Effective Capacity and Delay Optimization in Cognitive Radio Networks

Mai Abdel-Malek[1(✉)], Karim Seddik[2], Tamer ElBatt[1,3], and Yahya Mohasseb[1,4]

[1] Wireless Intelligent Networks Center (WINC), Nile University, Giza, Egypt
m.elkady@nileu.edu.eg, {telbatt, mohasseb}@ieee.org
[2] ECNG Department, American University in Cairo, New Cairo 11835, Egypt
kseddik@aucegypt.edu
[3] Department of EECE, Faculty of Engineering, Cairo University, Giza, Egypt
[4] Department of Communications, The Military Technical College, Cairo, Egypt

Abstract. In this paper, we study the fundamental trade-off between delay-constrained primary and secondary users in cognitive radio networks. In particular, we characterize and optimize the trade-off between the secondary user (SU) effective capacity and the primary user (PU) average packet delay. Towards this objective, we employ Markov chain models to quantify the SU effective capacity and average packet delay in the PU queue. Afterwards, we formulate two constrained optimization problems to maximize the SU effective capacity subject to an average PU delay constraint. In the first problem, we use the spectrum sensing energy detection threshold as the optimization variable. In the second problem, we extend the problem and optimize also over the transmission powers of the SU. Interestingly, these complex non-linear problems are proven to be quasi-convex and, hence, can be solved efficiently using standard optimization tools. The numerical results reveal interesting insights about the optimal performance compared to the unconstrained PU delay baseline system.

Keywords: Cognitive radios · Effective capacity · Delay constraints · Optimization · Quality of service (QoS)

1 Introduction

The rapid evolution and ubiquity of wireless connectivity, as well as the wide proliferation of smartphones and powerful hand-held devices, mandate the handling of a huge amount of data in applications such as wireless multimedia. These applications are characterized by high bandwidth requirements and relatively stringent delay constraints. The limited wireless spectrum presents a major challenge and adds to the complexity of the problem.

This publication was made possible by NPRP 4-1034-2-385, NPRP 09-1168-2-455 and NPRP 5-782-2-322 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

M. Weichold et al. (Eds.): CROWNCOM 2015, LNICST 156, pp. 30–42, 2015.
DOI: 10.1007/978-3-319-24540-9_3

The concept of cognitive radios was originally introduced by J. Mitola III in 1999 as a paradigm shift due to the severe under-utilization of the spectrum in some bands [1]. Cognitive radios enable opportunistic, or secondary users (SUs), to share part of the spectrum with licensed spectrum owners, referred to as the primary users (PUs), and possibly coexist with them in some paradigms.

Striking a balance between improving the performance of the opportunistic SUs and maintaining QoS requirements for the PUs is crucial for cognitive radio systems. This is particularly true for PUs running multimedia applications with stringent QoS constraints, which are particularly sensitive to potential performance degradation caused by the SUs. A significant portion of research in the cognitive radios arena has focused on improving the ability of the SUs to communicate over their, inherently, unreliable links. This is achieved either through improvements in SU data encoding schemes, e.g., [2], [3] or by proper adaptation of the SU power and rate in response to the time-varying channel conditions to achieve SU QoS requirements, e.g., [4], [5], [6].

Physical-layer channel models cannot be easily linked to QoS metrics. Therefore, in [7], the notion of "Effective Capacity" was originally introduced to express the maximum constant arrival rate that can be supported by a given channel service process while satisfying a statistical QoS requirement as specified by the QoS exponent, θ. The effective capacity may be thought of as the dual wireless concept to the "Effective Bandwidth" notion originally introduced in [8].

Applied to cognitive radio systems, the effective capacity has been employed to characterize the performance of the SU in [9]. The authors derived expressions for the effective capacity of the SU for combinations of, both, fixed and adaptive power and rate scenarios. To enhance the prediction of the PU channel state, the authors in [10] proposed a feedback model, where the SU leverages the overheard primary ARQ message and uses information gleaned from the PU feedback channel to improve its sensing. It was shown, analytically, and using simulations that such side information can potentially increase the effective capacity of the SU.

