Joint Control and Compressed Sensing for Dynamic Spectrum Access in Agile Wireless Networks

Nicolò Michelusi and Urbashi Mitra

Abstract—In this paper, a cross-layer framework for joint distributed sensing, estimation and control in agile wireless networks is presented. A network of secondary users (SUs) opportunistically accesses portions of the spectrum left unused by a licensed network of primary users (PUs). A central controller (CC) schedules the spectrum bands detected as idle for access by the SUs, based on compressed measurements acquired by the SUs. The sparsity in the spectrum occupancy dynamics is exploited: leveraging the spectrum occupancy estimate in the previous slot, the CC needs to estimate only a sparse residual uncertainty vector via sparse recovery techniques, so that only few measurements suffice. The sensing probability of the SUs and the spectrum scheduling are adapted over time by the CC, based on the current spectrum occupancy estimate, and jointly optimized so as to maximize the SU throughput, under constraints on the PU throughput degradation and the sensing cost incurred by the SUs. A compact state space representation and decoupling of the state estimator from the CC are proposed: the estimator provides a maximum-a-posteriori spectrum estimate, as well as false-alarm and mis-detection error probabilities for the bins detected as busy and idle, respectively, based on which the CC performs scheduling and sensing decisions. Simulation results demonstrate improvements up to 11% in the SU throughput over a static sensing scheme.

I. INTRODUCTION

The proliferation of mobile devices of late has been exponential in numbers as well as heterogeneity [1]. Thus, tools for the design and optimization of agile wireless networks, comprised of heterogeneous classes of wireless users with the ability to adapt their operation based on the state of the network, are needed. However, network design needs to explicitly consider the resource constraints typical of wireless systems, and their impact in the availability of network state information, which enables network control.

In this paper, we consider a wireless network composed of a licensed network of primary users (PUs) dynamically accessing a spectrum with $F$ frequency bins, and an agile network of secondary users (SUs) which opportunistically access the frequency bins left unused by the PUs [2]. Inference of the spectrum occupancy state is performed by a central controller (CC), which collects compressed noisy measurements from the SUs via distributed sensing and, accordingly, schedules a subset of bins for opportunistic spectrum access by the SUs. In order to cope with the resource constraints and the system dynamics, the CC adapts the sensing probability of the SUs as well as the spectrum scheduling over time, based on the current spectrum occupancy belief. The adaptive sensing-scheduling policy is jointly optimized so as to maximize the SU throughput, under a constraint on the throughput degradation incurred to the PU network, and sensing cost constraints for the SUs. Simulation results demonstrate the superiority of the proposed adaptive joint sensing-scheduling policy over a static one.

The contributions of this paper are as follows. We propose a framework which captures the interplay between sensing and scheduling, and makes it possible to trade off the cost of acquisition of network state information and the overall network performance. Distributed sensing provides measurements to infer the network state, whereas control exploits the available state information to schedule actions. This is in contrast to standard formulations based on partially observable Markov decision processes (POMDPs) [3], where observations are passively generated by control actions, rather than actively controlled via sensing. In order to tackle the high dimensionality of the POMDP formulation and of the dynamic programming (DP) algorithm [4], we build on [5] to propose a compact representation of the belief state, and the decoupling of the estimator from the CC: the estimator is treated as a black-box, which provides only a maximum-a-posteriori (MAP) estimate of the underlying state of the system and false-alarm and mis-detection error probabilities for the bins detected as busy and idle, respectively, based on which the CC performs spectrum scheduling and sensing decisions.

Moreover, we design an adaptive compressive sensing (CS) framework, which effectively exploits the sparse networks dynamics typical of wireless networks. In the spectrum sensing context analyzed in this paper, sparsity is induced by the fact that only few PUs join or leave the spectrum at any time, so that the spectrum occupancy state exhibits sparse time variations. 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 only few compressed spectrum measurements suffice to accurately estimate the spectrum occupancy in the new slot. Additionally, this sparsity and the compact belief representation are exploited to design a state estimator based on sparse recovery algorithms.

Although the focus of this paper is on spectrum sensing in agile wireless networks, this framework can be generalized to more general networked systems, where the state of the system is a collection of features, rather than spectrum bins (e.g., buffer state of all wireless nodes, or local channel quality), which evolve sparsely over time. These state features can be tracked by collecting few compressed measurements via distributed sensing, enabling more informed network control.

Active sensor scheduling and adaptation [6] encompass applications such as target tracking [7], physical activity...
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 $b_k = (b_{k,1}, b_{k,2}, \ldots, b_{k,F})^T$ be the $F$-dimensional spectrum occupancy (column) vector at time $k$.

