

Throughput Maximization in Multi-Channel Cognitive Radio Systems with Delay Constraints

Ahmed Ewaisha and Cihan Tepedelenlioğlu

School of Electrical, Computer, and Energy Engineering, Arizona State University, USA.

Email: {ewaisha, cihan}@asu.edu

Abstract—We study the throughput-vs-delay trade-off in an overlay multi-channel single-secondary-user cognitive radio system. Due to the limited sensing capabilities of the cognitive radio user, channels are sensed sequentially. Maximizing the throughput in such a problem is well-studied in the literature. Yet, in real-time applications, hard delay constraints need to be considered besides throughput. In this paper, optimal stopping rule and optimal power allocation are discussed to maximize the secondary user's throughput, subject to an average delay constraint. We provide a low complexity approach to the optimal solution of this problem. Simulation results show that this solution allows the secondary user to meet the delay constraint without sacrificing throughput significantly. It also shows the benefits of the optimal power allocation strategy over the constant power allocation strategy.

I. INTRODUCTION

Cognitive Radio (CR) systems have been proposed to help overcome spectral inefficiency by allowing unlicensed users, called the Secondary Users (SU), to access the spectrum with minimal or with no degradation of the licensed users' performance. To guarantee this, the SU radios are equipped with sensing units capable of detecting the presence of the Primary Users' (PU) transmission activities over the spectrum. When no activity is detected at some frequency band, the channel is deemed free and the SU is allowed to use it.

If the PU can tolerate some interference, the SU will be allowed to transmit on busy frequency channels as long as the interference does not exceed this tolerance value. This scenario has been studied in [1] for a single channel case. The authors found an optimal closed-form expression to the amount of power that the SU can transmit under this interference constraint. In a multi-channel system, the problem becomes more interesting since the SU faces heterogeneous interference constraints. In this case, the SU will have to decide how much power should be allocated to each channel. This problem has been studied in [2] with the objective of maximizing the SU's throughput under per-channel interference constraints.

In some scenarios the SU may have limited sensing capabilities so that only one channel can be sensed at a time. In [3] the authors considered this problem and showed how to optimally select this channel to maximize the SU's throughput. In a time-slotted system, if a channel is detected busy, then the time-slot is wasted. Thus, in [4] the authors allowed multiple channels to be sensed in the same time-slot sequentially. Because the duration of sensing a channel may be large, the SU is allowed to stop sensing at some

channel and begin transmission. The authors discussed the optimal stopping rule and proposed a dynamic programming algorithm to find the optimal sensing order that maximizes the SU's throughput. The channels were assumed to have heterogeneous channel availability probabilities but had independent and identically distributed (i.i.d.) fading. Optimal Stopping Rule has been considered in [5] as well, where the authors generalized the setup to heterogeneous fading channels and devised a polynomial-time algorithm to find the optimal stopping rule.

In real-time applications, such as video and audio streaming, delay is an important factor in evaluating the performance of a communication system, because packets are expected to arrive at the destination before some delay constraint. In this work we study the fundamental trade-off between throughput and delay in a CR context. We propose a channel sensing and access scheme and find the optimal stopping rule that maximizes a single SU's throughput while maintaining delay bounds on the packets of the CR user. To the best of our knowledge, throughput-delay trade-off has not been considered before in multichannel CR systems.

The rest of this paper is organized as follows. We present our system model in Section II. In Section III we formulate the problem mathematically. The optimal solution is discussed in Section IV by solving a special case of the problem to get some insights, then we present the optimal solution to the original problem along with a discussion of our low-complexity approach. Simulation results are presented in Section V, and the paper is concluded in Section VI.

II. SYSTEM MODEL

We adopt a time-slotted system with slot duration of T seconds. The PU has M licensed frequency channels to access. We assume that the PU's transmission begins only at the beginning of a time-slot, and stops at the end of this time-slot with the opportunity to transmit on subsequent time-slots independently. PU's transmission is independent across channels and across time-slots. On the other hand, we have a single SU trying to access one of these M channels. Before a time-slot begins, the SU is assumed to have ordered the channels according to some sequence¹. We denote the first channel in this sequence as channel 1, the second channel as 2, and so on until the last channel being M . Before the SU attempts to transmit its packet over channel i , it senses

¹The method of ordering the channels is outside the scope of this work. The reader is referred to [4] for further details about channel ordering.

