Robust Threshold Design for Cooperative Sensing in Cognitive Radio Networks

Shimin Gong, Ping Wang, Jianwei Huang

***Abstract*—The successful coexistence of cognitive radio systems and licensed systems requires the secondary users to have the capability of sensing and keeping track of primary transmissions. While existing spectrum sensing methods usually assume known distributions of the primary signals, such an assumption is often not true in practice. As a result, applying existing sensing methods directly will often lead to unreliable detection performance in practical networks. In this paper, we try to improve the sensing performance under the distribution uncertainty of primary signals. We formulate the optimal sensing design as a robust optimization problem, and propose an iterative algorithm to determine the optimal decision threshold for each user. Extensive simulations demonstrate the effectiveness of our proposed algorithm.**

***Index Terms*—Cognitive radio, spectrum sensing, robust optimization, distribution uncertainty.**

I. Introduction

Traditionally, wireless spectrum has been statically allocated to license holders for exclusive use, and the spectrum available for new applications has become scarce and pricey. Cognitive Radio (CR) [1] has been proposed to improve spectrum utilization by allowing secondary unlicensed users (SUs) to access the licensed spectrum bands that are temporarily unoccupied by primary licensed users (PUs). There have been extensive studies on the coexistence of SUs and PUs (e.g., [2], [3]). In this paper, we focus on the robust design of spectrum sensing, which is one of the basis for spectrum access and sharing.

Spectrum sensing aims to determine the presence of PUs based on channel characteristics. Spectrum sensing is usually formulated as a hypothesis testing problem where an SU makes a decision on whether the PU is present by comparing the measured channel samples with a pre-designed threshold. The detection performance is measured by the receiver operating characteristic (ROC), which denotes the tradeoff between detection probability (i.e., probability of sensing a busy channel as busy) and false alarm probability (i.e., probability of sensing an idle channel as busy). The detection probability is largely studied using a Q-function (tail probability of the standard Gaussian distribution) assuming a known Gaussian distribution for the primary signal [4], [5]. Unfortunately, precise prior knowledge regarding the primary signals' probability distribution function is a very strong assumption that often does not hold in practice.

This work is partially supported in part by the General Research Fund (Project Number 412511) established under the University Grant Committee of the Hong Kong Special Administrative Region, China.

Shimin Gong and Ping Wang are with Nanyang Technological University, Email: {gong0012,wangping}@ntu.edu.sg; Jianwei Huang is with The Chinese University of Hong Kong, Email: jwhuang@ie.cuhk.edu.hk.

In this paper, we will study robust spectrum sensing under distribution uncertainty of the primary signals. Specifically, we consider the case where the statistical features (i.e., mean and variance) of the primary signals are contaminated by estimation errors, and the actual probability density function (pdf) does not match what is often assumed in literature. To capture the distribution uncertainty, we define an uncertainty set that contains all possible signal distributions whose mean and variance are close to a nominal one, and then formulate the sensing design into a robust optimization problem.

Few prior results considered the issues of uncertainty and robustness in cognitive radio networks [6], [7], assuming of a known probability distribution for either noise or primary signals. As far as we know, our paper is the first work that tackles the uncertainty issue from a viewpoint of robust optimization. Specifically, we formulate the robust design for cooperative spectrum sensing as a max-min robust optimization problem, and propose an iterative algorithm that decomposes the problem into a series of semidefinite programs, which can be solved easily by conic optimization tools.

The rest of the paper is organized as follows. Section II describes the basic model, then Section III presents the robust design in cooperative spectrum sensing. We show some numerical results in Section IV and conclude in Section V.

II. System Model

A. Energy Detection Structure and Decision Function

We consider a CR network with one primary channel and N SUs. Let $\mathcal{N} = \{1, ..., N\}$ denote the set of all SUs. Each SU is equipped with a local energy detector, and all SUs work together to perform cooperative spectrum sensing [5], [8], [9]. We consider a hard combination scheme [4] at the fusion center and define a decision function as $h(\boldsymbol{\lambda}, \boldsymbol{\xi}, k) = I\left(\sum_{i=1}^{N} I(\xi_i - \lambda_i \geq 0) \geq k\right)$, where $\boldsymbol{\xi} = \{\xi_i\}_{i=1}^{N}$ and $\boldsymbol{\lambda} = \{\lambda_i\}_{i=1}^{N}$ denote the received signal and decision threshold at each user, respectively. Indicator $I(A)$ equals 1 if event A is true and 0 otherwise. Here k specifies the combining rule, e.g., $k = 1$ is the OR-rule and $k = \lceil N/2 \rceil$ represents the majority-rule. When $h(\boldsymbol{\lambda}, \boldsymbol{\xi}, k) = 1$, the fusion center reports the presence of the primary user (which is denoted as hypothesis H_1), otherwise reports an idle channel (which is denoted as hypothesis H_0).

