

Received 7 April 2023, accepted 24 April 2023, date of publication 1 May 2023, date of current version 9 May 2023. Digital Object Identifier 10.1109/ACCESS.2023.3271996

RESEARCH ARTICLE

Geodesic Affinity Propagation Clustering Based on Angle-Based Outlier Factor

CHAOJIE WANG¹⁰ AND JIAQI JU

School of Mathematics and Statistics, Beijing Technology and Business University, Beijing 100048, China Corresponding author: Chaojie Wang (wangcj2019@btbu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 12001022.

ABSTRACT The affinity propagation (AP) clustering algorithm has received a lot of attention over the past few years. AP is efficient and insensitive to initialization, and generates clustering results with lower error and in less time. However, there are still two key limitations: AP-related algorithms cannot identify outliers in clusters. And they are usually not ideal for processing nonlinear data. To address the above issues, we propose a geodesic affinity propagation clustering algorithm based on angle-based outlier factor (ABOF-GAP). First, outliers are identified according to the value of angle-based outlier factor. Besides, Euclidean distance is replaced with geodesic distance to measure similarity. Experiments on synthetic data and real data illustrate the effectiveness of the ABOF-GAP algorithm.

INDEX TERMS Affinity propagation (AP), geodesic distances, outlier identification, angle-based outlier factor (ABOF).

I. INTRODUCTION

Cluster analysis is a critical step in data science [1], [2], [3]. As a classical clustering algorithm based on similarity measure, K-means algorithm starts with a random set of initial cluster centers and minimize iterately the sum of squared errors between the data point and the corresponding examplars [4], [5], [6]. However, this algorithm is very sensitive to the setting of initial cluster centers. And it usually reruns many times for different initializations. In 2007, Frey and Dueck proposed Affinity Propagation (AP) [7], a clustering method which simultaneously considers all data as potential exemplars. It takes as input the similarities between data points. Then the messages between data points are continuously passed until a set of high-quality exemplars appear. By contrast, AP does not need to initialize the cluster centers and the number of clusters. AP is shown to perform more efficiency (with lower error and in less time) than k-means algorithm [8], [9], [10].

Many improved versions of AP have been proposed in recent years. In 2013, Wang et al presented a multi-exemplar affinity propagation (MEAP) algorithm to extend the

The associate editor coordinating the review of this manuscript and approving it for publication was Ines Domingues¹⁰.

single-exemplar model [11]. In 2017, Li et al proposed an adjustable preference affinity propagation algorithm (APAP). It computes the value of each element preference based on the data distribution in the initial stage, and the preference will adjust automatically during iteration process [12]. Besides, many other similarity measurement method were also developed to improve the Euclidean distance used in AP. In 2017, Liu et al used the negative sine-squared value of the acute angle between discontinuity unit normal vectors as similarity measurement [13]. In 2020, Wang proposed an IMFSE-AP method, which measures similarities between signals based on the sample entropy of IMFs [14].

Although many achievements have been made in AP-related researches, the performance of AP or AP-based algorithms is vulnerable to outliers. The existence of outliers will affect the accuracy of clustering results. In some practical problems, the identification of outliers is even critical. As for this issue, some improved methods have been developed. In 2012, Sushravya et al addressed this limitation through the use of an asymmetric dissimilarity metric and a density-based outlier detection technique [15]. In 2018, Lei and Li proposed a Semi-supervised Affinity Propagation Clustering Algorithm based on outlier pruning (LOF-SAP), which used the local outlier factor algorithm (LOF) to identify

outliers [16]. But there still is a problem that these methods are hard to apply to other kinds of data.

The identification of cluster outliers is a big challenge for the application and development of AP algorithms. In addition, the clustering performance of AP-related algorithms may get worse when dealing with nonlinear data. Calculating similarity between data points using Euclidean distance ignores the non-linear structure of the data, and may bring unexpected errors in clustering. The main contribution of this paper is to propose an improved AP algorithm, geodesic affinity propagation clustering based on angle-based outlier factor (ABOF-GAP). In our algorithm, outliers can be identified and eliminated more accurately based on the anglebased outlier factor. Besides, Euclidean distance is replaced with geodesic distance to measure similarity between data points. This is more consistent with the nonlinear structure of data. Experiments on both synthetic data and real data prove that the proposed ABOF-GAP algorithm can obtain better performance.

The remaining of this paper is organized as follows. In Section II, we give a brief review of AP algorithm, geodesic distance and angle-based outlier factor. In Section III, we propose the ABOF-GAP algorithm. In Section IV, we show the experimental results on simulation data and real-world data. At last, Section V states conclusions.

II. AFFINITY PROPAGATION METHOD AND RELATED PROBLEM

A. AP

Affinity Propagation is an exemplar-based clustering method, which considers all data points as potential exemplars simultaneously and identify clusters automatically. Compared with other clustering method, AP can avoid poor clustering results caused by unlucky initialization and hard decision, and find clusters with much lower error and in less time. AP takes as input measures of similarity between pairs of data points, then real-valued messages are transmitted between data points recursively until a high-quality set of exemplars and corresponding clusters gradually emerges. For data points i and data point k, their similarity s(i, k) is set to a negative Euclidean distance, which reflects how well data point k is suited to be the exemplar for data points *i*. In particular, AP takes as input a real number s(k, k) called preference for each data point k, and the data point with larger value of s(k, k) is more likely to be chosen as an exemplar. In APbased algorithms, all the values of preferences are commonly set as a constant, the median of the input similarities generally, which could results in a moderate number of clusters.

In the iterative process, two kinds of message, responsibility r(i, k) and availability a(i, k), are exchanged between data points. The responsibility r(i, k) sent from *i* to *k*, reflects the priority for data point *k* to be chosen as the exemplar of point *i*, compared with other potential exemplars for point *i*. The availability a(i, k) sent from *k* to *i*, reflects the fitness for point *k* to be the exemplar of point *i*, with the support from



FIGURE 1. Geodesic distance.

other points that point k can serve as an exemplar. Messages are updated on the basis of simple formulas that search for minima of an appropriately chosen energy function. At any point in time, the magnitude of each message reflects the current affinity that one data point has for choosing another data point as its exemplar. The concrete algorithm is shown as follows:

Input the matrix of similarities, S, where S(i, k) is the similarity s(i, k) between point i and point k. For point x_i and x_k,

$$s(i,k) = -||x_i - x_k||^2.$$
 (1)

and s(k, k) is commonly set to be the median of the input similarities.

2) Compute responsibility r(i, k), where

$$r(i,k) = s(i,k) - \max_{k' \neq k} \{a(i,k') + s(i,k')\}.$$
 (2)

To begin with, the availabilities are set to zero in the first iteration.

3) Compute availability a(i, k), where

$$a(i,k) = \min\left\{0, r(k,k) + \sum_{i' \notin \{i,k\}} \max\{0, r(i',k)\}\right\},\$$

$$a(k,k) = \sum_{i' \neq k} \max\{0, r(i',k)\}.$$
(4)

 Combine availabilities and responsibilities to identify exemplars and corresponding clusters, and terminate the algorithm: (i) after a fixed number of iteration, (ii) after changes in the messages fall below a threshold, or (iii) after the local decisions stay constant for some number of iterations.

B. GEODESIC DISTANCE

In AP and AP-based algorithms, each similarity between data points is set to be a negative Euclidean distance. However, it is difficult to obtain a good clustering effect when the dimension of dataset increases gradually. Geodesic distance (GD) is an important concept in mathematical morphology. The way of calculating distance is not only determined by the spatial location of two samples, but also by the spatial distribution of the dataset. For example, in graph theory, geodesic distance is the distance of the shortest path between two points in the graph, which is different from Euclidean distance. In Figure 1, the Euclidean distance of the two black points should be the length of the line segment represented by the dashed line d_{15} . But geodesic distance is the shortest distance of the actual path. Its distance should be the minimum value of the sum of the distances of the solid line segments along the way, namely $d_{12} + d_{23} + d_{34} + d_{45}$.

In order to make better use of the manifold structure information in data with complex structure, this paper will adopt the negative geodesic distance as the similarity measure between data points.

C. ABOF

The existence of outliers in original data influences the performance of AP-related methods. We use ABOF algorithm to solve this problem. ABOF algorithm is an outlier detection method based on angle [17]. As the dimensionality of the data increases, comparing distances becomes less and less meaningful. The ABOF algorithm not only uses the distance between points in a vector space, but also mainly considers the difference between the angles of the vectors. Figure 2 shows a 2D dataset with distinct in-cluster points, boundary points, and outliers. It can be found that the difference in the angle between the vectors formed by the points in the cluster and other points is very large. The difference in the angle between the vectors formed by the boundary points and other points is smaller. The difference in the angle between the vectors formed by the outliers and other points is very small.



FIGURE 2. Point distribution.

The ABOF algorithm uses the cosine between the vectors to measure the size of the angle, and uses the variance to measure the size of the angle difference. For all data points, each data has an ABOF value, to detect outliers. For dataset *D*, the concrete descriptions of ABOF algorithm are as follows: 1) Cosine of the angle between two vectors:

$$\cos\left\langle \overrightarrow{AB}, \overrightarrow{AC} \right\rangle = \frac{\left\langle \overrightarrow{AB}, \overrightarrow{AC} \right\rangle}{\left| \overrightarrow{AB} \right| \cdot \left| \overrightarrow{AC} \right|} \tag{5}$$

2) Angle-based outlier factor of point *A*:

$$ABOF\left(\vec{A}\right) = VAR_{\vec{B},\vec{C}\in D}\left(\frac{\cos\left\langle \vec{AB},\vec{AC}\right\rangle}{\|AB\|^{2}\cdot\|AC\|^{2}}\right) \quad (6)$$

The essence of the ABOF algorithm is the weighted variance of the angle cosine. So the smaller the ABOF value is, the more likely the object is an outlier.

III. THE PROPOSED CLUSTERING METHOD

In this section, we propose an improved AP algorithm.

A. OUTLIER DETECTION

In the sample space, points inconsistent with the general behavior or characteristics of other sample points are called outliers. The outliers is generally caused by the error of calculation or operation, or the variability or elasticity of the data itself. The existence of outliers often leads to calculation errors and even some false information. But on the other hand, outliers can sometimes provide important information that researchers are concerned about. For example, outliers in medical data are usually more important than normal data. Therefore, the identification of outliers is very important in clustering problems. We give a simple example to illustrate the problem of AP algorithm on outlier detection.



We design a dataset with 11 data points in 2D space, which is shown in Figure 3. In this dataset, there are 1 outlier (in red), and 3 categories with 5 (in green), 3 (in gold), and 2 (in blue) data points respectively. It is visually obvious that point n1 is isolated and should be considered as an outlier. However, under AP algorithm, point n1 is assigned to the cluster with point x5 as its "cluster center." And point x5 is the closest exemplar to point n1. In this example, the clustering results of AP algorithm are inconsistent with the real distribution of



FIGURE 4. Flowchart of ABOF-GAP algorithm.

the data. This is because AP cannot identify outliers automatically, which would have a negative effect on the accuracy of clustering results.

B. ABOF-GAP

In order to make up for the shortcoming that AP cannot identify outliers, we propose an improved AP algorithm based on geodesic distance and the ABOF method.

For each data point x_i , we calculate its angle-based outlier factor $ABOF(x_i)$ based on Formula (6). According to ABOF principle, the smaller the ABOF value is, the more likely the object is an outlier. So we remove outliers from original dataset with relatively small ABOF, and get a pruned dataset.

