Energy Detection based Spectrum Sensing for Cognitive Radios Over Time-Frequency Doubly Selective Fading Channels

Bin Li, Mengwei Sun, Xiaofan Li, A. Nallanathan, *Senior Member, IEEE* and Chenglin Zhao

***Abstract*—Cognitive radios may operate in practice under various adverse environments. For typical mobile and short-range scenarios, wireless links may tend to be time and frequency selective, i.e., the multipath propagations with time-varying fading coefficients will be inevitable. To cope with the encountered doubly-selective channels, in this paper we present a new spectrum sensing algorithm for distributed applications. Firstly, a dynamic discrete state-space model is established to characterize sensing process, where the occupancy state of primary band and the time-varying multipath channel are treated as two hidden states, while the summed energy is adopted as the observed output. With this new paradigm, spectrum sensing is realized by acquiring primary states and time-dependent multipath channel jointly. For the formulated problem, unfortunately, Bayesian statistical inference may be impractical due to the absence of likelihoods and involved non-stationary distributions. To remedy this problem, an iterative algorithm is further designed by resorting to sequential importance sampling techniques, thus the dynamic non-Gaussian multipath channel and primary states are estimated recursively. Another critical challenge, e.g., the noise uncertainty, is also considered, which may be incorporated conveniently into this sensing diagram and, furthermore, addressed effectively by the designed algorithm. Simulations validate the proposed algorithm. While classical schemes fail to deal with doubly-selective channels, the new sensing scheme can exploit the underlying channel memory and operate well, which provides a great promise to realistic applications.**

***Index Terms*—Spectrum sensing, time-frequency doublely selective channels, dynamic state-space model, noise uncertainty**

I. Introduction

THE general interests on cognitive radios (CRs) are driven essentially by the scarcity of spectrum resources and increasing demands of wide frequency-bands [1]. In order to accomplish the opportunistic usage of allocated spectrum, the secondary users (SUs) are endowed with the ability of detecting or sensing the unoccupied band [2]. Based on the real-time awareness of current wireless environments, then CRs devices intelligently adapt its functionalities and dynamically access authorized bands of primary users (PUs) [3].

Manuscript received January 9, 2014; revised May 27, 2014 and August 27, 2014; accepted September 2, 2014.

Bin Li, Mengwei Sun and Chenglin Zhao are with the School of Information and Communication Engineering (SICE), Beijing University of Posts and Telecommunications (BUPT), Beijing, 100876 China. (Email: stonebupt@gmail.com).

Xiaofan Li is with is with the State Radio monitoring center Testing Center (SRTC), Beijing, China. (Email: lixiaofan@srtc.org.cn).

A. Nallanathan is with the Department of Informatics, King's College London, London, WC2R2LS, United Kingdom. (Email: nallanathan@ieee.org).

As a key ingredient to promote frequency utilization and mitigate interferences, CR has enjoyed high favor among commercial communications recently, including IEEE 802.22, the long term evolution (LTE) systems [4] and wireless local area networks (WLANs) (i.e. IEEE 802.11af) [5]. Given the emerging services, it takes little imagination to foresee that CRs will operate in various adverse environments. Firstly, attributed to the relative movement, wireless propagations will be characterized intrinsically by time-varying dynamics [6], i.e., the time selective fading may be encountered. Secondly, the frequency selective multipath propagations will be inevitable, which is aroused by rich reflectors in typical short-range applications as well as the good temporal resolution of CRs receiver with a broad reception bandwidth [7], [8]. What it comes down is that, in practice, the design of CRs should take the complex time-frequency double selective fading channels (TF-DSFCs) into full accounts [9].

One of the fundamental issues to be addressed in CRs is spectrum sensing, which aims to identify the working state (i.e. active or sleep) of PUs and, therefore, makes CR users ready for the opportunistic use of vacant license bands [10]–[12]. Traditional techniques for spectrum sensing may be distributed into three classes, i.e., cyclostationary detection (CD) [13], matched filtering detection (MFD) [14] and energy detector (ED) [15], which may have different advantages and requirements [10]–[12]. Despite a moderate sensing performance, ED excludes any *a priori* information of primary signals and is robust and simple, which is hence of great interest to CRs. Recent new techniques designed for spectrum sensing include wavelet analysis and compressive sensing [16]. Another sensing scheme is developed based on a covariance matrix [17], which exploits the statistical information (i.e. either temporal or spatial correlations) of primary signals [10]. In refs. [18] and [19], the probabilistic property of PUs' states is properly utilized to either design the sensing algorithm or optimize the sensing schedule strategy.

It is noteworthy that, however, most existing sensing methods are tailored to a *static* single-path channel model. That is, wireless propagations from PUs to SUs are assumed to be invariant and, simultaneously, the multipath components (MPCs) are also ignored for simplicity. More importantly, the widespread *memory* distributed in time-correlated channels has been ignored unfortunately, which could be fully exploited to promote the detection performance. Therefore, the sensing performance of these methods (i.e. ED) may become less competitive in TF-DSFCs. Ref. [15] investigated the sensing

performance under Rayleigh fading based on the statistical probability distribution function (PDF), which, however, can only characterize the *instantaneous* random behaviors of the single-path gain and may fail to track its time variations, let alone exploit the underlying channel dynamics. Besides, multipath fading will also contaminate the received signal dramatically and, therefore, results in substantial performance degradations. In our previous work, a spectrum sensing scheme is designed for the time-varying flat fading (TVFF) channel, nevertheless, only with a single-path [20]. For the encountered more complex TF-DSFCs, the development of efficient sensing schemes remains still as an unexplored area.

To the best of our knowledge, there are few works in the literature reported on the design of spectrum sensing schemes in the presence of TF-DSFCs. In this paper, we present a promising spectrum detection algorithm for realistic time-varying multipath flat-fading (TVMFF) propagations. A dynamic discrete-state model (DSM), involving two hidden states (i.e., the working state of PUs and the unknown multipath channel), is established firstly. Spectrum sensing is thereby realized by blindly estimating both the real-time TF-DSFCs response and unknown PUs states. Relying on a Bayesian statistical inference framework and premised on a simulated Monte-Carlo approach, an iterative algorithm is designed to estimate the non-stationary *a posteriori* probability. To sum up, the main contributions of this work are two-fold.

A. Dynamic State-space Model of Spectrum Sensing. A novel DSS is established, which will effectively characterize the spectrum sensing process over TF-DSFCs and may exploit the underlying memory property. In the new stochastic model, the PU state and unknown time-varying multipath gains are considered as two hidden states to be estimated. As both of them will evolve dynamically along time, a two-state Markov chain (TSMC) is used to model the dynamics of PU's states, while another finite-state Markov chain (FSMC) is used to depict the random transitions of multipath gains. As in most schemes [10]–[12], [15], [19], [20], the summed energy is adopted as an observation of the formulated stochastic DSM for the simplicity of implementations.

B. Joint Estimations based Spectrum Sensing. With this new DSS paradigm, the time-varying multipath gain will be estimated sequentially on reception of new observations. As far as the two-hypothesis based spectrum sensing procedure is concerned, nevertheless, a direct joint estimation may become impractical due to the coupling relationship between the multipath response and PU states as well as the intractable computation. For this reason, an iterative estimation scheme is further designed, with which the estimations of two hidden states are refined successively. The new sensing algorithm involves three steps, i.e., coarse detection, multipath coefficients estimation and PU state update. In order to obtain the recurrence propagation of time-varying posterior densities, the Monte-Carlo random finite sampling based particle filtering (PF) is suggested furthermore. Thus, the non-analytic posterior probability can be derived recursively by using all information available up to the time. Besides, a special estimation mechanic is integrated to calculate *unknown* noise variance. With the assistance of the recovered multipath gains and by exploiting the dynamic of time-dependent channels, spectrum sensing can be implemented effectively even in realistic TF-DSFCs. The new scheme, which can be extended to combat other realistic challenges (e.g., the noise uncertainty or non-quiet sensing), may provide a great promise for future CR applications.

The rest of the paper is structured as following. The stochastic DSM is established in Section II. On this basis, in Section III, a new spectrum sensing scheme is presented based on a joint estimation framework. An iterative estimation algorithm relying on SIS is designed. Simulation results and performance evaluations are provided in Section IV. Finally, we conclude the whole investigation in Section V.

The notations used in this work are summarized as following: the $n \times 1$ dimensional vector is denoted by $\mathbf{x}_{n\times 1}$, and the $M \times N$ matrix is $\mathbf{X}_{M\times N}$; $\mathbb{C}^{n_x\times 1}$ (or $\mathbb{R}^{n_x\times 1}$) is the complex-valued (or real-valued) space of n_x dimensions. $\|\mathbf{x}\|_2^2 = \sum_n x^2(n)$ is the l_2-norm of vector $\mathbf{x}$. $(\cdot)^T$ denotes the transpose and $(\cdot)^H$ is the conjugate transpose (or Hermitian); $x_{0:n} \triangleq \{x_0, x_1, x_2, \cdots, x_n\}$ represents the variable trajectory till the nth time index; $\otimes$ denotes the Hadamard multiplication, $\oplus$ accounts for a Boolean complementary operation between two binary variable; $\lfloor x \rfloor$ is the floor of x; $diag(\cdot)$ gives a block matrix having the arguments along its main diagonal; $\mathsf{E}(\cdot)$ denotes the ensemble average; $\mathsf{E}_{\mathcal{H},\mathcal{B}}(\cdot)$ is the statistical expectation on the joint set $(\mathcal{H}, \mathcal{B})$.

II. SYSTEM MODEL OF SPECTRUM SENSING

In this section, the time-varying dynamics of multipath fading channels are given full considerations, and a comprehensive DSM is formulated. A prevailing advantage of the new stochastic DSM is that, while traditional methods can only concentrate on *time-invariant* statistical distribution of the single-path fading channel, it may elegantly incorporate realistic unknown TF-DSFCs into the spectrum sensing.

A. PUs Working State

The periodic sensing strategy is adopted [21], i.e., a fixed frame duration T_F is assumed and the sensing duration is T_S. As suggested, for most wireless services, the evolution of primary states $s_n \in \mathcal{S} = \{0, 1\} \subset \mathbb{Z}$ (i.e. a set of integer) may be described by an alternating renewal process [22], [23]. The PDFs of both busy and idle states are assumed to follow the negative exponential distribution, i.e.

$$f_1(n) = \mu \times exp(-\mu n) \tag{1a}$$

$$f_0(n) = \lambda \times exp(-\lambda n) \tag{1b}$$

where μ and λ denote two transitional rates of busy-to-idle and idle-to-busy, respectively. n is the index of discrete time slots. With the help of *Komogorov* Equation, the probability of the idle state (i.e. s_n ="0") remaining unchanged during q successive slots, which is denoted by $p_{00}(q) \triangleq \Pr\big(s(n) = 0 | s(n-j) = 0\big)(1 \leqslant j < q)$, is given by

$$p_{00}(q) = \frac{\mu}{\mu+\lambda} + \frac{\lambda}{\mu+\lambda} exp\big[-q(\mu+\lambda)\big]. \tag{2}$$

Similarly, the probability of the active state lasting for q slots, i.e., $p_{11}(q) \triangleq \Pr(s(n)=1|s(n-j)=1)$ $(1 \leqslant j < q)$, may be written to

$$p_{11}(q) = \frac{\lambda}{\mu+\lambda} + \frac{\mu}{\mu+\lambda} exp\big[-q(\mu+\lambda)\big]. \tag{3}$$

Thus, if the busy state is identified initially at the time index n_0 with a probability of $\Pr(s_{n_0}=1)$, then the probability that this PU state stays always in busy within q sensing slots is obtained from

$$\begin{aligned}\Pr\left(s_{n_0+q}=1|s_{j\in[n_0+1,n_0+q-1]}=1\right) \\ = p_{11}(q) \times \Pr(s_{n_0}=1).\end{aligned} \tag{4}$$

Accordingly, the probability of PU state transiting into idle in the qth detection period, after staying in "1" for $(q-1)$ periods, can be expressed as $1-p_{11}(q) \times \Pr(s_{n_0}=1)$.

B. Time-varying Multipath Channel

In this work, we consider a realistic channel more common to future wireless applications, i.e., the TVMFF channel.

1) Frequency-selective fading: Owing to rich reflectors involved in the typical scenario (e.g. indoor) and the improved time resolution of receivers, multipath propagations are observed usually in broad band wireless communications [8]. The channel impulse response (CIR) is given by

$$h_n = \sum_{l=0}^{L-1} h_{n,l} exp(-j\theta_{n,l}) \delta\left(t - t_{n,l}\right), \tag{5}$$

where $\delta(\cdot)$ denotes the Direct function. $h_{n,l}$ is the amplitude of the lth path at the nth time index, $\theta_{n,l}$ denotes the MPC's phase which is assumed to be uniformly distributed, i.e., $\theta_{n,l} \sim \mathcal{U}[0, 2\pi]$; $t_{n,l}$ corresponds to the time of arrivals (ToAs). L accounts for the length of complex multipath channel. For simplicity, the multipath gain vector is denoted by $\mathbf{h}_n \triangleq [h_{n,0}\ h_{n,1}\ \cdots\ h_{n,L-1}]^T$.

The amplitudes of independent MPCs are assumed to follow *a priori* log-normal distribution. The distribution mean is $\mathbf{m}_h = [m_0\ m_1\ \cdots\ m_{L-1}]^T$ with $m_l \triangleq \ln(\mathbf{h}_n(l))$, and the variance is $\mathbb{V}\{\ln(\mathbf{h}_n(l))\} = \sigma_h^2$ $(l = 0, 1, \cdots, L-1)$, which are both considered to be static. Thus, the joint distribution of negative amplitudes (i.e. $p(\mathbf{h}_n)$) is given by

$$p(\mathbf{h}_n) = \prod_{l=1}^{L-1} \frac{1}{\sigma_h\sqrt{2\pi}} \exp\left\{-\left[\ln(\mathbf{h}_n(l)) - \mathbf{m}_h(l)\right]^2 / 2\sigma_h^2\right\}. \tag{6}$$

2) Time-selective fading: Further considering the relative movement in radio environments, the multipath response of realistic CR links is also a *time-correlated* random process. The time-varying dynamics of fading coefficients can be described by two classical models, i.e., the autoregressive (AR) model [24] and the Clarke's model. Rather, in this analysis, a FSMC model is considered alternatively [25], which is proven to be a good match to a Clarke's model and may also reflect the dynamic transitional property.

