Information Geometry-Based Fuzzy-C Means Algorithm for Cooperative Spectrum Sensing

YONGWEI ZHANG [1], **CANFENG MA** [1], **YONGHUA WANG** [1], **(Member, IEEE),**
SHUNCHAO ZHANG [1], **AND PIN WAN** [1,2]

[1]School of Automation, Guangdong University of Technology, Guangzhou 510006, China
[2]Hubei Key Laboratory of Intelligent Wireless Communications, South-Central University for Nationalities, Wuhan 430074, China

Corresponding author: Yonghua Wang (wangyonghua@gdut.edu.cn)

ABSTRACT In this paper, the spectrum sensing problem is investigated under the context of information geometry and a novel clustering algorithm based spectrum sensing scheme is developed to obtain a classifier to estimate the channel state of primary user (PU). In order to enhance the sensing performance at complex environment, the empirical mode decomposition (EMD) algorithm is applied to wipe off the noise component of the received signals from all secondary users (SUs). Subsequently, a signal matrix composed of the reconstructed signals is constructed. Based on the information geometry (IG) theory, the covariance matrix of the signal matrix is mapped into a point on a manifold. Then, the sample points on the manifold are collected as a data set. Moreover, a novel clustering algorithm, namely Riemannian distance based Fuzzy-c means clustering (RDFCM) algorithm, is developed to cluster the samples on manifold for obtaining a classifier, which is employed to decide the PU state. The simulation results show that compared with other spectrum sensing methods, the proposed scheme improves the performance of detection.

INDEX TERMS Cognitive radio, spectrum sensing, information geometry, empirical mode decomposition.

I. INTRODUCTION

With the rapid development of the wireless communication technology, the amount of wireless devices and services has increased dramatically over the past decade, which requires a great deal of spectrum resources. However, according to the investigation report of Federal Communications Commission, the utilization of spectrum resources is poor [1]–[3]. In order to conquer this drawback, the spectrum sensing which is a vital technique in cognitive radio (CR), has been extensively developed to obtain awareness about the spectrum usage and existence of primary users (PUs) in an area [4]–[7]. Several conventional spectrum sensing approaches, such as energy detection (ED) [8], [9], matched filtering detection (MFD) [10], [11] and cyclostationary feature detection (CFD) [12], [13] have been developed. The ED approach is a comparatively simple technique, which obtains the sensing decision by calculating the energy of the signal and comparing it to a threshold. However, it is vulnerable to noise

The associate editor coordinating the review of this manuscript and approving it for publication was Hayder Al-Hraishawi [ORCID].

uncertainty. By analyzing the spectral correlation function of the signal, the CFD method is competent to differentiate the signal and noise, but the computational burden is enormous. The MFD algorithm requires the priori knowledge of the PU, such as bandwidth and operating frequency, which is restricted in practical application.

A. RELATED WORK

The aforementioned works are based on signal secondary user (SU), which are susceptible to multipath fading, shadowing and receiver uncertainty. To conquer these bottlenecks, the cooperative spectrum sensing (CSS) which is an indispensable framework in CR, has achieved widespread attention to improve the sensing performance by integrating the information received from multiple SUs [14]–[17]. Compared with signal SU sensing, the SUs in CSS transmit their sensing information to a fusion center (FC) to make a global decision, which is more precise than individual decision. Most of the CSS approaches are based on random matrix theory (RMT) [18]–[20]. The FC establishes a signal

TABLE 1. Weaknesses of different methods.

References	Methods	Weaknesses
Ref. [8], [10], [12], [18]	Traditional methods	The fixed decision threshold is required
Ref. [22], [23]	IG-based methods	The fixed decision threshold is required
Ref. [24]–[27]	Machine learning-base methods	Only work in linear space

matrix based on the signals received from all the SUs. Then, the statistical eigenvalue of the covariance matrix is calculated and compared it with a threshold to determine the state of PU. On the other hand, the information geometry (IG) based CSS methods are gradual investigated [21]–[23]. In [22], a novel wideband spectrum sensing method based on Riemannian distance and Riemannian mean was proposed. Based on IG theory, a Riemannian distance detector (RDD) which is blind to noise statistical characteristics and priori knowledge of PU, was designed to detect the spectrum hole. In [23], the covariance matrices of the received signal matrix were mapped into a series of coordinates on the statistical manifold, and the signal features were constructed by the geodesic distances between the coordinates and the corresponding Riemannian mean. Then, a constant false alarm rate detector was applied to determine the state of PU. In the view of IG theory, the statistical features are located in the manifold, rather than the linear space theory, which provides a novel perspective to analyze the spectrum sensing problem.

It should be emphasised that all the aforementioned works need to deduce a precisely threshold to decide the state of channel, which is intractable in complicated environment. In recent years, machine leaning has been widely applied in various fields due to its capacity to employ mathematical calculations to analyze and interpret patterns and structures in data. In spectrum sensing community, several machine leaning based schemes have been proposed [24]–[28]. In [24], newly CSS approaches based on machine learning algorithms, such as K-means clustering, Gaussian mixture model and K-nearest-neighbor, were developed. The vector of the signal energy is considered as a feature vector to train a classifier to decide whether the PU is presence or not. In [25], a spectrum sensing method based on K-means clustering and support vector machine algorithms was presented. To reduce the computational burden, a low-dimensional probability vector was designed as the feature vector and used it to obtain a classifier to achieve spectrum sensing. In [26], the statistical eigenvalue of the covariance matrix, such as maximum-minimum eigenvalue (MME), was treated as the feature vector to train a classifier by using the K-means and K-medoids algorithms, respectively. In general, the machine learning based spectrum sensing approaches use a feature vector, such as energy statistic, probability vector and statistical eigenvalue of the covariance matrix, to train a classifier by using supervised or unsupervised learning algorithm. Therefore, choosing appropriate eigenvalue and machine learning algorithm are crucial to obtain desired sensing performance.

B. MOTIVATIONS

The weaknesses of different methods are displayed in Table 1. To overcome these problems, a new spectrum sensing method based on IG theory is developed in this paper. In order to realize clustering directly in manifold space, under the context of IG theory, a novel spectrum sensing scheme is developed in this paper by using a newly designed Riemannian distance based Fuzzy-c means clustering (RDFCM) algorithm. In real scenario, the received information of SUs are adulterated by noise, which degrades the detection performance. To address this issue, the empirical mode decomposition (EMD) algorithm which treated as a forceful technique to deal with the nonstationary and nonlinear signal, is adopted to eliminate noise component [29]–[31]. Then, based on the IG theory, the covariance matrices of the signal matrices are mapped into a series of coordinate points on the manifold. Subsequently, the RDFCM algorithm is designed, which clustering directly on the manifold space. Finally, a classifier is trained to perceive the state of PU by using the coordinate points on the manifold.