The SUs opportunistically access the spectrum based on the spectrum scheduling output of a CC\(^1\) (e.g., a base station), which selects the binary action $u_{k,i} \in \{0,1\}$ for each bin $i$, where $u_{k,i}=0$ and $u_{k,i}=1$ denote, respectively, whether the $i$th bin cannot or can be used for secondary spectrum access. Let $u_k=(u_{k,1}, u_{k,2}, \ldots, u_{k,F})^T \in \{0,1\}^F$ be the $F$-dimensional scheduling decision of the CC. Then, the CC broadcasts, at the beginning of slot $k$, the set of spectrum bins scheduled for SU access, denoted as $\mathcal{I}_k \equiv \{i \in \{1,2,\ldots,F\} : u_{k,i} = 1\}$. Each SU then transmits its own packets with probability $q_S$, independently in one of the available spectrum bins $\mathcal{I}_k$. The SUs are backlogged. 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 $1-P_S$ (for the SU) and $1-P_P$ (for the PU), where $P_S$ and $P_P$ are the corresponding transmission failure probabilities. Herein, we choose $q_S=1/N_S$, which maximizes the throughput for the SUs, under the collision model employed [22]. Moreover, we use the approximation $N_S \to \infty$. The following discussion can be generalized to $N_S < \infty$.

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

$$P_{\text{succ}}(u_{k,i}) = (u_{k,i}e^{-1} + 1 - u_{k,i})(1 - \rho_P),$$

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

$$P_{\text{succ}}(b_{k,i}) = (1 - b_{k,i})(1 - \rho_S)e^{-1},$$

(2) where we have used the fact that the probability that one and only one SU transmits is $N_S q_S (1 - q_S)^{N_S-1} \to e^{-1}$, 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. On the other hand, if a given spectrum bin is idle, then it is occupied by a new PU with probability $\zeta \in (0,1)$ and it remains idle otherwise. The evolution of $b_k$ is affected by the spectrum scheduling decision $u_k$, and by PUs occupying and leaving the spectrum. We denote the transition probability from state $b_{k,i}=b$ to state $b_{k+1,i}=1$, under the scheduling decision $u_{k,i}\in\{0,1\}$, as $P_B(b,u) = \mathbb{P}(b_{k+1,i}=1|b_{k,i}=b, u_{k,i}=u)$. This is given by

$$P_B(b,u) = (1-b)\zeta + b \left[1 - (1-\zeta)P_{\text{succ}}(u)\right].$$

(3) In fact, the $i$th bin is occupied in the next slot if and only if a new PU arrives, with probability $\zeta$, or a retransmission occurs. Therefore, each component of $b_k$ evolves as an independent two state controlled Markov chain, with transition probability to state $b_{k+1,i}=1$ given by $P_B(b_{k,i},u_{k,i})$, depending on the scheduling decision of the CC.

The spectrum occupancy state $b_k$ is inferred by collecting noisy measurements from the $N_S$ SUs\(^2\). In particular, in slot $k$, the $n$th SU collects the compressed measurement

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

(4) where $z_{n,k} \sim \mathcal{N}(0,\sigma^2_{n,k})$ is Gaussian noise, i.i.d. over time and across SUs, $a_{n,k}^T$ is the measurement vector, and the superscript “$T$” denotes the matrix or vector transpose. This

\(^1\)We assume that the CC coordinates only the SU network, but does not have control over the PU network, hence $b_k$ is unknown.

\(^2\)We assume that the measurements are collected by the SUs. However, the analysis can be extended to the case where the sensors collecting the measurements and the SUs performing spectrum access do not coincide.
observation model is the result of filtering operations performed at each SU across the spectrum bins, 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 the entries of $a_{n,k}$ take values from a Gaussian distribution with zero mean and variance $\sigma_A^2$, and are known to the CC.

**Remark 1** Note that each SU can, in principle, estimate the spectrum occupancy state $b_k$ based on its own measurement $y_{nk}$. However, if $F$ is large, or the measurement is very noisy ($\sigma_Z^2/\sigma_A^2 \gg 1$), the estimation accuracy may be very poor. On the other hand, by collecting measurements from a large number of SUs, the CC can estimate $b_k$ accurately.

Let $v_S(b, u)$ and $v_P(b, u)$ be the expected number of successful transmissions per bin for the SU and PU systems, respectively, when $u_{k,i} = u$ is scheduled by the CC and the state is $b_{k,i} = b$. From (1) and (2), we have

\[ v_S(b, u) = uP_{\text{succ}}^{(\text{SU})}(b), \quad v_P(b, u) = bP_{\text{succ}}^{(\text{PU})}(u). \]  

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

\[ V_X(b_k, u_k) = \sum_{i=1}^{F} v_S(b_{k,i}, u_{k,i}), \]  

where $X \in \{S, P\}$ denotes the SU or PU label, respectively.

### III. Joint sensing-estimation-control

If the spectrum occupancy state $b_k$ is perfectly known at the CC, it can be shown that the optimal scheduling decision is $u_k = 1 - b_k$, i.e., a spectrum bin is scheduled for secondary access if and only if it is idle. In fact, SU transmissions in busy channels create interference to the PU system, and are harmful if and only if they are available. In fact, SU transmissions in busy channels create interference to the PU system, and are harmful if and only if it is idle. In fact, SU transmissions in busy channels create interference to the PU system, and are harmful if and only if it is idle. In fact, SU transmissions in busy channels create interference to the PU system, and are harmful if and only if it is idle. In fact, SU transmissions in busy channels create interference to the PU system, and are harmful if and only if it is idle. In fact, SU transmissions in busy channels create interference to the PU system, and are harmful if and only if it is idle. In fact, SU transmissions in busy channels create interference to the PU system, and are harmful if and only if it is idle. In fact, SU transmissions in busy channels create interference to the PU system, and are harmful if and only if it is idle.

