

# Spectrum Sensing and Primary User Localization in Cognitive Radio Networks via Sparsity

Lanchao Liu<sup>1</sup>, Zhu Han<sup>1,\*</sup>, Zhiqiang Wu<sup>2</sup>, Lijun Qian<sup>3</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA

<sup>2</sup>Department of Electrical Engineering, Wright State University, Dayton, Ohio, USA

<sup>3</sup>Department of Electrical and Computer Engineering, Prairie View A&M University, Prairie View, TX, USA

## Abstract

The theory of compressive sensing (CS) has been employed to detect available spectrum resource in cognitive radio (CR) networks recently. Capitalizing on the spectrum resource underutilization and spatial sparsity of primary user (PU) locations, CS enables the identification of the unused spectrum bands and PU locations at a low sampling rate. Although CS has been studied in the cooperative spectrum sensing mechanism in which CR nodes work collaboratively to accomplish the spectrum sensing and PU localization task, many important issues remain unsettled. Does the designed compressive spectrum sensing mechanism satisfy the Restricted Isometry Property, which guarantees a successful recovery of the original sparse signal? Can the spectrum sensing results help the localization of PUs? What are the characteristics of localization errors? To answer those questions, we try to justify the applicability of the CS theory to the compressive spectrum sensing framework in this paper, and propose a design of PU localization utilizing the spectrum usage information. The localization error is analyzed by the Cramér-Rao lower bound, which can be exploited to improve the localization performance. Detail analysis and simulations are presented to support the claims and demonstrate the efficacy and efficiency of the proposed mechanism.

Received on 30 September 2013; accepted on 14 November 2013; published on 11 April 2014

**Keywords:** Compressive sensing, decentralized spectrum sensing, localization, cognitive radio networks.

Copyright © 2014 L. Liu *et al.*, licensed to ICST. This is an open access article distributed under the terms of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/ws.1.1.e2

## 1. Introduction

Today, with rapid developments of wireless communication technologies and emergence of new wireless services, radio spectrum resource are facing the crisis of "spectrum drought" [1]. On the other hand, the spectrum resource are underutilized because of the current fixed spectrum allocation policy. Due to its capability of improving network spectrum utilization efficiency, cognitive radio (CR) [2], which enables the dynamical use of licensed spectrum bands from licensed/primary users (PUs)[3], has become a promising technique to resolve the dilemma of spectrum resource shortage versus underutilization.

The success of CR networks relies on the spectrum sensing technique [4, 5] that detects the unoccupied spectrum bands for opportunistic spectrum access. In CR networks, the CR nodes sense the spectrum and the locally collected measurements can be used collaboratively to detect the unused spectrum bands. These cooperative approaches significantly enhance the sensing reliability and efficiency. Traditionally, the CR nodes used to sense one band at a time, which results in a long sensing delay and high computational complexity in order to have the full picture of spectrums. Recent advances in compressive sensing (CS)[8, 9] make it possible to obtain the wide-band signals at a sub-Nyquist sampling rate by taking random projections, which can relax the analog-to-digital (ADC) requirements as well as shorten the sensing time. The reduced number of measurements also saves the energy for communication and reduces the signalling for the CR nodes. Moreover, the

\*Please ensure that you use the most up to date class file, available from ICST at <http://icst.org/icst-transactions/>

\*Corresponding author. Email: zhan2@uh.edu

proliferation of CR devices has fostered the demand for lots of context-aware applications, in which the location information of PUs is viewed as one significant context. As a byproduct, the compressive spectrum sensing scheme exploits the spatial sparsity of PU locations to estimate their locations, which facilitates the spatial frequency reuse and improves spectrum resource utilization.

Despite many benefits of compressive spectrum sensing, several significant issues in this mechanism have not been fully addressed. Firstly, in order to guarantee a successful recovery of the original sparse signal, the CS sensing/sampling matrix is required to have certain properties. In particular, for the recovery to be robust to sensing noises, the matrix needs to satisfy the restricted isometry property (RIP) [10]. Does the adopted sampling mechanism in compressive spectrum sensing possess this property? Secondly, the scheme of localization via spatial sparsity imposes a requirement on the (spatial) incoherence of the channel gain matrix, which is unnecessary for the purpose of spectrum sensing. Instead of estimating the spectrum occupation and PU locations at the same time, can the two processes be separated, and can the spectrum sensing results be exploited to reduce the requirement on the channel gain matrix that is needed for the stable recovery of PU locations? Finally, what are the error characteristics of the localization errors in this mechanism, and how can we utilize those characteristics in the deployed system to control the errors? To guarantee a stable performance, all those issues in the compressive spectrum sensing scheme should be carefully discussed and analyzed.

In this paper, while we try to answer the above questions, the framework of compressive spectrum sensing is redesigned and a novel mechanism is proposed to accomplish the spectrum sensing and PUs localization tasks. Specifically, a compressive sampling mechanism is utilized by each CR node to take the measurements of the received signal spectrum. By exploiting the spectrum resource underutilization, the CS techniques can acquire a sparse signal at a sampling rate much lower than the Nyquist rate by taking a small set of random linear projections. The measurements collected at the CR nodes are processed in a collaborative way to reconstruct the original sparse signal, and we propose algorithms based on the alternating direction method of multipliers [6, 7] to recover the received signal spectrum. Then, we utilize the spectrum sensing results to localize the PUs by a Bayesian inference scheme and improve the algorithm based on analyzing the Cramer-Rao lower bound (CRLB). The main contributions of this paper are listed as follows:

1. The RIP of the cooperative compressive spectrum sensing mechanism is justified mathematically, which guarantees a successful reconstruction of the original sparse signal with high probability. Three different algorithms, Lasso, group Lasso, and distributed group Lasso with feature splitting are used to recover the spectrum usage information. Compared to Lasso, group Lasso is an improved centralized algorithm exploiting joint sparsity, and distributed group Lasso with feature splitting is an approach suitable for distributed implementation. Hence, we validate the applicability of CS to spectrum sensing both theoretically and numerically.
2. The two tasks – spectrum sensing and PU localization – are done one after another. Instead of recovering the transmitted signal spectrum at PUs, the signal spectrum at CR nodes is recovered. After the recovery, the results are used for the localization of PUs. The benefits of this treatment are two-fold: the incoherence requirement of the channel gain matrix required for stable spectrum and PU location recoveries is relaxed, and the spectrum sensing results can be exploited to reduce the computation complexity and improve the performance of the localization step.
3. The error characteristics of localization are analyzed by calculating the Cramér-Rao lower bound, which is a lower bound on the covariance of any unbiased location estimate. We analyze the localization error characteristics in detail and improve the localization algorithm design by exploiting the revealed error trends associated with the deployed system.

The rest of this work is organized as follows: Section 2 discusses the related work. The system model and problem formulation are given Section 3. The proposed approaches for spectrum sensing and PU location are detailed in Section 4 and Section 5, respectively. Section 6 presents the simulation results, and finally Section 7 concludes this paper.

## 2. Related Work

Closely related to this paper is the existing work on compressive spectrum sensing and PU localization taking advantages of sparsity and CS. CS has been first proposed in wideband CR in [11] for spectrum hole identification. Exploiting the sparsity of the signal spectrum, CS samples the wideband signal at a sub-Nyquist sampling rate, and thus, significantly reduces the ADC cost. [12] adopts CS in combination with decentralized recovery algorithms in cooperative CR networks, which the spectral estimates from all CR

nodes are fused together to determine the spectrum availability. The spatial diversity gain induced by the distributed CR nodes greatly improves the sensing performance under fading channel environments.

Recently, CS has been used to solve the spectrum sensing problem in cooperative CR networks in [13–16]. In [13], the CS technique has been adopted for sampling wideband channel information at a sub-Nyquist rate at CR nodes. A small portion of the compressed measurements of received signals spectrum are reported to the fusion center, which recovers the spectrum usage information by matrix completion or joint sparsity reconstruction. In this centralized approach, the sampling rate at CR nodes is reduced, which saves time and energy for spectrum sensing. Also, the significantly compressed measurements save the energy consumed by transmission from the CR nodes to fusion center. [14] uses a distributed computing approach instead. Each CR node obtains its compressed individual samplings of received signal spectrum, and only communicates with the neighboring CR nodes with the one-hop range. A consensus constraint is imposed on each CR to ensure convergence, and the global convergence can be obtained by several rounds of information exchange. Upon convergence, all CR nodes obtain the fused spectrum sensing results. This approach gets rid of the fusion center at an expense of a certain amount of communication overhead. The locations of PUs are taken into consideration in [16], which exploits both the spectrum sparsity and PUs' spatial sparsity. A basis expansion model of the power spectral density (PSD) map in space and frequency is utilized to model the transmitted signals in CR networks. Each CR node collects the attenuated transmitted signals by the channel and work collaboratively to estimate the signal spectrum. The PSD of the transmitted signal is estimated by the Lasso [17] algorithm implemented via distributed online iterations, which can sense the spectrum occupation as well as recover the PU locations in the network. This kind of localization via spatial sparsity is formally defined in [18], and explicitly uses the CS theory in the localization algorithm. To guarantee a successful recovery of the source locations in the network, the requirement on the incoherence of the channel gain matrix is given therein. [19] further develops [18], which introduces a Bayesian framework for the localization problem and provides sparse approximations to its optimal solution.

Despite the early emergence and wide range of applications of the CS theory to compressive spectrum sensing, the applicability of the CS theory to this application scenario has not been rigorously justified yet. For example, the channel gain matrix is required to be sufficiently incoherent for reliable PU localization, but this requirement is often unexamined.

