG^2)$ multiplications are

needed when indexes corresponding to occupied subbands are all selected. Residual detection (RD) based iteration control uses $\mathcal{O}(M_s GN)$ multiplications to construct the test statistics. When $N \gg G$, the computational cost of the SNSS I and the RERD can be approximated as $\mathcal{O}(NG^2)$ and $\mathcal{O}(M_s NG^2)$, respectively. The computational complexity of the proposed RERD algorithm is higher than the SNSS I method. Nonetheless, from (6) and Parseval's theorem, we know that the gth, $g = 1, 2, \ldots, G$ diagonal elements of autocorrelation matrix $\hat{\mathbf{Q}}$ is equal to frequency domain energy of the gth sampling channel. The RERD algorithm can also use CMD and CMC operations to reduce computational complexity of the index selection in the OMP. Therefore, the RERD algorithm can be implemented with a slight increase in computational cost compared with the SNSS I methods to ensure autonomous and appropriate iteration control.

V. Conclusions

In this paper, a blind sub-Nyquist spectrum sensing algorithm, referred to as the Residual Energy Ratio based Detector (RERD), is proposed. The RERD algorithm is an MWC-based SNSS scheme, involving a series of binary hypothesis testing and residual detection for support recovery. By exploiting the stochastic properties of the discrete Fourier transform of compressive measurements from the MWC, the expressions for the false alarm probability and decision threshold of residual detection are derived. Simulation results show that, even with no prior knowledge, the RERD algorithm can achieve high correct detection performance at low false alarm probability for a wide range of SNR in additive white noise channel, as well as robust recovery performance with different numbers of occupied bands and undersampling rates.

Appendix
Proof of Lemma 1

Lemma 1: Under $\mathcal{H}_{0,\tau}$, the mean and variance of $|\mathbf{Y}_{g,k}^{\tau}|^2$ are

$$E\left[|\mathbf{Y}_{g,k}^{\tau}|^2\right] = LN\sigma_w^2, D\left[|\mathbf{Y}_{g,k}^{\tau}|^2\right] = L^2N^2\sigma_w^4, \tag{23}$$

respectively, where $g = 1, 2, \ldots, G$, and $k = 0, 1, \ldots, N-1$. The covariance of the power spectrum bins $|\mathbf{Y}_{h,u}^{\tau}|^2$ and $|\mathbf{Y}_{i,v}^{\tau}|^2$ for different values of (h, i) and (u, v) is

$$\mathrm{Cov}[|\mathbf{Y}_{h,u}^{\tau}|^2, |\mathbf{Y}_{i,v}^{\tau}|^2] = \begin{cases} 0, & h = g, i = g, u \neq v \\ L^2N^2\sigma_w^2\rho_g^{\tau}, & h = g, i = g+1, u = v \\ 0, & h = g, i = g+1, u \neq v \end{cases} \tag{24}$$

where $g = 1, 2, \ldots, G$, $u, v = 0, 1, \ldots, K-1$, ρ_g^{τ} is defined in Theorem 1.

Proof: Under $\mathcal{H}_{0,\tau}$, according to the result $E[\Re(\mathbf{W}_{l,k})] = E[\Im(\mathbf{W}_{l,k})] = 0$, $D[\Re(\mathbf{W}_{l,k})] = D[\Im(\mathbf{W}_{l,k})] = NL\sigma^2/2$ and $\mathrm{Cov}[\Re(\mathbf{W}_{l,k}), \Im(\mathbf{W}_{l,k})] = 0$ [46], the mean and variance of $\Re(\mathbf{Y}_{g,k}^{\tau})$ and $\Im(\mathbf{Y}_{g,k}^{\tau})$ can be obtained easily as

$$E\left[\Re(\mathbf{Y}_{g,k}^{\tau})\right] = E\left[\Im(\mathbf{Y}_{g,k}^{\tau})\right] = 0, \tag{25a}$$

$$D\left[\Re(\mathbf{Y}_{g,k}^{\tau})\right] = D\left[\Im(\mathbf{Y}_{g,k}^{\tau})\right] = LN\sigma_w^2/2, \tag{25b}$$

and the covariance of $\Re(\mathbf{Y}_{g,k}^{\tau})$ and $\Im(\mathbf{Y}_{g,k}^{\tau})$ can be calculated as