In our paper, we take $k = 1$ (OR-rule) and denote $h(\boldsymbol{\xi}, \boldsymbol{\lambda}, 1)$ as $h(\boldsymbol{\xi}, \boldsymbol{\lambda})$ for simplicity. Let $\boldsymbol{X} = [X_1, \ldots, X_N]$ denote the signal received by N SUs, and let $S \subset \mathcal{R}^N$ denote the feasible set of $\boldsymbol{X}$. Thus the joint pdf for $\boldsymbol{X}$ under hypothesis H_1 (H_0, respectively) is given by $f^1_{\boldsymbol{X}}(\boldsymbol{\xi})$ ($f^0_{\boldsymbol{X}}(\boldsymbol{\xi})$, respectively). Finally, the detection and false alarm probabilities can be represented as $Q_d = \int_{\xi\in S} h(\xi,\lambda) f_X^1(\xi)\, d\xi$ and $Q_f = \int_{\xi\in S} h(\xi,\lambda) f_X^0(\xi)\, d\xi$, respectively.

B. Distribution Uncertainty and Robust Detector

In practice, the signals transmitted by a PU are hard to know precisely in advance by an SU. But we can obtain the primary signals' long-time averages in terms of the mean $\boldsymbol{\mu}$ and variance Σ, which are regarded as the *nominal* statistics. Then we estimate the range of short-time fluctuations empirically according to a specific spectrum environment. It is reasonable to assume that the instant fluctuations of mean and variance should be within small ranges to their nominal values. Therefore, we define an uncertainty set for distribution f_X^1 using the nominal values of mean and variance in a similar way as [10],

$$\mathcal{U} = \left\{ f_X^1 \;\middle|\; \begin{array}{c} \mathcal{E}\mathcal{E}^T \le \gamma_1 \Sigma, \\ \int_{\xi\in S} (\xi - \mu)(\xi - \mu)^T f_X^1(\xi)\, d\xi \le \gamma_2 \Sigma, \\ \int_{\xi\in S} f_X^1(\xi)\, d\xi = 1. \end{array} \right\}, \tag{1}$$

where $\mathcal{E} = \int_{\xi\in S} \xi f_X^1(\xi)\, d\xi - \mu$ denotes the deviation of instant mean from its nominal value μ. Parameters γ_1 and γ_2 regulate the uncertainty size and provide a way to evaluate the confidence in nominal μ and Σ, respectively.

When there is no transmission from PUs, the background noise is also time varying [6], [7]. However, it only affects the false alarm probability but not the detection probability when decision threshold is fixed[1]. In this paper we focus on the choice of decision threshold and its impact on the detection probability, thus we will assume that noise follows a deterministic Gaussian distribution f_0, i.e., $f_X^0(\xi) = f_0(\xi)$. The distribution parameters of $f_0(\xi)$ can be obtained through field measurement or empirical data.

III. Robust Design of Cooperative Sensing

Our goal is to design a decision threshold λ_i for each SU i that maximizes the worst-case detection probability of the collaborative sensing, i.e., $\max_\lambda \min_{f_X^1(\xi)\in\mathcal{U}} \int_{\xi\in S} h(\xi,\lambda) f_X^1(\xi)\, d\xi$, subject to a false alarm probability requirement, i.e., $\int_{\xi\in S} h(\lambda,\xi) f_0(\xi)\, d\xi \le \alpha$. This constrained optimization is hard to analyze directly, since it involves the manipulation of function integrations (i.e., constraints on mean and variance). To ease the analysis, we propose an equivalent transformation that eliminates the function integrations associated with the distribution uncertainty, and turns the max-min problem into a maximization problem:

Theorem 1. *The proposed max-min problem is equivalent to (2a)-(2e) as follows,*

$$\max_{\lambda,Z,Q,r} \quad (\mu\mu^T - \gamma_2\Sigma) \otimes Q - \Sigma \otimes P + 2\mu^T p - \gamma_1 s - r \tag{2a}$$

$$s.t. \quad 1 + r - (p + Q\mathbf{u})^T Q^{-1} (p + Q\mu) \ge 0 \tag{2b}$$

$$\xi^T Q \xi - 2\xi^T (p + Q\mu) + r \ge 0, \quad \forall \xi \le \lambda \tag{2c}$$

$$Q \ge 0, \quad Z = \begin{bmatrix} P & p \\ p^T & s \end{bmatrix} \ge 0 \tag{2d}$$

$$\int_{\xi\in S} h(\lambda,\xi) f_0(\xi)\, d\xi \le \alpha, \tag{2e}$$

where Z and Q are symmetric positive semi-definite matrices.

