Adaptive Underlay Cognitive Radios with Imperfect CSI and Probabilistic Interference Constraints

Antonio G. Marques*, Luis M. Lopez-Ramos*, Georgios B. Giannakis†, and Javier Ramos*
*Dept. of Signal Theory and Communications, King Juan Carlos University, Madrid, Spain (Email: see http://www.tsc.urjc.es)
†Dept. of Electrical and Computer Engineering, University of Minnesota, Minneapolis, USA (Email: georgios@umn.edu)

***Abstract*—Efficient design of cognitive radios requires secondary users implementing adaptive resource allocation schemes that exploit knowledge of the channel state information (CSI), and limit interference to the primary system. In this paper, stochastic resource allocation algorithms are developed for underlay cognitive radios to maximize the sum-rate of secondary users while adhering to "average power" and "probability of interference" constraints. The latter guarantee that during most of the time the power interfering with primary receivers stays below a certain pre-specified level. Although the resultant optimization problem is non-convex, it exhibits zero-duality gap and can be efficiently solved. The optimal schemes are a function of the link quality of the secondary network, the activity of the primary users, and the Lagrange multipliers associated with the considered constraints. The focus is on developing algorithms that: i) employ stochastic approximation tools to estimate the multipliers; and ii) are able to cope with imperfections present in the CSI of the primary network.**

***Index Terms*—Cognitive radio, resource management, stochastic approximation, imperfect channel state information.**

I. Introduction

The perceived spectrum under-utilization along with the proliferation of new wireless services have prompted recent research on dynamic spectrum management and wireless cognitive radios (CRs), which are capable of sensing and accessing the spectrum dynamically. One of the most promising scenarios for the deployment of CRs entails *underlay* systems, where CR users (referred to as secondary users) adapt their transmission powers to limit the interference inflicted to receivers (referred to as primary users) holding the spectrum licence. Secondary CRs must intelligently sense the radio spectrum with two goals in mind: G1) Estimate the channels between secondary transmitters and primary receivers so that interference is kept under control; and G2) Estimate CR link gains so that fading can be mitigated, and secondary users can take advantage of "good" channel realizations. Based on the measurements obtained through sensing, secondary users will adapt their available resources (here power, rate, and scheduling coefficients) to the channel conditions.

The merits of adaptive schemes which exploit knowledge of statistical and instantaneous channel *state information* (SI) to optimally allocate transmitter resources are well documented in traditional wireless systems; see e.g., [3, Chap. 9]. However, for channel-adaptive schemes to be deployed in CR scenarios [4], [7], [9], [10], several challenges not present in traditional wireless networks need to be addressed. Major *design challenges* include: DC1) The need for adaptive schemes to satisfy additional constraints to keep the interference low; DC2) The quick variation and knowledge of the *statistical* channel SI which may not be available in CR settings; and DC3) The difficulty to acquire instantaneous channel SI, which has heterogenous quality; meaning that the CR is expected to have better knowledge of the *SI of the secondary network* (SISN) than that of the *SI of the primary network* (SIPN).

Different alternatives have been proposed to cope with these design challenges. To keep the interference low (DC1), some works limit the interfering power at the primary receivers, either by imposing instantaneous or average interference power constraints which are better suited for fading channels; see e.g., [4]. Recent designs pursue a probabilistic approach to limit the probability of interfering the primary transmissions [9], [10]. An alternative to deal with DC2 consists of using stochastic tools to solve the resource allocation problem; see [6], [10] for examples in the context of CRs. Dual stochastic algorithms have been successfully employed to allocate resources over wireless networks (see references in [10], [6]) because such algorithms do not require knowledge of the channel statistics, are robust to channel non-stationarities, and have affordable complexity. Regarding DC3, different types of SI imperfections have been considered. Most include noisy SI (where SI is corrupted with noise [7]), or, quantized SI (where only a coarse description of the channel SI is available [6]). However, in the context of CR only a few works have considered SI that is noisy and also outdated [2]; or have incorporated those imperfections into the design of the resource allocation algorithms.

