IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received August 21, 2017, accepted September 25, 2017, date of publication October 5, 2017, date of current version November 7, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2758848

Spectrum Sensing for Cognitive Radios With Unknown Noise Variance and Time-variant Fading Channels

SHENGHONG LI[1], MENGWEI SUN[2], YING-CHANG LIANG[4], (Fellow, IEEE), BIN LI[3], AND CHENGLIN ZHAO[3]

[1]Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[2]School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K.
[3]School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
[4]School of Electrical and Information Engineering, University of Sydney, Sydney, NSW 2006, Australia

Corresponding author: Mengwei Sun (mengwei.sun@hw.ac.uk)

This work was supported in part by the National Natural Science Foundation of China under Grant 61379016, Grant 61471061, Grant 61271316, Grant 61571100, and Grant 61631005, and in part by the Key Laboratory for Shanghai Integrated Information Security Management Technology Research.

ABSTRACT The unknown noise variance and time-variant fading channels make the spectrum sensing design a challenging task for cognitive radios. Most existing sensing methods suffer from the information uncertainty and can hardly acquire promising performances in the adverse situations. To address this challenge, in this paper, we first formulate a dynamic state-space model for spectrum sensing, in which the unknown noise variance and time-variant flat fading channels are all taken into considerations. The dynamic behaviors of both primary user states and fading channels are characterized by two discrete state Markov chains. Based on this model, a novel spectrum sensing scheme is designed to recursively estimate the occupancy state of primary users, by estimating the time-variant fading channel gain and noise parameters jointly. The joint estimation is primarily premised on a maximum *a posteriori* probability criterion and the marginal particle filtering schemes. Simulation results are provided to demonstrate the advantages of our proposed method, which can significantly improve the sensing performance over time-variant flat fading channels, even with unknown noise variance.

INDEX TERMS Spectrum sensing, unknown noise parameter, time-variant flat fading channel, joint estimation.

I. INTRODUCTION

Due to limited availability of spectrum resources, the conventional paradigm of static spectrum management will not be able to accommodate the ever-growing demands of future wireless communications [1]. As an innovative technique, cognitive radio (CR) supports the secondary users (SUs) to utilize the spectrum assigned to the primary users (PUs) opportunistically [2]–[6], thus, it has the potential to alleviate the spectrum scarcity problem. According to Federal Communication Commission (FCC) [1], CR is "a radio or system that senses its operational electromagnetic environment and dynamically and autonomously adjusts its radio operating parameters to modify system operation, to maximize throughput, mitigate interference, facilitate inter-operability, access secondary markets." In this regards, spectrum sensing is one of the fundamental and critical elements in CRs [7].

The main purpose of spectrum sensing is to identify the occupancy status of PUs, i.e., whether the spectrum of interest is occupied by the PUs [7]. The commonly used methods include energy detection (ED) [8], [9], matched filtering detection (MFD) [10], [11], cyclo-stationary feature detection [12] and waveform-based sensing [13]. Among these, MFD yields the optimal detection performance under the assumption that the received primary signals are *perfectly* known and there is no information uncertainty. Thus, besides the waveforms transmitted from the PU, the channel state information (CSI) from the PU to SU should also be acquired.

Since there is usually no coordination between PU and SU, the noise variance and the CSI between PU and SU can be hardly estimated. In order to overcome the impact due to noise uncertainty, there have been several approaches proposed in the literature. Chen et al. proposed a method

combining cooperative spectrum sensing with adaptive multi-threshold selection [14], [15]. The use of cooperative sensing, to some extent, increases the deployment complexity and may also introduce other optimization/feedback problems. A sensing method based on multi-antennas is proposed in [16] using the generalized likelihood-ratio test (GLRT) paradigm. Yet, the deployment of multi-antennas will pose the strict requirement on sensing equipments. Zeng et al. [17] proposed a sensing algorithm based on the difference of statistical covariance between the received signals and that of the white noises. This method, however, requires that the received signals are temporally or spatially correlated. When such correlation is low, the performance of this method will be degraded.

As for the time-variant CSI, two classical techniques have been proposed to combat the unfavorable effects. The first one employs cooperative techniques as in [18]. The second one, on the other hand, relies on the statistical properties of the time-variant fading channels [9]. This method focuses on the instantaneous random distribution of the fading channel, but fails to exploit the underlying *time-correlation* of the channel. A recent approach was suggested by Li et. al. [19], in which the time-variant channel is estimated/tracked when performing spectrum sensing. While this method manages to jointly estimate PU's occupancy states and time-variant fading channel, the noise variance is assumed to be perfectly known in [19].

To combat these imperfections, we focus on the spectrum sensing design for CRs with unknown noise variance as well as time-variant channel in this paper, and we propose an novel sensing method, which will suppress the information uncertainty and improve the sensing performance of MFD. With the accurately acquired fading channels, the time-correlation (or dynamic property) can be exploited to further promote the sensing performance. By estimating both unknown channel fading and noise variance, the new sensing scheme will mitigate the information uncertainty to the minimum, and therefore, the sensing performance could even approach an ideal MFD with the complete *a priori* information. In general, the main contributions are summarized as follows.

First, we formulate a new dynamic state-space model (DSM) to characterize the spectrum sensing process with unknown noise and time-variant flat fading (TVFF) channel. In the unified stochastic model, the PU's state, the dynamic TVFF channel as well as the unknown noise variance are considered as three hidden states to be estimated. In particular, the dynamic transitions of both the PU state and fading channel are considered, which are assumed to evolve respectively as two finite states Markov chains (FSMC).

Second, a new sensing method for single-node is proposed to cope with various link uncertainties, which jointly detects/estimates three hidden states in real time, relying on the Bayesian statistics inference. Traditional Bayesian schemes, such as the expectation maximization (EM) algorithm, may be also applied to solve this problem. Due to the unavailability of likelihood functions (e.g., in the absence of PU), however the EM scheme can hardly address the error propagation in detection/estimation. Thus, such schemes tend to be less appealing in the context of information uncertainties. In order to deal with this challenging problem, rather than a simple combination or direct application of existing Bayesian methods, e.g., maximum *a posteriori* probability (MAP) or maximum likelihood (ML) method, a novel three-stage joint sensing and estimation algorithm is proposed. A promising marginalized particle filtering (MPF) technology is integrated to track multiple unknown states [20], which will be coupled with each other. By tracking realtime fading channels and estimating unknown noise variance, the spectrum sensing is recursively implemented. It is noteworthy that the formulated DSM and the designed sequential estimation scheme can also be generalized to another kind of observations, e.g., the non-coherent ED. Simulation results are provided to validate our new sensing scheme. Except for the improved sensing performance, the estimated channel gain as well as noise variance will further facilitate the subsequent resource allocation.

The rest of this paper is organized as follows. In Section II, we establish a unified DSM for spectrum sensing, which is characterized by time-variant fading channels and unknown noise variance. On this basis, a joint-estimation based sensing paradigm is designed in Section III, and an iterative algorithm is also proposed. Numerical simulations and performance analyses are provided in Section IV. Finally, we conclude the whole work in Section V.

II. SYSTEM MODEL

In this section, a unified dynamic state-space model is formulated to characterize spectrum sensing, where three hidden states, i.e., the PU emission signal $\mathbf{x}_n$, the PU-SU channel state a_n and unknown noise variance Σ, need to be tracked based on the observation y_n. In contrast to traditional schemes focusing on the instantaneous random behaviors, the time-variant dynamic of fading channel will be fully exploited. Our dynamic state-space model is given by:

$$\mathbf{x}_n = \Phi\left(\mathbf{x}_{n-1}\right), \tag{1}$$

$$a_n = \Psi\left(a_{n-1}\right), \tag{2}$$

$$y_n = \Omega\left(\mathbf{x}_n, a_n, \mathbf{z}_n\right). \tag{3}$$

Two hidden states, i.e., the PU emission signal $\mathbf{x}_n$ and the channel state a_n, dynamically evolve according to independently transitional functions $\Phi(\cdot)$ and $\Psi(\cdot)$, respectively. The unknown variance of measurement noise, i.e., Σ, is static (or at least, in a long period). The noisy observation y_n is specified via the measurement function $\Omega(\cdot)$. In other words, $\Omega(\cdot)$ gives the coupling relationship between the measurement y_n and three hidden states (i.e., $\mathbf{x}_n$, a_n and Σ in Gaussian noise $\mathbf{z}_n$).