After getting rid of the outliers, we calculate geodesic distance to evaluate the similarity between data points. For data point x_i and data point x_j , their similarity s(i, k) is set to a negative geodesic distance. Then we set s(k,k) to the median of the input similarities and get the similarity matrix S.

At last we input similarity matrix S and perform AP clustering algorithm. Responsibility r(i, k) and availability

a(i, k) are initialized and calculated iteratively via Formula (2)-(4). When meeting termination condition, we combine availabilities and responsibilities to identify exemplars and their corresponding clusters.

Algorithm 1 ABOF-GAP

- 1) Calculate the angle-based outlier factor of each data point by (6) in turn to get *ABOF* (x_i).
- 2) Remove outliers with relatively small ABOF.
- Calculate the geodesic distance s(i,k) between data points.
- 4) Set s(k,k) to the median of the input similarities and get similarity matrix S.
- 5) Perform AP clustering algorithm.
- 6) Output clusters and their exemplars.

This algorithm is summarized in Algorithm 1. And the flowchart for the proposed ABOF-GAP algorithm is shown in Figure 4.

TABLE 1.	The	clustering	results	on the	synthetic	datasets
----------	-----	------------	---------	--------	-----------	----------

Datasets	Algorithm	RI	NMI	Precision	Recall	F-score
	AP [7]	0.7204	0.5313	0.1331	0.9777	0.2344
	LOF-SAP [16]	0.7224	0.5366	0.1391	0.9785	0.2436
	CLAP [23]	0.7213	0.5326	0.1383	0.8970	0.2647
data5d	SSAPEC [24]	0.7230	0.5446	0.1463	0.9758	0.2634
	DDAP [25]	0.7211	0.5472	0.1356	0.9755	0.2856
	GAP	0.7285	0.5534	0.1593	0.9764	0.2739
	ABOF-GAP	0.7330	0.5646	0.1809	0.9810	0.3055
	AP [7]	0.7897	0.6678	0.3503	0.9868	0.5170
	LOF-SAP [16]	0.7906	0.6700	0.3532	0.9928	0.5211
	CLAP [23]	0.7899	0.6845	0.3654	0.9901	0.5198
data25d	SSAPEC [24]	0.7900	0.6783	0.3763	0.9882	0.5324
	DDAP [25]	0.7901	0.6743	0.3515	0.9879	0.5209
	GAP	0.7974	0.6847	0.3750	0.9860	0.5434
	ABOF-GAP	0.8169	0.7081	0.4419	0.9966	0.6104
	AP [7]	0.8181	0.7090	0.4410	0.9847	0.6092
data50d	LOF-SAP [16]	0.8250	0.7292	0.4691	0.9856	0.6351
	CLAP [23]	0.8194	0.7135	0.4534	0.9884	0.6235
	SSAPEC [24]	0.8198	0.7241	0.4475	0.9893	0.6135
	DDAP [25]	0.8234	0.8235	0.4575	0.9904	0.6109
	GAP	0.8351	0.7334	0.4970	0.9801	0.6595
	ABOF-GAP	0.8905	0.8158	0.6752	0.9929	0.8001

C. COMPUTATIONAL COMPLEXITY

This section summarizes the computational complexity of the proposed ABOF-GAP algorithm. Assuming that there are N samples in the dataset, in ABOF algorithm, each data point in the dataset is combined with any two points other than itself to form a vector. The computational complexity of this step is $O(N^2)$, which means that the computational complexity of generating vectors for all objects in the data set is $O(N^3)$. In addition, the original AP clustering takes $N \times N$ similarity matrix S as input. Therefore, the computational complexity of the original AP clustering is $O(N^2T)$, where T is the number of iterations. Overall, the computational complexity of ABOF-GAP algorithm is $O(N^3 + N^2T)$.

IV. EXPERIMENTS

In this section, several numerical experiments are conducted to benchmark our algorithm.

A. EVALUATION INDEX

In order to make a reasonable evaluation of the results of each clustering algorithm, we use the following five evaluation metrics to analyze and compare the clustering performance.

Suppose U is the external evaluation criterion and V is the clustering result. TP is the number of pairs of data points of the same class in U and V; TN is the number of pairs of data points of the different class in U and V; FP is the number of pairs of data points of the different class in U but of the same

VOLUME 11, 2023

class in V; FN is the number of pairs of data points of the same class in U but of the different class in V.

The rand index (RI) can be considered as the ratio of correct clusters, which is a classic measure of the similarity between two data clusterings [18]. RI is defineded as:

$$RI = \frac{TP + TN}{TP + FP + FN + TN} \tag{7}$$

The Precision measures the proportion of sample points that are correctly assigned [19]. Precision is defined as:

$$Precision = \frac{TP}{TP + FP}$$
(8)

The Recall indicates the correctness of the clustering results [20]. Recall is defined as:

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

The F-score counts the harmonic mean of Precision and Recall [21]. F-score is defined as:

$$F - score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$
(10)

The normalized mutual information (NMI) evalutes the similarity between two clusters from an information theory perspective [22]. NMI is defined as:

$$NMI = \frac{\sum_{i=1}^{c} \sum_{j=1}^{c} N_{i,j} log \frac{N \cdot N_{i,j}}{N \cdot N_j}}{\sqrt{\sum_{i=1}^{c} N_i log \frac{N_i}{N} \cdot \sum_{j=1}^{c} N_j log \frac{N_j}{N}}}$$
(11)

43623

TABLE 2. The clustering results on the synthetic datasets.