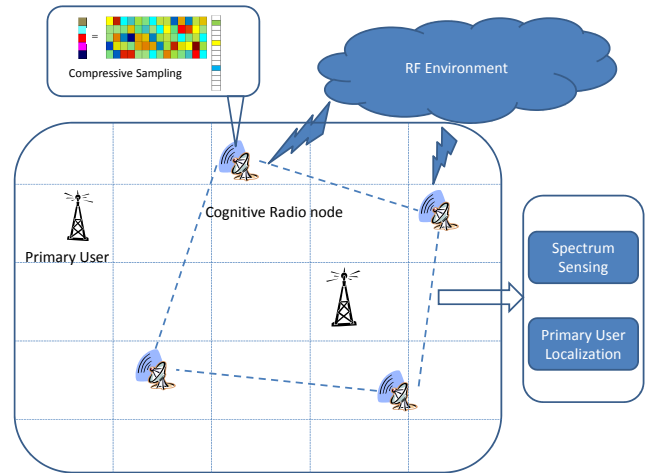


Figure 1. An illustration of system model

Furthermore, the characteristics of the localization errors have not been explicitly analyzed before. All these issues are the motivations for this paper.

### 3. System Model and Problem Formulation

Consider a CR network with  $K$  PUs and  $J$  CR nodes. The CR nodes locate randomly in the network with known locations  $\beta = \{\beta_1, \beta_2, \dots, \beta_J\}$ . However, for PUs, neither the number of the active PUs  $K$  nor their locations are known. We define a finite position space  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_I\}$  as the candidate positions of the PUs in the planar deployment area. Note that the PUs locate sparsely in the network and only present at  $K$  out of  $I$  positions. Each CR node in the network receives a superposition of the transmitted signals of the PUs, and our objective is to infer the spectrum occupation and locations of the PUs based on the received signals. An illustration of the system model is shown in Fig. 1.

In our model, we adopt a slotted frequency segmentation model to describe the PSD of the transmitted signal. The whole bandwidth is divided into  $B$  non-overlapping narrowband slots centered at  $f_b (b = 1, \dots, B)$  with full-width  $W$ . Suppose the PUs transmitted signals  $x_i(t)$  are stationary over  $t$  and mutually uncorrelated. Here  $i$  means candidate position  $\alpha_i$ . The PSD of the transmitted signal  $x_i(t)$  for the PU at position  $\alpha_i$  can be approximated by:

$$s_i(f) = \sum_{b=1}^B \theta_{i,b} \text{rect} \left( \frac{f - f_b}{W} \right), \quad i = 1, 2, \dots, I, \quad (1)$$

where  $\theta_{i,b}$  is the PSD for the corresponding narrowband slot centered at  $f_b$ . Apparently, the locations of those bases are known in our model, but the PSD levels are dynamic varying and need to be determined. If the PSD levels for certain bases are under a predefined threshold, those temporarily idle frequency bands are

called spectral holes and can be opportunistically accessed by the CR nodes. We use vector  $\mathbf{s}_i = [\theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,B}]$  to represent the PSD for PU at  $\alpha_i$ . Note that only  $K$  out of  $I$  positions have PUs and many of  $\mathbf{s}_i$  are all-zero vectors.

The transmitted signals are attenuated by the multipath channels between the PUs and CR nodes. The received signal of the CR node at  $\beta_j$  is a summation of amplitudes and delays of the multiple arriving signals from all PUs, and can be expressed as:

$$y_j(t) = \sum_{i=1}^I h(d_{i,j}; t) * x_i(t) + w_j(t), \quad j = 1, 2, \dots, J, \quad (2)$$

where  $*$  represents the convolution,  $h(d_{i,j}; t)$  is the channel response that is a function of distance  $d_{i,j}$  between  $\alpha_i$  and  $\beta_j$ , and  $w_j(t)$  is the noise. We assume that the signals propagate in a frequency-selective fading channel, and the channel responses are stationary w.r.t.  $t$ . Hence, we can calculate the PSD of the channel response by taking the Fourier transform of the autocorrelation of  $h(d_{i,j}; t)$ , i.e.,  $H_{i,j}(f) = \mathcal{F}(\mathbf{R}(\tau))$ , where  $\mathbf{R}(\tau) = \mathbf{E}[h(d_{i,j}; t)h(d_{i,j}; t - \tau)]$ . The channel gain  $H_{i,j}(f)$  can be obtained through extensive measurement campaigns or through the path loss model<sup>1</sup>.

At the receiver end, the received PSD can be expressed as the product of the PSD  $s_i(f)$  of the transmitted signal and channel gain  $H_{i,j}(f)$ :

$$\begin{aligned} r_j(f) &= \sum_{i=1}^I H_{i,j}(f)s_i(f) + \sigma_j(f) \\ &= \sum_{i=1}^I H_{i,j}(f) \sum_{b=1}^B \theta_{i,b} \text{rect}\left(\frac{f - f_b}{W}\right) + \sigma_j(f), \end{aligned} \quad (3)$$

where  $r_j(f)$  is the PSD of the CR nodes at position  $\beta_j$ , and can be estimated by the Fourier Transform of the received time domain signal  $y_j(t)$  traditionally. However, since the transmitted bands occupied by the PUs are quite narrow compared with the whole available band, we can exploit the spectral sparsity and utilize CS to effectively sample the signal.

To clarify this point, we first simplify the expression of (3) to a matrix-vector form:

$$\mathbf{r}_j = \mathbf{H}_{\alpha \rightarrow \beta_j} \mathbf{s} + \sigma_j, \quad j = 1, 2, \dots, J, \quad (4)$$

where vector  $\mathbf{r}_j$  denotes the received PSD of CR node at  $\beta_j$ , and  $\mathbf{s} = [\mathbf{s}_1^\top, \mathbf{s}_2^\top, \dots, \mathbf{s}_I^\top]^\top$  denotes the transmitted PSD at all  $I$  candidate points. Mapping operator  $\mathbf{H}_{\alpha \rightarrow \beta_j}$

stands for the corresponding channel gain matrix for all sub-bands between  $I$  candidate positions and CR node at  $\beta_j$ . The complete channel gain matrix from all candidate positions to all CR nodes can be written as  $\mathbf{H}_{\alpha \rightarrow \beta}$ .

By utilizing the CS theory, the compressed measurements at the CR node can be expressed as follows:

$$\mathbf{z}_j = \mathbf{\Phi}_j \mathcal{F}^{-1} \mathbf{r}_j + \tilde{\sigma}_j = \mathbf{\Phi}_j \mathcal{F}^{-1} \mathbf{H}_{\alpha \rightarrow \beta_j} \mathbf{s} + \tilde{\sigma}_j, \quad \forall j, \quad (5)$$

where matrix  $\mathbf{\Phi}_j \in \mathbb{R}^{M \times B}$ ,  $M \ll B$ . Matrix  $\mathbf{\Phi}_j$  can be designed sampling from a certain distribution, e.g., i.i.d. Gaussian distribution or i.i.d. Bernoulli distribution with  $\pm 1$ . This matrix can be implemented by the techniques used in CS for analog signals like Xampling [20, 21], random modulator [22] and frequency-selective surface [11]. The benefits of utilizing matrix  $\mathbf{\Phi}_j$  are two-fold: Firstly, it reduces the number of measurements of the CR nodes, which saves time and energy for spectrum sensing; Secondly, the reduced measurements obtained by the random projection are quite compressed compared with the original time domain signals. Since all CR nodes need to transmit the measurements back to the fusion center or exchange them with neighbors for spectrum sensing and localization, the reduction in the number of measurements saves the energy consumed by the transmission and signalling for the CR nodes.

We remark that the spectrum sensing problem is recover  $\{\mathbf{r}_j\}_{j=1}^J$  from  $\{\mathbf{z}_j\}_{j=1}^J$  for all CR nodes, and can be solved without the channel state information (CSI)  $\mathbf{H}_{\alpha \rightarrow \beta}$ . As long as the channel does not experience deep fadings, the received signals PSD  $\{\mathbf{r}_j\}_{j=1}^J$  provide sufficient information for spectrum usage. Moreover, note that the received signals PSD  $\{\mathbf{r}_j\}_{j=1}^J$  are also sparse. To clarify this point, we assume  $\tilde{\mathbf{s}} = \sum_{i=1}^I \mathbf{s}_i$ . The operation of summation neglects the location information of PUs and  $\tilde{\mathbf{s}}$  only contains the spectrum occupation information of transmitted signals. For each  $\mathbf{r}_j$  and  $\tilde{\mathbf{s}}$ , their supports, i.e., the positions of non-zero entries, are the same. The physical explanation is that every  $\mathbf{r}_j$  is only the attenuated duplicates of  $\tilde{\mathbf{s}}$  by channel effects. By exploiting these key observations, the received signals PSD are acquired by a small number of incoherent linear measurements in (5). As to localization, a Bayesian inference scheme can be adopted to localize the PUs from  $\{\mathbf{r}_j\}_{j=1}^J$ .

## 4. Compressive Spectrum Sensing

The spectrum sensing task is completed by recovering the received signals PSD, which can be implemented in two ways, centralized or distributed. An overview of CS is given first, and then the proposed algorithms is illustrated in detail.

<sup>1</sup> For a fading channel, the channel gain can be calculated by  $H_{i,j}(f) = d_{i,j}^{-\gamma/2} |h_{i,j}(f)|$ , where  $h_{i,j}(f)$  is the channel fading gain that can be obtained by averaging out the effect of the channel fading and  $\gamma$  is the path loss factor.

