Information Sciences 736 (2026) 123108

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Three-way clustering propelled by multi-scale uncertainty propagation

Caihui Liu[a, b, *], Xiying Chen[c], Wenjing Qiu[a, b], Duoqian Miao[c]

[a] Department of Mathematics and Computer Science, Gannan Normal University, Ganzhou 341000, Jiangxi, China
[b] Key Laboratory of Data Science and Artificial Intelligence of Jiangxi Education Institutes, Gannan Normal University, Ganzhou 341000, China
[c] Department of Computer Science and Technology, Tongji University, Shanghai, 201804, China

HIGHLIGHTS

- We provide a new perspective on evaluation function-based three-way clustering.
- We introduce two uncertainty measures by analyzing membership relationships between samples and clusters.
- We propose a multi-scale uncertainty propagation mechanism.
- We present a novel three-way clustering algorithm, Up3WC, based on the multi-scale uncertainty propagation mechanism.
- We carried out experiments on 16 UCI datasets. The results validate the effectiveness of the proposed algorithm.

ARTICLE INFO

Keywords:
Clustering
Three-way clustering
Multi-scale
Uncertainty data analysis

ABSTRACT

Guided by the three-way decision principles, three-way clustering methods effectively capture information uncertainty by characterizing cluster structures through cores and fringe regions. However, most existing approaches evaluate data uncertainty only from the perspective of density or distance, thus failing to comprehensively reflect the intrinsic structure of the data. To address this limitation, this paper proposes a multi-scale uncertainty propagation three-way clustering algorithm. First, by analyzing density-based and distance-based membership relationships between samples and clusters, two uncertainty measures, kernel density scores, and boundary uncertainty, are defined to jointly characterize data uncertainty through global density distribution and local geometric correlations. Subsequently, a multi-scale uncertainty propagation mechanism is developed to dynamically update the sample uncertainties through iterative propagation, enabling progressive information fusion and transmission. Finally, a dynamic three-way assignment strategy is designed to adaptively divide samples into three regions based on both distance and density information, and then a corresponding three-way clustering algorithm is constructed. In the experiments, the proposed algorithm is compared with eight other clustering methods on 16 datasets with varying dimensions, and its effectiveness is demonstrated through both qualitative and quantitative analysis.

[*] Corresponding author at: Department of Mathematics and Computer Science, Gannan Normal University, Ganzhou 341000, Jiangxi, China.
Email addresses: liucaihui@gnnu.edu.cn, liu_caihui@163.com (C. Liu), xiying_chen@gnnu.edu.cn (X. Chen), wenjingqiu@gnnu.edu.cn (W. Qiu), dqmiao@tongji.edu.cn (D. Miao).

https://doi.org/10.1016/j.ins.2026.123108
Received 10 October 2025; Received in revised form 9 January 2026; Accepted 10 January 2026
Available online 16 January 2026

1. Introduction

As one of the most prominent unsupervised learning tools, clustering analysis has evolved over six decades and found broad applications in many fields including bioinformatics [1], anomaly detection [2], community detection [3], and image segmentation [4]. Clustering analysis helps to find hidden structures and patterns inside datasets by means of dividing samples into separate clusters. Its primary objective is to maximize intra-cluster similarity and to minimize inter-cluster dissimilarity. Clustering methods can be classified into numerous types based on their underlying ideas and uses, including partition-based methods, hierarchical-based methods, and density-based approaches, etc.

Most clustering algorithms are designed based on the assumption that each sample exclusively belongs to a single cluster, thereby allowing only two possible relationships between samples and clusters. These categorical assignment methods, known as two-way clustering (2WC), emphasize clear and distinct cluster boundaries. Although well-defined boundaries help to examine and apply clustering findings, actual datasets usually include overlapping clusters, which complicates the creation of clear binary choices. This difficulty is especially evident in situations calling for multi-group membership. For instance, in culinary classification, a dish could fit both “vegetarian” and “Sichuan cuisine” classifications at once. Clearly, traditional 2WC methods fall short of capturing the inherent uncertainties in datasets, thus motivating the development of three-way clustering (3WC) algorithms.

Inspired by human cognitive processes, Granular Computing (GrC) [5–7] provides a fundamental paradigm in artificial intelligence for modeling and processing information under uncertainty, imprecision, and incompleteness. Building on this basis, three-way decisions (3WD) have developed into a new theoretical framework and approach for tackling difficult decision-making challenges, originally motivated by the need to formalize decision rules involving acceptance, rejection, and non-commitment under uncertainty [8,9]. The core concept of 3WD is to construct the universe as a triad and devise strategies to achieve the desired outcome. From a process perspective, this idea can be systematically organized into the TAO (Triading–Acting–Optimizing) framework, which characterizes the entire decision process from triadic modeling to decision execution and refinement [10]. Depending on the scope of research, existing studies on 3WD can be broadly categorized into two branches: narrow-sense 3WD, which focuses on probabilistic rough set models and their theoretical and applied developments, and broad-sense 3WD, which extends the triadic decision philosophy to general problem-solving and computing paradigms. In recent years, 3WD has inspired diverse research areas, including sequential three-way decisions [11], three-way multi-label classification [12], three-way group decision-making [13], and three-way time series analysis [14].

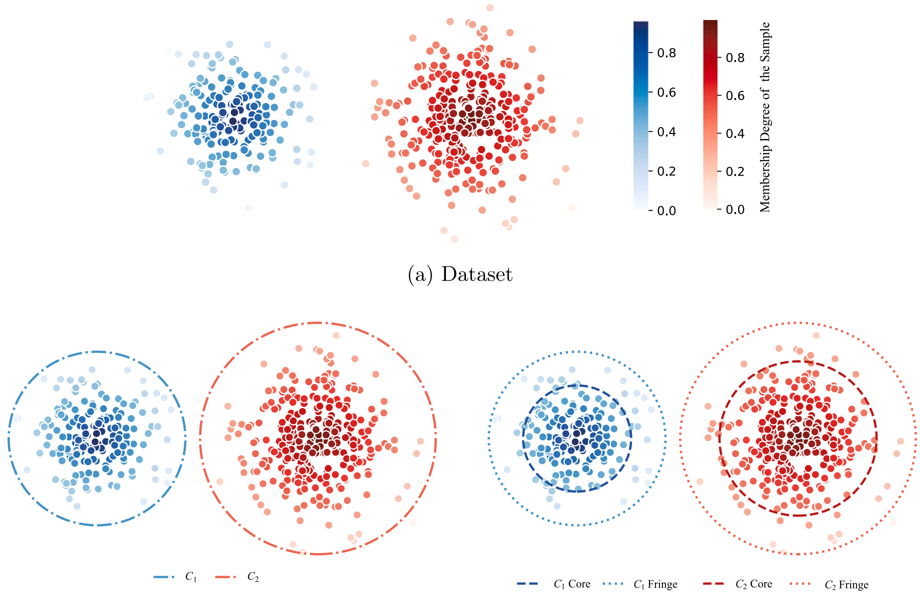
To address the problem of 2WC in capturing information uncertainty, three-way clustering applies 3WD theory to clustering analysis [15,16]. Unlike traditional methods, 3WC considers three types of relationships between samples and clusters: belonging, partial belonging, and non-belonging. Based on these three relationships, the data within each cluster is divided into three regions: core, fringe, and trivial regions, with clustering results represented by the core and fringe regions [17]. Fig. 1(a) illustrates a dataset with a Gaussian distribution, where the color intensity of the sample is positively correlated with its degree of membership to the corresponding cluster. In contrast to 2WC, where each point is uniquely assigned to a cluster (C_1 or C_2), as shown in Fig. 1(b), 3WC divides clusters into core and fringe regions, allocating low-membership points to the latter (Fig. 1(c)). This allocation strategy is consistent with the principle of risk minimization in three-way decision theory. By deferring ambiguous samples to the fringe region rather than forcing a hard assignment, 3WC reduces decision risk and enables more reliable structural analysis under uncertainty [18].

Since the introduction of 3WC, many researchers have contributed innovative methods within this framework, particularly those based on evaluation functions and thresholds. In these approaches, the evaluation function captures the membership relationship between each object and its corresponding cluster, while specific thresholds determine the criteria for regional assignments. For instance, Wang et al. [19] formulated the CE3 clustering model using mathematical morphology to establish thresholds via contraction and expansion, while Shah et al. [20] explored an analogous strategy using image blurring and sharpening techniques. By integrating three-way decision into partitional clustering, Wang et al. [21] proposed Three-way K-means (TWKM), which utilizes distance perturbation as the evaluation metric and separates core regions from supports based on allowable error thresholds. To better handle complex distributions, Guan et al. [22] developed a hierarchical approach (HC3) that employs kernel density estimation to evaluate sample significance, using a truncation density threshold to iteratively peel away fringe layers. Similarly, Han et al. [23] proposed a granular-ball method (3WG), which evaluates the spatial relationship between granular balls and their neighbors, delineating core and fringe regions based on adaptive radius-distance.

Beyond these foundational works, recent research has further extended 3WC toward more specialized data scenarios. Du et al. [24] introduced M3W, which employs a progressive erosion strategy to construct multilevel structures for sequential decision-making. Ju et al. [25] proposed D3WC, integrating deep learning with granular evidence fusion to manage uncertainty in high-dimensional representations. Similarly, Sun et al. [26] developed TWStream to adapt 3WC to data streams by buffering concept drift, while Yang et al. [27] presented graph-based approaches that utilize local density trends to characterize cluster backbones.

However, despite these advancements, notable gaps remain in both the mechanisms of uncertainty modeling and the applicability of the frameworks.

First, regarding the mechanism of uncertainty, most evaluation-based methods rely on a single perspective, specifically geometric distance or local density, combined with static evaluation. Distance-based methods assume metric compactness, failing on non-convex shapes, while density-based methods often overlook the intrinsic geometric structure, struggling to separate adjacent clusters with ambiguous boundaries. More importantly, these approaches employ a static determination strategy, assessing samples via rigid thresholds based on distance or density metrics. Consequently, they fail to capture the contextual dependency of uncertainty, overlooking how a sample’s ambiguity is structurally coupled with the neighboring samples.



(a) Dataset

(b) 2WC

(c) 3WC

Fig. 1. Illustration of two-way clustering (2WC) and three-way clustering (3WC).

Second, regarding applicability, advanced models often rely on complex deep architectures or specific structural assumptions such as hierarchical peeling, which limit their flexibility as a generic refinement tool. There is still a need for a flexible refinement strategy that can jointly exploit distance and density constraints to enhance two-way clustering results.

To bridge this gap, this paper proposes a multi-scale uncertainty propagation three-way clustering algorithm, called Up3WC. Unlike specialized deep models or static evaluation methods, Up3WC focuses on the fundamental coupling of distance and density, providing a generic, interpretable, and adaptive refinement module for standard clustering tasks. Specifically, first, kernel density estimation (KDE) is employed to calculate the uncertainty of each sample. KDE does not rely on specific data distribution assumptions and therefore is highly effective for handling complex or unknown distributions and better capturing the data uncertainty [28]. Next, to clearly identify the structural features of the dataset, geometric distance information is used to determine the initial uncertainty of each sample in its assigned cluster. By jointly considering density characteristics and geometric distance structure, an uncertainty propagation mechanism is constructed, which iteratively updates and determines the uncertainty of each sample. Finally, the core, fringe, and trivial regions are dynamically constructed by jointly exploiting density and distance information, thereby enabling more adaptive and accurate identification of cluster boundaries.