In essence, the scheme in [10] minimizes the probability of PU *re-transmission* failure. However, this does not automatically guarantee the PU's ability to satisfy delay constraints. In this paper, we incorporate an explicit QoS constraint on the PU, namely an average delay constraint, which is more relevant to users engaged in interactive or multimedia sessions. Thus, in this paper our prime objective is to maximize the SU effective capacity under an average delay constraint on the PU packets.

Our main contribution in this paper is two-fold. First, we develop a mathematical model that incorporates delay constraints into the optimization of the primary users and secondary users performance in cognitive radio networks. Second, we formulate, establish quasi-convexity and efficiently solve two optimization problems for maximizing the SU effective capacity subject to a constraint on the PU average packet delay. Towards this objective, we analyze the Markov chain models capturing the SU channel sensing model and the PU queue dynamics. Next, we formulate, assess complexity (establish quasi-convexity) and solve our first (basic) optimization problem which decides the optimal spectrum sensing energy detection threshold to maximize the effective capacity subject to the average PU delay constraint. Afterwards, we generalize the problem to jointly optimize over the energy detection threshold along with the SU transmission powers, yielding superior performance. Finally, we solve the problems numerically and demonstrate promising performance results for plausible scenarios.

The rest of the paper is organized as follows. First, we introduce the system model, assumptions and secondary user channel access model in Section 2. Afterwards, the basic and generalized optimization problems are formulated, to maximize the SU effective capacity under an average PU delay constraint, and examined for convexity in Section 3. Performance results confirming the fundamental trade-off and the optimal solution are quantified under plausible scenarios in Section 4. Finally, conclusions are drawn and potential directions for future work are pointed out in Section 5.

2 System Model

We focus on a simple cognitive radio network with one PU and one SU for mathematical tractability of the proposed model. We consider a time slotted system with slots of equal duration, T seconds. All channels are assumed to experience Rayleigh block fading where the channel remains constant over a time slot and changes independently from one slot to another. The PU queue has a Bernoulli packet arrival process with rate $0 \leq \lambda_p \leq 1$ per slot and the SU is assumed to be fully backlogged (always has a packet to transmit) at the beginning of each time slot.

We assume a hybrid (underlay/interweave) cognitive radio model as introduced in [11], whereby the SU senses the channel at the beginning of the slot for N seconds, where $N < T$, in order to determine the mode of channel access, as in the interweave model. Subsequently, the SU accesses the channel with high power, P_i, if the primary user is idle. Otherwise, the SU accesses the channel with lower power $P_b < P_i$ and correspondingly lower rate (as in the underlay model). The discrete-time SU channel input-output relation in the i^{th} symbol duration is given by

$$y(i) = \begin{cases} \text{PU idle :} & h_s(i)x(i) + n(i), \qquad\quad i = 1, 2, \cdots \\ \text{PU active :} & h_s(i)x(i) + s_p(i) + n(i), \quad i = 1, 2, \cdots, \end{cases} \tag{1}$$

where $x(i)$ is the complex channel input, $y(i)$ denotes the complex channel output and $h_s(i)$ is the channel fading coefficient between the secondary transmitter and receiver and $|h_s(i)|^2 = z_s(i)$. The primary transmitted signal, as perceived by the secondary receiver, is denoted $s_p(i)$, and $n(i)$ denotes the additive white Gaussian noise at the secondary receiver, with zero mean and variance of σ_n^2.

We adopt a simple energy detection spectrum sensing mechanism whereby the RF energy measured at the secondary transmitter is compared to an energy detection threshold, η, to decide whether the PU is active or idle. The channel sensing problem is known to be modeled as a binary hypothesis testing problem [12]. Under these assumptions, the optimal Neyman-Pearson detector is [9]:

$$Y = \frac{1}{NB} \Sigma_{i=0}^{NB} |y(i)|^2 \lessgtr_{\mathcal{H}_1}^{\mathcal{H}_0} \eta, \tag{2}$$

where NB denotes the number of complex symbols in the N (in seconds) sensing duration with B (in Hz) channel bandwidth. The test statistic Y is chi-square distributed with $2NB$ degrees of freedom. In this case, the probability of detection can be derived as follows

$$P_d = \mathcal{P}rob\{Y > \eta|\mathcal{H}_1\} = 1 - \frac{\gamma\left(\frac{NB\eta}{(\sigma_n^2+\sigma_{sp}^2)}, NB\right)}{\Gamma(NB)}, \tag{3}$$

where $\gamma(x,s)$ is the lower incomplete gamma function and $\Gamma(x)$ is the Gamma function. The probability of false alarm can be written as follows

$$P_f = \mathcal{P}rob\{Y > \eta|\mathcal{H}_0\} = 1 - \frac{\gamma\left(\frac{NB\eta}{\sigma_n^2}, NB\right)}{\Gamma(NB)}. \tag{4}$$