A. DYNAMIC OF PU STATES

The emission state of PU comes into two alternative forms: active and inactive. For clarity, H_0 and H_1 respectively denote

two hypotheses, i.e., the inactive and active state of PU. Once the PU switches to the active state H_1, it will transmit a sequence of preamble signals, i.e., $\mathbf{x}_n = \mathbf{x}_c$. Without lack of generality, the binary sequence is assumed (note that the generalization to complex formats is straightforward), with a transmission power $E_{\mathbf{x}} = \mathbb{E}\{\mathbf{x}_c\mathbf{x}_c^T\}$ and a length M. Otherwise, it stays in the inactive state and emits no signals, i.e., we have $\mathbf{x}_n = \mathbf{0}_{1\times M}$ under H_0. Such two states switch to each other with specific transitional probabilities. Thus, a two-state Markov chain is used to model the evolution of PU states, with its transitional probability matrix (TPM):

$$\mathbf{P}_{\mathbf{x}} = \begin{bmatrix} p_{inactive\to inactive} & p_{inactive\to active} \\ p_{active\to inactive} & p_{active\to active} \end{bmatrix}, \\ = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}. \tag{4}$$

B. TVFF CHANNEL

For dynamic wireless environments, the Rayleigh fading is further assumed to be time-variant [21]. To be specific, the current fading state will be related with previous states. The evolution of channel states will be abstracted as a Markov chain, which is also characterized by another specific TPM. That is, the fading channels assumed in the analysis is the TVFF channels with short-term memory, which will be commonly encountered (e.g., in mobile communications). The probability distribution function (PDF) of Rayleigh fading distribution, with a scale parameter σ_R^2, is:

$$p(a) = \begin{cases} \dfrac{a}{\sigma_R^2} \times \exp\left(-\dfrac{a^2}{2\sigma_R^2}\right), & 0 \le a \le \infty, \\ 0, & a < 0. \end{cases} \tag{5}$$

As far as the TVFF amplitude is considered, it will vary randomly across a wide range. For two adjacent times, however the fading amplitude will be highly correlated. I.e., the current fading state is related with a previous one. In order to accommodate such dynamic correlations, various channel models have been suggested, including the Clarke's model, autoregressive (AR) model and the FSMC model. Owing to its effectiveness in modelling fading time-correlations and the convenience of analysis [21], a FSMC model is adopted in this analysis.

In a FSMC model, the time-variant fading amplitude is partitioned to K non-overlapping regions, i.e., $a_n = A_k \in [v_k, v_{k+1}), k = 0, 1, \cdots, K-1$. Each fading region $[v_k, v_{k+1})$ is represented by one feasible state A_k with the stationary probability $\pi_k \triangleq \int_{v_k}^{v_{k+1}} f(a)da$. As discussed in [22], a common strategy in constructing FSMC is the equal partition rule, i.e., let $\pi_k = 1/K$. Thus, the region boundaries is calculated by:

$$v_k = \sqrt{-2\sigma_R^2 \times \ln\left(1 - \frac{k}{K}\right)}, \quad k = 0, 1, \ldots, K-1, \tag{6}$$

and accordingly, the representative fading state is:

$$A_k = \frac{\int_{v_k}^{v_{k+1}} af(a)\,da}{\pi_k}, \quad k = 0, 1, \ldots, K-1. \tag{7}$$

The set of representative fading states is $\mathsf{A} = \{A_0, A_1, \ldots, A_{K-1}\}$. Given the computational simplicity and the ease of analysis, we use the first-order Markov chain to model the evolution of channel states. Thus, the current fading state is only related to its previous one, but statistically independent of all other past and future states. I.e., for two adjacent slots $n-1$ and n, the fading state will either stay in the same state A_k, or transits to its immediate neighboring states A_{k-1} or A_{k+1} [22]. For slow-varying fading conditions, the first-order FSMC has been shown accurate in analysis [21].

Accordingly, the TPM of fading states becomes a tri-diagonal matrix as in (8), as shown at the bottom of this page. Each element of the TPM, i.e., $p_{A_k\to A_{k^\dagger}}$ $(k, k^\dagger \in \{0, \cdots, K-1\})$, species the transition probability from A_k of time $n-1$ to $A_{k^\dagger}$ of time n, i.e.,

$$p_{A_k\to A_{k^\dagger}} \triangleq \Pr\left(a_n = A_{k^\dagger} | a_{n-1} = A_k\right), \\ = \frac{\int_{v_{k^\dagger}}^{v_{k^\dagger+1}} \int_{v_k}^{v_{k+1}} f(a_{n-1}, a_n)\, da_{n-1} da_n}{\pi_k}, \tag{9}$$

where $f(a_{n-1}, a_n)$ denotes the bivariate Rayleigh joint probability density function [21]. Note that, the channel phase θ_n can be also cast to a FSMC, which will follow a uniform distribution, i.e., $f(\theta) = \frac{1}{2\pi}$. For example, with the binary signals, a channel phase θ_n may take 0 or π with the equal probability [21]. For complex signals, a similar set of phase, i.e., $\Theta = \{\Theta_0, \Theta_1, \ldots, \Theta_{K_1-1}\}$, may be constructed.

C. OBSERVATIONS

For the MFD-based sensing in the presence of unknown fading channels and noise variance, the observation conditioned on two hypothesis is:

$$y_n = \begin{cases} \mathbf{x}_c^* \otimes \mathbf{z}_n, & H_0, \\ \mathbf{x}_c^* \otimes (\epsilon_n a_n \mathbf{x}_c + \mathbf{z}_n), & H_1. \end{cases} \tag{10}$$

$$\mathbf{P}_a = \begin{bmatrix} p_{A_0\to A_0} & p_{A_0\to A_1} & 0 & \cdots & 0 & 0 & 0 \\ p_{A_1\to A_0} & p_{A_1\to A_1} & p_{A_1\to A_2} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & & & \\ 0 & 0 & 0 & \cdots & p_{A_{K-2}\to A_{K-3}} & p_{A_{K-2}\to A_{K-2}} & p_{A_{K-2}\to A_{K-1}} \\ 0 & 0 & 0 & \cdots & 0 & p_{A_{K-1}\to A_{K-2}} & p_{A_{K-1}\to A_{K-1}} \end{bmatrix}. \tag{8}$$

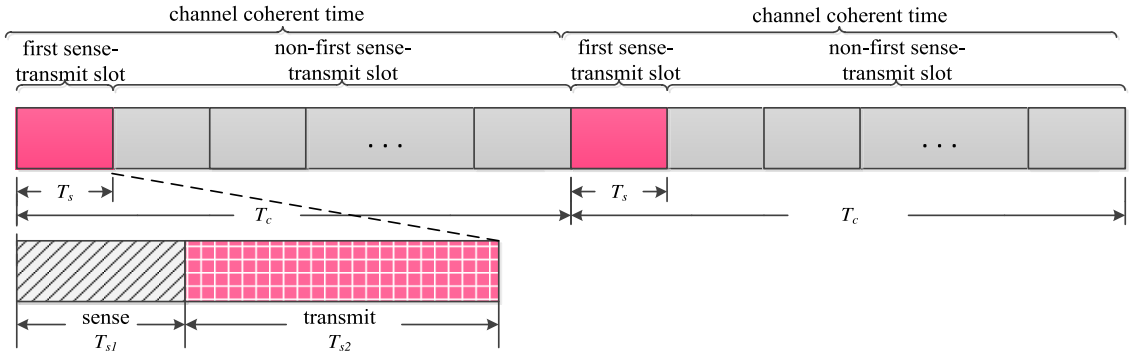


FIGURE 1. The frame structure of the spectrum sensing scheme.

Here, $\otimes$ in (10) denotes a convolution operation. The sampling size is M. $\mathbf{z}_n \triangleq [z_{n,0}, z_{n,1}, \ldots, z_{n,M-1}]^T$, and $z_{n,m}$ represents the additive white Gaussian noise (AWGN) with a mean μ and a variance Σ, i.e., $z_{n,m} \sim \mathcal{N}(\mu, \Sigma)$. The noise mean μ is assumed to be zero, whilst the noise variance Σ remains unknown but is static. $\epsilon_n \triangleq \exp(j\theta) \in \mathsf{B}$ accounts for unknown channel phase. For the binary sequences, it will be reduced to $\mathsf{B} \triangleq \{+1, -1\}$.[1] Due to the participation of unknown fading gains, it is seen from (10) that the received signal shows remarkable fluctuations.

As known, the above MFD detector requires both *a priori* PU's waveform (i.e., $\mathbf{x}_c$) and accurate CSI of PU-SU links (i.e., a_n and Σ), which will restrict its applications to some extents. The proposed scheme, premised on a unified DSM and a sequential estimation, is thereby designed to address realistic information uncertainty, which can be generalized to other observations, e.g., the non-coherent energy observation.

The frame structure is illustrated by Fig. 1, as in [25]. The channel states a_n are assumed to be invariant within several successive sense-transmit slots. That means, the coherence time of fading channels, i.e., $T_c \approx 1/f_D$, covers multiples sense-transmit slots with a duration of $T_s = T_{s1} + T_{s2}$ [23], [24]. Thus, we have $T_c = JT_s$ or $f_D T_s \approx T_s/T_c = 1/J < 1$. Here, f_D is the maximum Doppler frequency. To this end, a sense-transmission slot could be classified into two categories: first slot and non-first slot, and the channel state will possibly transit into another state only in the first slot.

III. JOINT ESTIMATION ALGORITHM

With an overall consideration of spectral utilization and mutual interference, the false alarm $p_f = p(\hat{\mathbf{x}}_n = \mathbf{x}_c|H_0) = 1 - p(\hat{\mathbf{x}}_n = \mathbf{0}|H_0)$ and the missing alarm $p_m = p(\hat{\mathbf{x}}_n = \mathbf{0}|H_1) = 1 - p(\hat{\mathbf{x}}_n = \mathbf{x}_c|H_1)$ are jointly considered. In this way, the total right detection probability is used as a metric of sensing performance, which is defined as [23]:

$$\begin{aligned} P_{TD} &\triangleq p_{\{D,0\}} + p_{\{D,1\}}, \\ &= p(\hat{\mathbf{x}}_n = \mathbf{0}|H_0)p(H_0) + p(\hat{\mathbf{x}}_n = \mathbf{x}_c|H_1)p(H_1), \end{aligned} \tag{11}$$

where $p_{\{D,0\}}$ represents the detection probability under H_0, i.e., $\mathbf{x}_n = \mathbf{0}$, and $p_{\{D,1\}}$ represents the detection probability under H_1, i.e., $\mathbf{x}_n = \mathbf{x}_c$.

