This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

Joint Spectrum Sensing and Resource Allocation in Multi-Band-Multi-User Cognitive Radio Networks

Xu Wang, *Student Member, IEEE,* Sabit Ekin, *Member, IEEE,* and Erchin Serpedin, *Fellow, IEEE*

***Abstract*—In this paper, the joint spectrum sensing and resource allocation problem is investigated in a multi-band-multi-user cognitive radio (CR) network. Assuming imperfect spectrum sensing information, our goal is to jointly optimize the sensing threshold and power allocation strategy such that the average total throughput of secondary users (SUs) is maximized. Additionally, the power of SUs is constrained to keep the interference introduced to primary users (PUs) under certain limits, which gives rise to a nonconvex mixed integer non-linear programming (MINLP) optimization problem. Our contribution in this paper is threefold. First, it is illustrated that the dimension of the nonconvex MINLP problem can be significantly reduced, which helps to re-formulate the optimization problem without resorting to integer variables. Second, it is demonstrated that the simplified formulation admits the canonical form of a monotonic optimization and an ϵ-optimal solution can be achieved using the polyblock outer approximation algorithm. Third, a practical low-complexity spectrum sensing and resource allocation algorithm is developed to reduce the computational cost. Finally, the effectiveness of proposed algorithms is verified by simulations.**

***Index Terms*—Cognitive radio, resource allocation, spectrum sensing, MINLP, monotonic optimization.**

I. Introduction

With the increasing demands for smart mobile devices and other bandwidth consuming wireless applications, radio frequency (RF) spectrum is becoming more and more crowded. Even though most of the available spectrum has already been licensed, in general, only a portion of the licensed spectrum is utilized simultaneously due to the varying demand and rate requirements from the licensed users. Recent reports, performed by many agencies such as Federal Communications Commission (FCC), have indicated that the RF spectrum is being used inefficiently [1], [2]. The concept of cognitive radio (CR) was proposed as a promising technology for the efficient utilization of RF spectrum [3]. CR networks generally can be classified into three broad categories [4]: underlay, overlay and interweave networks. In underlay networks, the simultaneous coexistence of primary users (PUs) and secondary users (SUs) is allowed as long as the quality of service (QoS) level at PUs is guaranteed. In overlay networks, SUs have access to the knowledge of the PUs' transmission scheme and use this knowledge to choose a transmission strategy that causes an acceptable amount of interference to PUs. In interweave networks, SUs sense the licensed spectrum and access the channel only if the channel is detected as being vacant. In our work, we will focus on the interweave network scheme, which represents also the original motivation for CR.

Spectrum sensing is of significant importance in CR networks since that's how SUs detect and access the available spectrum. Apparently, a well-designed spectrum sensing scheme can accurately find the spectrum holes and improve the performance of secondary network. In literature, several works studied the spectrum sensing problem by proposing strategies to select the sensing parameters to maximize the average throughput of secondary network under the constraint that PUs are sufficiently well protected [5]–[7]. In particular, the average throughput of secondary network is expressed as the multiplication of false alarm rate and achievable throughput, with the false alarm rate as a function of sensing parameters and the achievable throughput as fixed values. However, the achievable throughput of secondary network heavily depends on the allocation of the channels and power to SUs, and it represents an alternative way to enhance the SU's performance. References [8]–[10] focused on maximizing the throughput of secondary network assuming the perfect spectrum sensing information, i.e., no sensing errors are considered in these works. However, perfect spectrum sensing is extremely difficult to acquire in practical wireless networks. Considering the false alarm and misdetection rates as fixed values, the authors in [11], [12] studied the resource allocation problem in CR networks by maximizing the total throughput of SUs.

In a nutshell, both the selection of sensing parameters and the physical-layer resource allocation strategy have an impact on the performance of the secondary network. The aforementioned references either optimize the spectrum sensing scheme under a pre-defined SU resource layout or design a resource allocation strategy while fixing the sensing parameters. Therefore, joint spectrum sensing and resource allocation problems in CR networks have drawn great interest recently [13]–[17]. In the scenario of one single channel band, the authors in [13] studied the joint optimization of sensing time and power allocation. The goal was to maximize the average throughput of secondary network and minimize the outage probability of SUs' transmission. In [14], the sensing parameters and the transmission power of SUs are optimized to minimize the total power consumption by considering the cooperative sensing scheme. A recently published paper [15] proposed a prediction-and-sensing-based spectrum sharing model to solve a joint spectrum sensing and resource allocation problem for CRNs. An efficient iterative algorithm based on an in-

Xu Wang and Erchin Serpedin are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, U.S.A. e-mails: {xu.wang, eserpedin}@tamu.edu.

Sabit Ekin is with the Department of Electrical and Computer Engineering, Oklahoma State University, Stillwater, Oklahoma, U.S.A. e-mail: sabit.ekin@okstate.edu.

This work was supported by NSF EARS Award No. 1547447.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

ner approximation framework was developed for the highly-nonconvex problem. In terms of the multi-band setup, reference [16] formulated the joint optimization problem for one PU and one SU in OFDM-based CR networks. The problem was further generalized to the case of a non-cooperative power allocation game for two SUs. In [17], Fan *et al.* considered the joint optimization problem for multiple SUs in a multi-band CR network. The time-sharing property among SUs was assumed in [17], which allows multiple SUs to share a common channel. In both [16] and [17], the alternating optimization (AO) method was employed to maximize the average throughput of SUs by alternatively optimizing the sensing parameters and power variables while setting constant the other variables.

In the multi-band setup, both [16] and [17] utilized the AO method to iteratively optimize the sensing parameters and allocate transmission power using a random initial guess. However, the AO method can be easily trapped in a local minima/maxima near the starting point [18]. Thus, the converged solutions in [16] and [17] heavily depend on the initial point selected and the stability of the proposed AO algorithms is questionable. Moreover, for a multi-band-multi-user setup, the time-sharing property is not practical in some cases such as the OFDM scheme, in which an IFFT/FFT pair coupled with a cyclic prefix are used for modulation and de-modulation. Since the FFT-size is fixed in advance, no further subdivision of each frequency subchannel is possible [19].

In this paper, we address the joint spectrum sensing and resource allocation problem in a downlink multi-band-multi-user CR network. Our goal is to maximize the average throughput of secondary network subject to a peak power constraint of SUs and an average interference constraint to PUs. Furthermore, in order to prevent the inner interference among SUs, we assume that each channel in the frequency domain can be occupied by at most one SU, which leads to a nonconvex mixed integer nonlinear programming (MINLP) optimization problem. The contribution of this study is summarized below:

- The formulated MINLP optimization problem is NP-hard and computationally difficult even to obtain a suboptimal solution. However, we first show that under a general and practical assumption, the nonconvex MINLP problem can be reduced to a formulation which does not resort to any integer variable. As a result, the dimension of the original problem is significantly reduced and we end up with an equivalent nonconvex optimization problem.
- It is further illustrated that the reduced formulation represents the canonical form of a monotonic optimization and can be solved globally using the polyblock outer approximation algorithm. In general, even for pure resource allocation problems in literature, only a suboptimal solution can be achieved for multi-band-multi-user CR networks. In this paper, we provide an optimal optimization method based on monotonic optimization techniques. To our best knowledge, this is the first time an optimal algorithm (except for the exhaustive search method) is developed for the joint spectrum sensing and resource allocation problem in a multi-band-multi-user setup.
- Since the polyblock outer approximation algorithm can only be implemented efficiently in a small-size optimization problem, a computationally efficient suboptimal algorithm, which consists of a list of single-variable optimization problems, is proposed based on a modified version of the original nonconvex MINLP formulation.
- Numerical results of the proposed optimal and suboptimal algorithms are performed in comparison with the state-of-the-art AO method employed in [16] and [17]. The effectiveness of the proposed algorithms is verified.

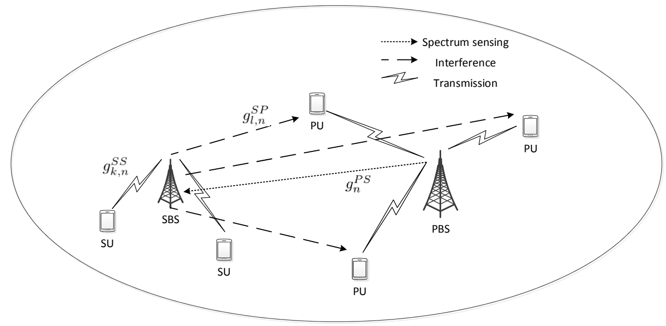

Fig. 1. A downlink CR network.

The rest of the paper is structured as follows. The system model and problem formulation are described in Section II. By exploring the special structure of the problem in Section II, an equivalent but significantly reduced in dimension formulation of the optimization problem is presented in Section III. The proposed algorithm exploits the knowledge of a monotonic optimization, and it is shown in Section IV to achieve an ϵ-optimal solution. In Section V, we develop a suboptimal spectrum sensing and power allocation algorithm with reduced computational complexity. Numerical results and discussions are presented in Section VI. Finally, conclusions are drawn in Section VII.

II. System Model and Problem Formulation

In order to make the rest of the paper easy to follow, a list of some frequently-used terminologies and their corresponding notations is illustrated in Table I. Thus, we will not explain the definitions of these notations in the remaining context.

We consider a downlink multi-band CR interweave network with a secondary base-station (SBS) coexisting in the vicinity of a primary base-station (PBS). It is assumed that the whole wideband spectrum has N total channels with equal bandwidth B. In addition, all the channels are licensed to PUs and the bands of PUs are not overlapping. There are L PUs licensed in the wideband and K SUs, as shown in Fig. 1. The CR model in the RF domain is depicted in Fig. 2. At a certain time instant, the SBS detects the spectrum [20], [21] and finds the spectrum holes which are licensed to but not utilized by PUs based on the joint optimization rule. These spectrum holes are referred to as the CR bands and are available for SUs' transmission. Note that only a portion of SUs can access the spectrum holes if the number of vacant channels is less than the total number of SUs.

