

# Efficient Spectrum Sensing for Cognitive Radio Networks via Joint Optimization of Sensing Threshold and Duration

Ling Luo and Sumit Roy

**Abstract**—Cognitive radio networks require fast and reliable spectrum sensing to achieve high network utilization by secondary users. Optimization approaches to spectrum sensing to-date have largely focused on maximizing throughput for secondary users while considering only a single parameter variable pertinent to sensing - notably the threshold or duration, but not both. In this work, we investigate the impact of true joint minimization under two performance criteria: a) minimization of the average time to detection of a spectrum hole and b) joint maximization of the aggregate opportunistic throughput. We show that the resulting non-convex problem is actually *biconvex* under practical conditions for which effective algorithms can be developed that yields reliable numerical procedures to solve the resulting optimization problem. The results show that the proposed approach can considerably improve system performance (in terms of the mean time to detect a spectrum hole and also the aggregate opportunistic throughput of both primary and secondary users), relative to the scenarios with only a single sensing variable or a sub-optimal ad-hoc optimization approach used for two variable case.

**Index Terms**—Cognitive radio, spectrum sensing, biconvex, joint optimization.

## I. INTRODUCTION

THE shift to digital television broadcasting in June 2009 has opened up “white spaces” [1] (temporally under-used spectrum) in the VHF and UHF bands. As a result, the Federal Communications Commission (FCC) is encouraging occupancy of these under-used spectrum by unlicensed users [2] to enhance overall spectral utilization. The TV broadcasting stations and receivers constitute the primary network that must be protected when unlicensed users attempt to use white spaces opportunistically. The unlicensed users must be *cognitive*, i.e., possess the attribute of *spectrum sensing* in order to detect available white spaces for occupancy and subsequent *spectrum sharing* (release of the white space when an incumbent re-appears).

Efficient spectrum sensing is critical to coexistence of primary and secondary users in licensed bands. Incorrect decisions at the sensing stage leads to two major consequences: a)

miss-detection of active primary users results in inadmissible interference to them when secondaries transmit as a result or b) missed transmission opportunistic when a primary channel is idle but the secondary fails to detect this. Thus, sensing performance is evaluated in terms of two probabilities: probability of false alarms (primary users present but detected as idle by secondary users) and probability of miss-detection (idle bands being detected as active by secondary users) that are functions of two design parameters: i) the sensing threshold and ii) the sensing duration.

Liang *et al.* [3] investigated the optimal sensing *duration* for detection that maximizes achievable throughput of the secondary network. Later, the authors extended this single frame optimization approach over multiple decision slots. Quan *et al.* [4] explored the scenario whereby the (wideband) primary user occupies multiple bands, that are sensed simultaneously, and sought to maximize the aggregate throughput as a function of sensing *thresholds* for each sub-band. *However no work to date considers optimization of sensing with respect to both threshold and integration time* - this is the main contribution of our work. Often, a single-parameter spectrum sensing formulation leads to a convex optimization problem with readily available algorithmic solutions. However, this ignores the potential from varying other available parameters (that are kept constant).

Our formulation of sensing as a double-parameter problem leads, at first sight, to a non-convex optimization; we later prove that these problems are actually biconvex, for which efficient numerical methods exist. This builds on our earlier work [5] where we investigated the various trade-offs involved in minimizing the mean time to detect a spectral hole, as a function of sensing architecture and algorithm, but kept the decision threshold fixed.

### A. Our Formulation and Relation to Existing Work

In this paper, we formulate *two* different approaches, both as a function of the two key variables (sensing threshold and duration) -

- Minimizing the mean time to detection, subject to constraints on the probabilities of miss-detection and false alarm;
- Maximizing the aggregate opportunistic throughput

We solve these two problems under a stochastic model where the probability that state of channel is (un)occupied is fixed,

Paper approved by M. R. Beuhrer, the Editor for Cognitive Radio and UWB of the IEEE Communications Society. Manuscript received October 5, 2010; revised August 8, 2011, January 14 and March 21, 2012.

L. Luo was with the University of Washington, Seattle. He is now with the Electric Power Research Institute, SMEPC, 200437, Shanghai, China (e-mail: luol@sh.sgcc.com.cn).

S. Roy is with the Department of Electrical Engineering, University of Washington, 98195, Seattle, USA (e-mail: sroy@u.washington.edu).

This work was supported in part by AFOSR Award FA-9550-09-1-0298. Digital Object Identifier 10.1109/TCOMM.2012.072612.100605

where the continuous spectrum sensing is used instead of one-shot sensing.

Although these problems are proved to be non-convex, we demonstrate in both cases how these can be converted into a *biconvex* optimization problem that can be efficiently solved numerically.

The relevant literature on optimization of spectrum sensing in cognitive radio networks can be classified along various dimensions - such as a) static or dynamic nature of the primary users' channel occupancy, b) the metric being optimized (e.g. mean time to detection of an available channel or aggregate throughput) and algorithmic implications. For example, Lee [6] investigated optimized sensing for a dynamic primary channel occupancy scenario, modeled by a birth and death process. In contrast, we assume a purely *static* channel scenario. Further, CR metrics can be broadly classified as PHY based - e.g. the mean time to detection of an available channel or MAC based - e.g. the aggregate throughput enhancement due to opportunistic use of the spectral holes by secondary users. Huang [7] used a threshold-based sensing transmission structure to maximize the secondary users' utility, using MAC layer utility maximization approach with static and dynamic channels. On the other hand, Peng [8] optimizes the sensing disruption by maximizing the average number of false detections, which can be treated as a PHY work.

Nearly all existing work employs one or the other formulations; here we consider optimization using both approaches to *highlight the important differences in the ensuing optimal points and ensuing design trade-offs*.

Most of the existing CR optimization formulations lead to single-parameter convex optimization, for which well-known approaches - such as interior-point method - exist. Since multi-parameter (joint) optimization is typically non-convex, prior work have tended to use *sub-optimal* approaches - such as sequential parameter-wise optimization in Quan [4]. In this paper, we will prove that our joint optimization problem is *biconvex* [9], that allows efficient algorithms - such as Alternate Convex Search [10]- to find *globally optimal solutions*.

The paper is organized as follows. Section II describes the system (signal and interference) model, and search strategy based on energy detection. Section III and IV respectively formulate the minimization of the mean detection time and the maximization of the aggregate opportunistic throughput. The proofs of biconvexity and the resulting algorithms to determine the optimal points are also given in these two sections. We present numerical results in Section V and conclude the paper with final remarks in Section VI.

## II. SYSTEM MODEL

### A. Channel Search

We consider a scenario where the band to be searched is composed of  $M$  contiguous channels, each of bandwidth  $B_c$ . Common sensing techniques include random and serial search [12]. In random search, a secondary user looks for a spectrum hole (available channel) by randomly choosing a channel for sensing and detecting, while serial search sequentially searches over the entire spectrum, as depicted in Fig. 1, until a free channel is found.

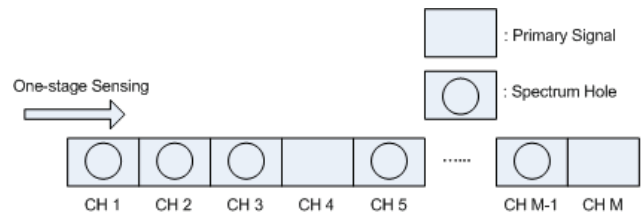


Fig. 1. Channel searching model.

We assume that the expected number of unused channels is  $L$  out of  $M$  ( $L \ll M$ ). Thus we can denote the probability of any channel being free by  $\lambda = L/M$  independent of all others. It was shown in [12] that the average number of search steps to detect an available channel for both random search and serial search for small  $L/M$  is given by (within a constant to the same order)

$$\overline{S}_{det} = \frac{1}{P_d \lambda}, \quad (1)$$

where  $P_d$  is the detection probability of a free channel, i.e. the product  $P_d \lambda$  is the aggregate probability that a channel is actually free, and that it is detected as such. Since the specific method of channel sensing (serial vs. random) is unimportant vis-a-vis search efficiency (at least for the assumed i.i.d model), we will not focus on this for the remainder of our work.

### B. Signal and Interference Model

We allow an underlay model, such that primary and secondary users can simultaneously transmit on the same channel, subject to adequate protection for the primary users. The reception model for any particular channel for this case is shown in Fig. 2.  $S_p$  and  $S_s$  respectively represent the transmitted power from a primary and a secondary user, while  $R_p$  and  $R_s$  represent the received power. In this work, we are interested in optimizing the sensing threshold and integration duration instead of selecting certain channels for multi-user communication. For the sake of simplicity, we assume that the channel response is time invariant. Further, the channel gains  $h_{pp}$ ,  $h_{ps}$ ,  $h_{sp}$  and  $h_{ss}$  from the primary (secondary) transmitters to primary (secondary) receivers are assumed to be the same over all the bands, consistent with the assumption of a flat fading channel scenario. The additive noise for primary and secondary users  $Z_p$  and  $Z_s$ , respectively, is assumed to be i.i.d, zero-mean Gaussian process with variance  $E[|Z_p|^2] = E[|Z_s|^2] = \sigma_n^2$ .

