Resource Allocation for Interweave and Underlay CRs Under Probability-of-Interference Constraints

Antonio G. Marques, *Member*, IEEE, Luis M. Lopez-Ramos, *Student Member*, IEEE, Georgios B. Giannakis, *Fellow*, IEEE, and Javier Ramos

***Abstract*—Efficient design of cognitive radios (CRs) calls for secondary users implementing adaptive resource allocation schemes that exploit knowledge of the channel state information (CSI), while at the same time limiting interference to the primary system. This paper introduces stochastic resource allocation algorithms for both interweave (also known as overlay) and underlay cognitive radio paradigms. The algorithms are designed to maximize the weighted sum-rate of orthogonally transmitting secondary users under average-power and probabilistic interference constraints. The latter are formulated either as short- or as long-term constraints, and guarantee that the probability of secondary transmissions interfering with primary receivers stays below a certain pre-specified level. When the resultant optimization problem is non-convex, it exhibits zero-duality gap and thus, due to a favorable structure in the dual domain, it can be solved efficiently. The optimal schemes leverage CSI of the primary and secondary networks, as well as the Lagrange multipliers associated with the constraints. Analysis and simulated tests confirm the merits of the novel algorithms in: i) accommodating time-varying settings through stochastic approximation iterations; and ii) coping with imperfect CSI.**

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

I. Introduction

The perceived spectrum under-utilization along with the proliferation of new wireless services have fueled the recent upsurge of research on dynamic spectrum management and wireless cognitive radios (CRs), which are capable of sensing and accessing the spectrum opportunistically [11], [26]. CR users –also referred to as secondary users (SUs)– adapt their transmissions to limit interference they inflict to primary users (PUs), which hold the licence of the spectrum band accessed. In the so-termed *interweave* (also termed *overlay* paradigm by some authors), CRs can use a frequency band only if no PU is active; whereas in the *underlay* paradigm CRs can access the channel even if PUs are active, provided they adjust their power so that interference at active PU sites remains below a pre-specified threshold [26], [10], [28].

Manuscript submitted on 5 January 2012; revised 16 May 2012. The work in this paper was partially supported by the Spanish MCIN grant No. TEC2009-12098, the Spanish FPU grant No. AP2010-1050, and the QNRF grant NPRP 09-341-2-128. Parts of this paper were presented at IEEE ICASSP'11 and IEEE CAMSAP'11.

A. G. Marques, L. M. Lopez-Ramos and J. Ramos are with the Dept. of Signal Theory and Comm., King Juan Carlos Univ., Camino del Molino s/n, Fuenlabrada, Madrid 28943, Spain (e-mails: see http://www.tsc.urjc.es).

G. B. Giannakis is with the Dept. of Electrical and Computer Eng., Univ. of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA (e-mail: georgios@umn.edu).

Digital Object Identifier 10.1109/JSAC.2012.121108.

Instrumental to controlling interference and also leveraging favorable link conditions, is knowledge of the CR-to-PU and CR-to-CR channels acquired during the sensing phase. Based on this, CRs adapt available resources, namely power, rate, and scheduling coefficients, to the intended channels. The merits of exploiting statistical or instantaneous channel state information (CSI) for adaptive resource allocation are well documented in wireless networking literature [9, Ch. 9]. But the CR paradigm faces the following additional *design challenges* (DC) [18], [13], [8], [14], [14], [23], [24], [2], [7]:

- DC1) Extra constraints are needed to effect interference control;
- DC2) CR volatility may render statistical CSI outdated; and also,
- DC3) Instantaneous CSI of the PU network is difficult or impossible to acquire.

In order to address DC1, existing works limit CR-inflicted interference either through instantaneous (short-term) and average (long-term) transmit-power constraints [14], [27], [28], [2]; or, by controlling the probability of interfering with PU transmissions, see, e.g., [23], [4], [5], [24], [1], [6]. For this second case, most works have focused on short-term constraints, which are relatively easier to handle. Stochastic resource allocation (RA) approaches [15], [24], offer viable means to deal with DC2. As with general wireless networks, dual stochastic algorithms are particularly attractive because they are computationally simple, do not require knowledge of channel statistics, and exhibit robustness to channel variations; see [25], [20] and references in [15], [24]. Regarding DC3, most prior CR works consider noisy or quantized CSI [18], [15], [22], [12]; a few consider *outdated* CSI for CRs [22], [18], [5]; and very few incorporate mechanisms to *predict* the actual CSI [4], [2].

The goal of the present paper is to develop stochastic RA algorithms for both interweave and underlay paradigms that optimize sum-rate performance of a CR network, limit the *probability* of interfering with PUs (both short-term and *long-term* limits are investigated), and jointly account for *outdated* and noisy CSI. Probabilistic long-term interference constraints are adopted not only because they lead to improved performance, but also because uncertain information on the CR-to-PU channels renders short-term interference constraints infeasible (if the constraint has to hold with probability one) or grossly suboptimal (if the constraint holds probabilistically). Instantaneous CSI of the CR-to-CR links is assumed perfect, while that of CR-to-PU channels can be noisy and outdated. A simple continuous first-order Markov model with additive white noise is used to capture such imperfections, but more complex models can be afforded too. Such models enable channel prediction and correction to track the CR-to-PU changing CSI, which is utilized by per-band orthogonal CR transmissions to adapt their power and rate loadings. The RA schemes are obtained as the solution of a weighted sum-average rate maximization subject to maximum "average power" and "probability of interference" constraints that come in two flavors: a short-term constraint ensuring that the probability of interference is kept below a pre-specified limit per time slot; and a novel long-term constraint guaranteeing the same for a fraction of time slots. Even though not all formulations are convex, it turns out that for all of them the duality gap is zero, meaning that the Langrangian relaxation is always optimal. Additionally, the operating conditions enable separation in the dual domain across users and frequency bands, which allows for optimal solvers with considerably reduced complexity. In all cases, the optimal RA scheme turns out to be a function of the *instantaneous* CSI of the CR-to-CR links, the (possibly outdated and noisy) CR-to-PU channels, and the optimum Lagrange multipliers, obtained via simple stochastic iterations that are robust to nonstationarities, and can even learn varying CSI on-the-fly – a highly desirable attribute for CR networks [11], [15]. Extensions to scenarios with more than one CR network (each with several users) are of interest, but go beyond the scope of this paper and are left as future work.

The rest of the paper is organized as follows. Section II presents the CSI model, means to account for CSI imperfections, and pertinent operating conditions. A simplified adaptive RA optimization problem, devoid of interference constraints, is formulated and solved in Section III. Incorporation of various interference constraints and design of the corresponding algorithms are the subjects of Section IV. Section V outlines the low-complexity stochastic iterations needed to estimate the multipliers. Numerical examples and conclusions in Sections VI and VII wrap-up this paper. [1]

II. Modeling

Consider a CR network of M SUs (indexed by m) transmitting opportunistically over K different frequency bands (indexed by k). For simplicity, suppose that: i) each band has identical bandwidth, and is licensed to a different PU; and ii) the CR network has a network controller (NC), which collects the CSI needed for channel-adaptive RA. Extensions to scenarios where those assumptions do not hold can be handled with a moderate increase in complexity.

A. Channel state information

Intuitively speaking, CSI in adaptive wireless systems entails channel-related information that must be: i) available to all users in the system; and, ii) relevant from an RA perspective. A key issue with CR systems is that CSI is heterogeneous, meaning that it is different for primary and secondary networks. The reason is twofold. First, CSI availability for links involving PU and/or CR users is different [cf. i)]. Second, the impact CSI has on the design of RA is different [cf. ii)]. The CSI for CR-to-CR links will be assumed stationary and perfectly known; that is, at every instant, the instantaneous gain of SU links will be deterministically available. For notational purposes, the channel's instantaneous power gain between the mth secondary transmitter and its intended receiver over the kth frequency band at instant n is denoted by $h_{k,2}^m[n]$. Subscript "2" is used to emphasize that the channel pertains to *secondary* transceivers. If PU transmitters are located far away from SU receivers, $h_{k,2}^m[n]$ represents the squared magnitude of the instantaneous fading coefficient divided by the noise power in the kth band. If this is not the case, $h_{k,2}^m[n]$ represents the squared magnitude of the instantaneous fading coefficient divided by the sum of the *noise* power plus the instantaneous *interference* power caused by the kth primary transmitter.

Regarding the CSI corresponding to the PU network, it will not be always assumed perfectly known; e.g., because not all frequency bands are sensed at every time instant. As a result, knowledge of the primary CSI will be probabilistic and time variant. This assumption is well suited for scenarios where sensing the PU network state costs much more than sensing the state of the CR links; e.g., because PUs are too many, or they are possibly located far away from the CRs, or they are simply not willing to collaborate. The CSI model adopted by the NC for the PU network is different for interweave and underlay settings. Each of the cases is described in detail next.

1) Perfect and imperfect primary CSI in interweave networks: In the interweave setup, the NC *only needs to know* whether each frequency band is occupied or not. To capture this occupancy, let the Boolean variable a_k represent the *activity* of the PU network on the kth band, so that $a_k[n] = 1$ if at instant n the kth PU is active, and zero otherwise. Only the 2×1 belief vector $\mathbf{f}_{\mathbf{a}k}[n] := [\Pr\{a_k[n] = 0\}, \Pr\{a_k[n] = 1\}]^T$ is available, where the probability mass of $a_k[n]$ is based on the history of the system up to n. The belief can be estimated either beforehand or in real time. Next, an example of imperfect CSI in the PU network is considered along with means of estimating the corresponding belief vector.