TABLE I
LIST OF TERMINOLOGIES AND CORRESPONDING NOTATIONS

L	Number of PUs
K	Number of SUs
N	Number of channels
B	Bandwidth of each channel
γ_n	Sensing threshold for the nth channel
M	Number of sensing samples
σ_ν^2	Noise power
g_n^{PS}	Power gain from the PBS to the SBS
σ_s^2	Average power of the transmitted signal from the PBS
$\mathcal{H}_n^0$	Hypothesis that the nth channel is vacant
$\mathcal{H}_n^1$	Hypothesis that the nth channel is occupied
d_n^0	Event that the nth channel is detected to be vacant
d_n^1	Event that the nth channel is detected to be occupied
P_n^D	Detection rate on the nth channel
P_n^F	False alarm rate on the nth channel
P_n^{MD}	Misdetection rate on the nth channel
P_n^0	Probability that the nth channel is vacant
P_n^1	Probability that the nth channel is occupied
$R_{k,n}$	Average throughput of the kth SU on the nth channel
$\bar{R}_{k,n}$	Approximated average throughput of the kth SU on the nth channel
$y_{k,n}$	Transmitted power from the SBS to the kth SU on the nth channel
$x_{k,n}$	Binary variable with $x_{k,n} = 1$ indicating that the nth channel is occupied by the kth SU and $x_{k,n} = 0$ otherwise
$I_{k,n}$	Interference from PUs introduced to the kth SU on the nth channel
$g_{k,n}^{SS}$	Power gain from the SBS to the kth SU on the nth channel
$g_{l,n}^{SP}$	Power gain from the SBS to the lth PU on the nth channel
$I_l^{\max}$	Average interference constraint to the lth PU
S_l	Licensed channels for the lth PU
P_n^T	Peak power limit of the SBS on the nth channel
$D_n^{\min}$	Lower bound for the probability of detection
$F_n^{\max}$	Upper bound for the probability of false alarm

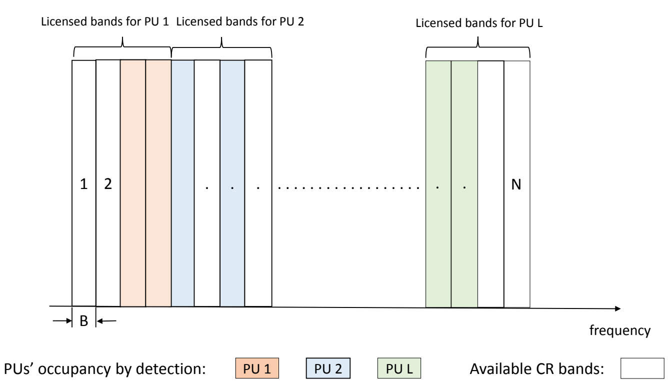

Fig. 2. RF spectrum sharing between PUs and SUs.

A. Spectrum Sensing

When the SBS senses the spectrum, two kinds of sensing errors typically exist. The first is termed as false alarm, and it occurs when the channel is detected to be occupied but it is actually vacant. The second kind is referred to as misdetection, and it represents the case when the SBS fails to detect the occupancy of PUs, i.e., the channel is detected to be vacant but it is actually occupied. We employ the sensing technique in [7], in which the false alarm and detection probabilities in the nth channel can be expressed as

$$P_n^F(\gamma_n) = P(d_n^1|\mathcal{H}_n^0) = Q\left(\frac{\gamma_n - M\sigma_\nu^2}{\sigma_\nu^2\sqrt{2M}}\right), \tag{1}$$

and

$$P_n^D(\gamma_n) = P(d_n^1|\mathcal{H}_n^1) = Q\left(\frac{\gamma_n - M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2)}{\sigma_\nu\sqrt{2M(\sigma_\nu^2 + 2g_n^{PS}\sigma_s^2)}}\right), \tag{2}$$

where $Q(\cdot)$ represents the tail probability of the standard normal distribution. Basically, the hypothesis $\mathcal{H}_n^1$ is chosen if the received signal level is above γ_n on the nth channel. On the other hand, the hypothesis $\mathcal{H}_n^0$ is selected if the level is below γ_n. Additionally, σ_s^2 and σ_ν^2 are both assumed to be equal at each channel. Similar to [7], we assume that M, σ_s^2, σ_ν^2, and g_n^{PS} are known *a priori* at the SBS. Particularly, the power gain g_n^{PS} can be learned during a period when the primary transmitter is known to be working [11].

The probability of misdetection is then given by

$$P_n^{MD}(\gamma_n) = P(d_n^0|\mathcal{H}_n^1) = 1 - P_n^D(\gamma_n). \tag{3}$$

It can be observed that the choice of sensing threshold results in a trade-off between the probabilities of false alarm and misdetection. As shown in (1) and (3), a higher sensing threshold leads to a smaller false alarm rate but a larger misdetection rate, which triggers the potential for a larger throughput as well as a larger interference. Similarly, a smaller sensing threshold yields a larger false alarm rate but a smaller misdetection rate, which turns out to be a potential for a smaller throughput and interference. Therefore, sensing threshold can be optimized to achieve the best trade-off between the throughput and interference [7].

B. Resource Allocation

The average throughput of the kth SU on the nth channel can be expressed via Shannon capacity formula as follows [5], [6]:

$$\begin{aligned} R_{k,n} = {} & P_n^0(1 - P_n^F)x_{k,n}\log\left(1 + \frac{g_{k,n}^{SS}p_{k,n}}{\sigma_\nu^2}\right) \\ & + P_n^1(1 - P_n^D)x_{k,n}\log\left(1 + \frac{g_{k,n}^{SS}p_{k,n}}{\sigma_\nu^2 + I_{k,n}}\right), \end{aligned} \tag{4}$$

where $P_n^0 = P(\mathcal{H}_n^0)$ and $P_n^1 = P(\mathcal{H}_n^1)$ are the priori probabilities, and they can be obtained via long-term estimation. Additionally, $g_{k,n}^{SS}$ is assumed to be known at the SBS. This information can be acquired at the SBS by measuring the uplink channel gain and utilizing the symmetry property between uplink and downlink channels [10]. Basically, SUs are only allowed to transmit on the nth channel when it is detected to be vacant, i.e., at the scenario d_n^0. In this case, if the detection is correct, then only noise exists. On the other hand, when it fails to detect the occupancy of PUs, there is interference from PUs to the SU. For simplicity, the logarithm functions are assumed to be natural logarithms in this paper.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

In order to formulate the optimization problem, the following constraints are considered:

- Interference introduced to PUs: To sufficiently protect PUs, an interference threshold $I_l^{\max}$, $l = 1, \cdots, L$, is enforced to constrain the average interference introduced at PUs by SUs:

$$\sum_{k=1}^{K}\sum_{n\in S_l} P_n^1(1-P_n^D)x_{k,n}p_{k,n}g_{l,n}^{SP} \le I_l^{\max},\ l=1,\cdots,L, \tag{5}$$

  where $g_{l,n}^{SP}$ can be periodically measured by a band manager or estimated by listening to a becon signal and then fed back to the SBS [22].

- Transmit power constraint: Instead of considering the average power constraint, which only guarantees that the power limit is not exceeded in an average sense, we consider the peak power constraint herein paper. Let P_n^T represent the maximum transmit power of the SBS on the nth channel. The power constraint reduces to

$$\sum_{k=1}^{K} x_{k,n}p_{k,n} \le P_n^T,\ n=1,\cdots,N. \tag{6}$$

  Additionally, constraint (6) also prevents the ill-conditioned situation in constraint (5). Specifically, if $P_n^1(1-P_n^D)$ admits a very small value on some channel n, then $\sum_{k=1}^{K} x_{k,n}p_{k,n}$ can even take a large value close to infinity and still make (5) satisfied.

- Interference constraint among SUs: To prevent the interference among SUs, we simply add a constraint requiring that each channel can be occupied by at most one SU [8]–[12], i.e.,

$$\sum_{k=1}^{K} x_{k,n} \le 1,\ n=1,\cdots,N. \tag{7}$$

- Sensing error constraint: In order to receive the best balance between false alarm and misdetection, the following sensing error constraints are added:

$$P_n^D(\gamma_n) \ge D_n^{\min}, P_n^F(\gamma_n) \le F_n^{\max},\ n=1,\cdots,N, \tag{8}$$

  where $D_n^{\min}$ and $F_n^{\max}$ are constants with $D_n^{\min} \ge 0.5$ and $F_n^{\max} \le 0.5$. Due to the monotonically decreasing property of Q function in (1) and (2), constraints (8) can be further expressed as

$$\gamma_n^{\min} \le \gamma_n \le \gamma_n^{\max},\ n=1,\cdots,N, \tag{9}$$

  where $\gamma_n^{\min} = Q^{-1}(F_n^{\max})\sigma_\nu^2\sqrt{2M} + M\sigma_\nu^2$, and $\gamma_n^{\max} = Q^{-1}(D_n^{\min})\sigma_\nu\sqrt{2M(\sigma_\nu^2+2g_n^{PS}\sigma_s^2)} + M(\sigma_\nu^2+g_n^{PS}\sigma_s^2)$.

- Basic constraints for $x_{k,n}$ and $p_{k,n}$: Since $x_{k,n}$ is a binary variable indicating whether the nth channel is occupied by the kth SU, it follows that

$$x_{k,n} \in \{0,1\},\ k=1,\cdots,K,\ n=1,\cdots,N. \tag{10}$$

  In addition, the non-negative property of the power variables is expressed as

$$p_{k,n} \ge 0,\ k=1,\cdots,K,\ n=1,\cdots,N. \tag{11}$$

Since we are investigating a CR network where the licensed spectrum is inefficiently utilized, for example, the TV white space where the spectrum occupancy is only around 10% to 30% [23]–[25] due to the switchover from the analog to digital TV , the PU activity probability P_n^1 is assumed to be very small, and it follows that $P_n^0 > P_n^1$. Moreover, based on (8), it turns out that $1-P_n^F > 1-P_n^D$. Therefore, combing these two observations and the fact that $\sigma_\nu^2 < \sigma_\nu^2 + I_{k,n}$, the first term in (4) dominates $R_{k,n}$, and $R_{k,n}$ can be further approximated as

$$\tilde{R}_{k,n} = P_n^0(1-P_n^F)x_{k,n}\log\left(1+\frac{g_{k,n}^{SS}p_{k,n}}{\sigma_\nu^2}\right). \tag{12}$$

The joint spectrum sensing and resource allocation problem, referred to as "P-J" herein paper, can be formulated as

$$\begin{aligned}\max_{x_{k,n},p_{k,n},\gamma_n}\quad & \sum_{k=1}^{K}\sum_{n=1}^{N}\tilde{R}_{k,n} \\ \text{s.t.}\quad & (5),(6),(7),(9),(10),(11).\end{aligned} \tag{P-J}$$

III. An Equivalent Formulation

Since the objective function of P-J is nonconvex with respect to $x_{k,n}, p_{k,n}$ and γ_n, it is a nonconvex MINLP problem, which is NP-hard and computationally complex to solve [26]. In this section, we illustrate that under a general and practical assumption, the dimension of the original nonconvex MINLP problem can be significantly reduced. Basically, the binary variable $x_{k,n}$ can be removed which results in an equivalent formulation only in terms of the power and sensing threshold variables.

