Dynamic Spectrum Sensing-Scheduling in Agile Networks with Compressed Belief Information

Nicolò Michelusi and Urbashi Mitra

Abstract—In this paper, we consider a network of primary users (PUs) opportunistically accessing a spectrum with $F$ frequency bins, and an agile network of secondary users (SUs) which attempt to access a licensed spectrum composed of $F$ frequency bins, and an agile network of secondary users (SUs) which opportunistically access the frequency bins left unused by the PUs [1]. Inference of the spectrum occupancy state is performed by a central controller (CC), based on distributed compressed spectrum measurements performed by the SUs and on local feedback information from the PUs. According to the resource constraints and the spectrum scheduling over time, based on the current spectrum occupancy belief, so as to maximize the SU throughput, under constraints on the PU throughput degradation and the sensing-transmission cost incurred by the SUs. The high dimensionality of the POMDP formulation is reduced by resorting to a compact state space representation via minimization of the Kullback-Leibler divergence. Simulation results demonstrate improvements up to 70% in the SU throughput, under constraints on the PU throughput degradation and the sensing-transmission cost incurred by the SUs. The high dimensionality of the POMDP formulation is reduced by resorting to a compact state space representation via minimization of the Kullback-Leibler divergence. Simulation results demonstrate improvements up to 70% in the SU throughput, under constraints on the PU throughput degradation and the sensing-transmission cost incurred by the SUs.

I. INTRODUCTION

In this paper, we consider a network of primary users (PUs) dynamically accessing a spectrum with $F$ frequency bins, and an agile network of secondary users (SUs) which attempt to access a licensed spectrum composed of $F$ frequency bins, and an agile network of secondary users (SUs) which opportunistically access the frequency bins left unused by the PUs [1]. Inference of the spectrum occupancy state is performed by a central controller (CC), based on distributed compressed spectrum measurements performed by the SUs and on local feedback information from the PUs. Accordingly, the CC schedules a subset of bins for opportunistic spectrum access by the SUs. In order to cope with the resource constraints and the network dynamics, the CC adapts the sensing probability of the SUs and the spectrum scheduling over time, based on the current spectrum occupancy belief, so as to maximize the SU throughput, under constraints on the throughput degradation incurred to the PUs and on the sensing cost for the SUs.

The contributions of this paper are as follows. We propose a framework which captures the interplay between sensing and scheduling, by trading off the cost of acquisition of network state information and the overall network performance. Spectrum sensing is done by collecting compressed spectrum measurements from distributed SUs and local measurements at the CC, based on which spectrum scheduling decisions are done. In order to tackle the high dimensionality of the POMDP formulation and of the dynamic programming (DP) algorithm [2], we build on [3], [4] and resort to Kullback-Leibler divergence (KLD) minimization to approximate the belief on a low-dimensional space. Based on the compressed belief, we design adaptive compressive sensing (CS) schemes, which effectively exploit the sparse networks dynamics typical of wireless networks, i.e., the fact that only few PUs join or leave the spectrum at any time, thus inducing sparse variations of the network state over time. Therefore, leveraging the estimate of the spectrum occupancy state in the previous slot, only a sparse residual uncertainty vector needs to be estimated, so that few measurements suffice to drive scheduling decisions.

Active sensor scheduling and adaptation [5] encompass applications such as target tracking [6], physical activity detection [7], and sequential hypothesis testing [8]. In [9], a centralized, non-adaptive CS framework [10] to track the support of a time-correlated Bernoulli-Gaussian signal is presented, and, in [11], sparse recovery algorithms for stream signals are designed. Distributed CS is studied in [12], [13], for a static signal model. Recovery of static binary sparse signals via CS has been investigated in [14]. Compressive spectrum sensing has been studied in [1], for a static setting, and in [15], for a dynamic setting with noiseless measurements, but without scheduling. All these prior works including ours [3], [16]–[19] assume that the underlying state is given by nature and is not controlled. In contrast, in this work, states are affected by scheduling decisions, and we build on our new results [20] to design joint distributed and adaptive sensing and scheduling schemes in wireless networks, which account for the cost of acquisition of state information and its impact on control. We extend the model in [4] to include local measurements and network feedback information at the CC, and recast the belief approximation within the framework of KLD minimization.

