Energy-Efficient Spectrum Sensing for Cognitive Radio Enabled Remote State Estimation Over Wireless Channels

Xianghui Cao, *Member, IEEE*, Xiangwei Zhou, *Member, IEEE*, Lu Liu, *Student Member, IEEE*, and Yu Cheng, *Senior Member, IEEE*

Abstract—The performance of remote estimation over wireless channels is strongly affected by sensor data losses due to interference. Although the impact of interference can be alleviated by applying cognitive radio technique which features in spectrum sensing and transmitting data only on clear channels, the introduction of spectrum sensing incurs extra energy expenditure. In this paper, we investigate the problem of energy-efficient spectrum sensing for remotely estimating the state of a general linear dynamic system, and formulate an optimization problem which minimizes the total sensor energy consumption while guaranteeing a desired level of estimation performance. We model the problem as a mixed integer nonlinear program and propose a simulated annealing based optimization algorithm which jointly addresses when to perform sensing, which channels to sense, in what order and how long to scan each channel. Simulation results demonstrate that the proposed algorithm well balances the sensing energy and transmission energy expenditure and can achieve the desired estimation performance.

Index Terms—Cognitive radio, energy efficiency, optimization, simulated annealing, spectrum sensing, state estimation.

I. INTRODUCTION

E STIMATING the states of dynamic processes is a fundamental task in many real-time applications such as environment monitoring, health-care, smart grid, industrial automation and wireless network operations [2]–[4]. Among existing estimation algorithms, we consider Kalman filtering [5], which has been widely applied for estimating the state of wireless channels [6], local power of a mobile station in cellular networks [7], and number of active terminals in wireless local area networks [8], [9], etc.

In many cases, sensor measurements are transmitted through wireless media to a remote estimator (or controller) which then performs state estimation and makes certain decisions based on the estimation results. However, due to factors such as interference, sensor packets may randomly get lost before successfully arriving at the estimator. Under packet losses, the estimator has intermittent observations of the dynamic process such that the amount of information for state estimation is decreased, which can significantly affect the estimation performance. Kalman filter has been shown optimal for estimating linear processes with lossy measurements; however, the estimation stability imposes strict constrains on the packet loss rate [5], [10], [11]. Many existing studies only consider one wireless channel, while practically we can have multiple accessible channels which may help reduce packet loss.

Multi-channel wireless communications can significantly reduce interference and improve the communication throughput [12]. To alleviate the impact of interference and improve estimation performance, a promising solution is to explore the availability of multiple channels using cognitive radio (CR) technology [13] for sensor data transmission. With the CR technology, before accessing a channel, a sensor performs spectrum sensing to assess the availability of that channel. In such a "listen before talk" manner, the sensor transmissions can avoid heavy interferences from other transmissions on the same channel and hence the sensor packets become less likely to get lost. As a result, the estimation performance is expected to be improved.

In the literature, a lot of studies focus on spectrum sensing efficiency to improve network throughput [14]–[17], while only a few consider the design and evaluation of spectrum sensing for the state estimation problem [18], [19]. However, studies in [18], [19] focus on actively applying spectrum sensing to improve estimation performance, without taking the critical and challenging energy efficiency issue into account. Since both data transmission and spectrum sensing are energy consuming, the system energy efficiency leads to a tradeoff between spectrum sensing and transmission. In this paper, we investigate the problem of energy-efficient spectrum sensing for state estimation. Specifically, with the objective to minimize

X. Cao is with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616 USA, and also with the Department of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: xh.cao@ieee.org).

X. Zhou is with the Department of Electrical and Computer Engineering, Southern Illinois University Carbondale, Carbondale, IL 62901 USA (e-mail: xzhou@engr.siu.edu).

L. Liu and Y. Cheng are with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616 USA (e-mail: lliu41@hawk.iit.edu; cheng@iit.edu).

energy consumption with guaranteed estimation performance, we investigate the following fundamental questions:

- When to perform spectrum sensing?
- Which channels to sense?
- In what order to scan the channels?
- How long to spend on scanning each channel?

The first question relates to the problem of scheduling sensor activities. The second and third questions concern about channel selection and channel sensing order. The last question arises because both sensing accuracy (and hence estimation performance) and energy consumption depend on the sensing time. Although there are some studies on each of the above questions [20]–[27], there lacks a holistic solution to these questions in the context of state estimation.

In this paper, we study the problem of energy-efficient spectrum sensing strategy for state estimation, aiming at systematically addressing the above four fundamental issues. Specifically, focusing on a general linear dynamic process, we consider the problem of minimizing the energy consumption of the sensor while guaranteeing a desired level of estimation performance. The main contributions of this paper can be summarized as follows. 1) We provide a cyber-physical model of the whole system for state estimation and formulate the above optimization problem as a mixed integer nonlinear program (MINLP), subjecting to an estimation performance constraint. 2) We first exploit the single-channel case and derive a condition under which the estimation error covariance is stable in mean sense. Since the mean estimation error covariance is usually a random value and may vary slightly but not converge along time, the explicit expression for the mean estimation error covariance is difficult to obtain. We thus resort to a close approximation of the constraint which results in an approximated optimization problem whose solution suffices the original problem. We also provide analytical results of the optimization solution. 3) We extend the approximate MINLP formulation into the general multi-channel case and propose an algorithm based on both Lagrange multiplier method and simulated annealing. The new algorithm can jointly determine all of the following: the optimal sensing schedule, the optimal set of channels to sense, the optimal channel sensing order and the optimal sensing time on each channel. 4) Extensive simulation results demonstrate that the proposed algorithm well balances the sensing energy and transmission energy expenditure and can achieve the optimal goal. To the best of our knowledge, we for the first time formulate and address the energy efficiency issue in spectrum sensing based state estimation over multiple wireless channels.

Some primary results have been presented in [1]. The remainder of the paper is organized as follows. Section II presents more related work. Section III presents system model and the optimization problem. The problem in the single-channel case is solved in Section IV. Extension to multi-channel cases is presented in Section V. Section VI presents simulation results, and Section VII concludes the paper.

II. RELATED WORK

The stability of Kalman filter under random packet losses has gained intensive studies recently. In the case that the packet losses are independently and identically distributed, the estimation error (in mean square sense) is stable only when the packet loss rate is below a certain bound [5]. Recently, there has been a large volume of literature investigating the problem of state estimation stability in various wireless communication situations, e.g., Markovian and semi-Markov packet losses [10], [28] and more general packet loss processes [11]. These results explicitly show that the estimation stability heavily depends on the packet loss process. However, most studies consider only one wireless channel for sensor data transmission.

Basically, there are two major challenges in the application of spectrum sensing for remote state estimation over multiple wireless channels. First, the spectrum sensing results may be inaccurate [26], which in turn affects the subsequent transmission successfulness and consequently the estimation performance. For example, spectrum sensing on a busy channel may mistakenly indicate that the channel is currently idle, and a subsequent transmission on it will get collided and probably lost. Considering that multiple channels can be accessed for sensor data transmission, Ma *et al.* derive an optimal estimation algorithm and analyze the estimation stability assuming a semi-Markov channel model and accurate spectrum sensing [18], [29]. With inaccurate spectrum sensing, the question of whether and to what extent the state estimation performance can be improved is addressed in [19].

Energy efficiency is an important design issue for today's wireless systems [30]–[33]. System energy efficiency is the second challenging issue which further introduces the sensor scheduling, channel selection, sensing order and sensing time optimization problems. It has been proved that the optimal data transmission schedule can be approximated by periodic (not necessarily strictly periodic) scheduling with arbitrarily close performance [34]. From the energy perspective, for a class of state estimation problems over a finite long time horizon, it also has been shown that the optimal sensor schedule is to distribute the data transmission time along the time horizon as uniform as possible [20]. These studies focus on scheduling the transmitting time; while in our case, we consider scheduling the spectrum sensing time. Aside from the fact that they only consider one channel, a major difference in our case is that, after carrying out sensing, the sensor may not transmit data if the sensing result indicates that the channel is unavailable.

The problems of channel selection and channel sensing order are mostly studied within the communication community. A three-step channel selection mechanism is proposed to maximize the channel efficiency of cognitive radio users [21]. In [22], a polynomial-time algorithm is designed to select channels out of a set of candidates to optimize user profit (e.g., throughput). The optimal channel sensing order in cognitive radio networks also has been investigated and effective searching algorithms have been designed in the literature [23], [25]. While many existing studies target at maximizing the overall throughput, in this paper, we emphasize on the reliability of each sensor transmission. Moreover, rather than continuously transmit a considerable amount of data after sensing an idle channel, the transmission in our case happens sporadically (as a consequence of spectrum sensing scheduling) and each time only the latest sensor packet is transmitted for state estimation.
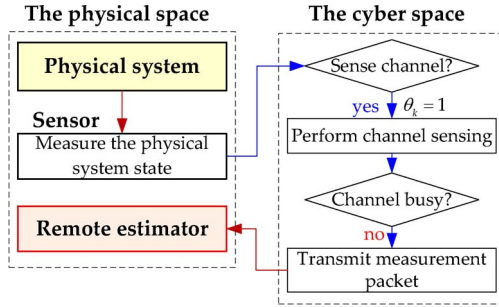
Fig. 1.   A cyber-physical view of the system architecture.

The optimal sensing time problem for spectrum efficiency maximization has been studied in [26]. Applying the cognitive radio technique to the communications between power consumers and the control unit in smart grid, the work in [27] shows that there exists a unique optimal sensing time which yields maximal system profit.

III. SYSTEM MODEL AND PROBLEM SETUP

In this paper, we focus on the problem of remotely estimating the state of a generic discrete-time linear dynamic system, one of the most widely studied dynamic system models. Examples are the channel coefficients dynamics [6] and the shadow process dynamics [7], where the system states are channel coefficients in MIMO networks and shadowing power fluctuations in cellular environment, respectively. Both of them also adopt linear measurement models (e.g., the measurement of received signal is modeled as a linear combination of channel coefficients and a channel thermal noise [6]). In other applications such as environment monitoring, mobile target tracking and industrial process control, linear systems are often adopted to model the dynamic processes therein as well as sensor measurements [2].

Consider that the process is periodically sampled by a wireless sensor and the measurement data will be transmitted to a remote estimator. At the beginning of each sampling period (also called step), the sensor first takes a measurement of the target dynamic system state and transmits this measurement packet to the remote estimator over a wireless channel. The sensor data transmission is augmented by the spectrum sensing technique. A transmission will be unsuccessful, i.e., a packet drop will happen, if the sensor transmits the packet when the channel is busy. The whole system can be described from a cyber-physical point of view. Fig. 1 illustrates the interactions between the cyber and physical spaces for state estimation over one wireless channel. The main notations used throughout this paper are listed in Table I. We use boldface lowercase and uppercase letters to represent column vectors and matrices, respectively. $\mathbb{E}[\cdot]$ and $\mathbb{P}[\cdot]$ denote the expectation and probability of a random variable, respectively. $(\cdot)^T$ denotes the transpose of either a vector or a matrix while $trace(\cdot)$ denotes the trace of a matrix.

A. Physical Space

The target dynamic system is modeled as follows (the discrete-time steps are determined by the sensor's sampling period T_s):

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{w}_k, \tag{1}$$

TABLE I
BASIC NOTATIONS

notation	Definition
$\mathbf{x}_k$	state of the target dynamic system at the beginning of step k
$\mathbf{y}_k$	sensor measurement about the dynamic system state in step k
γ_k	binary variable indicating whether $\mathbf{y}_k$ is successfully transmitted and received by the estimator in step k
T_s	Sensor's sampling period
t_x	transmission time of each measurement packet (the packet lengths are assumed the same)
CH_i	the i-th channel.
τ_i	channel sensing time over CH_i
$p_{B,i}, p_{I,i}$	probabilities that CH_i is busy and idle, respectively
$p_{d,i}, p_{f,i}$	correct and false detection probabilities on CH_i, respectively
$p_{tx,i}$	probability that the sensor transmit data over CH_i after the sensor senses it
$p_{c,i}$	probability of packet collision on CH_i

where $\mathbf{x} \in \mathbb{R}^{q_1}$ is the system state and q_1 is its dimension, $\mathbf{w}$ is the system noise with $\mathbb{E}[\mathbf{w}_k\mathbf{w}_k^T] = \mathbf{Q} \geq 0$.[1] $\mathbf{A}$ is a constant square matrix modeling the state dynamics in two successive time steps. Assume that $(\mathbf{A}, \mathbf{Q}^{\frac{1}{2}})$ is controllable [5].

The sensor measurement of the system state in kth step is modeled as

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k, \tag{2}$$

where $\mathbf{v}$ is the measurement noise with $\mathbb{E}[\mathbf{v}_k\mathbf{v}_k^T] = \mathbf{R} \geq 0$ and $\mathbb{E}[\mathbf{w}_i\mathbf{v}_j^T] = 0$. $\mathbf{y} \in \mathbb{R}^{q_2}$ (where q_2 is its dimension) and $\mathbf{C} \in \mathbb{R}^{q_2 \times q_1}$. Both $\mathbf{w}$ and $\mathbf{v}$ are assumed Gaussian with zero means. Assume that $\mathbf{C}$ has full column rank for simplicity.