#### 4.1. Compressive Sensing Overview

CS is an emerging signal processing technique for finding sparse solutions to under-determined linear systems. By utilizing the fact that a signal is sparse or compressible in a certain domain, the CS technique can powerfully acquire a signal from a small set of randomly projected measurements with a very low sampling rate. Suppose  $\mathbf{x}$  is an  $N \times 1$  vector with  $k$  nonzero entries, and  $\mathbf{y}$  is an  $M \times 1$  vector that  $k < M \ll N$ . If sensing matrix  $\Phi$  satisfies the RIP, the solution of the noisy under-determined system of equations  $\mathbf{y} = \Phi\mathbf{x} + \mathbf{w}$  can be reconstructed by

$$\min_x \|\mathbf{x}\|_1 + \lambda \|\mathbf{y} - \Phi\mathbf{x}\|_2^2, \quad (6)$$

where  $\|\cdot\|_1$  and  $\|\cdot\|_2$  are the  $l_1$  norm and  $l_2$  norm, respectively. The parameter  $\lambda$  is chosen according to the amount of noise in the measurements. The optimization problem in (6) is known as the least-absolute shrinkage and selection operator (Lasso) [17], a.k.a. the de-nosing basis pursuit for solving a sparse linear regression problem. A detail review of traditional and advanced methods to solve (6) can be found in [23] and the references therein. In this paper, we use the ADMM [24, 25] to solve (6). In particular, to be robust against noise, we require the sensing matrix to satisfy the RIP, which is defined as follows:

**Definition 1.** (restricted isometry property) [10] Let  $\Phi \in \mathbb{R}^{m \times n}$  matrix having unit  $l_2$ -norm columns. For each integer  $S \in \mathbb{N}$ , we say that  $\Phi$  satisfies the restricted isometry property of order  $K$  with the smallest restricted isometry constant  $\delta_K \in (0, 1)$ , and write  $\Phi \in \text{RIP}(K, \delta_K)$ , if

$$(1 - \delta_K)\|\theta\|_2^2 \leq \|\Phi\theta\|_2^2 \leq (1 + \delta_K)\|\theta\|_2^2, \quad \forall \theta : \|\theta\|_0 \leq K, \quad (7)$$

where  $\|\theta\|_0$  is the number of nonzero elements in  $\theta$ .

The same framework applies if signal  $\mathbf{x}$  is sparse in some transformed domain  $\Psi$  instead of the canonical domain. In this case,  $\mathbf{y} = \Phi\Psi\mathbf{x} + \mathbf{w}$ , and we recover  $\mathbf{x}$  from  $\mathbf{y}$  by:

$$\min_x \|\mathbf{x}\|_1 + \lambda \|\mathbf{y} - \Phi\Psi\mathbf{x}\|_2^2, \quad (8)$$

where  $\Phi\Psi$  is the sensing matrix and satisfy RIP if  $\Psi$  and  $\Phi$  are incoherent.

#### 4.2. A "Naive" Approach

The compressed measurements  $\{\mathbf{z}_j\}_{j=1}^J$  are cooperatively used to recover the spectral estimates  $\{\mathbf{r}_j\}_{j=1}^J$  of the received signals at all CR nodes. According to (5),  $\mathbf{z}_j = \Phi_j \mathcal{F}^{-1} \mathbf{r}_j + \tilde{\sigma}_j = \tilde{\Phi}_j \mathbf{r}_j + \tilde{\sigma}_j$ . When a fusion center is presented to collect all the compressed measurements  $\{\mathbf{z}_j\}_{j=1}^J$ , the fused measurements can be concatenated

together to form a new vector  $\mathbf{z} = [\mathbf{z}_1^\top, \mathbf{z}_2^\top, \dots, \mathbf{z}_J^\top]^\top$ . Conformably, the received PSD of CR nodes can be written as  $\mathbf{r} = [\mathbf{r}_1^\top, \mathbf{r}_2^\top, \dots, \mathbf{r}_J^\top]^\top$  and the sensing matrix is

$$\tilde{\Phi} = \begin{bmatrix} \tilde{\Phi}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \tilde{\Phi}_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \tilde{\Phi}_J \end{bmatrix}. \quad (9)$$

Here, we assume all CR nodes use the same random sampling matrix  $\{\Phi_j\}_{j=1}^J$ , denoted as  $\Phi$ , sampled from the i.i.d. Gaussian distribution, and thus,  $\Phi$  satisfies the RIP. Since the inverse Fourier transform matrix  $\mathcal{F}^{-1}$  is orthonormal, the RIP of  $\{\Phi_j\}_{j=1}^J$  (denoted as  $\tilde{\Phi}$ ) is unaffected. Formula (9) can be write as  $\tilde{\Phi} = \mathbf{I}_J \otimes \tilde{\Phi}$ .  $\otimes$  is the Kronecker product and  $\mathbf{I}_J$  is a  $J \times J$  identity matrix. The overall observation model can be expressed as:

$$\mathbf{z} = \tilde{\Phi}\mathbf{r} + \tilde{\sigma}, \quad (10)$$

where  $\tilde{\sigma} = [\tilde{\sigma}_1^\top, \tilde{\sigma}_2^\top, \dots, \tilde{\sigma}_J^\top]^\top$ .  $\mathbf{r}$  can be solved by:

$$\min_r \|\mathbf{r}\|_1 + \lambda \|\mathbf{z} - \tilde{\Phi}\mathbf{r}\|_2^2. \quad (11)$$

To guarantee a successful recovery of signal  $\mathbf{r}$ , the overall sensing matrix  $\tilde{\Phi}$  must satisfy the RIP, which can be justified by the following theorem.

**Theorem 1.** Let  $\mathbf{A}$  be an orthonormal basis, and let  $\mathbf{B}$  be the matrix with restricted isometry constant  $\delta_K(\mathbf{B})$ . The restricted isometry constant of the Kronecker product of  $\mathbf{A}$  and  $\mathbf{B}$  satisfies:

$$\delta_K(\mathbf{A} \otimes \mathbf{B}) = \delta_K(\mathbf{B}), \quad (12)$$

This theorem provides conservation of restricted isometry constants across the Kronecker product of an orthonormal matrix and a sensing matrix. The proof is given in Appendix A. For our case,  $\delta_K(\tilde{\Phi}) = \delta_K(\mathbf{I}_J \otimes \tilde{\Phi}) = \delta_K(\tilde{\Phi})$ , which implies that sensing matrix  $\tilde{\Phi}$  satisfies RIP.

The centralized fusion (11) results in the global optimal solution by incorporating the measurements from all the CR nodes. This approach to recover the PSD of the received signal is straight-forward and we use it as a simple illustration. To solve (11), we first rewrite it as:

$$\min_{\mathbf{r}, \mathbf{p}} \|\mathbf{p}\|_1 + \lambda \|\mathbf{z} - \tilde{\Phi}\mathbf{r}\|_2^2, \quad \text{s.t.} \quad \mathbf{r} - \mathbf{p} = \mathbf{0}. \quad (13)$$

The scaled augmented Lagrangian can be written as:

$$\mathcal{L}_\rho(\mathbf{r}, \mathbf{p}, \mathbf{u}) = \|\mathbf{p}\|_1 + \lambda \|\mathbf{z} - \tilde{\Phi}\mathbf{r}\|_2^2 + \rho/2 \|\mathbf{r} - \mathbf{p} + \mathbf{u}\|_2^2. \quad (14)$$

The objective function and variables are split into two parts. The separated variables  $\mathbf{p}$  and  $\mathbf{r}$  can be updated

in an alternating fashion as follows:

$$\mathbf{r}^{k+1} = \arg \min_{\mathbf{r}} (\lambda \|\mathbf{z} - \tilde{\Phi} \mathbf{r}\|_2^2 + \rho/2 \|\mathbf{r} - \mathbf{p}^k + \mathbf{u}^k\|_2^2), \quad (15)$$

$$\mathbf{p}^{k+1} = \arg \min_{\mathbf{p}} (\|\mathbf{p}\|_1 + \rho/2 \|\mathbf{r}^{k+1} - \mathbf{p} + \mathbf{u}^k\|_2^2), \quad (16)$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{r}^{k+1} - \mathbf{p}^{k+1}. \quad (17)$$

The  $\mathbf{r}$ -update process is a quadratically regularized least-square problem, and the  $\mathbf{p}$ -update can be solved in closed form by soft-thresholding. In this case, the fusion center collects the compressed measurements  $\{\mathbf{z}_j\}_{j=1}^J$  from all the CR nodes. The compressed measurements are computed centrally to recover the received signals PSD  $\{\mathbf{r}_j\}_{j=1}^J$ , and thus, the unoccupied spectrum bands are determined.

### 4.3. Spectrum Sensing via Joint Sparsity

Unlike the approach above, a more sophisticated approach further exploits the structure of the received signals PSD  $\{\mathbf{r}_j\}_{j=1}^J$  is proposed in the sequel. Since the signal ensembles  $\{\mathbf{r}_j\}_{j=1}^J$  are the received signals PSD at all the CR nodes, the sparse structures of  $\{\mathbf{r}_j\}_{j=1}^J$  are identical to  $\tilde{\mathbf{s}}$ . In other words, the occupied frequency bands of the transmitted signals and received signals are the same as long as the channel does not experience deep fading. Note that  $\{\mathbf{r}_j\}_{j=1}^J$  are the attenuated duplicates of  $\tilde{\mathbf{s}}$  and share a common support. This kind of jointly sparse signals can be solved by the proposed algorithm as follows,

$$\min \sum_{b=1}^B \|\tilde{\mathbf{r}}_b\|_2 + \lambda \|\mathbf{z} - \tilde{\Phi} \mathbf{r}\|_2^2, \quad (18)$$

where  $\tilde{\mathbf{r}}_b = \mathbf{R}(b, :)$ ,  $\mathbf{R}$  is a matrix formed by  $\{\mathbf{r}_j\}_{j=1}^J$ , and  $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_J]$ . The solution to (18) has a natural grouping of its components, and the components within the group are likely to be either all zeros or all non-zeros. Each  $\tilde{\mathbf{r}}_b$  corresponds to the received signals PSD in the  $b^{\text{th}}$  frequency band. We remark that the model in (18) incorporating joint sparsity outperforms that in (11), which recovers the spectra separately. In (18), the first term tends to select the same sparse structure of the received signals PSD, and thus, it can achieve a fast convergence and high accuracy compared with (11).

