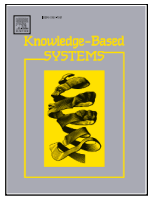
Knowledge-Based Systems 339 (2026) 115646

Contents lists available at ScienceDirect

Knowledge-Based Systems

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



Semidefinite program-inspired continuous relaxation robust multi-view clustering for large-scale data

Minghao Li [a,1], Huiying Xu [a,1,*], Ziying Wang [a], Hechuan Lin [a], Xiaolei Zhang [b,*], Xinzhong Zhu [a,*]

[a] College of Computer Science and Technology, Zhejiang Normal University, Jinhua, 321004, ZheJiang, China
[b] Academic Affairs Office, Zhejiang Normal University, Jinhua, 321004, ZheJiang, China

ARTICLE INFO

Keywords:
Multi-view clustering
Robustness
Large-scale
Semidefinite program relaxation
Iteration-free

ABSTRACT

In the modern era, with diverse and massive information, a large number of unlabeled large-scale multi-view datasets have emerged. Multi-view Clustering (MVC) has the potential to handle this type of dataset. However, in existing research, MVC methods have insufficient consideration of robustness and scalability. To address these issues, we propose a novel **S**emi-definite **P**rogram-inspired Continuous **R**elaxation Robust **M**ulti-view **C**lustering for **L**arge-**S**cale Data (SPRMC-LS) whose complexity is linear with the sample number. This model integrates the anchor strategy into the kernel function to obtain the view-specific similarity matrix. An iteration-free SDP-inspired continuous relaxation approach makes the indicator matrix able to distinguish outliers. Finally, we use a simple binary classification function to fuse the results of each view. Experiments conducted on several datasets demonstrate the superiority of our SPRMC-LS.

1. Introduction

With the advent of the vast data era, the volume of unprocessed data has reached unprecedented levels, and most of it has complicated and varied descriptions. On the one hand, the diversity of descriptions and sources breeds the concept of multi-view data [1]. A traditional Chinese parable called "*The blind man and the elephant* [2]" aptly illustrates the concept of multi-view data. On the other hand, the increasingly large and complex data requires algorithms with lower space-time complexity and higher robustness.

In recent years, **M**ulti-**V**iew **C**lustering (MVC) [3–5] is a popular research topic in the field of clustering. It aims to make maximum use of heterogeneous and complementary multi-view data and seeks a consistent clustering result by integrating information from different views. Compared with the traditional single clustering method, MVC has broader application prospects in many fields. Based on the different information fusion strategy, we can roughly divide current MVC approaches into four categories, i.e., **N**on-negative **M**atrix **F**actorization (NMF) [6–8], subspace clustering [9–12], co-training style [13,14], and graph-based [15–17]. Specifically, NMF-based methods use matrix decomposition to learn low-dimensional consensus representation from raw multi-view data. Nevertheless, it has poor efficiency when taking large-scale data, attributed to direct factorization of the original data observations. Subspace-based clustering methods premise that views are bred from the latent subspace and attempt to explore subspace representations. With the help of an alternating training strategy, co-training based methods combine multiple results to maximize the synergy between individual views. The objective of graph-based methods is to obtain a unified graph matrix to study the correlations between different views.

Despite the remarkable success, the high computational cost and the problem of outliers are still two inevitable problems for MVC, especially when handling large-scale data. To reduce the computational burden, the anchor graph-based strategy has received extensive attention due to its efficiency and practicability [18–20]. However, robustness to large-scale data is often overlooked by existing anchor-based MVC methods. In addition, most current MVC methods rely on multiple iterative optimization, which may require additional time to obtain a theoretical approximate solution. In the research on clustering about outliers, [21] shows that SDP-based clustering algorithms are strongly consistent without outliers, and robust to arbitrary outliers so long as the number of outliers increases slower than the number of nodes. Although many SDP-based clustering algorithms already exist [21–24], seldom of them center on large-scale multi-view clustering.

* Corresponding authors.
E-mail addresses: liminghao@zjnu.edu.cn (M. Li), xhy@zjnu.edu.cn (H. Xu), wzyszjnu@163.com (Z. Wang), linhechuan2@gmail.com (H. Lin), znuzxl@zjnu.cn (X. Zhang), zxz@zjnu.edu.cn (X. Zhu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.knosys.2026.115646
Received 16 January 2025; Received in revised form 22 January 2026; Accepted 24 February 2026
Available online 28 February 2026

Table 1
Main notations used throughout the paper.

Notation	Meaning
n	number of observations
m	number of selected anchors
v	number of views
c	number of clusters
$\lambda^{(v)}$	regularization coefficient for the i-th view
$\mathbf{E}_{nm}$	matrix full of ones in size of $n \times m$
$\mathbf{x}^{(v)}$	data matrix for the i-th view
$x_{i:}$ and $x_{:j}$	i-th row and j-th column of the matrix $\mathbf{x}$
$\mathbf{B}^{(v)}$	anchor selection matrix for the i-th view
$\mathbf{Z}^{(v)}$	cluster membership matrix for the i-th view
$\mathbf{S}$	consensus matrix

In light of this, a novel iteration-free MVC method called **S**emidefinite **P**rogram-inspired Continuous **R**elaxation Robust **M**ulti-view **C**lustering for **L**arge-**S**cale Data (SPRMC-LS) is proposed in this paper. We incorporate anchor strategy, kernel approach, and semidefinite program (SDP) into a unified framework. Specifically, we integrate anchor strategy into the kernel function for large-scale data. Then, after the data have been projected, we apply an SDP-based relaxation to obtain an indicator matrix, which can achieve better results when the dimensions are larger. The outliers can be directly distinguished by defining an auxiliary matrix consisting of the indicator matrix. Finally, a simple binary classification function is adopted to fuse the data from each view, improving the adaptability to large-scale data again. For clarity, the detailed process of SPRMC-LS is illustrated in Fig. 1. We summarize our contribution as follows:

1. We propose a novel robust MVC approach (SPRMC-LS). Unlike heuristic outlier removal in existing large-scale MVC methods, our SPRMC-LS achieves robustness against arbitrary outliers by integrating the SDP-inspired continuous relaxation into the anchor-based MVC framework. Outliers are identified as zero rows of the auxiliary matrix derived from the SDP structure.
2. Unlike most existing large-scale MVC approaches that rely on iterative optimization, we derive a closed-form solution via convex relaxation, making SPRMC-LS iteration-free and highly efficient. This closed-form solution stems from a novel optimization formulation that transforms the discrete clustering problem into a convex one, eliminating the need for iterative approximations.
3. Our SPRMC-LS achieves linear complexity in both time and memory with respect to the number of instances, which is a fundamental improvement over existing SDP-based clustering methods. This linear complexity is achieved without sacrificing robustness, as opposed to existing efficient MVC methods that often overlook theoretical robustness.

2. Related work

This section briefly introduce relaxed work in large-scale and SDP-based MVC methods. we summarize the key notations used in this paper in Table 1.

2.1. Large-scale MVC methods

The matrix factorization(MF)-based approaches [6,25,26], and the anchor graph style algorithms [18–20] are two major components of existing large-scale MVC methods. To handle large-scale multi-view data, the MF-based approaches attempt to factorize the original data matrix into sparse matrices with lower cost calculation. The work in [6] uses the structured sparsity-inducing norm to combine heterogeneous matrices of large-scale data. In [25], the main MF-based optimization problem is divided then conquered as smaller-size sub-problems that require much fewer matrix multiplications. To accelerate the processing speed of large-scale data, clustering tasks are run in parallel on rows and columns of the data matrix in [26].

The anchor graph-based strategy attempts to replace the original data with minority anchor representations. The smaller anchor graphs are constructed by taking advantage of the correlations between anchors and the original data, which reduce the computational burden. Li and He [19] propose an anchor graph-based MVC model, which fully uses the consensus information contained in multiple views. Sun et al. [20] combine anchor learning and graph construction into a unified optimization framework.