The contributions of this paper are summarized as follows:

(1) We provide a new perspective for evaluation-based three-way clustering by formulating it as a multi-scale uncertainty modeling problem. Compared with existing methods that rely on either distance or density alone, Up3WC jointly exploits distance-based boundary uncertainty and density-based similarity to characterize clustering uncertainty in a more comprehensive manner.
(2) We propose a multi-scale uncertainty propagation mechanism that iteratively diffuses boundary uncertainty over a density similarity structure built from two complementary measures. By coupling local geometric information with global density variation, this mechanism produces more stable and discriminative uncertainty estimates than using either component alone.
(3) We develop a novel three-way clustering algorithm, Up3WC, that can be used as a generic three-way refinement module on top of different two-way clustering algorithms. A dynamic three-way partitioning strategy is designed to adaptively construct core, fringe, and trivial regions within each cluster, overcoming the limitations of global, single-threshold evaluation schemes in existing three-way clustering methods.
(4) We conducted extensive experiments on 16 UCI datasets, along with ablation studies and qualitative analyses. The qualitative and quantitative results validate the effectiveness of the proposed algorithm and highlight the benefits of the multi-scale uncertainty propagation design.

The remainder of this paper is organized as follows. Section 2 reviews the background knowledge relevant to this study. Section 3 details the proposed algorithm. Section 4 presents experimental results on datasets of varying sizes to demonstrate the performance of the proposed algorithm. Finally, the summary and prospects are concluded in Section 5.

2. Preliminaries

2.1. Three-way clustering

Let $C = \{C_1, C_2, \dots, C_k, \dots, C_K\}$ represent the clustering result of the universe $\mathbb{U} = \{x_1, x_2, \dots, x_n\}$. Unlike two-way clustering, which uses a single set to represent clusters, three-way clustering divides a cluster into three regions: the core region $\hat{C}_k$, the fringe region $\tilde{C}_k$, and the trivial region $\dddot{C}_k$. If an object $x \in \hat{C}_k$, it definitely belongs to the cluster C_k; if $x \in \tilde{C}_k$, it indicates that x partially belongs to the cluster C_k; and if $x \in \dddot{C}_k$, then the object x does not belong to the cluster C_k.

Three-way clustering represents clustering results using a pair of nested sets:

$$C_k = \left(\hat{C}_k, \tilde{C}_k\right).$$

Clearly, $\dddot{C}_k = \mathbb{U} - \hat{C}_k - \tilde{C}_k$. Generally, the core and fringe regions satisfy the following properties:

(1) $\forall C_k \in C,\ \hat{C}_k \neq \emptyset$,

(2) $\bigcup_{k=1}^{K} \left(\hat{C}_k \cup \tilde{C}_k\right) = \mathbb{U}$,

(3) $\hat{C}_j \bigcap \hat{C}_k = \emptyset,\ j \neq k.$

Property (1) ensures that the core region of each cluster contains at least one object, ensuring the core region is non-empty. Property (2) ensures that all objects can be effectively partitioned through clustering. Property (3) ensures that the core regions of different clusters do not overlap.

Based on the above discussion, the three-way clustering result can be represented as:

$$\mathbb{C} = \{(\hat{C}_1, \tilde{C}_1), (\hat{C}_2, \tilde{C}_2), \dots, (\hat{C}_k, \tilde{C}_k), \dots, (\hat{C}_K, \tilde{C}_K)\}.$$

When $\tilde{C}_k = \emptyset$, the 3WC result is the same as that of 2WC, where C_k is represented only by $\hat{C}_k$. Hence, 2WC is a special case of three-way clustering. By introducing a deferment option for ambiguous data (i.e., the fringe region), three-way clustering offers an effective way to handle uncertainty and theoretically reduces the expected decision risk.

2.2. Kernel density estimation (KDE)

Kernel density estimation is a non-parametric method for estimating the probability density function of a random parameter [29]. Particularly, KDE uses a smooth kernel function to calculate the probability density function from the dataset and capture the distribution characteristics. For each sample in the universe $\mathbb{U} = \{x_1, x_2, \dots, x_n\}$, the probability density function $\hat{f}_h(x)$ in the feature space can be calculated as:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} F\left(\frac{x - x_i}{h}\right), \tag{1}$$

where $F(\cdot)$ is the kernel function, $h > 0$ is the bandwidth parameter controlling the smoothness of the estimate, and n is the total number of samples. KDE satisfies the following properties:

(1) Non-parametric: KDE does not assume any data distribution form, relying only on the universe $\mathbb{U} = \{x_1, x_2, \dots, x_n\}$.
(2) Consistency: When the sample size $n \to \infty$ and the bandwidth is appropriately chosen, $\lim_{n\to\infty} \hat{f}_h(x) = f(x)$.
(3) Regularity: For all x and $h > 0$, $\hat{f}_h(x) \geq 0$ and $\int_{-\infty}^{\infty} \hat{f}_h(x)\, dx = 1$.
(4) Tuning: There exists an optimal bandwidth h^* that minimizes the mean squared error, i.e., $h^* = \arg\min_h E\left[\int \left(\hat{f}_h(x) - f(x)\right)^2 dx\right]$, where $f(x)$ is the true density function, and $E[\cdot]$ is the expectation operator.

3. The proposed algorithm

This section first presents the steps of the Up3WC algorithm, shown in Fig. 2, followed by an analysis of its time complexity.

3.1. Multi-scale uncertainty propagation three-way clustering

In clustering analysis, accurately characterizing the relationship between samples and clusters is essential for improving clustering performance. As illustrated in Fig. 3, samples are typically located in either core or fringe regions of clusters, where both spatial position and local density structure jointly influence the certainty of their cluster memberships. Relying solely on the distance to the cluster center may ignore the continuity of density in the surrounding region; conversely, basing the judgment solely on density may fail to capture the sample's distance affiliation in the overall dataset. Thus, a systematic analysis of the sources of uncertainty is necessary. Most existing methods often employ single-scale evaluation strategies, which are inadequate for identifying the uncertainty characteristics of boundary samples. These samples typically exhibit two features: proximity to multiple cluster centers and location in transitional density regions. Neither distance nor density alone is adequate to determine their true affiliations. Accordingly, a

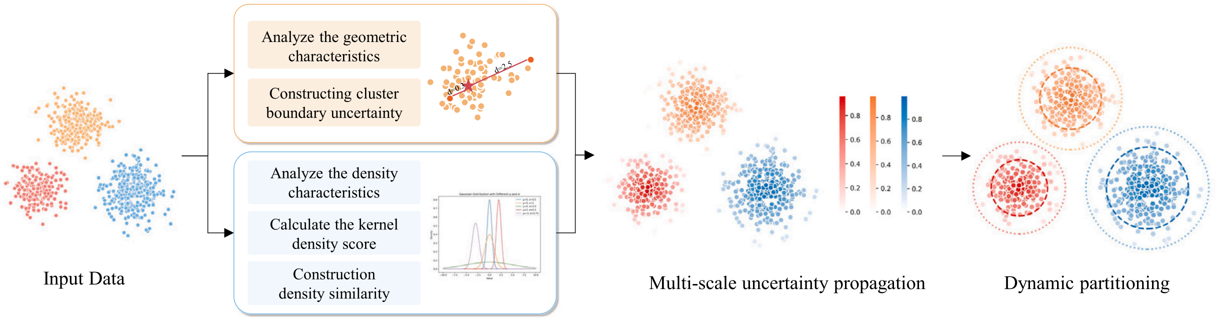

Analyze the geometric characteristics
Constructing cluster boundary uncertainty
d=0.3
d=2.5
Analyze the density characteristics
Calculate the kernel density score
Construction density similarity
Input Data
0.8
0.6
0.4
0.2
0.0
Multi-scale uncertainty propagation
Dynamic partitioning

Fig. 2. The framework of Up3WC. The method first computes distance-based boundary uncertainty and density-based similarity, then iteratively propagates uncertainty to obtain a stable multi-scale estimation, and finally performs dynamic three-way partitioning to generate core, fringe, and trivial regions.

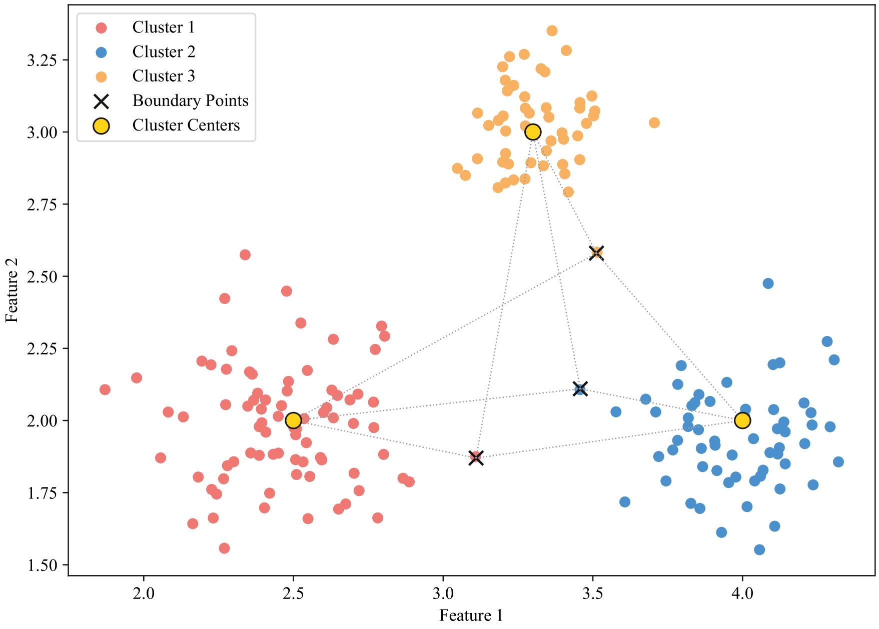


Fig. 3. The relationship between samples and clusters. The boundary samples are positioned near multiple cluster centers and in regions of lower local density, highlighting that uncertainty stems from the joint influence of distance and density.

sample's uncertainty should be viewed as the result of the combined influence of distance and density information, reflecting its ambiguous membership in the context of the entire dataset.

Kernel density estimation adaptively captures the local structure of data, making it suitable for modeling regional variations in heterogeneous density distributions. Among common kernels, the Gaussian kernel is adopted in this study for its C^∞ smoothness and infinite support. These properties ensure a globally connected density structure with smooth gradients, which are critical for the stability of the subsequent uncertainty propagation. The bandwidth h critically affects estimation precision; therefore, we adopt Scott's rule to automatically determine h based on sample size and dimensionality. This provides a robust, scale-adaptive bandwidth that balances smoothness and bias without manual intervention. Given that KDE is used here to modulate relative affinities rather than for exact density recovery, such a consistent bandwidth is sufficient to prevent the over-amplification of local fluctuations in fringe regions.

Definition 1 (Kernel Density Score, KDS). Let $\mathbb{U} = \{x_1, x_2, \ldots, x_n\}$ be a nonempty finite set of objects. For $\forall x_i \in \mathbb{U}$, its density value in the feature space can be obtained by the kernel density estimation method. Using the Gaussian kernel function, the kernel density score of $x_i \in \mathbb{U}$ is defined as:

$$\rho\left(x_i\right) = \frac{1}{nh\sqrt{2\pi}} \sum_{j=1}^{n} e^{-\frac{1}{2}(\frac{x_i - x_j}{h})^2}, h = n^{-\frac{1}{d+4}}, \tag{2}$$

where n is the number of samples, d is the dimensionality of the samples (i.e., the number of features), and h is the bandwidth of the Gaussian kernel.

The kernel density score quantifies the relative significance of each sample based on its density. Points in high-density regions, typically corresponding to core regions, are assigned higher KDS values, whereas points in low-density regions receive lower scores and are associated with fringe or trivial regions. As a non-parametric approach, kernel density estimation provides an accurate characterization of data distributions, particularly for datasets with complex structures. The resulting density estimates capture regional variations, forming the basis for both similarity calculations and uncertainty propagation.

As a non-parametric approach, kernel density estimation effectively characterizes complex data distributions, and the resulting density estimates capture regional variations that support both similarity calculations and uncertainty propagation.