The system operates in two phases [21]:

- **Sensing phase**: at the beginning of slot $k$, the CC collects noisy measurements from the $N_S$ SUs. The SUs share $B$ orthogonal feedback 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 $\alpha_k \in [0,1]$, incurring the sensing transmission cost $c_{ST} = 1$, and they remain idle otherwise, incurring no cost. The value of $\alpha_k$ is determined by the sensing action, which is broadcast to the CC by the SU.

- **Scheduling phase**: based on the measurements collected in the sensing phase, the CC schedules $u_k$, and the SUs then perform spectrum access as described in Sec. II.

In the sensing phase, we denote the number of SUs that successfully report their measurement to the CC as $R_k$, with distribution $P_R(|\alpha_k)$, and the corresponding set of SUs as $\mathcal{R}_k$. In [19], it has been shown that, when $N_S \rightarrow \infty$ with $\alpha_k = \kappa_k \frac{B}{N_S}$, where $\kappa_k$ denotes the normalized transmission probability per channel, $R_k$ has binomial distribution with $B$ trials and success probability $\kappa_k e^{-\kappa_k}$, i.e., $R_k \sim B_B(\kappa_k e^{-\kappa_k})$. Let $y_k$ be the $R_k \times 1$ measurement vector of measurements collected at the CC in slot $k$. From (4), this is given by

\[ y_k = A_k b_k + z_k, \]  

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

We denote the prior belief that the spectrum occupancy state takes value $b_k = b$, based on the history collected up to time $k$, as $\pi_k(b) = P(b_k = b | \text{history up to time } k)$. Similarly, we denote the posterior belief that $b_k = b$, based on the history collected up to time $k$, the measurement vector $y_k$ and measurement matrix $A_k$, as $\hat{\pi}_k(b) = P(b_k = b | y_k, A_k)$. Using (7), we have

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

where $\propto$ denotes proportionality up to a normalization factor. We thus define the function

\[ \hat{\pi}_k = \hat{\Pi}(\pi_k, A_k, y_k), \]  

which maps the prior belief in slot $k$ to the posterior belief, given $A_k, y_k$. Similarly, given the scheduling action $u_k = u$, we obtain the next prior belief as

\[ \pi_{k+1}(b) = \sum_b \hat{\pi}_k(b) \prod_{i=1}^{F} P_B(\hat{b}_i, u_i)^{b_i}(1 - P_B(\hat{b}_i, u_i))^{1 - b_i}. \]  

We thus define the function

\[ \pi_{k+1} = \hat{\Pi}(\hat{\pi}_k, u), \]  

which maps the posterior belief $\hat{\pi}_k$ to the next prior belief, given $u_k = u$. We have the following lemma, proved in [21].

**Lemma 1** The prior belief $\pi_k$ is a sufficient statistic to choose the sensing action $\alpha_k$ in slot $k$. The posterior belief $\hat{\pi}_k$ is a sufficient statistic to schedule $u_k$ in slot $k$.

Then, we define the stationary and deterministic sensing policy $\alpha_k = \alpha(\hat{\pi}_k)$ and scheduling policy $u_k = \mu(\hat{\pi}_k)$. Herein, we restrict $\mu$ to threshold policies, i.e., letting $\beta(\pi) = \mathbb{E}[b_k | \pi]$ be the expected spectrum occupancy with respect to the belief $\pi$,

\[ \mu(\hat{\pi}_k) = \chi(\beta(\hat{\pi}_k) \leq \beta_{th}(\hat{\pi}_k)), \]  

where $\chi(\cdot)$ is the entry-wise indicator function and $\beta_{th}(\hat{\pi}_k) \in [0,1]$ is the scheduling threshold, so that only the channels which are more likely to be idle are scheduled for SU access. We denote the joint sensing-scheduling policy as $(\alpha, \beta_{th})$. Given $(\alpha, \beta_{th})$, we have the following steps in each slot:

1) **Sensing action broadcast**: the CC broadcasts $\alpha_k = \alpha(\hat{\pi}_k)$.

2) **Sensing-transmission phase**: $R_k \sim P_R(\alpha_k)$ measurements are collected at the CC, with measurement vector $y_k$ and matrix $A_k$; the expected sensing cost $\tilde{\alpha}_k$ is incurred.

3) **Posterior belief update**: as $\hat{\pi}_k = \hat{\Pi}(\pi_k, A_k, y_k)$.
4) **Scheduling phase:** $u_k = \chi(\beta(\hat{\pi}_k) \leq \beta_{th}(\hat{\pi}_k))$; the expected SU/PU throughputs $V(X(b_k, u_k), X \in \{S, P\})$ are collected.

5) **Prior belief update** as $\pi_{k+1} = \Pi(\hat{\pi}_k, u_k)$.