The novelty and contribution of this work are summarized as follows:

1) This paper develops a new clustering algorithm, namely RDFCM algorithm. It achieves clustering directly on manifold space by applying Riemannian distance to measure the distance between two points on the manifold. The original FCM algorithm is suitable for linear space only. It cannot cluster the points which lie on the manifold space. Therefore, the developed RDFCM algorithm is the extension of the FCM algorithm.

2) Different from existing methods [20], [21], this paper extends the RDFCM clustering algorithm to deal with the spectrum sensing problem. Unlike the traditional methods, this method is realized on manifold space and the decision threshold is not required. Moreover, based on the sample points on the manifold and the developed RDFCM algorithm, a classifier is trained to decide whether the PU is present or not.

3) This scheme integrates the EMD algorithm to exclude the noise component of the signals received by SUs, which can be considered as a data preprocessing process to obtain superior feature. In simulation part, the availability of the developed scheme is analyzed under the SUs at different signal to noise ratio (SNR).
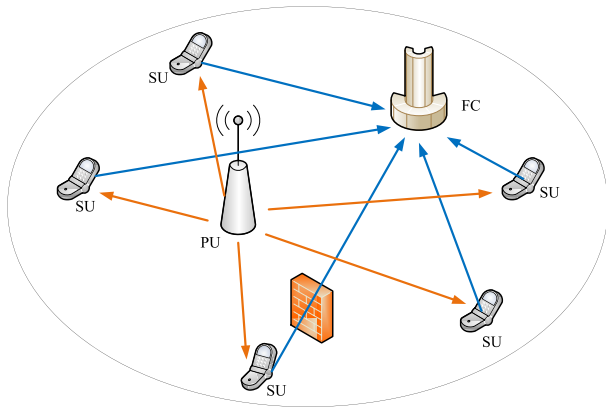
FIGURE 1. The scenario of CSS.

The rest of this paper is organized as follows. Section II introduces the scenario of CSS. Section III proposes an EMD-based RDFCM approach (EMD-RDFCM) for CSS. In Section IV, the effectiveness of the EMD-RDFCM approach is verified under different conditions. In Section V, the conclusion is given.

II. COOPERATIVE SPECTRUM SENSING

Consider a CR system with one PU, one FC and L SUs, as displayed in Figure 1. In CSS scenario, SUs share their sensing data to FC for acquiring a global decision to decide the presence or absence of the PU. Typically, the spectrum sensing can be considered as a binary hypotheses testing problem, which can be expressed as

$$\mathcal{H}_0 : y_l(n) = z_l(n), n = 1, \ldots, N,$$
$$\mathcal{H}_1 : y_l(n) = h_l(n)x(n) + z_l(n), n = 1, \ldots, N, \quad (1)$$

where $y_l(n)$ and $x_l(n)$ represent the received and transmitted signals of the lth SU, N is the number of sample points. $z_l(n)$ is the Gaussian white noise and satisfies $z_l(n) \sim \mathcal{N}(0, \sigma_z^2)$, $h_l(n)$ denotes the channel gain, $\mathcal{H}_0$ and $\mathcal{H}_1$ stand for the absence and presence of PU, respectively. Hence, the signal matrix can be established as

$$\mathbf{Y} = \begin{bmatrix} y_1(1) & y_1(2) & \cdots & y_1(N) \\ y_2(1) & y_2(2) & \cdots & y_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ y_L(1) & y_L(2) & \cdots & y_L(N) \end{bmatrix}. \quad (2)$$

In order to reflect the sensing performance, the probabilities of detection P_d and false alarm P_f are formulated as

$$P_d = P[\mathcal{H}_1|\mathcal{H}_1], \quad (3)$$
$$P_f = P[\mathcal{H}_1|\mathcal{H}_0]. \quad (4)$$

III. COOPERATIVE SPECTRUM SENSING BASED ON EMD-RDFCM APPROACH

The structure of EDM-RDFCM approach is given in Figure 2. To begin with, the EMD algorithm is employed to eliminate

the noise component in the received signal. Subsequently, the reconstructed signal matrix is established and the corresponding covariance matrix is calculated. Moreover, based on the IG theory, the covariance matrices are mapped to coordinate points on the manifold and the new RDFCM clustering algorithm is developed. Finally, a data set is prepared to train a classifier to determine the state of the channel.

A. EMD-BASED SIGNAL DENOISING
In practical sensing scenario, SUs will be disturbed by noise, which degrades the sensing performance. To address this conundrum, the EMD algorithm which is considered as an effective technique for nonlinear and nonstationary modal decomposition, is employed to remove the noise component in the noisy signal that collected by SUs. The basic idea of EMD algorithm is decomposing the signal into a series of intrinsic mode functions (IMFs). The major advantage of the EMD is that the basis functions are derived from the signal itself. Hence, the analysis is adaptive in contrast to the traditional methods where the basis functions are fixed. It is noticed that most of the noise components are concentrated on the higher frequency ones. Hence, by reconstructing the low-frequency modes, the noise part can be eliminated. Assume that $y_l(t)$ represents the received noisy signal for lth SU. The decomposition process of EMD algorithm is given in Algorithm 1.

After EMD decomposition, the received noisy signal $y_l(t)$ for lth SU can be represented by

$$y_l(t) = \sum_{j=1}^{C} \text{IMF}_j(t) + r_C(t), \quad (5)$$

where C is the number of modes, $r_C(t)$ is a residual. In order to remove the noise components in $y_l(t)$, the IMF components at low frequency should be reconstructed. Hence, we need to find a mode index j_s, after which, IMF components are dominated by the signal. In this paper, the consecutive MSE (CMSE) is adopted to determine the j_s. Define $\tilde{y}_l^{j_s}(t)$ is the reconstruction of the original signal, i.e.,

$$\tilde{y}_l^{j_s}(t) = \sum_{j=j_s}^{C} \text{IMF}_j(t) + r_C(t). \quad (6)$$

Then, the CMSE is defined as

$$\text{CMSE}\left(\tilde{y}_l^{j_s}(t), \tilde{y}_l^{j_s+1}(t)\right) \triangleq \frac{1}{N} \sum_{i=1}^{N} \left(\tilde{y}_l^{j_s}(t_i) - \tilde{y}_l^{j_s+1}(t_i)\right)^2$$
$$\triangleq \frac{1}{N} \sum_{i=1}^{N} \left(\text{IMF}^{j_s}(t_i)\right)^2. \quad (7)$$