Theorem 1 can be proved using the Lagrangian duality theory. We omit the details due to space limit. Here $\otimes$ denotes the Frobenius inner product[2]. Variables Z, Q and r are penalties to the constrained optimization problem if corresponding moment constraints in the original max-min problem are violated. The main difficulty of solving (2a)-(2e) is that constraint (2e) defines a non-convex set. This motivates us to solve (2a)-(2e) using an iterative algorithm. For each iteration, we will separate (2e) from the solvable structure (2a)-(2d), i.e., we solve sub-problem (2a)-(2d) with a *fixed* decision threshold vector λ. Then we will iteratively improve the objective function value by adjusting decision threshold in the feasible set $\mathcal{D} = \{\lambda \mid \int_{\xi\in S} h(\lambda,\xi) f_0(\xi)\, d\xi \le \alpha\}$ defined by (2e). The algorithm continues until a stable point is reached. Although we can not prove that such a stable point is always the global optimal solution, simulation results show that such result leads to a much robust performance comparing with the algorithms where the uncertainty is not explicitly taken into consideration.

A. Initial Decision Threshold (IDT)

In order to solve problem (2a)-(2d), we need to first set an IDT in the feasible set (2e). Choosing a conservative (i.e., too large) IDT will start the iterative algorithm in an interior point of the set (2e), leading to the false alarm probability strictly below α and thus suboptimal detection probability. Hence we prove the following proposition that will assist our choice for a good initial decision threshold λ.

Proposition 1. *An optimal decision threshold λ^* to the problem (2a)-(2e) satisfies (2e) with equality, that is*

$$\int_{\xi\le\lambda^*} f_0(\xi)\, d\xi = 1 - \alpha. \tag{3}$$

Note that (3) is a necessary but not a sufficient condition for the optimality of Problem (2a)-(2e).

Proof of Proposition 1 is given in Appendix A. It implies that an IDT can be chosen from set $\mathcal{D}_c \triangleq \{\lambda \mid \int_{\xi\le\lambda} f_0(\xi)\, d\xi = 1 - \alpha\}$. Moreover, we can guarantee the feasibility of (2e) by updating the decision threshold λ within the same set $\mathcal{D}_c$ during every iteration. In this way we separate constraint (2e) from the problem (2a)-(2d). Note that set $\mathcal{D}_c$ is not empty, since we can always find a decision threshold that balances equation (3).

For convenience, we choose an equal IDT[3] that assigns the same value for all SUs' decision thresholds, i.e., $\lambda_1 = \cdots = \lambda_N = \lambda_0$. We further assume that the noise at each cooperative SU is i.i.d., then (3) turns into $\int_{\xi\le\lambda_0} f_0(\xi)\, d\xi = (1-\alpha)^{1/N}$. Since we have assumed the noise pdf $f_0(\xi)$ to be a known Gaussian distribution with cdf $F_0(\lambda_0) = \int_{\xi\le\lambda_0} f_0(\xi)\, d\xi$, then the equal IDT will set $\lambda_0 = F_0^{-1}\left((1-\alpha)^{1/N}\right)$ to all SUs. The values

[1] Given the decision threshold, $Q_f = \int_{\xi\in S} h(\xi,\lambda) f_X^0(\xi)\, d\xi$ only relates to noise distribution f_X^0, while $Q_d = \int_{\xi\in S} h(\xi,\lambda) f_X^1(\xi)\, d\xi$ is independent of noise distribution.

[2] The Frobenius inner product is the component-wise inner product of two matrices, e.g., $A \otimes B = trace(AB^T)$, A and B are two matrices.

[3] The IDT can be set in different ways. However, an equal IDT is easy to implement and does not affect the algorithm convergence, which will be demonstrated in Section IV through simulations.

of IDT will be adjusted iteratively in the algorithm, such that objective (2a) is improved and the new threshold is still within set $\mathcal{D}_c$. In other words, such adjustments always satisfy the α-Preserving condition.

α-Preserving Condition: *Consider that two (and only two) SUs change their decision thresholds by Δ_i and Δ_j, respectively. If $\Delta_j = F_0^{-1}\left(F_0(\lambda_j)\frac{F_0(\lambda_i)}{F_0(\lambda_i+\Delta_i)}\right)-\lambda_j$, then the new threshold vector preserves the same false alarm probability. We call it an α-Preserving adjustment.*

Proof: In order to maintain the same false alarm probability, the decision threshold adjustments of any two different users i and j should satisfy $F_0(\lambda_i)F_0(\lambda_j) = F_0(\lambda_i+\Delta_i)F_0(\lambda_j+\Delta_j)$, where Δ_i and Δ_j are changes of λ_i and λ_j, respectively. Since $F_0(\cdot)$ denotes a known cdf function of the background noise, we get the result directly. ■

