JOINT SENSING AND RESOURCE ALLOCATION FOR UNDERLAY COGNITIVE RADIOS

Luis M. Lopez-Ramos, Antonio G. Marques, and Javier Ramos

King Juan Carlos University (Madrid, Spain), Dept. of Signal Theory and Communications

ABSTRACT

Effective operation of cognitive radios (CRs) requires sensing the spectrum and dynamic adaptation of the available resources according to the sensed information. Although sensing and resource allocation are coupled, most existing designs optimize each of the tasks separately. This work optimizes them jointly for an underlay CR paradigm. The formulation considers that secondary users adapt their power and rate based on the available imperfect channel state information, while taking into account the cost associated with acquiring such an information. The objective of the optimization is twofold: maximize the (sum-rate) performance of the CR and protect the primary users through an average interference constraint. Designing the sensing in our underlay paradigm amounts to decide what channel/frequency slots are sensed at every time instant. Partial observability of the channel state (due to noisy and outdated information) calls for (Bayesian) sequential estimators to keep track of the interference channel gains, as well as for dynamic programming tools to design the optimal schemes. Together with the optimal schemes, a simple approximate solution is also developed.

***Index Terms*—** Cognitive radio, underlay paradigm, sensing, dual decomposition, sequential estimation, dynamic programming.

1. INTRODUCTION

Cognitive radios (CRs) are a key technology to alleviate spectrum scarcity. When CRs are deployed, secondary users (SUs) have to sense their radio environment to optimize their communication performance while controlling the interference to primary users (PUs) [1]. In underlay CRs, sensing amounts to acquire the *channel state information* (CSI) needed to limit the power transmitted by the SU, so that the interference inflicted to the PU is kept under a prescribed limit. This typically amounts to acquiring the (fading) gains of the SU-to-PU channels (also known as interference channels). Due to lack of collaboration mechanisms between PU and CR systems, accurately estimating the SU-to-PU channels requires considerable effort [2]. However, *the cost of acquiring* the CSI is often not taken into account in the modeling; see [3, 4, 5, 6, 7] for some exceptions. As a result, a careful design of the sensing policy is critical to guarantee an efficient operation of the CR. This paper optimizes the sensing and resource allocation (RA) tasks for an underlay CR model *jointly*. Uncertainties on the sensed CSI and sensing cost will be taken into account during the RA, while the actual benefit of the CSI for the SUs will be taken into account during the sensing phase.

Some important challenges to optimize the sensing and RA jointly are: (C1) the need of the RA algorithms to deal with imperfect CSI (noisy and outdated) that renders the exact interference caused to PUs uncertain; (C2) the inability to estimate the totality of the PU-SU-channel-time grid, due to the scarcity of resources (power, time or hardware availability); and (C3) the coupling between sensing and transmission *resources*. To deal with (C1), several works aim to control the average interference using a probabilistic representation of the state information of the primary network (SIPN) [2, 7]. Adaptive stochastic algorithms provide robustness to non-stationarities and lack of knowledge of the channel distributions. See also [8] for a different approach to cope with estimation errors. To deal with (C2), advanced sensing schemes aim at optimally selecting the subset of sensed channels [9, 10, 7]. Moreover, when the SIPN exhibits time correlation, the information acquired can be reused ahead on time (accounting for the fact that information gets outdated). These schemes are usually designed using dynamic programming (DP) tools such as partially observable Markov decision processes (POMDPs) [7, 9]. Regarding (C3), RA in underlay CRs has been extensively investigated. In [11] an RA framework that considers both interference constraints for PUs and QoS constraints for SUs is presented; power is optimized jointly with admission control. The optimal RA strategies to achieve the ergodic and outage capacity of the SU fading channel is studied in [12] under different types of power constraints and fading channel models. See, e.g., [13, 8, 14, 2, 15] for other relevant setups. All those works consider that the sensing is given and, at best, account for the SIPN uncertainties (quantized, noisy, outdated) by making the RA aware of such imperfections. The number of works that aim to globally optimum RA and sensing by implementing a joint optimization is much smaller; see, e.g., [9, 16, 10, 6, 7, 17], all for interweave setups. When a joint design is implemented, the decision of what time instants/users/channels to sense has to take into account what the RA is going to do with such information, as well as the impact on CR performance for current and future time instants. As a result, the analytical complexity of the problem and the computational burden to obtain the optimal schemes increase considerably.

