Cognitive Radio Based State Estimation in Cyber-Physical Systems

Xianghui Cao *Member, IEEE*, Peng Cheng *Member, IEEE*, Jiming Chen *Senior Member, IEEE*, Shuzhi Sam Ge *Fellow, IEEE*, Yu Cheng *Senior Member, IEEE* and Youxian Sun

***Abstract*—We investigate the state estimation problem in cyber-physical systems (CPS) where the dynamical physical process is measured by a wireless sensor and the measurements are transmitted to a remote state estimator. It has been shown that the estimation performance strongly depends on the wireless communication quality. To enhance the estimation performance, we apply the cognitive radio technique to the system and propose a CHAnnel seNsing and switChing mEchanism (CHANCE) to explore opportunistic accessibility of multiple channels. We consider two types of wireless channels, i.e., one unlicensed channel which can be accessed freely and several licensed channels which have been pre-assigned to primary users. For the single-licensed-channel case, we develop a necessary condition for the estimation stability based on the physical process dynamics, channel quality and the channel sensing accuracy. This condition becomes also sufficient under certain conditions. We also derive the conditions under which the estimation performance is guaranteed to be improved by CHANCE. The above results are then extended to multi-licensed-channel cases. Simulations based on a particular linear system show that, the long-run mean estimation error covariance with CHANCE is at least 63% less than that without CHANCE. It is also shown that CHANCE outperforms the existing RANDOM mechanism in terms of estimation performance.**

***Index Terms*—Cognitive radio; cyber physical system; state estimation; stability condition; performance bounds**

I. Introduction

The emerging cognitive radio (CR) technology which allows dynamic spectrum reuses can significantly improve spectrum utilization efficiency and mitigate the situation of spectrum scarcity [1], [2]. With CR, unlicensed users can opportunistically use holes of the licensed spectrum, by which the communication quality of the unlicensed users is expected to be improved. A considerable amount of literature has been published that studies various aspects of the CR networks such as spectrum sensing algorithms, network protocols, resource allocation, and network security [3]–[5].

In this paper, we study the CR technology from a cyber physical system (CPS) perspective. CPS systematically integrate physical processes with sensors, actuators and efficient computation and communication modules to provide real-time monitoring, control and other services. CPS facilitate intimate interactions between human beings and the physical world in a wide range of applications such as environment surveillance, vehicular networks, industrial automation and smart grid [6]–[10]. The close interplay between the cyber and physical spaces often requires real-time operation, efficient system resource utilization and security guarantee [7], [8].

In many CPS applications, sensor data often need to be transmitted through the wireless media to remote units for state estimation which plays a vital role in provisioning real-time monitoring and control. For example, in a smart grid, field sensors report real-time voltage and phase observations to remote power generators which estimate the real voltage and take certain actions in order to stabilize it within some desired range [6]. Due to a variety of factors such as bandwidth limitation, interferences from other wireless devices or environment, and channel fading, wireless communications usually undergo packet losses and delay which impose strict requirements on the communication protocols as well as estimation and control policies [11], [12]. Since the estimation mechanism relies on the information received from sensors, packet losses strongly affect the estimation performance. When a loss occurs, the estimator will have to "guess" current state of the physical process based on all historically received data, which will introduce error to the estimation performance. Illustrations of such effect are shown in Fig. 1 where we consider a simple i.i.d. packet loss model. The first figure shows that a packet loss will cause the estimation deviate from the one without loss, and the deviation is enlarged when packets are lost successively; the other figure demonstrates that the estimation performance degrades as the loss rate increases.

The problem of state estimation stability under lossy wireless communications has attracted intensive research attentions [13]–[16]. Considering the i.i.d packet loss model, Sinopoli *et al.* prove the existence of a critical value for the packet loss rate below which the estimation error covariance would converge in the mean square sense [13]. The estimation stability under a more complicated channel model with Markovian packet losses has been studied in [14]. A class of more general packet loss models based on semi-Markov chains is investigated in [15]. In order to preserve the estimation performance, sensor scheduling and transmission power control methods have been proposed to improve the communication quality [17], [18]. Basically, due to bandwidth limitation, a single channel may be insufficient for estimating fast changing physical processes. For state estimation with multiple sensors,

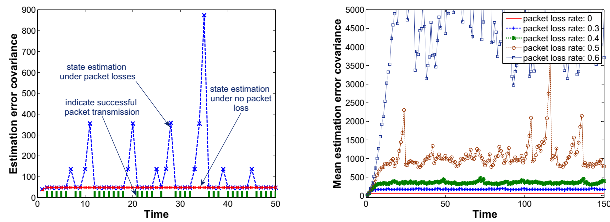
Manuscript received April 3, 2013; revised August 20, 2013. The work was partially supported by the Natural Science Foundation of China under Grants 61203036, 61222305 and 61290325-02. The work of Y. Cheng was also supported in part by NSF grants CNS-1053777 and CNS-1117687.

X. Cao, P. Cheng (correspondence author), J. Chen and Y. Sun are with the State Key Laboratory of Industrial Control Technology, Department of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China. Email: {xhcao,pcheng,jmchen,yxsun}@iipc.zju.edu.cn.

S.S. Ge is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. Email: samge@nus.edu.cn.

Y. Cheng is with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, IL, USA. Email:cheng@iit.edu.

Chiuso *et al.* prove that the optimal estimation performance under packet loss cannot be reached using a single bandwidth-limited channel [19]. Liu *et al.* study the estimation stability over multiple wireless channels by dividing the measurement data and then transmit them through the channels separately [20]. However, they focus on fixed communication structures without considering channel dynamics.



(a) Effect of packet losses (b) Effect of packet loss rates

Fig. 1. Impact of communication quality on estimation performance, where the physical process is the same as that in Section V.

We apply CR over multiple channels to enhance the state estimation performance in CPS. While it is well noticed that CR can effectively improve spectrum efficiency in the communication regime, its ability has not been fully exploited in the CPS domain. In the literature, only a few works have investigated the estimation and control problems with CR. In [21], Ma *et al.* propose an optimal estimation and control algorithm and analyze the system stability conditions focusing on a single licensed channel. The results have been extended to multi-channel cases [22]. However, in their work, only one licensed channel is sensed and opportunistically used for each transmission, which is different from ours where both licensed and unlicensed channels can be used. In addition, they do not consider channel sensing inaccuracy which is a fundamental problem for practical cognitive radio systems [23].

We propose a CR based CHAnnel seNsing and switChing mEchanism (CHANCE) for state estimation for a class of generic CPS with linear state dynamics. We aim to answer whether and how the state estimation can be improved by the new mechanism. Different from classical CR theories which focus on network throughput, we emphasize on the communication reliability and state estimation performance. Our main contributions can be summarized as follows: (1) We propose a new state estimation framework by taking the advantage of the CR technology, and identify a novel packet loss model. (2) For the case that only one licensed channel is available, we derive a necessary condition for the estimation stability which takes factors from both physical and cyber spaces into account. This condition is also sufficient when the observation matrix has full column rank. Moreover, this condition is a generalization of existing results for either i.i.d. or Markov packet loss models. (3) We obtain the conditions under which the proposed mechanism can certainly improves the estimation performance. Moreover, a pair of upper and lower bounds for the mean square estimation error is determined. The performance improvement in terms of performance ratio is analyzed and the worst case ratio is derived. (4) We generalize the above results to accommodate multiple licensed channels. Corresponding stability conditions and performance bounds are also derived. Numeric results are provided to demonstrate the effectiveness of our proposed mechanism.

The remainder of this paper is organized as follows. The whole system is modeled in Section II. For the single-licensed-channel case, the state estimation stability conditions and performance are studied in Section III. Extensions to multi-licensed-channel cases are provided in Section IV. Section V presents simulations results. Finally, Section VI concludes the paper. To be concise, we present all main results in form of theorems to which the proofs can be found in the Appendices.

II. SYSTEM MODELING

We consider a class of generic CPS where the physical process has linear state dynamics. As shown in Fig. 2, a wireless sensor periodically samples the physical process and transmits its measurement to a remote estimator for state estimation. There is an unlicensed channel, CH_0, which is likely to be used by a large amount of wireless devices and thus packet collision may happen frequently. Alternatively, the sensor can opportunistically use the spectrum holes of several licensed channels (sub-bands divided from a wide-band channel), $\{CH_i|i=1,\dots,m\}$, each of which is restricted to only a primary user (PU) that may not always occupy it. The basic idea of CHANCE is to apply CR to use spectrum holes of licensed channels for measurement transmission. Here, we assume all the channels are mutually independent [24], [25]. We adopt the "listen before talk" strategy such that the sensor performs channel sensing right before each data reporting time.

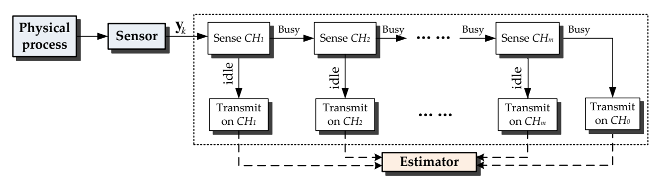


Fig. 2. Overview of CHANCE.

In the following, we present system models, where the main notations are listed in Table I. We adopt the following convention: $\mathbb{R}^n$ is the n-dimensional Euclidean space. $\mathbb{E}[\cdot]$ and $\mathbb{P}\{\cdot\}$ denote the expectation and probability of a random variable, respectively. We use boldface letters in lower case to denote vectors, and capital letters to denote matrices. For any matrix/vector M, M' and M^H denote the transpose and Hermitian transpose of M, respectively. $[\cdot]_i$ denotes the i-th entry of a vector, while $[\cdot]_{ij}$ denotes the (i,j)-th entry of a matrix. $\mathrm{Tr}(\cdot)$ and $\rho(\cdot)$ denote the trace and spectral radius of a square matrix, respectively. Denote $\mathbb{S}_+^n$ as the set of positive semi-definite matrices of dimension $n\times n$. $\forall X\in\mathbb{S}_+^n$, $\lambda_{\min}(X)$ and $\lambda_{\max}(X)$ are its minimum and maximum eigenvalues, respectively. $\forall X,Y\in\mathbb{S}_+^n$, $X\geq Y$ if $X-Y\in\mathbb{S}_+^n$.

A. The Physical Process and State Estimation Algorithm

We consider that the physical process runs with the following discrete-time state dynamics:

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

where $\mathbf{x}_k\in\mathbb{R}^n$ is the system state, $\mathbf{a}_k\in\mathbb{R}^{n'}$ is the control action, $\mathbf{w}_k\in\mathbb{R}^n$ is the system noise. As a typical and generic

Table I
NOTATIONS

Notation	Description
$\mathbf{x}_k, \mathbf{y}_k$	System state and measurement of the physical process
CH_0, CH_i	The original channel and the i-th licensed channel, respectively. $i \in \{1, \ldots, m\}$
ℓ_0, ℓ_i	Packet loss rate on CH_0 and the loss rate on CH_i when CH_i is idle
T, τ	Sensor's sampling period and channel sensing time
$s_{i,k}$, $o_{i,k}$	The channel state and channel sensing result of CH_i.
α_i, β_i	CH_i's state transition probabilities. See (5)
$p_{d,i}, p_{f,i}$	Correct detection and false detection probabilities, respectively, for sensing CH_i

process model, Eq. (1) is widely adopted for describing the state dynamics of power systems, smart grid infrastructures, building automation systems, etc. [8], [22], [26]. The sensor's measurement $\mathbf{y}_k \in \mathbb{R}^{n''}$ is modeled by

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

where $\mathbf{v}_k \in \mathbb{R}^{n''}$ is the measurement noise. In the above, A, B and C are constant matrices with compatible dimensions. For simplicity, we assume that: $\mathbf{w}_k \in \mathbb{R}^n$, $\mathbf{v}_k \in \mathbb{R}^{n''}$, and the initial state $\mathbf{x}_0$ are mutually independent white Gaussian noises with means $0, 0, \bar{\mathbf{x}}_0$ and covariance matrices $Q > 0, R > 0, P_{0|0} > 0$, respectively; the pair $(A, \sqrt{Q})$ is controllable and the pair (A, C) is observable.

Kalman Filter based State Estimation: In each step, we define a random variable $\gamma_k \in \{0, 1\}$ to model the wireless communication reliability. Specifically, $\gamma_k = 1$ if the measurement packet is successfully received by the estimator during step k, and $\gamma_k = 0$ if the packet is lost. 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. Denote the predicted and estimated error covariance by $P_{k|k-1} \triangleq \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})']$ and $P_{k|k} \triangleq \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})']$, respectively. The Kalman filter under packet losses runs recursively as follows [13].

$$\begin{cases} \hat{\mathbf{x}}_{k|k-1} & \triangleq & A\hat{\mathbf{x}}_{k-1|k-1} + B\mathbf{a}_k \\ P_{k|k-1} & = & AP_{k-1|k-1}A' + Q \\ \hat{\mathbf{x}}_{k|k} & = & \hat{\mathbf{x}}_{k|k-1} + \gamma_k K_k(\mathbf{y}_k - C\hat{\mathbf{x}}_{k|k-1}) \\ K_k & \triangleq & P_{k|k-1}C'(CP_{k|k-1}C' + R)^{-1} \\ P_{k|k} & = & (I - \gamma_k K_k)P_{k|k-1} \end{cases} \tag{3}$$

Based on (3), the predicted error covariance satisfies the following modified algebraic Riccati equation.

$$P_k = AP_{k-1}A' - \gamma_k AP_{k-1}C'(CP_{k-1}C' + R)^{-1}CP_{k-1}A' + Q. \tag{4}$$

where we define $P_k \triangleq P_{k|k-1}$ for convenience. We assume that[1] $P_0 \in \mathbb{S}_Q \triangleq \{S | S \in \mathbb{S}_+^n, S \geq Q\}$. We call the estimation process is *mean square error (MSE)* stable if $\sup_{k\geq 0} \mathbb{E}[\mathrm{Tr}(P_k)] < \infty$, where the expectation is taken over the packet loss process $\{\gamma_k\}_{k\geq 0}$.

If the communications are perfect, i.e., $\gamma_k \equiv 1$, P_k converges to a unique value [13]. The convergence may not hold in the presence of packet losses. In the literature, the sequence $\{\gamma_k\}_{k\geq 0}$ is usually assumed comply with either an i.i.d. or a Markov chain distribution [13], [14], [27]. Nevertheless, with the CHANCE mechanism, the packet loss process is neither i.i.d. nor Markovian as discussed in Remark 1. Therefore, we derive new necessary and sufficient conditions for the MSE stability of the state estimation.

B. CHANCE in the Cyber Space

The sensor is equipped with one antenna that can sense the signals in the licensed and unlicensed channels. To handle multi-channel sensing, sequential sensing (SS) and wideband sensing (WS) are two candidate methods [3]. With SS, the sensor scans the channels one-by-one until finding an idle channel and then transmit its data on it. In this way, the sensor can apply low-complexity spectrum sensing strategy, e.g., energy detection. In contrast, WS is able to handle multiple channels at the same time; however, it usually requires the sensor to be enough powerful for carrying out high-complexity calculations [3], [28]. In addition, since SS only sense a subset of all the channels in each round, it is likely to consume less energy than WS. Therefore, we apply SS in our study.