To solve (18), we first rewrite it as:

$$\begin{aligned} & \min \sum_{b=1}^B \|\tilde{\mathbf{r}}_b\|_2 + \lambda \|\mathbf{z} - \tilde{\Phi} \mathbf{p}\|_2^2 \\ & \text{s.t. } \tilde{\mathbf{r}}_b - \tilde{\mathbf{p}}_b = \mathbf{0}, \quad b = 1, 2, \dots, B, \end{aligned} \quad (19)$$

where  $\{\tilde{\mathbf{r}}_b\}_{b=1}^B$  are the received signals PSD at each frequency bands, and  $\mathbf{p}$  is the globe variable cascading

all received signals PSD. In this case,  $\tilde{\mathbf{p}}_b$  is the corresponding part of  $\mathbf{r}_b$  in  $\mathbf{p}$ . To solve (19), we first write down the scaled augmented Lagrangian:

$$\mathcal{L}_\rho(\tilde{\mathbf{r}}_b, \mathbf{p}, \mathbf{u}) = \sum_{b=1}^B (\|\tilde{\mathbf{r}}_b\|_2 + \rho/2 \|\tilde{\mathbf{r}}_b - \tilde{\mathbf{p}}_b + \mathbf{u}_b\|_2^2) + \lambda \|\mathbf{z} - \tilde{\Phi} \mathbf{p}\|_2^2. \quad (20)$$

Similarity, the separate variables can be updated in an alternating fashion:

$$\tilde{\mathbf{r}}_b^{k+1} = \arg \min_{\tilde{\mathbf{r}}_b} (\|\tilde{\mathbf{r}}_b\|_2 + \rho/2 \|\tilde{\mathbf{r}}_b - \tilde{\mathbf{p}}_b^k + \mathbf{u}_b^k\|_2^2), \quad (21)$$

$$\mathbf{p}^{k+1} = \arg \min_{\mathbf{p}} \left( \sum_{b=1}^B \rho/2 \|\tilde{\mathbf{r}}_b - \tilde{\mathbf{p}}_b^{k+1} + \mathbf{u}_b^k\|_2^2 + \lambda \|\mathbf{z} - \tilde{\Phi} \mathbf{p}\|_2^2 \right), \quad (22)$$

$$\mathbf{u}_b^{k+1} = \mathbf{u}_b^k + \tilde{\mathbf{r}}_b^{k+1} - \tilde{\mathbf{p}}_b^{k+1}. \quad (23)$$

The update rule is similar to the above with the  $\tilde{\mathbf{r}}_b$ -update replaced by block soft-thresholding. This approach exploits the structure of joint sparsity which further reduces the number of measurements and improves recovery quality. Once  $\{\mathbf{r}_b\}_{b=1}^B$  are recovered, the availability of the spectrum resource in the network is determined.

### 4.4. Distributed Spectrum Sensing via Joint Sparsity

The centralized approaches are costly and impractical to implement when a fusion center is absent. To address this situation, a distributed spectrum sensing algorithm is proposed in the sequel.

Note that  $\tilde{\Phi}$  is a block diagonal matrix, and the second term of (18) can be split into  $J$  least squares terms corresponding to the measurements from all CR nodes. The optimization problem can be equivalently expressed as:

$$\min \sum_{b=1}^B \|\tilde{\mathbf{r}}_b\|_2 + \sum_{j=1}^J \lambda_j \|\mathbf{z}_j - \tilde{\Phi} \mathbf{r}_j\|_2^2. \quad (24)$$

The second least-square term is naturally separable across  $j$ . At the same time, the first term enforcing group sparsity of the received signals PSD and couples all CR nodes. Assume the CR nodes are connected to each other in the network. In this case, the spectrum occupation can be estimated in a cooperative way by the CR nodes through one-hop communication. To design a decentralized implementation of (24), we first define a function  $f(\mathbf{R}) = \sum_{b=1}^B \|\tilde{\mathbf{r}}_b\|_2$ , which calculates the sum of row vector norm of matrix  $\mathbf{R}$ . The formula in (24) can be rewritten as:

$$\min \sum_{j=1}^J (f(\mathbf{R}) + \lambda_j \|\mathbf{z}_j - \tilde{\Phi} \mathbf{r}_j\|_2^2). \quad (25)$$

To distributively solve the above problem,  $J$  identical copies  $\{\mathbf{R}_j\}_{j=1}^J$  of  $\mathbf{R}$  are induced and the above problem can be reformulated as:

$$\begin{aligned} \min \sum_{j=1}^J (f(\mathbf{R}_j) + \lambda_j \|\mathbf{z}_j - \tilde{\Phi} \mathbf{p}_j\|_2^2) \\ \text{s.t. } \mathbf{p}_j - \mathbf{R}_j(:, j) = 0, \quad \mathbf{R} - \mathbf{R}_j = 0, \quad j = 1, 2, \dots, J. \end{aligned} \quad (26)$$

The basic idea of this approach is to update the received signals PSD at each CR nodes locally, and a synchronized update procedure on the coupled variable  $\mathbf{R}$  is performed by using the consensus averaging technique. Specifically, the variables  $\{\mathbf{R}_j\}$ ,  $\{\mathbf{p}_j\}$  and  $\{\mathbf{R}\}$  can be grouped into two sets, one set is  $\{\mathbf{R}_j\}$  and the other one is  $(\{\mathbf{p}_j\}, \{\mathbf{R}\})$ . After separation, we can update the two sets in an alternative fashion. When given  $\{\mathbf{p}_j\}$  and  $\{\mathbf{R}\}$ , the updates of  $\mathbf{R}_j$  are separable. When given  $\{\mathbf{R}_j\}$ , we can update  $(\{\mathbf{p}_j\}, \{\mathbf{R}\})$  separatively. The scaled augmented Lagrangian of this problem is given as follows:

$$\begin{aligned} \mathcal{L}(\mathbf{p}_j, \mathbf{R}, \mathbf{R}_j, \mathbf{u}_{1j}, \mathbf{u}_{2j}) = \sum_{j=1}^J (f(\mathbf{R}_j) + \lambda_j \|\mathbf{z}_j - \tilde{\Phi} \mathbf{p}_j\|_2^2 \\ + (\rho_1/2) \|\mathbf{p}_j - \mathbf{R}_j(:, j) + \mathbf{u}_{1j}\|_2^2 \\ + (\rho_2/2) \|\mathbf{R}(:, j) - \mathbf{R}_j(:, j) + \mathbf{u}_{2j}\|_2^2), \end{aligned} \quad (27)$$

During the iterative updating procedure, each CR node maintains a copy  $\mathbf{R}_j$  of the global variable  $\mathbf{R}$  and its local variable  $\mathbf{p}_j$ . At a specific iteration step,  $\mathbf{R}_j$  is updated locally by each CR node:

$$\begin{aligned} \mathbf{R}_j^{k+1} = \arg \min f(\mathbf{R}_j) + (\rho_1/2) \|\mathbf{p}_j^k - \mathbf{R}_j(:, j) + \mathbf{u}_{1j}^k\|_2^2 \\ + (\rho_2/2) \|\mathbf{R}(:, j) - \mathbf{R}_j(:, j) + \mathbf{u}_{2j}^k\|_2^2, \end{aligned} \quad (28)$$

where  $\mathbf{R}$  is the consensus average of the  $\mathbf{R}_j$  from all the CR nodes, which can be obtained by the one-hop communication between a CR node and its neighbors. The global variable  $\mathbf{R}$  is updated in a synchronous fashion and disseminated throughout the network to all CR nodes, as follows:

$$\mathbf{R}^{k+1} = \arg \min (1/J) \sum_{j=1}^J (\mathbf{R}_j(:, j)^{k+1} + \mathbf{u}_{2j}^k), \quad (29)$$

and the local received signal PSD  $\mathbf{p}_j$  is updated locally at each CR node by calculating

$$\begin{aligned} \mathbf{p}_j^{k+1} = \arg \min \lambda_j \|\mathbf{z}_j - \tilde{\Phi} \mathbf{p}_j\|_2^2 \\ + (\rho_1/2) \|\mathbf{p}_j - \mathbf{R}_j(:, j)^{k+1} + \mathbf{u}_{1j}^k\|_2^2. \end{aligned} \quad (30)$$

The parameters are updated locally at the end of each iteration step as follows:

$$\begin{aligned} \mathbf{u}_{1j}^{k+1} &= \mathbf{u}_{1j}^k + \mathbf{p}_j^{k+1} - \mathbf{R}_j(:, j)^{k+1}, \\ \mathbf{u}_{2j}^{k+1} &= \mathbf{u}_{2j}^k + \mathbf{R}(:, j)^{k+1} - \mathbf{R}_j(:, j)^{k+1}. \end{aligned} \quad (31)$$

Upon the convergence, every CR nodes in the network obtain the same global variable  $\mathbf{R}$  and its own local received signal PSD. Hence, the spectrum occupancy is determined locally at each CR node.

## 5. Primary User Localization

This section illustrates the PU localization mechanism. The spectrum sensing results are utilized to relax the incoherence requirement on  $\mathbf{H}_{\alpha \rightarrow \beta}$ , and then the CRLB is provided to give a lower bound of the errors of the location estimates and to improve the performance of the proposed algorithm.

### 5.1. Incoherence Requirement on $\mathbf{H}_{\alpha \rightarrow \beta}$

Before analyzing the incoherence requirement on channel gain matrix  $\mathbf{H}_{\alpha \rightarrow \beta}$ , we first recall the definition of the coherence of a matrix as follows:

**Definition 2.** (Coherence)[26] Let  $\psi_1, \dots, \psi_n$  be the columns of matrix  $\Psi \in \mathbb{R}^{m \times n}$ . The coherence of matrix  $\Psi$  is defined as the maximum absolute value of the cross-correlation between any two columns  $\psi_i, \psi_j$  for  $\forall i, j, 1 \leq i \neq j \leq n$ , i.e.,

$$\mu(\Psi) = \max_{1 \leq i \neq j \leq n} \frac{\langle \psi_i, \psi_j \rangle}{\|\psi_i\|_2 \|\psi_j\|_2}. \quad (32)$$

By definition, we have  $\mu(\Psi) < 1$  for all  $\mu(\Psi)$ . On the other hand, if  $\Psi$  has orthonormal columns  $\mu(\Psi) = 0$ . Actually, we can find the lower bound of  $\mu(\Psi)$  for  $\Psi$  such that  $\mu(\Psi) \in [\sqrt{\frac{n-m}{m(n-1)}}, 1]$ .