The estimator applies the following modified Kalman Filter [5] to estimate the system state $\mathbf{x}$ recursively. Given the system functions as shown in (1) and (2), define $\hat{\mathbf{x}}_{k|k-1}$ and $\hat{\mathbf{x}}_{k|k}$ as the prediction and estimate of the system state at step k, respectively. Define $\mathbf{P}_{k|k-1} \overset{\Delta}{=} \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T]$ and $\mathbf{P}_{k|k} \overset{\Delta}{=} \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T]$ as the covariance of the prediction and estimation errors, respectively. The prediction can be calculated based on the system model in (1) as follows.

$$\begin{cases} \hat{\mathbf{x}}_{k|k-1} &= \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1}, \\ \mathbf{P}_{k|k-1} &= \mathbf{A}\mathbf{P}_{k-1|k-1}\mathbf{A}^T + \mathbf{Q}. \end{cases} \tag{3}$$

As mentioned above, packets may be dropped. Let $\gamma_k \in \{0,1\}$ represent whether the measurement packet is dropped or successfully received by the estimator in step k, i.e., $\gamma_k = 1$ if received and $\gamma_k = 0$ otherwise. $\mathbb{P}[\gamma_k = 1]$ characterizes the *successful packet transmission rate* which will be elaborated in the next subsection. In the case of $\gamma_k = 1$, according to the standard Kalman filter design, the state estimate can be updated as follows:

$$\begin{cases} \hat{\mathbf{x}}_{k|k} = & \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1}), \\ \mathbf{K}_k \overset{\Delta}{=} & \mathbf{P}_{k|k-1}\mathbf{C}^T \left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R}\right)^{-1}, \\ \mathbf{P}_{k|k} = & (\mathbf{I} - \mathbf{K}_k)\mathbf{P}_{k|k-1}, \end{cases} \tag{4}$$

with a given initial value $\mathbf{P}_{1|0} \geq 0$, where $\mathbf{I}$ is an identity matrix of compatible dimension. Otherwise, without receiving the measurement of the current system state from the sensor, the estimator has to use the prediction to update its estimate [5],

---

[1]For any square matrix $\mathbf{M}$, by $\mathbf{M} \geq 0$ (or $\mathbf{M} > 0$) we mean that $\mathbf{M}$ is a semipositive (or positive) definite matrix. For any two square matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ of the same dimension, $\mathbf{M}_1 \geq \mathbf{M}_2$ means that $\mathbf{M}_1 - \mathbf{M}_2 \geq 0$.

i.e., $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1}$. Obviously, the estimation performance relies on the random variables $\{\gamma_1, \ldots, \gamma_k\}$.

B. Cyber Space

The main task of the cyber sub-system is to decide whether to sense the channel and to transmit the measurement packet in each step. For ease of exposition, we focus on a single channel in this section. Extension to the multi-channel case is presented in Section V.

1) Channel Model: Let t_I and t_B represent the idle and busy periods of the channel, respectively. They model the activities of all users other than the sensor on that channel over a long time. Assume that t_I and t_B obey certain distributions as $t_I \sim \Gamma_I(\cdot)$ and $t_B \sim \Gamma_B(\cdot)$, where Γ_I and Γ_B are the cumulative distribution functions of t_I and t_B, respectively. Let $\mathbb{E}[t_I]$ and $\mathbb{E}[t_B]$ be the average idle and busy periods, respectively. Define

$$\rho = \frac{\mathbb{E}[t_B]}{\mathbb{E}[t_I]}, \tag{5}$$

which is an important characteristic of the channel. The probabilities that the channel is idle and busy are respectively

$$p_I = \frac{1}{1+\rho} \text{ and } p_B = \frac{\rho}{1+\rho}. \tag{6}$$

We assume that the sensing time τ is bounded within $[0, \bar{\tau}]$ and is small enough such that the channel state does not change during the sensing time. The sampling period $T_s \gg \max(\mathbb{E}[t_B], \mathbb{E}[t_I])$, so that the packet drop rate in the current sampling period is irrelevant with that in previous steps. Based on this, the measurement packet drop rate, i.e., $\mathbb{P}[\gamma_k = 0]$, also can be deemed time-independent.

Once the channel is idle at time t and the sensor transmits a packet from t, the transmission may still be unsuccessful if the channel enters a busy period during $[t, t + t_x]$. Let $\tilde{t}_I$ be the time that the channel remains in idle state, given that it is currently idle. According to [35], the probability density function of $\tilde{t}_I$ can be expressed as $\frac{1}{\mathbb{E}[t_I]}[1 - \Gamma_I(\tilde{t}_I)]$. Then, the probability that $\tilde{t}_I$ will last for at least t_x period of time (i.e., the transmission during $[t, t + t_x]$ is successful) is

$$\eta = \mathbb{P}[\tilde{t}_I > t_x] = 1 - \mathbb{P}[\tilde{t}_I < t_x]$$
$$= 1 - \frac{1}{\mathbb{E}[t_I]} \int_0^{t_x} [1 - \Gamma_I(t)] \, dt. \tag{7}$$

As a popular channel model [26], [36], both the busy and idle periods are assumed following Poisson distributions with mean values $\mathbb{E}[t_B] = \frac{1}{\beta}$ and $\mathbb{E}[t_I] = \frac{1}{\alpha}$, respectively. In this case, $\Gamma_I(t) = 1 - e^{-\alpha t}$, $\Gamma_B(t) = 1 - e^{-\beta t}$ and η can be simplified as

$$\eta = 1 - \alpha \int_0^{t_x} e^{-\alpha t} \, dt = e^{-\alpha t_x}.$$

2) Channel Sensing: Before transmitting a packet, the sensor must check the channel state and transmit packet only when the channel is in its idle state. The sensing process lasts for a period of τ, during which the sensor mainly analyzes signals received from the channel and then decides whether the channel is idle or not. Among channel sensing strategies, *energy detection* is one of the most popular methods [26], [37]. An important drawback of this technique is that the sensing results may be inaccurate, e.g., the sensing process may mistakenly report idle state of the channel when it is actually busy. Let s_c be the sensing outcome (with 0 indicates idle and 1 otherwise) and define following two probabilities.

$$p_d = \mathbb{P}[s_c = 0 | \text{ channel idle}] = Q\left((1 - \varepsilon_d)\sqrt{\tau W}\right), \tag{8}$$

$$p_f = \mathbb{P}[s_c = 0 | \text{ channel busy}] = Q\left((1 - \varepsilon_f)\sqrt{\tau W}\right), \tag{9}$$

where $\varepsilon_d > \varepsilon_f > 0$, W is the channel bandwidth, and $Q(z) \triangleq \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-\frac{\tau^2}{2}} d\tau$. p_d and p_f are called the correct and false detection probabilities, respectively. Equations (8) and (9) and more details about energy detection are presented in Appendix A.

After sensing the channel, the sensor will transmit packet only if the sensing result indicates an idle channel. Thus, the transmission probability is

$$p_{tx} = p_I p_d + p_B p_f = \frac{1}{1+\rho}(p_d + \rho p_f). \tag{10}$$

Define a sequence of variables $\{\theta_k \in \{0, 1\}\}_{k \geq 1}$ as

$$\theta_k = \begin{cases} 1, & \text{if to sense the channel in step } k, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Under the case that $\theta_k = 1$, the successful packet transmission rate can be given as follows.

$$\gamma \triangleq \mathbb{P}[\gamma_k = 1 | \theta_k = 1, s_{c,k} = 0] = p_I \eta p_d = \frac{\eta}{1+\rho} p_d, \tag{12}$$

where $s_{c,k}$ is the channel sensing result at step k.

Let $\Theta \triangleq \{k | \theta_k = 1\}$ be the *channel sensing schedule*. Define $\bar{\Theta} = \lim_{t \to \infty} \frac{1}{t} \sum_{k=1}^t \theta_k$.

C. Problem Formulation

Let e_s and e_{tx} denote the amounts of energy consumed by the sensor for sensing the channel and transmitting a measurement packet, respectively. If channel sensing is performed in step k, the average amount of energy consumed by the sensor in this step is $\varphi_s = \tau e_s + p_{tx} e_{tx}$. Therefore, under schedule Θ, the average energy consumption in a single step is

$$\bar{\varphi} = \frac{1}{\bar{\Theta}} \varphi_s = \frac{1}{\bar{\Theta}}(\tau e_s + p_{tx} e_{tx}). \tag{13}$$

As can be seen, we take both spectrum sensing energy and data transmission energy into account. Since we focus on the energy consumption of the sensor, the receiving energy at the remote estimator (which is supposed to have much richer energy resource) is not considered in our optimization problem. For the sensor, according to (13) and (10), the probability that it transmits measurement packets depends on the spectrum sensing results and further on the sensing time τ, in which sense the transmission energy depends on how much energy is spent on sensing the channels. Another fact is that the spectrum sensing energy is the energy expenditure for the sensor to receive and determine whether the channel state by means of energy detection. Since the communication energy often

dominates others, our optimization objective function in terms of the sum of spectrum sensing energy and transmission energy can be viewed as the total energy that the sensor spends for both receiving and transmitting.

The estimation performance can be characterized by the error covariance $\mathbf{P}_{k|k-1}$. Hereafter, $\mathbf{P}_k \triangleq \mathbf{P}_{k|k-1}$ for ease of presentation. Based on the estimation process in Section III-A, we can see that $\mathbf{P}_k$ is a function of the random variable γ_k; hence it is both random and time-varying and may not converge along an infinite horizon. Therefore, we consider the long-time average of the expected $\mathbf{P}_k$, i.e., $\frac{1}{L}\sum_{k=1}^{L}\mathbb{E}[\mathbf{P}_k]$, where L is a sufficiently large number. We aim to bound the average value below a user defined threshold $\bar{\mathbf{P}}$. Our optimization problem can be formulated as follows.

Problem 1: Find the optimal schedule Θ and channel sensing time τ to

$$\begin{cases} \min_{\Theta,\tau} & \bar{\phi} = \frac{1}{\bar{\Theta}}(\tau e_s + p_{tx}e_{tx}) \\ s.t. & \frac{1}{L}\sum_{k=1}^{L}\mathbb{E}[\mathbf{P}_k] \leq \bar{\mathbf{P}}, \\ & 0 \leq \tau \leq \bar{\tau}. \end{cases} \quad (14)$$

In the above, the objective is a function of the channel sensing time τ, the sensing schedule Θ_n, the channel sensing accuracy (p_d and p_f) and the characteristics of the channel (p_I and p_B). Meanwhile, as discussed in the physical space model, the estimation performance depends on the successfulness of packet transmissions $\{\gamma_k\}$ which further depends on the sensing accuracy and channel characteristics. Therefore, both the objective and constraints of Problem 1 depend on the models of both cyber and physical spaces of the whole cognitive radio enabled state estimation system.

In this paper, we restrict our attention to strict periodical sensing schedule, i.e., $\Theta = \Theta_n \triangleq \{0, n, 2n, \ldots\} = \{k_i | k_i = in, i \in \mathbb{N}^+ \cup \{0\}\}$. In this case, the objective reduces to

$$\min_{n,\tau} \quad \bar{\phi} = \frac{1}{n}(\tau e_s + p_{tx}e_{tx}). \quad (15)$$

For an arbitrary schedule Θ, the optimization problem becomes significantly complicated since the solution space for Θ is infinitely large even if $\bar{\Theta}$ is small. As presented in Section II, prior studies show that periodic schedules can approximate the optimal state estimation performance with arbitrary accuracy, which encourages us to focus on periodic schedule. Another reason for considering periodic schedule lies in its simplicity and ease of implementation.

As can be seen, Problem 1 is an MINLP. The problem in the case of multiple channels involves extra decision variables such as the set of channels to be sensed and the sensing order and is formulated in Section V.

IV. OPTIMIZATION IN SINGLE-CHANNEL CASE

Since we impose strict bound on the mean square estimation error in the optimization problem, a necessary condition is that the mean square estimation error is finite (i.e., the estimation is stable), to ensure the problem feasibility. Therefore, in this section, we first examine the stability of the estimation error in mean square sense. Then we present problem approximation which ensures the problem tractability.

