RESEARCH

Open Access

Resource allocation for secondary users with chance constraints based on primary links control feedback information

Kenan Zhou[*] and Tat Ming Lok

Abstract

Resource allocation for secondary users is an important issue in cognitive radio networks. In our article, we introduce a resource allocation scheme for secondary users to share spectrum in a cognitive radio network. Secondary users can exploit the spectrum owned by primary links when the interference level does not exceed certain requirements. Uncertainties of channel gains pose a great impact on the allocation scheme. Since there are uncertainties about the channel states, we apply chance constraints to represent the interference level requirements with uncertainties. Secondary users can exceed the interference level with a predefined small probability level. Since chance constraints are generally difficult to solve and full information about the uncertain variables is not available due to the fading effects of wireless channels, we reformulate the constraints into stochastic expectation constraints. With sample average approximation method, we propose stochastic distributed learning algorithms to help secondary users satisfy the constraints with the feedback information from primary links when maximizing the utilities.

Keywords: Power control, Cognitive radio networks, Probabilistic constraints, Sample average approximation

1 Introduction

Today's wireless networks are faced with an increase in spectrum demand while radio spectrum is a limited and valuable resource [1,2]. On the contrary, some licensed frequency bands are unoccupied for most of the time and some other bands are just partially used while unlicensed users have needs to get access to the spectrum [3]. Spectrum utilization can be improved significantly by allowing secondary users to exploit the spectrum unoccupied by primary users [4]. Cognitive radio has been proposed as the method to promote the efficient utilization of the spectrum by allowing secondary users to access the spectrum when it is unoccupied [4,5]. In cognitive radio networks, primary users may get some encouragements to provide some feedback information such that their own transmission would not be severely affected by secondary users once they want to share the spectrum [6]. Some monitoring nodes in the networks may also assume the tasks to provide feedback information in some practical scenarios [6,7]. The spectrum then can be better utilized while

primary users' transmissions would not be affected, which is one main motivations for cognitive radio networks. Moreover, resource allocation with time varying channels is an interesting research area in wireless communication. Uncertainties of the channel states trigger many important research issues since the uncertainties would pose an impact on the resource allocation decisions [3,4].

Power control with spectrum sharing is an important topic in cognitive radio networks. Some researchers propose a framework for dynamic spectrum sharing in cognitive radio networks to facilitate the usage of bandwidth only under certain channel states [8,9]. Spectrum sharing for unlicensed bands is also taken into consideration and optimal spectrum allocation is given under the condition that all channel information is available to all users [10]. Moreover, perfect channel state information (CSI) is assumed to be known by secondary users when they have access to the spectrum [11,12]. However, in most practical scenarios, there is an uncertainty in obtaining the channel information. Uncertainties of channel gains would affect the power allocation decisions of secondary users when they try to exploit the spectrum owned by primary

*Correspondence: zhou.kenan@yahoo.com.cn
The Chinese University of Hong Kong, Hong Kong, Hong Kong

users [13]. Secondary users want to maximize their own utilities while not impairing primary users' transmission quality even under the cases with imperfect CSI [13].

In some scenarios, secondary users can exceed the required interference level with some small probability instead of staying strictly below the required level. Chance constraints are widely applied in tackling uncertainties [6,14]. Some researchers analyze the chance constraint with known distribution by reformulating into a deterministic form with some knowledge of distributions of random variables [14,15]. Others approximate the chance constraint with convex approximation, which still requires the knowledge of the distribution even for Monte Carlo simulation [16]. Chance constraints are generally difficult to solve especially when we do not have the full information about the distributions of the random variables due to the fading effects of wireless channels [17,18]. Furthermore, it is difficult to obtain the exact information about the distributions in general. Sample average approximation method can be applied to help the secondary users learn the uncertain channel information [18,19]. However, direct measurements of channel gains are difficult and transmission of such large amount of information would occupy a lot of resources. Some apply the average approximation methods in a robust way [20]. Some consider subgradient learning method while transforming chance constraints into stochastic programming problems [21-23]. Some researchers also consider the long term ergodic optimization problems with learning processes [24,25].

In our article, we introduce a resource allocation scheme for secondary users to share spectrum in a cognitive radio network. Secondary users can exploit the spectrum owned by primary links while their interference levels can exceed the given requirements with a small predefined probability. To tackle the chance constraints, we demonstrate a stochastic approach with sample average approximation method for secondary users to obtain the feedback information from primary links when channel gains are uncertain. Secondary users do not need to have the full information about the channel gains and they just need to know the outage information from the primary links. We propose two stochastic distributed learning algorithms (SDLA) to optimize the utilities with given feedback information.

The rest of this article is organized as follows: Section 2 introduces the system model. Section 3 illustrates deterministic approaches with chance constraints when full information about the distributions is available. Section 4 illustrates our stochastic approaches with chance constraints when full information is not available and demonstrates our SDLAs. Section 5 gives our numerical results and discussion. The last section contains our conclusions.

2 System model

We consider a cognitive radio network with N primary links and M secondary users. We define the sets $\mathcal{M} = \{1, 2, \ldots M\}$ and $\mathcal{N} = \{1, 2, \ldots N\}$. Secondary users would like to transmit to the secondary base station when accessing the spectrum. We assume that every primary link represents a channel which is orthogonal to each other. To make the notation simple, we assume that the ith primary link denotes the ith channel. Secondary users can exploit the channels when the interference caused by them would not exceed a certain level to pose a great threat for the primary link transmission. We consider the channel gains are time varying such that there are some uncertainties.

Without loss of generality, we consider the kth secondary user. We use $u_k(\boldsymbol{P_k})$ to represent the utility function for the kth secondary user. We define $\boldsymbol{P_k} = [P_k^1 \ P_k^2 \ldots \ P_k^i \ldots P_k^N]$ to represent the power vector for the kth secondary user. P_k^i is the allocated power of the kth secondary user for the ith channel. In our article, we assume that $u_k(\boldsymbol{P_k})$ is a concave and monotonic increasing function with respect to every element P_k^i. Utility functions with such properties are widely applied, e.g, $u_k = \sum_{i=1}^N \log\left(1 + \frac{P_k^i G_k^i}{N_0 + \sum_{j=1, j\neq k}^M P_j^i G_j^i}\right)$, where G_k^i is the channel gain between the kth secondary user and the secondary base station at the ith channel. The choice of utility functions is open and there are other examples of utility functions [26,27].

The total power that can be allocated to the kth secondary user serves as a constraint and we have

$$\sum_{i=1}^N P_k^i \leq P_T \tag{1}$$

P_T is the total power the secondary user can use. For simplicity, we assume that all secondary users have the same total power constraint.

We define the feasible set for power allocation as

$$X = \left\{\boldsymbol{P_k} : \sum_{i=1}^N P_k^i \leq P_T, \ P_k^i \geq 0, \quad \forall i \in \mathcal{N}\right\} \tag{2}$$

Moreover, secondary users cannot exceed certain interference level when exploiting the channel. In some practical scenarios, the primary links may allow secondary users exceed the interference constraint with a certain probability α instead of staying strictly below it. In this case, we have the chance constraints as follows:

$$Pr\left\{P_k^i G_{k,p}^i - V > 0\right\} \leq \alpha, \quad \forall i \in \mathcal{N}, \tag{3}$$

where V represents the predefined interference level. $G_{k,p}^i$ denotes the channel gain between the kth secondary user

and the primary receiver with respect to the ith channel. To make the analysis simple, we assume all primary links have the same level requirement.