As shown in Fig. 2, CHANCE runs as follows. In each step, the sensor first scans the licensed channels following an ordered channel index set $\mathcal{Q}_m$. It will stop sensing and transmit packet through the first channel that is sensed idle (unoccupied by its PU). If all the licensed channels in $\mathcal{Q}_m$ are sensed busy, the sensor will transmit over the original channel CH_0. During scanning, the channel sensing time spent on channel CH_i is τ_i. Assume that τ_i is small enough such that the maximum scanning time will not exceed one step, i.e., $\sum_{i=1}^m \tau_i < T$, and that the channel state does not change during τ_i. Without loss of generality, let $\mathcal{Q}_m = (1, \ldots, m)$.

For each $i \in \mathcal{Q}_m$, denote the state of CH_i at the time when the sensor senses CH_i in step k as $s_{i,k} \in \{0, 1\}$. $s_{i,k} = 1$ if CH_i is busy and $s_{i,k} = 0$ otherwise. Denote $o_{i,k} \in \{0, 1\}$ as the sensing result such that $o_{i,k} = 0$ if CH_i is sensed idle; in this case, the sensor will switch to CH_i for transmission[2]. Conversely, $o_{i,k} = 1$ indicates that CH_i is sensed busy. Once a transmission on CH_i is completed, the sensor switches back to CH_0 in order to minimize possible interference to the PU.

Channel Modeling: We assume that transmissions on CH_0 suffer from i.i.d. packet losses with average loss rate ℓ_0. For each licensed channel CH_i, define its *idle state loss rate* ℓ_i as the loss rate of packets transmitted on CH_i when CH_i is idle, to account for factors such as path fading and environmental noises. The PU's activities on CH_i are modeled as a continuous-time Markov process: the busy and idle periods of CH_i obey exponential distributions with means $\frac{1}{\omega_{i,1}}$ and $\frac{1}{\omega_{i,0}}$, respectively. In case $\omega_{i,1} = 0$ (or $\omega_{i,0} = 0$), CH_i becomes always busy (or idle). Under the hypothesis that $\omega_{i,1}$ and $\omega_{i,0}$ are known parameters, performing channel sensing on a constantly busy (idle) channel is unnecessary. Therefore, we assume $\omega_{i,1}\omega_{i,0} > 0$ in the sequel. The following proposition derives the *probability transition matrix* Φ_i of CH_i where $[\Phi_i]_{jl} =: \mathbb{P}\{s_{i,k+1} = l - 1 | s_{i,k} = j - 1\}$. Its proof can be found in Appendix A.

[1] Since $P_{0|0} > 0$, by the second line of (3), $P_k \in \mathbb{S}_Q, \forall k > 0$. It is without loss of generality to assume $P_0 \in \mathbb{S}_Q$.

[2] We assume that there is another common control channel for the sensor to inform the estimator to receive on CH_i [29].

Proposition 1: $\{s_{i,k}\}_{k\geq 0}$ forms a homogeneous Markov chain with transition probability matrix

$$\Phi_i = \begin{bmatrix} 1-\alpha_i & \alpha_i \\ \beta_i & 1-\beta_i \end{bmatrix} \tag{5}$$

where $\alpha_i = \frac{\omega_{i,0}}{\omega_{i,1}+\omega_{i,0}}\left[1 - e^{-(\omega_{i,1}+\omega_{i,0})T}\right]$ and $\beta_i = \frac{\omega_{i,1}}{\omega_{i,1}+\omega_{i,0}}\left[1 - e^{-(\omega_{i,1}+\omega_{i,0})T}\right]$.

Remark 1: Since $\omega_{i,1}\omega_{i,0} > 0$, $0 < \alpha_i + \beta_i < 1$. For finite sampling period T, the sequence $\{(s_{i,k}, \gamma_k) : k \geq 0\}$ forms a hidden Markov process with $s_{i,k}$ as the hidden state and γ_k as the observation [30]. Due to imperfect channel sensing (see below), present packet loss rate may depend on all historical data. In other words, the packet loss process under CHANCE is generally neither an i.i.d. process nor a Markov chain, making many existing results in the literature inapplicable.

Remark 2: Above we describe a simple model for CH_i. However, the results obtained in this paper also hold for more general models, providing that $\{s_{i,k}\}_{k\geq 0}$ is a homogeneous Markov chain with known transition probability matrix Φ_i and that $0 < \alpha_i + \beta_i < 1$. For example, the semi-Markov model presented in [22] can be adopted to allow any given homogeneous distributions of the busy/idle periods. In this case, the transition probability matrix Φ can be obtained by the method proposed in [31]. We will compare the performance of CHANCE and the strategy proposed in [22] in Section V.

Channel Sensing: There are a few channel sensing techniques available, such as energy detection (which is simple but has relatively low accuracy) and feature detector (which is more accurate but requires prior knowledge of the PU's signals) [32]. To characterize the channel sensing accuracy, we define two conditional probabilities, i.e., $p_{d,i} = \mathbb{P}\{o_{i,k} = 0 | s_{i,k} = 0\}$ and $p_{f,i} = \mathbb{P}\{o_{i,k} = 0 | s_{i,k} = 1\}$, as the *correct detection* and *false detection probabilities*, respectively. For perfect channel sensing, $p_{d,i} = 1$ and $p_{f,i} = 0$. Generally, $p_{d,i}$ and $p_{f,i}$ highly depend on the sensing technique. It is worth noticing that our results do not rely on any particular formulas of the above two probabilities.

III. The Single-Licensed-Channel Case

We begin with the simple single-licensed-channel case (i.e., $m = 1$) while extensions to general cases are presented in the next section. In this section, we first study the conditions for the estimation stability. Then we show how CHANCE can expand the estimation stability region. We analyze the estimation performance bounds and find the conditions when CHANCE certainly improves the performance.

Suppose the licensed channel is CH_1. The suffix 1 indicating CH_1 in the above defined variables is dropped in this section since the indication is clear. Denote the *conditional packet loss probabilities* on the state of CH_1 by $\psi_1 \triangleq \mathbb{P}\{\gamma_k = 0 | s_k = 0\}$ and $\psi_2 \triangleq \mathbb{P}\{\gamma_k = 0 | s_k = 1\}$. In case that CH_1 is idle, the measurement packet will get lost with probability equals to either ℓ_1 (if the channel sensing result is $o_k = 0$ and thus the packet is delivered over CH_1) or ℓ_0 (if $o_k = 1$ and thus packet is transmitted on CH_0). Therefore, ψ_1 can be expressed below. The expression for ψ_2 is similar.

$$\begin{cases} \psi_1 = & p_d\ell_1 + (1-p_d)\ell_0 \triangleq \psi_1(p_d), \\ \psi_2 = & p_f + (1-p_f)\ell_0 \triangleq \psi_2(p_f) \end{cases} \tag{6}$$

Note that the two equations in (6) are asymmetric because a packet will get lost at rate ℓ_1 under correct detection, but will certainly lose under false detection. Define the *sensing matrix* $\Psi \triangleq Diag\{\psi_1, \psi_2\}$ and the *critical matrix* $\Phi\Psi = \begin{bmatrix} \psi_1(1-\alpha) & \psi_2\alpha \\ \psi_1\beta & \psi_2(1-\beta) \end{bmatrix}$. Let $\sigma_2 \geq \sigma_1$ be the eigenvalues of the critical matrix. $[\Phi\Psi]_{ij}$ represents the conditional probability $\mathbb{P}\{\gamma_k = 0, s_k = j-1 | s_{k-1} = i-1\}$.

A. Stability Conditions

Theorem 1: A necessary condition for the MSE stability of the estimation applied with CHANCE is:

$$\sigma_2 \rho(A)^2 < 1. \tag{7}$$

Theorem 2: If C has full column rank, condition (7) is both necessary and sufficient.

Remark 3: The above result is more general than some existing conditions. For example,

1) in the special case with $p_d = p_f = 0$, CH_1 will never be accessed, i.e., all the transmissions are carried out over CH_0 with i.i.d. packet loss rate ℓ_0. In this case $\psi_1 = \psi_2 = \ell_0, \sigma_2 = \ell_0$, and condition (7) reduces to $\ell_0\rho(A)^2 < 1$ which recovers existing result in [13].
2) in the special case with $\ell_1 = 0$ and $p_f = p_d = 1$, all the transmissions are carried out over CH_1 and the packet loss process obeys Markov distribution with recovery rate $\mathbb{P}\{\gamma_k = 1 | \gamma_{k-1} = 0\} = \mathbb{P}\{s_k = 0 | s_{k-1} = 1\} = \beta$. In this case, $\psi_1 = 0, \psi_2 = 1$ and $\sigma_2 = 1 - \beta$. As a result, condition (7) reduces to $(1-\beta)\rho(A)^2 < 1$ which recovers the condition proposed in [14].

Remark 4: Condition (7) provides a joint condition for both physical and cyber spaces. It can be interpreted as follows. If we define $\tilde{s}_k = (\gamma_k = 0, s_k)$ which combines packet loss and channel state, then the critical matrix is indeed the probability transition matrix for the joint state $\tilde{s}_k$. σ_2 can be viewed as a measure of how frequently the measurement packets get lost, while $\rho(A)$ is a measure of how fast the system evolves [33]. $1 - \sigma_2\rho(A)$ can be roughly interpreted as the rate of real-time system information transmitted from the sensor to the estimator, and hence condition (7) means that rate is positive.

For high dimensional systems, the matrix C of a single sensor may be a "wide" matrix and may not have full column rank. However, our results still can be valid in some practical cases. For example, if the sensor under concern is the head of a cluster of sensors, it can gather measurements from them and report the total information to the estimator. In this case, the whole observation matrix C will have enough rows to have full column rank. For general C, it is mathematically difficult to derive necessary and sufficient conditions for the MSE stability. However, for a class of non-degenerate systems, one can follow the methods in [14], [34] to derive the stability conditions. The results are omitted due to space limitations.

In the rest of this section, we assume that the quality of CH_1 in its idle state is better than that of CH_0, i.e., $\ell_1 < \ell_0$. If otherwise, the sensor better stays on CH_0, which reduces the system to the conventional single channel case. Also assume that C has full column rank for simplicity. We use superscript † on the variables to indicate the case without CHANCE.

B. Stability Gain Analysis

Define the *stability region* Ω such that $\forall A \in \Omega$, the estimation process is MSE stable. According to Theorem 2, $\Omega = \{A|\sigma_2\rho(A)^2 < 1\}$. Define the stability region radius as $|\Omega| \triangleq \max_{A\in\Omega}\{\rho(A)\}$ and the stability gain as $\eta \triangleq \frac{|\Omega|}{|\Omega^\dagger|}$. It is easy to see that $|\Omega^\dagger| = \frac{1}{\sqrt{\ell_0}}$ and $|\Omega| = \frac{1}{\sqrt{\sigma_2}}$ according to Theorem 2. The following theorem characterizes the stability gain for which the proof is omitted due to space limitations.

***Theorem** 3:* The stability gain η is monotonically increasing as the correct detection probability p_d increases or the false detection probability p_f decreases. Moreover,

$$\eta = \sqrt{\frac{\ell_0}{\sigma_2}} \le \min\left\{\frac{1}{\sqrt{1-\beta}}, \sqrt{\frac{\ell_0}{\ell_1}}\right\}. \tag{8}$$

The theorem shows that, in order to expand the stability region (i.e., $\eta > 1$), we must have $\sigma_2 < \ell_0$. Since $\ell_1 < \ell_0$, by Lemma 1, we know that σ_2 is monotonically decreasing as either p_d increases or p_f decreases. Therefore, the condition $\sigma_2 < \ell_0$ yields a lower (or an upper) bound of p_d (or p_f). The upper bound in (8) is achievable if CH_1 is a constant channel (i.e., it is always busy or idle) and channel sensing is perfect (i.e., $p_d = 1, p_f = 0$). If CH_1 is always busy, the sensing results will always suggest to transmit on CH_0 (with loss rate ℓ_0) and thus $\eta = 1$; similarly, if CH_1 is constantly idle, CH_1 will be always used and $\eta = \sqrt{\frac{\ell_0}{\ell_1}}$.

C. Performance Analysis

In this part, we analyze the estimation performances in terms of the mean estimation error covariance $\mathrm{Tr}(\mathbb{E}[P_k])$. We assume that both ℓ_0 and σ_2 are less than $\frac{1}{\rho(A)^2}$ since otherwise either $\mathrm{Tr}(\mathbb{E}[P_k])$ or $\mathrm{Tr}(\mathbb{E}[P_k^\dagger])$ will grow to infinity based on Theorem 2, making the performance comparison trivial.

***Theorem** 4:* If $\psi_2(1-\beta)+\psi_1\beta \le \ell_0$ and the initial conditions satisfy $P_0^\dagger = P_0$, then $\forall k \ge 0$, $\mathrm{Tr}(\mathbb{E}[P_k]) \le \mathrm{Tr}(\mathbb{E}[P_k^\dagger])$. Moreover, if $\psi_2(1-\beta)+\psi_1\beta < \ell_0$, the inequality is strict.

By (6), the condition $\psi_2(1-\beta)+\psi_1\beta < \ell_0$ is equivalent to $p_f\beta(1-\ell_0) < p_d(1-\beta)(\ell_0-\ell_1)$, which yields an upper (or lower) bound of p_f (or p_d). For energy detection mentioned above, this condition can be satisfied by appropriately tuning the channel sensing time and detection threshold [3]. Hence, Theorem 4 reveals that, under a mild condition, the estimation performance is definitely improved by CHANCE. Note that the improvement is guaranteed in very step.

In general, even though the mean error covariance $\mathbb{E}[P_k]$ is upper bounded, it does not necessarily converge to a unique steady state. Consider two sequences: $\{\bar{P}_k|\bar{P}_k = (1-\gamma_k)A\bar{P}_{k-1}A' + \bar{W}(1-\gamma_k), k \ge 0, \bar{P}_0 = P_0\}$ and $\{\underline{P}_k|\underline{P}_k = (1-\gamma_k)A\underline{P}_{k-1}A' + \underline{W}(1-\gamma_k), k \ge 0, \bar{P}_0 = P_0\}$, where $\bar{W}(\cdot)$ and $\underline{W}(\cdot)$ are defined in Definition 1. Below we show that these sequences are upper and lower bounds of $\{P_k\}$, and both of them converge to their unique steady states in mean sense. Therefore, these two sequences facilitate us to study the behavior of $\mathbb{E}[P_k]$ over an infinite horizon.

***Definition** 1:* $\forall q \in [0, \frac{1}{\rho(A)^2})$, let us define the following functions: $\mathcal{G}_A(q) = \sum_{k=0}^{\infty} q^k A'^k A^k$, $\bar{W}(q) = (1-q)A(C'R^{-1}C)^{-1}A' + Q$, and $\underline{W}(q) = (1-q)A(Q^{-1} + C'R^{-1}C)^{-1}A' + Q$, where the first function presents the unique solution of Lyapunov equation[3] $X = qA'XA + I$. The rest two functions are useful in determining upper and lower performance bounds.