Datasets Ang [7] 0.7435 0.05768 0.0255 0.9807 0.03265 D=10 SLOF-SAP [16] 0.7441 0.5768 0.02105 0.9833 0.3361 SSAPEC [24] 0.7443 0.5892 0.02200 0.9799 0.3331 DDAP [25] 0.7441 0.5892 0.2276 0.9768 0.3640 ABOF-GAP 0.7610 0.6124 0.2679 0.5882 0.4271 D=20 CLAP [23] 0.7785 0.6592 0.3276 0.9746 0.4773 SSAPEC [24] 0.7786 0.6521 0.3286 0.4817 BDP-GAP 0.8029 0.6542 0.3380 0.9860 0.5559 DAP [25] 0.7786 0.6521 0.3386 0.9860 0.5569 DAP [25] 0.8034 0.7340 0.3332 0.9889 0.5569 DAP [27] 0.8017 0.6961 0.4106 0.9899 0.5569 DAP [27] 0.8131 0.7140 0.4358 0.9887 0.6202	Detecto	Algorithm	DI	NIMI	Provision	Docoll	Facoro
AP [7] 0.7445 0.6785 0.2055 0.9807 0.3286 D=10 CLAP [23] 0.7442 0.6881 0.2105 0.9333 0.3361 DDAP [25] 0.7441 0.6896 0.2200 0.9796 0.3640 ABOF-GAP 0.7610 0.6124 0.2676 0.9796 0.3640 LOF-SAF [16] 0.7792 0.6451 0.3141 0.9853 0.4699 LOF-SAF [16] 0.7792 0.6478 0.3155 0.9864 0.4773 SAPEC [24] 0.7786 0.6531 0.3155 0.9868 0.4817 ABOF-GAP 0.8029 0.6842 0.3984 0.5667 DLAP [25] 0.7860 0.6531 0.3356 0.9869 0.5569 DAP [25] 0.8031 0.7382 0.9981 0.5663 DAP [25] 0.8034 0.7332 0.9981 0.5663 DAP [25] 0.8034 0.7342 0.3354 0.9884 0.5643 DAP [25] 0.8034 0.7343 0.3989 </td <td>Datasets</td> <td>Angorithin</td> <td>n1</td> <td></td> <td>r recision</td> <td>necan</td> <td>r-score</td>	Datasets	Angorithin	n1		r recision	necan	r-score
Di-SAP [16] 0.7431 0.0581 0.2105 0.9883 0.03361 DDAP [25] 0.7441 0.6582 0.2210 0.9290 0.3361 DDAP [25] 0.7441 0.6582 0.2230 0.9799 0.3531 DDAP [25] 0.7441 0.6520 0.02679 0.9862 0.4071 ABOF-GAP 0.6521 0.2276 0.9746 0.4779 D=20 CLAP [23] 0.7785 0.6520 0.3380 0.9861 0.4476 DAP [25] 0.7786 0.6531 0.3155 0.9880 0.5566 DAF [7] 0.8011 0.6642 0.3984 0.5940 0.5533 DAF [25] 0.6341 0.7366 0.9890 0.5667 CLAP [23] 0.8017 0.6966 0.4048 0.3980 0.5580 DAF [25] 0.6324 0.7392 0.4459 0.9870 0.6123 DAF [25] 0.6314 0.7302 0.3380 0.9899 0.5569 DAF [25] 0.6463 0.7712 <td></td> <td>AP [7]</td> <td>0.7435</td> <td>0.5768</td> <td>0.2055</td> <td>0.9807</td> <td>0.3286</td>		AP [7]	0.7435	0.5768	0.2055	0.9807	0.3286
D=10 CLAP [23] 0.7442 0.5892 0.2230 0.9799 0.3337 DDAP [25] 0.7441 0.5896 0.2076 0.9796 0.3640 ABOF-GAP 0.7610 0.6124 0.2679 0.9862 0.4071 D=20 CLAP [23] 0.7785 0.6550 0.3315 0.9904 0.4749 DAP [25] 0.7786 0.6561 0.3155 0.9884 0.4773 SAPEC [24] 0.7786 0.6531 0.3155 0.9896 0.4817 ABOF-GAP 0.8019 0.6842 0.3984 0.5940 0.5556 DAP [25] 0.8014 0.6937 0.3996 0.5667 DAP [25] 0.8014 0.6937 0.3939 0.9889 0.5569 ABOF-GAP 0.8463 0.7142 0.4333 0.9891 0.5573 DAP [25] 0.8014 0.7937 0.4333 0.9891 0.5573 DAP [25] 0.8150 0.7077 0.4333 0.9891 0.5529 DAP [25]		LOF-SAP [16]	0.7451	0.5811	0.2105	0.9833	0.3361
DAT SSAPEC [24] 0.7453 0.5822 0.2230 0.9796 0.3634 DDAP [25] 0.7461 0.5826 0.2076 0.9796 0.3640 ABOF-GAP 0.7782 0.6451 0.3141 0.9853 0.4679 D=20 CLAP [23] 0.7785 0.6592 0.3276 0.9746 0.4773 SSAPEC [24] 0.7785 0.6592 0.3286 0.9861 0.4876 DDAP [25] 0.7786 0.6501 0.3380 0.9858 0.4817 ABOF-GAP 0.8029 0.6842 0.3984 0.5506 D=30 CLAP [23] 0.8011 0.6961 0.4006 0.9899 0.5663 D=40 CLAP [23] 0.8017 0.6961 0.4006 0.9899 0.5663 D=40 CLAP [23] 0.810 0.7149 0.4333 0.9891 0.5693 D=40 CFAP [7] 0.8124 0.7008 0.4229 0.9851 0.5694 D=40 SAPEC [24] 0.8181 0.7149	D=10	CLAP [23]	0.7442	0.5832	0.2140	0.9280	0.3497
DDAP [25] 0.7441 0.5896 0.2679 0.9862 0.4071 AP [7] 0.7782 0.6451 0.3141 0.9853 0.4071 D=20 CLAP [23] 0.7785 0.6552 0.3276 0.9746 0.4779 D=20 CLAP [23] 0.7785 0.6562 0.3380 0.9861 0.4476 DDAP [25] 0.7786 0.6531 0.3155 0.9858 0.4817 ABOF-GAP 0.8029 0.6842 0.3984 0.5960 0.5539 D=30 CLAP [23] 0.8011 0.6843 0.3966 0.4980 0.5667 D=40 CLAP [23] 0.8019 0.69661 0.4006 0.9894 0.5569 D=40 CLAP [23] 0.8019 0.7712 0.5552 0.9951 0.66648 DDAP [25] 0.8034 0.7740 0.4333 0.9889 0.5529 D=40 CLAP [23] 0.8181 0.7774 0.4333 0.9887 0.6023 D=410 CLAP [23] 0.8181	D-10	SSAPEC [24]	0.7453	0.5892	0.2230	0.9799	0.3531
ABOF-GAP 0.7610 0.6124 0.2879 0.9862 0.4711 AP [7] 0.7782 0.6413 0.3141 0.9853 0.4699 DF-SAP [16] 0.7792 0.6478 0.3175 0.9904 0.4479 SSAPEC [24] 0.7785 0.6592 0.3276 0.9746 0.4773 DDAP [25] 0.7786 0.6531 0.3984 0.9940 0.4773 DDAP [25] 0.7786 0.6531 0.3384 0.9940 0.5536 DDAP [25] 0.7781 0.6661 0.4006 0.9890 0.5613 DSFAPEC [24] 0.8011 0.66843 0.7512 0.3990 0.9880 0.5649 DDAP [25] 0.8034 0.7710 0.4358 0.9887 0.6028 CLAP [23] 0.8135 0.7077 0.4358 0.9887 0.6028 D=40 CLF-SAP [16] 0.8181 0.7090 0.4110 0.9876 0.6321 D=40 CLF-SAP [16] 0.8181 0.7072 0.4333 0.9890		DDAP [25]	0.7441	0.5896	0.2076	0.9796	0.3640
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$		ABOF-GAP	0.7610	0.6124	0.2679	0.9862	0.4071
LOF-SAP [16] 0.7792 0.6478 0.3175 0.9904 0.4749 D=20 CLAP [23] 0.7785 0.6592 0.3276 0.9746 0.4773 DDAP [25] 0.7786 0.6531 0.3155 0.9861 0.4876 DDAP [25] 0.7786 0.6531 0.3155 0.9861 0.4876 ABOF-GAP 0.8029 0.6842 0.3984 0.9940 0.5566 AP [7] 0.8011 0.6666 0.4048 0.9889 0.5569 DAP [25] 0.8034 0.7310 0.3939 0.9886 0.5667 DAP [25] 0.8034 0.7312 0.5322 0.9951 0.6663 DAP [23] 0.8121 0.7174 0.4459 0.9870 0.6123 D=40 LOF-SAP [16] 0.8181 0.7179 0.4333 0.9891 0.5979 D=40 SAPEC [24] 0.8181 0.7149 0.4333 0.9890 0.5929 AP [7] 0.8181 0.7149 0.4333 0.9846 0.6323		AP [7]	0.7782	0.6451	0.3141	0.9853	0.4699
D=20 CLAP [23] 0.7785 0.6592 0.3276 0.9746 0.4773 SSAPEC [24] 0.7788 0.6560 0.3380 0.9861 0.4876 DDAP [25] 0.7786 0.6531 0.3155 0.9858 0.4817 BOF-GAP 0.8029 0.6842 0.3984 0.9940 0.5556 CLAP [23] 0.8011 0.6643 0.3966 0.9894 0.5661 D=30 CLAP [23] 0.8017 0.6961 0.4048 0.9894 0.5663 D=40 ABOF-GAP 0.8463 0.7512 0.5352 0.9951 0.66863 D=40 CLAP [23] 0.813 0.7174 0.4435 0.9887 0.6028 D=40 CLAP [23] 0.8181 0.7174 0.4433 0.9881 0.5793 D=40 CLAP [23] 0.8181 0.7193 0.4363 0.9887 0.6028 D=40 CLAP [23] 0.8181 0.7193 0.4363 0.9890 0.5793 D=50 CLAP [23]		LOF-SAP [16]	0.7792	0.6478	0.3175	0.9904	0.4749
D=20 SSAPEC [24] 0.7788 0.6530 0.3380 0.9861 0.4876 DDAP [25] 0.7786 0.6531 0.3155 0.9858 0.4817 ABOF-GAP 0.8029 0.6842 0.3984 0.9940 0.5536 LOF-SAP [16] 0.8044 0.6937 0.3996 0.9890 0.5667 CLAP [23] 0.8017 0.6961 0.4006 0.9894 0.5613 SSAPEC [24] 0.8019 0.6966 0.4048 0.9886 0.5648 DDAP [25] 0.8034 0.7340 0.3939 0.9889 0.5569 ABOF-GAP 0.8463 0.7512 0.5352 0.9951 0.6863 DDAP [25] 0.8124 0.7008 0.4229 0.9851 0.5908 LOF-SAP [16] 0.8181 0.7174 0.4459 0.9870 0.6123 CLAP [23] 0.8135 0.7077 0.4338 0.9887 0.6028 SSAPEC [24] 0.8118 0.7174 0.4459 0.9870 0.6123 DDAP [25] 0.8167 0.7937 0.4333 0.9891 0.5929 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.7622 ABOF-GAP 0.8805 0.8136 0.6752 0.9994 0.6103 DDAP [25] 0.8194 0.7135 0.4534 0.9884 0.6235 SSAPEC [24] 0.8198 0.7241 0.4475 0.9893 0.6135 DDAP [25] 0.8234 0.8235 0.4575 0.9994 0.6109 ABOF-GAP 0.8805 0.8158 0.6752 0.9992 0.6601 ABOF-GAP 0.8805 0.8158 0.6752 0.9992 0.6403 DDAP [25] 0.8234 0.8235 0.4575 0.9994 0.6202 SSAPEC [24] 0.8292 0.7403 0.4738 0.9902 0.6439 DDAP [25] 0.8328 0.7199 0.4570 0.9904 0.6109 DDAP [25] 0.8328 0.7199 0.4570 0.9904 0.6220 SSAPEC [24] 0.8222 0.7403 0.4738 0.9902 0.6439 DDAP [25] 0.8328 0.7199 0.4560 0.9904 0.6220 SSAPEC [24] 0.8232 0.7412 0.4928 0.9852 0.6609 SSAPEC [24] 0.8232 0.7403 0.4738 0.9902 0.6439 DDAP [25] 0.8376 0.8733 0.5023 0.9924 0.6271 LOF-SAP [16] 0.8232 0.7403 0.4738 0.9902 0.6439 DDAP [25] 0.8376 0.8733 0.5023 0.9924 0.6673 DDAP [25] 0.8384 0.7654 0.5007 0.9850 0.6695 SSAPEC [24] 0.8386 0.7554 0.4934 0.9933 0.4644 DDAP [25] 0.8384 0.7564 0.5007 0.9856 0.6695 SSAPEC [24] 0.8386 0.7552 0.4991 0.6743 DDAP [25] 0.8384 0.8761 0.5048 0.9911 0.6743 DDAP [25] 0.8384 0.8761 0.5048 0.9911 0.6743 DDAP [25] 0.8384 0.8761 0.5048 0.9916 0.6433 ABOF-GAP 0.9370 0.8686 0.8220 0.9932 0.9318 ABOF-GAP 0.9370 0.8686 0.8220 0.9936 0.5577 DDAP [25] 0.8344 0.8672 0.4942 0.9916 0.6433 ABOF-		CLAP [23]	0.7785	0.6592	0.3276	0.9746	0.4773
DIAP [25] 0.7786 0.6303 0.3155 0.9858 0.4817 ABOF-GAP 0.8029 0.6642 0.3984 0.9940 0.5539 LOF-SAP [16] 0.8044 0.6633 0.3866 0.9890 0.5667 D304 CLAP [23] 0.8017 0.6961 0.4006 0.9894 0.5613 DAAP [25] 0.8034 0.7340 0.3335 0.9886 0.5569 DADAP [25] 0.8034 0.7512 0.5522 0.9951 0.6863 DAFSAP[16] 0.8181 0.7174 0.4459 0.9880 0.5579 DAPA [7] 0.8135 0.7077 0.4333 0.9891 0.5973 DAPA [23] 0.8135 0.7077 0.4363 0.9887 0.6028 DAPCAP [24] 0.8135 0.7937 0.4363 0.9891 0.5973 DAP [25] 0.8276 0.7937 0.4363 0.9847 0.6028 D=50 CLAP [23] 0.8194 0.7135 0.4534 0.8235 0.5572	D=20	SSAPEC [24]	0.7788	0.6560	0.3380	0.9861	0.4876