$$\begin{aligned} &\mathrm{Cov}\left[\Re(\mathbf{Y}_{g,k}^{\tau}), \Im(\mathbf{Y}_{g,k}^{\tau})\right] \\ &= E\left[\Re(\mathbf{Y}_{g,k}^{\tau})\Im(\mathbf{Y}_{g,k}^{\tau})\right] - E\left[\Re(\mathbf{Y}_{g,k}^{\tau})\right] E\left[\Im(\mathbf{Y}_{g,k}^{\tau})\right] \\ &= E\left\{\left[\sum_{l=0}^{L}\left(\Re\left(\mathbf{A}_{g,l}^{\tau}\right)\Re\left(\mathbf{W}_{l,k}\right) - \Im\left(\mathbf{A}_{g,l}^{\tau}\right)\Im\left(\mathbf{W}_{l,k}\right)\right)\right]\right. \\ &\qquad \left.\times\left[\sum_{l=0}^{L}\left(\Re\left(\mathbf{A}_{g,l}^{\tau}\right)\Im\left(\mathbf{W}_{l,k}\right) + \Im\left(\mathbf{A}_{g,l}^{\tau}\right)\Re\left(\mathbf{W}_{l,k}\right)\right)\right]\right\} \\ &= 0. \end{aligned} \tag{26}$$

Hence, $|\mathbf{Y}_{g,k}^{\tau}|^2 = \left[\Re(\mathbf{Y}_{g,k}^{\tau})\right]^2 + \left[\Im(\mathbf{Y}_{g,k}^{\tau}\right]^2$ can be treated as a sum of two independent and identically distributed Gaussian random variables, and its mean and variance can be calculated from a central chi-square distribution with degree 2,

$$E\left[|\mathbf{Y}_{g,k}^{\tau}|^2\right] = 2\left(\sqrt{LN\sigma_w^2/2}\right)^2 = N\sigma_w^2, \tag{27a}$$

$$D\left[|\mathbf{Y}_{g,k}^{\tau}|^2\right] = 4\left(\sqrt{LN\sigma_w^2/2}\right)^4 = L^2N^2\sigma_w^4, \tag{27b}$$

which is consistent with (23).

To prove (24), we need to analyze the covariance between the power spectrum with different frequency bins from the same channel, the covariance between power spectrum with the same frequency bins from adjacent channels, and the covariance between the power spectrum with different frequency bins from adjacent channels. Since the DFT $\mathbf{W}_{l,u}$ and $\mathbf{W}_{l,v}$, $u \neq v$ are uncorrelated [47], $\mathbf{Y}_{g,k}^{\tau} = \sum_{l=1}^{L}\mathbf{A}_{g,l}\mathbf{W}_{l,k}$ and $\mathbf{Y}_{g,k}^{\tau} = \sum_{l=1}^{L}\mathbf{A}_{g,l}\mathbf{W}_{l,k}$ are linear combinations of $\mathbf{W}_{l,u}$ and $\mathbf{W}_{l,v}$, separately, we can deduce that $\mathrm{Cov}[\mathbf{Y}_{g,u}^{\tau}, \mathbf{Y}_{g,v}^{\tau}] = 0$. Moreover, since $|\mathbf{Y}_{g,u}^{\tau}|^2$ and $|\mathbf{Y}_{g,v}^{\tau}|^2$ are continuous functions of $\mathbf{Y}_{g,u}^{\tau}$ and $\mathbf{Y}_{g,v}^{\tau}$, it can be directly verified that

$$\mathrm{Cov}\left[|\mathbf{Y}_{g,u}^{\tau}|^2, |\mathbf{Y}_{g,v}^{\tau}|^2\right] = 0, \tag{28}$$

where $g = 1, 2, \ldots, G$, and $u, v = 0, 1, \ldots, N-1$, $u \neq v$.