***Theorem** 5:* $\forall k \ge 0$, $\underline{P}_k \le P_k \le \bar{P}_k$ and $\mathbb{E}[\underline{P}_k] \le \mathbb{E}[P_k] \le \mathbb{E}[\bar{P}_k]$. Moreover,

$$\begin{aligned}
&\lim_{k\to\infty} \mathrm{Tr}(\mathbb{E}[\underline{P}_k]) = \mathrm{Tr}(\mathbb{E}[\underline{P}_\infty]) \\
=&\delta\mathrm{Tr}(\underline{W}(\sigma_1)\mathcal{G}_A(\sigma_1)) + (1-\delta)\mathrm{Tr}(\underline{W}(\sigma_2)\mathcal{G}_A(\sigma_2)) \quad (9)\\
&\lim_{k\to\infty} \mathrm{Tr}(\mathbb{E}[\bar{P}_k]) = \mathrm{Tr}(\mathbb{E}[\bar{P}_\infty]) \\
=&\delta\mathrm{Tr}(\bar{W}(\sigma_1)\mathcal{G}_A(\sigma_1)) + (1-\delta)\mathrm{Tr}(\bar{W}(\sigma_2)\mathcal{G}_A(\sigma_2)) \quad (10)
\end{aligned}$$

where $\delta = \frac{\sigma_2 - \bar{\gamma}}{\sigma_2 - \sigma_1} < 1$ and $\bar{\gamma} = \frac{1}{\alpha+\beta}(\beta\psi_1 + \alpha\psi_2)$. In case $\rho(A) > 1$, both bounds $\mathrm{Tr}(\mathbb{E}[\underline{P}_\infty])$ and $\mathrm{Tr}(\mathbb{E}[\bar{P}_\infty])$ decrease monotonically as the correct detection probability p_d increases.

Theorem 5 shows that the bounds relate to many parameters involved in the physical process model, the channel model and the channel sensing mechanism. It also gives a useful guidance that increasing the correct detection probability is always beneficial. Intuitively, increasing p_d means the sensor can accurately seize more opportunities to utilize the idle portion of CH_1. Hence, the packet loss probability can be reduced (since $\ell_1 < \ell_0$) and the estimation performance can be improved. Based on this theorem, we are able to further quantitatively study how much the performance is improved by using CHANCE. With the above bounds, we study the worst case performance ratio $\varphi_\infty \triangleq \frac{\mathrm{Tr}(\mathbb{E}[\bar{P}_\infty])}{\mathrm{Tr}(\mathbb{E}[\underline{P}^\dagger_\infty])}$.

***Theorem** 6:* If $\sigma_2 < \ell_0$ and $P_0^\dagger = P_0$, $\mathrm{Tr}(\mathbb{E}[\underline{P}_\infty]) < \mathrm{Tr}(\mathbb{E}[\underline{P}^\dagger_\infty])$, $\mathrm{Tr}(\mathbb{E}[\bar{P}_\infty]) < \mathrm{Tr}(\mathbb{E}[\bar{P}^\dagger_\infty])$, and the worst case performance ratio is

$$\varphi_\infty = \delta\frac{\mathrm{Tr}\left(\bar{W}(\sigma_1)\mathcal{G}_A(\sigma_1)\right)}{\mathrm{Tr}\left(\underline{W}(\ell_0)\mathcal{G}_A(\ell_0)\right)} + (1-\delta)\frac{\mathrm{Tr}\left(\bar{W}(\sigma_2)\mathcal{G}_A(\sigma_2)\right)}{\mathrm{Tr}\left(\underline{W}(\ell_0)\mathcal{G}_A(\ell_0)\right)}. \tag{11}$$

Moreover, φ_∞ is a monotonic decreasing function of p_d.

The bounds in (9) and (10) are generally conservative mainly because of the gap between $\bar{W}(\cdot)$ and $\underline{W}(\cdot)$ due to the difference between $A(C'R^{-1}C)^{-1}A'$ and $A(Q^{-1} + C'R^{-1}C)^{-1}A'$. However, since $A[(C'R^{-1}C)^{-1} - (Q^{-1} + C'R^{-1}C)^{-1}]A' = A(C'R^{-1}C)^{-1}(I + C'R^{-1}CQ)^{-1}A'$, if either the measurement noise covariance R is sufficiently small or the system noise covariance Q is large, the above gap shrinks and thus the bounds become tight. In the extreme case where R is zero, $\bar{W}(\cdot) = \underline{W}(\cdot) = Q$. Thus, the two bounds coincide and the performance ratio can be determined. To illustrate this, consider a class of special scalar systems with $R = 0$. Both P_k and $P_k^\dagger$ converge, and if $P_0^\dagger = P_0$, the performance ratio is $\varphi_\infty = \frac{P_\infty}{P^\dagger_\infty} = \delta\frac{1-\ell_0A^2}{1-\sigma_1A^2} + (1-\delta)\frac{1-\ell_0A^2}{1-\sigma_2A^2} \le \frac{1-\ell_0A^2}{1-\sigma_2A^2}$ which is clearly less than 1 if $\sigma_2 < \ell_0$.

IV. Extension to Multi-Licensed-Channel Case

In this section, we consider that the sensor can opportunistically access multiple licensed channels. We first update the critical matrix, based on which the stability conditions and estimation performance analysis are then updated.

[3]Let $\tilde{A} = \sqrt{q}A$, we get the standard form of Lyapunov equation. Since $\rho(\tilde{A}) < 1$, the finiteness and uniqueness of X is guaranteed.

A. The Critical Matrix $\Phi\Psi$

Define mapping $\mathcal{B}_m(\cdot)$ as follows. $\forall b \in \{1,\ldots,2^m\}$, there exists a unique binary vector $\mathcal{B}_m(b) = [b_1,\ldots,b_m]'$ such that $b = \sum_{j=1}^{m} 2^{j-1} b_j + 1$. In other words, $\mathcal{B}_m(b)$ is the binary representation of $b-1$. Obviously, the inverse mapping $\mathcal{B}_m^{-1}(\cdot)$ also exists and is unique. For the m licensed channels, define the channel state vector $\mathbf{s}_k = [s_{1,k},\ldots,s_{m,k}]'$. Since each channel has two states, $\mathbf{s}_k$ has totally 2^m possibilities, each of which corresponds to a unique number in $\{1,\ldots,2^m\}$ via the mapping $\mathcal{B}_m^{-1}(\mathbf{s}_k)$. Define the channel state probability vector $\mathbf{p}_k = [p_{1,k},\ldots,p_{2^m,k}]'$, where $p_{b,k} = \mathbb{P}\{\mathbf{s}_k = \mathcal{B}_m(b)\} = \prod_{i=1}^{m} \mathbb{P}\{s_{i,k} = [\mathcal{B}_m(b)]_i\}$, $\forall b \in \{1,\ldots,2^m\}$. Obviously, $\sum_{b=1}^{2^m} p_{b,k} = 1$. Based on $\mathbf{p}_k$, we can define the new transition probability matrix Φ whose dimension is $2^m \times 2^m$. Due to the hypothesis of inter-channel independencies, each entry of Φ is

$$
\begin{aligned}
[\Phi]_{ij} &\triangleq \mathbb{P}\big\{\mathbf{s}_{k+1} = \mathcal{B}_m(j) | \mathbf{s}_k = \mathcal{B}_m(i)\big\} \\
&= \prod_{l=1}^{m} \mathbb{P}\Big\{s_{l,k+1} = [\mathcal{B}_m(j)]_l | s_{l,k} = [\mathcal{B}_m(i)]_l\Big\}, \qquad (12)
\end{aligned}
$$

where $i,j \in \{1,\ldots,2^m\}$. Thus, one can easily prove that

$$
\Phi = \Phi_m \otimes \Phi_{m-1} \otimes \cdots \otimes \Phi_1 \triangleq \bigotimes_{i=1}^{m} \Phi_i, \qquad (13)
$$

where "$\otimes$" represents the Kronecker product.

The sensing matrix Ψ is in the form $Diag\{\psi_1,\ldots,\psi_{2^m}\}$, where $\forall b \in \{1,\ldots,2^m\}$, $\psi_b \triangleq \mathbb{P}\{\gamma_k = 0 | \mathbf{s}_k = \mathcal{B}_m(b)\}$. Denote ℓ_i^s as the packet loss rate on CH_i when it is in state s: $\ell_i^s = 1$ if $s = 1$ and $\ell_i^s = \ell_i$ otherwise. Thus,

$$
\begin{aligned}
&\psi_b = \mathbb{P}\left\{\gamma_k = 0 | [s_{1,k},\ldots,s_{m,k}]' = \mathcal{B}_m(b)\right\} \\
=&\mathbb{P}\{o_{1,k} = 0 | s_{1,k} = b_1\}\ell_1^{b_1} + \mathbb{P}\{o_{1,k} = 1 | s_{1,k} = b_1\} \\
&\times \mathbb{P}\{o_{2,k} = 0 | s_{2,k} = b_2\}\ell_2^{b_2} + \cdots + \mathbb{P}\{o_{1,k} = 1 | s_{1,k} = b_1\} \\
&\times \cdots \times \mathbb{P}\{o_{m-1,k} = 1 | s_{m-1,k} = b_{m-1}\} \\
&\times \mathbb{P}\{o_{m,k} = 0 | s_{m,k} = b_m\}\ell_m^{b_m} \\
&+ \mathbb{P}\{o_{1,k} = 1 | s_{1,k} = b_1\} \cdots \mathbb{P}\{o_{m,k} = 1 | s_{m,k} = b_m\}\ell_0 \\
=& \sum_{i=1}^{m}\left(\prod_{j=1}^{i-1}\mathbb{P}\{o_{j,k} = 1 | s_{j,k} = b_j\}\right)\mathbb{P}\{o_{i,k} = 0 | s_{i,k} = b_i\}\ell_i^{b_i} \\
&+ \ell_0 \prod_{i=1}^{m}\mathbb{P}\{o_{i,k} = 1 | s_{i,k} = b_i\} \qquad (14)
\end{aligned}
$$

where the second equality is based on the definition of sensing sequence $\mathcal{Q}_m$. $\mathbb{P}\{o_{i,k} = 0 | s_{i,k} = b_i\}$ equals to $p_{d,i}$ if $b_i = 0$ and to $p_{f,i}$ if $b_i = 1$. $\mathbb{P}\{o_{j,k} = 1 | s_{j,k} = b_j\} = 1 - \mathbb{P}\{o_{j,k} = 0 | s_{j,k} = b_j\}$. Thus, Ψ and the critical matrix are obtained.

B. Stability Analysis

***Theorem** 7:* For system (1) with the above channel sensing schedule $\mathcal{Q}_m$, a necessary condition for the MSE stability of the estimation process applied with CHANCE is

$$
\rho\left(\Phi\Psi\right)\rho(A)^2 < 1. \qquad (15)
$$

Moreover, (15) is also sufficient if C has full column rank.

In the following, we assume that C has full column rank for simplicity. As mentioned before, extensions to a special class of non-degenerate systems are possible but algebraically complicated and thus are omitted due to space limitations.

C. Performance Analysis

***Theorem** 8:* Let $\ell_{\min} = \min\{\ell_0,\ell_1,\ldots,\ell_m\}$. With the sensing schedule $\mathcal{Q}_m$, the stability gain η satisfies

$$
\eta = \sqrt{\frac{\ell_0}{\rho\left(\Phi\Psi\right)}} \le \min\left\{\frac{1}{\sqrt{\prod_{i=1}^{m}(1-\beta_i)}}, \sqrt{\frac{\ell_0}{\ell_{\min}}}\right\}, \qquad (16)
$$

Let $\tilde{\ell}^* \triangleq \max\{\mathbf{e}_i'\Phi\Psi\mathbf{u}, \forall i \in \{1,\ldots,2^m\}\}$, where $\mathbf{e}_i$ is a $2^m \times 1$ vector with $[\mathbf{e}_i]_i = 1$ and $[\mathbf{e}_i]_j = 0$ for all $j \neq i$, $\mathbf{u} = [1,\ldots,1]'_{1\times 2^m}$. Similar to Theorem 4, the following theorem establishes the relationship between $\mathrm{Tr}(\mathbb{E}[P_k])$ and $\mathrm{Tr}(\mathbb{E}[P_k^\dagger])$ for multi-licensed-channel cases. We focus on the nontrivial cases that $\ell_0 < \frac{1}{\rho(A)^2}$ and (15) are satisfied.

***Theorem** 9:* If $\tilde{\ell}^* \le \ell_0$ and the initial conditions satisfy $P_0^\dagger = P_0$, then $\forall k \ge 0$, $\mathrm{Tr}(\mathbb{E}[P_k]) \le \mathrm{Tr}(\mathbb{E}[P_k^\dagger])$.

***Theorem** 10:* $\forall k \ge 0$, $\underline{P}_k \le P_k \le \bar{P}_k$ and $\mathbb{E}[\underline{P}_k] \le \mathbb{E}[P_k] \le \mathbb{E}[\bar{P}_k]$ where $\{\underline{P}_k\}$ and $\{\bar{P}_k\}$ are defined above Theorem 5. Moreover,

$$
\mathrm{Tr}(\mathbb{E}[\underline{P}_\infty]) = \sum_{i=0}^{\infty}\left[\boldsymbol{v}'\left(\Phi\Psi\right)^i \mathbf{u}\mathrm{Tr}\left(\underline{W}(\varrho)A'^iA^i\right)\right], \qquad (17)
$$

$$
\mathrm{Tr}(\mathbb{E}[\bar{P}_\infty]) = \sum_{i=0}^{\infty}\left[\boldsymbol{v}'\left(\Phi\Psi\right)^i \mathbf{u}\mathrm{Tr}\left(\bar{W}(\varrho)A'^iA^i\right)\right], \qquad (18)
$$

where $\boldsymbol{v}$ is the stationary state of the vector $\mathbf{p}_k$, $\varrho = 1 - \boldsymbol{v}'(I - \Phi\Psi)\mathbf{u}$. If $P_0^\dagger = P_0$,

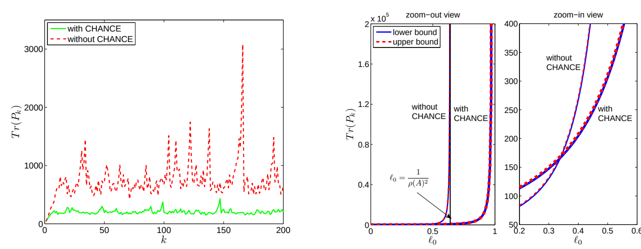
$$
\varphi_\infty \le \frac{\mathrm{Tr}\left(\bar{W}(\varrho)\mathcal{G}_A\left(\rho(\Phi\Psi)\right)\right)}{\mathrm{Tr}\left(\underline{W}(\ell_0)\mathcal{G}_A(\ell_0)\right)}. \qquad (19)
$$

Remark 5: We focus on the scenario with one unlicensed channel and multiple licensed channels; however, our results are applicable to a wide range of scenarios. For example, **(1)** CH_0 **is Markovian**: consider another channel sensing order which differs from $\mathcal{Q}_m$ in that the sensor will directly transmit packet through CH_m if CH_{m-1} is sensed busy (i.e., $p_{f,m} = p_{d,m} = 1$). If we deem CH_m as the original channel, Theorem 7 is directly applicable to that the original channel has Markovian packet losses. Moreover, suppose $m = 1$ and deem CH_m as the original channel; the scenario that the original channel is Markovian and CHANCE is not applied can be viewed as a special single-licensed-channel case. Then, the estimation performance can be analyzed in the way shown in Section III, and the results can be used to update Theorems 8-10. **(2) Sensing multiple unlicensed channels**: if the state dynamics of the unlicensed channels can be modeled by homogeneous Markov chains as shown in Proposition 1, one can see that our results are seamlessly applicable. In particular, an i.i.d. channel model is a special Markov model with $\alpha + \beta = 1$; in this case, it can be easily seen that the proofs of our theorems are still valid, and hence the theorems hold when multiple i.i.d. unlicensed channels are sensed by CHANCE.