Some models for cognitive radio networks [11] assume sharing of causal side information between primary and secondary transmitters that allows joint pre-coding. We do not entertain this possibility; and assume that the respective multi-access interference components are treated as noise. Hence if both primary and secondary users are transmitting at the same time through the same channel, the SINR at primary and secondary receivers are respectively given by:

$$SINR_p = \frac{S_p h_{pp}^2}{S_s h_{sp}^2 + \sigma_n^2},$$

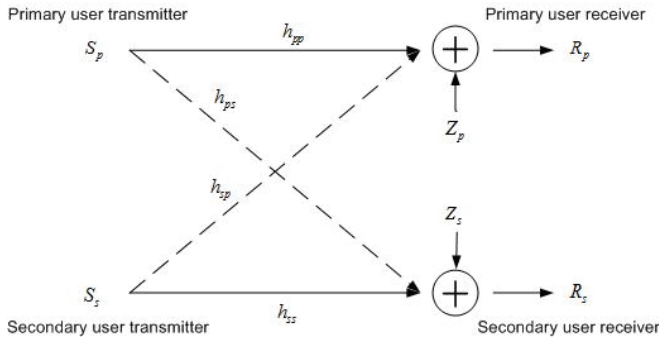


Fig. 2. Interference model of a primary and a secondary occupying the same channel.

$$SINR_s = \frac{S_s h_{ss}^2}{S_p h_{ps}^2 + \sigma_n^2}.$$

Then the capacities for both primary and secondary systems are given by:

$$C_{PI} = B_c \log \left( 1 + \frac{\gamma_{pp}}{1 + \gamma_{sp}} \right), \quad (2)$$

$$C_{SI} = B_c \log \left( 1 + \frac{\gamma_{ss}}{1 + \gamma_{ps}} \right), \quad (3)$$

where  $\gamma_{pp} = \frac{S_p h_{pp}^2}{\sigma_n^2}$ ,  $\gamma_{sp} = \frac{S_s h_{sp}^2}{\sigma_n^2}$ ,  $\gamma_{ss} = \frac{S_s h_{ss}^2}{\sigma_n^2}$  and  $\gamma_{ps} = \frac{S_p h_{ps}^2}{\sigma_n^2}$ .

On the other hand, if there are only primary (secondary) users transmitting without co-channel interference, their respective capacities can be written as:

$$C_P = B_c \log(1 + \gamma_{pp}), \quad (4)$$

$$C_S = B_c \log(1 + \gamma_{ss}). \quad (5)$$

### C. Energy Detector

We consider a binary hypothesis test for spectrum sensing of primary occupation as follows:

$$\begin{aligned} H_0 &: R_d = Z_s, \\ H_1 &: R_d = h_{ps} S_p + Z_s, \end{aligned}$$

where  $R_d$  is the received signal at the detector of the secondary user.

The channel status is determined based on energy detection with  $N$  samples. Thus, the sensing duration is  $T_i = N/B_c$ , and the test statistic for the energy detector is:

$$Y = \frac{1}{N} \sum_{i=1}^N |R_d(i)|^2. \quad (6)$$

we denote the event “the secondary user claims that the channel is idle” as  $D_0$  and the event “the secondary user claims that the channel is busy” as  $D_1$ . We then denote the probability of successful detection of an available channel (hypothesis  $H_0$ ) as  $P_d$  and corresponding false alarm probability  $P_{fa}$  (incorrect determination of a busy channel as free under hypothesis  $H_1$ ). Therefore we have

$$\begin{aligned} Pr(D_0|H_0) &= P_d, \\ Pr(D_1|H_0) &= 1 - P_d, \end{aligned}$$

$$\begin{aligned} Pr(D_0|H_1) &= P_{fa}, \\ Pr(D_1|H_1) &= 1 - P_{fa}. \end{aligned}$$

For a large value of  $N$ , by applying Central Limit Theorem, we approximate  $Y$  to be a Gaussian random variable whose mean and variance under the two hypothesis ( $H_0$  corresponds to channel idle and  $H_1$  denotes channel busy) are given by:

$$\begin{aligned} E[Y|H_0] &= \sigma_n^2, \\ E[Y|H_1] &= (\gamma + 1)\sigma_n^2, \end{aligned} \quad (7)$$

$$\begin{aligned} Var[Y|H_0] &= \frac{2}{N}\sigma_n^4, \\ Var[Y|H_1] &= \frac{2}{N}(\gamma + 1)^2\sigma_n^4, \end{aligned} \quad (8)$$

where  $\gamma$  is the per-sample signal-to-noise ratio (SNR) at the energy detector of the secondary user.

For a given sensing threshold  $x$ ,  $P_d$  and  $P_{fa}$  are given by:

$$\begin{aligned} P_d(x, N) &= Pr(Y < x | H_0) \\ &= 1 - Q\left(\frac{x - \sigma_n^2}{\sigma_n^2 \sqrt{2}} \sqrt{N}\right), \end{aligned} \quad (9)$$

$$\begin{aligned} P_{fa}(x, N) &= Pr(Y < x | H_1) \\ &= 1 - Q\left(\frac{x - (\gamma + 1)\sigma_n^2}{\sigma_n^2 (\gamma + 1) \sqrt{2}} \sqrt{N}\right), \end{aligned} \quad (10)$$

where the Q-function is:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt.$$

### III. OPTIMIZATION APPROACH I: MINIMIZING MEAN DETECTION TIME

In this section, we seek to minimize the average time to detect an available (unused by primary users) channel by secondary users as a function of two parameters of the energy detector, *i.e.*, the sensing threshold and the integration time. Searching over multiple channels involves two components: a) channel switching time  $T_s$  largely determined by the settling time of the phase locked loop (PLL) used in the receiver circuitry, and b) integration time  $T_i$  needed by the energy detector to arrive at a reliable detection [5], which in turn is determined by the sensing bandwidth and the number of sensing samples.

Thus, the net mean time to detect an available channel is given by

$$\begin{aligned} \overline{T}_{det} &= \overline{S}_{det}(T_s + T_i), \\ &= \frac{1}{P_d \lambda} (T_s + \frac{N}{B_c}). \end{aligned} \quad (11)$$

Considering Eq. (11), we note an immediate trade-off between the average number of detection steps and the integration time: a longer integration time leads to a better  $P_d - P_{fa}$  performance, hence reducing the mean number of detection steps (and consequently reducing the detection time), and vice versa. Moreover, for any given integration time, the energy detector with higher sensing threshold results in higher detection and false alarm probabilities. Hence, we focus on jointly optimizing the sensing threshold and the number of sensing samples in order to minimize the average detection time.

Therefore, the joint optimization problem is mathematically formulated as below:

$$\begin{aligned} \min_{x, N} \quad & \overline{T_{det}} = \frac{1}{P_d \lambda} (T_s + \frac{N}{B_c}) \\ \text{subject to} \quad & 1) P_d(x, N) \geq \alpha; \\ & 2) P_{fa}(x, N) \leq \beta \end{aligned} \quad (12)$$

where the detection  $P_d(\cdot, \cdot)$  and false alarm  $P_{fa}(\cdot, \cdot)$  probabilities are functions of sensing threshold ( $x$ ) and the number of integration samples ( $N$ ). The parameters  $\alpha$  and  $\beta$  denote appropriate constraints on detection and false alarm probabilities.

The formulation Eq. (12) leads generally to a *non-convex* optimization problem. For practical systems, a high detection probability  $\alpha$  and a low false alarm probability  $\beta$  are desired. Next, we will prove that the original non-convex problem can be transformed into a biconvex optimization problem [10] when  $\alpha > 0.5$  and  $\beta < 0.5$ .

*Defn:* A function  $f(x, y) : X \times Y \rightarrow \mathbb{R}$  is biconvex if and only if  $f(x, y)$  is convex in  $y$  given  $x \in X$  and convex in  $x$  given  $y \in Y$ . A brief tutorial of related definitions and properties of biconvex functions and optimization problems is presented in Appendix I.