In some practical scenarios, α represents the quality of service (QoS) that the primary links demand facing interference from secondary users, e.g., when $\alpha = 0$, the primary links cannot tolerate any exceeded interference; when $\alpha = 0.2$, the primary links allow some exceeded interference with a small probability.

The interference constraints on a single secondary user may also be used in the area of admission control. In some cases, when secondary users exceed the required interference level, the primary link may block the secondary user out. Due to the uncertainty of channel gain $G_{k,p}^i$, secondary users need to satisfy the constraints to share the spectrum. Secondary users also want to maximize their utilities while satisfying the constraints (3).

3 Power allocation with complete information about probabilistic constraints

In order to maintain the interference level below the requirement, the kth secondary user needs to adjust his or her power allocation under the uncertainties of channel gains. The kth secondary user wants to maximize his or her utility and the optimization problem is formulated as follows

$$\max_{P_k^i} \ u_k$$
$$\text{s.t} \ \ \boldsymbol{P_k} \in X$$
$$Pr\left\{P_k^i G_{k,p}^i - V > 0\right\} \leq \alpha, \quad \forall i \in \mathcal{N} \tag{4}$$

We can see that there are probabilistic constraints in the optimization problem. To obtain an optimal solution, the chance constraints themselves need to be reformulated into some deterministic forms.

Chance constraints are generally non-convex. However, in our system model, since the dimension in every chance constraint is one with respect to P_k^i and $G_{k,p}^i$, it can be shown that the constraints represent convex sets when full information about the distribution is available. Given full information about the distribution, (4) is a convex optimization problem.

Lemma 1. Suppose that channel gain $G_{k,p}^i$ $\forall i$ are bounded and we have the cumulative density functions (CDF) $F_k^i(.)$, respectively. The optimization problem is a convex optimization problem.

Proof. Based on the assumptions in our system model, the objective function is concave for every P_k^i. Thus it is to

maximize a concave function. For the constraint set, the probabilistic constraint is equivalent to the following

$$F_k^i\left(\frac{V}{P_k^i}\right) \geq 1 - \alpha,$$

where we define $\frac{V}{P_k^i} = +\infty$ when $P_k^i = 0$. Moreover, since $F_k^i(.)$ is monotonic non-decreasing, the above constraint is equivalent to

$$\frac{V}{P_k^i} \geq Q_k^i$$

where we have

$$Q_k^i = \min\left\{q_k^i : F_k^i\left(q_k^i\right) = 1 - \alpha\right\}$$

Therefore, we can reformulate the constraint set as follows

$$0 \leq P_k^i \leq \frac{V}{Q_k^i}, \quad \forall i \in \mathcal{N}$$

$$\sum_{i=1}^{N} P_k^i \leq P_T$$

The constraint set is clearly convex. Thus we are maximizing a concave function with the convex set. Our optimization problem is a convex optimization problem. □

Given the complete information about the distribution, we define the feasible set for the kth secondary user as follows

$$X_{k,\text{com}} = \left\{\boldsymbol{P_k} : \sum_{i=1}^{N} P_k^i \leq P_T, P_{k,\max}^i \geq P_k^i \geq 0, \quad \forall i \in \mathcal{N}\right\} \tag{5}$$

where $P_{k,\max}^i$ is the maximum power for the kth secondary user at the ith channel. It can be calculated based on the CDF of the respective distribution. We also define the set $X_{\text{com}} = X_{1,\text{com}} \times X_{2,\text{com}} \times \cdots \times X_{k,\text{com}} \times \cdots \times X_{M,\text{com}}$.

The optimization problem can be rewritten as follows

$$\max_{P_k^i} \ u_k$$
$$\text{s.t} \ \ \boldsymbol{P_k} \in X_{k,\text{com}} \tag{6}$$

As an example, when $u_k = \sum_{i=1}^{N} \log\left(1 + \frac{P_k^i G_k^i}{N_0 + \sum_{j=1, j\neq k}^{M} P_j^i G_j^i}\right)$, the optimal solution can be obtained as follows

$$P_k^i = \left[\eta_k - \frac{N_0 + \sum_{j=1, j\neq k}^{M} P_j^i G_j^i}{G_k^i}\right]_0^{P_{k,\max}^i} \tag{7}$$

where $[.]_0^{P_{k,\max}^i}$ is the projection into the interval $[0, P_{k,\max}^i]$ and η_k is a Lagrangian variable.

Under this example, secondary users can optimize power allocation with others' choices. It can be shown

that a game $\mathcal{G} = [M, u_k(.), \boldsymbol{P_k}]$ is formed. It can also be shown that with feasible set X_{com} for all secondary users, the game is a potential game. Thus a Nash equilibrium (NE) exists in our game.

Proposition 1. With full information and for the game $\mathcal{G} = [M, u_k(.), \boldsymbol{P_k}]$, there exists a NE in our game model.

Proof. It is obvious that the set X_{com} is a convex, nonempty and compact set for all $k \in \mathcal{M}$. Noting that u_k is concave with respect to P_k^i, the existence of NE thus follows the standard results in game theory in [28,29]. □

Proposition 2. With full information and for the game $\mathcal{G} = [M, u_k(.), \boldsymbol{P_k}]$, it is a potential game.

Proof. We define a function as follows:

$$\Phi(\boldsymbol{P_k}) = \sum_{i=1}^{N} \log \left(N_0 + \sum_{k=1}^{M} P_k^i G_k^i \right) \tag{8}$$

It can be easily verified that

$$\Phi(\boldsymbol{P_k}, \boldsymbol{P_{-k}}) - \Phi\left(\boldsymbol{P'_k}, \boldsymbol{P_{-k}}\right) = u_k(\boldsymbol{P_k}, \boldsymbol{P_{-k}}) - u_k\left(\boldsymbol{P'_k}, \boldsymbol{P_{-k}}\right) \tag{9}$$

for any feasible solution $\boldsymbol{P_k}$ and $\boldsymbol{P'_k}$ and fixed $\boldsymbol{P_{-k}}$ where $\boldsymbol{P_{-k}}$ represents all other secondary users' strategies. Therefore, $\Phi(\boldsymbol{P_k})$ is a potential function for the game $\mathcal{G} = [M, u_k(.), \boldsymbol{P_k}]$. Due to the properties of potential game, we know that our game is a potential game [30]. □

We propose an algorithm to help secondary users update the choices of power allocation (Algorithm 1).

3.1 Algorithm 1: Updating choice of power consumption for each user (complete information)
1: **for** $k = 1$ to M **do**
2:      Given other users' information
3:      The kth secondary user updates power
         allocation according to the result in (7)
4: **end for**
5: Until all secondary users meet the NE.

Remark 1. We can see that with full information about the distribution of $G_{k,p}^i, \forall k, \forall i$, the result for a single secondary user is a typical water-filling solution. The algorithm is similar to iterative water filling algorithm (IWFA). The solutions and algorithms are commonly found in multiple access channels and potential games [31,32]. Therefore, with complete information, the chance constraints in our model can be formulated into convex ones and the problem can be solved readily.

4 A stochastic approximation approach based on the outage event
The reformulation of probabilistic constraints into deterministic convex sets requires the full knowledge of the distributions of the channel gains. Even without exact information about the distributions, we still need to measure the channel such that approximations can be applied to reformulate the probabilistic constraints. However, measurement is difficult in some scenarios.

To reduce the complexity of measurement, instead of approximating the information of the distribution for the specified random variable, we can consider an outage event as follows

Definition 1. When $P_k^i G_{k,p}^i > V$, we define that an outage occurs for the kth user at the ith channel. When $P_k^i G_{k,p}^i \leq V$, no outage occurs.