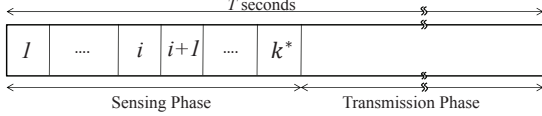


Fig. 1. Sensing and transmission phases in one time-slot. In this example, transmission begins after sensing the k^* th frequency channel.

this channel to determine its “state” which is described by a Bernoulli random variable b_i with parameter θ_i . If $b_i = 1$ (which happens with probability θ_i), then channel i is free and the SU can transmit over it until the on-going time-slot ends. If $b_i = 0$, channel i is busy, and the SU proceeds to sense channel $i + 1$.

We assume that the SU’s sensing errors are negligible but the SU has limited capabilities in the sense that no two channels can be sensed simultaneously. This may be the case when considering radios having a single sensing module with a fixed bandwidth, so that it can be tuned to only one frequency channel at a time. Therefore, at the beginning of a given time-slot, the SU selects a channel, say channel 1, senses it for τ seconds ($\tau \ll T/M$), and transmits over this channel if it is free. Otherwise, the SU skips this channel and senses channel 2, and so on until it finds a free channel. If all channels are busy (i.e. the PU has transmission activities on all M channels) then this time-slot will be considered as “wasted”. In this case, the SU waits for the following time-slot and begins sensing following the same channel sensing sequence.

The physical channel between the SU transmitter and corresponding receiver is assumed to be flat fading with independent, identically distributed (i.i.d.) channel gains among the M channels. To achieve higher data rates, the SU uses adaptive modulation. This means that the SU adapts the data rate according to the instantaneous power gain of the channel before beginning transmission on this channel. To do this, once the SU finds a free channel, say channel i , the gain γ_i is probed. The data rate will be proportional to $\log(1 + P_i(\gamma_i)\gamma_i)$, where $P_i(\gamma_i)$ (or simply P_i) is the power transmitted by the SU at channel i as a function of the instantaneous gain. Fig. 1 shows a potential scenario where the SU senses k^* channels, skips the first $k^* - 1$, and uses the k^* th channel for transmission until the end of this on-going time-slot. In this scenario the SU “stops” at the k^* th channel. Clearly, $k^* \in \{1, \dots, M\}$ is a random variable that changes from one time-slot to another. It depends on the states $[b_1, \dots, b_M]$ along with the gains of each channel $[\gamma_1, \dots, \gamma_M]$. To understand why, consider that the SU senses channel i , finds it free and probes its gain γ_i . If γ_i is found to be low, then the SU skips channel i (although free) and senses channel $i + 1$. This is to take advantage of the possibility that $\gamma_{i+1} \gg \gamma_i$. On the other hand, if γ_i is sufficiently large, the SU stops at channel i and begins transmission. In that latter case $k^* = i$. Defining the two random vectors $\underline{b} = [b_1, \dots, b_M]^T$ and $\underline{\gamma} = [\gamma_1, \dots, \gamma_M]^T$, k^* is a deterministic function of \underline{b} and $\underline{\gamma}$.

We define the stopping rule by defining a threshold $\gamma_{\text{th}}(i)$ to which each γ_i is compared when the i th channel is found

free. If $\gamma_i > \gamma_{\text{th}}(i)$, the SU stops and transmits at channel i . Otherwise, channel i is skipped and channel $i + 1$ sensed. If $\gamma_{\text{th}}(i) = 0$, then whenever channel i is sensed free, the SU will not skip it. Increasing $\gamma_{\text{th}}(i)$ allows the SU to skip channel i whenever $\gamma_i < \gamma_{\text{th}}(i)$, to search for a better channel, thus potentially increasing the throughput. Setting $\gamma_{\text{th}}(i)$ too large allows channel i to be skipped even if γ_i is high. This constitutes the trade-off on choosing the values of $\gamma_{\text{th}}(i)$ ’s. The optimal values of $\gamma_{\text{th}}(i)$ $i = 1, \dots, M$, determine the optimal stopping rule.