A. Estimation Stability

To satisfy the constraints in (14), $\{\mathbb{E}[\mathbf{P}_k]\}$ must be stable, i.e., $\mathbb{E}[\mathbf{P}_k] < \infty, \forall k \geq 1$. For any $k \geq 1$, if $\theta_k = 1$, based on the estimation process in Section III-A, we have

$$\begin{aligned} \mathbf{P}_k &= \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q} \\ &\quad - \gamma_k \mathbf{A}\mathbf{P}_{k-1}\mathbf{C}^T\left(\mathbf{C}\mathbf{P}_{k-1}\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\mathbf{P}_{k-1}\mathbf{A}^T \\ &= (1-\gamma_k)\mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q} \\ &\quad + \gamma_k \mathbf{A}\left(\mathbf{P}_{k-1}^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\right)^{-1}\mathbf{A}^T \\ &= (1-\gamma_k)\mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q} + \gamma_k \mathbf{A}\Upsilon_{k-1}\mathbf{A}^T, \quad (16) \end{aligned}$$

where $\Upsilon_{k-1} \triangleq (\mathbf{P}_{k-1}^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C})^{-1}$. Since $\mathbf{P}_{k-1}^{-1} \geq 0$, Υ_{k-1} is upper-bounded by $(\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C})^{-1}$ (notice that $\mathbf{C}$ has full column rank) [19]. Otherwise, $\theta_k = 0$, which is similar to the case that the measurement packet gets lost. That is, $\mathbf{P}_k = \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q}$. Therefore, we can write compactly that $\mathbf{P}_k = h_{\theta_k\gamma_k}(\mathbf{P}_{k-1})$ with

$$\begin{aligned} h_{\theta_k\gamma_k}(\mathbf{P}_{k-1}) &\triangleq \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q} \\ &\quad - \theta_k\gamma_k \mathbf{A}\mathbf{P}_{k-1}\mathbf{C}^T\left(\mathbf{C}\mathbf{P}_{k-1}\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\mathbf{P}_{k-1}\mathbf{A}^T. \quad (17) \end{aligned}$$

We can obtain the following condition for the stability of $\{\mathbb{E}[\mathbf{P}_k]\}$ which is both necessary and sufficient. The proof is provided in Appendix B.

Theorem 1: $\forall n \geq 1$, $\{\mathbb{E}[\mathbf{P}_k]\}$ is stable if and only if

$$(1-\gamma)\lambda_{\max}^{2n}(\mathbf{A}) < 1, \quad (18)$$

where $\lambda_{\max}(\mathbf{A})$ is the maximum absolute values of matrix $\mathbf{A}$'s eigenvalues.

Since $p_d \leq 1$, (12) gives that $\gamma \leq \frac{\eta}{1+\rho}$. Therefore, an upper bound of n can be obtained based on (18) as follows.

$$n \leq \bar{n}_1 = \begin{cases} \left\lceil \frac{\ln(1+\rho) - \ln(1+\rho-\eta)}{2\ln(\lambda_{\max}(\mathbf{A}))} \right\rceil - 1, & \text{if } \lambda_{\max}(\mathbf{A}) > 1, \\ \infty, & \text{otherwise.} \end{cases} \quad (19)$$

B. Problem Approximation

As shown in (16), since $\mathbf{P}_{k-1}$ appears in the inverse term of Υ_{k-1}, $\mathbb{E}[\mathbf{P}_k]$ will depend on all possible values of the random sequence $\{\gamma_k\}_{k\geq 1}$. In this case, $\mathbb{E}[\mathbf{P}_k]$ may not necessarily converge to a steady value.[2] As a result, it is mathematically difficult to obtain the long-term average of $\mathbb{E}[\mathbf{P}_k]$. Instead, for the problem tractability, we resort to an upper bound of $\mathbb{E}[\mathbf{P}_k]$ to sufficiently satisfy the constraint in Problem 1. Based on Lemma 1 in [5], $h_\gamma(\cdot)$ is a concave function. Thus, applying Jensen's inequality and noticing that $\mathbb{E}[\mathbf{P}_k] = h_0(\mathbb{E}[\mathbf{P}_{k-1}])$ if $\theta_k = 0$, we get that $\mathbb{E}[\mathbf{P}_k] \leq h_{\theta_k\gamma}(\mathbb{E}[\mathbf{P}_{k-1}])$. Define a new sequence $\{\mathbf{Y}_k\}$ with

$$\mathbf{Y}_k \triangleq h_{\theta_k\gamma}(\mathbf{y}_{k-1}). \quad (20)$$

Then, $\mathbb{E}[\mathbf{P}_k] \leq \mathbf{Y}_k$ if we let $\mathbf{Y}_0 = \mathbf{P}_0$. Thus, $\{\mathbf{Y}_k\}$ serves as an upper-bound sequence of $\{\mathbb{E}[\mathbf{P}_k]\}$.

---

[2] An example in Section VI shows that $\mathbb{E}[\mathbf{P}_k]$ always varies along time.
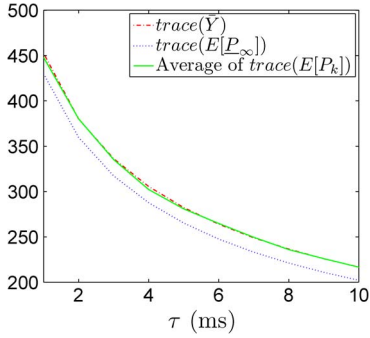
Fig. 2. Approximation error. We use matrix trace to produce scalar values as measures of the estimation performance.

Lemma 1: If condition (18) holds, there exists a unique value $\bar{\mathbf{Y}}(\gamma, n)$ such that

$$\lim_{L \to \infty} \frac{1}{L} \sum_{k=1}^{L} \mathbf{Y}_k = \bar{\mathbf{Y}}(\gamma, n). \tag{21}$$

In addition, $\bar{\mathbf{Y}}(\gamma, n)$ is monotonically decreasing as either γ increases or n decreases.

Therefore, for a sufficiently large L,

$$\frac{1}{L} \sum_{k=1}^{L} \mathbb{E}[\mathbf{P}_k] \leq \frac{1}{L} \sum_{k=1}^{L} \mathbf{Y}_k \to \bar{\mathbf{Y}}. \tag{22}$$

In this sense, the constraint in Problem 1 can be approximated as $\bar{\mathbf{Y}}(\gamma, n) \leq \bar{\mathbf{P}}$.

Remark 1 (Approximation Error): The approximation error is mainly caused by the application of Jensen's inequality. To characterize the approximation error, we consider the following lower bound of the long-term average of $\mathbb{E}[\mathbf{P}_k]$. Define a new sequence $\{\underline{\mathbf{P}}_k\}$ with

$$\underline{\mathbf{P}}_k = (1 - \theta_k \gamma_k) \mathbf{A} \underline{\mathbf{P}}_{k-1} \mathbf{A}^T$$
$$+ \theta_k \gamma_k \mathbf{A} \left(\mathbf{Q}^{-1} + \mathbf{C}^T \mathbf{R} \mathbf{C} \right)^{-1} \mathbf{A}^T + \mathbf{Q} \},$$

and $\underline{\mathbf{P}}_0 = \mathbf{P}_0$. By Theorem 5 in [19], we can prove that $\underline{\mathbf{P}}_k \leq \mathbf{P}_k, \forall k$, and that

$$\mathbb{E}[\underline{\mathbf{P}}_\infty] \triangleq \lim_{k \to \infty} \mathbb{E}[\underline{\mathbf{P}}_k]$$
$$= \sum_{\kappa=0}^{\infty} (1 - \gamma)^\kappa \mathbf{A}^\kappa \left[\gamma \mathbf{A} (\mathbf{Q}^{-1} + \mathbf{C}^T \mathbf{R} \mathbf{C})^{-1} \mathbf{A}^T + \mathbf{Q} \right] (\mathbf{A}^T)^\kappa,$$

where the infinite summation yields a finite value since $(1 - \gamma) \lambda_{\max}^2(\mathbf{A}) < 1$ which is ensured by condition (18) as in Theorem 1. In fact, $\mathbb{E}[\underline{\mathbf{P}}_\infty]$ is the unique solution of the Lyapunov equation $\mathbf{X} = (1 - \gamma) \mathbf{A} \mathbf{X} \mathbf{A}^T + \gamma \mathbf{A} (\mathbf{Q}^{-1} + \mathbf{C}^T \mathbf{R} \mathbf{C})^{-1} \mathbf{A}^T + \mathbf{Q}$. Therefore, the approximation error is bounded by

$$\frac{1}{L} \sum_{k=1}^{L} \mathbf{Y}_k - \frac{1}{L} \sum_{k=1}^{L} \mathbb{E}[\mathbf{P}_k] \leq \frac{1}{L} \sum_{k=1}^{L} (\mathbf{Y}_k - \underline{\mathbf{P}}_k) \to \bar{\mathbf{Y}} - \mathbb{E}[\underline{\mathbf{P}}_\infty].$$

To show the approximation error, we conduct simulations using the same settings as in Section VI. The results are shown in Fig. 2. As illustrated, both the true approximation error and its bound as presented above are very low.

Due to the monotonicity of $\bar{\mathbf{Y}}(\gamma, n)$ in γ, it is equivalent to say that $\gamma \geq \underline{\gamma}(n)$ where $\underline{\gamma}(n)$ is the unique solution of γ to $\bar{\mathbf{Y}}(\gamma, n) = \bar{\mathbf{P}}$. On the other hand, since $\gamma \leq \frac{\eta}{1+\rho}$, the inequality $\bar{\mathbf{Y}}(\frac{\eta}{1+\rho}, n) \leq \bar{\mathbf{Y}}(\gamma, n) \leq \bar{\mathbf{P}}$ yields another upper bound on n:

$$n \leq \bar{n}_2 = \max \left\{ \tilde{n} \middle| \bar{\mathbf{Y}} \left(\frac{\eta}{1+\rho}, \tilde{n} \right) \leq \bar{\mathbf{P}} \right\} < \infty. \tag{23}$$

Therefore, we obtain an approximation of Problem 1 as below.

Problem 2: Find the optimal schedule Θ_n and channel sensing time τ to

$$\begin{cases} \min_{n, \tau} & \bar{\varphi} = \frac{1}{n} (\tau e_s + p_{tx} e_{tx}) \\ s.t. & \gamma \geq \underline{\gamma}(n), \\ & n \leq \bar{n} = \min\{\bar{n}_1, \bar{n}_2\}, \\ & 0 \leq \tau \leq \bar{\tau}. \end{cases} \tag{24}$$

The solution to the above problem will be presented in Section V since Problem 2 is a special case of Problem 1 for multi-channel cases. For the single-channel optimization problem, according to the analysis below, the objective is either convex or concave but monotone in τ while the constraint is monotone. Therefore, the optimal solutions can be obtained by considering the four scenarios in the next sub-section.

C. Analysis of the Optimal Solution

Given any n, Problem 2 reduces to a subproblem with τ as the only decision variable. Since $n < \bar{n}$, the optimal n^* and τ^* can be obtained by solving $\bar{n}$ such subproblems. In the following, we analyze the optimal solution τ_n^* under any given n. We focus on that $\rho < 1$, while the case that $\rho \geq 1$ can be analyzed in the same way. For ease of analysis, we assume τ is continuous. Given n, the subproblem has following properties.

$$\frac{\partial \gamma}{\partial \tau} = (1 - p_1) \frac{\partial p_d}{\partial \tau}$$
$$= \frac{(1 - p_1) \sqrt{W}}{2\sqrt{2\pi\tau}} (\varepsilon_d - 1) e^{-\frac{(1-\varepsilon_d)^2}{2} W\tau}, \tag{25}$$

$$\frac{\partial \bar{\varphi}}{\partial \tau} = \frac{1}{n} \left(e_s + \frac{e_{tx}}{1+\rho} \left(\frac{\partial p_d}{\partial \tau} + \rho \frac{\partial p_f}{\partial \tau} \right) \right)$$
$$= \frac{1}{n} \left(e_s + \frac{e_{tx} \sqrt{W}}{2(1+\rho)\sqrt{2\pi\tau}} \Lambda_\tau \right), \tag{26}$$

$$\Lambda_\tau \triangleq (\varepsilon_d - 1) e^{-\frac{(1-\varepsilon_d)^2}{2} W\tau} - \rho(1 - \varepsilon_f) e^{-\frac{(1-\varepsilon_f)^2}{2} W\tau}. \tag{27}$$

Depending on the values of ε_d and ε_f (note that $\varepsilon_f < \varepsilon_d$), the shapes of the γ and $\bar{\varphi}$ curves are described as follows.

1) If either $\varepsilon_d \geq 1$ and $\varepsilon_f \geq 1$ or $\rho \leq \frac{\varepsilon_d - 1}{1 - \varepsilon_f} \leq 1$ and $\varepsilon_f < 1$, it is easy to see that $\frac{\partial \gamma}{\partial \tau} \geq 0$ and $\frac{\partial \bar{\varphi}}{\partial \tau} \geq 0$, which means that both γ and $\bar{\varphi}$ are increasing as τ increases. This corresponds to case 1 as shown in Fig. 3(a).

2) If $\frac{\varepsilon_d - 1}{1 - \varepsilon_f} > 1$ and $\varepsilon_f < 1$, since $e^{-\frac{(1-\varepsilon_d)^2}{2} W\tau} < e^{-\frac{(1-\varepsilon_f)^2}{2} W\tau}$, Λ_τ varies from positive infinite to a negative value and finally converges to 0. Depending on the parameters such as e_s and e_{tx}, the shape of $\bar{\varphi}$ will be in the form of either case 1 or case 2 as shown in Fig. 3(b).
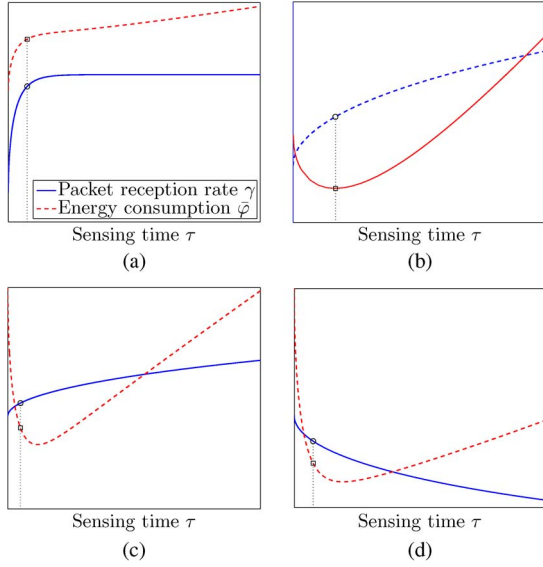
Fig. 3. Illustrations of the optimal τ under different ε_d and ε_f. (a) Case 1. (b) Case 2. (c) Case 3. (d) Case 4.