2.2. SDP-based robust clustering

Recently, several works aim to develop robust algorithms based on semidefinite programming [27–30]. Many clustering studies are usually conducted for a mixture model with well-separated centers, for which the Sub-Gaussian Mixture Models (SGMMs) are typical. Kushagra et al. [22] improve a robust k-means clustering formulation proposed in [24] and provide recovery guarantees under arbitrary. By minimizing the objective function with explicit cardinality constraints correlated to clusters and outliers, [27] obtain a robust SDP-based clustering model.

In addition to the SGMMs condition, [29] expresses a joint training problem with SDP relaxations to discuss the robustness of support vector machines. [21] discuss the robustness of SDP-based kernel and spectral clustering methods. The more recent work is [30], which proposes an SDP-based graph clustering in the stochastic block model for both large and small clusters.

3. Methodology

3.1. Model proposal

For the v-th view of observations $\left\{\mathbf{x}_1^{(v)}, \mathbf{x}_2^{(v)}, \ldots, \mathbf{x}_n^{(v)}\right\}$, we define the correlation between i-th and j-th data points as following Gaussian kernel matrix

$$\mathbf{K}_{ij}^{(v)} = \exp\left(-\frac{\left\|\mathbf{x}_i^{(v)} - \mathbf{x}_j^{(v)}\right\|^2}{2\theta^2}\right) \tag{1}$$

for the scaling parameter θ. Then we replace the whole data points with the $m(m \ll n)$ anchor points $\left\{\mathbf{a}_1^{(v)}, \mathbf{a}_2^{(v)}, \ldots, \mathbf{a}_m^{(v)}\right\}$ as

$$\mathbf{D}_{ij}^{(v)} = \exp\left(-\frac{\left\|\mathbf{x}_i^{(v)} - \mathbf{a}_j^{(v)}\right\|^2}{2\theta^2}\right), \tag{2}$$

where anchor points are generated by random sampling. So the computational complexity can be reduced by operating on $\mathbf{D} \in \mathbb{R}^{n\times m}$ instead of $\mathbf{K} \in \mathbb{R}^{n\times n}$.

To better explain the robustness of our method, we first discuss the clustering problem with equal-sized clusters and no outliers.

$$\begin{aligned} &\max_{\mathbf{Z}^{(v)}} \left\langle \mathbf{D}^{(v)}, \mathbf{Z}^{(v)}\mathbf{Z}^{(v)\top}\mathbf{B}^{(v)} \right\rangle \\ &s.t. \sum_j z_{:j}^{(v)} = 1, \sum_i z_{i:}^{(v)} = \frac{n}{c}, \end{aligned} \tag{3}$$

where c denotes the number of clusters, $\mathbf{Z}^{(v)} \in \{0,1\}^{n\times c}$ denoted the cluster membership matrix and $\mathbf{B}^{(v)} \in \{0,1\}^{n\times m}$ denoted the anchor selection matrix that $b_{ij} = 1$ represent i-th data point is selected as j-th anchor. Based on the assignment constraint that each data point belongs to exactly one cluster, and cardinality constraint that all clusters are exactly equal-sized with $\frac{n}{c}$, Eq. (3) aims to maximize the inside similarity of each cluster. However, clustering with exactly equal-sized clusters is a rare particular case in practice. If we remove the cardinality constraint

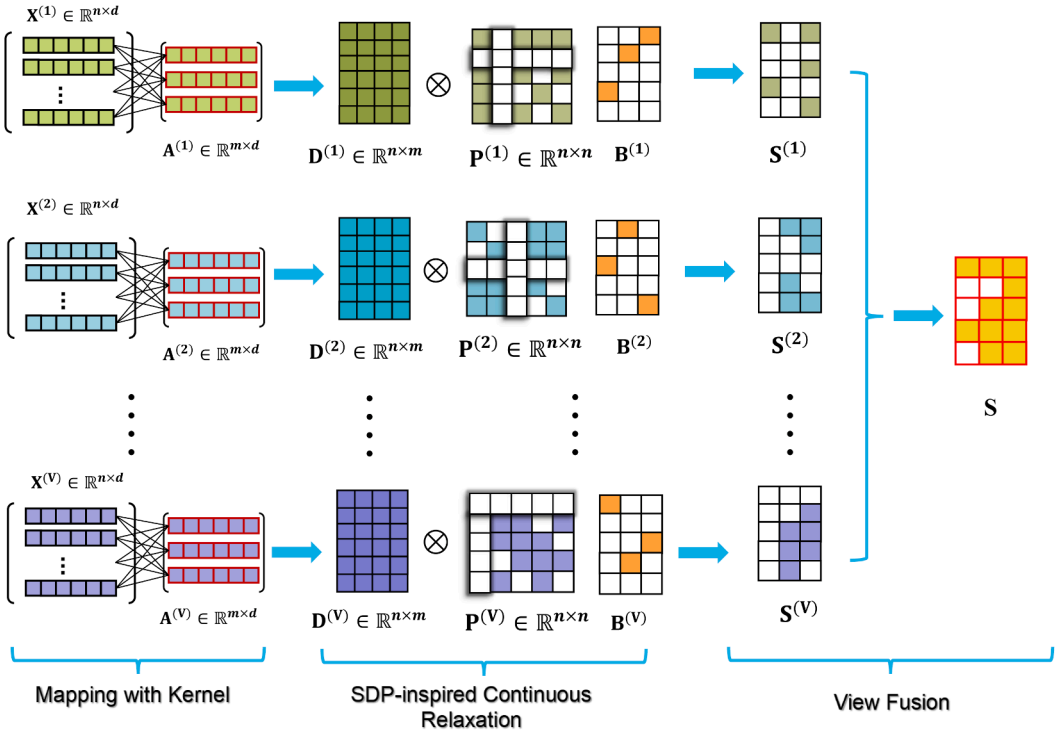


Fig. 1. The illustration of SPRMC-LS. We combine the multiple view observations $\{\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(V)}\}$ and the corresponding anchors $\{\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(V)}\}$ with an anchor-based kernel method. After obtaining the anchor graphs $\{\mathbf{D}^{(1)}, \ldots, \mathbf{D}^{(V)}\}$ by kernel method, we define an auxiliary matrix $\mathbf{P}^{(v)} = \mathbf{Z}^{(v)}\mathbf{Z}^{(v)\top}$ to transform the problem into a SDP suitable form. Outliers can be identified as zero rows in the auxiliary matrix $\mathbf{P}^{(v)}$. The final clustering results are obtained by performing spectral clustering on consensus matrix $\mathbf{S}$.

directly, the optimal solution will divide all points into the same cluster. To overcome this issue, a feasible way is

$$\begin{aligned} &\max_{\mathbf{Z}^{(v)}} \left\langle \mathbf{D}^{(v)} - \lambda^{(v)}\mathbf{E}_{nm}, \mathbf{Z}^{(v)}\mathbf{Z}^{(v)\top}\mathbf{B}^{(v)} \right\rangle \\ &s.t. \sum_j z_{:j} = 1, \end{aligned} \tag{4}$$

where $\lambda^{(v)} \in (0, 1)$ is the regularization coefficient and $\mathbf{E}_{nm} \in \mathbb{R}^{n\times m}$ is a matrix full of ones. It is worth mentioning that the penalty item $\left\langle \mathbf{E}_{nm}, \mathbf{Z}^{(v)}\mathbf{Z}^{(v)\top}\mathbf{B}^{(v)} \right\rangle$ achieves minimum value when all the clusters are equal-sized, so Eq. (4) essentially attempts to make the cluster as balanced as possible.

