Deep Sensing for Next-Generation Dynamic Spectrum Sharing: More Than Detecting the Occupancy State of Primary Spectrum

Bin Li, Shenghong Li, Amurugam Nallanathan, *Senior Member, IEEE*,
Yijiang Nan, Chenglin Zhao, and Zheng Zhou, *Member, IEEE*

***Abstract*—In this paper, spectrum sensing is investigated and a new detection framework, namely, deep sensing (DS), is proposed for more challenging scenarios of future dynamic spectrum sharing. In contrast to existing methods, the DS scheme is designed to proactively recover and exploit some other informative states associated with realistic cognitive links (e.g., fading gains), except detecting the occupancy of primary-band. A unified mathematical model, relying on the dynamic state-space approach, is formulated, in which the Bernoulli random finite set (RFS) is further exploited to theoretically characterize complex DS procedures. A Bernoulli filter algorithm is suggested to recursively estimate unknown PU states accompanying related link information, which is implemented by particle filtering based on numerical approximations. The proposed DS algorithm is applied to detect primary users under time-varying fading channel, which may increase the observation uncertainty and, therefore, deteriorate the sensing performance. With this new framework, the time-varying fading gain, modeled as a stochastic discrete-state Markov chain (DSMC), is estimated along with unknown PU states. Simulations demonstrate that, by exploiting the underlying dynamic fading property, the sensing performance will surpass other traditional schemes. The DS scheme may be conveniently generalized to other applications, which will promote sensing performance and provides a new paradigm for next-generation spectrum sharing.**

***Index Terms*—Spectrum sensing, deep sensing, dynamic state-space model, joint estimation, random finite set, time-variant flat fading.**

I. Introduction

By accessing the unoccupied spectrum of licensed band, cognitive radio (CR) based dynamic spectrum sharing (DSS) is initially intended to alleviate the most challenging problems of future wireless communications, namely, spectrum scarcity [1]. With a real-time perception of surroundings and bandwidth availability, CR devices may dynamically use the vacant spectrum and perform opportunistic transmissions, by adapting its functionality intelligently to accommodate current wireless environments [2]. Without causing harmful interference to primary users (PUs), the frequency utilization can be promoted significantly and, to some extent, the tensions of spectrum scarcity may be eased. Therefore, the DSS is of great promise to the developments of next-generation broadband communications, e.g., LTE in unlicensed band (LTE-U) and IEEE 802.11af.

As a key building block of DSS or CR, spectrum sensing is specially designed to identify the occupancy state of authorized spectrum [3]. Various sensing algorithms have attached extensive investigations [4]–[6]. Traditional techniques, with a fundamental target of detecting unknown PU states, include energy detector (ED) [7], matched filter detection [8] and cyclostationary feature detection [9], which may in practice have different advantages and requirements [5], [6]. Recently, wavelet analysis is adopted [10] and compressive sensing is further introduced to multi-band sensing [11]. By properly integrating the statistical information of PU signals, the covariance matrix [12], [13] and the probabilistic property based sensing techniques [14], [15] have also been developed, which are proven to effective in realistic applications.

Relying on the sensing result, the access strategies (e.g., transmission or silence) and recourse allocations (e.g., power and spectrum) of secondary users (SU) would be optimized accordingly. Such a widely recommended *sensing-and-acting* infrastructure, which is currently approved by the first worldwide CR-based wireless standard, i.e., IEEE 802.22, provides a naturally feasible approach for implementing DSS [1], [16]. As an evolving and enabling technique of next-generation 5G communications [17], however, there come also new characteristics and requirements for mobile DSS applications. Taking LTE-U for example, the mobile network operators (MNOs) are emerged to offload part of the traffic onto some unlicensed spectrum, in order to accommodate the orders-of-magnitude increase in mobile data volume per area. Such potential scenarios, on one hand, create the urgent need for spectral coexistence when various uncoordinated networks operate on the same frequency (e.g., mobile and distributed heterogeneous networks), and on the other hand, pose new challenges to traditional processing mechanisms especially in the attendant of dynamic or time-varying environments. Based on the following

Manuscript received July 17, 2014; revised February 3, 2015 and May 14, 2015; accepted June 3, 2015. Date of publication June 9, 2015; date of current version July 13, 2015. This work was supported by Natural Science Foundation of China (NSFC) under Grant 61471061. The associate editor coordinating the review of this paper and approving it for publication was M. Buehrer.

B. Li, Y. Nan, C. Zhao, and Z. Zhou are with the School of Information and Communication Engineering (SICE), Beijing University of Posts and Telecommunications (BUPT), Beijing 100876, China (e-mail: stonebupt@gmail.com; nanyj@bupt.edu.cn; clzhao@bupt.edu.cn; zz@bupt.edu.cn).

S. Li is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: shli@stju.edu.cn).

A. Nallanathan is with the Department of Informatics, King's College London, London WC2R 2LS, U.K. (e-mail: nallanathan@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2015.2443041

two considerations, the existing sensing algorithms seem to be less attractive to future DSS applications under more adverse conditions [17], e.g., mobile or vehicle-to-vehicle scenarios.

First, the sensing performance will be affected significantly by unknown fading channel. Despite a latent component of sensing procedure, the fading channel will increase the uncertainty of observations, especially in some emerging scenarios with time-varying fading gains [18], [19], e.g., involving mobile devices (or relative movements) in next-generation wireless local area networks (WLANs) and LTE-U. Even if some existing schemes have taken the statistical property of fading effects into accounts [7], [20], unfortunately, these methods relying on static probability density functions (PDFs) fail to track time-correlated fading gains and, therefore, cannot exploit the underlying dynamic property to further enhance the sensing performance.

Second, there is no doubt that the *cognitive-link state information* (CLSI), e.g., the time-varying fading gain, is of great promise to future DSS applications, e.g., interference evaluations, energy efficiency or distributed techniques for coexistence. With a deeper cognition and more detailed information of PU-SU links, subsequent reconfigurations may be definitely promoted. Although various feasible approaches have been developed which may exclude such CLSI by learning wireless environments adaptively [20], the computational complexity and less attractive performance seem to be a main stumbling block limiting their wide applications, especially in the time-variant scenarios. It seems that a direct estimation of CLSI, when performing spectrum sensing, may be more promising in such adverse environments.

The above considerations, nowadays, have become an inducement to the conceptual innovation of current spectrum sensing schemes. For the next-generation DSS, it is supposed to probe the other informative states associated with cognitive links, which are of great significance to enhance sensing performance or to facilitate following CR operations. In this paper, a new sensing framework, referred to as deep sensing (DS), is developed to overcome the encountered difficulties. In sharp contrast to classical sensing schemes, in the new DS the dynamic fading channel will also be estimated blindly based on the observed signals, at the same time of detecting unknown PU states. To sum up, the main contributions of this work are two-folds.

1) A novel *DS framework* is proposed. A major innovation of such a DS framework is that both unknown fading channel and PU's emission state are treated as two hidden states to be estimated. A dynamic state-space model (DSM), in the presence of time-varying flat-fading (TVFF) channel, is established in the context of more complex DS. These two states are assumed to evolve with time according to different Markov properties. I.e., both of them are modeled as discrete states Markov chains (DSMCs). The summed-energy, as the observed output for ease of implementation, is utilized to estimate two hidden states. It is noteworthy that the formulated DS may be easily generalized to other DS applications. That is, any important features of PU-SU interactive links, which are coupled with the observed signals, e.g., time-varying fading gains, dynamic power-levels, modulation formats, unknown distance of PUs, can be cast into the new DSM. Thus, the DS may be ready to various different scenarios, if an adapting estimation algorithm can be available.

For the formulated problem, it is noted that traditional sensing schemes, unfortunately, may become invalid to this considered problem. For one thing, as mentioned, most existing techniques can hardly deal with time-correlated fading channel and, consequently, their sensing performance will be deteriorated significantly [20]. For another, even with the objective of estimating the associative CLSI jointly, no likelihood information on the fading channel will be available in observations when the PU signal is absence. In the previous work, an iteratively implemented scheme has been designed to combat this major challenge [21], which may estimate dynamic fading gains premised on an initial estimation of PU states. Due to its inflexibility, nevertheless, the three-step iterative algorithm may be inefficient in DS applications where two hidden states are supposed to be acquired jointly, rather than via a *quasi-joint* approach as in [21]. Recently, another joint detection and estimation scheme is proposed by [22], relying on a *supervised* manner, which requires both synchronization and pilot signaling of PU and, unfortunately, may be impractical in realistic non-coordinated scenarios.

2) The second contribution, accordingly, is that a flexible *estimation algorithm* for DS is developed. The DS procedure, involving an occupancy state (i.e., PU's state) and another associated CLSI (i.e., the time-varying fading gain), is modeled theoretically as one special random variable, i.e., Bernoulli random finite set (BRFS). With the new formulation, the varying cardinality of BRFS is viewed as a token of the occupancy state of PU, with which dynamic fading states are closely coupled. A sequential estimation scheme, first predict and then update, is designed to estimate the BRFS cardinality (i.e., PU state) and the coupled CLSI simultaneously. The DS algorithm is recursively implemented on reception of new observations, which derives the posterior densities within a Bayesian inference framework. Two new mechanisms are further integrated to the specific application where fading gains keep invariant in several sensing slots. Particle filtering (PF) is suggested to approximate non-analytical distributions numerically, and a Bernoulli PF (BPF) scheme is thereby adopted to realize DS. With the assistance of recovered fading gains, the sensing performance will be improved in realistic TVFF channel, compared to traditional methods designed for the time-invariant fading channel.

The reason why the new scheme is called "deep sensing" is that, by acquiring and exploiting CLSI effectively, it indeed senses spectrum in a *deep cognition* mode. It estimates blindly unknown PU's states and varying fading gains *as a whole* and, therefore, remains also dramatically different from other existing joint estimation-based sensing concepts, e.g., estimating PU' states and fading gains separately [21] or utilizing the pilot signaling from a coordinated PU [22]. The potential benefits of the proposed DS scheme are the following three areas:

1) The time-correlated fading gain is recovered blindly without posing any infeasible requirements on PUs. Thus, the underlying dynamic property could be fully exploited to promote the sensing performance.
2) The estimated fading gains, as an extra gift of sensing process, may be utilized to optimize CR operations and, furthermore, maximize the functionality of CRs. This new

TABLE I
THE USED ACRONYMS

CR	cognitive radio	DSS	dynamic spectrum sharing
PUs	primary users	ED	energy detector
DSM	dynamic state-space model	MNOs	mobile network operators
WLANs	wireless local area networks	CLSI	cognitive-link state information
DS	deep sensing	TVFF	time-varying flat-fading
DSMCs	discrete states Markov chains	RFS	random finite set
PF	particle filtering	TPM	transitional probability matrix
LCR	level crossing rate	MAP	maximum *a posteriori*
FISST	finite set statistics	PDF	probability density function

DS framework, therefore, may lead to more flexible or even simplified DSS solutions.

3) The new DS scheme, which can be extended to other important scenarios, may put an insight into spectrum sensing and hence provides a brand-new idea for 5G spectrum sharing.

The rest of the paper is structured as following. In Section II, a new general system model, based on a dynamic state-space approach, is formulated in realistic TVFF propagation. Subsequently, a Bayesian sequential estimation is introduced briefly in Section IV. Then, a promising DS algorithm which jointly estimates the fading gain and PU's states is proposed, by suggesting a flexibly numerical scheme. In Section V, comprehensive numerical simulations and performance analysis are provided. Finally, we conclude this investigation in Section VI. Table I lists commonly used acronyms.

II. SYSTEM MODEL OF DEEP SENSING

A main feature of the proposed DS scheme is that, except for identifying the occupancy of vacant PU bands, it manages also to recover other important states associated with PUs that may be of use to promote the sensing performance or resource reconfiguration. In this section, by taking the time-varying property of fading channel into full account, a unified DSM is used to comprehensively characterize the DS process in TVFF conditions.

