

Dynamic Evidential Clustering Algorithm

Zuowei Zhang, Zhe Liu, Arnaud Martin, Zhun-Ga Liu, Kuang Zhou

▶ **To cite this version:**

Zuowei Zhang, Zhe Liu, Arnaud Martin, Zhun-Ga Liu, Kuang Zhou. Dynamic Evidential Clustering Algorithm. Knowledge-Based Systems, 2021, 213. hal-03080938

HAL Id: hal-03080938
https://hal.science/hal-03080938

Submitted on 18 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dynamic Evidential Clustering Algorithm

Zuo-wei Zhang[a,b], Zhe Liu[c], Arnaud Martin[b], Zhun-ga Liu[a], Kuang Zhou[d]
a. School of Automation, Northwestern Polytechnical University, Xi'an, China.
b. Univ Rennes, CNRS, IRISA, Rue E. Branly, 22300 Lannion, France.
c. School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, China.
d. Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an, China.

***Abstract*—In this paper, a dynamic evidential clustering algorithm (DEC) is introduced to address the computational burden of existing methods. To derive such a solution, an FCM-like objective function is first employed and minimized to obtain the support levels of the real singletons (specific) clusters to which the query objects belong, and then the query objects isinitially adaptively assigned to outlier, precise or imprecise one via a new rule-based on the conflicts between the different support levels. For each imprecise object, it is finally reassigned to the singleton clusters or related meta-cluster by partial credal redistribution with the corresponding dynamic edited framework to reduce the computational burden. The proposed method can reduce the complexity to the level similar to that of the fuzzy and possibilistic clustering, which can effectively extend the application of evidential clustering, especially in big data. The effectiveness of the DEC method is tested by four experiments with artificial and real datasets.**

Keywords: Dynamic evidential clustering, credal partition, uncertainty, belief functions, unsupervised classification.

I. Introduction

Clustering is one of the important branches of unsupervised pattern recognition and machine learning, and it has been widely used in various fields [1] including financial analysis, medical diagnosis, image cutting and information fusion, etc. The goal of clustering is to group a set of n objects into c clusters with the edited framework $\Omega = \{\omega_1, ...\omega_c\}$ where the members (objects) in the same cluster are similar in some way. A variety of methods have been developed for clustering object and relational data [2]-[4]. In general, the object data is described with p-dimensional attributes while the relational data is described by similarity or dissimilarity between objects. This paper mainly focuses on object data.

There exists a number of clustering methods designed for object data, and they can be broadly classified into two main families: hard and soft clustering [5]-[7]. Hard clustering allows each object to belong to only one cluster, *i.e.* the nearest cluster to the object will have it completely, and C-means algorithm [5] can be regarded as the representative of hard clustering. In soft clustering, the object will be assigned to different clusters with various levels of support, where the most popular fuzzy c-means algorithm (FCM) [6], [8] is the representative. However, some works [9]$-$[13] point out that it may produce some phenomena contrary to intuition sometimes although FCM has achieved good results, and they have given some solutions. In [9], for instance, a possibilistic (PCM) clustering algorithm is introduced by modifying the objective function to be minimized. A possibilistic fuzzy c-means clustering algorithm (PFCM) is also proposed in the literature [11]. The shortcomings of fuzzy partition are analyzed in detail and new solutions are put forward. In [12], an improved method of PFCM is proposed and the advantages in MRI image segmentation are analyzed. It mainly includes two aspects: considering the relationship between clusters denoted by intraclass and interclass similarity, and using local label information to supplement spatial constraints. To reasonably characterize the outlier, the concept of noise clustering (NC) [14], [15] is given to make fuzzy and possibilistic partition more robust via adding a "noise" cluster denoted as $\emptyset$. Once the object is assigned to the noise cluster, it indicates that the object is too far away from all singleton (specific) clusters, and naturally regarded as noise or outlier. That is, there are $c+1$ clusters and the object will be assigned to the noise cluster if the minimum distance between the object and the c known cluster centers is greater than a given distance δ. The parameter δ thereby controls the number of outliers.

Recently, a new evidential c-means (ECM) [3] clustering method, based on the notion of "credal partition" [16], [17], is proposed under the theoretical framework of belief functions [18]$-$[24], which can reasonably model a few objects, named imprecise objects[1]. In such case, it will greatly increase the risk of errors if they are forced to be assigned to singleton clusters. Specifically, credal partition extends the traditional probability framework $\Omega = \{\omega_1, ..., \omega_c\}$ to the power-set 2^Ω, and this additional flexibility allows us to gain a deeper insight into the data and to improve robustness with respect to outliers [3], [25]. For instance, $\Omega = \{\omega_1, \omega_2, \omega_3\}$ with $c = 3$, then $2^\Omega = \{\emptyset, \omega_1, \omega_2, \omega_3, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}, \Omega\}$. In other words, credal partition can produce three kinds of cluster: singleton (specific) clusters (*e.g.* ω_1), meta-clusters (*e.g.* $\{\omega_2, \omega_3\}$) defined by the disjunction (union) of several singleton clusters, and the noise cluster (outlier) represented by $\emptyset$. The meta-cluster is also regarded as a separate cluster and can be considered as a transition cluster among these included

[1]In the clustering problem, one object simultaneously close to several singleton (specific) clusters can be difficult to correctly classify since these close singleton (specific) clusters appear not distinguishable for this object. In such a case, we consider that the object can belong to the union of these singleton clusters, called meta-cluster, and considered as a new cluster, we thereby say that the object is imprecise. In other words, we have an imprecise object if the object is assigned to the meta-cluster. In contrast, we can say a precise object if we are sure that the object belongs to one singleton cluster only, *i.e.* the object is obviously close to one and only one singleton cluster. The object, thereby, is naturally assigned to this singleton cluster.

singleton clusters. Once assigned into meta-cluster, it indicates that the object may belong to any singleton clusters included in the meta-cluster and it is regarded as an imprecise one.

In ECM, the object can be assigned not only to singleton clusters including the noise cluster but also to any meta-clusters in 2^{Ω} depending on the "mass of beliefs", denoted as $m(\cdot)$, which is similar to the membership degree in FCM. The ECM thereby can be regarded as an evidential version of the FCM and NC. The difference, however, is that the ECM introduces the notion of meta-cluster A_j and defines the center $\overline{\boldsymbol{v}}_j$ as the average value of that of the singleton clusters included in the meta-cluster A_j. The distance d_{ij} between the object $\boldsymbol{x}_i$ and the center $\overline{\boldsymbol{v}}_j$ is taken as the main criterion to judge whether $\boldsymbol{x}_i$ belongs to the A_j or not. Whereas it may produce unreasonable results if the meta-cluster center is close to the centers of some singleton clusters. Therefore, a credal c-means (CCM) clustering method [25] is proposed to overcome this limitation. In CCM, the object should be not only close to the center of the meta-cluster but also close to the associated singleton clusters included in the meta-cluster if the object is assigned to the meta-cluster. By doing this, the CCM [25] can produce good results even if the center of meta-cluster is close to that of the singleton clusters.

Since credal partition has the prominent advantage in characterizing imprecision and uncertainty, it has been extended to many fields [26]-[28]. These methods, however, are time-consuming because they spend a lot of meaningless calculations. This seriously restricts the scope of applications because we are now in an era of big data. As illustrated in Fig. 1, the

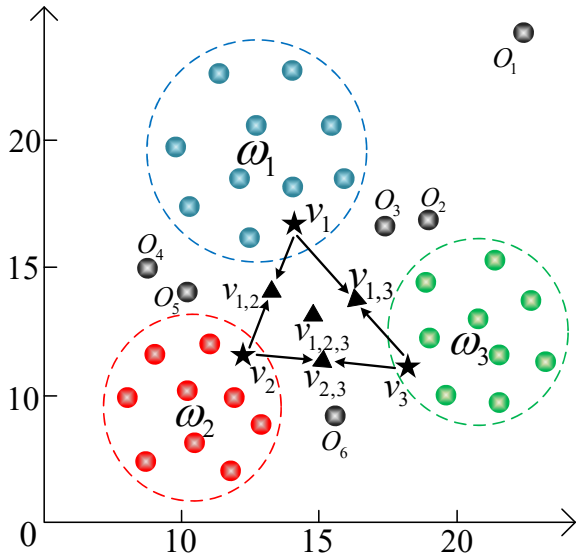


Figure 1: Illustration of the complexity in credal partition.

centers (*i.e.* black triangles) of meta-clusters are determined by the centers (*i.e.* black pentagrams) of related singleton clusters in credal partition. In other words, if the centers of singleton clusters are not real, that of meta-clusters is also unreliable, for instance, $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$ are only temporary cluster centers, not the real centers. Thus, the centers (*e.g.* $\boldsymbol{v}_{1,2}$ and $\boldsymbol{v}_{2,3}$) of meta-clusters are also not exactly in such a case, as shown in Fig. 1. Therefore, it doesn't make much sense to calculate the distances between the object and (large number of) meta-clusters in each iteration (except the last one). In addition, one can find that only a few imprecise objects (*e.g.* O_2, O_3) are difficult to accurately be assigned to the singleton cluster (*e.g.* ω_1 or ω_3), and they are only in a few related clusters. If one assigns each object under the power-set 2^{Ω}, it will also bring a lot of invalid calculations. For example, one only needs to assign the object O_5 under the sub-frame $\{\omega_1, \omega_2, \omega_{1,2}\}$ and O_6 under the sub-frame $\{\omega_2, \omega_3, \omega_{2,3}\}$, where $\omega_{i,k} \triangleq \{\omega_i, \omega_k\}$ is denoted for conciseness.

The ECM and CCM also try to reduce the computational complexity. In general, ECM can generate 2^c clusters, for example, for the framework of discernment $\Omega = \{\omega_1, ..., \omega_c\}$. However, if there are too many elements in the pow-set 2^{Ω}, it will cause a huge execution burden since the complexity of ECM, *i.e.* $\mathcal{O}(n \cdot 2^c)$, will increase exponentially with the increase of c. Thus, the authors consider limiting the cardinality of the elements in the meta-cluster to 2 (except for Ω), which can better reduce the computational complexity of ECM. Of course, this hypothesis is applicable in most cases, but it may produce unreasonable results in a few cases, for example, some objects may be indistinguishable among more than two but less than c singleton clusters. The CCM adopts a strategy similar to ECM.

To reduce complexity of credal partition as far as possible, a method, called dynamic evidential clustering (DEC), is proposed in this paper, which contains two steps: 1) preliminary adaptive credal partition; 2) partial credal redistribution. The main contributions lie in the following points.

1) An FCM-like objective function is employed to obtain the support degrees of each object belonging to the real singleton (specific) clusters, which are then used to preliminary assign the object as the outlier, precise or imprecise one based on a given adaptive rule.

2) A CCM-like objective function is then minimized to reassign the imprecise object to one singleton cluster or related meta-cluster again by partial credal redistribution with the corresponding dynamic edited framework. In other words, the specific dynamic edited framework will replace the power-set 2^{Ω} to reassign each imprecise object.