To the best of our knowledge, no previous work has addressed the joint design of sensing and RA for *underlay* CRs. The underlay setup is more challenging than the interweave setup, which only requires to know whether a frequency band is occupied or not. Not only the variables to estimate are continuous in the underlay setup, but their number is much higher (all SU-to-PU pairs in all frequency bands). Uncertainty and time correlation in the SIPN call for (Bayesian) sequential estimators to keep track of the interference channel and DP/POMDP tools to design the optimal schemes.

Our design approach is similar to one followed in our previous work [7, 17] for interweave CRs. We first design the RA for any sensing scheme and, then, design the optimal sensing taking into account the optimal RA. Since the modifications in the RA to account for the sensing cost are relatively simple, the main novelty is on the design of the sensing schemes. The main contributions of this paper are: the formulation of a joint optimization of RA and sensing for an underlay CR; the design of an algorithm that, leveraging dual decomposition and DP/POMDP tools, solves the joint optimization;

This work was partially supported by the Spanish Ministry of Science (FPU Grant AP2010-1050) and the EU-FP7 (ICT-2011-9-TUCAN3G).

and the design of a low-complexity algorithm that, using a greedy (myopic) approach, approximates the optimal solution. Our paper must be viewed as a first step to developing low-complexity approximations to the optimal solution.

The paper is organized as follows. Section 2 presents the system setup, SIPN and state information of the secondary network (SISN) models, design variables, and the constraints to be satisfied. The problem is formulated in Section 3. Section 4 solves the problem and analyzes its complexity. Section 5 presents a simple suboptimal solution. Preliminary numerical simulations evaluating the performance of the algorithms are provided in Section 6.

2. SYSTEM MODEL

A CR with M SUs (indexed by m) is considered. The frequency band used by the CR is divided into K frequency-flat orthogonal subchannels (indexed by k), so that if a SU is transmitting, no other SU can be active in the same subchannel. No constraints are imposed on the number of channels that can be accessed by a user. For simplicity, we assume that there is always exactly one active PU per channel. Extensions to scenarios where these assumption(s) do not hold can be handled with a moderate increase in complexity. Each SU can obtain (imperfect) measurements of the channel gain between itself and the PUs. More precisely, at every time slot (indexed by n) the following three tasks are run *sequentially* by the CR: T1) the SISN is acquired; T2) based on the output of T1 (and previous measurements) a set of users are selected to measure their interference links; T3) the outputs of T1 and T2 are used to find the optimal RA for instant n. This section describes the model for the SISN and SIPN; the variables to be designed; and the constraints that such variables need to satisfy.

Starting with the SISN, the instantaneous fading coefficient of the channel between the mth secondary transmitter-receiver pair in the kth channel at time n is denoted as $h_{k,2}^m[n]$. This variable is normalized with respect to noise and PU interference. Regarding the SIPN, the noise-normalized instantaneous fading coefficient of the interference channel between the mth SU and the kth PU is denoted as $h_{k,1}^m[n]$. Every time that the mth SU is required to obtain measurements from its interference channels, it has to pay a *power cost* denoted by q^m (other sensing costs can also be accommodated into our formulation [7]). The instantaneous value of $h_{k,1}^m[n]$ will not be assumed perfectly known because of: i) outdated information (to save power, the interference channels are not sensed at every n); and ii) errors due to noisy measurements. As a consequence, instead of the true value of the channel gain (perfect SIPN), only statistical information about it is available (probabilistic SIPN).