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$			0.7796	0.6591	0.3355	0.0001	0.4017
ABP-GAP 0.8022 0.8022 0.3984 0.9890 0.5539 D=30 LOF-SAP [16] 0.8011 0.66937 0.3996 0.9890 0.5667 CLAP [23] 0.8017 0.6961 0.4006 0.9894 0.5613 DAAP [25] 0.8034 0.7340 0.3392 0.9880 0.5569 ABOF-GAP 0.8463 0.7512 0.5352 0.9951 0.6863 LOF-SAP [16] 0.8181 0.7174 0.4435 0.9880 0.5573 D=40 CLAP [23] 0.8135 0.7077 0.4358 0.9887 0.6028 SAPEC [24] 0.8138 0.7149 0.4333 0.9891 0.5973 DADA [25] 0.8161 0.7093 0.4661 0.9856 0.6321 D=50 CLAP [23] 0.8181 0.7093 0.4533 0.9894 0.6225 D=50 CLAP [23] 0.8198 0.7222 0.4661 0.9856 0.6331 D=50 CLAP [23] 0.8217 0.7149 0.4323<		ADDAF [20]	0.1100	0.0001	0.3133	0.9606	0.4017
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $			0.0011	0.0042	0.3964	0.9940	0.5590
D=30 CLAP [23] 0.8047 0.6937 0.4066 0.9894 0.5613 D=30 CLAP [23] 0.8017 0.6966 0.4008 0.9886 0.5648 DDAP [25] 0.8034 0.7340 0.3339 0.9889 0.5569 AB [7] 0.8124 0.7008 0.4229 0.9951 0.6683 D=40 SAP [7] 0.8124 0.7008 0.4229 0.9851 0.5508 LOF-SAP [16] 0.8181 0.7174 0.4455 0.9987 0.6023 DAP [25] 0.8167 0.7377 0.4338 0.9887 0.6022 AP [7] 0.8181 0.7077 0.4338 0.9891 0.5929 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.7622 AP [7] 0.8181 0.7090 0.4110 0.9847 0.6293 D=50 CLAP [23] 0.8250 0.7573 0.4391 0.6133 DDAP [25] 0.8237 0.4575 0.9994 0.6103 D		AP [7]	0.8011	0.0843	0.3866	0.9860	0.5539
D=30 CLAP [23] 0.8017 0.6861 0.4006 0.9894 0.5613 DDAP [25] 0.8034 0.7340 0.3939 0.9889 0.5569 ABOF-GAP 0.8463 0.7512 0.5352 0.9951 0.6683 LOF-SAP [16] 0.8181 0.7174 0.4459 0.9887 0.6123 D=40 CLAP [23] 0.8135 0.7077 0.4358 0.9896 0.5292 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.7692 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.6625 SAPEC [24] 0.8181 0.7190 0.4410 0.9847 0.66092 LOF-SAP [16] 0.8250 0.7292 0.4691 0.9856 0.63351 DD=50 CLAP [23] 0.8234 0.8235 0.4575 0.9904 0.6109 ABOF-GAP 0.8905 0.8158 0.6752 0.9929 0.8001 D=50 CLAP [23] 0.8234 0.7422 0.4928 0.98		LOF-SAP [16]	0.8044	0.6937	0.3996	0.9899	0.5667
SSAPEC [24] 0.6019 0.6066 0.4048 0.9886 0.5569 ABOF-GAP 0.8463 0.7310 0.3339 0.9889 0.5569 ABOF-GAP 0.8463 0.7512 0.5352 0.9951 0.6863 D=40 CLAP [23] 0.8131 0.7174 0.4459 0.9870 0.6123 D=40 SAPEC [24] 0.8138 0.7149 0.4333 0.9891 0.5973 DDAP [25] 0.8167 0.7937 0.4363 0.9899 0.5929 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.7622 AP [7] 0.8181 0.7090 0.4410 0.9847 0.6092 LOF-SAP [16] 0.8250 0.7292 0.4619 0.9835 0.6135 DAP [25] 0.8234 0.8235 0.4475 0.9944 0.6220 DAP [26] 0.8322 0.7412 0.4475 0.9940 0.6109 DDAP [25] 0.8328 0.7173 0.4569 0.9947 0.6221	D=30	CLAP [23]	0.8017	0.6961	0.4006	0.9894	0.5613
DDAP [25] 0.8034 0.7340 0.3939 0.9889 0.5569 ABOF-GAP 0.8463 0.7512 0.5352 0.9951 0.6863 AP [7] 0.8124 0.7008 0.4229 0.9851 0.5908 LOF-SAP [16] 0.8181 0.7174 0.4459 0.9870 0.6123 DAAP [23] 0.8135 0.7077 0.4353 0.9891 0.5929 DDAP [25] 0.8167 0.7937 0.4363 0.9899 0.5929 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.7622 AP [7] 0.8181 0.7090 0.4410 0.9847 0.6092 D-50 CLAP [23] 0.8194 0.7135 0.4534 0.9884 0.6235 SAPEC [24] 0.8198 0.7221 0.4691 0.9847 0.6620 DAP [25] 0.8234 0.8235 0.4575 0.9994 0.6135 DDAP [25] 0.8232 0.7142 0.4912 0.9847 0.6221 DAP [7]		SSAPEC [24]	0.8019	0.6966	0.4048	0.9886	0.5648
ABOF-GAP 0.8463 0.7512 0.5352 0.9951 0.6863 LOF-SAP [16] 0.8124 0.7008 0.4229 0.9851 0.5908 LOF-SAP [16] 0.8135 0.7077 0.4358 0.9887 0.6023 D=40 CLAP [23] 0.8135 0.7077 0.4358 0.9896 0.5973 DDAP [25] 0.8167 0.7937 0.4363 0.9899 0.5973 DDAP [25] 0.8181 0.7090 0.4410 0.9847 0.6023 D=50 CLAP [23] 0.8194 0.7135 0.4534 0.9893 0.6135 DDAP [25] 0.8234 0.8235 0.4575 0.9904 0.6109 DDAP [25] 0.8221 0.7142 0.4475 0.9949 0.6220 D=60 CLAP [23] 0.8212 0.7422 0.9474 0.6220 D=74 0.8217 0.7149 0.4573 0.9940 0.6199 D=50 CLAP [23] 0.8221 0.7143 0.9477 0.6220		DDAP [25]	0.8034	0.7340	0.3939	0.9889	0.5569
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		ABOF-GAP	0.8463	0.7512	0.5352	0.9951	0.6863
D=40 LOF-SAP [16] 0.8181 0.7174 0.4459 0.9870 0.6123 D=40 CLAP [23] 0.8135 0.7077 0.4358 0.9887 0.6028 SAPEC [24] 0.8138 0.7149 0.4333 0.9891 0.5929 DAPAP [25] 0.8167 0.7937 0.4363 0.9890 0.5929 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.7622 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.6331 D=50 CLAP [23] 0.8194 0.7135 0.4475 0.9893 0.6135 DDAP [25] 0.8234 0.8235 0.4575 0.9904 0.6109 ABOF-GAP 0.8905 0.8158 0.6752 0.9929 0.8001 D=60 CLAP [23] 0.8222 0.7403 0.4573 0.9902 0.6439 DAP [25] 0.8328 0.7173 0.4569 0.9904 0.6392 SAPEC [24] 0.8231 0.7173 0.4569 0.9904 </td <td></td> <td>AP [7]</td> <td>0.8124</td> <td>0.7008</td> <td>0.4229</td> <td>0.9851</td> <td>0.5908</td>		AP [7]	0.8124	0.7008	0.4229	0.9851	0.5908
D=40 CLAP [23] 0.8135 0.7077 0.4358 0.9887 0.6028 SSAPEC [24] 0.8138 0.7149 0.4333 0.9891 0.5973 DDAP [25] 0.8167 0.7937 0.4363 0.9899 0.5929 AB0F-GAP 0.8758 0.7943 0.6285 0.9936 0.7622 LOF-SAP [16] 0.8250 0.7292 0.4691 0.9847 0.6092 LOF-SAP [23] 0.8194 0.7135 0.4534 0.9847 0.6092 DDAP [25] 0.8234 0.8235 0.4575 0.9929 0.8011 DDAP [25] 0.8234 0.8235 0.4575 0.9929 0.8011 DE60 CLAP [23] 0.8224 0.7422 0.4475 0.9929 0.8011 D=60 CLAP [23] 0.8238 0.7149 0.4573 0.9902 0.6439 DAP [7] 0.8217 0.7143 0.4928 0.9830 0.6322 D=60 CLAP [23] 0.8226 0.7240 0.4738 0.992		LOF-SAP [16]	0.8181	0.7174	0.4459	0.9870	0.6123
D=40 SSAPEC [24] 0.8138 0.7149 0.4333 0.9891 0.5973 DDAP [25] 0.8167 0.7937 0.4363 0.9899 0.5929 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.7622 LOF-SAP [16] 0.8250 0.7292 0.4691 0.9856 0.6331 D=50 CLAP [23] 0.8194 0.7135 0.4534 0.9884 0.6235 SSAPEC [24] 0.8198 0.7241 0.4475 0.9893 0.6135 DDAP [25] 0.8234 0.8235 0.4575 0.9904 0.6109 ABOF-GAP 0.8805 0.8158 0.6752 0.9929 0.8001 D=60 CLAP [23] 0.8221 0.7403 0.4573 0.9929 0.6439 DDAP [25] 0.8234 0.7149 0.4470 0.9906 0.6392 SAPEC [24] 0.8222 0.720 0.4670 0.9906 0.6639 DAF [25] 0.8231 0.7173 0.4569 0.9847 0.62	D (0	CLAP [23]	0.8135	0.7077	0.4358	0.9887	0.6028
DDAP [25] 0.8167 0.7937 0.4363 0.9899 0.5929 ABOF-GAP 0.8758 0.7943 0.6285 0.9936 0.7622 AP [7] 0.8181 0.7090 0.4410 0.9847 0.6092 LOF-SAP [16] 0.8250 0.7292 0.4691 0.9856 0.6351 DE50 SSAPEC [24] 0.8198 0.7211 0.4475 0.9983 0.6135 DAP [25] 0.8234 0.8235 0.4575 0.9904 0.6109 ABOF-GAP 0.8905 0.8158 0.6752 0.9929 0.8001 AP [7] 0.8217 0.7149 0.4523 0.9847 0.6229 D=60 CLAP [23] 0.8238 0.7199 0.4670 0.9902 0.6392 SAPEC [24] 0.8292 0.7403 0.4738 0.9902 0.6392 D=70 CLAP [23] 0.8245 0.7220 0.4715 0.9913 0.6643 D=70 CLAP [24] 0.8346 0.7722 0.4715 0.9913	D=40	SSAPEC [24]	0.8138	0.7149	0.4333	0.9891	0.5973
ABOF-GAP 0.8758 0.7943 0.6225 0.9936 0.7622 D=50 AP [7] 0.8181 0.7090 0.4410 0.9856 0.6351 D=50 CLAP [23] 0.8194 0.7135 0.4541 0.9893 0.6135 SSAPEC [24] 0.8198 0.7241 0.4475 0.9893 0.6135 DDAP [25] 0.8234 0.8235 0.4575 0.9904 0.6109 ABOF-GAP 0.8905 0.8158 0.6752 0.9929 0.8011 D=60 CLAP [23] 0.8234 0.8235 0.4475 0.9902 0.66392 SSAPEC [24] 0.8292 0.7422 0.4928 0.9852 0.6699 DAP [25] 0.8328 0.8567 0.4874 0.9915 0.6392 DAP [25] 0.8328 0.8567 0.4874 0.9915 0.6639 DAP [25] 0.8326 0.7220 0.4715 0.9913 0.6444 SSAPEC [24] 0.8365 0.7527 0.4941 0.9990 0.6673 <td></td> <td>DDAP [25]</td> <td>0.8167</td> <td>0.7937</td> <td>0.4363</td> <td>0.9899</td> <td>0.5929</td>		DDAP [25]	0.8167	0.7937	0.4363	0.9899	0.5929
$\begin{array}{c c c c c c c c c c c c c c c c c c c $		ABOF-GAP	0.8758	0.7943	0.6285	0.9936	0.7622
$D=50 \begin{array}{ c c c c c c c c c c c c c c c c c c c$		AP [7]	0.8181	0.7090	0.4410	0.98/17	0.6002