2.1 The Secondary User Access Model

We employ a classic SU access model adopted earlier in the literature, e.g., [1,11], which is represented by the Markov chain in Fig.1. The prime objective of this model is to capture the secondary user access decision and transmission parameters, namely, power and rate, depending on the spectrum sensing outcome and the instantaneous capacity of the secondary user channel. This is instrumental in characterizing the secondary user effective capacity for a statistical QoS constraint, θ, as shown later in the sequel. Thus, the secondary user transmission parameters have four cases, depending on the sensing outcome:

1. The PU channel is busy and the SU detects it as such: SU transmits with the lowest acceptable power P_b and rate r_b.
2. The PU channel is idle while the SU detects it busy (false alarm (FA)): the SU sends its packet with power P_b and rate r_b as in case (1).
3. The PU channel is busy while the SU detects it idle (mis-detection (MD)): the SU sends with power P_i and rate r_i, where $P_i > P_b$ and $r_i > r_b$.
4. The PU channel is idle and the SU detects it as such: the SU sends with power P_i and rate r_i.

On the other hand, the secondary user channel has two states, OFF and ON, depending on whether the secondary user transmission rate exceeds the instantaneous channel capacity or not, respectively. This is caused by the time-varying fluctuations in the SU channel.

Next, we present the SU effective capacity (EC) subject to statistical QoS constraints. It has been established in [7] that the EC for a given QoS exponent, θ, is given by

$$EC = -\lim_{t\to\infty} \frac{1}{\theta t} \log_e \mathbb{E}\left\{e^{-\theta S(t)}\right\} = -\frac{\Lambda(-\theta)}{\theta}, \tag{5}$$

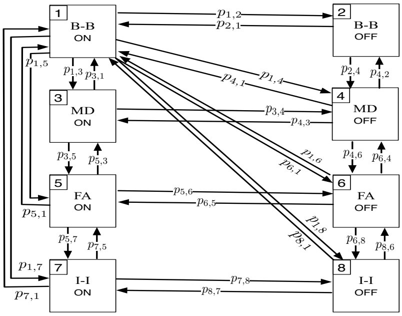

Fig. 1. SU access and transmission model.

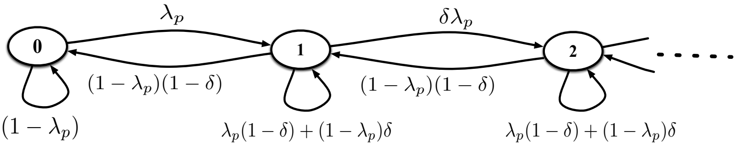

Fig. 2. A discrete-time Markov chain modeling the PU queue evolution.

where $S(t) = \sum_{k=1}^{t} r(k)$ represents the time accumulated service process and $\{r(k),\ k = 1, 2, ...\}$ is the discrete and ergodic stochastic service process.

For Markov modulated processes, like the model in Fig.1 , it has been shown in [13] that the effective capacity can be reduced to

$$EC = \frac{-1}{\theta TB} \log_e \left(\mathbf{sp}(\mathbf{\Phi}(-\theta)\mathbf{R}) \right), \tag{6}$$

where $\mathbf{sp}(.)$ denotes the spectral radius of a matrix, $\mathbf{R}$ represents the transition probability matrix of the secondary user Markov chain and θ is the delay exponent. $\mathbf{\Phi}(\theta) = diag(\phi_1(\theta), \ldots, \phi_M(\theta))$ is a diagonal matrix whose elements are the moment generating functions of the processes in the M states.