After that, we relax the assignment constraint to allow each data point to be either an inlier or an outlier. In this way, we extend Eq. (4) to consider the clustering problem with possible outliers.

$$\begin{aligned} &\max_{\mathbf{Z}^{(v)}} \left\langle \mathbf{D}^{(v)} - \lambda^{(v)}\mathbf{E}_{nm}, \mathbf{Z}^{(v)}\mathbf{Z}^{(v)\top}\mathbf{B}^{(v)} \right\rangle \\ &s.t.\ \mathbf{Z}^{(v)}\mathbf{1}_c \le \mathbf{1}_n. \end{aligned} \tag{5}$$

The formulation in Eq. (5) is an NP-hard discrete optimization problem that requires maximizing a quadratic, non-convex objective function over a binary matrix. Thus, we aim to transform it into a convex problem with theoretical guarantees on robustness by introducing an auxiliary matrix $\mathbf{P}^{(v)} = \mathbf{Z}^{(v)}\mathbf{Z}^{(v)\top}$ and then leveraging SDP relaxation theory [21,23]. As shown in [21], the SDP relaxation for clustering exhibits strong consistency and robustness to outliers when the number of outliers grows slower than the number of inliers. We first rewrite Eq. (5) using auxiliary matrix as follows

$$\begin{aligned} &\max_{\mathbf{P}^{(v)}} \left\langle \mathbf{D}^{(v)} - \lambda^{(v)}\mathbf{E}_{nm}, \mathbf{P}^{(v)}\mathbf{B}^{(v)} \right\rangle \\ &s.t.\ \mathbf{P}^{(v)} \succ \mathbf{0}, rank(\mathbf{P}^{(v)}) \le c, \end{aligned} \tag{6}$$

where $\mathbf{P}^{(v)} \in \{0, 1\}^{n\times n}$. The following lemma explains the equivalence between Eqs. (5) and (6).

Lemma 1 ([23]). *If P^* is an optimal solution to Eq. (5), there exists a decomposition $P^* = H^*H^{*\top}$ and an orthogonal matrix $U \in \mathbb{R}^{c\times c}$, so that $Z^* = P^*U$ is an optimal solution for Eq. (6) with the same objective function value.*

Using the auxiliary matrix $\mathbf{P}^{(v)}$, we can recognize the outliers directly by verifying the zero row vector in $\mathbf{P}^{(v)}$. However, Eq. (6) is still a non-convex optimization problem attribute to the two constraints. To transform the problem into convex optimization form, we adopted two convex relaxations for the problem. Specifically, we first relax the discrete solution to a continuous solution and then remove the rank constraint. Thus the SDP formulation of the problem can be described as follows:

$$\begin{aligned} &\max_{\mathbf{P}^{(v)}} \left\langle \mathbf{D}^{(v)} - \lambda^{(v)}\mathbf{E}_{nm}, \mathbf{P}^{(v)}\mathbf{B}^{(v)} \right\rangle \\ &s.t.\ 0 \le P_{ij}^{(v)} \le 1, \mathbf{P}^{(v)} \succ \mathbf{0}. \end{aligned} \tag{7}$$

The semidefinite constraint is tedious to take, so we relax it to simplify the solving procedure and improve Eq. (7) as

$$\begin{aligned} &\max_{\mathbf{P}^{(v)}} \left\langle \mathbf{D}^{(v)} - \lambda^{(v)}\mathbf{E}_{nm}, \mathbf{P}^{(v)}\mathbf{B}^{(v)} \right\rangle \\ &s.t.\ 0 \le P_{ij}^{(v)} \le 1. \end{aligned} \tag{8}$$

To facilitate the fusion of views, we define $\mathbf{S}^{(v)} = \mathbf{P}^{(v)}\mathbf{B}^{(v)}$ and specify the final formulation of a single view as follows:

$$\begin{aligned} &\max_{\mathbf{S}^{(v)}} \left\langle \mathbf{D}^{(v)} - \lambda^{(v)}\mathbf{E}_{nm}, \mathbf{S}^{(v)} \right\rangle \\ &s.t.\ 0 \le S_{ij}^{(v)} \le 1. \end{aligned} \tag{9}$$

It is obvious that the relationship between $\mathbf{S}^{(v)}$ and $\mathbf{P}^{(v)}$ is linear, thus the elements in $\mathbf{S}^{(v)}$ are the subset of elements in $\mathbf{P}^{(v)}$. In other words, Eqs. (8) and (9) are equivalent. It should be noted that the derivation of Eq. (9) involves two sequential relaxation steps, which is a double-edged sword. This approach significantly enhances computational efficiency by streamlining the solution objective. However, it also introduces theoretical approximate error and imposes stricter requirements

on the quality of the constructed anchor graphs. For the former, similar formulations of SDP relaxation and their approximation errors have been discussed in the community detection and clustering literature [23,31,32]. For the latter, to improve the quality of the acquired anchor graphs, techniques such as anchor alignment and consistent anchors [33–35] offer a promising solution.

From Eq. (9), note that the optimal solution of $\mathbf{S}^{(v)}$ can be obtained directly by thresholding $\mathbf{D}^{(v)}$ with $\lambda^{(v)}$ as

$$S_{ij}^{(v)} = \begin{cases} 1 & D_{ij}^{(v)} - \lambda^{(v)} > 0, \\ 0 & otherwise. \end{cases} \tag{10}$$

This closed-form solution is derived from the convex relaxation and does not require any iterative optimization. This is a key advantage over existing MVC methods that rely on iterative algorithms such as alternating minimization or gradient descent.

In the end, taking 0.5 as the threshold, we dichotomize the mean of results from each view to acquire the consensus matrix $\mathbf{S}$ specified below:

$$S_{ij} = \begin{cases} 1 & \frac{1}{V}\sum_{v=1}^{V} S_{ij}^{(v)} > 0.5, \\ 0 & otherwise, \end{cases} \tag{11}$$

where V is the number of views. This strategy fuses the view data while simultaneously achieving anomaly filtering. While alternative fusion strategies based on models or data-driven methods can be more effective, the proposed strategy provides greater interpretability and computational efficiency. The final clustering result is generated by performing the spectral clustering on $\mathbf{S}$.

In general, our SPRMC-LS has the following advantages:

- **SDP-inspired relaxation MVC algorithm**. A novel scheme is proposed to handle large-scale MVC tasks. The SDP-inspired continuous relaxation approach solves the non-convex optimization problem and improves the scalability of the algorithm. Although the SDP-inspired continuous relaxation clustering method has appeared in previous research, we first employ and improve it for MVC.
- **Identification of outliers in non-linear higher dimensional space**. SPRMC-LS integrates anchor-based kernel clustering and outlier detection. SPRMC-LS utilizes kernel function to obtain the bipartite graph structure from a high dimensional perspective, then identifies the outliers by an auxiliary matrix, thus the outliers can be easily separated.
- **High computational efficiency**. SPRMC-LS is solved by an iteration-free algorithm. Theoretical analysis shows that SPRMC-LS has linear complexity in both time and memory space complexity. Thus, it can be applied effectively to large-scale and scalable clustering tasks.

3.2. Parameter selection

The selection of parameters λ and θ is crucial for our algorithm. In our experiments, we employ a statistical approach that relies on two preset hyper-parameters, α and β, to determine the parameters.