3) If $0 \leq \varepsilon_d - 1 < \rho(1 - \varepsilon_f)$, one can verify that $\frac{\partial^2}{\partial^2 \tau} \Lambda_\tau > 0$; hence, $\frac{\partial \bar{\varphi}}{\partial \tau}$ increases from negative infinite to a positive value. Therefore, as shown in Fig. 3(c), $\bar{\varphi}$ is a convex function.

4) Otherwise, $\varepsilon_d < 1$. Then, $\varepsilon_f < 1$ either. Consequently, $\frac{\partial \gamma}{\partial \tau} < 0$ and $\frac{\partial}{\partial \tau} \Lambda_\tau < 0$. As shown in Fig. 3(d), the objective function is convex.

As shown in the figure, in case 1, the optimal τ_n^* is the smaller one between $\bar{\tau}$ and the point where $\gamma = \underline{\gamma}(n)$. In the other cases, let $\tau_{n,\bar{\varphi}}$ and $\tau_{n,\gamma}$ be the solution points for $\frac{\partial \bar{\varphi}}{\partial \tau} = 0$ and $\gamma = \underline{\gamma}(n)$, respectively. In case 2, τ_n^* is among $\{0, \tau_{n,\bar{\varphi}}, \tau_{n,\gamma}, \bar{\tau}\}$. In the other cases, $\tau_n^* \in \{\tau_{n,\bar{\varphi}}, \tau_{n,\gamma}, \bar{\tau}\}$.

V. Optimization in Multi-Channel Case

If the sensor is able to sense multi-channels, say $\{CH_i | i \in C = \{1, 2, \ldots, m\}\}$, it may gain more opportunities to successfully deliver packets to the estimator, but at the cost of spending more sensing energy. In such a multi-channel case, the sensor needs to not only choose channels and determine the sensing time to be spent on each of them before transmitting a packet, but also decide the channel sensing order.

Suppose l channels will be sensed in each step. Denote $O_l = \{o_1, o_2, \ldots, o_l\} \subseteq C$ as the *channel sensing order* as described below. As shown in Fig. 4, in each step, the sensor starts sensing the channels one-by-one from CH_{o_1} to CH_{o_l}. If currently CH_{o_i} is found idle (with sensing time τ_{o_i} and detection probability p_{d,o_i}), the sensor will transmit packet over CH_{o_i}. Otherwise, it will change to sense the next channel. If no channel is found idle, the sensor will drop the packet. For simplicity, we assume that $\sum_{i=1}^{l} \bar{\tau}_{o_i} < T_s$. To save energy, we also assume that the sensor keeps in sleep mode if it is neither sensing a channel nor transmitting a packet.

A. The Optimization Problem

For any O_l and any sensing time $\tau = \{\tau_{o_1}, \tau_{o_2}, \ldots, \tau_{o_l}\}$, the probability that the sensor will transmit packet, i.e., at least one
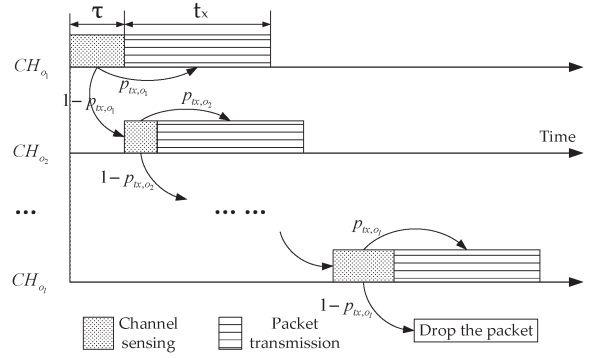

Fig. 4. Channel sensing order in one sampling period.

channel in O_l is found idle, can be given by

$$p_{tx}(O_l, \tau) = 1 - \prod_{i=1}^{l} (1 - p_{tx,o_i}), \qquad (28)$$

where p_{tx,o_i} is defined in (10). Let p_{c,o_i} be the probability of packet collision[3] if the sensor senses CH_{o_i}. By definition, p_{c,o_i} can be obtained by

$$\begin{aligned} p_{c,o_i} &= p_{B,o_i} p_{f,o_i} + p_{I,o_i} (1 - \eta_{o_i}) p_{d,o_i} \\ &= \frac{1}{1 + \rho_{o_i}} [\rho_{o_i} p_{f,o_i} + (1 - \eta_{o_i}) p_{d,o_i}]. \end{aligned} \qquad (29)$$

Then, the overall packet collision probability is

$$p_c(O_l, \tau) = p_{c,o_1} + \sum_{i=2}^{l} p_{c,o_i} \prod_{j=1}^{i-1} (1 - p_{tx,o_j}), \qquad (30)$$

and the successful packet transmission rate can be computed as

$$\begin{aligned} \gamma(O_l, \tau) &= p_{tx}(O_l) - p_c(O_l) \\ &= 1 - p_{c,o_1} - \sum_{i=2}^{l+1} p_{c,o_i} \prod_{j=1}^{i-1} (1 - p_{tx,o_j}) \\ &= \sum_{i=1}^{l} (p_{tx,o_i} - p_{c,o_i}) \prod_{j=1}^{i-1} (1 - p_{tx,o_j}), \end{aligned} \qquad (31)$$

where, for ease of presentation, we define $p_{c,o_{l+1}} = 1$ and $\prod_{j=1}^{0} (1 - p_{tx,o_j}) = 1$. Substituting (29) and (10) into the above equation, we get

$$\begin{aligned} \gamma(O_l, \tau) &= \sum_{i=1}^{l} \frac{\eta_{o_i} p_{d,o_i}}{1 + \rho_{o_i}} \prod_{j=1}^{i-1} \frac{1 - p_{d,o_i} + \rho_{o_i}(1 - p_{f,o_j})}{1 + \rho_{o_i}} \\ &\leq \sum_{i=1}^{l} \frac{p_{d,o_i}}{1 + \rho_{o_i}} \prod_{j=1}^{i-1} \frac{1 - p_{d,o_j} + \rho_{o_j}}{1 + \rho_{o_j}} \\ &= \sum_{i=1}^{l} \left[1 - \left(1 - \frac{p_{d,o_i}}{1 + \rho_{o_i}} \right) \right] \prod_{j=1}^{i-1} \left(1 - \frac{p_{d,o_j}}{1 + \rho_{o_j}} \right) \\ &= \sum_{i=1}^{l} \prod_{j=1}^{i-1} \left(1 - \frac{p_{d,o_j}}{1 + \rho_{o_j}} \right) - \sum_{i=1}^{l} \prod_{j=1}^{i} \left(1 - \frac{p_{d,o_j}}{1 + \rho_{o_j}} \right) \\ &= 1 - \prod_{i=1}^{l} \left(1 - \frac{p_{d,o_i}}{1 + \rho_{o_i}} \right) \\ &\leq 1 - \prod_{i=1}^{l} \frac{\rho_{o_i}}{1 + \rho_{o_i}}, \end{aligned}$$

[3] A collision occurs when the sensor transmits a packet during a busy period of channel CH_{o_i}.

where we have used the fact that $p_{f,o_i} \geq 0$ and $\eta_{o_i} \leq 1$ in obtaining the first inequality, and that $p_{d,o_i} \leq 1$ in the last inequality. According to Theorem 1, we must have $(1 - \gamma(O_l, \tau))\lambda_{\max}^{2n}(\mathbf{A}) < 1$. Combining the above inequality, we can obtain an upper bound of n as follows.

$$
n \leq \bar{n}_1(C_l) = \begin{cases} \left\lceil \frac{\Sigma_{i=1}^{l} \ln\left(1 + \frac{1}{\rho_{o_i}}\right)}{2\ln(\lambda_{\max}(\mathbf{A}))} \right\rceil - 1, & \text{if } \lambda_{\max}(\mathbf{A}) > 1, \\ \infty, & \text{otherwise,} \end{cases} \tag{32}
$$

where C_l is the index set of channels to be sensed. If given either O_l or τ, C_l can be easily determined. Moreover, similar to (23), we can obtain another bound on n as follows.

$$
n \leq \bar{n}_2(C_l) = \max\left\{ \tilde{n} \middle| \bar{\mathbf{Y}} \left(1 - \prod_{i=1}^{l} \frac{\rho_{o_i}}{1 + \rho_{o_i}}, \tilde{n}\right) \leq \bar{\mathbf{P}} \right\} < \infty. \tag{33}
$$

We assume the sensor uses the same channel sensing energy e_s and the same transmitting energy e_{tx} on these channels. The average energy consumption in the sensing step is

$$
\varphi(O_l, \tau) = \tau_{o_1} e_s + p_{tx,o_1} e_{tx} + (1 - p_{tx,o_1})(\tau_{o_2} e_s + p_{tx,o_2} e_{tx})
$$

$$
+ \ldots + \prod_{i=1}^{l-1}(1 - p_{tx,o_i})(\tau_{o_l} e_s + p_{tx,o_l} e_{tx})
$$

$$
= \sum_{i=1}^{l}(\tau_{o_i} e_s + p_{tx,o_i} e_{tx}) \prod_{j=1}^{i-1}(1 - p_{tx,o_j}). \tag{34}
$$

With the above formulation, Problem 2 can be updated as:

Problem 3: Find the optimal channel sensing schedule Θ_n^*, the optimal channel sensing order O_l^* and the optimal channel sensing time $\tau^* = \{\tau_{o_1}^*, \tau_{o_2}^*, \ldots, \tau_{o_l}^*\}$ to

$$
\begin{cases} \min_{n, O_l, \tau} & J = \frac{1}{n}\varphi(O_l, \tau) \\ s.t. & \gamma(O_l, \tau) \geq \underline{\gamma}(n), \\ & n \leq \bar{n}(C_l) = \min\{\bar{n}_1(C_l), \bar{n}_2(C_l)\}, \\ & 0 \leq \tau_{o_i} \leq \bar{\tau}_{o_i}, \quad i = 1, \ldots, l, \\ & l \leq m. \end{cases} \tag{35}
$$

Note that the set of all possible O_l is the power set of C, i.e., 2^C, where $|2^C| = 2^m$. In this sense, the above problem is a mixed integer program with both n and O_l chosen from two finite sets, respectively. For each given n and O_l, both the objective function and constraints of the above problem are continuous with respect to the decision variable τ. Moreover, since each τ_i is constrained within a finite but nonempty interval, by the extreme value theorem, the solution set of the global optimization of Problem 3 under given n and O_l is nonempty. Therefore, since the choices of n and O_l are finite, the solution set of the global optimization of Problem 3 is nonempty as well.

B. Algorithm Design

To address Problem 3, we apply two techniques: (1) a Lagrange multiplier is applied to the first constraint to transform Problem 3 into its dual, by which means searching the large power set 2^C is avoided; and (2) a simulated annealing (SA) based method is applied to search for the optimal τ where the

SA technique has the potential to escape from local optima and achieve a good approximation of the global optimal solution.

