JetBGC: Joint Robust Embedding and Structural Fusion Bipartite Graph Clustering

Liang Li, *Graduate Student Member, IEEE*, Yuangang Pan, Junpu Zhang, Pei Zhang, Jie Liu, Xinwang Liu, *Senior Member, IEEE*, Kenli Li, *Senior Member, IEEE*, Ivor W. Tsang, *Fellow, IEEE*, and Keqin Li, *Fellow, IEEE*

***Abstract*—Bipartite graph clustering (BGC) has emerged as a fast-growing research in the clustering community. Despite BGC has achieved promising scalability, most variants still suffer from the following concerns: a) Susceptibility to noisy features. They construct bipartite graphs in the raw feature space, inducing poor robustness to noisy features. b) Inflexible anchor selection strategies. They usually select anchors through heuristic sampling or constrained learning methods, degrading flexibility. c) Partial structure mining. Existing methods are mainly built upon Linear Reconstruction Paradigm (LRP) from subspace clustering or Locally Linear Paradigm (LLP) from manifold learning, which partially exploit linear or locally linear structures, lacking a unified perspective to integrate global complementary structures. To this end, we propose a novel model, termed Joint Robust Embedding and Structural Fusion Bipartite Graph Clustering (JetBGC), which focuses on three aspects, namely robustness, flexibility, and complementarity. Concretely, we first introduce a robust embedding learning module to extract latent representation that can reduce the impact of noisy features. Then, we optimize anchors via a constraint-free strategy that can flexibly capture data distribution. Furthermore, we revisit the consistency and specificity of LRP and LLP, and design a new unified structural fusion strategy to integrate both linear and locally linear structures from a global perspective. Therefore, JetBGC unifies robust representation learning, flexible anchor optimization, and structural bipartite graph fusion in a framework. Extensive experiments on synthetic and real-world datasets validate our effectiveness against existing baselines.**

***Index Terms*—Latent embedding learning, structural fusion, bipartite graph learning, multi-view clustering.**

Received 15 October 2023; revised 20 April 2025; accepted 22 June 2025. Date of publication 27 June 2025; date of current version 24 July 2025. This work was supported in part by National Natural Science Foundation of China under Grant 62325604, Grant 62441618, and Grant 62276271, in part by National Key Research and Development Program of China under Grant 2021YFB0300101, and in part by China Scholarship Council under Grant 202306110001. Recommended for acceptance by M. Zhang. *(Corresponding author: Xinwang Liu.)*

Liang Li, Junpu Zhang, Pei Zhang, and Xinwang Liu are with the College of Computer Science and Technology, National University of Defense Technology, Changsha 410073, China (e-mail: liliang1037@gmail.com; zhangjunpu@nudt.edu.cn; jeaninezpp@gmail.com; xinwangliu@nudt.edu.cn).

Yuangang Pan and Ivor W. Tsang are with the Center for Frontier AI Research, Agency for Science, Technology and Research (A∗STAR), Singapore 138632 (e-mail: yuangang.pan@gmail.com; ivor.tsang@gmail.com).

Jie Liu is with the Science and Technology and Parallel and Distributed Processing Laboratory, National University of Defense Technology, Changsha 410073, China, and also with the Laboratory of Software Engineering for Complex Systems, National University of Defense Technology, Changsha 410073, China (e-mail: liujie@nudt.edu.cn).

Kenli Li is with the College of Computer Science and Electronic Engineering, Hunan University, Changsha 410073, China, and also with the Supercomputing and Cloud Computing Institute, Hunan University, Changsha 410073, China (e-mail: lkl@hnu.edu.cn).

Keqin Li is with the Department of Computer Science, State University of New York, New Paltz, NY 12561 USA (e-mail: lik@newpaltz.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TKDE.2025.3583718, provided by the authors.

The code is provided at https://github.com/liliangnudt/JetBGC.