#### A. Biconvexity

To show that the optimization problem (12) is biconvex, we introduce variables  $\varepsilon = \frac{x - \sigma_n^2}{\sigma_n^2 \sqrt{2}}$ ,  $\eta = \sqrt{N}$  and let  $\tau = T_s B_c$ . We denote  $\mathbb{E}$  and  $\mathbb{H}$  as the feasible sets for  $\varepsilon$  and  $\eta$ . Combining Eqs. (9) and (10), the original optimization problem is rewritten as:

$$\begin{aligned} \min_{\varepsilon, \eta} \quad & \overline{T_{det}}(\varepsilon, \eta) = \frac{1}{B_c \lambda} \frac{\tau + \eta^2}{1 - Q(\varepsilon \eta)} \\ \text{subject to} \quad & 1) Q(\varepsilon \eta) \leq 1 - \alpha; \\ & 2) -Q\left(\frac{\sqrt{2}\varepsilon - \gamma}{\sqrt{2}(\gamma + 1)} \eta\right) \leq \beta - 1; \\ & 3) \eta \geq 1 \end{aligned} \quad (13)$$

*Lemma 1:* Given  $\varepsilon = \tilde{\varepsilon} \in \mathbb{E}$ , the function  $Q(\tilde{\varepsilon} \eta)$  is convex in  $\eta$  if  $P_d > 1/2$ .

*Proof:* See Appendix II.

*Lemma 2:* Given  $\varepsilon = \tilde{\varepsilon} \in \mathbb{E}$ , the function  $-Q\left(\frac{\sqrt{2}\tilde{\varepsilon} - \gamma}{\sqrt{2}(\gamma + 1)} \eta\right)$  is convex in  $\eta$  if  $P_{fa} < 1/2$ .

*Proof:* See Appendix III.

*Lemma 3:* Given  $\varepsilon = \tilde{\varepsilon} \in \mathbb{E}$ , the function  $\overline{T_{det}}(\tilde{\varepsilon}, \eta)$  is convex in  $\eta$ , if  $P_d > 1/2$ .

*Proof:* See Appendix IV.

*Theorem 1:* The optimization problem (13) is biconvex for  $\varepsilon \in \mathbb{E}$  and  $\eta \in \mathbb{H}$ , if  $P_d > 1/2$  and  $P_{fa} < 1/2$ .

*Proof:* Combining Lemmas 1-3 we can conclude that problem (13) is convex in  $\eta$  given  $\varepsilon = \tilde{\varepsilon} \in \mathbb{E}$ , if  $P_d > 1/2$  and  $P_{fa} < 1/2$ . Thus we are left to show that optimization problem (13) is convex in  $\varepsilon$  given  $\eta = \tilde{\eta} \in \mathbb{H}$ , if  $P_d > 1/2$  and  $P_{fa} < 1/2$ .

It is easy to prove that  $Q(\varepsilon \tilde{\eta})$  and  $-Q\left(\frac{\sqrt{2}\varepsilon - \gamma}{\sqrt{2}(\gamma + 1)} \tilde{\eta}\right)$  are all convex in  $\varepsilon$  under the constraints  $P_d > 1/2$  and  $P_{fa} < 1/2$  by using a similar technique. Then taking the second order

partial derivative of  $\overline{T_{det}}(\varepsilon, \tilde{\eta})$ , we have:

$$\frac{\partial^2 \overline{T_{det}}(\varepsilon, \tilde{\eta})}{\partial \varepsilon^2} = \frac{\tau + \tilde{\eta}^2}{B_c \lambda} [\tilde{\eta}^2 G(\tilde{\eta}, \varepsilon) + 2G(\tilde{\eta}, \varepsilon)^2 F(\varepsilon, \tilde{\eta})^{-3}],$$

which is clearly positive. Hence the convexity in  $\varepsilon$  given  $\eta = \tilde{\eta} \in \mathbb{H}$  follows, and Theorem 1 is proved.  $\square$

$$\begin{aligned} \frac{\partial^2 \overline{T_{det}}(\varepsilon, \eta)}{\partial \varepsilon \partial \eta} &= F(\varepsilon, \eta)^{-2} [(\tau + \eta^2) \left( \frac{\varepsilon^2 \eta^2}{\sqrt{2\pi}} e^{-\varepsilon^2 \eta^2 / 2} \right. \\ &+ F(\varepsilon, \eta)^{-1} \frac{\varepsilon \eta}{\pi} e^{-\varepsilon^2 \eta^2} - \frac{1}{\sqrt{2\pi}} e^{-\varepsilon^2 \eta^2 / 2} \left. \right) \\ &- \frac{2\eta^2}{\sqrt{2\pi}} e^{-\varepsilon^2 \eta^2 / 2}]. \end{aligned}$$

It is tediously straightforward to check that the determinant of the Hessian matrix  $\frac{\partial^2 \overline{T_{det}}(\varepsilon, \eta)}{\partial \varepsilon^2} \frac{\partial^2 \overline{T_{det}}(\varepsilon, \eta)}{\partial \eta^2} - \left( \frac{\partial^2 \overline{T_{det}}(\varepsilon, \eta)}{\partial \varepsilon \partial \eta} \right)^2$  is not always non-negative, which shows that the problem (13) is non-convex.

#### B. Feasible Algorithms for the Optimal Points

Several feasible algorithms are outlined in [10] to determine the optimum points in general biconvex optimization problems. For our specific biconvex optimization problem, we will describe a fast and simple method to find the globally optimal  $x$  and  $N$  by exploiting some problem-specific information.

For any given  $\eta = \tilde{\eta} \in \mathbb{H}$ , because  $P_d(\varepsilon, \tilde{\eta}) = 1 - Q(\varepsilon, \tilde{\eta})$  is monotonic increasing in  $\varepsilon$ , we see that the objective function in the problem (13), i.e.,  $\overline{T_{det}} = \frac{\tau + \tilde{\eta}^2}{P_d(\varepsilon, \tilde{\eta}) B_c \lambda}$  is monotonic decreasing in  $\varepsilon$ . To minimize  $\overline{T_{det}}$  is thus equivalent to achieve the maximum  $\varepsilon$  under the given constraints. Note that for a given  $\eta = \tilde{\eta}$ , the detection probability constraint  $P_d \geq \alpha$  gives the lower bound for  $\varepsilon$  and the false alarm probability constraint  $P_{fa} = 1 - Q\left(\frac{\sqrt{2}\varepsilon - \gamma}{\sqrt{2}(\gamma + 1)} \tilde{\eta}\right) \leq \beta$  gives the upper bound for  $\varepsilon$ . This can be explained as follows: the system benefits from the high detection probability for a pre-defined sensing threshold, which in turn leads to the high false alarm probability. The system achieves the minimum average detection time when the detection probability reaches the highest feasible value, if and only if the false alarm probability reaches the upper bound. Let  $P_{fa}(\varepsilon_{opt}, \tilde{\eta}) = \beta$ , if  $Q(\varepsilon_{opt} \tilde{\eta}) \leq 1 - \alpha$ , then we reach the upper bound of the false alarm probability and

$$\varepsilon_{optimal} = \varepsilon_{opt} = \frac{Q^{-1}(1 - \beta)(\gamma + 1)}{\tilde{\eta}} + \frac{\gamma}{\sqrt{2}},$$

where  $Q^{-1}(\cdot)$  is the inverse Q-function. Otherwise, the upper bound of  $\varepsilon$  is achieved when  $Q(\varepsilon \tilde{\eta}) = 1 - \alpha$ . Note that  $P_{fa}(\varepsilon, \tilde{\eta}) < \beta$  in this case, and thus

$$\varepsilon_{optimal} = \frac{Q^{-1}(1 - \alpha)}{\tilde{\eta}}.$$

Because the problem (13) is biconvex, there exists a global optimal  $(\varepsilon_{optimal}, \eta_{optimal})$  corresponding to minimum average detection time. Thus we can derive the relationship of global optimal points for  $\eta$  and  $\varepsilon$ :

$$\varepsilon_{optimal} = \max\left(\frac{Q^{-1}(1 - \beta)(\gamma + 1)}{\eta_{optimal}} + \frac{\gamma}{\sqrt{2}}, \frac{Q^{-1}(1 - \alpha)}{\eta_{optimal}}\right). \quad (14)$$

In general, we claim as follows:

*Claim 1:* If  $Q^{-1}(1-\beta)(\gamma+1) + \eta_{optimal}\gamma/\sqrt{2} \geq Q^{-1}(1-\alpha)$ , the minimum average detection time is attained when the false alarm probability reaches its constraint.

*Claim 2:* If  $Q^{-1}(1-\beta)(\gamma+1) + \eta_{optimal}\gamma/\sqrt{2} \leq Q^{-1}(1-\alpha)$ , the minimum average detection time is achieved when the detection probability constraint is attained.