We then have the following

$$\zeta_k^i = \chi \left\{ P_k^i G_{k,p}^i > V \right\} \quad \forall i \in \mathcal{N} \tag{10}$$

where we have

$$\chi\{A\} = \begin{cases} 0 & \text{if } A \text{ is false} \\ 1 & \text{if } A \text{ is true} \end{cases}$$

ζ_k^i represents the random variable such that it would be one when an outage occurs and zero when there is no outage.

It can be seen that the probability constraints are also equivalent to the stochastic constraints.

$$E\left(\zeta_k^i\right) \leq \alpha, \quad \forall i \in \mathcal{N} \tag{11}$$

Mathematically, the chance constraints are reformulated as expectation constraints. Since we do not have the full information about the exact distribution of the outage event ζ_k^i, we would use the sample average to approximate the expectation. Given L samples of the events, we have

$$\widehat{\zeta_k^i} = \frac{1}{L} \sum_{j=1}^{L} \chi \left\{ P_k^i \left(G_{k,p}^i \right)_j > V \right\} \tag{12}$$

where $(G_{k,p}^i)_j$ denotes the jth sample of the channel gains. Equivalent, the event $(\widehat{\zeta_k^i})_j = \chi\{P_k^i(G_{k,p}^i)_j > V\}$ denotes the jth sample of the outage event based on the jth sample of the channel gain. We assume that every sample of $(G_{k,p}^i)_j$ obtained is independent and identically distributed (iid).

We also have

$$E\left[\widehat{\zeta_k^i}\right] = E\left[\zeta_k^i\right] \tag{13}$$

The expectation of the approximate equals that of the original.

4.1 Feasibility of the stochastic approximation method

When L is sufficiently large, the sample average value is used to approximate the expectation value. We have the following lemma.

Lemma 2. Given the value of P_k^i and V, when every sample of $G_{k,p}^i$ obtained is iid, let any $\sigma > 0$, we have

$$\Pr\left\{\left|E\left(\zeta_k^i\right) - \widehat{\zeta_k^i}\right| > \sigma\right\} \leq 2\exp(-2\sigma^2 L)$$

Proof. By Hoeffding Inequality, we can have

$$\Pr\left\{\left|E\left(\zeta_k^i\right) - \widehat{\zeta_k^i}\right| > \sigma\right\} \leq 2\exp\left(-\frac{2\sigma^2 L^2}{\sum_{j=1}^{L}(b_j - a_j)^2}\right)$$

where a_j and b_j represents the lower bound and upper bound of the random variable $(\widehat{\zeta_k^i})_j$ respectively with probability one. Since ζ_k^i can only take value of 0 or 1, we assign $a_j = 0$, $\forall j$ and $b_j = 1$, $\forall j$. We finally have

$$\Pr\left\{\left|E\left(\zeta_k^i\right) - \widehat{\zeta_k^i}\right| > \sigma\right\} \leq 2\exp(-2\sigma^2 L)$$

□

This lemma says that if we take the average of L random variables to be our estimate of the expected value, then the probability of being far away from the true value is small so long as L is large.

Moreover, since all parameters are nonnegative, we can reformulate the probability constraint as follows

$$\Pr\left\{\frac{V}{P_k^i} < G_{k,p}^i\right\} \leq \alpha,$$

where when $P_k^i = 0$, we denote $\frac{V}{P_k^i} = +\infty$. We also define the feasible set for the original problem with respect to i as follows

$$X_\alpha^i = \left\{\boldsymbol{P_k} \in X : \Pr\left\{\frac{V}{P_k^i} \geq G_{k,p}^i\right\} \geq 1 - \alpha\right\} \tag{14}$$

Thus the whole feasible set is defined as $X_\alpha = \bigcap_{i=1}^{N} X_\alpha^i$. The feasible region of the sample average approximation problem for $\gamma \in [0, 1)$ is

$$\widehat{X_\gamma^i} = \left\{\boldsymbol{P_k} \in X : \frac{1}{L}\sum_{j=1}^{L}\chi\left(\frac{V}{P_k^i} \geq \left(G_{k,p}^i\right)_j\right) \geq 1 - \gamma\right\} \tag{15}$$

Similarly, the whole feasible set for the sample average approximation is defined as $\widehat{X_\gamma} = \bigcap_{i=1}^{N} \widehat{X_\gamma^i}$.

In particular, when taking the sample average approximation, the secondary users would like to be more conservative. Given $\tau > 0$, we define

$$\widehat{X_{\gamma,\tau}^i} = \left\{\boldsymbol{P_k} \in X : \frac{1}{L}\sum_{j=1}^{L}\chi\left(P_k^i\left(G_{k,p}^i\right)_j - V + \tau \leq 0\right)\right.$$
$$\left. \geq 1 - \gamma\right\} \tag{16}$$

and $\widehat{X_{\gamma,\tau}} = \bigcap_{i=1}^{N}\widehat{X_{\gamma,\tau}^i}$.

Since $P_k^i G_{k,p}^i - V$ is Lipschitz Continuous, when the channel gain $G_{k,p}^i$ is bounded, we have

$$\left|\left(P_k^i G_{k,p}^i - V\right) - \left(\left(P_k^i\right)' G_{k,p}^i - V\right)\right| \leq c_k^i\left|P_k^i - \left(P_k^i\right)'\right|$$
$$\leq c_k^i\left|\left|\boldsymbol{P_k} - (\boldsymbol{P_k})'\right|\right|,$$
$$\forall \boldsymbol{P_k}, (\boldsymbol{P_k})' \in X \tag{17}$$

where, we define the Lipschitz constant as $c_k^i > 0$. It can be shown that this constant c_k^i can represent the bounded constant for the channel gain $G_{k,p}^i$. P_k^i and $(P_k^i)'$ are the ith element of the vector $\boldsymbol{P_k}$ and $(\boldsymbol{P_k})'$. It can be shown that the strategy set X is bounded and let D be the diameter of X. Then we have the following lemma.

Lemma 3. Let $\gamma \in [0, \alpha)$, $\beta \in (0, \alpha - \gamma)$ and $\tau > 0$, we have

$$Pr\left\{\widehat{X_{\gamma,\tau}} \subseteq X_\alpha\right\} \geq 1 - N\left\lceil\frac{1}{\beta}\right\rceil\left\lceil\frac{2c_{k,\max}D}{\tau}\right\rceil^m$$
$$\times \exp(-2L(\alpha - \gamma - \beta)^2)$$

where m is a finite constant and $c_{k,\max}$ represents the largest Lipschitz constant of c_k^i, $\forall i$.

Proof. See Appendix. □

The number of primary links N is finite in all practical scenarios. We can see that by adding $\tau > 0$ and applying the sample average approximation method, the new strategy set is still feasible with large probability when the sample number L is sufficiently large.

Remark 2. In some practical scenarios, secondary users want to guarantee that the constraints are satisfied even with the approximation. Thus they would have $\gamma < \alpha$ and introduce the variable $\tau > 0$ to make the requirements tighter than the original ones. With these guard measures, the new strategy set is feasible with a large probability as Lemma 3 states.