A. DSM for Deep Sensing

In the consideration of exploiting the dynamics property of TVFF channel, a stochastic DSM in the context of DS is established as follows.

$$s_n = F(s_{n-1}), \tag{1}$$

$$\alpha_n = H(\alpha_{n-1}), \tag{2}$$

$$y_n = G(\alpha_n, s_{n,m}, d_{n,m}, z_{n,m}). \tag{3}$$

Here, eqs. (1), (2) are referred to as *dynamic equations*, while eq. (3) is the *measurement equation*. The non-analytical transitional function $F(\cdot) : \mathbb{Z}^1 \to \mathbb{Z}^1$ specifies the stochastic evolution of the PU's states $s_{n,m} \in \mathcal{S} = \{0, 1\}$ ($\mathcal{S} \subset \mathbb{Z}^1$) of the mth discrete sampling time in the nth slot as a 1st order Markov process. Another dynamic function $H(\cdot) : \mathbb{R}^1 \to \mathbb{R}^1$ characterizes dynamic behaviors of the fading channel $\alpha_n \in \mathcal{A}$ ($\mathcal{A} \subseteq \mathbb{R}^1$). The observation function $G(\cdot) : \mathbb{R}^{M\times 1} \to \mathbb{R}^1$ then describes the coupling relationship between two hidden states (i.e., $s_{n,m}$ and α_n) and the measurement $y_n \in \mathcal{Y}$ ($\mathcal{Y} \subseteq \mathbb{R}^1$), where $d_{n,m}$ denotes the mth sample of PU signals at discrete time n. The random noise involved by the received signals, denoted by $z_{n,m} \in \mathbb{R}^1$, is assumed to be a zero-mean additive white Gaussian noise (AWGN). The noise variance of is σ_z^2, i.e., $z_{n,m} \sim \mathcal{N}(0, \sigma_z^2)$. In practice, it will also remain independent of two hidden states.

The details on each equation will be elaborated next in the context of TVFF channel. Before proceeding, it is necessary to make three important assumptions to this DSM model for the ease of analysis.

1) For the considered slowly varying case, a fading gain α_n is assumed to keep invariant within several successive sensing-transmission slots. The ratio between the lasting duration of α_n and the length of sensing-transmission slot T_F, denoted by L, is practically related with the maximum Doppler frequency shift f_D. Specifically, a larger f_D leads to a smaller L.
2) The fading channel gain α_n is assumed to remain invariant during each sensing slot. Such a presumption becomes practically valid for the slowly-varying channel [21].
3) The PU's working state is assumed to be unchanged in each sensing slot, i.e., $s_{n,m} = s_n$.

B. Dynamics of PU State

We firstly study the dynamic behavior of an unknown PU state. It has been widely suggested that the evolution of PU states over time may be modeled as a finite state machine $\mathcal{S} = \{S_0, S_1\}$, i.e., a two-state Markov transition procedure [23]. If a PU is in active state S_1 at the current sensing slot n, it will then stay in S_1 with a *survival* probability

$$p_s \triangleq \Pr\{s_{n+1} = 1 | s_n = 1\},$$

and jump into the sleep state S_0 with a probability $1 - p_s$ in the next slot $n + 1$. Then, if it is currently in the sleep state S_0, it will change to S_1 with a *birth* probability

$$p_b \triangleq \Pr\{s_{n+1} = 1 | s_n = 0\},$$

and similarly, may stay in S_0 in next slot $n + 1$ with a probability of $1 - p_b$.

In practice, it may be further considered that the above probabilistic property is time-homogeneous. I.e., the dynamical transition is related only with specific wireless services while keeping invariant for a given application in all the time (or at least, in a long period). The statistical transitional probability matrix (TPM) of PU's emission states is known as *a priori*, which is given by:

$$\boldsymbol{P} = \begin{bmatrix} (1 - p_b) & p_b, \\ (1 - p_s) & p_s. \end{bmatrix}. \tag{4}$$

Thus, the dynamic equation $F(\cdot)$ is inexplicitly specified by the stochastic transitions from s_{n-1} to s_n, i.e., $F(s_{n-1}) = s_n$ with a probability $\Pr(s_n = j | s_{n-1} = i),\ i, j = 0, 1$. Also note that, different from a traditional sense of dynamic equations, there contains no driven noise in eq. (1) (and also eq. (2)).

Governed by an underlying finite-states Markov chain, however the states s_n will randomly transit even without an external driven term. Thus, we may still refer it (and eq. (2)) as a special dynamic equation.

C. Dynamics of Fading Channels

An adverse propagation condition, i.e., the time-correlated fading channel, is specially concerned by this work, which may be very common to emerging future wireless applications involving mobile devices (or relative movements), e.g., LTE-A or 5G communications [17], [24], [25].

The dynamic behaviors of time-dependent fading gains may be modeled properly by another DSMC. As suggested, the DSMC may reflect the dynamic nature (i.e., evolving along times) of time-varying fading gains effectively, which match also the statistical models sufficiently, e.g., Clarke's model [21], [25]. Given the wide-sense-stationary uncorrelated scattering, the time-correlation property may tend to be ergodic and stationary [19], [25]. In realistic scenarios, therefore, the DSMC model is stationary [25]–[27] and the transitional probabilities are independent of the time n.

The indecomposable DSMC is mainly considered here. I.e., if we denote the stationary probability vector $\boldsymbol{\pi} = [\pi_0, \pi_1, \cdots, \pi_{K-1}]^T$ with $\pi_k \triangleq \Pr(\alpha_n = A_k)$, then we have $\boldsymbol{\Pi}^T\boldsymbol{\pi} = \boldsymbol{\pi}$, where $\boldsymbol{\Pi}_{K\times K} = \{\pi_{k_1\to k_2}, k_1, k_2 \in [0, 1, \cdots, K-1]\}$ denotes the TPM of the fading gains, i.e.,

$$\boldsymbol{\Pi}_{K\times K} = \begin{bmatrix} \pi_{0\to 0} & \pi_{0\to 1} & \cdots & \pi_{0\to(K-1)} \\ \pi_{1\to 0} & \pi_{1\to 1} & \cdots & \pi_{1\to(K-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{(K-1)\to 0} & \pi_{(K-1)\to 1} & \cdots & \pi_{(K-1)\to(K-1)} \end{bmatrix}, \tag{5}$$

where each element $\pi_{k_1\to k_2}$ accounts for the transitional probability from the state k_1 at time index $(n'-1)$ to the state k_2 at time index n':

$$\pi_{k_1\to k_2} \triangleq \Pr\left(\alpha_{n'} = A_{k_2} | \alpha_{n'-1} = A_{k_1}\right), \tag{6}$$

where the discrete state of the transitional time $n' = \lfloor n/L \rfloor$ is denoted by $A_k \in \mathcal{A},\ k \in \{0, 1, \cdots, K-1\}$, which is regarded as the output of a specific DSMC [25], [26]. That is, the nonnegative fading amplitude α_n (or the received SNR) may be partitioned into K non-overlapping regions, denoted by $\mathbb{V}$. If we further specify $\nu_0 = 0$ and $\nu_K = \infty$, we have:

$$\mathbb{V} = \{[\nu_0, \nu_1), [\nu_1, \nu_2), \cdots, [\nu_{K-1}, \nu_K)\}.$$

Suppose the PDF of fading channel is denoted by $f(\alpha)$, then the steady probability the varying fading gains residing in the kth discrete region $[\upsilon_k, \upsilon_{k+1})$ and its representative state are derived from:

$$\pi_k = \int_{\nu_k}^{\nu_{k+1}} f(\alpha) d\alpha, \quad A_k = \frac{1}{\pi_k} \times \int_{\nu_k}^{\nu_{k+1}} \alpha f(\alpha) d\alpha, \tag{7}$$

respectively.

With an equi-probable partition rule, i.e., $\pi_k \triangleq 1/K$, we may easily derive the partitioning bounds by $\upsilon_k = \sqrt{-2\sigma^2 \ln(1-k/K)}$ [25], [28]. Then, the transitional probability $\pi_{k_1\to k_2}$ can be easily determined by:

$$\begin{aligned} \pi_{k_1\to k_2} &= \Pr\left\{\alpha_{n'} \in \left[\nu_{k_2}, \nu_{k_2}+1\right) \middle| \alpha_{n'-1} \in \left[\nu_{k_1}, \nu_{k_1}+1\right)\right\} \\ &= \frac{1}{\pi_{k_1}} \int_{\nu_{k_2}}^{\nu_{k_2}+1} \int_{\nu_{k_1}}^{\nu_{k_1}+1} f(\alpha_{n'-1}, \alpha_{n'}) d\alpha_{n'-1} d\alpha_{n'}, \end{aligned} \tag{8}$$

where $f(\alpha_{n'-1}, \alpha_{n'})$ is the bivariate Rayleigh joint PDF [29].

For simplicity, a first-order DSMC is considered, which has been proven to be sufficiently applicable to the slowly varying fading channel. With the first-order DSMC, then the current fading state will be independent of all other past and future fading states when the interval between two states is larger than 1, i.e., $\pi_{k_1\to k_2} = 0$ for $|k_2 - k_1| > 1$. Accordingly, the TPM $\boldsymbol{\Pi}_{K\times K}$ will be further simplified to the equation (shown at the bottom of the page).

In order to calculate the transitional probabilities via numerical approximations, in practice, the level crossing rate (LCR) N_k may be utilized [25]–[27]. LCR refers to the number of times per second that the fading gain crosses υ_k in a downward direction. For the commonly time-varying Rayleigh fading discussed in the analysis, we have

$$N_k \triangleq \int_0^\infty \dot{\alpha} f(\alpha, \dot{\alpha}) d\dot{\alpha}\bigg|_{\alpha=\nu_k} = \sqrt{2\pi} f_D \frac{\nu_k}{\sigma_\alpha} \exp\left(-\frac{\nu_k^2}{\sigma_\alpha^2}\right), \tag{9}$$

where $f(\alpha, \dot{\alpha})$ is the joint PDF of the signal envelop α and its time deviation $\dot{\alpha}$ [23], [24], and σ_α^2 denotes the distribution variance of the Rayleigh fading. Relying on such a LCR, the transitional probability may be numerically approximated by $\pi_{k_1\to k_2} \simeq N_{k_2}/R_{k_1}$. Here, $R_{k_1} = \pi_{k_1}/T_F$ accounts for the average number of sensing slots per second in the state i.

Given the numerically derived $\boldsymbol{\Pi}_{K\times K}$, another dynamic equation $H(\cdot)$ will be similarly determined via a group of stochastic transitions from $\alpha_{n'-1}$ to $\alpha_{n'}$, i.e., $H(\alpha_{n'-1} = A_{k_1}) = \{\alpha_{n'} = A_{k_2}\}$ with a probability $\pi_{k_1\to k_2}$ when $|k_1 - k_2| \le 1$ holds.

$$\boldsymbol{\Pi}_{K\times K} = \begin{bmatrix} \pi_{0\to 0} & \pi_{0\to 1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ \pi_{1\to 0} & \pi_{1\to 1} & \pi_{1\to 2} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \pi_{(K-2)\to(K-3)} & \pi_{(K-2)\to(K-2)} & \pi_{(K-2)\to(K-1)} \\ 0 & 0 & 0 & 0 & \cdots & 0 & \pi_{(K-1)\to(K-2)} & \pi_{(K-1)\to(K-1)} \end{bmatrix}.$$

D. Observation

As a fundamental sensing technique, the energy-based scheme has been widely recommended for ease of implementations [3]–[7]. This work will establish a general DS model based on ED. Note that, however, the generalization to other observations (i.e., the matched filter) is straightforward. Before proceeding, it is necessary to briefly illustrate ED scheme, which is formulated as a two-hypothesis problem:

$$y_n \triangleq \begin{cases} \sum_{m=1}^{M} z_{n,m}^2, & H_0 \text{ or } s_n = \mathcal{S}_0, \quad (10a) \\ \sum_{m=1}^{M} (\alpha_n d_{n,m} + z_{n,m})^2, & H_1 \text{ or } s_n = \mathcal{S}_1, \quad (10b) \end{cases}$$

where $M = T_S \times f_s$ is the length of samples, and f_s is the sampling frequency. H_0 and H_1 denote two hypotheses corresponding to the absence and presence of PU signals (i.e., S_0 and S_1), respectively; $u_{n,m} = \alpha_n d_{n,m} + z_{n,m}$ is the mth received sample of the discrete time n, and $y_n = \sum_{m=1}^{M} u_{n,m}^2$ denotes the energy observation. $d_{n,m} = s_n \times b_{n,\lfloor m/N_s \rfloor} g(m - N_s/2)$ is the received baseband signal of the nth slot, where $\{b_{n,i}\}$ denote the PU's information symbols and $g(k)$ $(k = 0, 1, \cdots, N_s - 1)$ is the pulse-shaping response, and N_s is the length of the pulse. For simplicity, the real-valued BPSK signal is considered, e.g., $b_{n,i} \in \{+1, -1\}$. It is noteworthy that the extension of the measurement function in eq. (10) to other modulated signals is also feasible.