This paper is organized as follows. The basics of belief functions is briefly introduced in Section II. The dynamic evidential clustering algorithm (DEC) is detailed in Section III. Four experiments using artificial and real datasets are given in Section IV to illustrate the validity of the DEC with other respectable methods. Managerial implication is provided in Section V, followed by Conclusions.

II. Brief Review of Belief Functions

The theory of belief functions, also called Dempster-Shafer theory (DST) [18]-[20], which has also been widely used in classification [29]-[31] and decision-making [23], [32]-[35]. In belief functions, a set of finite mutually exclusive and exhaustive elements $\Omega = \{\omega_1, \omega_2, \cdots, \omega_c\}$ is defined as the framework of discernment of the problem under study. The set of all subsets of Ω is called the power-set of Ω, which is represented as 2^{Ω} and contains $2^{|\Omega|}$

elements. For example, if $\Omega = \{\omega_1, \omega_2, \omega_3\}$, then $2^\Omega = \{\emptyset, \omega_1, \omega_2, \omega_3, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}, \Omega\}$. The singleton cluster (*e.g.* ω_i) is also called specific cluster. The disjunction of several singleton elements (*e.g.* $\{\omega_1, \omega_2\}, \{\omega_2, \omega_3\}$) represents partial ignorance, and they are called meta-clusters.

Credal partition based on the theory of belief functions has become one of the research hotspots since it extends the existing concepts of hard, fuzzy (probabilistic), and possibilistic partition by allocating, for each object, not only to singleton clusters but also to meta-clusters. This additional flexibility allows us to gain a deeper insight in the data and to improve robustness with respect to outliers [3]. A credal partition [3], [16] is defined as the n-tuple $M = (m_1, \ldots, m_n)$, where m_i is the basic belief assignment (BBA) of the object $\boldsymbol{x}_i \in \mathcal{X}$, $i = 1, 2, \ldots, n$ associated with the different elements of the power-set 2^Ω. The BBA or mass of belief on the framework of discernment Ω is a function $m(.)$ from 2^Ω to $[0, 1]$, and satisfies the following conditions:

$$\begin{cases} \sum\limits_{A \in 2^\Omega} m(A) = 1 \\ m(\emptyset) = 0 \end{cases} \tag{1}$$

All the elements $A \in 2^\Omega$ such that $m(A) > 0$ are called the focal elements of $m(.)$. In credal partition, the $m(.)$ is considered to be related to the distance between the object and different clusters, and the mass of belief that an object $\boldsymbol{x}_i$ belongs to a meta-cluster depends not only on the distance between $\boldsymbol{x}_i$ and the center of the meta-cluster (*i.e.* the mean value of the centers of the singleton clusters contained in the meta-cluster), but also on the distance between the object $\boldsymbol{x}_i$ and the prototype vector of the singleton clusters included in the meta-cluster [25]. Since the credal partition is regarded as the direct extension of FCM under belief functions, the convergence is similar to that of FCM.

In credal partition, the belief function $Bel(.)$ and plausibility function $Pl(.)$, corresponding to the lower and upper bounds of imprecise probability associated with BBAs, can also be used for decision-making support when adopting pessimistic or optimistic attitudes if necessary, and they are given for all $A \in 2^\Omega$ by:

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{2}$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \tag{3}$$

In this paper, one of them can be employed as criteria for preliminary credal partition, since they have a straight corresponding relationship in such particular BBAs structure.

III. Dynamic Evidential Clustering

To reduce the computational complexity of existing evidential clustering methods, a new alternative version, named dynamic evidential clustering (DEC), is proposed in this section. It is based on the following two assumptions:

1) For the same query set, the centers obtained in the fuzzy (possibilistic) partition and that of singleton clusters obtained by the credal partition are very similar. This means that the meta-clusters can be ignored in the initial iterations because the centers of meta-clusters are defined based on the instant information of the related singleton clusters.

2) Only a part of objects in the query set is difficult to be accurately assigned to singleton (specific) clusters. They are then assigned to the related meta-clusters composed of only several close singleton clusters. Thus, it is not necessary to expose all the objects under the power-set 2^Ω.

By the above assumptions, the DEC method can be summarized as two steps: 1) preliminary credal partition; 2) partial credal redistribution.

A. Preliminary credal partition

The purpose of this subsection is to preliminary assign each object in the query set as the outlier, precise or imprecise one adaptively. To derive such a proposal, let's consider a query set $\mathcal{X}$ including n objects in p-dimensions with $\Omega = \{\omega_1, ..., \omega_c\}$. The support degrees of each object belonging to different singleton (specific) clusters and the noise cluster, called the mass of beliefs in credal partition, can be minimized by an FCM-like objective function at first. There exist many methods to obtain the mass of beliefs. For example, the noise clustering method [14] can be applied for the query set, and we have modified it as the version of the credal partition to facilitate the presentation. The objective function can be expressed as follows:

$$J_{DEC-NC}(M_1, V_1) = \sum_{i=1}^{n} \sum_{j=1}^{c} m_{ij}^{\beta} \cdot d_{ij}^2 + \sum_{i=1}^{n} \delta^2 \cdot {m_{i\emptyset}}^{\beta} \tag{4}$$

with the constraint:

$$\sum_{j=1}^{c} m_{ij} + m_{i\emptyset} = 1, \quad \forall i = 1, n \tag{5}$$

where $M_1 = (m_1, ..., m_n) \in \mathbb{R}^{n \times (|\Omega|+1)}$ is the mass of belief matrix for n objects in $\mathcal{X}$, and $V_1 \in \mathbb{R}^{c \times p}$ is the matrix of the centers of singleton clusters. d_{ij} is the Euclidean distance between the object $\boldsymbol{x}_i$ and the center of singleton cluster ω_j. Parameters β, δ are adjustable with the same meanings as those in NC [14], ECM [3] and CCM [25]. Since it has the same structure as NC, we directly give the mass of beliefs of the query object $\boldsymbol{x}_i$ belonging to different clusters (*i.e.* the noise and singleton clusters) as follows:

$$m_{ij} = \frac{d_{ij}^{-2/(\beta-1)}}{\sum\limits_{k=1}^{c} d_{ik}^{-2/(\beta-1)} + \delta^{-2/(\beta-1)}}, \tag{6}$$

and

$$m_{i\emptyset} = 1 - \sum_{j=1}^{c} m_{ij}, \quad \forall i = 1, n. \tag{7}$$

where m_{ij} is the mass of belief (*i.e.* the support degree) of the object $\boldsymbol{x}_i$ belonging to the cluster ω_j, and $m_{i\emptyset}$ represents the possibility of belonging to the noise cluster (*i.e.* $\emptyset$).

By doing this, we can assign the query object to different clusters according to the mass of belief matrix. Here the maximum of belief function $Bel(.)$ given in (2) is used as

the criteria for the decision-making of the cluster which is strongly supported by the various mass of beliefs. Specifically, the object $\boldsymbol{x}_i$ is directly assigned to the noise cluster if $Bel(\emptyset)$ is larger than the others, and defined as:

$$Bel_i(\emptyset) > Bel_i(\omega_j),\ j = 1, ..., c \tag{8}$$

where $Bel_i(\emptyset) = m_{i\emptyset}$ and $Bel_i(\omega_j) = m_{ij}$ since $Bel(.)$ has a straight corresponding relationship in such particular BBAs structure [31]. In such a case, we don't further assign the object since it is obviously considered as the outlier. Whereas if the cluster ω_j is the most believed cluster for $\boldsymbol{x}_i$, defined by:

$$Bel_i(\omega_j) = \max[Bel_i(\omega_1), ..., Bel_i(\omega_c)]. \tag{9}$$

one needs to further judge the possibility that the object belongs to ω_j, because $Bel_i(\omega_j)$ maybe not significantly different from others sometimes. A numerical example is given to explicitly illustrate this.

Table I: The numerical example.

	$\emptyset$	ω_1	ω_2	ω_3
m_1	1.0	0	0	0
m_2	0	0.9	0	0.1
m_3	0	0.01	0.5	0.49
m_4	0	0.43	0.42	0.15

Example 1: The 4-tuple $M = (m_1,\ m_2,\ m_3,\ m_4)$ in Table I is an example of preliminary credal partition. One can see that the objects $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ likely belong to the outlier $\emptyset$ and the singleton cluster ω_1, respectively. In contrast, the objects $\boldsymbol{x}_3$ and $\boldsymbol{x}_4$ will be assigned to the ω_2 and ω_1 respectively from the probabilistic perspective, but it has the risk of error since this slight difference may be caused by multiple reasons. In fact, the objects like $\boldsymbol{x}_3$ and $\boldsymbol{x}_4$ need to be further assigned, because they are likely to be reasonably assigned to the meta-clusters $\{\omega_2, \omega_3\}$ and $\{\omega_1, \omega_2\}$ respectively in credal partition.

In practice, it can also happen that the belief $Bel_i(\omega_s)$ of the other cluster $\omega_{\rm s}$ ($s \in [1, c],\ s \neq j$) can be very close (or equal) to the $Bel_i(\omega_j)$ of the strongest cluster ω_j. In such a case, the object can also potentially belong to ω_s with a high likelihood, and we should adopt a more cautious strategy in the preliminary credal partition. That is, we need to consider all the very likely specific clusters as the potential solution for $\boldsymbol{x}_i$. The set of these potential singleton clusters can be called $\mathcal{M}_i$ and it is defined by:

$$\mathcal{M}_i = \{\omega_j \cup, ..., \cup \omega_s\},\ Bel_i(\omega_j) - Bel_i(\omega_s) \leq \varphi \tag{10}$$

where $\varphi \in [0, 1]$ is a chosen meta-cluster threshold. In fact, φ is a very interesting threshold, which determines the several frameworks of clustering as follows.

Fuzzy/possibilistic partition: We have a fuzzy/possibilistic partition with $\varphi = 0$ since all objects will be assigned directly to specific (singleton) clusters in the Bayesian framework. We don't need to partial credal redistribution for imprecise objects since there is no meta-cluster at this time. In contrast, the query set will be clustered under credal partition with $\varphi > 0$.

Traditional credal partition: We have a traditional credal partition with $\varphi = 1$, which means that all objects are imprecise ones in preliminary credal partition. All query objects need to be credal redistributed under the power-set 2^{Ω}.

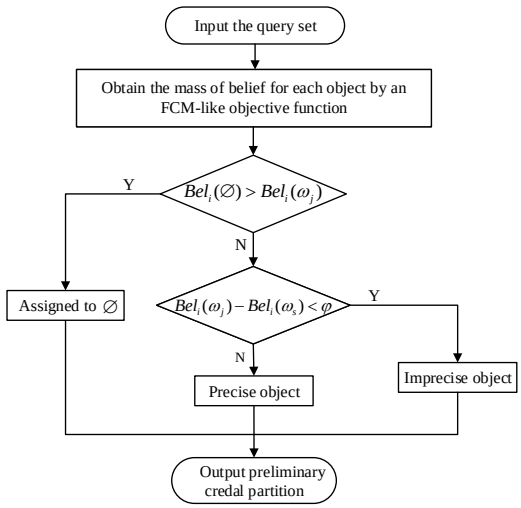


Figure 2: Flowchart of the adaptive credal partition.