Based on the FSMC, the representative state of the lth fading gain at time index n (i.e. $\mathbf{h}_n(l)$), which serves as an output of one specific Markov chain, is denoted by $\mathcal{A}_{n,l} = \mathcal{R}_k$, $k \in \{0, 1, \cdots, K-1\}$, $\mathcal{R}_k \subseteq \mathbb{R}^1$. For the lth path, the state evolutions at time n are specified by a transitional probability matrix (TPM) $\mathbf{P}_{n,l} = \{P_{k_1\to k_2,n,l},\ k_1, k_2 \in 0, 1, \cdots, K-1\}$.

$$\mathbf{P}_{n,l} = \begin{bmatrix} P_{0\to0,n,l} & P_{0\to1,n,l} & \cdots & P_{0\to(K-1),n,l} \\ P_{1\to0,n,l} & P_{1\to1,n,l} & \cdots & P_{1\to(K-1),n,l} \\ \vdots & \vdots & \ddots & \vdots \\ P_{(K-1)\to0,n,l} & P_{(K-1)\to1,n,l} & \cdots & P_{(K-1)\to(K-1),n,l} \end{bmatrix}. \tag{7}$$

Here, $P_{k_1\to k_2,n,l}$ specifies the prior probability of the lth path transiting from state k_1 to state k_2, i.e.

$$P_{k_1\to k_2,n,l} \triangleq \Pr\left(\mathcal{A}_{n,l} = \mathcal{R}_{k_2} | \mathcal{A}_{n-1,l} = \mathcal{R}_{k_1}\right). \tag{8}$$

For convenience, the first-order FSMC model is considered which usually coincides with the popular statistical fading models (e.g., Clarke's model) [35]. Thus, the current state of fading channels is only associated with the previous state, while keeps statistically independent of all other past and future fading states. More specifically, we may now have $P_{k_1\to k_2,n,l} = 0$ for $|k_1 - k_2| > 1$. So, the TPM $\mathbf{P}_{n,l}$ may be further simplified to eq. (9), with its elements specified in Appendix 1.

3) TVMFF channel: Taking both the time-selectivity and frequency-selectivity into accounts, the resolvable MPCs are also assumed to evolve independently according to

$$\begin{aligned} p(\mathbf{h}_n|\mathbf{h}_{n-1}) &\triangleq \Pr\left(\mathbf{h}_n \to \mathcal{H}_n \middle| \mathbf{h}_{n-1} \to \mathcal{H}_{n-1}\right) \\ &= \prod_{l=0}^{L-1} \Pr\left(\mathcal{A}_{n,l} = \mathcal{R}_{k_2} \middle| \mathcal{A}_{n-1,l} = \mathcal{R}_{k_1}\right) \\ &= \prod_{l=0}^{L-1} P_{k_1\to k_2,n,l}, \end{aligned} \tag{10}$$

where $P_{k_1\to k_2,n,l}$ may vary with different time indexes n. Given the latest fading set $\mathcal{H}_{n-1} = \{\mathcal{A}_{n-1,l} = \mathcal{R}_{k_1}\}$, the feasible state set of time n, i.e. $\mathcal{H}_n$, is specified by

$$\mathcal{H}_n = \{\mathcal{A}_{n,l} = \mathcal{R}_{k_2}, |k_2 - k_1| < 2,\ l = 0, 1, \cdots, L-1\}.$$

For convenience, the dynamic of $\mathbf{h}_n$ is further assumed to be stationary [27], i.e., the probabilistic transitions are

$$\mathbf{P}_{n,l} = \begin{bmatrix} P_{0\to0,n,l} & P_{0\to1,n,l} & 0 & 0 & \cdots & 0 & 0 & 0 \\ P_{1\to0,n,l} & P_{1\to1,n,l} & P_{1\to2,n,l} & 0\cdots & 0 & 0 & 0 & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & P_{(K-2)\to(K-3),n,l} & P_{(K-2)\to(K-2),n,l} & P_{(K-2)\to(K-1),n,l} \\ 0 & 0 & 0 & 0 & \cdots & 0 & P_{(K-1)\to(K-2),n,l} & P_{(K-1)\to(K-1),n,l} \end{bmatrix}. \tag{9}$$

independent of n. So, we may further have

$$p(\mathbf{h}_n|\mathbf{h}_{n-1}) = \prod_{l=0}^{L-1} P_{k_1 \to k_2, l}. \quad (11)$$

C. Observation

For ease of implementations, in practice, ED has been widely recommended as a fundamental sensing technique [10]–[12]. This work will establish a general DSM of spectrum sensing based on the summed energy. Before proceeding, it is necessary to briefly illustrate ED in the situation of TF-DSFCs, which is formulated to the following two hypotheses

$$y_n = \begin{cases} \sum_{m=1}^{M} z_{n,m}^2, & H_0, \quad (12a) \\ \sum_{m=1}^{M} \left| \left[\mathbf{h}_n \otimes exp(-j\boldsymbol{\theta}_n)\right]^H \mathbf{b}_{n,m} + z_{n,m} \right|^2, & H_1, \quad (12b) \end{cases}$$

where $M = T_S \times f_s$ is the samples size. In order to alleviate the hardware complexity and also ensure the recovery of MPCs, f_s is assumed to the Nyquist-rate. $\otimes$ denotes the Hadamard multiplication between two vectors. H_0 and H_1 denote two hypotheses corresponding to the absence and presence of PU signals, respectively. y_n is the summed-energy in SU's devices, while $\mathbf{b}_{n,m} = [b_{n,m}\ b_{n,m-1}\ \cdots\ b_{n,m-L+1}]^T$ is the mth information vector of PUs in the nth slot. In practice, $\mathbf{b}_{n,m} \in \mathcal{B} \subseteq \mathbb{C}^{L\times 1}$ can be either the single-carrier (SC) format or orthogonal frequency division multiplexing (OFDM) signals. Without losing generality, $\{b_{n,m}\}$ is assumed to be OFDM modulated signal, which follows *a priori* Gaussian distribution, i.e., $b_{n,m} \sim \mathcal{N}(0, \sigma_b^2)$. The additive white Gaussian noise (AWGN) noise $z_{n,m}$ is zero-mean Gaussian distributed, with a variance of σ_z^2. It is shown that the noise variance σ_z^2 may usually become uncertain in practice [10].

D. DSM of Spectrum Sensing in TF-DSFCs

In order to characterize spectrum sensing with TF-DSFCs, a stochastic DSM is established as follows.

$$s_n = F(s_{n-1}) \quad (13)$$

$$\mathbf{h}_n = H(\mathbf{h}_{n-1}) \quad (14)$$

$$\mathbf{b}_{n,m} = T(\mathbf{b}_{n,m-1}) \quad (15)$$

$$y_n = G(\mathbf{h}_n, s_n, \mathbf{b}_{n,m}, z_{n,m}) \quad (16)$$

Here, (13)-(15) are referred to as the *dynamic equations* and (16) is the *measurement equation*. Two hidden states to be estimated, i.e. $s_{n,m} \in \mathcal{S}$ and $\mathbf{h}_n \in \mathcal{H}$, are evolved independently according to transitional functions $F(\cdot) : \mathbb{Z}^1 \to \mathbb{Z}^1$ and $H(\cdot) : \mathbb{C}^{L\times 1} \to \mathbb{C}^{L\times 1}$, respectively. For the independent information source, we have $\mathbf{b}_{n,m} = \mathbf{T}\mathbf{b}_{n,m-1} + \mathbf{u}_{n,m} \in \mathcal{B}$, where the elements of linear transitional matrix $\mathbf{T}_{L\times L}$ are all zeros expect for $diag(\mathbf{T}_{2:L,2:L}) = [1\ 1 \cdots\ 1]^T$, and $\mathbf{u}_{n,m} = [b_{n,m}\ 0\ \cdots\ 0]^T$. The nonlinear and noisy observations y_n are derived from the function $G(\cdot) : \mathbb{C}^{M\times 1} \to \mathbb{R}^1$ in eq. (16). Note that, in this DSM, one auxiliary variable $\mathbf{b}_{n,m}$ is used, which, however, will not be resolved by SUs.

For the convenience of analysis, two important aspects are assumed to the established DSM.

1) This work focuses on a slow-fading case. The multipath amplitude $\mathbf{h}_n$, therefore, is assumed to be invariant within several successive slots. The length of static sensing slots, where $\mathbf{h}_n$ remains unchanged, is denoted by N_f. In practice, the static period $N_f \times T_F$ is associated with the maximum Doppler frequency shift f_D. Specifically, we have $f_D = \frac{1}{N_f T_F}$, i.e., $N_f \propto 1/f_D$.
2) The dynamic transitions of multipath channels $\mathbf{h}_n$ occur possibly at the edge of each sensing slot $n' = \lfloor n/N_f \rfloor$ [20]. So, the TF-DSFCs response may be dealt as a constant within each sensing slot.

III. Spectrum Sensing in TF-DSFCs

Contaminated both by TF-DSFCs and additive noise, in practice, the received signals will show remarkable random fluctuations and CR devices may fail to identify the actual occupancy state of PU bands. As a consequence, most sensing schemes (e.g. ED), which average out the unfavorable random fading effects by using statistical distributions [15], will become less attractive in this situation.

To evaluate the sensing performance, the false alarm probability P_f is defined by $P(H_1|H_0)$, while the missed detection probability P_m is $P(H_0|H_1)$ and the detection probability is given by $1 - P_m$. For realistic applications, the false alarm and missing detection have different implications. A low P_f is meaningful to maintain the high spectral utilization, while the missed detection probability P_m measures the interferences from SUs to PUs which should be restricted. Thus, the main metric of practical designs, under a Neyman-Pearson criterion, is either to minimize the missed probability for a target false alarm probability, or minimize the false alarm probability for a target miss probability [10]–[12]. Based on a consideration of both the spectral utilization to unused bands and the potential interference to PUs, alternatively the total error detection probability is considered by this work [28], [29], i.e.

$$\Omega = P_m P(H_1) + P_f P(H_0). \quad (17)$$

Or equivalently, we will focus on the *total probability of detections* defined by $P_D := 1 - \Omega = 1 - p(H_1) - \frac{\mu}{\mu+\lambda} \times P_f + \frac{\lambda}{\mu+\lambda} \times P_d$ in the following analysis. Note that, this compound probability P_D is fundamentally different from the single detection probability P_d.

A. MAP Detection

In order to obtain joint estimations of unknown PU states and time-varying multipath channels, a Bayesian approach, i.e. maximum *a posteriori* (MAP) detection, is specially designed,

$$\left(\hat{\mathbf{h}}_n, \hat{s}_n\right) = \arg \max_{\mathbf{h}_n \in \mathcal{H}, s_n \in \mathcal{S}, \mathbf{b}_{n,m} \in \mathcal{B}} p(\mathbf{h}_n, s_n, |y_n) \big| y_n, p(\mathbf{h}_n|\mathbf{h}_{n-1}), p(s_n|s_{n-1}), p(\mathbf{b}_{n,m}|\mathbf{b}_{n,m-1}), \quad (18)$$

which will rely on the posterior probability $p(\mathbf{h}_n, s_n|y_n)$. Thus, the main objective of our new scheme can be expressed into eq. (18), where the *a posteriori* density is given by

$$\begin{aligned} p(\mathbf{h}_n, s_n|y_n) &\propto p(y_n|\mathbf{h}_n, s_n) \times p(\mathbf{h}_n, s_n) \\ &\overset{(a)}{=} \int\limits_{\mathbf{b}_{n,m}\in\mathcal{B}} p(y_n|\mathbf{h}_n, \mathbf{b}_{n,m}, s_n) p\left(\mathbf{b}_{n,m}|\mathbf{h}_n, s_n\right) d\mathbf{b}_{n,m} p(\mathbf{h}_n) p(s_n) \\ &\overset{(b)}{=} \int\limits_{\mathbf{b}_{n,m}\in\mathcal{B}} p(y_n|\mathbf{h}_n, \mathbf{b}_{n,m}, s_n) p\left(\mathbf{b}_{n,m}|s_n\right) d\mathbf{b}_{n,m} p(\mathbf{h}_n) p(s_n). \end{aligned} \tag{19}$$

Here, (a) holds for the two independent states $\mathbf{h}_n$ and s_n, and (b) holds because $\mathbf{b}_{n,m}$ is independent of channels. The conditional *a priori* distribution of $\mathbf{b}_{n,m}$ is given by

$$p\left(\mathbf{b}_{n,m}|s_n\right) = \begin{cases} \prod\limits_{m=0}^{L-1} \frac{1}{\sqrt{2\pi}\sigma_b} exp[-b_{n,m}^2/\sigma_b^2], s_n = 1, & \text{(20a)} \\ 0, \qquad\qquad\qquad\qquad s_n = 0. & \text{(20b)} \end{cases}$$

As in ED, the likelihood function, i.e. $p\left(y_n\middle|\mathbf{h}_n, s_n\right)$, then conditionally follows a central chi-square distribution with $2M$ degrees of freedom (DoF) under H_0, i.e.

$$p\left(y_n\middle|\mathbf{h}_n, s_n = 0\right) \sim \chi_{2M}^2. \tag{21}$$

And, in the case of H_1, we may have the following remarks.