Conditioned on the fading gain α_n and PU's emission state s_n, the likelihood density $p(y_n|\alpha_n, s_n)$ follows a central chi-square distribution with M degrees of freedom (DoF) under H_0, and a non-central chi-square distribution with M degrees under H_1, i.e.,

$$\varphi(y_n|\alpha_n, s_{n,m} = 0) \sim \chi_M^2, \tag{11a}$$

$$\varphi(y_n|\alpha_n, s_{n,m} = 1) \sim \chi_M^2(\varrho_n), \tag{11b}$$

where $\varrho_n = \sum_{m=1}^{M} |\alpha_n d_{n,m}|^2 \simeq M E_d \alpha_n^2$ denotes the non-central parameter, which is related with the time-dependent fading gain α_n and the average power of PU signals, i.e., $E_d \triangleq \mathsf{E}\{|d_{n,m}|^2\} = \sigma_d^2$.

Note that, despite the same observation (i.e., the summed energy), classical ED schemes essentially ignore the time-variant CLSI. For such threshold-based techniques, a proper decision threshold γ_{ED} would be configured according to various different criterions, e.g., the *Neyman-Pearson* criterion [5]. Thus, the false alarm probability P_f is defined by $\Pr(y_n > \gamma_{ED}|H_0)$, while the detection probability P_d is $\Pr(y_n > \gamma_{ED}|H_1)$.

$$P_f \triangleq \Pr(H_1|H_0) = \frac{\Gamma(M, \gamma_{ED}/2)}{\Gamma(M/2)}, \tag{12a}$$

$$P_d \triangleq \Pr(H_1|H_1) = Q_M\left(\sqrt{2\tau}, \sqrt{\gamma_{ED}}\right), \tag{12b}$$

where $\tau = \alpha_n^2 \sigma_d^2 / \sigma_z^2$ is the instantaneous SNR at the nth sensing slot, and σ_d^2 is the variance of PU signals. $\Gamma(a, x)$ is the incomplete gamma function, and $\Gamma(a)$ denotes the gamma function. $Q_{M/2}(a, x)$ is the generalized Marcum Q-function. Even if the PDF of channel fading (e.g., Rayleigh fading) may be considered [7], most threshold-based schemes fail still to cope with time-correlated dynamic fading, which may be unattractive in the presence of TVFF propagations.

III. Deep Sensing

A. MAP Estimation of DS

In sharp contrast to traditional sensing schemes, a Bayesian stochastic approach is designed for DS, which recursively estimates the joint posterior PDF of two hidden states based on the formulated DSM. Relying on the maximum *a posteriori* (MAP) criterion other than the aforementioned *Neyman-Pearson* criterion, the estimations of unknown PU states, accompanying the dynamical CLSI (i.e., the fading gain), are derived from eq. (13), shown at the bottom of the page, where $y_{0:n} \triangleq \{y_0, y_1, \cdots, y_n\}$ denotes the trajectory of observations until the nth time slot.

For clarity, two hidden terms are denoted by one state vector $\mathbf{s}_n \triangleq \{s_n, \alpha_n\}$. Given the assumed first-order Markov chain and *a priori* distribution of $\mathbf{s}_0$, i.e., $p(\mathbf{s}_0)$, the joint posterior distribution is then propagated sequentially via the well-known two-step procedure, i.e., the prediction followed by an update.

1) Predict Stage: Let $p_{n-1|n-1}(\mathbf{s}_{n-1}|y_{0:n-1})$ be the posterior density at time $n-1$. With the help of the Chapman-Kolmogorov equation, the one-step prediction $p_{n|n-1}(\mathbf{s}_n|y_{0:n-1})$ is obtained via:

$$p_{n|n-1}(\boldsymbol{s}_n|y_{1:n-1}) = \int \phi_{n|n-1}(\boldsymbol{s}_n|\boldsymbol{s}_{1:n-1}) p_{n-1|n-1}(\boldsymbol{s}_{n-1}|y_{1:n-1}) d\boldsymbol{s}_{n-1}, \tag{14}$$

where $\boldsymbol{s}_{0:n} \triangleq \{\boldsymbol{s}_0, \boldsymbol{s}_1, \cdots, \boldsymbol{s}_n\}$ denotes the trajectory of hidden states until the nth sensing slot, and the traditional density $\phi_{n|n-1}(\boldsymbol{s}_n|\boldsymbol{s}_{1:n-1})$ is determined by:

$$\phi_{n|n-1}(\boldsymbol{s}_n|\boldsymbol{s}_{0:n-1}) = \Pr(s_n = j|s_{n-1} = i) \times \Pr\left(\alpha_{n'} = A_{k_2}|\alpha_{n'-1} = A_{k_1}\right).$$

where $i, j \in \{0, 1\}$ and $k_1, k_2 \in \{0, 1, \cdots, K-1\}$.

2) Update Stage: By taking full advantage of the current measurement y_n accompanying the likelihood density $\varphi_n(y_n|\boldsymbol{s}_n)$, the Bayesian rule is then applied to refine the predicted density in eq. (14), i.e.,

$$p_{n|n}(\boldsymbol{s}_n|y_{0:n}) = \frac{\varphi_n(y_n|\boldsymbol{s}_n) p_{n|n-1}(\boldsymbol{s}_n|y_{0:n-1})}{\int \varphi_n(y_n|\boldsymbol{s}_n) p_{n|n-1}(\boldsymbol{s}_n|y_{0:n-1}) d\boldsymbol{s}_n}. \tag{15}$$

Thus, the posterior distribution $p_{n|n}(\boldsymbol{s}_n|y_{0:n})$ can be computed from the above two-stage process.

$$\left(\hat{\alpha}_n, \hat{s}_n\right) = \arg \max_{\alpha_n \in \mathcal{A}, s_n \in \mathcal{S}} p(\alpha_n, s_n|y_{0:n}) \big| \hat{\alpha}_{n-1}, \hat{s}_{n-1}, p(\alpha_n|\alpha_{n-1}), p(s_n|s_{n-1}), \tag{13}$$

For the DS scenario with the time-varying fading channel, however the traditional sequential estimation framework may become infeasible. First, the fading channel α_n would disappear completely from observations when a PU device turns *off*, i.e., $s_n = \mathcal{S}_0$ or H_0, as noted from eq. (10a). Thus, unlike classical joint estimations of unknown parameters where the likelihood is always available, the likelihood of fading gains, which is now coupled with s_n, will unfortunately be sporadic or intermittent in such a situation, making the Bayesian inference very tough. Second, even if the likelihood function could be utilized when the PU state is *on*, the involved marginal integration seems still to be analytically intractable or computationally prohibitive to many practical uses [21], [30].

B. Bernoulli Random Finite States

To address the first problem, one has to resort to other more efficient approaches to deal with the complex DS procedure. As noted from DSM, except for the dynamic fading state (i.e., α_n), the appearance/disappearance of PU also is an important random process, which is used to be treated directly as another *separate* random state (i.e., s_n) [5], [6], [21]. From a more attractive perspective, the dynamic behaviors of two hidden states could be structured into *one* unified random process $\mathcal{F}$, which is referred to as RFS [31]. In this new formulation, the cardinality (i.e., the number of random elements) of the RFS $\mathcal{F}$, which is denoted by $d = |\mathcal{F}|$, is further specified by a random density $\kappa(d) = \Pr\{|\mathcal{F}| = d\},\ d \in \mathbb{N}_0$. That is, an RFS may be considered, in essence, to be a random variable which takes values as unordered finite sets [31], [34].

It is noted that, in sharp contrast to classical random variables, another cardinality distribution $\kappa(d)$ is exploited to characterize an RFS, apart from the d-elements joint distribution $p(x_1, \cdots, x_d)$, $\{x_1, \cdots, x_d\} \in \mathcal{A}^d \subset \mathbb{R}^d$. Similar to other random variables, a finite set statistics (FISST) PDF is used to describe an RFS probabilistically, which is denoted by $p(\mathcal{F})$. Due to its convenience, the Mahler's approach has been widely suggested to define a FISST PDF $p(\mathcal{F})$) [31], which is given by:

$$p\left(\mathcal{F} = \{x_1, \cdots, x_d\}\right) \triangleq d!\kappa(d)p(x_1, \cdots, x_d). \quad (16)$$

By using a set integration on eq. (16), it is seen that $p(\mathcal{F})$ indeed may be adopted as a PDF, i.e.,

$$\int_{\mathcal{F}} p(\mathcal{F})\delta\mathcal{F} = p(\varnothing) + \sum_{d=1}^{\infty} \frac{1}{d!} \int_{\{x_1,\cdots,x_d\}\in\mathcal{A}^d} p(x_1, \cdots, x_d)dx_1 \cdots dx_d \equiv 1.$$

Here, the notation $\int_{\mathcal{F}} p(\mathcal{F})\delta\mathcal{F}$ denotes the *set integration* on $\mathcal{F}$, rather than the random distribution marginalization. In order to fully describe an RFS $\mathcal{F}$, from the above elaboration, it is necessary to study a cardinality density $\kappa(d)$, an RFS PDF $p(\mathcal{F})$ and related transitional densities, respectively.

1) Cardinality Distribution: As far as the DS application is concerned, we have either $\mathcal{F} = \{\alpha_n \in \mathcal{A}\}$ or $\mathcal{F} = \varnothing$, i.e., $d = 1$. When the PU exists at time n (i.e., $s_n = \mathcal{S}_1$), we have $|\mathcal{F}| = 1$, and otherwise, $|\mathcal{F}| = 0$ (i.e., $s_n = \mathcal{S}_0$). Thus, d (or s_n) is a random *Bernoulli* variable, which may either be empty with probability $1 - q$ or have a single element with probability q. Thus, the random cardinality of $\mathcal{F}$ is:

$$\kappa(d) = \begin{cases} 1 - q, & \text{if } \mathcal{F}_n = \varnothing \text{ or } s_n = 0, & (17a) \\ q, & \text{if } \mathcal{F}_n = \{\alpha_n\} \text{ or } s_n = 1. & (17b) \end{cases}$$

2) RFS PDF: With the TVFF effect, the singleton state associate with the PUs' appearance (i.e., $|\mathcal{F}_n| = 1$) accounts for the dynamic fading gain α_n. The FISST PDF, depending on a cardinality distribution $\kappa(d)$ and the associative distribution $p(\alpha)$, is thereby given by:

$$p(\mathcal{F}_n) = \begin{cases} 1 - q, & \text{if } \mathcal{F}_n = \varnothing, & (18a) \\ q \times p(\alpha), & \text{if } \mathcal{F}_n = \{\alpha_n\}. & (18b) \end{cases}$$

For other cases the cardinality d is larger than 1, we have $p(\mathcal{F}) = 0$. Checking eq. (18), the set integration on $\mathcal{F}$ satisfies $\int_{\mathcal{F}} p(\mathcal{F})\delta\mathcal{F} = (1 - q) + q \times \int p(\alpha)d\alpha \equiv 1$.

3) Transitional Densities: Given the established DSM in eqs. (1)–(3), the transitional behaviors of the Bernoulli RSF $\mathcal{F}$ follows a first-order Markov process. As the cardinality (i.e., PU's states) and the fading states evolve independently, the *a priori* transitional densities of $\mathcal{F}$ are written to:

$$\phi_{n|n-1}(\mathcal{F}|\varnothing) = \begin{cases} 1 - p_b, & \text{if } \mathcal{F} = \varnothing, & (19a) \\ p_b b_{n|n-1}(\alpha_n), & \text{if } \mathcal{F} = \{\alpha_n\}, & (19b) \end{cases}$$

and

$$\phi_{n|n-1}(\mathcal{F}|\{\alpha_n\}) = \begin{cases} 1 - p_s, & \text{if } \mathcal{F} = \varnothing, & (20a) \\ p_s \pi_{n|n-1}(\alpha_n|\alpha_{n-1}), & \text{if } \mathcal{F} = \{\alpha_n\}, & (20b) \end{cases}$$

respectively. Here, $b_{n|n-1}(\alpha_n)$ refers to as the *birth* density, which represents an initial distribution of the singleton state $\{\alpha_n\}$ after the PU is *re-birthed*. This birth distribution should be properly specified as *a priori* in order to achieve good performance, which will be discussed in subsequent Section III-F.