Dynamic credal partition: We have a dynamic (partial) credal partition with $\varphi \in (0, 1)$ in this paper because we consider that only a part of objects in the query set is difficult to be accurately assigned to singleton clusters.

Based on the above analysis, all clusters in $\mathcal{M}_i$ may very likely correspond to the real (unknown) clusters for $\boldsymbol{x}_i$, and they appear indistinguishable with respect to the threshold φ. That is, the imprecise object $\boldsymbol{x}_i$ needs to be further assigned under the new frame $\mathcal{M}_i^{\Omega}$, and defined as follows:

$$\mathcal{M}_i^{\Omega} = \{\omega_j, ..., \omega_s, \{\omega_j \cup \omega_s\}, ..., \mathcal{M}_i\} \tag{11}$$

By doing this, one can easily find that only a few objects need to be further reassigned. It can greatly reduce the computational complexity, and each imprecise object has a specific dynamic edited framework. A simple example is given to explicitly illustrate this.

Example 2: Let us consider a chosen meta-cluster threshold $\varphi = 0.1$, and the objects $\boldsymbol{x}_3$, $\boldsymbol{x}_4$ in Table I are employed here to illustrate the specific dynamic edited framework. For the object $\boldsymbol{x}_3$, one can obtain that ω_2 is the most likely cluster with $Bel_3(\omega_2) = \max[Bel_3(\omega_1), Bel_3(\omega_2), Bel_3(\omega_3)]$. However, it does not mean that $\boldsymbol{x}_3$ is directly assigned to the singleton cluster ω_2 since there is another cluster ($i.e.\ \omega_3$) that the object may belong to under the chosen meta-cluster threshold $\varphi = 0.1$ with $\mathcal{M}_3 = \{\omega_2 \cup \omega_3 \,|Bel_3(\omega_2) - Bel_3(\omega_3) \leq 0.1\}$. The object $\boldsymbol{x}_3$ thereby is considered as an imprecise object by preliminary credal partition, and it will be reclustered by partial credal redistribution with the specific dynamic edited framework $\mathcal{M}_3^{\Omega} = \{\omega_2, \omega_3, \{\omega_2, \omega_3\}\}$. One can also deduce that the imprecise object $\boldsymbol{x}_4$ will be reassigned under the specific dynamic edited framework $\mathcal{M}_4^{\Omega} = \{\omega_1, \omega_2, \{\omega_1, \omega_2\}\}$. The flowchart is presented in Fig. 2 to explicitly show how the preliminary credal partition works.

B. Partial credal redistribution

In DEC, the center $\overline{\boldsymbol{v}}_j$ of the meta-cluster A_j is also defined as mean value of that of the singleton clusters included in A_j, which is similar to ECM [3] and defined mathematically by:

$$\overline{\boldsymbol{v}}_j = \frac{1}{|A_j|}\sum_{k=1}^{|A_j|} s_{kj}\boldsymbol{v}_k, \text{ with } s_{kj} = \begin{cases} 1, \text{ if } \omega_k \in A_j; \\ 0, \text{ otherwise.} \end{cases} \tag{12}$$

where $\boldsymbol{v}_k$ is center of the singleton cluster ω_k, and $|A_j|$ denotes the cardinality of A_j. Whereas the distance between the object $\boldsymbol{x}_i$ and the meta-cluster A_j depends not only on the distance between the object $\boldsymbol{x}_i$ and the center $\overline{\boldsymbol{v}}_j$, but also on the distance between the object $\boldsymbol{x}_i$ and the center $\boldsymbol{v}_k$ of all the singleton clusters included in A_j. Thus, the distance D_{ij}^2 from the object to different clusters can be defined as follows:

$$D_{ij}^2 = \begin{cases} d_{ij}^2, & \text{if } |A_j| = 1, \\ \frac{\sum_{A_k \in A_j} d_{ik}^2 + d_{ij}^2}{|A_j|+1}, & \text{if } |A_j| > 1. \end{cases} \tag{13}$$

where d_{ij} is the Euclidean distance between the object $\boldsymbol{x}_i$ and the center $\overline{\boldsymbol{v}}_j$ of the cluster A_j. If A_j is a singleton cluster, $\overline{\boldsymbol{v}}_j$ is the center of A_j. In contrast, $\overline{\boldsymbol{v}}_j$ is the mean value of the included singleton cluster centers if A_j is a meta-cluster.

One can find that the distance between the object and the meta-cluster is similar to that in CCM, but they are different. In DEC, $|A_j|+1$ in Eq. (13) denotes the number of distances from the object to different clusters. Whereas $|A_j|+\gamma$ varies with the change of γ of CCM, it may lead to some specific objects being assigned to meta-clusters and then increase the number of objects in meta-clusters unreasonably. Once $\mathcal{M}_i^\Omega$ is obtained, we only need to calculate the distance between the object and the related clusters in one iteration. This is the reason that the DEC can greatly reduce the computing time and produce similar clustering results with ECM and CCM.

For the imprecise object $\boldsymbol{x}_i$ with $\mathcal{M}_i^\Omega$, we need to update its mass of beliefs $m(\cdot)$ belonging to different clusters under the edited framework $\mathcal{M}_i^\Omega$. Assuming that there are q imprecise objects after preliminary credal partition. Since the centers of singleton and meta-clusters have been obtained, we can directly update the mass of belief $m(\cdot)$ for each imprecise object. Inspired by ECM [3] and CCM [25], the update formula of mass values can be derived by minimizing the other sub-objective function denoted as follows:

$$J_{DEC-CR}(M_2, V_2) = \sum_{i=1}^{q}\sum_{j/A_j \in \mathcal{M}_i^\Omega} m_{ij}^\beta D_{ij}^2, \tag{14}$$

subject to:

$$\sum_{j/A_j \in \mathcal{M}_i^\Omega} m_{ij} = 1. \tag{15}$$

where $V_2 \in \mathbb{R}^{|\mathcal{M}^\Omega| \times p}$ is the center matrix of singleton clusters and meta-clusters, and D_{ij}^2 can be obtained from Eq. (13). Although Eq. (14) is inspired by ECM and CCM, the noise cluster is no longer considered here since noise has been well identified in preliminary credal partition.

Since V_2 is known, lagrange multipliers λ_i are used to solve the constrained minimization problem with respect to M_2 as follows:

$$\begin{aligned}\mathcal{L}(M_2, \lambda_1, \ldots, \lambda_n) =& J_{DEC-CR}(M_2, V_2) \\ &- \sum_{i=1}^{q}\lambda_i\Big(\sum_{j=1}^{c} m_{ij} - 1\Big).\end{aligned} \tag{16}$$

By differentiating the Lagrangian with respect to the m_{ij} and λ_i and setting the derivatives to zero, we obtain:

$$\frac{\partial \mathcal{L}}{\partial m_{ij}} = \beta m_{ij}^{\beta-1} D_{ij}^2 - \lambda_i = 0, \tag{17}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = \sum_{j/A_j \in \mathcal{M}_i^\Omega} m_{ij} - 1 = 0. \tag{18}$$

From Eq. (17), we thus have:

$$m_{ij} = \left(\frac{\lambda_i}{\beta}\right)^{\frac{1}{(\beta-1)}} \left(\frac{1}{D_{ij}^2}\right)^{\frac{1}{(\beta-1)}}, \tag{19}$$

Using Eqs. (18) - (19):

$$\left(\frac{\lambda_i}{\beta}\right)^{\frac{1}{(\beta-1)}} = \frac{1}{\sum\limits_{j/A_j \in \mathcal{M}_i^\Omega} D_{ij}^{\frac{-2}{(\beta-1)}}}. \tag{20}$$

Returning in Eq. (19), one obtains the necessary condition of optimality for M_2:

$$m_{ij} = \frac{D_{ij}^{\frac{-2}{(\beta-1)}}}{\sum\limits_{k/A_k \in \mathcal{M}_i^\Omega} D_{ik}^{\frac{-2}{(\beta-1)}}}. \tag{21}$$

One can easily find that only one iteration is needed to obtain the mass of beliefs for each imprecise object $\boldsymbol{x}_i$ because V_2 is known. Once the imprecise object is reassigned to a specific cluster, it indicates that the object is separable under the current knowledge. In contrast, if the object is final assigned to the meta-cluster, it indicates that the object may belong to any singleton clusters included in the meta-cluster depending on the known information.

1) Guideline for the parameters: There are three parameters involved in the DEC method: the meta-cluster threshold φ, the weighting exponent β and the outlier threshold δ. The meta-cluster threshold φ can control the number of objects assigned in meta-clusters, and the larger the φ, the more objects in meta-clusters. Thus, φ can be adjusted according to the number of objects in meta-clusters that one can accept, and $\varphi = 0.3$ is the default value. The use of the tuning parameter β is similar to ECM [3] and CCM [25], and $\beta = 2$ can be used as default value. The parameter δ is strongly dependent on the dataset to be clustered and controls the amount of data considered as outliers which has been discussed in NC [14].

2) Discussion for the complexity: In order to reduce the computational complexity of existing evidential clustering, the DEC first uses an FCM-like objective function to iteratively obtain the real centers of singleton clusters and preliminary adaptively assign the query object as the outlier, precise or imprecise one by the preliminary credal partition, and then the results of the partial credal redistribution for imprecise

objects are used as the final decision-making. Since the partial credal redistribution is only for a part of objects and only needs to be iterated once, the computational complexity of DEC mainly depends on the preliminary credal partition. Thus, the computational complexity is $\mathcal{O}(n \cdot c + q)$ which is much lower than that of ECM ($\mathcal{O}(n \cdot 2^c)$) where n is the number of the objects in the query set and q is the number of imprecise objects, and c is the real number of singleton clusters.

The pseudo-code shown in **Algorithm 1** is used to clearly explain the working principle of the DEC method.

Algorithm 1 Dynamic Evidential Clustering Algorithm

Require: Data to cluster: $\mathcal{X} = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\}$ in $\mathbb{R}^p$;
Parameters: c, β, δ, φ
Ensure: Cluster decision results
Step 1
Construct the objective function J_{DEC-NC} without meta-clusters using Eqs. (4)-(5);
Iterate the mass of beliefs for each object by Eqs. (6)-(7);
for the *1-th* to *n-th* query object
Preliminary assign the object as the outlier, precise or imprecise one using Eqs. (8)-(10);
end
Subreturn: Preliminary credal partition.
Step 2
Calculate the centers of meta-cluster using Eq. (12);
Reconstruct the objective function J_{DEC-CR} for credal redistribution using Eqs. (14)-(15);
Reiterate the mass of beliefs for the q imprecise objects using Eqs. (13), (21);
for the *1-th* to *q-th* imprecise object
Reassign the object to singleton cluster or meta-cluster.
end
Subreturn: Partial credal redistribution.
Return: Dynamic credal partition.

