Journal of Communication and Computer 10 (2013) 1007-1018

Channel Estimation and Effective Capacity over Imperfect Rayleigh Channel Environment in Cognitive Radio Networks

Mohamed Elalem and Lian Zhao

Department of Electrical and Computer Engineering, Ryerson University, Toronto M5B 2K3, Canada

Received: February 07, 2013 / Accepted: March 11, 2013 / Published: July 31, 2013.

Abstract: In cognitive radio networks, a SU (secondary user) can share the same frequency band with the PU (primary user) as long as the interference introduced to the latter is below a predefined threshold. In this paper, the transmission performance in cognitive radio networks is studied assuming imperfect channel estimation, taking QoS (quality of service) constraints into consideration. It is assumed that the cognitive transmitter can perform channel estimation and send the data at two different rates and power levels depending on the activity of the PU. The existence of the PU can be detected by channel sensing. A two-state Markov chain process is used to model the existence of the PUs. The cognitive transmission is also configured as a state transition model depending on whether the rates are higher or lower than the instantaneous rate values. The maximum capacity of the SU under the delay constraint is investigated. The concept of effective capacity of the channel is applied. An optimization problem for rate and power allocation under interference and power constraints is formulated and solved. Numerical results are presented to illustrate the average effective capacity optimization and the impact of other system parameters.

Key words: Cognitive radio, effective capacity, PAT, MMSE, spectrum sensing.

1. Introduction

Recent research in spectrum-sharing techniques has enabled different wireless communications technologies to coexist and cooperate towards achieving a better gain from the limited spectrum resources. This started when spectrum utilization measurements showed that most of the allocated spectrum experiences low utilization [1]. Certain administrative authorities, as the FCC (federal communications commission) and the NTIA (national telecommunications and information administration), for radio spectrum regulation, divide the radio spectrum into many frequency bands. They also license the often exclusive usage of these bands that are provided to operators, typically for a long time such as one or two decades. Frequency bands are often idle in many areas, and inefficiently used. The concept of spectrum sharing occurred [2], as one device may transmit while others in the area are idle. Moreover, radio systems can dynamically use and release spectrum wherever and whenever they are available.

There are two different cognitive radio strategies [3]. When the SUs can use the PU's band only if not currently used by the owner PU, the scheme is known as overlay. The existence of the PU can be obtained through spectrum sensing. In the second approach, the SUs are allowed to avail the band even with the PU existence, but should control their interference powers to a tolerable threshold to not harm the PUs. This scheme is known as underlay. This paper adapts the second approach which, obviously, provides more spectrum efficiency [3, 4].

The main challenge for the SU is to control their

Corresponding author: Mohamed Elalem, Ph.D., research field: cognitive radio. E-mail: melalem@ryerson.ca.

interference levels not to exceed the limit where it may introduce harmful impact to the PU. For this reason, interference should be carefully controlled under the assumption of imperfect channel estimation and under the probabilities of getting false alarms and/or misdetections in channel sensing process. The SU should also guarantee its own quality of service requirements by transmitting at certain power for desired rates and by limiting the delay encountered by the transmission in the buffers [5].

Wireless channel conditions vary over time due to changing environment and mobility. The imperfect channel fading coefficients are possible to be estimated through training techniques, which is critical for the successful deployment of cognitive radio systems in practice.

In addition to channel estimation, activities of PUs should be detected through channel sensing. Hence, more challenging scenario may face the developers. There are certain interdependencies between these tasks of channel estimation and sensing. A mistake in channel sensing may lead to errors in the estimation of the channel coefficients. If the PUs are in the network but can not be detected, the channel estimate may be worse. Studying the transmission performance of cognitive radio in a practical scenario in which SUs perform channel sensing, channel estimation, and operate under QoS requirements is the main motivation for recent research.

Some early research in the channel estimation was studied by an analytical approach to the design of pilot-assisted techniques [6]. PAT (pilot assisted transmission), in which a known training symbol is multiplexed with the data symbols, may be used to estimate the channel state and to adapt the receiver parameters accordingly [7, 8].

For practical wireless networks, considering the QoS delay requirement deterministically is unrealistic because of the time-varying feature of wireless channels. Discussing the statistical case for QoS delay becomes more practical. Effective capacity is an appropriate technique in evaluating the capability of a time-varying wireless channel to support data transmissions. The concept of effective capacity has been introduced to support QoS requirement [9]. It can be defined as the maximum constant arrival-rate that can be supported by the time-varying service process where the QoS delay requirement of the system is satisfied [5, 10]. The authors in Ref. [11] studied the effective capacity of cognitive radio networks in the existence of statistical QoS constraints assuming the availability of perfect channel side information at the two cognitive radio sides.

In this paper, we investigate the concept of the effective capacity in cognitive radio channels, and identify the performance limits under imperfect channel estimation and quality of service constraint. The cognitive radio initially performs channel sensing, then the channel fading coefficients are estimated in the training phase of the transmission. Finally, data transmission is performed. The activity of PU is modeled by a Markov process. In this work, we jointly evaluate and optimize the training symbol and data symbol powers and transmission rates of the SUs.

The rest of the paper is organized as follows. In Section 2, the cognitive channel model is given. Channel sensing expressions are formulated. Section 3 discusses the channel training with pilot symbols and derives the MMSE channel estimation technique. In Section 4, data transmission phase and its performance is studied, and a state transition model for cognitive radio transmission is introduced. In Section 5, we formally define the effective capacity in terms of QoS constraints and identify the optimum throughput that the SU can achieve. We provide numerical results in Section 6, and conclude the paper in Section 7.

2. Channel Model and Spectrum Sensing

Fig. 1 depicts the proposed frame model for the cognitive transmission. Initially, the SU performs channel sensing which lasts τ seconds of a frame of total duration T seconds. We assume that pilot

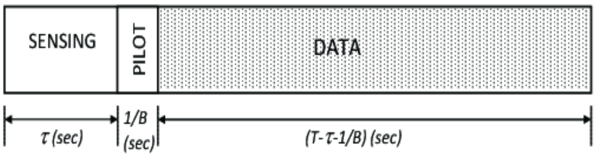

Fig. 1 Transmission frame model consisting of channel sensing, a single symbol as a pilot, and data transmission.

symbols are employed in the system to facilitate the sensing of channel fading coefficients. This will make the receiver able to track the time-varying channel. Since the MMSE estimate only depends on the training energy, not depend on the training duration [12], it can be claimed that transmission of a single pilot at every T seconds is enough and optimal [12, 13]. Instead of increasing number of pilot symbols, a single symbol with relatively high power is used as a pilot. With this, a decrease in the duration of the data transmission can be avoided. Consequently, it is assumed that the transmission is over time-selective flat fading channel in which fading remains constant in each frame.