V. Simulations

In this section, we evaluate the performances of the proposed mechanism based on a CPS with the physical process described by (1). We adopt the same parameters as in [13]: $A = \begin{bmatrix} 1.25 & 0 \\ 1 & 1.1 \end{bmatrix}$ and $Q = 20I_{2\times 2}$. The sensor measurements are described by (2), where $C = I_{2\times 2}$, $R = 2.5 I_{2\times 2}$, and the sampling period $T = 1$. The average loss rate on the original channel CH_0 is $\ell_0 = 0.5$. On licensed channel CH_1, the average busy and idle periods are $\frac{1}{\omega_{i,1}} = 0.35$ and $\frac{1}{\omega_{i,0}} = 0.3$, respectively. By Lemma 2, we have $\alpha = 0.537$ and $\beta = 0.461$. The measurement packets transmitted through CH_1 during its idle state will experience a loss rate $\ell_1 = 0.05$. The correct and false detection probabilities are $p_d = 0.8$ and $p_f = 0.3$, respectively. Thus, the largest eigenvalue of the critical matrix $\Phi\Psi$ is $\sigma_2 \approx 0.415 < \ell_0 < \frac{1}{\rho(A)^2}$. Hence, the Kalman filter based state estimation is stable in mean square sense either with or without CHANCE. However, the two cases achieve different performance in terms of $\mathrm{Tr}(\mathbb{E}[P_k])$. As shown in Fig. 3(a), CHANCE achieves much better estimation performance.

Fig. 3 shows the performance bounds obtained in Theorem 5. First, we observe that the bounds are tight in our simulation case. Without CHANCE, it is well known that the estimation error covariance diverges if ℓ_0 is larger than $\frac{1}{\rho(A)^2} = 0.64$. In contrast, with CH_1 and CHANCE, the critical value increases significantly, i.e., the requirement for estimation stability is relaxed. From Fig. 3(b) we observe that CHANCE improves the estimation performance when $\ell_0 \geq 0.34$. In fact, when $\ell_0 \geq 0.34$, both $\sigma_2 < \ell_0$ and $\psi_2(1-\beta) + \psi_1\beta \leq \ell_0$ are true. Hence, by Theorem 4 and Theorem 6, the estimation performance is guaranteed to be improved. However, the figure also shows that there is a performance degradation when $\ell_0 < 0.34$ such that the quality of CH_0 is much better than that of CH_1 (the packet loss rate on CH_0 is lower than that of jointly using CH_0 and CH_1). This can be solved by improving the channel sensing accuracy as implied by Theorem 6.

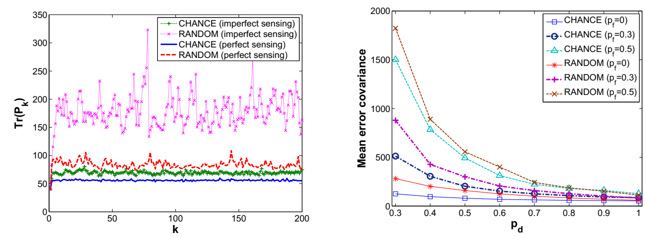


(a) Performance comparison (b) Performance bound comparison

Fig. 3. Performance comparison between with and without CHANCE.

More licensed channels will provide higher opportunities for the sensor to successfully transmit its measurement packets. To illustrate this, we introduce m licensed channels to the system and each of them has the same property as CH_1 mentioned above. The correct and false detection probabilities over them are also set the same. The performance bounds and performance ratio obtained in Theorem 10 are shown in Table II, where $m = 0$ indicates the case without CHANCE. We can observe that: 1) with CHANCE, the upper bound and the relaxed upper bounds are quite close in all cases. 2) The worst case performance ratio is less than 36.68%; in other words, the mean square estimation error is reduced by at least 63.32%. 3) The performance is obviously improved by introducing to CHANCE more licensed channels. Whereas, such performance improvement becomes less significant as m increases. Therefore, if the cost (e.g., energy expenditure) for channel sensing and switching is a concern, the sensor only needs to sense a few channels before each transmission.

We also conduct simulations to compare CHANCE and the mechanism (which we call as RANDOM) proposed in [22]. With RANDOM, in each step, the sensor chooses the same channel to sense if the channel is found idle in previous step; otherwise, it randomly chooses a different licensed channel to sense. To be comparable with CHANCE, RANDOM is enhanced by adding the original channel in the way that the sensor transmits data on the current channel if it is sensed idle and on the original channel otherwise. From Fig. 4 (where $m = 3$), we can see that CHANCE outperforms RANDOM. This is because, in each step, RANDOM senses only one channel while CHANCE may sense more; hence CHANCE gets more opportunity for transmitting during the channels' idle periods. Moreover, as expected, the estimation error is reduced monotonically as the correct detection probability increases in both mechanisms, which is shown in Fig. 4(b).



(a) Estimation error covariance. (b) Impact of sensing accuracy.

Fig. 4. Performance comparison between CHANCE and RANDOM [22].

VI. Conclusion

We have studied the state estimation problem in a class of CPS with linear process state dynamics. Based on the cognitive radio technology, we propose the CHANCE mechanism for the sensor to opportunistically access licensed spectrum for data transmissions. We develop new necessary and sufficient conditions for the estimation stability in mean square sense. For a special class of systems, we show that the estimation stability region is expanded by CHANCE. We also obtain a pair of upper and lower bounds for the estimation performance, and prove that they are monotonically decreasing as the correct detection probability increases. Explicit mathematical expression for the worst case performance ratio is derived. We also extend our results to accommodate any number of licensed channels. Simulation results show that the estimation performance is dramatically improved by CHANCE.

Appendix

A. Proof of Proposition 1

Due to the Markov property of the PU's activities on CH_i, the sequence $\{s_{i,k}\}_{k\geq 0}$ is a homogeneous Markov chain, and the steady distributions are $\mathbb{P}\{s_{i,k} = 1\} = \frac{\omega_{i,1}}{\omega_{i,1}+\omega_{i,0}}$ and $\mathbb{P}\{s_{i,k} = 0\} = \frac{\omega_{i,0}}{\omega_{i,1}+\omega_{i,0}}$, respectively. Consider the continuous-time channel state $s_i(t)$. According to Kolmogorov's Backward Equations, $\frac{\partial \Phi_i}{\partial t}(t) = G\Phi_i(t)$, where $[\Phi_i(t)]_{jl} \triangleq \mathbb{P}\{s_i(t) = l | s_i(0) = j\}$, and the generator matrix is $G = \begin{bmatrix} -\omega_{i,0} & \omega_{i,0} \\ \omega_{i,1} & -\omega_{i,1} \end{bmatrix}$ according to [35]. Then, $\Phi_i = \Phi_i(T) = e^{GT}$, which finally yields (5).

Table II
PERFORMANCE BOUNDS AND WORST-CASE RATIO.

Number of licensed channels m	0	1	2	3	4	5	6	7	8
Lower bound $\mathrm{Tr}(\mathbb{E}[\underline{P}_\infty])$	847.8	306.27	215.93	187.10	175.64	170.65	168.39	167.35	166.86
Upper bound $\mathrm{Tr}(\mathbb{E}[\overline{P}_\infty])$	857.89	310.78	219.37	190.18	178.57	173.52	171.23	170.17	169.68
Worst case performance ratio by (19)	100%	36.68%	25.90%	22.45%	21.08%	20.48%	20.21%	20.09%	20.03%

B. Proof of Theorem 1

We first introduce several useful lemmas.

Lemma 1: $\sigma_2 = \rho(\Phi\Psi)$ and $\sigma_{1,2} = \frac{1}{2}[\psi_1(1-\alpha) + \psi_2(1-\beta) \mp \sqrt{[\psi_1(1-\alpha) - \psi_2(1-\beta)]^2 + 4\psi_1\psi_2\alpha\beta}]$. Moreover, both σ_1 and σ_2 are monotonically increasing as either ψ_1 or ψ_2 increases. If $\ell_1 \leq \ell_0$, $\ell_1 \leq \psi_1 \leq \ell_0 \leq \psi_2 \leq 1$.

The following lemma models the packet loss process under CHANCE. $\forall k_1 > 0$ and $\forall k_2 \geq k_1$, denote successive packet losses as $\gamma(k_1, k_2) = 0$ which means $\{\gamma_{k_1} = 0, \gamma_{k_1+1} = 0, \ldots, \gamma_{k_2} = 0\}$. Define matrix M_k with each entry $[M_k]_{ij} \triangleq \mathbb{P}\{s_k = j-1, \gamma(1,k) = 0 | s_0 = i-1\}$.

Lemma 2: $M_k = (\Phi\Psi)^k$, $\forall k \geq 0$. Moreover, $\forall k_1 > 0$ and $\forall k_2 \geq k_1$, $\exists \varepsilon_{1,k_1} \in [-\bar{\varepsilon}_1, \bar{\varepsilon}_1]$ and $\varepsilon_{2,k_1} \in [\underline{\varepsilon}_2, \bar{\varepsilon}_2]$ such that the successive packet loss probability is $\mathbb{P}\{\gamma(k_1,k_2) = 0\} = \varepsilon_{1,k_1}\sigma_1^{k_2+1-k_1} + \varepsilon_{2,k_1}\sigma_2^{k_2+1-k_1}$, where $\bar{\varepsilon}_1, \underline{\varepsilon}_2$ and $\bar{\varepsilon}_2$ are all finite positive numbers.

Proof: $\forall k > 0$, from the definition of matrix M_k,

$$
\begin{aligned}
&[M_k]_{11} = \mathbb{P}\left\{s_k = 0, s_{k-1} = 0, \gamma(1,k) = 0 | s_0 = 0\right\} \\
&\quad + \mathbb{P}\left\{s_k = 0, s_{k-1} = 1, \gamma(1,k) = 0 | s_0 = 0\right\} \\
&= \mathbb{P}\left\{\gamma_k = 0, s_k = 0 \middle| s_{k-1} = 0, \gamma(1,k-1) = 0, s_0 = 0\right\} \\
&\quad \times [M_{k-1}]_{11} + [M_{k-1}]_{12} \\
&\quad \times \mathbb{P}\left\{\gamma_k = 0, s_k = 0 \middle| s_{k-1} = 1, \gamma(1,k-1) = 0, s_0 = 0\right\} \\
&= \mathbb{P}\left\{\gamma_k = 0, s_k = 0 \middle| s_{k-1} = 0\right\}[M_{k-1}]_{11} \\
&\quad + \mathbb{P}\left\{\gamma_k = 0, s_k = 0 \middle| s_{k-1} = 1\right\}[M_{k-1}]_{12} \\
&= \psi_1[\Phi]_{11}[M_{k-1}]_{11} + \psi_1[\Phi]_{21}[M_{k-1}]_{12}. \qquad (20)
\end{aligned}
$$

where the third equality is because of the Markov property of the channel state. The other three entries of M_k can be similarly obtained. Then, we have $M_k = M_{k-1}\Phi\Psi$. By definition, $M_0 = I$. Hence, $M_k = (\Phi\Psi)^k$. Moreover, there exists an invertible matrix F_σ such that $\Phi\Psi = F_\sigma \Lambda_\sigma F_\sigma^{-1}$, where $\Lambda_\sigma = Diag\{\sigma_1, \sigma_2\}$. Therefore, $M_k = F_\sigma \Lambda_\sigma^k F_\sigma^{-1}$. Define the channel state probability vector $\mathbf{p}_k = [p_{0,k}, p_{1,k}]'$, where $p_{b,k} = \mathbb{P}\{s_k = b | b \in \{0,1\}\}$. Then, the successive packet loss probability can be written as $\mathbb{P}\{\gamma(k_1,k_2) = 0\} = \mathbb{P}\{\gamma(k_1,k_2) = 0 | s_{k_1-1} = 1\}p_{1,k_1-1} + \mathbb{P}\{\gamma(k_1,k_2) = 0 | s_{k_1-1} = 0\}p_{0,k_1-1}$. Thus,

$$
\begin{aligned}
&\mathbb{P}\{\gamma(k_1,k_2) = 0\} = \mathbb{P}\{\gamma(1, k_2+1-k_1) = 0 | s_0 = 1\}p_{1,k_1-1} \\
&\qquad\qquad + \mathbb{P}\{\gamma(1, k_2+1-k_1) = 0 | s_0 = 0\}p_{0,k_1-1} \\
&= \left([M_{k_2+1-k_1}]_{21} + [M_{k_2+1-k_1}]_{22}\right)p_{1,k_1-1} \\
&\quad + \left([M_{k_2+1-k_1}]_{11} + [M_{k_2+1-k_1}]_{12}\right)p_{0,k_1-1} \\
&= \mathbf{p}'_{k_1-1}M_{k_2+1-k_1}\mathbf{u} = \mathbf{p}'_{k_1-1}F_\sigma\Lambda_\sigma^{k_2+1-k_1}F_\sigma^{-1}\mathbf{u} \qquad (21) \\
&= \frac{\psi_1 z - \sigma_1 + (\psi_2 - \psi_1)zp_{0,k_1-1}}{\sigma_2 - \sigma_1}\sigma_1^{k_2+1-k_1} \\
&\quad + \frac{\sigma_2 - \psi_1 z - (\psi_2 - \psi_1)zp_{0,k_1-1}}{\sigma_2 - \sigma_1}\sigma_2^{k_2+1-k_1} \\
&= \varepsilon_{1,k_1}\sigma_1^{k_2+1-k_1} + \varepsilon_{2,k_1}\sigma_2^{k_2+1-k_1} \qquad (22)
\end{aligned}
$$

where $\mathbf{u} = [1,1]'$ and $z = 1 - \alpha - \beta$. The boundedness of ε_{1,k_1} and ε_{2,k_1} is clear since $p_{0,k-1} \in [0,1]$. ■

Lemma 3: $\forall A \in \mathbb{R}^{n\times n}$ and $\forall k > n$, (1) there exist $\zeta_1 > 0$ and $\zeta_2 > 0$ such that $\zeta_1\rho(A)^{2k} \leq \mathrm{Tr}\left(A^k A'^k\right) < \zeta_2 k^{2n}\rho(A)^{2k}$. (2) $\forall\sigma > 0$ and $\forall S \in \mathbb{S}_Q$ where $\mathbb{S}_Q$ is defined in Section II-A, $\exists\zeta_3, \zeta_4 > 0$ such that

$$
\begin{aligned}
&\sum_{i=0}^{k-1}\mathrm{Tr}\left(\sigma^i A^i S A'^i\right) \\
&\geq \begin{cases} \zeta_3 k, & \text{if } \sigma\rho(A)^2 = 1, \\ \zeta_4\left|\sigma^k\rho(A)^{2k} - \sigma^n\rho(A)^{2n}\right|, & \text{if } \sigma\rho(A)^2 \neq 1 \end{cases} \qquad (23)
\end{aligned}
$$