IV. EXPERIMENT APPLICATIONS

In this section, four experiments are conducted to evaluate the performance of the proposed DEC method compared with C-means [5], FCM [6], NC [14], ECM [3], CCM [25], BPEC [26]. *Experiment 1*, *Experiment 2*, and *Experiment 3* are all particular artificial data to illustrate the use of DEC and the limitations of FCM, ECM, CCM, and BPEC. *Experiment 4* with nine real datasets from UCI repository is presented to evaluate the performance of the DEC method.

Except for the parameters set in the experiments, the others are all defaults. Since the introduction of meta-cluster in ECM [3], the error rate and imprecision rate [25] are used as the indicators of different methods. The error rate denoted by Re (in %) is calculated by $Re = n_e/n$, where n_e is the number of clustering errors, and n is the number of objects under the test. The imprecision rate denoted by Ri (in %) is calculated by $Ri = n_i/n$, where n_i is number of objects assigned to meta-clusters. Moreover, we also evaluate the clustering resuts using Credal Rand Index (CRI) [36], which is employed to measure the closeness of credal partition and ground truth. The upper bound of CRI is 1, a larger value corresponds to a better clustering result. In order to compare the computational complexity of all algorithms more fairly, in ECM and BPEC, the authors consider to limit the number of elements in meta-clusters is equal to 2 except for Ω. In CCM, the number of focal elements in meta-cluster is limited to 2 by a given threshold $t_c = 2$. For more detailed explanation, please refer to [3], [25], [26]. In addition to the above indexes, pieces of literature [37]-[39] put forward some other optimality criteria. Here we employ the objective function value and the number of iterations to evaluate the performance of the proposed method. In the tests, we report the mean of the above-mentioned indexes based on all methods that have been run 10 times. The elapsed time denoted by T (in second, s).

A. Experiment 1

This experiment is mainly to explain the use of the DEC in clustering on two 3-class datasets clearly, which we called Gaussian Data and Round Shape Data, respectively.

(1) Gaussian Data: Let's consider a 3-class gaussian dataset, as shown in Fig. 3, to test ECM and DEC. The set has $3 \times 100 = 300$ data points from three $2D$ Gaussian distributions with the mean variance parameters as follows:

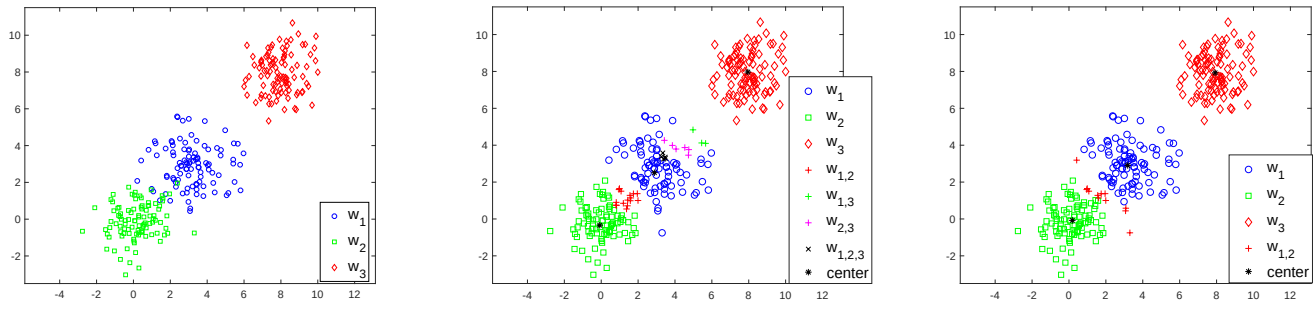
$$\mu_1 = (3, 3), \Sigma_1 = 1.5\boldsymbol{I}$$
$$\mu_2 = (0, 0), \Sigma_2 = \boldsymbol{I}$$
$$\mu_3 = (8, 8), \Sigma_3 = \boldsymbol{I}$$

where μ_i $(i = 1, 2, 3)$ is the mean vector, Σ_i $(i = 1, 2, 3)$ is the covariance matrix, and $\boldsymbol{I}$ is the identity matrix.

In ECM and DEC, the noise threshold is set as $\delta = 5$, and other parameters are the defaults. The clustering results of different methods are shown in Fig. 3 (b)-(c).

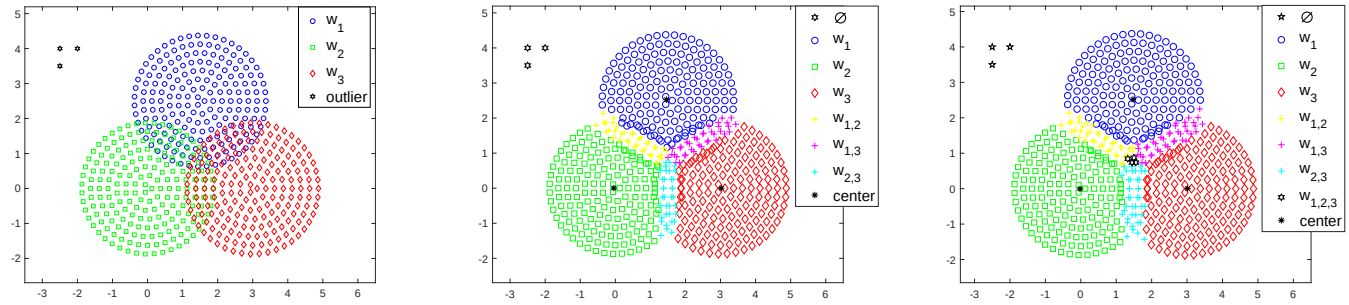
One can see from Fig. 3 (a) that there are some objects in the overlapping zones of different clusters ω_1 and ω_2, which are difficult to be correctly distinguished for the clusters. The clusters ω_2 and ω_3 are far apart, however, many of the objects belonging to ω_1 are incorrectly assigned to the meta-cluster $\omega_{2,3}$ labeled by pink plus in ECM, as shown in Fig. 3 (b). In addition, lots of objects belonging to ω_1 are assigned to the total ignorant cluster $\omega_{1,2,3}$ labeled by the black multiplier. Meanwhile, some objects belonging to ω_1 are assigned to the meta-cluster $\omega_{1,3}$ labeled by the green plus, although there is no overlapping zone between ω_1 and ω_3. These unreasonable clustering results are mainly caused by the close proximity of different cluster centers (*i.e.* $\boldsymbol{v}_1, \overline{\boldsymbol{v}}_{2,3}, \overline{\boldsymbol{v}}_{1,2,3}$). In contrast, there is no object in DEC assigned to the meta-clusters $\omega_{2,3}$ and $\omega_{1,2,3}$, as shown in Fig. 3 (c), and the objects in the overlapping zone are reasonably assigned to the corresponding meta-cluster $\omega_{1,2}$. The DEC has obtained reasonable results, which not only reduces the clustering error rate but also reduces the clustering imprecision compared with ECM.

(2) Round Shape Data: We consider a particular 3-class dataset in the round shape as shown in Fig. 4 (a), and some objects are in the overlapping zones. The dataset consists of 594 data points, including 3 outlier data points. The radius of the round is $r = 2$ and the centers of the three rounds are given by the points $c_1 = (1.5, 2.5)$, $c_2 = (0, 0), c_3 = (3, 0)$.



(a) (b) (c)

Figure 3: Clustering results of 3-class of gaussian dataset by different methods. (a) Original gaussian data. (b) Clustering result of ECM ($Re = 3.00, Ri = 8.00, T = 0.3011s$). (c) Clustering result of DEC ($Re = 1.00, Ri = 4.00, T = 0.0028s$).



(a) (b) (c)

Figure 4: Clustering results of 3-class of round shape dataset by different methods. (a) Original artificial data. (b) Clustering result of CCM ($Re = 7.07, Ri = 17.34, T = 0.5897s$). (c) Clustering result of DEC ($Re = 4.23, Ri = 20.81, T = 0.0072s$).

Here CCM and DEC are applied to clustering the dataset with $\delta = 3$. Fig. 4 (b)-(c) show the clustering results of CCM and DEC respectively. The error rate (in %), the imprecision rate (in %), and computation time (s) of these methods are given in the caption of each subfigure.

As we can see from Fig. 4 (a), the objects in the clusters ω_1, ω_2, and ω_3 overlap partially at their edges, and these objects in overlapping zones are difficult to cluster accurately and clearly. In Fig. 4 (b), we can see that CCM assigns data points in the overlapping zones of different clusters to the corresponding meta-clusters: $\omega_{1,2}, \omega_{1,3}, \omega_{2,3}$. However, some objects in the middle zone of these three clusters are not accurately assigned. This unreasonable assignment is mainly because the meta-cluster threshold is selected as $t_c = 2$ in CCM. This also shows that it may be unreasonable for ECM and CCM to reduce the computational burden by limiting the number of elements in the meta-cluster in a few cases.

In Fig. 4 (c), the DEC assigns the objects in the overlapping zones of ω_1 and ω_2, ω_1 and ω_3, ω_2 and ω_3 to meta-clusters $\omega_{1,2}, \omega_{1,3}$ and $\omega_{2,3}$. Meanwhile, the DEC can reasonably assign the objects in the overlapping zone of these three clusters to the meta-cluster $\omega_{1,2,3}$ because these objects are totally unrecognized. Three objects labeled with the black hexagram are far away from the others, the CCM and DEC have better detection, and all of them are regarded as the outliers. Meanwhile, the execution time of DEC is much shorter than that of CCM, which indicates that the computational complexity of DEC is significantly less than CCM.

B. *Experiment 2*

In this experiment, we consider a 4-class square dataset including 1000 data points in each cluster with $c = 4$ as shown in Fig. 5, and there are two outliers in the dataset. The distribution intervals of the 4-class of dataset is as follows:

	x-label interval	y-label interval
ω_1	(-1, 1)	(-1, 1)
ω_2	(0.5, 2.5)	(0.5, 2.5)
ω_3	(2, 4)	(-1, 1)
ω_4	(0.5, 2.5)	(-2.5, -0.5)

C-means, FCM, NC, ECM, CCM, BPEC, and DEC are applied to clustering the dataset. NC, ECM, CCM, BPEC and DEC use the same outlier threshold $\delta = 2$. In BPEC, we take $K = 400$. In DEC, we take $\varphi = 0.2$ and $\varphi = 0.3$ to test the effect of parameter adjustment in different methods on the results. The clustering results of seven methods are shown in Fig. 5 (b)-(i). The error rate (in %), the imprecision rate (in %), and computation time (s) of these methods are given in the caption of each subfigure.