This paper is organized as follows. In Sec. II, we present the system model. In Sec. III, we propose a compact belief state based on KLD minimization. In Sec. IV, we present numerical results, and, in Sec. V, we conclude the paper.

II. SYSTEM MODEL

We consider a network of $N_S$ SUs with sensing capability, which attempt to access a licensed spectrum composed of $F$ frequency bins, represented in Fig. 1. The occupancy state of the $i$th bin in slot $k$ is denoted as $b_{k,i} \in \{0, 1\}$, where $b_{k,i} = 0$ if the bin is idle and $b_{k,i} = 1$ if it is occupied by a PU. Let $\mathbf{b}_k = (b_{k,1}, b_{k,2}, \ldots, b_{k,F})^T$ be the $F$-dimensional spectrum occupancy (column) vector at time $k$.

The system operates in two phases [20]: a sensing phase, of duration $\delta$ (we assume that one slot has duration 1), during which the SUs collect compressed distributed measurements of the spectrum occupancy state and report them to a CC (e.g.,
a base station) (Sec. II-B); followed by a scheduling phase, of duration $1 - \delta$, where the SUs access the spectrum based on the spectrum scheduling decision of the CC (Sec. II-A).

A. Spectrum Scheduling

The SUs opportunistically access the spectrum based on the spectrum scheduling decision $(u_k,q_{S,k})$ broadcasted by the CC, where $u_k = (u_{k,1}, u_{k,2}, \ldots, u_{k,F})^T \in \{0,1\}^F$, $u_{k,i} = 0$ and $u_{k,i} = 1$ denote, respectively, whether the $i$th bin cannot or can be used for SU spectrum access. Each SU, assumed to be backlogged, then transmits its own packet with probability $q_{S,k} \sum u_{k,i} / N_{S,1}$ independently in one of the scheduled spectrum bins in the set $I_k \equiv \{ i \in \{1,2,\ldots,F \} : u_{k,i} = 1 \}$, where $q_{S,k} \in [0,1]$ denotes the normalized transmission probability.\footnote{Note that the marginal transmission probability of each SU in a given scheduled channel is $q_{S,k}/N_s$.} We employ a collision channel model, i.e., if more than one terminals (either SUs or PUs) transmit on the same channel, those packets cannot be decoded correctly at the corresponding receiver and are lost. Otherwise, if one and only one user transmits, then the transmission is successful with probability $1 - \rho_S$ (for the SU) and $1 - \rho_P$ (for the PU), where $\rho_S$ and $\rho_P$ are the corresponding transmission failure probabilities. Herein, we use the approximation $N_s \to \infty$ to derive the transition probabilities and performance of the system. The following discussion can be generalized to $N_s \to \infty$.

If $b_{k,i} = 1$, the success probability for the PU as a function of $u_{k,i}$ and of $q_{S,k}$, denoted as $P_{\text{suc}}^{(PU)}(u_{k,i},q_{S,k})$, is given by

$$P_{\text{suc}}^{(PU)}(u_{k,i},q_{S,k}) = (1 - \rho_P)\left(u_{k,i}e^{-q_{S,k}} + 1 - u_{k,i}\right), \quad (1)$$

where we have used the fact that the probability of no collisions from the SUs satisfies $(1 - q_{S,k}/N_s)^{N_S} \to e^{-q_{S,k}}$ for $N_S \to \infty$. Similarly, if $u_{k,i} = 1$, the probability of successful transmission in the $i$th bin for the SU bin, as a function of $b_{k,i}$ and $q_{S,k}$, denoted as $P_{\text{suc}}^{(SU)}(b_{k,i},q_{S,k})$, is given by

$$P_{\text{suc}}^{(SU)}(b_{k,i},q_{S,k}) = (1 - b_{k,i})(1 - \rho_S)q_{S,k}e^{-q_{S,k}}, \quad (2)$$

where we have used the fact that the probability that one and only one SU transmits satisfies $q_{S,k}(1 - q_{S,k}/N_s)^{N_S-1} \to q_{S,k}e^{-q_{S,k}}$, and, if the channel is busy, the transmission fails.