Motivated by these considerations, stochastic resource allocation algorithms are developed here to optimize CR performance, limit the probability of interfering with primary users, and consider outdated as well as noisy SIPN. Probabilistic interference constraints are considered because: SIPN imperfections render deterministic interference constraints infeasible (or grossly suboptimal), and they allow secondary users to take advantage of very good channel realizations even if interference is caused. We recently investigated this problem for *overlay* CR, where secondary users can only transmit if the band is not occupied by a primary user [5]. Differently, the focus here is on *underlay* CR, where secondary users adapt their powers to keep interference below a prespecified threshold. This is a challenging task because the resultant optimization problem is non-convex. Fortunately, the formulated problem has zero-duality gap, so that dual methods can be used to find the overall optimum allocation.

The work in this paper was supported by the Spanish Ministry of Science and Innovation grant No. TEC2009-12098.

Operating conditions of the considered CRs are as follows. It is assumed that the available SISN is instantaneous and error free, while the available SIPN is outdated and noisy. A simple continuous first-order Markov model plus additive white noise is used to characterize such imperfections, but the results also hold for more complex models. Consideration of imperfections on the SISN is also of interest, but exceeds the scope of this work. Secondary users capable of adapting their power and rate loadings access orthogonally a set frequency bands, which are primarily assigned to licensed users. Orthogonality here means that if a secondary user is transmitting, no other secondary user can be active over the same band. The resource allocation schemes are then obtained as the solution of a sum-average rate maximization subject to maximum "average power" and "probability of interference" constraints. The optimal resource allocation scheme turns out to be a function of the *instantaneous* SISN, the (possibly outdated and noisy) SIPN, and the optimum Lagrange multipliers associated with the optimization problem. The optimization problem structure facilitates separability, so that optimization of the primary variables, although non-convex, can be efficiently performed. To obtain the multipliers low-complexity stochastic schemes are developed for estimating the multiplier values online.

Section 2 describes the operating conditions of the network. The optimization problem is formulated in Section 3. Stochastic algorithms for solving the formulated problem are developed in Section 4. Numerical examples and conclusions wrap-up this paper. [1]

II. Model description

Consider a CR with M secondary users (indexed by m) transmitting opportunistically over K different frequency bands (indexed by k). For simplicity, we consider that each band has the same bandwidth and is occupied by a different primary user. The secondary network has an access point (AP) which is the destination of all secondary users. The AP acts as a central scheduler which collects the CSI and makes the resource allocation decisions. It is assumed that at every instant, the instantaneous gain of the secondary links is available. Differently, the gain between the secondary and the primary users is not sensed at every time instant. These assumptions are well suited for scenarios where sensing the state of the primary network entails a cost much higher than that of sensing the state of the secondary links (because primary users are too many, are located far away from secondary users, or do not collaborate with the sensing process). The channel's instantaneous power gain between the mth secondary user and the AP in the kth frequency band is denoted by $h_{k,2}^m$; and represents the noise-normalized squared magnitude of the fading coefficient (subindex 2 emphasizes that the channel pertains to *secondary* users). Similarly, the instantaneous normalized power gain between the mth secondary transmitters and the kth *primary* receiver is denoted by $h_{k,1}^m$. Channels are assumed to be ergodic.

[1] *N*otation: x^* denotes the optimal value of variable x; $\wedge$ the boolean "and" operator; $\mathbb{1}_{\{\cdot\}}$ the indicator function ($\mathbb{1}_{\{x\}} = 1$ if x is true and zero otherwise); and $[x]_+$ the projection of x onto the non-negative orthant, i.e., $[x]_+ := \max\{x, 0\}$.