Digital Object Identifier 10.1109/TKDE.2025.3583718

I. Introduction

THE dramatic growth of data highlights the necessity to develop unsupervised or self-supervised learning to reduce reliance on costly human annotations [1]. As a fundamental task in unsupervised learning, clustering plays a critical role in uncovering the inherent grouping structures [2], [3], [4], [5], [6]. Owing to the flexibility of graph in representing complex relationships [7], [8], [9], graph clustering has become an active field of research [10], [11], [12]. To further exploit complementary information from diverse sources or perspectives, multi-view graph clustering (MVGC) [13], [14], as a popular subfield of multi-view clustering (MVC) [15], [16], has been widely applied in data mining [17], knowledge graph [18], and computer vision [19].

However, existing MVGC methods require to compute pairwise similarities to build full graphs, which incurs quadratic space and cubic time complexity w.r.t. instance number [20], [21]. This limits their scalability when dealing with large-scale data. In particular, computing and storing a similarity matrix for over 100,000 nodes often results in out-of-memory errors or unacceptable running time.

To improve scalability, multi-view bipartite graph clustering (MVBGC) [22], [23] instead to merely build the memberships between a few representative anchors/landmarks and all instances, achieving linear complexity. Fig. 1 plots two popular paradigms for bipartite graph construction: subspace clustering based *Linear Reconstruction Paradigm (LRP)* and manifold learning based *Locally Linear Paradigm (LLP)*. Built upon LRP, Kang et al. [24] selected anchors via k-means and concatenated view-specific bipartite graphs to fuse multi-view structures. Sun et al. [25] incorporated anchor into optimization, avoiding sampling anchors. Wang et al. [26] further extended a parameter-free version. Li et al. [27] designed a feature self-attention mechanism to reduce noisy features. Built upon LLP, Wang et al. [28] designed a semi-unsupervised single-view model with constrained Laplacian rank, anchors are selected via k-means. Li et al. [29] further extended [28] into multi-view clustering scenarios, and learned a neighbor bipartite graph. Nie et al. [30] and Chen et al. [31] introduced feature re-weighting and selection methods to preserve useful features. Li et al. [32] developed a multi-view bipartite graph fusion framework, which introduced a heuristic anchor selection method and connectivity constraint, enforcing the bipartite graph holds clear component structures. Lu et al. [33] further designed a structure-diversity fusion to refine the graph.

Despite achieving favorable performance, existing MVBGC models still encounter the following limitations: a) Susceptibility to noisy features: most methods estimate anchor-instance correlations in the raw feature space, disregarding the impact of noisy features. b) Inflexible anchor strategy: anchors are typically pre-selected via k-means or random sampling and remain fixed thereafter. Although a learnable anchor strategy has recently been proposed, it requires additional constraints, limiting flexibility. c) Partial structure mining: most variants are derived from LRP or LLP paradigms, capturing only linear or locally linear structures, lacking a unified perspective to explore global complementarity.

To this end, we consider designing a novel MVBGC model that aims to enhance bipartite graph clustering from three aspects: robustness, flexibility, and complementarity. Concretely, a) To improve robustness, we introduce a robust feature extraction module to learn robust embeddings from the input raw feature space. b) To maintain flexibility, we propose to flexibly acquire anchors through constraint-free optimization, unlike the existing inflexible k-means, unstable random sampling, or constrained learning methods. c) To achieve complementarity, we generalize LRP and LLP into the latent space, and subtly integrate them into a unified form to fuse linear and locally linear structures.

We summarize our contributions as follows:

1) We develop a novel MVBGC model, named JetBGC, which integrates robust embedding learning, constraint-free anchor optimization, and structural bipartite graph fusion into a unified framework.
2) We revisit the consistency and specificity of two popular BGC paradigms, and design a new structural bipartite graph fusion strategy that integrates linear and locally linear structures from a global perspective. Furthermore, we establish a theoretical connection between our model and the existing LLP paradigm by showing that the proposed structural fusion strategy is a generalization of LLP under a newly defined η-norm.
3) We design an ADMM solver with linear algorithm complexity w.r.t. instance number. Extensive experiments on synthetic and real-world datasets verify the superiority. Detailed ablation analysis validate the effectiveness of the robust embedding learning, flexible anchor selection, structural fusion modules.