B. Solving Sub-problem (2a)-(2d) with Fixed Thresholds

Sub-problem (2a)-(2d) can be solved in polynomial time if and only if a separation problem [10] can be solved in polynomial time. In the separate problem, we will check whether (2c) is satisfied for all $\boldsymbol{\xi} \leq \boldsymbol{\lambda}$ when (Q, p, r) is fixed, if not we need to provide a violated case. However, the separation problem is proved to be NP-hard in [11]. The main difficulty is that we need to consider a free variable $\boldsymbol{\xi}$ when judging the non-negativity of the quadratic polynomial in LHS of (2c). The proposed α-Preserving condition provides a way to eliminate $\boldsymbol{\xi}$ and transform the separation problem (2c) into the following form:

Proposition 2. *α-Preserving condition ensures that (2c) is equivalent to $\boldsymbol{\lambda}^T Q\boldsymbol{\lambda} - 2\boldsymbol{\lambda}^T(p + Q\boldsymbol{\mu}) + r \geq 0$ if $\boldsymbol{\lambda}$ is optimal to problem (2a)-(2e).*

The proof is given in Appendix B. Proposition 2 removes $\boldsymbol{\xi}$ from the sub-problem, and thus reduces (2c) to a linear constraint w.r.t. (Q, p, r). Furthermore, we reformulate (2a)-(2d) into a semidefinite program, which can be efficiently solved by the interior-point method [12].

Theorem 2. *Sub-problem (2a)-(2d) has an equivalent form in (4a)-(4e) as follows:*

$$\begin{aligned}
\max_{\boldsymbol{\lambda},Q,p,q,r,z} \quad & z && \text{(4a)}\\
s.t. \quad & \boldsymbol{\lambda}^T Q\boldsymbol{\lambda} - 2\boldsymbol{\lambda}^T(p + Q\boldsymbol{\mu}) + r \geq 0, && \text{(4b)}\\
& (\boldsymbol{\mu}\boldsymbol{\mu}^T - \gamma_2\Sigma)\otimes Q + 2\boldsymbol{\mu}^T p \geq z + r + q, && \text{(4c)}\\
& q \geq 2\sqrt{\gamma_1}\|Lp\|_2, && \text{(4d)}\\
& \begin{bmatrix} r+1 & (p+Q\mu)^T \\ (p+Q\mu) & Q \end{bmatrix} \geq 0. && \text{(4e)}
\end{aligned}$$

Proof: Let $z = (\boldsymbol{\mu}\boldsymbol{\mu}^T - \gamma_2\Sigma)\otimes Q - \Sigma\otimes P + 2\boldsymbol{\mu}^T p - \gamma_1 s - r$, and $q \geq 2\sqrt{\gamma_1 p^T\Sigma p} = 2\sqrt{\gamma_1}\|Lp\|_2$, where L is a triangular matrix denoting the Cholesky decomposition of Σ, i.e., $\Sigma = L^T L$. Since $Z = \begin{bmatrix} P & p \\ p^T & s \end{bmatrix} \geq 0$, i.e., $P \geq \frac{1}{s}pp^T$, we have

$$\begin{aligned}
z &\leq (\boldsymbol{\mu}\boldsymbol{\mu}^T - \gamma_2\Sigma)\otimes Q - \frac{1}{s}p^T\Sigma p + 2\boldsymbol{\mu}^T p - \gamma_1 s - r\\
&\leq (\boldsymbol{\mu}\boldsymbol{\mu}^T - \gamma_2\Sigma)\otimes Q + 2\boldsymbol{\mu}^T p - r - 2\sqrt{\gamma_1}\|Lp\|_2\\
&\leq (\boldsymbol{\mu}\boldsymbol{\mu}^T - \gamma_2\Sigma)\otimes Q + 2\boldsymbol{\mu}^T p - r - q.
\end{aligned}$$

Therefore, the objective in (2a) can be written as (4a) with a linear constraint (4c) and a quadratic constraint (4d). Constraint (2b) is written in its matrix form (4e), and (2c) is equivalent to (4b) according to Proposition 2. ■

When $\boldsymbol{\lambda}$ is fixed, (4a)-(4e) is a conic optimization problem in the standard dual form, and can be easily solved by a conic optimization package such as SeDuMi [13]. We denote its solution by $y = [vec(Q), p^T, q, r, z]^T$, where $vec(Q)$ transforms the variable matrix Q into a row vector.

C. Update Decision Threshold

Given the solution $y = [vec(Q), p^T, q, r, z]^T$ from a semidefinite program, we show how the decision thresholds $\boldsymbol{\lambda}$ can be updated iteratively such that the objective (4a) is improved.