Instead of directly analyzing P-J with respect to $x_{k,n}, p_{k,n}$ and γ_n, we alternatively tackle the problem with a fixed value of γ_n satisfying the constraint (9). In this way, the resulting problem can be expressed as

$$\begin{aligned}\max_{x_{k,n},p_{k,n}}\quad & \sum_{k=1}^{K}\sum_{n=1}^{N}\tilde{R}_{k,n} \\ \text{s.t.}\quad & (5),(6),(7),(10),(11).\end{aligned} \tag{13}$$

It can be observed (13) is still a nonconvex MINLP problem. However, by exploring the special structure in the dual domain, it is demonstrated that the optimal solution of (13) can be derived using standard convex optimization tools under a general and practical assumption. We first introduce a new variable $y_{k,n}$, where $y_{k,n} = x_{k,n}p_{k,n}$, to replace the variable $p_{k,n}$. Thus, the resulting problem resumes to

$$\begin{aligned}\max_{x_{k,n},y_{k,n}}\quad & \sum_{k=1}^{K}\sum_{n=1}^{N} P_n^0(1-P_n^F)x_{k,n}\log\left(1+\frac{g_{k,n}^{SS}y_{k,n}}{\sigma_\nu^2 x_{k,n}}\right) \\ \text{s.t.}\quad & \sum_{k=1}^{K}\sum_{n\in S_l} P_n^1(1-P_n^D)y_{k,n}g_{l,n}^{SP} \le I_l^{\max},\ l=1,\cdots,L, \\ & \sum_{k=1}^{K} y_{k,n} \le P_n^T,\quad \sum_{k=1}^{K} x_{k,n} \le 1,\quad n=1,\cdots,N, \\ & x_{k,n}\in\{0,1\},\ k=1,\cdots,K,\ n=1,\cdots,N, \\ & y_{k,n} \ge 0,\ k=1,\cdots,K,\ n=1,\cdots,N.\end{aligned} \tag{14}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

Notice that any feasible point $(x_{k,n}, p_{k,n})$ in (13) represents a feasible solution $(x_{k,n}, y_{k,n})$ for (14), and vice versa. In addition, the objective function in the limit case admits the following expression:

$$\begin{aligned}&\lim_{x_{k,n}\to 0} x_{k,n}\log\left(1+\frac{g_{k,n}^{SS}y_{k,n}}{\sigma_\nu^2 x_{k,n}}\right)\\ =&\lim_{x_{k,n}\to 0} x_{k,n}\log\left(x_{k,n}+g_{k,n}^{SS}y_{k,n}/\sigma_\nu^2\right)-x_{k,n}\log\left(x_{k,n}\right)\\ =&\ 0,\end{aligned} \tag{15}$$

due to the fact that $\lim_{x\to 0} x\log(x) = 0$. Therefore, it can be claimed that optimization problems (13) and (14) are equivalent. In other words, the optimal solution $(x^*_{k,n}, y^*_{k,n})$ for (14) is also optimal for (13), and vice versa.

The time-sharing method is then used by relaxing the binary variable $x_{k,n} \in \{0, 1\}$ in (14) to $0 \le x_{k,n} \le 1$. Even though the time-sharing method is employed herein paper, unlike reference [17], we do not allow multiple SUs to share a common channel. The time-sharing method is just utilized to conduct the continuous relaxation of the MINLP problem such that the convex optimization tools can be implemented. Thus, the problem (14), or equivalently (13) is relaxed to

$$\begin{aligned}\max_{x_{k,n},y_{k,n}}\quad&\sum_{k=1}^{K}\sum_{n=1}^{N}P_n^0(1-P_n^F)x_{k,n}\log\left(1+\frac{g_{k,n}^{SS}y_{k,n}}{\sigma_\nu^2 x_{k,n}}\right)\\ \text{s.t.}\quad&\sum_{k=1}^{K}\sum_{n\in S_l}P_n^1(1-P_n^D)y_{k,n}g_{l,n}^{SP}\le I_l^{\max},\ l=1,\cdots,L\\ &\sum_{k=1}^{K}y_{k,n}\le P_n^T,\quad \sum_{k=1}^{K}x_{k,n}\le 1,\quad n=1,\cdots,N\\ &x_{k,n}\ge 0, y_{k,n}\ge 0,\quad k=1,\cdots,K,\quad n=1,\cdots,N\end{aligned} \tag{16}$$

Since constraint $\sum_{k=1}^{K} x_{k,n} \le 1$ implicitly implies that $x_{k,n} \le 1$, the relaxation condition is simply expressed as $x_{k,n} \ge 0$ instead of $0 \le x_{k,n} \le 1$. It can be readily inferred that $x_{k,n}\log(1 + g^{SS}_{k,n}y_{k,n}/(\sigma^2_\nu x_{k,n}))$ is a concave function with respect to $x_{k,n}$ and $y_{k,n}$, since it is the perspective function of the concave logarithm function. The objective function of (16), being the positive summation of concave functions, is also concave. Hence, the optimization problem (16) is a convex optimization problem due to the fact that all its constraints are affine.

We tackle the problem in the dual domain, which leads to the adoption of the Lagrange function

$$\begin{aligned}Lg=&\sum_{k=1}^{K}\sum_{n=1}^{N}P_n^0(1-P_n^F)x_{k,n}\log\left(1+\frac{g_{k,n}^{SS}y_{k,n}}{\sigma_\nu^2 x_{k,n}}\right)\\ &+\sum_{n=1}^{N}\lambda_n\left(P_n^T-\sum_{k=1}^{K}y_{k,n}\right)+\sum_{n=1}^{N}\eta_n\left(1-\sum_{k=1}^{K}x_{k,n}\right)\\ &+\sum_{l=1}^{L}\mu_l\left(I_l^{\max}-\sum_{k=1}^{K}\sum_{n\in S_l}P_n^1(1-P_n^D)y_{k,n}g_{l,n}^{SP}\right),\end{aligned} \tag{17}$$

where the last constraint in (16) is absorbed in the Karush-Kuhn-Tucker (KKT) conditions. Applying KKT conditions [27] leads to

$$\begin{aligned}\frac{\partial Lg}{\partial y_{k,n}}=&P_n^0(1-P_n^F)\frac{g_{k,n}^{SS}x_{k,n}}{\sigma_\nu^2 x_{k,n}+g_{k,n}^{SS}y_{k,n}}-\lambda_n\\ &-P_n^1(1-P_n^D)\mu_l g_{l,n}^{SP}\begin{cases}=0, & y_{k,n}>0,\\ <0, & y_{k,n}=0,\end{cases}\end{aligned}$$

with $n \in S_l$. It turns out that

$$y_{k,n}=\left(\psi_n-\frac{\sigma_\nu^2}{g_{k,n}^{SS}}\right)^+ x_{k,n}, \tag{18}$$

where $\psi_n = \frac{P^0_n(1-P^F_n)}{\lambda_n+P^1_n(1-P^D_n)\mu_l g^{SP}_{l,n}}$ with $n \in S_l$, and the operator $(\cdot)^+ = \max(\cdot, 0)$. Moreover, taking the derivative of the Lagrange function with respect to $x_{k,n}$ and plugging (18) into the result yields

$$\frac{\partial Lg}{\partial x_{k,n}}=J_{k,n}-\eta_n\begin{cases}=0, & x_{k,n}>0,\\ <0, & x_{k,n}=0,\end{cases} \tag{19}$$

where

$$\begin{aligned}J_{k,n}=&P_n^0(1-P_n^F)\left[\log\left(1+\frac{g_{k,n}^{SS}\left(\psi_n-\sigma_\nu^2/g_{k,n}^{SS}\right)^+}{\sigma_\nu^2}\right)\right.\\ &\left.-\frac{g_{k,n}^{SS}(\psi_n-\sigma_\nu^2/g_{k,n}^{SS})^+}{\sigma_\nu^2+g_{k,n}^{SS}(\psi_n-\sigma_\nu^2/g_{k,n}^{SS})^+}\right].\end{aligned} \tag{20}$$

Based to the aforementioned equations, the following result can be established.

Lemma 1. *All the channels are assigned to SUs. That is to say, for any channel n, there exist some SUs such that $x_{k,n} > 0$. In the case that the number of SUs is less than the number of available channels, then some SUs will be assigned multiple channels. In addition, $\eta_n = \max_k J_{k,n},\ n = 1, \cdots, N$.*

Proof. See Appendix A. □

Theorem 1. *Under the assumption that for any channel n, $g^{SS}_{k,n}$'s are all distinct for $k = 1, \cdots, K$, the problem* (16) *always achieve a binary optimal solution for the user assignment variable $x_{k,n}$. Particularly, each channel is solely assigned to the SU with the largest $g^{SS}_{k,n}$, i.e.,*

$$x_{k_n^*,n}=1|_{k_n^*=\arg\max_k g_{k,n}^{SS}},\quad x_{k,n}=0|_{k\ne k_n^*},\quad n=1,\cdots,N, \tag{21}$$

Proof. See Appendix B. □

Generally, SUs are randomly located around the SBS and the probability that two SUs present exactly the same power gain $g^{SS}_{k,n}$ is almost zero. Therefore, the optimization problem (16) always achieve a binary optimal solution for the user assignment variable. It can be further claimed that the optimal binary solution in (21) for problem (16) is also optimal for problem (13). This argument can be proved in two steps: first, the optimal objective value of (16) serves as an upper-bound for (13) due to the fact that (16) is a relaxed version of (13);

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

second, since the optimal solution (21) for (16) is binary, it represents a feasible solution of (13) and consequently achieves a lower-bound for the optimal objective value of (13).

Towards this end, it turns out that for a fixed γ_n, the optimization problem (13) has an optimal solution (21) for the user assignment variable $x_{k,n}$. Since this statement holds for any feasible γ_n, by plugging (21) into the formulation of P-J, it can be concluded that the original joint optimization P-J can be simplified to

$$\max_{p_{k_n^*,n},\gamma_n} \quad \sum_{n=1}^{N} P_n^0(1-P_n^F)\log\left(1+\frac{g_{k_n^*,n}^{SS}p_{k_n^*,n}}{\sigma_\nu^2}\right) \quad \text{(P-JE)}$$

$$\text{s.t.} \quad \sum_{n\in S_l} P_n^1(1-P_n^D)p_{k_n^*,n}g_{l,n}^{SP} \le I_l^{\max},\ l=1,\cdots,L \quad (22)$$

$$0 \le p_{k_n^*,n} \le P_n^T,\ n=1,\cdots,N \quad (23)$$

$$\gamma_n^{\min} \le \gamma_n \le \gamma_n^{\max},\ n=1,\cdots,N, \quad (24)$$

where $k_n^* = \arg\max_k g_{k,n}^{SS},\ n=1,\cdots,N$. Since the above optimization problem is equivalent to P-J, it is termed as "P-JE" herein paper. For the rest of this paper, we focus on addressing P-JE since its dimension is significantly less than that of P-J and the integer variable $x_{k,n}$ has been removed.