Given the time-variant fading and ambient noises with unknown variance, most sensing schemes (e.g. ED and MFD) may be less attractive, i.e., by only averaging out the unfavorable uncertainty via its statistical distribution. Thus, a joint estimation paradigm is suggested, which will fully exploit the coupling relations between hidden states and observations. Thus, the main objective is to maximize the total right detection probability P_{TD}, by acquiring three hidden states, i.e.,

$$\begin{aligned} (\hat{\Sigma}, \hat{a}_{0:n}, \hat{\mathbf{x}}_{0:n}|y_{0:n}) &= \arg\max\,[P_{TD}], \\ &= \arg\max\left[p_{\{D,0\}} + p_{\{D,1\}}\right]. \end{aligned} \tag{12}$$

From a Bayesian perspective, the joint estimation will be implemented via the MAP criterion, i.e.,

$$(\hat{\Sigma}, \hat{a}_{0:n}, \hat{\mathbf{x}}_{0:n})^{\mathrm{MAP}} = \arg \max_{\substack{a_n \in \mathsf{A} \\ \mathbf{x}_n \in \{\mathbf{0}, \mathbf{x}_c\}}} [p(\Sigma, a_{0:n}, \mathbf{x}_{0:n}|y_{0:n})]. \tag{13}$$

With an objective of mixed detection and estimation in (13), it will be ineffective (or even infeasible) to apply classical Bayesian methods (e.g., MAP or ML) directly to solve this complicated problem. In (13), the mutual interruption of three unknown parameters will be inevitable, which makes most Bayesian methods infeasible. For example, if a PU is inactive (H_0), then there is no useful likelihood information available to estimate fading channels. Conversely, if there is no accurate link information (i.e., fading channels), then the PU cannot be detected accurately. More importantly, such two processes may affect each other, and the detection (or estimation) error will lead to the wrong estimation (or detection) in next slot.

In order to deal with the mixed detection and estimation problem, a recursive Bayesian algorithm is designed,

[1] Note that, since the time-correlation of phases is insignificant (i.e., following a uniform distribution), it will not be estimated in a recursive manner, while its effects will be considered lately when evaluating likelihood distributions. For complex format signals, alternatively the channel phase can be directly taken into a DSM and the joint estimations. In both cases, the subsequent estimation scheme will be similarly applicable.

as in (14), as shown at the bottom of this page. Based on an intuitive decomposition of (14), our proposed joint sensing method will consist of three phases. Firstly, the channel state is estimated based on MAP criterion; secondly, the PU state is detected using the particle filtering (PF) technology [26]; and finally, the noise variance is updated via a marginalization concept [27].

A. ESTIMATION OF TIME-VARIANT CHANNEL

From eq. (10), the likelihood information of fading channels will completely disappear under H_0, i.e., $\mathbf{x}_n = \mathbf{0}$, thus the Bayesian estimation will becomes impractical in this situation. Therefore, an estimation strategy is related with detection results. On the other hand, the fading channel is assumed to be slow-varying, which can be utilized to make the estimation result more accurate by using historical observations. Based on the above considerations, we further integrate coarse detection and accumulative modification. A schematic diagram of channel state estimation is illustrated in Fig. 2.

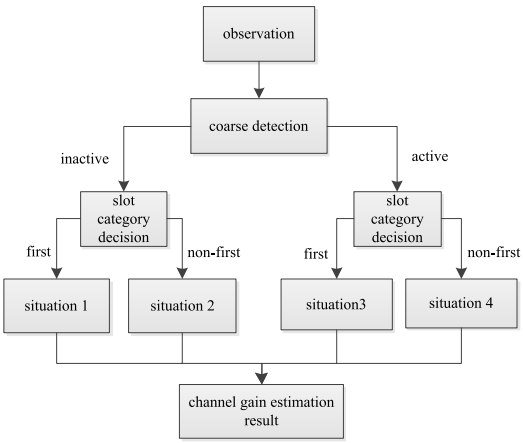


FIGURE 2. Diagram of the channel gain estimation.

1) COARSE DETECTION

The purpose of coarse detection is to obtain a rough estimation of PU states, which will facilitate different strategies in channel estimation. Although the accuracy of coarse detection is relatively low, subsequent processing will modify the erroneous results and accomplish the accurate detection. According to (10), the observation y_n under H_0 is related only with random noises $\mathbf{z}_n$, which determines the variance of y_n. The observation y_n under H_1 is related to the transmitting power $E_{\mathbf{x}}$, the temporal fading state a_n and the noise $\mathbf{z}_n$. Note that, in each sensing slot, the transmitting power $E_{\mathbf{x}}$ is fixed and known. For the adopted FSMC model, a_n belongs to a discrete representative states set A, which is hence also fixed but needs to be estimated. Thus, the first term in (10) will become a deterministic but unknown term, which has no contribution to the variance of y_n. Therefore, y_n follows normal distributions with the same variance but different expectations under H_0 and H_1, respectively, i.e.,

$$y_n \sim \begin{cases} \mathcal{N}(0, \|\mathbf{x}_c\|^2\Sigma), & H_0, \\ \mathcal{N}(\epsilon_n a_n E_{\mathbf{x}}, \|\mathbf{x}_c\|^2\Sigma), & H_1. \end{cases} \tag{15}$$

The initial estimation of PU states is derived via MAP criterion:

$$\mathbf{x}_n^{\dagger} = \arg \max_{\mathbf{x}_n \in \{\mathbf{0}, \mathbf{x}_c\}} \left[p\left(\mathbf{x}_n \mid y_n, a^{\dagger}, \hat{\Sigma}_{n-1}\right)\right]. \tag{16}$$

Premised on conservative estimation, $a^{\dagger}$ is the minimum of channel gain set and can be computed by $a^{\dagger} = \min(\mathsf{A})$. The posterior probability in (16) can be expressed as:

$$p\left(\mathbf{x}_n \mid y_n, a^{\dagger}, \hat{\Sigma}_{n-1}\right) = \frac{p\left(y_n \mid \mathbf{x}_n, a^{\dagger}, \hat{\Sigma}_{n-1}\right) p\left(\mathbf{x}_n\right)}{\sum_{\mathbf{x}_n \in \{\mathbf{0}, \mathbf{x}_c\}} p\left(y_n \mid \mathbf{x}_n, a^{\dagger}, \hat{\Sigma}_{n-1}\right) p\left(\mathbf{x}_n\right)}, \tag{17}$$

and the likelihood function can be computed as:

$$p\left(y_n \mid \mathbf{x}_n, a^{\dagger}, \hat{\Sigma}_{n-1}\right) = \begin{cases} \dfrac{1}{\sqrt{2\pi \|\mathbf{x}_c\|^2 \hat{\Sigma}_{n-1}}} \exp\left[-\dfrac{\|y_n\|^2}{2\|\mathbf{x}_c\|^2 \hat{\Sigma}_{n-1}}\right], & \mathbf{x}_n = \mathbf{0}, \\ \max\limits_{\epsilon_n \in \mathsf{B}} \dfrac{1}{\sqrt{2\pi \|\mathbf{x}_c\|^2 \hat{\Sigma}_{n-1}}} \exp\left[-\dfrac{\left\|y_n - \epsilon_n E_{\mathbf{x}} a^{\dagger}\right\|^2}{2\|\mathbf{x}_c\|^2 \hat{\Sigma}_{n-1}}\right], & \mathbf{x}_n = \mathbf{x}_c. \end{cases} \tag{18}$$

$$\begin{aligned} &\left(\hat{\Sigma}_n, \hat{a}_n, \hat{\mathbf{x}}_n\right)^{\mathrm{MAP}} \\ &= \arg \max_{\substack{a_n \in \mathsf{A} \\ \mathbf{x}_n \in \{\mathbf{0}, \mathbf{x}_c\}}} \left[p\left(\Sigma_n, a_n, \mathbf{x}_n | \hat{\Sigma}_{0:n-1}, \hat{a}_{0:n-1}, \hat{\mathbf{x}}_{0:n-1}, y_{0:n}\right)\right], \\ &= \arg \max_{\substack{a_{0:n} \in \mathsf{A} \\ \mathbf{x}_{0:n} \in \{\mathbf{0}, \mathbf{x}_c\}}} \left[\underbrace{p\left(a_n | \hat{\Sigma}_{0:n-1}, \hat{a}_{0:n-1}, \hat{\mathbf{x}}_{0:n-1}, y_{0:n}\right)}_{\text{MAP criterion}} \underbrace{\underbrace{p\left(\mathbf{x}_n | \hat{\Sigma}_{0:n-1}, \hat{a}_{0:n}, \hat{\mathbf{x}}_{0:n-1}, y_{0:n}\right)}_{\text{PF technology}} \underbrace{p\left(\Sigma_n | \hat{\Sigma}_{0:n-1}, \hat{a}_{0:n}, \hat{\mathbf{x}}_{0:n}, y_{0:n}\right)}_{\text{marginalization concept}}}_{\text{MPF}} \right] \end{aligned} \tag{14}$$

2) SLOT CATEGORY DECISION

As mentioned, the static fading time covers several sense-transmit periods. The transition of fading states occurs only in some switching times, i.e., $\lfloor nT_s/T_c \rfloor$, which is refereed as the first sense-transmit slot. More specifically, the first (or non-first) sense-transmit slots are determined by:

$$\mathrm{mod}(nT_s, T_c) \begin{cases} = 0, & \text{first slot}, \\ \neq 0, & \text{non}-\text{first slot}, \end{cases} \quad n = 0, 1, \ldots, N-1. \tag{19}$$

3) CHANNEL ESTIMATION FOR DIFFERENT SITUATIONS

For various coarse detection results (i.e., $\mathbf{x}_n^\dagger$) and different slots (e.g., first switching or non-first), the channel estimations will be implemented respectively according to the following four cases, as in Fig. 2.