To characterize pairwise similarity and reveal the underlying density structure, an adjacency matrix $A \in \mathbb{R}^{n\times n}$ is constructed based on the kernel density score $\rho(x_i)$. The similarity between two samples is defined as the minimum of their respective density values, i.e., $\min(\rho(x_i), \rho(x_j))$, which represents the common density level shared by both samples. This conservative measure assigns high similarity only when both samples lie in high-density regions, while automatically suppressing connections involving low-density or boundary points. Consequently, it prevents high-density samples from dragging sparse ones across cluster boundaries and enhances the discriminability between core, fringe, and trivial regions.

Definition 2 (Density Similarity). Let $\mathbb{U} = \{x_1, x_2, \dots, x_n\}$ be a nonempty finite set of objects. For $\forall x_i,\ x_j \in \mathbb{U}$, the density similarity between sample x_i and sample x_j is defined as:

$$A_{ij} = \begin{cases} \min\left(\rho\left(x_i\right), \rho\left(x_j\right)\right) & if\ i \neq j \\ 0 & if\ i = j \end{cases}, \tag{3}$$

where $\rho\left(x_i\right)$ represents the KDS of the sample x_i, and $\min(\cdot)$ is the function that returns the minimum value.

From a theoretical perspective, our choice of the minimum operator is motivated by boundary preservation in density-based uncertainty propagation. For positive density scores ρ_i, ρ_j (normalized to $[0, 1]$), the following inequality holds: $\min(\rho_i, \rho_j) \le \mathrm{HM}(\rho_i, \rho_j) \le \mathrm{GM}(\rho_i, \rho_j) \le \mathrm{AM}(\rho_i, \rho_j)$, where HM, GM, and AM denote the harmonic, geometric, and arithmetic means, respectively (with equality when $\rho_i = \rho_j$). Consider a typical boundary case connecting a high-density core sample ($\rho_c \approx 1$) and a low-density boundary or noise sample ($\rho_n \approx 0$). Averaging operators such as AM, GM, or HM allow the large ρ_c to partially compensate for the small ρ_n, yielding a non-negligible edge weight and thus risking "density leakage" across sparse regions. In contrast, $\min(\rho_c, \rho_n)$ enforces a bottleneck constraint: the connection strength is upper-bounded by the lower-density point. As a result, strong links are formed only when both samples lie in high-density areas, which suppresses cross-boundary connections and improves the discriminability between core and fringe structures during uncertainty propagation.

Based on the definition of density similarity, the adjacency matrix $A = [A_{ij}]$ is an $n \times n$ symmetric matrix, where each element A_{ij} denotes the similarity between samples x_i and x_j. Note that we set the diagonal entries to zero, i.e., $A_{ii} = 0$, because the self-uncertainty of each sample is already preserved by the restart term $(1-\alpha)U^{(0)}$ in the propagation update (Eq. 6). This avoids redundant self-reinforcement and allows A to focus on encoding inter-sample relations. Specifically, the matrix has the following properties:

- Symmetry: $A_{ij} = A_{ji},\ \forall i, j \in \{1, \dots, n\}$, indicating that the similarity between samples x_i and x_j is mutual.
- Zero diagonal: $A_{ii} = 0,\ \forall i \in \{1, \dots, n\}$, meaning that self-similarity is ignored so that the adjacency reflects only interactions among distinct samples.
- Non-negativity: $A_{ij} \ge 0,\ \forall i, j \in \{1, \dots, n\}$, ensuring the validity of the similarity measure.

To ensure stability and rationality, the adjacency matrix is normalized. Let D be the degree matrix, where each diagonal element is defined as $D_{ii} = \sum_j A_{ij}$. The normalized weight matrix W is then computed as:

$$W = D^{-1/2} A D^{-1/2}, \tag{4}$$

where $D^{-1/2}$ represents the inverse square root of the degree matrix. After normalization, the elements of W represent the weighted similarity between samples. This process mitigates the influence of degree imbalance and scale variations in pairwise similarities, ensuring that each sample contributes relatively equally to the final result. The symmetric normalization further confines the spectrum of W within $[0, 1]$, stabilizing diffusion and improving the convergence behavior of the propagation.

The objective of three-way clustering is to partition similar samples within the same cluster into core and fringe regions. Adopting a strategy that favors lower density values ensures that, during the clustering process, samples in high-density regions maintain high similarity even if their actual densities differ slightly. Meanwhile, in low-density regions, samples that are spatially close are prevented from being mistakenly grouped into high-density areas, thereby achieving more accurate cluster boundaries.

However, relying solely on static similarity matrices is insufficient to distinguish between core and fringe regions of samples accurately. To further improve the accuracy and stability of clustering results, an uncertainty propagation mechanism is introduced. This mechanism iteratively updates the uncertainty of each sample by integrating both distance and density information. Typically, points situated farther from their cluster centers exhibit higher uncertainty, as they are likely located near the cluster boundaries or in regions close to neighboring clusters. These points are assigned higher initial uncertainty values. As shown in Fig. 4, sample x_1 is within the core region of the cluster, with a distance of 0.6 from the cluster center, while sample x_2 is in the fringe region with a distance of 2.8 from the center. Clearly, points like x_1, closer to the cluster center, demonstrate greater stability, whereas points near the cluster boundary, like x_2, exhibit higher uncertainty and are more likely to be influenced by neighboring clusters. Therefore, this

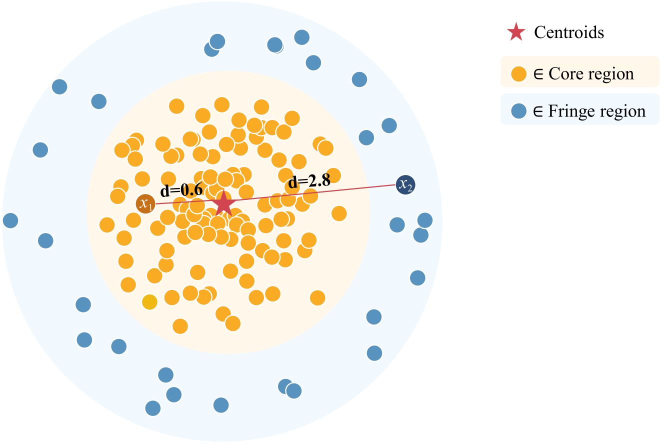


Fig. 4. The uncertainty of the geometric distance between the samples and the cluster.

study combines the inherent uncertainty of the data with the relative position of the cluster centers. Each sample's initial uncertainty value, or boundary uncertainty, is calculated by measuring the distance to its corresponding cluster center.

Definition 3 (Boundary Uncertainty). Let $\mathbb{U} = \{x_1, x_2, \dots, x_n\}$ be a nonempty finite set of objects. For $\forall x_i \in \mathbb{U}$, its boundary uncertainty $\Upsilon(x_i)$ is defined as:

$$\Upsilon(x_i) = \left\| x_i - \mu_{k_i} \right\|_2, \tag{5}$$

where $\mu_{k_i} = \frac{1}{|C_k|} \sum_{x_i \in C_k} x_i$ is the cluster center of the sample x_i, $\Upsilon(x_i) \geq 0$.

A larger value of $\Upsilon(x_i)$ indicates that the sample is farther from the cluster center, implying it lies in the fringe region with higher uncertainty. Conversely, a smaller value suggests that the sample is closer to the center, belonging to the core region with lower uncertainty. The calculated boundary uncertainty is then assigned as the initial uncertainty value $U^{(0)}$.

After obtaining the initial boundary uncertainty, the uncertainty of each sample is further refined by propagating uncertainty among density-similar samples. Concretely, the adjacency matrix A is constructed from kernel density scores and then symmetrically normalized to obtain the weight matrix W, where each entry W_{ij} measures how much the uncertainty of x_j contributes to updating x_i. Let $U^{(t)}$ denote the uncertainty vector at the t-th iteration, with $U^{(0)}$ as initialization. The uncertainty is updated iteratively via the following propagation rule:

$$U^{(t+1)} = (1 - \alpha)U^{(0)} + \alpha W U^{(t)}, \tag{6}$$

where $WU^{(t)}$ acts as a weighted aggregation of uncertainty from density-similar samples. This update progressively aligns local boundary uncertainty with the underlying density structure, leading to a more stable and discriminative estimation. The propagation coefficient $\alpha \in [0, 1]$ controls the balance between retaining the local initialization and incorporating propagated context.

Here, α balances the contributions from the initial distance-derived uncertainty $U^{(0)}$ and the density-diffused uncertainty $WU^{(t)}$. A larger α allows uncertainty to propagate more extensively across similar samples, while a smaller α preserves local specificity. The convergence of this linear iteration is theoretically guaranteed because the symmetric-normalized weight matrix $W = D^{-1/2}AD^{-1/2}$ has a spectral radius $\rho(W) \leq 1$, where $\rho(W)$ denotes the largest absolute eigenvalue of W. Consequently, the iterative update is a linear contraction with rate $\alpha\rho(W) < 1$, ensuring convergence to a unique fixed point and numerical stability. The number of iterations t mainly depends on the spectral gap of W rather than the data size n, providing numerical stability and scalability.

When α approaches 0, the clustering is more influenced by the initial distance information, whereas when α approaches 1, the results depend more on the density information. The propagation iterates until convergence, which is satisfied when:

$$\|U^{(t+1)} - U^{(t)}\| \leq \varepsilon, \tag{7}$$

where ε is the convergence threshold, a sufficiently small number.

The uncertainty propagation process diffuses local uncertainty information across the dataset by integrating distance-based initial uncertainty with density-based neighborhood relationships. This mechanism enables each sample's uncertainty to reflect both its local context and the global data structure, thereby achieving a balanced and comprehensive representation of uncertainty. Through iterative diffusion, the propagated uncertainties converge smoothly, enhancing the overall stability and robustness of the clustering results.

Based on the final uncertainty scores after the convergence of uncertainty propagation and the distance characteristics of each object, this study proposes a dynamic three-way classification strategy. To mitigate the influence of boundary outliers and ensure a stable cluster-wise benchmark, we select the median as the threshold. Compared with the arithmetic mean or fixed quantiles, the median is robust to skewness and insensitive to extreme uncertainty values, which are prevalent near cluster boundaries. Moreover, unlike parametric adaptive strategies, the median offers a parameter-free adaptive solution, avoiding the introduction of additional complexity. Let $\mathbb{U} = \{x_1, x_2, \dots, x_n\}$ denote a non-empty finite set, where n is the number of objects. Given the two-way clustering result $C = \{C_1^{twi}, C_2^{twi}, \dots, C_k^{twi}, \dots, C_K^{twi}\}$ of the universe, for all samples $x_i \in C_k^{twi}$, the uncertainty benchmark U_k^* is defined as:

$$U_k^* = median(U_i^{(T)} \mid x_i \in C_k^{twi}), \tag{8}$$

where $U_i^{(T)}$ is the convergence uncertainty of sample x_i, $median(S)$ is the median function, which returns the middle value of the set S when the values are arranged in order. If there is an even number of elements, the median is the average of the two middle values. The effective distance benchmark ed_k^* is defined as:

$$ed_k^* = median(ed_i \mid x_i \in C_k^{twi}), \tag{9}$$

where $ed_i = \|x_i - \mu_k\|_2$ is the distance from the sample x_i to the center μ_k of its assigned cluster C_k^{twi}.

After the propagation converges, the dynamic three-way assignment iteratively refines cluster structures. In each iteration, samples are reassigned using the current uncertainty U and effective distances, then cluster centroids μ_k are updated based only on the core regions $\hat{C}_k$ to reduce the influence of boundary noise. The effective distances $ed_i = \|x_i - \mu_k\|_2$ are recalculated accordingly, and the process repeats until assignments remain unchanged. This iterative mechanism ensures that centroid updates consistently rely on the most stable core-based structure.