IV. Optimal Joint Spectrum Sensing and Resource Allocation Algorithm

Due to the nonconvex property of the objective function, P-JE is still a nonconvex optimization problem with respect to $p_{k_n^*,n}$ and γ_n. However, it can be readily found that P-JE is a convex optimization problem with regard to $p_{k_n^*,n}$ while fixing γ_n. Moreover, it also admits the formulation of a convex optimization in terms of γ_n while fixing $p_{k_n^*,n}$. This is due to the fact that P_n^F and $1-P_n^D$ are convex functions with respect to γ_n under the constraint (8), or equivalently (24) [7]. In literature, the nonconvex optimization with such kind of formulation is typically solved by means of the AO method [16], [17], where the sensing threshold and power variable are iteratively optimized while fixing the other. However, as discussed earlier, the convergence of the AO method heavily depends on the initial point and the iterative algorithm can be trapped at a local maximum near the starting point. Therefore, in this section, we propose an algorithm by exploiting the knowledge of a monotonic optimization to solve the joint optimization problem P-JE. It is shown that the problem admits the canonical form of a monotonic optimization, which can be addressed in a finite number of steps with an ϵ-optimal solution[1].

A. Some Preliminaries on Monotonic Optimization

To prepare the background, some basic preliminaries on monotonic optimization are presented herein subsection. The interested readers can refer to [28], [29] for more details and proofs.

[1] $\bar{\mathbf{x}}$ is an ϵ-optimal solution if $f(\mathbf{x}^*)-\epsilon \le f(\bar{\mathbf{x}}) \le f(\mathbf{x}^*)$, where $\mathbf{x}^*$ represents the optimal solution.

Definition 1. *A function f ($\mathcal{R}_+^m \to \mathcal{R}$) is increasing if $f(\mathbf{x}) \le f(\mathbf{y})$ when[2] $\mathbf{x} \le \mathbf{y}$.*

Definition 2. *A set $\mathcal{G} \subset \mathcal{R}_+^m$ is normal if $\mathbf{x}\in\mathcal{G} \Rightarrow [\mathbf{0},\mathbf{x}] \subset \mathcal{G}$.*

Definition 3. *A set $\mathcal{H} \subset \mathcal{R}_m^+$ is conormal in $[\mathbf{0},\mathbf{b}]$ if $\mathbf{x}\in\mathcal{H} \Rightarrow [\mathbf{x},\mathbf{b}] \subset \mathcal{H}$.*

Definition 4. *Monotonic optimization is referred to as a problem with the following structure:*

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{G}\cap\mathcal{H}, \end{aligned} \quad (25)$$

where $f(\mathbf{x})(\mathcal{R}_+^m \to \mathcal{R})$ is an increasing function, $\mathcal{G} \subset [\mathbf{0},\mathbf{b}] \subset \mathcal{R}_m^+$ is a normal set with nonempty interior, and $\mathcal{H}$ is a closed conormal set in $[\mathbf{0},\mathbf{b}]$.

B. Polyblock Outer Approximation Algorithm

Let $[\mathbf{p},\boldsymbol{\gamma}] = [p_{k_1^*,1},\cdots,p_{k_N^*,N},\gamma_1,\cdots,\gamma_N]$ be a vector with length $2N$, the objective function for P-JE is an increasing function with respect to $[\mathbf{p},\boldsymbol{\gamma}]$ since $1-P_n^F$ is an increasing function of γ_n and the logarithm is an increasing function of $p_{k_n^*,n}$. Moreover, in the region

$$\left[\mathbf{0},[P_1^T,\cdots,P_N^T,\gamma_1^{\max},\cdots,\gamma_N^{\max}]\right], \quad (26)$$

the normal set $\mathcal{G}$ and the conormal set $\mathcal{H}$ can be expressed as

$$\mathcal{G} = \left\{(\mathbf{p},\boldsymbol{\gamma})|(22),(23),0\le\gamma_n\le\gamma_n^{\max}\right\},$$
$$\mathcal{H} = \left\{(\mathbf{p},\boldsymbol{\gamma})|p_{k_n^*,n}\ge 0,\gamma_n\ge\gamma_n^{\min}\right\}.$$

The convergence of the polyblock outer approximation algorithm is ensured under mild conditions [28], namely, $f(\mathbf{x})$ is upper semicontinuous, $\mathcal{G}$ has a nonempty interior, and $\mathcal{H}$ is strictly bounded away from $\mathbf{0}$, i.e., there exists a positive vector $\mathbf{a}$ such that

$$\mathbf{0} < \mathbf{a} \le \mathbf{x},\ \forall \mathbf{x}\in\mathcal{H}. \quad (27)$$

The first two conditions can be readily proved since semicontinuity is a weaker requirement than continuity and the interior of $\mathcal{G}$ is clearly nonempty. However, it can be observed that the 3rd condition does not hold for our formulation due to the fact that $p_{k_n^*,n}$ can be zero in $\mathcal{H}$. To address this issue, we set a small enough number δ as a lower-bound for $p_{k_n^*,n}$. In this way, the resulting problem satisfies the condition (27) while differing only slightly from the original problem [29]. In addition, since numerical simulations can only achieve finite precision, it is a general custom to set up a small enough number δ, i.e., $p_{k_n^*,n} \ge \delta$, to enforce the validity of the constraint $p_{k_n^*,n} \ge 0$. Hence, the simulation error for this approximation can be neglected. Alternatively, one can change the variable $p_{k_n^*,n}$, for example, with $p'_{k_n^*,n} = p_{k_n^*,n}+1$. As a result, $p'_{k_n^*,n}$ is bounded away from zero. This change of variable action can be also regarded as shifting the origin $[0,\cdots,0]$ of $p_{k_n^*,n}$ to a negative coordinate $[-1,\cdots,-1]$ such that the conormal set $\mathcal{H}$ is strictly bounded away from the new origin. Therefore, this method is named as "shift of origin" in [28], [30].

[2] For two vectors $\mathbf{x},\mathbf{y}\in\mathcal{R}^m$, we say $\mathbf{x}\le\mathbf{y}$ if $x_i \le y_i,\ i=1,\cdots,m$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

The proposed polyblock outer approximation algorithm for P-JE is summarized in Algorithm 1. The basic idea of polyblock outer approximation algorithm is to enclose the feasible set $\mathcal{G} \cap \mathcal{H}$ as closely as possible by a nested sequence of "polyblocks", starting with the polyblock $\mathcal{P}_1$ in (26). The starting polyblock $\mathcal{P}_1$ is illustrated in Fig. 3(a) where a $2N$ dimensional figure is plotted in a two-dimension plane for simplicity. Among all the vertices of the polyblock (in $\mathcal{P}_1$, we only have one vertex $\mathbf{z}_1$), we choose the one with the largest objective value (Line 3) and project this vertex onto the upper-boundary of $\mathcal{G}$ (Line 4). The projection process is mathematically defined as the following single-variable optimization problem $\alpha = \arg\max \alpha|_{\alpha>0, \alpha\mathbf{z}_i \in \mathcal{G}}$, which can be solved efficiently by the bisection search algorithm due to the normality of $\mathcal{G}$. The projection of the best vertex is termed as $\pi_{\mathcal{G}}(\mathbf{z}_i)$ in Algorithm 1. From Line 5 to Line 11, the algorithm determines the current best solution and the current best value (CBV), and store them as $[\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]$ and CBV, respectively. In Line 12, a new polyblock $\mathcal{P}_{i+1}$ with vertices $\mathcal{T}_{i+1}$ is generated by cutting off a cone that is in the infeasible set. In particular, the best vertex $\mathbf{z}_i$ is removed and $2N$ new vertices $\{\mathbf{v}_i,\ i = 1, \cdots, 2N\}$ are added. The newly-added vertices together with the remaining vertices $\mathcal{T}_i \backslash \mathbf{z}_i$ generate the new polyblock. For illustration, the new polyblock $\mathcal{P}_2$, which comes from the polyblock $\mathcal{P}_1$, is depicted in Fig. 3(b). In Line 13, improper vertices[3] are removed to accelerate the convergence speed and save memory. Due to the monotonically increasing property of the objective function, removing improper vertices does not affect the shape of the polyblock [28]. This procedure is repeated until the current best value is close to the best vertex value, i.e., $|f(\mathbf{z}_i) - \text{CBV}_i| \le \epsilon$. In this way, a sequence of polyblocks

$$\mathcal{P}_1 \supset \mathcal{P}_2 \supset \cdots \supset \mathcal{P}_i \supset \cdots \supset \mathcal{G} \cap \mathcal{H}$$

is constructed and the current best solution $[\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]$ is guaranteed to be ϵ-optimal.

The dimension of the problem, which directly determines the size of the polyblock vertex set, is a key factor for the computational complexity of the monotonic optimization. Additionally, the form of constraints also affects the complexity since the feasibility is required to be checked in each iteration. As an under-explored research area, limited computational experience has shown that polyblock algorithms work very well for monotonic optimization problems that present a small dimension m, generally $m \le 10$ [14]. The reader can find more details about monotonic optimization as well as its applications in communication and networking systems in [28] and references therein.

The problem P-JE, at first glance, has a dimension $2N$. However, since the licensed channels for PUs are non-overlapping, i.e., $S_i \cap S_j = 0,\ i \ne j$ and all the channels are licensed to PUs, i.e., $S_1 \cup \cdots \cup S_l = \{1, \cdots, N\}$, P-JE in fact can be decomposed into L independent optimization problems

[3] Let $\mathcal{T}$ be the vertex set of a polyblock, a vertex $\mathbf{v} \in \mathcal{T}$ is proper if there does not exist $\mathbf{v}' \ne \mathbf{v}$ and $\mathbf{v}' \in \mathcal{T}$ such that $\mathbf{v}' > \mathbf{v}$. A vertex is improper if it is not proper.

Algorithm 1 Polyblock outer approximation algorithm for problem P-JE

Input: Function f: The objective function in P-JE.

$$\mathcal{G} = \{(\mathbf{p}, \boldsymbol{\gamma}) | (22), (23), 0 \le \gamma_n \le \gamma_n^{\max}\},$$
$$\mathcal{H} = \left\{(\mathbf{p}, \boldsymbol{\gamma}) | p_{k_n^*, n} \ge \delta, \gamma_n \ge \gamma_n^{\min}\right\}$$

Output: An ϵ-optimal solution $[\mathbf{p}^*, \boldsymbol{\gamma}^*]$

Initialization: Error tolerance $\epsilon > 0$. $\text{CBV}_0 = -\infty$. $i = 0$. $[\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_0 = \mathbf{0}$. The vertex set $\mathcal{T}_1 = [P_1^T, \cdots, P_N^T, \gamma_1^{\max}, \cdots, \gamma_N^{\max}]$.