Intuitively, the PU location can be obtained by recovering  $\mathbf{s} = \{\mathbf{s}_i\}_{i=1}^J$  from the measurements  $\mathbf{z} = \{\mathbf{z}_j\}_{j=1}^J$  of all CR nodes,

$$\mathbf{z} = \tilde{\Phi} \mathbf{H}_{\alpha \rightarrow \beta} \mathbf{s} + \tilde{\sigma}. \quad (33)$$

Although straightforward, this approach imposes the incoherence requirement on  $\mathbf{H}_{\alpha \rightarrow \beta}$ , which can be explained by the following theorem.

**Theorem 2.** [18, 27] For a  $K$ -sparse signal  $\theta$  with exact  $K$  non-zero entries. Let  $\Psi \in \mathbb{R}^{L \times N}$  be a matrix with coherence  $\mu(\Psi)$ , and  $\tilde{\Phi} = \mathbf{I} \otimes \Phi \in \mathbb{R}^{M \times L}$  with  $\Phi \in \text{RIP}(K, \delta_K)$ . Then with high probability, any  $K$ -sparse signal  $\theta$  can be reconstructed from the measurement  $y = \tilde{\Phi} \Psi \theta$  by the  $l_1$  minimization if

$$K \leq 1 + \frac{1}{16\mu(\Psi)}, \quad (34)$$

and the number of measurements  $M$  obeys the relation  $M = O(K \log(N/K))$ .

For the inequality in (34), when there are more than one PU in the CR network, the coherence of  $\mathbf{H}_{\alpha \rightarrow \beta}$  should be quite small, which limits the application scope of this algorithm. To recover multiple PUs in the CR network, channel matrix  $\mathbf{H}_{\alpha \rightarrow \beta}$  should be incoherent enough to separate the PUs and recover their locations. The coherence property of  $\mathbf{H}_{\alpha \rightarrow \beta}$  depends on the geographical separation of the PUs, channel frequency response, and the number of CR nodes in the network. An incoherent matrix  $\mathbf{H}_{\alpha \rightarrow \beta}$  can separate multiple PUs in the network to implement the localization task. This may require the well spaced PUs in the geographical field, detailed profile of the channel information, or a large number of CR nodes in the network, which may be impractical to satisfy.

If we view the problem from a different perspective, however, the incoherence requirement on  $\mathbf{H}_{\alpha \rightarrow \beta}$  can be delicately removed. We remark that the PUs are naturally separated in the frequency domain due to the avoidance of communication interference. So at different frequency band  $\{f_b\}_{b=1}^B$ , the geographical positions of the transmitted signal can be estimated. Note that there is only one transmitted signal at a specific frequency  $f_b$ , which can greatly relax the incoherence requirement on  $\mathbf{H}_{\alpha \rightarrow \beta}$  for localization. Once the geographical position of the transmitted signal is obtained, the corresponding PU locations are determined. Capitalizing this key observation, a localization algorithm based on the Bayesian inference at each frequency band is proposed in the sequel.

## 5.2. The Localization Inference Problem

Once the spectrum sensing results  $\{\mathbf{r}_j\}_{j=1}^J$  are obtained, the occupied spectral bands are determined. For a specific occupied frequency band  $f_b$ , the geographical location of the transmitted signal can be estimated through a Bayesian inference. Specifically, the posterior density of  $\alpha_i$  can be determined by  $p(\alpha_i|\tilde{\mathbf{r}}_b) \propto p(\tilde{\mathbf{r}}_b|\alpha_i)p(\alpha_i)$ , where  $p(\alpha_i)$  is the prior probability of being at location  $\alpha_i$ . Without loss of generality, we assume  $p(\alpha_i)$  subject to a uniform distribution in our problem. The conditional probability  $p(\tilde{\mathbf{r}}_b|\alpha_i)$  is calculated as:

$$p(\tilde{\mathbf{r}}_b|\alpha_i) = \prod_{j=1}^J p(\mathbf{r}_b(j)|\alpha_i), \quad (35)$$

and can be obtained during the off-line training. For a specific position  $\alpha_i$ , the probability  $p(\tilde{\mathbf{r}}_b|\alpha_i)$  for the PSD  $\tilde{\mathbf{r}}_b$  at frequency band  $f_b$  subjects to  $\mathcal{N}(\mathbf{u}_b, \Sigma_b)$ . The localization problem can be solved as:

$$\alpha_i = \arg \max p(\alpha_i|\tilde{\mathbf{r}}_b) = \arg \max p(\tilde{\mathbf{r}}_b|\alpha_i). \quad (36)$$

In this case, the  $\alpha_i$  with the largest  $p(\alpha_i|\tilde{\mathbf{r}}_b)$  is chosen as the geographical position of the transmitted signal. The centroid of  $K$  locations with the largest  $p(\alpha_i|\tilde{\mathbf{r}}_b)$  can also be used as the final estimate of the transmitted signal geographical position, and this approach is called KNN ( $K$ -Nearest Neighbors).

## 5.3. Localization CRLB

Understanding the characteristics of the localization error is an essential step to the error control and performance improvement [28]. In CR networks, the localization error depends on variety of network configuration parameters. In this part, the CRLB is calculated to specify this dependency and understand the error characteristics of network localization.

The CRLB provides a lower bound on the error covariance of any unbiased estimator. It is independent on particular estimation methods as long as the statistical model of observations on the variable is specified. It is a benchmark to evaluate various estimators that provide estimates equal to the ground truth if averaged over enough realizations.

In our case, the parameters to be estimated include the coordinates of  $\alpha_i$  and can be denoted as  $\alpha_i = (x_i, y_i)^\top$ . Its corresponding estimation is  $\hat{\alpha}_i = (\hat{x}_i, \hat{y}_i)^\top$ . The covariance matrix of the estimate  $\hat{\alpha}_i$  can be written as

$$\text{Cov}(\hat{\alpha}_i) = E_{\alpha_i}\{(\hat{\alpha}_i - \alpha_i)(\hat{\alpha}_i - \alpha_i)^\top\} = \begin{bmatrix} \sigma_{\hat{x}_i}^2 & \sigma_{\hat{x}_i \hat{y}_i} \\ \sigma_{\hat{y}_i \hat{x}_i} & \sigma_{\hat{y}_i}^2 \end{bmatrix}, \quad (37)$$

where  $E_{\alpha_i}\{\cdot\}$  stands for the expectation value, and the mean square error of the location estimate  $\hat{\alpha}_i$  is  $\text{var}(\hat{\alpha}_i) = \sigma_{\hat{x}_i}^2 + \sigma_{\hat{y}_i}^2$ .

Let  $f(P|\alpha_i)$  denote the probability function of observations  $P$  conditioned on parameter  $\alpha_i$ . The Fisher information matrix (FIM) for exponential family  $f(P|\alpha_i)$  is defined as:

$$J(\alpha_i) = -E \left\{ \frac{\partial^2 \ln f(P|\alpha_i)}{\partial \alpha_i^2} \right\} = \begin{bmatrix} J_{x_i x_i}(\alpha_i) & J_{x_i y_i}(\alpha_i) \\ J_{y_i x_i}(\alpha_i) & J_{y_i y_i}(\alpha_i) \end{bmatrix}. \quad (38)$$

By definition, the CRLB is the inverse of FIM, which is a lower limit for the covariance matrix, i.e.,  $\text{Cov}(\hat{\alpha}_i) \geq \{J(\alpha_i)\}^{-1}$ . Without loss of generality, the received signal of CR  $j$  at frequency  $f_b$  can be calculated using the shadowing path loss model, as follows:

$$\mathbf{r}_b(j) = \overline{PL(d_0)} - 10\gamma \log\left(\frac{d_{ij}}{d_0}\right) + \mathcal{X}_{\sigma_{b_j}}, \quad (39)$$

where  $d_0$  is the reference distance.  $\gamma$  is the path loss factor,  $d_{ij}$  is the distance between the transmitted signal at  $\alpha_i$  and CR node at  $\beta_j$ , and  $\mathcal{X}_{\sigma_{b_j}} \sim \mathcal{N}(0, \sigma_{b_j}^2)$  stands for the shadowing effect. The lower bound of the mean

square of the location estimate  $\hat{\alpha}_i$  is given by the following theorem.

**Theorem 3.** The lower bound of the mean square error of the location estimate  $\hat{\alpha}_i$  is given by:

$$\text{var}(\hat{\alpha}_i) = \frac{\sum_{j=1}^J (\eta_{1j}\eta_{2j})^2 + \sum_{j=1}^J (\eta_{1j}\eta_{3j})^2}{\sum_{j=1}^J (\eta_{1j}\eta_{2j})^2 \sum_{j=1}^J (\eta_{1j}\eta_{3j})^2 - (\sum_{j=1}^J ((\eta_{1j})^2 \eta_{2j}\eta_{3j})^2)}, \quad (40)$$

where  $\eta_{1j} = \frac{10\gamma}{\sigma_{b_j} \ln 10}$ ,  $\eta_{2j} = \frac{\cos \phi_{ij}}{d_{ij}}$ , and  $\eta_{3j} = \frac{\sin \phi_{ij}}{d_{ij}}$ .  $\phi_{ij}$  is the angle between the PU and CR node as illustrated in Fig. 2(a).

The CRLB of the location estimate depends on many factors in the network [29], which can be classified into three categories: the system resource, the environmental parameters, and the geographical layout of the CR nodes. The system resource are the number of CR nodes that can performance the localization task in the network. The environmental parameters are  $\sigma_{b_j}$  and  $\gamma$  in the path loss propagation model. The geographical layout of the CR nodes determines the value of the angle  $\phi_{ij}$  and distance  $d_{ij}$ . A simple illustration is given in Fig. 2. In Fig. 2(a), three CR nodes, A, B and C are evenly separated around PU O. This configuration achieves the highest accuracy and called the geographic dilution of precision [30]. To see the effect of the distance  $d_{ij}$  on the CRLB, CR node A is moving towards PU O along line AO while angle  $\phi$  is unchanged. The ratio of the new CRLB while moving A is compared to the original one. In Fig. 2(b) we can see that the CRLB decreases as the distance between CR node A and PU O decreases.