Situation 1: In the case, the fading state may transit to another state or stay invariant. Yet, due to the coarse detection result $\mathbf{x}_n^\dagger = \mathbf{0}$, there is little information of observation can be utilized to estimate the channel gain, and the MAP estimation will be infeasible. So, we have to obtain the estimation based on the prior transition property, i.e.,

$$\hat{a}_n = \arg\max_{a_n \in \mathsf{A}} \left[p(a_n | \hat{a}_{pre}) \right], \tag{20}$$

where $\hat{a}_{pre}$ denotes the estimated channel of the previous time, i.e., $\hat{a}_{pre} = \hat{a}_{\lfloor nT_s/T_c \rfloor - 1}$. Each prior transition probability has been specified by the TPM in (8).

Situation 2: In the case, there is neither useful likelihood information (e.g., H_0) nor the transition of fading channels (i.e., the non-first slot). Thus, the estimated channel state remains as the same with the previous slot, i.e.,

$$\hat{a}_n = \hat{a}_{n-1}. \tag{21}$$

Situation 3: In this case, the observation and the related likelihood information will be exploited. Based on an MAP criterion, the fading channel is estimated by (22), as shown at the bottom of this page. Here, $\hat{\Sigma}_{n-1}$ denotes the estimation of noise variance in the previous slot. Recall that the evolution of channel states is independent of PU states, noise and observation, which is characterized by a first-order FSMC. Thus, the current channel gain a_n is only related to its previous state $\hat{a}_{pre}$. Furthermore, we may simplify the prior probability of channel states, i.e., $p\left(a_n | \hat{\Sigma}_{n-1}, \hat{a}_{0:n-1}, \hat{\mathbf{x}}_{0:n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, y_{0:n-1}\right) = p\left(a_n | \hat{a}_{pre}\right)$.

For the MF observation, the likelihood function $p(y_n|\cdot)$ follows a normal distribution, i.e.,

$$\begin{aligned} & p\left(y_n | \hat{\Sigma}_{n-1}, \hat{a}_{0:n-1}, a_n, \hat{\mathbf{x}}_{0:n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, y_{0:n-1}\right) \\ &= \max_{\epsilon_n \in \mathsf{B}} p(y_n | \hat{\Sigma}_{n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, a_n, \epsilon_n), \\ &= \max_{\epsilon_n \in \mathsf{B}} \frac{1}{\sqrt{2\pi \|\mathbf{x}_c\|^2 \hat{\Sigma}_{n-1}}} \exp\left[-\frac{\|y_n - \epsilon_n E_{\mathbf{x}} a_n\|^2}{2\|\mathbf{x}_c\|^2 \hat{\Sigma}_{n-1}} \right] \end{aligned} \tag{23}$$

Note that, in order to mitigate the effects from unknown channel phases, a similar ML concept is used here.

In each first slot, the accumulative operation will be activated. That is, we will configure the accumulation observation and counter to $Y_n = y_n$ and $M_n = 1$, respectively. Such two variables (M_n and Y_n) are introduced in order to run accumulative mechanisms. Superficially, the accumulation counter M_n is used to record the event of $\mathbf{x}_n^\dagger = \mathbf{x}_c$ until the current n-th slot within the static period (i.e., $0 \leqslant M_n \leqslant J$), while the accumulation observation Y_n collect the historical information corresponding to these events.

Situation 4: Within the static time, the historical information will be utilized to further modify the estimation of channel states. First, the accumulation observation and counter are updated via:

$$Y_n = Y_{n-1} + y_n, \tag{24}$$

$$M_n = M_{n-1} + 1. \tag{25}$$

Then, the channel state will be estimated, conditioned on the updated accumulation observation and counter, i.e.,

$$\begin{aligned} \hat{a}_n &\propto \arg\max_{a_n \in \mathsf{A}} [p(Y_n | \hat{\Sigma}_{n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, a_n) \times p(a_n | \hat{a}_{pre})], \\ &= \arg\max_{a_n \in \mathsf{A}} \{ \max_{\epsilon_n \in \mathsf{B}} \frac{1}{\sqrt{2\pi M_n \|\mathbf{x}_c\|^2 \hat{\Sigma}_{n-1}}} \\ &\quad \times \exp\left(-\frac{\|Y_n - \epsilon_n M_n E_{\mathbf{x}} a_n\|^2}{2M_n \|\mathbf{x}_c\|^2 \hat{\Sigma}_{n-1}} \right) \times p(a_n | \hat{a}_{pre}) \}. \end{aligned} \tag{26}$$

B. DETECTION OF PU STATES

In the absence of channel information, the coarse detection will be inaccurate. Based on the PF scheme, an iterative scheme is developed to further modify the inaccurate

$$\begin{aligned} \hat{a}_n &= \arg\max_{a_n \in \mathsf{A}} \left[p\left(a_n | \hat{\Sigma}_{n-1}, \hat{a}_{0:n-1}, \hat{\mathbf{x}}_{0:n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, y_{0:n}\right) \right], \\ &= \arg\max_{a_n \in \mathsf{A}} \left[\frac{p\left(y_n | \hat{\Sigma}_{n-1}, \hat{a}_{0:n-1}, a_n, \hat{\mathbf{x}}_{0:n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, y_{0:n-1}\right) p\left(a_n | \hat{\Sigma}_{n-1}, \hat{a}_{0:n-1}, \hat{\mathbf{x}}_{0:n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, y_{0:n-1}\right)}{\sum_{a_n \in \mathsf{A}} p\left(y_n | \hat{\Sigma}_{n-1}, \hat{a}_{0:n-1}, a_n, \hat{\mathbf{x}}_{0:n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, y_{0:n-1}\right) p\left(a_n | \hat{\Sigma}_{n-1}, \hat{a}_{0:n-1}, \hat{\mathbf{x}}_{0:n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, y_{0:n-1}\right)} \right], \\ &\propto \arg\max_{a_n \in \mathsf{A}} \left[p\left(y_n | \hat{\Sigma}_{n-1}, \hat{a}_{0:n-1}, a_n, \hat{\mathbf{x}}_{0:n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, y_{0:n-1}\right) p\left(a_n | \hat{\Sigma}_{n-1}, \hat{a}_{0:n-1}, \hat{\mathbf{x}}_{0:n-1}, \mathbf{x}_n^\dagger = \mathbf{x}_c, y_{0:n-1}\right) \right]. \end{aligned} \tag{22}$$

$$\begin{aligned}
&p\left(y_n|\mathbf{x}_{n-1}^{(i)}=\mathbf{0},\hat{a}_n,\hat{\Sigma}_{n-1}\right)\\
&\quad=p\left(y_n|\mathbf{x}_n=\mathbf{0},\hat{a}_n,\hat{\Sigma}_{n-1}\right)p\left(\mathbf{x}_n=\mathbf{0}|\mathbf{x}_{n-1}^{(i)}=\mathbf{0}\right)+p\left(y_n|\mathbf{x}_n=\mathbf{x}_c,\hat{a}_n,\hat{\Sigma}_{n-1}\right)p\left(\mathbf{x}_n=\mathbf{x}_c|\mathbf{x}_{n-1}^{(i)}=\mathbf{0}\right),\\
&\quad=\frac{1}{\sqrt{2\pi\|\mathbf{x}_c\|^2\hat{\Sigma}_{n-1}}}\exp\left(-\frac{y_n^2}{2\|\mathbf{x}_c\|^2\hat{\Sigma}_{n-1}}\right)p_{00}+\max_{\epsilon_n\in\mathsf{B}}\frac{1}{\sqrt{2\pi\|\mathbf{x}_c\|^2\hat{\Sigma}_{n-1}}}\exp\left(-\frac{\left\|y_n-\epsilon_nE_{\mathbf{x}}\hat{a}_n\right\|^2}{2\|\mathbf{x}_c\|^2\hat{\Sigma}_{n-1}}\right)p_{01},
\end{aligned}\tag{30}$$

$$\begin{aligned}
&p\left(y_n|\mathbf{x}_{n-1}^{(i)}=\mathbf{x}_c,\hat{a}_n,\hat{\Sigma}_{n-1}\right)\\
&\quad=p\left(y_n|\mathbf{x}_n=\mathbf{0},\hat{a}_n,\hat{\Sigma}_{n-1}\right)p\left(\mathbf{x}_n=\mathbf{0}|\mathbf{x}_{n-1}^{(i)}=\mathbf{x}_c\right)+p\left(y_n|\mathbf{x}_n=\mathbf{x}_c,\hat{a}_n,\hat{\Sigma}_{n-1}\right)p\left(\mathbf{x}_n=\mathbf{x}_c|\mathbf{x}_{n-1}^{(i)}=\mathbf{x}_c\right),\\
&\quad=\frac{1}{\sqrt{2\pi\|\mathbf{x}_c\|^2\hat{\Sigma}_{n-1}}}\exp\left(-\frac{y_n^2}{2\|\mathbf{x}_c\|^2\hat{\Sigma}_{n-1}}\right)p_{10}+\max_{\epsilon_n\in\mathsf{B}}\frac{1}{\sqrt{2\pi\|\mathbf{x}_c\|^2\hat{\Sigma}_{n-1}}}\exp\left(-\frac{\left\|y_n-\epsilon_nE_{\mathbf{x}}\hat{a}_n\right\|^2}{2\|\mathbf{x}_c\|^2\hat{\Sigma}_{n-1}}\right)p_{11}.
\end{aligned}\tag{31}$$

detection results. Specifically, the posterior probability of interest $p(\mathbf{x}_n|y_{0:n},\mathbf{x}_{0:n-1},a_{0:n})$, which is unfortunately non-analytical, is approximated by a group of particles with associated weights [28]. Then, the MAP estimation is numerically obtained, i.e.,

$$\hat{\mathbf{x}}_n^{(\mathrm{MAP})}=\arg\max_{\mathbf{x}_n\in\{\mathbf{0},\mathbf{x}_c\}}\left\{\sum_{i=1}^{I}\delta\left(\mathbf{x}_n-\mathbf{x}_n^{(i)}\right)\times w_n^{(i)}\right\}.\tag{27}$$

The implementation of PF, for the considered scenarios, will basically involve the following three steps [24], [29], [30].