To reduce the region of  $\eta$  for initiating the search for optimal points, we can observe from (13) that:

$$\frac{Q^{-1}(1-\alpha)}{\gamma+1} - \frac{\gamma\eta}{\sqrt{2}(\gamma+1)} \leq \frac{\varepsilon\eta}{\gamma+1} - \frac{\gamma\eta}{\sqrt{2}(\gamma+1)} \leq Q^{-1}(1-\beta).$$

Thus the achievable region for  $\eta$  is:

$$\eta \geq \left( \frac{Q^{-1}(1-\alpha)}{\gamma+1} - Q^{-1}(1-\beta) \right) \frac{\sqrt{2}(\gamma+1)}{\gamma}.$$

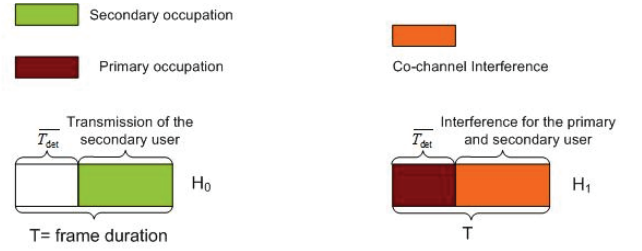
Using Eq. (14), we can convert the problem (13) into a single-parameter optimization problem, since it is convex for any given  $\varepsilon = \tilde{\varepsilon} \in \mathbb{E}$ . Therefore, we can numerically solve the global optimal values of  $\eta$  and  $\varepsilon$  by using penalty method in terms of variable  $\eta$ .

#### IV. OPTIMIZATION APPROACH 2: AGGREGATE OPPORTUNISTIC THROUGHPUT

Minimizing the mean time to detection of a spectral hole is an appropriate *physical* layer metric for optimization a cognitive network. However, this ignores the impact of incorrect link layer decisions; for example, if there is a false alarm for a busy channel, the primary and the secondary users will simultaneously transmit and the co-channel interference will significantly decrease the throughput from both primary and secondary systems. Therefore, it is more appropriate to use a MAC layer metric - notably the aggregate opportunistic throughput of both primary and secondary users, so as to uncover newer system trade-offs not captured by the earlier method.

##### A. Average Aggregate Opportunistic Data Rate

We assume that secondary users periodically sense the spectrum frame by frame. The duration of a frame is of **fixed** length  $T$ , and thus the average transmission time for a secondary user is  $T - \overline{T_{det}}$  per frame. On the other hand, a primary user uses whole frame duration to transmit its signal when occupying a certain channel. We assume the presence of a central controller that searches for available channels and then assigns any such channels to the secondary user. Two situations of successful detection and false alarm are shown in Fig. 3, leading to different aggregate throughput. In the rest of the paper, we use the term ' $TH$ ' to denote the opportunistic throughput per frame per channel. Note that the sensing scheme in our paper is NOT one-shot. If the scheme misses a free channel (with probability  $1 - P_d$ ), it will NOT simply give up but go on sensing, which means that the search scheme will ultimately find a channel for the secondary user, either in hypothesis  $H_0$  (channel vacant) or  $H_1$  (channel occupied by a primary signal). The searching time of the wasted chances due to the miss-detection probability is related to calculating the average detection time.



(a). Successful detection of a free channel (b). False alarm of a busy channel

Fig. 3. Two cases for secondary occupation of the channel.

Fig. 3(a) shows the successful occupation of a free channel by a secondary user. The average aggregate opportunistic throughput in this case is thus given by:

$$\overline{TH|H_0} = \left(1 - \frac{\overline{T_{det}}}{T}\right) C_S. \quad (15)$$

When a channel is occupied, the false alarm of busy channel status as shown in Fig. 3(b) will result in secondary transmissions on the same channel, resulting in co-channel interference. The opportunistic throughput obtained by the secondary user per frame per channel in the presence of the primary is given by

$$\overline{TH_{sec}|H_1} = \left(1 - \frac{\overline{T_{det}}}{T}\right) C_{SI}$$

However, the *average* throughput of the primary user is composed of the time-average of that without and with secondary interference, respectively:

$$\overline{TH_{pri}|H_1} = \frac{\overline{T_{det}}}{T} C_P + \left(1 - \frac{\overline{T_{det}}}{T}\right) C_{PI}.$$

The pertinent quantity is the *change* (reduction) in throughput of the primary user due to secondary interference, given by

$$\overline{TH'_{pri}|H_1} = \overline{TH_{pri}|H_1} - C_P = \left(1 - \frac{\overline{T_{det}}}{T}\right) (C_{PI} - C_P).$$

We introduce a factor to signal protection of the primary user throughput from secondary interference by weighting it higher in the aggregate throughput<sup>1</sup>, i.e. the average additional aggregate throughput under primary channel occupancy is now modified as follows:

$$\begin{aligned} \overline{TH|H_1} &= \overline{TH_{sec}|H_1} + J \cdot \overline{TH'_{pri}|H_1} \\ &= \left(1 - \frac{\overline{T_{det}}}{T}\right) (C_{SI} + JC_{PI} - JC_P) \end{aligned} \quad (16)$$

Note that the only situation when the sensing scheme stops is that it claims a free channel is detected, although this case might be a false alarm with probability  $P_{fa}$ . Thus, the throughput is  $\overline{TH|H_0}$  under hypothesis  $H_0$ , while the throughput is  $\overline{TH|H_1}P_{fa}$  under hypothesis  $H_1$ . Combining Eqns. (15) and (16), the average aggregate opportunistic throughput can be re-written as:

$$\overline{TH} = \overline{TH|H_0}Pr(H_0) + \overline{TH|H_1}P_{fa}Pr(H_1)$$

<sup>1</sup> $J = 1$  is the direct sum of the throughput a) increase by opportunistic secondary occupation and b) decrease seen by the primary user due to secondary interference.

$$= (1 - \frac{\overline{T}_{det}}{T})(\phi - \varphi P_{fa}), \quad (17)$$

where  $\phi = \lambda C_S$  and  $\varphi = (1 - \lambda)(JC_P - JC_{PI} - C_{SI})$ .

### B. Maximizing Average Aggregate Opportunistic Throughput

Following the constraints setup in the previous section, to maximize the average aggregate opportunistic throughput, we can formulate the problem as below:

$$\begin{aligned} & \max_{x, N} \quad \overline{TH} \\ & \text{subject to} \quad 1) P_d(x, N) \geq \alpha; \\ & \quad \quad \quad 2) P_{fa}(x, N) \leq \beta; \end{aligned} \quad (18)$$

It can be observed that  $\overline{TH} = \phi - [\frac{\overline{T}_{det}}{T}(\phi - \varphi P_{fa}) + \varphi P_{fa}]$ . Because  $T$  and  $\phi$  are of constant positive values, we can then re-write the original optimization problem (18) from Eq. (13) as:

$$\begin{aligned} & \min_{\varepsilon, \eta} \quad K = \frac{\overline{T}_{det}}{T}(\phi - \varphi P_{fa}) + \varphi P_{fa} \\ & \text{subject to} \quad 1) Q(\varepsilon, \eta) \leq 1 - \alpha; \\ & \quad \quad \quad 2) -Q(\frac{\sqrt{2}\varepsilon - \gamma}{\sqrt{2}(\gamma + 1)}\eta) \leq \beta - 1; \\ & \quad \quad \quad 3) \eta \geq 1; \end{aligned} \quad (19)$$

### C. Biconvexity under Constraints

The equivalent optimization problem (19) is still non-convex and seemingly not biconvex. However, we will prove that it is biconvex under the constraints below:

$$J(C_P - C_{PI}) \geq C_{SI}; \quad (20)$$

$$\phi - \varphi P_{fa} \geq 0; \quad (21)$$

$$\tau \ll \eta^2. \quad (22)$$

Firstly we will discuss the physical meaning of the constraints above. Constraint (20) is equivalent to  $\varphi \geq 0$ . It can be achieved when the system desires a strong protection of primary users, i.e.,  $J \gg 1$ . In this case, the interference from a secondary user in the same channel leads to the decrease in aggregate opportunistic throughput. Constraint (21) is achieved when  $P_{fa}$  is relatively small, implying that the system requires a small probability that secondary users interfere with primary users. Constraint (22) is achieved when PLL settling time is very small, compared with the integration time in the energy detector. Under this constraint,  $T_s = \tau/B_c \ll \eta^2/B_c = N/B_c$ . The PLL design [13] ( $T_s = 0.009\mu s$ ) in TV channel detection ( $B_c = 6MHz$ ) is an example of this case. Then we say  $\overline{T}_{det}$  approximately equals:

$$\overline{T}_{det} = \frac{\tau + \eta^2}{1 - Q(\varepsilon, \eta)} \approx \eta^2 F(\varepsilon, \eta)^{-1}.$$

*Theorem 2:* The optimization problem (19) is biconvex for  $\varepsilon \in \mathbb{E}$  and  $\eta \in \mathbb{H}$  under constraints (20)-(22), if  $P_d > 1/2$  and  $P_{fa} < 1/2$ .