With the energy-based observation, the likelihood distribution, denoted by $\varphi(y_n|\mathcal{F})$, is easily obtained, i.e.,

$$\varphi(y_n|\mathcal{F}) = \begin{cases} \chi_M^2, & \text{if } \mathcal{F}_n = \varnothing, & (21a) \\ \chi_M^2(\varrho), & \text{if } \mathcal{F}_n = \{\alpha_n\}. & (21b) \end{cases}$$

C. Sequential MAP Detection

It is recognized that, different strikingly from an intuitive formulation involving two independent hidden states (e.g., s_n and α_n), RFS deals only with a unified random variable $\mathcal{F}_n$, which, however, will dynamically turn *on-off*. A principal benefit of such a new RFS is that, for this *mixed* detection and estimation problem, two hidden states need to be estimated simultaneously unlike the previous works [21], [37]. By casting the stochastic DS problem into an RFS, the dynamically associative CLSI (i.e., fading state) could be estimated along with a Bernoulli random cardinality, which actually reveals the appearance/disappearance of PU's signals. Thus, the first problem in designing DS algorithm is addressed.

To accomplish this, a similar two-stage recursive framework may be suggested. Note that, a major difference between the densities propagation of RFS and the general sequential estimations is that, rather than the distribution integration as in eqs. (14), (15), the set integration (i.e., $\delta\mathcal{F}_n$) will be applied to RFS. That is,

$$p_{n|n-1}(\mathcal{F}_n|y_{1:n-1}) = \int_{\mathcal{F}_n} \phi_{n|n-1}(\mathcal{F}_n|\mathcal{F}_{1:n-1})p_{n-1|n-1}(\mathcal{F}_{n-1}|y_{1:n-1})\delta\mathcal{F}_{n-1}, \tag{22}$$

$$p_{n|n}(\mathcal{F}_n|y_{1:n}) = \frac{\varphi_n(y_n|\mathcal{F}_n)p_{n|n-1}(\mathcal{F}_n|y_{1:n-1})}{\int_{\mathcal{F}_n}\varphi_n(y_n|\mathcal{F}_n)p_{n|n-1}(\mathcal{F}_n|y_{1:n-1})\delta\mathcal{F}_n}. \tag{23}$$

D. Bernoulli Filter

Note that, most existing schemes may fail to handle the above complex DS process. A particular stochastic algorithm is thereby suggested to derive the recursive estimations of the formulated Bernoulli RFS $\mathcal{F}_n$, which is known as the Bernoulli filter scheme [31], [34].

In order to obtain the FISST PDF $p_{n|n}(\mathcal{F}_n|y_{1:n})$, two related distributions may be of great importance from eq. (18). I.e., (1) the posterior density of PU's appearance which is denoted by:

$$q_{n|n} \triangleq \Pr\{\mathcal{F}_n|y_{1:n}\}, \tag{24}$$

and (2) the *a posteriori* spatial PDF of the single state

$$f_{n|n}\left(\mathcal{F}_n \triangleq \{\alpha_n\}\right) = p\{\alpha_n|y_{1:n}\}. \tag{25}$$

With the assistance of the above prediction-and-update framework, then the Bernoulli filter will propagate the above two posterior terms recursively.

1) Prediction Stage: During the one-step prediction stage, the predicted terms $q_{n|n-1}$ and $f_{n|n-1}(\alpha_n)$ will be derived based on the posterior densities $q_{n-1|n-1}$ and $f_{n-1|n-1}(\alpha_{n-1})$ of time $(n-1)$. By applying eq. (18), the predicted posterior FISST PDF at time n is expressed as:

$$\begin{aligned} p_{n|n-1}(\mathcal{F}_n|y_{1:n-1}) &= \int_{\mathcal{F}_n} \phi_{n|n-1}(\mathcal{F}_n|\mathcal{F}_{n-1})p_{n-1|n-1}(\mathcal{F}_{n-1}|y_{1:n-1})\delta\mathcal{F}_{n-1} \\ &= \phi_{n|n-1}(\mathcal{F}_n|\varnothing)p_{n-1|n-1}(\varnothing|y_{1:n-1}) \\ &\quad + \int \phi_{n|n-1}(\mathcal{F}_n|\alpha_{n-1})p_{n-1|n-1}(\alpha_{n-1}|y_{1:n-1})d\alpha_{n-1}. \end{aligned} \tag{26}$$

Remark 1: The one-step prediction density of the appearance of PU is

$$q_{n|n-1} = p_b \times (1 - q_{n-1|n-1}) + p_s \times q_{n-1|n-1}, \tag{27}$$

and the predicted spatial PDF of associative CLSI is

$$f_{n|n-1}(\alpha_n) = \frac{p_b(1-q_{n-1|n-1})b_{n|n-1}(\alpha_n)}{q_{n|n-1}} + \frac{p_s q_{n-1|n-1} \times \int_{\alpha_{n-1}} \pi_{n-1|n-1}(\alpha_n|\alpha_{n-1})f_{n-1|n-1}(\alpha_{n-1})d\alpha_{n-1}}{q_{n|n-1}}. \tag{28}$$

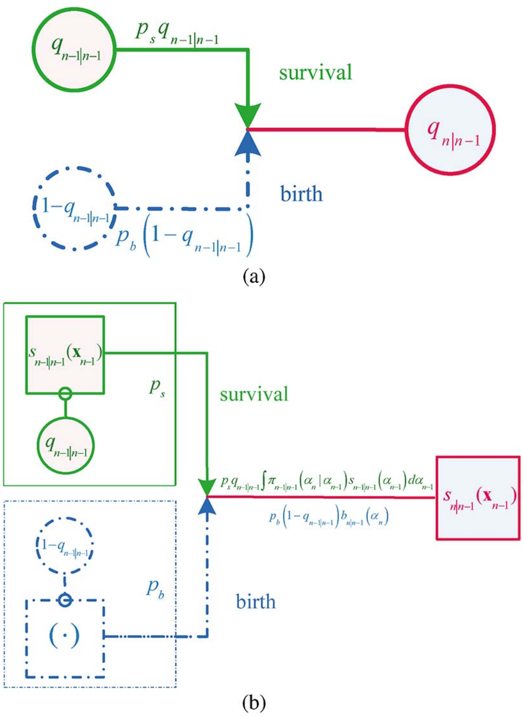


Fig. 1. An illustration of the predicting process in the recursive estimation algorithm. (a) The one-step prediction of the existence density $q_{n|n-1}$. (b) The one-step prediction of the spatial density $f_{n|n-1}(\alpha_n)$.

Generally, the above two prediction densities are derived by checking respectively two complementary situations, i.e., $\mathcal{F}_n = \varnothing$ and $\mathcal{F}_n = \{\alpha_n\}$, and by further interpreting the terms of eqs. (17)–(20) with eq. (26).

For more detailed mathematical derivation on RFS, one may refer to [31] and [34]. As the communication practitioners, here we focus on some intuitive explanations on the derived predicting densities, especially in the context of the considered DS application. As illustrated by Fig. 1(a), the predicted density of PU's appearance is a result of the combined action of two dynamic switches.

1) It is seen from eq. (27) that the predicted existence probability $q_{n|n-1}$ involves two *complementary* terms. The first term, i.e., $p^b_{n|n-1} \triangleq p_b \times (1-q_{n-1|n-1})$, implies a *birth* procedure of a dynamic PU. Recall that, in eq. (17), the posterior density $q_{n-1|n-1} = \Pr\{|\mathcal{F}_{n-1}| = 1|y_{1:n-1}\}$ testifies the presence of PU at time $(n-1)$, while the term $(1-q_{n-1|n-1})$ indicates the absence of PU. So, the term $p_b(1-q_{n-1|n-1})$, further given the one-step Markov transition property of PU states, actually manifests the re-birth of a disappeared PU at time n.

2) The second term, i.e., $p^s_{n|n-1} \triangleq p_s \times q_{n-1|n-1}$, accounts for the *survival* procedure of a dynamic PU. As the continuation of an appeared PU at time $n-1$, the term $p_s \times q_{n-1|n-1}$ reflects the sustained exitance of a PU at next time n.

Similarity, the predicted special density in eq. (28) also involves two terms [34], whose physical meanings are given as below.

1) The first term, i.e., $p_b(1-q_{n-1|n-1}) \times b_{n|n-1}(\alpha_n)$, characterizes essentially the PU's dynamic *birth* procedure. Once a PU is re-birthed at time n, then the *a priori* spatial density $b_{n|n-1}(\alpha_n)$ is required to predict the distribution of associative

fading states, even in the absence of any past likelihood information. Thus, the birth component of the predicted spatial density, conditioned on a birth probability $p_b(1-q_{n-1|n-1})$, will be easily derived.

2) The second term, i.e., $p_s q_{n-1|n-1} \times \int_{\mathcal{A}} \pi_{n-1|n-1}(\alpha_n|\alpha_{n-1}) f_{n-1|n-1}(\alpha_{n-1})d\alpha_{n-1}$, reveals the PU's dynamic *survival* process. It is easily found that, for a survival PU already appeared at time $n-1$, the one-step predicted density will be given by $\int_{\alpha_{n-1}\in\mathcal{A}} \pi_{n-1|n-1}(\alpha_n|\alpha_{n-1}) f_{n-1|n-1}(\alpha_{n-1})d\alpha_{n-1}$. Recall that the predicted survival probability is given by $p_s q_{n-1|n-1}$, then the survival component of the spatial density is relatively straightforward.

According to the total probability formula, the predicted spatial density will be contributed consequently by the above two complementary procedures [34]. An illustration of the predicted spatial density is shown by Fig. 1(b).

2) Update Stage: Based on the predicted density $p_{n|n-1}(\mathcal{F}_n|y_{1:n-1})$, the denominator term of eq. (26) becomes to eq. (29), by applying the set integration.

$$\begin{aligned} p(y_n|y_{1:n-1}) &= \int_{\mathcal{F}_n} \varphi_n(y_n|\mathcal{F}_n) p_{n|n-1}(\mathcal{F}_n|y_{1:n-1})\delta\mathcal{F}_n \\ &= \varphi_n(y_n|\varnothing)p_{n|n-1}(\varnothing|y_{1:n-1}) \\ &\quad + \int_{\alpha_n\in\mathcal{A}} \varphi_n(y_n|\alpha_n)p_{n|n-1}(\alpha_n|y_{1:n-1})d\alpha_n. \end{aligned} \tag{29}$$

Remark 2: The updated posterior density of the PU's appearance is given by:

$$\begin{aligned} q_{n|n} &= q_{n|n-1} \times \int_{\alpha_n\in\mathcal{A}} \varphi_n\left(y_n|\{\alpha_n\}\right) f_{n|n-1}(\alpha_n)d\alpha_n \\ &\div \Bigg[(1-q_{n|n-1})\varphi_n(y_n|\varnothing) \\ &\quad + q_{n|n-1} \int_{\alpha_n\in\mathcal{A}} \varphi_n\left(y_n|\{\alpha_n\}\right) f_{n|n-1}(\alpha_n)d\alpha_n\Bigg], \end{aligned} \tag{30}$$

and the updated spatial PDF of associative CLSI is

$$f_{n|n}(\alpha_n) = \frac{r_n\left(y_n|\{\alpha_n\}\right) f_{n|n-1}(\alpha_n)}{\int_{\alpha_n\in\mathcal{A}} r_n\left(y_n|\{\alpha_n\}\right) f_{n|n-1}(\alpha_n)d\alpha_n}, \tag{31}$$

where $r_n(y_n|\alpha_n)$ accounts for the ratio of likelihoods between two hypothesis, i.e.,

$$r_n\left(y_n|\{\alpha_n\}\right) \triangleq \varphi_n\left(y_n|\{\alpha_n\}\right)/\varphi_n(y_n|\varnothing). \tag{32}$$

Similar to eqs. (27), (28), some intuitive explanations (maybe not rigorous) are given to the derived densities $q_{n|n}$ and $f_{n|n}(\alpha_n)$, see refs [31], [34].

1) In the eq. (30), the numerator is represented by the multiplication between the prior predicted density $q_{n|n-1}$ and the likelihood $\int_{\alpha_n\in\mathcal{A}} \varphi_n(y_n|\{\alpha_n\}) f_{n|n-1}(\alpha_n)d\alpha_n$ (marginalized on the predicted spatial density $f_{n|n-1}(\alpha_n)$), leading to the updating on $q_{n|n-1}$. The updated term is further normalized by the total probability of the denominator.