Firstly, the sequential importance sampling (SIS) [26] process is applied, i.e., a group of random particles are simulated from an important distribution function, i.e., $\mathbf{x}_n^{(i)}\sim\pi(\mathbf{x}_n|\mathbf{x}_{0:n-1}^{(i)},\hat{a}_n,\hat{\Sigma}_{n-1})$. Here, the important density is:

$$\begin{aligned}
&\pi(\mathbf{x}_n|\mathbf{x}_{0:n-1}^{(i)},\hat{a}_n,\hat{\Sigma}_{n-1}),\\
&\quad\triangleq p\left(\mathbf{x}_n|\mathbf{x}_{0:n-1}^{(i)},\hat{a}_n,y_n,\hat{\Sigma}_{n-1}\right),\\
&\quad\propto p\left(y_n|\mathbf{x}_n,\hat{a}_n,\hat{\Sigma}_{n-1}\right)\times p\left(\mathbf{x}_n|\mathbf{x}_{n-1}^{(i)}\right),
\end{aligned}\tag{28}$$

where $p\left(\mathbf{x}_n|\mathbf{x}_{n-1}^{(i)}\right)$ denotes the prior transition probability of PU states.

Secondly, after sampling from the importance distribution, the associated importance weights are updated via:

$$w_n^{(i)}=w_{n-1}^{(i)}\times p\left(y_n|\mathbf{x}_{n-1}^{(i)},\hat{a}_n,\hat{\Sigma}_{n-1}\right),\tag{29}$$

where likelihood functions are calculated by (30) and (31), as shown at the top of this page.

Finally, the MAP estimation is derived, premised on the simulated particles and the associated weights, as in (27). If $\hat{\mathbf{x}}_n=\mathbf{x}_c$, the PU transmitter is active at time n, and the hypothesis H_1 is true. Otherwise, $\hat{\mathbf{x}}_n=\mathbf{0}$, then H_0 is true.

C. ESTIMATION OF NOISE VARIANCE

After acquiring of fading channel gain and PU state, the noise variance will be then updated via the following two steps.

Firstly, relying on a Bayesian framework and the conception of conjectured prior, we assume one suitable prior distribution for unknown noise variance and, therefore, the posterior distribution can be derived then. Further, the hyper-parameters of this posterior distribution will be updated based on updated particles. Secondly, the estimation of noise variance will be obtained relying on a marginalization technique.

Without losing generality, the conjectured prior of unknown noise variance is assumed to be an Inverse-Gamma distribution [27], and its PDF is expressed as:

$$p(\Sigma)=\frac{\beta_0^{\gamma_0}}{\Gamma(\gamma_0)}\left(\frac{1}{\Sigma}\right)^{\gamma_0+1}\times\exp\left(-\frac{\beta_0}{\Sigma}\right),\quad\Sigma>0,\tag{32}$$

where β_0 and γ_0 denote two hyper-parameters of an Inverse-Gamma distribution.

With the help of conjectured prior information, the posterior distribution $p(\Sigma|\cdot)$ follows also an Inverse-Gamma distribution. A detailed derivation is given by (33), see the next page.

With some manipulation, the statistical density of unknown noise variance is expressed as:

$$\begin{aligned}
&\Sigma|\mathbf{x}_{0:n}^{(i)},\hat{a}_{0:n},y_{0:n}\\
&\sim\mathcal{IG}\left(\gamma_{n-1}^{(i)}+\frac{1}{2},\beta_{n-1}^{(i)}+\frac{\min_{\epsilon_n\in\mathsf{B}}\left\|y_n-\epsilon_n\hat{a}_n\mathbf{x}_n^{(i)}\mathbf{x}_c^T\right\|^2}{2\|\mathbf{x}_c\|^2}\right).
\end{aligned}\tag{34}$$

Based on (34), the hyper-parameters of the posterior distribution will be updated via:

$$\gamma_n^{(i)}=\gamma_{n-1}^{(i)}+\frac{1}{2},\tag{35}$$

$$\beta_n^{(i)}=\beta_{n-1}^{(i)}+\frac{\min_{\epsilon_n\in\mathsf{B}}\left\|y_n-\epsilon_n\hat{a}_n\mathbf{x}_n^{(i)}\mathbf{x}_c^T\right\|^2}{2\|\mathbf{x}_c\|^2}.\tag{36}$$

On this basis, the mean of noise variance is then written as:

$$\mathbb{E}\left(\Sigma|\mathbf{x}_{0:n}^{(i)},\hat{a}_{0:n},y_{0:n}\right)=\frac{\beta_n^{(i)}}{\gamma_n^{(i)}-1},\quad\gamma_n^{(i)}>1.\tag{37}$$

$$
\begin{aligned}
p\left(\Sigma|\mathbf{x}_{0:n}^{(i)}, \hat{a}_{0:n}, y_{0:n}\right) &\propto p\left(\mathbf{x}_{0:n}^{(i)}, \hat{a}_{0:n}, y_{0:n}|\Sigma\right) p\left(\Sigma\right), \\
&= \frac{1}{\left(2\pi\|\mathbf{x}_c\|^2\Sigma\right)^{n/2}} \exp\left[-\frac{\sum_{n^\dagger=0}^{n}\left(\min_{\epsilon_n\in\mathsf{B}}\left\|y_{n^\dagger}-\epsilon_n\hat{a}_{n^\dagger}\mathbf{x}_{n^\dagger}^{(i)}\mathbf{x}_c^T\right\|^2\right)}{2\|\mathbf{x}_c\|^2\Sigma}\right]\frac{\beta_0^{\gamma_0}}{\Gamma\left(\gamma_0\right)}\left(\frac{1}{\Sigma}\right)^{\gamma_0+1}\exp\left(-\frac{\beta_0}{\Sigma}\right), \\
&\propto \frac{\beta_0^{\gamma_0}}{\Gamma\left(\gamma_0\right)}\left(\frac{1}{\Sigma}\right)^{\gamma_0+1+n/2}\exp\left[-\frac{\beta_0+\sum_{n^\dagger=1}^{n}\left(\min_{\epsilon_n\in\mathsf{B}}\left\|y_{n^\dagger}-\epsilon_n\hat{a}_{n^\dagger}\mathbf{x}_{n^\dagger}^{(i)}\mathbf{x}_c^T\right\|^2\right)\Big/\left(2\|\mathbf{x}_c\|^2\right)}{\Sigma}\right], \\
&= \frac{\beta_0^{\gamma_0}}{\Gamma\left(\gamma_0\right)}\left(\frac{1}{\Sigma}\right)^{\gamma_{n-1}^{(i)}+1/2}\exp\left[-\frac{\beta_{n-1}^{(i)}+\left(\min_{\epsilon_n\in\mathsf{B}}\left\|y_n-\epsilon_n\hat{a}_n\mathbf{x}_n^{(i)}\mathbf{x}_c^T\right\|^2\right)\Big/\left(2\|\mathbf{x}_c\|^2\right)}{\Sigma}\right].
\end{aligned}
\tag{33}
$$

By resorting to the marginalization technique, the marginal posterior of Σ at time n is computed via:

$$
\begin{aligned}
p\left(\Sigma|y_{0:n}\right) &= \int p\left(\Sigma|\mathbf{x}_{0:n}, \hat{a}_{0:n}, y_{0:n}\right) d\mathbf{x}_{0:n}, \\
&\approx \sum_{i=1}^{I} p\left(\Sigma|\mathbf{x}_{0:n}^{(i)}, \hat{a}_{0:n}, y_{0:n}\right) \times w_n^{(i)}.
\end{aligned}
\tag{38}
$$

Finally, the unknown noise variance will be estimated via the following unbiased estimation, i.e.,

$$
\begin{aligned}
\hat{\Sigma}_n &= \mathbb{E}(\Sigma|y_{0:n}), \\
&\approx \sum_{i=1}^{I}\mathbb{E}(\Sigma|\mathbf{x}_{0:n}^{(i)}, \hat{a}_{0:n}, y_{0:n}) \times w_n^{(i)}, \\
&= \sum_{i=1}^{I}\frac{\beta_n^{(i)}}{\gamma_n^{(i)}-1} \times w_n^{(i)}.
\end{aligned}
\tag{39}
$$

It is noteworthy that our study focuses on a mixed detection and estimation problem, while the works of [26] and [27] deal only with a pure estimation problem, by using PF and MPF technologies. Thus, the developed scheme can be regarded as a general joint estimation and detection framework, which utilizes MPF to estimate unknown noise.