Therefore, the index j_s is obtained by

$$j_s = \underset{1 \leq j_s \leq C-1}{\arg \min} \text{CMSE}\left(\tilde{y}_l^{j_s}(t), \tilde{y}_l^{j_s+1}(t)\right). \quad (8)$$
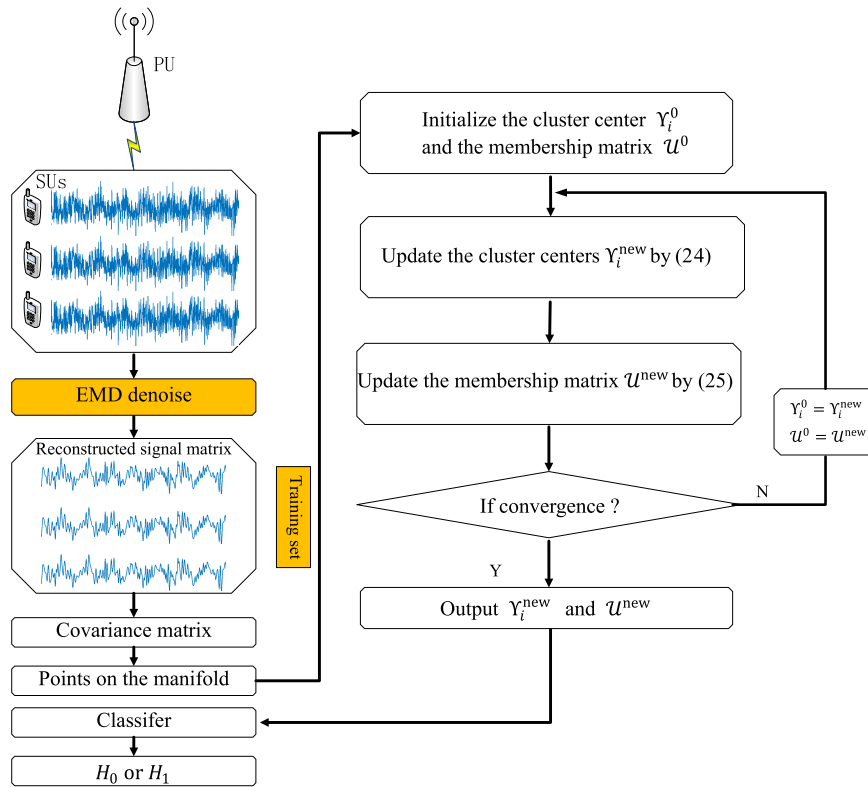
FIGURE 2. The structure of EMD-RDFCM approach.

After EMD denoising, the received signal for lth SU is expressed by

$$\bar{y}_l(t) = \sum_{j=j_s}^{C} \text{IMF}_j(t) + r_C(t). \tag{9}$$

Then, the signal matrix is represented as

$$\mathbf{Y} = \begin{bmatrix} \bar{y}_1(1) & \bar{y}_1(2) & \cdots & \bar{y}_1(N) \\ \bar{y}_2(1) & \bar{y}_2(2) & \cdots & \bar{y}_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ \bar{y}_L(1) & \bar{y}_L(2) & \cdots & \bar{y}_L(N) \end{bmatrix}. \tag{10}$$

B. SPECTRUM SENSING BASED ON INFORMATION GEOMETRY

Based on (10), the covariance matrix of signal matrix is calculated as

$$\mathbf{C} = \frac{1}{N} \mathbf{Y} \mathbf{Y}^{\mathrm{T}} \tag{14}$$

Then, the binary hypotheses can be formulated as

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{C} = \sigma_z^2 \mathbf{I}, \\ \mathcal{H}_1 &: \mathbf{C} = \mathbf{C}_p + \sigma_z^2 \mathbf{I}, \end{aligned} \tag{15}$$

where $\mathbf{C}_p \in \mathbb{R}^{L \times L}$ is the PU signal matrix, $\mathbf{I} \in \mathbb{R}^{L \times L}$ is the identity matrix. From [22], [32], we know that $\mathbf{C}$ follows the Wichter distributions $\mathbb{W}(L, \sigma_z^2 \mathbf{I})$ and $\mathbb{W}(L, \mathbf{C}_p + \sigma_z^2 \mathbf{I})$ in the case of $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively.

Algorithm 1 EMD Algorithm

Step 1: Extract all the extreme points on $y_l(t)$.

Step 2: Calculate the upper and lower envelopes $y_l^{\max}(t)$ and $y_l^{\min}(t)$ by cubic spline interpolation approach. Then, the envelope mean is calculated by

$$m_1(t) = \frac{y_l^{\max}(t) + y_l^{\min}(t)}{2}. \tag{11}$$

Step 3: Calculate the first component by

$$h_1(t) = y_l(t) - m_1(t). \tag{12}$$

If $h_1(t)$ satisfies the IMF conditions, then goto Step 4. Otherwise, repeat Steps 1 and 2 on $h_1(t)$.

Step 4: Let $\text{IMF}_1 = h_1(t)$ and calculate the residual signal by

$$r_C(t) = y_l(t) - \text{IMF}_1. \tag{13}$$

Step 5: If $r_C(t)$ is a monotonous function or a constant, then let $r_C(t)$ be the final residual signal. Otherwise, let $r_C(t)$ be the original signal and repeat Steps 1–5.

IG is the application of differential geometry in statistics. The main idea is using parameterized probability distribution clusters to construct statistical manifold and then transform the statistical problem into a geometric problem. Spectrum sensing is a kind of signal detection problem, which can be solved by analyzing the probability distribution of

detection data. Based on IG theory, the spectrum sensing problem can be transformed into a geometric problem on manifold space and the properties of probability distribution function clusters can be analyzed more intuitively by using geometric method. In the context of IG theory, the statistical manifold $\mathcal{S}_m$ can be defined as

$$\mathcal{S}_m = \{p(x|\theta)|\theta \in \Theta\}, \tag{16}$$

where $x \in \mathbb{R}^n$ is the random variable, $\theta \in \mathbb{R}^m$ is the parameter vector, $p(x|\theta)$ is the probability density function, Θ is the probability distribution space. In IG theory, θ can be considered as the coordinate on the manifold. From the above analysis, the Wichter distributions $\mathbb{W}(L, \sigma_z^2 \mathbf{I})$ and $\mathbb{W}(L, \mathbf{C}_p + \sigma_z^2 \mathbf{I})$ can be mapped into the statistical manifold and the covariance matrices $\sigma_z^2 \mathbf{I}$ and $\mathbf{C}_p + \sigma_z^2 \mathbf{I}$ can be regarded as the corresponding coordinates. The structure of the statistical manifold is shown in Figure 3.
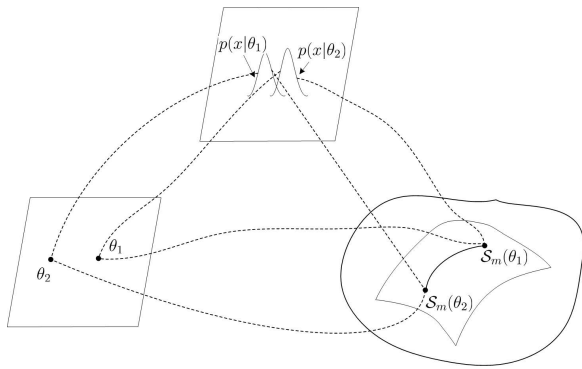

FIGURE 3. The structure of the statistical manifold.