Proof: (1) Let $\lambda_1, \ldots, \lambda_q$ be the distinct eigenvalues (may be complex) of A where $\rho(A) = |\lambda_1| > |\lambda_2| > \cdots > |\lambda_q|$. There exists an invertible matrix F_A such that $A = F_A J F_A^{-1}$, where J is the Jordan canonical form of A with each block $J_i \in \mathbb{C}^{n_i \times n_i}$ corresponds to λ_i. Moreover, $\forall k \geq n_i$, we have

$$
J_i^k = \begin{bmatrix} \lambda_i^k & \cdots & \binom{k}{n_i-1}\lambda_i^{k-n_i+1} \\ & \ddots & \vdots \\ & & \lambda_i^k \end{bmatrix}
$$

$$
n_i|\lambda_i|^{2k} \leq \mathrm{Tr}((J_i^k)^H J_i^k) < n_i^2 k^{2n_i}|\lambda_i|^{2(k-n_i+1)}
$$

Let $\tilde{\zeta}_1 = \lambda_{\min}((F_A)^H F_A)\lambda_{\min}(F_A^{-1}(F_A^{-1})^H) > 0$ and $\tilde{\zeta}_2 = \lambda_{\max}((F_A)^H F_A)\lambda_{\max}(F_A^{-1}(F_A^{-1})^H) > 0$. Then, $\mathrm{Tr}(A^k A'^k) = \mathrm{Tr}(F_A J^k F_A^{-1}(F_A^{-1})^H (J^k)^H (F_A)^H) \geq \tilde{\zeta}_1\sum_{i=1}^q \mathrm{Tr}((J_i^k)^H J_i^k) \geq \zeta_1\rho(A)^{2k}$ where $\zeta_1 > 0$. Since $\mathrm{Tr}\left((J_i^k)^H J_i^k\right) < \frac{n_i^2}{|\lambda_i|^{2(n_i-1)}}k^{2n_i}|\lambda_i|^{2k} < \frac{n^2}{|\lambda_q|^{2(n_i-1)}}k^{2n}|\lambda_1|^{2k}$ for $|\lambda_i| > 0$, $\mathrm{Tr}(A^k A'^k) \leq \tilde{\zeta}_2\sum_{i=1}^q \mathrm{Tr}(J_i'^k J_i^k) < \tilde{\zeta}_2 k^{2n}|\lambda_1|^{2k}\sum_{i=1}^q \frac{n^2}{|\lambda_{\bar{q}}|^{2(n_i-1)}} < \zeta_2 k^{2n}\rho(A)^{2k}$, where $\zeta_2 > 0$ and $|\lambda_{\bar{q}}| = \min\{|\lambda_i| > 0 | i = 1, \ldots, q\}$.

(2) $\Gamma_S\Gamma_S' > 0$ where $\Gamma_S = [\sqrt{S}, A\sqrt{S}, \ldots, A^{n-1}\sqrt{S}]$. $\forall k > n$, $\exists h \geq 1$ such that $k \in [hn, hn+n)$. By Lemma 3, $\sum_{i=0}^{k-1}\mathrm{Tr}(\sigma^i A^i S A'^i) \geq \sum_{j=0}^{h-1}\mathrm{Tr}(\sigma^{jn}A^{jn}\Gamma_S\Gamma_S' A'^{jn}) \geq \lambda_{\min}(\Gamma_S\Gamma_S')\zeta_1\sum_{j=0}^{h-1}\sigma^{jn}\rho(A)^{2jn}$. If $\sigma\rho(A)^2 = 1$, $\exists\zeta_3 > 0$, $\sum_{i=0}^{k-1}\mathrm{Tr}(\sigma^i A^i S A'^i) \geq h\lambda_{\min}(\Gamma_S\Gamma_S')\zeta_1 \geq \zeta_3 k$. Otherwise, if $\sigma\rho(A)^2 < 1$, $\sum_{i=0}^{k-1}\mathrm{Tr}(\sigma^i A^i S A'^i) \geq \lambda_{\min}(\Gamma_S\Gamma_S')\zeta_1\frac{1-\sigma^{hn}\rho(A)^{2hn}}{1-\sigma^n\rho(A)^{2n}} \geq \zeta_4(\sigma^n\rho(A)^{2n} - \sigma^k\rho(A)^{2k})$, where $\zeta_4 > 0$. Similarly, if $\sigma\rho(A)^2 > 1$, $\sum_{i=0}^{k-1}\mathrm{Tr}(\sigma^i A^i S A'^i) \geq \zeta_4[\sigma^k\rho(A)^{2k} - \sigma^n\rho(A)^{2n}]$. ■

Lemma 4: $\forall A \in \mathbb{R}^{n\times n}$ and $\forall S \in \mathbb{S}_Q$, under the packet loss process $\{\gamma_k\}_{k\geq 0}$ described in Section II-A, the sequence $\{X_k | X_k = (1-\gamma_k)AX_{k-1}A' + S, \forall k \geq 0, X_0 \in \mathbb{S}_Q\}$ is stable in mean sense iff. $\sigma_2\rho(A)^2 < 1$.

Proof: $\forall k > 0$, we have $X_k = (1-\gamma_k)AX_{k-1}A'] + S = \prod_{i=1}^k(1-\gamma_i)A^k X_0 A'^k + \sum_{i=1}^{k-1}\prod_{j=k+1-i}^k(1-\gamma_j)A^i S A'^i + S$. Taking expectation at both sides yields

$$
\begin{aligned}
\mathbb{E}[X_k] =& \sum_{i=1}^{k-1} \mathbb{P}\{\gamma(k+1-i,k)=0\} A^i S A'^i \\
&+ \mathbb{P}\{\gamma(1,k)=0\} A^k X_0 A'^k + S \\
=& \sum_{i=1}^{k-1} \left(\varepsilon_{1,k+1-i}\sigma_1^i + \varepsilon_{2,k+1-i}\sigma_2^i\right) A^i S A'^i \\
&+ \left(\varepsilon_{1,1}\sigma_1^k + \varepsilon_{2,1}\sigma_2^k\right) A^k X_0 A'^k + S
\end{aligned} \tag{24}
$$

1) Sufficiency: By Lemma 2, we have $\mathrm{Tr}(\mathbb{E}[X_k]) \leq \mathrm{Tr}(S) + (\bar{\varepsilon}_1 + \bar{\varepsilon}_2)[\sigma_2^k \mathrm{Tr}\left(A^k X_0 A'^k\right) + \sum_{i=1}^{k-1} \sigma_2^i \mathrm{Tr}\left(A^i S A'^i\right)] \leq \mathrm{Tr}(S) + (\bar{\varepsilon}_1 + \bar{\varepsilon}_2)[\mathrm{Tr}(S) \sum_{i=1}^{k-1} i^{2n} (\sigma_2 \rho(A)^2)^i + \mathrm{Tr}(X_0) k^{2n} (\sigma_2 \rho(A)^2)^k]$. If $\sigma_2 \rho(A)^2 < 1$, $\exists \varepsilon \in (\sigma_2\rho(A)^2, 1)$ such that $k^{2n} < \frac{\varepsilon}{\sigma_2 \rho(A)^2}(k-1)^{2n}$ holds for all $k > \underline{k} = \left\lceil \frac{1}{\sqrt[2n]{\frac{\varepsilon}{\sigma_2\rho(A)^2}} - 1} \right\rceil + 1 > 1$. Thus, $\forall k > \underline{k}$, $k^{2n}(\sigma_2\rho(A)^2)^k < \varepsilon(k-1)^{2n}(\sigma_2\rho(A)^2)^{k-1} < \cdots < \varepsilon^{k-\underline{k}} \underline{k}^{2n} (\sigma_2\rho(A)^2)^{\underline{k}} < \infty$ and $\sum_{i=1}^{k-1} i^{2n}(\sigma_2\rho(A)^2)^i < \sum_{i=1}^{\underline{k}} i^{2n}(\sigma_2\rho(A)^2)^i + \underline{k}^{2n}(\sigma_2\rho(A)^2)^{\underline{k}} \sum_{i=\underline{k}+1}^{k-1} \varepsilon^{i-\underline{k}} < \sum_{i=1}^{\underline{k}} i^{2n}(\sigma_2\rho(A)^2)^i + \underline{k}^{2n}(\sigma_2\rho(A)^2)^{\underline{k}} \frac{\varepsilon}{1-\varepsilon} < \infty$. Therefore, $\sigma_2\rho(A)^2 < 1$ is a sufficient condition for $\sup_{k\geq 0} \mathrm{Tr}(\mathbb{E}[X_k]) < \infty$.

2) Necessity: Above we know that $\varepsilon_{2,k} \geq \underline{\varepsilon}_2 > 0$ and also $\sigma_2 > |\sigma_1| \geq 0$. Hence $\exists \underline{k}' > 0$ and $\varepsilon' > 0$ such that $\forall k > \underline{k}'$, $-\bar{\varepsilon}_1(\frac{\sigma_1}{\sigma_2})^k + \underline{\varepsilon}_2 > \varepsilon'$. Hence, $\mathrm{Tr}(\mathbb{E}[X_k]) \geq \sum_{i=1}^{k-1}(\varepsilon_{1,k+1-i}\sigma_1^i \varepsilon_{2,k+1-i}\sigma_2^i)\mathrm{Tr}(A^i S A'^i) \geq \sum_{i=1}^{k-1}(-\bar{\varepsilon}_1\sigma_1^i + \underline{\varepsilon}_2\sigma_2^i)\mathrm{Tr}(A^i S A'^i) > \underline{\Sigma} + \varepsilon' \sum_{i=0}^{k-1} \mathrm{Tr}(\sigma_2^i A^i S A'^i)$, where $\underline{\Sigma} = \sum_{i=1}^{\underline{k}'-1}\left(-\bar{\varepsilon}_1\sigma_1^i + \underline{\varepsilon}_2\sigma_2^i - \varepsilon'\sigma_2^i\right) \mathrm{Tr}\left(A^i S A'^i\right) - \varepsilon' \mathrm{Tr}(S)$. $\sup_{k\geq 0} \mathrm{Tr}(\mathbb{E}[X_k]) > \varepsilon' \zeta_3 \sup_{k\geq 0} k + \underline{\Sigma} = \infty$ if $\sigma_2\rho(A)^2 = 1$. Otherwise, $\sup_{k\geq 0} \mathrm{Tr}(\mathbb{E}[X_k]) > \underline{\Sigma} + \varepsilon'\zeta_4 \sup_{k\geq 0} \left|\sigma_2^k \rho(A)^{2k} - \sigma_2^n \rho(A)^{2n}\right\lceil$. Apparent, the finiteness of the righthand side of the above inequality requires $\sigma_2\rho(A)^2 < 1$. Therefore, $\sigma_2\rho(A)^2 < 1$ is a necessary condition for $\sup_{k\geq 0} \mathrm{Tr}(\mathbb{E}[X_k]) < \infty$. ■

The predicted error covariance P_k is lower bounded by [13]

$$
P_k \geq (1-\gamma_k) A P_{k-1} A' + Q. \tag{25}
$$

Then, based on Lemma 4, Theorem 1 can be proved.

C. Proof of Theorem 2

First, it is easy to verify that $(P_{k-1}^{-1} + C'R^{-1}C)^{-1}C'R^{-1} = P_{k-1}C'(CP_{k-1}C' + R)^{-1}$. Then, after simple matrix manipulations, P_k can be written as $(1-\gamma_k)AP_{k-1}A' + \gamma_k A(P_{k-1}^{-1} + C'R^{-1}C)^{-1}A' + Q$. If C has full column rank, $(P_k^{-1} + C'R^{-1}C)^{-1} \leq (C'R^{-1}C)^{-1} < \infty$ since $P_k > 0$. Therefore,

$$
P_k \leq (1-\gamma_k) A P_{k-1} A' + \widetilde{W} \tag{26}
$$

with $\widetilde{W} = A(C'R^{-1}C)^{-1}A' + Q$. By Lemma 4, $\sigma_2\rho(A)^2 < 1$ is a sufficient condition for MSE stability of the Kalman filter. Combined with Theorem 1, the proof is thus completed.

D. Proof of Theorem 4

Define following two maps $\mathcal{L}$, $\mathcal{R}$: $\mathcal{S}_+^n \to \mathcal{S}_+^n$. $\mathcal{L}(X) = AXA' + Q$ and $\mathcal{R}(X) = AXA' + Q - AXC'(CXC' + R)^{-1}CXA'$. $\forall \phi \geq 0$, by definitions of $\psi_{1,2}, \alpha$ and β,

$$
\begin{aligned}
&\mathbb{P}\{P_{k+1} > \phi I\} \\
=&\mathbb{P}\{\mathcal{L}(P_k) > \phi I, s_k = 0\}\psi_1 + \mathbb{P}\{\mathcal{L}(P_k) > \phi I, s_k = 1\}\psi_2 \\
&+ \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_k = 0\}(1-\psi_1) \\
&+ \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_k = 1\}(1-\psi_2) \\
=&\mathbb{P}\{\mathcal{L}(P_k) > \phi I, s_{k-1} = 0\}[\psi_1(1-\alpha) + \psi_2\alpha] \\
&+ \mathbb{P}\{\mathcal{L}(P_k) > \phi I, s_{k-1} = 1\}[\psi_2(1-\beta) + \psi_1\beta] \\
&+ \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 0\}[(1-\psi_1)(1-\alpha) + (1-\psi_2)\alpha] \\
&+ \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 1\}[(1-\psi_2)(1-\beta) + (1-\psi_1)\beta] \\
=&\mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 0\} + \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 1\} \\
&+ [\psi_1(1-\alpha) + \psi_2\alpha]\{\mathbb{P}\{\mathcal{L}(P_k) > \phi I, s_{k-1} = 0\} \\
&- \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 0\}\} \\
&+ [\psi_2(1-\beta) + \psi_1\beta]\{\mathbb{P}\{\mathcal{L}(P_k) > \phi I, s_{k-1} = 1\} \\
&- \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 1\}\},
\end{aligned}
$$

Since $\alpha + \beta < 1$ (by Remark 1) and $\psi_1 \leq \psi_2$ (by Lemma 1), $\psi_1(1-\alpha) + \psi_2\alpha = \psi_1 + (\psi_2 - \psi_1)\alpha \leq \psi_1 + (\psi_2 - \psi_1)(1-\beta) = \psi_2(1-\beta) + \psi_1\beta \leq \ell_0$. Also noticing that, $\forall X \in \mathcal{S}_+^n$, $\mathcal{L}(X) \geq \mathcal{R}(X)$, we have

$$
\begin{aligned}
&\mathbb{P}\{P_{k+1} > \phi I\} \\
\leq&\mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 0\} + \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 1\} \\
&+ \ell_0\{\mathbb{P}\{\mathcal{L}(P_k) > \phi I, s_{k-1} = 0\} \\
&- \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 0\}\} \\
&+ \ell_0\{\mathbb{P}\{\mathcal{L}(P_k) > \phi I, s_{k-1} = 1\} \\
&- \mathbb{P}\{\mathcal{R}(P_k) > \phi I, s_{k-1} = 1\}\} \\
=&\ell_0\mathbb{P}\{\mathcal{L}(P_k) > \phi I\} + (1-\ell_0)\mathbb{P}\{\mathcal{R}(P_k) > \phi I\},
\end{aligned} \tag{27}
$$

We now apply an induction argument to show that $\mathbb{P}\{P_k > \phi I\} \leq \mathbb{P}\{P_k^\dagger > \phi I\}$. Clearly, it holds for $k = 0$. Suppose it holds for $k > 0$, then, $\mathbb{P}\{P_{k+1} > \phi I\} \leq \ell_0 \mathbb{P}\{\mathcal{L}(P_k^\dagger) > \phi I\} + (1-\ell_0)\mathbb{P}\{\mathcal{R}(P_k^\dagger) > \phi I\} = \mathbb{P}\{P_{k+1}^\dagger > \phi I\}$, where the inequality holds due to the monotonicity of both $\mathcal{L}(\cdot)$ and $\mathcal{R}(\cdot)$ (see Proposition 1 in [27]). Therefore, $\mathbb{P}\{P_k > \phi I\} \leq \mathbb{P}\{P_k^\dagger > \phi I\}$ and hence $\mathbb{P}\{\mathrm{Tr}(P_k) > \phi\} \leq \mathbb{P}\{\mathrm{Tr}(P_k^\dagger) > \phi\}$ hold for all $k \geq 0$. Furthermore, according to [27], $\mathbb{E}[\mathrm{Tr}(P_k)] = \int_0^\infty \mathbb{P}\{\mathrm{Tr}(P_k) > \phi\} d\phi$. It is not difficult to see that $\mathbb{E}[\mathrm{Tr}(P_k)] \leq \mathbb{E}[\mathrm{Tr}(P_k^\dagger)]$ which proves the first part of this theorem. For the second part, if $R > 0, Q > 0$, we have $P_k > 0$ and $\mathcal{L}(P_k) > \mathcal{R}(P_k)$. Combined with that $\psi_2(1-\beta) + \psi_1\beta < \ell_0$, the inequality in (27) becomes strict, which finally leads to that $\mathbb{E}[\mathrm{Tr}(P_k)] < \mathbb{E}[\mathrm{Tr}(P_k^\dagger)]$.