Let $s_k[n]$ denote a Boolean variable which equals one if the kth band is sensed at instant n, and zero otherwise. Moreover, let $\tilde{a}_k[n]$ be the (perhaps noisy) measurement of $a_k[n]$ obtained at instant n, if $s_k[n] = 1$. Two main types of imperfect CSI are: i) outdated CSI (for the instants n when $s_k[n] = 0$); and ii) noisy CSI (due to errors in the sensing process that render $a_k[n] \neq \tilde{a}_k[n]$). To cope with outdated CSI, a model is needed to capture the dynamics of $a_k[n]$ across time which, for simplicity, are assumed here to follow a first-order Markov process [2], [4]; see, e.g., [29] for alternative models. Define the transition probability matrix $\mathbf{Q}$ with (i, j)th entry $Q_{ij} := \Pr\{a_k[n]=i \;|a_k[n-1]=j\}$, for $i, j = 0, 1$. In order to account for sensing errors, consider further the probabilities of miss detection and false alarm, namely $P_{MD} := \Pr\{\tilde{a}_k[n]= 0 \;|a_k[n]=1\}$ and $P_{FA} := \Pr\{\tilde{a}_k[n]=1 \;|a_k[n]=0\}$; and use them to form the 2×1 vectors $\mathbf{q_1} := [1 - P_{FA},\ P_{MD}]^T$ and $\mathbf{q_0} := [P_{FA},\ 1 - P_{MD}]^T$.

[1]Notation: ^{T} denotes vector transposition; x^* the optimal value of variable x; $[\mathbf{x}]_l$ the lth entry of vector $\mathbf{x}$; $\wedge$ ($\vee$) the boolean "and" ("or") operator; $\mathbb{1}_{\{\cdot\}}$ the indicator function ($\mathbb{1}_{\{x\}} = 1$ if x is true and zero otherwise); and $[x]_a^b$ the projection of the scalar x onto the interval $[a, b]$, i.e., $[x]_a^b := \min\{\max\{x, a\}, b\}$.

Clearly, the CSI measurements are the observed states of a Hidden Markov Model (HMM), so that recursive Bayesian estimation can be implemented to obtain the instantaneous belief (posterior probability mass function of the unobserved states). In particular, the belief $\mathbf{f}_{\mathbf{a_k}}[n]$ is updated as follows:

- If $s_k[n]=0$, then $\mathbf{f}_{\mathbf{a_k}}[n]=\mathbf{Q}\,\mathbf{f}_{\mathbf{a_k}}[n-1]$.
- If $s_k[n]=1$ and $\tilde{a}_k[n]=0$, then predict the belief vector as $\hat{\mathbf{f}}_{\mathbf{a_k}}[n]:=\mathbf{Q}\,\mathbf{f}_{\mathbf{a_k}}[n-1]$; and using that $\tilde{a}_k[n]=0$, correct $\hat{\mathbf{f}}_{\mathbf{a_k}}[n]$ via Bayes' rule to obtain ($[\cdot]_l$ stands for the lth entry of a vector)

$$[\mathbf{f}_{\mathbf{a_k}}[n]]_l=([\mathbf{q_0}]_l[\hat{\mathbf{f}}_{\mathbf{a_k}}[n]]_l)/(\mathbf{q_0}^T\hat{\mathbf{f}}_{\mathbf{a_k}}[n])\ . \tag{1}$$

- If $s_k[n]=1$ and $\tilde{a}_k[n]=1$, predict as before, and subsequently correct to find

$$[\mathbf{f}_{\mathbf{a_k}}[n]]_l=([\mathbf{q_1}]_l[\hat{\mathbf{f}}_{\mathbf{a_k}}[n]]_l)/(\mathbf{q_1}^T\hat{\mathbf{f}}_{\mathbf{a_k}}[n])\ . \tag{2}$$

Note that the described procedure resembles other recursive Bayesian models, such as the prediction-correction steps of a Kalman filter (only prediction if $s_k[n]=0$, and prediction followed by correction when $s_k[n]=1$). Different prediction-correction steps will be required if the model for the sensing error changes, the transition matrix $\mathbf{Q}$ is unknown or, if the dynamics of $a_k[n]$ are modeled differently. To be more specific about the latter, let τ_k denote the time passed between two changes of $a_k[n]$. Experimental studies, see [29] and references therein, have shown that heavy-tailed distributions are proper alternatives to model τ_k (in contrast with Markov occupancy models, which give rise to exponentially distributed τ_k). Both *Pareto* and *lognormal* distributions are investigated in [29]. Clearly, in those cases $a_k[n]$ is no longer Markovian and (1)-(2) are not optimal any more. However, the joint process $\{a_k[n],t_k[n]\}$, where $t_k[n]$ represents the time passed since the last time the value of $a_k[n]$ changed, can be modeled as Markovian, so that recursive Bayesian estimation can be employed again. These alternatives will be briefly explored through simulations in Section VI.

2) Perfect and imperfect primary CSI in underlay networks: In the underlay setup, the NC also needs to know the gains of the CR-to-PU channels. This implies that the primary CSI model in this case is different. Specifically, CSI here comprises information about the instantaneous squared fading coefficient between the mth CR and the kth PU divided by the noise power, which is denoted by $h_{k,1}^m$ (subscript "1" is used to emphasize that the link involves primary receivers). Note that $h_{k,2}^m$ accounts for the interference power, while $h_{k,1}^m$ does not. The reason is that while the interfering power generated by the PUs is a state variable, the one generated by the SU is a design variable. Clearly, if this CSI is perfect, then $h_{k,1}^m[n]$ is deterministically known at instant n. If imperfections are present, only the distribution of $h_{k,1}^m[n]$ (conditioned on all previous measurements) is available. The belief state then consists of the cumulative and the probability density function (PDF) denoted by $F_{h_{k,1}^m[n]}(h)$ and $f_{h_{k,1}^m[n]}(h)$, respectively. Depending on the operating conditions, the belief can be known beforehand or estimated over time. As in the interweave setup, the ensuing example highlights CSI imperfections in the underlay scenario, and the corresponding adaptive schemes to estimate the belief vector.

Define a Boolean variable $s_k^m[n]$ taking value 1 if $h_{k,1}^m$ is sensed at instant n, and 0 otherwise. Moreover, let $\tilde{h}_{k,1}^m[n]$ be the (possibly noisy) measurement of $h_{k,1}^m[n]$ obtained if $s_k^m[n]=1$. Paralleling the previous example, two types of imperfections are possible: i) outdated CSI (for the instants n when $s_k^m[n]=0$); and ii) noisy CSI (due to errors in the sensing process that cause $\tilde{h}_{k,1}^m[n]\neq h_{k,1}^m[n]$). The time evolution of $h_{k,1}^m[n]$ is assumed Markovian with $q_k^m(h_{new},h_{hold})$ denoting 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 then follows that $f_k^m(h,n+1)=\int_{\forall x}q_k^m(h,x)f_k^m(x,n)dx$. Next, in order to account for sensing errors, the following memoryless additive noise model is assumed: $\tilde{h}_{k,1}^m[n]=h_{k,1}^m[n]+v_k^m[n]$, where $v_k^m[n]$ stands for white noise with known PDF $f_{v_k^m}(v)$ independent of $h_{k,1}^m[n]$.

With these operating conditions, the observations follow again an HMM. Hence, the belief $f_{h_k^m[n]}(h)$ can be found using recursive Bayes estimates according to the following cases:

- If $s_k^m[n]=0$, then $f_{h_k^m[n+1]}(h)=\int_{\forall x}q_k^m(h,x)f_{h_k^m[n]}(x)dx$.
- If $s_k^m[n]=1$, then predict as $\hat{f}_{h_k^m[n+1]}(h)=\int q_k^m(h,x)f_{h_k^m[n]}(x)dx$, and use $\tilde{h}_k^m[n]$ to correct via Bayes' rule as

$$f_{h[n+1]}(h)=\frac{\hat{f}_{h_k^m[n+1]}(h)f_{v_k^m}(h-\tilde{h})}{\int_{\forall x}\hat{f}_{h_k^m[n+1]}(x)f_{v_k^m}(x-\tilde{h})dx}\ . \tag{3}$$

Because in this case the number of unobserved HMM states is infinite (the channel is a continuous variable), the denominator in the update equation (3) is an integral. This is in contrast with (2), where the denominator was a finite sum, reflecting the fact that in the previous section the number of unobserved states is finite. From a practical perspective there are a few cases where those integrals can be found in closed form (e.g. Gaussian channels). For the remaining cases, an approximate technique (such as grid-based Bayesian estimators or particle filters) should be used.

Before moving to the proposed RA approach, it is worth reiterating the main points so far. The CSI model adopted by the NC is distinct for the primary and secondary networks. The secondary CSI consists of the CR-to-CR link gains, which account for primary interference; whereas the primary CSI is formed either by the PU activity vector alone (interweave setup), or, it is augmented by CR-to-PU channel gains (underlay setup), which do not account for secondary interference. Moreover, secondary CSI is assumed perfectly known, so that information about the instantaneous realization is deterministic; whereas primary CSI is allowed to be uncertain, so that information (belief state) about the instantaneous realization is probabilistic.