2.2 Modeling the Primary User Queue

The primary user queue evolution is modeled formally by the discrete-time Markov chain shown in Fig. 2 with Bernoulli arrival rate λ_p. The primary user service rate (i.e. queue length decreases by one), μ_p, depends only on two parameters, namely, the arrival rate and the primary channel outage probability, denoted δ, as follows,

$$\mu_p = (1 - \lambda_p)(1 - \delta). \tag{7}$$

As the SU always uses the spectrum with different powers and rates depending on the PU activity, the primary channel outage may occur in two cases: first, if the SU attempts to transmit on the PU channel with power P_b and rate r_b, which in addition to the PU rate, r, brings the total rate on the channel above the instantaneous capacity. Alternatively, PU channel outage occurs when the SU attempts to transmit on the PU channel with power P_i on a busy channel due to mis-detection. Therefore, we can express the primary outage probability as follows:

$$\delta = P_d \mathcal{P}rob\{z_p(i) < \phi_1\} + (1 - P_d)\mathcal{P}rob\{z_p(i) < \phi_2\}, \tag{8}$$

where $z_p(i) = |h_p(i)|^2$ and $h_p(i)$ is the fading coefficient of the PU transmission channel, $\phi_1 = \frac{2^{\frac{r}{B}}-1}{SINR_1}$ and $\phi_2 = \frac{2^{\frac{r}{B}}-1}{SINR_2}$. $SINR_1$ and $SINR_2$ denote the average signal to interference plus noise ratios for the PU in each outage case.

Solving the Markov chain in Fig. 2 which is M/M/1 queue with steady-state probability of the PU queue having i packets, π_i, then using simple queuing analysis (see [14]), π_0 can be derived as

$$\pi_0 = \frac{\delta(\mu_p - \delta\lambda_p)}{\mu_p(1+\delta) - \delta^2\lambda_p}, \tag{9}$$

and the utilization factor of the primary user queue, ρ, is

$$\rho = \frac{\delta\lambda_p}{\mu_p}, \tag{10}$$

and, finally, the expected delay for the PU packets can be derived as [14]

$$D_p = \frac{\rho(\pi_0 - \delta(1-\rho)^2)}{\delta^2\lambda_p(1-\rho)^2}. \tag{11}$$

3 SU Effective Capacity Optimization Under PU Average Delay Constraint

In this section, we study the problem of balancing the conflicting QoS requirements for the primary and secondary users in cognitive radio networks. In particular, we attempt to maximize the SU effective capacity (defined for a given statistical QoS exponent, θ) subject to an average PU delay constraint, denoted D_{max}. Towards this objective, we formulate two optimization problems. The first (basic) problem solves for the optimum energy detection threshold, η, that maximizes the SU effective capacity subject to the aforementioned PU constraint. The key insight guiding the choice of the energy detection threshold, η, is that it affects the power level used by the SU which, in turn, affects the channel outage probability. For this reason, increasing η always degrades the PU delay, yet, enhances the SU effective capacity up to some point beyond which it starts decreasing, as shown in Fig. 3. The explanation of this behavior is that when η exceeds the sensed PU RF energy, the SU will always detect an idle channel, and use high rate and power, hence, causing higher interference. On the other hand, when $0 \leqslant \eta \leqslant E_{s_p}(i)$ ($E_{s_p}(i)$ is the PU energy as perceived by the SU), the probability of miss-detection decreases with increasing η; as η gets closer to

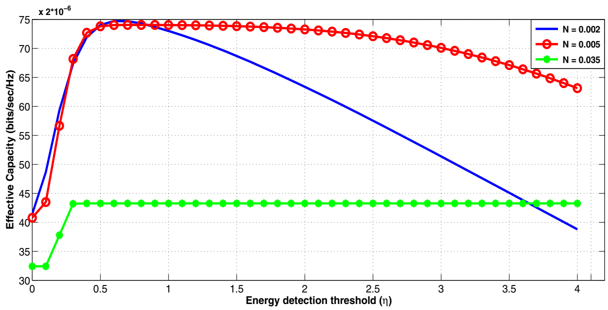

Fig. 3. The effect of the energy detection threshold, η, on the SU effective capacity for different sensing durations, N.

the sensed PU RF energy, the SU sensing is more reliable, hence, its effective capacity increases.