Selection of the parameter θ**.** Inspired by [36], it should be at least a part β of adjacent points that can be mapped by the kernel function for $(1-\alpha)\times n$ of the data points. Thus we select the optimal value of θ in low dimension problem as followed:

$$\theta = \frac{(1-\alpha)\ quantile\ of\ \{l_1, l_2, \cdots, l_n\}}{\sqrt{(1-\alpha)\ quantile\ of\ \chi_d^2}}, \tag{12}$$

where l_i is the β quantile of the distance between point i and other points in the dataset, χ_d^2 is a chi-square distribution with d freedom. A sufficiently high value of β can ensure that the points of the same cluster have a high degree of similarity, and a small value of α corresponds to high similarity between non-noise points. ***Selection of the parameter*** λ**.** The robustness of our algorithm is heavily influenced by the selection of λ. A value of λ approach to 1 or 0 will generate an all-ones or all-zeros matrix of $\mathbf{P}^{(v)}$ after rounding, invalidating the denoising process. Since anchor graph $\mathbf{D}^{(v)}$ is constructed by the Gaussian kernel function, we select the value of λ by setting the distance in the kernel function equal to the $(1-\alpha)\ quantile\ of\ \{l_1, l_2, \cdots, l_n\}$. Specially, we define λ as $\lambda = e^{-\frac{t_\alpha}{2}}$, where $t_\alpha = (1-\alpha)\ quantile\ of\ \chi_d^2$. Thus, the threshold λ is theoretically motivated by the quantile of the chi-square distribution. This approach links the regularization parameter to the inlier probability of data points, thereby enhancing robustness.

Table 2
The detailed information of multi-view datasets in experiments.

Dataset	Size(n)	Views(v)	Class(k)	Dimensionality(d)
ProteinFold	694	12	27	27/27/.../27
100Leaves	1600	3	100	64/64/64
UCI-digit	2000	3	10	64/76/216
Handwritten	2000	6	10	6/47/64/76/216/240
WiKi	2866	2	10	10/128
ALOI-100	10800	4	100	77/13/64/125
Cifar10	50000	3	10	512/1024/2048
Mnist	60000	3	10	54/342/1024
YouTubeFace	101499	5	31	64/64/512/647/838
CoverType	581012	2	7	10/44

3.3. Computational complexity analysis

The computational complexity of SPRMC-LS includes three main parts without iterative optimization. It cost $\mathcal{O}(nmd)$ to compute the improved kernel matrix $\mathbf{D}^{(v)}$. Computing the auxiliary matrix $\mathbf{S}^{(v)}$ needs $\mathcal{O}(nm)$, and perform SVD on consensus matrix $\mathbf{S}$ to obtain the spectral embedding then generate the final clustering labels by k-means [37] needs $\mathcal{O}(nm^2)$. So the total time cost of SPRMC-LS is $\mathcal{O}(n(md + m + m^2)$, which is linear complexity $\mathcal{O}(n)$.

4. Experiments

In this section, we demonstrate the superiority of SPRMC-LS by executing sufficient experiments on different scale datasets. All datasets are available from the public website.[1]

4.1. Datasets

To obtain a more fair and comprehensive evaluation of the experiment results. We adopte several frequently-used muti-view benchmark datasets, including six small and media-scale datasets: proteinFold [38], 100Leaves, UCI-digit, Handwritten [39], WIkI, ALOI-100 and four large-scale datasets: Mnist, Cifar10, YoutubeFace [40] and CoverType. Table 2 shows the detailed statistical information of the above datasets.

ProteinFold: It was originally collected by researchers at the University of California, San Diego, which aims to promote the development of protein folding prediction. The characteristic of the dataset is abundant and diverse protein sequence information.

100Leaves: It is a plant leaves classification dataset that includes 1600 samples and is composed of 16 colorful leaf images. The 100 categories add diversity to the dataset. In our experiment, each image is expressed as three features with shape descriptors, fine-scale boundaries, and texture histograms.

UCI-digit: It is a Handwritten digit recognition dataset collected from the Dutch utility map. The 2000 samples are divided into ten classes and indicated by six features. We adopt three of them in our experiments: profile correlations (216 dimensions), Fourier coefficients in

[1] https://github.com/wangsiwei2010/large_scale_multi-view_clustering_datasets

Table 3
Comparison of ACC, NMI and PUR among different MVC methods on small and media-scale datasets. The best results are marked in bold, and the second-bset results are marked in underlined.

Dataset	Metric	BMVC	MVCC	GMC	TBGL	LMVSC	EOMSC	MVSC-HFD	FSMSC	LAIMVC	OURS
ProteinFold	ACC	20.17	16.57	23.78	20.31	26.95	30.83	27.67	31.70	**34.73**	32.85
	NMI	26.71	23.50	19.69	12.36	36.75	38.26	36.01	40.60	**43.77**	40.85
	PUR	25.79	22.05	26.66	20.76	45.53	35.44	34.01	38.18	41.21	**94.96**
100Leaves	ACC	82.81	14.75	82.38	64.06	65.75	42.56	40.94	40.94	64.56	**87.75**
	NMI	91.94	47.56	90.25	71.92	83.93	76.31	73.33	81.40	71.76	**94.84**
	PUR	85.13	14.75	85.06	66.62	76.94	45.06	42.44	66.69	51.01	**92.31**
UCI-digit	ACC	72.95	50.55	84.95	84.55	89.35	91.10	81.30	89.55	85.10	**94.25**
	NMI	64.74	48.75	84.47	87.29	83.15	83.55	80.26	80.62	74.99	**88.68**
	PUR	76.45	52.00	87.40	87.05	89.35	91.10	82.01	89.55	85.10	**94.25**
Handwritten	ACC	75.95	50.95	88.20	87.10	84.20	76.00	64.40	86.90	68.35	**95.85**
	NMI	73.86	42.23	89.32	87.30	88.96	82.55	70.08	76.98	61.60	**91.30**
	PUR	75.95	54.80	88.25	87.10	88.00	76.45	65.60	86.90	68.45	**95.85**
WiKi	ACC	15.84	47.24	19.47	31.37	57.64	56.35	58.06	26.24	56.66	**61.10**
	NMI	3.72	34.85	4.54	21.53	53.92	53.27	52.85	10.57	54.30	**54.11**
	PUR	20.10	52.44	20.24	34.26	66.12	61.79	61.51	28.37	62.88	**66.26**
ALOI-100	ACC	54.32	43.26	**72.07**	42.80	65.14	23.4	35.25	63.20	57.42	66.26
	NMI	71.11	63.43	74.36	55.64	68.72	57.91	64.65	76.73	75.10	**76.89**
	PUR	57.87	45.48	**73.03**	52.69	66.81	24.7	38.45	64.56	60.44	71.13

character shapes (76 dimensions), and Karhunen-Love coefficients (64 dimensions).

Handwritten: It is constituted by 2000 Handwritten digit images, and those images are divided into ten categories represent digit from 0 to 9. Each digit image is described by six features.

WiKi: It comprises 2866 multimedia documents collected from Wikipedia articles, categorized into 10 subjects. We adopted two views, TF-IDF of the titles (10 dimensions) and TF-IDF of the content (128 dimensions) in our experiment.

ALOI-100: ALOI is a classic multi-view image dataset. It contains numerous images of small objects taken under diverse lighting conditions and rotation angles. We use the ALOI subset of 10,080 samples across 100 categories. Each sample is represented by four feature views: color similarity, Haralick texture features, HSV color space characteristics, and RGB color histogram.

Cifar10: It is a large-scale image dataset widely used in machine learning research. The dataset contains 50,000 color images in ten categories. Each image is 32×32 pixels, extracted by the neural network to form a three-view dataset with 512, 2048, and 1024.

Mnist: It is a large-scale Handwritten digits dataset consisting of 60.000 samples. Each sample is a 28×28 pixel grayscale image from 250 different people.

YouTubeFace: It is a face video database established for researching the field of unconstrained face recognition. The entire sample of the database are collected from YouTube, containing 3,425 videos of 1595 different people.

CoverType: This dataset contains 581,012 samples of seven vegetation coverage types, collected from four wilderness areas within the Roosevelt National Forest in northern Colorado, USA.

4.2. Compared methods and evaluation metrics

To compare the performance of our SPRMC-LS, we select seven MVC methods as the baseline and measure them by three widely used clustering metrics, *i,e.*, accuracy (ACC), normalized mutual information (NMI) and purity (PUR). The higher value of those metrics means the better clustering results. We briefly introduce the compared methods as follows.