B. Resources at the secondary network

This subsection introduces the design variables to be adapted as a function of the *overall* CSI that is collectively denoted by $\mathbf{h}$. Define further a Boolean scheduling variable w_k^m taking the value 1, if the mth CR is scheduled to transmit over the kth band, and 0 otherwise. When $w_k^m=1$, let p_k^m denote the instantaneous power transmitted over the kth band by the mth CR. Under bit error rate or capacity constraints, instantaneous rate and power variables are coupled. This rate-power coupling will be represented by the function $C_k^m(h_{k,2}^m, p_k^m)$. It will be assumed throughout that $C_k^m(h_{k,2}^m, \cdot)$ is given by Shannon's capacity formula $\log(1 + h_{k,2}^m p_k^m/\kappa_k^m)$, where κ_k^m represents the SNR-gap that depends on the coding scheme implemented [9]. For systems that implement a relatively small number of adaptive modulation and coding (AMC) modes, the last formula can be replaced with a piecewise linear function combining the rates achieved by the modes (see, e.g., [15], for details).

The secondary network operates in a block-by-block fashion, where the duration of each block corresponds to the coherence time of the fading channel. This way, per time slot n the NC uses the current CSI vector $\mathbf{h}$ to find w_k^m and p_k^m. Since $\mathbf{h}$ depends on n and $\{w_k^m, p_k^m\}$ depend on $\mathbf{h}$, $\{w_k^m, p_k^m\}$ will clearly vary across time. Henceforth, $\mathbf{h}$, $w_k^m(\mathbf{h})$, and $p_k^m(\mathbf{h})$ will be replaced by $\mathbf{h}[n]$, $w_k^m[n]$, and $p_k^m[n]$, whenever time dependence is to be stressed.

For this CR configuration, the goal is to develop adaptive RA algorithms leveraging the instantaneous secondary CSI and the generally uncertain primary CSI to determine which CR should transmit per band, and at what rate and power. An optimization problem will be formulated and solved in the ensuing section, first without interference constraints. Those will be incorporated in Section IV.

III. The Optimization Problem for Adaptive RA

To formulate the optimization problem associated with the novel RA approach, it is prudent to identify: i) the variables to be optimized, ii) the metric to be optimized, and iii) the constraints that must be satisfied. Section II-B identified $\{w_k^m, p_k^m\}$ as optimization variables. The metric to be optimized is the CRs' weighted sum-average rate given by $\bar{c} := \sum_{k,m} \mathbb{E}_{\mathbf{h}}\left[\beta^m w_k^m(\mathbf{h}) C_k^m(h_{k,2}^m, p_k^m(\mathbf{h}))\right]$, where $\mathbb{E}_{\mathbf{h}}$ stands for expectation over all CSI realizations, and $\beta^m > 0$ represents a user-dependent priority coefficient. Note that only the rate of CR user-channel pairs for which $w_k^m(\mathbf{h}) = 1$ participate in forming $\bar{c}$. Other objective functions such as sum-utility rate could be used without changing the basic structure of the solution; see, e.g., [25], [17] for further details. Regarding the constraints, $\{p_k^m\}$ must be obviously nonnegative, while $\{w_k^m\}$ must belong to the set $\{0, 1\}$. Moreover, since at most one CR transmits over each band k, it must hold that

$$\sum\nolimits_k w_k^m(\mathbf{h}) \leq 1, \quad \forall k. \tag{4}$$

If the left hand side (LHS) of (4) equals one, then one user accesses the channel (orthogonal access); otherwise, no user transmits either because all CR-to-CR channels are poor, or, because excessive interference is inflicted to the PU. The maximum average (long-term) power the mth CR can transmit is upper bounded; that is,

$$\mathbb{E}_{\mathbf{h}}\left[\sum\nolimits_k w_k^m(\mathbf{h}) p_k^m(\mathbf{h})\right] \leq \check{p}^m, \quad \forall m. \tag{5}$$

Under these considerations, the optimal RA emerges as the solution of the following problem:

$$\bar{c}^* := \max_{\{w_k^m(\mathbf{h}), p_k^m(\mathbf{h})\}} \sum\nolimits_k \mathbb{E}_{\mathbf{h}}\left[\beta^m w_k^m(\mathbf{h}) C_k^m(h_{k,2}^m, p_k^m(\mathbf{h}))\right] \tag{6a}$$

$$\text{s. to}: (4),\ (5),\ w_k^m(\mathbf{h}) \in \{0,1\},\ \text{and}\ p_k^m(\mathbf{h}) \geq 0; \tag{6b}$$

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

A. Optimal RA without interference constraints

Although the problem in (6) is non-convex, it can be trivially transformed (relaxed) into a convex one with identical Karush-Kuhn-Tucker (KKT) conditions[2]. In fact, the problem in (6) is a weighted sum-rate optimization of an uplink channel with orthogonal access. With π^m denoting the Lagrange multiplier associated with the constraint in (5), it has been shown that the solution of such a problem is (see, e.g., [16])

$$\varphi_k^m(p_k^m[n]) := \beta^m C_k^m(h_{k,2}^m[n], p_k^m[n]) - \pi^m[n] p_k^m[n], \tag{7}$$

$$p_k^{m*}[n] := \left[\arg\max_{p_k^m[n]} \varphi_k^m(p_k^m[n])\right]_0^\infty \tag{8}$$

$$= \left[\frac{\beta^m}{\pi_m[n]} - \frac{\kappa_k^m}{h_{k,2}^m}\right]_0^\infty \tag{9}$$

$$w_k^{m*}[n] := \mathbb{1}_{\{(m = \arg\max_l \varphi_k^l((p_k^{l*}[n]))) \wedge (\varphi_k^m(p_k^{m*}[n]) > 0)\}}. \tag{10}$$

Key to understanding the solution of (6) is the definition of the functional $\varphi_k^m(\cdot)$ in (7). Intuitively, (7) can be interpreted as a user-quality indicator where the rate is a reward, the power a cost, and β^m and $\pi^m[n]$ their corresponding prices. Analytically, $\varphi_k^m(x)$ represents the contribution to the *Lagrangian* of (6) if the transmit-power is $p_k^m[n] = x$ and $w_k^m[n] = 1$.

Based on the definition of $\varphi_k^m(p_k^m[n])$, equation (8) reveals that $p_k^{m*}[n]$ is found separately for each of the CR user-channel pairs. Similarly, (10) shows that finding the optimal scheduling variables $\{w_k^{m*}[n]\}_{m=1}^M$ per channel k, requires no information from channels other than k. These attractive properties hold thanks to the assumed orthogonal access in the secondary network and the definition of the objective in (6), both of which render the optimization problem in the dual domain separable across users and channels. Delving into the nuts-and-bolts of the optimal RA, consideration of a logarithmic rate-power function implies that (9) follows the well-known waterfilling solution [9]; and (10) manifests that the user scheduling is opportunistic (as desired) and greedy (only the user with *highest* quality must be scheduled per band).

Finally, it is worth emphasizing that although traditionally $\pi^m[n]$ is set to a constant value π^{m*}, corresponding to the

[2]There are two sources of non-convexity in (6). The first comes from $w_k^m \in \{0,1\}$, but can be relaxed to $w_k^m \in [0,1]$. As w_k^m only appears in linear terms, this relaxed solution coincides with the original one [16]. The second source corresponds to the monomials $w_k^m p_k^m$ and $w_k^m C_k^m$, for which one can introduce dummy implicit variables $\tilde{p}_k^m := w_k^m p_k^m$ in (6), and establish convexity using the properties of the perspective function. The resulting problem yields the same KKT conditions as those of (6); it is convex; and can be solved using a dual approach. Proofs are omitted due to space limitations, see e.g., [16], for details.

value that maximizes the dual function associated with (6) [3], alternative (stochastic) methods can be used. Such an alternative is attractive especially for the CR setup considered here, and will be explored in Section V.

IV. Interference Constraints

Different interference constraints are considered in this section, along with ways the optimal RA algorithms must be modified to account for such constraints. Attention is centered around constraints that limit the *probability of* CR transmitters to *interfere* with PU receivers. Other interference constraints (such as limiting the average interference power, or the rate loss for the primary network) could also be considered. Note that probabilistic constraints naturally account for CSI imperfections and, depending on their formulation, they can even exploit CSI variability.

When constraints on the probability of interference are included, there are two factors that significantly affect the design of optimum adaptive RA. The first is whether the interference constraints are formulated as instantaneous (short-term) or as average (long-term) constraints. The former require a certain probability of interference to hold for each and *every time instant*, while the latter allow PUs to be interfered at most over a maximum *fraction of time*. Clearly, instantaneous constraints are more restrictive than their average counterparts, which can exploit the so-called "cognitive diversity" of the primary CSI [27], [28]. As a result, the total rate transmitted by the CR transmitters will be higher in the latter case. On the other hand, optimization problems under instantaneous interference constraints are easier to solve because such constraints are amenable to simplification. Differently, average interference constraints cannot be easily simplified, and a dual approach is often invoked to deal with them. The second factor is whether an interweave or an underlay setup is in operation. The definition of interference in each setup is different. In fact, it will be shown that underlay formulations will render the problem non-convex and thus, challenge the development of an efficient solver able to achieve optimal performance. Remarkably, for the formulations in this paper, the optimization problem for the underlay setup exhibits zero duality gap, and the optimal solution can still be found with a moderate increase in terms of computational complexity.