D. IMPLEMENTATION

Based on the elaborations above, the new joint detection algorithm is summarized as follows.

E. COMPLEXITY

Based on the above analysis, we may evaluate the complexity in terms of the multiplication operations. Firstly, in order to get observations, $\mathcal{O}(M)$ multiplication operations is required. Secondly, the complexity of sequential estimations is independent of a sampling size M, but proportional to the number of particle, and the complexity is $\mathcal{O}(QI)$, where the multiplication operation in calculating likelihood is denoted by Q. To sum up, the complexity will be measured by $\mathcal{O}(M+QI)$.

IV. NUMERIC SIMULATIONS

In the following numerical simulation and performance evaluations, a counterpart method, i.e., traditional MFD sensing, is used. In the context of unknown fading channels and noise variance, its decision threshold will be determined by also maximizing P_{TD}, i.e.,

$$
\tau_{\text{MFD}} = \arg\max_{\tau_{\text{MFD}}}\left[p_{\{D,0\}}+p_{\{D,1\}}\right], \tag{40}
$$

where only the partial information of CSI can be available, i.e., $\bar{a} = \mathbb{E}\{\mathsf{A}\}$ and $\bar{\theta} = \mathbb{E}\{\theta\}$ will be used in classical MF-based sensing.

In following simulations, we equivalently focus on the uncertainty of SNR aroused by unknown noise variance, which is randomly ranged in $[-\varepsilon, \varepsilon]$ dB. For clarity, the true value and its initial estimation of signal-to-noise ratio (SNR) are denoted by SNR and SNR_0, respectively. According to [14], when the noise uncertainty is involved, then the initial estimation Σ_0 will be distributed randomly around its true value Σ, i.e. $\Sigma_0 \sim [(1/\rho)\Sigma, \rho\Sigma]$. Here, $\rho > 1$ is a parameter that quantifies the level of uncertainty. When it comes to the SNR metric, i.e., $SNR = 10\log(E_{\mathbf{x}}/\Sigma)$, SNR_0 will be ranged randomly in $[SNR-\varepsilon, SNR+\varepsilon]$, and $\varepsilon = 10\log(\rho)$.

It is considered that the proposed algorithm is not relevant to the signal symbol transmission rate, and hence, we use a normalized sensing slot time, and in each sense-transmit slot total M samples are assumed. In practice, this sensing time can be either set to 1 ms or even to 30 ms, thus the sampling rate will correspond to 1KHz or 30KHz when $M = 10$.

As mentioned, the coherent time of time-variant fading channels, denoted by T_c, will cover many sensing slots with a interval T_s, i.e., $T_c = J \times T_s$. For the slow-varying fading

Algorithm 1

For $n = 0 : N - 1$

1) Estimate the fading state based on MAP criterion:
   - Perform the coarse detection and simultaneously determine whether the current time is the first switching slot;
   - For various cases, the fading state is updated via the specific mechanism. Here, the noise variance in calculating likelihood functions is the estimation result of the previous slot (recall that the noise variance is assumed to be static).

2) Based on the estimation of time-dependent fading states, the PU's emission state is identified via PF:
   - Draw discrete particles by sampling from an importance function;
   - Update their associated weights recursively;
   - Obtain the MAP estimation result via numerical approximations.

3) Update the noise variance estimation:
   - Update the hyper-parameters in *a posteriori* distribution of noise variance;
   - Obtain the estimation of noise variance via the marginalization concept.

End

channels [31], the static length $J = 1/f_d T_s$ may range from 10 to 50, and correspondingly, the normalized Doppler frequency shift $f_D T_s$ will range from 0.02 to 0.1. In practice, if the sensing time is set to 1ms, the coherent time of channel will range from 10ms to 50ms. The statistical property of fading channel is set to be Rayleigh distribution with variance $\sigma_R^2 = 1.5$. Given the total number of representative states is $K = 8$, then the feasible states A and the corresponding TPM $\mathbf{P}_a$ will be calculated based on (7)-(9).

A. DETECTION PERFORMANCE

We firstly study the effect on sensing performance from various value of $f_D T_s$, the uncertainty boundary ε and the partitioning number K, respectively. Three typical configurations of ε are adopted, i.e., $\varepsilon = 3,\ 5,\ 10$, while the sample size M is 15. In Fig. 3, it is seen that, compared with traditional MFD scheme, the sensing performance of our new algorithm will be improved significantly. For example, when the total right detection probability P_{TD} is 0.9, a rough detection gain around 3dB will be obtained by the new algorithm.

Then, we compared the sensing performance of our joint estimation scheme and other existing methods of [19], in which only the time-variant fading channel is estimated while the noise uncertainty is not considered. In the simulation, the noise variance in the scheme of [19] was replaced by one inaccurate estimation Σ_0. From the simulation results in Fig. 4, we find that the noise variance will have little effects on sensing performance, given the uncertain range $\varepsilon < 10$ dB. However, as the uncertain range ε increases, the

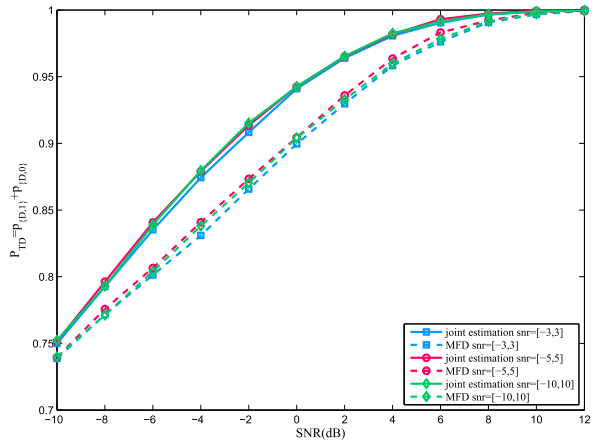


FIGURE 3. Sensing performance of the proposed joint estimation algorithm and traditional MFD method under different float rang boundary ε.

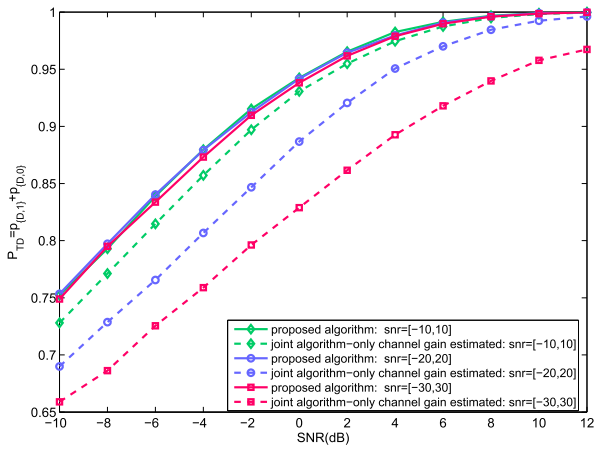


FIGURE 4. Sensing performance of the proposed algorithm and our past method under different float rang boundary ε.

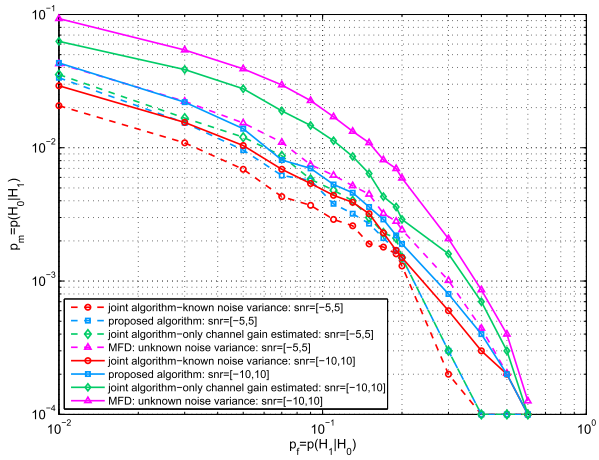


FIGURE 5. ROC curves of the proposed algorithm and our past method under different float rang boundary ε.

advantage of the new scheme, by jointly estimating unknown noise variance, will become remarkable.

In Fig. 5, the receiver operation character (ROC) curve are further provided. Here, both the proposed scheme and its counterpart (i.e., MFD without estimating unknown fading)

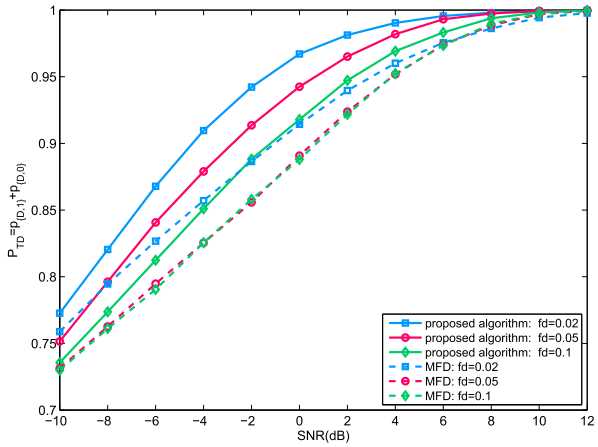


FIGURE 6. Sensing performance of the proposed algorithm and traditional MFD method under different maximum Doppler frequency shifts f_D.