Remark 1. Given TF-DSFCs response $\mathbf{h}_n$, the likelihood density is Gaussian distributed under H_1 (see Appendix 2).

$$\begin{aligned} p\left(y_n\middle|\mathbf{h}_n, s_n = 1\right) =& \mathsf{E}_{\mathbf{b}_{n,m}\in\mathcal{B}}\left\{p\left(y_n\middle|\mathbf{h}_n, \mathbf{b}_{n,m}, s_n = 1\right)\right\} \\ \sim& \mathcal{N}\left(\mu_n, \sigma_n^2\right), \end{aligned} \tag{22}$$

with the mean μ_n and variance σ_n^2 specified by

$$\mu_n = 2M \times \left[\mathsf{E}_{\mathcal{B}}\left\{\left|\mathbf{h}_n^H \mathbf{b}_{n,m}\right|^2 \middle| \mathbf{h}_n\right\} + \sigma_z^2\right], \tag{23}$$

and

$$\sigma_n^2 = 2M \times \left(\sigma_b^4 ||\mathbf{h}_n||_2^4 + \sigma_z^4 + 2\sigma_z^2\sigma_b^2 ||\mathbf{h}_n||_2^2\right), \tag{24}$$

respectively, where the expectation term is given by

$$\begin{aligned} &\mathsf{E}_{\mathcal{B}}\left\{\left|\mathbf{h}_n^H \mathbf{b}_{n,m}\right|^2 \middle| \mathbf{h}_n\right\} = \mathsf{E}_{\mathcal{B}}\left\{\left|\left[\mathbf{h}_n \otimes exp(j\boldsymbol{\theta}_n)\right]^H \mathbf{b}_{n,m}\right|^2 \middle| \mathbf{h}_n\right\} \\ &\triangleq \int_{\mathbf{b}_{n,m}\in\mathcal{B}} \left|\mathbf{h}_n^H \mathbf{b}_{n,m}\right|^2 p\left(\mathbf{b}_{n,m}|s_n = 1\right) d\mathbf{b}_{n,m} \\ &= \sigma_b^2 ||\mathbf{h}_n||_2^2. \end{aligned} \tag{25}$$

It should be noteworthy that, from eqs. (22)-(25), the MPCs gain vector $\mathbf{h}_n$ is of significance in practice, while the random phase (i.e. $\boldsymbol{\theta}_n$) may not matter much to *non-coherent* ED schemes. For realistic CR networks with the non-cooperative PU, in fact, the interference evaluation or capacity optimization will also be related primarily with the gain (or l_2-norm) of multipath channels.

With Remark 1, it is also seen that both the mean and variance are all associated with $\mathsf{E}_{\mathcal{B}}\left\{\left|\mathbf{h}_n^H \mathbf{b}_{n,m}\right|^2 \middle| \mathbf{h}_n\right\}$ (i.e. ${\sigma_b}^2||\mathbf{h}_n||_2^2$). The likelihood function, marginalized on the multipath channel $\mathbf{h}_n$, is thereby the weighted sum of a group Gaussian distributions [30], and the occurrence probability of each component distribution is specified by $p(\mathbf{h}_n \to \mathcal{H}_n)$.

Remark 2. For the formulated DSM, the marginal likelihood function under H_1, which is of importance to evaluate the overall sensing performance of TF-DSFCs, is a mixed Gaussian process, i.e.,

$$p(y_n\big|s_n = 1) \sim \sum_{i=1}^{N_h} p(\mathbf{h}_n \to \mathcal{H}_n) \times \mathcal{N}\left(\mu_i, \sigma_i^2\right), \tag{26}$$

with its total mean and variance given by

$$\mathsf{E}(y_n) = \sum_{i=1}^{N_h} \mu_i p(\mathbf{h}_n \to \mathcal{H}_n), \tag{27}$$

$$\mathsf{E}\left\{\left[y_n - \mathsf{E}(y_n)\right]^2\right\} = \sum_{i=1}^{N_h} p(\mathbf{h}_n \to \mathcal{H}_n) \left\{\left[\mathsf{E}(y_n) - \mu_i\right]^2 + \sigma_i^2\right\}, \tag{28}$$

respectively, where $N_h = K^L$ is the size of feasible states set. Notice that, $\mathsf{E}_{\mathcal{B}}\left\{\left|\mathbf{h}_n^H \mathbf{b}_{n,m}\right|^2 \middle| \mathbf{h}_n\right\}$ is also a random variable given different $\mathbf{h}_n$. With some manipulations, the first term in eq. (28) can be further expressed to

$$\begin{aligned} &\sum_{i=1}^{N_h} p(\mathbf{h}_n \to \mathcal{H}_n) \left[\mathsf{E}(y_n) - \mu_i\right]^2 \\ =& \sum_{i=1}^{N_h} p(\mathbf{h}_n \to \mathcal{H}_n) \\ &\qquad \times \left\{\mathsf{E}_{\mathcal{H},\mathcal{B}}\left\{\left|\mathbf{h}_n^H \mathbf{b}_{n,m}\right|^2\right\} - \mathsf{E}_{\mathcal{B}}\left\{\left|\mathbf{h}_n^H \mathbf{b}_{n,m}\right|^2 \middle| \mathbf{h}_n\right\}\right\}^2 \\ \triangleq& \mathsf{V}_{\mathcal{H}}\left\{\mathsf{E}_{\mathcal{B}}\left\{\left|\mathbf{h}_n^H \mathbf{b}_{n,m}\right|^2\right\}\right\}. \end{aligned} \tag{29}$$

B. Sequential MAP Detection

From a point of view of sequential inference, the estimations should be acquired based on the new observation and all historical information available to the time n. Thus, $p(\mathbf{h}_{0:n}, s_{0:n}|y_{0:n})$ is of special interest to spectrum sensing in TF-DSFCs. Here, $y_{0:n} \triangleq \{y_0, y_1, \cdots, y_n\}$ represents the observation trajectory till the nth sensing slot, $\mathbf{h}_{0:n}$ and $s_{0:n}$ denote two hidden states trajectories, respectively. Given the independent noise samples, then the update of $p(\mathbf{h}_{0:n}, s_{0:n}|y_{0:n})$ is given by eq. (30), where $p(\mathbf{h}_0)$ and $p(s_0)$ are assumed to be known as *a priori*.

Notice that, for the considered alternating renewal process, the transitional probabilities $p(s_n|s_{n-1})$ tend to be non-stationary or time-dependent. For clarity, it is further replaced by $p_{n,q}\left(s_n|s_j, s_{j\in[n-q+1,n-1]} = 1 \oplus s_{n-q}\right)$, which is related with the lasting intervals (i.e. q) of the current PUs state, where $1 \oplus s_{n-q}$ accounts for the complementary state of s_{n-q}. Given the flat-fading channel with a static period of $N_f T_F$, the posterior probability can be written to eq. (31).

In addition, the evolution of multipath channel $\mathbf{h}_n$ in eq. (31) remains also asynchronous with s_n. Thus, the involved two heterogeneous states, in collusion with the non-stationary property, further challenge the DSM-based joint estimations. More than this, for realistic CR networks, it is also desirable to evaluate the posterior probability and accomplish spectrum sensing in real-time on reception of new observations.

To this end, the recurrence estimation seems to be a promising way to tackle above difficulties, which estimates posterior densities sequentially by incorporating the innovation information of new observations. Following the Chapman-Kolmogorov equation and the 1st-order Markov process, $p(\mathbf{h}_{0:n}, s_{0:n}|y_{0:n})$ will be updated via eq. (32).

Attributed to the involved time-dependent dynamics of the specific DSM, and the intermittent disappearance of likelihoods (e.g. H_1) as well as the non-Gaussian marginalization in the denominator of eq. (32), unfortunately the analytical posterior density of interest cannot be obtained [20], [31]. Thus, the above recurrence procedure is only a conceptual solution of the Bayesian statistical inference and, ultimately, becomes intractable in realistic applications.

C. Particle Filtering

As a feasible approach that may effectively deal with realistic complex distributions (e.g., non-Gaussian and non-stationary situations), the SIS inspired PF may be of promise to spectrum sensing in TF-DSFCs.

Relying on a numerical mechanic, SIS obtains the consistent estimation of *a posteriori* probability via a group of discrete random measures (i.e. particles) $\mathbf{x}^{(i)}$ with different probability masses (or weights) $w^{(i)}$ $(i = 1, 2, \cdots, I)$, where I is the particles size [32], [33]. For clarity, $\mathbf{x}$ represents the hidden states to be estimated here. The particles are simulated samples and drawn from an unknown state space which is associated with the target probability $p(\mathbf{x})$. Thus, the target distribution is approximated by

$$\tilde{p}(\mathbf{x}) = \sum_{i=1}^{I} w^{(i)} \delta\left(\mathbf{x} - \mathbf{x}^{(i)}\right), \tag{33}$$

It is shown from eq. (33) that the numerically approximated distribution $\tilde{p}(\mathbf{x})$ may converge to $p(\mathbf{x})$ when the particle size I is large. Note that, it is infeasible to sample directly from $p(\mathbf{x})$. So, a proposal distribution (or importance function) $\pi(\mathbf{x})$ is designed from which discrete particles are drawn, i.e., $\mathbf{x}^{(i)} \sim \pi(\mathbf{x})$. Thus, the importance weights $w^{(i)}$ are determined by

$$w^{(i)} = p\left(y_{0:n}|\mathbf{x}_{0:n}^{(i)}\right) p\left(\mathbf{x}_{0:n}^{(i)}\right) \Big/ \pi\left(\mathbf{x}_{0:n}^{(i)}|y_{0:n}\right). \tag{34}$$

As an approximation of the realistic probability distribution, the importance weights should be normalized to 1, i.e., $w^{*(i)} = w^{(i)} \big/ \sum_{i=1}^{I} w^{(i)}$. Usually, $\pi(\mathbf{x})$ may be factored as

$$\pi(\mathbf{x}_{0:n}|y_{0:n}) = \pi(\mathbf{x}_n|\mathbf{x}_{0:n-1}, y_{0:n})\pi(\mathbf{x}_{0:n-1}|y_{0:n-1}). \tag{35}$$

With the particles $\mathbf{x}_k^{(i)} (i = 1, 2, \cdots, I)$ sequentially sampled from $\pi\left(\mathbf{x}_n|\mathbf{x}_{0:n-1}^{(i)}, y_{0:n}\right)$ as each new observation arrives, $w_k^{(i)}$ may be updated by

$$w_k^{(i)} \propto \frac{p\left(y_n|\mathbf{x}_n^{(i)}\right) p\left(\mathbf{x}_n^{(i)}|\mathbf{x}_{n-1}^{(i)}\right)}{\pi\left(\mathbf{x}_n^{(i)}|\mathbf{x}_{0:n-1}^{(i)}, y_{0:n}\right)} w_{n-1}^{(i)}. \tag{36}$$

The importance function $\pi\left(\mathbf{x}_n|\mathbf{x}_{0:n-1}^{(i)}, y_{0:n}\right)$ has a significant impact on estimation performances, which should be carefully designed in accordance with realistic situations. In the analysis, the optimal importance function is adopted [31], which is specified by $p\left(\mathbf{x}_n|\mathbf{x}_{0:n-1}^{(i)}, y_{0:n}\right)$ and may minimize the one-step variance of importance weights. Accordingly, the importance weights $w^{(i)}$ are recursively updated by

$$w_n^{(i)} = p\left(y_n|\mathbf{x}_{n-1}^{(i)}\right) \times w_{n-1}^{(i)}. \tag{37}$$

To sum up, two steps are involved in PF. (1) Draw the discrete measures by sampling the importance distribution, i.e., $\mathbf{x}_n^{(i)} \sim \pi\left(\mathbf{x}_n|\mathbf{x}_{0:n-1}^{(i)}, y_{0:n}\right)$; and (2) update the weights $w_n^{(i)}$ by using (37). One can refer to refs. [31]–[33] for details.

D. Iterative Estimation Based Spectrum Sensing

In contrast to classical joint detections of wireless communications, note that, there are two formidable challenges in this considered problem. (1) The fading channel $\mathbf{h}_n$, unfortunately, will disappear completely from observations y_n when $s_n = 0$, while in classical situations the multipath response may be always enclosed in received signals. (2) As observed from eq. (32), the direct recurrence estimation of both multipath fading gains and unknown PU states is analytically intractable and computationally prohibitive, which may involve at least $2^L \times K^L$ dimensional non-Gaussian marginal integrations.

To cope with above problems, we design an effective sensing algorithm for more complex TF-DSFCs. A recursive scheme is developed, which iteratively estimates the multipath response and PU states. The proposed algorithm mainly contains three steps, i.e. (a) coarse detection; (b) MAP estimation of channel gains and (c) PF-based PU state estimation. Note that, (b) and (c) will be iteratively performed to further refine the estimations.

$$p(\mathbf{h}_{0:n}, s_{0:n}|y_{0:n}) \propto p(\mathbf{h}_0)p(\mathbf{b}_{0,0})p(s_0) \times \prod_{i=1}^{n} \underbrace{\int_{\mathbf{b}_{i,m}\in\mathcal{B}} p(y_i|\mathbf{h}_i, \mathbf{b}_{i,m}, s_i)p(\mathbf{b}_{i,m}|s_n)d\mathbf{b}_{i,m}}_{p(\mathbf{h}_{0:i}, s_{0:i}|y_{0:i})} p\left(\mathbf{h}_i|\mathbf{h}_{i-1}\right) p\left(s_i|s_{i-1}\right). \tag{30}$$

$$p(\mathbf{h}_{0:n}, s_{0:n}|y_{0:n}) \propto p(\mathbf{h}_0)p(s_0) \times \prod_{i=1}^{n} p(\mathbf{h}_{0:i}, s_{0:i}|y_{0:i}) p\left(\mathbf{h}_{\lfloor i/N_f \rfloor}|\mathbf{h}_{\lfloor i/N_f \rfloor - 1}\right) p_{i,q}\left(s_i|s_{i-1}, s_{j\in[i-q+1,i-1]} = 1 \oplus s_{i-q}\right). \tag{31}$$

$$p(\mathbf{h}_{0:n}, s_{0:n}|y_{0:n}) = p(\mathbf{h}_{0:n-1}, s_{0:n-1}|y_{0:n-1}) \times \frac{p(y_n|\mathbf{h}_n, s_n)p(\mathbf{h}_n|\mathbf{h}_{n-1})p(s_n|s_{n-1})}{\iint_{\mathbf{h}_n\in\mathcal{H}, s_n\in\mathcal{S}} p(y_n|\mathbf{h}_n, s_n)p(\mathbf{h}_n, s_n|y_{0:n-1})d\mathbf{h}_n ds_n}. \tag{32}$$

1. Coarse Detection

This procedure is designed to obtain a rough estimation of unknown PU states and, on this basis, subsequent estimation methodologies of $\mathbf{h}_{n'}$ in accordance with different PU states may be determined. To accomplish this, the MAP criterion would be applied to derive the initial estimation of PU states, which is denoted by i_n, i.e.

$$i_n = \arg\max_{s_n \in \mathcal{S}} p\left(y_n | s_n, \hat{\mathbf{h}}_{0:n-1}\right) \times p(s_n | \hat{s}_{0:n-1}). \quad (38)$$