We define the average discounted sensing-transmission cost per SU, under the joint sensing-scheduling policy $(\alpha, \beta_{th})$, as

$$C_S(\alpha, \beta_{th}; \pi_0) \triangleq (1 - \gamma)E \left[ \sum_{k=0}^{\infty} \gamma^k \alpha(\pi_k) \right] \pi_0,$$  
(13)

and the discounted average SU/PU throughputs as

$$\bar{V}_X(\alpha, \beta_{th}; \pi_0) \triangleq (1 - \gamma)E \left[ \sum_{k=0}^{\infty} \gamma^k \sum_{b_k} \bar{v}_k(b_k) V_X(b_k, u_k) \right] \pi_0$$  
(14)

for $X \in \{S, P\}$, where the expectation is with respect to the realization of $\{b_k, A_k, y_k, u_k\}$, induced by $(\alpha, \beta_{th})$. The objective is to determine $(\alpha^*, \beta_{th}^*)$ so as to optimize a trade-off between the SU/PU throughputs $V_X(\alpha, \beta_{th}; \pi_0), X \in \{S, P\}$, and the sensing cost $C_S(\alpha, \beta_{th}; \pi_0)$, i.e.,

$$(\alpha^*, \beta_{th}^*) = \arg \max_{(\alpha, \beta_{th})} \xi \bar{V}_S(\alpha, \beta_{th}; \pi_0) + (1 - \xi) \bar{V}_P(\alpha, \beta_{th}; \pi_0) - \lambda C_S(\alpha, \beta_{th}; \pi_0),$$  
(15)

where the scalar non-negative parameters $\lambda \geq 0$ and $\xi \in (0, 1)$ capture the desired trade-off. The intuition behind (15) is that, in order to schedule the optimal control action $u_k = 1 - b_k$ under perfect state information and minimize scheduling errors, the CC needs to infer accurately the underlying state $b_k$. However, the acquisition of such state information is costly, and its cost is captured by the Lagrangian term $-\lambda C_S(\alpha, \beta_{th}; \pi_0)$.

The policy $(\alpha^*, \beta_{th}^*)$ can be determined via DP. However, the DP algorithm suffers from the curse of dimensionality. In fact, the optimal policy operates based on the prior/posterior belief, over a $2^F$ dimensional state space (the number of possible realizations of $b_k$), resulting in huge complexity. Moreover, the DP algorithm and the operation of the system requires knowledge of the functions $I$ and $\bar{I}$. While these can be determined exactly via (10) and (8), their exact computation is a combinatorial problem, requiring marginalization over the $2^F$ possible realizations of $b_k$. In the next section, we propose a compact belief representation which enables a more efficient optimization and operation of the system, and a scalable state estimator based on sparse recovery algorithms.

**IV. COMPACT BELief REPRESENTATION**

As a first approximation, we assume that the posterior belief, in each slot $k$, is statistically independent across spectrum bins, i.e., we approximate $\hat{\pi}_k$ with the posterior expected spectrum occupancy $\beta_k = \beta(\hat{\pi}_k)$. Given $\hat{\pi}_k = \chi(\beta(\hat{\pi}_k) \leq \beta_{th}(\hat{\pi}_k))$, the next prior belief is thus also independent across spectrum bins, as can be seen from (10), and is thus univocally expressed by the prior expected spectrum occupancy $\beta_{th,k+1}$, with $i$th component

$$\beta_{k+1,i} = \beta_{k,i} P_B(1, u_k,i) + (1 - \beta_{k,i}) P_B(0, u_k,i).$$  
(16)

The prior belief can thus be updated with linear complexity with respect to $F$, rather than exponential as in (10).

**Remark 2** In general, updating the next prior belief based on this independence approximation may lead to sub-optimal performance. However, note that, from (14) and (6) and given $\beta_k$, the expected SU/PU throughput in slot $k$ is given by

$$\sum_{b_k} \bar{v}_k(b_k) V_X(b_k, u_k) = \sum_{i=1}^{F} \left[ \hat{\beta}_{k,i} v_X(1, u_k,i) + (1 - \hat{\beta}_{k,i}) v_X(0, u_k,i) \right],$$  
(17)

and is thus a function of $\hat{\beta}_k$ only, so that no loss of optimality is incurred in the instantaneous reward, but only in the prior-posterior belief updates.

The system thus operates based on $\beta_k$ and $\hat{\beta}_k$, rather than $\pi_k$ and $\hat{\pi}_k$. Thus, we redefine the sensing policy as $\alpha(\beta_k)$, and the scheduling policy as $\beta_{th}(\hat{\beta}_k)$. However, determining $\beta_k$ from $(\beta_k, A_k, y_k)$ has exponential complexity, and a high-dimensional argument space $(\beta_k, A_k, y_k)$. In order to overcome the dependence on $A_k, y_k$ in Sec. IV-A we propose to decouple the state estimator from the spectrum scheduling.