Different formulations are considered for the probability of interference constraints because they will give rise to novel optimal resource allocation schemes. But also because, upon comparing the different solutions, it will be possible to understand the differences among the considered alternatives, both theoretically and from a performance perspective. The first formulation considered is the one involving instantaneous interference constraints for both interweave and underlay setups. Subsequently, the interweave and underlay setups will be investigated separately under average interference constraints. In all formulations, the schemes will be designed assuming imperfect CSI and then specialized for the case of perfect CSI. To simplify derivations, schemes for the underlay setup will be developed assuming that the PU is always active. The minor modifications required when this assumption does not hold are discussed in the closing remark of Section IV.

A. Short-term interference constraints

To keep the interference to the primary network under control, a *maximum probability of interference*, call it $\check{o}_k \in (0,1)$, is placed per band. Since this subsection focuses on short-term (instantaneous) interference constraints, such a limit is enforced $\forall n$.

1) Interweave networks: In this setup, interference occurs when $a_k[n] = 1$ (kth PU active over the kth band), and $\sum_m w_k^m[n] = 1$ (one CR transmits over the kth band). Then, the constraint on the probability can be formulated as $\Pr\{a_k[n]\sum_m w_k^m[n] = 1 \;|n\} \leq \check{o}_k \;\forall n$. At time n, the only random quantity in the previous expression is $a_k[n]$. Hence, the constraint can be written as

$$\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]\sum_m w_k^m[n]=1\}}\right] \leq \check{o}_k \;. \tag{11}$$

Taking into account that $\sum_m w_k^m[n]$ is Boolean and deterministically known at time n, the constraint can be rewritten as $\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right]\sum_m w_k^m[n] \leq \check{o}_k$. Clearly, the expectation on the LHS corresponds to the second entry of the belief vector $[\mathbf{f}_{\mathbf{a_k}}[n]]_2$. Thus, $\sum_m w_k^m[n] = 1$ only if $[\mathbf{f}_{\mathbf{a_k}}[n]]_2 \leq \check{o}_k$. This in turn implies that: i) there is no need to dualize the constraint, and therefore the expression for the link-quality indicator in (7) does not change; ii) the power allocation plays no role in the definition of the interference, and hence (8) still holds; and iii) to satisfy the interference constraint the optimal scheduling is now

$$w_k^{m*}[n] := \mathbb{1}_{\{[\mathbf{f}_{\mathbf{a_k}}[n]]_2 \leq \check{o}_k\}} \cdot \mathbb{1}_{\{(\varphi_k^m[n]=\max_l \varphi_k^l[n]) \;\wedge\; (\varphi_k^m[n]>0)\}}. \tag{12}$$

In words, the "winner CR" can transmit only if the probability of the channel being occupied is less than $\check{o}_k$. When the primary CSI is noisy and outdated, such a probability depends on the previous measurements and the accuracy of the sensor [cf. (1)-(2)]. On the other hand, if the primary CSI is perfect, $[\mathbf{f}_{\mathbf{a_k}}[n]]_2$ is either one or zero, and therefore transmissions can be allowed only if the channel is not occupied, i.e., it holds that $w_k^{m*}[n] := \mathbb{1}_{\{a_k[n]=0\}} \cdot \mathbb{1}_{\{(\varphi_k^m[n]=\max_l \varphi_k^l[n]) \;\wedge\; (\varphi_k^m[n]>0)\}}$.

2) Underlay networks: In this case, interference occurs when the received power at the PU due to CR transmissions exceeds a threshold Γ_k; i.e., if $w_k^m[n] > 0$ and $p_k^m[n]h_{k,1}^m[n] > \Gamma_k$, the constraint to be satisfied at every time n is $\Pr\{p_k^m[n]h_{k,1}^m[n] > \Gamma_k \;|n\} \leq \check{o}_k$. Since at time n the only random quantity is now $h_{k,1}^m[n]$, the constraint can be rewritten as

$$E_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}\right] \leq \check{o}_k, \quad \forall k \tag{13}$$

or equivalently, $E_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{h_{k,1}^m[n]<\Gamma_k/p_k^m[n]\}}\right] \geq 1-\check{o}_k$. Using the belief for the primary CSI at time n, it follows that $F_{h_{k,1}^m[n]}(\Gamma_k/p_k^m[n]) \geq 1-\check{o}_k$. Upon defining $p_k^{m'}[n]$ as the root of $\check{o}_k = F_{h_{k,1}^m[n]}(p_k^{m'}[n]/\Gamma_k)$, the last *inequality* amounts to the bound $p_k^m[n] \leq p_k^{m'}[n]$. In words, the interference constraint can be rewritten as a maximum (peak) power constraint.

This is very convenient, because while the original constraint in (13) is not convex, the maximum peak power constraint is convex. Since no multiplier is introduced to enforce the constraint, the Lagrangian remains the same, and thus $\varphi_k^m(p_k^m[n])$ is identical to that in (7). Similarly, the scheduling does not play a role in defining the interference so that the expression for $w_k^{m*}[n]$ in (10) holds true too. On the other hand, the expression for the optimum power in (8) needs to be updated because it has to satisfy the constraint $p_k^m[n] \leq p_k^{m'}[n]$. Such a (box) constraint can be easily handled by a scalar projection, which readily yields

$$p_k^{m*}[n] := \left[\arg\max_{p_k^m[n]} \ \varphi_k^m(p_k^m[n])\right]_0^{p_k^{m'}[n]}. \quad (14)$$

When the CSI is perfect, there is no uncertainty regarding $h_{k,1}^m[n]$; hence, the upperbound on the transmit-power is $p_k^{m'}[n] := h_{k,1}^m[n]/\Gamma_k$, and no interference is inflicted to the PU.

B. Long-term interference constraints in interweave systems

The previous subsection demonstrated that short-term interference constraints are easy to handle. In fact, for the interweave case the difficulty does not lie in how to satisfy the constraint, which is straightforward, but in estimating the probability of the PU being active. Here, the long-term probability of interfering with PUs is considered for the interweave setup. Since there is no easy way to enforce such a constraint, a dual relaxation will be used instead. It will be argued that regardless of CSI imperfections, the augmented optimization problem is convex and thus exhibits the following two properties: i) it can be tackled optimally using a dual approach, i.e., the duality gap is zero; and, ii) it is efficiently solvable.

Starting with the constraint formulation, recall that limiting the short-term probability of interference in an interweave setup consists in satisfying $\Pr\{\sum_m w_k^m[n]a_k[n] = 1 \ |n\} \leq \breve{o}_k$, or equivalently, $E_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]\sum_m w_k^m[n]=1\}}\right] \leq \breve{o}_k$ [cf. (11)]. In this section, the interest is in a long-term constraint so that all time instants are jointly considered. In this case, $\breve{o}_k$ can be viewed as an upperbound on the fraction of time instants for which interference occurs. This implies that the solution needs to satisfy the following condition

$$\mathbb{E}_{\mathbf{h}}\left[\mathbb{1}_{\{a_k\sum_m w_k^m(\mathbf{h})=1\}}\right] \leq \breve{o}_k, \quad \forall k. \quad (15)$$

Unlike (11), the expectation in (15) takes into account all CSI realizations. Note also that the LHS of (15) represents the joint probability of the PU being active *and* the NC scheduling one CR transmission. If one wants to limit the probability of one CR being active *provided that* the PU is active, then $\breve{o}_k$ must be multiplied (re-scaled) by the stationary probability of the kth band being occupied by the corresponding PU.

When (15) is incorporated into (6), the augmented problem is still convex because: i) (15) can be rewritten as $\mathbb{E}_{\mathbf{h}}\left[\sum_m w_k^m(\mathbf{h})\mathbb{1}_{\{a_k=1\}}\right] \leq \breve{o}_k$; and ii) the last inequality is convex (in fact linear) with respect to (w.r.t.) the only primary variable involved (i.e., w.r.t. w_k^m). As already mentioned, the approach to deal with the long-term interference constraint is to dualize it. To this end, let θ_k denote the Lagrange multiplier associated with the kth constraint in (15). The introduction of a new multiplier implies that the link-quality indicator needs to be redefined as

$$\begin{aligned}\varphi_k^m(p_k^m[n]) \ &:= \ \beta^m C_k^m(h_{k,2}^m[n], p_k^m[n]) - \pi^m[n]p_k^m[n] \\ &- \ \theta_k[n]\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right]. \quad (16)\end{aligned}$$

If the primary CSI is imperfect, then $\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right] = [\mathbf{f}_{\mathbf{a_k}}[n]]_2$; when perfect, it is simply $a_k[n]$. The only difference between the definitions of the quality indicator in (7) and (16) is that on top of considering the trade-off between rate and power, (16) also penalizes CR transmissions that are likely to cause interference whose "price" is multiplied by the instantaneous (short-term) probability of interference. The structure of the indicator in (16) also shows the role of the secondary CSI in the RA (first term in the sum), the role of the primary CSI (third term in the sum), as well as the impact of CSI imperfections (specific expression for $\mathbb{E}_{a_k[n]}[\mathbb{1}_{\{a_k[n]=1\}}]$).