*Proof:* See Appendix V.

### D. Algorithm of Solving Biconvex Problem

To solve this biconvex optimization problem, we use an Alternate Convex Search [10], which is a special case of

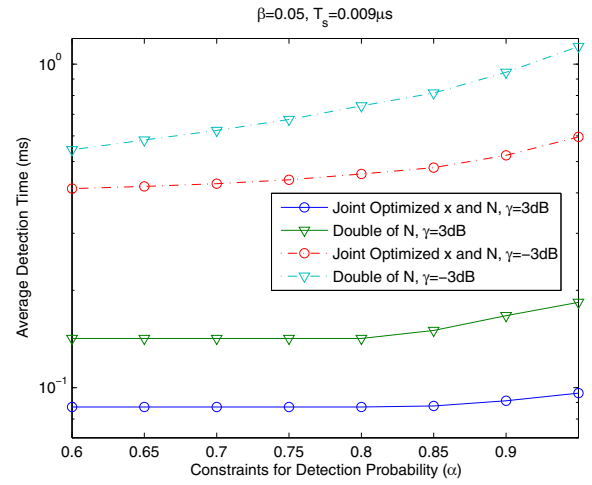


Fig. 4. Minimized average detection time: varying  $\alpha$ .

the Block-Relaxation Methods proposed by Leeuw [18]. The algorithm is as follows:

Step 1: Suppose that  $B$  is the convex set for  $\varepsilon$  and  $\eta$  as described in Appendix I. Choose an arbitrary available starting point  $(\varepsilon_i, \eta_i)$  and set  $i = 0$ .

Step 2: Solve for fixed  $\eta_i$  the convex optimization problem:

$$\min\{K(\varepsilon, \eta_i), \varepsilon \in B_{\eta_i}\}.$$

If there is an optimal point  $\varepsilon' \in B_{\eta_i}$  then set  $\varepsilon_{i+1} = \varepsilon'$ , otherwise stop.

Step 3: Solve for fixed  $\varepsilon_{i+1}$  the convex optimization problem:

$$\min\{K(\varepsilon_{i+1}, \eta), \eta \in B_{\varepsilon_{i+1}}\}.$$

If there is an optimal point  $\eta' \in B_{\varepsilon_{i+1}}$  then set  $\eta_{i+1} = \eta'$ , otherwise stop.

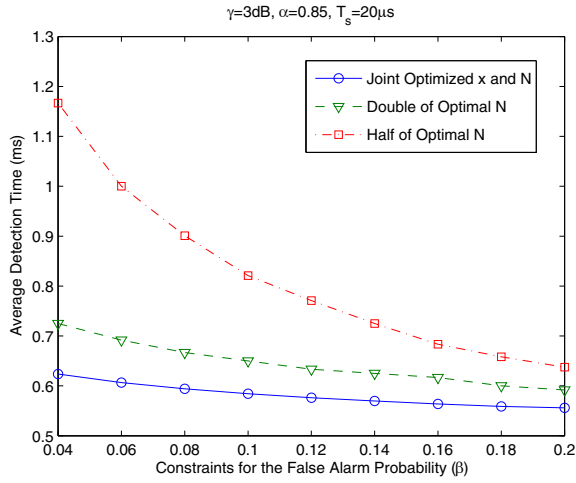
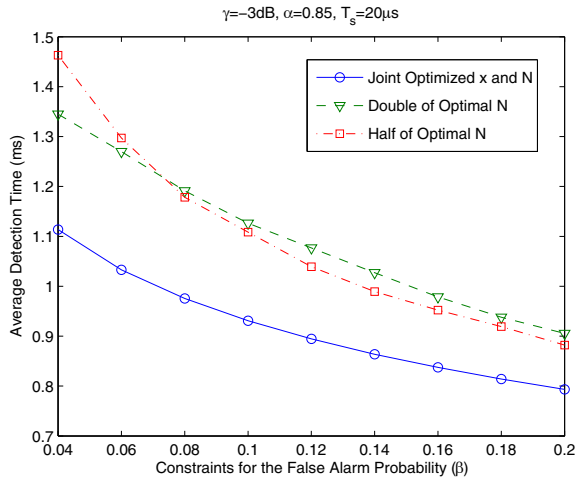
Step 4: If  $(\varepsilon_{i+1}, \eta_{i+1})$  is a stopping criterion, then stop. Otherwise, go back to Step 2.

## V. FINDING OPTIMAL POINTS: NUMERICAL RESULTS

Iterative algorithms for finding the global optimum points in the biconvex optimization problems are discussed in [10]. Here we use the algorithm described in Section III-C to solve the minimized average detection time. We present results for two scenarios: for fixed  $\beta$  and varying  $\alpha$ , and vice versa. We let  $M = 100, L = 4$  and  $B_c = 6MHz$  conforming to the bandwidth of a TV channel.

We note from the equivalent problem (13) that both  $\tau$  (proportional to  $T_s$ ) and  $\gamma$  (SNR) play important roles in determining the minimum average detection time. Thus, we consider two SNR cases ( $\gamma = \pm 3dB$ ) and also two values of switching time  $T_s = 0.005\mu s$  (type I PLL) and  $T_s = 20\mu s$  (type II PLL). From Eq. 11 we can notice that the channel switching time significantly impact the the minimized detection time. If  $T_s$  is very small (type I PLL), minimizing  $\overline{T}_{det}$  is equivalent to minimizing  $P_d\lambda$ . On the other hand,  $N/B_c$  will be a dominant factor when  $T_s$  is relatively large (type II PLL).

Fig. 4 shows the results for type I circuit, where the circuit switching time is much smaller than the integration time. The


 Fig. 5. Minimized average detection time: varying  $\beta$ ,  $\gamma = 3dB$ .

 Fig. 6. Minimized average detection time: varying  $\beta$ ,  $\gamma = -3dB$ .

curves of 'Joint Optimized  $x$  and  $N$ ' present the results of our joint optimization method. The curves of 'Double of  $N$ ' shows the independent optimization of  $x$  when the number of sensing samples is the twice of the optimized value from our method. The average detection time increases with  $\alpha$  as expected. We also note that the minimum average detection time is approximately constant for  $0.6 \leq P_d \leq 0.85$ . This is because the local optimal point  $\varepsilon_{optimal} = \frac{Q^{-1}(1-\beta)(\gamma+1)}{\eta_{optimal}} + \frac{\gamma}{\sqrt{2}}$  is invariant w.r.t  $\alpha$  while satisfying the constraint  $Q(\varepsilon_{optimal}\eta_{optimal}) \leq 1 - \alpha$ . Thereafter, the optimal mean detection time is strictly increasing for  $P_d > 0.85$ , because  $\varepsilon_{optimal} = \frac{Q^{-1}(1-\alpha)}{\eta_{optimal}}$ .

Figs. 5-6 shows the results for type II circuit, where the circuit switching time is close to the integration time. We compare the results of joint optimization with the results of independent optimized  $x$  when  $N$  is half or double of the optimized values. It can be noticed that all curves are going down as  $\alpha$  increases, because the higher constraints for false alarm probability leads to less integration time, hence decreasing the minimum average detection time. It can be also observed that the optimal values of detection time are at least