D=50 CLAP [23] 0.8194 0.7135 0.4534 0.9884 0.6235 SSAPEC [24] 0.8198 0.7241 0.4475 0.9893 0.6135 DDAP [25] 0.8234 0.8235 0.4575 0.9904 0.6109 ABOF-GAP 0.8905 0.8158 0.6752 0.9929 0.8001 ABOF-GAP 0.8905 0.8158 0.6752 0.9929 0.8001 CLOF-SAP [16] 0.8222 0.7422 0.4928 0.9852 0.6609 CLAP [23] 0.8238 0.7199 0.4670 0.9906 0.6392 SSAPEC [24] 0.8292 0.7403 0.4738 0.9902 0.6439 DDAP [25] 0.8328 0.8567 0.4874 0.9915 0.6389 ABOF-GAP 0.9109 0.8468 0.7394 0.9930 0.8577 AP [7] 0.8231 0.7173 0.4569 0.9847 0.6271 LOF-SAP [16] 0.8322 0.7220 0.4715 0.9913 0.6444 SSAPEC [24] 0.8365 0.7527 0.4941 0.9909 0.6673 DDAP [25] 0.8376 0.8733 0.5023 0.9920 0.6528 ABOF-GAP 0.9268 0.8711 0.7899 0.9931 0.9030 AP [7] 0.8160 0.7054 0.4342 0.9847 0.6015 LOF-SAP [16] 0.8314 0.7408 0.4901 0.9852 0.6580 CLAP [23] 0.8194 0.7135 0.4534 0.9844 0.6235 SSAPEC [24] 0.8387 0.7564 0.5002 0.9911 0.6743 DDAP [25] 0.8387 0.7564 0.5002 0.9911 0.6743 DDAP [25] 0.8386 0.8751 0.5048 0.9921 0.6552 ABOF-GAP 0.9327 0.8800 0.8082 0.9932 0.9195 SSAPEC [24] 0.8386 0.6755 0.4194 0.9884 0.6235 SSAPEC [24] 0.8386 0.6755 0.4194 0.9884 0.6235 ABOF-GAP 0.9327 0.8800 0.8082 0.9932 0.9195 SSAPEC [24] 0.8386 0.6755 0.4194 0.9830 0.5842 SSAPEC [24] 0.8386 0.6755 0.4194 0.9838 0.5323 DDAP [25] 0.8384 0.7552 0.4982 0.9916 0.6743 DDAP [25] 0.8344 0.8622 0.4924 0.9916 0.6435 ABOF-GAP 0.9370 0.8866 0.6822 0.9932 0.9195 SSAPEC [24] 0.8386 0.7552 0.4982 0.9916 0.6713 DDAP [25] 0.8344 0.6627 0.4924 0.9916 0.6435 ABOF-GAP 0.9370 0.8866 0.6822 0.9932 0.9318 AP [7] 0.7909 0.6639 0.3549 0.9845 0.5119 DDAP [25] 0.8344 0.6674 0.4112 0.9867 0.5720 CLAP [23] 0.7925 0.6740 0.3695 0.9750 0.5266 SAPEC [24] 0.8336 0.7477 0.4860 0.9906 0.6579 DDAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249 ABOF-GAP 0.9386 0.8888 0.8266 0.9932 0.9359		LOESAP [16]	0.8250	0.7202	0.4601	0.0856	0.6351
D=50 SSAPEC [24] 0.8194 0.7135 0.4354 0.5354 0.6253 SSAPEC [24] 0.8198 0.7241 0.4475 0.9893 0.6135 DDAP [25] 0.8234 0.8235 0.4575 0.9904 0.6109 ABOF-GAP 0.8905 0.8158 0.6752 0.9929 0.8001 LOF-SAP [16] 0.8322 0.7422 0.4928 0.9852 0.6609 CLAP [23] 0.8238 0.7199 0.4670 0.9906 0.6392 SSAPEC [24] 0.8292 0.7403 0.4738 0.9902 0.6439 DDAP [25] 0.8328 0.8567 0.4874 0.9915 0.6389 ABOF-GAP 0.9109 0.8468 0.7394 0.9930 0.8577 AP [7] 0.8231 0.7173 0.4569 0.9847 0.6271 LOF-SAP [16] 0.8326 0.7465 0.5007 0.9850 0.6695 CLAP [23] 0.8252 0.7220 0.4715 0.9913 0.6444 SSAPEC [24] 0.8365 0.7527 0.4941 0.9909 0.6673 DDAP [25] 0.8376 0.8733 0.5023 0.9920 0.6528 ABOF-GAP 0.9268 0.8711 0.7899 0.9931 0.9030 AP [7] 0.8160 0.7054 0.4342 0.9847 0.6015 LOF-SAP [16] 0.8314 0.7408 0.4901 0.9852 0.6580 CLAP [23] 0.8194 0.7135 0.4534 0.9884 0.6235 SSAPEC [24] 0.8367 0.7564 0.5002 0.9911 0.6743 DDAP [25] 0.8384 0.8761 0.5048 0.9921 0.6552 ABOF-GAP 0.9327 0.8800 0.8082 0.9932 0.9195 D=90 CLAP [23] 0.8059 0.6888 0.4025 0.9846 0.5657 LOF-SAP [16] 0.8218 0.7234 0.4586 0.9858 0.6236 CLAP [23] 0.8059 0.6888 0.4025 0.9846 0.5657 LOF-SAP [16] 0.8218 0.7234 0.4586 0.9858 0.6236 ABOF-GAP 0.9327 0.8800 0.8082 0.9932 0.9195 AP [7] 0.8059 0.6639 0.3549 0.9845 0.5119 DDAP [25] 0.8344 0.8622 0.4924 0.9916 0.6713 DDAP [25] 0.8344 0.8622 0.4924 0.9916 0.6713 DDAP [25] 0.8344 0.8622 0.4924 0.9916 0.6719 DDAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249 ABOF-GAP 0.9385 0.6886 0.8220 0.9932 0.9318 AP [7] 0.7909 0.6639 0.3549 0.9845 0.5119			0.0200	0.7292	0.4524	0.0004	0.0001
DAP [25] 0.8234 0.8235 0.4473 0.9994 0.6109 ABOF-GAP 0.8905 0.8158 0.6752 0.9929 0.8001 D=60 CLAP [2] 0.8217 0.7149 0.4523 0.9847 0.6220 LOF-SAP [16] 0.8322 0.7422 0.4928 0.9852 0.6609 CLAP [23] 0.8238 0.7199 0.4670 0.9906 0.6392 DDAP [25] 0.8328 0.8567 0.4874 0.9915 0.6389 DDAP [25] 0.8328 0.8567 0.4874 0.9915 0.6695 ABOF-GAP 0.9109 0.8468 0.7394 0.9909 0.6673 DAP [25] 0.8326 0.7527 0.4911 0.9909 0.6673 DDAP [25] 0.8376 0.7533 0.5023 0.9920 0.6528 ABOF-GAP 0.9268 0.8711 0.7899 0.9311 0.6444 SSAPEC [24] 0.8346 0.7654 0.4342 0.9847 0.6015	D=50	CLAI [25]	0.0194	0.7133	0.4334	0.9004	0.0200
DDAP [25] 0.8234 0.8235 0.4375 0.9904 0.6109 ABOF-GAP 0.8905 0.8158 0.6752 0.9929 0.8001 AP [7] 0.8217 0.7149 0.4523 0.9827 0.6629 LOF-SAP [16] 0.8322 0.7422 0.4928 0.9852 0.6609 DDAP [25] 0.8328 0.7199 0.4670 0.9906 0.6392 SSAPEC [24] 0.8292 0.7403 0.4738 0.9902 0.6439 DDAP [25] 0.8328 0.8567 0.4874 0.9915 0.6389 ABOF-GAP 0.9109 0.8468 0.7394 0.9930 0.8577 AP [7] 0.8231 0.7173 0.4569 0.9847 0.6271 LOF-SAP [16] 0.8346 0.7465 0.5007 0.9850 0.6695 DDAP [25] 0.8365 0.7527 0.4941 0.9913 0.6444 SAPEC [24] 0.8365 0.7523 0.9920 0.6573 DDAP [25] 0.8344 0.7		SSAPEC [24]	0.0190	0.7241	0.4475	0.9893	0.0155
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		DDAP [25]	0.8234	0.8235	0.4575	0.9904	0.6109
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		ABOF-GAP	0.8905	0.8158	0.6752	0.9929	0.8001
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		AP [7]	0.8217	0.7149	0.4523	0.9847	0.6220
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		LOF-SAP [16]	0.8322	0.7422	0.4928	0.9852	0.6609
D=00 SSAPEC [24] 0.8292 0.7403 0.4738 0.9902 0.6439 DDAP [25] 0.8328 0.8567 0.4874 0.9915 0.6389 ABOF-GAP 0.9109 0.8468 0.7394 0.9930 0.8577 AP [7] 0.8231 0.7173 0.4569 0.9847 0.6271 LOF-SAP [16] 0.8326 0.7220 0.4715 0.9913 0.6444 SSAPEC [24] 0.8365 0.7527 0.4941 0.9090 0.6673 DDAP [25] 0.8376 0.8733 0.5023 0.9920 0.6528 ABOF-GAP 0.9268 0.8711 0.7899 0.9931 0.9030 D=80 AP [7] 0.8160 0.7054 0.4342 0.9847 0.6015 LOF-SAP [16] 0.8314 0.7108 0.4901 0.9852 0.6580 D=80 CLAP [23] 0.8194 0.7135 0.4534 0.9844 0.6235 D=80 CLAP [24] 0.8387 0.7564 0.5022 0.9911 </td <td>D - 60</td> <td>CLAP [23]</td> <td>0.8238</td> <td>0.7199</td> <td>0.4670</td> <td>0.9906</td> <td>0.6392</td>	D - 60	CLAP [23]	0.8238	0.7199	0.4670	0.9906	0.6392
DDAP [25] 0.8328 0.8567 0.4874 0.9915 0.6389 ABOF-GAP 0.9109 0.8468 0.7394 0.9930 0.8577 AP [7] 0.8231 0.7173 0.4569 0.9847 0.6271 LOF-SAP [16] 0.8346 0.7465 0.5007 0.9850 0.6695 CLAP [23] 0.8252 0.7220 0.4715 0.9913 0.6444 SSAPEC [24] 0.8365 0.7527 0.4941 0.9909 0.6673 DDAP [25] 0.8376 0.8733 0.5023 0.9920 0.6528 ABOF-GAP 0.9268 0.8711 0.7899 0.9931 0.9030 D=80 CLAP [23] 0.8160 0.7054 0.4342 0.9847 0.6015 LOF-SAP [16] 0.8314 0.7135 0.4534 0.9847 0.6015 LOF-SAP [16] 0.8314 0.7135 0.4534 0.9847 0.6235 SSAPEC [24] 0.8387 0.7564 0.5002 0.9911 0.6743 D	D-00	SSAPEC [24]	0.8292	0.7403	0.4738	0.9902	0.6439
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		DDAP [25]	0.8328	0.8567	0.4874	0.9915	0.6389
$ \begin{array}{c} & AP\left[7\right] & 0.8231 & 0.7173 & 0.4569 & 0.9847 & 0.6271 \\ LOF-SAP\left[16\right] & 0.8346 & 0.7465 & 0.5007 & 0.9850 & 0.6695 \\ CLAP\left[23\right] & 0.8252 & 0.7220 & 0.4715 & 0.9913 & 0.6444 \\ SSAPEC\left[24\right] & 0.8365 & 0.7527 & 0.4941 & 0.9099 & 0.6673 \\ DDAP\left[25\right] & 0.8376 & 0.8733 & 0.5023 & 0.9920 & 0.6528 \\ \hline & ABOF-GAP & 0.9268 & 0.8711 & 0.7899 & 0.9931 & 0.9030 \\ AP\left[7\right] & 0.8160 & 0.7054 & 0.4342 & 0.9847 & 0.6015 \\ LOF-SAP\left[16\right] & 0.8314 & 0.7408 & 0.4901 & 0.9852 & 0.6580 \\ CLAP\left[23\right] & 0.8194 & 0.7135 & 0.4534 & 0.9884 & 0.6235 \\ SSAPEC\left[24\right] & 0.8387 & 0.7564 & 0.5002 & 0.9911 & 0.6743 \\ DDAP\left[25\right] & 0.8384 & 0.8761 & 0.5048 & 0.9921 & 0.6552 \\ \hline & ABOF-GAP & 0.9327 & 0.8800 & 0.8082 & 0.9932 & 0.9195 \\ \hline & AP\left[7\right] & 0.8059 & 0.6888 & 0.4025 & 0.9846 & 0.5657 \\ LOF-SAP\left[16\right] & 0.8218 & 0.7234 & 0.4586 & 0.9858 & 0.6236 \\ CLAP\left[23\right] & 0.8085 & 0.6975 & 0.4194 & 0.9830 & 0.5842 \\ SSAPEC\left[24\right] & 0.8380 & 0.7552 & 0.4982 & 0.9910 & 0.6719 \\ DDAP\left[25\right] & 0.8344 & 0.8622 & 0.4924 & 0.9916 & 0.6435 \\ \hline & ABOF-GAP & 0.9370 & 0.8866 & 0.8220 & 0.9932 & 0.9318 \\ \hline & AP\left[7\right] & 0.7909 & 0.6639 & 0.3549 & 0.9845 & 0.5119 \\ LOF-SAP\left[16\right] & 0.8074 & 0.6974 & 0.4112 & 0.9867 & 0.5720 \\ CLAP\left[23\right] & 0.7925 & 0.6740 & 0.3695 & 0.9750 & 0.5266 \\ SSAPEC\left[24\right] & 0.8336 & 0.7477 & 0.4860 & 0.9906 & 0.6579 \\ DDAP\left[25\right] & 0.8281 & 0.8401 & 0.4724 & 0.9909 & 0.6249 \\ \hline & ABOF-GAP & 0.9385 & 0.8888 & 0.8266 & 0.9932 & 0.9359 \\ \hline & ABOF-GAP & 0.9385 & 0.8888 & 0.8266 & 0.9932 & 0.9359 \\ \hline & ABOF-GAP & 0.9370 & 0.8866 & 0.8220 & 0.9932 & 0.9318 \\ \hline & AP\left[7\right] & 0.7909 & 0.6639 & 0.3549 & 0.9845 & 0.5119 \\ \hline & LOF-SAP\left[16\right] & 0.8074 & 0.6974 & 0.4112 & 0.9867 & 0.5720 \\ \hline & CLAP\left[23\right] & 0.7925 & 0.6740 & 0.3695 & 0.9750 & 0.5266 \\ \hline & SSAPEC\left[24\right] & 0.8336 & 0.7477 & 0.4860 & 0.9906 & 0.6579 \\ \hline & DDAP\left[25\right] & 0.8281 & 0.8401 & 0.4724 & 0.9909 & 0.6249 \\ \hline & ABOF-GAP & 0.9385 & 0.8888 & 0.8266 & 0.9932 & 0.9359 \\ \hline & ABOF-GAP & 0.9385 & 0.8888 & 0.8266 & 0.9932 & 0.9359 \\ \hline & ABOF-GAP & 0.9385 & 0.8888 & 0.8266 & 0.9932 & 0.9359 \\ \hline & ABOF$		ABOF-GAP	0.9109	0.8468	0.7394	0.9930	0.8577