In a cluster, the core region typically includes samples with high density and low uncertainty that are located closer to the cluster center, whereas the fringe region consists of samples with low density or greater distance from the center and thus exhibits higher uncertainty. Accordingly, the samples in cluster C_k are classified into three distinct regions: the core region $\hat{C}_k$, the fringe region $\tilde{C}_k$, and the trivial region $\dddot{C}_k$. The specific classification strategy is as follows:

(1) For any object x_i in cluster C_k, if its uncertainty $U_i^{(T)}$ is less than or equal to the uncertainty benchmark U_k^*, then x_i is assigned to the core region $\hat{C}_k$.
(2) For any object x_i in cluster C_k, if its uncertainty $U_i^{(T)}$ is greater than the uncertainty benchmark U_k^*, but its effective distance ed_i is less than or equal to the effective distance benchmark ed_k^*, then x_i is assigned to the fringe region $\tilde{C}_k$.
(3) For any object x_i in cluster C_k, if its uncertainty $U_i^{(T)}$ is greater than the uncertainty benchmark U_k^* and its effective distance ed_i is greater than the effective distance benchmark ed_k^*, it is assigned to the trivial region $\dddot{C}_k$.

According to the above criteria, $\hat{C}_k$ contains confident members, $\tilde{C}_k$ highlights uncertain boundary samples, and $\dddot{C}_k$ mainly captures noise or highly ambiguous samples. Such a three-way result enables risk-aware usage: trust $\hat{C}_k$, inspect or defer $\tilde{C}_k$, and filter $\dddot{C}_k$ when robust clustering is required.

3.2. Algorithm and complexity analysis of Up3WC

In summary, the proposed multi-scale uncertainty propagation three-way clustering algorithm is detailed in Algorithm 1, and its flowchart is illustrated in Fig. 5. Its computational complexity is analyzed as follows.

Line 1 computes the kernel density $\rho(x_i)$ for each sample, requiring $O(n)$ operations. Lines 2–7 construct the symmetric adjacency matrix. Since only the upper triangular part needs to be computed, $\frac{n(n+1)}{2}$ pairwise similarities are evaluated, leading to a time complexity of $O(n^2)$. After normalization (line 8), the adjacency matrix is transformed into the weight matrix W using Eq. (4), which is represented in a sparse form with complexity $O(n+m)$, where m denotes the number of nonzero elements. Lines 9–11 calculate the cluster boundary uncertainty for all samples based on Eq. (5), with a complexity of $O(K \times n)$, where K is the number of clusters. Lines 12–14 perform uncertainty propagation according to Eq. (6), requiring $O(t \times n)$ operations, where t is the number of iterations until convergence. Although a maximum iteration limit is predefined, the actual number of iterations t before convergence mainly depends on the spectral gap of W rather than the data size n, providing numerical stability and scalability. Finally, lines 15–26 execute the dynamic three-way assignment process, whose complexity is $O(n)$. Considering all components together, the overall time complexity of the Up3WC algorithm is $O(n^2 + t \times n)$.

4. Experimental design and analysis

In this section, comparative experiments are conducted to evaluate the effectiveness of the proposed Up3WC algorithm by comparing it with 8 algorithms on 16 UCI datasets.

4.1. Experimental setup

The experiments were conducted in Python on WIN 11; Intel(R) Core(TM) i7-12700H 2.30 GHz CPU; 16 GB RAM. For performance evaluation, 16 datasets were selected from the UCI Machine Learning Data Repository (https://archive.ics.uci.edu). These datasets differ in the number of clusters, sample size, and feature dimensions. Detailed descriptions are provided in Table 1.

The proposed Up3WC algorithm is designed to refine uncertainty characterization by combining distance and density information, independent of the specific two-way clustering (2WC) backbone. As the original 2WC procedure requires no modification, Up3WC

Algorithm 1: Proposed Up3WC algorithm.

```
Input: Two-way clustering result $C = \{C_1^{twi}, C_2^{twi}, \dots, C_k^{twi}, \dots, C_K^{twi}\}$, number of clusters $K$, propagation coefficient $\alpha$
Output: Three-way clustering result $\mathbb{C} = \{(\hat{C}_1, \tilde{C}_1), (\hat{C}_2, \tilde{C}_2), \dots, (\hat{C}_k, \tilde{C}_k), \dots, (\hat{C}_K, \tilde{C}_K)\}$
Calculate the kernel density score $\rho\left(x_i\right)$ for each sample $x_i$ according to Eq. (2).
for $i = 1 : n$ do
    for $j = i : n$ do
        Compute the adjacency matrix using Eq. (3);
        $A[j, i] = A[i, j]$.
    end
end
Compute the normalized adjacency matrix to obtain the weight matrix $W$ by Eq. (4).
for $i = 1 : n$ do
    Calculate the cluster boundary uncertainty for sample $x_i$ by Eq. (5).
end
repeat
    Propagate uncertainty according to Eq. (6), and the updated uncertainty $U^{(t+1)}$ is obtained.
until $\left\| U^{(t+1)} - U^{(t)} \right\| \le \varepsilon$;
for $k = 1 : K$ do
    Calculate the uncertainty benchmark $U_k^*$ and effective distance benchmark $ed_k^*$ for $C_k^{twi}$ by Eqs. (8) and (9).
    for $x_i \in C_k^{twi}$ do
        if $U_i^{(T)} \le U_k^*$ then
            $\hat{C}_k = \hat{C}_k \bigcup \{x_i\}$;
        else if $U_i^{(T)} > U_k^*$ and $ed_i \le ed_k^*$ then
            $\tilde{C}_k = \tilde{C}_k \bigcup \{x_i\}$;
        else if $U_i^{(T)} > U_k^*$ and $ed_i > ed_k^*$ then
            $\dddot{C}_k = \dddot{C}_k \bigcup \{x_i\}$.
        end
    end
end
```

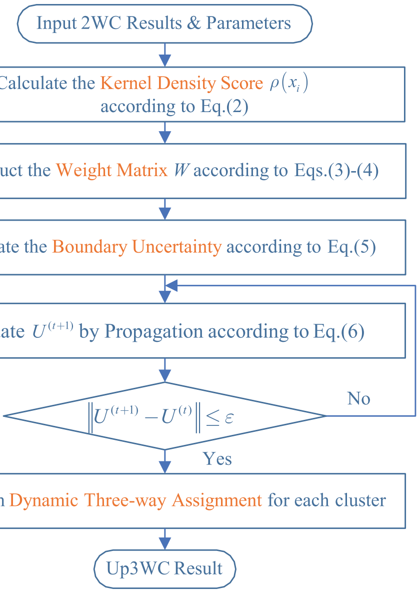

Input 2WC Results & Parameters
Calculate the Kernel Density Score $\rho(x_i)$ according to Eq.(2)
Construct the Weight Matrix W according to Eqs.(3)-(4)
Calculate the Boundary Uncertainty according to Eq.(5)
Update $U^{(t+1)}$ by Propagation according to Eq.(6)
$\left\| U^{(t+1)} - U^{(t)} \right\| \le \varepsilon$
No
Yes
Perform Dynamic Three-way Assignment for each cluster
Up3WC Result

Fig. 5. The flowchart of Up3WC.

serves as a generic three-way refinement framework capable of directly accepting results from various algorithms, thereby ensuring broad applicability. To verify its effectiveness, the framework is integrated with two representative algorithms, K-means and density peaks clustering (DPC), constructing two three-way variants denoted as K-means-Up3WC (KUp3WC) and DPC-Up3WC (DUp3WC).

Table 1
UCI datasets used in experiments.

ID	Datasets	Number of samples	Dimensions	Number of classes
1	Iris	150	4	3
2	Wine	178	13	3
3	Seeds	210	7	3
4	Glass	214	9	6
5	Winequality	215	5	3
6	Heartdisease	303	13	5
7	Liver	345	6	2
8	Libras	360	90	15
9	WDBC	569	30	2
10	Balancescale	625	4	3
11	Vehicle	946	18	4
12	Vowel	990	14	11
13	Maternalhealthrisk	1014	5	3
14	Bank	1372	4	2
15	Segment	2310	19	7
16	Waveform	5000	21	3

Table 2
Parameter settings.

Algorithms	Parameter value 1	Parameter value 2	Parameter value 3
FCM	$m = 2$	error = 0.0005	max iterations = 1000
TWKM	$\alpha = 0.02$	$\beta = 0.00023 * N$	——
DBSCAN	$eps = 0.3$	min_samples = 4	——
CE3	$q = 8$	——	——
3WG	$\omega_{co} = 0.5$	——	——
Up3WC	$\alpha \in [0, 1]$	$\varepsilon = 10^{-6}$	——

For a comprehensive performance evaluation, KUp3WC and DUp3WC are compared against eight established clustering algorithms: 3WG [23], CE3 k-means (CE3) [19], TWKM [21], DPC, density-based spatial clustering of applications with noise (DBSCAN), Fuzzy C-means (FCM), K-medoids, and K-means. 3WG, CE3, and TWKM are three-way clustering algorithms, where CE3 and TWKM are improvements of the three-way clustering method based on K-means, while 3WG is a three-way clustering method based on granular ball. DPC and DBSCAN are density-based two-way clustering algorithms capable of handling clusters with non-linear shapes. FCM is a benchmark soft clustering algorithm, and K-medoids and K-means are two representative two-way clustering algorithms.

In comparative experiments, five widely used clustering metrics are employed to evaluate the performance of Up3WC and the comparison algorithms: Adjusted Rand Index (ARI) [30], Davies-Bouldin Index (DBI) [31], Normalized Mutual Information (NMI) [32], Adjusted Mutual Information (AMI) [33], and F1 Score (F1) [34]. Specifically, both ARI and AMI take values in $[-1, 1]$, where larger values indicate higher consistency between the clustering results and the ground-truth labels. NMI and F1 take values in $[0, 1]$, with 0 and 1 representing poor and perfect clustering, respectively. DBI is non-negative, i.e., DBI $\in [0, +\infty)$, and smaller values signify higher intra-cluster compactness and better clustering performance.

4.2. Experimental results

To objectively evaluate algorithm performance, each algorithm was run within its parameter range, and the best clustering result for each dataset was recorded. For all algorithms except DBSCAN, the number of clusters K was pre-set to the number of ground-truth labels of each dataset. To ensure a fair and reproducible comparison, the parameters for baseline algorithms were set to the standard values recommended in their original literature, as detailed in Table 2. For the proposed Up3WC, the propagation coefficient α was determined through a systematic search in the range $[0, 1]$ for each dataset, which is further analyzed in Section 4.4. Regarding the propagation mechanism (Eq. 6), the convergence threshold was fixed at $\varepsilon = 10^{-6}$ to guarantee numerical precision. Theoretically, since the normalized propagation matrix W satisfies $\rho(W) \leq 1$, the propagation coefficient $\alpha \in (0, 1)$ ensures a stable fixed-point iteration for uncertainty propagation. Empirically, Up3WC converges within 4–5 iterations across datasets of different sizes (tested up to $n = 5000$ on Waveform), indicating that the required iteration number is largely insensitive to sample size. While the runtime increases with n due to matrix operations, the convergence behavior remains consistently efficient.

In this experiment, each feature was scaled to the unit interval through min-max normalization to eliminate scale differences among features while maintaining their original value distributions. Each experiment was repeated 30 times, and the average value was recorded as the experimental result of the algorithm to compare the overall performance. The detailed comparison results are reported in Tables 3–5, which correspond to the five evaluation metrics: F1, DBI, NMI, ARI, and AMI. In each table, the best performance for each dataset is highlighted in bold. Furthermore, Fig. 6 visualizes the performance variations of the 10 clustering algorithms across the 16 benchmark datasets with respect to these five metrics (Tables 6 and 7).

Tables 3–5 demonstrate that the proposed KUp3WC algorithm achieves superior performance across all evaluation metrics on 8 of the 16 datasets. Moreover, it obtains optimal values for DBI, ARI, AMI, and NMI metrics on at least 10 datasets, confirming the

Table 3
Performance comparison of F1 score on 16 UCI datasets (Higher is better).