Let $\tilde{h}_{k,1}^m[n]$ denote the observation (output of the sensing task, possibly corrupted by noise) of $h_{k,1}^m[n]$. The CR relies on the dynamics of $h_{k,1}^m[n]$ to track the SIPN. Let us define the Boolean variable $s^m[n]$, which is 1 if at time n the mth SU takes measurements $\tilde{h}_{k,1}^m[n]$, and 0 otherwise. While $s^m[n]$ will depend on measurements acquired in the past [cf. T1], the power allocation and transmission scheduling will also leverage the newly available measurements [cf. T3]. Let $f_k^m(h_{k,1}^m[n]|n-1)$ denote the information about $h_{k,1}^m$ before sensing (prediction density). Similarly, let $f_k^m(h_{k,1}^m[n]|n)$ denote the information about $h_{k,1}^m$ after sensing (filtering density)[1]. To use a compact notation, these densities will be assumed to belong to the same family and will be represented by their parameters. Let $\hat{F}_k^m[n]$ be the parameter vector of the prediction density at time n, and $F_k^m[n]$ that of the filtering density. This way, $f_k^m(h_{k,1}^m[n]|n-1) := f(h_{k,1}^m[n];\hat{F}_k^m[n])$ and $f_k^m(h_{k,1}^m[n]|n) := f(h_{k,1}^m[n];F_k^m[n])$. The stochastic filter that tracks the SIPN works as follows [19]. The prediction density parameters at time n are deterministically computed at the *prediction step* from the previous filtering density parameters:

$$\hat{F}_k^m[n] := \mathcal{P}(F_k^m[n-1]). \tag{1}$$

The filtering density at time n will depend on $s^m[n]$. If $s^m[n]=0$, then $F_k^m[n]=\hat{F}_k^m[n]$; if $s^m[n]=1$, then

$$F_k^m[n] := \mathcal{U}(\hat{F}_k^m[n], \tilde{h}_{k,1}^m[n]). \tag{2}$$

There exist different alternatives to model the stochastic process $h_{k,1}^m$. Here, the time dynamics of the complex-valued secondary-primary channel gain $h_{k,1}^m[n]$ are described by an auto-regressive (AR) model with circularly-symmetric complex normal (CSCN) innovations and CSCN noise. As a result, the parameter vectors correspond to the mean and variance of the densities, and the prediction and correction steps of the channel estimation can be effected by a standard Kalman filter [19]. Since the time variability of the SISN is considered faster than that of the SIPN, $h_{k,2}^m[n]$ will be considered i.i.d. across time. As stated in [15], such a heterogeneous system information model is well suited for scenarios where the mobility of the PUs is low and sensing the SIPN is more difficult than sensing the SISN.

Next, we introduce the design variables $w_k^m[n]$ (scheduling coefficients), $p_k^m[n]$ (transmit power), and $s^m[n]$ (sensing decision, already described). Coefficients $w_k^m[n]$ effect the orthogonal access among SUs. Specifically, $w_k^m[n]$ is 1 if the mth SU is scheduled to transmit into the kth band at time n and 0 otherwise. Moreover, if $w_k^m[n]=1$, $p_k^m[n]$ denotes the instantaneous nominal power transmitted over the kth band by the mth SU. This means that power $p_k^m[n]$ is consumed when $w_k^m[n]=1$. Under bit error rate or capacity constraints, instantaneous rate and power variables are coupled. This rate-power coupling will be represented by the non-decreasing function $C_k^m(h_k^m[n], p_k^m[n])$ and β^m will denote the benefit (price) associated with the rate.

The last step is to describe the constraints that the aforementioned variables need to satisfy. The sensing decision variable is binary, so that $s^m[n] \in \{1,0\}$. Powers are non-negative, so that $p_k^m[n] \geq 0$. Moreover, orthogonal access requires

$$w_k^m[n] \in \{0,1\} \text{ and } \textstyle\sum_m w_k^m[n] \leq 1. \tag{3}$$

The average (long-term) power the mth SU can consume (including the power devoted to transmit and the power devoted to estimate the interference channel gains) is upper bounded, that is, $\forall m$

$$\lim_{N\to\infty}\sum_{n=1}^{N}\gamma^n\mathbb{E}\left[q^m s^m[n] + \sum_k w_k^m[n]p_k^m[n]\right] \leq \lim_{N\to\infty}\sum_{n=1}^{N}\gamma^n\check{p}^m, \tag{4}$$

where $0<\gamma<1$ is a discount factor that is typically included in infinite horizon formulations to facilitate the design of the optimal schemes and accommodate potential non-stationarities [18]. Note also that the right hand side of (4) is equivalent to $\frac{\check{p}^m}{1-\gamma}$. The allocated power will generate interference to PUs. Since an underlay setup is considered, each time a SU transmits in channel k, the interference generated at the PU receiver is $h_{k,1}^m[n]p_k^m[n]$. To protect the PUs, a limit on the average (long-term) interference at each PU is enforced. This amounts to require for all k

$$\lim_{N\to\infty}\sum_{n=1}^{N}\gamma^n\mathbb{E}\left[\sum_m w_k^m[n]h_{k,1}^m[n]p_k^m[n]\right] \leq \lim_{N\to\infty}\sum_{n=1}^{N}\gamma^n\check{o}_k, \tag{5}$$