Fig. 3 shows the SU effective capacity for different values of the sensing duration, N. We notice that the relation between effective capacity and energy threshold can always be divided into the two modes discussed above; one increasing and the other is monotonically decreasing/non-increasing. We also notice that if we increase N from 0.002 sec to 0.005 sec, the SU effective capacity increases due to the more accurate estimate of the PU channel activity. On the other hand, when we increase N beyond 0.005 sec, the SU effective capacity decreases because the slot portion dedicated to data transmission, $T - N$, decreases linearly with N.

3.1 Basic Problem Formulation

In this section, we maximize the SU effective capacity with respect to the spectrum sensing energy detection threshold, η, subject to the PU average delay constraint. Thus, the basic optimization problem can be formulated as

$$\mathbf{P1} : \max_{\eta} \frac{-1}{\theta T B} \log_e(sp(\mathbf{\Phi}(-\theta)\mathbf{R})) \\ s.t.\ D_p \leq D_{max}, \tag{12}$$

where D_{max} is the maximum allowable average PU delay which represents the predefined QoS for the PU. Next, we establish an important complexity reduction result for **P1** which facilitates an efficient solution using standard optimization solvers.

Theorem 1. **P1** *is a quasi-convex problem in* η.

Proof. In order to establish this result, we must assess the quasi-concavity of the SU effective capacity as well as the convexity of the delay constraint.

We first establish the convexity of the constraint with respect to the optimization variable, η. From (8-11), the expected PU delay, D_p, is given by

$$D_p = \frac{-1}{(1-\lambda_p)(1-\delta)} + \frac{(1-\lambda_p)(1-\delta)}{(1-\lambda_p-\delta^2)(1-\lambda_p-\delta)}. \quad (13)$$

Since D_p consists of two terms, it suffices to prove that each term is convex in η. This follows from the fact that the summation of convex functions is convex. This can be validated be examining the Hessian of each term. Performing partial differentiation of each term with respect to δ, then from (3) and (8), we get the partial derivative of δ in terms of η. The Hessian of the first term, denoted f, is given by:

$$f'' = \frac{qe^{-x}}{(1-\delta)^2} x^{(NB-x)} \left(1 - NB - x\right), \quad (14)$$

where $x = \frac{\eta NB}{\sigma_n^2 + \sigma_{sp}^2}$, q is a positive constant. It turns out from (14) that the Hessian is positive definite only under the condition that

$$\eta \leqslant (\sigma_n^2 + \sigma_{sp}^2)\left(1 - \frac{1}{NB}\right). \quad (15)$$

For a stable system ($0 < \rho < 1$ and $0 \leqslant P_d \leqslant 1$), η will never violate this condition, then the first term of the constraint function is convex over the optimization domain. Similarly, we can prove that the second term is also convex over the optimization domain, and, hence, the constraint in (12) is convex.

Second, we establish the quasi-concavity of the objective function, namely the SU effective capacity. To prove this, we must show that the function inside the logarithm (in (6)) is quasi-convex. This is based on the fact that if an arbitrary function U is quasi-convex and a function g is monotonically non-decreasing, then the function f defined as $f(x) = g(U(x))$ is also quasi-convex [15]. Since the log function is non-decreasing, then it suffices to prove that its interior function is quasi-convex to establish the desired result. From (3), (4), (6) and ((28) in [9]), the objective function can be written as in (16) where $z_s = |h_s|^2$ and $\alpha_1, \alpha_2, \alpha_3$ and α_4 are the fading channel thresholds for no outage for the four SU sensing outcome cases stated in Section 2.1.

$$\begin{aligned} EC &= \tfrac{-1}{\theta TB} \log_e \Big\{ e^{-\theta r_b (T-N)} \left[\rho\left(1-\kappa_2\right)\mathbf{p}(z_s > \alpha_1) + (1-\rho)\left(1-\kappa_1\right)\mathbf{p}(z_s > \alpha_3)\right] \\ &+ e^{-\theta r_i (T-N)} \left[\rho\kappa_2 \mathbf{p}(z_s > \alpha_2) + (1-\rho)\kappa_1 \mathbf{p}(z_s > \alpha_4)\right] + \rho\left(1-\kappa_2\right)\mathbf{p}(z_s < \alpha_1) \\ &+ (1-\rho)\kappa_2 \mathbf{p}(z_s < \alpha_2) + (1-\rho)\left(1-\kappa_1\right)\mathbf{p}(z_s < \alpha_3) + (1-\rho)\kappa_1 \mathbf{p}(z_s < \alpha_4) \Big\} \end{aligned} \quad (16)$$