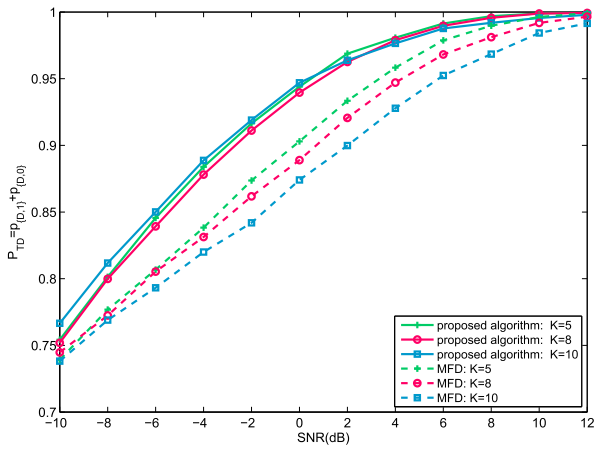


FIGURE 7. Sensing performance of the proposed algorithm and our past method under different channel state number K.

are assumed to have no *priori* knowledge on the unknown noise variance. In the analysis, SNR is set to 4dB, $f_D T_s = 0.05$ and $M = 15$. From the numerical result of ROC curves, it is demonstrated that the new algorithm can significantly outperform a classical MFD sensing scheme, by acquiring unknown noise variance and time-variant channel state jointly (when performing spectrum sensing).

Furthermore, the effects from various fading channels, with different maximum Doppler shifts, are studied. In the simulations, ε is 5, and K is set to 8. The normalized maximum Doppler shift is set to be $f_D T_s \in \{0.1, 0.05, 0.02\}$, i.e., $J \in \{10, 20, 50\}$. From Fig. 6, the sensing performance of the new method will decrease as $f_D T_s$ increases. This is easy to understand, i.e., the estimation of fading states will be refined via the accumulative modifications. To this end, the slower the channel varies, the more information of coherent slots will be utilized to promote the estimations. In comparison, if a channel changes too fast, then the refinement will be inadequate. As a result, the estimation of PU states will be affected by the coarse detection and tends to be inaccurate.

Finally, we study the effects on sensing performance from the representative states number K. It is suggested that,

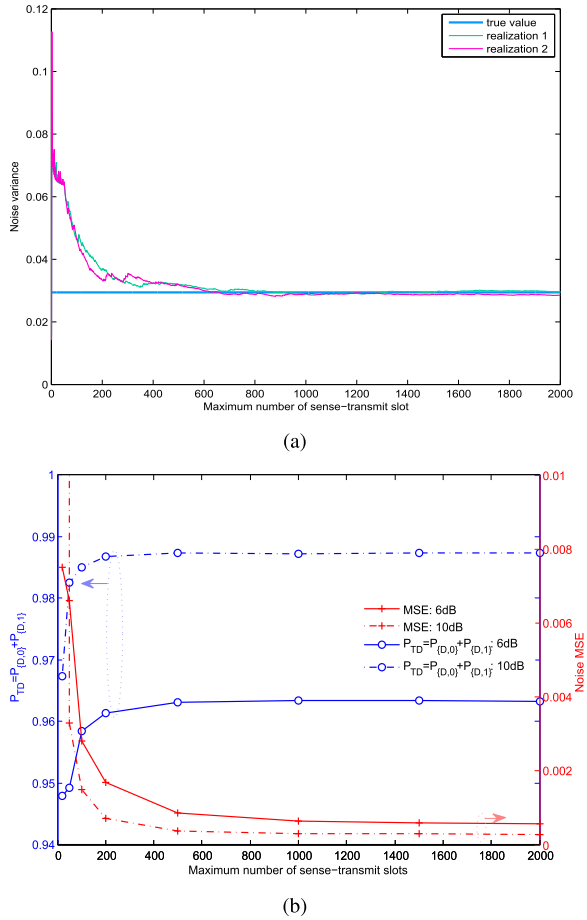


FIGURE 8. MSE vs Iteration number of sense-transmit slots. (a) different realization. (b) influence on different performance from estimation error.

the larger K is, the higher representation accuracy a FSMC model has, and also the more complicated the estimation algorithm is [21]. In order to have the balance between accuracy and complexity, the number of representative states K can be set from 6 to 64, as suggested in [21]. In Fig. 7, different values (e.g., $K = 5, 8, 10$) are investigated via numerical simulations, which are shown to have little influences on sensing performances (recall that the fading states have been jointly estimated by our proposed scheme).

B. ESTIMATION PERFORMANCE

It is shown by Fig. 8(a) that the estimation of unknown noise variance will be more accurate, as the time slots n increases. In the simulations, we configure $f_D T_s = 0.1$, $M = 15$ and $SNR = 6$dB. It is found that the required numbers of sensing slots, in order to achieve the convergence of noise estimation, may range from 500 to 1000. Thus, despite the time-variant fading states, the static unknown variance can be tracked via the proposed scheme. In Fig. 8(b), we further demonstrated the change of detection performance, as the number of sensing slots increases. As noted, the residual errors of noise estimation will be reduced, when the sensing slots increases. Simultaneously, we find that the detection performance will be also affected by different residual errors.

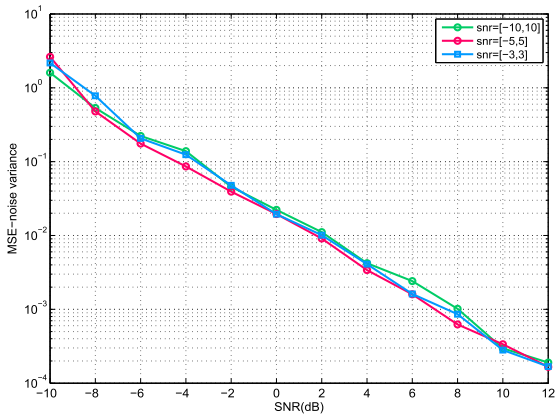


FIGURE 9. MSE of estimated variance for the noise.

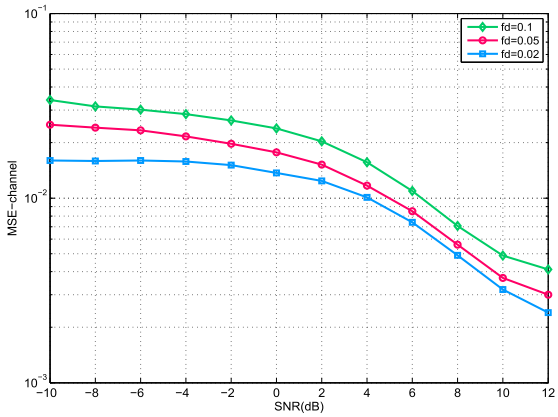


FIGURE 10. MSE of estimated time-variant channel gain.

For a low *SNR* (e.g., 6dB), the detection performance will be remarkably affected by residual errors. However, when the *SNR* is relatively high (e.g., 10dB), then the detection performance may be affected slightly by such residual errors. From the sensing performance point of view, the necessary number of sensing slots will be relatively small, e.g., around 400.

Then, the estimation MSE of unknown noise variance is also evaluated, under different uncertain boundaries ε. Numerical results are shown in Fig. 9. It seems that the uncertain boundary ε may have little influences on the estimation MSE of unknown noise variance. In other words, in different cases the unknown noise variance will be estimated accurately by the proposed joint estimation scheme.

The estimation MSE performance of time-variant fading channels, under different maximum Doppler shifts, is plotted in Fig. 10. It is seen from numerical results that the estimation accuracy may be degraded, as the maximum Doppler shift $f_D T_s$ increases. As analyzed from the previous Fig. 6, the slower fading channel permits the more sufficient accumulation of historical information, and thereby produces the more accurate estimation of fading gains.

V. CONCLUSIONS

A new spectrum sensing scheme is designed to address realistic challenges aroused by time-variant fading channels and unknown noise variance. We formulate a novel DSM to model the sensing process, in the absence of exact noise variance, which gives also the full considerations to dynamic PU states and channel fading. With the new model, spectrum sensing is implemented by acquiring time-variant channel states as well as unknown noise variance jointly. Simulation results are provided to validate our new algorithm. The formulated DSM provides a powerful tool for other signal processing of CR, e.g., spectrum sensing in mobile scenarios. Except for the improved sensing performance, the estimated noise variance and fading channels, as the additional link information, may greatly facilitate subsequence network optimizations. In conclusion, our sensing scheme will be of great promise to CR networks. Future works includes the designing of sensing schemes in more complicated environments, e.g., the time-variant noise variances.