$ \begin{array}{c} \text{LOF-SAP} \left[16 \right] & 0.8346 & 0.7465 & 0.5007 & 0.9850 & 0.6695 \\ \text{CLAP} \left[23 \right] & 0.8252 & 0.7220 & 0.4715 & 0.9913 & 0.6444 \\ \text{SSAPEC} \left[24 \right] & 0.8365 & 0.7527 & 0.4941 & 0.9099 & 0.6673 \\ \text{DDAP} \left[25 \right] & 0.8376 & 0.8733 & 0.5023 & 0.9920 & 0.6528 \\ \hline & \textbf{ABOF-GAP} & \textbf{0.9268} & \textbf{0.8711} & \textbf{0.7899} & \textbf{0.9931} & \textbf{0.9030} \\ \hline & \textbf{AP} \left[7 \right] & 0.8160 & 0.7054 & 0.4342 & 0.9847 & 0.6015 \\ \text{LOF-SAP} \left[16 \right] & 0.8314 & 0.7408 & 0.4901 & 0.9852 & 0.6580 \\ \text{CLAP} \left[23 \right] & 0.8194 & 0.7135 & 0.4534 & 0.9884 & 0.6235 \\ \text{SSAPEC} \left[24 \right] & 0.8387 & 0.7564 & 0.5002 & 0.9911 & 0.6743 \\ \text{DDAP} \left[25 \right] & 0.8384 & 0.8761 & 0.5048 & 0.9921 & 0.6552 \\ \hline & \textbf{ABOF-GAP} & \textbf{0.9327} & \textbf{0.8800} & \textbf{0.8082} & \textbf{0.9932} & \textbf{0.9195} \\ \hline & \textbf{AP} \left[7 \right] & 0.8059 & 0.6888 & 0.4025 & 0.9846 & 0.5657 \\ \text{LOF-SAP} \left[16 \right] & 0.8218 & 0.7234 & 0.4586 & 0.9858 & 0.6236 \\ \text{CLAP} \left[23 \right] & 0.8085 & 0.6975 & 0.4194 & 0.9830 & 0.5842 \\ \hline & \textbf{SAPEC} \left[24 \right] & 0.8380 & 0.7552 & 0.4982 & 0.9910 & 0.6719 \\ \text{DDAP} \left[25 \right] & 0.8344 & 0.8622 & 0.4924 & 0.9916 & 0.6435 \\ \hline & \textbf{ABOF-GAP} & \textbf{0.9370} & \textbf{0.8866} & \textbf{0.8220} & \textbf{0.9323} & \textbf{0.9318} \\ \hline & \textbf{AP} \left[7 \right] & 0.7909 & 0.6639 & 0.3549 & 0.9845 & 0.5119 \\ \text{LOF-SAP} \left[16 \right] & 0.8074 & 0.6974 & 0.4112 & 0.9867 & 0.5720 \\ \hline & \textbf{CLAP} \left[23 \right] & 0.7925 & 0.6740 & 0.3695 & 0.9750 & 0.5266 \\ \hline & \textbf{SSAPEC} \left[24 \right] & 0.8336 & 0.7477 & 0.4860 & 0.9906 & 0.6579 \\ \hline & \textbf{DDAP} \left[25 \right] & 0.8281 & 0.8401 & 0.4724 & 0.9909 & 0.6249 \\ \hline & \textbf{ABOF-GAP} & \textbf{0.9385} & \textbf{0.8888} & \textbf{0.8266} & \textbf{0.9932} & \textbf{0.9318} \\ \hline & \textbf{AP} \left[7 \right] & 0.7925 & 0.6740 & 0.3695 & 0.9750 & 0.5266 \\ \hline & \textbf{SSAPEC} \left[24 \right] & 0.8336 & 0.7477 & 0.4860 & 0.9906 & 0.6579 \\ \hline & \textbf{DDAP} \left[25 \right] & 0.8281 & 0.8401 & 0.4724 & 0.9909 & 0.6249 \\ \hline & \textbf{ABOF-GAP} & \textbf{0.9385} & \textbf{0.8888} & \textbf{0.8266} & \textbf{0.9932} & \textbf{0.9359} \\ \hline & \textbf{ABOF-GAP} & \textbf{0.9385} & \textbf{0.8888} & \textbf{0.8266} & \textbf{0.9932} & \textbf{0.9359} \\ \hline & \textbf{ABOF-GAP} & \textbf{0.9385} & \textbf{0.8888} & \textbf{0.8266} & \textbf{0.9932} & \textbf{0.9359} \\ \hline & \textbf{ADF-GAP} & \textbf{0.9385} & 0.88$		AP [7]	0.8231	0.7173	0.4569	0.9847	0.6271
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		LOF-SAP [16]	0.8346	0.7465	0.5007	0.9850	0.6695
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		CLAP [23]	0.8252	0.7220	0.4715	0.9913	0.6444
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	D=70	SSAPEC [24]	0.8365	0.7527	0.4941	0.9909	0.6673
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		DDAP [25]	0.8376	0.8733	0.5023	0.9920	0.6528
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		ABOF-GAP	0.9268	0.8711	0.7899	0.9931	0.9030
$D=80 \begin{array}{c} D=80 \\ D=100 \\ D$		AP [7]	0.8160	0.7054	0.4342	0.9847	0.6015
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		LOF_SAP [16]	0.831/	0 7/08	0.4001	0.9852	0.6580
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$			0.0014	0.7400	0.4594	0.0002	0.0000
DAP [25] 0.8384 0.7394 0.5002 0.9911 0.6743 DDAP [25] 0.8384 0.8761 0.5048 0.9921 0.6552 ABOF-GAP 0.9327 0.8800 0.8082 0.9932 0.9195 AP [7] 0.8059 0.6888 0.4025 0.9846 0.5657 LOF-SAP [16] 0.8218 0.7234 0.4586 0.9830 0.5842 DAP [23] 0.8085 0.6975 0.4194 0.9830 0.5842 SSAPEC [24] 0.8380 0.7552 0.4982 0.9910 0.6719 DDAP [25] 0.8344 0.8622 0.4924 0.9916 0.6435 ABOF-GAP 0.9370 0.8866 0.8220 0.9932 0.9318 D=100 AP [7] 0.7909 0.6639 0.3549 0.9845 0.5119 D=100 CLAP [23] 0.7925 0.6740 0.3695 0.9750 0.5266 SAPEC [24] 0.8336 0.7477 0.4860 0.9906 0.6579	D=80	CLAI [20]	0.0194	0.7133	0.4004	0.2004	0.0200
DDAF [25] 0.8384 0.8701 0.5048 0.9921 0.6552 ABOF-GAP 0.9327 0.8800 0.8082 0.9932 0.9195 AP [7] 0.8059 0.6888 0.4025 0.9846 0.5657 LOF-SAP [16] 0.8218 0.7234 0.4586 0.9858 0.6236 CLAP [23] 0.8085 0.6975 0.4194 0.9830 0.5842 SSAPEC [24] 0.8380 0.7552 0.4982 0.9910 0.6719 DDAP [25] 0.8344 0.8622 0.4924 0.9916 0.6435 ABOF-GAP 0.9370 0.8866 0.8220 0.9932 0.9318 D=100 AP [7] 0.7909 0.6639 0.3549 0.9845 0.5119 D=100 CLAP [23] 0.7925 0.6740 0.3695 0.9750 0.5266 SAPEC [24] 0.8336 0.7477 0.4860 0.9906 0.6579 D=100 CLAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249<		33AF EU [24]	0.0000/	0.7004	0.0002	0.9911	0.0743
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		DDAP [20]	0.0004	0.8701	0.5048	0.9921	0.0002
$D=90 \begin{array}{ c c c c c c c c c c c c c c c c c c c$		ABOF-GAP	0.9327	0.8800	0.8082	0.9932	0.9195
$ \begin{array}{c} \text{LOF-SAP} \ [16] \\ \text{D=90} \\ \begin{array}{c} \text{LOF-SAP} \ [16] \\ \text{CLAP} \ [23] \\ \text{SSAPEC} \ [24] \\ \text{D=90} \end{array} \begin{array}{c} 0.8085 \\ \text{SSAPEC} \ [24] \\ \text{DAP} \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [25] \ [2$	D=90	AP [7]	0.8059	0.6888	0.4025	0.9846	0.5657
$ \begin{array}{c} D=90 \\ D=90 \\ D=90 \\ DAP [23] \\ DAP [25] \\ DDAP [25] \\ DDAP [25] \\ DDAP [25] \\ DAP [25] \\ D$		LOF-SAP $[16]$	0.8218	0.7234	0.4586	0.9858	0.6236
SSAPEC [24] 0.8380 0.7552 0.4982 0.9910 0.6719 DDAP [25] 0.8344 0.8622 0.4924 0.9916 0.6435 ABOF-GAP 0.9370 0.8866 0.8220 0.9932 0.9318 AP [7] 0.7909 0.6639 0.3549 0.9845 0.5119 LOF-SAP [16] 0.8074 0.6974 0.4112 0.9867 0.5720 D=100 CLAP [23] 0.7925 0.6740 0.3695 0.9750 0.5266 SSAPEC [24] 0.8336 0.7477 0.4860 0.9909 0.6249 DDAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249 ABOF-GAP 0.9385 0.8888 0.8266 0.9932 0.9359		CLAP [23]	0.8085	0.6975	0.4194	0.9830	0.5842
DDAP [25] 0.8344 0.8622 0.4924 0.9916 0.6435 ABOF-GAP 0.9370 0.8866 0.8220 0.9932 0.9318 AP [7] 0.7909 0.6639 0.3549 0.9845 0.5119 LOF-SAP [16] 0.8074 0.6974 0.4112 0.9867 0.5720 D=100 CLAP [23] 0.7925 0.6740 0.3695 0.9750 0.5266 SSAPEC [24] 0.8336 0.7477 0.4860 0.9906 0.6579 DDAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249 ABOF-GAP 0.9385 0.8888 0.8266 0.9932 0.9359		SSAPEC [24]	0.8380	0.7552	0.4982	0.9910	0.6719
ABOF-GAP 0.9370 0.8866 0.8220 0.932 0.9318 AP [7] 0.7909 0.6639 0.3549 0.9845 0.5119 LOF-SAP [16] 0.8074 0.6974 0.4112 0.9867 0.5720 D=100 CLAP [23] 0.7925 0.6740 0.3695 0.9750 0.5266 SSAPEC [24] 0.8336 0.7477 0.4860 0.9906 0.6579 DDAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249 ABOF-GAP 0.9385 0.8888 0.8266 0.9932 0.9359		DDAP [25]	0.8344	0.8622	0.4924	0.9916	0.6435
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		ABOF-GAP	0.9370	0.8866	0.8220	0.9932	0.9318
LOF-SAP [16] 0.8074 0.6974 0.4112 0.9867 0.5720 D=100 CLAP [23] 0.7925 0.6740 0.3695 0.9750 0.5266 SSAPEC [24] 0.8336 0.7477 0.4860 0.9906 0.6579 DDAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249 ABOF-GAP 0.9385 0.8888 0.8266 0.9932 0.9359	D=100	AP [7]	0.7909	0.6639	0.3549	0.9845	0.5119
D=100 CLAP [23] 0.7925 0.6740 0.3695 0.9750 0.5266 SSAPEC [24] 0.8336 0.7477 0.4860 0.9906 0.6579 DDAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249 ABOF-GAP 0.9385 0.8888 0.8266 0.9932 0.9359		LOF-SAP [16]	0.8074	0.6974	0.4112	0.9867	0.5720
D=100 SSAPEC [24] 0.8336 0.7477 0.4860 0.9906 0.6579 DDAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249 ABOF-GAP 0.9385 0.8888 0.8266 0.9932 0.9359		CLAP [23]	0.7925	0.6740	0.3695	0.9750	0.5266
DDAP [25] 0.8281 0.8401 0.4724 0.9909 0.6249 ABOF-GAP 0.9385 0.8888 0.8266 0.9932 0.9359		SSAPEC [24]	0.8336	0.7477	0.4860	0,9906	0.6579
ABOF-GAP 0.9385 0.8888 0.8266 0.9932 0.9359		DDAP [25]	0,8281	0.8401	0.4724	0,9909	0.6249
		ABOF-GAP	0.9385	0.8888	0.8266	0.9932	0.9359