The PUs implement a retransmission mechanism in case of transmission failure. Retransmissions are performed in the same bin in the next slot. If the transmission is successful, then the PU occupying the spectrum bin has a new data packet to transmit in the next slot in the same spectrum bin with probability $\theta$, and leaves the spectrum otherwise. An idle spectrum bin is occupied by a new PU with probability $\zeta \in (0,1)$ and it remains idle otherwise. Therefore, the state of each spectrum bin $b_{k,i} \in \{0,1\}$ is a two-state controlled Markov chain. The transition probability from state $b_{k,i} = 0$ to $b_{k,i+1} = b'$, given $u_{k,i} = 1$ and $q_{S,k} = u$, denoted as $P_B(b'|b,u,q) = P(b_{k,i} = b_{k,i} = b, u_{k,i} = u, q_{S,k} = q)$, is given by

$$P_B(1|b,u,q) = (1 - \delta)\zeta + \delta \left[1 - (1 - \theta)(1 - \zeta)P_{\text{suc}}^{(PU)}(u,q)\right]$$

and $P_B(0|b,u,q) = 1 - P_B(1|b,u,q)$. In fact, the $i$th bin is occupied in the next slot if and only if the transmission of the SU is successful and it has a new data packet to transmit, with probability $\theta$, a new PU arrives, with probability $\zeta$, or a retransmission needs to be done.

The CC, at the end of the slot, may overhear the PU acknowledgments of correct (ACK) or incorrect (NACK) reception of the packets, fed back by the PU receivers on each channel, denoted as $y_{CC,k,i} \in \{\text{ACK},\text{NACK},\emptyset\}$, where $y_{CC,k,i} = \emptyset$ if either an erasure occurs (the ACK/NACK message cannot be detected by the CC) or the $i$th bin was idle. We denote the erasure probability as $\epsilon \in (0,1)$. Therefore, the joint probability of $y_{CC,k} = y$ and $b_{k+1} = b'$, given $b_{k,i} = b$, $u_{k,i} = 1$ and $q_{S,k} = q$, denoted as $P_B(y_{CC,k} = y, b', u, q)$, is given by

$$P_B(y_{CC,k} = y, b', u, q) = (1 - \epsilon)b'\left(1 - P_{\text{suc}}^{(SU)}(u,q)\right)$$

$$= (1 - \epsilon)b'(\theta + (1 - \theta)\zeta) + (1 - \epsilon)\left(\theta + (1 - \theta)\zeta\right)P_{\text{suc}}^{(SU)}(u,q),$$

where $y_{CC,k} = y$ and $y_{CC,k} = \emptyset$ in the first term, and $y_{CC,k} = \emptyset$ in the second term.

From which we obtain the conditional probability

$$P_B(b'|b,u,q,y) = P(b_{k+i} = b'|b_{k,i},u_{k,i},q_{S,k},y_{CC,k+i}) \approx (u,b,y,q) = \frac{P_B(y_{CC,k} = y, b', u, q)}{P_B(y_{CC,k} = y, b', u, q) + P_B(y_{CC,k} = \emptyset, b', u, q)}, \quad (3)$$

We define the aggregate expected throughput for the SU and PU systems, respectively, given $(b_k, u_k, q_{S,k})$, as

$$V_S(b_k, u_k, q_{S,k}) = \sum_{i=1}^{F} u_{k,i}P_{\text{suc}}^{(SU)}(b_{k,i}, q_{S,k}), \quad (4)$$

$$V_P(b_k, u_k, q_{S,k}) = \sum_{i=1}^{F} b_{k,i}P_{\text{suc}}^{(SU)}(u_{k,i}, q_{S,k}). \quad (5)$$