A. Average Power

Now we will formulate the problem under some average power constraint. Let P_i denote the power used at the i th channel, which is zero unless $i = k^*$. Defining $c_i := 1 - \frac{i\tau}{T}$ which is the fraction of the time-slot remaining for the SU’s transmission if the SU transmits on the i th channel in the sensing sequence. The average power constraint is $\mathbb{E}[c_{k^*} P_{k^*}(\gamma_{k^*})] \leq P_{\text{avg}}$, where the expectation is over \underline{b} and $\underline{\gamma}$. This expectation can be calculated recursively. Let

$$S_i(\mathbf{\Gamma}_i^{\text{th}}, \mathbf{P}_i) = \theta_i c_i \int_{\gamma_{\text{th}}(i)}^{\infty} P_i f_{\gamma_i}(\gamma) d\gamma + [1 - \theta_i \bar{F}_{\gamma_i}(\gamma_{\text{th}}(i))] S_{i+1}(\mathbf{\Gamma}_{i+1}^{\text{th}}, \mathbf{P}_{i+1}), \quad (1)$$

$i = 1, \dots, M$, with $S_{M+1}(\mathbf{\Gamma}_{M+1}^{\text{th}}, \mathbf{P}_{M+1}) := 0$, while $\mathbf{P}_i := [P_i, \dots, P_M]^T$, $\mathbf{\Gamma}_i^{\text{th}} := [\gamma_{\text{th}}(i), \dots, \gamma_{\text{th}}(M)]^T$, and $f_{\gamma_i}(\gamma)$ is the Probability Density Function (PDF) of the gain γ_i of channel i , and $\bar{F}_{\gamma_i}(x) := \int_x^{\infty} f_{\gamma_i}(\gamma) d\gamma$. The first term in (1) is the average power transmitted at channel i given that channel is chosen for transmission (i.e. given that $k^* = i$). The second term represents the case where channel i is skipped and channel $i + 1$ is sensed. It can be shown that

$$S_1(\mathbf{\Gamma}_1^{\text{th}}, \mathbf{P}_1) = \mathbb{E}[c_{k^*} P_{k^*}]. \quad (2)$$

We note that we will drop the index i from the subscript of $f_{\gamma_i}(\gamma)$ and $\bar{F}_{\gamma_i}(\gamma)$ since channels suffer i.i.d. fading.

B. Throughput

The SU’s average throughput is defined as $\mathbb{E}[c_{k^*} \log(1 + P_{k^*} \gamma_{k^*})]$. Similarly, we denote the expected throughput as $U_1(\mathbf{\Gamma}_1^{\text{th}}, \mathbf{P}_1)$ which can be derived using the following recursive formula

$$U_i(\mathbf{\Gamma}_i^{\text{th}}, \mathbf{P}_i) = \theta_i c_i \int_{\gamma_{\text{th}}(i)}^{\infty} \log(1 + P_i \gamma) f_{\gamma}(\gamma) d\gamma + \theta_i U_{i+1}(\mathbf{\Gamma}_{i+1}^{\text{th}}, \mathbf{P}_{i+1}) \int_0^{\gamma_{\text{th}}(i)} f_{\gamma}(\gamma) d\gamma + (1 - \theta_i) U_{i+1}(\mathbf{\Gamma}_{i+1}^{\text{th}}, \mathbf{P}_{i+1}), \quad (3)$$

$i = 1, \dots, M$, with $U_{M+1} := 0$.

C. Delay

If the SU skips all channels, either due to being busy or due to their low gain, then this time-slot is wasted and a *blocking* event occurs. The SU has to wait for the following time-slot to begin searching for a free channel again. This results in a delay to the SU’s transmitted packet. In real-time applications, there may exist some average delay requirement \bar{D}_{max} on each packet. That is, the average packet delay may not exceed \bar{D}_{max} .

Define the delay D as the number of time-slots the SU wastes, due to blocking, while trying to transmit a given

packet. Here, we put a constraint on the expected value of D , $\mathbb{E}[D]$. Since the availability of each channel is independent across time-slots, D follows a geometric distribution having $\mathbb{E}[D] = (\text{Pr}[\text{Success}])^{-1}$ where $\text{Pr}[\text{Success}] = 1 - \text{Pr}[\text{Blocking}]$ is the probability of not wasting a time-slot, i.e. no blocking occurs. In other words, $\text{Pr}[\text{Success}]$ is the probability that the SU finds a free channel having a high gain that makes it not skipped, out of the M possible channels. It is given by $\text{Pr}[\text{Success}] := p_1(\mathbf{\Gamma}_1^{\text{th}})$ which is calculated recursively using the following equation

$$p_i(\mathbf{\Gamma}_i^{\text{th}}) = \theta_i \bar{F}_\gamma(\gamma_{\text{th}}(i)) + [1 - \theta_i \bar{F}_\gamma(\gamma_{\text{th}}(i))] p_{i+1}(\mathbf{\Gamma}_{i+1}^{\text{th}}),$$

$i = 1, \dots, M$, where $p_{M+1} := 0$. Here, $p_i(\mathbf{\Gamma}_i^{\text{th}})$ is the probability that the SU transmits on channel i , $i + 1, \dots$, or M .