1) Dual Function: We can introduce a Lagrange multiplier $\mu \geq 0$ associated with constraint $\gamma(O_l, \tau) \geq \underline{\gamma}(n)$ and define the Lagrange dual function of Problem 3 as

$$
\mathcal{L}(\mu) = \min_{O_l, \tau, n} \left\{ \frac{1}{n}\varphi(O_l, \tau) - \mu\left(\gamma(O_l, \tau) - \underline{\gamma}(n)\right) \right\}
$$

$$
= \min_{\tau}\{\mathcal{L}(\mu, \tau)\}, \tag{36}
$$

where

$$
\mathcal{L}(\mu, \tau) \triangleq \min_{n < \bar{n}(C_l)} \left\{ \min_{O_l}\{\mathcal{M}(O_l, \mu, \tau, n)\} + \mu\underline{\gamma}(n) \right\},
$$

$$
\mathcal{M}(O_l, \mu, \tau, n) = \frac{1}{n}\varphi(O_l, \tau) - \mu\gamma(O_l, \tau)
$$

$$
= \sum_{i=1}^{l} \left[\frac{\tau_{o_i} e_s + p_{tx,o_i} e_{tx}}{n} - \mu(p_{tx,o_i} - p_{c,o_i})\right] \prod_{j=1}^{i-1}(1 - p_{tx,o_j})
$$

$$
= \sum_{i=1}^{l} \left[\frac{\tau_{o_i} e_s}{n} + \left(\frac{e_{tx}}{n} - \mu\right)p_{tx,o_i} + \mu p_{c,o_i}\right] \prod_{j=1}^{i-1}(1 - p_{tx,o_j}).
$$

Then, the original problem can be transformed into

$$
\max_{\mu} : \quad \mathcal{L}(\mu). \tag{37}
$$

We can adopt the following method to iteratively approach the optimal μ.

$$
\mu_{t+1} = \max\left\{0, \mu_t + \delta\frac{\partial}{\partial\mu_t}\mathcal{L}(\mu)|_{O_l^*, \tau^*, n^*}\right\}, \tag{38}
$$

where O_l^*, τ^* and n^* are the optimizers of (36) as determined in the next subsection. δ is the step size of each iteration. However, $\frac{\partial}{\partial\mu_t}\mathcal{L}(\mu)$ is difficult to obtain since O_l^*, τ^* and n^* are all dependent on μ. Instead, we resort to a discrete approximated approach to find $\frac{\partial}{\partial\mu_t}\mathcal{L}(\mu)$ by using a sufficiently small step size δ. In this way, μ is updated as follows.

$$
\mu_{t+1} = \max\left\{0, \mu_t + \delta\frac{\mathcal{L}(\mu_t) - \mathcal{L}(\mu_{t-1})}{\mu_t - \mu_{t-1}}\right\}. \tag{39}
$$

2) Obtaining $\mathcal{L}(\mu, \tau)$: Given any μ and τ, $\mathcal{L}(\mu, \tau)$ is calculated by iterating n from 1 to at most $\bar{n}(C_l)$ and selecting the optimal n that yields the minimum value. In each iteration, the optimal sensing order is obtained based on Theorem 2 as below. Then, $\mathcal{L}(\mu, \tau)$ is determined at the optimal n.

Theorem 2: For any $\mu > 0$ and any sensing time τ, there will be l channels (i.e., $\{CH_i | \tau_i \geq 0\}$) to be sensed and the optimal sensing order is

$$
O_l^* = \{o_1, o_2, \ldots, o_l | \forall i \in \{1, \ldots, l-1\}, \tau_{o_i} \geq 0,
$$

$$
\mathcal{H}(o_i, \tau_{o_i}) \leq \mathcal{H}(o_{i+1}, \tau_{o_{i+1}})\}, \tag{40}
$$

where $\mathcal{H}(o_i, \tau_{o_i}) \triangleq \left(\frac{e_s}{n}\tau_{o_i} + \mu p_{c,o_i}\right)/p_{tx,o_i}$.

Proof: Consider another sensing order $\tilde{O}_l = \{\tilde{o}_1, \tilde{o}_2, \ldots, \tilde{o}_l\}$ which is a reordering of O_l. For any $i \in \{1, \ldots, l-1\}$, if we

swap $\tilde{o}_i$ and $\tilde{o}_{i+1}$ in $\tilde{O}_l$ and let $\breve{O}_l$ be the new sensing order, then the difference in $\mathcal{M}$ is

$$\mathcal{M}(\breve{O}_l, \mu, \tau, n) \leq \mathcal{M}(\tilde{O}_l, \mu, \tau, n)$$

$$\Leftrightarrow \frac{1}{n}\varphi(\breve{O}_l, \tau) - \mu\gamma(\breve{O}_l, \tau) \leq \frac{1}{n}\varphi(\tilde{O}_l, \tau) - \mu\gamma(\tilde{O}_l, \tau)$$

$$\Leftrightarrow \frac{\tau_{\tilde{o}_{i+1}} e_s + p_{tx,\tilde{o}_{i+1}} e_{tx}}{n} - \mu(p_{tx,\tilde{o}_{i+1}} - p_{c,\tilde{o}_{i+1}})$$

$$+ \left[\frac{\tau_{\tilde{o}_i} e_s + p_{tx,\tilde{o}_i} e_{tx}}{n} - \mu(p_{tx,\tilde{o}_i} - p_{c,\tilde{o}_i})\right](1 - p_{tx,\tilde{o}_{i+1}})$$

$$\leq \frac{\tau_{\tilde{o}_i} e_s + p_{tx,\tilde{o}_i} e_{tx}}{n} - \mu(p_{tx,\tilde{o}_i} - p_{c,\tilde{o}_i}) + (1 - p_{tx,\tilde{o}_i})$$

$$\times \left[\frac{\tau_{\tilde{o}_{i+1}} e_s + p_{tx,\tilde{o}_{i+1}} e_{tx}}{n} - \mu(p_{tx,\tilde{o}_{i+1}} - p_{c,\tilde{o}_{i+1}})\right]$$

$$\Leftrightarrow \left(\frac{e_s}{n}\tau_{\tilde{o}_{i+1}} + \mu p_{c,\tilde{o}_{i+1}}\right) p_{tx,\tilde{o}_i} \leq \left(\frac{e_s}{n}\tau_{\tilde{o}_i} + \mu p_{c,\tilde{o}_i}\right) p_{tx,\tilde{o}_{i+1}}$$

$$\Leftrightarrow \mathcal{H}(\tilde{o}_{i+1}, \tau_{\tilde{o}_{i+1}}) \leq \mathcal{H}(\tilde{o}_i, \tau_{\tilde{o}_i}). \tag{41}$$

(41) implies that the swapping will reduce the objective $\mathcal{M}(\tilde{O}_l, \mu, \tau, n)$. In other words, if we deem $\mathcal{H}(o_i, \tau_{o_i})$ as a measure of the channel quality, (41) suggests that, for any pair of channels, the one with better quality should be sensed first. Continuing applying the "better channel first" rule, we will finally get the optimal sensing order as in (40). ∎

3) Optimal Spectrum Sensing Algorithm (OSSA): We propose a simulated annealing based optimal spectrum sensing algorithm (OSSA) to randomly search the sensing time τ. To avoid searching across an uncountable space constructed by the continuous intervals $\{[0, \bar{\tau}_i]\}$, we assume that each τ_i can only be chosen from a finite discrete set. Although such an assumption may affect optimality of the obtained results, the solutions are more practically implementable since the discrete sensing time is compatible with the discrete sampling rate used for channel sensing as in the energy detection technique as presented in Appendix A. In fact, the sensor applies a certain sampling rate (depending on the channel bandwidth W_i) to sense each channel. That is, the sensing time τ_i can be viewed as a discrete value with the unit time depending on W_i. Without loss of generality, let $\tau_i \in \{0, 1, \ldots, \bar{\tau}_i\}$ be the number of the time units spent on sensing channel CH_i. In particular, we define $\tau_i = 0$ if CH_i will not be sensed and we do not allow blind transmission on an un-sensed channel.

The pseudo codes of the proposed algorithm are shown in Algorithm 1. Given μ_t, Algorithm 1 contains $K_{\max}$ annealing processes and each annealing process has $\kappa_{\max}$ iterations. It optimizes τ by letting each τ_i walk randomly with caution on the set $\{0, 1, \ldots, \bar{\tau}_i\}$. By "caution" we mean that each move should be meaningful. Specifically, in each walk step (iteration), a tentative sensing time $\hat{\tau}$ is generated randomly, and the consequent value of $\mathcal{L}(\mu_t, \hat{\tau})$ is evaluated. If $\hat{\tau}$ is better than the previous τ, i.e., $\Delta\mathcal{L} = \mathcal{L}(\mu_t, \hat{\tau}) - \mathcal{L}(\mu_t, \tau) \leq 0$, $\hat{\tau}$ will be accepted (i.e., τ will move to $\hat{\tau}$). Otherwise, $\hat{\tau}$ will be accepted with a probability which depends on both $\Delta\mathcal{L}$ and an annealing

temperature, not to have $\mathcal{L}(\mu_t, \tau)$ be trapped at local optima. By gradually decreasing the annealing temperature by a factor $\zeta > 0$ after the completion of each annealing process, $\mathcal{L}(\mu_t, \tau)$ will finally converge to one of the global optima. Finally, we iteratively update μ_t until it converges, meaning that the dual problem is solved and hence the original optimization is also solved.

Algorithm 1: OSSA: optimal spectrum sensing algorithm.

1 **Initial settings**:
   $t = 1, \mu_0 > 0, \Delta > 0, K_{\max} > 0, \kappa_{\max} > 0$ and $\zeta > 0$.
2 **while** $|\mu_t - \mu_{t-1}| > \Delta$ **do**
3   choose initial sensing time τ and annealing temperature $Temp_1$ and set $k = 1$;
4   calculate $\mathcal{L}(\mu_t, \tau)$ (refer to line 10-12);
5   **while** $k \leq K_{\max}$ **do**
6   **while** $\kappa \leq \kappa_{\max}$ **do**
7   randomly generate $\Delta\tau = [\Delta\tau_1, \ldots, \Delta\tau_m]^T$ with $\Delta\tau_i \in \{-1, 0, 1\}, i = 1, \ldots, m$, and $\Delta\tau \neq 0$;
8         $\hat{\tau}_i \leftarrow \max\{-1, \min\{\tau_i + \Delta\tau_i, \bar{\tau}_i\}\}, i = 1, \ldots, m$;
9         $\hat{\tau} \leftarrow [\hat{\tau}_1, \hat{\tau}_2, \ldots, \hat{\tau}_m]^T$;
10   **foreach** $n < \bar{n}(C_l)$ **do** determine $\mathcal{O}_l^*$ based on $\hat{\tau}$ according to Theorem 2;
11        $n^* \leftarrow \underset{n < \bar{n}(C_l)}{\text{argmin}}\{\mathcal{M}(\mathcal{O}_l^*, \mu_t, \hat{\tau}, n) + \mu_t\underline{\gamma}(n)\}$;
12        $\mathcal{L}(\mu_t, \hat{\tau}) \leftarrow \mathcal{M}(\mathcal{O}_l^*, \mu_t, \hat{\tau}, n^*) + \mu_t\underline{\gamma}(n^*)$;
13        $\Delta\mathcal{L} \leftarrow \mathcal{L}(\mu_t, \hat{\tau}) - \mathcal{L}(\mu_t, \tau)$;
14   **if** $\Delta\mathcal{L} \leq 0$ **then** $\tau \leftarrow \hat{\tau}$;
15   **else**
16          $\xi \leftarrow random[0, 1], \epsilon \leftarrow \exp(-\frac{\Delta\mathcal{L}}{Temp_k})$;
17   **if** $\xi < \epsilon$ **then** $\tau \leftarrow \hat{\tau}$;
18   **end**
19        $\kappa \leftarrow \kappa + 1$;
20   **end**
21       $k \leftarrow k + 1, Temp_{k+1} \leftarrow \zeta Temp_k$;
22   **end**
23     $\mathcal{L}(\mu_t) \leftarrow \mathcal{L}(\mu_t, \tau)$;
24   update μ_t based on (38);
25     $t \leftarrow t + 1$;
26 **end**

Remark 2: OSSA is derived from the standard simulated annealing algorithm; however, our major novelty lies in that, instead of randomly generating all the decision variables (i.e., channel sensing order O_l, sensing time τ and n which characterizes the sensing schedule), we only generate the sensing time at each iteration. Once τ is given, the optimal O_l is obtained based on Theorem 2, while the optimal n is found by searching in a finite set. Thus, at each iteration, by only randomly generating a $\hat{\tau}$, we can evaluate the objective function $\mathcal{L}(\mu_t, \hat{\tau})$ and determine whether this $\hat{\tau}$ will be accepted or not based on the improvement $\Delta\mathcal{L}$.

Remark 3 (Computation Complexity Analysis): In each iteration t of OSSA, for each pair of (k, κ) (there are $K_{\max}\kappa_{\max}$ of them), the computation complexity is $O(\bar{n}(C)m\log m)$.

Specifically, generating $\hat{\tau}$ has complexity $O(m)$; for each n, determining O_l^* has complexity $O(m \log m)$ and calculating $\mathcal{M}(O_l^*, \mu_t, \hat{\tau}, n)$ has complexity $O(m)$; thus obtaining n^* and consequently determining $\Delta \mathcal{L}$ have complexity $O(\bar{n}(C) m \log m)$. Therefore, the complexity in each iteration is in the order $O(K_{\max} \kappa_{\max} \bar{n}(C) m \log m)$.

Remark 4: OSSA is an off-line algorithm, which can be executed at either the sensor or some computing center, e.g., the estimator. The output results are fixed over time so that the algorithm only needs to run once. Therefore, a possible implementation can be as follows: the estimator runs the algorithm and finds the solution off-line. It sends the solution to the sensor which operates according to the solution from then on.

VI. SIMULATION RESULTS

In this section, we conduct simulations to demonstrate the effectiveness of the proposed method. We consider a linear system (1) with $\mathbf{A} = \begin{bmatrix} 1.05 & 0 \\ 1 & 0.9 \end{bmatrix}$, $\mathbf{C} = \mathbf{I}$, $\mathbf{Q} = \mathbf{I}$ and $\mathbf{R} = 0.8\mathbf{I}$, where $\mathbf{I}$ is the 2-by-2 identity matrix. The sensor samples the system every $T_s = 1$ second and the transmission time of each measurement packet is $t_x = 50$ ms. The estimation performance requirement is set as $\bar{\mathbf{P}} = \bar{\mathbf{Y}}(0.7, 6)$, where $\bar{\mathbf{Y}}(\gamma, n)$ is defined in Lemma 1.