Dataset	K-means	K-medoids	FCM	DBSCAN	DPC	TWKM	CE3	3WG	DUp3WC	KUp3WC
Segment	0.5650	0.6887	0.6664	0.4448	0.5749	0.5806	0.5714	0.6240	**0.6980**	0.6883
Glass	0.4640	0.5125	0.4911	0.4189	0.3414	0.4737	0.5316	0.5332	**0.6302**	0.5562
Libras	0.4001	0.3110	0.0576	0.0530	0.2574	0.4204	0.4154	0.3951	0.3523	**0.5556**
Bank	0.5752	0.3966	0.6083	0.3966	0.3966	0.5749	0.4957	0.5741	0.4243	**0.6886**
WDBC	0.9294	0.9381	0.9309	0.7677	0.5462	0.9334	0.8941	0.9711	0.6158	**0.9859**
Vehicle	0.3334	0.3237	0.3298	0.2134	0.3131	0.3332	0.3374	0.3444	0.3820	**0.4285**
Waveform	0.4181	0.4384	0.4269	0.1731	0.4864	0.4150	0.4183	0.4186	**0.5861**	0.4700
Balancescale	0.6062	0.6660	0.5877	0.2907	0.6248	0.6206	0.6201	0.6584	**0.7155**	0.6817
Liver	0.4255	0.4255	0.5009	0.3947	0.5159	0.4173	0.4499	0.4255	**0.5586**	0.4799
Seeds	0.8881	0.8905	0.9002	0.6156	0.7857	0.8899	0.8518	0.9621	0.8418	**1.0000**
Iris	0.8503	0.8992	0.8926	0.6029	0.8593	0.8049	0.8120	0.8853	0.8732	**0.9742**
Wine	0.9192	0.9021	0.9488	0.6804	0.7567	0.9699	0.9192	0.9436	0.7768	**1.0000**
Vowel	0.2425	0.2017	0.1487	0.2818	0.2358	0.2600	0.2476	0.2836	0.2993	**0.4169**
Winequality	0.4875	0.4580	0.4483	0.2822	0.4802	0.4964	0.4915	0.5055	0.4879	**0.5595**
Heartdisease	0.4619	0.4905	0.4799	0.3896	0.5084	0.4769	0.5087	0.4845	0.5218	**0.5911**
Maternal	0.4724	0.4307	0.5369	0.3508	0.2290	0.5247	0.4695	0.4720	0.2897	**0.5741**
Average	0.5649	0.5608	0.5597	0.3973	0.4945	0.5745	0.5646	0.5926	0.5658	**0.6657**

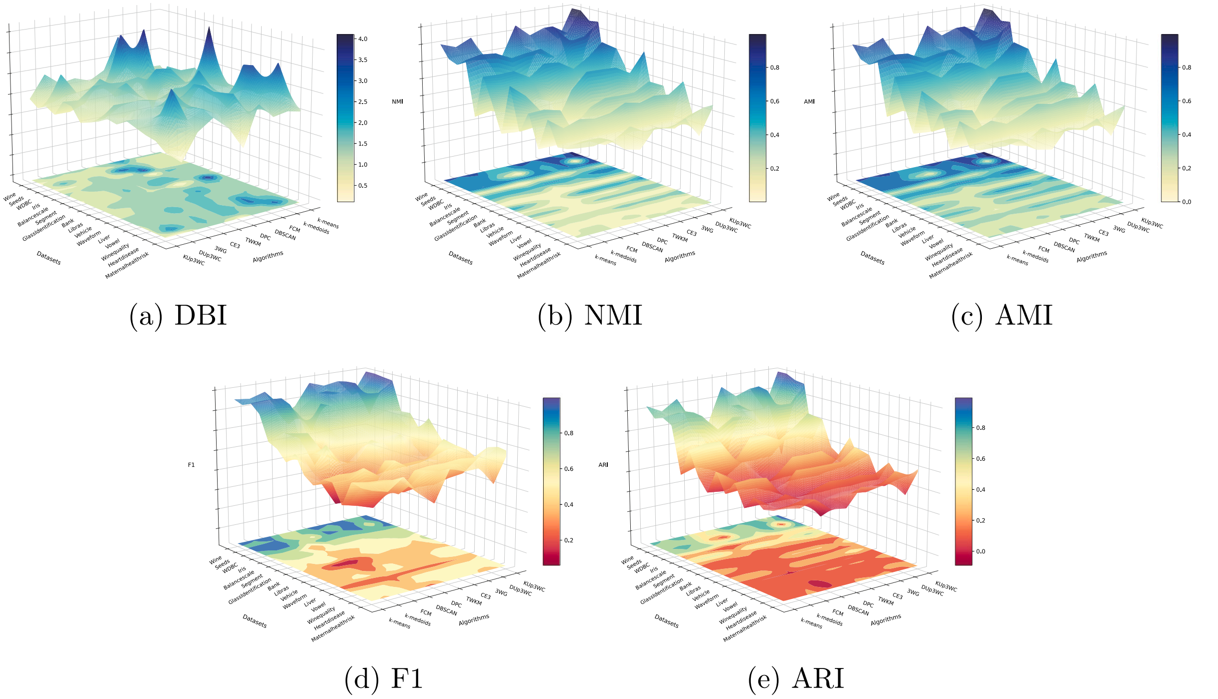
Table 4
Performance comparison of DBI on 16 UCI datasets (Lower is better).

Dataset	K-means	K-medoids	FCM	DBSCAN	DPC	TWKM	CE3	3WG	DUp3WC	KUp3WC
Segment	1.2833	1.2360	1.2159	1.9225	1.7579	1.2069	1.6151	1.1889	**1.0004**	1.1791
Glass	1.1514	1.4398	1.5344	2.3242	2.4379	1.1586	2.0238	1.0364	0.9319	**0.5457**
Libras	1.3091	1.9162	1.7764	1.7033	1.7763	1.3254	1.4915	1.5046	1.5240	**0.9036**
Bank	1.1909	4.2685	1.1981	0.0000	1.1734	1.1550	1.3138	1.1911	0.7686	**0.7103**
WDBC	1.1213	1.1358	1.1327	2.4470	3.6807	1.1073	1.2150	1.1203	2.1953	**0.7332**
Vehicle	1.4292	1.5477	1.5268	2.6491	1.4865	1.3852	1.5348	1.3695	**1.0683**	1.1654
Waveform	1.4971	1.5051	1.5457	1.6488	2.2400	1.4588	1.4972	1.4970	1.1672	**1.0717**
Balancescale	1.7235	1.7200	2.0422	0.0000	1.7780	1.6562	1.6964	1.7348	**1.1778**	1.2428
Liver	**1.3141**	1.8333	1.6190	1.5517	2.2945	1.3438	2.3201	1.3955	1.3338	1.3823
Seeds	0.8766	0.8773	0.8767	1.7295	1.0967	0.8727	0.9386	0.8759	0.6425	**0.5059**
Iris	0.7791	0.7686	0.7746	3.9499	0.8468	0.7748	1.8185	0.7610	**0.4774**	0.4845
Wine	1.3292	1.3385	1.3181	3.1543	1.5806	1.1860	1.4360	1.3096	1.8816	**0.8168**
Vowel	1.6291	1.6402	3.1507	1.2271	2.1705	1.5748	1.7090	1.6302	1.4500	**1.2161**
Winequality	1.6160	2.2515	3.2395	1.7298	2.2395	1.5508	1.8620	1.4454	1.8879	**1.0507**
Heartdisease	1.5266	3.4021	1.7132	1.2408	2.3433	1.5205	1.9249	1.5617	3.1524	**0.9469**
Maternal	1.1074	1.2591	1.1378	1.8623	0.0000	1.0763	1.6784	1.1844	0.0000	**0.5835**
Average	1.3052	1.7587	1.6126	1.8213	1.8064	1.2721	1.6297	1.3004	1.2912	**0.9086**

Table 5
Performance comparison of NMI on 16 UCI datasets (Higher is better).

Dataset	K-means	K-medoids	FCM	DBSCAN	DPC	TWKM	CE3	3WG	DUp3WC	KUp3WC
Segment	0.6052	0.5972	0.6086	0.5014	0.6316	0.6178	0.5078	0.6106	**0.7673**	0.6620
Glass	0.3309	0.2590	0.2960	0.3793	0.1325	0.3371	0.3419	0.4622	**0.4737**	0.4691
Libras	0.5936	0.4505	0.2402	0.0695	0.4193	0.5958	0.5912	0.5831	0.4877	**0.7120**
Bank	0.0169	0.0012	0.0329	0.0000	0.0086	0.0164	0.0178	0.0163	0.0412	**0.1068**
WDBC	0.6352	0.6568	0.6293	0.2746	0.0005	0.6511	0.5376	0.8016	0.0001	**0.9008**
Vehicle	0.1088	0.1275	0.0986	0.0371	0.1283	0.1126	0.0959	0.1593	0.1697	**0.1903**
Waveform	0.3641	0.3721	0.3302	0.0008	0.2939	0.3693	0.3642	0.3642	**0.4604**	0.3789
Balancescale	0.1029	0.1358	0.0833	0.0000	0.1847	0.1132	0.1202	**0.3232**	0.2843	0.2296
Liver	0.0000	0.0022	0.0007	0.0004	0.0025	0.0006	0.0058	0.0003	**0.0134**	0.0001
Seeds	0.6699	0.6873	0.6911	0.4905	0.5235	0.6730	0.6292	0.8597	0.7578	**1.0000**
Iris	0.7185	0.7578	0.7433	0.6644	0.6933	0.7045	0.5750	0.7419	0.7768	**0.9168**
Wine	0.8123	0.7507	0.8336	0.5257	0.5361	0.8975	0.7511	0.8155	0.6535	**1.0000**
Vowel	0.3615	0.2342	0.2962	0.2698	0.2942	0.3783	0.3430	0.3923	0.3709	**0.4901**
Winequality	0.0972	0.0878	0.0866	0.0079	0.0674	0.1042	0.0965	0.1042	0.0763	**0.1292**
Heartdisease	0.1180	0.1658	0.1644	0.0911	0.2222	0.1192	0.1306	0.0841	0.1997	**0.2533**
Maternal	0.1644	0.1501	0.1813	0.0392	0.0000	0.2047	0.1578	0.2375	0.0000	**0.3198**
Average	0.3562	0.3398	0.3323	0.2095	0.2587	0.3685	0.3291	0.4097	0.3458	**0.4849**

overall robustness of the proposed framework. In particular, KUp3WC maintains low DBI scores while accurately classifying all core region samples on the Seeds and Wine datasets. Compared to other three-way clustering algorithms based on K-means, such as 3WG, CE3, and TWKM, KUp3WC consistently outperforms them across the 16 datasets. When compared to the DPC algorithm, DUp3WC surpasses DPC on all 16 datasets, demonstrating the generalizability of the Up3WC framework.



(a) DBI

(b) NMI

(c) AMI

(d) F1

(e) ARI

Fig. 6. Comparative visualization of evaluation metrics on UCI datasets.

Table 6
Performance comparison of ARI on 16 UCI datasets (Higher is better).