BMVC [41]: BMVC represents the multi-views data by the collaborative binary codes then obtains the clustering results by executing matrix decomposition on binary matrix.

MVCC [42]: MVCC reach a consensus of multiple views to better use the consistent and complementary information from multi-view data. it adopt a local manifold regularization to maintain locally geometrical structure. **GMC** [43]: GMC fuses the data from each view by a uniform graph matrix and then extracts eigenvectors of the matrix as clustering results directly.

TBGL [44]: TBGL integrates the tensor scheme with the bipartite graph learning to explore the structure of multiple views.It provides a novel perspective of bipartite graph learning.

LMVSC [45]: LMVSC is an efficient MVC algorithm for large-scale data with liner time complexity. It aims to address large-scale MVC issue by combining the anchor graph with subspace MVC.

EOMSC [46]: EOMSC combines anchor learning with graph fusing and then implements the one-pass process by imposing a rank constraint on the consensus matrix to obtain the clustering results directly without post-processing.

MVSC-HFD [47]: MVSC-HFD attempts to address the discrepancy among views by searching a unified subspace and assuming the structure of different views is the same in the subspace. The integration of the bipartite graph well balances the dimensional projection and efficiency.

FCMVSC [48]: FCMVSC is a subspace-based large-scale MVC method. By integrating view-shared anchor learning and global-guided-local self-guidance, the method ensures high efficiency and strong discriminability in modeling fuzzy clustering structures.

LAIMVC [49]: LAIMVC is a lightweight continual MVC method that employs a elegant technique seamlessly integrate anchor learning into an incremental framework.

We downloaded the source codes for all compared methods from publicly available repositories provided by their respective authors.

4.3. Experimental setting

The parameter settings of these comparison methods strictly follow the description in the corresponding paper. For the parameter settings of our SPRMC-LS, we search the best hyperparameters α in range of $[0.01, 0.02, \ldots, 0.08, 0.09]$ and β in range of $[0.05, 0.06, \ldots, 012, 0.13]$. Meanwhile, we explore the optimal anchor number by using grid search in the range of $\left[c, 2c, 3c, 4n^{\frac{1}{2}}, 5n^{\frac{1}{2}}, 6n^{\frac{1}{2}}, 7n^{\frac{1}{2}}, 8n^{\frac{1}{2}}\right]$ for large-scale datasets, and in the range of $\left[\frac{n}{20}, \frac{2n}{20}+, \frac{3n}{20}+, \ldots, \frac{19n}{20}, n\right]$ for small and media-scale datasets.

Table 4
Comparison of ACC, NMI and PUR among different MVC methods on large-scale datasets. The best results are marked in bold, and the second-best results are marked in underlined.'-'means out of the CPU memory.

Dataset	Metric	BMVC	LMVSC	EOMSC	MVSC-HFD	FSMSC	LAIMVC	OURS
Cifar10	ACC	53.08	98.96	99.10	99.09	**99.15**	87.38	98.91
	NMI	38.28	97.21	97.53	97.51	**97.67**	74.43	97.01
	PUR	53.08	98.96	99.10	99.09	99.15	87.38	**99.98**
Mnist	ACC	27.26	98.99	**99.10**	98.03	98.88	93.13	98.78
	NMI	18.79	96.93	94.75	94.65	96.58	86.59	**98.26**
	PUR	30.15	98.99	98.08	98.03	98.88	93.13	**99.99**
YouTubeFace	ACC	8.28	14.60	26.66	13.80	25.39	18.66	**26.89**
	NMI	4.22	12.98	10.40	10.50	**21.89**	17.38	7.59
	PUR	26.62	22.58	27.40	27.65	34.22	29.34	**99.99**
CoverType	ACC	-	42.68	-	-	-	25.53	**48.76**
	NMI	-	**15.42**	-	-	-	10.36	6.06
	PUR	-	68.20	-	-	-	56.85	**99.99**

Table 5
Running time comparison of different algorithms on small and media-scale datasets. (in seconds).

Dataset	BMVC	MVCC	GMC	LMVSC	TBGL	EMOSC	MVSC-HFD	FSMSC	LAIMVC	OURS
ProteinFold	0.051	44.831	2.011	1.442	60.010	1.555	1.596	18.637	0.005	0.046
100Leaves	0.201	47.839	2.542	0.876	114.903	9.154	6.806	28.886	0.015	0.102
UCI-digit	0.193	40.231	5.876	0.959	222.244	2.768	15.214	26.824	0.019	0.150
Handwritten	0.286	77.886	7.418	1.929	378.558	3.107	2.858	17.436	0.002	0.293
WiKi	0.264	19.321	10.635	0.638	301.784	3.474	3.383	22.017	0.004	0.182
ALOI-100	2.622	162.880	461.434	8.296	23680.302	55.195	56.496	221.099	0.041	6.784

Table 6
Running time comparison of different algorithms on large-scale datasets. (in seconds) '-'means out of the CPU memory.

Dataset	BMVC	LMVSC	EMOSC	MVSC-HFD	FSMSC	LAIMVC	OURS
Cifar10	31.422	20.039	86.290	360.081	884.374	2.952	19.088
Mnist	44.143	20.716	116.530	412.027	1013.623	0.132	0.746
YouTubeFace	207.161	66.463	462.911	1064.600	2632.030	4.290	3.989
CoverType	-	178.550	-	-	-	1.149	0.696

All experiments are performed on a PC with 11th Gen Intel(R) Core(TM) i7-11700K CPU and 64 GB RAM with MATLAB R2021b.

4.4. Experimental results analysis

Tables 3 and 4 demonstrate the experimental results between different MVC methods with three metrics on small and media-scale and large-scale datasets respectively. We have the following conclusions from the two tables.