Both powers of pilot and data symbols, and transmission rates are related to the channel sensing results. Let S_b and r_b be the average transmission power and rate if the PU is detected as busy, respectively, while, they are S_d and r_d, if the channel is detected as idle. The secondary transmitter should terminate the transmission, i.e., $S_b = 0$, when the detection process senses the existence of the PU. The input-output relation between the cognitive transmitter and receiver in the i^{th} symbol duration can be expressed as

$$y_i = \begin{cases} h_i x_{1i} + \mathrm{n}_i, & PU \text{ is idle} \\ h_i x_{2i} + n_i + \zeta_i, & PU \text{ is busy} \end{cases} \quad (1)$$

where, x_{1i} and x_{2i} are the secondary transmitted signal when the channel is idle and busy respectively. y_i denotes the channel output signal, h_i represents the fading coefficient between the cognitive transmitter and receiver, modeled as Rayleigh random distribution with rms value of α. The statistical values of the Rayleigh distribution random variable is Ref. [14]:

$$\begin{cases} E[h] = \alpha\sqrt{\frac{\pi}{2}} \\ \sigma_h^2 = \alpha^2(2 - \frac{\pi}{2}) \end{cases} \quad (2)$$

The term $\{n_i\}$ in Eq. (1) is random noise samples at the cognitive receiver, that are zero-mean Gaussian distributed with variance σ_n^2 for all i. The term ζ_i represents the sum of active PUs' signals received at the cognitive receiver with a variance of σ_ζ^2.

Among different spectrum sensing schemes for reliably identifying the spectrum holes, Energy Detection incurs a very low implementation cost and is hence widely used [15]. It has a good resistance against fast time varying radio environment where none a priori knowledge about the PU is required (non-coherent detector). In order to identify the presence of PU with unknown frequency locations, energy detector serves as the optimal sensing scheme since it only needs to measure the power of the received signal [15, 16].

Spectrum sensing is to decide between the following two hypotheses:

$$\mathcal{H}_0: \quad z_i = n_i \quad i = 1, 2 \cdots \tau B,$$
$$\mathcal{H}_1: \quad z_i = n_i + \zeta_i \quad i = 1, 2 \cdots \tau B.$$

Since the bandwidth is B, we have τB symbols in a duration of τ seconds. By the assumption that $\{\zeta_i\}$ signal samples are $i.i.d.$, the optimal detector response for this hypothesis problem is given in Ref. [17] by

$$\mathcal{Z} = \frac{1}{\tau B}\sum_{i=1}^{\tau B} |z_i|^2 \gtrless_{\mathcal{H}_1}^{\mathcal{H}_0} \delta \quad (3)$$

where, δ is a pre-designed threshold. The cognitive radio assumes that the primary system is in operation if $\mathcal{Z} \geq \delta$ i.e., $\mathcal{H}_1$. Otherwise, it assumes $\mathcal{H}_0$. Assuming τB is sufficiently high, $\mathcal{Z}$ can be approximated, using Central Limit Theorem, as a Gaussian random variable with mean and variance

$$\mathbb{E}[\mathcal{Z}] = \begin{cases} \sigma_n^2 & PU \text{ is idle} \\ \sigma_\zeta^2 + \sigma_n^2 & PU \text{ is busy} \end{cases} \quad (4)$$

and

$$\sigma_{\mathcal{Z}}^2 = \begin{cases} \sigma_n^4/(\tau B) & \text{idle} \\ \frac{(2(\sigma_\zeta^4+\sigma_n^4)-(\sigma_\zeta^2-\sigma_n^2)^2)}{\tau B} & \text{busy}, \end{cases} \quad (5)$$

respectively. The probabilities of detection, false alarm and missing of energy detector (miss detection occurs when the PU is in operation but the cognitive radio fails to sense it) are given as follows [18]

$$P_d = Pr\{\mathcal{Z} > \delta|\mathcal{H}_1\}$$
$$= Q\left(\frac{\delta-\sigma_\zeta^2-\sigma_n^2}{\sqrt{\frac{2\left((\sigma_\zeta^4+\sigma_n^4)-(\sigma_\zeta^2-\sigma_n^2)^2\right)}{\tau B}}}\right), \quad (6)$$

$$P_f = Pr\{\mathcal{Z} > \delta|\mathcal{H}_0\} = Q\left(\frac{\delta-\sigma_n^2}{\sqrt{\sigma_n^4/\tau B}}\right), \quad (7)$$

$$P_m = Pr\{\mathcal{H}_0|\mathcal{H}_1\} = 1 - P_d, \quad (8)$$

where, $Q(\cdot)$ represents the complementary distribution function of the standard Gaussian distribution [19].

Regarding the channel sensing result, the cognitive radio network has the following four cases:

Correct detection: with two possible cases;

Channel is busy, detected as busy, (BB);

Channel is idle, detected as idle, (DD);

Miss detection: channel is busy, detected as idle (BD);

False alarm: Channel is idle, detected as busy (DB).

3. Pilot Power Analysis

In PAT (pilot aided (or assisted) transmission), a known symbol is embedded in the data transmitted stream to facilitate the receiver to estimate the channel fading coefficients [20]. The cognitive transmitter sends one pilot symbol after the processing of channel sensing to make the receiver able to estimate the channel coefficients. Obviously, this estimation will be affected by channel sensing results.

As mentioned in Section 2, with the assumption of constant fading within a frame, one pilot symbol is adequate to provide estimations. The first τ seconds of a frame with duration T is reserved for sensing process, while sending a single pilot with relatively high power is optimal [13]. This increases the duration of the data transmission. After channel sensing and pilot symbol transmission phases, the rest $(T-\tau)B-1$ symbols are devoted for data transmitting. The average input power in each frame can be written as

$$S_l = \sum_{i=\tau B+1}^{TB} \mathbb{E}[\ |x_{li}|^2];\ i = 0, 1, \cdots, \quad l = 1, 2 \quad (9)$$

where, x_{li} is defined in Eq. (1). Thereby, the total power assigned to the pilot and data symbols in a frame is limited by $\mathrm{S_b}$ when the channel is sensed as busy, or by S_d when the channel is sensed as idle. For the possible two cases mentioned above in which the channel is busy (i.e., BB and BD), the cognitive transmitter transmits with an average power $\mathrm{S_b}$ for the case of BB. While for the case of BD, the cognitive transmitter transmits with an average power S_d, making the active PU suffer from interference introduced by the SU. It is assumed that depending on the capabilities of the transmitters and the energy resources they are equipped with, there exists peak constraints on both average powers, say: S^m.