**A. Decoupled estimator and spectrum scheduling**

In this formulation, the estimator is treated as a black-box with input $(\beta_k, A_k, y_k)$, which outputs a MAP estimate $b_k^{(MAP)}$ and the number of measurements received, $R_k$. The CC, on the other hand, has access only to the prior belief $\beta_k$, and the output from the estimator, $(\hat{\beta}_k^{(MAP)}, R_k)$, but not to the realizations of $(A_k, y_k)$. Given $(\beta_k, b_k^{(MAP)}, R_k)$, the CC then computes the posterior false-alarm probability $\hat{P}_{E,k}^{(1)}$ and the posterior mis-detection probability $\hat{P}_{E,k}^{(0)}$ for the bins detected as busy and idle, respectively. This characterization inherently assumes that the number of measurements received $R_k$ is the main parameter that drives the spectrum estimation performance, rather than the specific realization of $(A_k, y_k)$, so that the latter are treated as nuisance parameters. Hence, $\hat{P}_{E,k}^{(0)}$ and $\hat{P}_{E,k}^{(1)}$ are obtained by marginalization, as shown in (22). Given $\beta_k$, $b_k^{(MAP)}$, $\hat{P}_{E,k}^{(0)}$ and $\hat{P}_{E,k}^{(1)}$, the CC thus approximates the posterior belief as

$$\hat{\beta}_k = \left[ P(b_k,i = 1 | b_k^{(MAP)}, \hat{\beta}_k, \hat{P}_{E,k}^{(0)}, \hat{P}_{E,k}^{(1)}) \right]_{vi} = b_k^{(MAP)} \left( 1 - \hat{P}_{E,k}^{(1)} \right) + (1 - b_k^{(MAP)}) \hat{P}_{E,k}^{(0)},$$  
(18)

based on which it schedules $u_k = \chi(\beta(\hat{\pi}_k) \leq \beta_{th}(\hat{\pi}_k))$ and it updates the next prior belief as in (16). This approach suggests a compact state space representation of the prior and posterior beliefs $\beta_k$ and $\hat{\beta}_k$. In particular, $\beta_k$ can be represented by the tuple $\beta_k \triangleq (\nu_k, P^{(0)}_{E,k}, P^{(1)}_{E,k})$, where $\nu_k \triangleq \sum_i b_k^{(MAP)}$ represents the number of bins detected as busy (this representation is unique up to ordering of the bins, i.e., up to a permutation matrix $P$). From (16), the scheduling threshold $\beta_{th,k} = \beta_{th}(\hat{\beta}_k)$ the prior belief of the $\nu_k$ bins detected as busy ($\beta_k^{(MAP)} = 1$), is updated as

$$\beta_{k+1,i} = (1 - \hat{P}_{E,k}^{(1)}) P_B(1, \chi(\hat{P}_{E,k}^{(1)} \geq 1 - \beta_{th,k})) + \hat{P}_{E,k}^{(1)} P_B(0, \chi(\hat{P}_{E,k}^{(1)} \geq 1 - \beta_{th,k})) \equiv 1 - \hat{P}_{E,k}^{(1)}.$$  
(19)

where we have defined the prior false-alarm probability $\hat{P}_{E,k}^{(1)} \in [0, 1]$ (with respect to the MAP estimate from
slot \( k \)). This is a function of \( (\hat{P}^{(1)}_{E,k}; \beta_{th,k}) \), denoted as \( P^{(1)}_{E,k+1} \). Similarly, the prior belief of the \( F - \nu_k \) bins detected as idle \( (\beta^{(MAP)}_{k,i} = 0) \) is updated as

\[
\hat{\beta}_{k+1,i} = \hat{P}^{(0)}_{E,k} P_B(1, \chi(\hat{P}^{(0)}_{E,k} \leq \beta_{th,k})) + (1 - \hat{P}^{(0)}_{E,k}) P_B(0, \chi(\hat{P}^{(0)}_{E,k} \leq \beta_{th,k})) = P^{(0)}_{E,k+1},
\]

where we have defined the prior mis-detection probability \( P^{(0)}_{E,k+1} \in [0, 1] \). This is a function of \( (\hat{P}^{(0)}_{E,k}; \beta_{th,k}) \), denoted as \( \hat{P}^{(0)}_{E,k+1} = P^{(0)}_{E,k}(\hat{P}^{(0)}_{E,k}; \beta_{th,k}) \). Combining the two cases \( \beta^{(MAP)}_{k,i} \in \{0, 1\} \), the next prior belief \( \beta_{k+1} \) is given by

\[
\beta_{k+1} = \hat{b}^{(MAP)}_{k} (1 - \hat{P}^{(0)}_{E,k+1}) + (1 - \hat{b}^{(MAP)}_{k} \hat{P}^{(0)}_{E,k+1}),
\]

so that it can be compactly represented by the tuple \( s_{k+1} \triangleq (\nu_k, P^{(0)}_{E,k+1}, P^{(1)}_{E,k+1}) \) up to ordering of the elements in \( \beta_{k+1} \).