Under $\mathcal{H}_{0,\tau}$, the covariance between $|\mathbf{Y}_{g,k}^{\tau}|^2$ and $|\mathbf{Y}_{g+1,k}^{\tau}|^2$ can be expanded as

$$\begin{aligned} &\mathrm{Cov}\left[|\mathbf{Y}_{g,k}^{\tau}|^2, |\mathbf{Y}_{g+1,k}^{\tau}|^2\right] = \\ &\mathrm{Cov}\left[\left[\Re\left(\mathbf{Y}_{g,k}^{\tau}\right)\right]^2 + \left[\Im\left(\mathbf{Y}_{g,k}^{\tau}\right)\right]^2, \left[\Re\left(\mathbf{Y}_{g+1,k}^{\tau}\right)\right]^2\right. \\ &\left.+\left[\Im\left(\mathbf{Y}_{g+1,k}^{\tau}\right)\right]^2\right] = \mathrm{Cov}\left[\left[\Re\left(\mathbf{Y}_{g,k}^{\tau}\right)\right]^2\left[\Re\left(\mathbf{Y}_{g+1,k}^{\tau}\right)\right]^2\right] \\ &+ \mathrm{Cov}\left[\left[\Re\left(\mathbf{Y}_{g,k}^{\tau}\right)\right]^2\left[\Im\left(\mathbf{Y}_{g+1,k}^{\tau}\right)\right]^2\right] \\ &+ \mathrm{Cov}\left[\left[\Im\left(\mathbf{Y}_{g,k}^{\tau}\right)\right]^2\left[\Re\left(\mathbf{Y}_{g+1,k}^{\tau}\right)\right]^2\right] \\ &+ \mathrm{Cov}\left[\left[\Im\left(\mathbf{Y}_{g,k}^{\tau}\right)\right]^2\left[\Im\left(\mathbf{Y}_{g+1,k}^{\tau}\right)\right]^2\right]. \end{aligned} \tag{29}$$

For further simplification, we first compute the covariance between $\Re(\mathbf{Y}^{\tau}_{g,k})$ and $\Re(\mathbf{Y}^{\tau}_{g+1,k})$

$$\begin{aligned}
&\mathrm{Cov}\left[\Re(\mathbf{Y}^{\tau}_{g,k}), \Re(\mathbf{Y}^{\tau}_{g+1,k})\right] \\
&= E\left[\Re(\mathbf{Y}^{\tau}_{g,k})\Re(\mathbf{Y}^{\tau}_{g+1,k})\right] - E\left[\Re(\mathbf{Y}^{\tau}_{g,k})\right]E\left[\Re(\mathbf{Y}^{\tau}_{g+1,k})\right] \\
&= E\left[\left(\sum_{l=0}^{L}\left(\Re\left(\mathbf{A}^{\tau}_{g,l}\right)\Re\left(\mathbf{W}_{l,k}\right) - \Im\left(\mathbf{A}^{\tau}_{g,l}\right)\Im\left(\mathbf{W}_{l,k}\right)\right)\right)\right. \\
&\left.\times\left(\sum_{l'=0}^{L}\left(\Re\left(\mathbf{A}^{\tau}_{g+1,l'}\right)\Re\left(\mathbf{W}_{l',k}\right) - \Im\left(\mathbf{A}^{\tau}_{g+1,l'}\right)\Im\left(\mathbf{W}_{l',k}\right)\right)\right)\right] \\
&= \frac{LN\sigma_w^2\rho_g^{\tau,2}}{2}. \qquad (30)
\end{aligned}$$

As the variance of $\Re(\mathbf{Y}^{\tau}_{g,k})$ and $\Re(\mathbf{Y}^{\tau}_{g+1,k})$ are $LN\sigma_w^2/2$, we can infer that the correlation coefficient of $\Re(\mathbf{Y}^{\tau}_{g,k})$ and $\Re(\mathbf{Y}^{\tau}_{g+1,k})$ is $\rho_g^{\tau,2}$, and then the correlation coefficient of $[\Re(\mathbf{Y}^{\tau}_{g,k})]^2$ and $[\Re(\mathbf{Y}^{\tau}_{g+1,k})]^2$ is $(\rho_g^{\tau,2})^2$[44]. Thus, the covariance between $[\Re(\mathbf{Y}^{\tau}_{g,k})]^2$ and $[\Re(\mathbf{Y}^{\tau}_{g+1,k})]^2$ is

$$\mathrm{Cov}\left[\left[\Re\left(\mathbf{Y}^{\tau}_{g,k}\right)\right]^2\left[\Re\left(\mathbf{Y}^{\tau}_{g+1,k}\right)\right]^2\right] = \frac{L^2N^2\sigma_w^4}{2}(\rho_g^{\tau,2})^2, \qquad (31)$$