E. Proof of Theorem 5

We first introduce two lemmas, following which the proof of Theorem 5 is straightforward. $\forall W \in \mathbb{S}_+$ and $\forall q \in [0, \frac{1}{\rho(A)^2})$, define $\mathcal{H}_A(q, W) = \sum_{k=0}^\infty q^k A^k W (A^k)'$. Obviously, $\mathcal{G}_A(q) = \mathcal{H}_{A'}(q, I)$, where $\mathcal{G}_A(q)$ is defined in Definition 1. $\mathcal{H}_A(q, W)$ is the unique solution of Lyapunov equation $X = qAXA' + W$. $\forall z \in (0,1)$, define another series $\mathcal{Z}_k(z, q, W) = \sum_{i=0}^k z^{k+1-i} q^i A^i W A'^i$.

***Lemma** 5:* $\mathcal{H}_A(q, W)$ and $\mathcal{Z}_k(z, q, W)$ have following properties: (1) $\mathcal{H}_A(q, W)$ is finite and monotonically increases as q increases. (2) If $\rho(A) > 1$, $\forall W_1, W_2 \in \mathbb{S}_+$, $\mathcal{H}_A(q, (1-$

$q)W_1 + W_2)$ is monotonically increasing as q increases. (3) $\lim_{k\to\infty} \mathcal{Z}_k(z,q,W) = 0$.

Proof: (1) Since $q \in [0, \frac{1}{\rho(A)^2})$, the first property is obvious. (2) Let $\tilde{X} = (1-q)\mathcal{H}_A(q,W)$. First, by a similar technique as in the proof of Lemma 4, we can show that $\frac{\partial \tilde{X}}{\partial q}$ is finite. It can be seen that $\tilde{X}$ is the unique solution of $\tilde{X} = qA\tilde{X}A' + (1-q)W$, which gives $\tilde{X} - W = qA(\tilde{X} - W)A' + q(AWA' - W)$. Since $AWA' - W > 0$, $\tilde{X} - W > 0$ and hence $A\tilde{X}A' - W > 0$. Taking derivatives at both sides of the above Lyapunov equation for $\tilde{X}$, we have $\frac{\partial \tilde{X}}{\partial q} = qA\frac{\partial \tilde{X}}{\partial q}A' + A\tilde{X}A' - W$. Therefore, $\frac{\partial \tilde{X}}{\partial q} = \sum_{k=0}^{\infty} q^k A^k (A\tilde{X}A' - W) A'^k > 0$, which proves the monotonicity of $\tilde{X}$. Then, we can prove the statement by noticing that $\mathcal{H}_A(q,(1-q)W_1 + W_2) = (1-q)\mathcal{H}_A(q,W_1) + \mathcal{H}_A(q,W_2)$. (3) Since $z \in (0,1)$, $\mathcal{Z}_k(z,q,W) < \mathcal{H}(q,W) < \infty$. $\mathcal{Z}(z,q,W) = z\mathcal{Z}_k(z,q,W) + zq^{k+1}A^{k+1}WA'^{k+1}$. As $k \to \infty$, $zq^{k+1}A^{k+1}WA'^{k+1} \to 0$ and $\mathcal{Z}_\infty(z,q,W) = z\mathcal{Z}_\infty(z,q,W)$. Therefore, $\mathcal{Z}_\infty(z,q,W) = 0$. ∎

With a little abuse of notation, we define sequence $\{X_k | X_k = (1-\gamma_k)AX_{k-1}A' + \gamma_k W + Q, W \geq 0, X_0 \in \mathbb{S}_Q\}$.

Lemma 6: For the above sequence $\{X_k\}$, we have: (1) $\{X_k\}$ is stable in mean sense iff. $\sigma_2\rho(A)^2 < 1$. (2) If $\sigma_2\rho(A)^2 < 1$, $\{\mathbb{E}[X_k]\}$ converges to a unique value $\mathbb{E}[X_\infty]$, with $\mathrm{Tr}(\mathbb{E}[X_\infty]) = \delta\mathrm{Tr}([(1-\sigma_1)W + Q]\mathcal{G}_A(\sigma_1)) + (1-\delta)\mathrm{Tr}([(1-\sigma_2)W + Q]\mathcal{G}_A(\sigma_2))$. (3) If $\rho(A) > 1$, both $\mathbb{E}[X_\infty]$ and $\mathrm{Tr}(\mathbb{E}[X_\infty])$ are monotonically decreasing as p_d increases.

Proof: (1) It follows directly from Lemma 4.

(2) Given any $p_{0,0} \in [0,1]$, $\forall k > 0$, by the definition of Φ, we have $p_{0,k} = (1-\alpha)p_{0,k-1} + \beta p_{1,k-1} = zp_{0,k-1} + \beta$, where $z, p_{0,k}$ and $p_{1,k}$ are defined above (22) in the proof of Lemma 2. By induction, $p_{0,k} = (p_{0,0} - \frac{\beta}{\alpha+\beta})z^k + \frac{\beta}{\alpha+\beta}$ and $p_{0,\infty} = \frac{\beta}{\alpha+\beta}$. Substituting $p_{0,\infty}$ into (22) and let $k_2 = k_1 = k$, by several simple manipulations, we get $\bar{\gamma} = \lim\limits_{k\to\infty} \mathbb{P}\{\gamma_k = 0\} = \frac{1}{\alpha+\beta}(\beta\psi_1 + \alpha\psi_2)$. Substituting the above equation for $p_{0,k}$ into (22) yields $\varepsilon_{1,k_1} = \delta + \delta' z^{k_1}$ and $\varepsilon_{2,k_1} = 1 - \delta - \delta' z^{k_1}$, where $\delta' = \frac{(\psi_2 - \psi_1)(p_{0,0} - \frac{\beta}{\alpha+\beta})}{\sigma_2 - \sigma_1}$. Similar to (24), we have

$$
\begin{aligned}
&\mathbb{E}[X_k] = \mathbb{E}[(1-\gamma_k)AX_{k-1}A'] + \mathbb{E}[\gamma_k]W + Q \\
&= \mathbb{P}\{\gamma(1,k) = 0\}A^k X_0 A'^k + \mathbb{E}[\gamma_k]W + Q \\
&\quad + \sum_{i=1}^{k-1} \big[\mathbb{P}\{\gamma(k+1-i,k) = 0\}A^i Q A'^i + \\
&\quad (\mathbb{P}\{\gamma(k+1-i,k) = 0\} - \mathbb{P}\{\gamma(k-i,k) = 0\})\, A^i W A'^i\big] \\
&= \left(\varepsilon_{1,1}\sigma_1^k + \varepsilon_{2,1}\sigma_2^k\right) A^k X_0 A'^k + (1-\bar{\gamma})W + Q \\
&\quad + \sum_{i=1}^{k-1} \big[\left(\varepsilon_{1,k+1-i}\sigma_1^i + \varepsilon_{2,k+1-i}\sigma_2^i\right) A^i Q A'^i + (\varepsilon_{1,k+1-i}\sigma_1^i \\
&\quad + \varepsilon_{2,k+1-i}\sigma_2^i - \varepsilon_{1,k-i}\sigma_1^{i+1} - \varepsilon_{2,k-i}\sigma_2^{i+1}) A^i W A'^i\big] \quad (28)
\end{aligned}
$$

As $k \to \infty$, $\left(\varepsilon_{1,1}\sigma_1^k + \varepsilon_{2,1}\sigma_2^k\right) A^k X_0 A'^k$ converges to 0 since $|\sigma_1|\rho(A)^2 \leq \sigma_2\rho(A)^2 < 1$. Consider the following series where $S \in \{Q, W\}$: $\sum_{i=1}^{k-1}(\varepsilon_{1,k+1-i}\sigma_1^i)A^i S A'^i = \delta\sum_{i=0}^{k-1}\sigma_1^i A^i S A'^i + \delta'\sum_{i=0}^{k-1} z^{k+1-i}\sigma_1^i A^i S A'^i - (\delta + \delta' z^{k+1})S$. As $k \to \infty$, $\sum_{i=0}^{\infty}\sigma_1^i A^i S A'^i = \mathcal{H}_A(\sigma_1, S)$ (see Definition 1). By Lemma 5, $\lim_{k\to\infty}\sum_{i=0}^{k} z^{k+1-i}\sigma_1^i A^i S A'^i$. Thus, the above series converges to $\delta(\mathcal{H}_A(\sigma_1, S) - S)$. Similarly, the series $\sum_{i=0}^{k}\left(\varepsilon_{2,k+1-i}\sigma_2^i\right) A^i S A'^i$ converges to $(1-\delta)(\mathcal{H}_A(\sigma_2, S) - S)$ as k tends to ∞. Then, after several manipulations, (24) tends to $\mathbb{E}[X_\infty] = \delta\mathcal{H}_A(\sigma_1, (1-\sigma_1)W + Q) + (1-\delta)\mathcal{H}_A(\sigma_2, (1-\sigma_2)W + Q)$ as $k \to \infty$. Taking traces at both sides of the above equation and noticing the fact that $\mathrm{Tr}(\mathcal{H}_A(q,W)) = \mathrm{Tr}(W\sum_{k=0}^{\infty} q^k A'^k A^k) = \mathrm{Tr}(W\mathcal{G}_A(q))$, we can prove the second statement of this lemma.

(3) First, $\mathbb{E}[X_\infty]$ is monotonically increasing as ψ_1 increases, i.e., $\frac{\partial}{\partial\psi_1}\mathbb{E}[X_\infty] \geq 0$. One can prove this by using the second statement of Lemma 5 and the facts that $\frac{\partial\delta}{\partial\psi_1} \leq 0$, $\frac{\partial\sigma_1}{\partial\psi_1} \geq 0$ and $\frac{\partial\sigma_2}{\partial\psi_1} \geq 0$. Then, since ψ_1 is monotonically decreasing as p_d increases, $\mathbb{E}[X_\infty]$ (and also $\mathrm{Tr}(\mathbb{E}[X_\infty])$) is monotonically decreasing as p_d increases. ∎

Similar to (26), if C has full column rank,

$$
\begin{aligned}
P_k =& (1-\gamma_k)AP_{k-1}A' + \gamma_k A(P_{k-1}^{-1} + C'R^{-1}C)^{-1}A' + Q \\
\leq& (1-\gamma_k)AP_{k-1}A' + \gamma_k A(C'R^{-1}C)^{-1}A' + Q. \quad (29)
\end{aligned}
$$

Thus, $P_k \leq \bar{P}_k$ and $\mathbb{E}[P_k] \leq \mathbb{E}[\bar{P}_k]$. The lower bound $\underline{P}_k$ of P_k can be proved similarly by noticing that $P_k \geq Q$. Then, Theorem 5 can be proved by applying Lemma 6.

F. Proof of Theorem 6

According to Remark 3, the case without CHANCE (where packet losses obey i.i.d. distribution) is equivalent to the case with CHANCE and $p_d = p_f = 0$. In this case, $\psi_1^\dagger = \psi_2^\dagger = \sigma_2^\dagger = \bar{\gamma}^\dagger = \ell_0$, $\sigma_1^\dagger = \ell_0 z$, and $\delta^\dagger = 0$. Thus, the bounds of $\mathrm{Tr}(\mathbb{E}[P_\infty])$ as in (9) and (10) reduce to

$$
\begin{cases}
\mathrm{Tr}(\mathbb{E}[\underline{P}_\infty^\dagger]) = & \mathrm{Tr}(\underline{W}(\ell_0)\mathcal{G}_A(\ell_0)) \\
\mathrm{Tr}(\mathbb{E}[\bar{P}_\infty^\dagger]) = & \mathrm{Tr}(\bar{W}(\ell_0)\mathcal{G}_A(\ell_0))
\end{cases} \quad (30)
$$

If $\sigma_2 < \ell_0$, $\mathrm{Tr}(\mathbb{E}[\underline{P}_\infty]) = (1-\delta)\mathrm{Tr}(\underline{W}(\sigma_2)\mathcal{G}_A(\sigma_2)) + \delta\mathrm{Tr}(\underline{W}(\sigma_1)\mathcal{G}_A(\sigma_1)) \leq \mathrm{Tr}(\underline{W}(\sigma_2)\mathcal{G}_A(\sigma_2)) < \mathrm{Tr}(\mathbb{E}[\underline{P}_\infty^\dagger])$, by Lemma 5. Similarly, $\mathrm{Tr}(\mathbb{E}[\bar{P}_\infty]) < \mathrm{Tr}(\bar{W}(\ell_0)\mathcal{G}_A(\ell_0))$. Then, (11) can be proved based on (30) and (10).

G. Proof of Theorem 7

Lemma 7: With the channel sensing schedule $\mathcal{Q}_m$, the sequence $\{X_k | X_k = (1-\gamma_k)AX_{k-1}A' + S, \forall k \geq 0, X_0 \in \mathbb{S}_Q\}$ is stable in mean sense iff. $\rho(\Phi\Psi)\rho(A)^2 < 1$.

Proof: Define matrix $M_k \in \mathbb{R}^{2^m \times 2^m}$, where $\forall i, j \in \mathcal{M}$, $[M_k]_{ij} \triangleq \mathbb{P}\{\mathbf{s}_k = \mathcal{B}_m(j), \gamma(1,k) = 0 \big| \mathbf{s}_0 = \mathcal{B}_m(i)\}$. Following the very similar way as in (20), we can prove that $[M_k]_{ij} = \sum_{l=1}^{2^m}[M_{k-1}]_{il}[\Phi]_{lj}\psi_j$. Therefore, $M_k = M_{k-1}\Phi\Psi$. Moreover, since $M_0 = I$, $M_k = (\Phi\Psi)^k$. Similar to (21), $\mathbb{P}\{\gamma(k_1,k_2) = 0\} = \mathbf{p}'_{k_1-1}(\Phi\Psi)^{k_2+1-k_1}\mathbf{u}$ holds true for all $k_1 > 0$ and $k_2 \geq k_1$.