B. Spectrum Sensing

At the beginning of slot $k$, $b_k$ is inferred by collecting noisy compressed spectrum measurements by the SUs,

$$y_{n,k} = a_{n,k}^T b_k + z_{n,k}, \quad \forall n = 1,2,\ldots,N_S, \quad (6)$$

where $z_{n,k} \sim \mathcal{N}(0, \sigma_z^2)$ is Gaussian noise, i.i.d. over time and across SUs, $a_{n,k}$ is the measurement vector, and the superscript "T" denotes the matrix or vector transpose. Eq. (6) is the result of filtering over the spectrum band, so that $a_{n,k}$
denotes the filtering coefficient vector, which includes also the signal attenuation between the PU and the SU. We assume that \( \alpha_{n,k} \sim \mathcal{N}(0, \sigma_k^2 I_r) \), where \( I_r \) is the \( n \times n \) identity matrix, and is known to the CC. The SUs share \( B \) orthogonal control channels to report their measurements, resulting in packet losses if more than one SU transmit on the same channel. The SUs operate in a decentralized fashion, i.e., they decide to sense and transmit their measurement to the CC, randomly in one of the \( B \) available channels, with common probability \( \alpha_{n,k} = \kappa_k B / N_S \in [0, 1] \), where \( \kappa_k \) denotes the normalized transmission probability per channel [18], incurring the sensing-transmission cost \( c_S \), and they remain idle otherwise, incurring no cost. The value of the sensing action \( \kappa_k \) is broadcast by the CC to the SUs.

We denote the set of SUs that report successfully their measurement to the CC as \( R_k \), and their cardinality as \( N_k \). When \( N_k \to \infty \), it was shown in [18] that \( R_k \) has binomial distribution with \( B \) trials and success probability \( \kappa_k e^{-\kappa_k} \). Let \( y_k \) be the \( R_k \)-dimensional vector of measurements collected at the CC in slot \( k \).

\[
y_k = A_k^T b_k + z_k, \tag{7}
\]

where \( A_k = [a_{n,k}]_{n \in R_k} \) is the measurement matrix, known to the CC, and \( z_k = [z_{n,k}]_{n \in R_k} \) is the noise column vector.

C. Policy definition

We denote the prior belief that the spectrum occupancy state takes value \( b_k = b \), based on the history collected up to time \( k \) and before the sensing phase, as \( \pi_k(b) = \mathbb{P}(b_k = b | \text{history up to time } k) \). In the sensing phase, \( \kappa_k \) is chosen by the CC according to the stationary sensing policy \( \kappa_k = \kappa(\pi_k) \) [20]. Similarly, we denote the posterior belief that \( b_k = b \), given \( \pi_k \), the measurement vector \( y_k \) and measurement matrix \( A_k \), as \( \hat{\pi}_k(b) = \mathbb{P}(b_k = b | y_k, A_k) \). Using (7),

\[
\hat{\pi}_k(b) \propto \pi_k(b) \exp\left\{-\frac{1}{2 \sigma_k^2} \left\| y_k - A_k^T b \right\|^2 \right\}, \tag{8}
\]

where \( \propto \) denotes proportionality up to a normalization factor, so that we can write \( \hat{\pi}_k = \Pi(\pi_k, A_k, y_k) \), for a proper function \( \Pi(\cdot) \). In the scheduling phase, \( u_k \) is chosen according to the stationary scheduling policy \( u_k = u(\hat{\pi}) \) [20]. Furthermore, we assume that only the first \( F \) bins most likely to be idle are scheduled for spectrum access, where \( F \in \{0, 1, \ldots, F\} \). Specifically, letting \( \beta(\pi) = \mathbb{E}[b_k | \pi] \) be the expected spectrum occupancy with respect to the belief \( \pi \), and \( m_k(i) \) be a permutation of the spectrum in non-decreasing order of occupancy, \( i.e., \beta(\pi)_{m_k(1)} \leq \beta(\pi)_{m_k(2)} \leq \ldots \leq \beta(\pi)_{m_k(F)} \), \( u_k \) is given by

\[
u_{u_k, m_k(i)} = \chi(i \leq \phi_{k}), \tag{9}
\]

where \( \chi(\cdot) \) is the indicator function. The pair \( (\phi_k, q_k) \) is chosen according to the scheduling policy \( (\phi_k, q_k, s_k) = \mu(\hat{\pi}_k) \), and the scheduling vector \( u_k \) is determined via (9).