Proposition 3. *Given the parameters Q, p, q, and r in (4a)-(4e), the objective z will be improved if $\boldsymbol{\lambda}(t+1) = \boldsymbol{\lambda}(t)+\rho(t)G(t)$, where $\rho(t) > 0$ is a small step size and $G(t) = Q\cdot\boldsymbol{\lambda}(t) - Q\cdot\boldsymbol{\mu} - p$.*

Proof: From (4b) and (4c), we have $z \leq g(\boldsymbol{\lambda})$ and $g(\boldsymbol{\lambda}) \triangleq \boldsymbol{\lambda}^T Q\boldsymbol{\lambda} - 2\boldsymbol{\lambda}^T(p+Q\boldsymbol{\mu}) + (\boldsymbol{\mu}\boldsymbol{\mu}^T - \gamma_2\Sigma)\otimes Q + 2\mu^T p - q$. The last term of $g(\boldsymbol{\lambda})$ does not depend on $\boldsymbol{\lambda}$ and can be viewed as a constant with respect to $\boldsymbol{\lambda}$. In order to find another $\boldsymbol{\lambda}$ that improves the objective z, we can update $\boldsymbol{\lambda}$ in its gradient direction, i.e., $\boldsymbol{\lambda}(t+1) = \boldsymbol{\lambda}(t) + \rho(t)G(t)$. Positive $\rho(t)$ is a step size, which should be small enough to avoid overshot, and $G(t)$ is the gradient given by $Q\cdot\boldsymbol{\lambda}(t) - Q\cdot\boldsymbol{\mu} - p$. ■

We assume that user can accept a small error θ in their final decision threshold, thus we further require $\rho(t) \leq \theta/\max_i G_i(t)$[4]. However, from the proof of Proposition 2 we know that $\boldsymbol{\lambda} \leq \boldsymbol{\mu} + Q^{-1}p$, which implies a non-positive gradient $G(t) \leq 0$. If the decision threshold $\boldsymbol{\lambda}$ is adjusted as in Proposition 3, then it will decrease in every iteration. This is intuitive since lowering the decision thresholds can increase detection probability in hypothesis testing, but the false alarm probability cannot be maintained at the prescribed level. This motivates us to increase the decision threshold of the user i^- who corresponds to the minimum term in $G(t)$. We choose this term because we prefer minimum adjustment to the decision threshold, and this term impacts the objective (4a) mostly if the same adjustment is applied to all SUs. Meanwhile we apply an α-Preserving adjustment to the user i^+ who corresponds to the largest term in $G(t)$, such that the false alarm probability is maintained at the same level. As shown in Algorithm 1, parameter Δ_{i^-} denotes the difference between the maximum and minimum terms in $G(t)$, and is used to control the convergence. The algorithm terminates when Δ_{i^-} is small enough or the number of iterations exceeds a predefined threshold Γ.

IV. Numerical and Simulation Results

In the simulation, we consider $N = 3$ SUs and they receive primary signals with different nominal mean values (denoted by the vector $\boldsymbol{\mu}$). We set $\boldsymbol{\mu} = [3.7,\ 4.0,\ 4.3]$ and $\alpha = 0.1$. The signals received at the SUs are independent with each other with variances $\Sigma = 2I_N$, where $\boldsymbol{I}_N$ is an identity matrix with

[4]We use G_i to indicate the i-th terms in vector G.

Algorithm 1 Decision Threshold Adjustment Algorithm

initialization:
  Set stopping criterion $\Gamma = 200$ and $\epsilon = 10^{-4}$
  Set initial $\lambda_i(t) = F_0^{-1}\left((1-\alpha)^{1/N}\right)$ for all SUs
while $\Delta_{i^-} \geq \epsilon$ and $t \leq \Gamma$ **do**
  Solve sub-problem (4a)-(4e) with $\lambda(t)$ fixed and return the optimizer $y = [vec(Q), p^T, q, r, z]$
  Calculate direction $G(t)$ and step size $\rho(t)$
  Set $i^- = \arg\min_i G_i(t)$, and $i^+ = \arg\max_{i \neq i^-} \lambda_i(t)$
  Update $\lambda_i(t+1) = \lambda_i(t) + \Delta_i$, $i \in \{i^-, i^+\}$ where
    $\Delta_{i^-} = \max_i G_i(t) - \min_j G_j(t)$
    $\Delta_{i^+} = F_0^{-1}\left(F_0(\lambda_{i^+}(t))\frac{F_0(\lambda_{i^-}(t))}{F_0(\lambda_{i^-}(t)+\Delta_{i^-})}\right)$
  Set $t = t + 1$
end while

size N. The distribution uncertainty parameters in set (1) are $\gamma_1 = 0.02$ and $\gamma_2 = 1.2$.