A. Single-Channel Scenario

For the channel under consideration, we assume its bandwidth $W = 2$ MHz and noise power $\sigma_n = 1$ mW. The signal-to-noise ratio of the signal received by the sensor during the channel's busy periods is -3 dB. We adopt the Poisson channel model as described in Section III-B1 with the average channel busy and idle rates $\alpha = 5$ and $\beta = 30$, respectively. The sensing parameter $\varepsilon_d = 1.2$, the maximum sensing time is $\bar{\tau} = 250$ units with the unit channel sensing time as 0.1 ms. The per-second energy cost of sensing and transmitting is $e_s = e_{tx} = 100$ unit.

Fig. 5(a) shows an example trace of the estimation error covariance along time, where the sensor conducts sensing every step (i.e., $n = 1$) and the sensing time on the channel is fixed at 1 ms. The curve of $\mathbb{E}[\mathbf{P}_k]$ is obtained by averaging the results of 2000 independent simulation runs. As discussed before and also shown in this figure, $\mathbb{E}[\mathbf{P}_k]$ does not converge. However, the figure indicates that the upper bound curve $\mathbf{Y}_k$ is a good approximation of the long-term average of $\mathbb{E}[\mathbf{P}_k]$, which justifies the approximation method in Section IV-B.

The optimal solutions of Problem 2 are depicted in Fig. 5(b), where we vary the channel idle probability p_I by gradually increasing β. The results show that, under a certain n, the optimal sensing time τ^* drops quickly as the idle probability increases, which results in the decrease of the average energy consumption $\bar{\varphi}$. In fact, as the channel quality becomes better, less sensor energy will be wasted for conducting sensing and transmitting during the channel's busy periods. Meanwhile, when p_I increases from 0.3 to 1, the optimal n increases piecewise, which means that the sensor conducts spectrum sensing and packet transmission less frequently. Therefore, generally speaking, the energy consumption decreases as p_I increases.
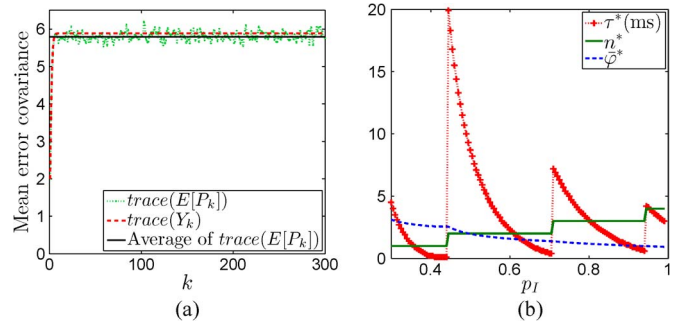

Fig. 5. Performance in the single-channel scenario. (a) Estimation performance over time. (b) Optimal sensing strategy.
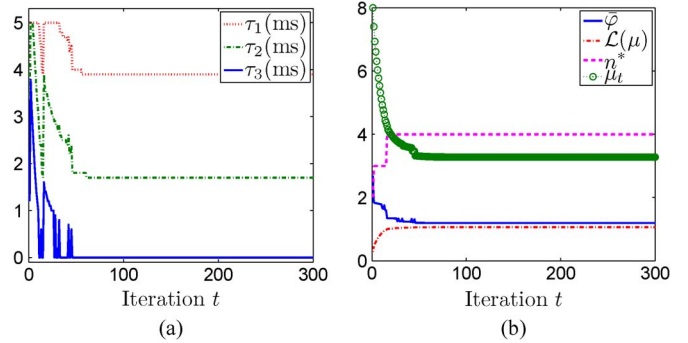

Fig. 6. Performance of the proposed algorithm in the multi-channel scenario. (a) Sensing time on each channel. (b) Some important variables.

B. Multi-Channel Scenario

In this scenario, the sensor data can be transmitted to the remote estimator via one of $m = 3$ orthogonal channels. The average busy and idle rates for channel CH_i are $\alpha_i = 5i$ and $\beta_i = 30$, respectively, where $i = 1, 2, 3$. All the other parameters of these channels are set the same as those in the single-channel scenario above. To run the proposed OSSA algorithm, we set $K_{\max} = 200$, $\kappa_{\max} = 20$ and $\zeta = 0.94$ [38].

Fig. 6 demonstrates the performance of the proposed algorithm OSSA. First, the two figures confirm the convergence of our algorithm. As μ_t changes according to (39), $\mathcal{L}(\mu_t)$ increases as expected from (37). Consequently, the objective function J decreases and finally converges to a steady value as the optimization process evolves. Similar behavior can be observed in Fig. 6(a). Our simulation results indicate that the optimal channel scanning strategy is always in the order of CH_1, CH_2 and then CH_3, which is the same as the decreasing order of the channel quality measured by the idle probability p_I (we have $p_{I,1} = 85.7\%$, $p_{I,2} = 75\%$ and $p_{I,3} = 66.7\%$). In fact, as can be seen from the two figures, τ_i is relatively small, which makes the term $\mathcal{H}$ in (40) approximately become $\mu \frac{p_{c,i}}{p_{tx,i}}$. Based on Theorem 2, the optimal order is then roughly determined by $\frac{p_{c,i}}{p_{tx,i}}$, which can be viewed as the conditional probability that a transmit attempt results in a collision. In this case, a channel with better quality will have lower conditional collision probability and subsequently lower value of $\mathcal{H}$, and hence should be sensed first according to Theorem 2.

The optimal solutions with the proposed algorithm are demonstrated in Figs. 7 and 8. In Fig. 7, we compare the
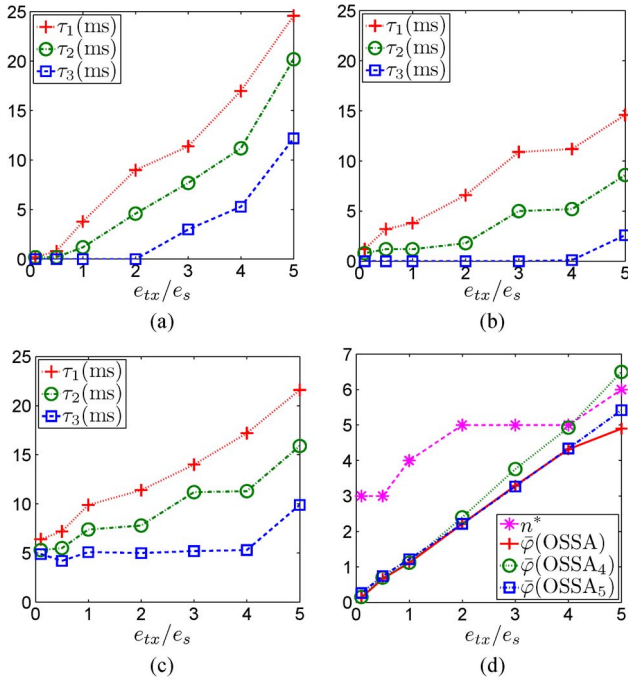
Fig. 7. Optimal solutions by OSSA with respect to e_{tx}/e_s where $e_s = 100$. (a) Sensing time under OSSA. (b) Sensing time under OSSA$_4$. (c) Sensing time under OSSA$_5$. (d) Performance comparison.
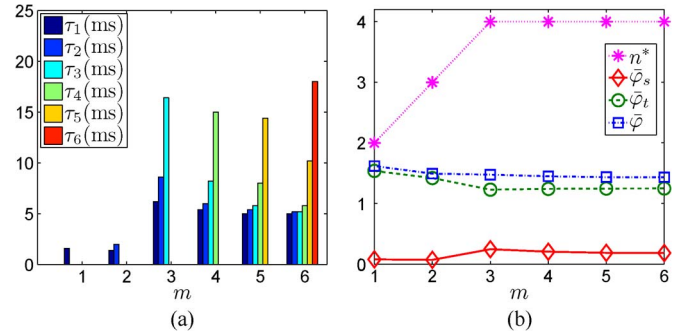

Fig. 8. Solutions by OSSA with respect to number of available channels. $\bar{\varphi}_{tx}$ and $\bar{\varphi}_s$ are the average transmission and sensing energy expenditure, respectively. (a) Sensing time. (b) Energy consumption.
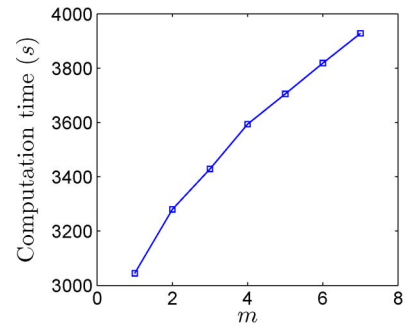

Fig. 9. Algorithm running time.

performance of OSSA with its two variants OSSA$_4$ and OSSA$_5$ (which runs OSSA under a fixed n. $n = 4$ and $n = 5$ for OSSA$_4$ and OSSA$_5$, respectively). Firstly, the comparison in Fig. 7(d) shows that OSSA$_4$ (or OOSSA$_5$) achieves almost the same performance as OSSA if the optimal n obtained by OSSA is 4 (or 5); otherwise, OSSA outperforms the two variants, which to some extent justifies the optimality of OSSA. Secondly, as e_{tx}/e_s becomes larger, i.e., the transmission energy becomes to dominate the total energy $\bar{\varphi}$, the sensor intends to transmit data less frequently (i.e, to use larger n) to save energy. This requires each transmission to be more reliable to still satisfy the estimation performance constraint. The reason is that, based on Lemma 1, an increase of n requires to increase γ to maintain $\bar{\mathbf{Y}}$. As a result, increasing γ means the amount of energy waste due to collisions can be reduced, which saves the total energy expenditure. Therefore, under all these three algorithms, when e_{tx}/e_s increases, the sensor will use a larger n and spend more sensing time to increase the sensing accuracy, which are clearly demonstrated in Fig. 7.

We show by Fig. 8 that using more channels can improve the system performance. Here we assume different settings of the channels: the average busy and idle rates for CH_i are $\alpha_i = 5 + i$ and $\beta_i = 30$, respectively, where $i = 1, \ldots, 6$. As the number of available channels increases, the sensor tends to use more channels, and the optimal sensing order always follows the order of $CH_1, CH_2, \ldots, CH_m$. As explained before, since n^* increases from 2 to 4 as m changes from 1 to 3, the sensor will spend more sensing time on each channel, as shown in Fig. 8(a). When n^* stays at 4 and $m \geq 3$, the sensors try to save both sensing and transmission energy by appropriately allocating sensing time on each channel. Specifically, as m increases by 1, the old channels will be sensed with slightly shorter time to save sensing energy, while the new one with lowest quality

will be sensed a long time to improve the sensing accuracy on it and thus save transmission energy. From Fig. 8(b), we observe that the average energy consumption of the sensor decreases as m increases; however, the amount of decrease becomes unnoticeable as $m \geq 4$, which suggests that we only need a limited number of channels to (roughly) achieve the optimal goal.

The running time of the proposed algorithm is characterized by many parameters such as the number of channels m, the number of annealing processes (i.e., K_{max}) and the length of each annealing process (i.e., κ_{max}). As shown in Fig. 9, with fixed K_{max} and κ_{max} and when the number of channels m is small, the running time is roughly linear to m. For a large m, as discussed in Remark 3, the computation complexity of each iteration of the proposed algorithm is $O(m \log m)$. If we fix the number of iterations, the total running time of the proposed algorithm will also be $O(m \log m)$.

VII. CONCLUSION

We have studied the energy-efficient spectrum sensing problem for remote state estimation over multiple wireless channels. We formulated it as a mixed integer nonlinear program and proposed a simulated annealing based optimization algorithm OSSA which jointly addresses the problems such as when to perform sensing, which channels to sense, in what order and how long to scan each channel. Simulation results show that OSSA can achieve the optimal goal and balance between sensing energy and transmission energy expenditure. Specifically, as the transmission energy becomes dominant over the sensing energy, the sensor will conduct sensing less frequently

but use longer sensing time to increase the sensing accuracy, reduce energy waste due to collisions and hence save total transmission energy. We also demonstrated that the overall performance is improved by introducing more channels, though the improvement becomes less significant as the number of channels continues to increase.