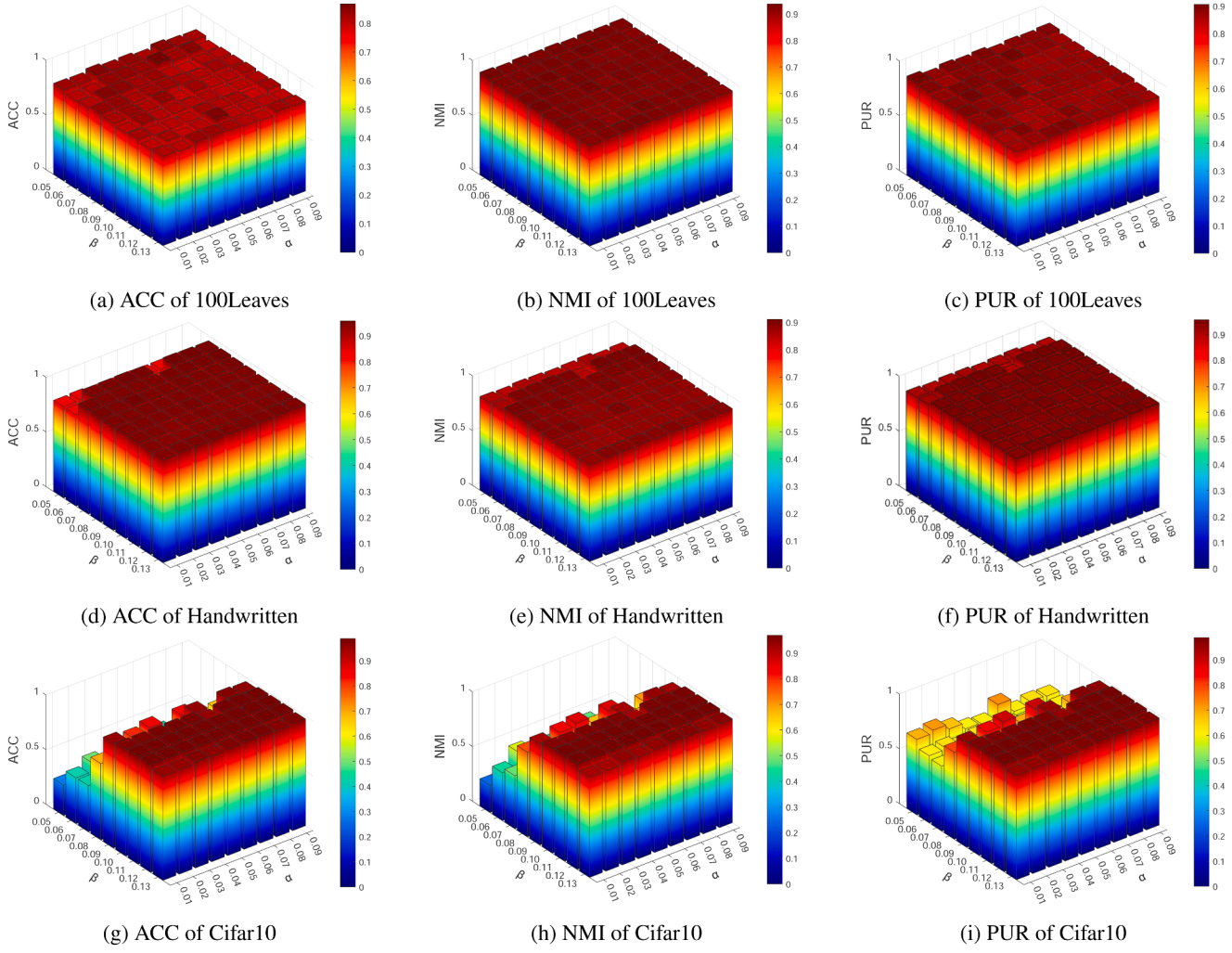
- In almost all datasets, our SPRMC-LS achieves the best performance in the case of three metrics against all compared methods. On small and media-scale dataset such as Handwritten, our SPRMC-LS outperforms the GMC, which achieves the second-best results, with improvements 8.67%, 2.21% and 8.61% in three metrics. In large-scale datasets such as Mnist, although our SPRMC-LS is only 0.2% less than the best result achieved by LMVSC on ACC, it achieves the best on NMI and PUR.
- Compared methods with tedious calculation processes like GMC, and TBGL are troubled by out-of-memory when dealing with large-scale data. However, LMVSC, EOMSC and our SPRMC-LS are able to handle large-scale datasets with superior performance.
- Our SPRMC-LS shows excellent performance on PUR for all datasets, demonstrating its high efficiency in clustering. It is worth mentioning that our SPRMC-LS presents the situation of high PUR but low ACC and NMI in the YouTubeFace dataset. We consider it caused by the specificity of the datasets. Specifically, the dataset has imbalanced distributed categories. Thus, points belonging to a big category are clustered in the same way as points of other categories, which has a greater impact on the final result.

In summary, there experimental results indicate that SPRMC-LS has superior clustering performance on different-scale datasets and can well deal with large-scale datasets with over 500,000 instances.

4.5. Computational time analysis

To further explain the efficiency of our approach, we demonstrate the running time comparison of different methods on Table 5 and 6. In Table 6, we only report the running time results of BMVC, LMVSC, and EMOSC since other methods are out-of-memory or overlong running time for large-scale datasets.

It can be observed that our SPRMC-LS has a absolute advantage on both small and media-scale and large-scale datasets. Specifically, Table 5 shows that BMVC has less time expenditure than other compared methods since its simple binary clustering structure, but our SPRMC-LS is better than BMVC in running time. LMVSC has been known as the fastest MVC method for large-scale data. Table 6 demonstrates that compared with LMVSC, our SPRMC-LS has a comparable running time in the Cifar10 dataset and almost saves half the time on the Mnist and YouTubeFace datasets.Thus, our SPRMC-LS can achieve great clustering performance with short running time, making it well-suited for handling practical clustering.

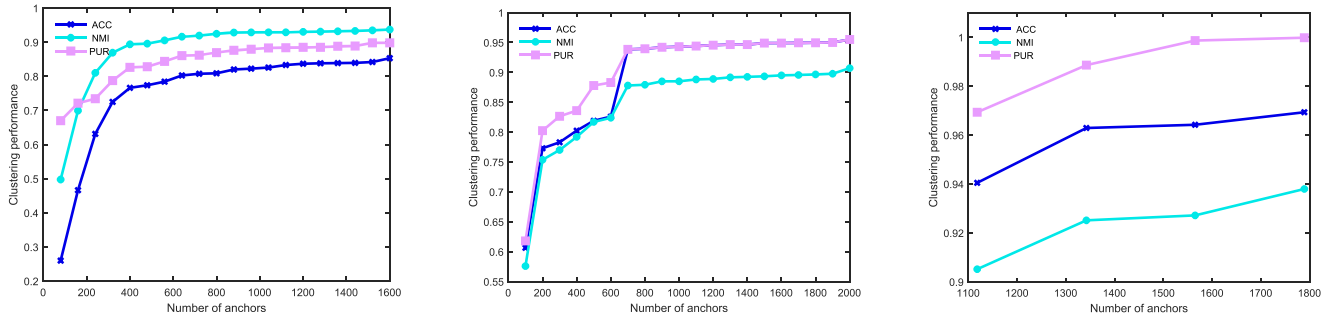


(a) ACC of 100Leaves (b) NMI of 100Leaves (c) PUR of 100Leaves

(d) ACC of Handwritten (e) NMI of Handwritten (f) PUR of Handwritten

(g) ACC of Cifar10 (h) NMI of Cifar10 (i) PUR of Cifar10

Fig. 2. Sensitivity analysis of α and β for our SPRMC-LS on the three datasets 100Leaves, Handwritten, and Cifar10.



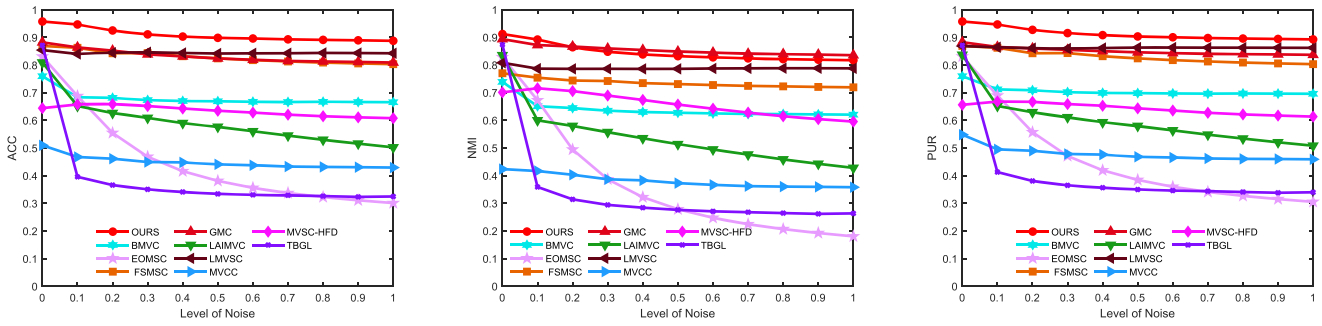
(a) 100Leaves (b) Handwritten (c) Cifar10

Fig. 3. Clustering performance with different numbers of anchors for our SPRMC-LS on the 100Leaves, Handwritten, and Cifar10 datasets.

4.6. Parameter sensitivity analysis

Three hyper-parameters (α, β, and the number of anchors), with precise settings, influence the efficiency of the proposed SPRMC-LS. To explore the sensitivity of the parameters for SPRMC-LS, we take the 100Leaves, Handwritten and Cifar10 datasets as the example to conduct a comparative experiment. The results with different setting of α and β in term of clustering ACC, NMI, and PUR are shown in Fig. 2. It can be seen from these figures that our SPRMC-LS is basically unaffected by α and β. Furthermore, the clustering results are stable within the parameters we set in the 100Leaves and Handwritten dataset, with a narrower stability range of β in the Cifar10 dataset. The value β represents the range of mapping of adjacent points using the kernel function. Thus, the experiments explain that the large-scale data require the kernel function to map more adjacent points.