Given the scheduling decision \( u_k \) and \( q_k, s_k \), and the ACK/NACK feedback message \( y_{CC,k} = y \), the CC updates the next prior belief as \( \pi_{k+1} = \Pi(\hat{\pi}_k, u_k, q_k, y_{CC,k}) \), where

\[
\pi_{k+1}(b) = \sum_b \hat{\pi}_k(b) \sum_{i=1}^{F} P_B(b_i | b_k, u_i, q_i, y_{CC,i}). \tag{10}
\]

We denote the joint sensing-scheduling policy as \((\kappa, \mu)\), and we define the average discounted network cost as

\[
\bar{C}_{ST}(\kappa, \mu; \pi_0) \triangleq \mathbb{E} \left[ \sum_{k=0}^{\infty} \gamma^k \kappa_k B c_S + q_S k | \phi_{CTX} \big| \pi_0 \right], \tag{11}
\]

where \( c_S \) is the unitary sensing-transmission cost in the sensing phase, and \( c_{CTX} \) is the unitary transmission cost in the scheduling phase; and the discounted average SU/PU throughputs as

\[
\bar{V}_X(\kappa, \mu; \pi_0) \triangleq \mathbb{E} \left[ \sum_{k=0}^{\infty} \sum_{b_k} \sum_{b_k} \hat{\pi}_k(b_k) V_X(b_k, u_k, q_k) \big| \pi_0 \right], \tag{12}
\]

for \( X \in \{S, P\} \), where the expectation is with respect to the realization of \( \{b_k, A_k, y_k, u_k, q_k, y_{CC,k}\} \), induced by \((\kappa, \mu)\). The goal is to determine \((\kappa^*, \mu^*)\) such that

\[
(\kappa^*, \mu^*) = \arg \max_{(\kappa, \mu)} \mathbb{E} \left[ \xi \bar{V}_S(\kappa, \mu; \pi_0) \right] + (1 - \xi) \bar{V}_P(\kappa, \mu; \pi_0) - \lambda \bar{C}_{ST}(\kappa, \mu; \pi_0), \tag{13}
\]

where the parameters \( \lambda > 0 \) and \( \xi \in [0, 1] \) capture the desired trade-off between achieving high PU/SU throughput and incurring low cost of acquisition of state information at the CC. Note that, by letting \( \epsilon = 1 \) (no ACK/NACK feedback observations at the CC), \( \theta = 0, c_{TX} = 0, q_S k = 1 \), we obtain the model investigated in [4] as a special case.

The policy \((\kappa^*, \mu^*)\) can be determined via DP. However, the DP algorithm and computation of the prior/posterior beliefs via (8) and (10) have huge complexity, since the beliefs are defined over a \( 2^F \) dimensional state space. In the next section, we propose a compact belief representation based on the minimization of the KLD, which enables a more efficient optimization and operation of the system, and a scalable state estimator based on sparse recovery algorithms.