APPENDIX

A. Energy Detection

By integrating the received signal from a channel in a pre-defined bandwidth W over the sensing period τ, the sensor gets the detected energy, denoted by E_d, and compare it with a pre-defined threshold, say E_{th}, to decide whether the channel is currently busy or not. In this case, the sensing result can be denoted as

$$s_c = \begin{cases} 1, & \text{if } E_d > E_{th}, \\ 0, & \text{otherwise.} \end{cases}$$

When the channel is idle, the detected energy follows a Gaussian distribution as $E_d \sim \mathcal{N}(N\sigma_n^2, N\sigma_n^4)$, where N is the number of samples. σ_n^2 and σ_s^2 are the variances of the channel noise and received signal from the channel, respectively. According to Nyquist sampling theorem, the sampling rate should be at least $2W$, where W is the bandwidth of the channel. Thus, we just assume the sampling rate be $2W$ and that $N = 2\tau W$. Reasonably and for simplicity, we also assume that $E_{th} = \varepsilon_d N\sigma_n^2 = 2\varepsilon_d \tau W \sigma_n^2$, where $\varepsilon_d > 0$. Notice that, since E_d is random, the outcome s_c may be different from the true channel state. The correct detection probability p_d can be obtained straightforward as follows:

$$\begin{aligned} p_d &= \mathbb{P}[E_d < E_{th}|\text{channel is idle}] \\ &= 1 - \mathbb{P}[E_d > E_{th}|\text{channel is idle}] \\ &= 1 - Q\left(\frac{E_{th} - 2\sigma_n^2\tau W}{2\sigma_n^2\sqrt{\tau W}}\right) \\ &= Q\left((1 - \varepsilon_d)\sqrt{\tau W}\right). \end{aligned}$$

Similarly, $E_d \sim \mathcal{N}(2\tau W(\sigma_s^2 + \sigma_n^2), 2\tau W(\sigma_s^2 + \sigma_n^2)^2)$ when the channel is busy. We can obtain p_f as in (9), where $\varepsilon_f = \varepsilon_d / (1 + \frac{\sigma_s^2}{\sigma_n^2})$.

B. Proof of Theorem 1

Consider the schedule Θ_n. According to (17), we have

$$\mathbf{P}_{k_i-1} = \mathbf{A}\mathbf{P}_{k_i-2}\mathbf{A}^T + \mathbf{Q} = \dots$$

$$= \mathbf{A}^{n-1}\mathbf{P}_{k_i-1}\left(\mathbf{A}^T\right)^{n-1} + \sum_{t=0}^{n-2}\mathbf{A}^t\mathbf{Q}\left(\mathbf{A}^T\right)^t. \quad (42)$$

Substituting the above equation into (16) and taking expectation at both sides yield that

$$\mathbb{E}[\mathbf{P}_{k_i}] = (1-\gamma)\left[\mathbf{A}^n\mathbb{E}[\mathbf{P}_{k_i-1}]\left(\mathbf{A}^T\right)^n + \sum_{t=1}^{n-1}\mathbf{A}^t\mathbf{Q}\left(\mathbf{A}^T\right)^t\right]$$

$$+ \mathbf{Q} + \gamma\mathbf{A}\mathbb{E}[\Upsilon_{k_i-1}]\mathbf{A}^T. \quad (43)$$

Since n is finite, according to (42), we can easily verify that the stability of $\{\mathbb{E}[\mathbf{P}_k]\}$ is equivalent to that of $\{\mathbb{E}[\mathbf{P}_{k_i}]\}$. Thus, we only need to prove the stability of $\{\mathbb{E}[\mathbf{P}_{k_i}]\}$ where $\{k_i\}$ indicates the channel sensing steps. Then, according to (43), due to the finiteness of n and Υ_{k_i-1}, there must exist a matrix $\bar{M} \geq \mathbf{Q}$ such that

$$(1-\gamma)\mathbf{A}^n\mathbb{E}[\mathbf{P}_{k_i-1}](\mathbf{A}^T)^n + \mathbf{Q}$$

$$\leq \mathbb{E}[\mathbf{P}_{k_i}] \leq (1-\gamma)\mathbf{A}^n\mathbb{E}[\mathbf{P}_{k_i-1}](\mathbf{A}^T)^n + \bar{M}.$$

Based on the stability of Lyapunov functions, we know that (18) is both sufficient and necessary for the stability of both the lower and upper bound sequences above. Therefore, (18) is both sufficient and necessary for the stability of $\{\mathbb{E}[\mathbf{P}_k]\}$, which completes the proof of this theorem.

C. Proof of Lemma 1

Consider the subset $\{\mathbf{Y}_{k_i}|\theta_{k_i} = 1\}$. Suppose $\mathbf{Y}_{k_i} = g(\mathbf{Y}_{k_i-1})$. Similar to (43), we have

$$g\left(\mathbf{Y}_{k_i-1}\right) = (1-\gamma)\left[\mathbf{A}^n\mathbf{Y}_{k_i-1}\left(\mathbf{A}^T\right)^n + \sum_{t=1}^{n-1}\mathbf{A}^t\mathbf{Q}\left(\mathbf{A}^T\right)^t\right]$$

$$+ \mathbf{Q} + \gamma\mathbf{A}\tilde{\Upsilon}_{k_i-1}\mathbf{A}^T, \quad (44)$$

$$\tilde{\Upsilon}_{k_i-1}^{-1} = \left[\mathbf{A}^{n-1}\mathbf{Y}_{k_i-1}\left(\mathbf{A}^T\right)^{n-1} + \sum_{t=0}^{n-2}\mathbf{A}^t\mathbf{Q}\left(\mathbf{A}^T\right)^t\right]^{-1}$$

$$+ \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}.$$

Since $\tilde{\Upsilon}_{k_i-1}$ is bounded, the sequence $\{\mathbf{Y}_{k_i}\}$ is stable if and only if condition (18) is satisfied. Below we show that it converges to a unique value.

First, (44) indicates that $g(\cdot)$ is an increasing function. Suppose the initial value is $\mathbf{Y}_0 = \mathbf{P}_0 = 0$. By induction, we can see that $\{\mathbf{Y}_{k_i}\}$ is increasing and finally reaches a steady value, say $\bar{\mathbf{Y}}_0$, since the sequence has already been proved stable. In other words, $\bar{\mathbf{Y}}_0$ is the fix point of the iteration (44), i.e., $\bar{\mathbf{Y}}_0 = g(\bar{\mathbf{Y}}_0)$.

Next, $\forall \mathbf{Y}_0 > \bar{\mathbf{Y}}_0$, since $g(\cdot)$ is increasing, by induction, we always have $\mathbf{Y}_{k_i} > \bar{\mathbf{Y}}_0$. Let us introduce a useful inequality as below. Let $\tilde{\mathbf{R}} = \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}, \tilde{\mathbf{Q}} = \sum_{t=0}^{n-2}\mathbf{A}^t\mathbf{Q}(\mathbf{A}^T)^t, \mathbf{Z}_1, \mathbf{Z}_2$ be four positive definite matrices of compatible dimensions with $\mathbf{Z}_1 > \mathbf{Z}_2$. We have

$$\left[\tilde{\mathbf{R}} + (\mathbf{Z}_1 + \tilde{\mathbf{Q}})^{-1}\right]^{-1} - \left[\tilde{\mathbf{R}} + (\mathbf{Z}_2 + \tilde{\mathbf{Q}})^{-1}\right]^{-1}$$

$$= \left\{\left[\tilde{\mathbf{R}} + (\mathbf{Z}_1 + \tilde{\mathbf{Q}})^{-1}\right]^{-1}\left[\tilde{\mathbf{R}} + (\mathbf{Z}_2 + \tilde{\mathbf{Q}})^{-1}\right] - \mathbf{I}\right\}$$

$$\times \left[\tilde{\mathbf{R}} + (\mathbf{Z}_2 + \tilde{\mathbf{Q}})^{-1}\right]^{-1}$$

$$= \left[\tilde{\mathbf{R}} + (\mathbf{Z}_1 + \tilde{\mathbf{Q}})^{-1}\right]^{-1}\left[(\mathbf{Z}_2 + \tilde{\mathbf{Q}})^{-1} - (\mathbf{Z}_1 + \tilde{\mathbf{Q}})^{-1}\right]$$

$$\times \left[\tilde{\mathbf{R}} + (\mathbf{Z}_2 + \tilde{\mathbf{Q}})^{-1}\right]^{-1}$$

$$= \left[\tilde{\mathbf{R}} + (\mathbf{Z}_1 + \tilde{\mathbf{Q}})^{-1}\right]^{-1}(\mathbf{Z}_1 + \tilde{\mathbf{Q}})^{-1}\left[(\mathbf{Z}_1 + \tilde{\mathbf{Q}}) - (\mathbf{Z}_2 + \tilde{\mathbf{Q}})\right]$$

$$\times (\mathbf{Z}_2 + \tilde{\mathbf{Q}})^{-1}\left[\tilde{\mathbf{R}} + (\mathbf{Z}_2 + \tilde{\mathbf{Q}})^{-1}\right]^{-1}$$

$$= \left[(\mathbf{Z}_1 + \tilde{\mathbf{Q}})\tilde{\mathbf{R}} + \mathbf{I} \right]^{-1} (\mathbf{Z}_1 - \mathbf{Z}_2) \left[\tilde{\mathbf{R}}(\mathbf{Z}_2 + \tilde{\mathbf{Q}}) + \mathbf{I} \right]^{-1}$$

$$< \left[(\mathbf{Z}_2 + \tilde{\mathbf{Q}})\tilde{\mathbf{R}} + \mathbf{I} \right]^{-1} (\mathbf{Z}_1 - \mathbf{Z}_2) \left[\tilde{\mathbf{R}}(\mathbf{Z}_2 + \tilde{\mathbf{Q}}) + \mathbf{I} \right]^{-1}, \quad (45)$$

where $\mathbf{I}$ represents identity matrices of compatible dimensions. Define $\tilde{\mathbf{F}} = \left[\left(\mathbf{A}^{n-1}\bar{\mathbf{Y}}_0(\mathbf{A}^T)^{n-1} + \tilde{\mathbf{Q}} \right) \tilde{\mathbf{R}} + \mathbf{I} \right]^{-1} \mathbf{A}^{n-1}$. Based on the above inequality and the fact that $\mathbf{Y}_{k_i} > \bar{\mathbf{Y}}_0$, we get

$$\mathbf{Y}_{k_{i+1}} - \bar{\mathbf{Y}}_0 = g(\mathbf{Y}_{k_i}) - g(\bar{\mathbf{Y}}_0)$$

$$= (1-\gamma)\mathbf{A}^n(\mathbf{Y}_{k_i} - \bar{\mathbf{Y}}_0)\left(\mathbf{A}^T\right)^n + \gamma\mathbf{A}\left[\tilde{\Upsilon}_{k_i} - \tilde{\Upsilon}(\bar{\mathbf{Y}}_0)\right]\mathbf{A}^T$$

$$< (1-\gamma)\mathbf{A}^n(\mathbf{Y}_{k_i} - \bar{\mathbf{Y}}_0)\left(\mathbf{A}^T\right)^n + \gamma\mathbf{A}\tilde{\mathbf{F}}(\mathbf{Y}_{k_i} - \bar{\mathbf{Y}}_0)\tilde{\mathbf{F}}^T\mathbf{A}^T$$

$$\triangleq \mathcal{F}(\mathbf{Y}_{k_i} - \bar{\mathbf{Y}}_0).$$

where the above inequality can be verified by substituting $\mathbf{Z}_1 = \mathbf{A}^{n-1}\mathbf{Y}_{k_{i-1}}$ and $\mathbf{Z}_2 = \mathbf{A}^{n-1}\bar{\mathbf{Y}}_0(\mathbf{A}^T)^{n-1}$ into (45). With the newly defined matrix $\tilde{\mathbf{F}}$, we can obtain another property of $\bar{\mathbf{Y}}_0$ as below.

$$\tilde{\mathbf{F}}\bar{\mathbf{Y}}_0\tilde{\mathbf{F}}^T = \left[\left(\mathbf{A}^{n-1}\bar{\mathbf{Y}}_0\left(\mathbf{A}^T\right)^{n-1} + \tilde{\mathbf{Q}} \right)\tilde{\mathbf{R}} + \mathbf{I} \right]^{-1} \mathbf{A}^{n-1}\bar{\mathbf{Y}}_0$$

$$\times \left(\mathbf{A}^T\right)^{n-1} \left[\tilde{\mathbf{R}}\left(\mathbf{A}^{n-1}\bar{\mathbf{Y}}_0\left(\mathbf{A}^T\right)^{n-1} + \tilde{\mathbf{Q}} \right) + \mathbf{I} \right]^{-1}$$

$$= \tilde{\Upsilon}(\bar{\mathbf{Y}}_0) \left[\mathbf{A}^{n-1}\bar{\mathbf{Y}}_0\left(\mathbf{A}^T\right)^{n-1} + \tilde{\mathbf{Q}} \right]^{-1} \mathbf{A}^{n-1}\bar{\mathbf{Y}}_0$$

$$\times \left(\mathbf{A}^T\right)^{n-1} \left[\mathbf{A}^{n-1}\bar{\mathbf{Y}}_0\left(\mathbf{A}^T\right)^{n-1} + \tilde{\mathbf{Q}} \right]^{-1} \tilde{\Upsilon}(\bar{\mathbf{Y}}_0)$$

$$< \tilde{\Upsilon}(\bar{\mathbf{Y}}_0) \left[\mathbf{A}^{n-1}\bar{\mathbf{Y}}_0\left(\mathbf{A}^T\right)^{n-1} + \tilde{\mathbf{Q}} \right]^{-1} \tilde{\Upsilon}(\bar{\mathbf{Y}}_0)$$

$$< \tilde{\Upsilon}(\bar{\mathbf{Y}}_0) \left\{ \tilde{\mathbf{R}} + \left[\mathbf{A}^{n-1}\bar{\mathbf{Y}}_0\left(\mathbf{A}^T\right)^{n-1} + \tilde{\mathbf{Q}} \right]^{-1} \right\} \tilde{\Upsilon}(\bar{\mathbf{Y}}_0)$$

$$= \tilde{\Upsilon}(\bar{\mathbf{Y}}_0).$$

Then, based on (44), it is easy to see that $\bar{\mathbf{Y}}_0 = g(\bar{\mathbf{Y}}_0) > \mathcal{F}(\bar{\mathbf{Y}}_0)$. According to Lemma 3 in [5], the iteration $\mathbf{Y}_{k_{i+1}} - \bar{\mathbf{Y}}_0 = \mathcal{F}(\mathbf{Y}_{k_i} - \bar{\mathbf{Y}}_0)$ converges to 0, which means that $\{\mathbf{Y}_{k_i}\}$ converges to $\bar{\mathbf{Y}}_0$.