As shown in Fig. 3, the clustering performance improves and then gradually stabilizes with the increase in the number of anchors, demonstrating that our model is able to obtain efficient and stable results by choosing an appropriate number of anchors in the recommended range.



(a) ACC (b) NMI (c) PUR

Fig. 4. Comparison of ACC, NMI and PUR on Handwritten with different noise level. The noise level ranges in $\{0.1, 0.2 \ldots, 1\}$.

Table 7
Ablation study on four datasets. Bold texts indicate the best results.

Methods	Handwritten				Cifar10			
	ACC	NMI	PUR	Time(s)	ACC	NMI	PUR	Time(s)
w/o anchor	93.10	87.14	93.10	0.558	98.24	96.91	99.24	1575.501
w/o SDP	70.10	65.24	71.60	0.162	98.22	97.93	99.22	5.032
w/o fusion	71.64	68.36	76.02	**0.105**	39.26	37.48	**99.98**	**4.445**
SPRMC-LS	**95.85**	**91.30**	**95.85**	0.293	**98.91**	**97.01**	**99.98**	19.089
Methods	100Leaves				ALOI-100			
	ACC	NMI	PUR	Time(s)	ACC	NMI	PUR	Time(s)
w/o anchor	77.19	89.92	83.31	0.198	62.24	72.23	66.48	12.937
w/o SDP	11.01	42.10	94.19	0.065	16.41	39.84	18.72	3.523
w/o fusion	48.94	72.12	61.73	**0.014**	41.34	61.55	47.78	**4.614**
SPRMC-LS	**87.75**	**94.84**	**92.31**	0.102	**66.26**	**72.89**	**71.13**	6.784

4.7. Ablation study

We conduct an ablation study on four datasets to demonstrate the effects of the anchor strategy, the SDP-inspired relaxation approach, and the binary fusion function. The variants without these components are denoted as w/o anchor, w/o SDP and w/o fusion, respectively. For w/o fusion, we simply average the clustering results from each view to obtain the final result. The experimental results are shown in Table 7. It shows that, compared with SPRMC-LS, w/o anchor has similar clustering performance but takes longer. This weakness is particularly evident when evaluated on the large-scale dataset Cifar10, which proves the validity of the anchor strategy. Moreover, across most datasets, the performance enhancement from w/o SDP to SPRMC-LS is significantly greater than from w/o fusion to SPRMC-LS. This phenomenon demonstrates the effectiveness of our SDP-inspired relaxation approach and binary fusion function, while the latter still holds significant potential for further improvement.

4.8. Robustness analysis

We analyze the robustness of the proposed algorithm by conducting experiments with noises on the Handwritten dataset. Specifically, we add standard Gaussian noises with a multipliing factor α in even-valued views while the remaining views are unchanged. Fig. 4 illustrates the experimental results with α ranges in $\{0, 0.1, 0.2, \ldots, 1\}$ and the red curve denotes the proposed algorithm. All experiments were repeated 20 times to mitigate the influence of randomness. It can be observed that the proposed method superior to all the comparison methods. As seen, the performance of TBGL and EOMSC decreases rapidly as the noise increases, while slight variation in other methods. Unlike the curve of MVCC, MVSC-HFD, and BMVC, which fluctuates significantly, the LMVSC and GMC have relatively stable performance, but they are all inferior to the proposed method. Moreover, the robustness of SPRMC-LS

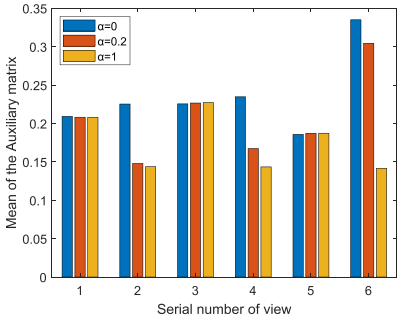


Fig. 5. Mean of the indicator matrix $\mathbf{P}^{(v)}$ of proposed SPRMC-LS on Handwritten with different noise level.

can be theoretically explained by the SDP relaxation property. As shown in Fig. 5, the mean of the auxiliary matrix $\mathbf{P}^{(v)}$ is significantly decreasing with noise level in even-numbered views, while basically unchanged in odd-numbered views. This behavior is consistent with the theoretical guarantee that the SDP-based clustering can identify outliers as zero rows in auxiliary matrix $\mathbf{P}^{(v)}$. Therefore, the experimental results validate the theoretical robustness of our method.

5. Conclusion

In this paper, we proposed SPRMC-LS, a novel iteration-free MVC method that achieves both high efficiency and strong robustness. The key innovations of SPRMC-LS are threefold: (1) It is a novel, large-scale MVC method that extends a provably robust SDP-inspired continuous relaxation into an anchor-based MVC framework to effectively handle outliers. (2) It achieves linear time and memory complexity with a closed-form, iteration-free solution, which is a breakthrough

compared to existing SDP-based clustering and iterative MVC methods. (3) It introduces a novel optimization formulation that transforms the discrete MVC problem into a convex one, naturally embedding outlier detection. The extensive experiments conducted on multiple datasets demonstrate the efficiency and robustness of the proposed SPRMC-LS.

CRediT authorship contribution statement

Minghao Li: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation; **Huiying Xu:** Validation, Supervision, Investigation; **Ziying Wang:** Resources, Methodology, Investigation, Data curation, Conceptualization; **Hechuan Lin:** Writing – review & editing, Resources, Data curation, Conceptualization; **Xiaolei Zhang:** Writing – review & editing, Validation, Conceptualization; **Xinzhong Zhu:** Validation, Supervision, Investigation.

Data availability

Data will be made available on request.

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 document.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (62376252); the Zhejiang Province Province-Land Synergy Program (2025SDXT004-3) and the Zhejiang Province Leading Geese Plan (2025C02025, 2025C01056).