In IG theory, the Riemannian distance which is the shortest curve to connect the two coordinate points on manifold, can be used to indicate the similarity of two distributions. Let Ξ_1 and Ξ_2 be the two points on the manifold. Then, the Riemannian distance between Ξ_1 and Ξ_2 can be calculated by

$$
\begin{aligned}
\mathcal{D}^2(\Xi_1, \Xi_2) &= \left\| \log(\Xi_1^{-\frac{1}{2}} \Xi_2 \Xi_1^{-\frac{1}{2}}) \right\|^2 \\
&= \left\| \log(\Xi_1^{-1} \Xi_2) \right\|^2 \\
&= \operatorname{tr}\left(\log^2(\Xi_1^{-1} \Xi_2) \right) \\
&= \sum_{i=1}^{L} \log^2(\lambda_i),
\end{aligned}
\tag{17}
$$

where $\|\cdot\|$ is the Frobenius norm and λ_i is the i eigenvalues of the matrix $\Xi_1^{-1} \Xi_2$.

C. COOPERATIVE SPECTRUM SENSING BASED ON RDFCM CLUSTERING ALGORITHM

FCM algorithm is a vital clustering approach in unsupervised machine learning, which groups the data points into a specific number of clusters by minimizing the distance between

the data and the cluster centers of their fuzzy memberships [33]–[36]. Let $\mathcal{S} = \{s_1, \ldots, s_k\}$ be a data set containing k data points, q be the number of clusters and satisfies $2 \leq q < k$, $C_F = \{c_1, \ldots, c_q\}$ be a set of cluster centers. Then, FCM algorithm performs clustering by addressing

$$\text{minimize } \mathcal{J}(U, C_F) = \sum_{i=1}^{q} \sum_{j=1}^{k} \mu_{ij}^m \|s_j - c_i\|^2,$$

$$\text{subject to } \sum_{i=1}^{q} \mu_{ij} = 1, \sum_{j=1}^{k} \mu_{ij} > 0, \tag{18}$$

where $\mu_{ij}^m \in [0, 1]$ is the membership of s_j in class i, $m \in [1, \infty)$ is the degree of fuzzification, $U = \{\mu_{ij}\} \in \mathbb{R}^{q \times k}$ is the membership matrix. The constraints guarantee that each data has the same overall weight in the data set and none of the clusters is empty. By using the alternating optimization approach, the update equations at p iterate are obtained by

$$
\mu_{ij}^{p+1} = \begin{cases}
\left(\sum_{h=1}^{q} \left[\dfrac{\|s_j - c_i^p\|^2}{\|s_j - c_h^p\|^2} \right]^{\frac{1}{m-1}} \right)^{-1}, & \text{if } \mathcal{I}_j = \emptyset \\
\dfrac{1}{|\mathcal{I}_j|}, & \text{if } \mathcal{I}_j \neq \emptyset, i \in \mathcal{I}_j \\
0, & \text{if } \mathcal{I}_j \neq \emptyset, i \notin \mathcal{I}_j
\end{cases}
$$

$$
c_i^{p+1} = \frac{\sum_{j=1}^{k} (\mu_{ij}^{p+1})^m s_j}{\sum_{j=1}^{k} (\mu_{ij}^{p+1})^m},
$$

where $\mathcal{I}_j = \{i | i \in [1, q], s_j = c_i^p\}$. It is noticed that traditional FCM algorithm uses Euclidean distance to measure the distance between the sample point and the cluster center, which only available in linear space. To achieve clustering on the manifold, a modified FCM algorithm, namely RDFCM method is developed, which uses Riemannian distance instead of Euclidean distance to measure the distance between two elements on manifold.

Under the spectrum sensing scenario, RDFCM needs to divide the sampled data located on the manifold into two groups, i.e., the PU is idle and the PU is active. Before training the classifier, the data set $\mathcal{T}$ which contains M covariance matrices, needs to obtain, such that

$$\mathcal{T} = \{\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_M\}. \tag{19}$$

Let $\mathbf{T}_j \in \mathbb{R}^{L \times L}$ be the jth sample in $\mathcal{T}$. Denote Υ_i is the ith center of the clusters. Then, the objective function of RDFCM algorithm is designed as

$$\text{minimize } \mathcal{J}_{\text{RDFCM}} = \sum_{j=1}^{M} \sum_{i=1}^{q} \mu_{ij}^m \mathcal{D}^2(\mathbf{T}_j, \Upsilon_i),$$

$$\text{subject to } \sum_{i=1}^{q} \mu_{ij} = 1, \sum_{j=1}^{M} \mu_{ij} > 0, \tag{20}$$

Algorithm 2 RDFCM

Step 1: Input ε, $\mathcal{T}$ and let $t = 0$.

Step 2: Initialize the cluster center Υ_i^0 and the membership matrix $\mathcal{U}^0 = \{\mu_{ij}\}_{1 \leq i \leq q, 1 \leq j \leq M}$.

Step 3: Calculate the equation (24) and update the cluster centers.

Step 4: Calculate the equation (25) and update the membership matrix.

Step 5: If $\|\mathcal{J}_{RDFCM}^{t+1} - \mathcal{J}_{RDFCM}^t\|^2 \leq \varepsilon$, then goto Step 3. Otherwise, let $t = t + 1$, go back to Step 2 and continue.

Step 6: Output Υ_i and $\mathcal{U}$.