Additionally, in order to mitigate the average interference and protect the PU, the following constraint on S_b and S_d must be imposed:

$$P_d S_b + P_m S_d \le S^m \quad (10)$$

where, $\mathrm{P_d}$ and $\mathrm{P_m} = (1-\mathrm{P_d})$ are the detection and miss-detection probability defined in Eqs. (6) and (8), respectively. Now, the average interference experienced by the PU can be expressed as

$$\mathbb{E}\{\mathrm{P_d S_b}|\mathrm{h_{cp}}|^2 + \mathrm{P_m S_d}|\mathrm{h_{cp}}|^2\}$$
$$= (P_d S_b + P_m S_d)\mathbb{E}\{|\mathrm{h_{cp}}|^2\} \le I^m \quad (11)$$

where, h_{cp} denotes the fading coefficient between the cognitive transmitter and primary receiver, and I^m is the average interference constraint. Note that h_{cp} is not known at the cognitive transmitter and hence the cognitive transmitter can not adapt its transmission according to it. However, if the statistics of this coefficient ($\mathbb{E}\{|\mathrm{h_{cp}}|^2$) is known, then in order to satisfy Eq. (11), the cognitive transmitter can choose $S^m = \frac{I^m}{\mathbb{E}\{|h_{cp}|^2\}}$.

The pilot symbol power is also related to the sensing result. Let the power of pilot symbol be $S_{pb} = \mu_b S_b$ if the PU is detected, while, it is $S_{pd} = \mu_d S_d$ when PU is not detected, where μ_b and μ_d are fractions of the total power assigned to the pilot symbol and data when channel is detected as busy and idle, respectively. Since we assume that the fading coefficients $\{h_i\}$ been constant within each frame, the index i will be omitted.

The received signal in the pilot phase in a certain frame (i.e., y_p), can be written as

$$y_p = \begin{cases} h(S_{pb})^{1/2} + n + \zeta & BB\ \ case \\ h\ (S_{pd})^{1/2} + n & DD\ \ case \\ h\ (S_{pd})^{1/2} + n + \zeta & BD\ \ case \\ h\ (S_{pb})^{1/2} + n & DB\ \ case\ . \end{cases} \quad (12)$$

If the receiver employs MMSE (minimum mean-square error) estimator to obtain the estimated fading coefficients, then the estimates can be found using MMSE estimation [13, 21] as

$$\hat{h} = \begin{cases} \frac{\sqrt{S_{pb}}\sigma_h^2}{S_{pb}\sigma_h^2+\sigma_n^2+\sigma_\zeta^2} y_p & BB\ \&\ DB \\ \frac{\sqrt{S_{pd}}\sigma_h^2}{S_{pd}\sigma_h^2+\sigma_n^2} y_p & BD\ \&\ DD\ . \end{cases} \quad (13)$$

It is essentially to know that the MMSE estimates given above are related to the channel sensing results. $\hat{h}$ in Eq. (13) is the estimated channel fading, which is a circularly symmetric, complex, Gaussian random variable with zero mean and variance $\sigma_{\hat{h}}^2$ (i.e., $\hat{h}\sim\mathcal{CN}(0,\ \sigma_{\hat{h}}^2)$. It can be expressed as $\hat{h} = \sigma_{\hat{h}} w$, where w is a standard complex Gaussian random variable, (i.e., $w\sim\mathcal{CN}(0, 1)$. Thus the fading coefficient can now be expressed as follows [22]:

$$\hat{h} = h + \varepsilon, \quad (14)$$

where ε is the estimate error in the fading coefficient h, and $\varepsilon\sim\mathcal{CN}(0,\ \sigma_\varepsilon^2)$ [6, 22, 23].

Now, the input-output relationship for data phase in Eq. (1) can be rewritten as:

$$\hat{y} = \begin{cases} \hat{h}x_1 + n + \zeta & \text{channel is busy}, \\ \hat{h}x_2 + n & \text{channel is idle}, \end{cases} \quad (15)$$

The estimation of the channel variance is Ref. [20]

$$\sigma_{\hat{h}}^2 = \begin{cases} \frac{S_{pb}\sigma_h^4}{S_{pb}\sigma_h^2+\sigma_n^2+\sigma_\zeta^2} & BB \\ \frac{S_{pd}\sigma_h^4}{S_{pd}\sigma_h^2+\sigma_n^2} & DD \\ \frac{S_{pd}\sigma_h^4}{(S_{pd}\sigma_h^2+\sigma_n^2)^2}(S_{pd}\sigma_h^2+\sigma_n^2+\sigma_\zeta^2) & BD \\ \frac{S_{pb}\sigma_h^4}{(S_{pb}\sigma_h^2+\sigma_n^2+\sigma_\zeta^2)^2}(S_{pb}\sigma_h^2+\sigma_n^2) & DB\ , \end{cases} \quad (16)$$

Using Eqs. 2 and 16 , the variance of the estimation error σ_ε^2 can be written as

$$\sigma_\varepsilon^2 = \sigma_{\hat{h}}^2 + \alpha^2(1+\frac{\pi}{2}), \quad (17)$$

by assuming that there is no correlation between the error and its estimation. We have omitted the subscript i in above equation because of the block fading assumption, and because of the assumed identicalness property of the fading coefficient and its estimates random variables in each frame.

4. Data Transmission Phase

Finding the capacity of the channel in Eq. (15) is not an easy task. A lower bound capacity is generally obtained by considering the estimate error ε as another source of Gaussian noise, i.e., by considering the term $(\varepsilon\, x_l + n)$; $l = 1, 2$, in Eq. (15) as Gaussian distributed noise uncorrelated with the input [5].

The channel can be modeled as a two Morkov chain states (i.e., ON and OFF), for the state when target transmission rate is greater than or less than the instantaneous rate that the channel can support, respectively. These two states are possible in each of the four cases discussed above. Hence, totally there are eight states (2 × 4).