III. COMPACT BELIEF STATE VIA KLD MINIMIZATION

In this section, we propose a compact belief representation by resorting to the KLD. In particular, we approximate the belief \( \pi(b) \) with the following factorized model:

\[
\pi(b) \approx \hat{\pi}(b) = \prod_{i} \frac{b_i^{d_{f(i)}^i} (1 - b_i^{d_{f(i)}^i})^{1 - b_i^{d_{f(i)}^i}}}{b_i^{d_{f(i)}^i}}, \tag{14}
\]

where \( b_1, b_2 \in \{0, 1\} \) with \( b_1 \leq b_2 \) are probability levels, and \( f : \{1, 2, \ldots, F\} \to \{1, 2\} \) is a function which maps the \( i \)th spectrum bin to indices corresponding to one of the levels \( b_1 \) or \( b_2 \). Note that this approximation assumes that the spectrum bins are statistically independent of each other, and that their probability of being occupied takes two possible values, \( b_1 \) or \( b_2 \). The approximate belief \( \hat{\pi}(b) \), termed compressed belief state (CBS), is parameterized by \( (b_1, b_2, f(\cdot)) \). We thus write \( \hat{\pi} = g(b_1, b_2, f) \). The KLD between \( \pi \) and \( \hat{\pi} \) is thus given by

\[
D(\pi, \hat{\pi}) = \mathbb{E} \left[ \log \frac{\pi(b)}{\hat{\pi}(b)} \right] = \sum_{b 
\in \{0, 1\}^F} \pi(b) \log \left( \frac{\pi(b)}{\hat{\pi}(b)} \right). \tag{15}
\]

The goal is to find parameters \( b_1^*, b_2^*, f^* \) such that

\[
\arg \min_{b_1, b_2, f} D(\pi, b_1, b_2, f) = \arg \max_{b_1, b_2, f} \left[\beta(\pi) \log (b_{f(i)}^f) + (1 - \beta(\pi)) \log (1 - b_{f(i)}^f)\right], \tag{16}
\]
where we have used (14) and $\beta(\pi) = \mathbb{E}[b_k | \pi]$. In the following theorem, we determine the solution of (16).

**Theorem 1** Given $\beta = \mathbb{E}[b | \pi]$, the parameters of the approximate belief, solution of (16), are given by

$$b_1^*(\pi) = \frac{1}{n^*(\pi)} \sum_{i=1}^{n^*(\pi)} \beta_{m(i)}, \quad b_2^*(\pi) = \frac{1}{F - n^*(\pi)} \sum_{i=n^*(\pi) + 1}^{F} \beta_{m(i)};$$

where $m : \{1, 2, \ldots, F\} \rightarrow \{1, 2, \ldots, F\}$ is a permutation of the entries of $\beta$ in increasing order; i.e., such that $\beta_{m(1)} \leq \beta_{m(2)} \leq \cdots \leq \beta_{m(F)}$, and $n^*(\pi)$ solves

$$n^*(\pi) = \arg \min_{n \in \{1, 2, \ldots, F-1\}} nH_2(b_1(n)) + (F - n)H_2(b_2(n)).$$

The function $f^*$ is given by $f^*(\pi; m(i)) = 1$ for $i \leq n^*(\pi)$ and $f^*(\pi; m(i)) = 2$ otherwise.

$n^*(\pi)$ (17) can be determined via exhaustive search, with linear complexity in $F$. The system is then operated as follows:

1) **Sensing phase**: given $\pi_k$, the parameters $b_{1,k}, b_{2,k}, n_k$ of the prior CBS are determined via Theorem 1. The sensing action is then chosen as $\kappa_k = \kappa(b_{1,k}, b_{2,k}, n_k)$ (rather than as a function of $\pi_k$). Given $(y_k, A_k)$ and $\hat{n}_k = g(b_{1,k}, b_{2,k}, n_k)$, the posterior belief $\hat{\pi}_k$ is updated as in (8).

2) **Scheduling phase**: given $\hat{\pi}_k$, the parameters $b_{1,k}, b_{2,k}, \hat{n}_k$ of the posterior CBS are determined via Theorem 1, and the scheduling action is selected as $(\phi_k, q_{s,k}) = \mu(b_{1,k}, b_{2,k}, \hat{n}_k)$ (rather than as a function of $\hat{\pi}_k$). Given the CBS $\hat{\pi}_k = (y_k, A_k, \hat{n}_k)$ and the local ACK/NACK feedback $y_{CC,k}$, the prior belief for the next slot is then given by $\pi_{k+1}(b) = \prod_{i} \beta_{k+1,i}^{b_i}(1 - \beta_{k+1,i})^{1-b_i}$, where

$$\beta_{k+1,i} = \sum_{b \in \{0,1\}} \beta_{k,i}(1 - \hat{\beta}_{k,i})^{1-b_i} \mathbb{P}_B(1|b, u_{k,i}, q_{S,k}, y_{CC,k,i})$$

and $\hat{\beta}_k = \mathbb{E}[b_k | \hat{\pi}_k]$. The prior belief is thus updated with linear complexity in $F$, rather than exponential as in (10).