2) In the eq. (31), the numerator is expressed as the multiplication between the prior predicted density $f_{n|n-1}(\alpha_n)$ and the likelihood ratio $r_n(y_n|\{\alpha_n\})$, also suggesting an updating procedure. Then, the updated term is normalized by the total probability of the denominator.

Up to now, the joint estimations of two posterior densities, i.e., the probability existence $q_{n|n}$ and the spatial PDF $f_{n|n}(\alpha_n)$, have been computed recursively. Further combined with two predict distributions, i.e., $q_{n|n-1}$ and $f_{n|n-1}(\alpha_n)$, the Bernoulli filter may be implemented then and, consequently, the Bernoulli RFS $\mathcal{F}_n$ would be estimated.

Note from the above analysis, with the single unified RFS $\mathcal{F}_n$, the unknown cardinality of BRFS (i.e., $q_{n|n}$) will be detected along with the associative CLSI (i.e., α_n). The DS scheme, therefore, is in sharp contrast to another existing three-step estimation scheme [21], [35], in which two hidden terms are treated as two *separate* quantities and are estimated successively. Except for its generality and flexibility, the unified DS framework is supposed also to have more competitive sensing performance by acquiring two hidden states simultaneously.

E. Designing Considerations

An illustration of the state evolutions is given by Fig. 2(a). It is seen that, for the formulated DS problem with slowly-varying fading channels, the transitions of two states may be mutually asynchronous. I.e., after the transition of fading states, then the static channel gain α_n will remain invariant or *homologous* in L successive slots, whilst the PU's existence state s_n may switch dynamically among different slots n. Note that, this may remain somewhat different from the classical RFS. Recall that, traditionally, the cardinality (or *on/off* state) and its associated states are assumed to evolve simultaneously, or as in typical object position tracking scenarios [34], [40], the on/off state s_n will be invariant for a long while. The DS estimation algorithm, accordingly, should be refined or re-designed to accommodate this specific situation, as the complementary to [40]. In the section, we will integrate two mechanisms to maximize the functioning of sequential estimation algorithm which thereby may produce the promising detection and estimation performance.

1) Observations Accumulation: Since the slowly varying fading channel is temporarily unchanged during successive L slots, the historical measurements could be utilized to promote the estimation performance furthermore. Denote the latest time index the fading channel switches by $n' = \lfloor n/L \rfloor$, in subsequent slots $n = Ln' + l$ $(0 < l < L)$, then the observations will be accumulated to refine the estimation in the case of $q_{n|n} \geq \gamma$. The underlying principle behind this observation accumulation mechanism is that, with more likelihood information exploited, the estimation of fading gains accompanying unknown PU states would be reinforced.

To be specific, if the total l_1 $(0 \leq l_1 \leq l)$ former $q_{n|n}$ is larger than a threshold γ (indicating the estimated PU state $\hat{s}_n = 1$), then the observation of the time $n = n'L + l$ will be redefined as:

$$y_{n'L+l_1} = y_{n'L} + \sum_{t=1}^{l_1} y_{n'L+t}, \tag{33}$$

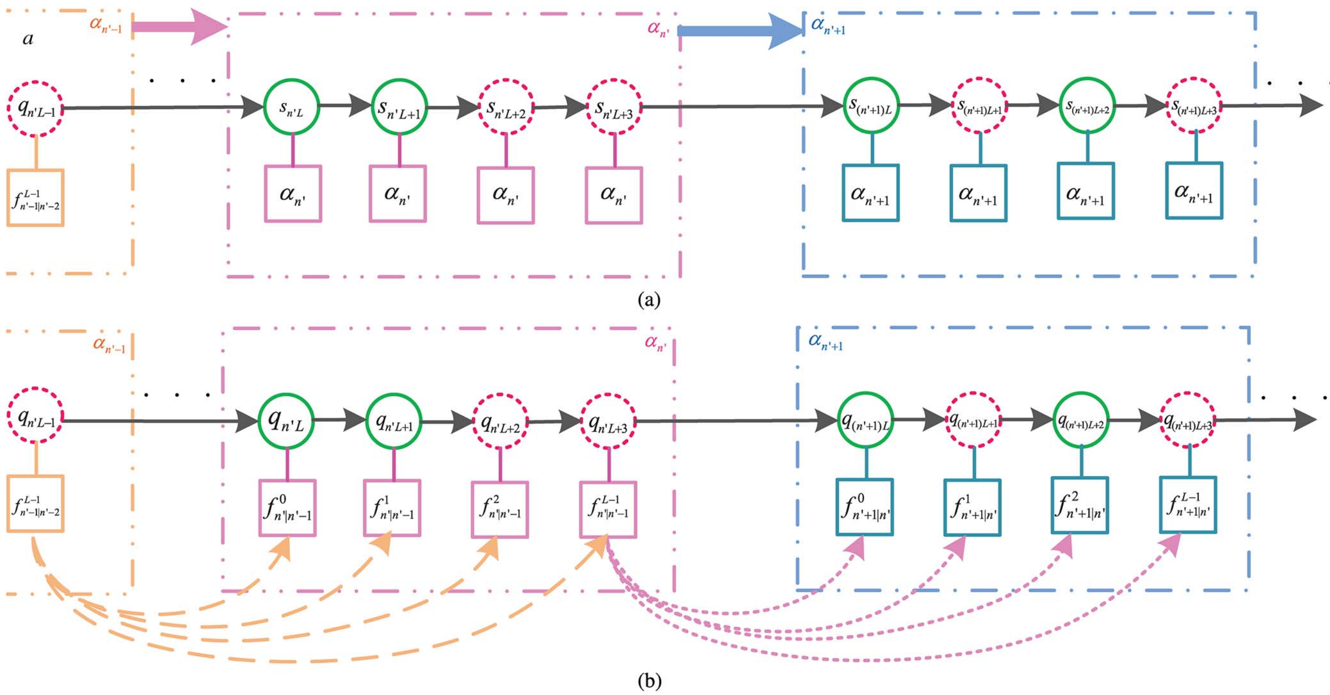


Fig. 2. (a) Illustration of multiple transitional states in the DS process. (b) Illustration of the densities transitions in the DS process. The *solid* green circle denotes a PU is in sleep state, while the *dotted* red circle denotes a PU is in active state. Here, the time-varying fading channel of a static length $L = 4$ is taken for example. For clarity, the term $f^{l}_{n'|n'-1}$ accounts for $f_{n|Ln'-1}(\alpha_{n'-1})$ with $n = Ln' + l$ $(l = 0, 1, 2, \cdots, L-1)$.

with an evolving likelihood density

$$\varphi_n\left(y_{n'L+l_1}|\alpha_n\right) = \frac{\left(y_{n'L+l_1}/\varrho'\right)^{(M-2)/4}}{2\sigma^2} \times \exp\left[-\left(y_{n'L+l_1}+\varrho'\right)/2\sigma'^2\right] \times I_{M/2-1}\left(\sqrt{\varrho' y_{n'L+l_1}}/\sigma'^2\right), \tag{34}$$

where $I_{M/2-1}(\cdot)$ is the Bessel function of the first kind. Here, the involved parameters are given by:

$$\varrho' = \sum_{m=1}^{l_1 M} |\alpha_n d_{n,m}|^2 = M l_1 \alpha_n^2 \sigma_d^2, \tag{35}$$

$$\sigma'^2 = \alpha_n^2\sigma_d^2 + \sigma_z^2. \tag{36}$$

2) Homologous Estimation: Rather than utilizing the distribution of the previous time $(n-1)$, in the current slot $n = Ln' + l$ $(l > 1)$, the spatial density of the last slot in the previous static period $n'-1$, i.e., the time slot $L(n'-1)+L-1 = Ln'-1$, will be used in both the predicting and updating procedures, as shown by Fig. 2(b). Thus, the eqs. (28) and (31) will become to:

$$\begin{aligned} f_{n|Ln'-1}(\alpha_n) &= \frac{p_s q_{n-1|n-1}}{q_{n|n-1}} \times \int_{\alpha_{n'-1}} \pi_{n-1|n-1}(\alpha_n|\alpha_{n'-1}) f_{Ln'-1|Ln'-1}(\alpha_{n'-1}) d\alpha_{n'-1} \\ &\quad + \frac{p_b(1-q_{n-1|n-1}) \times b_{n|n-1}(\alpha_n)}{q_{n|n-1}}, \end{aligned} \tag{37}$$

and

$$f_{n|n}(\alpha_n) = \frac{r_n\left(y_n|\{\alpha_n\}\right) f_{n|Ln'-1}(\alpha_n)}{\int_{\alpha_n\in\mathcal{A}} r_n\left(y_n|\{\alpha_n\}\right) f_{n|Ln'-1}(\alpha_n) d\alpha_n}. \tag{38}$$

Premised on such a *homologous* propagation strategy, the predicting and updating stages would be implemented in accordance with the state transitions in Fig. 2(a). Note that, with this suggested mechanism, now the predicting and updating processes utilize the previous spatial density $f_{Ln'-1|Ln'-1}(\alpha_{n'-1})$ and $f_{n|Ln'-1}(\alpha_n)$, respectively. The motivation of adopting the term $f_{Ln'-1|Ln'-1}(\alpha_{n'-1})$ is that, with the first accumulation mechanism, the spatial distribution of time $Ln'-1$ may become very accurate based on the former L_1 innovative observations. In practice, we have $\mathsf{E}(L_1) = L \times \Pr(s_n = 1)$.

F. PF-Based Implementations

It is observed from eq. (37) the the prediction of the spatial density $f_{n|n-1}(\alpha_n)$ involves the complex computation that is prohibitive to practical uses. The sequential importance sampling (SIS) based PF, as an alternative, shows great promise to the DS problem, which will greatly facilitate the Bayesian inference via a simulated Monte-Carlo approach.

1) PF: PF obtains the consistent estimation of a complex density via $f_n(x) \simeq \sum_{i=0}^{I} w_n^{(i)} \times \delta\left(x - x_n^{(i)}\right)$, relying on I randomly discrete measures (or particles) $x_n^{(i)}$ with evolving probability masses (or weights) $w_n^{(i)}$ $(i = 1, 2, \cdots, I)$ [30], [34], [35], [37]. Here, $\delta(x)$ is the Dirac-delta mass function. The particles are drawn from a related space, i.e., $x_n^{(i)} \sim \rho(x_n)$, and then the particle weights are updated by:

$$w_n^{(i)} = p\left(y_{0:n}\left|x_{0:n}^{(i)}\right.\right) p\left(x_{0:n}^{(i)}\right) / \rho(x_n). \tag{39}$$

2) Implementations: For the DS scenario, since the predicted spatial density involves two terms as in eq. (37), two groups particles will be simulated accordingly from [34], [40]:

$$x_{n|n-1}^{(i)} = \begin{cases} \xi\left(x_{n|n-1}\left|x_{n-1|n-1}^{(i)}, y_{1:n-1}\right.\right), \; i = 1, \cdots, N, & \text{(40a)} \\ \beta(x_{n|n-1}|y_{1:n-1}), \; i = N+1, \cdots, N+B, & \text{(40b)} \end{cases}$$

where the first N particles drawn from a proposal survival density $\xi\left(x_{n|n-1}|x_{n-1|n-1}^{(i)}, y_{1:n-1}\right)$ are used to approximate the first term in (37); while the latter B particles simulated from a birth density $\beta(x_{n|n-1}|y_{1:n-1})$ are used to evaluate the second term. Then, the importance weights would be updated via eq. (39).