---

where $\mathcal{D}^2(T_j, \Upsilon_i)$ is the Riemannian distance between T_j and Υ_i, μ_{ij}^m is the membership and satisfies $m = 2$. The Lagrange multiplier method is employed to minimize the objective function. Hence, we can obtain

$$L(\mu_{ij}, \Upsilon_i, \lambda_i) = \sum_{j=1}^{M} \Big(\sum_{i=1}^{q} \mu_{ij}^m \mathcal{D}^2(T_j, \Upsilon_i)$$

$$+ \lambda_i \Big(1 - \sum_{i=1}^{q} \mu_{ij}\Big)\Big). \quad (21)$$

The partial derivatives of L with respect to μ_{ij} and Υ_i can be calculated as

$$\frac{\partial L}{\partial \mu_{ij}} = m\mathcal{D}^2(T_j, \Upsilon_i)\mu_{ij}^{m-1} - \lambda_i, \quad (22)$$

$$\frac{\partial L}{\partial \Upsilon_i} = -\sum_{j=1}^{M} 2\mu_{ij}^m(T_j - \Upsilon_i). \quad (23)$$

Then, we can final obtain

$$\Upsilon_i = \frac{\sum_{j=1}^{M} \mu_{ij}^m T_j}{\sum_{j=1}^{M} \mu_{ij}^m}, \quad (24)$$

$$\mu_{ij} = \sum_{h=1}^{q} \left(\frac{\mathcal{D}^2(T_j, \Upsilon_i)}{\mathcal{D}^2(T_j, \Upsilon_h)} \right)^{-\frac{1}{m-1}}. \quad (25)$$

From (24) and (25), we know that Υ_i and μ_{ij} are related to each other. Therefore, an iterative approach is adopted to obtain the optimal solution. The specific process is given in Algorithm 2. After the training is accomplished, a classifier for spectrum sensing is obtained by

$$\mathcal{F}(\hat{T}) = \frac{\mathcal{D}^2(\Upsilon_1, \hat{T})}{\mathcal{D}^2(\Upsilon_2, \hat{T})}, \quad (26)$$

where $\hat{T}$ is the data on the manifold that need to be classified. If $\mathcal{F}(\hat{T}) > \xi$, it means that the PU is absence and the SUs are not allowed to be accessed. If $\mathcal{F}(\hat{T}) \leq \xi$, it indicates that PU is presence, then the SUs can use the licensed spectrum. ξ is a designed parameter that uses to control the P_d and P_f.

D. COMPLEXITY ANALYSIS

The complexity of the EMD-RDFCM-based CSS method consists of two parts, i.e., the training phase and the sensing phase. In training phase, the complexity is expressed

as $O((M \times q) \times I_k)$, where M is the number of data set, q is the number of cluster centers, and I_k is the number of iterations. It is worth mentioning that the complexity of the EMD-RDFCM-based CSS approach in training phase can be rewritten as $O(M)$ as the M is increase. In sensing phase, the classifier has been obtained and can be used directly. Therefore, comparing with the training phase, the complexity of the sensing phase is low and can be ignored.

Remark 1: The differences between the EMD-RDFCM method and existing methods lies in that: 1) Different from existing methods [8], [10], [12], [18], [22], [23], the fixed decision threshold and the fixed reference are not required in EMD-RDFCM method. Therefore, it is more adaptive in different scenarios. 2) Different from existing methods [24]–[27], this method is realized on manifold space by using the novel RDFCM clustering algorithm. Moreover, the EMD algorithm is adopted to exclude the noise component of the signals received by SUs, which can be considered as a data preprocessing process to obtain superior feature.

IV. SIMULATION ANALYSIS

In this section, the effectiveness of the EMD-RDFCM-based CSS method is verified.

A. SUs IN SAME SNR

In this part, the performance of the EMD-RDFCM-based CSS method is analyzed when SUs in same SNR. Before training the classifier, a date set $\mathcal{T} = \{T_1, \ldots, T_{2000}\}$ which contains 2000 samples is generated. The number of SU and the number of sample points are selected as $L = 2$ and $N = 1000$, respectively. Under the conditions of SNR $= [-10dB, -10dB]$, SNR $= [-12dB, -12dB]$, SNR $= [-14dB, -14dB]$, and SNR $= [-16dB, -16dB]$, the comparison results between the EMD-RDFCM-based CSS approach and other spectrum sensing methods, such as DARDMM, IQDMM, and IQRMET [26], [27], are revealed in Figures 4–7, respectively. It is obvious that the RDFCM method is better than traditional ones. Similarly, the EMD-RDFCM achieve superior detection performance than RDFCM, which means that the EMD algorithm removes the noise component effectual. The detailed data is shown in Table 2. In SNR $= [-10dB, -10dB]$, $P_f = 0.1$, compared with conventional methods, the detection probability of EMD-RDFCM approach is enhanced

TABLE 2. Detection probability of different methods when SUs in same SNR.

	EMD-RDFCM	RDFCM	DARDMM	DARIG	IQDMM
SNR = [-10dB, -10dB], P_f = 0.1	0.993	0.937	0.885	0.876	0.862
SNR = [-10dB, -10dB], P_f = 0.3	0.998	0.978	0.941	0.932	0.929
SNR = [-12dB, -12dB], P_f = 0.1	0.952	0.914	0.737	0.636	0.533
SNR = [-12dB, -12dB], P_f = 0.3	0.987	0.970	0.866	0.800	0.746
SNR = [-14dB, -14dB], P_f = 0.1	0.684	0.577	0.390	0.320	0.274
SNR = [-14dB, -14dB], P_f = 0.3	0.940	0.873	0.649	0.548	0.534
SNR = [-16dB, -16dB], P_f = 0.1	0.350	0.299	0.184	0.202	0.161
SNR = [-16dB, -16dB], P_f = 0.3	0.769	0.706	0.456	0.441	0.426

TABLE 3. Detection probability of different methods when SUs in different SNR.

|	EMD-RDFCM	RDFCM	DARDMM	DARIG	IQDMM
SNR = [-10dB, -10.5dB], P_f = 0.1	0.981	0.945	0.845	0.832	0.831
SNR = [-10dB, -10.5dB], P_f = 0.3	0.998	0.989	0.930	0.927	0.926
SNR = [-12dB, -12.5dB], P_f = 0.1	0.954	0.901	0.676	0.539	0.471
SNR = [-12dB, -12.5dB], P_f = 0.3	0.987	0.966	0.839	0.745	0.717
SNR = [-14dB, -14.5dB], P_f = 0.1	0.659	0.570	0.349	0.273	0.258
SNR = [-14dB, -14.5dB], P_f = 0.3	0.904	0.845	0.627	0.526	0.507
SNR = [-16dB, -16.5dB], P_f = 0.1	0.328	0.282	0.197	0.189	0.163
SNR = [-16dB, -16.5dB], P_f = 0.3	0.727	0.700	0.470	0.423	0.412
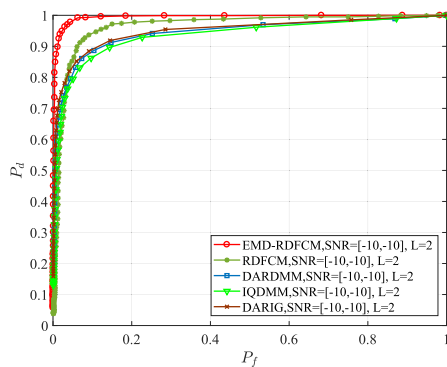
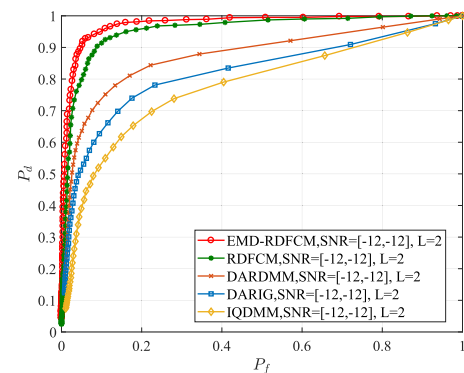**FIGURE 4.** The ROC curves of different methods when SNR = [−12dB, −12dB].