Dataset	K-means	K-medoids	FCM	DBSCAN	DPC	TWKM	CE3	3WG	DUp3WC	KUp3WC
Segment	0.4572	0.5021	0.4999	0.2238	0.4809	0.4812	0.3853	0.5023	**0.6525**	0.5482
Glass	0.1737	0.1439	0.1614	0.2388	0.0384	0.1822	0.2011	0.2406	**0.3098**	0.2233
Libras	0.3091	0.1865	0.0721	0.0032	0.1395	0.3093	0.3041	0.3132	0.1752	**0.4137**
Bank	0.0212	−0.0003	0.0452	0.0000	0.0007	0.0209	0.0233	0.0205	0.0196	**0.1402**
WDBC	0.7384	0.7674	0.7427	0.2598	0.0025	0.7521	0.6327	0.8778	0.0003	**0.9439**
Vehicle	0.0798	0.0928	0.0745	0.0013	0.0775	0.0824	0.0742	0.1048	0.1165	**0.1648**
Waveform	0.2535	0.2657	0.2436	0.0000	0.2611	0.2546	0.2535	0.2536	**0.4020**	0.2677
Balancescale	0.1194	0.1780	0.1074	0.0000	0.1261	0.1338	0.1384	0.3091	**0.3186**	0.2554
Liver	−0.0011	0.0002	−0.0049	−0.0048	−0.0024	−0.0048	0.0033	−0.0046	**0.0090**	−0.0009
Seeds	0.6991	0.7064	0.7266	0.4143	0.5056	0.7031	0.6330	0.8941	0.6773	**1.0000**
Iris	0.6834	0.7430	0.7287	0.5464	0.6648	0.6477	0.5165	0.7163	0.7132	**0.9264**
Wine	0.8183	0.7262	0.8498	0.4264	0.4521	0.9136	0.7685	0.8368	0.5492	**1.0000**
Vowel	0.1580	0.0357	0.1228	0.0253	0.1186	0.1680	0.1491	0.1618	0.1372	**0.2424**
Winequality	0.0611	0.0319	0.0302	−0.0020	0.0416	0.0695	0.0657	0.0753	0.0266	**0.0854**
Heartdisease	0.0894	0.1101	0.1474	−0.0981	0.0841	0.0957	0.0723	0.1072	**0.2513**	0.1564
Maternal	0.1529	0.1171	0.1779	0.0280	0.0000	0.2097	0.1597	0.1859	0.0000	**0.2878**
Average	0.3008	0.2879	0.2953	0.1289	0.1869	0.3137	0.2738	0.3497	0.2724	**0.4159**

Specifically, on the Segment dataset, DUp3WC delivers significantly better results than all comparison algorithms, with KUp3WC ranking second. Although both clustering algorithms based on the Up3WC framework do not achieve the highest values for all evaluation metrics on the Glass, Liver, and Balance Scale datasets, they demonstrate the best performance in F1 and ARI metrics while maintaining competitive performance on the remaining measures. On the Libras dataset, KUp3WC improves the F1 score by over 10% compared to the second-best method and achieves a marked reduction in DBI.

In addition, on the WDBC dataset, KUp3WC slightly outperforms 3WG by 0.01 in F1, while showing a significant improvement in the DBI metric. On the Maternal dataset, both the DPC and DUp3WC algorithms fail to produce meaningful results, but KUp3WC achieves a substantial improvement in DBI. The results on the Bank dataset show that KUp3WC exceeds the second-best algorithm by more than 5% in DBI, NMI, ARI, and AMI, and by 10% in F1. Additionally, on the Bank and Balancescale datasets, DBSCAN with the fixed parameter setting $\varepsilon = 0.3$ (Table 2) assigns all samples to a single cluster. This degenerate one-cluster partition leads to an F1 score greater than zero but all other external indices equal to zero, illustrating an over-connectivity effect that can arise in single-scale density methods under a global threshold. In contrast, the Up3WC-based methods utilize distance-based boundary uncertainty to maintain separated clusters on these datasets. On the Vehicle, Waveform, Heartdisease, and Iris datasets, both KUp3WC and DUp3WC

Table 7
Performance comparison of AMI on 16 UCI datasets (Higher is better).

Dataset	K-means	K-medoids	FCM	DBSCAN	DPC	TWKM	CE3	3WG	DUp3WC	KUp3WC
Segment	0.6036	0.5956	0.6070	0.4960	0.6301	0.6161	0.5058	0.6090	**0.7652**	0.6592
Glass	0.3009	0.2284	0.2662	0.3579	0.0962	0.3062	0.3132	**0.4293**	0.4281	0.4248
Libras	0.5384	0.3764	0.2137	0.0441	0.3466	0.5356	0.5357	0.5285	0.3645	**0.6293**
Bank	0.0163	−0.0002	0.0324	0.0000	0.0080	0.0159	0.0172	0.0158	0.0401	**0.1059**
WDBC	0.6347	0.6563	0.6288	0.2703	−0.0008	0.6506	0.5370	0.8010	−0.0028	**0.9005**
Vehicle	0.1053	0.1240	0.0951	0.0237	0.1247	0.1089	0.0924	0.1560	0.1630	**0.1840**
Waveform	0.3639	0.3719	0.3299	0.0003	0.2936	0.3691	0.3639	0.3640	**0.4600**	0.3784
Balancescale	0.1000	0.1330	0.0799	0.0000	0.1818	0.1102	0.1174	**0.3210**	0.2791	0.2246
Liver	−0.0027	0.0001	−0.0015	−0.0022	0.0004	−0.0020	0.0033	−0.0022	**0.0092**	−0.0054
Seeds	0.6669	0.6846	0.6884	0.4827	0.5191	0.6701	0.6259	0.8577	0.7533	**1.0000**
Iris	0.7149	0.7547	0.7401	0.6591	0.6895	0.7006	0.5695	0.7387	0.7707	**0.9147**
Wine	0.8102	0.7480	0.8318	0.5204	0.5310	0.8963	0.7485	0.8135	0.6450	**1.0000**
Vowel	0.3470	0.2140	0.2801	0.2089	0.2778	0.3634	0.3282	0.3565	0.3385	**0.4648**
Winequality	0.0920	0.0825	0.0813	0.0007	0.0618	0.0985	0.0912	0.0989	0.0664	**0.1207**
Heartdisease	0.1000	0.1487	0.1547	0.0380	0.2065	0.1009	0.1098	0.0608	0.1635	**0.2199**
Maternal	0.1628	0.1486	0.1798	0.0339	0.0000	0.2032	0.1561	0.2348	0.0000	**0.3172**
Average	0.3472	0.3292	0.3255	0.1959	0.2479	0.3590	0.3197	0.3990	0.3277	**0.4712**

Table 8
Average runtime (seconds) of all methods across the 16 datasets.

Dataset	K-means	K-medoids	FCM	DBSCAN	DPC	TWKM	CE3	3WG	DUp3WC	KUp3WC
Iris	0.0340	0.0112	0.0104	0.0117	0.0075	0.0344	0.0744	0.0172	0.0163	0.0533
Wine	0.0161	0.0170	0.0167	0.0097	0.0088	0.0472	0.0800	0.0130	0.0204	0.0676
Glass	0.0196	0.0125	0.0210	0.0180	0.0130	0.1582	0.0913	0.0231	0.0237	0.0742
Seeds	0.0266	0.0163	0.0232	0.0162	0.0179	0.1231	0.0649	0.0127	0.0230	0.0682
Libras	0.0381	0.0266	0.0844	0.0858	0.0200	1.0349	0.1773	0.0727	0.0689	0.1547
Segment	0.0364	0.1611	0.1947	0.0604	0.2709	9.1607	0.9908	0.1332	1.7793	1.9959
Balancescale	0.0217	0.0241	0.2311	0.0143	0.0347	0.7608	0.1204	0.0193	0.1407	0.2411
Liver	0.0191	0.0175	0.0216	0.0196	0.0196	0.1857	0.0815	0.0136	0.0488	0.1105
WDBC	0.0199	0.0221	0.0347	0.0243	0.0313	0.2649	0.1739	0.0244	0.1242	0.2008
Vowel	0.0471	0.0505	0.5363	0.0695	0.0721	4.6724	0.2704	0.0766	0.3401	0.4635
Vehicle	0.0243	0.0292	0.0265	0.0270	0.0597	1.1144	0.2958	0.0429	0.2566	0.3549
Maternal	0.0225	0.0442	0.0805	0.0704	0.0547	1.2982	0.2318	0.0369	0.3478	0.4719
Waveform	0.0300	0.6003	0.2567	0.1788	1.2372	8.0980	2.5989	0.1329	9.7360	9.7257
Winequality	0.0325	0.0845	0.3542	0.1088	0.2502	6.5678	0.4269	0.0863	0.9378	0.9690
Bank	0.0249	0.0588	0.0331	0.0485	0.0928	1.0357	0.3936	0.0275	0.6181	0.6971
Heartdisease	0.0240	0.0172	0.0820	0.0224	0.0189	0.4041	0.0878	0.0297	0.0405	0.1075
Average	0.0273	0.0746	0.1254	0.0491	0.1381	2.1850	0.3850	0.0476	0.9076	0.9847

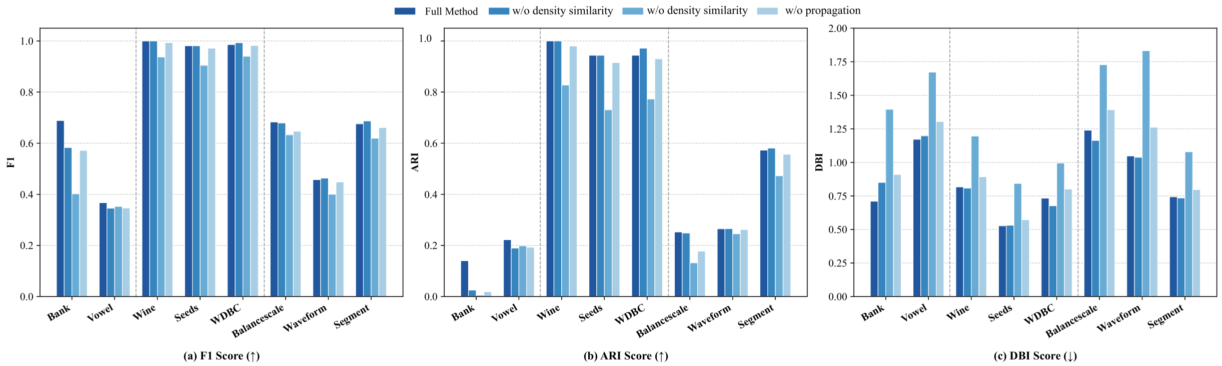
significantly outperform other algorithms, while on the Vowel and Winequality datasets, KUp3WC demonstrates clear superiority over all other methods.

Through comparative analysis, both KUp3WC and DUp3WC demonstrate significant improvements in clustering metrics across most datasets, particularly in reducing DBI. This confirms that the Up3WC algorithm effectively captures the intrinsic data structure by evaluating sample uncertainty at both density and distance scales. However, their specific effectiveness is intrinsically linked to the geometric characteristics of the dataset. KUp3WC, inheriting the variance-minimization property of K-means, demonstrates superior stability and efficiency on datasets with compact, spherical clusters, such as Wine and Bank. In contrast, DUp3WC leverages density peaks to effectively handle non-convex or manifold structures, such as Segment, capturing complex topologies that K-means-based methods might misinterpret. In practice, users are recommended to select KUp3WC for tasks involving data with regular, convex distributions, whereas DUp3WC is the preferred choice for datasets with complex, irregular shapes or varying densities. Beyond these specific instances, a major strength of Up3WC lies in its flexibility and generalizability. Whether applied with K-means, DPC, or extended to methods such as spectral clustering, Up3WC can incorporate an additional three-way classification step to enhance clustering results. This adaptability broadens its universality and scalability across diverse clustering scenarios. In conclusion, the proposed Up3WC algorithm effectively improves clustering accuracy and yields robust clustering results.

4.3. Runtime and memory consumption

To evaluate the practical efficiency of Up3WC, we report the average runtime of all methods on the 16 datasets under exactly the same experimental configuration as in Table 2. For each algorithm on each dataset, we ran 30 independent trials and recorded the mean runtime to mitigate randomness from system scheduling and initialization. The results are summarized in Table 8.

The additional cost of integrating Up3WC mainly arises from constructing and normalizing the density-similarity matrix, while the propagation itself converges within only a few iterations. This behavior can be clearly observed in representative datasets of different scales.



(a) F1 Score (↑)

(b) ARI Score (↑)

(c) DBI Score (↓)

Fig. 7. Ablation study illustrating the contribution of density similarity, boundary uncertainty, and uncertainty propagation to KUp3WC, evaluated by F1, ARI, and DBI.