In practice, the noise uncertainty (NU) may have bad impacts on the sensing performance. So, a special mechanic is further integrated to estimate the unknown noise variance recursively. In order to obtain likelihood densities, firstly the estimated noise variance of time $(n-2)$ may be used. Thus, the likelihoods, conditioned on $\hat{\mathbf{h}}_{0:n-1}$, are given by

$$p\left(y_n | s_n, \hat{\mathbf{h}}_{0:n-1}\right) = \begin{cases} \mathcal{N}\left\{M\hat{\sigma}_z^2(n-2), 4M\hat{\sigma}_z^4(n-2)\right\}, & H_0, \quad (39a) \\ \mathcal{N}\left\{\mathsf{E}_{\mathcal{H},\mathcal{B}}\left(\left|\mathbf{h}_{n'}^H \mathbf{b}_{n,m}\right|^2 \middle| \hat{\mathbf{h}}_{n'-1}\right), \hat{\sigma}_n^2\right\}, & H_1, \quad (39b) \end{cases}$$

where the variance term is given by

$$\hat{\sigma}_n^2 = 2M \times \left[\sigma_b^4 ||\hat{\mathbf{h}}_{n'-1}||_2^4 + \hat{\sigma}_z^4(n-2) + 2\hat{\sigma}_z^2(n-2)\sigma_b^2 ||\hat{\mathbf{h}}_{n'-1}||_2^2\right], \quad (40)$$

and the conditional expectation of receiving power $\mathsf{E}_{\mathcal{H},\mathcal{B}}\left\{\left|\mathbf{h}_{n'}^H \mathbf{b}_{n,m}\right|^2 \middle| \hat{\mathbf{h}}_{n'-1}\right\}$ is specified by

$$\begin{aligned} &\mathsf{E}_{\mathcal{H},\mathcal{B}}\left\{\left|\mathbf{h}_n^H \mathbf{b}_{n,m}\right|^2 \middle| \hat{\mathbf{h}}_{n-1}\right\} \\ &\triangleq \iint_{\mathbf{h}_n \in \mathcal{H}, \mathbf{b}_{n,m} \in \mathcal{B}} \left|\hat{\mathbf{h}}_n^H \mathbf{b}_{n,m}\right|^2 p\left(\mathbf{h}_n | \hat{\mathbf{h}}_{n-1}\right) p\left(\mathbf{b}_{n,m}\right) d\mathbf{b}_{n,m} d\mathbf{b}_n \\ &\overset{(a)}{=} \sum_{\mathbf{h}_n \in \hat{\mathcal{H}}_n} p\left(\mathbf{h}_n | \hat{\mathbf{h}}_{n-1}\right) \left\{\sum_{l=0}^{L} \mathsf{E}_{\mathcal{B}}\left\{h_{n,l}^2 b_{n,m-l}^2\right\} + \right. \\ &\left. \sum_{l_1=0}^{L} \sum_{l_2=0, l_1 \neq l_2}^{L} h_{n,l_1} h_{n,l_2} \mathsf{E}_{\mathcal{B}}\left\{b_{n,m-l_1} b_{n,m-l_2}\right\}\right\} \\ &= \sum\nolimits_{\mathbf{h}_n \in \hat{\mathcal{H}}_n} p\left(\mathbf{h}_n | \hat{\mathbf{h}}_{n-1}\right) \parallel \mathbf{h}_n \parallel_2^2 \sigma_b^2, \end{aligned} \quad (41)$$

where $\hat{\mathcal{H}}_{n'}$ refers to the feasible set of $\mathbf{h}_{n'}$ at the time $n' = \lfloor n/N_f \rfloor$, which is predicted from the estimated multipath gain (i.e. $\hat{\mathbf{h}}_{n'-1}$). In (41), (a) holds due to the fact $\mathbf{h}_{n'}$ and $\mathbf{b}_{n,m}$ are two independent terms.

Secondly, the noise variance of time $(n-1)$ will be estimated. As suggested, the unknown noise variance may practically follows *a priori* Inverse-Gamma distribution [34], [35], i.e., $\sigma_z^2 \sim \mathcal{IG}(\alpha, , \beta)$. Based on a Bayesian criterion, the estimation of noise variance, i.e. $\hat{\sigma}_z^2(n-1)$, would be recursively derived by (see Appendix 3)

$$\hat{\sigma}_z^2(n-1) = \beta_{n-1}/(\alpha_{n-1} - 1), \quad (42)$$

where the numerator β_{n-1} and the denominator term α_{n-1}, conditioned on different estimated PU states $\hat{s}_{n-1}$, will be updated recursively by

$$\beta_{n-1} = \begin{cases} \beta_{n-2}, & \hat{s}_{n-1} = 0, \quad (43a) \\ \beta_{n-2} + y_{n-1}/2, & \hat{s}_{n-1} = 1, \quad (43b) \end{cases}$$

and

$$\alpha_{n-1} = \begin{cases} \alpha_{n-2}, & \hat{s}_{n-1} = 0, \quad (44a) \\ \alpha_{n-2} + M/2, & \hat{s}_{n-1} = 1. \quad (44b) \end{cases}$$

2. Estimation of Fading Gain

As far as the multipath fading is concerned, the estimation methodology will rely on different initial estimations of s_n [29], [36]. First, the MAP estimation of $\mathbf{h}_n$ may become infeasible in H_0 (i.e. $\hat{s}_n = 0$), due to the unavailability of the related likelihood information. Second, for the slow varying channels with a static response in N_f successive sensing slots, the estimated fading channel can be further modified sequentially in H_1 (i.e. $\hat{s}_n = 1$).

(1) If we have $i_n = 0$, expect for *a priori* state transitional probabilities, there is little innovation information can be utilized in the sensing slot where the fading channel may change (i.e. $n' = \lfloor n/N_f \rfloor$). In such a case, the estimation of $\mathbf{h}_{n'}$ is directly derived by

$$\hat{\mathbf{h}}_{n'} = \arg\max_{\mathbf{h}_{n'} \in \hat{\mathcal{H}}_n} p\big(\mathbf{h}_{n'} | \hat{\mathbf{h}}_{n'-1}\big). \quad (45)$$

Then, in subsequent sensing slots $n = (n' + J)$ $(1 \leq J < N_f)$, the estimated multipath response will remain unchanged, i.e. $\hat{\mathbf{h}}_{n'+J} = \hat{\mathbf{h}}_{n'}$.

(2) If we have $i_n = 1$, in the sensing slot $n' = \lfloor n/N_f \rfloor$ where the multipath channel may evolve, the observations and the related likelihood function can be fully exploited. Thus, the posterior estimation of $\mathbf{h}_{n'}$ is obtained from

$$\begin{aligned} \hat{\mathbf{h}}_{n'} &= \arg\max_{\mathbf{h}_{n'} \in \hat{\mathcal{H}}_n} p\big(\mathbf{h}_{n'} | y_{n'}, i_{n'} = 1, \hat{\mathbf{h}}_{n'-1}\big) \\ &= \arg\max_{\mathbf{h}_{n'} \in \hat{\mathcal{H}}_n} p\big(y_{n'} | \mathbf{h}_{n'}, i_{n'} = 1\big) p\big(\mathbf{h}_{n'} | \hat{\mathbf{h}}_{n'-1}\big). \end{aligned} \quad (46)$$

Then, in subsequent sensing slots with $n = (n' + J)$ $(1 \leq J < N_f)$, we may entirely utilize those past observations to further refine the estimation of $\mathbf{h}_{n'}$. More specifically, during the $(n' + J)$th subsequent slot, with the historic information $y_{n'+J}$ we may redefine the accumulated observation via

$$\begin{aligned} y_{n'+J} = &\sum_{j_n=1}^{J-1} y_{n'+j_n} \\ &+ \sum_{m=1}^{M} \left|\left[\mathbf{h}_n \otimes exp(-j\boldsymbol{\theta}_n)\right]^H \mathbf{b}_{n,m} + z_{n,m}\right|^2. \end{aligned} \quad (47)$$

Correspondingly, the accumulated observation $y_{n'+J}$ still follows a Gaussian distribution, nevertheless, with a new mean $J\mu_n$ and variance $J\sigma_n^2$, i.e.,

$$p\big(y_{n'+J} | \mathbf{h}_{n'}, i_{n'} = 1\big) \sim \mathcal{N}\left(J\mu_n, J\sigma_n^2\right). \quad (48)$$

Notice from eq. (48) that, with more information from observations being exploited, it is expected the estimation of fading channel will become more accurate.

3. Particle Filtering based PU State Estimation

Although an initial estimation of PU states (i.e. i_n) is obtained by exploiting the 1-step prediction of the estimated multipath response, such an estimation ignores the dynamic evolutions of multipath fading and, therefore, may become inaccurate. Thus, this estimated PU state, which is treated as a *soft-output*, will be employed to further refine multipath coefficients iteratively. After R iterations, the sensing result will be derived finally.

Relying on an SIS framework, the posterior probability is derived numerically via a group of discrete particles $\{x_{0:n}^{(i)}\}$ and associative weights $\{w_{0:n}^{(i)}\}$. Thus, a more accurate estimation of s_n may be obtained by maximizing the approximated posterior density. With the sequential Bayesian mechanic, the innovation information carried by new observations will be incorporated with historical knowledge.

In practice, given the new observation and the estimation of multipath fading gains $\hat{\mathbf{h}}_{n'}$, we may then derive new particles from the importance distribution. Here, the optimal importance distribution is specified by

$$\begin{aligned}&\pi(x_n|x_{0:n-1}^{(i)},\hat{\mathbf{h}}_{n'},y_n) \triangleq p\left(x_n\middle|x_{0:n-1}^{(i)},\hat{\mathbf{h}}_{n'},y_n\right)\\&= p\left(y_n\middle|\hat{\mathbf{h}}_{n'},x_n\right)\\&\quad\times p_{n,q}\left(x_n\middle|x_{n-j}^{(i)},x_{j\in[n-q+1,n-1]}=1\oplus x_{n-q}\right).\end{aligned}\tag{49}$$

After sampling from the importance distribution, the associative weights of these random measures, which correspond to the probability pass, will be updated by

$$w_n^{(i)} = w_{n-1}^{(i)} \times p\left(y_n\middle|x_{n-1}^{(i)},\hat{\mathbf{h}}_n\right). \tag{50}$$

The likelihood densities, conditioned on estimated multipath gains and new particles, are then given by eq. (51) and (52), respectively, where $\Gamma(x)$ denotes the gamma function.

With the measures randomly sampled from a relevant distribution, thus the estimation of s_n is finally derived by eq. (53), which is a MAP estimation when $I \to \infty$.

$$\hat{s}_n = \arg\max_{s_n\in\mathcal{S}} p\big(s_n|y_{0:n},s_{0:n-1},\hat{\mathbf{h}}_{n'}\big), \tag{53}$$

where the *a posteriori* probability are approximated numerically by eqs. (54) and (55).

$$\begin{aligned}&p\left(s_n=0\middle|y_{0:n},s_{0:n-1},\hat{\mathbf{h}}_{n'}\right)\\&\simeq\int_{\mathbf{b}_{n,m}\in\mathcal{B}} p\left(s_n=0\middle|y_{0:n},s_{0:n-1}^{(i)},\hat{\mathbf{h}}_{n'},\mathbf{b}_{n,m}\right)d\mathbf{b}_{n,m}\\&=\frac{\sum_{i\in\mathcal{X}_0} w_n^{(i)}}{\sum_{i=1}^{I} w_n^{(i)}},\quad \mathcal{X}_0=\left\{i\middle|x_n^{(i)}=0\right\},\end{aligned}\tag{54}$$

$$\begin{aligned}&p\left(s_n=1\middle|y_{0:n},s_{0:n-1},\hat{\alpha}_{n'}\right)\\&\simeq\int_{\mathbf{b}_{n,m}\in\mathcal{B}} p\left(s_n=1\middle|y_{0:n},s_{0:n-1}^{(i)},\hat{\mathbf{h}}_{n'},\mathbf{b}_{n,m}\right)d\mathbf{b}_{n,m}\\&=\frac{\sum_{i\in\mathcal{X}_1} w_n^{(i)}}{\sum_{i=1}^{I} w_n^{(i)}},\quad \mathcal{X}_1=\left\{i\middle|x_n^{(i)}=1\right\}.\end{aligned}\tag{55}$$

E. Implementations

Based on the above elaborations, the schematic flow of the proposed algorithm is illustrated by Fig. 1. In realistic implementations, it is worth noting that there contains two counters in the proposed algorithm. (1) The first counter J is involved in the second phase (i.e. the estimation of fading channel), which is reset to 0 every time the channel state transition occurs. (2) The second counter q is used in the third phase (i.e. the estimation of PUs states). If we have $\hat{s}_n = \hat{s}_{n-1}$, we may set $q = q+1$; and otherwise we reset $q = 0$.