1: **while** $|f(\mathbf{z}_i) - \text{CBV}_i| \le \epsilon$ **do**
2: $\quad i = i + 1$
3: $\quad$ From $\mathcal{T}_i$, select $\mathbf{z}_i = \arg\max f(\mathbf{z})|_{\mathbf{z} \in \mathcal{T}_i}$.
4: $\quad$ Find $\pi_{\mathcal{G}}(\mathbf{z}_i) = \alpha \mathbf{z}_i,\ \alpha = \arg\max \alpha|_{\alpha>0, \alpha\mathbf{z}_i \in \mathcal{G}}$.
5: $\quad$ **if** $\pi_{\mathcal{G}}(\mathbf{z}_i) = \mathbf{z}_i$, i.e., $\mathbf{z}_i \in \mathcal{G}$ **then**
6: $\quad\quad [\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_i = \mathbf{z}_i$ and $\text{CBV}_i = f(\mathbf{z}_i)$.
7: $\quad$ **else if** $\pi_{\mathcal{G}}(\mathbf{z}_i) \in \mathcal{H}$ and $f(\pi_{\mathcal{G}}(\mathbf{z}_i)) \ge \text{CBV}_{i-1}$ **then**
8: $\quad\quad [\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_i = \pi_{\mathcal{G}}(\mathbf{z}_i)$ and $\text{CBV}_i = f(\pi_{\mathcal{G}}(\mathbf{z}_i))$.
9: $\quad$ **else**
10: $\quad\quad [\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_i = [\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_{i-1}$ and $\text{CBV}_i = \text{CBV}_{i-1}$.
11: $\quad$ **end if**
12: $\quad \mathcal{T}_{i+1} = (\mathcal{T}_i \backslash \mathbf{z}_i) \cup \{\mathbf{v}_i,\ i = 1, \cdots, 2N\}$, where $\mathbf{v}_i$ is obtained by replacing the ith entry of $\mathbf{z}_i$ by the ith entry of $\pi_{\mathcal{G}}(\mathbf{z}_i)$.
13: $\quad$ Remove from $\mathcal{T}_{i+1}$ all the improper vertices.
14: **end while**
15: Let $[\mathbf{p}^*, \boldsymbol{\gamma}^*] = [\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_i$ and terminate the algorithm.

whose dimensions are $2|S_1|, \cdots, 2|S_L|$, respectively. Specifically, these L independent optimization problems are

$$\begin{aligned}
\max_{p_{k_n^*, n}, \gamma_n} \quad & \sum_{n \in S_l} P_n^0 (1 - P_n^F) \log\left(1 + \frac{g_{k_n^*, n}^{SS} p_{k_n^*, n}}{\sigma_\nu^2}\right) \\
\text{s.t.} \quad & \sum_{n \in S_l} P_n^1 (1 - P_n^D) p_{k_n^*, n} g_{l,n}^{SP} \le I_l^{\max} \\
& 0 \le p_{k_n^*, n} \le P_n^T,\ n \in S_l \\
& \gamma_n^{\min} \le \gamma_n \le \gamma_n^{\max},\ n \in S_l,
\end{aligned}$$

with $l = 1, \cdots, L$. In this way, as long as the largest number of channels occupied by licensed PUs is small, the monotonic optimization technique can be used to achieve an ϵ-optimal solution.

V. Low-complexity Suboptimal Algorithm

Algorithm 1 provides an ϵ-optimal solution for the joint spectrum sensing and resource allocation problem P-JE. However, the proposed polyblock outer approximation algorithm is typically suitable for a small-dimension monotonic optimization problem and the computational complexity is very high in a CR network with a large number of channels, such as the OFDM wireless system. In this section, we propose a low-complexity suboptimal algorithm by first modifying the constraint of P-JE and then solving a list of single-variable optimization problems.

Instead of limiting the total interference to PUs, we alternatively impose a threshold for the interference to each licensed

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

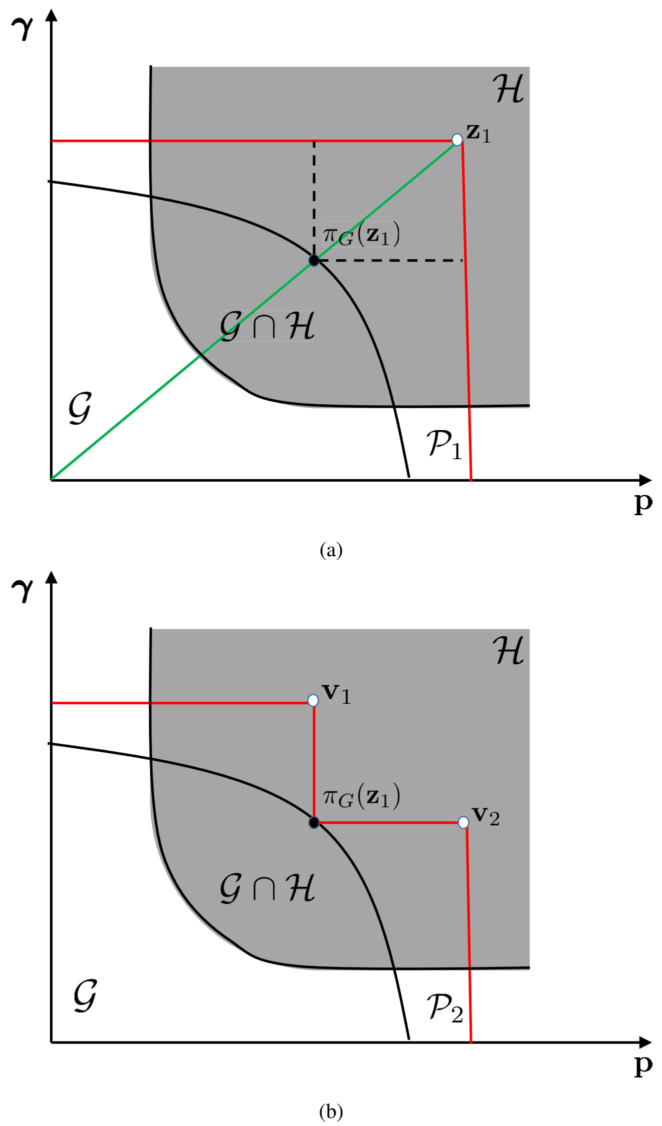

Fig. 3. (a) The initialized polyblock $\mathcal{P}_1$; (b) The polyblock $\mathcal{P}_2$ generated from $\mathcal{P}_1$.

channel of PUs. Mathematically, the modified interference constraint is expressed as

$$P_n^1(1-P_n^D)p_{k_n^*,n}g_{l,n}^{SP} \leq I_l^{\max}/|S_l|, \ \forall n \in S_l, \ l = 1, \cdots, L. \tag{28}$$

Thus, the joint spectrum sensing and resource allocation problem can be reformulated as

$$\begin{aligned} \max_{p_{k_n^*,n},\gamma_n} \quad & \sum_{n=1}^{N} P_n^0(1-P_n^F)\log\left(1+\frac{g_{k_n^*,n}^{SS}p_{k_n^*,n}}{\sigma_\nu^2}\right) \\ \text{s.t.} \quad & (28), (23), (24) \end{aligned} \tag{29}$$

It can be inferred that constraint (28) is more stringent compared to constraint (22). Thus, (29), being a modified version of P-JE, provides a lower-bound for P-JE. Moreover, with the help of the modified constraint (28), (29) can be decomposed into N independent subproblems as shown below

$$\begin{aligned} \max_{p_{k_n^*,n},\gamma_n} \quad & P_n^0(1-P_n^F)\log\left(1+\frac{g_{k_n^*,n}^{SS}p_{k_n^*,n}}{\sigma_\nu^2}\right) \\ \text{s.t.} \quad & (28), (23), (24) \end{aligned} \tag{P-JEM}$$

with $n = 1, \cdots, N$. In this way, we only need to solve N optimization problems, each of which presents two variables $p_{k_n^*,n}$ and γ_n.

To solve P-JEM, it is first claimed that either constraint (28) or (23) achieves the equality at the optimal solution. This claim can be easily verified via a proof by contradiction. First, it is assumed that neither constraint achieves the equality at the optimal solution. Then, we can definitely add more power to this channel to achieve the equality at either constraint while obtaining a larger objective value, which completes the proof by contradiction. Consequently, $p_{k_n^*,n}$ can be written in terms of γ_n as follows:

$$\begin{aligned} p_{k_n^*,n} &= \min\left(\frac{I_l^{\max}}{|S_l|P_n^1(1-P_n^D)g_{l,n}^{SP}}, P_n^T\right) \\ &= \frac{I_l^{\max}}{\max\left(|S_l|P_n^1(1-P_n^D)g_{l,n}^{SP}, I_l^{\max}/P_n^T\right)}, \end{aligned} \tag{30}$$

and P-JEM resumes to a single-variable optimization problem

$$\begin{aligned} \max_{\gamma_n} \quad & f(\gamma_n) - g(\gamma_n) \\ \text{s.t.} \quad & \gamma_n^{\min} \leq \gamma_n \leq \gamma_n^{\max}, \end{aligned} \tag{31}$$

with

$$f(\gamma_n) = P_n^1(1-P_n^F)\log(\sigma_\nu^2 + g_{k_n^*,n}^{SS}I_l^{\max}),$$

and

$$\begin{aligned} g(\gamma_n) =& P_n^1(1-P_n^F)\times \\ & \log\left[\sigma_\nu^2 \cdot \max\left(|S_l|P_n^1(1-P_n^D)g_{l,n}^{SP}, \frac{I_l^{\max}}{P_n^T}\right)\right]. \end{aligned}$$

Problem (31) can be solved using a one-dimensional brute-force algorithm. To avoid the exhaustive search, in the rest of this section, we show that it can also be addressed using the monotonic optimization technique described in Section IV.

Both P_n^D and P_n^F are monotonically decreasing functions in terms of γ_n. As a result, the cost function in (31) is the difference between two increasing functions $f(\gamma_n)$ and $g(\gamma_n)$. It is shown in Appendix C that problem (31) can be rewritten as the canonical form of a monotonic optimization by introducing a new parameter t:

$$\begin{aligned} \max \quad & f(\gamma_n) + t \\ \text{s.t.} \quad & (\gamma_n, t) \in \mathcal{G} \cap \mathcal{H}, \end{aligned} \tag{32}$$

with

$$\begin{aligned} \mathcal{G} =& \Big\{(\gamma_n, t)\Big| 0 \leq \gamma_n \leq \gamma_n^{\max}, 0 \leq t \leq g\left(\gamma_n^{\max}\right) - g\left(\gamma_n^{\min}\right), \\ & g\left(\gamma_n\right) + t \leq g\left(\gamma_n^{\max}\right)\Big\}, \end{aligned} \tag{33}$$

and

$$\mathcal{H} = \{(\gamma_n, t)|\gamma_n \geq \gamma_n^{\min}, t \geq 0\}. \tag{34}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

We can implement the same technique as in Section IV to achieve an ϵ-optimal solution for variable γ_n in problem (32), or equivalently P-JEM. Then the power solution can be obtained using (30).