In general, we thus compactly represent the posterior and prior beliefs by the tuple \( s = (\nu, P_0, P_1) \in \{0, 1, \ldots, F\} \times \{0, 1\}^2 \), termed the compressed belief state (CBS), and restrict the system to operate based on the CBS rather than the original prior and posterior beliefs \( \pi_k \) and \( \bar{s}_k \). Accordingly, we redefine the sensing policy as \( \alpha(s) \) and the scheduling policy as \( \beta_{th}(s) \).

B. CBS transition probabilities

Note that \( \hat{b}^{(MAP)}_{k} \) can be written as a deterministic function of the prior CBS \( s_k \), and of \( (P^i_k, A_k, y_k, R_k) \), where \( P^i_k \) is a permutation matrix, which accounts for the ordering of the prior beliefs. We denote such function in the ith bin as \( \hat{b}^{(MAP)}_{i} \). Given \( s_k, \nu_k, R_k \), the CC thus computes the posterior mis-detection and false-alarm probabilities as

\[
\hat{P}^{(0)}_{E,k} = P^{(0)}_{E,k}(s_k; \nu_k, R_k) \triangleq \mathbb{E} \left[ \sum_{i=1}^{F} b_{k,i} \left[ 1 - \text{MAP}_i(s_k, P^{(1)}_k A_k, y_k, R_k) \right] \right],
\]

\[
\hat{P}^{(1)}_{E,k} = P^{(1)}_{E,k}(s_k; \nu_k, R_k) \triangleq \mathbb{E} \left[ \sum_{i=1}^{F} b_{k,i} \left[ \text{MAP}_i(s_k, P^{(1)}_k A_k, y_k, R_k) \right] \right]
\]

where the expectation is with respect to the realization of \( P^{i}_k, A_k, y_k \), given \( s_k, \nu_k, R_k \).

Note that \( \nu_k = \sum_i \text{MAP}_i(s_k, P^{(1)}_k A_k, y_k, R_k) \) is a random variable, which depends on the realization of \( P_k A_k, y_k, R_k \) and on the estimator employed. We denote its distribution as \( P^{(i)}_{\nu, \nu, R}(\nu_k | s_k, r) = P(\nu_k | s_k, R_k = r) \), after marginalization with respect to \( (P^{(1)}_k A_k, y_k) \). Therefore, the transition probability from the prior CBS \( s_k \) to the posterior CBS \( \hat{s}_k = (\nu_k, \hat{P}^{(0)}_{E,k}(s_k; \nu_k, r), \hat{P}^{(1)}_{E,k}(s_k; \nu_k, r)) \), given the sensing action \( \alpha_k = \alpha(s_k) \), is

\[
\mathbb{P}(\hat{s}_k | s_k, \alpha_k) = P_E(r | \alpha_k) P^{(0)}_{\nu, \nu, R}(\nu_k | s_k, r).
\]

Given \( \hat{s}_k = (\nu_k, \hat{P}^{(0)}_{E,k}, \hat{P}^{(1)}_{E,k}) \) and the scheduling threshold \( \beta_{th,k} = \beta_{th}(s_k) \), the next prior CBS is given by

\[
s_{k+1} = (\nu_k, P^{(0)}_{E,k}(\hat{P}^{(0)}_{E,k}, \beta_{th,k}), P^{(1)}_{E,k}(\hat{P}^{(1)}_{E,k}, \beta_{th,k})),
\]

denoted as the function \( s_{k+1} = s(\hat{s}_k, \beta_{th,k}) \). Hence, the transitions probabilities of the CBS as a function of the joint sensing scheduling policy \( (\alpha, \beta_{th}) \) have been characterized.

The functions \( P^{(i)}_{\nu, \nu, R}(\nu_k | s_k, r) \), \( i \in \{1, 2\} \) and \( P^{(i)}_{\psi, \psi, R}(\nu_k | s_k, r) \) cannot be determined analytically. Therefore, we recur to Monte-Carlo simulation to approximate them.

C. Optimization based on CBS

The sensing policy \( \alpha(s) \) and the scheduling policy \( \beta_{th}(s) \) can be determined numerically via DP, which solves, in each stage:

Algorithm 1 (DP stage \( i \) over CBS)

1) Scheduling optimization stage: compute, for each posterior CBS \( \bar{s} = (\nu, \hat{P}^{(0)}_{E}, \hat{P}^{(1)}_{E}) \),

\[
M^{(i)}_1(\bar{s}) = \max_{x \in \{0, 1\}} \gamma M^{(i-1)}_2(\bar{s}(x, x))
\]

\[
+ \nu (1 - \hat{P}^{(1)}_{E})(\chi[s_1, u_1(x)] + (1 - \chi) v_P(1, u_1(x))
\]

\[
+ \nu \hat{P}^{(1)}_{E}(\chi[s_0, u_0(x)] + (1 - \chi) v_P(0, u_0(x))
\]

\[
+ (F - \nu)(1 - \hat{P}^{(0)}_{E})(\chi[s_0, u_0(x)] + (1 - \chi) v_P(0, u_0(x))],
\]

where \( u_1(x) = \chi(x \geq 1 - \hat{P}^{(1)}_{E}) \) and \( u_0(x) = \chi(x < \hat{P}^{(0)}_{E}) \); the maximizer is the optimal scheduling threshold in the ith stage of the algorithm, denoted as \( \beta^{(i)}_{th}(s) \).