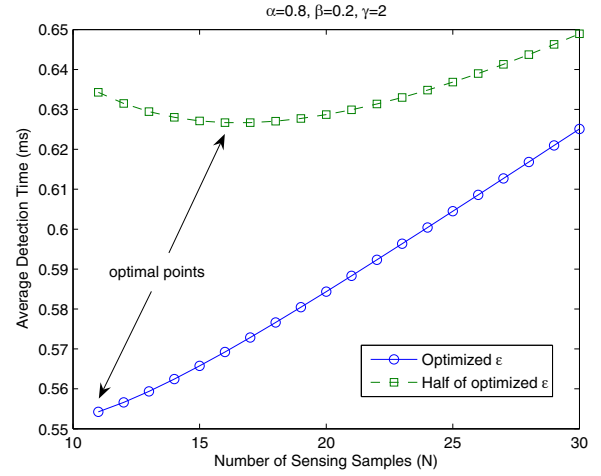
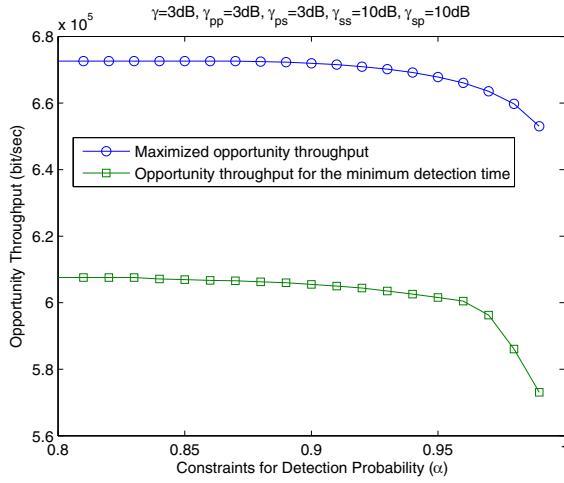
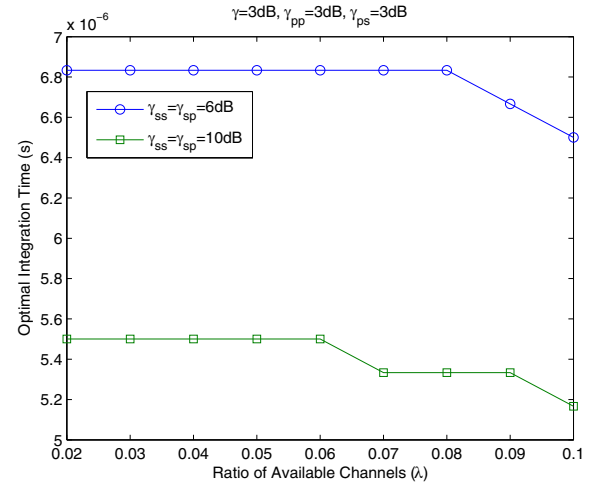
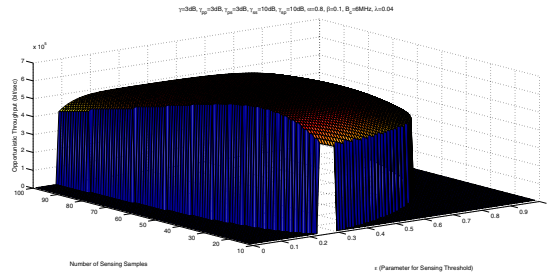


Fig. 7. Performance of joint optimization and single parameter optimization.

20% lower than taking half or double of the optimal number of sensing samples.

To have a better sense of the advantage of joint optimization over single parameter optimization, we change the value of optimized sensing threshold and take various values of  $N$  ( $N = \eta^2$ ) to explore the minimized average detection time. Performance of two different approaches in the given scenario (type II PLL,  $\lambda = 0.04$ ,  $\alpha = 0.8$ ,  $\beta = 0.2$ ,  $\bar{\gamma} = 3dB$ ) is shown in Fig. 7. Because of the constraints  $1 - Q(\varepsilon\eta) \geq 0.8$  and  $1 - Q(\frac{\sqrt{2}\varepsilon - \bar{\gamma}}{\sqrt{2}(\bar{\gamma}+1)}\eta) \leq 0.2$ , we can calculate that  $N_{min} = 11$ . The circle line in the figure gives the average detection time for the  $\varepsilon_{optimal}$ . By taking various values of the number of sensing samples, we can find that  $N_{optimal} = 11$  and  $\bar{T}_{det} = 0.554ms$  for joint optimization. The square line shows the performance when we take half of  $\varepsilon_{optimal}$  to set the sensing threshold. It can be observed that the optimal number of sensing samples is 16 in this case. The minimized average detection time of this sensing threshold by applying single parameter optimization is  $0.628ms$ , which is 13% higher than the value of joint optimization.

The maximized opportunistic throughput is shown in Fig. 8, where we set  $\lambda = 0.04$ ,  $B_c = 6MHz$ ,  $\beta = 0.05$ ,  $T = 1ms$ ,  $J = 2$ . The SNR of secondary transmission at both primary and secondary receivers is 10dB, and the SNR of primary transmission at primary receivers, secondary receivers and energy detectors is 3dB. The experiment needs 1000 iterations to converge. It can be observed from the graph that the maximized opportunistic throughput decreases as the constraints becomes tighter. When  $\alpha \leq 0.87$ , the maximized opportunistic throughput keeps the same, because the optimal points ( $N_{optimal} = 16$ ) is inside the region determined by the constraints of detection and false alarm probability. On the other hand, the maximized opportunistic throughput is going down for  $\alpha \geq 0.88$ , since the optimal points are achieved when the system reaches the constraint of detection probability. For example,  $N_{optimal} = 26$  for the case of  $\alpha = 0.99$ . Intuitively, the less detection time leads to the larger throughput. However, taking the false alarm probability into account from Eq. (17), there is no guarantee that the optimal points of the minimum

Fig. 8. Maximized opportunistic throughput: varying  $\alpha$ .Fig. 10. Optimal integration time: varying  $\lambda$ .Fig. 9. Opportunistic throughput:  $\alpha = 0.9$  and  $\beta = 0.1$ .

average detection time result in the maximum opportunistic throughput. Shown in Fig. 8, the opportunistic throughput when the system achieves the minimum average detection time is about 10% lower than the optimal value.

Fig. 9 shows the 3-D plot of the aggregate opportunistic throughput in the same scenario while  $\alpha = 0.8$  and  $\beta = 0.1$ . It can be observed that the maximized aggregate opportunistic throughput matches the results we have. We can also see that there is only one saddle point in the achievable region. Although the objective function has some other local minima, they lead to either higher false alarm probability or lower detection probability compared to the constraints for  $\alpha$  and  $\beta$ . In this scenario, the maximized aggregate opportunistic throughput is  $6.73 \times 10^5$  (bit/sec), and it is achieved when  $N_{optimal} = 33$  and  $\varepsilon_{optimal} = 0.2$ .

By changing the ratio of available channels ( $\lambda$ ), we present optimal integration time for the maximum opportunistic throughput in Fig. 10. We set the SNR of primary transmission at primary receivers, secondary receivers and energy detectors to 3dB. The higher SNR of secondary transmissions, the less integration time to achieve a good  $P_d - P_{fa}$  performance. Hence the optimal integration time is smaller than the lower SNR case. For a higher ratio of available channels, the mean time to detect a free channel is shorter, implying that the optimal integration time will not be longer than the low  $\lambda$  case. This can be also observed from Fig. 10.

## VI. CONCLUSION

This paper proposes a biconvex optimization approach for spectrum sensing in cognitive radio networks. We formulate a joint optimization problem regarding both the sensing threshold and the number of sensing samples in the energy detector, to minimize the average detection time of an available channel. Although this problem is non-convex, we prove that it can be transformed into a biconvex problem under practical conditions, followed by a feasible fast algorithm to solve the optimal points. We then formulate a joint optimization problem to maximize the average aggregate opportunistic throughput and prove that it is also biconvex under some conditions. Numerical results show that the proposed optimal approach can significantly decrease the mean time to detect a spectrum hole and increase the average aggregate opportunistic throughput of both primary and secondary systems.

## APPENDIX A BICONVEXITY

Let  $X \subseteq \mathbb{R}$  and  $Y \subseteq \mathbb{R}$  be two non-empty convex sets over the real line; and  $B \subseteq X \times Y$ . Let  $B_x$  and  $B_y$  be  $x$ - and  $y$ -sections of  $B$ :

$$\begin{aligned} B_x &:= \{y \in Y : (x, y) \in B\}, \\ B_y &:= \{x \in X : (x, y) \in B\} \end{aligned}$$

*Definition 1:* Set  $B$  is a biconvex set on  $X \times Y$  if  $B_x$  is convex for every  $x \in X$  and  $B_y$  is convex for every  $y \in Y$ .

A convex set is by definition, biconvex, but the reverse is not true. Fig. 11 gives an example of biconvex set  $B$ , which is non-convex.

$$\begin{aligned} B &= \{(x, y) \in \mathbb{R}^2 : 1 \leq x, y \leq 2\} \\ &\cup \{(x, y) \in \mathbb{R}^2 : 1 \leq x \leq 2, 0 \leq y \leq 1\} \\ &\cup \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 1 \leq y \leq 2\}. \end{aligned}$$

It can be easily checked from the graph that  $B_x$  is convex for all  $y \in Y$  and  $B_y$  is convex for all  $x \in X$ . For example, given  $y = 0.5$ ,  $B_x = \{(x, y) \in \mathbb{R}^2 : y = 0.5, 1 \leq x \leq 2\}$ .