On small-scale datasets such as Iris and Wine, the absolute overhead is minor: KUp3WC costs 0.053 s and 0.068 s, respectively, compared with 0.034 s and 0.016 s for the K-means backbone. In other words, Up3WC introduces only several tens of milliseconds of extra time on UCI-small tasks, remaining comparable to other lightweight three-way baselines.

On medium-scale datasets with more heterogeneous density, such as Bank and Winequality, the quadratic similarity construction becomes visible but still practical. KUp3WC runs in 0.697 s (Bank) and 0.969 s (Winequality), while DUp3WC requires 0.618 s and 0.938 s. Both are below 1 s and are consistently faster than TWKM (1.036 s and 6.568 s) and in the same order of magnitude as CE3, indicating that the proposed refinement is computationally competitive among three-way methods.

For the largest dataset Waveform ($n = 5000$), the runtime increases to about 9–10 s, which is consistent with the dense similarity construction. Importantly, even at this scale the propagation converges in a few iterations, so the runtime growth is dominated by matrix construction rather than iterative propagation. Therefore, Up3WC remains feasible for moderate-scale applications, while delivering substantial accuracy gains over its two-way backbones.

In terms of memory usage, the dominant cost stems from storing the density-similarity matrix A, which is then normalized to obtain W. Under a dense implementation, this requires $O(n^2)$ space, consistent with the time-complexity analysis. The uncertainty vectors and the three-way assignment procedure require only $O(n)$ additional memory. However, if W is stored in a sparse format with m nonzero entries, the space complexity reduces to $O(n+m)$. In particular, for large-scale data ($n > 10^4$), acceleration strategies such as kd-trees or other approximate nearest-neighbor search can be employed to retain only the top-k density-similar neighbors, yielding $m \approx kn$ and thus an $O(kn)$ memory footprint for W. This indicates that the $O(n^2)$ memory requirement arises from a dense similarity representation rather than an inherent limitation of Up3WC, whose propagation can be implemented on sparse neighborhoods without changing the update form.

4.4. *Ablation studies*

To better understand the contribution of each component in Up3WC, we conduct an ablation study by selectively removing or modifying key modules of the framework. In this study, K-means is adopted as the backbone, so the **Full Method** corresponds to KUp3WC. In particular, we focus on three core ingredients: (i) the density-based similarity, (ii) the distance-based boundary uncertainty, and (iii) the uncertainty propagation mechanism. We compare the Full Method against three variants:

(i) **w/o density similarity (Euclidean Only):** Replace the density-similarity adjacency with a purely Euclidean one, $A_{ij} = \exp\left(-\frac{\|x_i - x_j\|_2^2}{2\sigma^2}\right)$, while keeping $U^{(0)} = \|x_i - c_{k(i)}\|_2$ and the same propagation update. This variant isolates the role of density similarity and tests whether purely Euclidean distance information alone can support reliable uncertainty propagation.

(ii) **w/o boundary uncertainty (Kernel Only):** Keep the density-based adjacency $A_{ij} = \min(\rho_i, \rho_j)$, but initialize uncertainty using density only, $U_i^{(0)} = 1 - \frac{\rho_i - \min(\rho)}{\max(\rho) - \min(\rho) + \epsilon}$, without centroid distances; propagation remains unchanged. This variant isolates the contribution of distance-based boundary uncertainty, i.e., whether density information alone is sufficient to anchor three-way clustering results.

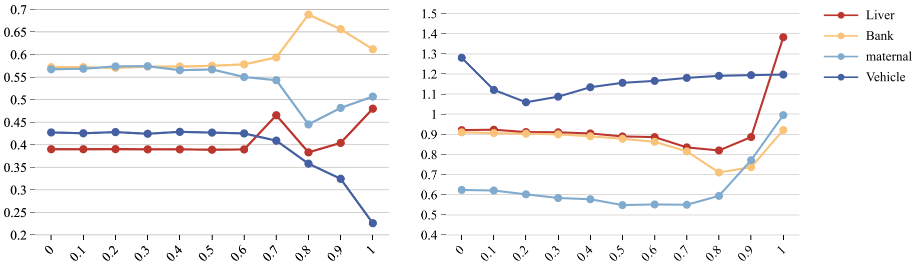
(iii) **w/o Propagation (No Propagation):** Disable iterative propagation and directly use the initial uncertainty for partitioning, $U^{(\text{final})} = U^{(0)}$. This variant removes the diffusion step to evaluate the performance contributed by uncertainty propagation beyond the local initialization.

We report the ablation results on 8 representative datasets spanning diverse structural properties, including density-heterogeneous data (Bank, Vowel), compact/convex clusters (Wine, Seeds, WDBC), and datasets with complex geometry or high boundary ambiguity (Balance Scale, Waveform, Segment). Performance is evaluated using F1, ARI, and DBI, and the comparative results are summarized in Fig. 7.

Effect of density similarity. Compared with the variant *w/o density similarity*, the full method achieves clear gains on *Bank* and *Vowel*, evidenced by higher F1 and ARI together with lower DBI. This suggests that density similarity is essential for suppressing

Table 9
Optimal propagation coefficient α of KUp3WC identified via grid search.

Dataset	Segment	Glass	Libras	Bank	WDBC	Vehicle	Waveform	Balancescale
α	1.0	0.8	0.6	0.8	0.8	0.4	0.9	0.8
Dataset	Liver	Seeds	Iris	Wine	Vowel	Winequality	Heartdisease	Maternal
α	1.0	0.9	0.9	0.8	0.9	0.5	0.0	0.3



(a) F1

(b) DBI

Fig. 8. Sensitivity analysis of KUp3WC with respect to α on four representative datasets (evaluated by F1 and DBI).

spurious Euclidean-based connections across low-density gaps, thereby preserving non-linear cluster boundaries during uncertainty propagation.

Effect of boundary uncertainty. When boundary uncertainty is removed (*w/o boundary uncertainty*), the performance on *Wine*, *Seeds*, and *WDBC* deteriorates noticeably, with reduced F1 and ARI and increased DBI. This indicates that distance-based boundary uncertainty provides a necessary geometric reference for three-way partitioning, particularly on compact and convex clusters where centroid distances reliably capture the core structure.

Effect of propagation. Without propagation (*w/o propagation*), DBI becomes consistently higher on most datasets, especially *Balancescale*, *Waveform*, *Segment*, and *Bank*, while F1 and ARI also tend to decline. This supports the idea that iterative diffusion is crucial for transforming local boundary uncertainty into a globally consistent uncertainty pattern, yielding more compact clusters and a more reliable core–fringe separation.

Overall, the ablation study confirms that each module contributes to the final effectiveness of Up3WC, and their integration yields the most stable and discriminative uncertainty estimation.

4.5. Analysis of the propagation coefficient α

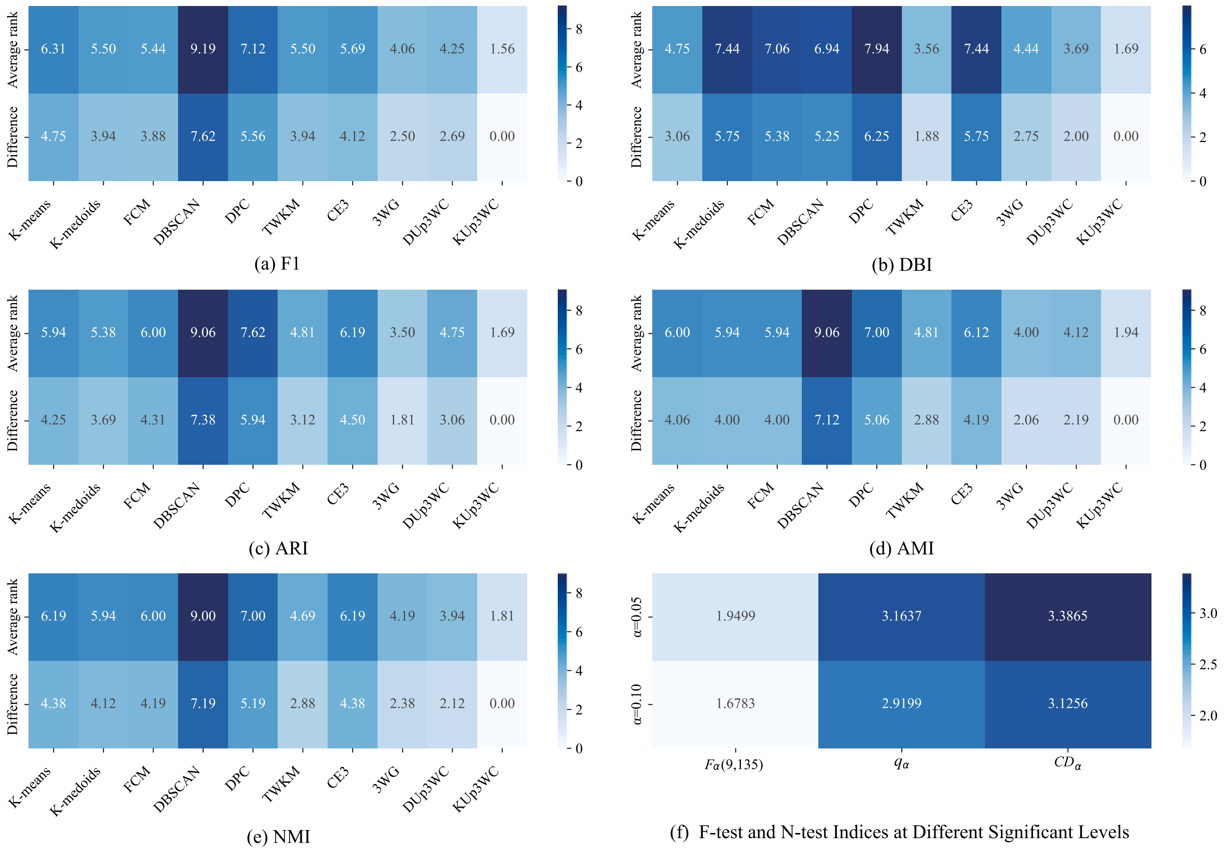
The propagation coefficient α controls the strength of uncertainty propagation in Up3WC, i.e., the balance between retaining the locally initialized boundary uncertainty and incorporating density-guided propagated information. A larger α encourages stronger diffusion along density-similar structures, whereas a smaller α keeps the three-way partitioning closer to the original distance-based initialization. Therefore, α plays a central role in adjusting Up3WC to datasets with different geometric and density characteristics.

Selection protocol. To objectively evaluate the performance of Up3WC and ensure fair comparison against baselines, we adopted a grid-search protocol. Specifically, α was determined by searching the interval $[0, 1]$ with a coarse step size of 0.1 for each dataset. For the comparison experiments, we selected the optimal α value yielding the highest F1 score from this grid search to benchmark the maximal potential of Up3WC. These optimal values, listed in Table 9, provide empirical insight into how different data structures favor different balances between distance and density. While this systematic protocol ensures comparability and reproducibility, practical applications can follow the heuristic guidelines proposed below to avoid extensive tuning.

Sensitivity analysis. To further examine the impact of α, we selected four representative datasets (Liver, Bank, Maternal, and Vehicle) covering distinct structural patterns. Fig. 8 plots the variations in F1 and DBI across the full range of α. The trends confirm that while Up3WC is sensitive to α, the performance changes are generally smooth. Crucially, the optimal α values align well with the intrinsic structure of the data: datasets requiring strong density connectivity (e.g., Liver) peak at high α, while those with compact geometry (e.g., Maternal) favor low α.

Practical guidelines. Based on the sensitivity results and theoretical analysis, we provide the following data-driven heuristics to guide the selection of α in practice:

- Low α ($0.0 \le \alpha \le 0.4$): Recommended for datasets with compact, spherical, or convex clusters (e.g., Maternal). In these cases, the centroid-based geometry is reliable, and the initial distance-based uncertainty dominates. A lower α preserves this geometric prior, avoiding unnecessary diffusion.
- High α ($0.6 \le \alpha \le 1.0$): Recommended for datasets with heterogeneous densities (e.g., Liver, Bank). Here, Euclidean distance to a centroid may be misleading. A higher α allows uncertainty to be refined along high-density pathways, compensating for local geometric ambiguity.