4.2 SDLA-I

To solve the optimization problem with sample average approximation, we can apply the general stochastic Arrow–Hurwciz algorithm [33]. First of all, we apply the Lagrangian with probabilistic constraints as follows

$$F_k = u_k + \sum_{i=1}^{N} \lambda_k^i \left(\alpha - E\left(\xi_k^i \right) \right)$$

and we define $\Theta_k(\boldsymbol{P_k}) = \sum_{i=1}^{N} (\alpha - E(\xi_k^i))$. We have the update of P_k^i and λ_k^i in vector form as follows

$$\boldsymbol{P_k(n+1)} = \Pi_X \big[\boldsymbol{P_k(n)} + \beta(n)(\nabla u_k(n)$$
$$+ \boldsymbol{\lambda_k(n)} \odot \nabla \Theta_k(n)) \big]$$

$$\boldsymbol{\lambda_k(n+1)} = \left[\boldsymbol{\lambda_k(n)} - \beta(n) \left(\boldsymbol{\alpha} - E\left(\boldsymbol{\xi_k^i} \right) \right) \right]^+ \qquad (18)$$

where $[x]^+ = \max(x, 0)$ and $\odot$ denotes the multiplication of two elements in the two vectors, respectively. $\Pi_X[.]$ represents the Euclidean projection to the set X. $\beta(n)$ is the step size for the nth iteration. However, in our case, it is very difficult to obtain the gradient of $E(\xi_k^i)$ with respect to P_k^i, that is, it is difficult to obtain the exact $\nabla \Theta_k$. To apply this algorithm, we need to approximate the gradient for this indicator function. From [33], we can choose the following result as the approximate gradient.

$$\nabla \Theta_k = \left[g_k^1 \, g_k^2 \, \cdots \, g_k^i \, \cdots \, g_k^N \right],$$

where we have

$$g_k^i \left(P_k^i, G_{k,p}^i \right) = -\frac{1}{r} h \left(\frac{P_k^i G_{k,p}^i - V}{r} \right) G_{k,p}^i,$$

where $r > 0$ and the function $h(.)$ has the following properties

$$\forall x, \ h(x) \geq 0, \ h(x) = h(-x), \ \int_{-\infty}^{+\infty} h(x) dx = 1$$

and $h(x)$ has a unique maximum at $x = 0$.

There are many choices about $h(x)$ [33]. From [33], we choose the following function

$$h(x) = \frac{3}{4}(1 - x^2)B(x),$$

where we define

$$B(x) = \begin{cases} 1 & \text{if } -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Remark 3. We can see that since r is a small positive constant as defined in [33], when $P_k^i G_{k,p}^i$ is much smaller than V, the approximate gradient is zero such that the power would increase through updating because ∇u_k is positive. When $P_k^i G_{k,p}^i$ is very close to V, the approximate

gradient represents a negative large constant such that power may decrease. Therefore, the update scheme can help power increase while the level of $P_k^i G_{k,p}^i$ is not close to V and help power decrease when the level of $P_k^i G_{k,p}^i$ would be higher than V.

Remark 4. Moreover, since the approximate gradient is not the exact value, the initial values of power vector $\boldsymbol{P_k}$ for the algorithm should be small. These kind of initial values are common in gradient update schemes and are widely applied in gradient or sub-gradient methods [34].

It can be shown that this approximate gradient just represents an indicator for whether subtracting λ_k in the power update equations. When the expectation constraint is satisfied, the approximate gradient is zero such that power is increased through updating. Otherwise the approximate gradient is negative such that power is decreased through updating. In some practical scenarios, it can be further simplified as we will show in our simulation results.

We can see that to update the power level, secondary users need to have the feedback information about λ and the approximate gradient. The secondary user updates power levels as primal update scheme while primary links would observe the outage events and feedback the respective information through λ as dual update scheme. The result is a primal-dual update scheme in a stochastic setting.

Moreover, we also need to apply the sample average method to tackle the value of $\boldsymbol{\alpha} - E(\boldsymbol{\xi_k^i})$. The algorithm scheme can be summarized as follows and we denote it as SDLA-I:

- Power level would be adjusted based on a gradient-like algorithm.
- Step 1: Between the nth iteration and the $(n + 1)$th iteration, the kth secondary user obtains the feedback information of the outage event regarding the power allocation $\boldsymbol{P_k(n)}$.
- Step 2: The primary links calculate the outage with regard to all past history. When an outage can occur given the channel information of the jth iteration, $\forall j \leq n$, we denote $O_k^i(j) = 1$. Otherwise it is zero. We obtain an outage probability as $\frac{N_k^i(n)}{n}$ for the nth iteration. $N_k^i(n) = \sum_{j=1}^{n} O_k^i(j)$ represents the sum value for the outages that can happen during all previous iterations where we have $0 \leq N_k^i(n) \leq n$.
- Step 3: Based on the observed outage probability, the primary link would update the respective λ based on the dual update scheme and calculate the approximate gradient. The primary link would feed back such information to the kth secondary user.

- Step 4: The kth secondary user would then update the power level for the $(n+1)$th iteration based on the primal update scheme.
- Step 5: The power level and the respective dual variable λ would be updated iteratively.

The update of our algorithm is as follows

$$P_k(n+1) = \Pi_X\big[P_k(n) + \beta(n)(\nabla u_k(n) + \lambda_k(n) \odot \nabla\Theta_k(n))\big]$$

$$\lambda_k(n+1) = \left[\lambda_k(n) - \beta(n)(\alpha - \frac{N_k^i(n)}{n})\right]^+. \qquad (19)$$

We apply learning algorithm to learn the event ζ_k^i such that the constraint is satisfied in the long term. The secondary users can obtain some feedbacks to update the choices of power. Moreover, in the learning process, the secondary users just need to know something about the instantaneous outage events and do not need to know the distributions of specific channel gains. Also, secondary users do not need to handle the measurements. Thus the kth secondary users would just focus on the randomness of the term ζ_k^i instead of the specified channel gains.

Remark 5. Chance constraints are generally difficult to tackle even when they are reformulated into expectation constraints, let alone we do not have the information about distributions. The reformulated expectation constraints are also different from general expectation constraints as discussed in stochastic programming areas, which pose great difficulties for the exact proof of convergence of the algorithms. In [35] and [36], researchers discuss the complexity of stochastic programming with chance constraints and illustrate that few analytical results can be provided even under sample average approximations, since there are some severe difficulties when facing chance constraints.

Remark 6. In our system model, for SDLA-I, $E(\zeta_k^i)$ is close to $\widehat{\zeta_k^i}$ with large probability from Lemma 2 given a large L. Suppose that there exists a large number n_0 and for $L \geq n_0$, we can treat $E(\zeta_k^i) = \widehat{\zeta_k^i}$ for relaxation. Then, in this case, for $n \geq n_0$, our SDLA-I can be written as follows

$$P_k(n+1) = \Pi_X\big[P_k(n) + \beta(n)(\nabla u_k(n) + \lambda_k(n) \odot \nabla\Theta_k(n))\big]$$

$$\lambda_k(n+1) = \left[\lambda_k(n) - \beta(n)\left(\alpha - E\left(\zeta_k^i\right)\right)\right]^+ \qquad (20)$$

where $E(\zeta_k^i) = \widehat{\zeta_k^i}$ can be obtained from previous iterations with approximations. Given a close deterministic approximations of gradient $\Theta_k(n)$, the above SDLA-I can be

treated as a deterministic primal-dual update algorithms. Such primal-dual algorithms have been discussed in some areas [25,37].

4.3 SDLA-II

In SDLA-I, though secondary users do not need to measure the channel gain $G_{k,p}^i$, the calculation of λ_k and the approximate gradient are still required in the primary links. We can further reduce the burden of such calculation such that just simple calculation and feedback information is needed for both primary links and secondary users, which is desirable in practical scenarios.

We propose another scheme to adjust power levels with observed outage probability. Instead of updating power level with information about λ and the approximate gradient, the secondary user can update the power level just with the feedback information about the observed outage probability. In this case, only primal update is realized without dual variables.

We denote it as SDLA-II. The algorithm scheme is similar to that in SDLA-I but we change some steps as follows.