Since the CSI for the secondary network is assumed to be perfect, at every instant n the coefficients $h_{k,2}^m[n]$ are deterministically known. SIPN imperfections imply that at instant n the value of $h_{k,1}^m[n]$ is not deterministically known, rather the distribution of $h_{k,1}^m[n]$ conditioned on all previous measurements is available. This distribution will be referred to as belief. When written as a probability density function, the believe is denoted as $f_{h_{k,1}^m[n]}(h)$ and when written as a cumulative density function, it is denoted as $F_{h_{k,1}^m[n]}(h)$.

Let $s_k^m[n]$ denote a boolean variable which is one if the channel $h_{k,1}^m$ is sensed at instant n and zero otherwise. Moreover, $\tilde{h}_{k,1}^m[n]$ let denote the (possibly noisy) measurement of $h_{k,1}^m[n]$ obtained if $s_k^m[n] = 1$. The two main sources of imperfections in the SIPN are: i) outdated SIPN (for the instants n such that $s_k^m[n] = 0$); and ii) noisy SIPN (due to errors in the sensing process that render $\tilde{h}_{k,1}^m[n] \neq h_{k,1}^m[n]$).

We start by modeling the time dynamics of $h_{k,1}^m[n]$ which, for simplicity, are assumed to follow a first-order Markov process. Let $q_k^m(h_{new}, h_{old})$ denote the probability mapping which measures the probability of having $h_{k,1}^m[n+1] = h_{new}$ given that $h_{k,1}^m[n] = h_{old}$. Moreover, let $f_k^m(h, n)$ denote the pdf of $h_{k,1}^m[n] = h$. It follows readily that $f_k^m(h, n+1) = \int q_k^m(h, x) f_k^m(x, n) dx$.

Next, in order to incorporate the sensing errors into our model, we assume a simple additive model so that $\tilde{h}_{k,1}^m[n] = h_{k,1}^m[n] + v_k^m[n]$ where $v_k^m[n]$ is a white noise with known pdf $f_{v_k^m}(v)$ which does not depend on $h_{k,1}^m[n]$.

Once the model is described, we are ready to describe how the belief $f_{h_k^m[n]}(h)$ is found (updated). The following cases need to be considered :

- $s_k^m[n] = 0$: It holds for this case that $f_{h_k^m[n+1]}(h) = \int q_k^m(h, x) f_{h_k^m[n]}(x) dx$.
- $s_k^m[n] = 1$: We first update the belief of the previous instant to get the prediction $\tilde{f}_{h_k^m[n+1]}(h) = \int q_k^m(h, x) f_{h_k^m[n]}(x) dx$. Then, we use the measurement $\tilde{h}_k^m[n]$ to correct $\tilde{f}_{h_k^m[n+1]}(h)$ using Bayes' rule:
$$f_{h[n+1]}(h) = \frac{\tilde{f}_{h_k^m[n+1]}(h) f_{v_k^m}(h - \tilde{h})}{\int_{\forall x} \tilde{f}_{h_k^m[n+1]}(x) f_{v_k^m}(x - \tilde{h}) dx}$$

The described procedure clearly resembles the prediction-correction steps of a Kalman filter (only prediction if $s_k^m[n] = 0$, and both prediction and correction when $s_k^m[n] = 1$).

The value of coefficients $h_{k,2}^m[n]$ $\forall(k, m)$ constitutes the SISN, while the beliefs $f_{h_k^m[n]}(h)$ $\forall(k, m)$ constitute the SIPN. The overall SI of the system will be denoted by $\mathbf{h}$.

A. Resources at the secondary network

Now, we introduce the variables to be designed. Let w_k^m denote a boolean variable such that $w_k^m = 1$ if the mth user is scheduled to transmit into the kth band and $w_k^m = 0$ otherwise. Provided that $w_k^m = 1$, let p_k^m denote the instantaneous power transmitted over the kth band by the mth user. Under bit error rate or capacity constraints, instantaneous rate and power variables are coupled. This rate-power coupling will be represented by the function $C_k^m(h_k^m, p_k^m)$. Throughout this paper it is assumed that the rate-power function $C_k^m(h_k^m, \cdot)$ is given by Shannon's capacity formula $\log(1 + h_k^m p_k^m / \Gamma_k^m)$, where Γ_k^m represents the SNR-gap which depends on the coding scheme implemented [3].