As we can see from Fig. 5 (a), the edges of different clusters are overlapping partially, and the objects in the zones are difficult to be clustered accurately and clearly. However, the

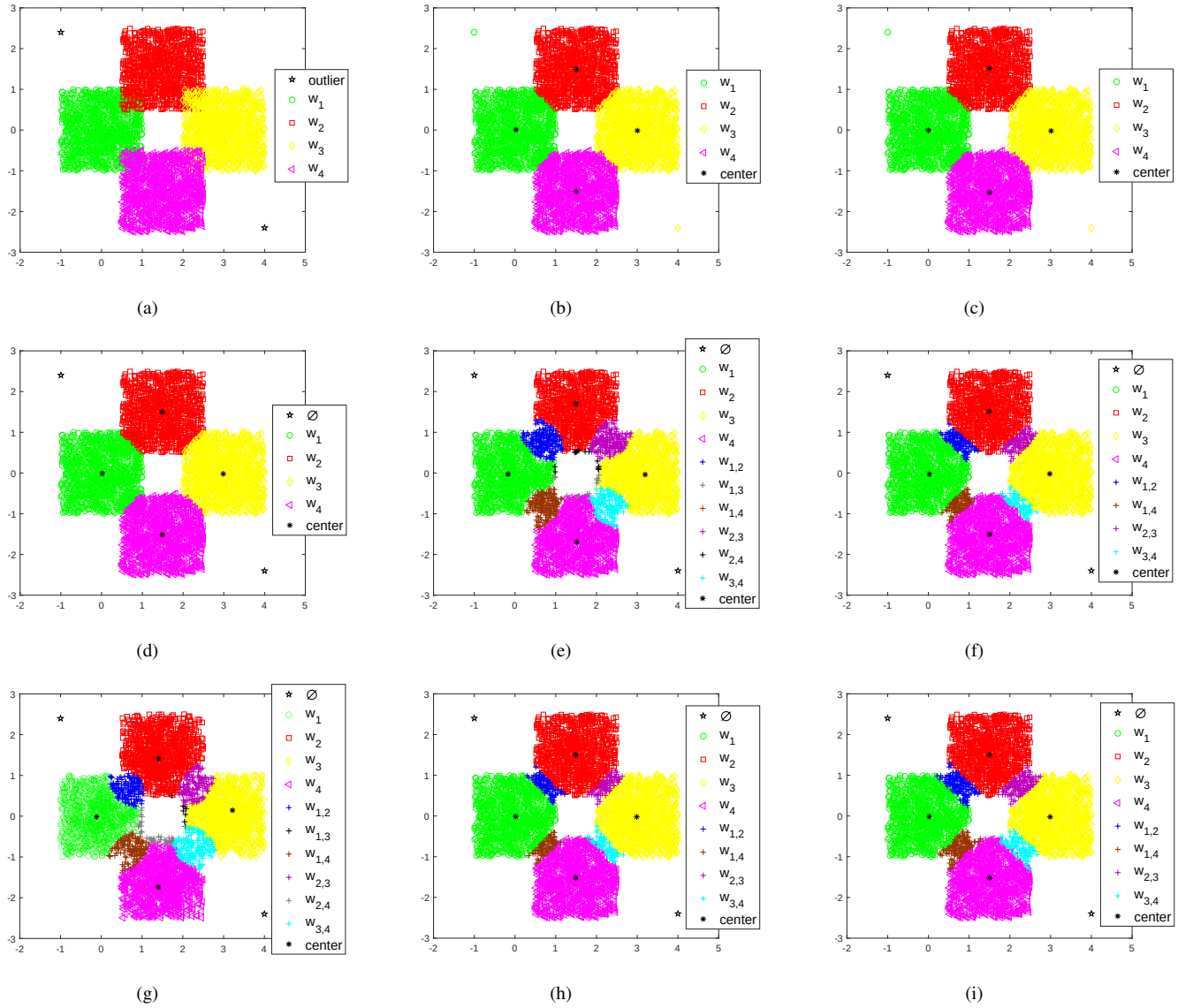


Figure 5: Clustering results of 4-class of square dataset by different methods. (a) Original square data. (b) Clustering result of C-means ($Re = 4.32, T = 0.0113s$). (c) Clustering result of FCM ($Re = 4.37, T = 0.0321s$). (d) Clustering result of NC ($Re = 4.38, T = 0.0445s$). (e) Clustering result of ECM ($Re = 0.30, Ri = 18.79, T = 6.83s$). (f) Clustering result of CCM ($Re = 1.12, Ri = 9.72, T = 5.35s$). (g) Clustering result of BPEC ($Re = 0.32, Ri = 15.17, T = 10.22s$). (h) Clustering result of DEC with $\varphi = 0.2$ ($Re = 1.50, Ri = 8.38, T = 0.0799s$). (i) Clustering result of DEC with $\varphi = 0.3$ ($Re = 0.78, Ri = 12.50, T = 0.1090s$).

C-means and the FCM, as shown in Fig. 5 (b) and (c), cluster these objects into only four singleton clusters ω_1, ω_2, ω_3 and ω_4 based on the probability framework, and the objects in overlapping zones are assigned to different singleton clusters, which is obviously unreasonable. In Fig. 5 (a), two objects labeled black hexagram are farther away from other objects, which also called outliers, C-means and FCM cannot detect these outliers, and assign them to ω_1 and ω_3 respectively. The NC in Fig. 5 (d) can effectively assign the two objects to outliers, but there are still some errors for those objects in different overlapping zones.

In ECM and BPEC, some objects from the specific clusters are unreasonably assigned to the meta-clusters. This is because the centers of some meta-clusters and that of singleton clusters are very close locally. For example, several objects in the cluster ω_1 and ω_3 are incorrectly assigned to the meta-cluster $\omega_{2,4}$ labeled with the black plus. However, the CCM and DEC can address this problem well. One can find that only the objects in overlapping zones that are difficult to be clustered clearly and accurately are assigned to meta-clusters. For instance, the objects in the overlapping zones of ω_1, ω_2, and ω_4 are reasonably assigned to meta-clusters $\omega_{1,2}$ and $\omega_{1,4}$. In addition, with the increases of φ from $\varphi = 0.2$ to $\varphi = 0.3$, increasingly objects will be assigned to the meta-clusters, as shown in Fig. 5 (h)-(i). Although there is little difference in the error rate between CCM and DEC, the execution time

of the DEC is obviously lower than that of CCM, and the DEC will be a good choice when the real-time requirement for computing is high in engineering applications.

C. Experiment 3

In the third experiment, we test the methods on the dataset consisting of 180000 points generated by different Gaussian distributions where each class contains 30000 data points. The dataset is shown in Fig. 6 (a). Here we take $\delta = 10$, and the meta-cluster threshold $\varphi = 0.2$ and $\varphi = 0.3$ in DEC. The results obtained by different methods are specified in the caption of each subfigure.

In Fig. 6 (a), one can clearly see that the objects in the overlapping zones are difficult to be accurately assigned to singleton clusters. The ECM and BPEC obtain similar clustering results in Fig. 6 (b) and (d), and they assign the imprecise objects to the corresponding meta-clusters with credal partition, for instance, the objects lying in the overlapping zones of ω_1 and ω_2 are assigned to the meta-cluster $\omega_{1,2}$. However, ECM and BPEC also unreasonably assign the objects that originally belong to singleton clusters to the unrelated meta-clusters. In Fig. 6 (a), ω_2 and ω_6 are completely separated, for example, ECM and BPEC still assign some objects belonging to ω_1 to the meta-cluster $\omega_{2,6}$ labeled by purple plus.

CCM and DEC overcome the limitation since the mass of belief of the object belonging to the meta-cluster is related not only to the distance from the object to the meta-cluster center but also to the distance between the object and the centers of the singleton clusters included in the meta-cluster. Therefore, in CCM and DEC, only imprecise objects in overlapping zones of different clusters are assigned to appropriate meta-clusters as shown in Fig. 6 (c) and (e)-(f). Whereas CCM and DEC are different in some cases, especially the computational complexity of DEC is much lower than that of CCM, which greatly expands the use of DEC especially in big data. From Fig. 6, one can see that the DEC take only $T = 85.31(s)$ and $T = 110.90(s)$ under different meta-clusters threshold, while ECM, CCM and BPEC take $T = 4354.21(s)$, $T = 3833.51(s)$ and $T = 10534.04(s)$ respectively. In DEC, it will lead to a slight increase in the number of objects assigned to meta-clusters with the increase of φ ($e.g.$, $\varphi = 0.2$ and $\varphi = 0.3$). In the applications, one can adjust the parameter φ according to the acceptable imprecision rate.

D. Experiment 4

In this experiment, nine real-world datasets are selected from the UCI repository to test the performance of DEC with respect to ECM, CCM, and BPEC. The basic information of nine datasets including the number of clusters (#Clus.), attributes (#Attr.), and instances (#Inst.) are shown in Table II, and all the detailed information can be found at http://archive.ics.uci.edu/ml/.

Table II: Basic information of the used datasets.

Name	#Clus.	#Attr.	#Inst.
Iris	3	4	150
Seeds	3	7	210
Haberman	2	3	306
Wine	3	13	178
Magic	2	10	19020
Contraceptive	3	9	1573
Vehicle	4	18	846
Satimage	7	36	6435
Glass	7	9	214

For the Iris and Seeds datasets, the outlier thresholds of ECM, CCM, BPEC, and DEC are set as $\delta = 10$, and the Haberman and Appendicitis datasets with $\delta = 100$, while for other datasets, the outlier thresholds are $\delta = 1000$. Furthermore, the number of nearest neighbors and quantile of these nearest neighbors are set for BPEC as follows, for the nine datasets from Iris to Appendicitis: (30,0.5), (20,0.9), (20,0.9), (15,0.5), (300,0.9), (115,0.9), (100, 0.9), (500, 0.9) and (5, 0.5). The other parameters in this experiment are default. The clustering results of these datasets by different methods are shown in Table III.

For most of the datasets, CCM and DEC usually obtain similar results, they provide fewer errors than ECM and BPEC. For the Wine dataset, the error rate of ECM and BPEC is lower than that of DEC, whereas its imprecision is much higher, and even some objects are assigned to total ignorant cluster Ω. In parallel, DEC has the highest CRI values in most situations, which truly shows the results obtained by the DEC are much more reasonable. The execution time of DEC is much lower than that of ECM, CCM, and BPEC, especially in the case of a large number of instances or clusters, which indicates that the DEC can still ensure good performance.

Furthermore, we visualize the clustering result of Iris real dataset in Fig. 7. Four features are measured from each object. They are the length and width of sepal and petal. From the original dataset shown in Fig. 7 (a) and (g), we can see that ω_1 is significantly distinct from ω_2 and ω_3, but ω_2 and ω_3 are partially overlapped. These four methods can easily separate ω_1 from ω_2 and ω_3. However, it is difficult to separate ω_2 and ω_3 in the overlapping zone. In ECM, CCM, and BPEC, although several objects in the overlapping zone are assigned to the meta-clusters, there are still some objects that are misclassified. In DEC, the error rate of clustering is significantly reduced when the meta-cluster threshold is from $\varphi = 0.2$ to $\varphi = 0.3$. More precisely, the number of objects with error clustering is reduced from 11 to 8, while the number of objects in the meta-clusters is increased from 10 to 13. The experiment shows that some objects in the overlapping zone, which are misclassified by other methods, are regarded as imprecise objects and then carefully assigned to the corresponding meta-clusters in DEC.