[1] The prediction and filtering densities play the role of the *pre-decision* (prior) and *post-decision* (posterior) beliefs, respectively [17, 7, 18, 19].

Here, we controlled interference by limiting the *average interfering power* at the PU receiver [11]. This keeps the modeling simple and leads to a convex constraint. Alternative metrics can be used to control interference (e.g. outage probability [15]), provided that the increase in computational complexity can be afforded.

3. PROBLEM FORMULATION

The last step to formulate the optimization problem is to identify the metric to be maximized. In this work, the average sum rate achieved by the secondary network will be maximized. With $\mathcal{X} := \{s^m[n], w_k^m[n], p_k^m[n] \,|\forall m, k, n\}$, the optimal joint design is then

$$\max_{\mathcal{X}} \lim_{N\to\infty} \sum_{n=1}^{N} \gamma^n \mathbb{E}\left[\sum\nolimits_{k,m} \beta^m w_k^m[n] C_k^m(h_{k,2}^m[n], p_k^m[n])\right] \quad \text{(6a)}$$

$$\text{s. to : (3), (4), (5), } p_k^m[n] \geq 0,\ s^m[n] \in \{0,1\}. \quad \text{(6b)}$$

The two main issues that render this problem challenging to solve are: i) The design variables $w_k^m[n]$ and $s^m[n]$ are binary, so that the complexity to optimize over them is combinatorial; and ii) The value of some design variables at time n has an impact on the state variables at instants $n' \geq n$ (specifically, $s^m[n]$ has an impact on future beliefs through $F_k^m[n]$) – as a consequence, solving (6) optimally requires using DP tools.

Regarding the first challenge, the combinatorial complexity associated with optimizing over $w_k^m[n]$ can be bypassed by relaxing the binary constraint to its convex counterpart $w_k^m[n] \in [0,1]$. Such a relaxation can be shown optimal because $\{w_k^m[n]\}$ are present only in linear terms and because $\{w_k^m[n]\}$ do not have an impact on the future state variables; see, e.g., [15] for details. Unfortunately, that is not true for $s^m[n]$ and, hence, the associated complexity remains combinatorial. The optimal solution is presented in the next section, while Section 5 presents a low-complexity approximation.

4. OPTIMAL SOLUTION

After dualizing the long-term constraints (4) and (5), the optimization of $\{w_k^m[n]\}$ and $\{p_k^m[n]\}$ can be separated across time and channels. This fact, together with other properties of (6) will be leveraged to decrease the computational complexity required to solve the DP. The critical step is to tackle the optimization in two stages: i) finding the optimal $\{w_k^m[n]\}$ and $\{p_k^m[n]\}$ for *any* sensing policy; and ii) substituting the output of (i) into (6) and solving for the optimal $\{s^m[n]\}$. Note that this does not entail a loss of optimality because the solution in (i) is a function of the sensing policy, which is later optimized in (ii). Mathematically, for a generic function $f(x,y)$, the approach amounts to find $(x^*, y^*) = \arg\min_{x,y} f(x,y)$ as follows: i) $x^*(y) := \arg\min_x f(x,y)$ and ii) $y^* = \arg\min_y f(x^*(y), y)$.