In the above elaboration, $\left\{x_{n-1|n-1}^{(i)}, w_{n-1|n-1}^{(i)}\right\}_{i=1}^{I=N+B}$ correspond to the posterior density $\hat{f}_{n-1|n-1}(\alpha_n) = \sum_{i=1}^{N+B} w_{n-1|n-1}^{(i)} \delta\left(x - x_{n-1|n-1}^{(i)}\right)$. In order to draw B birth particles and N survival particles, a birth density $\beta(x_{n|n-1}|y_{1:n-1})$ and another proposal survival density $\xi\left(x_{n|n-1}|x_{n-1|n-1}^{(i)}, y_{1:n-1}\right)$ should be designed.

a) Proposal survival-density: In order to be consistent with eq. (37), in the current slot $n = Ln' + l$ $(0 \le l \le L-1)$, the survival density will be derived directly from one-step transitions of N *survival* particles of the time $Ln' - 1$, i.e.,

$$\begin{aligned} \xi\left(x_{n|n-1}^{(i)}\right) &\triangleq \xi\left(x_{n|Ln'-1}^{(i)}\right), \\ &= \int_{x' \in \mathcal{A}} \pi_{n|n-1}(x|x') \times f_{Ln'-1|Ln'-1}(x') dx', \\ & i = 1, 2, \cdots, N, \end{aligned} \tag{41}$$

where the previous survival density is constructed by:

$$\begin{aligned} & f_{Ln'-1|Ln'-1}(x) \\ & \quad \simeq \sum_{i=1}^{N} \delta\left\{x - x_{Ln'-1|Ln'-1}^{(i)}\right\} \times w_{Ln'-1|Ln'-1}^{(i)}. \end{aligned} \tag{42}$$

The new weights $w_{n|Ln'-1}^{(i)}$ will be updated according to eq. (39). Note that, a re-sample procedure will be applied then [30], [34], [40].

b) Proposal birth-density: The proposal birth density is specified as:

$$\begin{aligned} x_{n|Ln'-1}^{(i)} &\sim \beta(x_{n|n'-1}) \triangleq \pi_{n'|n'-1}(\alpha_{n'}|\hat{\alpha}_{n'-1}), \\ & i = N+1, N+2, \cdots, N+B. \end{aligned} \tag{43}$$

Note that, the predicted particle weights is specified by:

$$w_{n|Ln'-1}^{(i)} = 1/B, \quad i = N+1, N+2, \cdots, N+B. \tag{44}$$

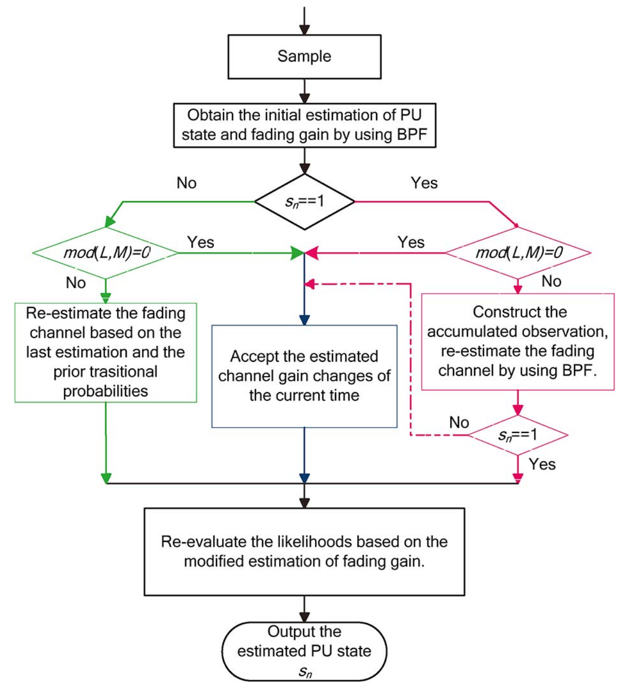


Fig. 3. Schematic flow of the DS algorithm with observation accumulations.

Thus, the predict process in eq. (37) will be approximated by PF via two group of discrete particles. Thereafter, the particle weights are updated based on the new observation y_n. Finally, both the spatial density and unknown PU states will be derived. One may refer to [34], [40] for more details.

With the Bayesian rule, the threshold may be empirically configured to $\gamma = 0.5$, thus we have:

$$\hat{s}_n = \begin{cases} 1, & \text{if } q_{n|n} \ge \gamma, \quad \text{(45a)} \\ 0, & \text{if } q_{n|n} < \gamma. \quad \text{(45b)} \end{cases}$$

Given the observation accumulation mechanism, the survival particles would be re-determined based on the accumulated observation $y_{Ln'+l_1}$ $(0 \le l_1 \le l)$ in the case of $\hat{s}_{Ln'+l} = 1$, as in eq. (46). It is seen that, as a consequence, the estimation accuracy will be improved as the time goes on.

$$\begin{aligned} & w_{Ln'+l|Ln'+l}^{(i)} \propto w_{n|L(n'-1)}^{(i)} \times r\left(y_{Ln'+l_1}\left|x_{n|Ln'-1}^{(i)}\right.\right), \\ & \qquad i = 1, \cdots, N, N+1, \cdots, N+B. \end{aligned} \tag{46}$$

After the above weights update, the new survival particles of the time $(Ln' + l)$ will be derived. Notice that, a similar resample process will be used after updating the weights.

3) Algorithm Flow: Based on the above elaborations, the implementation flow of the proposed DS algorithm, which acquires the varying fading gains along with PU states, is demonstrated by Figs. 3 and 4. Firstly, (1) based on the current observation y_n, the posterior existence probability $q_{n|n}$ and the spatial PDF $f_{n|n}(x)$ are estimated via the prediction and update stages; (2) in the case of $q_{n|n} > 1$, the observation will be accumulated to further promote the estimations; (3) and finally, the DS is realized by estimating unknown PU states along with the associated fading states.

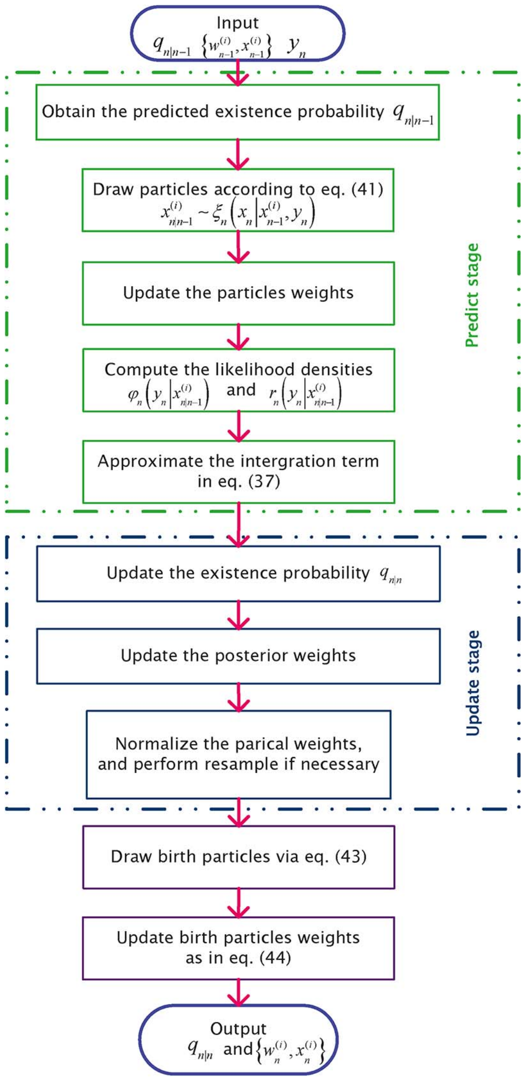


Fig. 4. Sub-algorithm flow of the DS scheme implemented by PF.

With the accumulation procedure, it is worth noting a counter, namely l_1, is adopted by the proposed DS algorithm. Such a counter l_1 is mainly used to record the numbers of accumulated observations (as the sensing slot index n increases), which is of great importance to determine the DoF of an accumulated observation $y_{Ln'+l_1}$. Such a counter is reset to 0 every time the state transition of fading gains occurs, i.e., when $mod(n, L) = 0$. Subsequently, after the evaluation of $y_{Ln'+l_1}$ and the estimation refinement of fading gains, the counter l_1 will be updated by $l_1 = l_1 + 1$.

The algorithm complexity can be measured roughly by the total numbers of multiplications. Firstly, in order to obtain the summed-energy observation, $\mathcal{O}(M)$ multiplications are required. Then, for the subsequent BPF, the multiplication operations in implementing PF are basically proportional to the size of particles, i.e., $I = N + B$. Denote the multiplications in the computation of likelihood density by ϑ, which is practically related with the representative precision of numbers and the adopted algorithms, then the involved complexity is about $\mathcal{O}\{(K_1 + I)\vartheta\}$, where K_1 is the number of transitional states. Thus, the total complexity of the proposed DS algorithm will be measured by $\mathcal{O}\{M + (K_1 + I)\vartheta\}$. Note that, K_1 may decrease to 3 for a first-order DSMC. When the sample size M is large, the complexity of the DS algorithm is comparable with that of ED, i.e., $\mathcal{O}(M)$.

IV. Simulation Results

Based on experimental simulations, the presented DS algorithm will be evaluated in realistic TVFF condition. In practice, a false alarm P_f and a missing detection P_m may have different effects to CR. P_f measures the spectral utilization, while the missed detection probability P_m controls the interferences from CR users to PUs. With a classical NP criterion, a main objective is either to minimize P_m for a target P_f, or to minimize the P_f for a target P_m. It is noteworthy that the DS scheme is designed specially under the Bayesian criterion and, accordingly, the spectral utilization to unused bands and the potential interference to PUs will be considered jointly. Thus, the total detection probability is adopted as a performance metric as in [21], [36]–[38], i.e.,

$$P_D \triangleq 1 - p(H_1)P_m - p(H_0)P_f. \tag{47}$$

Another performance measurement of the DS process, i.e., the MSE of the estimated fading channel, is defined as:

$$MSE \triangleq \mathsf{E}\left\{\sum_{n=1}^{N} \left|\hat{\alpha}_n - \alpha_n\right|^2 / \sum_{n=1}^{N} |\alpha_n|^2\right\}. \tag{48}$$

A. Sample Size M

In the first experiment, the sensing performance with different sample sizes M is investigated. The static length of fading channel is configured to $L = 20$, the transitional probability (p_b, p_s) are set to (0.2, 0.8). Note that, when the TPM of eq. (4) is unknown in practice, the estimation scheme of [39] can be applied to acquire the transitional parameters. The discrete states of DSMC channel is set to $K = 5$. From the numerical results shown by Fig. 5(a), it is seen that the estimation MSE of fading gains will be decreased with the increasing of SNR, which would finally become convergent. Meanwhile, the MSE may be decreased by increasing the sample size M. Taking SNR $= -2$ dB for example, the MSE value in the case of $M = 50$ is about 0.068. When the sample size is increased to $M = 200$, the MSE will be decreased to 0.027. Similar trends can be observed from the sensing performance demonstrated by Fig. 5(b). Specifically, when the total detection probability P_D is 0.8, then a detection gain of 2 dB may be achieved by increasing the sample size M from 100 to 200. It is comprehensible that, with more independent samples accumulated, the statistical information of observations will be more accurate and, therefore, the estimation performance will be enhanced. Note that, however a compromise should be made, since a larger M results also in the prolonged sensing time and reduced transmission efficiency.

B. Static Length L

In the second experiment, the effects from various static length L are investigated. In the simulation, we configure the

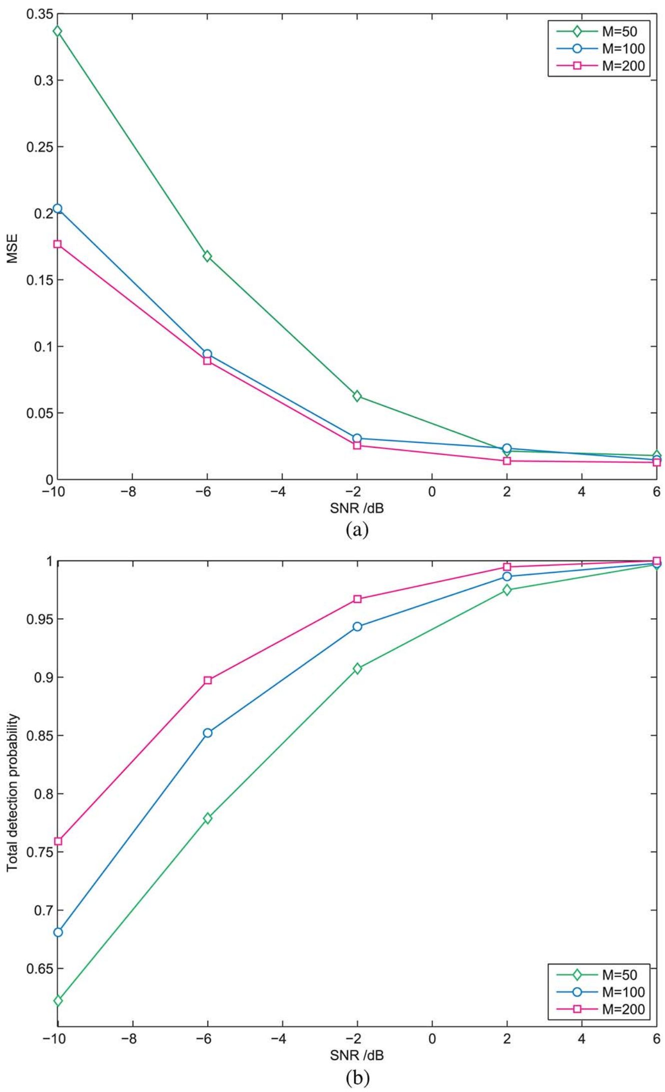


Fig. 5. Deep sensing performance under differen sample size M. (a) The estimation MSE performance with $L = 20$. (b) The detection performance with $L = 20$.