(a) 5 dimensions and 610 data points.



(b) 25 dimensions and 610 data points.



(c) 50 dimensions and 610 data points.

FIGURE 5. Rand Index(RI) of clustering result on synthetic data sets for 5, 25 and 50 dimensions.

where, N is the number of data, N_i and N_j denote the number of data in category *i* and cluster *j* respectively, $N_{i,j}$ denotes the number of data in category *i* as well as in cluster *j*.

The above five methods have a range of value [0,1]. As the value is higher, the clustering quality is better. The value is close to 1 if the clustering results perfectly match the category labels. And the value is close to 0 if data are randomly partitioned.

B. SYNTHETIC DATA EXPERIMENTS

The synthetic dataset for simulation evaluation consists of 600 normal data points and 10 outliers. Normal data points are from three different Gaussian distributions. Outliers are randomly generated for each data set. To exactly evaluate the behavior of our new method for different dimensionalities, we generated multiple data sets in 5, 25 and 50 dimensions.



FIGURE 6. Rand Index(RI) of clustering result on synthetic data sets for 10-100 dimensions.

TABLE 3. Characteristics of datebases

Dateset	Size	Class	Dimensions	
Wine	178	3	13	
Iris	150	4	4	
Glass	178	7	9	
Segmnt	210	7	19	
PenDigits	10992	10	16	
Semeion	1593	10	256	

Besides, we also constructed an "intermediate" algorithm, geodesic affinity propagation (GAP) for comparison. The GAP method uses geodesic distance instead of Euclidean distance to measure similarity between data points. The GAP method can be taken as a reference for the proposed ABOF-GAP algorithm. This is in a view that the GAP does not include the outlier elimination process based on ABOF. For comparison, AP [7], LOF-SAP [16], CLAP [23], SSAPEC [24], DDAP [25] and GAP algorithm are all performed on these three datasets for comparison. The results are shown in Table 1.

As shown in Table 1, all five indices of ABOF-GAP are significantly higher than the other six algorithms. To better show the difference between the new algorithm and other algorithms in different dimensions, we make RI for each clustering result, which are shown in Figure 5. As the dimension increases, the gap between ABOF-GAP and other algorithms in indexes also increases. The above results also indicate that, as the dimension increases, distance gradually loses its value as a measure of outlier degree. However, this does not mean that the more outliers are eliminated, the better effect of the clustering could be got. Two indexes of ABOF-GAP reach the highest values when 3 outliers are removed, instead of all outliers being pruned. This may because the relative positions of all data points will change after an outlier is removed.

TABLE 4. The clustering results on the UCI datasets.

Datasets	Algorithm	RI	NMI	Precision	Recall	F-score
	AP [7]	0.7141	0.4805	0.1690	0.1998	0.1832
	LOF-SAP [16]	0.7228	0.5077	0.2006	0.2486	0.2221
	CLAP [23]	0.7265	0.5194	0.1843	0.2683	0.2597
Wine	SSAPEC [24]	0.7354	0.5423	0.2298	0.3743	0.2790
	DDAP [25]	0.7234	0.5254	0.2156	0.2943	0.2609
	GAP	0.7537	0.5717	0.3022	0.4147	0.3496
	ABOF-GAP	0.7555	0.5725	0.3103	0.4265	0.3593
	AP [7]	0.7528	0.5638	0.2713	0.3610	0.3098
	LOF-SAP [16]	0.7589	0.5782	0.2745	0.3792	0.3294
	CLAP [23]	0.7578	0.5723	0.2853	0.3764	0.3268
Iris	SSAPEC [24]	0.7612	0.5843	0.2963	0.3901	0.3367
	DDAP [25]	0.7583	0.5798	0.2784	0.3854	0.3298
	GAP	0.7641	0.5989	0.3127	0.4359	0.3641
	ABOF-GAP	0.7698	0.6295	0.3235	0.4629	0.3808
	AP [7]	0.7521	0.4072	0.1571	0.1646	0.1607
	LOF-SAP [16]	0.7523	0.4089	0.1571	0.1650	0.1626
	CLAP [23]	0.7522	0.4087	0.1579	0.1661	0.1630
Glass	SSAPEC [24]	0.7527	0.4103	0.1610	0.1690	0.1639
	DDAP [25]	0.7530	0.4091	0.1598	0.1654	0.1629
	GAP	0.7534	0.4158	0.1567	0.1651	0.1608
	ABOF-GAP	0.7567	0.4126	0.1620	0.1718	0.1668
	AP [7]	0.7909	0.4531	0.4234	0.3423	0.3786
	LOF-SAP $[16]$	0.8523	0.5523	0.4398	0.3490	0.3892
	CLAP [23]	0.8432	0.5743	0.4313	0.3487	0.3856
Segment	SSAPEC [24]	0.8619	0.5762	0.4367	0.3502	0.3887
	DDAP [25]	0.8532	0.5773	0.4298	0.3498	0.3857
	GAP	0.8623	0.5872	0.4412	0.3505	0.3907
	ABOF-GAP	0.8712	0.5902	0.4443	0.3512	0.3923
	AP [7]	0.8022	0.6733	0.3870	0.4589	0.4199
	LOF-SAP [16]	0.8032	0.6873	0.3994	0.4679	0.4309
	CLAP [23]	0.8040	0.6856	0.3954	0.4634	0.4267
Pendigits	SSAPEC [24]	0.8052	0.6878	0.4034	0.4680	0.4333
	DDAP [25]	0.8024	0.6867	0.3921	0.4656	0.4257
	GAP	0.8043	0.6879	0.4038	0.4678	0.4335
	ABOF-GAP	0.8102	0.6890	0.4189	0.4786	0.4468
	AP [7]	0.7021	0.3819	0.2198	0.3843	0.2797
	LOF-SAP [16]	0.7235	0.4280	0.2219	0.3927	0.2836
	CLAP [23]	0.7124	0.4156	0.2226	0.3910	0.2837
Semeion	SSAPEC [24]	0.7467	0.4298	0.2210	0.3925	0.2828
	DDAP $[25]$	0.7398	0.4198	0.2204	0.3819	0.2795
	GAP	0.7465	0.4296	0.2209	0.3920	0.2826
	ABOF-GAP	0.7562	0.4376	0.2234	0.3935	0.2850

To further show the change of clustering effect of ABOF-GAP algorithm under different dimensions, another experiment is performed on datasets generated from three

different Gaussian distributions. Each dataset consists of 600 normal data points and 10 outliers. The dimension of these datasets varies from 10 to 100. The number of outliers

TABLE 5. The clustering results on the BBC dataset.

Datasets	Algorithm	RI	NMI	Precision	Recall	F-score
	AP [7]	0.7728	0.5146	0.0817	0.9031	0.1498
	LOF - SAP [16]	0.7724	0.5142	0.0815	0.9027	0.1494
	CLAP [23]	0.7731	0.5153	0.0823	0.9021	0.1616
BBC10d	SSAPEC [24]	0.7767	0.5264	0.0956	0.9036	0.1792
	DDAP [25]	0.7751	0.5257	0.0876	0.9025	0.1724
	GAP	0.7787	0.5366	0.1095	0.8994	0.1952
	ABOF-GAP	0.7792	0.5369	0.1098	0.8992	0.1958
	AP [7]	0.7704	0.4998	0.0701	0.9129	0.1301
	LOF-SAP [16]	0.7702	0.5001	0.0705	0.9133	0.1309
	CLAP [23]	0.7707	0.5010	0.0812	0.9128	0.1308
BBC20d	SSAPEC [24]	0.7785	0.5491	0.1045	0.9139	0.2073
	DDAP [25]	0.7769	0.5278	0.1008	0.9087	0.2052
	GAP	0.7805	0.5482	0.1169	0.9022	0.2070
	ABOF-GAP	0.7825	0.5543	0.1199	0.9131	0. 2119
	AP [7]	0.7698	0.4982	0.0708	0.8781	0.1310
	LOF-SAP $[16]$	0.7692	0.4983	0.0688	0.8835	0.1277
	CLAP [23]	0.7723	0.5113	0.0892	0.8910	0.1421
BBC25d	SSAPEC [24]	0.7812	0.5342	0.1072	0.9097	0.1873
	DDAP [25]	0.7793	0.5279	0.1013	0.8978	0.1762
	GAP	0.7803	0.5519	0.1138	0.9182	0.2024
	ABOF-GAP	0.7852	0.5657	0.1371	0.9079	0.2382
	AP [7]	0.7653	0.4757	0.0473	0.9034	0.0898
BBC50d	LOF-SAP $[16]$	0.7652	0.4783	0.0480	0.9065	0.0911
	CLAP [23]	0.7689	0.4827	0.0579	0.9024	0.0956
	SSAPEC [24]	0.7793	0.5176	0.0873	0.9102	0.1980
	DDAP [25]	0.7726	0.5092	0.0690	0.9049	0.1846
	GAP	0.7812	0.5259	0.1398	0.8104	0.2384
	ABOF-GAP	0.7832	0.5308	0.1463	0.8276	0.2487

to be removed in ABOF-GAP algorithm is set to 3. The complete experimental results are shown in Table 2. ABOF-GAP achieves better clustering performance on all datasets. Figure 6 compares ABOF-GAP and AP, LOF-SAP, CLAP, SSAPEC, DDAP by RI. As the increase of dimension, the difference between other algorithms and ABOF-GAP's RI becomes more obvious. After the dimension exceeds 80, the RI of AP, CLAP and LOF-SAP clustering especially has a downward trend, which can be concluded that ABOF-GAP can better overcome the negative impact of dimension increment on clustering.