F. Complexity

For the presented algorithm, the complexity comes essentially from three parts. In the 1st phase, the number of multiplications in obtaining the observation and the rough threshold is about $\mathcal{O}(M + L \times K_1^L)$. Note that, from eqs. (42)-(44), the complexity of updating unknown noise variance may be basically ignored. In the 2nd phase, there are K_1^L feasible transitions of dynamic multipath response, and for each candidate, $\mathcal{O}(L)$ multiplications are required in calculating the related likelihoods. Correspondingly the number of

$$\begin{aligned}&p\left(y_n|\hat{\mathbf{h}}_n,x_{n-1}^{(i)}=0\right)\propto p\left(y_n|x_n=0\right)p\left(x_n=0\middle|x_{n-1}^{(i)}=0\right)+p\left(y_n|x_n=1,\hat{\mathbf{h}}_n\right)p\left(x_n=1\middle|x_{n-1}^{(i)}=0\right)\\&=\frac{1}{2^M\Gamma(M)}y_n^{M-1}exp\left(-y_n/2\right)\times p_{00}(q)\\&\quad+\frac{1}{2\sqrt{\pi M\left(\sigma_b^4||\hat{\mathbf{h}}_n||_2^4+\sigma_z^4+2\sigma_z^2\sigma_b^2||\hat{\mathbf{h}}_n||_2^2\right)}}exp\left\{\frac{-\left|y_n-2M\times\left(\|\hat{\mathbf{h}}_n\|_2^2\,\sigma_b^2+\sigma_z^2\right)\right|^2}{2M\left(\sigma_b^4||\hat{\mathbf{h}}_n||_2^4+\sigma_z^4+2\sigma_z^2\sigma_b^2||\hat{\mathbf{h}}_n||_2^2\right)}\right\}\times\big[1-p_{00}(q)\big],\end{aligned}\tag{51}$$

$$\begin{aligned}&p\left(y_n|\hat{\mathbf{h}}_n,x_{n-1}^{(i)}=1\right)\propto p\left(y_n|x_n=1,\hat{\mathbf{h}}_n\right)p\left(x_n=1|x_{n-1}^{(i)}=1\right)+p\left(y_n|x_n=0\right)p\left(x_n=0|x_{n-1}^{(i)}=1\right)\\&=\frac{1}{2\sqrt{\pi M\left(\sigma_b^4||\hat{\mathbf{h}}_n||_2^4+\sigma_z^4+2\sigma_z^2\sigma_b^2||\hat{\mathbf{h}}_n||_2^2\right)}}exp\left\{\frac{-\left|y_n-2M\times\left(\|\hat{\mathbf{h}}_n\|_2^2\,\sigma_b^2+\sigma_z^2\right)\right|^2}{2M\left(\sigma_b^4||\hat{\mathbf{h}}_n||_2^4+\sigma_z^4+2\sigma_z^2\sigma_b^2||\hat{\mathbf{h}}_n||_2^2\right)}\right\}\times p_{11}(q)\\&\quad+\frac{1}{2^M\Gamma(M)}y_n^{M-1}exp\left(-y_n/2\right)\times\big[1-p_{11}(q)\big].\end{aligned}\tag{52}$$

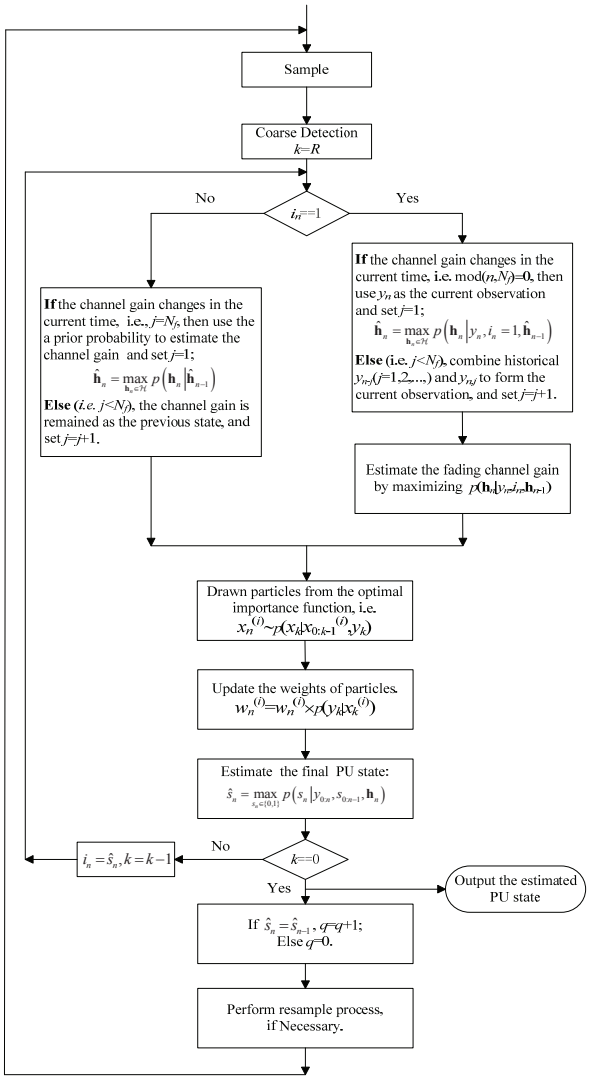


Fig. 1. Algorithm flow of the proposed spectrum sensing algorithm. For clarity, the estimation process of unknown noise is not shown.

multiplications are $\mathcal{O}(K_1^L \times L)$. Finally, the complexity of SIS process is proportional to $\mathcal{O}(I)$. Further given the R iterations, the total complexity can be roughly measured by $\mathcal{O}\{M + K_1^L L + R \times (K_1^L L + I)\}$. In practice, we may have $L = 3$ and, for the 1st order FSMC, $K_1 = 3$.

IV. Numerical Simulations and Performance Evaluations

Based on numerical simulations, in this section the proposed sensing algorithm will be evaluated in realistic TF-DSFCs. For convenience, in the analysis the memory length of multipath channel is set to $L = 3$, the number of discrete-states is $K = 8$. The variance of log-normal distribution is $\sigma_h^2 = 0.1$. The iterative time of implementing the new scheme is configured to 3, which is practically considered to be sufficient in the specific application. We consider three kinds of multipath channels,

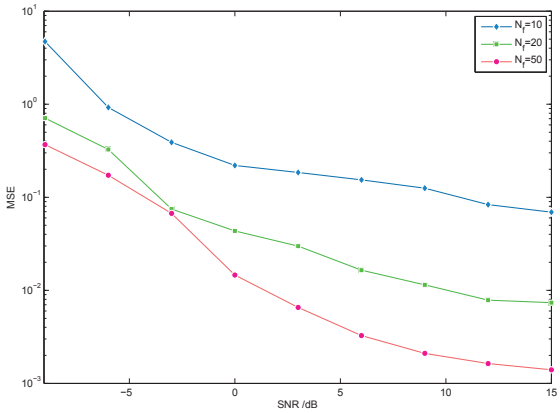


Fig. 2. MSE performance under different static length N_f. The sample size is M=100 and the Type-2 channel is used.

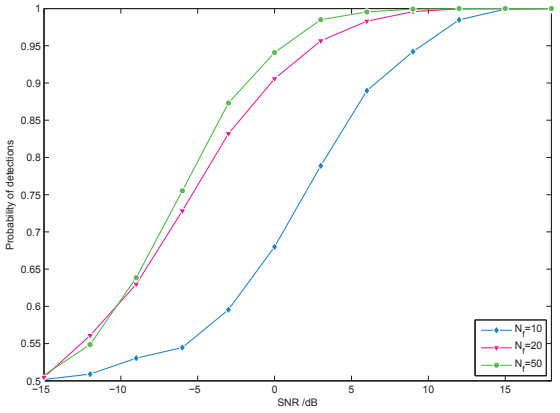


Fig. 3. Detection performance under different static length N_f. The static length is M=100 and the Type-2 channel is used.

with the mean-energy vectors $\mathsf{E}\{\mathbf{h}\}$ specified by eq. (56).

$$\text{Type-1}: \mathsf{E}\{|\mathbf{h}|\} = [0.2\ 0.5\ 1]^T, \tag{56a}$$

$$\text{Type-2}: \mathsf{E}\{|\mathbf{h}|\} = [0.05\ 0.2\ 1]^T, \tag{56b}$$

$$\text{Type-3}: \mathsf{E}\{|\mathbf{h}|\} = [0.0001\ 0.01\ 1]^T. \tag{56c}$$

As suggested, even for a blind case of distributed CRs, the time-varying multipath channel would be jointly estimated. In the simulations, we will also investigate the mean-square error (MSE) performance of estimated multipath channel $\mathbf{h}_n$, i.e.

$$MSE \triangleq \mathsf{E}\left\{\frac{1}{N}\sum_{n=1}^{N}\frac{\|\hat{\mathbf{h}}_n - \mathbf{h}_n\|_2^2}{\|\mathbf{h}_n\|_2^2}\right\} \times 100\%. \tag{57}$$

A. Static Length

The MSE performance of TF-DSFCs estimations, with various the maximum Doppler frequency f_D, is shown by Fig. 2. In this experimental simulation, the sample size is configured to $M = 100$ and the Type-2 response is used. For realistic TF-DSFCs, the maximum Doppler frequency shift is relatively large and, accordingly, three typical configurations are adopted in the analysis, i.e. $N_f = 10$, 20 and 50 (corresponding to $f_D T_F = 0.1,\ 0.05,\ 0.02$). We may observe from Fig. 2

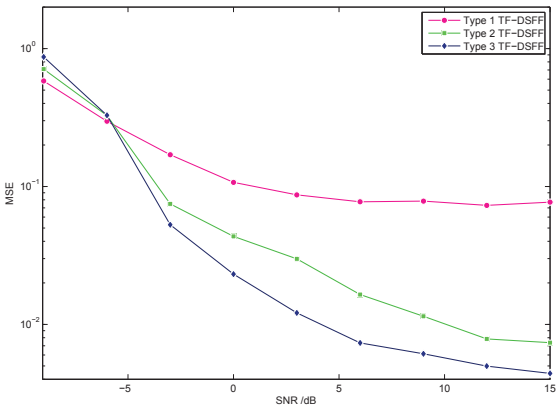


Fig. 4. MSE performance under different type of TF-DSFCs. The sample size is M=100, the static length is N_f=20.

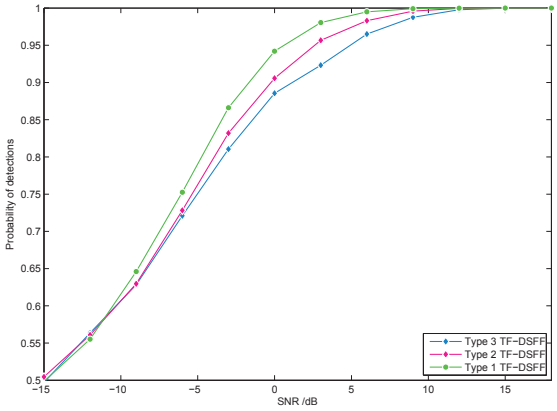


Fig. 5. Detection performance under different type TF-DSFCs. The sample size is M=100, the static length is N_f=20.

that the static length N_f may affect the MSE performance significantly. Taking the SNR=15dB for example, the MSE is about 1.14×10^{-3} when N_f=50, and it will increase to 6.92×10^{-2} when N_f=10. This is mainly attributed to the increment estimation of multipath gains by thoroughly exploiting the accumulated observations. To be specific, with a larger N_f, the multipath channel will remain unchanged in more sensing slots and, therefore, more historical observations can be accumulated to refine the estimations.

Much similar to MSE curves of different static length N_f, the better detection performance may be achieved by a larger N_f as in Fig. 3. When the detection probability P_D is 0.9, a detection gain of 6.5dB can be obtained if the static length N_f is increased from 10 to 20. When N_f is further increased from 20 to 50, however the achieved gain is only 1.7dB. That is, with a large N_f accompanying the more accumulated information, the estimation of time-varying multipath response may become sufficiently accurate. Accordingly, the improvement on detection performances may become limited for a large N_f ($N_f > 20$).

B. Multipath Channels

The MSE performance of TF-DSFCs estimations under various time-varying responses has been shown by Fig. 4. In the experiment, we configure the sample size to $M = 100$ and the static length $N_f = 20$. It is seen that the MSE of channel Type-3 is much less than that of Type-1. Taking SNR=15dB for example, the MSE is 4.41×10^{-3} for Type-3 and 7.70×10^{-2} for Type-1. In fact, the channel of Type-3 may be approximately treated as a *single-path* channel and, accordingly, its frequency selectivity is not obvious. So, it is expected Type-3 would achieve a better MSE, i.e., the single-path channel is usually easier to estimate compared with the multipath case due to inter-symbol interferences.

It is seen from Fig. 5 that, despite a worse MSE performance, the Type-3 channel may achieve the better detection performance compared with Type-1. It should be noteworthy that, as far as the detection performance of unknown PUs is concerned, the marginal likelihood function, i.e., $p(y_n|s_n = 1)$, is of great significance on the top of the estimation accuracy of TF-DSFCs, which is related with the multipath property characterized by a mixed Gaussian distribution as in eqs. (27)-(30). It is seen that the total variance $\mathsf{V}_{\mathcal{H}}\left\{E_{\mathcal{B}}\left\{\left|\mathbf{h}_n^H\mathbf{b}_{n,m}\right|^2\right\}\right\}$ may vary with different multipath gains (e.g. different $\mathsf{E}\{|\mathbf{h}|\}$). For the realistic discrete TF-DSFC, a larger $\mathsf{V}_{\mathcal{H}}\left\{E_{\mathcal{B}}\left\{\left|\mathbf{h}_n^H\mathbf{b}_{n,m}\right|^2\right\}\right\}$ comes also with a wider likelihood distribution $p(y_n|s_n = 1)$. For the considered three TF-DSFCs, the corresponding variance terms are evaluated and listed by eq. (59). Thus, the total detection probability of Type-1 is supposed to be better than Type-2, while Type-2 is superior to Type-3. The above analysis has also been verified by numerical results shown in Fig. 5.

C. Different Modulation Signals

The detection performance of various modulated PU signals with unknown constellations are plotted by Fig. 6. In this analysis, the Type-2 channel is adopted, we configure $N_f = 20$ and $M = 100$. The sensing performance is evaluated in the context of both OFDM and SC modulated signals. Fron numerical results, the detection probability of OFDM-QPSK signals is relatively inferior to that of SC-QPSK signals when the same modulation scheme are used (i.e. QPSK).

Even for SC signals, different modulation schemes have different performance. To be specific, the detection probability of SC-QPSK signals is superior to SC-BPSK signals. Given the same PU signal variance and SNR, it is shown from eq. (12) that the DoF of SC-QPSK signals is $2M$, while the DoF of SC-BPSK signals is only M. In other words, even with the same sample size $M = 100$, the performance of SC-QPSK signals would become comparable to SC-BPSK signals

$$\mathsf{V}_{\mathcal{H}}\left\{\mathsf{E}_{\mathcal{B}}\{\left|\mathbf{h}_n^H\mathbf{b}_{n,m}\right|^2|\text{Type-1}\}\right\} < \mathsf{V}_{\mathcal{H}}\left\{\mathsf{E}_{\mathcal{B}}\{\left|\mathbf{h}_n^H\mathbf{b}_{n,m}\right|^2|\text{Type-2}\}\right\} < \mathsf{V}_{\mathcal{H}}\left\{\mathsf{E}_{\mathcal{B}}\{\left|\mathbf{h}_n^H\mathbf{b}_{n,m}\right|^2|\text{Type-3}\}\right\}. \quad (59)$$

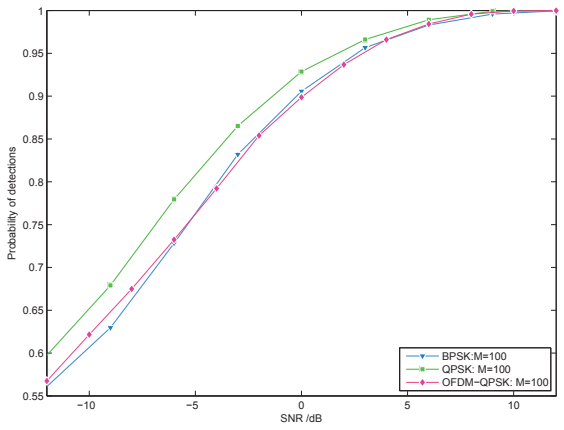


Fig. 6. Detection performance under various different modulated signals. The sample size is $M = 100$, the static length is $N_f = 20$ and the Type-2 channel is used.