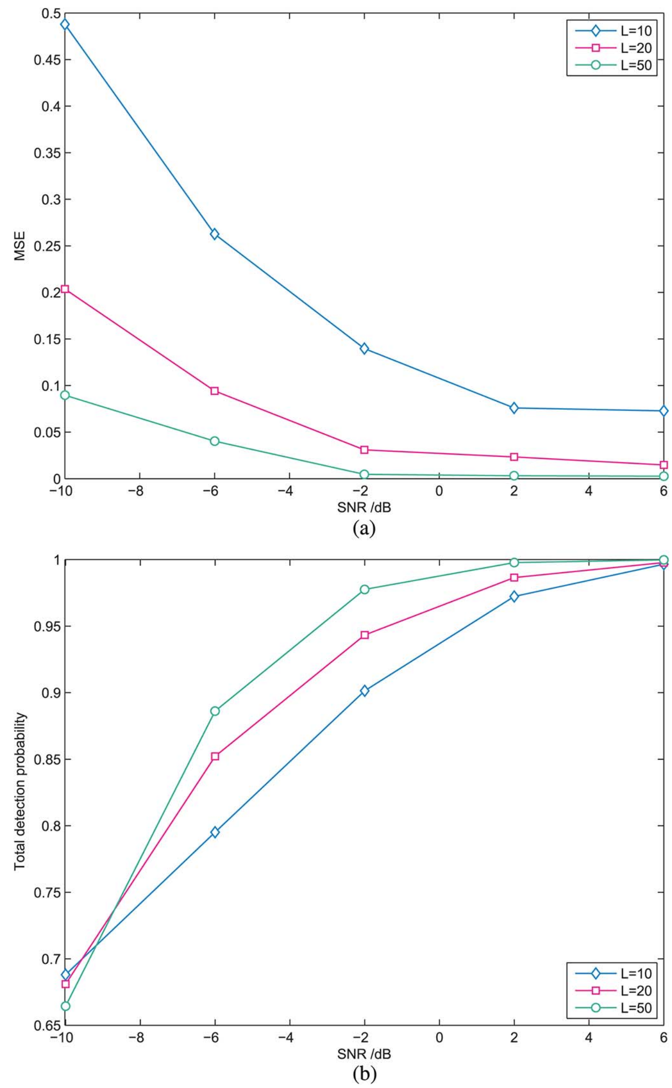


Fig. 6. Deep sensing performance under differen static length L. (a) The estimation MSE performance with $M = 50$. (b) The detection performance with $M = 50$.

sample size M to 100 and *a priori* transitional probability (p_b, p_s) to (0.2, 0.8).

In realistic applications, the encountered TVFF channel is slowly varying, i.e., f_D is relatively small. Thus, three typical configurations are considered, i.e., the static length L is 10, 20 and 50, respectively. From the MSE curves of the DS scheme shown by Fig. 6(a), it is seen that the estimation accuracy may be improved dramatically with the increasing of a static length L. For example, if the SNR is configured to 2 dB, the MSE value in the case of $L = 10$ is about 0.077. In comparison, the MSE of $L = 20$ will be noticeably decreased to 0.023. In the presence of a larger L, more historical information would be exploited and more profound refinement on previous inaccurate estimations will be ensured. In contrary, with a faster varying channel, further refinements on the erroneous estimation of fading states via the L_1-segment observations may be weakened due to a shortened static period.

As far as the sensing performance is concerned, a similar trend may be observed from the provided results in Fig. 6(b). That is, the larger the static length L, the higher the total detection probability P_D. For example, a detection gain of 3.5 dB may be obtained by a configuration of $L = 20$ (e.g., $P_D > 0.95$), compared with the faster time-varying fading of $L = 10$.

C. Comparative Performance Analysis

In this experiment, we will compare the performance of the proposed DS scheme with the existing counterpart, i.e., the joint estimation-based three-step method. In numerical experiments, the sample size is set to $M = 100$ and the static fading length is $L = 20$. The particle size of both the DS scheme and the other joint-estimation scheme is configured to $I = 200$. Simulation results are plotted in Fig. 7. It is seen that the DS scheme is superior to the other joint estimation algorithm. To be specific, a detection gain of 1$\sim$2 dB is achieved by the DS scheme from Fig. 7(a). This is because the joint estimation scheme in [35] is sub-optimal, where the PU state and dynamic fading gains are estimated independently via a successive manner. The DS

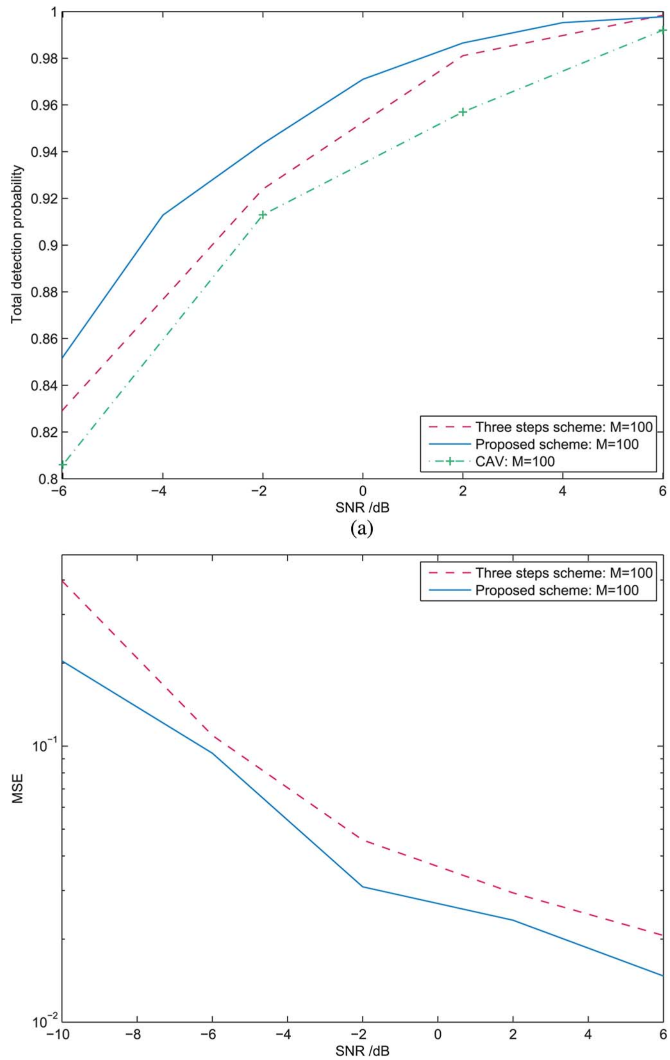


Fig. 7. (a) The total probability detection P_D of the proposed scheme and the other existing estimation schemes. (b) The MSE performance of the proposed scheme and the existing 3-step joint estimation scheme. Here, the static length of time-varying fading channel is configured to $L = 20$.

scheme, in contrast, acquires two unknown states *as a whole*, which thereby exploits more entirely the underlying dynamics of the CLSI. Except for the advantage of sensing performance, the channel MSE of the DS scheme is also superior to the joint estimation method. As shown by Fig. 7(b), an estimation gain of 2 dB may be observed in high SNRs (e.g., >0 dB).

We further compare the new DS scheme with another covariance absolute value (CAV) based algorithm, which relies on the statistical covariance of PU signals. As a widely recommended technique, the CAV scheme essentially utilizes the temporal/spatial correlation of received signals [13], which is relatively robust to the time-varying fading. For the CAV method, we assume the over-sampling technique is used (e.g., for narrow-band PU signals). A Gaussian pulse shaper of $N_s = 11$ samples is used, and the smoothing length is 10. We note from Fig. 7(a) that a detection gain around 3∼4 dB will be achieved by the new scheme, when M is set to 100 and the total detection probability $P_D > 0.95$. Although the two methods are all premised on the statistical information, the DS scheme

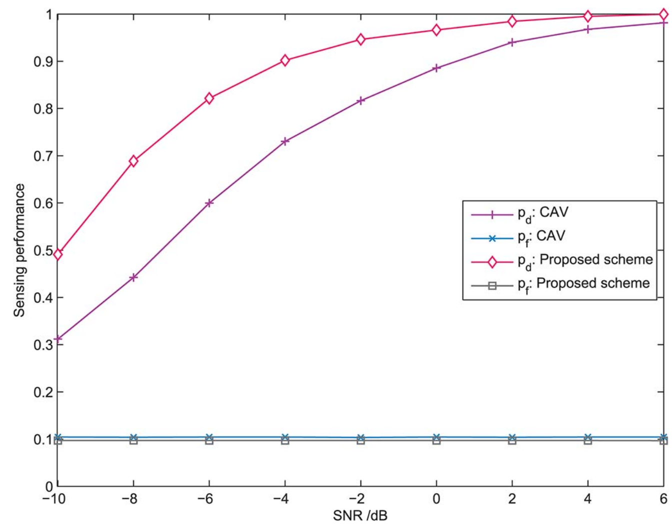


Fig. 8. The probability of detection P_d for a specified probability of false-alarm $P_f = 0.1$.

would track fading states dynamically and, therefore, exploits the underlying channel memory to further enhance the sensing performance. The CAV method focuses merely on the correlation of PU signals, by ignoring the time-varying property of fading gains. Notice that, in some special cases, unfortunately the correlation of either PU signals or noise samples may tend to be very weak. The performance of the CAV method, which essentially relies on PU's signal correlations, will be then deteriorated seriously [35], [37]. The DS method, fortunately, is immune to different PU's signals, which may have more widespread applications.

D. Practical Considerations

It is noteworthy that, by maximizing *a posteriori* probability, our DS scheme belongs essentially to a Bayesian approach. Accordingly, the total detection probability $P_D = P(H_0) \times (1 - P(H_1|H_0)) + P(H_1) \times P(H_1|H_1)$, which is dramatically different from the single *probability of detection* $P_d \triangleq P(H_1|H_1)$, is evaluated. A merit of the compound metric is that, as suggested by [35], [36], both the utilization of primary bands and the interference to PUs will be organically considered. In classical threshold-based techniques (e.g., ED), another Neyman-Pearson criterion is applied by prescribing a single probability of false-alarm $P_f \triangleq P(H_1|H_0)$, which, however, may be no longer suitable to this DS scheme. As demonstrated, with the MAP approach and an objective of maximizing the compound probability P_D, the false-alarm probability of the DS scheme will be not fixed [35].

Even so, the performance of the DS scheme could still be evaluated under the Neyman-Pearson criterion. Note that, in this case, the estimated fading channels can be hardly exploited by a simple threshold coinciding with a target P_f. The detection probability P_d of the new method, accompanying the joint estimation scheme [35], will be comparable to ED (recall that the observation in the new DSM is actually the summed-energy). In this experiment, the static fading length is $L = 20$ and the sample size is $M = 100$. It is found from Fig. 8 that, as far as the single P_d is concerned, the DS scheme is more attractive than the CAV method for the target $P_f = 0.1$. More importantly, as another extra gift of the sensing procedure, the recovered fading channels (e.g., with the estimation MSE $= 0.014$ when SNR $=$ 6 dB and $M = 100$) are of great significance to subsequent DSS strategies optimization.

Compared with traditional techniques, the new DS scheme creates a unified sensing paradigm for future DSS. As a major advantage of the suggested scheme over other joint estimation methods, the recursive algorithm would be extended conveniently to any other important CLSI, e.g., PU's unknown modulation format or its moving positions [38]. By providing both the accurate sensing report and such additional CLSI, the DS framework and the resulting deep cognition is more attractive to the DSS of next-generation communications, by enabling more flexible accessing strategies and more effective resource allocations.