Considering the channel estimation results and interference generated by the PU ζ, we have the following lower bounds instantaneous channel capacities in a frame for the four scenarios described above:

$$C_k^l = C_o\log_2(1+\eta_k|w|^2), \qquad k = 1, 2, 3, 4 \quad (18)$$

where $C_o = \frac{(T-\tau)B-1}{T}$, and

$$\eta_k = \begin{cases} \frac{S_{db}\sigma_{\hat{h}}^2}{S_{db}\sigma_\varepsilon^2+\sigma_n^2+\sigma_\zeta^2} & k = 1\ ;\ \text{BB} \\ \frac{S_{dd}\sigma_{\hat{h}}^2}{S_{dd}\sigma_\varepsilon^2+\sigma_n^2} & k = 2\ ;\ \text{DD} \\ \frac{S_{dd}\sigma_{\hat{h}}^2}{S_{dd}\sigma_\varepsilon^2+\sigma_n^2+\sigma_\zeta^2} & k = 3\ ;\ \text{BD} \\ \frac{S_{db}\sigma_{\hat{h}}^2}{S_{db}\sigma_\varepsilon^2+\sigma_n^2} & k = 4\ ;\ \text{DB} \end{cases} \quad (19)$$

C_k^l is the frame's lower band capacity of each scenario k, which are obtained by assuming the factors $\{\varepsilon \cdot x\}$ and $\{\zeta\}$ in Eq. (15) as worst case noises whereas these noises are considered as Gaussian distributed [10]. S_{db} is the data symbol power when the channel is detected as busy, while S_{dd} is the data symbol power when the channel is detected as idle. These two powers are related to the cognitive average powers, as $S_{db} = S_b(1-\mu_b)/TC_o = S_b(1-\mu_b)/(T-\tau)B-1)$, and $S_{dd} = S_d(1-\mu_d)/(T-\tau)B-1)$.

Since $w \sim \mathcal{CN}(0,1)$, the magnitude $|w|$ will have the *Rayleigh* distribution and the squared magnitude $|w|^2$ will have *Exponential* distribution with unity mean. The transmitter will transmit the information at the desired rate regardless of the channel conditions. We assume that the transmitter will send its data at fixed rate r_b if the channel is sensed as busy, and at r_d if it is sensed as idle. If these rates are below the instantaneous capacity values, i.e., when $r_b < C_1^l, C_4^l$ or $r_d < C_2^l, C_3^l$, the transmission can be assessed to be in the ON state and, so, the target rates can be achieved. While , if $r_b \geq C_1^l, C_4^l$ or $r_d \geq C_2^l, C_3^l$, the channel is in the OFF state, where reliable communication can not be achieved.

The activity of the PU between the frames can be also modeled as a two-state Markov model. Busy state indicates that the PU occupied the channel, and iDle state indicates the absent of the PU in the channel, as can be seen in Fig. 2. Switching from busy state to idle state and from idle state to busy state is with probability b and d, respectively. The state transition is assumed to occur every T s.

Taking into account the four possible cases related to the channel sensing results jointly with the reliability of the transmissions states, the cognitive radio transmission can be represented by state transition model as P(8× 8) transition matrix denoted as

$$\boldsymbol{P} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 \\ p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 \\ p'_1 & p'_2 & p'_3 & p'_4 & p'_5 & p'_6 & p'_7 & p'_8 \\ p'_1 & p'_2 & p'_3 & p'_4 & p'_5 & p'_6 & p'_7 & p'_8 \\ p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 \\ p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 \\ p'_1 & p'_2 & p'_3 & p'_4 & p'_5 & p'_6 & p'_7 & p'_8 \\ p'_1 & p'_2 & p'_3 & p'_4 & p'_5 & p'_6 & p'_7 & p'_8 \end{bmatrix} \quad (20)$$

The transition probabilities depend on channel coefficients, sensing probabilities, transmission rates, and the two state Markov model in Fig. 2. Table 1 summarizes the entries of the matrix P, where λ_k in the table is defined as:

$$\lambda_k = \begin{cases} \frac{2^{(r_b/C_O)}-1}{\eta_i}, & k=1,4; \\ \frac{2^{(r_d/C_O)}-1}{\eta_i}, & k=2,3. \end{cases} \quad (21)$$

The details are provided in Appendix A. According to the entries of the matrix P displayed in Table 1, the rank of this matrix is 2. All the transition probabilities are functions of λ_k which is specified in Eq. (21).

5. Effective Capacity Optimization CU

5.1 Preliminary on Effective Capacity

The E_C (Effective Capacity) (or Effective Bandwidth) theory is a powerful approach to evaluate the capability of a wireless channel to support data transmissions with diverse statistical QoS (quality of service) guarantees [9, 10, 24, 25]. It is defined as the maximum constant arrival rate that the channel can support while meeting the QoS requirement [10].

In particular, the statistical QoS guarantee can be characterized by a metric called QoS exponent

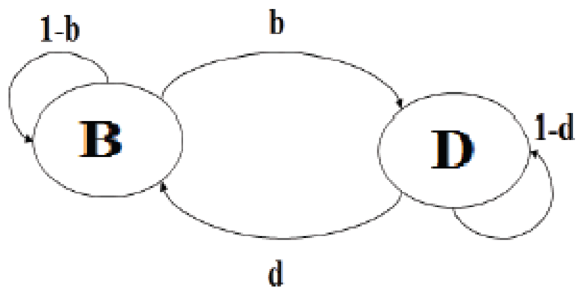

Fig. 2 PU activity between two states: busy and idle.

Table 1 The transition probabilities of Matrix *P*.

$l = 1,2,5,6$		$n = 3,4,7,8$	
$p_{l1} = (1-b)P_d e^{-\lambda_1}$	$=p_1$	$p_{n1} = dP_d e^{-\lambda_1}$	$=p'_1$
$p_{l2} = (1-b)P_d(1-e^{-\lambda_1})$	$=p_2$	$p_{n2} = dP_d(1-e^{-\lambda_1})$	$=p'_2$
$p_{l3} = b(1-P_f)e^{-\lambda_2}$	$=p_3$	$p_{n3} = (1-d)(1-P_f)e^{-\lambda_2}$	$=p'_3$
$p_{l4} - b(1-P_f)(1-e^{-\lambda_2})$	$-p_4$	$p_{n4} - (1-d)(1-P_f)(1-e^{-\lambda_2})$	$-p'_4$
$p_{l5} = (1-b)P_m e^{-\lambda_3}$	$=p_5$	$p_{n5} = dP_m e^{-\lambda_3}$	$=p'_5$
$p_{l6} = (1-b)P_m(1-e^{-\lambda_3})$	$=p_6$	$p_{n6} = dP_m(1-e^{-\lambda_3})$	$=p'_6$
$p_{l7} = bP_f e^{-\lambda_4}$	$=p_7$	$p_{n7} = (1-d)P_f e^{-\lambda_4}$	$=p'_7$
$p_{l8} = bP_f(1-e^{-\lambda_4})$	$=p_8$	$p_{n8} = (1-d)P_f(1-e^{-\lambda_4})$	$=p'_8$

denoted by θ $(0 < \theta < \infty)$ [9]. The QoS exponentθ characterizes the exponentially decaying rate of the violation probability against the queue-length threshold [10]. With the pair (Effective Capacity E_C and QoS exponent θ), it can be observed that there is tradeoff between the QoS requirement and the system rate. Higher θ represents more stringent delay QoS requirements, and vice versa.