VI. Numerical Results and Discussions

For the numerical results presented in this section, we consider two different kinds of scenarios, namely a CR network with a small number of channels and one with a large number of channels. For the former scenario, we compare the simulation performance among three algorithms: i) The proposed optimal monotonic optimization algorithm in Section IV; ii) The proposed suboptimal monotonic optimization algorithm in Section V; and iii) The state-of-the-art AO method in literature, which is briefly discussed in the beginning of Section IV. For the later scenario, only the simulation for the second and third algorithms aforementioned is carried out, since the optimal monotonic optimization algorithm does not work very well for a large size problem. However, notice that the suboptimal solution for P-JEM can be further enhanced by plugging it back into P-JE as an initial point of the AO method. Therefore, this scheme is termed as "suboptimal-enhanced" and included as a benchmark for the simulation in a large-dimension CR network. As discussed earlier in the paper, some references studied the CR problem with a fixed sensing parameter or a fixed physical layer resource layout. We do not consider these set-ups in this section since the joint spectrum sensing and resource allocation problem obviously outperforms these set-ups from a mathematical point of view.

In order to easily follow the numerical results, the parameter settings are summarized in Table II for each figure in this section. Specifically, the simulation performed for Figs. 4 - 8 are drawn in Subsection VI-A while Figs. 9 and 10 present the numerical results in Subsection VI-B. The channel bandwidth is assumed to be 100kHz. With the noise spectral density being assumed at 10^{-13}W/Hz, the noise power σ_ν^2 is set to -50dBm. Furthermore, we assume $D_n^{\min} = F_n^{\max} = 0.5$ to achieve the best trade-off between the false alarm and misdetection rates.

A. A Small-Dimension CR Network

In this subsection, the licensed channels for these three PUs are 1-2, 3-4 and 5-6, respectively. The PU activity probability P_n^1 is set to be 0.2 for all the channels. The channels from the SBS to SUs, as well as the channels from the SBS to PUs, are assumed to be Rayleigh fading with an average channel power gain -4dB. Thus, $g_{k,n}^{SS}$ and $g_{l,n}^{SP}$ are both distributed exponentially with mean -4dB.

Since the channel gains are randomly generated, an average result of 100 independent Monte-Carlo runs is carried out for calculating the average total throughput of the secondary network. In Fig. 4, the average total throughput of SUs is simulated with a power limit P_n^T ranging from -30dBm to -10dBm for all channels and an interference threshold $I_l^{\max} = -50$dBm, $l = 1, \cdots, L$. Moreover, we consider the sensing signal-to-noise ratio (SNR) at the SBS in two cases. Specifically, the average sensing SNR, $\Gamma = \mathrm{E}[g_n^{PS}]\sigma_s^2/\sigma_\nu^2$, at the SBS are set to be 0dB and 6dB for all channels, also

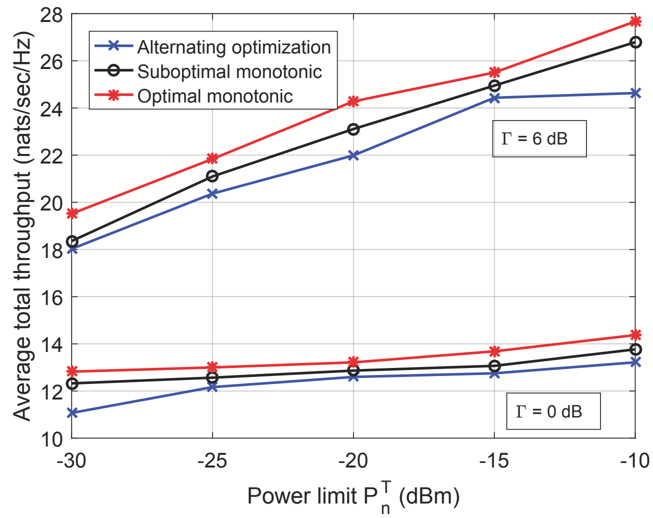

Fig. 4. Average total throughput of SUs in terms of the power limit and the sensing SNR ($N = 6$).

with Rayleigh fading assumption. It can be seen that with the power limit increasing, the average total SU throughput also increases. In addition, the suboptimal monotonic optimization algorithm outperforms the AO method and is very close to the optimal solution in both sensing SNRs. The same trend can be observed in Fig. 5 with $P_n^T = -15$dBm and $I_l^{\max}$ ranging from -50dBm to -30dBm. For both Fig. 4 and Fig. 5, the error tolerance ϵ for the monotonic algorithms is set to 0.05, which only allows a 0.05 difference from the optimal objective value. The convergence analysis of the proposed monotonic algorithms is depicted in Fig. 6. It can be observed that the number of iterations for the optimal monotonic algorithm grows up dramatically as the error tolerance decreases. On the other hand, since the suboptimal monotonic algorithm resumes to solving a list of monotonic optimization problems (32) with only γ_n and an auxiliary variable t, the convergence time is significantly less than the optimal monotonic algorithm.

In Fig. 7 and Fig. 8, the simulation with one realization of channel gains is performed coupled with parameters $P_n^T = -15$dBm, $I_l^{\max} = -50$dBm and $\Gamma = 6$dB. Particularly, in Fig. 7, the probabilities of misdetection and false alarm, together with the allocated power, for each individual channel are illustrated. For better illustration, the unit of power in this figure is modeled as watts (W) instead of dBm. For the 1st, 3rd, 5th and 6th channel where the probabilities of misdetection and false alarm are low, the maximum power P_n^T is assigned. In contrast, for the 2nd channel where the sensing errors are relatively high, a very low power is allocated to limit the inference introduced to PUs. Intuitively, a channel with high probabilities of misdetection and false alarm can be regarded as an unreliable channel. Therefore, a very low or no power is allocated to this channel to prevent the highly potential interference. From a mathematical perspective, the coefficients associated with the logarithm terms are $P_n^0(1 - P_n^F)$. As a result, the power is more likely to be allocated to the channel with a small false alarm rate. Moreover, based on (22), allocating more power to a channel with high misdetection

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

TABLE II
PARAMETER SETTINGS

Parameter	Fig. 4	Fig. 5	Fig. 6	Fig. 7	Fig. 8	Fig. 9	Fig. 10
L	3	3	3	3	3	4	4
K	3	3	3	3	3	10	10
N	6	6	6	6	6	40	40
M	10	10	10	10	10	10	10
σ_ν^2 (dBm)	-50	-50	-50	-50	-50	-50	-50
Γ (dB)	6/0	6/0	6	6	6	6/0	6/0
$g_{k,n}^{SS}$ (Exp with mean in dB)	-4	-4	-4	-4	-4	-4	-4
$g_{l,n}^{SP}$ (Exp with mean in dB)	-4	-4	-4	-4	-4	-4	-4
$I_l^{\max}$ (dBm)	-50	[-50,-30]	-50	-50	-50	-50	[-50,-30]
P_n^T (dBm)	[-30,-10]	-15	-15	-15	-15	[-30,-10]	-15
$D_n^{\min}$	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$F_n^{\max}$	0.5	0.5	0.5	0.5	0.5	0.5	0.5
P_n^1	0.2	0.2	0.2	0.2	0.2	0.2	0.2
ϵ	0.05	0.05	[0.01,0.3]	0.05	0.05	0.05	0.05

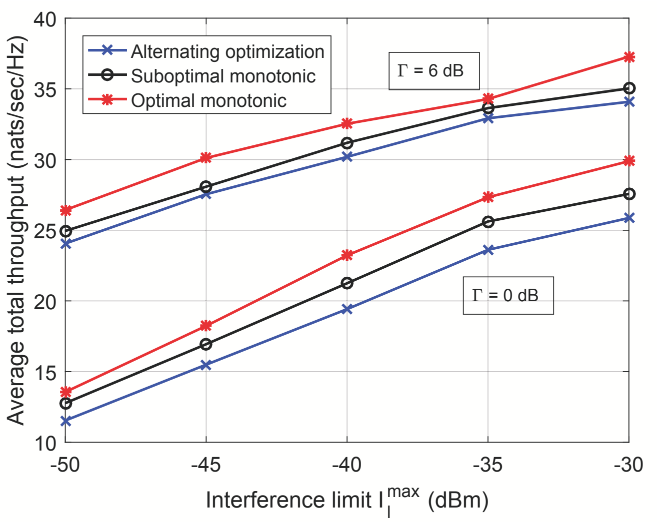

Fig. 5. Average total throughput of SUs in terms of the interference threshold and the sensing SNR ($N = 6$).

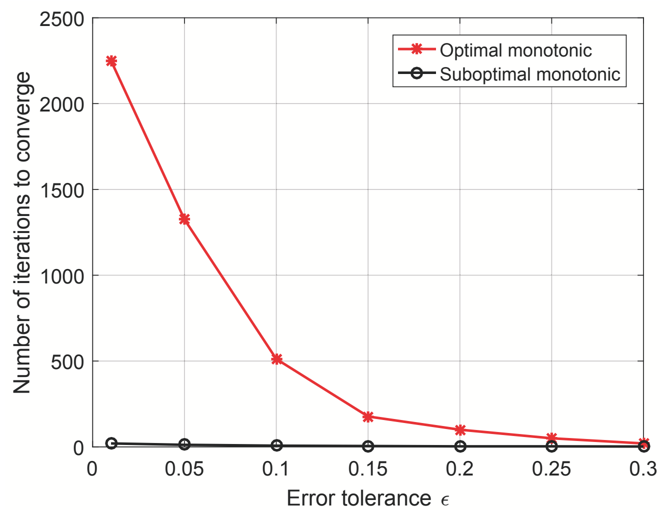

Fig. 6. Convergence of the optimal and suboptimal monotonic algorithms

probability may yield a large interference to the primary user. In Fig. 8, 10 independent simulations are performed of these algorithms with a fixed realization of channel gains. Since both the optimal monotonic and the suboptimal monotonic algorithms achieve a solution at most ϵ away from its own optimal solution, their performances keep consistent. On the other hand, the performance of AO method varies a lot due to different selections of the initial guess.

B. A Large-Dimension CR Network

A simulation parameter setup similar to the one in Subsection VI-A is implemented for a large-dimension CR network. In particular, the licensed channels for four PUs are 1-10, 11-20, 21-30 and 31-40, respectively.