where $\kappa_1 = \gamma\left(\frac{\eta NB}{\sigma_n^2}, NB\right)/\Gamma(NB)$ and $\kappa_2 = \gamma\left(\frac{\eta NB}{\sigma_n^2+\sigma_{sp}^2}, NB\right)/\Gamma(NB)$. From (16), we have eight terms that can classified into three categories. First category is

$$f_1 = q_1\left(1 - \kappa_1\right), \quad (17)$$

second category is of the form

$$f_2 = q_2 \rho = q_2 \frac{c_1 - c_2 P_d}{k_1 + k_2 P_d}, \quad (18)$$

and the last category is of the form

$$f_3 = q_3 \rho P_d \ \text{ or } f_3 = q_4 \rho P_f, \tag{19}$$

where $q_1, q_2, c_1, c_2, k_1, k_2, q_3$ and q_4 are positive constants. For the first category

$$f_1'' = q\left(e^{-\frac{\eta NB}{\sigma_n^2}}\right)\left(\frac{\eta NB}{\sigma_n^2}\right)^{\left(NB-\frac{\eta NB}{\sigma_n^2}\right)}\left(1 - NB - \frac{\eta NB}{\sigma_n^2}\right), \tag{20}$$

which is convex only on the region indicated in (15), thus it is quasi-convex. Similarly, we can prove that all other terms are quasi-convex. Hence, their summation is quasi-convex. Thus, the effective capacity is quasi-concave in η since it is the negative of a quasi-convex function.

3.2 The Generalized Optimization Problem

In this section, we generalize the basic problem **P1** to optimize over the energy detection threshold, η, along with the SU transmission powers, P_b and P_i. As expected, the expanded policy space for the generalized problem yields noticeable performance improvement, as confirmed by the numerical results in Section 4, compared to the basic problem where the transmission powers are given and fixed throughout. Thus, the generalized optimization problem can be formulated as follows

$$\begin{aligned}\mathbf{P2}: \max_{\eta, P_b, P_i} \frac{-1}{\theta TB} \log_e(sp(\mathbf{\Phi}(-\theta)\mathbf{R}))\\ s.t.\ D_p \leq D_{max}\\ 0 \leqslant P_i \leqslant P\\ 0 \leqslant P_b \leqslant P,\end{aligned} \tag{21}$$

where P is the maximum SU transmission power. In preparation for our main result in this section establishing the quasi-convexity of **P2**, we utilize the following two definitions from optimization theory.

Definition 1. *The Hessian of a multi-variate function $f(\mathbf{x})$ is:*

$$\mathbf{H}_n(\mathbf{x}) = \begin{bmatrix} f''(x)_{11} & \cdots & f''(x)_{1n} \\ \vdots & \vdots & \vdots \\ f''(x)_{n1} & \cdots & f''(x)_{nn} \end{bmatrix}, \tag{22}$$

where $f''(\mathbf{x})_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. The function $f(\mathbf{x})$ is convex if its *Hessian* $\mathbf{H_n}(\mathbf{x})$ satisfies $\mathbf{H_n}(\mathbf{x}) \succeq 0$ which means that the Hessian matrix is positive semi-definite.

Definition 2. *The bordered Hessian of a multi-variate function $f(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, is given by:*

$$\mathbf{H}_n^b(\mathbf{x}) = \begin{bmatrix} 0 & f'(x)_1 & \cdots & f'(x)_n \\ f'(x)_1 & f''(x)_{11} & \cdots & f''(x)_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ f'(x)_n & f''(x)_{n1} & \cdots & f''(x)_{nn}, \end{bmatrix} \tag{23}$$

where $f'(\mathbf{x})_i = \frac{\partial f}{\partial x_i}$. If $|\mathbf{H}_n^b(x)| \leq 0$ for all n and x, then the function is quasi-convex (Note that the condition that $|\mathbf{H}_1^b| \leq 0$ is automatically satisfied).

Theorem 2. P2 *is a quasi-convex problem in the optimization variables* η, P_b, P_i

Proof. For problem **P2**, $n = 3$ as we have three optimization variables. We first start by establishing the convexity of the constraint with respect to the three optimization variables. Using (13) and (22), we can show that the delay constraint is convex and the other two constraints are affine in the optimization variables[1].