Upon substituting (16) into (8) and (10), the *expressions* for the optimal power in (8) and the optimal scheduling in (10) still apply. However, this does not mean that *actual allocation* of resources is the same. While in the previous section transmissions never took place when $\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right] > \breve{o}_k$ [cf. (12)], the allocation in (16) allows for transmissions when the probability of interfering is high, provided that $\max_m\{\beta^m C_k^m(h_{k,2}^m[n], p_k^{m*}[n]) - \pi^m[n]p_k^{m*}[n]\}_{m=1}^M > \theta_k[n]\mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right]$. In other words, even if the scheduler knows that $a_k[n] = 1$, the secondary network can access the channel if the reward for the winner CR is high enough to exceed the cost of interfering represented by $\theta_k[n]$. Clearly, $\theta_k[n]$ is tuned to enforce that the percentage of interfering transmissions does not exceed the limit set by $\breve{o}_k$ (a higher price for interfering means that secondary transmissions will be less frequent). Finally, since the new term in $\varphi_k^m(p_k^m[n])$ does not depend on $p_k^m[n]$, the equivalence between optimum power in (8) and the waterfilling interpretation is still valid [cf. (9)]. Hence, an important difference between the short-term and the long-term solutions for the interweave paradigm is the way in which scheduling decisions are made. Optimal scheduling for the short-term formulation does not take into account the benefit for the winner SU. Focus is placed first on the PU. Only if the interference caused to the PU is below a threshold, the winner SU can transmit [cf. (12) and (7)]. Differently, optimal scheduling for the long-term formulation is more flexible and weights both the benefit for the SU and the harm caused to the PU [cf. (10) and (16)].

C. Long-term interference constraints in underlay systems

As in Section IV-B, the approach to deal with a long-term constraint on the probability of interfering with PUs in the underlay setup, is to dualize it. Regardless of CSI imperfections, the interference constraints here render the optimization problem non-convex. However, the problem at hand has two attractive features: i) since the functions causing non-convexity are averaged across time, existing results can be adapted to show that the duality gap is zero; and, ii) the problem can still be separated in the dual domain, so that minimization of the Lagrangian can be efficiently performed. More details will be given soon.

To formulate the interference constraint, recall that limiting the short-term probability of interference in an underlay setup amounts to bounding $\Pr\{p_k^m[n]h_{k,1}^m[n] > \Gamma_k \ |n\} \leq \check{o}_k$, or equivalently, $\mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}\right] \leq \check{o}_k$. For such a long-term bound, all channel realizations (time instants) must be accounted for, along with the CR causing interference. This can be accomplished by writing the constraint as

$$\mathbb{E}_{\mathbf{h}}\left[\sum\nolimits_m w_k^m(\mathbf{h})\mathbb{1}_{\{p_k^m(\mathbf{h})h_{k,1}^m(\mathbf{h})>\Gamma_k\}}\right] \leq \check{o}_k, \quad \forall k. \qquad (17)$$

Similar to (15), averaging over all $\mathbf{h}$ in (17) clearly implies that the constraint need not be satisfied for every CSI realization $\mathbf{h}$, but only on the average.

When (17) is incorporated into (6), the augmented problem is non-convex and thus challenging to solve. Remarkably, since the functions responsible for the non-convexity are averaged across time, existing results can be leveraged to show that the duality gap is zero. A rigorous proof can be obtained after adapting the results in either [21] or [19, App. A] for the problem at hand[3]. The fact of having zero-duality gap implies that dual methods can be used to relax the constraints without loss of optimality. However, the (unconstrained) Lagrangian is still non-convex and thus challenging to minimize. Next, the optimal RA for this scenario is developed and shown how it allows for an efficient minimization of the Lagrangian. Let ϑ_k denote the Lagrange multiplier associated with the kth constraint in (17). As in the interweave case, introduction of a new multiplier modifies the Lagrangian structure, and thus the link-quality indicator has to be modified accordingly as

$$\begin{aligned}\varphi_k^m(p_k^m[n]) \;&:=\; \beta^m C_k^m(h_{k,2}^m[n], p_k^m[n]) - \pi^m[n]p_k^m[n] \\ &-\; \vartheta_k[n]\mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}\right]. \qquad (18)\end{aligned}$$

As its counterpart in (16), the quality indicator in (18) considers both secondary and primary CSI and trades off rate reward with power and the cost of interfering. Indeed, the only difference between (16) and (18) is the expression for the instantaneous probability of interfering. Here, $\mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}\right]$ corresponds to $1 - F_{h_{k,1}^m[n]}(\Gamma_k/p_k^m[n])$ when CSI is imperfect, and to $\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}$ when CSI is perfect. As in Section IV-B, upon replacing (7) with (18), the expressions for the optimal power in (8) and the optimal scheduling in (10) remain the same. However, the equivalence between (8) and (9) no longer holds. This is because the third term in (18) depends on the transmit power and the optimal power in (9) is found by optimizing only the two first terms. In fact, the power optimization when (18) is substituted into (8) is challenging because the third term renders $\varphi_k^m(\cdot)$ non-concave. However, since optimizing $\varphi_k^m(\cdot)$ involves a single (scalar) variable, efficient methods to solve the optimization can be employed. Once $\{p_k^{m*}[n]\}_{m=1}^M$ are obtained, finding $\{w_k^{m*}[n]\}_{m=1}^M$ just requires the evaluation of closed-form expressions [cf. (10)]. In other words, because in the dual domain the problem can be separated across users and channels, optimizing the Lagrangian does not require optimizing a non-convex problem over a $2MK$-dimensional space; but instead, MK closed forms and MK *one-dimensional* non-convex problems must be solved. Recall that the factors enabling separability in the dual domain were the orthogonal access adopted by SUs within the CR network, and the definition of the metric to be optimized (summation across users) under the long-term constraints.

Indeed, when CSI is perfect, power optimization is straightforward and proceeds as follows. Let $p_k^{m''}[n] := h_{1,k}^m[n]/\Gamma_k$ be the maximum transmit-power by which interference is avoided; and let $\hat{p}_k^{m*}[n]$ denote the optimal power in (9), which ignores the interference constraint. Then, it holds that

$$p_k^{m*}[n] := \begin{cases} \hat{p}_k^{m*}[n] & \text{if } \ (\hat{p}_k^{m*}[n] < p_k^{m''}[n]) \ \vee \\ & (\varphi_k^m(\hat{p}_k^{m*}[n]) > \varphi_k^m(p_k^{m'}[n])) \\ p_k^{m''}[n] & \text{otherwise}\ . \end{cases} \qquad (19)$$

In words, if the cost of interfering is too high, transmit-power is constrained not to exceed $p_k^{m''}[n]$. However, if the cost of interfering is low enough (or the reward of the CR transmission is high enough), $p_k^{m*}[n]$ is allowed to exceed the upperbound.

When the primary CSI is imperfect, evaluating $F_{h_k^m[n]}$ dominates the complexity of power optimization. Unless $f_{h_k^m[n]}$ (which is the derivative of $F_{h_k^m[n]}$) is monotonic, the optimization is non-convex. However, if the number of stationary points of $f_{h_k^m[n]}$ is small (which holds true for most practical distributions), the number of local optima of $\varphi_k^m(\cdot)$ will be small too. In this case, all of them can be found, and the global optimum can be subsequently selected.

Remark 1: The schemes for the underlay setup in Sections IV-A.2 and IV-C have been developed under the assumption that PUs are always active, meaning that $a_k[n] = 1$. If this is not the case, interference only occurs if $p_k^m[n]h_{k,1}^m[n] > \Gamma_k$ and $a_k[n] = 1$. Assuming that $h_{k,1}^m[n]$ and $a_k[n]$ are independent, the only modification required is to replace the instantaneous probability of interference $\mathbb{E}_{h_{k,1}^m[n]}[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}]$ with $\mathbb{E}_{h_{k,1}^m[n]}[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}]\mathbb{E}_{a_k[n]}[\mathbb{1}_{\{a_k[n]=1\}}]$.

V. ESTIMATING THE OPTIMUM LAGRANGE MULTIPLIERS

Different methods can be used to estimate $\pi^m[n]$, $\theta_k[n]$, and $\vartheta_k[n]$. Since the duality gap is zero, one approach is to set $\pi^m[n] = \pi^{m*}$, $\theta_k[n] = \theta_k^*$ and $\vartheta_k[n] = \vartheta_k^*$, where $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$ are the values which optimize the dual function associated with (6). Clearly, the RA resulting after substituting those values into (7)-(19) would be the optimal solution for (6) [3]. The main limitations of this approach are that: i) $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$ need to be found through numerical search[4] which, at every step, requires averaging over all possible states of $\mathbf{h}$ (including channel imperfections); and ii) every time channel statistics or the number of users change, $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$ must be recomputed. Recently, alternative approaches that rely on stochastic approximation iterations have been proposed to obtain the multipliers [15], [24]. These

[3]The basic idea is that non-convexity comes from a constraint of the form $\mathbb{E}_{\mathbf{x}}[g(\mathbf{y}, \mathbf{x})]$, where $g(\mathbf{y}, \mathbf{x})$ is a non-convex function w.r.t. $\mathbf{y}$, and $\mathbf{x}$ is a random process with infinite support. Here $\mathbf{y}$ is the power; $\mathbf{x}$ is the CSI; and $g(\mathbf{y}, \mathbf{x})$ is $\mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}\right]$. The proof is omitted due to space limitations, but the reader is referred to [21] and [19, App. A] for further details.