For any $\mathbf{Y}_0 < \bar{\mathbf{Y}}_0$, $\mathbf{Y}_{k_i}$ is always bounded within $[\underline{\mathbf{Y}}_{k_i}, \bar{\mathbf{Y}}_{k_i}]$, where $\{\underline{\mathbf{Y}}_{k_i}\}$ and $\{\bar{\mathbf{Y}}_{k_i}\}$ are sequences evolving according to (20) with initial values 0 and $\bar{\mathbf{Y}}_0$, respectively. Since we already show above that the two sequences converge to $\bar{\mathbf{Y}}_0$, $\{\mathbf{Y}_{k_i}\}$ also converges to $\bar{\mathbf{Y}}_0$. In other words, $\lim_{i\to\infty}\mathbf{Y}_{k_i} = \bar{\mathbf{Y}}_0$ regardless of $\mathbf{Y}_0$.

Finally, consider the whole consequence $\{\mathbf{Y}_k\}$. With $\lim_{i\to\infty}\mathbf{Y}_{k_i} = \bar{\mathbf{Y}}_0$, $\forall j \in \{1, \ldots, n-1\}$, the sequence $\{\mathbf{Y}_{k_i+j} = \mathbf{A}^j\mathbf{Y}_{k_i}(\mathbf{A}^T)^j + \sum_{t=0}^{j-1}\mathbf{A}^t\mathbf{Q}(\mathbf{A}^T)^t\}$ also converges to a unique value $\bar{\mathbf{Y}}_j = \mathbf{A}^j\bar{\mathbf{Y}}_0(\mathbf{A}^T)^j + \sum_{t=0}^{j-1}\mathbf{A}^t\mathbf{Q}(\mathbf{A}^T)^t$. Therefore,

$$\lim_{L\to\infty}\frac{1}{L}\sum_{k=1}^{L}\mathbf{Y}_k = \frac{1}{n}\sum_{j=0}^{n-1}\bar{\mathbf{Y}}_j \triangleq \bar{\mathbf{Y}}(\gamma, n).$$

Moreover, based on (44), we can conclude that $\bar{\mathbf{Y}}_j$ and also $\bar{\mathbf{Y}}(\gamma, n)$ are monotonically decreasing as either γ increases or n decreases.

REFERENCES

[1] X. Cao, X. Zhou, and Y. Cheng, "Energy efficient spectrum sensing for state estimation over a wireless channel," in *Proc. IEEE GlobalSIP*, Atlanta, GA, USA, Nov. 2014.

[2] J. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, Jan. 2007.

[3] X. Lin, R. Lu, X. Shen, Y. Nemoto, and N. Kato, "SAGE: A strong privacy-perserving scheme against global eavesdropping for ehealth systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 4, pp. 365–378, May 2009.

[4] R. T. Sukhavasi and B. Hassibi, "The Kalman-like particle filter: Optimal estimation with quantized innovations/measurements," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 131–136, Jan. 2013.

[5] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[6] K. J. Kim, J. Yue, R. A. Iltis, and J. D. Gibson, "A QRD-M/Kalman filter-based detection and channel estimation algorithm for MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 710–721, Mar. 2005.

[7] T. Jiang, N. D. Sidiropoulos, and G. B. Giannakis, "Kalman filtering for power estimation in mobile communications," *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 151–161, Jan. 2003.

[8] G. Bianchi and I. Tinnirello, "Kalman filter estimation of the number of competing terminals in an IEEE 802.11 network," in *Proc. IEEE INFOCOM*, 2003, vol. 2, pp. 844–852.

[9] I. Kadota, A. Baiocchi, and A. Anzaloni, "Kalman filtering: Estimate of the numbers of active queues in an 802.11e EDCA WLAN," *Comput. Commun.*, vol. 39, pp. 54–64, Feb. 2014.

[10] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, Apr. 2007.

[11] E. Rohr, D. Marelli, and M. Fu, "A unified framework for mean square stability of Kalman filters with intermittent observations," in *Proc. 9th IEEE ICCA*, 2011, pp. 177–182.

[12] H. Li, Y. Cheng, C. Zhou, and P. Wan, "Multi-dimensional conflict graph based computing for optimal capacity in mr-mc wireless networks," in *Proc. IEEE 30th ICDCS*, 2010, pp. 774–783.

[13] E. Hossain, D. Niyato, and Z. Han, *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[14] E. Peh, Y.-C. Liang, Y. Guan, and Y. Zeng, "Optimization of cooperative sensing in cognitive radio networks: A sensing-throughput tradeoff view," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 5294–5299, Nov. 2009.

[15] X. Zhou, G. Y. Li, D. Li, D. Wang, and A. C. Soong, "Probabilistic resource allocation for opportunistic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2870–2879, Sep. 2010.

[16] S. Stotas and A. Nallanathan, "On the throughput and spectrum sensing enhancement of opportunistic spectrum access cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 97–107, Jan. 2012.

[17] S.-S. Tan, J. Zeidler, and B. Rao, "Opportunistic channel-aware spectrum access for cognitive radio networks with interleaved transmission and sensing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2376–2388, May 2013.

[18] X. Ma, S. M. Djouadi, and H. Li, "State estimation over a semi-Markov model based cognitive radio system," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2391–2401, Jul. 2012.

[19] X. Cao, P. Cheng, J. Chen, S. S. Ge, Y. Cheng, and Y. Sun, "Cognitive radio based state estimation in cyber-physical systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 489–502, Mar. 2014.

[20] L. Shi, M. Epstein, B. Sinopoli, and R. Murray, "Effective sensor scheduling schemes in a sensor network by employing feedback in the communication loop," in *Proc. IEEE Int. Conf. Control Appl.*, Singapore, 2007, pp. 1006–1011.

[21] X. Zhou, Y. Li, Y. H. Kwon, and A. Soong, "Detection timing and channel selection for periodic spectrum sensing in cognitive radio," in *Proc. IEEE GLOBECOM*, 2008, pp. 1–5.

[22] X. Fang, D. Yang, and G. Xue, "Taming wheel of fortune in the air: An algorithmic framework for channel selection strategy in cognitive

radio networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 783–796, Feb. 2013.

[23] H. Jiang, L. Lai, R. Fan, and H. V. Poor, "Optimal selection of channel sensing order in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 297–307, Jan. 2009.

[24] R. Fan and H. Jiang, "Channel sensing-order setting in cognitive radio networks: A two-user case," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 4997–5008, Nov. 2009.

[25] H. T. Cheng and W. Zhuang, "Simple channel sensing order in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 676–688, Apr. 2011.

[26] W.-Y. Lee and I. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3845–3857, Oct. 2008.

[27] R. Deng, J. Chen, X. Cao, Y. Zhang, S. Maharjan, and S. Gjessing, "Sensing-performance tradeoff in cognitive radio enabled smart grid," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 302–310, Mar. 2013.

[28] A. Censi, "Kalman filtering with intermittent observations: Convergence for semi-Markov chains and an intrinsic performance measure," *IEEE Trans. Autom. Control*, vol. 56, no. 2, pp. 376–381, Feb. 2011.

[29] X. Ma, H. Li, and S. Djouadi, "Networked system state estimation in smart grid over cognitive radio infrastructures," in *Proc. 45th Annu. CISS*, 2011, pp. 1–5.

[30] L. Liu, X. Cao, Y. Cheng, L. Du, W. Song, and Y. Wang, "Energy-efficient capacity optimization in wireless networks," in *Proc. IEEE INFOCOM*, 2014, pp. 1384–1392.

[31] W. Xu, A. Tajer, X. Wang, and S. Alshomrani, "Power allocation in miso interference channels with stochastic csit," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1716–1727, Mar. 2014.

[32] W. Xu, Q. Shi, X. Wei, Z. Ma, X. Zhu, and Y. Wang, "Distributed optimal rate-reliability-lifetime tradeoff in time-varying wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4836–4847, Sep. 2014.

[33] L. Liu, X. Cao, Y. Cheng, and Z. Niu, "Energy-efficient sleep scheduling for delay-constrained applications over WLANs," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2048–2058, Jan. 2014.

[34] Y. Mo, B. Sinopoli, L. Shi, and E. Garone, "Infinite-horizon sensor scheduling for estimation over lossy networks," in *Proc. IEEE 51st Annu. Conf. Decision Control*, 2012, pp. 3317–3322.

[35] H. Kim and K. G. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 5, pp. 533–545, May 2008.

[36] X. Zhou, J. Ma, G. Y. Li, Y. H. Kwon, and A. C. Soong, "Probability-based optimization of inter-sensing duration and power control in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4922–4927, Oct. 2009.

[37] W. S. Jeon, D. H. Lee, and D. G. Jeong, "Collaborative sensing management for cognitive radio networks with reporting overhead," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 595–605, Feb. 2013.

[38] X. Cao, P. Cheng, J. Chen, and Y. Sun, "An online optimization approach for control and communication codesign in networked cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 439–450, Feb. 2013.

Xiangwei Zhou (S'08–M'13) received the B.S. degree in communication engineering from Nanjing University of Science and Technology, Nanjing, China, in 2005, the M.S. degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2007, and the Ph.D. degree in electrical and computer engineering from Georgia Institute of Technology, Atlanta, GA, USA, in 2011. Since 2013, he has been with the Department of Electrical and Computer Engineering, Southern Illinois University Carbondale, as an Assistant Professor. Prior to that, he was a Senior Systems Engineer with Marvell Semiconductor, Santa Clara, CA, USA, from 2011 to 2013. His general research interests include wireless communications, statistical signal processing, and cross-layer optimization, with current emphasis on cognitive radio and spectrum coexistence. He won the Best Paper Award in the 2014 International Conference on Wireless Communications and Signal Processing.

Lu Liu (S'13) received the B.S. degree in automation from Tsinghua University, Beijing, China, in 2010 and the M.S. degree in electrical engineering from Illinois Institute of Technology, Chicago, IL, USA, in 2012. She is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL, USA. Her current research interest includes energy efficient networking and communication, resource allocation and protocol design of wireless networks.

Xianghui Cao (S'08–M'11) received the B.S. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2006 and 2011, respectively. During 2007–2009, he was a Visiting Scholar in the Department of Computer Science, University of Alabama, Tuscaloosa, AL, USA. Since 2012, he has been a Senior Research Associate with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL, USA. His research interests include cyber-physical systems, wireless network performance analysis, networked estimation and control, and network security. He received the Best Paper Runner-Up Award from ACM MobiHoc 2014. He serves as an Associate Editor for several journals including *KSII Transactions on Internet and Information Systems, Security and Communication Networks* (Wiley) and *International Journal of Ad Hoc and Ubiquitous Computing*, a publicity co-chair for ACM MobiHoc 2015, and a TPC member for a number of conferences including IEEE Globecom, IEEE ICC, and IEEE VTC.

Yu Cheng (S'01–M'04–SM'09) received the B.E. and M.E. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1995 and 1998, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, Ontario, Canada, in 2003. From September 2004 to July 2006, he was a Postdoctoral Research Fellow in the Department of Electrical and Computer Engineering, University of Toronto, Ontario, Canada. Since August 2006, he has been with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, Illinois, USA, where he is now an Associate Professor. His research interests include next-generation Internet architectures and management, wireless network performance analysis, network security, and wireless/wireline interworking. He received a Best Paper Award from the conferences QShine 2007 and IEEE ICC 2011, and the Best Paper Runner-Up Award from ACM MobiHoc 2014. He received the National Science Foundation (NSF) CAREER AWARD in 2011 and IIT Sigma Xi Research Award in the junior faculty division in 2013. He served as a Co-Chair for the Wireless Networking Symposium of IEEE ICC 2009, a Co-Chair for the Communications QoS, Reliability, and Modeling Symposium of IEEE GLOBECOM 2011, a Co-Chair for the Signal Processing for Communications Symposium of IEEE ICC 2012, a Co-Chair for the Ad Hoc and Sensor Networking Symposium of IEEE GLOBECOM 2013, and a Technical Program Committee (TPC) Co-Chair for WASA 2011 and ICNC 2015. He is a founding Vice Chair of the IEEE ComSoc Technical Subcommittee on Green Communications and Computing. He is an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and the New Books & Multimedia Column Editor for IEEE NETWORK.