III. PROBLEM STATEMENT

From equation (3) we can see that the SU's expected throughput, U_1 , is affected by the thresholds $\gamma_{\text{th}}(i)$'s. The goal is to find the optimum value of $\gamma_{\text{th}}(i)$'s that maximizes U_1 subject to an expected packet delay constraint. The delay constraint can be written as $\mathbb{E}[D] \leq \bar{D}_{\text{max}}$ or, equivalently, $p_1(\mathbf{\Gamma}_1^{\text{th}}) \geq \frac{1}{\bar{D}_{\text{max}}}$. Mathematically, the problem becomes

$$\begin{aligned} & \text{maximize} && U_1(\mathbf{\Gamma}_1^{\text{th}}, \mathbf{P}_1) \\ & \text{subject to} && S_1(\mathbf{\Gamma}_1^{\text{th}}, \mathbf{P}_1) \leq P_{\text{avg}} \\ & && p_1(\mathbf{\Gamma}_1^{\text{th}}) \geq \frac{1}{\bar{D}_{\text{max}}} \\ & \text{variables} && \mathbf{\Gamma}_1^{\text{th}}, \mathbf{P}_1, \end{aligned} \quad (4)$$

where the first constraint represents the average power constraint, while the second is a bound on the average packet delay.

IV. OPTIMAL SOLUTION

A. Two-Level Power Control: A Special Case

First we solve this problem assuming $P_i = 1$ for $i = k^*$ and $P_i = 0$ for $i \neq k^*$ as a special case to get some insights, then we generalize the solution to the more general case where $P_i \equiv P_i(\gamma_i)$ are functions of the channel gain. The optimization problem in (4) becomes

$$\begin{aligned} & \text{maximize} && U_1(\mathbf{\Gamma}_1^{\text{th}}, \mathbf{1}_M) \\ & \text{subject to} && p_1(\mathbf{\Gamma}_1^{\text{th}}) \geq \frac{1}{\bar{D}_{\text{max}}} \\ & \text{variables} && \mathbf{\Gamma}_1^{\text{th}}, \end{aligned} \quad (5)$$

where $\mathbf{1}_M$ is a vector with M ones. Let λ_D be the Lagrange multiplier associated with the delay constraint in (5). The Lagrangian associated with problem (5) is

$$L(\mathbf{\Gamma}_1^{\text{th}}, \lambda_D) = U_1(\mathbf{\Gamma}_1^{\text{th}}, \mathbf{1}_M) + \lambda_D \left(p_1(\mathbf{\Gamma}_1^{\text{th}}) - \frac{1}{\bar{D}_{\text{max}}} \right). \quad (6)$$

To solve problem (5) we find the KKT equations that represent necessary conditions for the optimal solution. The first equation among those KKT equations is obtained by differentiating (6) with respect to $\gamma_{\text{th}}(i)$ for each $i \in \{1, \dots, M\}$

and equating each derivative to zero. Thus we get

$$\gamma_{\text{th}}^*(i) = \left[\exp \left(\frac{U_{i+1}(\mathbf{\Gamma}_{i+1}^{\text{th}*}, \mathbf{1}_{M-i}) - \lambda_D^* \cdot (1 - p_{i+1}^*)}{c_i} \right) - 1 \right]^+, \quad (7)$$

$i = 1, \dots, M$, where $(x)^+ := \max(x, 0)$, $\mathbf{\Gamma}_{i+1}^{\text{th}*} := [\gamma_{\text{th}}^*(i + 1), \dots, \gamma_{\text{th}}^*(M)]^T$ while $p_{i+1}^* := p_{i+1}(\mathbf{\Gamma}_{i+1}^{\text{th}*})$, $p_{M+1}^* := 0$ and $U_{M+1}(\mathbf{\Gamma}_{M+1}^{\text{th}*}, \mathbf{1}_0) := 0$. The second KKT equation is $\lambda_D^* \cdot \left(p_1^* - \frac{1}{\bar{D}_{\text{max}}} \right) = 0$. Thus, λ_D^* is chosen to satisfy the inequality constraint in (5) with equality (i.e. $p_1^* = \frac{1}{\bar{D}_{\text{max}}}$). Hence, we substitute by $\gamma_{\text{th}}^*(i)$ in $p_1(\mathbf{\Gamma}_1^{\text{th}})$ to get p_1^* as a function of λ_D^* . Then we solve for λ_D^* to satisfy $p_1^* = \frac{1}{\bar{D}_{\text{max}}}$. The optimality of our proposed approach is discussed in the following theorem.