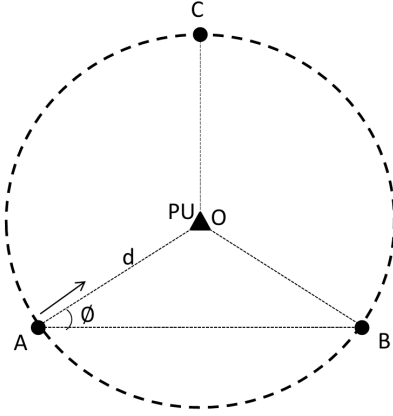
By the observation that CRLB of the location estimation decreases with distance  $d_{ij}$ , a weighted modification to (35) can improve the performance of the localization algorithm. The conditional probability  $p(\mathbf{r}_b|\alpha_i)$  can be rewritten as follows:

$$p(\mathbf{r}_b|\alpha_i) = \prod_{j=1}^J \frac{\frac{1}{d_{ij}^\xi} p(\mathbf{r}_b(j)|\alpha_i)}{\sum_{i=1}^I \frac{1}{d_{ij}^\xi}}, \quad (41)$$

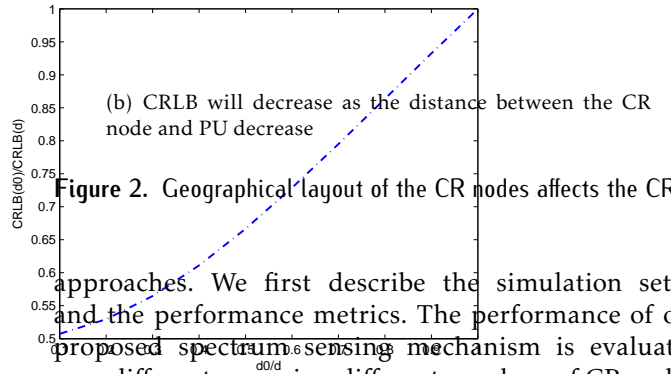
where we assign a weight  $\frac{1}{d_{ij}^\xi}$  to each  $p(\mathbf{r}_b(j)|\alpha_i)$ , and  $\sum_{i=1}^I \frac{1}{d_{ij}^\xi}$  is the normalizing factor calculated across all the possible locations  $\{\alpha_i\}_{i=1}^I$ . During the calculation in (41), the probability  $p(\mathbf{r}_b(j)|\alpha_i)$  provided by a nearer CR node is assigned to more importance, and thus, the precision of the localization mechanism is enhanced.

## 6. Simulation Results

In this section, the numerical simulation results are presented to demonstrate the performance of the proposed spectrum sensing and PU localization



(a) An illustration of geographical dilution of precision



**Figure 2.** Geographical layout of the CR nodes affects the CRLB approaches. We first describe the simulation setup and the performance metrics. The performance of our proposed spectrum sensing mechanism is evaluated over different scenarios: different numbers of CR nodes, different numbers of PUs, and different compressed sampling ratios. Finally, the localization performance is presented.

### 6.1. Simulation Setup and Performance Metrics

We assume the operational space is a  $1000m \times 1000m$  square field, which is divided into  $I = 25$  uniformly distributed grids.  $K$  PUs locate randomly on  $I$  candidate points. The wide bandwidth of the PUs' transmitted signal is equally divided into  $B = 64$  sub-channels. The transmitted signals experience frequency selective fading during transmission. Here, we assume Rayleigh channel fading in our simulation.  $J$  SUs are randomly deployed in the same field working collaboratively to implement the spectrum sensing and localization tasks. Each CR node samples  $M$  random projections of the original signal and the compressed sampling ratio is

defined as  $M/B$ . The SNR is defined as the ratio of the average received signal power to the noise power over the entire bandwidth.

We evaluate the performance of our algorithms using the following metrics. Denote the  $\hat{\mathbf{r}}$  as the recovered sparse vector, and  $\hat{\alpha}_i$  as the recovered PU locations.  $D$  is the distance between adjacent grids. For spectrum availability detection, the receiver operating characteristic (ROC) is considered, where the true positive rate and false alarm rate are defined respectively, as follows:

$$TPR = \frac{N_{Hit}}{N_{Hit} + N_{Miss}}, \quad FAR = \frac{N_{False}}{N_{False} + N_{Correct}}, \quad (4)$$

where  $N_{Hit}$  is the number of successful detections,  $N_{Miss}$  is the number of missed detections,  $N_{False}$  is the number of false alarms, and  $N_{Correct}$  is the number of correct reports of unoccupied spectrum bands.

The spectrum sensing and localization performance are evaluated by their normalized MSEs, which are defined as follows, respectively,

$$\mathbf{r}_{MSE} = E \left\{ \frac{\|\mathbf{r} - \hat{\mathbf{r}}\|_2}{\|\mathbf{r}\|_2} \right\}, \quad L_{MSE} = E \left\{ \frac{\sum_{i=1}^K (\|\alpha_i - \hat{\alpha}_i\|_2)}{K \times D} \right\}. \quad (43)$$

### 6.2. Performance v.s. Different Numbers of CR nodes

In this scenario, we set the number of active PUs  $K = 3$  and the compressed sampling ratio  $M/B = 25\%$ . The SNR is  $10dB$  and the number of CR nodes  $J$  varies from 3 to 7 to evaluate the performance. The simulation results are shown in Fig. 3 and Fig. 4.

From Fig. 3, we can see that the increase in the number of CR nodes can greatly improve the performance. This may mainly due to two reasons. On the one hand, the larger number of CR nodes will provide more measurements for the fusion center using the  $l_1$  minimization algorithm to recover the original sparse signal, which will enhance the probability of successful reconstruction. On the other hand, more CR nodes will better exploit the diversity in the measurements from different CR nodes, which will improve the detection correctness and reduce the reconstruction error. The centralized algorithm with group sparsity performs best. When the number of CR nodes is three, the "naive" approach degrades greatly, which validates the effectiveness of exploiting the joint sparsity in spectrum sensing problem. When few CR nodes present in the network, utilizing the joint sparsity in the received signals can guarantee the performance of reconstruction.

Fig. 4 shows the ROC curves at different numbers of CR nodes when sampling rate equals 25%. When the number of CR nodes is 3, the proposed mechanism

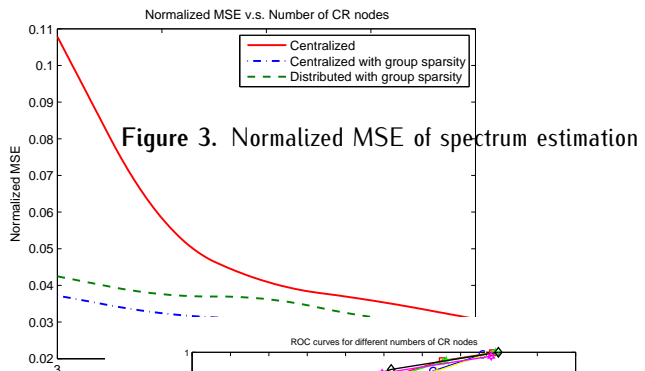


Figure 3. Normalized MSE of spectrum estimation

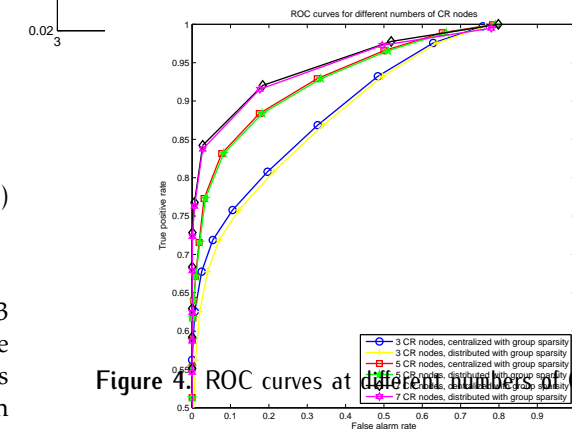


Figure 4. ROC curves at different numbers of CR nodes

can still achieve a moderate true positive rate at a low false positive rate. At the same false alarm rate, the true positive rate improves as the number of CR nodes increases. The centralized and distributed algorithms demonstrate similar performances for the spectrum occupancy detection task.

### 6.3. Performance v.s. Different Numbers of PUs

In this scenario, we set the number of CR nodes  $J = 4$  and the compressed sampling ratio  $M/N = 25\%$ . The SNR is  $10dB$  and the number of active PUs  $K$  varies from 3 to 7 to evaluate the performance. The simulation results are shown in Fig. 5 and Fig. 6.