1) Sufficiency: According to Lemma 3, $\exists\zeta'_2 > 0$ such that

$$
\begin{aligned}
&\mathbb{P}\{\gamma(k_1,k_2) = 0\} = \mathrm{Tr}(\mathbf{u}\mathbf{p}'_{k_1-1}M_{k_2+1-k_1}) \\
&\leq \sqrt{\mathrm{Tr}(\mathbf{p}_{k_1-1}\mathbf{u}'\mathbf{u}\mathbf{p}'_{k_1-1})}\sqrt{\mathrm{Tr}(M'_{k_2+1-k_1}M_{k_2+1-k_1})} \\
&\leq \sqrt{2^m(\sum_{b=1}^{2^m} p_{b,k_1-1})^2}\sqrt{\mathrm{Tr}(M'_{k_2+1-k_1}M_{k_2+1-k_1})} \\
&< 2^{\frac{m}{2}}\zeta'_2(k_2+1-k_1)^{2^m}\rho(\Phi\Psi)^{k_2+1-k_1}
\end{aligned}
$$

where the first inequality is based on the Cauchy-Schwarz inequality, while the last one is based on Lemma 3. Therefore,

$$\begin{aligned}
&\mathrm{Tr}(\mathbb{E}[X_k]) = \mathbb{P}\{\gamma(1,k)=0\}\mathrm{Tr}(A^k X_0 A'^k) \\
&\quad + \sum_{i=1}^{k-1}\mathbb{P}\{\gamma(k,k+1-i)=0\}\mathrm{Tr}(A^i S A'^i) + \mathrm{Tr}(S) \\
&< 2^{\frac{m}{2}}\zeta_2' k^{2^m}\rho(\Phi\Psi)^k\,\mathrm{Tr}(A^k X_0 A'^k) \\
&\quad + 2^{\frac{m}{2}}\zeta_2'\sum_{i=1}^{k-1} i^{2^m}\rho(\Phi\Psi)^i\,\mathrm{Tr}(A^i S A'^i) + \mathrm{Tr}(S) \\
&< 2^{\frac{m}{2}}\zeta_2'\zeta_2\mathrm{Tr}(X_0)k^{2^m+2n}\rho(\Phi\Psi)^k\rho(A)^{2k} \\
&\quad + 2^{\frac{m}{2}}\zeta_2'\zeta_2\mathrm{Tr}(S)\sum_{i=1}^{k-1} i^{2^m+2n}\rho(\Phi\Psi)^i\rho(A)^{2i} + \mathrm{Tr}(S).
\end{aligned}$$

Similar to the proof of Lemma 4, a sufficient condition for $\sup_{k\geq 0}\mathrm{Tr}(\mathbb{E}[X_k])<\infty$ is $\rho(\Phi\Psi)\rho(A)<1$.

2) Necessity: Because $\alpha_i+\beta_i<1$, $\mathbb{P}\{s_{i,k+1}=1\}=(1-\beta_i)\mathbb{P}\{s_{i,k}=1\}+\alpha_i\mathbb{P}\{s_{i,k}=0\}=(1-\alpha_i-\beta_i)\mathbb{P}\{s_{i,k}=1\}+\alpha_i\geq\alpha_i$ and also $\mathbb{P}\{s_{i,k+1}=0\}\geq\beta_i$. Therefore, $\exists p_{\min}>0$ such that $\forall k\geq 1, p_{b,k}\geq\prod_{i=1}^m\min\{\alpha_i,\beta_i\}=p_{\min}$. Then,

$$\begin{aligned}
&\mathbb{P}\{\gamma(k_1,k_2)=0\}\geq p_{\min}\sum_{b=1}^{B}\sum_{l=1}^{B}[M_{k_2+1-k_1}]_{lb} \\
&\geq p_{\min}\sqrt{\sum_{b=1}^{B}\sum_{l=1}^{B}[M_{k_2+1-k_1}]_{bl}^2} \\
&= p_{\min}\sqrt{\mathrm{Tr}\left(M'_{k_2+1-k_1}M_{k_2+1-k_1}\right)}\geq p_{\min}\zeta_1'\rho(\Phi\Psi)^{k_2+1-k_1},
\end{aligned}$$

where $\zeta_1'>0$ and the last inequality is from Lemma 3. Then, the necessity is obvious. ■

In view of the above lemma, Theorem 7 can be similarly proved based on (25) and (26) which have been used to prove Theorems 1 and 2.

H. Proof of Theorem 8

***Lemma** 8 ([36]):* For any two square matrices $X,Y\in\mathbb{R}^{n\times n}$, $\rho(X)\geq\rho(Y)$ if for all possible i,j, $[X]_{ij}\geq[Y]_{ij}\geq 0$; $\rho(X\otimes Y)=\rho(X)\rho(Y)$.

$\forall b\in\{1,\ldots,2^m\}$, if $b_m=1$, from (14) we have

$$\begin{aligned}
&\psi_b=[p_{f,m}+\ell_0(1-p_{f,m})]\prod_{j=1}^{m-1}\mathbb{P}\{o_{j,k}=1|s_{j,k}=b_j\} \\
&+\sum_{i=1}^{m-1}\Big(\prod_{j=1}^{i-1}\mathbb{P}\{o_{j,k}=1|s_{j,k}=b_j\}\Big)\mathbb{P}\{o_{i,k}=0|s_{i,k}=b_i\}\ell_i^{b_i} \\
&\geq\sum_{i=1}^{m-1}\Big(\prod_{j=1}^{i-1}\mathbb{P}\{o_{j,k}=1|s_{j,k}=b_j\}\Big)\mathbb{P}\{o_{i,k}=0|s_{i,k}=b_i\}\ell_i^{b_i} \\
&\quad+\ell_0\prod_{j=1}^{m-1}\mathbb{P}\{o_{j,k}=1|s_{j,k}=b_j\},
\end{aligned}$$

where the equality holds only when $p_{f,m}=0$. If $b_m=0$, in a similar way, we have

$$\begin{aligned}
\psi_b\geq&\sum_{i=1}^{m-1}\Big(\prod_{j=1}^{i-1}\mathbb{P}\{o_{j,k}=1|s_{j,k}=b_j\}\Big)\mathbb{P}\{o_{i,k}=0|s_{i,k}=b_i\}\ell_i^{b_i} \\
&+\min\{\ell_0,\ell_m\}\prod_{j=1}^{m-1}\mathbb{P}\{o_{j,k}=1|s_{j,k}=b_j\},
\end{aligned}$$

and the equality holds only when $p_{d,m}=0$ if $\ell_0<\ell_m$ or $p_{d,m}=1$ if $\ell_0\geq\ell_m$. Continuing applying this method to $\{b_{m-1},\ldots,b_1\}$, we finally get that $\psi_b\geq\ell_0$ if $b=2^m$ and $\psi_b\geq\ell_{\min}$ otherwise. Thus, $\rho(\Phi\Psi)\geq\rho(\Phi\ Diag\{0,\ldots,0,\ell_0\})=\ell_0\prod_{i=1}^m(1-\beta_i)$, if $\ell_{\min}=0$. In this case, the stability gain $\eta\leq\frac{1}{\sqrt{\prod_{i=1}^m(1-\beta_i)}}$. Otherwise, if $\ell_{\min}>0$, by Lemma 8, we have $\rho(\Phi\Psi)\geq\rho(\ell_{\min}\Phi)=\ell_{\min}\rho(\bigotimes_{\kappa=1}^m\Phi_\kappa)=\ell_{\min}\prod_{\kappa=1}^m\rho(\Phi_\kappa)=\ell_{\min}$. Thus, $\eta\leq\sqrt{\frac{\ell_0}{\ell_{\min}}}$. In sum, (16) holds and thus the theorem is proved.

I. Proof of Theorem 9

Similar to the proof of Theorem 4, we have

$$\begin{aligned}
&\mathbb{P}\{P_{k+1}>\phi I\}=\mathbb{P}\{\mathcal{L}(P_k)>\phi I,\gamma_k=0\} \\
&\quad+\mathbb{P}\{\mathcal{R}(P_k)>\phi I,\gamma_k=1\} \\
&=\sum_{b=1}^{2^m}\Big\{\mathbb{P}\{\mathcal{L}(P_k)>\phi I,\mathbf{s}_{k-1}=\mathcal{B}_m(b)\}\sum_{i=1}^{2^m}[M_1]_{bi} \\
&\quad+\mathbb{P}\{\mathcal{R}(P_k)>\phi I,\mathbf{s}_{k-1}=\mathcal{B}_m(b)\}[\sum_{i=1}^{2^m}([\Phi]_{bi}-[M_1]_{bi})]\Big\} \\
&=\sum_{b=1}^{2^m}\Big\{\mathbb{P}\{\mathcal{L}(P_k)>\phi I,\mathbf{s}_{k-1}=\mathcal{B}_m(b)\}\mathbf{e}_b'\Phi\Psi\mathbf{u} \\
&\quad+\mathbb{P}\{\mathcal{R}(P_k)>\phi I,\mathbf{s}_{k-1}=\mathcal{B}_m(b)\}[\sum_{i=1}^{2^m}[\Phi]_{bi}-\mathbf{e}_b'\Phi\Psi\mathbf{u}]\Big\} \\
&=\sum_{b=1}^{2^m}\Big\{\mathbb{P}\{\mathcal{R}(P_k)>\phi I,\mathbf{s}_{k-1}=\mathcal{B}_m(b)\} \\
&\quad+\mathbf{e}_b'\Phi\Psi\mathbf{u}\big[\mathbb{P}\{\mathcal{L}(P_k)>\phi I,\mathbf{s}_{k-1}=\mathcal{B}_m(b)\} \\
&\quad-\mathbb{P}\{\mathcal{R}(P_k)>\phi I,\mathbf{s}_{k-1}=\mathcal{B}_m(b)\}\big]\Big\} \\
&\leq\mathbb{P}\{\mathcal{R}(P_k)>\phi I\}+\tilde{\ell}^*\sum_{b=1}^{2^m}\big[\mathbb{P}\{\mathcal{L}(P_k)>\phi I,\mathbf{s}_{k-1}=\mathcal{B}_m(b)\} \\
&\quad-\mathbb{P}\{\mathcal{R}(P_k)>\phi I,\mathbf{s}_{k-1}=\mathcal{B}_m(b)\}\big] \\
&=\tilde{\ell}^*\mathbb{P}\{\mathcal{L}(P_k)>\phi I\}+(1-\tilde{\ell}^*)\mathbb{P}\{\mathcal{R}(P_k)>\phi I\} \\
&\leq\ell_0\mathbb{P}\{\mathcal{L}(P_k)>\phi I\}+(1-\ell_0)\mathbb{P}\{\mathcal{R}(P_k)>\phi I\},
\end{aligned}$$

where we have used the fact that $\sum_{i=1}^{2^m}[\Phi]_{bi}=1,\forall b\in\{1,\ldots,2^m\}$. In the same way as in the proof of Theorem 4, the proof of the current theorem can be achieved.

J. Proof of Theorem 10

In the proof of Lemma 6, we show that $\forall k\geq 0, p_{0,k}=(p_{0,0}-\frac{\beta}{\alpha+\beta})z^k+\frac{\beta}{\alpha+\beta}$, which can be generalized to the multi-licensed-channel case as $\mathbb{P}\{s_{i,k}=s\}=\xi_{i,s}+\tilde{\xi}_{i,s}z_i^k$, where $s\in\{0,1\}$, $\xi_{i,s}$ and $\tilde{\xi}_{i,s}$ are constants. Thus, $\forall b\in\{1,\ldots,2^m\}$, $p_{b,k}=\mathbb{P}\{\mathbf{s}_k=\mathcal{B}_m(b)\}=\prod_{i=1}^m(\xi_{i,[\mathcal{B}_m(b)]_i}+\tilde{\xi}_{i,[\mathcal{B}_m(b)]_i}z_i^k)=w_b+\tilde{w}_{b,k}z^k$, where $w_b\triangleq\prod_{i=1}^m\xi_{i,[\mathcal{B}_m(b)]_i}$,

$z \triangleq \max\{z_1, \ldots, z_m\}$. The expression of $\tilde{w}_{b,k}$ is complicated but it is easy to see that, $\tilde{w}_{b,k}$ is above bounded by some $\tilde{w}_b$. Since $z \in (0,1)$, $w_b = \lim_{k\to\infty} p_{b,k}$. Thus, the channel state vector can be rewritten as $\mathbf{p}_k = \boldsymbol{\upsilon} + \tilde{\boldsymbol{\upsilon}}_k z^k$, where $\boldsymbol{\upsilon}_k = [w_1, \ldots, w_{2^m}]'$, $\tilde{\boldsymbol{\upsilon}}_k = [\tilde{\boldsymbol{\upsilon}}_{1,k}, \ldots, \tilde{\boldsymbol{\upsilon}}_{2^m,k}]'$. Let $\tilde{\boldsymbol{\upsilon}} = [\tilde{\boldsymbol{\upsilon}}_1, \ldots, \tilde{\boldsymbol{\upsilon}}_{2^m}]'$. As shown in the proof of Lemma 7, the consecutive packet loss probability is $\mathbb{P}\{\gamma(k_1,k_2) = 0\} = \mathbf{p}'_{k_1-1}(\Phi\Psi)^{k_2+1-k_1}\mathbf{u} = (\boldsymbol{\upsilon} + \tilde{\boldsymbol{\upsilon}}_{k_1-1}z^{k_1-1})'(\Phi\Psi)^{k_2+1-k_1}\mathbf{u}$. Then, according to (28), the mean of the sequence $\{X_k | X_k = (1-\gamma_k)AX_{k-1}A' + \gamma_k W + Q, W \geq 0, X_0 \in \mathbb{S}_Q\}$ becomes

$$\begin{aligned}
\mathbb{E}[X_k] &= (\boldsymbol{\upsilon} + \tilde{\boldsymbol{\upsilon}}_0)'\,(\Phi\Psi)^k\,\mathbf{u}A^k X_0 A'^k \\
&\quad + \sum_{i=1}^{k-1}\big[(\boldsymbol{\upsilon} + \tilde{\boldsymbol{\upsilon}}_{k+1-i}z^{k+1-i})'\,(\Phi\Psi)^i\,\mathbf{u}A^i Q A'^i \\
&\quad + ((\boldsymbol{\upsilon} + \tilde{\boldsymbol{\upsilon}}_{k+1-i}z^{k+1-i})' \\
&\quad - (\boldsymbol{\upsilon} + \tilde{\boldsymbol{\upsilon}}_{k-i}z^{k-i})'\,(\Phi\Psi))\,(\Phi\Psi)^i\,\mathbf{u}A^i W A'^i\big] \\
&\quad + \left(1 - (\boldsymbol{\upsilon} + \tilde{\boldsymbol{\upsilon}}_{k-1}z^{k-1})'\,(\Phi\Psi)\,\mathbf{u}\right)W + Q.
\end{aligned}$$

If $\rho\,(\Phi\Psi)\,\rho(A)^2 < 1$, $(\Phi\Psi)^k \to 0$. Since $\Phi\Psi \leq \rho\,(\Phi\Psi)\,I$, $(\tilde{\boldsymbol{\upsilon}}_{k+1-i}z^{k+1-i})'\,(\Phi\Psi)^i\,\mathbf{u} \leq \tilde{\boldsymbol{\upsilon}}'_{k+1-i}\mathbf{u}z^{k+1-i}\rho\,(\Phi\Psi)^i \leq \tilde{\boldsymbol{\upsilon}}'\mathbf{u}z^{k+1-i}\rho\,(\Phi\Psi)^i$. With $\rho\,(\Phi\Psi) < \frac{1}{\rho(A)^2}$, by Lemma 5, $\lim\limits_{k\to\infty}\sum_{i=1}^{k-1} z^{k+1-i}\rho\,(\Phi\Psi)^i A^i S A'^i = 0$. Hence $\lim\limits_{k\to\infty}\sum_{i=1}^{k-1}(\tilde{\boldsymbol{\upsilon}}_{k+1-i}z^{k+1-i})'\,(\Phi\Psi)^i\,\mathbf{u}A^i S A'^i = 0$, where S can be either Q or W. Thus, $\mathbb{E}[X_k]$ converges to

$$\begin{aligned}
\mathbb{E}[X_\infty] &= \sum_{i=1}^{\infty}\Big[\boldsymbol{\upsilon}'\,(\Phi\Psi)^i\,\mathbf{u}A^i Q A'^i + (1 - \boldsymbol{\upsilon}'\Phi\Psi\mathbf{u})\,W + Q \\
&\qquad + \boldsymbol{\upsilon}'(I - \Phi\Psi)\,(\Phi\Psi)^i\,\mathbf{u}A^i W A'^i\Big] \\
&= \sum_{i=0}^{\infty}\boldsymbol{\upsilon}'\,(\Phi\Psi)^i\,\mathbf{u}A^i[Q + (1-\varrho)W]A'^i.
\end{aligned}$$

Then, the lower and upper bounds as in (17) and (18) of the current theorem are evident.