The secondary CR operates in a time-block fashion, where the duration of each block corresponds to the coherence time of the fading channel. This way, at every time n the AP will use the current SI vector $\mathbf{h}$ to find the (optimum) value of w_k^m and p_k^m. Since $\mathbf{h}$ depends on n and $\{w_k^m, p_k^m\}$ depend on $\mathbf{h}$, the value of the design variables $\{w_k^m, p_k^m\}$ will clearly vary across time. Through the manuscript, we will write $w_k^m(\mathbf{h})$ and $p_k^m(\mathbf{h})$, or $\mathbf{h}[n]$, $w_k^m[n]$ and $p_k^m[n]$, wherever is convenient to emphasize the corresponding dependence.

For this CR configuration, we wish to develop adaptive algorithms that use the instantaneous SISN and the outdated/noisy SIPN to determine the secondary users who transmit on each band and their corresponding rate and power loadings.

III. Problem formulation

We begin by identifying the constraints that the optimal schemes need to satisfy. Clearly, variables p_k^m are constrained to be nonnegative and the boolean variables w_k^m are constrained to belong to the set $\{0,1\}$. Moreover, to ensure that at most one user transmits into a given band k, we need

$$\sum\nolimits_k w_k^m(\mathbf{h}) \leq 1, \quad \forall k. \tag{1}$$

Note that the previous constraint does not prevent a secondary user to occupy more than one band. We also consider that the maximum average power the mth secondary user can transmit is $\check{p}^m$; hence,

$$\mathbb{E}\left[\sum\nolimits_k w_k^m(\mathbf{h}) p_k^m(\mathbf{h})\right] \leq \check{p}^m, \quad \forall m, \tag{2}$$

where expectations are taken over $\mathbf{h}$. Finally, to keep the interference to the primary network under control we allow a maximum probability of interference per band $\check{o}_k$. Interference occurs when the received power at the primary receiver due to secondary transmissions exceeds a threshold γ_k; i.e. if $p_k^m(\mathbf{h}) h_{k,1}^m(\mathbf{h}) > \gamma_k$ and $w_k^m(\mathbf{h}) > 0$. Hence, limiting the probability of interference amounts to bound $\Pr\{\sum_m w_k^m(\mathbf{h}) \mathbb{1}_{\{p_k^m(\mathbf{h}) h_{k,1}^m > \gamma_k\}} = 1\} \leq \check{o}_k$. Since w_k^m is a boolean variable, the previous inequality can be alternatively written as

$$\mathbb{E}\left[\sum\nolimits_m w_k^m(\mathbf{h}) \mathbb{1}_{\{p_k^m(\mathbf{h}) h_{k,1}^m > \gamma_k\}}\right] \leq \check{o}_k, \quad \forall k. \tag{3}$$

The second step to formulate the optimization problem consists of defining the metric to be optimized. In this paper, we are interested in maximizing the sum-average rate given by $\bar{c} := \sum_k \mathbb{E}\left[w_k^m(\mathbf{h}) \;\; C_k^m(h_k^m, p_k^m(\mathbf{h}))\right]$. Nevertheless, other objective functions such as weighted sum-rate or sum-utility rate can be used without changing the basic structure of the solution; see, e.g., [5] for further details.

Under all previous considerations, the optimal resource allocation is obtained as the solution of the following problem:

$$\bar{c}^* := \max_{w_k^m(\mathbf{h}), p_k^m(\mathbf{h})} \sum\nolimits_k \mathbb{E}\left[w_k^m(\mathbf{h}) C_k^m(h_k^m, p_k^m(\mathbf{h}))\right] \tag{4a}$$

$$\text{s. to}: (1),\ (2),\ (3),\ w_k^m(\mathbf{h}) \in \{0,1\},\ \text{and}\ p_k^m(\mathbf{h}) \geq 0. \tag{4b}$$

Where the dependence of the optimization variables on the SI vector $\mathbf{h}$ has been made explicit.