2) Sensing optimization stage: compute, for each prior CBS \( s \),

\[
M^{(i)}_2(s, r) = \sum_{\nu = 0}^{F} P^{(0)}_{\psi, \psi, R}(\nu | s, r) M^{(i)}_1(\nu, \hat{P}^{(0)}_{E}(s, \nu, r), \hat{P}^{(1)}_{E}(s, \nu, r))
\]

\[
M^{(i)}_2(s) = \max_{y \in \{0, 1\}} -\lambda y + \sum_{r = 0}^{F} P_R(r | y) M^{(i)}_1(s, r);
\]

the maximizer is the optimal sensing policy in the ith stage of the algorithm, denoted as \( \alpha^{(i)}(s) \).

The term \( M^{(i)}_1(s_k) \) represents the cost-to-go function, starting from the posterior CBS \( s_k \) and before \( u_k \) is scheduled, and includes the instantaneous expected SU/PU throughput accrued in slot \( k \); the term \( M^{(i)}_1(s_k, r) \) represents the cost-to-go function, starting from the prior CBS \( s_k \), given that \( R_k = r \) measurements are collected in the sensing phase; finally, the term \( M^{(i)}_2(s_k) \) represents the cost-to-go function, starting from the prior CBS \( s_k \) and before sensing, and includes the instantaneous expected sensing cost \( y \) incurred in slot \( k \) by each SU.

Note that in the scheduling optimization stage, it is sufficient to optimize the threshold \( x \) over the set \( \{0, \frac{1 - \hat{P}^{(1)}_{E} + \hat{P}^{(2)}_{E}}{2}, 1\} \). In fact, the posterior belief \( \hat{b}_{k,i} \) can only take values in the set \( \{1 - \hat{P}^{(1)}_{E}, \hat{P}^{(2)}_{E}\} \). In (25), we have used (17). The terms \( \nu (1 - \hat{P}^{(1)}_{E}) \) and \( (F - \nu)(1 - \hat{P}^{(0)}_{E}) \) are the expected number of bins correctly detected as busy and idle, respectively, and \( \nu \hat{P}^{(1)}_{E} \) and \( (F - \nu) \hat{P}^{(0)}_{E} \) are the expected number of bins erroneously detected as busy and idle, respectively.
D. MAP estimator via convex relaxation

Given the prior belief $\beta_k$ (a function of the CBS), the MAP estimate $b^{(MAP)}_k$ is given by the solution of

$$b^{(MAP)}_k = \arg \max_{b \in \{0,1\}^F} \mathbb{P}(b_k = b | \beta_k, A_k, y_k)$$

$$= \arg \min_{b \in \{0,1\}^F} \frac{1}{2} ||y_k - A_k^T b||_F^2 + \sigma^2 \sum_i b_i \ln \left( \frac{1 - \beta_{k,i}}{\beta_{k,i}} \right). \quad (27)$$

In particular, letting $b^{(MAP)}_k = \chi(\beta_k \geq 0.5)$ be the MAP estimate based on the prior belief $\beta_k$, we have

$$\hat{b}^{(MAP)}_k = b^{(MAP)}_k \oplus \epsilon^{(MAP)}_k,$$ \quad (28)

where $\oplus$ denotes the component-wise XOR operation, and $\epsilon^{(MAP)}_k$ is a residual uncertainty vector. In particular, substituting (28) into (27), it can be shown that $\epsilon^{(MAP)}_k$ is given by the solution of the optimization problem

$$\epsilon^{(MAP)}_k = \arg \min_{e \in \{0,1\}^F} \frac{1}{2} ||y_k - A_k^T e||_F^2 + \mu_T e, \quad (29)$$

where we have defined $\hat{y}_k = y_k - A_k^T b_k^{(MAP)}$ as the measurement vector corrected by removing the offset from the prior MAP estimate $b_k^{(MAP)}$, and, similarly, $\hat{A}_k = (I - 2\text{diag}(b_k^{(MAP)})) A_k$ is the corrected measurement matrix, and $\mu_k$ is a Lagrangian multiplier column vector with components $\mu_{k,i} = \sigma^2 \ln \left( \frac{1 - \beta_{k,i}}{\beta_{k,i}} \right)$. Note that the Lagrangian vector $\hat{A}_k$ weights the error components $e_i$ differently, based on their prior log-likelihood. Moreover, from the definition of $b_k^{(MAP)}$, we have that $\mu_{k,i} \geq 0$, with equality if and only if $\beta_{k,i} = 0.5$. Therefore, if the prior MAP estimate is good, the sparsity constraint satisfies $\mu_k \gg 1$, so that (29) induces a sparse vector $\epsilon^{(MAP)}_k$ exhibiting only few non-zero components.