4.1. Optimal RA

The optimization carried out in the first step yields a problem of the same form than that solved in [7]; for this reason, the optimal solution is given here directly. The optimal solution to the problem at hand consists in defining a link quality indicator (LQI) $\varphi_k^m(p)$, optimizing it with respect to the power for every user-channel pair, and selecting for transmission the SU with the highest LQI in each channel. The LQI for the problem at hand is:

$$\varphi_k^m(p) := \beta^m C(h_{k,2}^m[n], p) - (\pi^m + \theta_k \mu_k^m[n])\, p \qquad (7)$$

where $\mu_k^m[n] := \mathbb{E}\big[|h_{k,1}^m[n]|^2 \big| F_k^m[n]\big]$ is the expected power gain of the interference channel (according to the post-decision belief); and π^m and θ_k are the Lagrange multipliers associated with constraints (4) and (5), respectively[2]. To optimize the RA for instant n, select $p_k^{m,\star}[n] := \arg\max_p \varphi_k^m(p)$, and $w_k^{m,\star}[n] := \mathbb{1}_{\{\varphi_k^m(p) = \max_{q,p} \varphi_k^q(p)\}}$, where $\mathbb{1}_{\{\cdot\}}$ is the indicator function. Note that (7) can be expressed in closed form for several choices of $C(\cdot)$. For example, if $C(\cdot)$ is Shannon's capacity, then $p_k^{m,\star}$ takes the form of the water-filling solution [20, 12].

4.2. Optimal sensing

Leveraging the expressions for the optimal RA, we now solve for the optimal $s^m[n]$. First, we define the *instantaneous* reward $R[n]$, which accounts for the terms *at time* n that depend on $s^m[n]$:

$$R[n] := \sum\nolimits_k \max_m \varphi_k^{m,\star}(\mu_k^m[n]) - \sum\nolimits_m \pi^m q^m s^m[n], \qquad (8)$$

where $\varphi_k^{m,\star}(\mu_k^m[n])$ is the optimal value of $\varphi_k^m[n]$ for a given $\mu_k^m[n]$. Note that $\varphi_k^{m,\star}(\mu_k^m[n])$ depends on $\mathbf{s}[n]$ because the SIPN $\mu_k^m[n]$ depends on $\mathbf{s}[n]$ [cf. (2)].

After substituting the optimal RA and (8) into the Lagrangian of (6), the maximization boils down to

$$\max_{\{\mathbf{s}[n] \in 2^M\}\forall n} \lim_{N\to\infty} \sum_{n=1}^{N} \gamma^n \mathbb{E}\left[R[n] \middle| \mathbf{s}[n]\right]. \qquad (9)$$

To stress that the sensing decisions of all users have to be jointly optimized, the notation $s^m[n] \in \{0,1\}\ \forall m$ has been replaced with $\mathbf{s}[n] \in 2^M$. The coupling exists because the sensing decision for user m affects its probability (and hence, also the other users' probabilities) of being scheduled.

The main differences between (9) and the original formulation in (6) are that now: i) as a result of the Lagrangian relaxation of the DP, the objective has been augmented with the terms accounting for the dualized constraints; ii) the only remaining optimization variables are $\mathbf{s}[n]$; and iii) because the optimal RA fulfills the constraints (3)-(5) and $p_k^m[n] \geq 0$, the only remaining constraint is $\mathbf{s}[n] \in 2^M$ (standard DP algorithms usually assume countable action spaces).

The problem falls into the class of POMDP because state transitions and average rewards only depend on the current state-action pair, and the system state is not known perfectly. Only an observation (affected by noise or missing data) of the state is available instead [21]. In this model, the partially observable variable is $h_{k,1}^m[n]$. The belief variable required to solve this POMDP is constituted by the prediction and filtering densities associated with $h_{k,1}^m[n]$.

To solve for $\mathbf{s}^*[n]$, we derive the Bellman equations [18] associated with (9). The objective is split into present and future rewards, yielding

$$\mathbf{s}^*[n] = \underset{\mathbf{s}\in 2^M}{\arg\max}\Big\{\mathbb{E}\big[R[n]\big|\mathbf{s}[n]=\mathbf{s}\big] + \sum_{t=n+1}^{\infty} \gamma^{t-n}\mathbb{E}\big[R[t]\big|\mathbf{s}[n]=\mathbf{s}\big]\Big\}. \quad (10)$$

To account for the effect of current actions in future instants, the value function $V(\mathbf{h_2}[n], \hat{\mathbf{F}}[n])$ is introduced (where $\hat{\mathbf{F}}[n]$ collects $\hat{F}_k^m[n] \forall (k,m)$ and $\mathbf{h_2}[n]$ collects $h_{k,2}^m[n] \forall (k,m)$). It quantifies the expected sum reward for all future instants. Since the latter is an infinite horizon DP with $\gamma < 1$, the optimal value function is stationary and its existence is guaranteed [18]. Similarly to [7], since