Similarly, we can get

$$\mathrm{Cov}\left[\left[\Re\left(\mathbf{Y}^{\tau}_{g,k}\right)\right]^2\left[\Im\left(\mathbf{Y}^{\tau}_{g+1,k}\right)\right]^2\right] = \frac{L^2N^2\sigma_w^4}{2}(\rho_g^{\tau,0})^2, \quad (32)$$

$$\mathrm{Cov}\left[\left[\Im\left(\mathbf{Y}^{\tau}_{g,k}\right)\right]^2\left[\Re\left(\mathbf{Y}^{\tau}_{g+1,k}\right)\right]^2\right] = \frac{L^2N^2\sigma_w^4}{2}(\rho_g^{\tau,1})^2, \quad (33)$$

$$\mathrm{Cov}\left[\left[\Im\left(\mathbf{Y}^{\tau}_{g,k}\right)\right]^2\left[\Im\left(\mathbf{Y}^{\tau}_{g+1,k}\right)\right]^2\right] = \frac{L^2N^2\sigma_w^4}{2}(\rho_g^{\tau,3})^2. \quad (34)$$

Substituting equations (31)–(34) into equation (29), the equation (29) can be written as

$$\begin{aligned}
&\mathrm{Cov}\left[|\mathbf{Y}^{\tau}_{g,k}|^2, |\mathbf{Y}^{\tau}_{g+1,k}|^2\right] \\
&= \frac{(\rho_g^{\tau,0})^2 + (\rho_g^{\tau,1})^2 + (\rho_g^{\tau,2})^2 + (\rho_g^{\tau,3})^2}{2}L^2N^2\sigma_w^4. \qquad (35)
\end{aligned}$$

For the case $h = g$, $i = g + 1$ and $u \neq v$, in the same process, it is easy to obtain

$$\mathrm{Cov}\left[\left[\Re\left(\mathbf{Y}^{\tau}_{h,u}\right)\right]^2\left[\Re\left(\mathbf{Y}^{\tau}_{i,v}\right)\right]^2\right] = 0, \qquad (36a)$$

$$\mathrm{Cov}\left[\left[\Re\left(\mathbf{Y}^{\tau}_{h,u}\right)\right]^2\left[\Im\left(\mathbf{Y}^{\tau}_{i,v}\right)\right]^2\right] = 0, \qquad (36b)$$

$$\mathrm{Cov}\left[\left[\Im\left(\mathbf{Y}^{\tau}_{h,u}\right)\right]^2\left[\Re\left(\mathbf{Y}^{\tau}_{i,v}\right)\right]^2\right] = 0, \qquad (36c)$$

$$\mathrm{Cov}\left[\left[\Im\left(\mathbf{Y}^{\tau}_{h,u}\right)\right]^2\left[\Im\left(\mathbf{Y}^{\tau}_{i,v}\right)\right]^2\right] = 0, \qquad (36d)$$

which leads to

$$\mathrm{Cov}\left[|\mathbf{Y}^{\tau}_{h,u}|^2, |\mathbf{Y}^{\tau}_{i,v}|^2\right] = 0. \qquad (37)$$

With the result in (28), (35) and (37), we finally obtain (24) in Lemma 1.

REFERENCES

[1] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Select. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[2] Y. C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[3] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: Requirements, challenges and design trade-offs," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 32–39, Apr. 2008.

[4] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.

[5] N. Wang, Y. Gao, and X. Zhang, "Adaptive spectrum sensing algorithm under different primary user utilizations," *IEEE Commun. Lett.*, vol. 17, no. 9, pp. 1838–1841, Sep. 2013.

[6] H. Sun, A. Nallanathan, C. X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 74–81, Apr. 2013.

[7] D. Bao, L. De Vito, and S. Rapuano, "A histogram-based segmentation method for wideband spectrum sensing in cognitive radios," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 7, pp. 1900–1908, Jul. 2013.

[8] Y. Pei, Y. C. Liang, K. C. Teh, and K. H. Li, "How much time is needed for wideband spectrum sensing?" *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5466–5471, Nov. 2009.