The only remaining step towards establishing the proof is to show the quasi-concavity of the objective function, EC, with respect to the three optimization variables. This involves characterizing the determinants of the bordered Hessian of the EC, defined in (23). Hence, from (23) and (16) we can simply show that $|\mathbf{H}_2^b| \leq 0$ and $|\mathbf{H}_3^b| \leq 0$ are satisfied and, hence, all terms are quasi-convex over the same domain 1. Thus, the effective capacity is quasi-concave in η, P_b and P_i since it is the negative of a quasi-convex function.

4 Numerical Results

In this section, we present the numerical results which confirm the fundamental trade-off under investigation, between the secondary user effective capacity and the primary user average packet delay. Moreover, we characterize the optimal solution using standard optimization tools for convex problems, e.g., CVX [15]. The system parameters used throughout this section are as follows: $r_b = 1.4$ Mbps, $r_i = 5.7$ Mbps, $P_b = 1$ unit power, $P_{max} = 2$ unit power, $P_i = 3$ unit

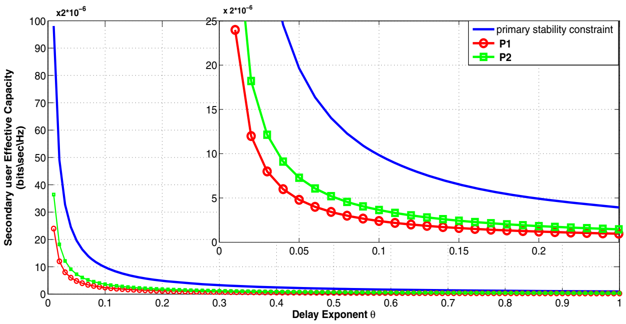

Fig. 4. Comparing the performance of the baseline (delay unconstrained system) to the optimal solutions of **P1** and **P2** ($D_{max} = 0.057$ sec)

[1] Details are omitted due to space limitations.

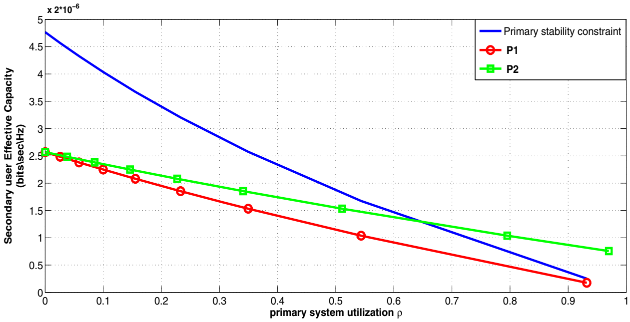

Fig. 5. The role of the primary user utilization, ρ.

power, $P = 3$ unit power, $B = 5$ MHz, $T = 0.1$ sec, $N = 0.002$ sec, $\lambda_p = 0.1$ and $\theta = 0.2$.

In Fig. 4, the effective capacity in (bit/sec/Hz) is plotted against the delay exponent, θ, for three different systems. The baseline is is a system with no constraints on the average PU packet delay and, hence in essence, it represents the case where the PU queue is stable "primary stability constraint". Thus the base line consider only maximizing the SU effective capacity which does not provide any delay guarantees for PU. On the other hand, **P1** represents the optimal solution for the basic problem which optimizes only over the energy detection threshold, η. **P2** represents the optimal solution for the generalized problem which optimizes over the three variables, η, P_b and P_i. A number of key observations are now in order. First, it is straightforward to observe that the SU effective capacity monotonically decreases under the three systems as the delay exponent increases, i.e. the statistical delay constraint becomes stricter. Second, the baseline (unconstrained PU delay) system achieves the highest SU effective capacity, as expected. However, it is also worth noting that the performance loss, due to the finite PU delay constraint in **P1** and **P2** diminishes as the SU statistical delay constraint, represented by the delay exponent, θ, becomes tighter. Third, we notice that the optimal solution for **P2** outperforms that of **P1** due to the expanded policy space under the generalized problem, **P2**. It is worth noting that, both, **P1** and **P2** are subject to a tight delay constraint on the PU average delay with $D_{max} = 0.057$ sec.