- Changes of Step 3: No feedback of dual variables from primary links. Secondary users check the feasibility of the constraints with the observed outage probability to see whether $\alpha > \frac{N_k^i(n)}{n}$.
- Change Step 4 as: If $\alpha > \frac{N_k^i(n)}{n}$, secondary users regard that the constraint is satisfied and would increase the power level for the next iteration. Otherwise they regard the power is so high that it needs to be reduced to satisfy the constraint.
- Changes of Step 5: There is no update of dual variables and approximate gradients. Secondary users only update the power levels iteratively.

Though the observe outage probability is not the exact true value of the expectation, it can approximate to the true value when n is sufficient large by the law of large numbers, that is, when the number of iterations is large enough.

We define

$$V_k^i(n) = \alpha - \frac{N_k^i(n)}{n}$$

and we define the vector form $V_k(n) = [\,V_k^1(n)\,V_k^2(n) \ldots V_k^i(n) \ldots V_k^N(n)]$.

We have the gradient method as follows

$$P_k^i(n+1) = \Pi_X\left[P_k^i(n) + \beta(n)\left(\alpha - \frac{N_k^i(n)}{n}\right)\frac{du_k(n)}{dP_k^i(n)}\right]$$

We can see that when the calculated outage level is smaller than α, P_k^i would increase according to the general gradient algorithm. When the calculated outage level

is larger than α, we regard that P_k^i is so large that it should be decreased.

In vector form, we have

$$P_k(n+1) = \Pi_X\big[P_k(n) + \beta(n)(V_k(n) \odot \nabla u_k(n))\big] \tag{21}$$

The secondary users update the power allocation based on the feedback of outage events such that in the long term, the probability constraints can be satisfied and the utility is maximized.

Remark 7. For SDLA-II, similar to the remark discussion above about SDLA-II, suppose that when $n \geq n_0$, the constraint $\alpha - \widehat{\zeta_k^i}$ is satisfied with large probability such that we can treat $\alpha \geq \widehat{\zeta_k^i}$ for future iterations. Our SDLA-II for the kth secondary user can be written as follows

$$P_k(n+1) = \Pi_X[P_k(n) + \beta(n)(\nabla u_k(n))] \tag{22}$$

we can see that it is a deterministic gradient update algorithm and such algorithms has been discussed in [34,38,39].

5 Numerical results and discussion

5.1 Examples of $u_k(.)$ for simulation

For simulation, first, we illustrate some examples about the objective function $u_k(P_k)$. We have one as the following

$$u_k(P_k) = \sum_{i=1}^{N} E_{G^i}\left(R_k^i\right),$$

where we have

$$R_k^i = \log\left(1 + \frac{P_k^i G_k^i}{N_0 + \sum_{j=1,j\neq k}^{M} P_j^i G_j^i}\right)$$

$G^i = [G_1^i \ G_2^i \ldots G_k^i \ldots G_M^i]$. Since G_k^i also changes every time slot, we want to maximize the expected value. $E_{G^i}(R_k^i)$ denotes the expected value with respect to G^i. We shall in the sequel drop the subscript to write $E(R_k^i)$ instead of $E_{G^i}(R_k^i)$ when not leading to confusion. It can be shown that this $u_k(P_k)$ is concave and monotonic increasing with respect to every element P_k^i.

In particular, when G_k^i would not change much during the learning process of $G_{k,p}^i$, that is, the channel gain G_k^i remains more or less the same as constant during the algorithm, we take the following example

$$u_k(P_k) = \sum_{i=1}^{N} R_k^i$$

There are some practical scenarios in which the channel gain G_k^i would not change much. We regard that it would be a constant in the objective function. This kind

of approximation is used in the stochastic modeling and algorithm in [6].

5.2 Simulation model

We consider a micro-cell area for wireless transmissions. We set that there are two secondary users allocated with $N = 20$ channels. Due to the environment setting, shadowing effects play an important role in the uncertainties of channel gains. We set that the channel gain G_k^i is generated according to the following

$$G_k^i = \kappa_k^i - 10\omega \log(d_k) + \epsilon_k^i (\text{dB}),$$

where we set $\omega = 3.3$. κ_k^i is generated uniformly from $[20, 25]$. d_k is generated uniformly from $[1, 20]$ m, which represents that transmissions of secondary users occur in small area. ϵ_k^i, which represents the shadowing effect, is generated from a gaussian distribution $N(0, (\sigma_k^i)^2)$. We set σ_k^i generated uniformly from the interval $[0.1, 2]$ in our simulation. The value of σ_k^i can be varied in an interval and what we choose here represents some common environment conditions of micro-cell with shadowing effects [40,41].

Similarly, to model the uncertainty of $G_{k,p}^i$, for the channel gain between the primary links and secondary users, we consider that $G_{k,p}^i$ is generated as follows

$$G_{k,p}^i = \kappa_{k,p} - 10\alpha_p \log(d_{k,p}) + \epsilon_{k,p}^i (\text{dB}),$$

where α_p is set to be 3.1. $d_{k,p}$ is generated uniformly from $[1, 15]$ m. $\kappa_{k,p} = 20$ and $\epsilon_{k,p}^i$ is generated from a gaussian distribution $N(0, (\sigma_{k,p}^i)^2)$. $\sigma_{k,p}^i$ is generated uniformly from the interval $[1, 8]$ for each i and k. The large interval for $\sigma_{k,p}^i$ represents uncertainties in these channel gains since secondary users may not be able to obtain the exact statistical knowledge of the channel gains based on the transmission environment.

We also set $\alpha = \gamma$ in our simulation. The choice of distribution of $G_{k,p}^i$ remains open and we just use this distribution to show some examples under a micro-cell with shadowing effects. For the step-size $\beta(n)$ in the algorithm, when $\beta(n) = \frac{1}{n}$, the result may finally converge to a stable point. However, in our simulation model, to speed up the gradient-like SDLA algorithms, we set $\beta(n)$ equal to one. The choice of step-size is also open. In our simulation, we just want to show some examples and try to speed up the learning process.