Theorem 1. *The expression for $\gamma_{\text{th}}^*(i)$ in (7), along with λ_D^* that solves $p_1(\mathbf{\Gamma}_1^{\text{th}}) = \frac{1}{\bar{D}_{\text{max}}}$, represent the optimal solution to problem (5).*

Proof Sketch: The proof depends on the fact that p_1^* is strictly monotonic in λ_D^* . This indicates that there exists a unique λ_D^* satisfying $p_1^* = \frac{1}{\bar{D}_{\text{max}}}$. Thus $\gamma_{\text{th}}^*(i)$ is unique. Consequently, the KKT equations are sufficient for optimality. \square

B. Optimal Power Control: General Case

Now we allow the power P_i to be an arbitrary function of γ_i and optimize over this function to maximize the throughput subject to average power and delay constraints. The optimization problem is expressed by (4). Let λ_P and λ_D be the dual variables associated with the constraints in problem (4). The Lagrangian for (4) becomes

$$\begin{aligned} L(\mathbf{\Gamma}_1^{\text{th}}, \mathbf{P}_1, \lambda_P, \lambda_D) &= U_1(\mathbf{\Gamma}_1^{\text{th}}, \mathbf{P}_1) - \\ & \lambda_P (S_1(\mathbf{\Gamma}_1^{\text{th}}, \mathbf{P}_1) - P_{\text{avg}}) + \\ & \lambda_D \left(p_1(\mathbf{\Gamma}_1^{\text{th}}) - \frac{1}{\bar{D}_{\text{max}}} \right). \end{aligned} \quad (8)$$

The KKT conditions [6] are necessary equations for the optimal point of problem (4). To get the KKT equations associated with this problem, we differentiate (8) with respect to (w.r.t.) each of the primal variables P_i and $\gamma_{\text{th}}(i)$ and equate the resulting derivatives each to zero. We also have the primal feasible, dual feasible and complementary slackness equations. Thus the KKT equations are

$$P_i^*(\gamma) = \left(\frac{1}{\lambda_P^*} - \frac{1}{\gamma} \right)^+, \quad i = 1, \dots, M, \quad (9)$$

$$\begin{aligned} & c_i \left(\log \left(\frac{\gamma_{\text{th}}^*(i)}{\lambda_P^*} \right) - \lambda_P^* \left(\frac{1}{\lambda_P^*} - \frac{1}{\gamma_{\text{th}}^*(i)} \right)^+ \right) = \\ & U_{i+1}^* - \lambda_P^* S_{i+1}^* - \lambda_D^* \cdot (1 - p_{i+1}^*), \quad i = 1, \dots, M, \end{aligned} \quad (10)$$

$$S_1^* \leq P_{\text{avg}}, \quad p_1^* \geq \frac{1}{\bar{D}_{\text{max}}} \quad (11)$$

$$\lambda_P^* \geq 0, \quad \lambda_D^* \geq 0 \quad (12)$$

$$\lambda_P^* \cdot (S_1^* - P_{\text{avg}}) = 0 \quad (13)$$

$$\lambda_D^* \cdot \left(p_1^* - \frac{1}{\bar{D}_{\text{max}}} \right) = 0, \quad (14)$$

where $\mathbf{P}_{i+1}^* := [P_{i+1}^*, \dots, P_M^*]^T$ is the optimal power allocation vector. We use $U_{i+1}^* := U_{i+1}(\mathbf{\Gamma}_{i+1}^{\text{th}*}, \mathbf{P}_{i+1}^*)$, $U_{M+1}^* := 0$, $S_{i+1}^* := S_{i+1}(\mathbf{\Gamma}_{i+1}^{\text{th}*}, \mathbf{P}_{i+1}^*)$, $S_{M+1}^* := 0$ and $p_{M+1}^* := 0$ in the sequel.