In order to overcome the exponential complexity of (29), we propose the following convex relaxation:

$$\tilde{\epsilon}_k = \arg \min_{e \in \{0,1\}^F} \frac{1}{2} ||\hat{y}_k - \hat{A}_k^T e||_F^2 + \mu_T^e e, \quad (30)$$

i.e., the optimization is over the convex set $[0,1]^F$, rather than the discrete one $\{0,1\}^F$, and can thus be solved using convex optimization techniques. In particular, it is a quadratic programming problem minimizing a least-squares term, plus an $\ell_1$ regularization term, which induces sparsity constraints on the optimal solution $\tilde{\epsilon}_k$, and can thus be solved via sparse recovery algorithms. Note that the optimal solution $\tilde{\epsilon}_k$ is not feasible with respect to the original optimization problem (29). A feasible point is thus obtained using, e.g., a minimum distance criterion $\hat{\epsilon}^{(MAP)}_k = \chi(\epsilon_k \geq 0.5)$. This solution, in general, does not coincide with (29), and it can thus be improved with local search algorithms, e.g., hill climbing [23].

V. NUMERICAL RESULTS

In this section, we present numerical results. We consider a scenario with the following system parameters: $F = 10$, $N_S = 200$, $B = 20$, $\rho_S = \rho_P = 0.1$, $\zeta = 0.3$, $\sigma^2 = 1$, $\sigma^2_\tilde{S} = 1/20$, $\gamma = 1$. We consider the following policies:

- No sensing (NS) policy, in which no sensing is performed by the SUs, i.e., $\alpha_k = 0, \forall k$.
- Static (or Non-Adaptive) sensing (NAS) policy, in which sensing occurs with a fixed probability $\alpha_k = a$ in each slot, independently of the CBS; this is obtained after 50 iterations of Algorithm 1 (Sec. IV-C.), where the maximization (26) in the sensing optimization stage is replaced with an evaluation of $M_2^{(i)}(s)$ under the non-adaptive sensing policy $\alpha_k = a$.
- Adaptive sensing (AS) policy, computed after 50 iterations of Algorithm 1 (Sec. IV-C.).
- Full sensing (FS) policy, in which sensing occurs with a fixed probability $\alpha_k = B/N_S$, which maximizes the throughput of the feedback channel.

Note that NS and FS yield a lower and upper bound in the SU-PU throughput trade-off, respectively. In fact, the choice of $\alpha_k$ for FS maximizes the success probability $\kappa_k e^{-\kappa_s}$, where $\alpha_k = B\kappa_k/N_S$, and thus saturates the capacity of the feedback channel when $N_S \to \infty$, i.e., any larger sensing probability $\alpha_k > B/N_S$ incurs a degradation in the number of measurements collected at the CC, hence worse performance. The parameters $\xi$ and $\lambda$ in Algorithm 1 are varied in order to obtain different operational points, in terms of SU-PU throughputs and sensing cost. These policies are evaluated via simulation, over a time-horizon of $T = 10^4$ slots, and the time-average metrics are computed.

Fig. 2 plots the resulting trade-off between the PU and SU throughputs. We note that the best performance is obtained by FS, followed by AS, NAS and NS. However, FS incurs a significantly larger sensing cost for the SUs (10-fold), which may be unfeasible in practical deployments. In contrast, a poor performance is attained by NS, whose sensing cost is 0. On the other hand, AS and NAS optimally balance the trade-off between the SU-PU throughputs and the sensing cost.

We notice that AS improves the SU throughput over NAS up to 11%. This result can be explained with the help of Fig. 3, which plots the normalized sensing probability $\kappa(s) = \alpha(s)N_S/B$, as a function of the CBS $s = (F/2, P_{E_0}^{(0)}, P_{E_1}^{(1)})$, where $P_{E_0}^{(0)} = P_{E_1}^{(1)} \in [0.5]$. We notice that, under AS, sensing is more frequent when the spectrum state $b_k$ is more uncertain, i.e., $P_{E_0}^{(0)} = P_{E_1}^{(1)}$ approaches 0.5, so that more measurements are received at the CC to estimate $b_k$. On the other hand, when
Figure 3. Normalized transmission probability per channel $\kappa_k = N_S \alpha_k / B$ for AS and NAS, as a function of the CBS $(F/2, p, p)$, $p \in [0, 0.5]$.

$P_E^{(0)} = P_E^{(1)}$ is small, the sensing probability approaches 0, so that the SUs remain idle and save energy. In this case, no measurements are collected at the CC, since $b_k$ can be reliably estimated based on the prior belief. In contrast, under NAS, sensing is independent of the CBS, resulting in waste of resources when the CBS is good.

**VI. Conclusions**

In this paper, we have presented a cross-layer framework for joint distributed sensing, estimation and control 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. We have proposed a method to effectively exploit the sparsity in the spectrum occupancy dynamics; by leveraging the spectrum occupancy estimate in the previous slot, the CC needs to estimate only a sparse residual uncertainty vector via sparse recovery techniques, so that only few measurements suffice. In order to reduce the dimensionality of the problem, we have proposed a compact state space representation, which enables an efficient optimization of the joint scheduling-sensing policy via DP, and decoupling of the state estimator from the CC, which enables the use of sparse recovery algorithms. Simulation results demonstrate improvements up to 11% in the SU throughput over a static scheme.

**REFERENCES**