FIGURE 5. The ROC curves of different methods when SNR = [−16dB, −16dB].

by 5.92%, 12.20%, 13.35%, and 15.19%, respectively. In SNR = [−12dB, −12dB], P_f = 0.1, the detection probability of EMD-RDFCM approach is increased by 4.15%, 29.17%, 49.68%, and 78.61%, respectively. Figure 12 shows the convergence of the loss value for EMD-RDFCM approach at SNR = [−12dB, −12dB]. It means that the developed algorithm is stable.

B. SUs IN DIFFERENT SNR
In this section, the effectiveness of the EMD-RDFCM-based CSS method is demonstrated in the condition of

SUs in different SNR. As can be seen from Figures 8–11, compared with traditional methods, the EMD-RDFCM approach can obtain the best detection results. The detection probability of different methods are provided in Table 3. In condition of SNR = [−10dB, −10.5dB], P_f = 0.1, we can calculate that the detection probability of EMD-RDFCM method is 0.981. However, the detection probability of RDFCM, DARDMM, DARIG, and IQDMM are 0.945, 0.845, 0.832 and 0.831, respectively. In condition of SNR = [−12dB, −12.5dB], P_f = 0.1, the detection probability of EMD-RDFCM method is 0.954. However, the
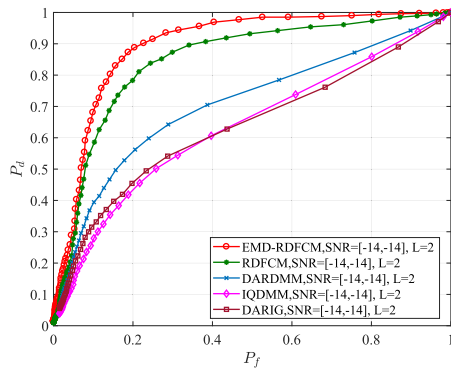
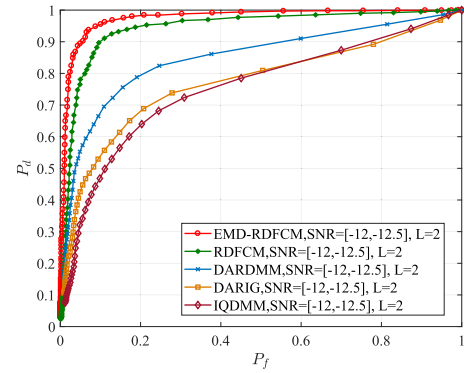**FIGURE 6.** The ROC curves of different methods when SNR = [−12dB, −12.5dB].

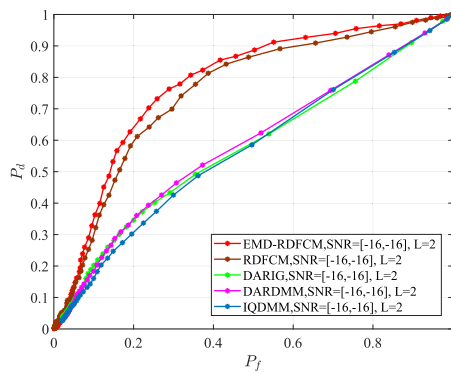**FIGURE 9.** The ROC curves of different sampling number.

FIGURE 7. The ROC curves of different methods when SNR = [−16dB, −16.5dB].
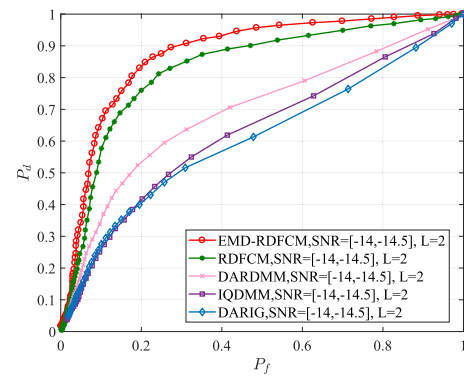

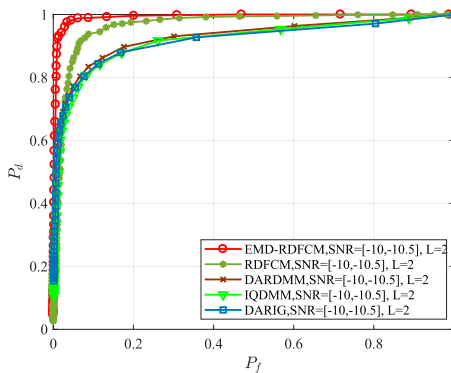**FIGURE 10.** The ROC curves of different methods when SNR = [−14dB, −14.5dB].

FIGURE 8. The Loss value of EMD-RDFCM method at SNR = [−12dB, −12.5dB].
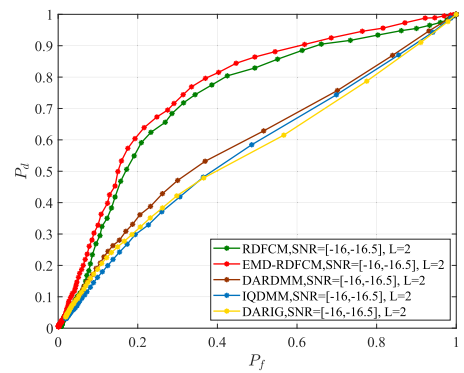

FIGURE 11. The ROC curves of different methods when SNR = [−16dB, −16.5dB].

detection probability of the other methods are 0.901, 0.676, 0.539 and 0.471, respectively. From section A and B, we can conclude that the EMD-RDFCM method can obtain the optimal sensing performance than the rest of approaches under two different scenarios. Under the condition of SNR = [−12dB, −12.5dB], the convergence of the loss value for the EMD-RDFCM method is shown in Figure 13.