IV. Optimal Resource Allocation

The main difficulty to solve (4) is the non-convexity[2] associated with (3). Remarkably, although (4) is a non-convex problem, it can be shown that: i) the global optimization of the Lagrangian can be performed efficiently, and ii) the problem has zero duality gap. The second fact is basically true due to the averages over the channel distribution in (4) and can be rigorously proved following the lines in [8]. Property ii) implies that dual methods can be used to find the global optimum solution of (4), provided that the global optimum of the Lagrangian is found. Property i) implies that such a task can be performed efficiently. Let π^m and θ_k denote the Lagrange multipliers associated with the constraints in (2) and (3), respectively. It can be shown that the optimal allocation in this case is

$$\begin{aligned} \phi_k^m(p_k^m[n]) \;&:= C_k^m(h_k^m[n], p_k^m[n]) - \pi^m[n] p_k^m[n] \\ &- \;\; \theta_k[n] F_{h_k^m[n]}(p_k^m[n]/\gamma_k), \end{aligned} \tag{5}$$

$$p_k^{m*}[n] \;:= \left[\arg\max_{p_k^m[n]} (\phi_k^m(p_k^m[n]))\right]_+, \tag{6}$$

$$w_k^{m*}[n] \;:= \mathbb{1}_{\{(m = \arg\max_l \phi_k^l(p_k^{l*}[n]))\ \wedge\ (\phi_k^m(p_k^{m*}[n]) > 0)\}} \tag{7}$$

The functional (5) can be readily interpreted as a user-channel quality indicator. Intuitively, $\phi_k^m(p_k^m[n])$ represents the potential reward for the system associated with the (m,k) user-channel pair (the rate is a reward, the power and the interference are costs, and the multipliers are the corresponding prices). Mathematically, $\phi_k^m(p_k^m[n])$ represents the contribution to the Lagrangian if user m is scheduled for transmission on channel k; i.e., if $w_k^m[n] = 1$.

Interestingly, (6) reveals that to find the value of $p_k^m[n]$ which maximizes the overall Lagrangian, it suffices to find the power $p_k^m[n]$ maximizing the quality of the user-channel combination. Similarly, (7) dictates that to find the scheduling in channel k that maximizes the overall Lagrangian, it suffices to select the coefficients maximizing the quality of the specific channel. That is a manifestation of the separability of the Lagrangian across channels and users (which follows from the structure of the optimization problem in (4)). Note finally that (7) establishes that the user scheduling is opportunistic and greedy (only the user with highest quality in a given band must be scheduled).

The maximization in (6) is challenging because $\phi_k^m(\cdot)$ is not strictly concave. The rate term is strictly concave and the power term is linear, but the interference term $-F_{h_k^m[n]}(p_k^m[n]/\gamma_k)$ is non-concave. However, (6) involves a single (scalar) optimization variable, opening the door to develop efficient methods to solve the optimization. For example, the number of local optima of $\phi_k^m(\cdot)$ depends on the number of local optima of $f_{h_k^m[n]}$, so that for most practical pdfs $\phi_k^m(\cdot)$

[2]There are two other sources of non-convexity in (4), but they can be easily handled. The first one is that $w_k^m \in \{0,1\}$ is a non-convex set. This can be solved by relaxing that set and allowing $w_k^m \in [0,1]$. Since w_k^m is a random variable that only appears in linear terms, it can be shown that this relaxed solution coincides with the original one. The second source is the presence of the monomials $w_k^m p_k^m$ and $w_k^m C_k^m$. This can be trivially solved by introducing dummy implicit variables $\tilde{p}_k^m := w_k^m p_k^m$ in (4) and showing convexity using the properties of the perspective function, see, e.g., [5].

has a small number of local optima. If all of them can be found, then the global optimum can be selected.