[4]A classical dual subgradient with diminishing stepsize [3, Ch. 6] would work here; see, e.g., [16] for a related case.

approaches do not aim at the optimal $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$, but estimates that are updated at every time instant, and remain sufficiently close to $\{\pi^{m*}, \theta_k^*, \vartheta_k^*\}$. The main advantages of these approaches, especially for CR settings, are: i) their computational complexity is very low; and, ii) they can cope with non-stationary channels. The latter is very convenient when the PU transmitters are close to the SU receivers. The price paid is that the resulting RA schemes are slightly suboptimal. Specifically, with μ_π, μ_θ and μ_ϑ denoting sufficiently small, constant stepsizes, the following iterations yield the desired multipliers $\forall n$

$$\pi^m[n+1] = \left[\pi^m[n] - \mu_\pi(\check{p}^m - \textstyle\sum_k w_k^{m*}[n]p_k^{m*}[n])\right]_0^\infty \tag{20}$$

$$\theta_k[n+1] = \Big[\theta_k[n] - \mu_\theta(\check{o}_k - \mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right] \textstyle\sum_m w_k^{m*}[n])\Big]_0^\infty \tag{21}$$

$$\vartheta_k[n+1] = \Big[\vartheta_k[n] - \mu_\vartheta(\check{o}_k - \mathbb{E}_{a_k[n]}\left[\mathbb{1}_{\{a_k[n]=1\}}\right] \mathbb{E}_{h_{k,1}^m[n]}\left[\mathbb{1}_{\{p_k^m[n]h_{k,1}^m[n]>\Gamma_k\}}\right] \textstyle\sum_m w_k^{m*}[n]\Big]_0^\infty. \tag{22}$$

Recall that the expression for the instantaneous probability of interference in (21) and (22) is different for the cases of perfect and imperfect CSI. From an optimization point of view, the updates in (20)-(22) form an *unbiased* stochastic subgradient of the dual function of (6); see [3]. Using also that the updates in (20)-(22) are *bounded*, it can be shown that the sample average of the stochastic RA: i) is feasible; and, ii) incurs minimal performance loss relative to the optimal solution of (6). Rigorously stated, define $\mu := \max\{\mu_\pi, \mu_\theta, \mu_\vartheta\}$; $\bar{p}^m[n] := \frac{1}{n}\sum_{l=1}^n \sum_k w_k^{m*}[l]p_k^{m*}[l]$; $\bar{c}[n] := \frac{1}{n}\sum_{l=1}^n \sum_{k,m} \beta^m w_k^{m*}[l]C_k^m (h_{k,2}^m[l], p_k^{m*}[l])$; and $\bar{o}_k[n] := \frac{1}{n} \sum_{l=1}^n \sum_m w_k^{m*}[l]\mathbb{1}_{\{a_k[l]=1\}}$ (interweave) or $\bar{o}_k[n] := \frac{1}{n} \sum_{l=1}^n \sum_m w_k^{m*}[l]\mathbb{1}_{\{a_k[n]=1\}} \mathbb{1}_{\{p_k^{m*}[l]h_{k,1}^m[l]>\Gamma_k\}}$ (underlay). It then holds with probability one that[5] 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 $\mu \to 0$.

VI. Simulated tests

The default simulation parameters are as follows: $M = 5$, $K = 10$, $\beta^m = 1$, $\check{p}^m = 2$, $\kappa_k^m = 1$, $\check{o}_k = 4\%$, and $\Gamma_k = 0.5$. Amplitudes of the secondary links are Rayleigh (so that $h_{k,2}^m[n]$ are exponential) distributed, and the average SNR for all users and bands is $\mathbb{E}_{\mathbf{h}}[h_{k,2}^m] = 9$. The primary CSI model is $h_{k,1}^m[n] = |H_{k,1}^m[n]|^2$, where $H_{k,1}^m[n]$ is low-pass equivalent, complex Gaussian distributed (CGD) with zero mean and unit variance. Real and imaginary parts are independent, so that the amplitude is Rayleigh (and likewise $h_{k,1}^m[n]$ is exponential) distributed. The time correlation model is $H_{k,1}^m[n] = \sqrt{\rho}H_{k,1}^m[n-1] + \sqrt{1-\rho}z_k^m[n]$, with $\rho = 0.95$ and $z_k^m[n]$ white, CGD with zero mean and unit variance. Measurement noise $v_k^m[n]$ is CGD, with zero mean and variance 0.01. The NC senses $H_{k,1}^m[n]$ every $N_h = 6$ slots. The PU activity model is simulated with the following parameters: $Q_{00} = 0.95$, $Q_{01} = 0.10$, $Q_{10} = 0.05$, and $Q_{11} = 0.90$; $P_{FA} = 3\%$ and $P_{MD} = 2\%$; and the NC senses $a_k[n]$ every $N_a = 3$ slots. Since the optimality and feasibility of the developed schemes has been established theoretically, the simulation parameters and test cases have been chosen to illustrate relevant properties of the developed schemes.

Test Case 1: optimality and feasibility. Table I lists the average weighted sum-rate, power, and interference probability for an *interweave* CR network implementing nine different RA schemes. The first three solve (6) under a short-term interference constraint (STIC): S1) is a genie-aided scheme in which the true CSI is known; S2) is the optimal one developed in this paper that accounts for CSI imperfections; and S3) is a scheme implementing the RA as if the imperfect CSI were error free, so that it sets $a_k[n+n_a] = \tilde{a}_k[n]$, and $h_{k,1}^m[n+n_h] = \tilde{h}_{k,1}^m[n]$, for $n_a = 0, 1, ..., N_a - 1$, $n_h = 0, ..., N_h - 1$. The following three S4), S5) and S6) are the counterparts of S1), S2) and S3) under a long-term interference constraint (LTIC). For further comparison, three more are considered: S7) a scheme with no instantaneous information of the primary CSI, since it relies only on statistical CSI [1], [2]; S8) a scheme that solves (6) ignoring the interference constraints [17]; and S9) a scheme that accounts for CSI imperfections, and solves (6) guaranteeing that the *average* interfering power at the PUs is less than Γ_k [18], [27].

The results corroborate the analytical claims and illustrate the advantages of the developed algorithms. The novel schemes satisfy the constraints, while those ignoring CSI errors violate them; and outperform the suboptimal schemes, especially the one based on statistical knowledge of the CR-to-PU channels. It is worth noticing how S2 (long-term constraint) yields a higher maximum than S1 (short-term). Indeed, S1 over-satisfies the long-term interference constraint, while S2 satisfies the constraint tightly. Finally, the results confirm that the probability of interference estimated by the novel algorithms using the stochastic updates of the belief state corresponds to the actual one. Since our results guarantee that the long-term constraints are satisfied as $n \to \infty$, small discrepancies may occur when the number of simulated time instants is not high enough.

Table II is the counterpart of Table I for an *underlay* system. The additional scheme S7' (which is the counterpart of S7 for the case of a STIC) is also tested. Such a scheme is not appropriate for an interweave setup, but it has been considered for underlay CR networks [1]. The results confirm the previous findings. The main observation is that the underlay schemes achieve higher sum-rate than the interweave ones. This is reasonable because CRs in underlay operation have more opportunities to transmit (secondary transmissions with sufficiently low transmit-power do not cause interference even if the PU is active). Results in Tables III-V summarize further numerical tests assessing performance of the novel schemes over a wide range of parameter values, including a non-Markov model for the PUs activity [29] in Table V. These not only confirm the previous conclusions, but also show that when a more demanding setup is simulated (pronounced CSI imperfections and/or strict interference constraints) then: i) the impact of CSI imperfections on $\bar{c}$ is larger (cf. S4, S5, and S6);

[5] A proof of this result provided can be derived following the lines of [20], [17].

TABLE I

INTERWEAVE CR WITH $N_a = 3$, $N_h = 6$, $P_{FA} = 1\%$, $P_{MD} = 2\%$, $\check{o}_k = 4\%$, $\mathrm{Var}\{v_k^m[n]\} = 0.01$, $\Gamma_k = 0.5$. MEANING OF CODES USED IN ROW "COMMENTS": C1=STIC ENFORCED, LONG-TERM $\bar{o}_k$ SHOWN FOR ILLUSTRATIVE PURPOSES; C2=STIC OFTEN VIOLATED; C3=LTIC VIOLATED.

	S1	S2	S3	S4	S5	S6	S7	S8	S9
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5
$\bar{c}$	16.3	6.9	16.8	17.6	16.9	17.8	5.3	23.1	17.6
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	3.6%	4.0%	4.0%	7.3%	4.0%	39.6%	39.9%
$(1/K)\sum_k \bar{o}_k$ (estimated)	0.0%	0.05%	0.0%	4.0%	4.0%	4.0%	4.0%	$--$	$--$
Comments	C1	C1	C1,C2			C3		C2, C3	C2, C3

TABLE II

UNDERLAY CR WITH $N_a = 3$, $N_h = 6$, $P_{FA} = 1\%$, $P_{MD} = 2\%$, $\check{o}_k = 4\%$, $\mathrm{Var}\{v_k^m[n]\} = 0.01$, $\Gamma_k = 0.5$. MEANING OF CODES USED IN ROW "COMMENTS": SEE TABLE I.