In Fig. 9, the average total SU throughput is simulated with a power limit P_n^T ranging from -30dBm to -10dBm for $n = 1, \cdots, N$ and an interference limit $I_l^{\max} = -50$dBm, $l = 1, \cdots, L$. Similar to Subsection VI-A, we consider the scenarios of $\Gamma = 0$dB and 6dB. As expected, the increase of power limit results in a higher average throughput for the secondary network. It is also found that the proposed suboptimal algorithm is comparable to the AO method with the enhanced-suboptimal scheme outperforming the others. The same trend can be found in Fig. 10 with $P_n^T = -15$dBm and $I_l^{\max}$ ranging from -50dBm to -30dBm. More simulation results reveal that with the suboptimal monotonic solution as the starting point, the AO method generally converges in less than 5 iterations.

C. Discussions

Based on the numerical results presented in this section, it can be claimed that the proposed optimal and suboptimal monotonic optimization algorithms outperform the AO method in terms of accuracy and stability in the presence of a small-dimension CR network. For a large-dimension CR network, the performance of the suboptimal algorithm is comparable to the AO method. Using the solution provided by the suboptimal

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2807432, IEEE Transactions on Communications

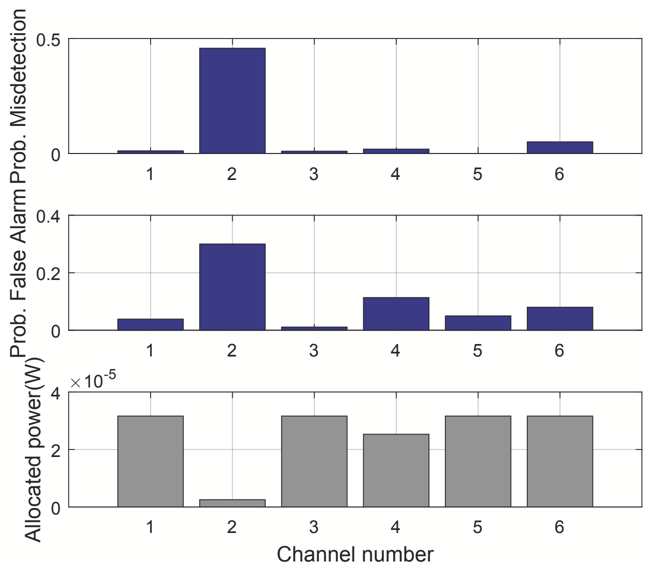

Fig. 7. The probabilities of misdetection and false alarm, and the allocated power on individual channels.

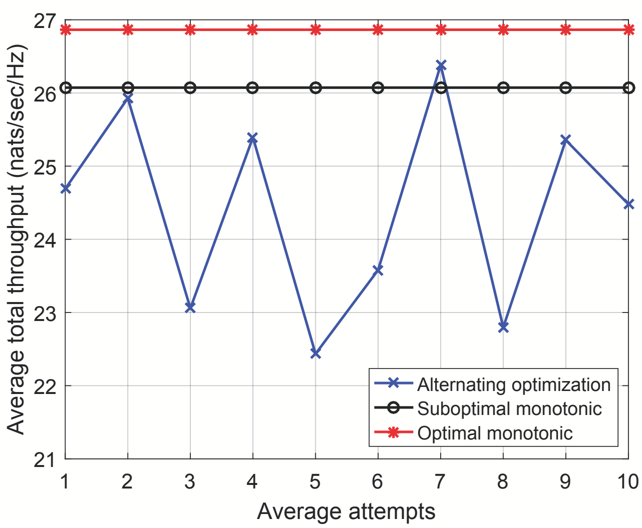

Fig. 8. Stability of algorithms.

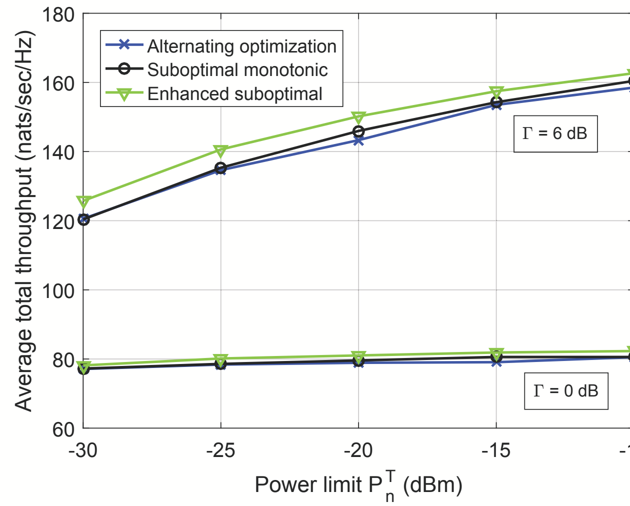

Fig. 9. Average total throughput of SUs in terms of the power limit and the sensing SNR ($N = 40$).

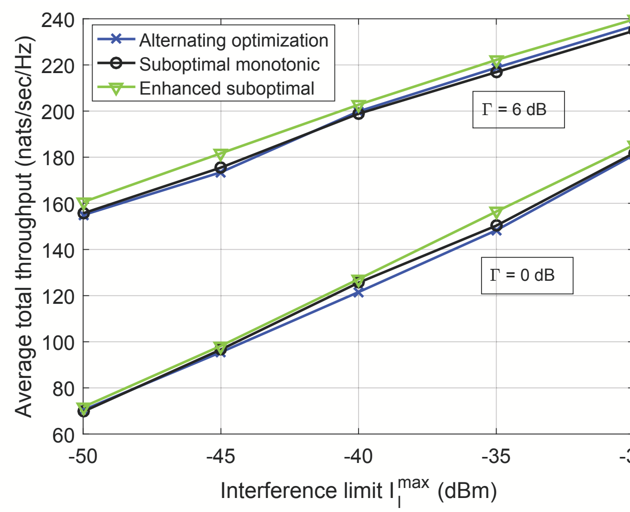

Fig. 10. Average total throughput of SUs in terms of the interference threshold and the sensing SNR ($N = 40$).

monotonic algorithm as the starting point, the suboptimal-enhanced algorithm achieves a higher objective value than the AO method. Moreover, it is also shown that the suboptimal-enhanced scheme converges in only a few iterations.

In terms of computational complexity, the state-of-the-art AO method is an iterative optimization method where spectrum sensing and resource allocation are optimized in turns. However, in each turn, the optimization problem presents a variable size same to the number of channels in the CR network, which results in a high computational burden in each iteration. Since the monotonic optimization is still under-explored, to our best knowledge, no closed-form complexity analysis exists in literature. However, as discussed earlier, limited computational experience shows that the algorithm is generally applicable for a problem whose size is less than 10. Therefore, such a framework is typically used as a benchmark for other heuristic algorithms or employed in a CR network with a very small number of channels. For the optimal algorithm proposed in Section IV, the monotonic optimization is employed to solve problem P-JE, whose variable size may reach a large number. Thus, the optimal algorithm is only implemented in the case of small CR networks. Finally, the suboptimal algorithm generally exhibits the lowest computational complexity. That is to say, no matter how large the size of problem P-JE is, the suboptimal algorithm decomposes the problem into a list of single-variable exhaustive search problems (31), or equivalently, two-dimension monotonic optimization problems (32), and thus significantly reduces the computational complexity. In addition, the convergence analysis in Fig. 6 also verifies the computational advantages of the low-complexity algorithm compared to the optimal algorithm.

In a nutshell, for a small-dimension CR network, the suboptimal monotonic algorithm yields a near-optimal solution and converges very fast. If the computational time is not considered significant, the optimal monotonic algorithm can also be chosen. In a large-dimension CR network, the proposed suboptimal monotonic algorithm achieves a reasonable performance with a low computational complexity. The suboptimal-enhanced scheme can also be employed to further improve the performance.

VII. Conclusions

This paper studied the problem of designing the sensing threshold and power allocation strategy that maximize the average throughput of a multi-band-multi-user cognitive radio (CR) network. The average interference and the peak power constraints, together with the sensing error constraint, are imposed to seek for the best balance between the protection of PUs and the throughput of CR network. Mathematically, the joint spectrum sensing and resource allocation problem is formulated as a nonconvex mixed-integer nonlinear programming (MINLP) optimization problem. The convex optimization technique is first employed to simplify the MINLP problem, which results in an equivalent formulation without any integer variables. Based on the knowledge of a monotonic optimization, we propose an algorithm to acquire the optimal sensing threshold and power allocation strategy for a CR network with a small number of channels. In addition, a low-complexity suboptimal algorithm is developed to reduce the computational cost, and it is applicable to both small and large dimension CR networks. Numerical results indicate that the proposed optimal and suboptimal algorithms consistently outperform the start-of-the-art alternating optimization (AO) method in a CR network with a small number of channels. For a large-dimension CR network, the suboptimal algorithm also represents a good choice due to its low computational complexity. As future research work, we plan to incorporate a minimum throughput constraint to guarantee the worst-case performance of SUs. Or alternatively, a logarithm function of the throughput can be employed as the objective function to ensure the fairness among SUs. Additionally, the joint optimization with respect to other important metrics such as the energy consumption/efficiency and the sensing time, may be analyzed via a framework similar to our paper. Moreover, the imperfect channel state information and the cooperative sensing strategy should be considered to make the work more promising for practical applications.

Appendix A

Assume that the nth channel is not assigned any SU, i.e., $x_{k,n} = 0, \forall k$. Due to (19), it follows that $J_{k,n} < \eta_n, \forall k$. It can be observed that $J_{k,n}$ is a monotonically increasing function with respect to $(\psi_n - \sigma_\nu^2/g_{k,n}^{SS})^+$ with its minimum value 0 attained at $(\psi_n - \sigma_\nu^2/g_{k,n}^{SS})^+ = 0$. Thus, $\eta_n > 0$. Based on this fact and the complementary slackness condition

$$\eta_n \left(1 - \sum_{k=1}^{K} x_{k,n}\right) = 0, \tag{35}$$

it turns out that $\sum_{k=1}^{K} x_{k,n} = 1$, which contradicts the original assumption that $x_{k,n} = 0, \forall k$. Therefore, it can be claimed that all the channels must be occupied by SUs. Furthermore, the statement $J_{k,n} < \eta_n, \forall k$, does not hold, which implies that $\max_k J_{k,n} \geq \eta_n$. From (19), we have $\max_k J_{k,n} \leq \eta_n$, and thus $\eta_n = \max_k J_{k,n}$, which completes the proof.