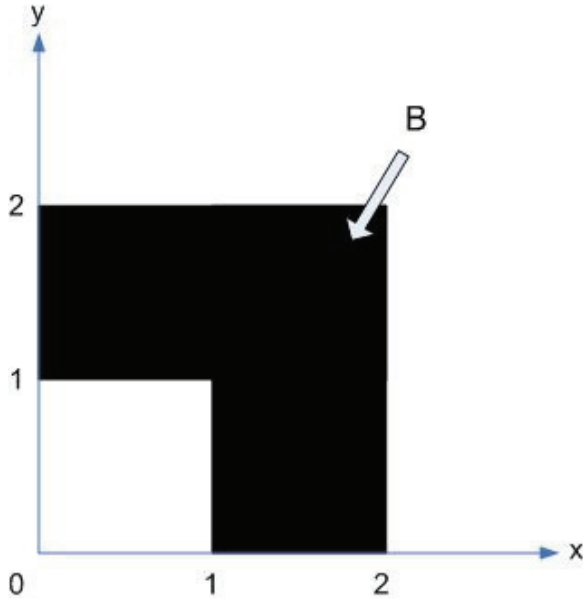


Fig. 11. An example of biconvex set.

**Definition 2:** A function  $f : B \rightarrow \mathbb{R}$  on a biconvex set  $B \subset X \times Y$  is a biconvex function on  $B$  if  $f_x(\cdot) := f(x, \cdot) : B_x \rightarrow \mathbb{R}$  is a convex function on  $B_x$  for every fixed  $x \in X$  and  $f_y(\cdot) := f(\cdot, y) : B_y \rightarrow \mathbb{R}$  is a convex function on  $B_y$  for every fixed  $y \in Y$  [10].

The set of convex functions is a sub-set of possible biconvex functions. A convex function must have a non-negative definite Hessian matrix, while a biconvex function might not. Luenberger [14] provides the example below:

$$f(x, y) = x^3 - x^2y + 2y^2 : x > 4, y \in [0, 10].$$

It is easy to check that the determinant of the Hessian matrix is  $4(-x^2 + 6x - 2y)$ , which is not always non-negative. This implies that the Hessian matrix is not positive semi-definite and thus  $f(x, y)$  is not convex. However, we have

$$\begin{aligned} \frac{\partial^2 f(x, y)}{\partial x^2} &= 6x - 2y > 0 \\ \frac{\partial^2 f(x, y)}{\partial y^2} &= 4 > 0 \end{aligned}$$

implying that  $f(x)$  is convex for all  $y \in Y$  and  $f(y)$  is convex for all  $x \in X$ . Thus,  $f(x, y)$  is biconvex. For this biconvex function, the global minimum is  $(x, y) = (4, 4)$ , and thus  $f_{\min}(x, y) = 32$ .

An optimization problem of the form:

$$\min \{f(x, y) : (x, y) \in B\}$$

is said to be a biconvex optimization problem if the feasible set  $B$  is convex on  $X \times Y$  and the objective function  $f$  is biconvex on  $B$ .

Distinct from convex optimization problems, biconvex problems may have a large number of local minima [10]. Regarding minimizing of a biconvex function  $f : X \times Y \rightarrow \mathbb{R}$  with a partial optimum  $(x^*, y^*) \in X \times Y$ , Wendell and Hurter [15] proved that  $f(x^*, y^*) \leq f(x, y) \forall x \in U(x^*) \forall y \in U(y^*)$ ,

where  $U(x^*)$  and  $U(y^*)$  respectively denote the set of optimal solutions with  $x = x^*$  and  $y = y^*$  to the following biconvex minimization problem:

$$\min \{f(x, y) : x \in X \subseteq \mathbb{R}, y \in Y \subseteq \mathbb{R}\}.$$

Based on this result, some methods and algorithms have been developed which exploit the convex substructures of a biconvex minimization problem to arrive at the global optimum, more efficiently than general global optimization methods. Floudas [16] and Goh [17] provide two such algorithms to solve for the global minimum of biconvex problems.

#### APPENDIX B PROOF OF LEMMA 1

Taking the second partial derivative of  $Q(\tilde{\varepsilon}\eta)$  with respect to  $\eta$  gives:

$$\frac{\partial^2 Q(\tilde{\varepsilon}\eta)}{\partial \eta^2} = \frac{\tilde{\varepsilon}^3 \eta}{\sqrt{2\pi}} e^{-(\tilde{\varepsilon}\eta)^2/2}.$$

Since  $Q(\cdot)$  is a monotonic decreasing function and  $Q(0) = 1/2$ , the condition that  $P_d = 1 - Q(\tilde{\varepsilon}\eta) > 1/2$  is equivalent to the condition that  $\tilde{\varepsilon}\eta > 0$ . Because  $\eta > 0$ , we have  $\tilde{\varepsilon} > 0$ , and this condition satisfies the positivity of  $\frac{\partial^2 Q(\tilde{\varepsilon}\eta)}{\partial \eta^2}$ . Thereby  $Q(\tilde{\varepsilon}\eta)$  is convex in  $\eta$ .  $\square$

#### APPENDIX C PROOF OF LEMMA 2

Taking the second partial derivative of the function with respect to  $\eta$  gives:

$$\begin{aligned} \frac{\partial^2}{\partial \eta^2} \left[ -Q\left(\frac{\sqrt{2}\tilde{\varepsilon} - \gamma}{\sqrt{2}(\gamma+1)}\eta\right) \right] \\ = -\frac{\eta}{\sqrt{2\pi}} \left(\frac{\sqrt{2}\tilde{\varepsilon} - \gamma}{\sqrt{2}(\gamma+1)}\right)^3 e^{-(\sqrt{2}\tilde{\varepsilon} - \gamma)^2 \eta^2 / [4(\gamma+1)^2]}. \end{aligned}$$

Similarly, because  $P_{fa} = 1 - Q\left(\frac{\sqrt{2}\tilde{\varepsilon} - \gamma}{\sqrt{2}(\gamma+1)}\eta\right) < 1/2$ , we have  $\frac{\sqrt{2}\tilde{\varepsilon} - \gamma}{\sqrt{2}(\gamma+1)} < 0$ . The positivity of the second partial derivative proves the convexity of  $-Q\left(\frac{\sqrt{2}\tilde{\varepsilon} - \gamma}{\sqrt{2}(\gamma+1)}\eta\right)$  in  $\eta$ .  $\square$

#### APPENDIX D PROOF OF LEMMA 3

Taking the second partial derivative of the function with respect to  $\eta$  gives:

$$\frac{\partial^2 \overline{T_{det}}(\tilde{\varepsilon}, \eta)}{\partial \eta^2} = \frac{\partial^2}{\partial \eta^2} \frac{M\tau}{B_c L} F(\tilde{\varepsilon}, \eta)^{-1} + \frac{\partial^2}{\partial \eta^2} \frac{M}{B_c L} \frac{\eta^2}{F(\tilde{\varepsilon}, \eta)}, \quad (23)$$

where  $F(\tilde{\varepsilon}, \eta) = 1 - Q(\tilde{\varepsilon}\eta) > 0$ .

Let  $G(\tilde{\varepsilon}, \eta) = \frac{\partial F(\tilde{\varepsilon}, \eta)}{\partial \eta} = \frac{\tilde{\varepsilon}}{\sqrt{2\pi}} e^{-\tilde{\varepsilon}^2 \eta^2 / 2}$ , and thus

$$\begin{aligned} \frac{\partial^2}{\partial \eta^2} \frac{M\tau}{B_c L} F(\tilde{\varepsilon}, \eta)^{-1} &= \frac{M\tau}{B_c L} F(\tilde{\varepsilon}, \eta)^{-3} G(\tilde{\varepsilon}, \eta) \cdot \\ &\quad [2G(\tilde{\varepsilon}, \eta) + F(\tilde{\varepsilon}, \eta)\tilde{\varepsilon}^2 \eta] \quad (24) \end{aligned}$$

$$\frac{\partial^2}{\partial \eta^2} \frac{M}{B_c L} \frac{\eta^2}{F(\tilde{\varepsilon}, \eta)} = \frac{2M}{B_c L} \frac{[F(\tilde{\varepsilon}, \eta) - \eta G(\tilde{\varepsilon}, \eta)]^2}{F(\tilde{\varepsilon}, \eta)^4} \quad (25)$$

As we have shown in Lemma 1 that  $\tilde{\varepsilon} > 0$ , hence  $G(\tilde{\varepsilon}, \eta) > 0$ . Because  $M$ ,  $\tau$ ,  $B_c$  and  $L$  are of positive values, (24) is

obviously positive and (25) is non-negative, leading to the positivity of (23). Thus  $\overline{T_{det}}(\tilde{\varepsilon}, \eta)$  is convex in  $\eta$ .  $\square$