V. Conclusion

A general DS paradigm is proposed for future DSS applications, which is intended to meet more adverse propagation conditions, e.g., time-varying fading channel in emerging mobile scenarios. In order to fully exploiting the established DSM, two complementary mechanisms are specially integrated into the DS estimation algorithm that is essentially based on a Bernoulli RFS. It is found that, by detecting the unknown PU states and estimating the associated fading gains as a whole, the underlying dynamic property of the time-correlated channel can be fully utilized to enhance the sensing performance. Compared to the main counterpart, i.e., the other joint estimation based 3-step method, both the sensing performance and the channel estimation MSE are improved by this DS scheme. Numerical simulations are provided to validate the DS scheme. Despite the context of TVFF channel, the formulated new DSM and the DS scheme can be extended to other scenarios, which puts an insight into spectrum sensing and may provide a new idea for the next-generation DSS of 5G communications. Future research directions may include (1) the generalization to the unknown TPM of time-varying fading channel, and (2) the theoretical analysis on the lower bound of the sensing or MSE performance.

References

[1] Y. C. Liang, A. T. Hoang, and H. H. Chen, "Cognitive radio on TV bands: A new approach to provide wireless connectivity for rural areas," *IEEE Wireless Commun.*, vol. 15, no. 3, pp. 16–22, Jun. 2008.

[2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[3] 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.

[4] L. Lu, X. W. Zhou, U. Onunkwo, and G. Y. Li, "Ten years of cognitive radio technology," *EURASIP J. Wireless Commun. Netw.*, vol. 28, pp. 1–16, 2012.

[5] 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.

[6] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.

[7] F. F. Digham, M. S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channel," in *Proc. IEEE ICC*, Anchorage, AK, USA, May 2003, vol. 5, pp. 3575–3579.

[8] H. S. Chen, W. Gao, and D. G. Daut, "Signature based spectrum sensing algorithms for IEEE 802.22 WRAN," in *Proc. IEEE ICC*, Glasgow, U.K., Jun. 2007, pp. 6487–6492.

[9] P. D. Sutton, K. E. Nolan, and L. E. Doyle, "Cyclostationary signature in practical cognitive radio applications," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 13–24, Jan. 2008.

[10] Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Proc. IEEE Conf. CROWNCOM*, Mykonos Island, Greece, Jun. 2006, pp. 1–5.

[11] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE ICASSP*, Honolulu, HI, USA, Apr. 2007, pp. 1357–1360.

[12] Y. H. Zeng and Y. C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784–1793, Jun. 2009.

[13] Y. H. 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.

[14] X. W. Zhou, Y. Li, Y. H. Kwon, and S. A. "Detection timing and channel selection for periodic spectrum sensing in cognitive radio," in *Proc. IEEE GLOBECOM*, New Orleans, LA, USA, Nov. 30–Dec. 4, 2008, pp. 1–5.

[15] J. Ma and Y. Li, "A probability-based spectrum sensing scheme for cognitive radio," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 3416–3420.

[16] *Draft Standard for Wireless Regional Area Networks Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands*, IEEE P802.22/D1.0, Apr. 2008.

[17] M. Mueck *et al.*, White Paper: Novel spectrum usage paradigms for 5G, IEEE Tech. Committee Cognitive Netw., New York, NY, USA, SIG CR 5G, Nov. 14, 2014.

[18] B. Sklar, "Rayleigh fading channel in mobile digital communication systems Part I: Characterization," *IEEE Commun. Mag.*, vol. 35, no. 7, pp: 90–100, Jul. 1997.

[19] J. G. Proakis, *Digital Communications*, 3rd ed. Singapore: McGraw-Hill, 1995.

[20] M. Bkassiny, L. Yang, and S. K. Jayaweera, "A survey on machine-learning techniques in cognitive radios," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1136–1159, Oct. 2012.

[21] B. Li, Z. Zhou, and A. Nallanathan, "Joint estimation based spectrum sensing for cognitive radios in time variant flat fading channel," in *Proc. IEEE GLOBECOM*, Atlanta, GA, USA, Dec. 2013, pp. 1–6.

[22] Y. Yılmaz, Z. Y. Guo, and X. D. Wang, "Sequential joint spectrum sensing and channel estimation for dynamic spectrum access," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 11, pp. 2000–2012, Nov. 2014.

[23] B. Vujitic, N. Cackov, and S. Vujitic. "Modeling and characterization of traffic in public safety wireless network," in *Proc. Int. Symp. Perform. Eval. Comput. Telecommun. Syst.*, Edinburgh, U.K., Jul. 2005, pp. 213–223.

[24] W. Haselmayr, D. Schellander, and A. Springer, "Iterative channel estimation and turbo equalization for time-varying channel in a coded OFDM-LTE system for 16-QAM and 64-QAM," in *Proc. IEEE 21st Int. Symp. PIMRC*, Istanbul, Turkey, Sep. 26th–29th, 2010, pp. 614–618.

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

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

[27] H. S. Wang and P. Chang, "On verifying the first order Markovian assumption for a Rayleigh fading channel model," *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 353–357, May 1996.

[28] F. Babich and G. Lombardi, "A Markov model for the mobile propagation channel," *IEEE Trans. Veh. Technol.*, vol. 49, no. 1, pp. 63–73, Jan. 2000.

[29] W. B. Davenport, Jr. and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*. New York, NY, USA: IEEE Press, 1958.

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

[31] R. Mahler, *Statistical Multisource Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2007.

[32] B. T. Vo and B. N. Vo, "A random finite set conjugate prior and application to multi-target tracking," in *Proc. 7th IEEE Int. Conf. ISSNIP*, Adelaide, Australia, Dec. 2011, pp. 431–436.

[33] B. T. Vo, B. N. Vo, and A. Cantoni, "Bayesian filtering with random finite set observations," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1313–1326, Apr. 2008.

[34] B. Ristic, B. T. Vo, B. N. Vo, and A. Farina, "A tutorial on Bernoulli filters: Theory, implementation and applications," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3406–3430, Jul. 2013.

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

[36] W. Zhang, R. K. Mallik, and K. B. Letaief, "Cooperative spectrum sensing optimization in cognitive radio networks," in *Proc. IEEE ICC*, Beijing, China, May 19–23, 2008, pp. 3411–3415.

[37] B. Li, M. W. Sun, C. L. Zhao, and A. Nallanathan, "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.

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

[39] Q. Q. Liang, M. Y. Liu, and D. F. Yuan, "Channel estimation for opportunistic spectrum access: Uniform and random sensing," *IEEE Trans. Mobile Comput.*, vol. 11, no. 8, pp. 1304–1316, Aug. 2012.

[40] B. Ristic and J. Sherrah, "Bernoulli filter for joint detection and tracking of an extended object in Clutter," *IET Radar, Sonar Navigat.*, vol. 7, no. 1, pp. 26–35, Jan. 2013.

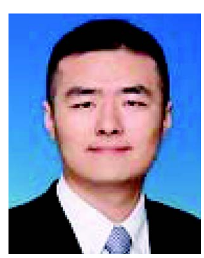
Bin Li received the bachelor's degree in electrical information engineering from Beijing University of Chemical Technology (BUCT), in 2007, and the Ph.D. degree in communication and information engineering from Beijing University of Posts and Telecommunications (BUPT), in 2013. Since then, he joined BUPT as a Lecturer at the School of Information and Communication Engineering (SCIE). His current research interests are focused on statistical signal processing algorithms for wireless communications, e.g., ultra-wideband (UWB), wireless sensor networks, millimeter-wave (mm-Wave) communications, and cognitive radios (CRs). He has published more than 60 journal and conference papers. He received the 2011 ChinaCom Best Paper Award and the 2010 and 2011 BUPT Excellent Ph.D. Student Award Foundation. 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 (IEEE-ICNC'16).

Shenghong Li received the B.S. and the M.S. degrees in electrical engineering from Jilin University of Technology, China, in 1993 and 1996, respectively, and the Ph.D. degree in radio engineering from Beijing University of Posts and Telecommunications (BUPT), China, in 1999. Since 1999, he has been working in Shanghai Jiao Tong University (SJTU), China, as Research Fellow, Associate Professor, and Professor, successively. In 2010, he worked as Visiting Scholar in Nanyang Technological University, Singapore. His research interests include information security, signal and information processing, cognitive radio networks, and artificial intelligence.

He has published more than 80 papers, co-authored 4 books, and holds 10 granted patents. In 2003, he received the 1st Prize of Shanghai Science and Technology Progress in China. In 2006 and 2007, he was elected for New Century Talent of Chinese Education Ministry and Shanghai Dawn Scholar.

Amurugam Nallanathan (S'97–M'00–SM'05) is a Professor of wireless communications in the Department of Informatics at King's College London (University of London). He served as the Head of Graduate Studies in the School of Natural and Mathematical Sciences at King's College London, in 2011 and 2012. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore from August 2000 to December 2007. His research interests include 5G technologies, millimeter wave communications, cognitive radio and relay networks. In these areas, he co-authored nearly 250 papers. He is a co-recipient of the Best Paper Award presented at the 2007 IEEE International Conference on Ultra-Wideband (ICUWB2007). He is a Distinguished Lecturer of IEEE Vehicular Technology Society.

He is an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and a Guest Editor for IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING (Special Issue on Advances in Mobile and Cloud Computing). He was an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2006–2011), IEEE WIRELESS COMMUNICATIONS LETTERS, and IEEE SIGNAL PROCESSING LETTERS. He served as the Chair for the Signal Processing and Communication Electronics Technical Committee of IEEE Communications Society, Technical Program Co-Chair (MAC track) for IEEE WCNC 2014, Co-Chair for the IEEE GLOBECOM 2013 (Communications Theory Symposium), Co-Chair for the IEEE ICC 2012 (Signal Processing for Communications Symposium), Co-Chair for the IEEE GLOBECOM 2011 (Signal Processing for Communications Symposium), Technical Program Co-Chair for the IEEE International Conference on UWB 2011 (IEEE ICUWB 2011), Co-Chair for the IEEE ICC 2009 (Wireless Communications Symposium), Co-Chair for the IEEE GLOBECOM 2008 (Signal Processing for Communications Symposium) and General Track Chair for IEEE VTC 2008. He currently serves as the Technical Program Co-Chair for IEEE ICNC'2016. He received the IEEE Communications Society SPCE Outstanding Service Award 2012 and IEEE Communications Society RCC Outstanding Service Award 2014.



Yijiang Nan received the B.S. degree in communication engineering from Tianjin University, China, in 2013. He is currently pursuing the master's degree at Beijing University of Posts and Telecommunications (BUPT). His research interests include cognitive radios and statistical signal processing.

Chenglin Zhao received the bachelor's degree in radio-technology from Tianjin University, in 1986 and the master's degree in circuits and systems and the Ph.D. degree in communication and information system from Beijing University of Posts and Telecommunications (BUPT), in 1993 and 1997, respectively. At present, he serves as a Professor in Beijing University of Posts and Telecommunications, Beijing, China. His research is focused on emerging technologies of short-range wireless communication, cognitive radios, and 60 GHz millimeter-wave communications.

Zheng Zhou (M'05) received the bachelor's degree in electrical engineering from the Harbin Institute of Military Engineering, Harbin, China, in 1967, and the master's and Ph.D. degrees in electrical engineering from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1982, and 1988, respectively. At present, he is a Professor with Beijing University of Posts and Telecommunications, Beijing, China. Supported by the Hong Kong Telecom International Postdoctoral Fellowship, he was a Visiting Research Fellow in Information Engineering Department of the Chinese University of Hong Kong, Hong Kong, from 1993 to 1995. He was also the Vice-Dean in the School of Telecommunication Engineering of BUPT from 1998 to 2003, and was the invited overseas researcher in Japan Kyocera DDI Future Communication Research Institute (supported by the Japan Key Technology Center) in 2000.

He was a member of the Technical Subcommittee on Cognitive Networks (TCCN), IEEE Communications Society, the International Steering Committee Member of IEEE ISCIT (International Symposium on Communications and Information Technologies) during 2003–2010, and the TPC Co-Chair of the IEEE ISCIT 2005. He was also the General Vice Chair of the IEEE ChinaCom 2006 (the first international conference on communications and networking in China), and the Steering Committee Member of IEEE ChinaCom 2007. He is a voting member and contributor of the IEEE 802.15 Task Group (TG3a and TG4a), Senior Member of China Institution of Communications (CIC), Radio Application and Management Technical Committee Member of CIC, Senior Member of China Computer Federation (CCF), Sensor Network Technical Committee Member of CCF, H Subcommittee Member of China Radio Interference Standard Technology Committee, the General Secretary of China UWB Forum, and the General Secretary of China Bluetooth Forum.