Fig. 8 shows the ROC curves for imprecision rates and error rates of ECM and DEC based on different datasets, where the x-axis denotes the mean of the imprecision rate, and the y-axis denotes the mean of the error rate with scale reversed. From Fig. 8, one can intuitively find that different clustering results can be obtained by adjusting parameters α and φ in ECM and DEC. More precisely, with the reduction of error rate, the imprecision rate will increment correspondingly. Interestingly, one can also see from Fig. 8 that the error rate of DEC, in most cases, is lower than that of ECM when

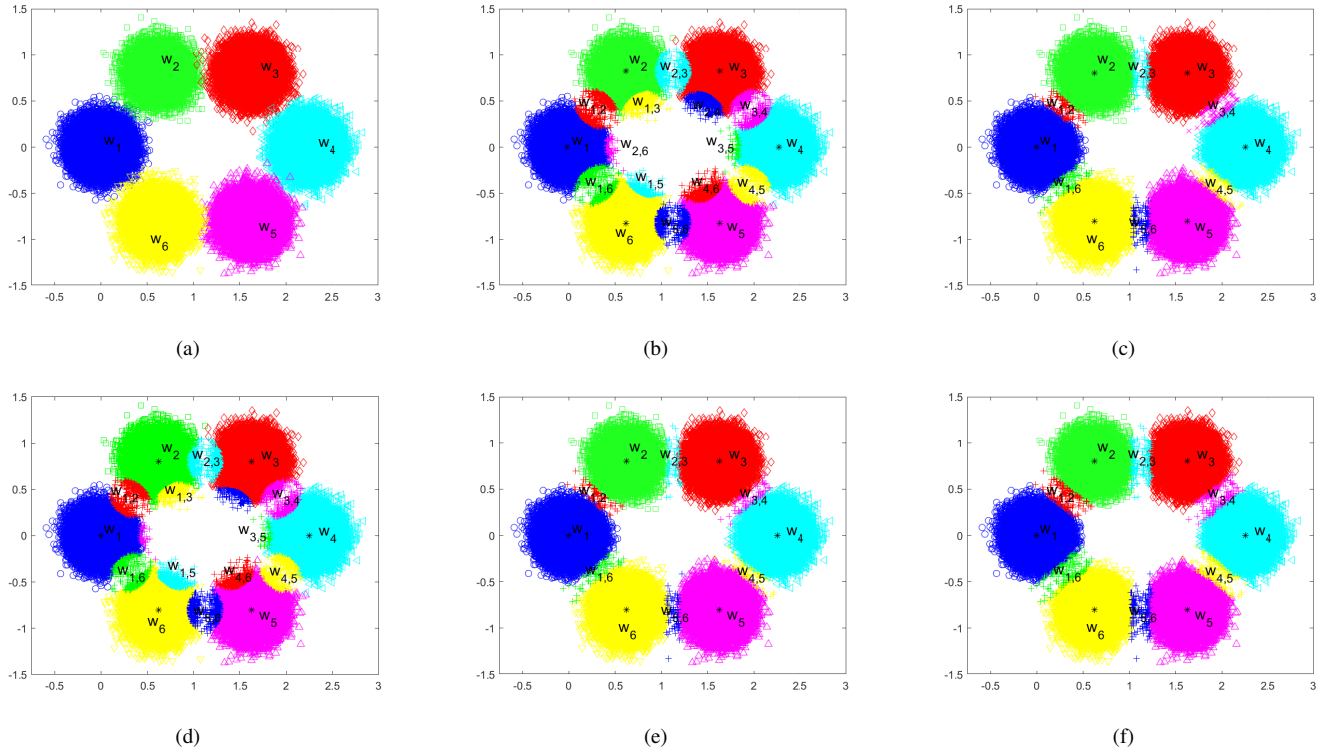


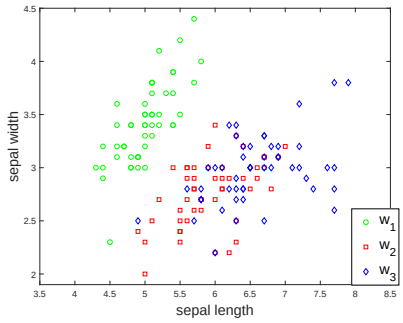
Figure 6: Clustering results of 6-class of dataset by different methods. (a) Original artificial data. (b) Clustering result of ECM ($Re = 1.13, Ri = 1.23, T = 4354.21s$). (c) Clustering result of CCM ($Re = 0, Ri = 0.30, T = 3833.51s$). (d) Clustering result of BPEC ($Re = 0.82, Ri = 1.14, T = 10534.04s$). (e) Clustering result of DEC with $\varphi = 0.2$ ($Re = 0, Ri = 0.21, T = 85.31s$). (f) Clustering result of DEC with $\varphi = 0.3$ ($Re = 0, Ri = 0.44, T = 110.90s$).

the methods have the same imprecision rate. In practical applications, both ECM and DEC can control the number of objects assigned to the meta-clusters by adjusting the parameters thereby controlling the imprecision rate. Moreover, the value of φ in DEC determines different partition methods. The DEC becomes a fuzzy/possibilistic partition when $\varphi = 0$, *i.e.* all objects are assigned to different singleton clusters, and DEC has a traditional credal partition when $\varphi = 1$, *i.e.* all objects are imprecise ones, which are assigned to singleton clusters or meta-clusters under the power-set 2^{Ω}. The result of DEC also present that the known information does not allow us to accurately classify the objects in the meta-clusters. If one wants to get more accurate results, some other (possibly expensive) techniques seem necessary to distinguish and classify these imprecise objects.

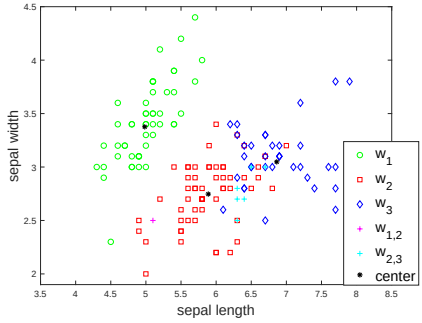
In order to verify the effectiveness of the DEC method from different perspectives, Table IV shows the objective function values and number of iterations of all datasets. One can see from Table IV that the number of iterations required for convergence of the DEC method is significantly less than that of ECM and CCM. For some datasets, however, the final objective function value of DEC is greater than that of ECM and CCM. The reason is that the objective function value of an object depends mainly on the mass of belief of its most believed cluster and its distance from the cluster center. We can find that although DEC, ECM, and CCM can assign objects to the most believed singleton cluster, the mass of beliefs for the objects belonging to the most believed cluster in DEC is greater than that of ECM and CCM. This is because most precise objects are directly assigned to the c singleton clusters in DEC without considering meta-clusters, which will result in a greater mass of beliefs compared to that of ECM and CCM. Thus, the objective function value of DEC is sometimes greater than that of ECM and CCM, especially when the number of clusters is large.

V. Managerial Implication

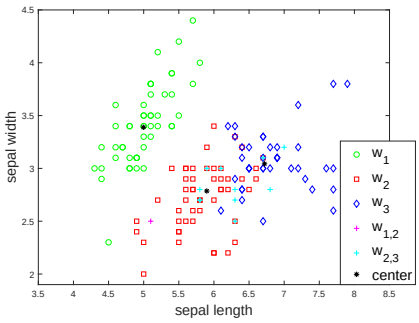
The management of parameters and sensitivity analysis are important components for the application of the proposed method [40]-[43]. In this paper, three parameters, $i.e.$ the meta-cluster threshold φ, the weighting exponent β and the outlier threshold δ, are involved in the proposed DEC method. Since the parameters β and δ has been discussed in a number of works, we mainly focus on the choice of the threshold φ in this part. One can find that the proposed DEC method can deal with fuzzy partition/classical credal partition/dynamic credal partition mode switching by controlling the threshold φ. This can help us to achieve fuzzy partition when imprecise information is not allowed in the results. However, sometimes users are more willing to obtain reliable imprecise results because the risk of incorrect clustering may be unbearable. In such a case, DEC can switch to dynamic credal partition, which greatly reduces the computational complexity compared to traditional credal partition.



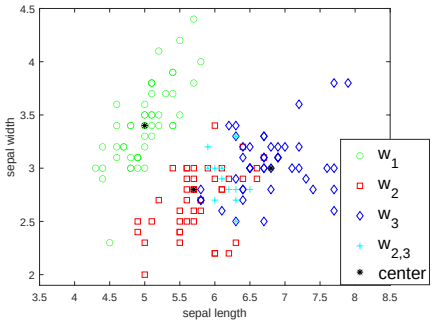
(a) Original dataset.



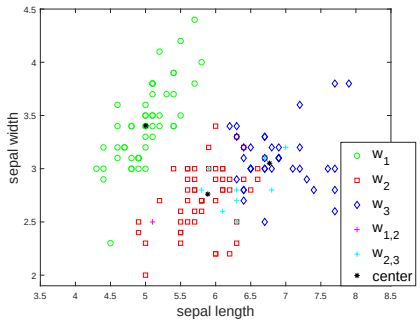
(b) Clustering results of ECM.



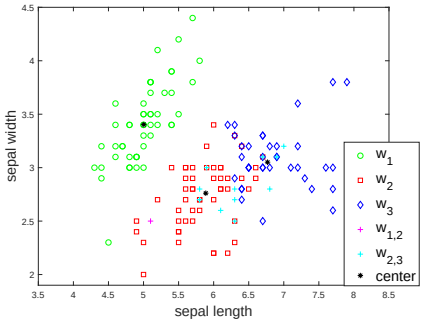
(c) Clustering results of CCM.



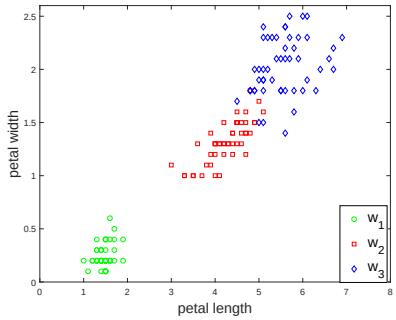
(d) Clustering results of BPEC.



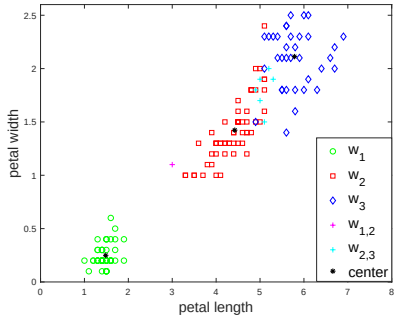
(e) Clustering results of DEC with $\varphi = 0.2$.