Though (18) can be used to compute the worst case performance ratio in the way similar to Theorem 6, the computation may be very complicated as the critical matrix can be nearly in any form (may not be diagonalizable). Instead, we derive a relaxed but much simpler bound for φ_∞. Since $\boldsymbol{\upsilon}'\,(\Phi\Psi)^i\,\mathbf{u} \leq \boldsymbol{\upsilon}'\rho(\Phi\Psi)^i\mathbf{u} = \rho(\Phi\Psi)^i$, $\mathbb{E}[X_\infty] \leq \sum_{i=0}^{\infty}\rho(\Phi\Psi)^i A^i[Q + (1-\varrho)W]A'^i = \mathcal{H}_A\left(\rho\,(\Phi\Psi)\,, [Q + (1-\varrho)W]\right)$. Thus, $\mathrm{Tr}(\mathbb{E}[\bar{P}_\infty]) \leq \mathrm{Tr}(\bar{W}(\varrho)\mathcal{G}_A(\rho(\Phi\Psi)))$. Together with (30), we can prove (19).

References

[1] S. Haykin. Cognitive radio: brain-empowered wireless communications. *IEEE J. Select. Areas Commun.*, 23(2):201–220, 2005.

[2] M. Pan, P. Li, and Y. Fang. Cooperative communication aware link scheduling for cognitive vehicular networks. *IEEE J. Select. Areas Commun.*, 30(4):760–768, 2012.

[3] W.-Y. Lee and I.F. Akyildiz. Optimal spectrum sensing framework for cognitive radio networks. *IEEE Trans. on Wireless Communications*, 7(10):3845–3857, Oct. 2008.

[4] J. Huang, Z. Han, M. Chiang, and H. V. Poor. Auction-based resource allocation for cooperative communications. *IEEE J. Select. Areas Commun.*, 26(7):1226–1237, 2008.

[5] Y. Pei, Y.-C. Liang, K. C. Teh, and K. H. Li. Energy-efficient design of sequential channel sensing in cognitive radio networks: optimal sensing strategy, power allocation, and sensing order. *IEEE J. Select. Areas Commun.*, 29(8):1648–1659, 2011.

[6] H. Li, L. Lai, and V. Poor. Multicast routing for decentralized control of cyber physical systems with an application in smart grid. *IEEE J. Select. Areas Commun.*, 30(6):1097–1107, 2012.

[7] S.-Y. Lien, S.-M. Cheng, S.-Y. Shih, and K.-C. Chen. Radio resource management for qos guarantees in cyber-physical systems. *IEEE Trans. Parallel and Distributed Systems*, 23(9):1752–1761, 2012.

[8] Y. Mo, T.-J. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli. Cyber–physical security of a smart grid infrastructure. *Proceedings of the IEEE*, 100(1):195–209, 2012.

[9] X. Cao, P. Cheng, J. Chen, and Y. Sun. An online optimization approach for control and communication co-design in networked cyber-physical systems. *IEEE Trans. Industrial Informatics*, 9(1):439–450, 2013.

[10] W. Shen, L. Liu, X. Cao, Y. Hao, and Y. Cheng. Cooperative message authentication in vehicular cyber-physical systems. *IEEE Trans. Emerging Topics in Computing*, 1(1):84–97, 2013.

[11] C. L Robinson and P. Kumar. Optimizing controller location in networked control systems with packet drops. *IEEE J. Select. Areas Commun.*, 26(4):661–671, 2008.

[12] S.L. Dai, H. Lin, and S.S. Ge. Scheduling-and-control codesign for a collection of networked control systems with uncertain delays. *IEEE Trans. on Control Systems Technology*, 18(1):66–78, 2010.

[13] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan, and S.S. Sastry. Kalman filtering with intermittent observations. *IEEE Trans. on Automatic Control*, 49(9):1453–1464, Sep. 2004.

[14] K. You, M. Fu, and L. Xie. Mean square stability for Kalman filtering with Markovian packet losses. *Automatica*, 47(12):2647–2657, 2011.

[15] A. Censi. Kalman filtering with intermittent observations: convergence for semi-Markov chains and an intrinsic performance measure. *IEEE Trans. on Automatic Control*, 56(2):376–381, 2011.

[16] E. Rohr, D. Marelli, and M. Fu. Kalman filtering for a class of degenerate systems with intermittent observations. In *IEEE Conference on Decision and Control*, 2011.

[17] D.E. Quevedo, A. Ahlén, and J. Østergaard. Energy efficient state estimation with wireless sensors through the use of predictive power control and coding. *IEEE Trans. on Signal Processing*, 58(9):4811–4823, 2010.

[18] L. Shi, P. Cheng, and J. Chen. Sensor data scheduling for optimal state estimation with communication energy constraint. *Automatica*, 47(8):1693–1698, 2011.

[19] A. Chiuso, L. Schenato. Information fusion strategies and performance bounds in packet-drop networks. *Automatica*, 47(7):1304–1316, July 2011.

[20] X. Liu, A. Goldsmith. Kalman filtering with partial observation losses. In *IEEE Conference on Decision and Control*, volume 4, pages 4180–4186, 2004.

[21] X. Ma, S.M. Djouadi, T.P. Kuruganti, J.J. Nutaro, and H. Li. Control and estimation through cognitive radio with distributed and dynamic spectral activity. In *American Control Conference (ACC), 2010*, pages 289–294, 2010.

[22] X. Ma, S.M. Djouadi, and H. Li. State estimation over a semi-Markov model based cognitive radio system. *IEEE Trans. on Wireless Communications*, 11(7):2391–2401, 2012.

[23] G. Taricco. Optimization of linear cooperative spectrum sensing for cognitive radio networks. *IEEE Journal of Selected Topics in Signal Processing*, 5(1):77–86, 2011.

[24] L. Lai, H. El Gamal, H. Jiang, and H.V. Poor. Cognitive medium access: Exploration, exploitation, and competition. *IEEE Trans. on Mobile Computing*, 10(2):239–253, 2011.

[25] Y. Cheng, H. Li, P.J. Wan. A theoretical framework for optimal cooperative networking in multiradio multichannel wireless networks. *IEEE Wireless Communications*, 19(2):66–73, 2012.

[26] X. Cao, J. Chen, Y. Xiao, and Y. Sun. Building-environment control with wireless sensor and actuator networks: centralized versus distributed. *IEEE Trans. on Industrial Electronics*, 57(11):3596–3605, 2010.

[27] Y. Mo, B. Sinopoli. Kalman filtering with intermittent observations: tail distribution and critical value. *IEEE Trans. on Automatic Control*, 57(3):677–689, 2012.

[28] Z. Quan, S. Cui, A.H. Sayed, and H.V. Poor. Wideband spectrum sensing in cognitive radio networks. In *Communications, 2008. ICC'08. IEEE International Conference on*, pages 901–906. IEEE, 2008.

[29] C. Cormio, K.R. Chowdhury. Common control channel design for cognitive radio wireless ad hoc networks using adaptive frequency hopping. *Ad Hoc Networks*, 8(4):430–438, 2010.

[30] Y. Ephraim and N. Merhav. Hidden Markov processes. *IEEE Trans. on Information Theory*, 48(6):1518–1569, 2002.

[31] G. Corradi, J. Janssen, R.Manca. Numerical treatment of homogeneous semi-Markov processes in transient case–a straightforward approach. *Methodology and Computing in Applied Probability*, 6(2):233–246, 2004.

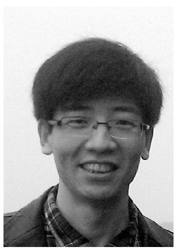
[32] A. Ghasemi and E.S. Sousa. Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs. *IEEE Communications Magazine*, 46(4):32–39, 2008.

[33] A. Tiwari, M. Jun, D.E. Jeffcoat, and R.M. Murray. Analysis of dynamic sensor coverage problem using Kalman filters for estimation. In *Proc. IFAC world congress*, 2005.

[34] Y. Mo and B. Sinopoli. A characterization of the critical value for kalman filtering with intermittent observations. In *IEEE Conference on Decision and Control (CDC)*, pages 2692–2697, 2008.

[35] S. Chen and L. Tong. Maximum throughput region of multiuser cognitive access of continuous time Markovian channels. *IEEE J. Select. Areas Commun.*, 29(10):1959–1969, 2011.

[36] R.E. Bellman. *Introduction to matrix analysis*, volume 19. Society for Industrial and Applied Mathematics, 1997.



Xianghui Cao (S'08-M'11) received his B.S. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2006 and 2011, respectively. During 2008-2010, he was a visiting scholar in the Department of Computer Science, the University of Alabama. Currently, he is a postdoctoral fellow in Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago. His research interests include wireless network performance analysis, energy efficiency of wireless networks, networked estimation and control, and network security.

Peng Cheng (M'10) received the B.E. degree in Automation, and the Ph.D. degree in Control Science and Engineering in 2004 and 2009 respectively, both from Zhejiang University, Hangzhou, P.R. China. Currently he is Associate Professor with Department of Control Science and Engineering, Zhejiang University. He serves as the publicity co-chair for IEEE MASS 2013. His research interests include networked sensing and control, cyber-physical systems, and robust control.

Jiming Chen (M'08-SM'11) received B.Sc degree and Ph.D degree both in Control Science and Engineering from Zhejiang University in 2000 and 2005, respectively. He was a visiting researcher at INRIA in 2006, National University of Singapore in 2007, and University of Waterloo from 2008 to 2010. Currently, he is a full professor with Department of Control Science and Engineering, and the coordinator of group of Networked Sensing and Control in the State Key laboratory of Industrial Control Technology at Zhejiang University, China. His research interests are estimation and control over sensor network, sensor and actuator network, coverage and optimization in sensor network. He currently serves associate editors for several international Journals including IEEE Transactions on Control of Network Systems, IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Industrial Electronics and IEEE Networks. He is a guest editor of IEEE Transactions on Automatic Control, Computer Communication (Elsevier), Wireless Communication and Mobile Computer (Wiley) and Journal of Network and Computer Applications (Elsevier). He also serves as a Co-chair for Ad hoc and Sensor Network Symposium, IEEE Globecom 2011, general symposia Co-Chair of ACM IWCMC 2009 and ACM IWCMC 2010, WiCON 2010 MAC track Co-Chair, IEEE MASS 2011 Publicity Co-Chair, IEEE DCOSS 2011 Publicity Co-Chair, IEEE ICDCS 2012 Publicity Co-Chair and TPC member for IEEE ICDCS 2010, IEEE MASS 2010, IEEE SECON 2011, IEEE INFOCOM 2011, IEEE INFOCOM 2012, IEEE ICDCS 2012 etc.

Shuzhi Sam Ge (S'90-M'92-SM'99-F'06) received the B.Sc. degree from Beijing University of Aeronautics and Astronautics, in 1986, the Ph.D. degree, and the Diploma of Imperial College (DIC) from the Imperial College of Science, Technology, and Medicine, London, U.K., in 1993. He is the Director of Social Robotics Lab of Interactive Digital Media Institute, and Professor of the Department of Electrical and Computer Engineering, the National University of Singapore, the Director of Robotics Institute, and Professor of School of Computer Science and Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China. He has (co)-authored four books, and over 300 international journal and conference papers. He is the Editor-in-Chief, International Journal of Social Robotics, Springer. He has served/been serving as an Associate Editor for a number of flagship journals including IEEE Transactions on Automatic Control, IEEE Transactions on Control Systems Technology, IEEE Transactions on Neural Networks, and Automatica. He also serves as a book Editor of the Taylor & Francis Automation and Control Engineering Series. At IEEE Control Systems Society, he served/serves as Vice President for Technical Activities, 2009-2010, Vice President of Membership Activities, 2011-2012, Member of Board of Governors of IEEE Control Systems Society, 2007-2009. He is also a Fellow of IEEE, and IFAC. His current research interests include social robotics, multimedia fusion, adaptive control, intelligent systems and artificial intelligence.

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, 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, USA, and 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 Postdoctoral Fellowship Award from the Natural Sciences and Engineering Research Council of Canada (NSERC) in 2004, and a Best Paper Award from the conferences QShine 2007 and ICC 2011. 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 TPC Co-Chair for WASA 2011. He is a founding Vice Chair of the IEEE ComSoc Technical Subcommittee on Green Communications and Computing. He is an Associated Editor for IEEE Transactions on Vehicular Technology and the New Books & Multimedia Column Editor for IEEE Network.

Youxian Sun received the Diploma from the Department of Chemical Engineering, Zhejiang University, China, in 1964. He joined the Department of Chemical Engineering, Zhejiang University, in 1964. From 1984 to1987, he was an Alexander Von Humboldt Research Fellow and Visiting Associate Professor at University of Stuttgart, Germany. He has been a full professor at Zhejiang University since 1988. In 1995, he was elevated to an Academician of the Chinese Academy of Engineering. His current research interests include modeling, control, and optimization of complex systems, and robust control design and its application. He is author/co-author of 450 journal and conference papers. He is currently the director of the Institute of Industrial Process Control and the National Engineering Research Center of Industrial Automation, Zhejiang University. He is President of the Chinese Association of Automation, also has served as Vice-Chairman of IFAC Pulp and Paper Committee and Vice-President of China Instrument and Control Society.