1) *Primal Variables:* We can see that equation (9) is the *water-filling* strategy that gives the closed-form optimum solution for the first M primal variables, namely P_i ($i = 1, \dots, M$). The remaining M primal variables, namely $\gamma_{\text{th}}^*(i)$ ($i = 1, \dots, M$), are found via the set of equations (10). Note that equations (10) indeed form a set of M equations, each solves for one of the $\gamma_{\text{th}}^*(i)$'s ($i = 1, \dots, M$). We refer to this set as the “ γ -finding” equations. We also note that for a given value of i , solving for $\gamma_{\text{th}}^*(i)$ needs the knowledge of only $\gamma_{\text{th}}^*(i+1)$ through $\gamma_{\text{th}}^*(M)$, and does not require knowing $\gamma_{\text{th}}^*(1)$ through $\gamma_{\text{th}}^*(i-1)$. In other words, the equation that finds $\gamma_{\text{th}}^*(i)$ is a function of only $\gamma_{\text{th}}^*(i)$ through $\gamma_{\text{th}}^*(M)$. Thus, these M equations can be solved using back-substitution. The optimal solution for the γ -finding equations is given by

$$\gamma_{\text{th}}^*(i) = \frac{-\lambda_{\text{P}}^*}{W_0 \left[-\exp \left(-\frac{[U_{i+1}^* - \lambda_{\text{P}}^* S_{i+1}^* - \lambda_{\text{D}}^* (1-p_{i+1}^*)]^+}{c_i} - 1 \right) \right]}, \quad (15)$$

$i = 1, \dots, M$, and $W_0[z]$ is called the principle value of the Lambert W function [7] and is given by

$$W_0[z] = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} z^n. \quad (16)$$

Although the γ -finding equations are necessary equations for the optimal point, Theorem 2 proves that the solution $\mathbf{\Gamma}_1^{\text{th}*} := [\gamma_{\text{th}}^*(1), \dots, \gamma_{\text{th}}^*(M)]^T$ is unique given the dual variables $(\lambda_{\text{P}}, \lambda_{\text{D}})$. This indicates that these equations are sufficient as well, proving that the solution to the γ -finding equations is optimum for problem (4) given the optimum dual variables $(\lambda_{\text{P}}^*, \lambda_{\text{D}}^*)$.

Theorem 2. *Given some value for $(\lambda_{\text{P}}^*, \lambda_{\text{D}}^*)$, the left-hand-side (l.h.s.) of equation (10) is monotonically increasing in $\gamma_{\text{th}}^*(i)$. Thus the solution of $\gamma_{\text{th}}^*(i)$ that satisfies equation (10) is unique.*

Proof: Let the l.h.s. of equation (10) be

$$Y(\gamma_{\text{th}}^*(i)) = c_i \left(\log \left(\frac{\gamma_{\text{th}}^*(i)}{\lambda_{\text{P}}^*} \right) - \lambda_{\text{P}}^* \left(\frac{1}{\lambda_{\text{P}}^*} - \frac{1}{\gamma_{\text{th}}^*(i)} \right)^+ \right). \quad (17)$$

If $\gamma_{\text{th}}^*(i) < \lambda_{\text{P}}^*$. Then $Y(\gamma_{\text{th}}^*(i)) = c_i \log \left(\frac{\gamma_{\text{th}}^*(i)}{\lambda_{\text{P}}^*} \right)$ which is an increasing function in $\gamma_{\text{th}}^*(i)$. On the other hand, if $\gamma_{\text{th}}^*(i) > \lambda_{\text{P}}^*$, $Y(\gamma_{\text{th}}^*(i))$ becomes

$$Y(\gamma_{\text{th}}^*(i)) = c_i \left(\log \left(\frac{\gamma_{\text{th}}^*(i)}{\lambda_{\text{P}}^*} \right) - \left(1 - \frac{\lambda_{\text{P}}^*}{\gamma_{\text{th}}^*(i)} \right) \right). \quad (18)$$

Differentiating $Y(\gamma_{\text{th}}^*(i))$ w.r.t. $\gamma_{\text{th}}^*(i)$ we get

$$\frac{\partial Y}{\partial \gamma_{\text{th}}^*(i)} = c_i \left(\frac{1}{\gamma_{\text{th}}^*(i)} - \frac{\lambda_{\text{P}}^*}{\gamma_{\text{th}}^2(i)} \right) \quad (19)$$

$$> c_i \left(\frac{1}{\gamma_{\text{th}}^*(i)} - \frac{\lambda_{\text{P}}^*}{\lambda_{\text{P}}^* \gamma_{\text{th}}^*(i)} \right) \quad (20)$$

$$= 0. \quad (21)$$

This means that $Y(\gamma_{\text{th}}^*(i))$ is monotonic in $\gamma_{\text{th}}^*(i)$. Hence, $\gamma_{\text{th}}^*(i)$ satisfying equation (10) is unique. \square

To solve any of the γ -finding equations, we assumed the knowledge of the dual variables λ_{P}^* and λ_{D}^* . We will discuss next how to get these dual variables.