C. DIFFERENT NUMBER OF SUs
In this portion, the sensing performance of the EMD-RDFCM method is analyzed under different number of SUs. The ROC curves and detection probability of different methods are displayed in Figure 14 and Table 4, respectively. It is clear that the detection performance is better as the number of the SUs increase.
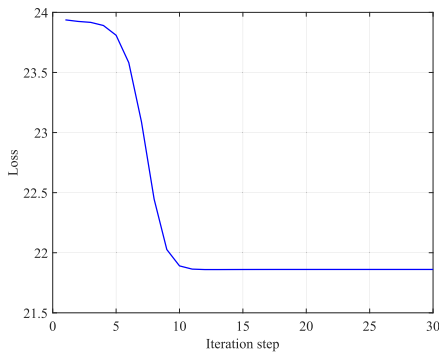
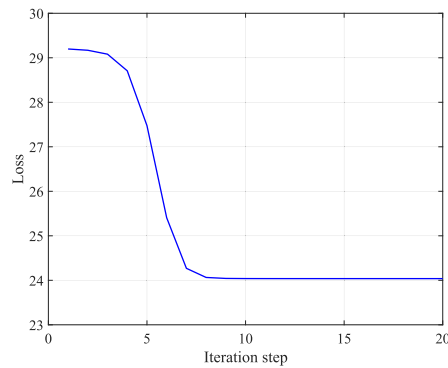**FIGURE 12.** The Loss value of EMD-RDFCM method at SNR = [−12dB, −12dB].

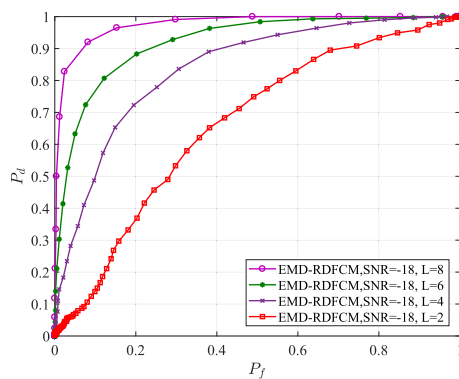**FIGURE 13.** The Loss value of EMD-RDFCM method at SNR = [−12dB, −12.5dB].

FIGURE 14. The ROC curves under different number of SUs.

TABLE 4. Detection probability under different number of SUs.

	L = 2	L = 4	L = 6	L = 8
SNR = [-18dB, -18dB], $P_f = 0.1$	0.539	0.676	0.901	0.954
SNR = [-18dB, -18dB], $P_f = 0.3$	0.745	0.839	0.966	0.987

D. DIFFERENT NUMBER OF SAMPLING POINT
In this portion, the availability of the EMD-RDFCM approach is verified under different number of sampling point. The ROC curves of EMD-RDFCM approach at
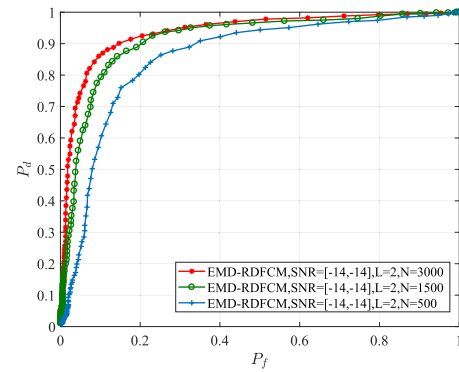

FIGURE 15. The ROC curves of different sampling number.

TABLE 5. Detection probability of different sampling number.

	N = 500	N = 1500	N = 3000
SNR = [-14dB, -14dB], $P_f = 0.1$	0.593	0.789	0.864
SNR = [-14dB, -14dB], $P_f = 0.3$	0.882	0.942	0.949

different sampling point are revealed at Figure 15. The detailed detection probability is provided in Table 5. According to the experimental results, the sensing performance is improved as the number of the sampling point increase.

V. CONCLUSION
In this paper, an EMD-RDFCM approach is developed to address the spectrum sensing problem in CR. To enhance the detection precision, the EMD algorithm is adopted to wipe off the noise portion of the received signal. A data set which contains the covariance matrix of the reconstructed signal is prepared to train a classifier. Then, an RDFCM clustering algorithm is proposed to cluster samples on manifold space. After the classifier is obtained, we use it to decide the state of the channel. Finally, the simulation part verifies the effectiveness of the EMD-RDFCM approach under two different conditions, i.e., the SUs in same SNR and the SUs in different SNR. The major contribution of this paper is on proposing an RDFCM clustering algorithm to obtain a classifier on manifold. Furthermore, the developed EMD-RDFCM approach obtains the better sensing performance in contrast to traditional methods. In future work, the delay, delivery ratio, fairness index and efficiency of the CR will be considered and extend the developed CSS method to real-world scenarios.

REFERENCES
[1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
[2] P. Qi, Y. Du, D. Wang, and Z. Li, "Wideband spectrum sensing based on bidirectional decision of normalized spectrum for cognitive radio networks," *IEEE Access*, vol. 7, pp. 140833–140845, 2019.

[3] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, part I: Two user networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2204–2213, Jun. 2007.

[4] A. Mariani, S. Kandeepan, and A. Giorgetti, "Periodic spectrum sensing with non-continuous primary user transmissions," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1636–1649, Mar. 2015.

[5] J. Zhuang, Y. Wang, S. Zhang, P. Wan, and C. Sun, "A multi-antenna spectrum sensing scheme based on main information extraction and genetic algorithm clustering," *IEEE Access*, vol. 7, pp. 119620–119630, 2019.

[6] I. Mustapha, A. Mohd, B. Ali, A. Sali, M. Rasida, and H. Mohamad, "An energy efficient reinforcement learning based cooperative channel sensing for cognitive radio sensor networks," *Pervas. Mobile Comput.*, vol. 35, pp. 165–187, Feb. 2017.

[7] A. Mohamedou, A. Sali, B. Ali, M. Othman, and H. Mohamad, "Bayesian inference and fuzzy inference for spectrum sensing order in cognitive radio networks," *Trans. Emerg. Telecommun. Technol.*, vol. 28, no. 1, p. e2916, Jan. 2017.

[8] S. Atapattu, C. Tellambura, and H. Jiang, "Energy detection based cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1232–1241, Apr. 2011.

[9] J.-Y. Wu, C.-H. Wang, and T.-Y. Wang, "Performance analysis of energy detection based spectrum sensing with unknown primary signal arrival time," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1779–1784, Jul. 2011.

[10] F. Salahdine, H. E. Ghazi, N. Kaabouch, and W. F. Fihri, "Matched filter detection with dynamic threshold for cognitive radio networks," in *Proc. Int. Conf. Wireless Netw. Mobile Commun. (WINCOM)*, Marrakech, Morocco, Oct. 2015, pp. 1–6.