(a) F1 (b) DBI (c) ARI (d) AMI (e) NMI (f) F-test and N-test Indices at Different Significant Levels

Fig. 9. The Friedman-Nemenyi statistical test results.

Table 10
Average ranking values and Friedman's test results of ten algorithms.

Performance	K-means	K-medoids	FCM	DBSCAN	DPC	TWKM	CE3	3WG	DUp3WC	KUp3WC	F_F	Prob>Chi-sq
F1	6.31	5.50	5.44	9.19	7.13	5.5	5.69	4.06	**4.25**	**1.56**	9.48	8.73E-09
DBI	4.75	7.44	7.06	6.94	7.94	3.56	7.44	4.44	**3.69**	**1.69**	14.49	1.08E-11
ARI	5.94	5.38	6.00	9.06	7.63	4.81	6.19	3.5	**4.75**	**1.69**	12.21	1.71E-10
AMI	6.00	5.94	5.94	9.06	7.00	4.81	6.13	4.00	**4.13**	**1.94**	9.79	5.35E-09
NMI	6.19	5.94	6.00	9.00	7.00	4.69	6.19	4.19	**3.94**	**1.81**	10.32	2.40E-09

- Balanced α (≈ 0.5): For datasets with mixed characteristics or unknown structures, setting $\alpha = 0.5$ offers a robust default baseline that balances both sources of information.

In summary, α acts as a geometric adaptability parameter. The proposed heuristics substantially narrow the search space, making the algorithm practical without introducing complex adaptive mechanisms.

4.6. Statistical test analysis

To determine whether the proposed algorithms exhibit statistically significant differences in clustering performance compared to comparison methods, we conducted the Friedman test and the Nemenyi test. In this analysis, KUp3WC was selected as the baseline algorithm for comparisons. Fig. 9 presents the results of the Friedman-Nemenyi statistical test, where lower rank value indicates better performance. First, the Friedman test was performed at significance levels of 0.05 and 0.1. Table 10 summarizes the average ranks of the 10 methods and the Friedman test results, including F_F and corresponding p-values. In this experiment, 10 comparison algorithms and 16 datasets were used, i.e., $k = 10$, $N = 16$, meaning the F-value follows an F-distribution with degrees of freedom (9, 135). By referring to Table 10 and Fig. 9(f), it is evident that $F_F > F_{0.05}(9, 135) = 1.9499$ and $F_F > F_{0.1}(9, 135) = 1.6783$, and the p-values for all metrics are far smaller than 0.05 and 0.1, indicating significant differences between the proposed algorithm and the comparison algorithms. Therefore, the Nemenyi post-hoc test was conducted to further identify if there are substantial differences between any two methods.

In the Nemenyi test, the critical distance is given by $CD_\alpha = q_\alpha\sqrt{\frac{N(N+1)}{6k}}$, where q_α is the critical value. If the difference in average ranks between two algorithms exceeds CD_α, the hypothesis that 'the two algorithms perform equally' is rejected at the corresponding significance level. For this experiment, with $k=10$ and $N=16$, the critical values are $q_{\alpha=0.05} = 3.1637$, $q_{\alpha=0.1} = 2.9199$. When the distance between two methods exceeds $CD_{\alpha=0.05} = 3.3865$ and $CD_{\alpha=0.1} = 3.1256$, the performance is considered significantly different. As shown in Fig. 9(a)–(e), KUp3WC ranks first in all metrics, and both KUp3WC and DUp3WC statistically outperform most of the other comparison methods. Therefore, the Up3WC algorithm demonstrates superior performance in clustering.

5. Conclusions and future work

This paper presents a multi-scale uncertainty propagation three-way clustering algorithm (Up3WC) to address the limitation that single-scale methods cannot fully capture data uncertainty, which often hinders the ability of clustering methods to reveal the intrinsic structure of data. By incorporating both data density and geometric distance information into the uncertainty propagation mechanism, Up3WC provides a more thorough representation of data uncertainty and effectively captures the underlying structure of the data. Furthermore, we extend the Up3WC algorithm to enhance two widely used two-way clustering methods, K-means and DPC, thereby providing a more comprehensive and flexible solution for clustering tasks. To validate the effectiveness of the Up3WC algorithm, we conduct experiments on 16 datasets with five evaluation metrics: F1 score, DBI, ARI, AMI, and NMI. Both qualitative and quantitative analyses indicate that Up3WC effectively enhances clustering performance.

Future work will focus on improving the propagation coefficient to enable the automatic selection of optimal values based on dataset characteristics. In addition, the proposed framework will be explored in the context of explicit overlapping clustering. Furthermore, we plan to extend the proposed framework to high-dimensional and large-scale datasets, exploring its adaptability and efficiency in more complex and heterogeneous data environments.

CRediT authorship contribution statement

Caihui Liu: Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Xiying Chen:** Writing – original draft, Validation, Software, Methodology, Conceptualization. **Wenjing Qiu:** Software, Data curation. **Duoqian Miao:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The research is supported by the National Natural Science Foundation of China under Grant Nos. 62566003, 62166001, Jiangxi Province Natural Science Foundation of China under Grant Nos. 20232ACB202013, and the Graduate Innovation Funding Program of Jiangxi Province under Grant No. YC2024-S816.

Data availability

Data will be made available on request.

References

[1] S. Wang, Y. Zhang, Y. Zhang, Y. Zhang, S. Pang, J. Su, Y. Liu, Graph attention autoencoder model with dual decoder for clustering single-cell RNA sequencing data 54 (6) 2024, 5136–5146.
[2] S. Perumal, P.K. Sujatha, S. Krishnaa, M. Krishnan, Clusters in chaos: a deep unsupervised learning paradigm for network anomaly detection, J. Netw. Comput. Appl. 235 (2025) 104083.
[3] H. Qing, Community detection by spectral methods in multi-layer networks, Appl. Soft Comput. 171 (2025) 112769.
[4] A. Kartakoullis, N. Caporaso, M.B. Whitworth, I.D. Fisk, Gaussian mixture model clustering allows accurate semantic image segmentation of wheat kernels from near-infrared hyperspectral images, Chemom. Intell. Lab. Syst. 259 (2025) 105341.
[5] Q. Liu, M. Cai, Q. Li, Low-dimensional representation-driven TSK fuzzy system for feature selection, IEEE Trans. Fuzzy Syst. 33 (11) (2025) 4044–4054.
[6] C. Wang, C. Wang, Y. Qian, Q. Leng, Feature selection based on weighted fuzzy rough sets, IEEE Trans. Fuzzy Syst. 32 (7) (2024) 4027–4037.
[7] X. Chen, C. Liu, D. Xie, D. Miao, Image thresholding segmentation method based on adaptive granulation and reciprocal rough entropy, Inf. Sci. 695 (2025) 121737.
[8] Y. Yao, Three-way decision and granular computing, Int. J. Approx. Reason. 103 (2018) 107–123.
[9] Y. Yao, Tri-level thinking: models of three-way decision, Int. J. Mach. Learn. Cybern. 11 (5) (2020) 947–959.
[10] Y. Yao, The Dao of three-way decision and three-world thinking, Int. J. Approx. Reason. 162 (2023) 109032.
[11] J. Xu, R. Wang, Y. Zhang, W. Ding, Adaptive sequential three-way decisions for dynamic time warping, Inf. Sci. 690 (2025) 121541.
[12] X. Che, D. Chen, J. Deng, J. Mi, Exploiting local label correlation from sample perspective for multi-label classification via three-way decision theory, Appl. Soft Comput. 149 (2023) 110950.
[13] J. Zhan, M. Cai, A cost-minimized two-stage three-way dynamic consensus mechanism for social network-large scale group decision-making: utilizing K-nearest neighbors for incomplete fuzzy preference relations, Expert Syst. Appl. 263 (2025) 125705.
[14] C. Jiang, Y. Duan, A novel three-way deep learning approach for multigranularity fuzzy association analysis of time series data, IEEE Trans. Fuzzy Syst. 32 (9) (2023) 4835–4845.
[15] Y.U. Hong, Three-way cluster analysis, Peak Data Sci. 5 (1) (2016) 31–35 (In Chinese with English Abstract).
[16] Y. Xie, J. Yang, Y. Luo, X. Zhang, J. Luo, Constructing intuitionistic neighborhood based on intuitionistic fuzzy sets for three-way clustering, Int. J. Approx. Reason. 186 (2025) 109512.

[17] Y. Chen, P. Zhu, Y. Yao, An axiomatic framework for three-way clustering, Inf. Sci. 675 (2024) 120761.

[18] P. Wang, X. Yang, W. Ding, J. Zhan, Y. Yao, Three-way clustering: foundations, survey and challenges, Appl. Soft Comput. 151 (2024) 111131.

[19] P. Wang, Y. Yao, CE3: A three-way clustering method based on mathematical morphology, Knowl.-Based Syst. 155 (2018) 54–65.

[20] A. Shah, N. Azam, E. Alanazi, J. Yao, Image blurring and sharpening inspired three-way clustering approach, Appl. Intell. 52 (15) (2022) 18131–18155.

[21] P. Wang, H. Shi, X. Yang, J. Mi, Three-way K-means: integrating K-means and three-way decision, Int. J. Mach. Learn. Cybern. 10 (2019) 2767–2777.

[22] W. Guan, P. Wang, W. Jiang, Y. Zhang, HC3: a three-way clustering method based on hierarchical clustering, Cogn. Comput. 17 (1) (2025) 1–14.

[23] X. Han, J. Zhu, Y. Meng, P. Wang, Three-way clustering based on granular-ball, J. Jiangsu Univ. Sci. Technol. (Nat. Sci. Ed.) 38 (5) (2024) 57–62 (In Chinese with English Abstract).

[24] M. Du, J. Zhao, J. Sun, Y. Dong, M3W: multistep three-way clustering, IEEE Trans. Neural Netw. Learn. Syst. 35 (4) (2022) 5627–5640.

[25] H. Ju, J. Guo, W. Ding, X. Yang, D3WC: deep three-way clustering with granular evidence fusion, Inf. Fusion. 114 (2025) 102699.

[26] J. Sun, M. Du, Z. Lew, Y. Dong, TWStream: three-way stream clustering, IEEE Trans. Fuzzy Syst. 32 (9) (2024) 4927–4939.

[27] H. Yang, W. Wang, J. Cai, J. Wang, Y. Li, Y. Xun, X. Zhao, Three-way clustering based on the graph of local density trend, Int. J. Approx. Reason. 182 (2025) 109422.

[28] D. Xu, Y. Wang, Density estimation for toroidal data using semiparametric mixtures, Stat. Comput. 33 (6) (2023) 140.

[29] T. Duong, KS: kernel density estimation and kernel discriminant analysis for multivariate data in R, J. Stat. Softw. 21 (2007) 1–16.

[30] L. Hubert, P. Arabie, Comparing partitions, J. Classif. 2 (1985) 193–218.

[31] C. Shao, R. Wan, D. Miao, J. Zhao, Three-way gaussian mixture clustering based on granular ball neighborhood rough sets, J. Zhengzhou Univ. (Nat. Sci. Ed.) (2024) 1–8 (In Chinese with English Abstract).

[32] M. Meilă, Comparing clusterings—an information based distance, J. Multivar. Anal. 98 (5) (2007) 873–895.

[33] X. Chen, C. Liu, B. Lin, J. Lai, D. Miao, AHA-3WKM: the optimization of K-means with three-way clustering and artificial hummingbird algorithm, Inf. Sci. 672 (2024) 120661.

[34] D. Pfitzner, R. Leibbrandt, D. Powers, Characterization and evaluation of similarity measures for pairs of clusterings, Knowl. Inf. Syst. 19 (2009) 361–394.