A. Estimating the optimum Lagrange multipliers

Different methods can be used to set the value of $\pi^m[n]$ and $\theta_k[n]$. Traditionally, $(\pi^m[n], \theta_k[n])$ are set to a constant value (π^{m*}, θ_k^*) corresponding to the value that maximize the dual function associated with (4) (recall that the duality gap is zero). This implies that if $\pi^m[n] = \pi^{m*}$ and $\theta_k[n] = \theta_k^*$ are substituted into (5)-(7), the resulting resource allocation is the optimal solution of (4) [1]. The main drawbacks associated with this approach are that: i) π^{m*} needs to be found through numerical search which, at every step, requires averaging over all possible states of $\mathbf{h}$; and ii) if the channel statistics or the number of users change, π^{m*} needs to be recomputed. Recently, alternative approaches that rely on stochastic approximation tools have been proposed to find the value of the multipliers [6], [10], [5]. These approaches do not try to find the optimal value of (π^{m*}, θ_k^*), but an estimate of it which is updated at every time instant and remains sufficiently close to (π^{m*}, θ_k^*). Main advantages of these approaches are: i) the computational complexity is very low, and ii) they can cope with channel non-stationarities. The drawback is that the resulting schemes are slightly suboptimal. Specifically, for the problem at hand the following iterations are proposed

$$\pi^m[n+1] = \left[\pi^m[n] - \mu(\check{p}^m - \textstyle\sum_k w_k^{m*}[n] p_k^{m*}[n])\right]_+ \quad (8)$$

$$\theta_k[n+1] = \left[\theta_k[n] - \mu(\check{o}_k - \textstyle\sum_m w_k^{m*}[n] F_{h_k^m[n]}(p_k^{m*}[n]/\gamma_k))\right]_+ \quad (9)$$

where μ denotes a small stepsize. Basically, the update in (8) is an unbiased stochastic subgradient of the dual function of (4); see [5], [1]. Assuming that the updates in (8) are bounded, it can be shown that the sample average of the stochastic resource allocation: i) is feasible and ii) entails a small loss of performance relative to the optimal solution of (4). To be rigorous, let us define $\bar{p}^m[n] := \frac{1}{n}\sum_{l=1}^{n}\sum_k w_k^{m*}[l] p_k^{m*}[l]$; $\bar{c}[n] := \frac{1}{n}\sum_{l=1}^{n}\sum_{k,m} w_k^{m*}[l] C_k^m\ (h_k^m[l], p_k^{m*}[l])$; and $\bar{o}^k[n] := \frac{1}{n}\sum_{l=1}^{n}\sum_m w_k^{m*}[l] \mathbb{1}_{\{p_k^{m*}[l] h_{k,1}^m[l] > \gamma_k\}}$. Then, it holds with probability one that as $n \to \infty$: i) $\bar{p}^m[n] = \check{p}^m$ and $\bar{o}^k[n] = \check{o}^k$, and ii) $\bar{c}[n] \geq \bar{c}^* - \delta(\mu)$, where $\delta(\mu) \to 0$ as $\mu \to 0$.

V. Numerical simulations

The simulation setup is the following: $M = 5$, $K = 10$, $\check{p}^m = 2$, $\Gamma_k^m = 1$, $\check{o}_k = 10\%$, and $\gamma_k = 0.5$. The amplitudes of the secondary links $(h_{k,2}^m[n])^{1/2}$ are Rayleigh distributed, and the average SNR for all users and bands is 3dB. Channel gains of $h_{k,1}^m[n]$ are Gaussian distributed (GD) and the time correlation model is $h_{k,1}^m[n] = \delta h_{k,1}^m[n-1] + (1-\delta) z_k^m[n]$, with $\delta = 0.95$ and $z_k^m[n]$ being white and GD with zero mean and variance $(1-\delta^2)^2$. The measurement noise $v_k^m[n]$ is GD, with zero mean and variance 0.005. The AP senses the primary channels every 16 slots.