5.3 Simulation results and discussions

In Figure 1, we consider the example when $u_k = \sum_{i=1}^{N} E(R_k^i)$. For the data rates with full information, we obtain the results by averaging the optimization results for each iteration with full information about $G_{k,p}^i$. For data
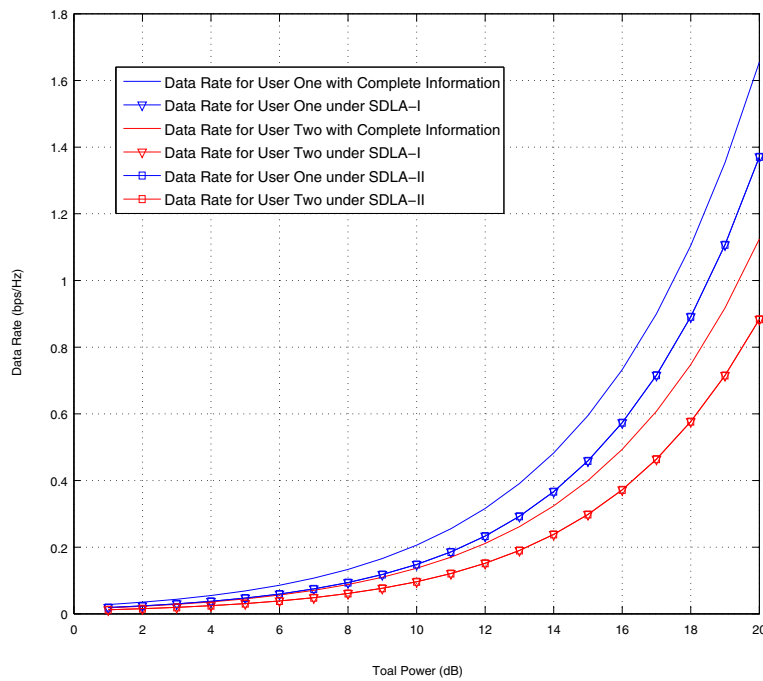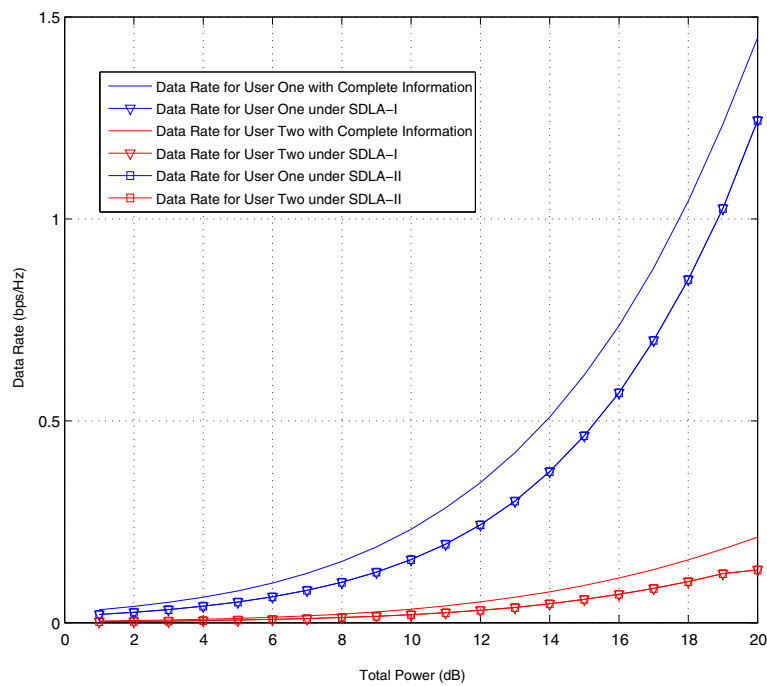
Figure 1 Data rates for secondary users under different resource allocation schemes ($u_k = \sum_{i=1}^{N} E(R_k^i)$, $\alpha = 0.2$, $V = 100$).

rates under SDLA, we obtain the final power allocations through iterations and apply the results to calculate the average value. To make the simulation simple, we set the approximate gradient with respect to power as minus one in our simulation.

It can be seen that in Figure 1, data rates generated by SLDA-I and SLDA-II are very close to those with full information about $G_{k,p}^i$. For both users, when the data rate is 0.6 bps/Hz, the algorithm with full information has only 0.7 dB gain over SDLA. When the data rate is 0.8 bps/Hz,
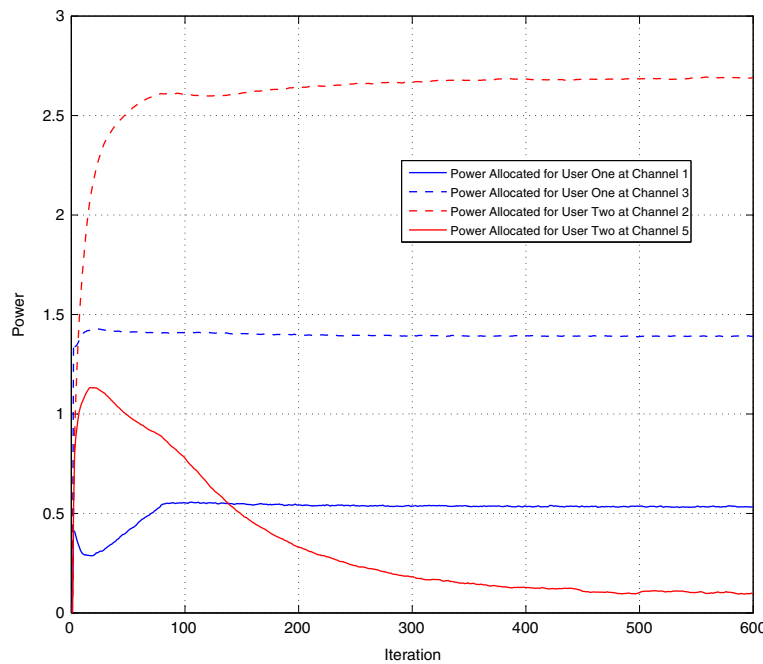
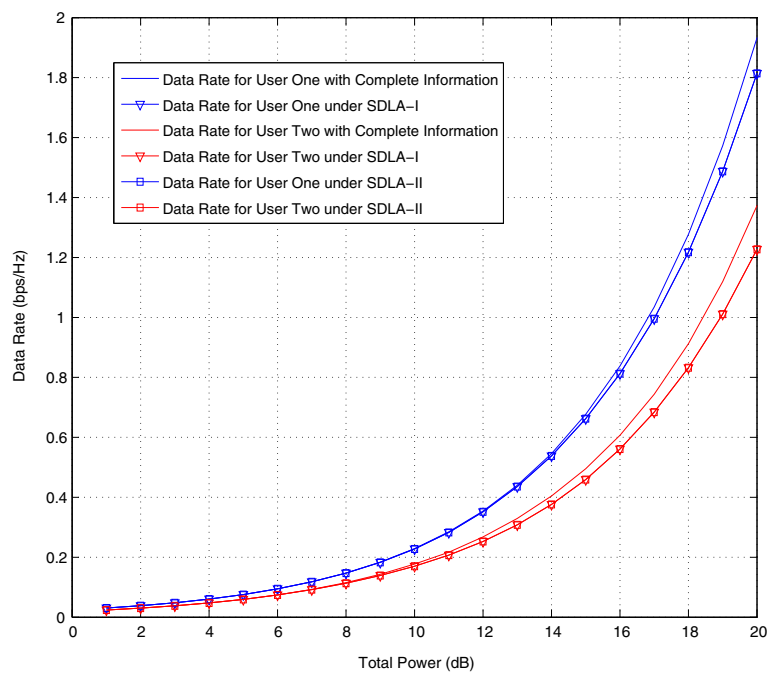**Figure 2 Power Allocated under SDLA over Iterations ($u_k = \sum_{i=1}^{N} E(R_k^i)$, $\alpha = 0.2$, $V = 100$).**

Figure 3 Data rates for secondary users under different resource allocation schemes ($u_k = \sum_{i=1}^{N} E(R_k^i), \alpha = 0.2, V = 1$).

the gain is just about 0.8 dB. When the total power ranges from 1 to 20 dB, the data rates under SDLA-I are very close to those under SDLA-II. When the data rate is 1 bps/Hz, for both users, SDLA-I has only 0.02 dB gain over SDLA-II. Notice that, for the algorithm with full information,

optimization needs to be done every time while our SDLA can learn the result with finite iterations and do not require optimization every time.

In Figure 2, we consider the allocated power over iterations. It can be seen that the power allocated for different

Figure 4 Data rates for secondary users under different resource allocation schemes ($u_k = \sum_{i=1}^{N} R_k^i, \alpha = 0.2, V = 100$).

channels remains almost constant after about 400 iterations. For user one, the value for channel one remains about 0.5 and that for channel three remains about 1.4.