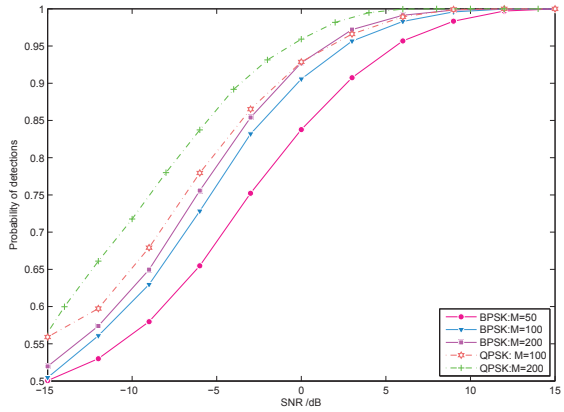


Fig. 7. Detection performance of SC signals (i.e. BPSK and QPSK) under different sample size M. The static length is N_f=20 and the Type-2 channel is used.

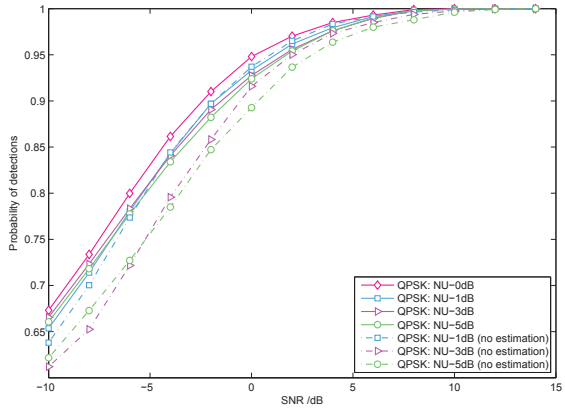


Fig. 8. Detection performance of SC-QPSK signals under realistic noise uncertainty. The sample size is $M = 100$, the static length is $N_f = 20$ and the Type-2 channel is used.

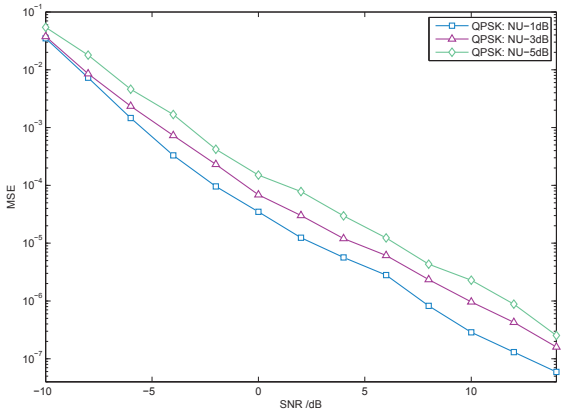


Fig. 9. MSE of the unknown noise variance. The SC-QPSK signals are considered with a sample size of $M = 100$. The static length is $N_f = 20$ and the Type-2 channel is used.

of $M = 200$. The above analysis has also been verified by the experimental results shown in Fig. 7.

D. Noise Uncertainty

As suggested, with the designed recursive mechanic, the noise variance may be estimated by the proposed scheme. In the experiment, the Type-2 channel is adopted, the static length is $N_f = 20$ and the sample size is $M = 100$. The noise variance is assumed to be uniformly distributed around the predefined SNR. Here, "NU-1dB" denotes the maximum derivation from the assume SNR is 1dB. From Fig. 8, we may see that the UN of 1dB may have a slight effect. As shown by experimental results, with the increase of the NUs, however the sensing performance will be degraded to some extent. Thus, even if a recursive estimation technique has been integrated, the designed scheme may slightly suffer from realistic UNs. Fig. 9 also plots the estimation accuracy of unknown noise variance. It is seen that the estimation MSE will be improved as the SNR increases.

E. Comparative Analysis

In order to thoroughly evaluate the proposed algorithm, another covariance absolute value (CAV) scheme is further investigated, which is also proven to be robust relatively even with time-varying fading effects or multipath propagations [17]. In the experiment, the static length is configured to $N_f = 20$. For the CAV method, the smoothing factor is chosen to 12. As has been mentioned, the Nyquist sampling rate is adopted for the ease of hardware implementations. Two multipath configurations are considered, i.e., the Type-1 and Type-2 channels. Firstly, from Fig. 10, the sensing performance of both two schemes will be improved by increasing the sample size M, as the statistical information may become more accurate with a large M. Secondly, for the CAV method, the signal correlation of Type-1 is much stronger than that of Type-2 and, accordingly, the sensing performance will be significantly enhanced with Type-1 multipath channels. A similar result can be observed to the new algorithm, which has also been analyzed in Section IV-C. It is demonstrated by Fig. 10 that, with realistic TF-DSFSc, the proposed method is in general superior to the CAV method. The main reason is that, although these two methods are all premised on statistical information, the CAV method unfortunately ignored the time-varying property accompanying the resulting *memory* of fading channels which could be utilized to further promote

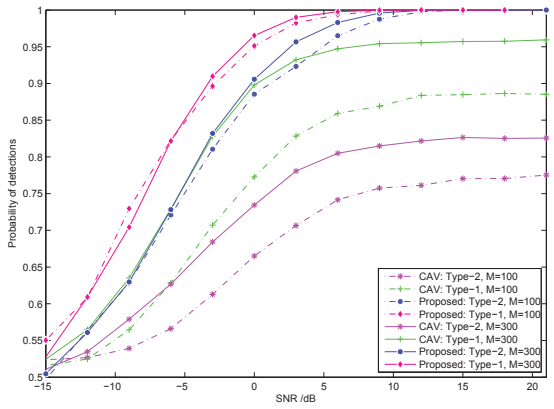


Fig. 10. Performance comparison of different sensing algorithms. The sample size is $M = 100$, the static length is $N_f = 20$ and the Type-2 channel is used.

detection performance [29]. With the formulated DSM and a joint estimation framework, in contrast, the new algorithm will estimate the TF-DSFC accurately and, therefore, improve the sensing performance by fully exploiting the underlying dynamics of time-varying multipath channels.

V. Conclusion

In order to address the formidable challenges posed by realistic time-varying multipath fading channels in distributed CR applications, a new non-coherent spectrum sensing paradigm is proposed in this paper. A novel DSM is established to thoroughly characterize spectrum sensing process, by fully considering the dynamic evolution behaviors of both PU states and time-variant multipath channels. On this basis, an iteratively joint estimation scheme is presented within a Bayesian statistical inference framework. By resorting to an SIS-based PF, spectrum sensing is effectively realized by estimating time-varying multipath gains and unknown PU states jointly. The designed algorithm can also deal with other critical challenges, e.g. noise uncertainty. Both the MSE of channel estimations and the detection probability of PU states are investigated. Simulation results validate the proposed algorithm. With a constructive exploitation of the dynamic of TF-DSFCs, the established new DSM and the designed joint estimation scheme provide a promising solution of spectrum sensing in realistic TF-DSFCs.

Acknowledgment

We greatly thank anonymous reviewers for their constructive comments that allowed us to improve the paper significantly. This work was supported by Natural Science Foundation of China (NSFC) under Grants 61471061 and 61379016, by the Doctoral Fund of Ministry of Education of China under Grant 20130005110016, by the Fundamental Research Funds for the Central Universities under Grant 2014RC0101.

Appendix 1 Transitional Probability of Log-normal Distributions

The Markov transition process, in practice, is assumed to be indecomposable and stationary [25], [36]. That is, if the stationary probability is denoted by $\boldsymbol{\pi}_l = [\pi_{0,l}, \pi_{1,l}, \cdots, \pi_{K-1,l}]^T$ with $\pi_{k,l} \triangleq \Pr(\mathcal{A}_{n,l} = \mathcal{R}_k)$ for any n, then we have $\mathbf{P}_l^T \boldsymbol{\pi}_l = \boldsymbol{\pi}_l$. To obtain discrete channel representations, the nonnegative gain $h_{n,l}$ is partitioned into K non-overlapping regions, with the representative state set denoted by $\mathbb{V}_l$. If we further specify $\nu_{0,l} = 0$ and $\nu_{K,L} = \infty$, we may have

$$\mathbb{V}_l = \{[\nu_{0,l}, \nu_{1,l}), [\nu_{1,l}, \nu_{2,l}), \cdots, [\nu_{K-1,l}, \nu_{K,l})\}.$$

Given the prior log-normal distribution, the steady probability that the lth fading coefficient resides in the kth region is derived by

$$\begin{aligned}\pi_{k,l} =& \int_{\nu_{k,l}}^{\nu_{k+1,l}} \frac{1}{\sigma_h\sqrt{2\pi}} exp\left\{-\left[\ln h_{n,l} - \mathbf{m}_h(l)\right]^2 / 2\sigma_h^2\right\} dh_{n,l} \\ =& F(\nu_{k+1,l}) - F(\nu_{k,l}),\end{aligned} \tag{60}$$

where $F(x) \triangleq \int_0^x \frac{1}{\sigma_h\sqrt{2\pi}} exp\left[-\left(\ln y - \mathbf{m}_h(l)\right)^2 / 2\sigma_h^2\right] dy$ is the cumulated distribution function. For the considered log-normal distribution, we further have $F(x) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left[-\left(\ln x - \mathbf{m}_h(l)\right) / \sqrt{2}\sigma_h\right]$.

Under an equiprobable partition, i.e. $\pi_{k,l} = 1/K$, the partitioning bounds, denoted by $v_{k,l}$, may be easily derived. Then, the transitional probability $P_{k_1 \to k_2, l}$ is determined by

$$\begin{aligned}&P_{k_1 \to k_2, l} \\ &\triangleq \Pr\left\{h_{n',l} \in [\nu_{k_2,l}, \nu_{k_2+1,l}) \middle| h_{n'-1,l} \in [\nu_{k_1,l}, \nu_{k_1+1,l})\right\} \\ &= \frac{1}{\pi_{k,l}} \int_{\nu_{k_2}}^{\nu_{k_2+1}} \int_{\nu_{k_1}}^{\nu_{k_1+1}} f(h_{n'-1,l}, h_{n',l}) dh_{n'-1,l} dh_{n',l}.\end{aligned} \tag{61}$$

For convenience, the level crossing rate (LCR) $N_{k,l}$, which refers to the number of times per second that the lth coefficient crosses $\nu_{k,l}$ in a downward direction, can be further utilized to evaluate the transitional probability. For the log-normal distribution, $N_{k,l}$ will be calculated from

$$\begin{aligned}N_{k,l} =& \frac{f_D \cdot f(\nu_{k,l})}{\sqrt{2\pi}\sigma_h} \times \int_0^\infty exp\left(-\frac{(lnr - m_l)^2}{2\sigma_h^2}\right) dr \\ \overset{(a)}{=}& \frac{f_D \cdot exp\left(m_l + \sigma_h^2/2\right)}{\sqrt{2\pi}\nu_{k,l}\sigma_h} \times exp\left[-\frac{(lnr - m_l)^2}{2{\sigma_h}^2}\right],\end{aligned} \tag{62}$$

where f_D denotes the maximum Doppler frequency shift. (a) holds for the substitution variable t=ln$r$ and $dt = \frac{1}{r}dt$. Relying on the LCR, $P_{k_1 \to k_2, l}$ will be approximately calculated by

$$P_{k_1 \to k_2, l} \simeq N_{k_2,l} / R_{k_1,l}, \tag{63}$$

where $R_{k_1,l} \triangleq \pi_{k_1,l} / T_F$ denotes the average number of frames per second in the state k_1.

Appendix 2 Proof of Remark 1

As in most applications, the OFDM signal $\{b_{n,m}\}$ is considered. Here, the modulation constellations of PU signals are also assumed to be unknown. In practice, $b_{n,m}$ may usually follow the Gaussian distribution, i.e. $b_{n,m} \sim \mathcal{N}(0, \sigma_b^2)$. As suggested, the Gaussian assumption may be justified by

OFDM signals with QPSK and MQAM modulations on subcarriers. The signal variance is specified by $\sigma_b^2 = \sigma_s^2 + \sigma_c^2$, where the variances of the quadrature and in-phase components are given by

$$\sigma_s^2 = \mathsf{E}[\Re\{b_{n,m}^2\}],\ \sigma_c^2 = \mathsf{E}[\Im\{b_{n,m}^2\}],\ \sigma_s^2 = \sigma_c^2 = \frac{\sigma_b^2}{2}, \quad (64)$$

where $\Re(x)$ and $\Im(x)$ represent the real and image parts of x, respectively.

In this situation of H_1, the received signals may be given by eq. (65), where the in-phase component y_s can be further reformatted into eq. (66).