The delay, which is a QoS measure, can be described through the probability that the occupancy of the buffer is higher than a specific value, say x, so the QoS exponent can be formulated as [10]

$$\theta = -\lim_{x\to\infty} \frac{\log Pr\{L>\mathrm{x}\}}{x}, \qquad (22)$$

where L is the cognitive queue length and follows the equilibrium queue-length distribution of the buffer at the source [9, 25]. When $\theta \to 0$, the user does not impose any delay constraints on the service process. On the other hand, $\theta \to \infty$ implies that any delay is not tolerable, and thus the effective capacity reduces to the minimum supportable service rate.

From Eq. (22), for large x, (x^m), the buffer violation probability can be approximated as

$$\Pr\{L > x^m\} \approx \exp(-x^m\theta). \qquad (23)$$

Smaller θ corresponds to looser constraints, and larger θ implies more strict QoS constraints.

5.2 Framework for Effective Capacity Optimization

In this subsection, we aim to analyze the maximum capacity that the cognitive radio channel can achieve. The effective capacity for a given θ is defined in Refs [9, 26] as

$$E_c = -\lim_{t\to\infty} \frac{1}{\theta t} \log\mathbb{E}(\exp(-\theta R(t)), \qquad (24)$$

where $R(t) = \sum_{i=1}^{t} r(i)$ is the time-accumulated service process. Here we assume that $r(i)$ is discrete time stationary and ergodic stochastic service process. $\mathbb{E}(\cdot)$ is the expectation with respect to the random variable r.

It can be noticed that the service rate is $r(i) = r_b T$ if the SU is in the state ON_1 or ON_7 at time i (the subscript number points to the state number). Similarly, the service rate is $r(i) = r_d T$ in the states ON_3 and ON_5. For the remaining states $(OFF_j, j = 2,4,6,8)$, the target transmission rate is greater than the instantaneous channel capacities and, so, communication can not be achieved. This leads to vanish all the service rates in these four even states.

Eq. (24) can be solved using the technique given in Ref. [24] as follows

$$E_c = \frac{1}{\theta} \log\left(\rho(M)\right) = \frac{1}{\theta} \log\rho(D \cdot P), \qquad (25)$$

where, $\rho(M)$ function is the spectral radius of the matrix M, $D = \mathrm{diag}(d_1(\theta), \cdots, d_N(\theta)$ is a diagonal matrix with elements equal to the moment generating functions of the processes in the N states [24] (here, we have 8 states). Spectral radius of a matrix is the maximum of the absolute values of its eigenvalues, i.e., $\rho(A) \overset{\mathrm{def}}{=} \max_i(|\omega_i|)$, ω_i'saretheeigenvalues of the matrix A [27]. P in Eq. (25) is the transition matrix given in Eq. (20). Note that in our assumptions, the transmission rates are deterministic and constants in each state, thus, the possible rates are: Tr_b, Tr_d, and 0 for which the moment generating functions $e^{T\theta r_b}$, $e^{T\theta r_d}$ and 1, respectively. Therefore, $\mathrm{D} = \mathrm{diag}(e^{T\theta r_b},\ 1,\ e^{T\theta r_d},\ 1,\ e^{T\theta r_d},\ 1,\ e^{T\theta r_b},\ 1)$. So the matrix M can be filled as

$$M = (D \cdot P) =$$

$$\begin{bmatrix} \tau_b p_1 & \tau_b p_2 & \tau_b p_3 & \tau_b p_4 & \tau_b p_5 & \tau_b p_6 & \tau_b p_7 & \tau_b p_8 \\ p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 \\ \tau_d p'_1 & \tau_d p'_2 & \tau_d p'_3 & \tau_d p'_4 & \tau_d p'_5 & \tau_d p'_6 & \tau_d p'_7 & \tau_d p'_8 \\ p'_1 & p'_2 & p'_3 & p'_4 & p'_5 & p'_6 & p'_7 & p'_8 \\ \tau_d p_1 & \tau_d p_2 & \tau_d p_3 & \tau_d p_4 & \tau_d p_5 & \tau_d p_6 & \tau_d p_7 & \tau_d p_8 \\ p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 \\ \tau_b p'_1 & \tau_b p'_2 & \tau_b p'_3 & \tau_b p'_4 & \tau_b p'_5 & \tau_b p'_6 & \tau_b p'_7 & \tau_b p'_8 \\ p'_1 & p'_2 & p'_3 & p'_4 & p'_5 & p'_6 & p'_7 & p'_8 \end{bmatrix} \qquad (26)$$

where $\tau_b = e^{T\theta r_b}$, and $\tau_d = e^{T\theta r_d}$ are the moment generating functions of the possible rates when the channel is busy and idle respectively.

It is easy to note that the matrix M has also a rank of 2. The characteristic polynomial of the matrix can be written as(see Appendix B for details):

$$Q(\omega) = \omega^2 - C_7\omega + C_6, \tag{27}$$

where the nonzero-eigenvalues ω can be found by solving the quadratic Eq. (27).

The effective capacity in Eq. (24) can be optimized by choosing the maximum values of r_b and r_d over the optimized power allocation constraints. This maximization is firstly done by choosing the maximum value of the eigenvalue of the matrix $(\mathrm{D} \cdot \mathrm{P})$ which maximizes the function $\rho(M)$ in Eq. (25). Another optimization should be done over the entire variables which will lead to the optimal effective capacity formula.

The effective capacity expression in Eq. (28) is obtained by choosing the largest value of the eigenvalues of the matrix M for a given sensing duration τ, detection threshold δ, and QoS exponent θ. One can note that if the sensing results are perfect with no errors, i.e., the detection probability $P_d = 1$, and so ($P_m = P_f = 0$), the transition probabilities in matrix P, $p_5 = p_6 = p_7 = p_8 = p'_5 = p'_6 = p'_7 = p'_8 = 0$ in the effective capacity expression in Eq. (28).

An analytical optimized solution for the problem Eq. (28) is possible whenever the generating function has an analytical expression [28. 29]. In the following section, we investigate the impact of several parameters on the effective capacity through numerical results.