References

[1] C. Xu, D. Tao, C. Xu, A survey on multi-view learning, (2013). arXiv preprint arXiv:1304.5634
[2] J.G. Saxe, The blind men and the elephant, Whittlesey House New York, 1963.
[3] M. Qian, C. Zhai, Unsupervised feature selection for multi-view clustering on text-image web news data, in: Proceedings of the 23Rd ACM International Conference on Conference on Information and Knowledge Management, pp. 1963–1966.
[4] C. Hou, F. Nie, H. Tao, D. Yi, Multi-view unsupervised feature selection with adaptive similarity and view weight, IEEE Trans. Knowl. Data Eng. 29 (2017) 1998–2011.
[5] C.-D. Wang, J.-H. Lai, S.Y. Philip, Multi-view clustering based on belief propagation, IEEE Trans. Knowl. Data Eng. 28 (2015) 1007–1021.
[6] X. Cai, F. Nie, H. Huang, Multi-view k-means clustering on big data, in: Twenty-Third International Joint Conference on Artificial Intelligence.
[7] X. Cao, C. Zhang, H. Fu, S. Liu, H. Zhang, Diversity-induced multi-view subspace clustering, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 586–594.
[8] F. Nie, J. Li, X. Li, et al., Self-weighted multiview clustering with multiple graphs, in: IJCAI, pp. 2564–2570.
[9] G.-Y. Zhang, Y.-R. Zhou, X.-Y. He, C.-D. Wang, D. Huang, One-step kernel multi-view subspace clustering, Knowl. Based Syst. 189 (2020) 105126.
[10] A. Dong, Z. Wu, H. Zhang, Multi-view subspace clustering based on adaptive search, Knowl. Based Syst. 289 (2024) 111553.
[11] M.-S. Chen, L. Huang, C.-D. Wang, D. Huang, Multi-view clustering in latent embedding space, in: Proceedings of the AAAI Conference on Artificial Intelligence, 34, pp. 3513–3520.
[12] P. Zhu, X. Yao, Y. Wang, B. Hui, D. Du, Q. Hu, Multiview deep subspace clustering networks, IEEE Trans. Cybern. (2024).
[13] A. Kumar, H. Daumé, A co-training approach for multi-view spectral clustering, in: Proceedings of the 28th International Conference on Machine Learning (ICML-11), pp. 393–400.
[14] W. Cai, H. Zhou, L. Xu, A multi-view co-training clustering algorithm based on global and local structure preserving, IEEE Access 9 (2021) 29293–29302.
[15] D. Zhao, J. Bian, H. Yin, Y. Huang, Y. Qin, Multi-view clustering via latent consistency multi-graph fusion, Knowl. Based Syst. (2024) 112106.
[16] F. Nie, G. Cai, X. Li, Multi-view clustering and semi-supervised classification with adaptive neighbours, in: Proceedings of the AAAI Conference on Artificial Intelligence, 31.
[17] W. Liang, X. Liu, S. Zhou, J. Liu, S. Wang, E. Zhu, Robust graph-based multi-view clustering, in: Proceedings of the AAAI Conference on Artificial Intelligence, 36, pp. 7462–7469.
[18] Y. Li, F. Nie, H. Huang, J. Huang, Large-scale multi-view spectral clustering via bipartite graph, in: Proceedings of the AAAI Conference on Artificial Intelligence, 29.
[19] L. Li, H. He, Bipartite graph based multi-view clustering, IEEE Trans. Knowl. Data Eng. 34 (2020) 3111–3125.
[20] M. Sun, P. Zhang, S. Wang, S. Zhou, W. Tu, X. Liu, E. Zhu, C. Wang, Scalable multi-view subspace clustering with unified anchors, in: Proceedings of the 29th ACM International Conference on Multimedia, pp. 3528–3536.
[21] B. Yan, P. Sarkar, On robustness of kernel clustering, Adv. Neural Inf. Process. Syst. 29 (2016).
[22] S. Kushagra, Y. Yu, S. Ben-David, Provably noise-robust, regularised k-means clustering, (2017). arXiv preprint arXiv:1711.11247
[23] P.R. Srivastava, P. Sarkar, G.A. Hanasusanto, A robust spectral clustering algorithm for sub-gaussian mixture models with outliers, Oper. Res. 71 (2023) 224–244.
[24] J. Peng, Y. Wei, Approximating k-means-type clustering via semidefinite programming, SIAM J. Optim. 18 (2007) 186–205.
[25] H. Wang, F. Nie, H. Huang, F. Makedon, Fast nonnegative matrix tri-factorization for large-scale data co-clustering, Citeseer.
[26] F. Nie, S. Shi, X. Li, Auto-weighted multi-view co-clustering via fast matrix factorization, Pattern Recognit. 102 (2020) 107207.
[27] N. Rujeerapaiboon, K. Schindler, D. Kuhn, W. Wiesemann, Size matters: cardinality-constrained clustering and outlier detection via conic optimization, SIAM J. Optim. 29 (2019) 1211–1239.
[28] D.G. Mixon, S. Villar, R. Ward, Clustering subgaussian mixtures by semidefinite programming, Inf. Inf. A J. IMA 6 (2017) 389–415.
[29] L. Xu, K. Crammer, D. Schuurmans, et al., Robust support vector machine training via convex outlier ablation, in: AAAI, 6, pp. 536–542.
[30] M. Zurek, Y. Chen, Gap-free clustering: Sensitivity and robustness of sdp, PMLR.
[31] A.A. Amini, E. Levina, On semidefinite relaxations for the block model, 2018.
[32] T.T. Cai, X. Li, Robust and computationally feasible community detection in the presence of arbitrary outlier nodes, 2015.
[33] S. Wang, X. Liu, S. Liu, J. Jin, W. Tu, X. Zhu, E. Zhu, Align then fusion: generalized large-scale multi-view clustering with anchor matching correspondences, Adv. Neural Inf. Process. Syst. 35 (2022) 5882–5895.
[34] H. Ma, S. Wang, S. Yu, S. Liu, J.-J. Huang, H. Wu, X. Liu, E. Zhu, Automatic and aligned anchor learning strategy for multi-view clustering, in: Proceedings of the 32Nd ACM International Conference on Multimedia, pp. 5045–5054.
[35] S. Liu, Q. Liao, S. Wang, X. Liu, E. Zhu, Robust and consistent anchor graph learning for multi-view clustering, IEEE Trans. Knowl. Data Eng. 36 (2024) 4207–4219.
[36] T. Shi, M. Belkin, B. Yu, Data spectroscopy: eigenspaces of convolution operators and clustering, The Annals of Statistics (2009) 3960–3984.
[37] S. Wang, X. Liu, X. Zhu, P. Zhang, Y. Zhang, F. Gao, E. Zhu, Fast parameter-free multi-view subspace clustering with consensus anchor guidance, IEEE Trans. Image Process. 31 (2021) 556–568.
[38] S. Zhou, X. Liu, J. Liu, X. Guo, Y. Zhao, E. Zhu, Y. Zhai, J. Yin, W. Gao, Multi-view spectral clustering with optimal neighborhood laplacian matrix, in: Proceedings of the AAAI Conference on Artificial Intelligence, 34, pp. 6965–6972.
[39] M. van Breukelen, R.P. Duin, D.M. Tax, J.D. Hartog, Handwritten digit recognition by combined classifiers, Kybernetika 34 (1998) 381–386.
[40] O. Madani, M. Georg, D.A. Ross, On using nearly-independent feature families for high precision and confidence, PMLR.
[41] Z. Zhang, L. Liu, F. Shen, H.T. Shen, L. Shao, Binary multi-view clustering, IEEE Trans. Pattern Anal. Mach. Intell. 41 (2018) 1774–1782.
[42] H. Wang, Y. Yang, T. Li, Multi-view clustering via concept factorization with local manifold regularization, IEEE
[43] H. Wang, Y. Yang, B. Liu, Gmc: graph-based multi-view clustering, IEEE Trans. Knowl. Data Eng. 32 (2019) 1116–1129.
[44] W. Xia, Q. Gao, Q. Wang, X. Gao, C. Ding, D. Tao, Tensorized bipartite graph learning for multi-view clustering, IEEE Trans. Pattern Anal. Mach. Intell. 45 (2022) 5187–5202.
[45] Z. Kang, W. Zhou, Z. Zhao, J. Shao, M. Han, Z. Xu, Large-scale multi-view subspace clustering in linear time, in: Proceedings of the AAAI Conference on Artificial Intelligence, 34, pp. 4412–4419.
[46] S. Liu, S. Wang, P. Zhang, K. Xu, X. Liu, C. Zhang, F. Gao, Efficient one-pass multi-view subspace clustering with consensus anchors, in: Proceedings of the AAAI Conference on Artificial Intelligence, 36, pp. 7576–7584.
[47] Q. Ou, S. Wang, P. Zhang, S. Zhou, E. Zhu, Anchor-based multi-view subspace clustering with hierarchical feature descent, Inf. Fusion 106 (2024) 102225.
[48] Z. Chen, X.-J. Wu, T. Xu, J. Kittler, Fast self-guided multi-view subspace clustering, IEEE Trans. Image Process. 32 (2023) 6514–6525.
[49] Q. Qu, X. Wan, W. Liang, J. Liu, Y. Feng, H. Xu, X. Liu, E. Zhu, A lightweight anchor-based incremental framework for multi-view clustering, in: Proceedings of the 32nd ACM International Conference on Multimedia, pp. 8652–8661.