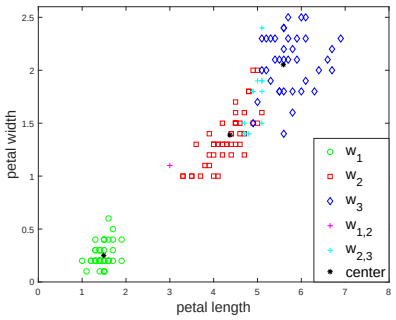
(f) Clustering results of DEC with $\varphi = 0.3$.



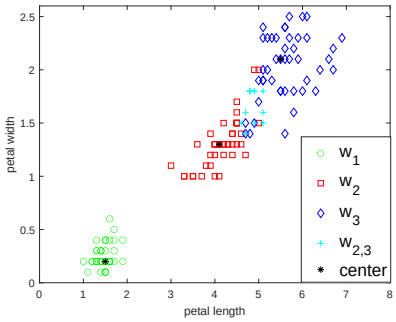
(g) Original dataset.



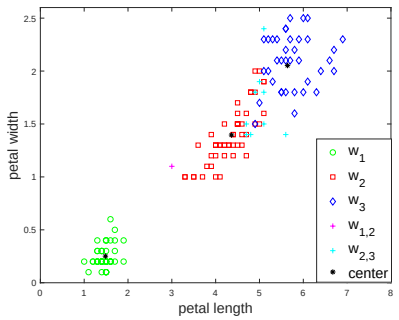
(h) Clustering results of ECM.



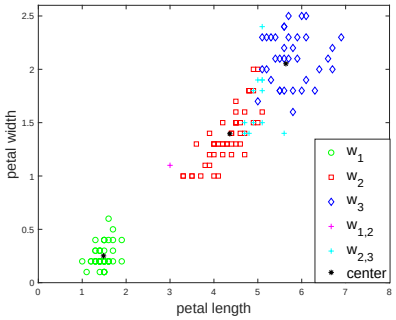
(i) Clustering results of CCM.



(j) Clustering results of BPEC.

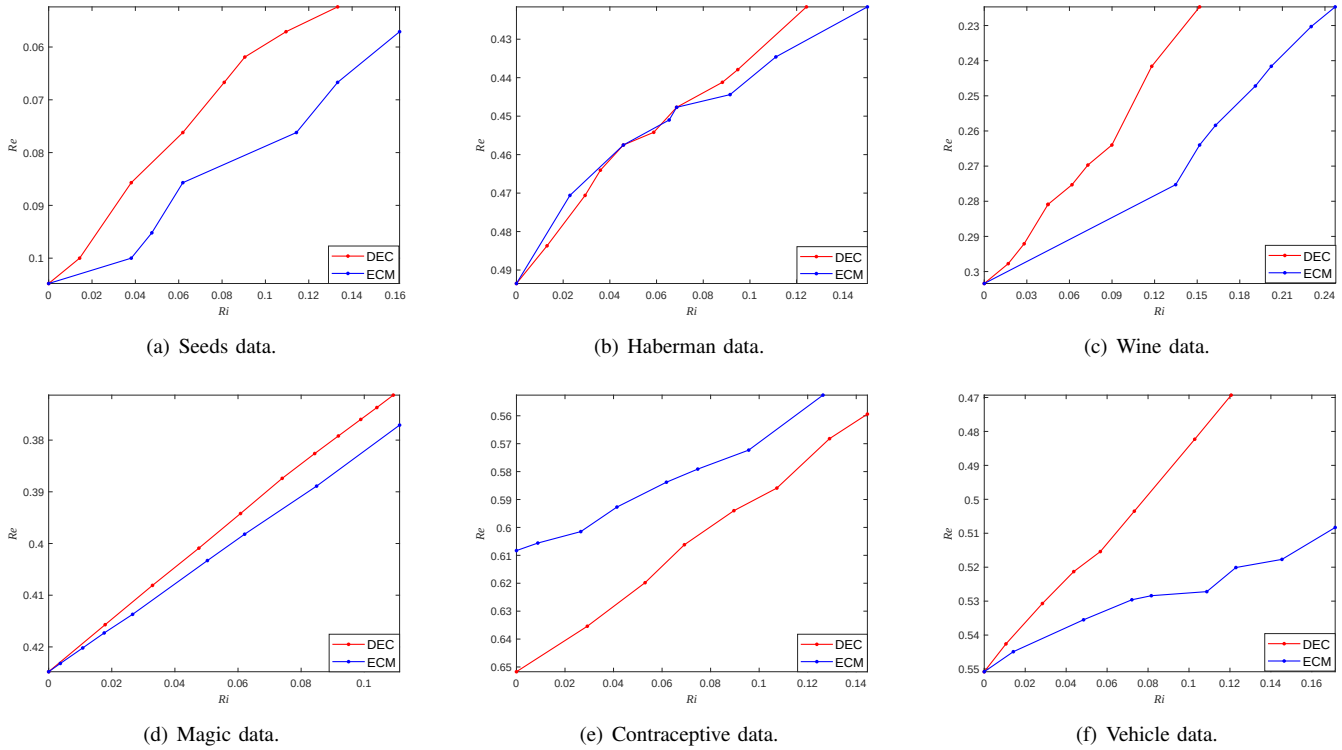


(k) Clustering results of DEC with $\varphi = 0.2$.



(l) Clustering results of DEC with $\varphi = 0.3$.

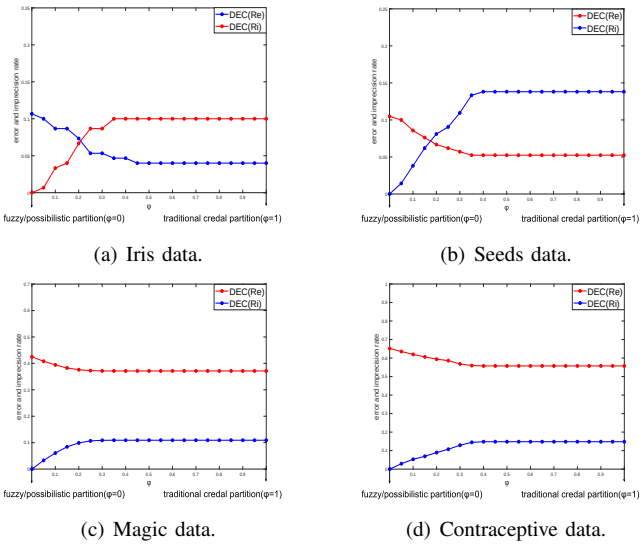
Figure 7: Clustering results of Iris dataset.



(a) Seeds data. (b) Haberman data. (c) Wine data.

(d) Magic data. (e) Contraceptive data. (f) Vehicle data.

Figure 8: ROC curves for imprecision rates and error rates of ECM and DEC based on different datasets.



(a) Iris data. (b) Seeds data.

(c) Magic data. (d) Contraceptive data.

Figure 9: Clustering results of different datasets for various φ.

Fig. 9 shows the statistical results of the error and imprecision rates based on the Iris, Seeds, Magic, and Contraceptive datasets as φ changes from 0 to 1. It is easy to see that as φ increases the error rate decreases and the inaccuracy rate increases, but this does not last forever. The method has a high sensitivity when $\varphi = 0.4$, as measured from the point of view of prudent decision-making, which is consistent with our intuition. This also demonstrates the validity of the method, $i.e.$ it is time-consuming and pointless to subject all query objects to the power-set 2^{Ω}. However, this does not imply that $\varphi = 0.4$ is the only value to be taken, as there are also applications where it is desirable to obtain precise clustering results. Therefore, we suggest that $\varphi \in [0, 0.4]$ and $\varphi = 0.3$ as the default. The user should set the threshold φ based on an acceptable imprecision rate.

VI. Conclusions

A new dynamic evidential clustering algorithm (DEC) has been developed for imprecise objects thanks to the credal partition, which provides a dynamic edited framework $\mathcal{M}_i^{\Omega}$ ($\mathcal{M}_i^{\Omega} \subset 2^{\Omega}$) including several close singleton clusters and related meta-clusters for each imprecise object to be assigned more reasonably. The DEC overcomes the shortcomings of ECM [3] and CCM [25] and greatly reduces the time-consuming because it eliminates the possible negative impact of other meta-clusters on imprecise objects and also avoids a large number of invalid computations compared with the framework of power-set 2^{Ω}. The proposed method (DEC) can effectively extend the applications of credal partition, especially in big data, because it reduces the computational complexity of credal partition to a level similar to that of the fuzzy and possibilistic partition (such as FCM [6], PCM [9], and NC [14]). Four experiments with artificial and real datasets have been done to verify the performance of the proposed DEC method with respect to other methods. It should be noted that in credal partition, once the object is assigned to the meta-cluster, the object may belong to any singleton cluster included in the meta-cluster under the known information, which is also one of the advantages of credal partition – it can reasonably characterize and represent the imprecision and

uncertainty caused by a variety of reasons. If one wants to get more precise results, some other (possibly costly) techniques or information sources must be developed and used for the imprecise objects rather than enforce assigning to increase the risk of errors. In addition, one can find that the imprecision rate increases as φ increases in a smaller range (not [0,1]), but it does not continue. This is consistent with our assumption that only part of the objects' cluster information is imprecise. Therefore, we recommend that $\varphi = 0.3$ is the default value. We will try to give a method of optimizing the threshold φ for different specific datasets in future work.

Table III: Clustering results of different datasets (in $\%$).

Data Set		ECM	CCM	BPEC	DEC
Iris	Re	8.00	5.33	6.00	5.33
	Ri	4.67	8.00	8.00	8.67
	CRI	0.8509	0.8631	0.8549	**0.9184**
	T	0.1600	0.0695	0.2949	**0.0050**
Seeds	Re	7.62	5.71	7.62	5.71
	Ri	11.43	10.00	10.00	10.95
	CRI	0.7977	0.7912	0.7760	**0.8367**
	T	0.0754	0.0730	0.3810	**0.0075**
Haberman	Re	43.46	42.16	39.87	42.16
	Ri	11.11	12.42	15.03	12.42
	CRI	0.6631	**0.7064**	0.6821	0.5494
	T	0.0912	0.0883	0.2943	**0.0104**
Wine	Re	17.98	27.53	18.54	26.97
	Ri	32.02	9.55	29.21	7.30
	CRI	0.8207	0.7343	0.8076	**0.8814**
	T	0.5020	0.3813	0.9520	**0.0086**
Magic	Re	43.32	39.33	40.85	37.15
	Ri	0.37	8.25	18.05	10.88
	CRI	0.6245	**0.7421**	0.6987	0.5754
	T	9.15	9.05	34.66	**0.6273**
Contraceptive	Re	57.91	56.48	54.38	56.82
	Ri	7.49	8.35	12.90	12.90
	CRI	0.6245	0.6419	0.6519	**0.7712**
	T	2.63	1.53	4.29	**0.0600**
Vehicle	Re	52.84	51.73	54.14	48.23
	Ri	8.16	6.03	6.38	10.28
	CRI	0.7490	**0.7973**	0.7558	0.7251
	T	2.77	1.69	3.92	**0.0626**
Satimage	Re	28.08	30.54	31.45	24.07
	Ri	5.44	8.66	4.83	12.56
	CRI	0.7279	0.7523	0.7101	**0.8482**
	T	23.72	18.17	44.58	**1.70**
Glass	Re	48.13	47.66	49.07	46.26
	Ri	2.34	6.07	2.34	13.08
	CRI	0.6470	0.7007	0.6411	**0.7805**
	T	0.6993	0.3673	2.15	**0.0678**

Table IV: Clustering results of different datasets.