6. Numerical Results

In this section, numerical results are presented to illustrate the impact of the sensing duration τ, detection threshold δ, and other factors on the effective capacity. Without loss of generality, we set all variances to unity ($\sigma_h = \sigma_n = \sigma_\zeta = 1$). We also assume the symbol rate $\mathrm{B} = 10{,}000$ symbol/sec, and the frame duration $T = 0.25$s. This means there are 2,500 symbols in the frame. Unless they are not variable, time allocated for sensing is set to 5 ms, and QoS exponent θ is assumed to be 10%. The maximum average power constraint $S_m = 20$ dB. The fraction assigned to the pilot symbol is 10% whether the channel is busy or idle (i.e., $\mu_b = \mu_d = 0.1$). To simplify the objective function of the effective capacity, we set the transition probabilities of the two-state Markov model in Fig. 2: b and d, such that $b = 1 - d$. It is further assumed that in each frame, PU activity does not change, while it may change frames. independently from one state to another across the frame.

$$\begin{aligned}
E_c^{opt} \;=\;& \max -\frac{1}{\theta TB}\log\{\frac{1}{2}(\tau_b(p_1+p'_7)+\tau_d(p_5+p'_3)+p_2+p_6+p'_4+p'_8)\\
+\;& \frac{1}{2}(\tau_b(p_1+p'_7)+\tau_d(p_5+p'_3)+p_2+p_6+p'_4+p'_8)^2\\
-\;& 4(\tau_b^2(p'_7\,p_1-p'_1\,p_7)+\tau_d^2(p_5\,p'_3-p_3\,p'_5)+\tau_b(p'_7\,p_2+p'_7\,p_6-p'_1\,p_4-p'_2\,p_7-p'_5\,p_7-p'_6\,p_7\\
-\;& p'_1\,p_8+p'_8\,p_1)+\tau_d(p'_3\,p_2-p'_2\,p_3-p_4\,p'_5+p_5\,p'_4-p_3\,p'_6+p_6\,p'_3-p'_5\,p_8+p'_8\,p_5)\\
+\;& \tau_b\tau_d\big(p'_3\,p_1+p'_7\,p_5-p'_1\,p_3\big)\\
-\;& p'_2\,p_4+p'_4\,p_2+p'_4\,p_1-p_4\,p'_6+p_6\,p'_4-p'_2\,p_8-p'_6\,p_8+p'_8\,p_6+p'_8\,p_2))^{\frac{1}{2}}\}
\end{aligned} \tag{28}$$

$$\begin{aligned}
S.t.\quad & 0 < S_b, S_d \le S^m\\
& 0 < \mu_b, \mu_d \le 1\\
& r_b, r_d \ge 0\\
& P_d S_b + P_m S_d \le S^m\\
& (P_d S_b + P_m S_d)\mathbb{E}(|h_{cp}|^2) \le I^m
\end{aligned}$$

In Fig. 3, the normalized effective capacity is plotted versus the delay QoS exponent (θ) for various interference-limit values. We observe that the capacity increases as θ decreases. However, the gain in the effective capacity decreases for higher values of θ. The figure shows that in the case with loose QoS restrictions (i.e., lower values of θ), the capacity benefits significantly, whereas in the case with higher values of θ (i.e., $\theta = 10(1/\text{bit})$, about 70% reduction in the capacity is seen.

Fig. 4 illustrates the effective capacity of the SU versus the interference limit that the PU can tolerate, I^m for various QoS exponent values. This figure reveals that the capacity gain that can be achieved under strict peak interference constraint is much lower than the one under released interference constraint. Also, similar to the conclusion in Fig. 3, for specific I^m value, the capacity increases as θ becomes lower which means loose QoS restrictions.

Fig. 5 studies the effect of the channel sensing duration (τ). It can be seen that for a short time reserved for sensing process, the SU is more likely to get a false alarm detecting the PU, whereas the detection probability approaches to one for long sensing duration.

In Fig. 6, we display the effective capacity (normalized value) as a function of the detection probability for different values of S_m. As expected, with increasing S_m, the effective capacity value increases. It can be seen from this figure that the maximum effective capacity points are achieved when P_d is close to 0.8. As P_d further increases and approaches 1, the SUs start to consider the channel as busy all the time and hence the performance degradation occurs. This is because of not being able to take advantage of idle channel states. The impact of the average power on the capacity is also clear in this figure: more average power means more relax constraints, which leads to more capacity.

In Fig. 7, we plot the normalized effective capacity as a function of the ratio of the power allocated to the pilot symbol to the total power when the channel is sensed as busy μ_b (here, we assume μ_b=μ_d). We consider two different interference thresholds.

The optimal ratio occurs at $\mu \simeq 10\%$. As the interference constraint increases which means more allowance to the SU to transmit, the effective capacity increases. This can be intuitively understood. More μ_b (or μ_d) means more power allocated to the pilot symbol and less power allocated to data transmission, which in turns reduces the capacity.

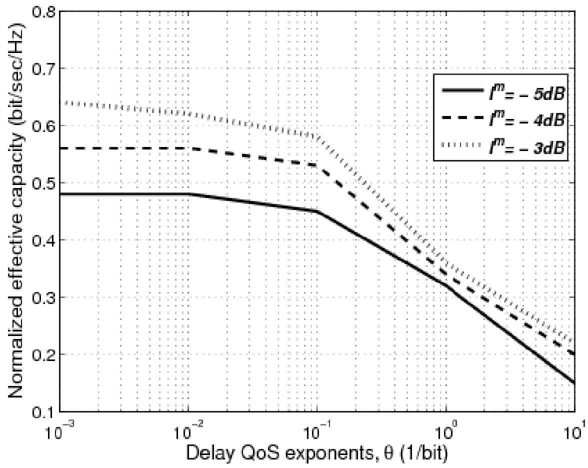

Fig. 3 Normalized effective capacity versus delay QoS exponents, for various interference-limit.

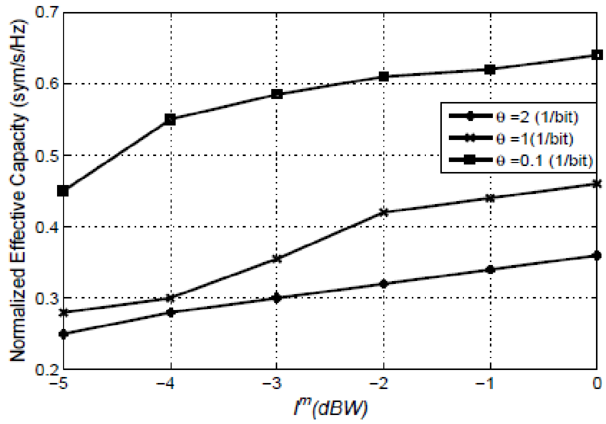

Fig. 4 Normalized effective capacity versus interference limit for various QoS exponent values, θ.

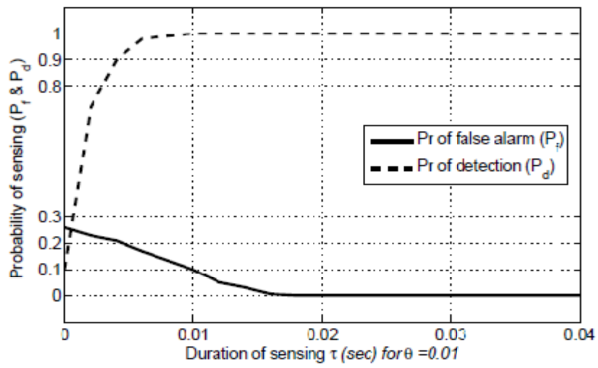

Fig. 5 Probabilities of sensing (P_d&P_f) versus channel sensing duration τ.