It can be seen that in Figure 3, when $V = 1$, data rates generated by SLDA-I and SLDA-II are very close to those with full information. For user one, when the data rate is 0.8 bps/Hz, the algorithm with full information has only 0.8 dB gain over SDLA. When the data rate is 1 bps/Hz, the gain is just about 0.7 dB. When the total power ranges from 1 to 20 dB, the data rates under SDLA-I are very close to those under SDLA-II. When the data rate is 1 bps/Hz, for user one, SDLA-I has only 0.02 dB gain over SDLA-II.

In Figure 4, we set $u_k = \sum_{i=1}^{N} R_k^i$. It can be seen that the data rates generated by both SDLA-I and SDLA-II are very close to those with full information about $G_{k,p}^i$. For user one, when the data rate is 0.6 bps/Hz, the algorithm with full information has only 0.1 dB gain over SDLA-I and SDLA-II. For user two, when the data rate is 0.4 bps/Hz, the algorithm with full information has only 0.5 dB gain over SDLA-I and SDLA-II. Moreover, the data rates under SDLA-I are very close to those under SDLA-II when the total power ranges from 1 to 20 dB. For both users, when the data rate is 0.8 bps/Hz, SDLA-I has about 0.03 dB gain over SDLA-II.

In Figure 5, when $\alpha = 0.02$ and $V = 1$, it can be seen that the data rates generated by both SDLA-I and SDLA-II are very close to those with complete information. For both users, when the data rate is 0.2 bps/Hz, the algorithm with full information has only 0.3 dB gain over

SDLA-I and SDLA-II. Also, the data rates under SDLA-I are very close to those under SDLA-II when the total power ranges from 1 to 20 dB. Moreover, for both users, when total power is larger than 20 dB, the data rates would not improve due to maximum power constraints. Results under SDLA-I and SDLA-II are still close to those with complete information.

Moreover, comparing Figure 1 with Figure 4, our SDLA algorithm can give good performance especially when the variation of G_k^i is small over iterations. In Figure 4, the channel gains in the utility functions remain the same over the learning processes. The small variation of G_k^i over the learning process is reasonable in some practical scenarios. For example, when the standard deviation of the shadowing effects of the channel gain is small, the distance between the transmitter and the receiver becomes the key factor such that the value of the channel gains would not change very much.

6 Conclusion

Resource allocation for secondary users with uncertainties is an important issue in cognitive radio networks. In our article, we introduce a resource allocation scheme for secondary users to share spectrum in a cognitive radio network. Secondary users can exceed the interference level with a predefined small probability level to share the spectrum. There is an uncertainty about the channel states between secondary users and primary links. We apply chance constraints to represent the interference
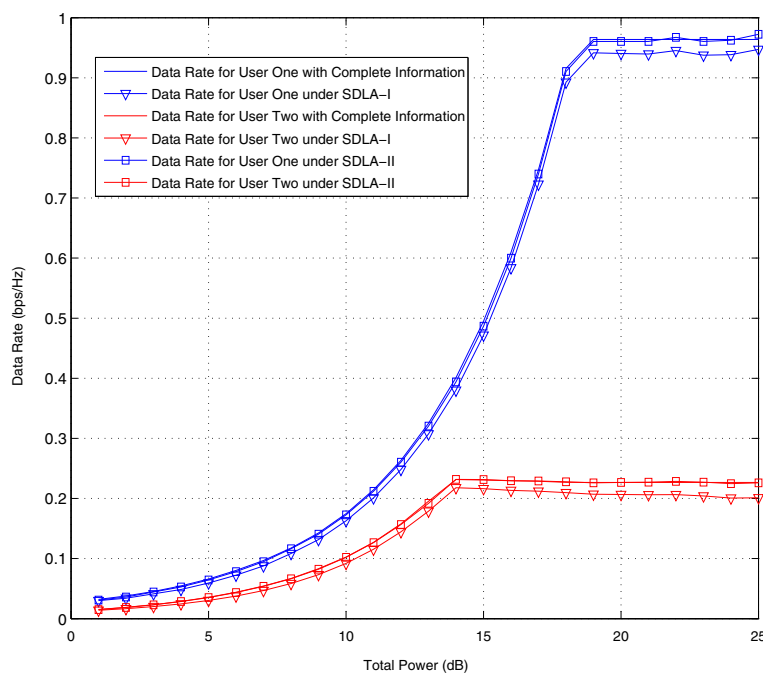

Figure 5 Data rates for secondary users under different resource allocation schemes ($u_k = \sum_{i=1}^{N} R_k^i$, $\alpha = 0.02$, $V = 1$).

level requirements with the uncertainties. Since chance constraints are generally difficult to solve and full information about the uncertain variable is not available, we reformulate them into stochastic expectation constraints. The secondary users can learn the outage feedback information instead of measuring certain channel gains to satisfy the constraints. We propose two SDLAs to help secondary users adjust the power to maximize the utility with only feedback information from primary links. Our simulation results show that the algorithms can give performance close to that with complete information.

Appendix
Proof of Lemma 3

According to the proof of Theorem 10 in [36], for $\beta \in (0, \alpha - \gamma)$, there exist finite sets $Z_i^\tau \subseteq X$ with

$$
\left| Z_i^\tau \right| \leq \left\lceil \frac{1}{\beta} \right\rceil \left\lceil \frac{2c_k^i D}{\tau} \right\rceil^m \tag{23}
$$

where $\lceil . \rceil$ denotes the upper integer part, and for any $\boldsymbol{P_k} \in \widehat{X_{\gamma,\tau}}$ and any i there exists $\boldsymbol{z} \in Z_i^\tau$ such that $\|\boldsymbol{z} - \boldsymbol{P_k}\| \leq \frac{\tau}{c_k^i}$. Using the finite set Z_i^τ we can define

$$
Z_\gamma^{\tau,i} = \left\{ \boldsymbol{P_k} \in Z_i^\tau : \frac{1}{L} \sum_{j=1}^{L} \chi \left(P_k^i \left(G_{k,p}^i \right)_j - V \right) \geq 1 - \gamma \right\} \tag{24}
$$

$$
Z_{\alpha-\beta}^{\tau,i} = \left\{ \boldsymbol{P_k} \in Z_i^\tau : Pr \left\{ P_k^i G_{k,p}^i - V \leq 0 \right\} \geq 1 - (\alpha - \beta) \right\} \tag{25}
$$

$$
Z_\gamma^\tau = \bigcap_{i=1}^{N} Z_\gamma^{\tau,i} \tag{26}
$$

$$
Z_{\alpha-\beta}^\tau = \bigcap_{i=1}^{N} Z_{\alpha-\beta}^{\tau,i} \tag{27}
$$

Moreover, from the proof of Theorem 10 in [36], for all i it holds that $Z_\gamma^{\tau,i} \subseteq Z_{\alpha-\beta}^{\tau,i}$ implies $\widehat{X_{\gamma,\tau}^i} \subseteq X_\alpha^i$. From the proof in Section 3.2.2 in [42], for the finite set, we have

$$
1 - Pr \left\{ Z_\gamma^\tau \subseteq Z_{\alpha-\beta}^\tau \right\} \leq N \left\lceil \frac{1}{\beta} \right\rceil \left\lceil \frac{2c_{k,\max} D}{\tau} \right\rceil^m
$$
$$
\exp(-2L(\alpha - \gamma - \beta)^2), \tag{28}
$$

where we apply the the Bonferroni inequality as follows

$$
Pr \left\{ \bigcap_{i=1}^{N} A_i \right\} \geq 1 - \sum_{i=1}^{N} (1 - Pr\{A_i\}),
$$

where A_i is an event.