Table 1 lists the values of the average sum-rate, power and interference probability for the following schemes: S1) the optimal scheme derived in this paper (Section 5.2); S2) a scheme that considers the SIPN is free of errors (it assumes that $h_{k,1}^m[n+n_0] = \tilde{h}_{k,1}^m[n]$ where $n_0 = 0,...,15$) and then optimally solves (4); S3) a scheme that optimizes (4) replacing the average interference constraint with an instantaneous interference constraint; S4) a scheme optimizing only the power allocation, but assumes that the channel allocation is fixed (10/5=2 channels per user) and S5) a genie-aided scheme that knows the actual value of the SIPN. The space limitations prevent us from a detailed discussion of the results. Nevertheless, the preliminary results suggest the theoretical claims are valid and illustrate the advantages of our algorithms.

TABLE I
SIMULATION RESULTS

	S1	S2	S3	S4	S5
$(\bar{c}, \bar{p})$	(12,2.0)	(13.5,2.0)	(10,0.9)	(7,2.0)	(12.5,2.0)
$(1/K)\sum_k \bar{o}_k$	10%	30%	10%	10%	10%

VI. Conclusions

We developed stochastic resource allocation algorithms for underlay secondary cognitive radios operating over wireless fading channels. The schemes were obtained as the solution of a sum-rate maximization problem subject to maximum "average power" and "probability of interference" constraints. The probabilistic interference constraint was tailored to account for imperfections on the sensing schemes (which rendered the primary CSI outdated and noisy). Although non-convex, the resultant problem had zero-duality gap and could be solved with low/moderate complexity. It was shown that the optimal schemes should maximize a quality link functional which weights: the quality of the secondary links (in terms of rate and power), the probability of interfering the primary users, and several Lagrange multipliers. The value of those multipliers depends on the history of the system and the requirements of the primary and secondary networks. Stochastic algorithms where proposed to: i) estimate and predict the probability of interference and ii) estimate the optimum value of the multipliers. Future (and current) work includes not only accounting for sensing imperfections, but jointly optimizing sensing and resource allocation.

References

[1] D. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, 2003.

[2] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Trans. Info. Theory*, pp. 2053-2071, May 2008.

[3] A. Goldsmith, *Wireless Communications*, Cambridge Univ. Press, 2005.

[4] X. Kang, Y. Liang, A. Nallanathan, H.K. Garg and R. Zhang, "Optimal power allocation for fading channels in cognitive radio networks: Ergodic capacity and outage capacity," *IEEE Trans. Wireless Commun.*, pp. 940-950, Feb. 2009.

[5] A.G. Marques, G. B. Giannakis, L.M. Lopez-Ramos, and J. Ramos, "Stochastic resource allocation for cognitive radio networks based on imperfect state information," *Proc. of Intl. Conf. on Acoustics, Speech and Signal Processing.*, Prague, Czech Rep., May. 22- 27, 2011.

[6] A. G. Marques, X. Wang, and G. B. Giannakis, "Dynamic resource management for cognitive radios using limited-rate feedback," *IEEE Trans. on Signal Processing.*, pp. 3651-3666, Sep. 2009.

[7] L. Musavian and S. Aissa, "Fundamental capacity limits of cognitive radio in fading environments with imperfect channel information," *IEEE Trans. Commun.*, pp. 3472-3480, Nov. 2009.

[8] A. Ribeiro and G. B. Giannakis, "Separation principles in wireless networking," *IEEE Trans. on Info. Theory*, pp. 4488-4505, Sep. 2010.

[9] R. Urgaonkar and M. Neely, "Opportunistic scheduling with reliability guarantees in cognitive radio networks," *IEEE Trans. Mobile Comp.*, pp.766-777, Jun. 2009.

[10] X. Wang, "Jointly Optimal Sensing Selection and Power Allocation for Cognitive Communications," *Proc. of Globecom Conf.*, Miami, FL, Dec. 6-10, 2010.