REFERENCES

[1] *Spectrum Policy Task Force Report*, document ET Docket no. 02-155, Federal Communication Commission, Nov. 2002.
[2] J. Mitola and G. Q. Maguire, Jr., "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Apr. 1999.
[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
[4] B. Wang and K. J. R. Liu, "Advances in cognitive radio networks: A survey," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 5–23, Feb. 2011.
[5] Y.-C. Liang, K.-C. Chen, G. Y. Li, and P. Mahonen, "Cognitive radio networking and communications: An overview," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3386–3407, Sep. 2011.
[6] J. Ma, G. Y. Li, and B. H. Juang, "Signal processing in cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 805–823, May 2009.
[7] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 1st Quart., 2009.
[8] M. Lopez-Benitez and F. Casadevall, "Improved energy detection spectrum sensing for cognitive radio," *IET Commun.*, vol. 6, no. 8, pp. 785–796, May 2012.
[9] F. F. Digham, M. S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Anchorage, AK, USA, May 2003, pp. 3575–3579.
[10] H.-S. Chen, W. Gao, and D. G. Daut, "Signature based spectrum sensing algorithms for IEEE 802.22 WRAN," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Glasgow, U.K., Jun. 2007, pp. 6487–6492.
[11] Z. Zhang, Q. Yang, L. Wang, and X. Zhou, "A novel hybrid Matched Filter structure for IEEE 802.22 standard," in *Proc. IEEE Asia–Pacific Conf. Circuits Syst. (APCCAS)*, Kuala Lumpur, Malaysia, Dec. 2010, pp. 652–655.
[12] S. L. Sabat, S. Srinu, A. Raveendranadh, and S. K. Udgata, "Spectrum sensing based on entropy estimation using cyclostationary features for cognitive radio," in *Proc. 4th Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Bangalore, India, Jan. 2012, pp. 1–6.
[13] H. Tang, "Some physical layer issues of wide-band cognitive radio systems," in *Proc. IEEE Int. Symp. New Front. Dyn. Spectr. Access Netw. (DySPAN)*, Baltimore, MD, USA, Nov. 2005, pp. 151–159.
[14] D. Chen, J. Li, and J. Ma, "Cooperative spectrum sensing under noise uncertainty in cognitive radio," in *Proc. Wireless Commun., Netw. Mobile Comput.*, Dalian, China, Oct. 2008, pp. 1–4.
[15] C. Song *et al.*, "Adaptive two thresholds based energy detection for cooperative spectrum sensing," in *Proc. Consum. Commun. Netw.*, Las Vegas, NV, USA, Jan. 2010, pp. 1–6.
[16] R. Zhang, T. J. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 84–88, Jan. 2010.
[17] Y. Zeng and Y.-C. Liang, "Spectrum-sensing algorithms for cognitive radio based on statistical covariances," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1804–1815, May 2009.
[18] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. IEEE Symp. New Front. Dyn. Spectr. Access Netw. (DySPAN)*, Baltimore, MD, USA, Nov. 2005, pp. 131–136.

[19] B. Li, C. L. Zhao, M. W. Sun, and A. Nallanathan, "Spectrum sensing for cognitive radios in time-variant flat-fading channels: A joint estimation approach," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2665–2680, Aug. 2014.

[20] T. B. Schön, F. Gustafsson, and P.-J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2279–2289, Jul. 2005.

[21] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state Markov modeling of fading channels—A survey of principles and applications," *IEEE Signal Process. Mag.*, vol. 25, no. 5, pp. 57–80, Sep. 2008.

[22] H. S. Wang and N. Moayeri, "Finite-state Markov channel—A useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163–171, Feb. 1995.

[23] M. Sun, B. Li, Q. Song, L. Zhao, and C. Zhao, "Joint detection scheme for spectrum sensing over time-variant flat fading channels," *IET Commun.*, vol. 8, no. 12, pp. 2064–2073, 2014.

[24] C. L. Zhao, M. W. Sun, B. Li, L. Zhao, and X. Peng, "Blind spectrum sensing for cognitive radio over time-variant multipath flat-fading channels," *EURASIP J. Wireless Commun. Netw.*, May 2014.

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

[26] P. M. Djuric *et al.*, "Particle filtering," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 19–38, Sep. 2003.

[27] S. Saha, E. Özkan, F. Gustafsson, and V. Šmídl, "Marginalized particle filters for Bayesian estimation of Gaussian noise parameters," in *Proc. 13th Conf. Inf. Fusion (FUSION)*, Edinburgh, U.K., Jul. 2010, pp. 1–8.

[28] J. Miguez and P. M. Djuric, "Blind equalization by sequential importance sampling," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2002, pp. I-845–I-848.

[29] B. Li, S. Li, A. Nallanathan, and C. Zhao, "Deep sensing for future spectrum and location awareness 5G communications," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1331–1344, Jul. 2015.

[30] B. Li, S. Li, Y. Nan, A. Nallanathan, C. Zhao, and Z. Zhou, "Deep sensing for next-generation dynamic spectrum sharing: More than detecting the occupancy state of primary spectrum," *IEEE Trans. Commun.*, vol. 63, no. 7, pp. 2442–2457, Jul. 2015.

[31] B. Li, M. Sun, X. Li, A. Nallanathan, and C. Zhao, "Energy detection based spectrum sensing for cognitive radios over time-frequency doubly selective fading channels," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 402–417, Jan. 2015.

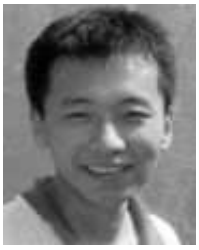
[32] K. E. Baddour and N. C. Beaulieu, "Autoregressive models for fading channel simulation," in *Proc. IEEE Global Commun. Conf. (Globecom)*, San Antonio, TX, USA, Nov. 2001, pp. 1187–1192.

SHENGHONG LI received the B.S. and M.S. degrees in electrical engineering from the Jilin University of Technology, China, in 1993 and 1996, respectively, and the Ph.D. degree in radio engineering from the Beijing University of Posts and Telecommunications, China, in 1999. Since 1999, he has been a Research Fellow, Associate Professor, and Professor with Shanghai Jiao Tong University, China, respectively. In 2010, he was Visiting Scholar with Nanyang Technological University, Singapore. His research interests include information security, signal, and information processing, cognitive radio networks, and artificial intelligence.

MENGWEI SUN received the B.S. degree in communication engineering from the Civil Aviation University of China, Tianjin, China, in 2011, and the Ph.D. degree in information and communication engineering from the Beijing University of Posts and Telecommunications, China, in 2017. She has been a Research Associate with the Heriot-Watt University. Her research interests include signal processing for cognitive radios and 60 GHz millimeter-wave communications, channel modelling, and satellite communications.

YING-CHANG LIANG (F'11) was a Principal Scientist and Technical Advisor with the Institute for Infocomm Research, Singapore. He is currently a Professor with the University of Sydney, Australia. He is also a Professor with the University of Electronic Science and Technology of China, China. His research interest lies in the general area of wireless networking and communications, with current focus on cognitive radio, and intelligent wireless networks.

Dr. Liang received the IEEE ICC Best Paper Award in 2017, IEEE ComSoc's TAOS Best Paper Award in 2016, IEEE Jack Neubauer Memorial Award in 2014, the First IEEE ComSoc's APB Outstanding Paper Award in 2012, and the *EURASIP Journal of Wireless Communications and Networking* Best Paper Award in 2010. He also received the Institute of Engineers Singapore Prestigious Engineering Achievement Award in 2007, and the IEEE Standards Association's Outstanding Contribution Appreciation Award in 2011, for his contributions to the development of the IEEE 802.22 standard. He is currently an Associate Editor-in-Chief of the *World Scientific Journal on Random Matrices: Theory and Applications*. He was a Distinguished Lecturer of the IEEE Communications Society and the IEEE Vehicular Technology Society, and has been a member of the Board of Governors of the IEEE Asia-Pacific Wireless Communications Symposium since 2009. He served as the Technical Program Committee Chair of CROWN'08 and DySPAN'10, Symposium Chair of ICC'12 and Globecom'12, General Co-Chair of ICCS'10 and ICCS'14. He serves as TPC Chair and Executive Co-Chair of Globecom'17 to be held in Singapore. He was Founding Editor-in-Chief of the IEEE Journal on Selected Areas in Communications—Cognitive Radio Series, key founder of the new journal IEEE Transactions on Cognitive Communications and Networking, and Chair of the IEEE Communications Society Technical Committee on Cognitive Networks. He has been the Guest/Associate Editor of the IEEE Transactions on Wireless Communications, IEEE Journal of Selected Areas in Communications, IEEE Transactions on Vehicular Technology, and the IEEE Signal Processing Magazine, and IEEE Transactions on Signal and Information Processing Over Network. He was recognized by Thomson Reuters as Highly Cited Researcher in 2014, 2015, and 2016.



BIN LI received the B.S. degree in electrical information engineering from the Beijing University of Chemical Technology in 2007, the Ph.D. degree in communication and information engineering from the Beijing University of Posts and Telecommunications in 2013, where he is currently an Associate Professor with the School of Information and Communication Engineering. His current research interests are focused on statistical signal processing for wireless communications, e.g. molecular communications, millimeter-wave (mm-Wave) communications and cognitive radios. He has published over 70 journal and conference papers. He received 2011 ChinaCom Best Paper Award, 2015 IEEE WCSP Best Paper Award, 2010 and 2011 BUPT Excellent Ph.D. Student Award Foundations. He serves as the Co-Chair of the Technical Program Committee of the Signal Processing for Communications Symposium of the 2016 IEEE International Conference on Computing, Networking and Communications.

CHENGLIN ZHAO received the B.S. degree in radio-technology from Tianjin University in 1986, the M.S. degree in circuits and systems from the Beijing University of Posts and Telecommunications in 1993, and the Ph.D. degree in communication and information system from the Beijing University of Posts and Telecommunications, in 1997. He currently serves as a Professor with the Beijing University of Posts and Telecommunications, Beijing, China. His research is focused on emerging technologies of short-range wireless communication, cognitive radios, 60GHz millimeter-wave communications.