A. Comparison with an Existing Method

First, we compare the performance of our proposed robust sensing algorithm with an existing method. In the existing scheme (denoted as Scheme E in Figure 1), SUs will optimize the sensing thresholds assuming that the primary received signals follow Gaussian distribution $\mathcal{G}_0$ with mean $\boldsymbol{\mu} = [3.7,\ 4.0,\ 4.3]$. However, the actually distribution may follow either $\mathcal{G}_0$ or $\mathcal{G}_1$, where $\mathcal{G}_1$ is another Gaussian distribution with the mean $\boldsymbol{\mu} = [3.0,\ 3.5,\ 4.0]$. For our proposed scheme (i.e., Scheme P in Figure 1), we consider such distribution uncertainty. As shown in Figure 1, Scheme E works very well when the actual distribution of primary signals is $\mathcal{G}_0$ (and thus matches with the assumption). However, when the actual distribution is $\mathcal{G}_1$ (and thus there is a mismatch), Scheme E's performance degrades significantly. While for Scheme P, though the detection performance is not as good as that of Scheme E in the matched case, it is more robust and has a much better performance when the distribution is mismatched. This shows that our proposed algorithm achieves a good tradeoff between robustness and performance. In fact, the optimization under our Scheme P is performed with respect to the uncertainty set in (1), instead on a particular distribution. In the following, we further investigate various properties of Scheme P.

B. Impact of the Initial Decision Threshold (IDT)

In Section III-A, we have assumed that our proposed algorithm starts with the equal IDT. Here we show that such an assumption will not affect the performance and convergence of the algorithm. We examine three cases with different IDTs: $\lambda_0^e = [1.82, 1.82, 1.82]$, $\lambda_0^1 = [2.22, 1.82, 1.60]$, and $\lambda_0^2 = [2.62, 1.82, 1.52]$. Note that the false alarm probability is kept at the same level of 0.1 for all cases. Figure 2 plots the dynamics of detection probability when IDTs are set to λ_0^e, λ_0^1 and λ_0^2, respectively. Though the initial detection probabilities are not the same, the differences due to different initializations quickly diminish as the iteration increases. In Figure 3, we plot the threshold dynamics of all SUs for two cases, i.e., λ_0^e and λ_0^2. Here λ_j^i denotes threshold dynamics for user j when the IDT is set to λ_0^i. It is clear that the eventual decision thresholds of all SUs do not depend on the initializations. From Figure 3 we also can see that the final decision thresholds follow an ordering, i.e., $\lambda_3^e > \lambda_2^e > \lambda_1^e$ (and $\lambda_3^2 > \lambda_2^2 > \lambda_1^2$), which is consistent with our assumption of $\mu_3 > \mu_2 > \mu_1$. This

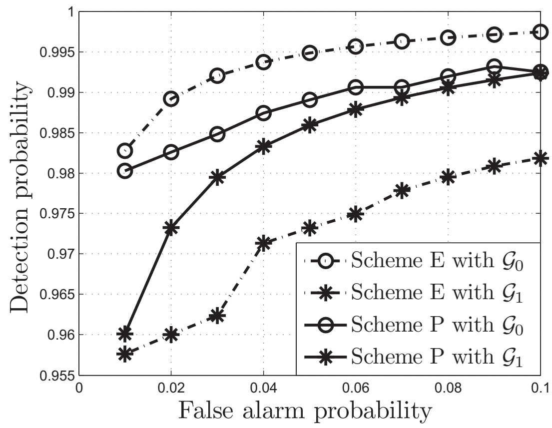

Fig. 1: Comparison between the Schemes E and P.

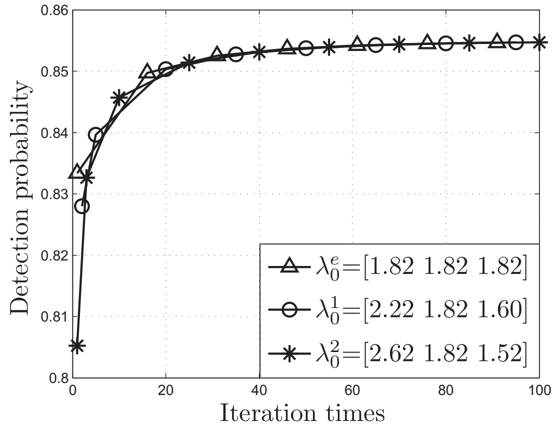

Fig. 2: Detection probabilities converge with different IDTs.

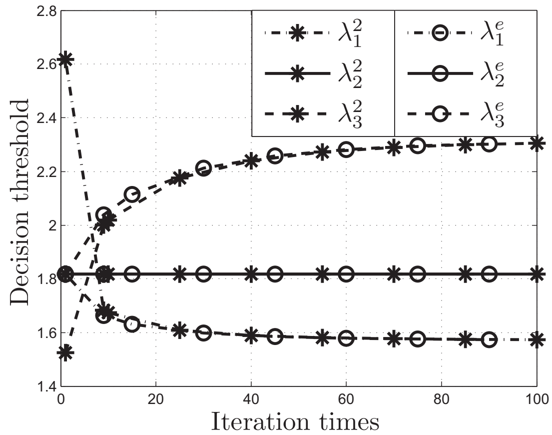

Fig. 3: Decision thresholds converge with different IDTs.

observation implies that an SU would have a higher decision threshold if the received primary signal is stronger.