[2] Since, after relaxation, the RA problem has zero duality gap, there exists a constant (stationary) optimal value for each multiplier [7]. In this work, the optimal values of $\{\pi^m, \theta_k\}$ are assumed to be known.

the $h^m_{k,2}[n]$ are considered i.i.d. across time and independent of $s^m[n]$, the Bellman equations that drive the optimal sensing can be expressed in terms of $\bar{V}(\hat{\mathbf{F}}[n]) := \mathbb{E}_{\mathbf{h_2}}[V(\mathbf{h_2}[n], \hat{\mathbf{F}}[n])]$:

$$\mathbf{s}^*[n] = \arg\max_{\mathbf{s}\in 2^M} \left\{ \mathbb{E}_{\tilde{\mathbf{h}}} \left[R[n] + \gamma \bar{V}(\hat{\mathbf{F}}[n+1]) \middle| \mathbf{s}[n]=\mathbf{s} \right] \right\} \quad (11)$$

$$\bar{V}(\hat{\mathbf{F}}[n]) = \mathbb{E}_{\mathbf{h}_2} \left[\max_{\mathbf{s}\in 2^M} \left\{ \mathbb{E}_{\tilde{\mathbf{h}}} \left[R[n] + \gamma \bar{V}(\hat{\mathbf{F}}[n+1]) \middle| \mathbf{s}[n]=\mathbf{s} \right] \right\} \right] \quad (12)$$

where $\mathbb{E}_{\tilde{\mathbf{h}}}$ is the expectation over the distribution of $\{\tilde{h}^m_{k,1}\} \forall (m,k)$. The only remaining step to design the sensing scheme is to design an algorithm to compute $\bar{V}(\hat{\mathbf{F}}[n])$. There exist different alternatives that exploit the recursive definition in (12) to accomplish this task [18]. Space limitations prevent us to delve into the details of such algorithms, but it is important to stress that (even after leveraging the problem structure) their computational complexity is very large.

5. APPROXIMATE SOLUTION

The two main sources of complexity to find $\mathbf{s}^*[n]$ are: i) during the initialization phase, the multidimensional function $\bar{V}(\cdot)$ needs to be estimated iteratively using a Monte Carlo approach and ii) at every time instant, an exhaustive search over 2^M needs to be implemented. Since (i) is run off line only once, we focus on reducing the online complexity in (ii). In particular, we use a greedy approach under which users are sequentially selected to measure the channel. We start by supposing that no user senses the channel and sequentially set $s^m[n] = 1$ for the SU that yields the highest (positive) expected reward. The algorithm stops either when none of the remaining SUs yields a positive reward, or when all users are scheduled to sense the channel. The approximation is well justified because channels across SUs are not correlated. Algorithm 1 lists the main steps of the algorithm, with $\mathbf{0}$ and $\mathbf{1}$ denoting the all-zeros and all-ones vectors and $\mathbf{e}_m$ the mth canonical $M \times 1$ vector.

Algorithm 1 Greedy approximation to the optimal sensing policy.

1: $\tilde{\mathbf{s}} \leftarrow \mathbf{0}$ and $\mathcal{M} \leftarrow \{1, \ldots, M\}$
2: **repeat**
3: $\quad m^\star \leftarrow \arg\max_{m\in\mathcal{M}} \mathbb{E}[R[n] + \gamma \bar{V}(\hat{\mathbf{F}}[n+1]) | \mathbf{s}[n] = \tilde{\mathbf{s}} + \mathbf{e}_m]$
4: $\quad \Delta R \leftarrow \mathbb{E}[R[n] + \gamma \bar{V}(\hat{\mathbf{F}}[n+1]) | \mathbf{s}[n] = \tilde{\mathbf{s}} + \mathbf{e}_{m^\star}] - \mathbb{E}[R[n] + \gamma \bar{V}(\hat{\mathbf{F}}[n+1]) | \mathbf{s}[n] = \tilde{\mathbf{s}}]$
5: $\quad$ **if** $\Delta R > 0$, **then** $\tilde{\mathbf{s}} \leftarrow \tilde{\mathbf{s}} + \mathbf{e}_{m^\star}$ and $\mathcal{M} \leftarrow \mathcal{M} \setminus \{m^*\}$
6: **until** $\Delta R < 0$ **or** $\tilde{\mathbf{s}} = \mathbf{1}$
7: **return** $\mathbf{s}[n] := \tilde{\mathbf{s}}$