#### APPENDIX E PROOF OF THEOREM 2

Firstly we will show that the problem (19) is convex in  $\eta$  under constraints (20)-(22), given  $\varepsilon = \tilde{\varepsilon} \in \mathbb{E}$ , if  $P_d > 1/2$  and  $P_{fa} < 1/2$ . Lemmas 1-2 provide that  $Q(\tilde{\varepsilon}, \eta)$ ,  $-Q(\frac{\sqrt{2}\tilde{\varepsilon}-\gamma}{\sqrt{2}(\gamma+1)}\eta)$  and  $\frac{M\tilde{\eta}^2}{B_c L}F(\tilde{\varepsilon}, \eta)^{-1}$  are all convex in  $\eta$ . We have also proved that  $\frac{\partial^2 \overline{T_{det}}}{\partial \eta^2} \geq 0$ ,  $\frac{\partial^2 P_{fa}}{\partial \eta^2} \geq 0$  and  $G(\frac{\sqrt{2}\tilde{\varepsilon}-\gamma}{\sqrt{2}(\gamma+1)}, \eta) \leq 0$  when  $P_d > 1/2 > P_{fa}$ . By taking the second order partial derivative of the objective function, we have:

$$\begin{aligned} \frac{\partial^2 K}{\partial \eta^2} &= \frac{\phi - \varphi P_{fa}(\tilde{\varepsilon}, \eta)}{T} \frac{\partial^2 \overline{T_{det}}(\tilde{\varepsilon}, \eta)}{\partial \eta^2} \\ &\quad + (1 - \frac{\overline{T_{det}}(\tilde{\varepsilon}, \eta)}{T}) \varphi \frac{\partial^2 P_{fa}(\tilde{\varepsilon}, \eta)}{\partial \eta^2} \\ &\quad - \frac{2\varphi}{T} G(\frac{\sqrt{2}\tilde{\varepsilon}-\gamma}{\sqrt{2}(\gamma+1)}, \eta) \frac{M\eta}{B_c L} \frac{2F(\tilde{\varepsilon}, \eta) - \eta G(\tilde{\varepsilon}, \eta)}{F(\tilde{\varepsilon}, \eta)^2} \\ &\geq -\frac{2\varphi M\eta}{T B_c L} F(\tilde{\varepsilon}, \eta)^{-2} G(\frac{\sqrt{2}\tilde{\varepsilon}-\gamma}{\sqrt{2}(\gamma+1)}, \eta) \cdot \\ &\quad [2F(\tilde{\varepsilon}, \eta) - \eta G(\tilde{\varepsilon}, \eta)]. \end{aligned} \quad (26)$$

Let  $I(x) = xe^{-x^2/2}$ , and then  $\frac{\partial I(x)}{\partial x} = (1-x^2)e^{-x^2/2}$ . It is easy to check that  $I(x)_{max} = I(1) \approx 0.6$  for  $x \in \mathbb{R}$ . Because  $P_d > 1/2$ , we have:

$$\begin{aligned} 2F(\tilde{\varepsilon}, \eta) - \eta G(\tilde{\varepsilon}, \eta) &= 2P_d(\tilde{\varepsilon}, \eta) - \frac{\tilde{\varepsilon}\eta}{\sqrt{2\pi}} e^{-\tilde{\varepsilon}^2 \eta^2 / 2} \\ &\geq 2P_d(\tilde{\varepsilon}, \eta) - \frac{I(x)_{max}}{\sqrt{2\pi}} \\ &> 0. \end{aligned} \quad (27)$$

Combining (26) and (27), we can conclude that  $\frac{\partial^2 K}{\partial \eta^2} \geq 0$ . Therefore, the convexity in  $\eta$  under the given constraints is proved.  $\square$

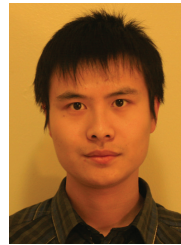
Next, we show that the optimization problem (19) is convex in  $\varepsilon$  under constraints (20)-(22), given  $\eta = \tilde{\eta} \in \mathbb{H}$ , if  $P_d > 1/2$  and  $P_{fa} < 1/2$ . Similarly we know that  $Q(\varepsilon, \tilde{\eta})$ ,  $-Q(\frac{\sqrt{2}\varepsilon-\gamma}{\sqrt{2}(\gamma+1)}\tilde{\eta})$  and  $\frac{M\tilde{\eta}^2}{B_c L}F(\varepsilon, \tilde{\eta})^{-1}$  are all convex in  $\varepsilon$ , and  $\frac{\partial^2 \overline{T_{det}}}{\partial \varepsilon^2} \geq 0$ ,  $\frac{\partial^2 P_{fa}}{\partial \varepsilon^2} \geq 0$ . The convexity in  $\varepsilon$  is proved, by taking second partial derivative of the objective function gives:

$$\begin{aligned} \frac{\partial^2 K}{\partial \varepsilon^2} &= \frac{2\varphi}{T} \sqrt{2} G(\tilde{\eta}, \frac{\sqrt{2}\varepsilon-\gamma}{\sqrt{2}(\gamma+1)}) \frac{M\tilde{\eta}^2}{B_c L} G(\tilde{\eta}, \varepsilon) F(\varepsilon, \tilde{\eta})^{-2} \\ &\quad + \frac{\phi - \varphi P_{fa}(\varepsilon, \tilde{\eta})}{T} \frac{\partial^2 \overline{T_{det}}(\varepsilon, \tilde{\eta})}{\partial \varepsilon^2} + (1 - \frac{\overline{T_{det}}(\varepsilon, \tilde{\eta})}{T}) \varphi \frac{\partial^2 P_{fa}(\varepsilon, \tilde{\eta})}{\partial \varepsilon^2} \\ &\geq 0. \end{aligned} \quad (28)$$

Since  $\varepsilon$  and  $\eta$  are independent, Theorem 2 is proved.  $\square$

#### REFERENCES

- [1] M. Nekovee, "Quantifying the availability of TV white spaces for cognitive radio operation in the UK," in *Proc. 2009 IEEE ICC*, pp. 1-5.
- [2] J. Mitola III, "Cognitive radio: an integrated agent architecture for software defined radio," Ph.D. thesis, Swedish Royal Institute of Technology, 2000.
- [3] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326-1337, Apr. 2008.
- [4] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128-1140, Mar. 2009.
- [5] L. Luo, N. M. Neihart, S. Roy, and D. J. Allstot, "A two-stage sensing technique for dynamic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3028-3037, June 2009.
- [6] W.-Y. Lee and I. F. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, Oct. 2008.
- [7] S. Huang, X. Liu, and Z. Ding, "Optimal sensing-transmission structure for dynamic spectrum access," in *Proc. 2009 IEEE INFOCOM*, pp. 2295-2303.
- [8] Q. Peng, P. C. Cosman, and L. B. Milstein, "Optimal sensing disruption for a cognitive radio adversary," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1801-1810, May 2010.
- [9] L. Luo, C. Ghosh, and S. Roy, "Joint optimization of spectrum sensing for cognitive radio networks," in *Proc. 2010 IEEE GLOBECOM*, pp. 1-5.
- [10] J. Gorski, F. Pfeuffer, and K. Klamroth, "Biconvex sets and optimization with biconvex functions: a survey and extensions," *Math. Meth. Oper. Res.*, vol. 66, no. 3, pp. 373-407, Dec. 2007.
- [11] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, May 2009.
- [12] L. Luo and S. Roy, "Modeling and analysis of detection time trade-offs for channel searching in cognitive radio networks," to be published.
- [13] R. Beek, D. Leenaerts, and G. van der Weide, "A fast-hopping single-PLL 3-band UWB synthesizer in 0.25 $\mu$ m SiGe BiCMOS," in *Proc. 2005 European Solid-State Circuits Conf.*, pp. 173-176.
- [14] D. Luenberger, *Linear and Nonlinear Programming*, 2nd edition. Addison-Wesley, 1984.
- [15] R. E. Wendell and A. P. Hurter, "Minimization of a non-separable objective function subject to disjoint constraints," *Operations Research*, vol. 24, no. 4, pp. 643-657, 1976.
- [16] C. A. Floudas and V. Visweswaran, "A global optimization algorithm (GOP) for certain classes of nonconvex NLPs—I: theory," *Computer Chem. Eng.*, vol. 14, no. 12, July 1990.
- [17] K. Goh, M. Safonov, and G. Papavassilopoulos, "Global optimization for the biaffine matrix inequality problem," *J. Global Optimization*, vol. 7, no. 4, pp. 365-380, Dec. 1995.
- [18] J. D. Leeuw and G. Michailidis, "Block relaxation algorithms in statistics," *J. Comput. Graphical Stat.*, vol. 9, pp. 26-31, 2000.



**Ling Luo** received the B.S and M.S. degrees from Zhejiang University, Hangzhou, China, in 2003 and 2006, respectively. He received the Ph.D degree from the Department of Electrical Engineering, University of Washington, Seattle, in 2010. At present, he is working as an engineer at the Electric Power Research Institute (SMEPC). His current research interests focus on smart grid, load forecasting, and power quality.