C. REAL-WORLD DATA EXPERIMENTS

The experimental data are from the UCI reference database and text dataset. UCI database is set up by the University of California Irvine as a database for machine learning, which derived from real life [26]. In the UCI database, datasets Wine, Iris, Glass, Segmnt, PenDigits and Semeion are tested in our experiments. The data characteristics are shown in Table 3.

We use the proposed algorithm ABOF-GAP to cluster the datasets Wine, Iris, Glass, Segmnt, PenDigits and Semeion. In order to reflect the effect of the algorithm, we also test AP, LOF-SAP, CLAP, SSAPEC, DDAP and GAP for comparison. The results are shown in Table 4. As shown in Table 4, RI, Precision, Recall and F-score of ABOF-GAP on the six UCI data sets are all higher than that of AP and other algorithms. The five indexes of ABOF-GAP is also higher than that of GAP on Wine and Iris. However, the improvement effect of ABOF on clustering by GAP method on Glass is not very significant compared to performing GAP method directly. Furthermore, NMI of ABOF-GAP is slightly lower than that of GAP on Glass. This may because there are no outliers in the Glass dataset. And removing the two points with the smallest ABOF slightly leads to worse clustering results.

In addition, based on the above data experiments, we compared the clustering results of AP, LOF-SAP, CLAP,



FIGURE 7. The clustering results on wine.

SSAPEC, DDAP and ABOF-GAP algorithm on the Wine dataset. The comparison results of the six algorithms are shown in Figure 7. The RI value of ABOF-GAP is significantly higher than that of the other algorithms. And the five indexes of ABOF-GAP reach the highest values when 3 outliers are removed.

Next we tested the effect of ABOF-GAP algorithm on text datasets. We chose BBC dataset in this work, which consists of 737 text documents, 4613 terms and 85576 words from the BBC website [27]. The data set can be divided into 5 categories: athletics, cricket, football, rugby and tennis. We use TF-IDF (Term Frequency-Inverse Document Frequency) to construct the feature vector of text [28]. In order to evaluate the influence of dimension on the clustering effect of the algorithm, the feature vector matrix of BBC text is processed into different dimensions through PCA (Principal Components Analysis) algorithm. We performed AP, LOF-SAP, CLAP, SSAPEC, DDAP, GAP and ABOF-GAP algorithm to cluster BBC dataset. As shown in Table 5, the 4 indexes of ABOF-GAP algorithm are higher than other algorithms except Recall. Figure 8 shows the RI achieved by AP, LOF-SAP, CLAP, SSAPEC, DDAP, GAP and ABOF-GAP algorithm, plotted against the dimension of text feature vector. We can find that with the increase of dimension, the clustering effect of the ABOF-GAP algorithm remains relatively stable while most of the other methods show a downtrend. And the gap of clustering effect between other algorithms and ABOF-GAP algorithm increases gradually.

The real data experimental results demonstrate that the proposed ABOF-GAP algorithm is an effective algorithm.



FIGURE 8. RI of clustering result on BBC dataset.

The overall performance of the proposed ABOF-GAP algorithm is better than that of other existing AP-related methods. The results also indicate that the NMI of ABOF-GAP should be improved for such a data set like Glass, which will be studied in our future work.

V. CONCLUSION

Considering the problems caused by outliers, and nonlinearity of data in most AP-based algorithms, an improved AP algorithm based on angle-based outlier factor and geodesic distance (ABOF-GAP) is proposed in this work. The distinguishing features of the proposed ABOF-GAP algorithm are that outliers is firstly identified and eliminated based on ABOF method. Then Euclidean distance is replaced by geodesic distance to more realistically measure the similarity. Experiments on six datasets in UCI and BBC dataset are carried out. Five evaluation indexes (RI, NMI, Precision, Recall and F-score) are adopted to evaluate the performance of clustering. The experimental results show that the proposed ABOF-GAP algorithm is an effective algorithm and its overall performance is better than that of other existing AP-related methods.

This work attempts to develop a new AP-based algorithm to conveniently and efficiently solve some problems in data processing, when the clustering effect is not ideal or the clustering performance needs to be improved. Since ABOF-GAP is an algorithm involving outlier problem, it is more suitable for clustering data with outliers. In addition, computational complexity and memory cost become a great bottleneck of AP when handling extremely large-scale datasets. In our future work, we will embark on solving the problem of AP clustering for extremely large-scalable and structurally complex datasets.

REFERENCES

- G. Wang, C. Bu, and Y. Luo, "Modified FDP cluster algorithm and its application in protein conformation clustering analysis," *Digit. Signal Process.*, vol. 92, pp. 97–108, Sep. 2019.
- [2] C. Wu and Z. Kang, "Robust entropy-based symmetric regularized picture fuzzy clustering for image segmentation," *Digit. Signal Process.*, vol. 110, Mar. 2021, Art. no. 102905.
- [3] X. Zhao, Y. Li, and Q. Zhao, "Mahalanobis distance based on fuzzy clustering algorithm for image segmentation," *Digit. Signal Process.*, vol. 43, pp. 8–16, Aug. 2015.
- [4] A. Khamparia, G. Saini, D. Gupta, A. Khanna, S. Tiwari, and V. H. C. de Albuquerque, "Seasonal crops disease prediction and classification using deep convolutional encoder network," *Circuits, Syst., Signal Process.*, vol. 39, no. 2, pp. 818–836, Feb. 2020.
- [5] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognit.*, vol. 36, no. 2, pp. 451–461, Feb. 2003.
- [6] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, Oakland, CA, USA, 1967, vol. 1, no. 14, pp. 281–297.
- [7] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 972–976, Feb. 2007.
- [8] K. Chehdi, M. Soltani, and C. Cariou, "Pixel classification of large-size hyperspectral images by affinity propagation," *J. Appl. Remote Sens.*, vol. 8, no. 1, Aug. 2014, Art. no. 083567.
- [9] C. Yang, L. Bruzzone, R. Guan, L. Lu, and Y. Liang, "Incremental and decremental affinity propagation for semisupervised clustering in multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 3, pp. 1666–1679, Mar. 2013.
- [10] C. Yang, L. Bruzzone, F. Sun, L. Lu, R. Guan, and Y. Liang, "A fuzzystatistics-based affinity propagation technique for clustering in multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2647–2659, Jun. 2010.
- [11] C.-D. Wang, J.-H. Lai, C. Y. Suen, and J.-Y. Zhu, "Multi-exemplar affinity propagation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2223–2237, Sep. 2013.
- [12] P. Li, H. Ji, B. Wang, Z. Huang, and H. Li, "Adjustable preference affinity propagation clustering," *Pattern Recognit. Lett.*, vol. 85, pp. 72–78, Jan. 2017.
- [13] J. Liu, X.-D. Zhao, and Z.-H. Xu, "Identification of rock discontinuity sets based on a modified affinity propagation algorithm," *Int. J. Rock Mech. Mining Sci.*, vol. 94, pp. 32–42, Apr. 2017.
- [14] C. Wang, "A sample entropy inspired affinity propagation method for bearing fault signal classification," *Digit. Signal Process.*, vol. 102, Jul. 2020, Art. no. 102740.
- [15] S. Raghunath, S. Rajagopalan, R. Karwoski, B. Bartholmai, and R. Robb, "Quantitative image analytics for stratified pulmonary medicine," in *Proc.* 9th IEEE Int. Symp. Biomed. Imag. (ISBI), May 2012, pp. 1779–1782.

- [16] L. Qi and L. Ting, "Active semi-supervised affinity propagation clustering algorithm based on local outlier factor," in *Proc. 37th Chin. Control Conf.* (CCC), Jul. 2018, pp. 9368–9373.
- [17] H.-P. Kriegel, M. Schubert, and A. Zimek, "Angle-based outlier detection in high-dimensional data," in *Proc. KDD*, 2008, pp. 444–452.
- [18] F. Cao, M. Estert, W. Qian, and A. Zhou, "Density-based clustering over an evolving data stream with noise," in *Proc. SIAM Int. Conf. Data Mining*, Apr. 2006, pp. 328–339.
- [19] J. Zhang, M. He, and Y. Dai, "Modified affinity propagation clustering," in *Proc. IEEE China Summit Int. Conf. Signal Inf. Process. (ChinaSIP)*, Jul. 2014, pp. 505–509.
- [20] T. Hu and S. Y. Sung, "Detecting pattern-based outliers," *Pattern Recognit. Lett.*, vol. 24, no. 16, pp. 3059–3068, Dec. 2003.
- [21] R. Guan, X. Shi, M. Marchese, C. Yang, and Y. Liang, "Text clustering with seeds affinity propagation," *IEEE Trans. Knowl. Data Eng.*, vol. 23, no. 4, pp. 627–637, Apr. 2011.
- [22] A. Strehl and J. Ghosh, "Cluster ensembles—A knowledge reuse framework for combining multiple partitions," *J. Mach. Learn. Res.*, vol. 3, no. 3, pp. 583–617, 2003.
- [23] H. Ge, L. Wang, H. Pan, Y. Zhu, X. Zhao, and M. Liu, "Affinity propagation based on structural similarity index and local outlier factor for hyperspectral image clustering," *Remote Sens.*, vol. 14, no. 5, p. 1195, Feb. 2022.
- [24] Q. Lei and T. Li, "Semi-supervised selective affinity propagation ensemble clustering with active constraints," *IEEE Access*, vol. 8, pp. 46255–46266, 2020.
- [25] Y. Li, C. Guo, and L. Sun, "Fast clustering by affinity propagation based on density peaks," *IEEE Access*, vol. 8, pp. 138884–138897, 2020.
- [26] (2022). UC Irvine Machine Learning Repository. [Online]. Available: http://archive.ics.uci.edu/ml/
- [27] A. I. Kadhim, Y.-N. Cheah, and N. H. Ahamed, "Text document preprocessing and dimension reduction techniques for text document clustering," in *Proc. 4th Int. Conf. Artif. Intell. with Appl. Eng. Technol.*, Dec. 2014, pp. 69–73.
- [28] K. S. Jones, "A statistical interpretation of term specificity and its application in retrieval," J. Document., vol. 28, no. 1, pp. 11–21, 1972.



CHAOJIE WANG received the M.Sc. and Ph.D. degrees in applied mathematics from Beihang University, in 2015 and 2019, respectively. From 2017 to 2018, he visited the University of Birmingham funded by the China Scholar Council. In 2019, he joined the Department of Mathematics and Statistics, Beijing Technology and Business University. His current research interests include digital signal and image processing, machine learning, and preconditioning iterative methods for large-scale linear systems.



JIAQI JU received the B.Sc. degree in chemical engineering and technology from Xiamen University, in 2016. She is currently pursuing the M.Sc. degree with the Department of Mathematics and Statistics, Beijing Technology and Business University. Her current research interests include data science and business statistics.

...