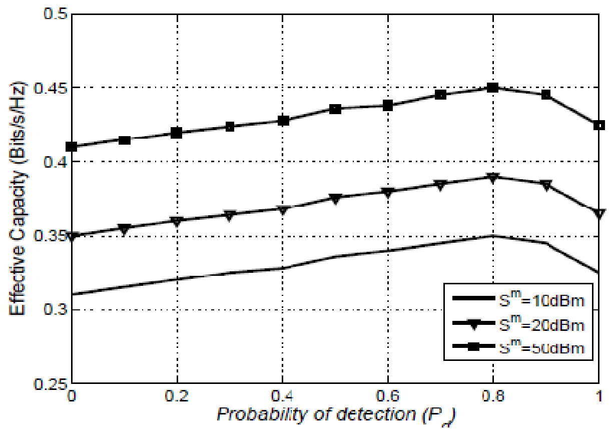

Fig. 6 Effective capacity versus probability of detection P_d for different values of S_m.

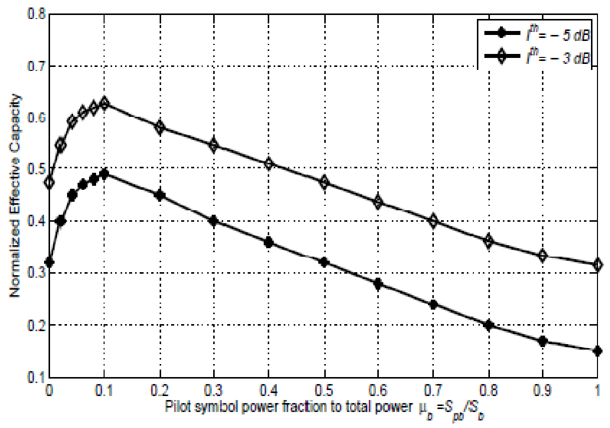

Fig. 7 Normalized effective capacity versus the ratio of power pilot symbol to the total power allocated for different values of I^{th}.

7. Conclusions

In this paper, the effective capacity of cognitive radio channels has been analyzed taking into account QoS constraints, imperfect channel information, and transmission power limitations. First, a system model is introduced in which the cognitive transmitter initially senses the channel in order to detect the activity of the PU. It then sends a pilot symbol for channel estimation followed by data transmission. An energy detector is adopted to perform channel sensing, which incurs a very low implementation cost and is widely used. The estimation of the channel fading coefficients is performed through pilot transmission in the training phase. The MMSE (minimum mean square error) estimator is assumed to be employed at the receiver. Through the study, the interrelation between channel sensing and estimation has been investigated. We have observed that degradation in the channel estimation is a result of faulty sensing. The cognitive transmitter is assumed to transmit data at fixed powers and rates according to the channel sensing results. For the SU, we have constructed a state-transition model taking into account the reliability of the transmission, channel sensing results, and the PU activity in the channel. We have formulated the transition probabilities for this model. A closed form for the effective capacity is obtained as a function of exponent delay constraint. Numerical results are provided to examine the impact of delay constraint, interference limit, channel sensing duration, threshold, and sensing probabilities on the effective capacity. Many insightful observations and investigations are presented.

Appendix A

$p_1 =$ Pr{the channel is being busy and it is detected as busy and $r_b < C_1^l(k)$ in the k^{th} frame given that the channel is being busy and it is detected as busy and $r_b < C_1^l(k-1)$ in the $(k-1)^{th}$ frame}

According to the chain rule in probability theorem, if there are four events: $\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3$ and A_4, then

$$\Pr(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3|\mathrm{A}_4) = \Pr(\mathrm{A}_1 \cap \mathrm{A}_2 \cap \mathrm{A}_3|\mathrm{A}_4)$$
$$= \Pr(\mathrm{A}_1|\mathrm{A}_4) \times \Pr(\mathrm{A}_2|\mathrm{A}_1 \cap \mathrm{A}_4) \times \Pr(\mathrm{A}_3|\mathrm{A}_1 \cap \mathrm{A}_2 \cap \mathrm{A}_4) \quad \text{A.1}$$

So

$p_{11} =$ Pr{channel is busy in i^{th} frame|channel is busy in $(\mathrm{i}-1)^{\mathrm{th}}$} × Pr{channel is busy in i^{th} frame | channel is busy in $(\mathrm{i})^{\mathrm{th}}$} × $\Pr\{\mathrm{r_b} < \mathrm{C}_1^{\mathrm{l}}(\mathrm{i})|\mathrm{r_b} < \mathrm{C}_1^{\mathrm{l}}(\mathrm{i}-1)\}$

$$p_{11} = (1-\mathrm{b})\mathrm{P_d}\Pr\{\mathrm{r_b} < \mathrm{C}_1^{\mathrm{l}}(\mathrm{i})|\mathrm{r_b} < \mathrm{C}_{\mathrm{i}-1}^{\mathrm{l}}(\mathrm{i}-1)\}$$
$$p_{11} = (1-\mathrm{b})\mathrm{P_d}\Pr\{\mathrm{z_i} > \lambda_1 |\mathrm{z_{i-1}} > \lambda_1\} \quad \text{A.2}$$
$$p_{11} = (1-\mathrm{b})\mathrm{P_d}\Pr\{\mathrm{z_i} > \lambda_1\} = (1-\mathrm{b})\mathrm{P_d}\Pr\{\mathrm{z} > \lambda_1\} = p_1$$

We omitted the index i in z_i due to the fact that z_i and z_{i-1} are independent due to the block fading assumption.

By the same manner, the transition probabilities from any state to state 1 can be expressed as

$$p_{l1} = p_{11} = p_{21} = p_{31} = p_{41}$$
$$= (1-b)P_d Pr\{z > \lambda_1\}$$
$$= (1-b)P_d e^{-\lambda_1} = p_1 \qquad \text{A.3}$$
$$p_{n1} = p_{51} = p_{61} = p_{71} = p_{81}$$
$$= dP_d \Pr\{z > \lambda_1\} = dP_d e^{-\lambda_1} = p'_1$$

Using the same modality, the full transition probabilities can be obtained, and they are listed in Table 1.