	S1	S2	S3	S4	S5	S6	S7	S7'	S8	S9
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.2	1.0	0.5
$\bar{c}$	21.4	20.3	21.5	22.8	21.5	22.8	13.1	11.2	23.1	17.6
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	3.2%	14.4%	4.0%	2.7%	17.0%	4.0%	3.9%	37.5%	14.1%
$(1/K)\sum_k \bar{o}_k$ (estimated)	0.0%	3.4%	0.0%	4.0%	4.0%	4.0%	4.0%	4.0%	$--$	$--$
Comments	C1	C1	C1,C2			C3		C1	C2, C3	C2, C3

TABLE III

INTERWEAVE AND UNDERLAY CR WITH $N_a = 5$, $N_h = 10$, $P_{FA} = 5\%$, $P_{MD} = 3\%$, $\check{o}_k = 2\%$, $\mathrm{Var}\{v_k^m[n]\} = 0.02$, $\Gamma_k = 0.25$. MEANING OF CODES USED IN ROW "COMMENTS": SEE TABLE I. THE RESULTS FOR THE INTERWEAVE SETUP ARE SHOWN IN THE FIRST (TOP) HALF OF THE TABLE AND THE ONES FOR THE UNDERLAY SETUP ARE SHOWN IN THE SECOND (BOTTOM) HALF.

	S1	S2	S3	S4	S5	S6	S7	S7'	S8	S9
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	$--$	1.0	0.5
$\bar{c}$	16.7	3.2	16.5	17.2	10.7	16.9	4.3	$--$	23.2	17.6
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	7.0%	2.0%	2.0%	8.8%	2.0%	$--$	39.6%	39.2%
$(1/K)\sum_k \bar{o}_k$ (estimated)	0.0%	0.05%	0.0%	2.0%	2.0%	2.0%	2.0%	$--$	$--$	$--$
Comments	C1	C1	C1,C2			C3		$--$	C2, C3	C2, C3
Setup	Interweave									
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	0.3	0.1	1.0	0.5
$\bar{c}$	19.3	11.4	19.2	22.1	15.7	22.1	6.4	5.9	23.1	17.7
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	1.9%	21.8%	2.0%	2.1%	22.0%	2.1%	2.0%	38.0%	26.4%
$(1/K)\sum_k \bar{o}_k$ (estimated)	0.0%	2.0%	0.0%	2.0%	2.0%	2.0%	2.0%	2.0%	$--$	$--$
Comments	C1	C1	C1,C2			C3		C1	C2, C3	C2, C3
Setup	Underlay									

TABLE IV

RESULTS FOR DIFFERENT SIMULATION SETUPS. THE CR PARADIGM AND THE PARAMETERS WHICH ARE DIFFERENT FROM THOSE IN THE DEFAULT TEST CASE ARE DESCRIBED IN ROW "SETUP".

	S1	S2	S3	S4	S5	S6	S7	S7'	S8	S9
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	$--$	1.0	0.5
$\bar{c}$	8.8	4.1	8.8	9.4	9.0	9.4	3.7	$--$	11.4	7.7
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	3.8%	4.0%	4.0%	7.2%	4.0%	$--$	40.0%	36.8%
Setup	Interweave: $\mathbb{E}_{\mathbf{h}}[h_{k,2}^m] = 2$									
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.2	1.0	0.5
$\bar{c}$	10.3	9.6	10.3	11.2	10.5	11.2	5.4	3.9	11.5	7.7
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	1.7%	14.2%	4.0%	3.9%	15.7%	4.0%	3.6%	38.0%	13.5%
Setup	Underlay: $\mathbb{E}_{\mathbf{h}}[h_{k,2}^m] = 2$									
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	$--$	1.0	0.3
$\bar{c}$	10.2	3.6	10.3	11.0	10.4	11.0	2.8	$--$	14.7	8.8
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	3.8%	4.0%	4.0%	7.3%	4.0%	$--$	39.7%	40.0%
Setup	Interweave: $M = 5$, $K = 5$									
$(1/M)\sum_m \bar{p}_m$	0.8	0.8	0.8	0.8	0.8	0.8	0.8	$--$	0.8	0.5
$\bar{c}$	15.2	6.1	15.3	16.3	15.6	16.4	5.0	$--$	21.1	17.4
$(1/K)\sum_k \bar{o}_k$ (actual)	0.0%	0.05%	3.8%	4.0%	4.0%	7.3%	4.0%	$--$	39.8%	40.1%
Setup	Interweave: $\check{p}^4 = \check{p}^5 = 0.5$									

TABLE V
INTERWEAVE CR WITH DIFFERENT MODELS FOR THE ACTIVITY OF THE PUs. MM REPRESENTS THE MARKOV MODEL CONSIDERED IN THIS PAPER. PM REPRESENTS THE PARETO MODEL CONSIDERED IN [29]. THE STATIONARY DISTRIBUTION OF a_k IS THE SAME IN BOTH CASES. THE SIMULATION SETUP IS THE SAME THAN THAT IN TABLE I. MEANING OF CODES USED IN ROW "COMMENTS": SEE TABLE I.

	MM: S4	MM: S5	MM: S6	PM: S4	PM: S5	PM: S6
$(1/M)\sum_m \bar{p}_m$	1.0	1.0	1.0	1.0	1.0	1.0
$\bar{c}$	22.8	21.5	22.8	22.8	21.3	22.8
$(1/K)\sum_k \bar{o}_k$ (actual)	4.0%	3.6%	17.0%	4.0%	3.4%	17.1%
$(1/K)\sum_k \bar{o}_k$ (estimated)	4.0%	4.0%	4.0%	4.0%	4.0%	4.0%
Comments			C3			C3

ii) the interference constraints are more difficult to be satisfied; iii) the performance gain of the LTIC schemes relative to the STIC ones is more pronounced; and iv) the performance gain of our underlay schemes (S2, S5) relative to their interweave counterparts is larger too.

Test Case 2: dynamic behavior of the stochastic schemes. The dynamic behavior of the stochastic iterates is analyzed in this simulation, focusing on S2 in Table I. Figure 1 comprises four subplots, each depicting the evolution over time of a different subset of variables. Subplot (a) corresponds to the average power consumption $\bar{p}_m[n]$, and subplot (b) to the long-term probability of interference $\bar{o}_k[n]$ (cf. Section V). Dashed lines mark the performance when the optimal multipliers are known, while solid lines correspond to the proposed stochastic RA algorithms. Subplots (c) and (d) depict the instantaneous value of the Lagrange multipliers $\pi^m[n]$ and $\theta_k[n]$, respectively (in this case dashed lines correspond to the optimum multiplier values).

Subplots (a) and (b) show that the considered constraints are satisfied (with equality), and the stochastic RA converges in a few hundreds iterations. The Lagrange multipliers plotted in (c) and (d) suggest that after an initial phase during which the multipliers approach the optimal value, they never converge but hover around the optimum value. This is reasonable because (c) and (d) are instantaneous estimates while (a) and (b) are running averages. Finally, it is worth noting that in order to attain comparable rate of convergence, the stepsizes used to update $\pi^m[n]$ and $\theta_k[n]$ are considerably different ($\mu_\pi = 10^{-2}$ versus $\mu_\theta = 10^{-1}$).

Space limitations prevent inclusion of additional simulations confirming the merits of our schemes, but interested readers can access them online, along with the corresponding *Matlab* codes, at the first author's webpage [Accessed: 05/05/12] http://www.tsc.urjc.es/∼amarques/simulations/NumSimulations14.html.

VII. CONCLUSIONS

Stochastic resource allocation algorithms were developed for wireless cognitive radios communicating over fading links in interweave and underlay settings. The schemes were obtained as the solution of a weighted sum-rate maximization problem subject to maximum "average power" and "probability of interference" constraints. The probabilistic interference constraint was tailored to account for imperfections present in the sensing and CSI acquisition phase. Two types of interference constraints were investigated. The first one was a short-term constraint that took into account CSI imperfections to guarantee that the probability of interfering at any instant does not exceed a given threshold. The second one was a long-term constraint that capitalized on the diversity of the interfering channel to guarantee that the fraction of time during which interference occurs does not exceed a threshold. Although not all formulated problems were convex, enticingly they all turned out to have zero-duality gap, and could thus be solved with manageable complexity. It was shown that the optimal schemes maximize a functional which accounts for the quality of the secondary links (in terms of rate), the transmission power, and the probability of interfering with primary users. Several of those terms were multiplied by Lagrange multipliers whose value depended on the history of the system and the requirements of the primary and secondary networks. Stochastic algorithms were introduced to: i) estimate and predict the instantaneous (short-term) probability of interference; and, ii) estimate the optimum value of the multipliers. Future directions include accounting for alternative interference constraints, distributed implementations, as well as jointly optimizing the sensing and resource allocation tasks.

REFERENCES

[1] J.A. Ayala Solares, Z. Rezki, and M. S. Alouini, "Optimal power allocation of a sensor node under different rate constraints," *Proc. of IEEE Intl. Conf. on Commun.*, Otawa, Canada, 2012.

[2] S. Barbarossa, A. Carfagna, S. Sardellitti, M. Omilipo, and L. Pescosolido, "Optimal radio access in femtocell networks based on Markov modeling of interferers' activity," *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Signal Process.*, Prague, Czech Rep., May. 22-27, 2011.

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

[4] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2053-2071, May 2008.

[5] Y. Chen, G. Yu, Z. Zhang, H.-H. Chen, and P. Qiu, "On cognitive radio networks with opportunistic power control strategies in fading channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2752–2761, Jul. 2008.

[6] E. Dall'Anese, S.-J. Kim, G. B. Giannakis, and S. Pupolin, "Power control for cognitive radio networks under channel uncertainty," *IEEE Trans. Wireless Comm.*, vol. 10, no. 10, pp. 3541–3551, Oct. 2011.

[7] P. Di Lorenzo and S. Barbarossa, "A bio-inspired swarming algorithm for decentralized access in cognitive radio," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6160 – 6174, Dec. 2011.