V. Conclusion

In this paper, we consider the robust design of cooperative sensing in cognitive radio systems. As solving the constrained robust optimization problem directly is very difficult, we propose an approximate algorithm that separates the false alarm probability constraint from the rest of the problem. More specifically, the algorithm starts from an initial choice of decision thresholds that satisfies the false alarm probability constraint, and solve the rest of the problem which is in the form of semidefinite program. Then during each iteration, the algorithm searches new decision thresholds in order to improve the system objective. The significance of our work is that we incorporate the distribution uncertainty into the formulation so that our solution is more robust and can improve the worst-case detection performance.

Appendix

A. Proof of Proposition 1

The feasible set for $\boldsymbol{\lambda}$ defined by (2e) can be denoted as $\mathcal{D} = \{\boldsymbol{\lambda}_c + \boldsymbol{o} \,|\, \boldsymbol{o} \geq 0, \, \boldsymbol{\lambda}_c \in \mathcal{D}_c\}$. We notice that the values of $\boldsymbol{\lambda}$ influence the objective (2a) indirectly by interacting with the dual variables Q, p and r through constraint (2c). Now we define a set depending on $\boldsymbol{\lambda}$ as follows:

$$\Omega(\boldsymbol{\lambda}) = \left\{(Q, p, r) \,|\, \boldsymbol{\xi}^T Q \boldsymbol{\xi} - 2\boldsymbol{\xi}^T (p + Q\boldsymbol{\mu}) + r \geq 0, \quad \forall \boldsymbol{\xi} \leq \boldsymbol{\lambda}\right\}.$$

Before further analysis, we have the following results: A decision threshold $\boldsymbol{\lambda}^*$ is optimal to problem (2a)-(2e) if $\Omega(\boldsymbol{\lambda}^*) \supseteq \Omega(\boldsymbol{\lambda})$, $\forall \boldsymbol{\lambda} \in \mathcal{D}$. This is because, for any combination $(Q, p, r) \in \Omega(\boldsymbol{\lambda})$ and $\Omega(\boldsymbol{\lambda}^*) \supseteq \Omega(\boldsymbol{\lambda})$, we always have $(Q, p, r) \in \Omega(\boldsymbol{\lambda}^*)$. Therefore $\Omega(\boldsymbol{\lambda}^*)$ is the largest set that contains the optimal combination (Q^*, p^*, r^*) to problem (2a)-(2e).

Further, we notice that set $\Omega(\boldsymbol{\lambda})$ has a quadratic form in $\boldsymbol{\xi}$, thus for any $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2 \in \mathcal{D}$, we will have $\Omega(\boldsymbol{\lambda}_1) \supseteq \Omega(\boldsymbol{\lambda}_2)$ if $\boldsymbol{\lambda}_1 \leq \boldsymbol{\lambda}_2$. That means that we need to find a $\boldsymbol{\lambda}^*$ such that $\boldsymbol{\lambda}^* \leq \boldsymbol{\lambda}$ for any $\boldsymbol{\lambda} \in \mathcal{D}$. Note that for any $\boldsymbol{\lambda}_c \in \mathcal{D}_c$, we can always find $\boldsymbol{\lambda} = \boldsymbol{\lambda}_c + \boldsymbol{o} \in \mathcal{D}$ and $\boldsymbol{\lambda} \geq \boldsymbol{\lambda}_c$, then we have $\Omega(\boldsymbol{\lambda}_c) \supseteq \Omega(\boldsymbol{\lambda})$. It implies $\boldsymbol{\lambda}^* \in \mathcal{D}_c$. However, set $\mathcal{D}_c$ is not singleton. Given any $\boldsymbol{\lambda} \in \mathcal{D}_c$, we can construct another $\boldsymbol{\lambda}' \in \mathcal{D}_c$ by increasing λ_i and decreasing λ_j $(j \neq i)$ at the same time. Thus, a threshold satisfying (3) is not necessarily optimal to problem (2a)-(2e).

B. Proof of Proposition 2

Note that λ only appears in constraint (2c). We denote this constraint in a more compact form $\Phi(\boldsymbol{\lambda}, Q, p, r) \geq 0$, where $\Phi(\boldsymbol{\lambda}, Q, p, r) = \min_{\boldsymbol{\xi} \leq \boldsymbol{\lambda}} \boldsymbol{\xi}^T Q \boldsymbol{\xi} - 2\boldsymbol{\xi}^T (p + Q\boldsymbol{\mu}) + r$. If vector $\boldsymbol{\xi}^*$ is an optimal solution to this problem and $\xi_i^* = \lambda_i$ for some $i \in \mathcal{N}$, we say that the i^{th} term (i.e., $\xi_i \leq \lambda_i$) of the box constraint[5] (i.e., $\boldsymbol{\xi} \leq \boldsymbol{\lambda}$) is active. If $\xi_i^* < \lambda_i$, we say it is inactive.