From the proof in Section 3.2.2 in [42], since $Z_\gamma^\tau \subseteq Z_{\alpha-\beta}^\tau$ implies $\widehat{X_{\gamma,\tau}} \subseteq X_\alpha$, we can obtain the result as follows

$$
Pr \left\{ \widehat{X_{\gamma,\tau}} \subseteq X_\alpha \right\} \geq Pr \left\{ Z_\gamma^\tau \subseteq Z_{\alpha-\beta}^\tau \right\}
$$
$$
\geq 1 - N \left\lceil \frac{1}{\beta} \right\rceil \left\lceil \frac{2c_{k,\max} D}{\tau} \right\rceil^m
$$
$$
\exp(-2L(\alpha - \gamma - \beta)^2) \tag{29}
$$

References
1. TS Rappaport, *Wireless Communications: Principles and Practice*. (Prentice Hall, New York, 2002)
2. J Proakis, *Digital Communications*. (McGraw-Hill, New York, 2001)
3. D Tse, P Viswanath, *Fundamentals of Wireless Communication*. (Cambridge University Press, Cambridge, 2005)
4. S Haykin, Cognitive radio: Brain-empowered wireless communications. IEEE J. Sel. Areas Commun. **23**(2), 201–220 (2005)
5. L Grokop, DN Tse, in *Proc. 27th IEEE International Conference on Computer Communications (INFOCOM'08)*. Spectrum sharing between wireless networks (Phoenix, AZ, USA, Apr 2008), pp. 201–205
6. S Huang, X Liu, Z Ding, Decentralized cognitive radio control based on inference from primary link control information. IEEE J. Sel. Areas Commun. **29**(2), 394–406 (2011)
7. G Zhao, GY Li, C Yang, Proactive detection of spectrum opportunities in primary systems with power control. IEEE Trans. Wirel. Commun. **8**(9), 4815–4823 (2009)
8. J Acharya, RD Yates, in *IEEE 2007 International Conference on Communications (ICC'07)*. A framework for dynamic spectrum sharing between cognitive radios (Glasgow, Britain, June 2007), pp. 5166–5171
9. R Cendrillon, W Yu, M Moonen, J Verlinden, Optimal multiuser spectrum balancing for digital subscriber lines. IEEE Trans. Commun. **54**(5), 922–933 (2006)
10. R Etkin, A Parekh, D Tse, Spectrum sharing for unlicensed bands. IEEE J. Sel. Areas Commun. **25**(3), 517–528 (2007)
11. JS Pang, G Scutari, DP Palomar, F Facchinei, Design of cognitive radio systems under temperature-interference constraints: a variational inequality approach. IEEE Trans. Signal Process. **58**(6), 3251–3271 (2010)
12. S Hayashi, ZQ Luo, Spectrum management for interference-limited multiuser communication systems. IEEE Trans. Inf. Theory. **55**(3), 1153–1175 (2009)
13. Y Xing, CN Mathur, MA Haleem, R Chandramouli, KP Subbalakshmi, Dynamic spectrum access with qos constraints and interference temperature. IEEE Trans. Mobile Comput. **6**(4), 423–433 (2007)
14. P-J Chung, H Du, J Gondzio, A probabilistic constraint approach for robust transmit beamforming with imperfect channel information. IEEE Trans. Signal Process. **59**(6), 2773–2782 (2011)
15. YJ Zhang, So AM-C, Optimal spectrum sharing in mimo cognitive radio networks via semidefinite programming. IEEE J. Sel. Areas Commun. **29**(2), 362–373 (2011)
16. A Nemirovski, A Shapiro, Convex approximation of chance constrained programs. J. Opt. **17**(4), 969–996 (2006)
17. A Ruszczynski, A Shapiro, in *Handbooks in Operations Research and Management Science*. 1st edn, ed. by A Shapiro. Stochastic Programming (Elsevier, Amsterdam, 2003)
18. H Robbins, S Monro, A stochastic approximation method. Ann. Math. Stat. **22**, 400–407 (1951)
19. BK Pagnoncelli, S Ahmed, A Shapiro, Sample average approximation method for chance constrained programming: Theory and applications. J. Opt. Theory Appl. **142**(2), 399–416 (2009)
20. A Nemirovski, A Juditsky, G Lan, A Shapiro, Robust stochastic approximation approach to stochastic programming. SIAM J. Opt. **19**(4), 1574–1609 (2009)

21. F Yousefian, A Nedic, UV Shanbhag, On stochastic gradient and subgradient methods with adaptive steplength sequences. Automatica. **48**(1), 56–67 (2012)
22. SS Ram, A Nedic, VV Veeravalli, Distributed stochastic subgradient projection algorithms for convex optimization. J. Opt. Theory Appl. **147**(3), 516–545 (2010)
23. X Wang, N Gao, Stochasticresourceallocationoverfadingmultiple access, broadcastchannels Trans. IEEE Inf. Theory. **56**(5), 2382–2391 (2010)
24. A Ribeiro, Ergodic stochastic optimization algorithms for wireless communication and networking. IEEE Trans. Signal Process. **58**(12), 6369–6386 (2010)
25. J Zhang, D Zheng, M Chiang, The impact of stochastic noisy feedback on distributed network utility maximization. IEEE Trans. Inf. Theory. **54**(2), 645–665 (2008)
26. CU Saraydar, NB Mandayam, DJ Goodman, Pricing and power control in a multicell wireless data network. IEEE J. Sel. Areas Commun. **19**(10), 1883–1892 (2001)
27. M Xiao, NB Shroff, EKP Chong, A utility based power control scheme in wireless cellular systems. IEEE/ACM Trans. Netw. **11**(2), 210–221 (2003)
28. D Fudenberg, J Tirole, *Game Theory*. (MIT Press, Cambridge, 1991)
29. MJ Osborne, A Rubinstein, *A Course in Game Theory*. (MIT Press, Cambridge, 1994)
30. D Monderer, L Shapley, Potential games. Games Econ. Behav. **14**(1), 124–143 (1996)
31. DP Palomar, M Chiang, A tutorial on decomposition methods for network utility maximization. IEEE J. Sel. Areas Commun. **24**(8), 1439–1451 (2006)
32. G He, L Cottatellucci, M Debbah, The waterfilling game theoretical framework for distributed wireless network information flow. EURASIP J. Wirel. Commun. Netw. **2010**, 13 (2010)
33. L Andrieu, G Cohen, FJ Vazquez-Abad, in *10th International Conference on Stochastic Programming*. Stochastic programming with probabilistic constraints (Tucson, Arizona, USA, Oct 2004), pp. 493–501
34. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge University Press, Cambridge, 2004)
35. A Nemirovski, A Shapiro, *Continuous Optimization: Current Trends and Applications*, vol. 99. (Springer , Berlin, Germany, 2005), pp. 111–146
36. J Luedtke, S Ahmed, A sample approximation approach for optimization with probabilistic constraints. SIAM J. Opt. **19**, 674–699 (2007)
37. T Voice, Stability of congestion control algorithms with multi-path routing and linear stochastic modeling of congestion control, Ph.D. dissertation, University of Cambridge, 2006
38. DP Bertsekas, *Nonlinear Programming*, 2nd edn. (Athena Scientific, Belmont, 2003)
39. DP Bertsekas, JN Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. (Prentice Hall, Englewood Cliffs, 1989)
40. JW Mark, W Zhuang, *Wireless Communications and Networking*. (Julie L Higle, ed.) (Prentice Hall, Englewood Cliffs, 2003)
41. B Sklar, Rayleigh channel in mobile digital communication systems part i: Characterization. IEEE Commun. Mag. **35**(7), 90–100 (1997)
42. M Branda, Stochastic Programming E-Print Series. Reformulation of General Chance Constrained Problems Using the Penalty Functions, vol. 2, pp. 1–18, May 2010