[9] R. G. Baraniuk, "More is less: Signal processing and the data deluge," *Science*, vol. 331, no. 6018, pp. 717–719, Feb. 2011.

[10] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[11] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[12] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Honolulu, HI, USA, Apr. 2007, pp. IV-1357–IV-1360.

[13] Z. Qin, Y. Gao, M. D. Plumbley, and C. G. Parini, "Wideband spectrum sensing on real-time signals at sub-Nyquist sampling rates in single and cooperative multiple nodes," *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3106–3117, Jun. 2016.

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

[15] S. Kirolos, T. Ragheb, J. Laska, M. F. Duarte, Y. Massoud, and R. G. Baraniuk, "Practical issues in implementing analog-to-information converters," in *Proc. 6th Int. Workshop Syst. Chip Real Time Appl.*, Cairo, Egypt, 2006, pp. 141–146.

[16] C. P. Yen, Y. Tsai, and X. Wang, "Wideband spectrum sensing based on sub-Nyquist sampling," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3028–3040, Jun. 2013.

[17] S. Ren, Z. Zeng, C. Guo, and X. Sun, "A low complexity sensing algorithm for wideband sparse spectra," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 92–95, Jan. 2017.

[18] H. Sun, W. Y. Chiu, J. Jiang, A. Nallanathan, and H. V. Poor, "Wideband spectrum sensing with sub-Nyquist sampling in cognitive radios," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 6068–6073, Nov. 2012.

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

[20] D. Adams, Y. C. Eldar, and B. Murmann, "A mixer front end for a four-channel modulated wideband converter with 62-dB blocker rejection," *IEEE J. Solid-State Circuits*, vol. 52, no. 5, pp. 1286–1294, May 2017.

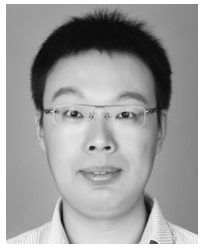
[21] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wideband spectrum sensing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Taipei, Taiwan, Apr. 2009, pp. 2337–2340.

[22] D. D. Ariananda and G. Leus, "Wideband power spectrum sensing using sub-Nyquist sampling," in *Proc. IEEE 12th Int. Workshop Signal Process. Adv. Wireless Commun.*, San Francisco, CA, USA, Jun. 2011, pp. 101–105.

[23] Y. Wang, Z. Tian, and C. Feng, "Sparsity order estimation and its application in compressive spectrum sensing for cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2116–2125, Jun. 2012.

[24] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. Roy. Statist. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.

[25] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Select. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, Dec. 2007.
[26] Z. Qin, Y. Gao, and C. G. Parini, "Data-assisted low complexity compressive spectrum sensing on real-time signals under sub-Nyquist rate," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1174–1185, Feb. 2016.
[27] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.
[28] Y. Ma, Y. Gao, Y. C. Liang, and S. Cui, "Reliable and efficient sub-Nyquist wideband spectrum sensing in cooperative cognitive radio networks," *IEEE J. Select. Areas Commun.*, vol. 34, no. 10, pp. 2750–2762, Oct. 2016.
[29] H. Sun, A. Nallanathan, S. Cui, and C. X. Wang, "Cooperative wideband spectrum sensing over fading channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1382–1394, Mar. 2016.
[30] M. Mishali *et al.*, "Xampling: Analog to digital at sub-Nyquist rates," *IET Circuits, Devices Syst.*, vol. 5, no. 1, pp. 8–20, Jan. 2011.
[31] M. Mishali and Y. C. Eldar, "Wideband spectrum sensing at sub-Nyquist rates [applications corner]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 102–135, Jul. 2011.
[32] D. Cohen and Y. C. Eldar, "Sub-Nyquist sampling for power spectrum sensing in cognitive radios: A unified approach," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3897–3910, Aug. 2014.
[33] D. Arianandа and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4775–4789, Sep. 2012.
[34] M. Mishali and Y. C. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4692–4702, Oct. 2008.
[35] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, Dec. 2006.
[36] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.
[37] D. L. Donoho, Y. Tsaig, I. Drori, and J. L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1094–1121, Feb. 2012.
[38] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[39] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Process.*, vol. 86, no. 3, pp. 572–588, Mar. 2006.
[40] M. Rashidi, K. Haghighi, A. Owrang, and M. Viberg, "A wideband spectrum sensing method for cognitive radio using sub-Nyquist sampling," in *Proc. IEEE Digit. Signal Process. Workshop IEEE Signal Process. Educ. Workshop*, Sedona, AZ, USA, 2011, pp. 30–35.
[41] A. Lavrenko, F. Römer, G. Del Galdo, R. Thomä, and O. Arikan, "An empirical eigenvalue-threshold test for sparsity level estimation from compressed measurements," in *Proc. 22nd Eur. Signal Process. Conf.*, Lisbon, Portugal, 2014, pp. 1761–1765.
[42] H. Wu and S. Wang, "Adaptive sparsity matching pursuit algorithm for sparse reconstruction," *IEEE Signal Process. Lett.*, vol. 19, no. 8, pp. 471–474, Aug. 2012.
[43] A. Lavrenko, F. Römer, G. Del Galdo, and R. S. Thomä, "Sparsity order estimation for sub-Nyquist sampling and recovery of sparse multiband signals," in *Proc. IEEE Int. Conf. Commun.*, London, U.K., 2015, pp. 4907–4912.
[44] A. H. Joarder, "Moments of the product and ratio of two correlated chi-square variables," *Statist. Papers*, vol. 50, no. 3, pp. 581–592, Jun. 2009.
[45] D. V. Hinkley, "On the ratio of two correlated normal random variables," *Biometrika*, vol. 56, no. 3, pp. 635–639, May 1969.
[46] T. Xiong, H. Li, P. Qi, Z. Li, and S. Zheng, "Predecision for wideband spectrum sensing with sub-Nyquist sampling," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6908–6920, Aug. 2017.
[47] A. V. Oppenheim, *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 2010