(a) When the sample size M is large (e.g., $M > 100$), according to the central limit theorem (CLT), it will suffice to say the first term $y_{s,1}(n)$ may tend to be Gaussian distributed, i.e., $y_{s,1}(n) \sim \mathcal{N}\{\mathsf{E}(y_{s,1}(n)|\mathbf{h}_n, H_1), \mathsf{V}(y_{s,1}(n)|\mathbf{h}_n, H_1)\}$. Conditioned on the multipath response $\mathbf{h}_n$ at time index n, then the distribution mean and variance can be given by eq. (67) and (68), respectively. In eq. (67), (a) holds for the fact that the quadrature component and in-phase component remain independent of each other, i.e.

$$\mathsf{E}\left[\Re(b_{n,m-1})\Im(b_{n,m-l})\right] = 0. \quad (69)$$

In order to derive a more compact expression of the distribution variance, we have to investigate the property of high-order moments of $b_{n,m}$. For such a zero-mean Gaussian variable, the 4th moment may be calculated from [27]

$$\mathsf{E}\left\{\Re(b_{n,l}^k)\right\} = \begin{cases} 1 \cdot 3 \cdots (k-1) \times \sigma_s^k, & k = 2^r, \quad (70a) \\ 0, & k = 2^r - 1. \quad (70b) \end{cases}$$

Thus, the variance term may be simplified to

$$\begin{aligned} \mathsf{V}(y_{s,1}(n)|\mathbf{h}_n, H_1) &= 2M\sigma_s^4 \times \sum_{l=1}^{L} \left[\Re(h_{n,l})^2 + \Im(h_{n,l})^2\right]^2 \\ &= 2M\sigma_s^4 ||\mathbf{h}_n||_2^4. \end{aligned} \quad (71)$$

(b) The second term is a weighted Gaussian variable, given the independent noise samples with *a priori* Gaussian distribution. Thus, it also becomes a Gaussian variable, i.e.,

$$\begin{aligned} &p\left(y_{s,2}(n)|\mathbf{h}_n, s_{n,m} = 1\right) \\ &= \int_{\mathbf{b}_{n,m} \in \mathcal{B}} p\left(y_{s,2}(n)|\mathbf{h}_n, s_{n,m} = 1\right) d\mathbf{b}_{n,m} \\ &\sim \mathcal{N}\left\{\mathsf{E}(y_{s,2}(n)|\mathbf{h}_n, H_1), \mathsf{V}(y_{s,2}(n)|\mathbf{h}_n, H_1)\right\}, \end{aligned} \quad (72)$$

with its mean and variance respectively given by

$$\begin{aligned} &\mathsf{E}\left(y_{s,2}(n)|\mathbf{h}_n, H_1\right), \mathsf{V}(y_{s,2}(n)) \\ &= 2M \times \mathsf{E}\left\{\sum_{l=1}^{L}\left[\Re(h_{n,l})\Re(b_{n,m-l}) - \Im(h_{n,l})\Im(b_{n,m-1})\right]\right\} \\ &= 2M \times \sum_{l=1}^{L}\left\{\Re(h_{n,l})\mathsf{E}\{\Re(b_{n,m-l})\} - \Im(h_{n,l})\mathsf{E}\{\Im(b_{n,m-1})\}\right\} \\ &= 0, \end{aligned} \quad (73)$$

and

$$\begin{aligned} &\mathsf{V}\left(y_{s,2}(n)|\mathbf{h}_n, H_1\right), \mathsf{V}(y_{s,2}(n)) \\ &= 4M\sigma_z^2\mathsf{E}\left\{\sum_{l=1}^{L}\left[\Re(h_{n,l})\Re(b_{n,m-l}) - \Im(h_{n,l})\Im(b_{n,m-1})\right]\right\}^2 \\ &= 4M\sigma_z^2 \times \sum_{l=1}^{L}\left\{\Re(h_{n,l})^2\mathsf{E}\{\Re^2(b_{n,m-l})\}\right. \\ &\qquad \left. + \Im(h_{n,l})^2\mathsf{E}\{\Im(b_{n,m-1}^2)\}\right\} \\ &= 4M\sigma_z^2\sigma_s^2||\mathbf{h}_n||_2^2. \end{aligned} \quad (74)$$

(c) The third term $y_{s,3}(n)$ is a summation of a group of central chi-square distributed variables with a DoF of M, i.e., $p\left(y_{s,2}(n)|\mathbf{h}_n, H_1 = 1\right) \sim \chi_M^2$. Similarly, consider the sample

$$\begin{aligned} y(n) &= \sum_{m=0}^{M-1} |\mathbf{h}_n^H \mathbf{b}_{n,m} + z_{n,m}|^2 = \sum_{m=0}^{M-1} \left| \sum_{l=1}^{L} [\Re(h_{n,l}) + j\Im(h_{n,l})] \times [\Re(b_{n,m-l}) + j\Im(b_{n,m-1})] + [\Re(z_{n,m}) + j\Im(z_{n,m})] \right|^2 \\ &= \sum_{m=0}^{M-1} \left| \sum_{l=1}^{L} [\Re(h_{n,l})\Re(b_{n,m-l}) - \Im(h_{n,l})\Im(b_{n,m-1})] + \Re(z_{n,m}) \right|^2 \\ &\quad + \sum_{m=0}^{M-1} \left| \sum_{l=1}^{L} [\Re(h_{n,l})\Im(b_{n,m-l}) + \Im(h_{n,l})\Re(b_{n,m-1})] + \Im(z_{n,m}) \right|^2 \\ &= y_s(n) + y_c(n). \end{aligned} \quad (65)$$

$$\begin{aligned} y_s(n) = &\underbrace{\sum_{m=0}^{M-1} \left\{ \sum_{l=1}^{L} \left[\Re(h_{n,l})^2\Re(b_{n,m-l})^2 + \Im(h_{n,l})^2\Im(b_{n,m-1})^2\right] - 2 \times \sum_{l=1}^{L} \left[\Re(h_{n,l})\Re(b_{n,m-l})\Im(h_{n,l})\Im(b_{n,m-1})\right] \right\}}_{y_{s,1}(n)} \\ &+ \underbrace{2 \times \sum_{m=0}^{M-1} \sum_{l=1}^{L} \left[\Re(h_{n,l})\Re(b_{n,m-l})\Im(h_{n,l})\Im(b_{n,m-1})\right] \times \Re(z_{n,m})}_{y_{s,2}(n)} + \underbrace{\sum_{m=0}^{M-1} \Re(z_{n,m})^2}_{y_{s,3}(n)}, \qquad H_1. \end{aligned} \quad (66)$$

number M is usually large in practice, the third term $y_{s,2}$ may also tend to be Gaussian distributed according to the CLT, i.e.

$$p\left(y_{s,3}(n)\right) \sim \mathcal{N}\left\{M\sigma_z^2, 4M\sigma_z^4\right\}. \tag{75}$$

It is seen that, as the PU signal is independent of noise samples, the first term is uncorrelated with the second and third terms. That is, only the cross-correlation function between $y_{s,2}(n)$ and $y_{s,3}(n)$, denoted by $R_{y_{s,2},y_{s,2}}(\Delta m)$, need to be investigated in order to obtained the distribution of y_n, which can be expressed into

$$\begin{aligned} &R_{y_{s,2},y_{s,2}}(\Delta m) \\ &= \mathsf{E}\left\{y_{s,2}^H(n) y_{s,3}(n-\Delta m)\right\} \\ &\simeq 2M^2 \times \mathsf{E}_{\mathcal{H},\mathcal{B}}\left\{\mathbf{h}_n^H \mathbf{b}_{n,m_1}\right\} \mathsf{E}\left\{z_{n,m_1}^H z_{n,m_2}^2\right\}, \end{aligned} \tag{76}$$

and with independent variables $z_{n,m}$, it is easily seen

$$\mathsf{E}\left\{z_{n,m_1}^H z_{n,m_2}^2\right\} \to 0, \quad \text{for} \quad \Delta m = |m_1 - m_2| > 0. \tag{77}$$

Thus, for any $\Delta m > 0$ we have $R_{y_{s,2},y_{s,2}}(\Delta m) \simeq 0$, i.e., $y_{s,2}(n)$ and $y_{s,2}(n)$ remain independent of each other.

Recall that the linear combination of independent Gaussian variables (i.e. $y_{s,2}(n)$ and $y_{s,2}(n)$) forms a Gaussian variable, we may have

$$y_s(n) \sim \mathcal{N}\left\{\mathsf{E}(y_s(n)|\mathbf{h}_n, H_1), \mathsf{V}(y_s(n)|\mathbf{h}_n, H_1)\right\} \tag{78}$$

with its mean and variance determined by:

$$\begin{aligned} \mathsf{E}(y_s(n)|\mathbf{h}_n, H_1) &= \mathsf{E}\left(y_{s,1}(n)|\mathbf{h}_n, H_1\right) + \mathsf{E}(y_{s,3}(n)) \\ &= M\sigma_z^2 ||\mathbf{h}_n||_2^2 + \sigma_z^2, \end{aligned} \tag{79}$$

and

$$\begin{aligned} &\mathsf{V}(y_s(n)|\mathbf{h}_n, H_1) \\ &= \mathsf{V}(y_{s,1}(n)|\mathbf{h}_n, H_1) + \mathsf{V}(y_{s,2}(n)|\mathbf{h}_n, H_1) + \mathsf{V}(y_{s,3}(n)) \\ &= 2M \times \left(\sigma_s^4 ||\mathbf{h}_n||_2^4 + \sigma_z^4 + 2\sigma_z^2\sigma_s^2 ||\mathbf{h}_n||_2^2\right), \end{aligned} \tag{80}$$

respectively.

Similarly, the term $y_c(n)$ is also a Gaussian variable, whose mean and variance is equal to that of $y_s(n)$. Given y_c and $y_s(n)$ involve a group of noise variables of the same distribution, then the likelihood density of the observations $y(n)$ is also a Gaussian distribution, i.e.,

$$y(n) \sim \mathcal{N}\left\{\mathsf{E}(y(n)|\mathbf{h}_n, H_1), \mathsf{V}(y(n)|\mathbf{h}_n, H_1)\right\}, \tag{81}$$

where the mean and variance are specified by eqs. (82) and (83), respectively.

$$\mathsf{E}(y(n)|\mathbf{h}_n, H_1) = 2 \times \mathsf{E}(y_s(n)|\mathbf{h}_n, H_1), \tag{82}$$

$$\mathsf{V}(y(n)|\mathbf{h}_n, H_1) = 2 \times \mathsf{V}(y_s(n)|\mathbf{h}_n, H_1). \tag{83}$$

Based on the above analysis and with little manipulations, we may finally obtain the conclusion in *Remark 1*.

Appendix 3 Estimation of Noise Variance

In the analysis, the unknown noise variance is assumed to follow the *a priori* Inverse-Gamma distribution, as in most investigations [35], [37]. That is, we may have $\sigma_z^2 \sim \mathcal{IG}(\alpha, \beta)$, where the shape parameter is α and the scale parameter is β. Thus, the PDF of σ_z^2 is given by

$$p(\sigma_z^2) = \frac{\beta^\alpha}{\Gamma(\alpha)} \times \left(\frac{1}{\sigma_z^2}\right)^{\alpha+1} \cdot exp\left(-\frac{\beta}{\sigma_z^2}\right). \tag{84}$$

$$\begin{aligned} &\mathsf{E}(y_{s,1}(n)|\mathbf{h}_n, H_1) \\ &= M \times \mathsf{E}\left\{\sum_{l=1}^{L}\left[\Re(h_{n,l})^2\Re(b_{n,m-l})^2 + \Im(h_{n,l})^2\Im(b_{n,m-1})^2\right] - 2 \times \sum_{l=1}^{L}\left[\Re(h_{n,l})\Re(b_{n,m-l})\Im(h_{n,l})\Im(b_{n,m-1})\right]\right\} \\ &= M \times \sum_{l=1}^{L}\left[\Re(h_{n,l})^2\mathsf{E}\{\Re(b_{n,m-l})^2\} + \Im(h_{n,l})^2\mathsf{E}\{\Im(b_{n,m-1})^2\}\right] - 2M \times \sum_{l=1}^{L}\left[\Re(h_{n,l})\Im(h_{n,l})\mathsf{E}\{\Re(b_{n,m-l})\Im(b_{n,m-1})\}\right] \\ &\overset{(a)}{=} M\sigma_s^2 \times \sum_{l=1}^{L}\left[\Re(h_{n,l})^2 + \Im(h_{n,l})^2\right] \\ &= M\sigma_s^2 ||\mathbf{h}_n||_2^2, \end{aligned} \tag{67}$$

$$\begin{aligned} &\mathsf{V}(y_{s,1}(n)|\mathbf{h}_n, H_1) \\ &= M \times \mathsf{E}\left\{\sum_{l=1}^{L}\left[\Re(h_{n,l})^2\Re(b_{n,m-l})^2 + \Im(h_{n,l})^2\Im(b_{n,m-1})^2\right] - 2\sum_{l=1}^{L}\left[\Re(h_{n,l})\Re(b_{n,m-l})\Im(h_{n,l})\Im(b_{n,m-1})\right] - M\sigma_s^2||\mathbf{h}_n||_2^2\right\}^2 \\ &= M \times \mathsf{E}\left\{\sigma_s^4\sum_{l=1}^{L}\left[\Re(h_{n,l})^4 + 2\Re(h_{n,l})^2\Im(h_{n,l})^2 + \Im(h_{n,l})^4\right] - 2\sigma_s^2\Re(h_{n,l})^2\sum_{l=1}^{L}\left[\Re(h_{n,l})^4 + 2\Re(h_{n,l})^2\Im(h_{n,l})^2 + \Im(h_{n,l})^4\right]\right. \\ &\qquad \left. + \mathsf{E}\left\{\Re(b_{n,l}^4)\right\}\sum_{l=1}^{L}\left[\Re(h_{n,l})^4 + \Im(b_{n,m-l})^4\right] + 6\sigma_s^4\sum_{l=1}^{L}\Re(h_{n,l})^2\Im(h_{n,l})^2\right\}, \end{aligned} \tag{68}$$

Given the *a priori* Gaussian distribution of the likelihood density, then the posterior density of unknown noise variance, which is conditioned on unknown PU state and the observations, is expressed as

$$\begin{aligned} & p(\sigma_z^2|y_{0:n}, s_{0:n}, \mathbf{h}_{0:n}) \\ & \propto p(y_{0:n}, s_{0:n}|\sigma_z^2, \mathbf{h}_{0:n}) p(\sigma_z^2|\mathbf{h}_{0:n}) \\ & = p(y_{0:n}|s_{0:n}, \sigma_z^2, \mathbf{h}_{0:n}) p(s_{0:n}) p(\sigma_z^2). \end{aligned} \tag{85}$$