In Fig. 5, the effective capacity of the SU is plotted versus the primary user utilization factor, ρ, for the three systems. The most interesting observation is that all systems exhibit noticeably different performance for low to moderate PU utilization factor. However, under high PU utilization factor The **P2** exhibits superior performance to both the unconstrained system and **P1**. This behavior is attributed to the fact that as the PU utilization increases, the PU is using the channel more frequently and the probability of interference increases and, hence, the SU EC decreases. However, if we optimize over the transmission powers as in **P2**, this helps sustaining high SU EC. Also, we notice that the difference between the unconstrained system and **P1** diminishes as the PU utilization factor increases which is also attributed to higher interference.

5 Conclusions

In this paper we investigate a fundamental trade-off in delay-constrained cognitive radio networks. In particular, we characterize and optimize the trade-off between the secondary user effective capacity and the primary user average delay. Towards this objective, we employ Markov chain models to characterize the secondary user effective capacity and the average packet delay in the primary user queue. First, we formulate two constrained optimization problems to maximize the secondary user effective capacity subject to an average primary user delay constraint. Afterwards, the two formulated problems are proven to be quasi-convex and, hence, can be solved efficiently using standard techniques. Finally, the numerical results reveal interesting insights about the optimal performance compared to the unconstrained PU delay baseline system studied earlier in the literature.

References

1. Mitola, J.: Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio. Ph.D. thesis, Royal Institute of Technology (KTH) (2000)
2. Chaoub, A., Ibn Elhaj, E., El Abbadi, J.: Multimedia traffic transmission over cognitive radio networks using multiple description coding. In: Advances in Computing and Communications, pp. 529–543. Springer, Heidelberg (2011)
3. Kushwaha, H., Xing, Y., Chandramouli, R., Heffes, H.: Reliable multimedia transmission over cognitive radio networks using fountain codes. Proc. of the IEEE Journal **96**(1), 155–165 (2007)
4. Liu, Q., Zhou, S., Giannakis, G.B.: Cross-layer modeling of adaptive wireless links for QoS support in multimedia networks. In: IEEE Quality of Service in Heterogeneous Wired/Wireless Networks Conf., pp. 68–75, October 2004
5. Tang, J., Zhang, X.: Quality-of-service driven power and rate adaptation for multichannel communications over wireless links. IEEE Trans. on Wireless Comm. 4349–4360 (2007)
6. Simeone, O., Bar-Ness, Y., Spagnolini, U.: Stable throughput of cognitive radios with and without relaying capability. IEEE Trans. on Comm. 2351–2360 (2007)
7. Wu, D., Negi, R.: Effective capacity: a wireless link model for support of quality of service. IEEE Trans. on Wireless Comm. **2**(4), 630–643 (2003)
8. Mohammadi, A., Kumar, S., Klymyshyn, D.: Characterization of effective bandwidth as a metric of quality of service for wired and wireless ATM networks. In: IEEE Int. Conf. on Comm. (ICC), pp. 1019–1024, June 1997
9. Akin, S., Gursoy, M.C.: Effective capacity analysis of cognitive radio channels for quality of service provisioning. IEEE Trans. on Wireless Comm. **9**(11), 3354–3364 (2010)
10. Anwar, A.H., Seddik, K.G., ElBatt, T., Zahran, A.H.: Effective capacity of delay constrained cognitive radio links exploiting primary feedback. In: Int. Symposium on Modeling Optimization in Mobile AdHoc Wireless Networks (WiOpt), pp. 412–419 (2013)

11. Chakravarthy, V., Xue, L., Zhiqiang, W.: Novel overlay/underlay cognitive radio waveforms using SD-SMSE framework to enhance spectrum efficiency- part I: theoretical framework and analysis in AWGN channel. IEEE Trans. on Comm. **57**(12), 3794–3804 (2009)
12. Ma, J., Li, Y.: Soft combination and detection for cooperative spectrum sensing in cognitive radio networks. In: IEEE Global Telecom. Conf., (GLOBECOM), pp. 3139–3143 (2007)
13. Chang, C.S.: Performance Guarantees in Communication Networks, chap. 7. Springer (1995)
14. Bertsekas, D.P., Gallager, R.G.: Data Networks, chap. 3. Prentice Hall (1992)
15. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge Unive. Press (2004)