The expectations in line 3 (which are taken over $\tilde{\mathbf{h}}$) can be run efficiently using a Monte Carlo method. Since the imperfections in $h^m_{k,1}[n]$ are independent across (m,k), each of the M expectations in line 3 can be implemented with complexity $O(MKN)$, where N is the number of random realizations per $\tilde{h}^m_{k,1}[n]$. Then, the online complexity of the overall algorithm is $O(M^3KN)$, because the *repeat* loop is executed at most M times. Note that Algorithm 1 assumes that $\bar{V}(\cdot)$ has been computed off line. If the associated burden cannot be afforded, further simplified algorithms can be developed either by approximating $\bar{V}(\cdot)$, or just by dropping it (myopic policy).[3].

[3]A myopic policy ignores the impact of the sensing decision on future time instants, focusing only on maximizing the instantaneous reward. Mathematically this amounts to setting $\bar{V}(\cdot) = 0$.

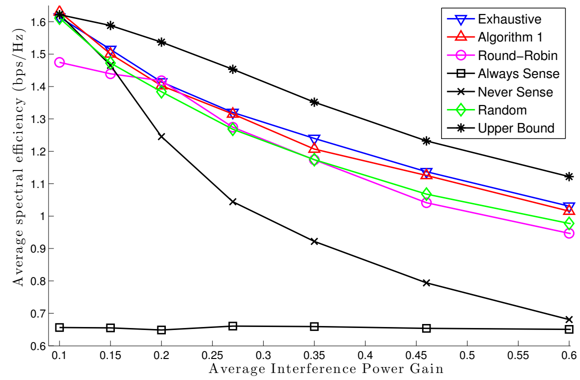

Fig. 1.1. Results for Test Case 1

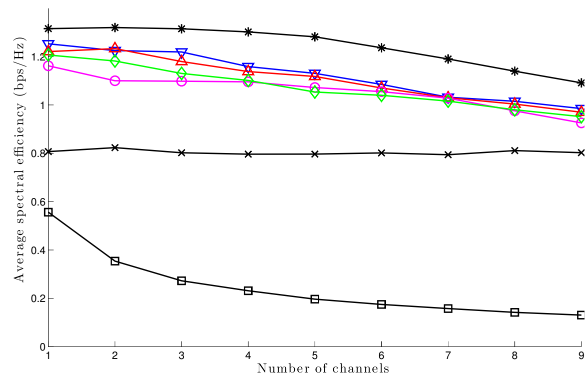

Fig. 1.2. Results for Test Case 2

6. NUMERICAL RESULTS

A CR with $M = 4$ SUs is simulated. The SISN follows a Rayleigh fading model with an average SNR of -5 dB. All users have the same priority, so that $\beta^m = 1\ \forall m$. The SIPN follows an AR-1 model with a coefficient of 0.95. The observed SIPN is corrupted by additive gaussian noise with a SNR of 3 dB. The power constraint at the SU transmitters is set to $[\check{p}^1, \ldots \check{p}^4] = [6.0, 7.2, 9.0, 12.0]$. The interference power constraint at the PU receivers is $\check{o}_k = 2.0\ \forall k$. The sensing cost parameter [cf. (4)] is $q^m = 5\ \forall m$. The Lagrange multipliers [cf. (7)] are computed using the method in [15].

Since we focus on the sensing policy, all tested schemes implement the optimal RA policy in Section 4.1. We are interested in comparing the performance of the myopic policy using the following schemes: i) an exhaustive search over $\mathbf{s}[n]$ (combinatorial complexity); ii) Algorithm 1 (proposed, polynomial complexity); iii) a round-robin scheme that sequentially selects a single different user at each n; iv) a scheme that randomly selects $s^m[n]$ mimicking the distribution of $s^m[n]$ at (ii); deterministic schemes that v) always sense, vi) never sense; and vii) an upper bound on the system performance (using the algorithm in (v) and setting $q^m = 0\ \forall m$).