Dataset	Method	Objective function value	Number of iterations
Iris	ECM	47.57	33
	CCM	42.44	18
	DEC	56.91	15
Seeds	ECM	294.99	27
	CCM	263.81	24
	DEC	371.24	16
Haberman	ECM	18071.77	20
	CCM	15519.06	23
	DEC	18860.38	17
Wine	ECM	796031.24	75
	CCM	1177975.85	86
	DEC	1647084.50	40
Magic	ECM	115788607.23	40
	CCM	91457280.53	66
	DEC	109617350.83	38
Contraceptive	ECM	12886.81	67
	CCM	11627.91	33
	DEC	16199.99	26
Vehicle	ECM	1460608.50	97
	CCM	1366180.72	87
	DEC	2182377.66	56
Satimage	ECM	4859854.19	140
	CCM	3120338.23	136
	DEC	5325419.18	63
Glass	ECM	96.97	31
	CCM	56.30	31
	DEC	76.45	29

VII. Acknowledgment

This work has been partially supported by the National Natural Science Foundation of China (No. U20B2067, No. 61790552, No. 61790554, No. 61701409), Shaanxi Science Fund for Distinguished Young Scholars (No. 2018JC-006), and Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University (No. CX201953).

References

[1] J. Liang, L Bai, C. Dang, The k-means-type algorithms versus imbalanced data distributions, IEEE Trans. Fuzzy Syst., vol. 20, no. 4, pp. 728-745, 2012.

[2] T. Denœux, S. Sriboonchitta, O. Kanjanatarakul. Evidential clustering of large dissimilarity data, Knowl.-Based Syst, vol. 106, pp. 179-195, 2016.

[3] M. H. Masson, T. Denœux, ECM: An evidential version of the fuzzy c-means algorithm, Pattern Recognit., vol. 41, no. 4, pp. 1384-1397, 2008.

[4] F. Li, S. Li, T. Denœux, k-CEVCLUS: Constrained evidential clustering of large dissimilarity data, Knowl.-Based Syst, vol. 142, pp. 29-44, 2018.

[5] A. K. Jain, Data clustering: 50 years beyond k-means, Pattern recognit. lett., vol. 31, no. 8, pp. 651-666, 2010.

[6] J. Bezdek, Pattern recognition with fuzzy objective function algorithm, Plenum Press, New-York, 1981.

[7] Q. Bui, B. Vo, H. N. Do, N. Q. V. Hung, V. Snasel, F-Mapper: a fuzzy mapper clustering algorithm, Knowl.-Based Syst, vol. 189, 2020.

[8] J. C. Bezdek, A convergence theorem for the fuzzy ISODATA clustering algorithms, IEEE Trans. Pattern Anal. Mach. Intell., vol. 2, no. 1, pp. 1-8, 1980.

[9] R. Krishnapuram, J. Keller, A possibilistic approach to clustering, IEEE Trans. Fuzzy Syst., vol. 1, no. 2, pp. 98-110, 1993.

[10] M. Barni, V. Cappellini, A. Mecocci, A possibilistic approach to clustering, IEEE Trans. Fuzzy Syst., vol. 4, no. 3, pp. 393-396, 1996.

[11] N. R. Pal, K. Pal, J. Keller, J. C. Bezdek, A possibilistic fuzzy cmeans clustering algorithm, IEEE Trans. Fuzzy Syst., vol. 13, no. 4, pp. 517-530, 2005.

[12] X. Bai, Y. Zhang, H. Liu, Z. Chen, Similarity measure-based possibilistic FCM with label information for brain MRI segmentation, IEEE Trans. Cybern., vol. 49, no. 7, pp. 2618-2630, 2019.

[13] M. Yang, Y. Nataliani, A feature-reduction fuzzy clustering algorithm based on feature-weighted entropy, IEEE Trans. Fuzzy Syst., vol. 26, no. 2, pp. 817-835, 2018.

[14] R. N. Davé, Clustering relational data containing noise and outliers, Pattern Recognition Lett. vol. 12, pp. 657-664, 1991.

[15] S. Sen, R. N. Davé, Clustering of relational data containing noise and outliers, Proc. IEEE Int. Conf. Fuzzy Syst., vol. 2, pp. 1411-1416, 1998.

[16] T. Denœux, M. H. Masson, EVCLUS: EVidential CLUStering of proximity data, IEEE Trans. Syst. Man Cybern. Part B, vol. 34, no. 1, pp. 95-109, 2004.

[17] M. H. Masson, T. Denœux, Clustering interval-valued data using belief functions, Pattern Recognition Lett. vol. 25, no. 2, pp. 163-171, 2004.

[18] G. Shafer, A mathematical theory of evidence, Princeton, NJ, USA: Princeton Univ. Press, 1976.

[19] P. Smets, The combination of evidence in the transferable belief model, IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, no. 5, pp. 447-458, 1990.

[20] P. Smets, Decision making in the TBM: the necessity of the pignistic transformation, Int. J. Approximate Reasoning, vol. 38, no. 2, pp. 133-147, 2005.

[21] Y. Yang, D. Q. Han, A new distance-based total uncertainty measure in the theory of belief functions, Knowl.-Based Syst., vol. 94, pp. 114-123, 2016.

[22] X. Deng, Analyzing the monotonicity of belief interval based uncertainty measures in belief function theory, Int. J. Intell. Syst., vol. 33, no. 9, pp. 1869-1879, 2018.

[23] T. Denœux, Decision-Making with belief functions: a review. Int. J. Approximate Reasoning, vol. 109, pp. 87-110, 2019.

[24] C. Lian, S. Ruan, T. Denœux, H. Li, P. Vera, Joint tumor segmentation in PET-CT images using co-clustering and fusion based on belief functions, IEEE Trans. Image Proc., vol. 28, no. 2, pp. 755-766, 2019.

[25] Z. G. Liu, Q. Pan, J. Dezert, G. Mercier, Credal c-means clustering method based on belief functions, Knowl.-Based Syst, vol. 74, pp.119-132, 2015.

[26] Z. G. Su, T. Denœux, BPEC: Belief-peaks evidential clustering. IEEE Trans. Fuzzy Syst., vol. 27, no, 1, pp. 111-123, 2019.

[27] K. Zhou, A. Martin, Q. Pan, Z.-G. Liu, Median evidential c-means algorithm and its application to community detection. Knowl.-Based Syst, vol. 74, pp.69-88, 2015.

[28] K. Zhou, A. Martin, Q. Pan, Z. G. Liu, ECMdd: Evidential c-medoids clustering with multiple prototypes. Pattern Recognit., vol. 60, pp. 239-257, 2016.

[29] T. Denœux, A k-nearest neighbor classification rule based on Dempster-Shafer theory. IEEE Trans. Syst. Man Cybern., vol. 25, no. 5, pp. 804-844, 1995.

[30] Z. G. Liu, L. Q. Huang , K. Zhou, T. Denoeux, Combination of Transferable Classification With Multisource Domain Adaptation Based on Evidential Reasoning, IEEE Trans. Neural. Netw. Learn. Syst., DOI: 10.1109/TNNLS.2020.2995862.

[31] X. Xu, J. Zheng, J. B. Yang, Data classification using evidence reasoning rule, Knowl.-Based Syst, vol. 116, pp. 144-151, 2017.

[32] Z. Luo, Y. Deng, A matrix method of basic belief assignment's negation in Dempster-Shafer theory, IEEE Trans. Fuzzy Syst., 2019, doi: 10.1109/TFUZZ.2019.2930027.

[33] Z. G. Liu, Y. Liu, J. Dezert, F. Cuzzolin, Evidence combination based on credal belief redistribution for pattern classification. IEEE Trans. Fuzzy Syst., vol. 28, no. 4, pp. 618-631, 2020

[34] Z. G. Liu, X. Zhang, J. W. Niu, J. Dezert, Combination of Classifiers With Different Frames of Discernment Based on Belief Functions, IEEE Trans. Fuzzy Syst., 2020. DOI 10.1109/TFUZZ.2020.2985332

[35] D. Han, J. Dezert, Y. Yang, Belief interval-based distance measures in the theory of belief functions, IEEE Trans. Syst. Man Cybern., vol. 48, no. 6, pp. 833-850, 2018.

[36] T. Denœux, S. Li, S. Sriboonchitta. Evaluating and comparing soft partitions: an approach based on Dempster-Shafer theory, IEEE Trans. Fuzzy Syst., vol. 26, no. 3, pp. 1231-1244, 2018.

[37] S. H. Pasandideh, S. T. Niaki, A. Gharaei, Optimization of a multiproduct economic production quantity problem with stochastic constraints using sequential quadratic programming, Knowl.-Based Syst., vol. 84, pp. 98-107, 2015.

[38] A. Gharaei, S. A. Shekarabi, M. Karimi, E. Pourjavad, A. Amjadian, An integrated stochastic EPQ model under quality and green policies: generalised cross decomposition under the separability approach, Int. J. Syst. Sci.: Oper. Logist., pp. 1-13, 2019.

[39] A. Gharaei, S. A. Shekarabi, M. Karimi, Modelling And optimal lot-sizing of the replenishments in constrained, multi-product and bi-objective EPQ models with defective products: Generalised Cross Decomposition. Int. J. Syst. Sci., pp. 1-13, 2019.

[40] A. Awasthi, H. Omrani, A goal-oriented approach based on fuzzy axiomatic design for sustainable mobility project selection. Int. J. Syst. Sci.: Oper. Logist., pp. 1-13, 2018.

[41] M. Noori-daryan, A. A. Taleizadeh, M. Rabbani, A hybrid robust possibilistic approach for a sustainable supply chain location-allocation network design. Int. J. Syst. Sci.: Oper. Logist., pp. 60-75, 2018.

[42] B. C. Giri, M. Masanta, Developing a closed-loop supply chain model with price and quality dependent demand and learning in production in a stochastic environment. Int. J. Syst. Sci.: Oper. Logist., pp. 147-163, 2018.

[43] B. C. Giri, S. Bardhan, Coordinating a supply chain with backup supplier through buyback contract under supply disruption and uncertain demand. Int. J. Syst. Sci.: Oper. Logist., pp. 193-204, 2014.