For the optimal decision threshold $\boldsymbol{\lambda}^*$ to problem (2a)-(2e), constraint (2c) requires $\Phi(\boldsymbol{\lambda}^*, Q, p, r) \geq 0$. If no term of the box constraint is active, i.e., $\boldsymbol{\xi}^* = Q^{-1} p + \boldsymbol{\mu} < \boldsymbol{\lambda}^*$, then we have $\Phi(\boldsymbol{\lambda}^*, Q, p, r) = r - (p + Q\boldsymbol{\mu})^T Q^{-1} (p + Q\boldsymbol{\mu}) \geq 0$. Let $z = (\boldsymbol{\mu}\boldsymbol{\mu}^T - \gamma_2 \Sigma) \otimes Q - \Sigma \otimes P + 2\boldsymbol{\mu}^T p - \gamma_1 s - r$, consequently,

$$\begin{aligned} z &\leq (\boldsymbol{\mu}\boldsymbol{\mu}^T - \gamma_2 \Sigma) \otimes Q + 2\mu^T p - r - 2\sqrt{\gamma_1} \|Lp\|_2 \\ &\leq -\left(\gamma_2 \Sigma \otimes Q + 2\sqrt{\gamma_1} \|Lp\|_2 + p^T Q^{-1} p\right) \leq 0. \end{aligned}$$

Since objective z denotes the optimal detection probability for cooperative sensing under distribution uncertainty, a negative result is meaningless in practice. Next consider the case where some but not all terms of the box constraint are active, i.e., $\exists\, i, j \in \mathcal{N}$, $\xi_i^* = \lambda_i^*$ and $\xi_j^* < \lambda_j^*$. In this case, threshold vector $\boldsymbol{\lambda}^*$ would not be optimal to problem (2a)-(2d), since we can get a different $\boldsymbol{\lambda}'$ by properly increasing λ_i^* and decreasing λ_j^* as dictated by the α-Preserving condition. This new decision threshold $\boldsymbol{\lambda}'$ will improve the objective. Thus, we reach the conclusion that for an optimal decision threshold $\boldsymbol{\lambda}^*$, all terms of the box constraint must be active, i.e., $\boldsymbol{\xi}^* = \boldsymbol{\lambda}^* \leq \boldsymbol{\mu} + Q^{-1} p$, then we have $\Phi(\boldsymbol{\lambda}^*, Q, p, r) = (\boldsymbol{\lambda}^*)^T Q \boldsymbol{\lambda}^* - 2(\boldsymbol{\lambda}^*)^T (p + Q\boldsymbol{\mu}) + r$.

[5] Box constraint refers to a series of simple inequality constraints on individual variables, e.g., $L_i \leq x_i \leq U_i$, $\forall\ 1 \leq i \leq N$.

References

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, February 2005.

[2] L. Duan, L. Gao, and J. Huang, "Contract-based cooperative spectrum sharing," in *Proceeding of IEEE Dynamic Spectrum Access Networks (DySPAN)*, Aachen, Germany, May 2011.

[3] L. Gao, X. Wang, Y. Xu, and Q. Zhang, "Spectrum trading in cognitive radio networks: A contract-theoretic modeling approach," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 843–855, April 2011.

[4] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proceeding of IEEE Dynamic Spectrum Access Networks (DySPAN)*, Nov. 2005.

[5] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 21–24, Jan. 2007.

[6] Y. Zhang, Q. Zhang, and S. Wu, "Entropy-based robust spectrum sensing in cognitive radio," *IET Communications*, vol. 4, no. 4, pp. 428–436, 5 2010.

[7] J. Lunden, S. A. Kassam, and V. Koivunen, "Robust nonparametric cyclic correlation-based spectrum sensing for cognitive radio," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 38–52, Jan. 2010.

[8] G. Taricco, "Optimization of linear cooperative spectrum sensing for cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 77–86, Feb. 2011.

[9] K. B. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 878–893, May 2009.

[10] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations Research*, vol. 58, pp. 595–612, May 2010.

[11] K. G. Murty and S. N. Kabadi, "Some NP-complete problems in quadratic and nonlinear programming," *Mathematical Programming*, vol. 39, pp. 117–129, 1987.

[12] Y. E. Nesterov and A. Nemirovski, *Interior-Point Polynomial Algorithms in Convex Programming (Studies in Applied and Numerical Mathematics)*. Society for Industrial Mathematics, 1994.

[13] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11–12, pp. 625–653, 1999.