Appendix B

The theorem has to be proved in two scenarios. For a channel n, on one hand, it might occur that $\psi_n - \sigma_\nu^2/g_{k,n}^{SS} \leq 0, \forall k$. In this case, $y_{k,n} = 0$, $\forall k$, and we can simply assign the channel to the SU with the largest $g_{k,n}^{SS}$ without affecting the objective value. On the other hand, if the set $\{k|\psi_n - \sigma_\nu^2/g_{k,n}^{SS} > 0\}$ is not empty, $\max_k J_{k,n}$ must be achieved among SUs in this set. In this way, the $(\cdot)^+$ can be removed and it follows that $J_{k,n}$ is a monotonically increasing function of $g_{k,n}^{SS}$ in the set $\{k|\psi_n - \sigma_\nu^2/g_{k,n}^{SS} > 0\}$. Thus, $\max_k J_{k,n}$ is solely achieved at the SU with the largest $g_{k,n}^{SS}$, say the k_n^*th SU, provided that $g_{k,n}^{SS}$'s are all distinct. Then, based on Lemma 1, $\eta = \max_k J_{k,n} > 0$, and it follows that $x_{k_n^*,n} > 0$ and $x_{k,n} = 0,\ k \neq k_n^*$ due to (19). Finally, the complementary slackness condition in (35) yields the result that $x_{k_n^*,n} = 1$.

Appendix C

To transform problem (31) into the canonical form of a monotonic optimization, notice that for every $\{\gamma_n|\gamma_n^{\min} \leq \gamma_n \leq \gamma_n^{\max}\}$ there exists a nonnegative parameter t such that $g(\gamma_n) + t = g(\gamma_n^{\max})$. In this way, problem (31) can be rewritten as

$$\begin{aligned} \max_{\gamma_n} \quad & f(\gamma_n) + t \\ \text{s.t.} \quad & \gamma_n^{\min} \leq \gamma_n \leq \gamma_n^{\max}, \quad g(\gamma_n) + t = g(\gamma_n^{\max}). \end{aligned}$$

Let $F(\gamma_n, t) = f(\gamma_n) + t$, the formulation above admits the canonical form

$$\begin{aligned} \max \quad & F(\gamma_n, t) \\ \text{s.t.} \quad & (\gamma_n, t) \in \mathcal{G} \cap \mathcal{H}, \end{aligned}$$

with $\mathcal{G}$ and $\mathcal{H}$ defined in (33) and (34), respectively. It is easy to check that in the region

$$\left[\mathbf{0}, [\gamma_n^{\max}, g(\gamma_n^{\max}) - g(\gamma_n^{\min})]\right],$$

$F(\gamma_n, t)$ is an increasing function, $\mathcal{G}$ and $\mathcal{H}$ are normal and conormal sets, respectively.

References

[1] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood, “Chicago spectrum occupancy measurements & analysis and a long-term studies proposal,” in *Proceedings of the First International Workshop on Technology and Policy for Accessing Spectrum*, Boston, MA, USA, 2006.

[2] F. C. Commission, “Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies,” *Et Docket*, no. 3-108, pp. 5–57, 2003.

[3] J. Mitola and G. Q. Maguire, “Cognitive radio: Making software radios more personal,” *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.

[4] E. Biglieri, “An overview of cognitive radio for satellite communications,” in *Satellite Telecommunications (ESTEL), 2012 IEEE First AESS European Conference on*, 2012, pp. 1–3.

[5] Y.-C. Liang, Y. Zeng, E. C. Peh, and A. T. Hoang, “Sensing-throughput tradeoff for cognitive radio networks,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, 2008.

[6] E. C. Y. Peh, Y.-C. Liang, Y. L. Guan, and Y. Zeng, "Optimization of cooperative sensing in cognitive radio networks: A sensing-throughput tradeoff view," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 9, pp. 5294–5299, 2009.

[7] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1128–1140, 2009.

[8] G. Bansal, M. J. Hossain, and V. K. Bhargava, "Optimal and suboptimal power allocation schemes for OFDM-based cognitive radio systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4710–4718, 2008.

[9] M. Ge and S. Wang, "Fast optimal resource allocation is possible for multiuser OFDM-based cognitive radio networks with heterogeneous services," *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1500–1509, 2012.

[10] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wen, and M. Tao, "Resource allocation in spectrum-sharing OFDMA femtocells with heterogeneous services," *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2366–2377, 2014.

[11] S. Wang, Z.-H. Zhou, M. Ge, and C. Wang, "Resource allocation for heterogeneous cognitive radio networks with imperfect spectrum sensing," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 3, pp. 464–475, 2013.

[12] H. Zhang, C. Jiang, X. Mao, and H.-H. Chen, "Interference-limited resource optimization in cognitive femtocells with fairness and imperfect spectrum sensing," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1761–1771, 2016.

[13] H. Yu, W. Tang, and S. Li, "Joint optimal sensing and power allocation for cooperative relay in cognitive radio networks," in *Communications (ICC), 2012 IEEE International Conference on*, 2012, pp. 1635–1640.

[14] N. Janatian, S. Sun, and M. Modarres-Hashemi, "Joint optimal spectrum sensing and power allocation in CDMA-based cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 9, pp. 3990–3998, 2015.

[15] V.-D. Nguyen and O.-S. Shin, "Cooperative prediction-and-sensing based spectrum sharing in cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, 2017.

[16] X. Huang and B. Beferull-Lozano, "Joint optimization of detection and power allocation for OFDM-based cognitive radios," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, 2010, pp. 1–5.

[17] R. Fan, H. Jiang, Q. Guo, and Z. Zhang, "Joint optimal cooperative sensing and resource allocation in multichannel cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 722–729, 2011.

[18] J. C. Bezdek and R. J. Hathaway, "Some notes on alternating optimization," in *AFSS International Conference on Fuzzy Systems*. Springer, 2002, pp. 288–300.

[19] W. Yu and J. M. Cioffi, "FDMA capacity of gaussian multiple-access channels with ISI," *IEEE Transactions on Communications*, vol. 50, no. 1, pp. 102–111, 2002.

[20] I. Harjula and A. Hekkala, "Spectrum sensing in cognitive femto base stations using welch periodogram," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2011 IEEE 22nd International Symposium on*, 2011, pp. 2305–2309.

[21] T. Wang, Y. Liao, B. Zhang, and L. Song, "Joint spectrum access and power allocation in full-duplex cognitive cellular networks," in *Communications (ICC), 2015 IEEE International Conference on*, 2015, pp. 3329–3334.

[22] S. M. Almalfouh and G. L. Stuber, "Interference-aware radio resource allocation in OFDMA-based cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1699–1713, 2011.

[23] P. Kumar, N. Rakheja, A. Sarswat, H. Varshney, P. Bhatia, S. R. Goli, V. J. Ribeiro, and M. Sharma, "White space detection and spectrum characterization in urban and rural india," in *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a*, 2013, pp. 1–6.

[24] M. R. Dzulkifli, M. R. Kamarudin, and T. A. Rahman, "Spectrum occupancy at uhf tv band for cognitive radio applications," in *RF and Microwave Conference (RFM), 2011 IEEE International*, 2011, pp. 111–114.

[25] A. Marţian, C. Vlădeanu, I. Marcu, and I. Marghescu, "Evaluation of spectrum occupancy in an urban environment in a cognitive radio context," *International Journal on Advances in Telecommunications*, vol. 3, no. 3, pp. 172–181, 2010.

[26] S. Burer and A. N. Letchford, "Non-convex mixed-integer nonlinear programming: A survey," *Surveys in Operations Research and Management Science*, vol. 17, no. 2, pp. 97–106, 2012.

[27] H. Hindi, "A tutorial on convex optimization ii: Duality and interior point methods," in *American Control Conference, 2006*. IEEE, 2006, pp. 11–pp.

[28] Y. J. A. Zhang, L. Qian, and J. Huang, "Monotonic optimization in communication and networking systems," *Foundations and Trends in Networking*, vol. 7, no. 1, pp. 1–75, 2013.

[29] H. Tuy, "Monotonic optimization: Problems and solution approaches," *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 464–494, 2000.

[30] L. P. Qian and Y. J. Zhang, "S-MAPEL: Monotonic optimization for non-convex joint power control and scheduling problems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1708–1719, 2010.

Xu Wang (S'15) was born in Dalian, China. He received his Ph.D. from Electrical and Computer Engineering, Texas A&M University, College Station, TX, 2017, the M.Sc. from Electrical Engineering, Southern Methodist University, TX, 2012, and the B.Sc. from Electrical and Electronics Engineering, Huazhong University of Science and Technology, China, 2010. Currently, he is a data scientist at LeanTaaS Inc., Santa Clara, CA. His main research interests are the studies of convex, non-convex, and variational optimization with applications in wireless communications, statistical signal processing, operational research and bioinformatics.

Sabit Ekin (M'12) has joined the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK, as an Assistant Professor in Fall 2016. He has four years of industrial experience from Qualcomm Inc. as a Senior Modem Systems Engineer at the Department of Qualcomm Mobile Computing. He received his Ph.D. from Electrical and Computer Engineering, Texas A&M University, College Station, TX, 2012, the M.Sc. from Electrical Engineering, New Mexico Tech, NM, 2008, and the B.Sc. from Electrical and Electronics Engineering, Eskisehir Osmangazi University, Turkey, 2006. He was working as a visiting research assistant in Electrical and Computer Engineering Program at Texas A&M University at Qatar (2008-2009). During the summer of 2012, he worked with the Femto-cell interference management team in the Corporate R&D at New Jersey Research Center, Qualcomm Inc. After his Ph.D. study, he joined in Qualcomm Inc., San Diego, CA, where he has received numerous Qualstar awards for his achievements/contributions on cellular modem receiver design. Dr. Ekin's research interests are in the areas of design and performance analysis of wireless communications systems in both theoretical and practical point of views, particularly; Interference Modeling, Management and Optimization in 5G, mmWave, HetNets, and Cognitive Radio Systems and Applications. In addition, he works in the areas of Satellite Communications, Visible Light Sensing, Communications and Applications, RF Channel Modeling, Noncontact Health Monitoring, Internet of Things (IoT) Applications.

Erchin Serpedin (F'13) is a professor with the Department of Electrical and Computer Engineering, Texas A&M University, College Station. He received the specialization degree in transmission and processing of information from Ecole Superieure D'Electricite (SUPELEC), Paris, in 1992, the M.Sc. degree from the Georgia Institute of Technology in 1992, and the Ph.D. degree from the University of Virginia in January 1999. Dr. Serpedin is the author of two research monographs, one textbook, 15 book chapters, 150 journal papers, and 250 conference papers. His current research interests include signal processing, machine learning, cyber security, smart grids, bioinformatics, and wireless communications. Dr. Serpedin is currently serving as an associate editor of the IEEE Signal Processing Magazine. He served as an associate editor for more than 12 journals, including journals such as the IEEE Transactions on Information Theory, IEEE Transactions on Signal Processing, IEEE Transactions on Communications, IEEE Signal Processing Letters, IEEE Communications Letters, IEEE Transactions on Wireless Communications, Signal Processing (Elsevier), Physical Communications (Elsevier), EURASIP Journal on Advances in Signal Processing, and as a Technical Chair for six major conferences.