In order to improve the accuracy, we may estimate the unknown noise variance only in this case of $s_n = S_0$. Denote the accumulated length of state S_0 with $n^\dagger$ ($n^\dagger < n$), the likelihood density of observation trajectory of silence states is given by

$$\begin{aligned} & p\left(y_{0:n^\dagger}|\sigma_z^2, s_{0:n^\dagger} \in \{S_0\}\right) \\ & = \frac{\left(\prod_{t=1}^{n^\dagger} y_t\right)^{M/2-1}}{(2\sigma_z^2)^{Mn^\dagger/2} \times \Gamma(M/2)^{n^\dagger}} \times exp\left(\sum_{t=1}^{n^\dagger} y_t \Big/ 2\sigma_z^2\right). \end{aligned} \tag{86}$$

With the Bayesian criterion, the posterior density of the noise variance will become

$$\begin{aligned} & p(\sigma_z^2|y_{0:n^\dagger}, s_{0:n^\dagger} \in \{S_0\}) \\ & \propto p\left(y_{0:n^\dagger}|\sigma_z^2, s_{0:n^\dagger} \in \{S_0\}\right) \times p(\sigma_z^2) \\ & = \frac{\left(\prod_{t=1}^{n^\dagger} y_t\right)^{M/2-1}}{(2\sigma_z^2)^{Mn^\dagger/2} \times \Gamma(M/2)^{n^\dagger}} \times exp\left(\sum_{t=1}^{n^\dagger} y_t \Big/ 2\sigma_z^2\right) \\ & \quad \times \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\sigma_z^2}\right)^{\alpha+1} exp\left(-\frac{\beta}{\sigma_z^2}\right) \\ & \propto \frac{1}{(\sigma_z^2)^{\alpha+1+Mn^\dagger/2}} \times exp\left(-\frac{\beta + \sum_{t=1}^{n^\dagger} y_t/2}{\sigma_z^2}\right). \end{aligned} \tag{87}$$

It is noted that the final term in eq. (87) will also become an inverse Gamma PDF (if normalized), i.e., $\mathcal{IG}(\alpha + 1 + Mn^\dagger/2, \beta + \sum_{t=1}^{n^\dagger} y_t/2)$. Inspired by this derived relationship, now the shape parameter α and the scale parameter β can be refined recursively via

$$\alpha_n = \alpha + Mn^\dagger/2 = \alpha_{n-1} + M/2, \qquad s_n = S_0, \tag{88}$$

$$\beta_n = \beta + \sum_{t=1}^{n^\dagger} y_t/2 = \beta_{n-1} + y_t/2, \qquad s_n = S_0. \tag{89}$$

When the PU state is present (i.e., $s_n = S_1$), note that both the shape and scale parameters, i.e., α_n and β_n, however will be remained unchanged, i.e.

$$\alpha_n = \alpha_{n-1}, \qquad s_n = S_1, \tag{90a}$$

$$\beta_n = \beta_{n-1}, \qquad s_n = S_1. \tag{90b}$$

In realizations, therefore, the two density parameters will be updated according to different rules based on the PU state s_n. With the recursively propagated shape parameter and scale parameter, then the estimation of noise variance could be derived from an MAP criterion, i.e.,

$$\hat{\sigma}_z^2(n) = \beta_n/(\alpha_n - 1).$$

With the designed recursive mechanic, the noise variance may be effectively estimated.

REFERENCES

[1] Y. C. Liang, K. C. Chen, G. Y. Li, and P. Mahonen, "Cognitive Radio Networking and Communications: An Overview," *IEEE Transactions on Vehicular Technology,* vol. 60, no. 7, Sep. 2011, pp. 3386-3407.

[2] J. Mitola and G. Q. Maguire, "Cognitive Radio: Making Software Radios More Personal," *IEEE Personal Communications,* vol. 6, no. 4, Aug. 1999, pp. 13-18.

[3] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," *IEEE Journal on Selected Areas in Communications,* vol. 23, no. 2, Feb. 2005, pp. 201-220.

[4] J. F. Xiao, R. Q. Hu, Y. Qian, L. Gong , et.al. "Expanding LTE network spectrum with cognitive radios: From concept to implementation," *IEEE Wireless Communications,* Vol. 20, No. 2, April 2013, pp. 12-19.

[5] C. Ghosh, S. Roy, D. Cavalcanti, "Coexistence challenges for heterogeneous cognitive wireless networks in TV white spaces," *IEEE Wireless Communications,* Vol. 18, No. 4, August 2011, pp. 22-31.

[6] X. Zhao, J. Kivinen, P. Vainikainen, and K. Skog, "Characterization of Doppler spectra for mobile communications at 5.3 GHz," *IEEE Trans. Veh. Technol.,* vol. 52, no. 1, Jan. 2003, pp. 14-23.

[7] J. Medbo and P. Schramm, *Channel models for HIPERLAN/2 in Different Indoor Scenarios,* ETSI/BRAN Document No. 3ERI085B.

[8] A. F. Molisch, *Wireless Communications*. Chichester (UK): Wiley, 2005.

[9] L. L. He, S. D. Ma, Y. C. Wu, et.al., "Pilot-Aided IQ Imbalance Compensation for OFDM Systems Operating Over Doubly Selective Channels," *IEEE Transactions on Signal Processing,* Vol. 59, No.5, 2011, pp: 2223-2233.

[10] J. Ma, G. Y. Li and B. H. Juang, "Signal Processing in Cognitive Radio," *The Proceedings of IEEE,* vol. 97, no. 5, May 2009, pp. 805-823.

[11] L. Lu, X. W. Zhou, U. Onunkwo and G. Y. Li, "Ten Years of Cognitive Radio Technology," *EURASIP Journal on Wireless Communications and Networking*, vol. 28, 2012, pp. 1-16.

[12] 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 Processing Magazine,* vol. 29, no. 3, May 2012, pp. 101-116.

[13] P. D. Sutton, K. E. Nolan, L. E. Doyle, "Cyclostationary Signature in Practical Cognitive Radio Applications," *IEEE Journal on Selected Areas in Communications,* vol. 2008, pp. 13-24.

[14] H. S. Chen, W. Gao, and D. G. Daut, "Signature Based Spectrum Sensing Algorithms for IEEE 802.22 WRAN," in *Proc. of IEEE International Conference on Communications (ICC),* Glasgow, Scotland, Jun. 2007, pp. 6487-6492.

[15] F. F. Digham, M. S. Alouini, M. K. Simon, "On the Energy Detection of Unknown Signals Over Fading Channels," *IEEE Transactions on Communications,* vol. 55, no. 1, Jan. 2007, pp. 21-24.

[16] Z. Tian and G. B. Giannakis, "Compressed Sensing for Wideband Cognitive Radios," *in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP),* Honolulu, HI, Apr. 2007, pp. 1357-1360.

[17] Y. H. Zeng, Y. C. Liang, "Eigenvalue-based Spectrum Sensing Algorithms for Cognitive Radio," *IEEE Transactions on Communications,* vol. 57, 2009, pp. 1784-1793.

[18] X. W. Zhou, Y. Li, Y. H. Kwon, S. A. "Detection Timing and Channel Selection for Periodic Spectrum Sensing in Cognitive Radio," *in Proc. of IEEE Global Telecommunications Conference (GLOBECOM 2008),* New Orleans, LA, Nov. 30th-Dec. 4th, 2008, pp. 1-5.

[19] J. Ma and Y. Li, "A Probability-based Spectrum Sensing Scheme for Cognitive Radio," *in Proc. IEEE International Conference on Communications (ICC),* Beijing, China, May 2008, pp. 1-5.

[20] B. Li, C. L. Zhao, Z. Zhou, A. Nallanathan, "Joint Estimation-Based Spectrum Sensing for Cognitive Radios in Time Variant Flat Fading Channel," in *Proc. of IEEE Global Telecommunications Conference (Globecom 2013)*, Atlanta, GA, USA, 9-13th December, 2013, pp. 1-6.

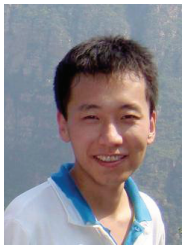
[21] Y. C. Liang, Y. H. Zeng, E. C. Y. Peh, T. H. Anh, "Sensing-Throughput Tradeoff for Cognitive Radio Networks," *IEEE Transactions on Wireless Communications,* vol. 7, no. 4, April 2008, pp. 1326-1337.

[22] B. Vujitic, N Cackov, S. Vujitic. "Modeling and Characterization of Traffic in Public Safety Wireless Network," *in Proc. Int. Symp. Performance Evaluation of Computer and Telecommunication Systems,* Edinburgh, UK, July 2005, pp. 213-223.

[23] B. Li, Z. Zhou, W. X. Zou, "Interference Mitigation between Ultra-Wideband Sensor Network and Other Legal Systems," *EURASIP Journal on Wireless Communications and Networking,* Article number: 290306, 2010, pp. 1-15.

[24] K. E. Baddour and N. C. Beaulieu, "Autoregressive Models for Fading Channel Simulation," in *Proc. IEEE Global Communication Conf. (GLOBECOM),* San Antonio, TX, Nov. 2001, pp. 1187-1192.

[25] P. Sadeghi, R. Kennedy, P. Rapajic, R. Shams, "Finite-state Markov Modeling of Fading Channels: A Survey of Principles and Applications," *IEEE Signal Processing Magazine,* vol. 25, n0. 5, 2008, pp. 57-80.
[26] H. S. Wang, N. Moayeri, "Finite-state Markov Channel: A Useful Model for Radio Communication Channels," *IEEE Transactions on Vehicular Technology,* vol. 44, no. 1, Feb. 1995, pp. 163-171.
[27] John G. Proakis. *Digital Communications 3rd ed.*. Singapore: McGraw-Hill Book Co., 1995.
[28] W. Zhang, R. K. Mallik, and K. B. Letaief, "Cooperative Spectrum Sensing Optimization in Cognitive Radio Networks," *in Proc. IEEE International Conference on Communications (ICC 2008),* Beijing, China, May 19-23, 2008, pp. 3411-3415.
[29] B. Li, C. L. Zhao, M. W. Sun, Z. Zhou, A. Nallanathan "Spectrum Sensing for Cognitive Radios in Time-Variant Flat-Fading Channels: A Joint Estimation Approach," *IEEE Transactions on Communications,* vol. 62, no. 8, 2014, pp. 2665-2680.
[30] D. A. Reynolds, R. C. Rose, "Robust Text-Independent Speaker Identification using Gaussian Mixture Speaker Models," *IEEE Transactions on Acoustics, Speech, and Signal Processing,* vol. 3, no. 1, 1995, pp.72C83.
[31] P. M. Djuric, J. H. Kotecha, J. Q Zhang, Y. F. Huang, T. Ghirmai, M. F. Bugallo, J. Miguez, "Particle Filtering," *IEEE Signal Processing Magazine,* vol. 20, no. 5, 2003, pp. 19-38.
[32] Z. G. Yang, X. D. Wang, "A Sequential Monte Carlo Blind Receiver for OFDM Systems in Frequency-selective Fading Channels," *IEEE Transactions on Signal Processing,* vol. 50, no. 2, 2002, pp. 271-280.
[33] J. Liu and R. Chen, "Blind Deconvolution via Sequential Imputations," *J. Amer. Statist. Assoc.*, vol. 90, 1995, pp. 567-576.
[34] C. P. Robert, *The Bayesian Choice*. New York: Springer-Verlag, 1996.
[35] X. D. Wang, R. Chen, "Blind Turbo Equalization in Gaussian and Impulsive Noise," *IEEE Transactions on Vehicular Technology*, Vol. 50, no. 4, July. 2001, pp. 1092-1105.
[36] B. Li, C. L. Zhao, M. W. Sun, A. Nallanathan, "Energy Detection based Spectrum Sensing In The Presence of Time-Frequency Double Selective Fading Propagations," in *Proc. of the IEEE Global Communications Conference (Globecom 2014)*, Austin, TX, USA, Dec. 2014, pp. 1-6.
[37] D. Arnaud, and X. D. Wang. "Monte Carlo Methods for Signal Processing: A Review in the Statistical Signal Processing Context," *IEEE Signal Processing Magazine*, vol. 22, no. 6, 2005, pp. 152-170.



Bin Li received the Bachelor's degree in electrical information engineering from Beijing University of Chemical Technology (BUCT) in 2007, the Ph.D. degree in communication and information engineering from Beijing University of Posts and Telecommunications (BUPT) in 2013. He joined BUPT since 2013, as a lecture of 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) and channel modeling. He has published more than 40 journal and conference papers. He received 2011 ChinaCom Best Paper Award, 2010 and 2011 BUPT Excellent Ph.D Student Award Foundation. He served as the regular reviewer for IEEE SIGNAL PROCESSING LETTER, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

Mengwei Sun received the Bachelor's degree in communication engineering from Civil Aviation University of China (CAUC) in 2011. She is currently working toward the Ph.D. degree with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications (BUPT), Beijing, China. Her current research interests focus on signal processing for cognitive radios and 60GHz millimeter-wave (mm-Wave) communications.

Xiaofan Li (M'10) is the Deputy Chief Engineer of the State Radio monitoring center and Testing Center (SRTC) Shenzhen Lab, received her B.S. degree and Ph.D degree from Beijing University of Posts and Telecommunication (BUPT) in 2007 and 2012. From 2010 to 2011, she was a visiting Ph.D student in University of Washington. She has involved in nearly 10 national foundation projects. Her research interests include interference analysis among different radio systems, testing method and technology on radio equipments, cooperative transmission, cognitive radio, internet of things, radio management policy, etc.

Arumugam 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, 2011/12. 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. He was an Editor for IEEE Transactions on Wireless Communications (2006-2011), IEEE Wireless Communications Letters, IEEE Signal Processing Letters and a Guest Editor for EURASIP Journal on Wireless Communications and Networks: Special issue on UWB Communication Systems-Technology and Applications (2006). He currently serves as the Chair for the Signal Processing and Communication Electronics Technical Committee of IEEE Communications Society. He served as the 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 received the IEEE Communications Society SPCE outstanding service award 2012 and IEEE Communications Society RCC outstanding service award 2014.

Chenglin Zhao received the Bachelor's degree in radio-technology from Tianjin University in 1986, and the Master's degree in circuits and systems from Beijing University of Posts and Telecommunications (BUPT) in 1993, and the Ph.D. degree in communication and information system from Beijing University of Posts and Telecommunications, in 1997. 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, 60GHz millimeter-wave communications.