[8] A. Ghasemi and E. S. Sousa, "Fundamental limits of spectrum-sharing in fading environments," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 649-658, Feb. 2007.

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

[10] 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, pp. 894–914, May 2009.

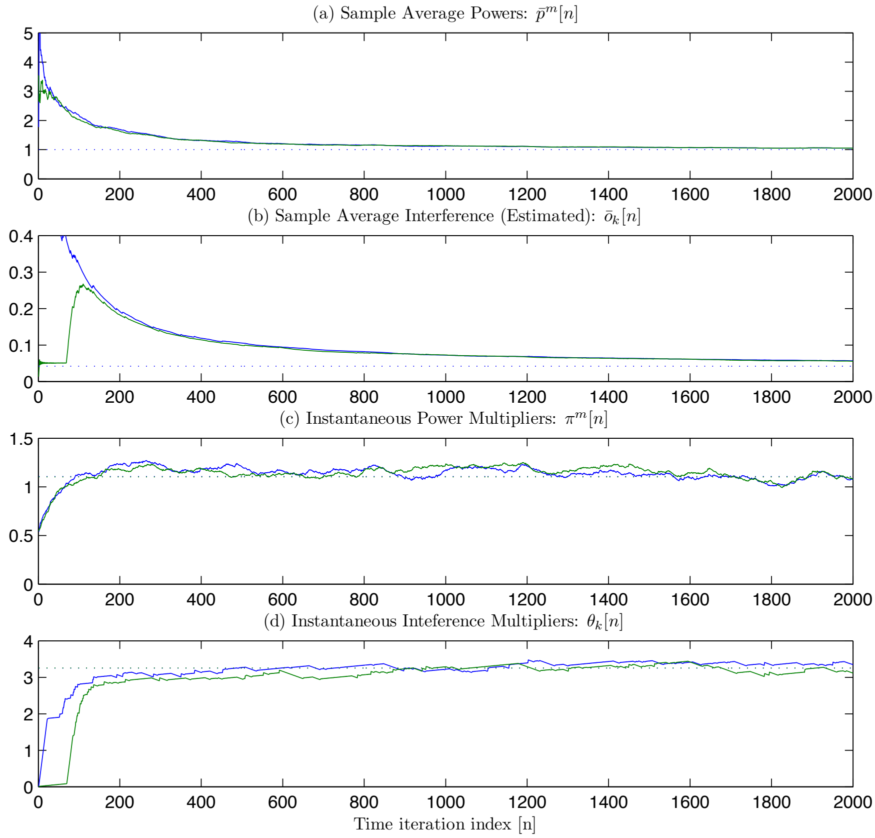


Fig. 1. Trajectories of different primal and dual variables for scheme S2 in Table I: (a) sample average power; (b) sample average interference; (c) (instantaneous) power multiplier; (d) (instantaneous) interference multiplier. To help visualization, only the multipliers of users $m = 1, 2$ and channels $k = 5, 6$ are plotted. Dashed lines correspond to the optimal (constant) values.

[11] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[12] Y. Y. He and S. Dey, "Power allocation in spectrum sharing cognitive radio networks with quantized channel information," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1644–1656, Jun. 2011.

[13] S. A. Jafar and S. Srinivasa, "Capacity limits of cognitive radio with distributed and dynamic spectral activity," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 529-537, Apr. 2007.

[14] X. Kang, Y.-C. 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.*, vol. 8, no. 2, pp. 940-950, Feb. 2009.

[15] A. G. Marques, X. Wang, and G. B. Giannakis, "Dynamic resource management for cognitive radios using limited-rate feedback," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3651–3666, Sep. 2009.

[16] A. G. Marques, G. B. Giannakis, and J. Ramos, "Optimizing orthogonal multiple access based on quantized channel state information," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 5023 – 5038, Oct. 2011.

[17] A. G. Marques, L. M. Lopez-Ramos, G. B. Giannakis, J. Ramos, and A. Caamano, "Optimal cross-layer resource allocation in cellular networks using channel and queue state information," *IEEE Trans. Vehic. Tech.*, vol. 61, no. 6, pp. 2789 – 2807, Jul. 2012.

[18] L. Musavian and S. Aissa, "Fundamental capacity limits of cognitive radio in fading environments with imperfect channel information," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3472–3480, Nov. 2009.

[19] K. Rajawat, N. Gatsis, and G. B. Giannakis, "Cross-Layer designs in coded wireless fading networks with multicast," *IEEE/AMC Trans. Networking*, vol. 19, no. 5, pp. 1276–1289, Oct. 2011.

[20] A. Ribeiro, "Ergodic stochastic optimization algorithms for wireless communication and networking," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6369–6386, Dec. 2010.

[21] A. Ribeiro and G. B. Giannakis, "Separation principles in wireless networking," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4488–4505, Sep. 2010.

[22] H. A. Suraweera, P. J. Smith, and M. Shafi, "Capacity limits and performance analysis of cognitive radio with imperfect channel knowledge," *IEEE Trans. Vehic. Tech.*, vol. 59, no. 4, pp. 1811 – 1822, May 2010.

[23] R. Urgaonkar and M. Neely, "Opportunistic scheduling with reliability guarantees in cognitive radio networks," *IEEE Trans. Mobile Comp.*, vol. 8, no. 6, pp.766–777, Jun. 2009.

[24] X. Wang, "Joint sensing-channel selection and power control for cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 958–967, Mar. 2011.

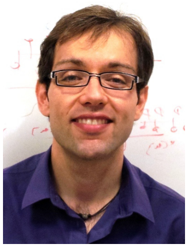
[25] X. Wang, G. B. Giannakis, and A. G. Marques, "A unified approach to QoS-guaranteed scheduling for channel-adaptive wireless networks," *Proc. IEEE*, vol. 95, no. 12, pp. 2410–2431, Dec. 2007.

[26] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, pp. 79-89, May 2007.

[27] R. Zhang, "On peak versus average interference power constraints for protecting primary users in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2112-2120, Apr. 2009.

[28] R. Zhang, Y.-C. Liang, and S. Cui, "Dynamic resource allocation in cognite radio networks: A convex optimization perspective," *IEEE Signal Process. Mag.*, vol. 27, no. 5, pp. 102–114, May 2010.

[29] X. Zhang and H. Su, "Opportunistic spectrum sharing schemes for CDMA-based uplink MAC in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 716-730, Apr. 2011.

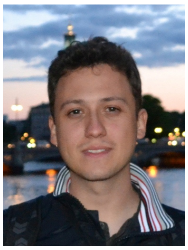


Antonio G. Marques (M'07) received the Telecommunication Engineering degree and the Doctorate degree (together equivalent to the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering), both with highest honors, from the Carlos III University of Madrid, Madrid, Spain, in 2002 and 2007, respectively.

In 2003, he joined the Department of Signal Theory and Communications, King Juan Carlos University, Madrid, Spain, where he currently develops his research and teaching activities as an Associate Professor. Since 2005, he has also been a Visiting Researcher at the Department of Electrical Engineering, University of Minnesota, Minneapolis.

His research interests lie in the areas of communication theory, signal processing, and networking. His current research focuses on stochastic resource allocation and network optimization, cognitive radios, and wireless ad hoc and sensor networks.

Dr. Marques' work has been awarded in several conferences and workshops.



Luis M. Lopez-Ramos (S'10) received the B.Sc. degree (with highest honors) in telecommunications engineering from the King Juan Carlos University of Madrid, Madrid, Spain, in 2010. He is currently working toward the M.Sc. degree in multimedia and communications with Carlos III University of Madrid. Since then, he has worked as a Research Assistant at the Department of Signal Theory and Communications, King Juan Carlos University. His research interests include signal processing for wireless networks, computer vision, cognitive radios and stochastic resource allocation schemes.

G. B. Giannakis (Fellow'97) received his Diploma in Electrical Engr. from the Ntl. Tech. Univ. of Athens, Greece, 1981. From 1982 to 1986 he was with the Univ. of Southern California (USC), where he received his MSc. in Electrical Engineering, 1983, MSc. in Mathematics, 1986, and Ph.D. in Electrical Engr., 1986. Since 1999 he has been a professor with the Univ. of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications in the ECE Department, and serves as director of the Digital Technology Center.

His general interests span the areas of communications, networking and statistical signal processing - subjects on which he has published more than 325 journal papers, 525 conference papers, 20 book chapters, two edited books and two research monographs. Current research focuses on compressive sensing, cognitive radios, cross-layer designs, wireless sensors, social and power grid networks. He is the (co-) inventor of 21 patents issued, and the (co-) recipient of 8 best paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, and the G. W. Taylor Award for Distinguished Research from the University of Minnesota. He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SP Society.

Javier Ramos received the B.Sc. and M.Sc. degrees in telecommunications engineering from the Polytechnic University of Madrid, Madrid, Spain, and the Ph.D. degree in 1995. Between 1992 and 1995, he was involved with several research projects with Purdue University, West Lafayette, IN, working on signal processing for communications. In 1996, he was a Postdoctoral Research Associate with Purdue University. From 1997 to 2003, he was an Associate Professor with Carlos III University of Madrid. Since 2003, has been with the King Juan Carlos University, Madrid, where he is currently a Professor and the Dean of the School of Telecommunications Engineering. His research interests include broadband wireless services and technologies and distributed sensing. Dr. Ramos received the Ericsson Award for the Best Ph.D. Dissertation on Mobile Communications in 1996.