[11] X. Zhang, R. Chai, and F. Gao, "Matched filter based spectrum sensing and power level detection for cognitive radio network," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Atlanta, GA, USA, Dec. 2014, pp. 1267–1270.

[12] P. Sepidband and K. Entesari, "A CMOS spectrum sensor based on Quasi-cyclostationary feature detection for cognitive radios," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 12, pp. 4098–4109, Dec. 2015.

[13] K. Choi, W. Jeon, and D. Jeong, "Sequential detection of cyclostationary signal for cognitive radio systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4480–4485, Sep. 2009.

[14] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 1st Quart., 2009.

[15] Y. Arjoune and N. Kaabouch, "A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions," *Sensors*, vol. 19, no. 1, p. 126, Jan. 2019.

[16] C.-H. Liu, A. Azarfar, J.-F. Frigon, B. Sanso, and D. Cabric, "Robust cooperative spectrum sensing scheduling optimization in multi-channel dynamic spectrum access networks," *IEEE Trans. Mobile Comput.*, vol. 15, no. 8, pp. 2094–2108, Aug. 2016.

[17] X. Liu, M. Jia, Z. Na, W. Lu, and F. Li, "Multi-modal cooperative spectrum sensing based on Dempster–Shafer fusion in 5G-based cognitive radio," *IEEE Access*, vol. 6, pp. 199–208, 2018.

[18] C. Liu and M. Yang, "A distance-weighed algorithm based on maximum-minimum eigenvalues for cooperative spectrum sensing," in *Proc. 7th Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Sep. 2011, pp. 1–4.

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

[20] V. Abed and A. Shahzadi, "A maximum-minimum eigenvalue detection simpler method based on secondary users locations for cooperative spectrum sensing," in *Proc. Int. Congr. Technol., Commun. Knowl. (ICTCK)*, Nov. 2014, pp. 1–4.

[21] K. Du, P. Wan, Y. Wang, X. Ai, and H. Chen, "Spectrum sensing method based on information geometry and deep neural network," *Entropy*, vol. 22, no. 1, p. 94, Jan. 2020.

[22] Q. Lu, S. Yang, and F. Liu, "Wideband spectrum sensing based on Riemannian distance for cognitive radio networks," *Sensors*, vol. 17, no. 4, pp. 661–679, Mar. 2017.

[23] Q. Chen, P. Wan, and Y. Wang, "Research on cognitive radio spectrum sensing method based on information geometry," in *Proc. Int. Conf. Cloud Comput. Secur.*, 2017, pp. 554–564.

[24] K. M. Thilina, K. Won Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative-spectrum sensing in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2209–2221, Nov. 2013.

[25] Y. Lu, P. Zhu, D. Wang, and M. Fattouche, "Machine learning techniques with probability vector for cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Doha, Qatar, Apr. 2016, pp. 1–6.

[26] Y. Zhang, P. Wan, S. Zhang, Y. Wang, and N. Li, "A spectrum sensing method based on signal feature and clustering algorithm in cognitive wireless multimedia sensor networks," *Adv. Multimedia*, vol. 2017, no. 4, pp. 1–10, Oct. 2017.

[27] Y. Wang, S. Zhang, and Y. Zhang, "A cooperative spectrum sensing method based on empirical mode decomposition and information geometry in complex electromagnetic environment," *Complexity*, vol. 2019, Feb. 2019, Art. no. 5470974.

[28] S. Zhang, Y. Wang, P. Wan, Y. Zhang, and X. Li, "A cooperative spectrum sensing method based on clustering algorithm and signal feature," in *Proc. Int. Conf. Cloud Comput. Secur.*, 2018, pp. 50–62.

[29] A.-O. Boudraa and J.-C. Cexus, "EMD-based signal filtering," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 6, pp. 2196–2202, Dec. 2007.

[30] L.-Q. Zuo, H.-M. Sun, Q.-C. Mao, X.-Y. Liu, and R.-S. Jia, "Noise suppression method of microseismic signal based on complementary ensemble empirical mode decomposition and wavelet packet threshold," *IEEE Access*, vol. 7, pp. 176504–176513, 2019.

[31] Y. Kopsinis and S. McLaughlin, "Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 1–13, Jan. 2008.

[32] S.-I. Amari, "Information geometry on hierarchy of probability distributions," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1701–1711, Jul. 2001.

[33] L. Groll and J. Jakel, "A new convergence proof of fuzzy c-Means," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 5, pp. 717–720, Oct. 2005.

[34] F. Zhao, Y. Chen, H. Liu, and J. Fan, "Alternate PSO-based adaptive interval type-2 intuitionistic fuzzy C-Means clustering algorithm for color image segmentation," *IEEE Access*, vol. 7, pp. 64028–64039, 2019.

[35] J. K. Parker and L. O. Hall, "Accelerating fuzzy-C means using an estimated subsample size," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 5, pp. 1229–1244, Oct. 2014.

[36] X. Yang, G. Zhang, J. Lu, and J. Ma, "A kernel fuzzy C-means clustering-based fuzzy support vector machine algorithm for classification problems with outliers or noises," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 1, pp. 105–115, Feb. 2011.

YONGWEI ZHANG received the B.S. degree from the School of Electronic and Information Engineering, Jiaying University, Meizhou, China, in 2016. He is currently pursuing the Ph.D. degree with the School of Automation, Guangdong University of Technology, Guangzhou, China. His current research interests include cognitive radio, spectrum sensing, adaptive dynamic programming, and optimal control.

CANFENG MA received the B.S. degree from Wuyi University, in 2019. He is currently pursuing the M.S. degree with the Guangdong University of Technology, Guangdong, China. His research interest includes spectrum sensing in cognitive radio.

YONGHUA WANG (Member, IEEE) received the B.S. degree in electrical engineering and automation from the Hebei University of Technology, in 2001, the M.S. degree in control theory and control engineering from the Guangdong University of Technology, in 2006, and the Ph.D. degree in communication and information system from Sun Yat-sen University, in 2009. He is currently with the School of Automation, Guangdong University of Technology.

PIN WAN received the B.S. degree in electronic engineering and the M.S. degree in circuit and system from Southeast University, in 1984 and 1990, respectively, and the Ph.D. degree in control theory and control engineering from the Guangdong University of Technology, in 2011. He is currently a Professor with the School of Automation, Guangdong University of Technology.

• • •

SHUNCHAO ZHANG received the B.S. degree from the School of Electrical and Information Engineering, Hunan Institute of Engineering, Xiangtan, China, in 2016, and the M.S. degree from the School of Automation, Guangdong University of Technology, Guangzhou, China, in 2019, where he is currently pursuing the Ph.D. degree. His current research interests include cooperative spectrum sensing, clustering algorithms, and adaptive dynamic programming.