Peihan Qi was born in Henan, China, in 1986. He received the B.S. degree in telecommunications engineering from Chang'an University, Xi'an, China, in 2006, and the M.S. degree in communication and information system and the Ph.D. degree in military communication from Xidian University, Xi'an, China, in 2011 and 2014, respectively. Since Jan. 2015, he has been an Assistant Professor with the School of Telecommunications Engineering, Xidian University. His research interests include compressed sensing, spectrum sensing in cognitive radio networks, and high-speed digital signal processing.

Zan Li was born in Shaanxi, China, in 1975. She received the B.S. degree in telecommunications engineering and the M.S. and Ph.D. degrees in communication and information system from Xidian University, Xi'an, China. She is currently a Professor with the School of Telecommunications Engineering, Xidian University. Her research interests include wireless communication system, cognitive radio networks, and digital signal processing.

Hongbin Li (M'99–SM'08) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, FL, USA, in 1999, all in electrical engineering. From July 1996 to May 1999, he was a Research Assistant with the Department of Electrical and Computer Engineering, University of Florida. Since July 1999, he has been with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, USA, where he became a Professor in 2010. He was a Summer Visiting Faculty Member at the Air Force Research Laboratory in the summers of 2003, 2004, and 2009. His general research interests include statistical signal processing, wireless communications, and radars. He was the recipient of the IEEE Jack Neubauer Memorial Award in 2013 from the IEEE Vehicular Technology Society, the Outstanding Paper Award from the IEEE AFICON Conference in 2011, the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for Excellence in Research in 2001 from Stevens Institute of Technology, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He has been a Member of the IEEE SPS Signal Processing Theory and Methods Technical Committee (TC) and the IEEE SPS Sensor Array and Multichannel TC, an Associate Editor for *Signal Processing* (Elsevier), the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE SIGNAL PROCESSING LETTERS, and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, as well as a Guest Editor for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING and the *EURASIP Journal on Applied Signal Processing*. He has been involved in various conference organization activities, including serving as a General Co-Chair for the 7th IEEE Sensor Array and Multichannel Signal Processing Workshop, Hoboken, NJ, June 17–20, 2012. He is a Member of Tau Beta Pi and Phi Kappa Phi.

Tianyi Xiong was born in Sichuan, China, in 1989. He received the B.S. degree in optical information science and technology from South China University of Technology, Guangzhou, China, in 2011. He is currently working toward the Ph.D. degree at the School of Telecommunications Engineering, Xidian University, Xi'an, China. His research interests include wireless communication, spectrum sensing, and cognitive radio networks.