Appendix B

Let A be an $\mathrm{n} \times \mathrm{n}$ matrix, the eigenvalues of the matrix A are the zeroes of its characteristic polynomial, $\det(\omega\mathrm{I} - \mathrm{A})$, which can be written as

$$Q(\omega) = \omega^n - C_{n-1}\omega^{n-1} + C_{n-2}\omega^{n-2} - \cdots (-1)^n C_0 \quad \text{B.1}$$

It is well known that the coefficients C_{n-1} and C_0 are, respectively, the $\mathrm{trace(A)}$(the sum of its diagonal entries) and the $\det(\mathrm{A})$. All other coefficients C_{n-k}, $k = 1,2,\cdots,$ can be expressed by the sum of the k −rowed principle minors of A. A k-rowed principal minor of an $\mathrm{n} \times \mathrm{n}$ matrix A is the determinant of a $\mathrm{k} \times \mathrm{k}$ submatrix of A whose entries, a_{ij}, have indices i and j that are the elements of the same k −element subset of $1, 2, \cdots, \mathrm{n}$.

With rank (the dimension of the largest square submatrix of A with nonzero determinant) r, where, $\mathrm{r} < n$. All nonzero eigenvalues of A are among the zeros of the polynomial [18]

$$Q(\omega) = \omega^r - C_{n-1}\omega^{r-1} + \cdots (-1)^r C_{n-r} \qquad \text{B.2}$$

For r = 2, we can write the characteristic equation of the system as in Eq. (27).

References

[1] C. Cordeiro, K. Challapali, D. Birru, N. Shankar, IEEE 802.22: The first worldwide wireless standard based on cognitive radio, in: First IEEE Int. Symp. New Frontiers in Dynamic Spectrum Access Networks, MD, Nov. 8-10, 2005, pp. 328-337.

[2] M. Marcus, Unlicensed cognitive sharing of TV spectrum: The controversy at the federal communications commission, IEEE Communication Magazine 43 (2005) 24-25.

[3] M. Khoshkholgh, K. Navaie, H. Yanikomeroglu, Access strategies for spectrum sharing fading environment: Overlay, underlay and mixed, IEEE Trans. Mobile Comput. 9 (2010) 1780-1793.

[4] F. Hou, J. Huang, Dynamic channel selection in cognitive radio network with channel heterogeneity, in: Proceedings of IEEE GLOBECOM, Hawaii, Nov.30-Dec.4, 2009.

[5] D. Qiao, M. Gursoy, S. Velipasalar, Energy efficiency of fixed-rate wireless transmissions under queueing constraints and channel uncertainty, in: Proceedings of IEEE GLOBECOM, Hawaii, Nov. 30-Dec. 4, 2009, pp. 24-29.

[6] J. Cavers, An analysis of pilot symbol assisted modulation for rayleigh fading channels, IEEE Trans. Vehicular Tech. 40 (1991) 686-693.

[7] B. Sadler, L. Tong, M. Dong, Pilot-assisted wireless transmissions, IEEE Signal Processing Magazine 21 (2004) 12-25.

[8] S. Akin, M. Gursoy, Ergodic capacity analysis in cognitive radio systems under channel uncertainty, in: 44th Annual Conference on Information Sciences and Systems (CISS) 2010, NJ, March 17-19, 2010.

[9] D. Wu, R. Negi, Effective capacity: A wireless link model for support of quality of service, IEEE Trans. Wireless Communication 2 (2003) 630-643.

[10] Q. Du, X. Zhang, Effective capacity optimization with layered transmission for multicast in wireless networks, Wireless Communications and Mobile Computing Conference IWCMC, Greece, August 6-8, 2008, pp. 267-272.

[11] S. Akin, M. Gursoy, Effective capacity analysis of cognitive radio channels for quality of service provisioning, IEEE Trans. Wireless Commun. 9 (2010) 3354-3364.

[12] B. Hassibi, B. Hochwald, How much training is needed in multiple-antenna wireless links, IEEE Trans. Inform. Theory 49 (2003) 951-963.

[13] M. Gursoy, An energy efficiency perspective on training for fading channels, in: Proceedings of IEEE International Symposium on Inf. Theory (ISIT) Nice, France, June 24-29, 2007.

[14] T. Rappaport, Wireless Communication; Principles and Practice, 2nd ed., Prentice Hall, NJ, 2002, pp. 172-174 (Chapter 4).

[15] F. Digham, M. Alouini, M. K. Simon, On the energy detection of unknown signals over fading channels, IEEE TR. on Comm. 55 (2007) 21-24.

[16] S. Herath, N. Rajatheva, C. Tellambura, Energy detection of unknown signals in fading and diversity reception, IEEE TR. on Comm. 59 (2011) 2443-2453.

[17] H. Poor, An Introduction to Signal Detection and Estimation, 2nd ed., Springer, New York, 1994, pp. 84-88.

[18] Y. Liang, Y. Zeng, E. Peh, A. Hong, Sensing-throughput tradeoff for cognitive radio networks, IEEE Trans. Wireless Commun. 7 (2008) 1326-1337.

[19] F. Olver, D. Lozier, R. Boisvert, C. Clark, NIST Handbook of Mathematical Functions, Cambridge University Press, UK, 2010.

[20] Akin, M. Gursoy, Performance analysis of cognitive radio systems under qos constraints and channel uncertainty, in: Proceedings of IEEE GLOBECOM, Florida, Dec. 6-10 2010.

[21] M. Sencan, M. Gursoy, Achievable rates for pilot-assisted transmission over rayleigh fading channels, in: 40th Annual Conference on Information Sciences and Systems (CISS), Princeton University, NJ, March 22-24, 2006.

[22] Y. Chen, G. Yu, On cognitive radio networks with opportunistic power control strategies in fading channels, IEEE Trans. on Wireless Communications 7 (2008) 2752-2761.

[23] Q. Sun, D. Cox, H. Huang, A. Lozano, Estimation of continuous flat fading mimo channels, IEEE Trans. Wireless Communications 1 (2002) 549-553.

[24] C. Chang, Performance Guarantees in Commucation Networks, Springer, London, 2000.

[25] L. Musavian, S. Aïssa, S. Lambotharan, Effective capacity for interference and delay constrained cognitive radio relay channels, IEEE Trans. on Wireless Comm. 9 (2010) 1698-1707.

[26] E. Jorswieck, R. Mochaourab, M. Mittelbach, Effective capacity maximization in multi-antenna channels with covariance feedback, IEEE Trans. Wireless Commun. 9 (2010) 2988-2993.

[27] J. Gentle, Matrix Algebra: Theory, Computations and Applications in Statistics, Springer, New York, 2010.

[28] S. Boyd, L. Vandenberghe, Convex Optimization, 7th ed., Cambridge University Press, UK, 2009.

[29] D. Bertsekas, Convex Analysis and Optimization, Athena Scientific, Belmont, MA, 2003.