3) These steps are repeated in each slot, by projecting the current belief into the CBS, based on which sensing-scheduling decisions are made.

Note that determining the posterior belief $\hat{\pi}_k$ in the sensing phase via (8) has exponential complexity. In order to achieve complexity reduction, we can decouple the estimator from the controller as in [4], i.e., the estimator is treated as a black-box which outputs a MAP estimate of $\mathbb{B}_k$ via sparse recovery algorithms, and posterior false-alarm and mis-detection probabilities for the bins detected as busy and idle, respectively.

The sensing-scheduling policy $(\kappa, \mu)$ can thus be determined efficiently via DP based on the CBS, by exploiting the low-dimensional CBS parameterization $(b_1, b_2, \phi)$, using a similar approach as in [4].

**IV. NUMERICAL RESULTS**

In this section, we present numerical results. We consider a scenario with the following system parameters: $F=10$, $N_S=200$, $B=20$, $\rho_S=\rho_P=0.1$, $\zeta=0.3$, $\sigma_1^2=1$, $\sigma_2^2=1/20$.

![Figure 2. Trade-off between PU throughput and SU throughput.](image)

$\gamma=1, \epsilon=0.9, \theta=0, c_S=1, c_{TX}=0, q_{S,k}=1$. We consider the following policies, evaluated via simulation over $10^3$ slots:

1) **Local sensing (LS)** policy, in which the CC relies only on the ACK/NACK feedback information from the PUs.

2) **Adaptive sensing (AS)** policy, computed after 50 iterations of the DP algorithm (see [4]). The parameters $\xi$ and $\lambda$ in (13) are varied in order to obtain different operational points, in terms of SU-PU throughputs and sensing cost.

3) **Full sensing (FS)** policy, in which sensing occurs with a fixed probability $\alpha_k = B/N_S$ ($\kappa_k = 1$), which maximizes the throughput of the feedback channel.

4) **Full network state information (FNSI)** policy, in which the state $b_k$ is perfectly known to the CC when $u_k$ is scheduled.

Fig. 2 plots the trade-off between the PU and SU throughputs. Despite achieving the best trade-off, FS incurs a significant sensing cost for the SUs (10-fold compared to AS). In contrast, a poor performance is attained by LS, whose sensing cost is 0. On the other hand, AS optimally balances the trade-off between the SU-PU throughputs and the cost of acquisition of state information from the SUs, with gains up to 70% in the SU throughput over LS, while achieving significant cost-savings compared to FS.

**V. CONCLUSIONS**

In this paper, we have presented a cross-layer framework for joint distributed spectrum sensing, estimation and scheduling in agile wireless networks, in which a network of SUs opportunistically accesses portions of the spectrum left unused by a licensed network of PUs. Inference of the underlying spectrum occupancy state is obtained by collecting compressed measurements at the CC from nearby SUs, and via local ACK/NACK feedback information from the PUs. In order to reduce the high operation and operational complexity due to the POMDP formulation, we have proposed a technique to project the belief state into a compressed representation via the minimization of the Kullback-Leibler divergence. The compressed belief state makes it possible to efficiently optimize the performance via DP. Simulation results demonstrate a significant performance gain on the PU/SU throughput trade-off over a local sensing scheme, at a fraction of the cost in the acquisition of state information with respect to a scheme where sensing is done in each slot by the SUs.

$^3$Its proof is not provided due to space constraints.
REFERENCES