2) *Dual Variables:* The optimum dual variable λ_{P}^* must satisfy equation (13). Thus if $\lambda_{\text{P}}^* > 0$, then we need $S_1^* - P_{\text{avg}} = 0$. λ_{P}^* that satisfies the latter equation gives the optimum solution. The same goes for λ_{D}^* ; if $\lambda_{\text{D}}^* > 0$, then $p_1^* - \frac{1}{D_{\text{max}}} = 0$. These two equations can be solved using a suitable root-finding algorithm (e.g. the bisection method [8]). To find $(\lambda_{\text{P}}^*, \lambda_{\text{D}}^*)$, we propose Algorithm 1 that executes two nested bisection algorithms. We note that Algorithm 1 can be systematically modified to call any other root-finding algorithm (e.g. the secant algorithm [8] converges faster than the bisection algorithm).

Algorithm 1 Finding $(\lambda_{\text{P}}^*, \lambda_{\text{D}}^*)$

- 1: Initialize $t := 1$, $\lambda_{\text{D}}^{(1)} := \lambda_{\text{D}}^{\text{initial}}$
 - 2: **while** $p_1^* - \frac{1}{D_{\text{max}}} > \epsilon$ **do**
 - 3: Initialize $n := 1$, $\lambda_{\text{P}}^{(1)} := \lambda_{\text{P}}^{\text{initial}}$
 - 4: **while** $S_1^* - \frac{1}{D_{\text{max}}} > \epsilon$ **do**
 - 5: Update $\lambda_{\text{P}}^{(n+1)}$ with the bisection algorithm's update equation [8]
 - 6: $n \leftarrow n + 1$
 - 7: **end while**
 - 8: Update $\lambda_{\text{D}}^{(t+1)}$ with the bisection algorithm's update equation [8]
 - 9: $t \leftarrow t + 1$
 - 10: **end while**
 - 11: $(\lambda_{\text{P}}^*, \lambda_{\text{D}}^*) := (\lambda_{\text{P}}^{(n)}, \lambda_{\text{D}}^{(t)})$
-

V. SIMULATION RESULTS

The two-level power control system was simulated considering a Rayleigh-fading channel between the secondary transmitter and the intended secondary receiver. The channel has an exponential gain distribution (i.e. $f_{\gamma}(\gamma) = \frac{1}{\bar{\gamma}} \exp(-\gamma/\bar{\gamma})$, where $\bar{\gamma}$ is the average channel gain of the transmitted assuming unity noise variance). We assumed $M = 10$ channels each having $\theta_i = 0.1$, $\tau/T = 0.05$. To select a suitable value for \bar{D}_{max} , we note that small values for \bar{D}_{max} may lead to an infeasible problem where the system is not able to satisfy this small delay constraint. On the other hand, large values for \bar{D}_{max} may yield a trivial solution that can be found by neglecting the delay constraint in our optimization problem. Thus, we set $\bar{D}_{\text{max}} = 1.54T$ which corresponds to the minimum delay that the system can achieve (i.e. $\frac{1}{1 - \prod_{i=1}^M (1 - \theta_i)}$). The performance of the

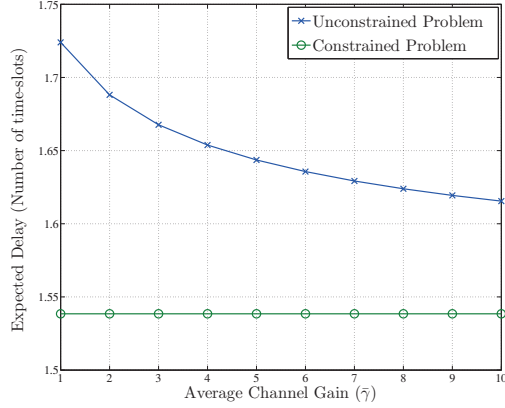


Fig. 2. The expected delay is bounded in the constrained maximization problem.