From Fig. 5, we can notice that the reconstruction errors of spectrum sensing and localization increase with the number of PUs. More PUs will make the original signal less sparse, which will result in the requirement of a larger number of measurements. However, the performances of the proposed mechanism are still acceptable. In the worst case, the normalized

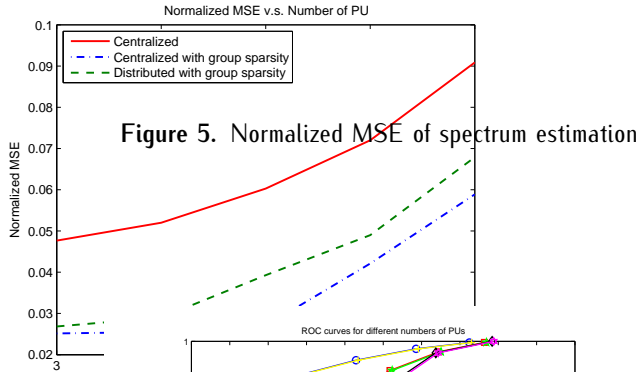


Figure 5. Normalized MSE of spectrum estimation

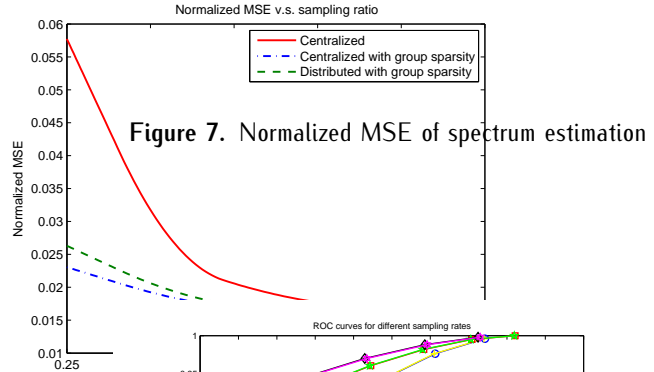


Figure 7. Normalized MSE of spectrum estimation

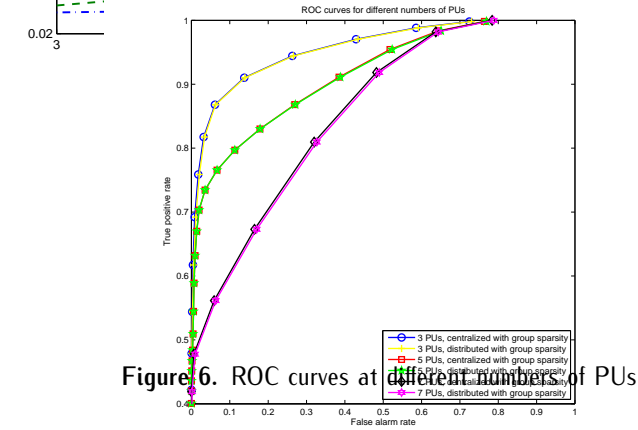


Figure 6. ROC curves at different numbers of PUs

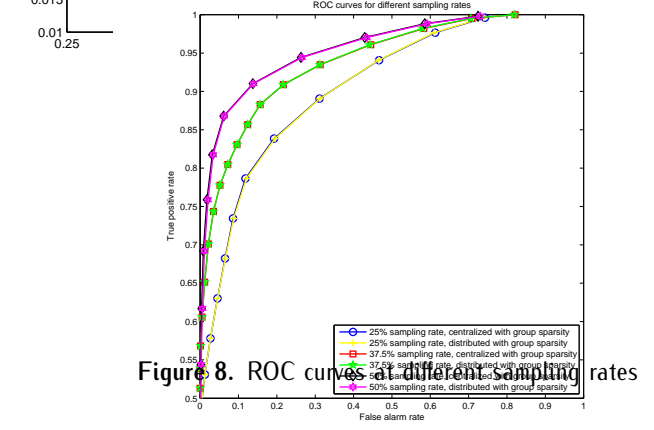


Figure 8. ROC curves at different sampling rates

MSE of spectrum sensing for 7 PU by the centralized algorithm with group sparsity is around 0.06.

Fig. 6 shows the ROC curves at different numbers of PUs when sampling rate equals 50%. The performance of spectrum holes detection deteriorates as the number of PUs increases. However, the proposed scheme can still achieve a satisfied performance for the detection of spectrum holes when 7 PUs present in the network.

#### 6.4. Performance v.s. Different Compressed Sampling Ratio

In this scenario, we set the number of CR nodes  $J = 4$  and the number of active PUs  $K = 3$ . The  $SNR$  is  $10dB$  and the compressed sampling ratio varies from 25% to 50% to evaluate the performance. The simulation results are shown in Fig. 7.

In Fig. 7, the normalized MSE of spectrum sensing decreases as the sampling ratio increases. A larger number of samples result in a smaller normalized MSE. The centralized algorithm with group sparsity perform best in this case. In a severe environment, the

reliable performance at the low sampling rate (25%) can significantly save the energy of the CR nodes.

Fig. 8 depicts the ROC curves of the proposed mechanism at different compressed sampling ratios. With a larger number of measurements, the performance of detecting occupied frequency band is better. When the compressed sampling ratio equals 50%, the performance is best, at the expense of more measurements taken by each CR node and processed centrally or distributively.

#### 6.5. Localization Performance

In this scenario, we set the number of CR nodes  $J = 5$ , the compressed sampling ratio  $M/N = 50%$ , the number of active PU  $K = 3$  and the  $SNR = 10dB$ . We first compare the performance of the proposed algorithms with the method by reconstructing  $s$  in (33), and then the performance improvement of the modified Bayesian location inference is demonstrated.

Fig. 9 shows the normalized localization error for different algorithms by varying the number of CR nodes from 4 to 7. Due to the incoherence requirement

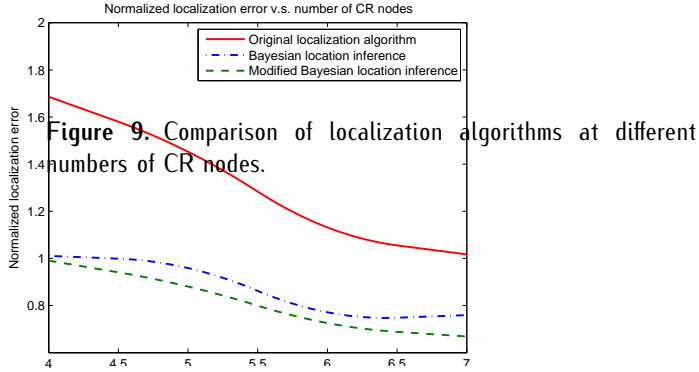


Figure 9. Comparison of localization algorithms at different numbers of CR nodes.

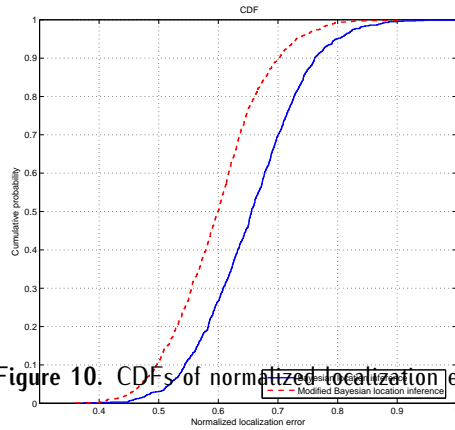


Figure 10. CDFs of normalized localization error

on  $\mathbf{H}_{\alpha \rightarrow \beta}$ , (33) results in significant error in location estimate when few CR nodes are present in the network. The performance is enhanced when the number of CR nodes is larger, which reduces the coherence of the channel gain matrix  $\mathbf{H}_{\alpha \rightarrow \beta}$ . However, the proposed Bayesian location inference approach outperforms (33) because of exploiting the spectrum sensing results. Different frequency bands naturally separate the PUs, and thus, the incoherence requirement on  $\mathbf{H}_{\alpha \rightarrow \beta}$  is significantly relaxed. Moreover, the modified algorithm improves the performance by utilizing the error characteristics revealing by the CRLB. The cumulative distribution functions of the normalized localization errors are shown in Fig. 10 for the case when the number of CR nodes equals to 7. The simulation results show that the accuracy of the localization algorithm is improved.

## 7. Conclusion

In this paper, we redesign the framework of compressive spectrum sensing and PU localization in CR networks. The applicability of CS to cooperative spectrum sensing is justified, and the spectrum recovery performance is guaranteed. We use three different algorithms, Lasso, group Lasso, and (distributed) group Lasso with feature splitting, to recover the received signal's PSD. The locations of PUs are estimated through a Bayesian inference approach based on the results of spectrum sensing. This two-step treatment greatly relaxes the incoherence requirement of the channel gain matrix, which is required for PU localization based on spatial sparsity. The error characteristics of localization are analyzed by computing the CRLB, and the localization algorithm design is improved by exploiting the revealed error trends associated with the system. Through numerical simulations, the efficacy and efficiency of the proposed mechanism are validated.

## Appendix A. Proof of Theorem 1

*Proof.* We follow the idea of [31] to prove the Theorem. Let matrix  $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$ . Suppose set  $\Omega$  is of cardinality  $K$ . We denote  $\mathbf{C}_\Omega$  as the sub-matrix of  $\mathbf{C}$  containing the columns  $c_t$ ,  $t \in \Omega$ . For a column  $t \in \Omega$  of matrix  $\mathbf{C}_\Omega$ , we have  $\mathbf{c}_t = \mathbf{a}_{t_1} \otimes \mathbf{b}_{t_2}$ , where  $\mathbf{a}_{t_1}$  is column  $t_1$  of matrix  $\mathbf{A}$ , and  $\mathbf{b}_{t_2}$  is column  $t_2$  of matrix  $\mathbf{B}$ .  $t_1 \in \Omega_1$  and  $t_2 \in \Omega_2$ . For each  $t \in \Omega$ , we can find the corresponding  $t_1 \in \Omega_1$  and  $t_2 \in \Omega_2$ . The cardinality product  $|\Omega_1||\Omega_2| \leq K^2$ . We have  $\Omega \subset \Omega_1 \cup \Omega_2$ . Since the range of singular values of a sub-matrix are interlaced inside those of the original matrix, we have:

$$\sigma_{\min}(\mathbf{A} \otimes \mathbf{B}) = \sigma_{\min}(\mathbf{A}_{\Omega_1})\sigma_{\min}(\mathbf{B}_{\Omega_2}) \leq \sigma_{\min}(\mathbf{C}_\Omega), \quad (\text{A.1})$$

$$\sigma_{\max}(\mathbf{A} \otimes \mathbf{B}) = \sigma_{\max}(\mathbf{A}_{\Omega_1})\sigma_{\max}(\mathbf{B}_{\Omega_2}) \geq \sigma_{\max}(\mathbf{C}_\Omega). \quad (\text{A.2})$$