Two test cases are run: TC1) the average power gain of the interference is fixed to -3 dB and the spectral efficiency is plotted vs. K in Fig. 1.1; TC2) $K = 4$ and the spectral efficiency is plotted vs. the average power gain of the interference in Fig 1.2. The average power and interference constraints are tightly satisfied in all cases.

Results show close performance of Algorithm 1 and exhaustive search for the simulated test cases. This suggests that Algorithm 1 can be a good option when M is large. Further, this motivates using the greedy approach to compute a suboptimal estimation of $\bar{V}(\cdot)$. Such schemes will be addressed in future work.

7. REFERENCES

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

[2] E. Dall’Anese, S.-J. Kim, G. Giannakis, and S. Pupolin, “Power control for cognitive radio networks under channel uncertainty,” *Wireless Communications, IEEE Transactions on*, , no. 99, pp. 1–11, 2011.

[3] Y.-C. Liang, Y. Zeng, E.C.Y. Peh, and A.T. Hoang, “Sensing-throughput tradeoff for cognitive radio networks,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, 2008.

[4] G. Xiong, S. Kishore, and A. Yener, “Cost constrained spectrum sensing in cognitive radio networks,” in *44th Conf. on Information Sciences and Systems (CISS)*, Princeton, NJ, Mar. 17–19, 2010.

[5] D. Xu and X. Liu, “Opportunistic spectrum access in cognitive radio networks: When to turn off the spectrum sensors,” in *4th Intl. Wireless Internet Conf. (WiCON)*, Maui, HW, Nov. 17–19, 2008.

[6] S.-J. Kim and G. Giannakis, “Sequential and cooperative sensing for multi-channel cognitive radios,” *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4239ñ4253, Aug. 2010.

[7] L. M. Lopez-Ramos, A. G. Marques, and J. Ramos, “Jointly optimal sensing and resource allocation for multiuser overlay cognitive radios,” *CoRR*, vol. abs/arXiv/1211.0954, 2012.

[8] Y. Chen, G. Yu, Z. Zhang, H.-H. Chen, and P. Qiu, “On cognitive radio networks with opportunistic power control strategies in fading channels,” *Wireless Communications, IEEE Transactions on*, vol. 7, no. 7, pp. 2752–2761, 2008.

[9] Q. Zhao, L. Tong, A. Swami, and Y. Chen, “Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework,” *Selected Areas in Communications, IEEE Journal on*, vol. 25, no. 3, pp. 589–600, 2007.

[10] 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.

[11] L. B. Le and E. Hossain, “Resource allocation for spectrum underlay in cognitive radio networks,” *Wireless Communications, IEEE Transactions on*, vol. 7, no. 12, pp. 5306–5315, 2008.

[12] 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,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 2, pp. 940–950, 2009.

[13] X. Gong, S. Vorobyov, and C. Tellambura, “Optimal bandwidth and power allocation for sum ergodic capacity under fading channels in cognitive radio networks,” *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1814ñ–1826, Apr. 2011.

[14] 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.

[15] A.G. Marques, L.M. Lopez-Ramos, G.B. Giannakis, and J. Ramos, “Resource allocation for interweave and underlay crs under probability-of-interference constraints,” *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 10, pp. 1922–1933, 2012.

[16] Y. Chen, Q. Zhao, and A. Swami, “Joint sensing-channel selection and power control for cognitive radios,” *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2053–ñ2071, May 2008.

[17] L.M. Lopez-Ramos, A.G. Marques, and J. Ramos, “Soft-decision sequential sensing for optimization of interweave cognitive radio networks,” in *Signal Processing Advances in Wireless Communications (SPAWC), 2013 IEEE 14th Workshop on*, 2013, pp. 235–239.

[18] W.B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, Wiley Series in Probability and Statistics. Wiley, 2011.

[19] J. V. Candy, *Bayesian Signal Processing: Classical, Modern and Particle Filtering Methods*, Wiley-Interscience, 2009.

[20] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

[21] D. Braziunas, “POMDP solution methods,” *University of Toronto*, 2003.