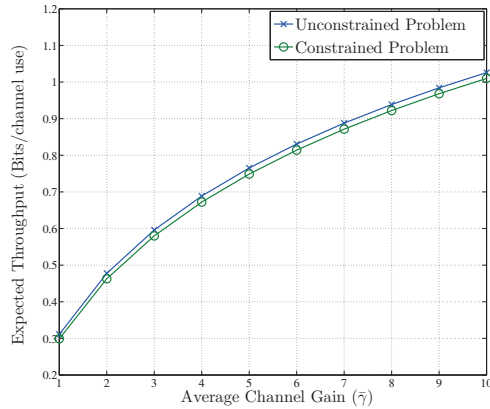


Fig. 3. Throughput of the constrained system compared to the unconstrained one. The proposed model does not sacrifice much throughput. The relative gap difference is less than 4% at $\bar{\gamma} \geq 1$.

system was compared to that of an unconstrained throughput-maximization problem for different values of $\bar{\gamma}$. Fig. 2 compares the expected delay versus the average channel gain $\bar{\gamma}$. In the unconstrained scenario, the expected delay is not controlled and exceeds the maximum average-delay-constraint \bar{D}_{\max} that the system can tolerate. On the other hand, when adding the delay constraint to the optimization problem, we guarantee that the delay will be bounded below \bar{D}_{\max} without sacrificing much throughput. The throughput for both cases is shown in Fig. 3 where the relative gap is less than 4% for $\bar{\gamma} = 1$ and decreases as $\bar{\gamma}$ increases.

For the optimum power control scenario, the same channel model was used. For a fair comparison with the two-level power control system, the average power was chosen to be equal to the expected power in the two-level power control system. Fig. 4 compares the expected throughput of the optimal power control system to that of the two-level power control system. At low average channel gain values the throughput increases by about 34% when allocating the power using the water-filling algorithm. Although the relative gap decreases with the average channel gain, the increase in throughput is still significant (more than 6% at $\bar{\gamma} = 10$).

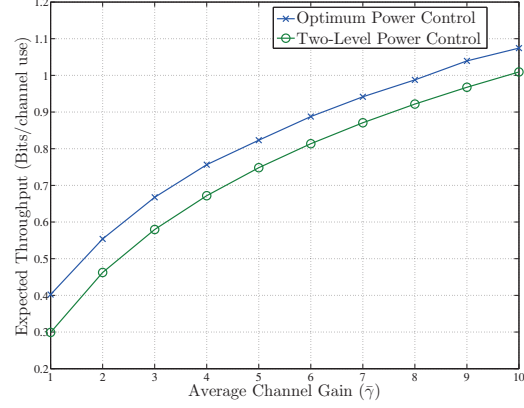


Fig. 4. Comparison of the throughput with and without optimal power control. Using optimal power control increases the throughput by more than 6% at $\bar{\gamma} \leq 10$.

VI. CONCLUSION AND FUTURE WORK

In this work, we formulate a throughput maximization problem constrained by a bound on the expected delay. We provide an optimal solution that has a closed-form expression for this optimization problem when we adopt the two-level power control strategy. Simulations show that constraining the delay will not degrade the throughput significantly, yet will guarantee average delay bounds to the packets of the CR user. Yet when the optimum power control strategy is adopted, we find an improvement in the throughput of about 34% when the average channel gain is 10.

In this paper, we neglected the sensing errors of the SU assuming perfect sensing. The problem becomes more interesting when the false-alarm and miss-detection probabilities are considered. This is because these errors result in higher delay since a false-alarm event leads to wasting a potential transmission opportunity, while a miss-detection event results in colliding with the PU's signal leading to the SU's receiver not being able to decode, thus wasting a time-slot.

REFERENCES

- [1] V. Asghari and S. Aissa, "Adaptive Rate and Power Transmission in Spectrum-Sharing Systems," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 10, pp. 3272–3280, Oct 2010.
- [2] Y. Zhang and C. Leung, "Resource Allocation in an OFDM-based Cognitive Radio System," *Communications, IEEE Transactions on*, vol. 57, no. 7, pp. 1928–1931, Jul 2009.
- [3] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *Selected Areas in Communications, IEEE Journal on*, vol. 25, no. 3, pp. 589–600, Apr 2007.
- [4] H. Jiang, L. Lai, R. Fan, and H. Poor, "Optimal Selection of Channel Sensing Order in Cognitive Radio," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 1, pp. 297–307, Jan. 2009.
- [5] L. Zappaterra, J. S. Gomes, A. Arora, and H.-A. Choi, "A Polynomial-Time Algorithm for Optimizing Channel Selection in Cognitive Radio Networks," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2013 9th International*, 2013, pp. 1559–1564.
- [6] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.
- [7] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W Function," in *ADVANCES IN COMPUTATIONAL MATHEMATICS*, 1996, pp. 329–359.
- [8] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3rd ed. New York, NY, USA: Cambridge University Press, 2007.