By definition of RIP, the  $\delta_K(\mathbf{A}_{\Omega_1})$ ,  $\delta_K(\mathbf{B}_{\Omega_1})$  and  $\delta_K(\mathbf{C}_\Omega)$  are the smallest constants that make the following inequalities holds:

$$(1 - \delta_K(\mathbf{A}_{\Omega_1})) \leq \sigma_{\min}(\mathbf{A}_{\Omega_1}) \leq \sigma_{\max}(\mathbf{A}_{\Omega_1}) \leq (1 + \delta_K(\mathbf{A}_{\Omega_1})), \quad (\text{A.3})$$

$$(1 - \delta_K(\mathbf{B}_{\Omega_1})) \leq \sigma_{\min}(\mathbf{B}_{\Omega_1}) \leq \sigma_{\max}(\mathbf{B}_{\Omega_1}) \leq (1 + \delta_K(\mathbf{B}_{\Omega_1})), \quad (\text{A.4})$$

$$(1 - \delta_K(\mathbf{C}_\Omega)) \leq \sigma_{\min}(\mathbf{C}_\Omega) \leq \sigma_{\max}(\mathbf{C}_\Omega) \leq (1 + \delta_K(\mathbf{C}_\Omega)). \quad (\text{A.5})$$

Substitute (A.3) and (A.4) into (A.1) and (A.2), we have:

$$\begin{aligned} (1 - \delta_K(\mathbf{A}_{\Omega_1}))(1 - \delta_K(\mathbf{B}_{\Omega_1})) &\leq \sigma_{\min}(\mathbf{C}_\Omega), \\ (1 + \delta_K(\mathbf{A}_{\Omega_1}))(1 + \delta_K(\mathbf{B}_{\Omega_1})) &\geq \sigma_{\max}(\mathbf{C}_\Omega). \end{aligned} \quad (\text{A.6})$$

Note that  $\delta_K(\mathbf{C}_\Omega)$  is the smallest constant that satisfies (A.5)

$$(1 + \delta_K(\mathbf{C})) \leq (1 + \delta_K(\mathbf{A}))(1 + \delta_K(\mathbf{B})). \quad (\text{A.7})$$

$\mathbf{A}$  is an orthonormal basis, and  $\delta_K(\mathbf{A}) = 0$ . So we have  $\delta_K(\mathbf{C}) \leq \delta_K(\mathbf{B})$ . By Theorem 3.7 in [32], we have  $\delta_K(\mathbf{C}) \geq \delta_K(\mathbf{B})$ . Hence,  $\delta_K(\mathbf{C}) = \delta_K(\mathbf{B})$ .  $\square$

## Appendix B. Proof of Theorem 2

The joint probability density distribution  $f(P|\alpha_i)$  can be expressed as,

$$f(P|\alpha_i) = \prod_{j=1}^J \frac{1}{\sqrt{2\pi}\sigma_{bj}} \times \exp \left\{ -\frac{\mathbf{r}_b(j) - \overline{PL(d_0)} + 10\gamma \log(\frac{d_{ij}}{d_0})}{2\sigma_{bj}^2} \right\} \quad (B.8)$$

After obtaining the log-likelihood of  $f(P|\alpha_i)$ , the entries of the FIM can be calculated according to (38), which can be expressed as follows:

$$J_{x_i x_i}(\alpha_i) = \sum_{j=1}^J \left( \frac{10\gamma}{\sigma_{bj} \ln 10} \right)^2 \left( \frac{\cos \phi_{ij}}{d_{ij}} \right)^2, \quad (B.9)$$

$$J_{x_i y_i}(\alpha_i) = J_{y_i x_i}(\alpha_i) = \sum_{j=1}^J \left( \frac{10\gamma}{\sigma_{bj} \ln 10} \right)^2 \left( \frac{\cos \phi_{ij}}{d_{ij}} \right) \left( \frac{\sin \phi_{ij}}{d_{ij}} \right), \quad (B.10)$$

$$J_{y_i y_i}(\alpha_i) = \sum_{j=1}^J \left( \frac{10\gamma}{\sigma_{bj} \ln 10} \right)^2 \left( \frac{\sin \phi_{ij}}{d_{ij}} \right)^2. \quad (B.11)$$

The definition of the angle  $\phi_{ij}$  is illustrated in Fig. 2(a). Since  $Cov(\hat{\alpha}_i) \geq \{J(\alpha_i)\}^{-1}$ , the lower bound of the location estimation variance  $var(\hat{\alpha}_i)$  can be calculated as:

$$\begin{aligned} var(\hat{\alpha}_i) &= \sigma_{\hat{x}_i}^2 + \sigma_{\hat{y}_i}^2 \geq \frac{J_{x_i x_i}(\alpha_i) + J_{y_i y_i}(\alpha_i)}{|J(\alpha_i)|} \\ &= \frac{J_{x_i x_i}(\alpha_i) + J_{y_i y_i}(\alpha_i)}{J_{x_i x_i}(\alpha_i)J_{y_i y_i}(\alpha_i) - J_{x_i y_i}(\alpha_i)J_{y_i x_i}(\alpha_i)}. \end{aligned} \quad (B.12)$$

By Substituting (B.9), (B.10) and (B.11) into (B.12), we can obtain the lower bound of the mean square error of the location estimate  $\hat{\alpha}_i$ .

## References

- [1] G. Staple and K. Werbach, "The end of spectrum scarcity," *IEEE Spectrum Archive*, vol. 41, no. 3, pp. 48-52, Mar. 2004.
- [2] S. Haykin, "Cognitive radio: brained-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201-220, Feb. 2005.
- [3] E. Hossain, D. Niyato, and Z. Han, "Dynamic Spectrum Access in Cognitive Radio Networks," *Cambridge University Press*, UK, 2009.
- [4] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 849-877, May. 2009.
- [5] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: a survey," *Physical Communication*, vol. 4, no. 1, pp. 40-62, Mar. 2011.
- [6] S. Boyd, N. Parikh, and E. Chu, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Now Publishers Inc*, Boston, 2011.
- [7] D. P. Bertsekas and J. N. Tsitsiklis, "Parallel and distributed computation: numerical methods," *Prentice Hall*, New Jersey, 1989.
- [8] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.
- [9] E. J. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406-5425, Dec. 2006.
- [10] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from in-complete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 9, pp. 1207-1223, Aug. 2006.
- [11] Z. Tian and G. Giannakis, "Compressed sensing for wideband cognitive radios," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Honolulu, HI, 2007.
- [12] Z. Tian, "Compressed wideband sensing in cooperative cognitive radio networks," *IEEE Global Telecommunications Conference*, New Orleans, LA, 2008.
- [13] J. Meng, W. Yin, H. Li, E. Hossain, and Z. Han, "Collaborative spectrum sensing from sparse observations in cognitive radio networks," *IEEE Journal on Selected Topics on Communications, special issue on Advances in Cognitive Radio Networking and Communications*, vol. 29, no. 2, pp. 327-337, Feb. 2011.
- [14] F. Zeng, C. Li, and Z. Tian, "Distributed compressive spectrum sensing in cooperative multi-hop wideband cognitive networks," *IEEE Journal of Selected Topics in Signal Processing, Special Issue on Signal Processing in Cooperative Cognitive Radio Systems*, vol. 5, no. 1, pp. 37-48, Feb. 2011.
- [15] L. Liu, Z. Han, Z. Wu, and L. Qian, "Collaborative compressive sensing based dynamic spectrum sensing and mobile primary user localization in cognitive radio networks," *IEEE Global Telecommunications Conference*, Houston, TX, 2011.
- [16] J. A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1847-1862, Mar. 2010.
- [17] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. Ser. B*, vol. 58, no. 1, pp. 267-288, 1996.
- [18] V. Cevher, M. F. Duarte, and R. G. Baraniuk, "Distributed target localization via spatial sparsity," *European Signal Processing Conference*, Lausanne, Switzerland, Aug. 2008.
- [19] V. Cevher, P. Boufounos, R. G. Baraniuk, A. C. Gilbert, and M. J. Strauss, "Near-optimal bayesian localization via incoherence and sparsity," *Int. Conf. on Information Processing in Sensor Networks*, San Francisco, CA, Apr. 2009.

- [20] M. Mishali, Y. C. Eldar, and A. Elron, "Xampling: signal acquisition and processing in union of subspaces," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4719-4734, Oct. 2011.
- [21] M. Mishali and Y. C. Eldar, "Sub-Nyquist sampling: bridging theory and practice," *IEEE Signal Processing Magazine*, vol. 28, no. 6, pp. 98-124, Nov. 2011.
- [22] J. Tropp, M. Wakin, M. Duarte, D. Baron, and R. Baraniuk, "Random filters for compressive sampling and reconstruction," *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Toulouse, France, May 2006.
- [23] Z. Han, H. Li, and W. Yin, "Compressive sensing for wireless networks," *Cambridge University Press*, UK, 2012.
- [24] J. Yand and Y. Zhang, "Alternating direction algorithms for  $l_1$ -problems in compressive sensing," *SIAM journal on Scientific Computing*, vol. 33, no. 1, pp. 250-278, Feb. 2011.
- [25] W. Deng, W. Yin, and Y. Zhang, "Group sparse optimization by alternating direction method," *Rice CAAM Report*, TR11-06, 2011.
- [26] E. J. Candès and Y. Plan, "Near-ideal model selection by  $l_1$  minimization," *Ann. Statist.*, vol. 27, no. 5, pp. 2145-2177, 2007.
- [27] H. Rauhut, K. Schnass, and P. Vandhergheynst, "Compressed sensing and redundant dictionaries," *IEEE Trans. Info. Theory*, vol. 54, no. 5, pp. 2210-2219, May 2008.
- [28] Y. Liu and Z. Yang, "Location, localization and localizability: location-awareness technology for wireless networks," *Springer*, NY, 2011.
- [29] A. M. Hossian and W. Soh, "Cramér-Rao bound analysis of localization using signal strength difference as location fingerprint," in *Proceedings of IEEE Conference on Computer Communications*, San Diego, CA, USA.
- [30] M. A. Spirito, "On the accuracy of cellular mobile station location estimation," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 3, pp. 674-685, May. 2001.
- [31] M. F. Duarte and R. G. Baraniuk, "Kronecker compressive sensing," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 494-504, Feb. 2012.
- [32] S. Jokar and V. Mehrmann, "Sparse representation of solutions of Kronecker product systems," *Linear Algebr. Appl.*, vol. 431, no. 12, pp. 2437-2447, Dec. 2009.