II. Related Research

Table I lists the notations used throughout this work.

A. Non-Negative Matrix Factorization (NMF)

NMF [34] is a popular matrix decomposition method, widely used in bioinformatics, image annotation, and social networks. Given the raw data $\mathbf{X} \in \mathbb{R}^{\widetilde{d} \times n}$, NMF factorizes it into two non-negative parts. Typically, the standard Frobenius-norm (F-norm) form is as follows:

$$\min_{\mathbf{U},\mathbf{V}} \quad \|\mathbf{X} - \mathbf{U}\mathbf{V}\|_{\mathrm{F}}^2, \quad \text{s.t. } \mathbf{U} \geq 0, \mathbf{V} \geq 0, \tag{1}$$

where $\mathbf{U}$ is base matrix and $\mathbf{V}$ is coefficient matrix.

TABLE I
Notations

Notation	Explanation
k, n, v	Number of clusters, samples, and views
d_p	Feature dimension for the p-th view
d	Latent feature dimension
m	Number of anchors
μ, σ	Penalty parameter, scaling factor
ϵ	Number of neighbors in initialization
$\boldsymbol{\gamma} \in \mathbb{R}^{v \times 1}$	View weights
$\mathbf{X}_p \in \mathbb{R}^{d_p \times n}$	Input data for the p-th view
$\mathbf{U}_p \in \mathbb{R}^{d_p \times d}$	Base matrix for input data $\mathbf{X}_p$
$\mathbf{V} \in \mathbb{R}^{d \times n}$	Consistent latent representation
$\mathbf{A}' \in \mathbb{R}^{d' \times m}$	Anchor matrix in raw feature space
$\mathbf{A} \in \mathbb{R}^{d \times m}$	Anchor matrix in latent space
$\mathbf{Z} \in \mathbb{R}^{n \times m}$	Bipartite graph matrix
$\mathbf{E}_p \in \mathbb{R}^{d_p \times n}$	Auxiliary variables in ADMM
$\mathbf{\Lambda}_p \in \mathbb{R}^{d_p \times n}$	ALM multipliers

Following this idea, many variants are proposed. Ding et al. [35] built the relationship between NMF and k-means clustering. Cai et al. [36] introduced spectral embedding and designed a graph regularization NMF. Kuang et al. [37] proposed symmetric NMF that builds connection of NMF and spectral clustering. Ding et al. [38] developed $\mathbf{V}$-orthogonal NMF to improve diversity as follows

$$\min_{\mathbf{U},\mathbf{V}} \quad \|\mathbf{X} - \mathbf{U}\mathbf{V}\|_{\mathrm{F}}^2, \text{s.t. } \mathbf{U} \geq 0, \mathbf{V} \geq 0, \mathbf{V}\mathbf{V}^\top = \mathbf{I}. \tag{2}$$

Instead of standard F-norm, Kong et al. [39] proposed a robust version [40] in which residuals are measured by $\ell_{2,1}$-norm, i.e.,

$$\min_{\mathbf{U},\mathbf{V}} \|\mathbf{X} - \mathbf{U}\mathbf{V}\|_{2,1}, \text{s.t. } \mathbf{U} \geq 0, \mathbf{V} \geq 0. \tag{3}$$

Based on $\ell_{2,1}$-norm, Huang et al. [41] further designed a robust graph regularization NMF. Li et al. [42] incorporated linear discriminant analysis into NMF. For more details, please refer to [43].

B. Bipartite Graph Construction

According to the construction manner, there are two representative strategies.

Linear Reconstruction Paradigm (LRP): LRP originates from subspace clustering [44], which assumes that data can be linearly reconstructed by anchors in the same subspace. Formally, LRP is defined as:

$$\begin{aligned} \min_{\mathbf{Z}} \quad & \left\|\mathbf{X} - \mathbf{A}\mathbf{Z}^\top\right\|_{\mathrm{F}}^2 + \lambda \|\mathbf{Z}\|_{\mathrm{F}}^2, \\ \text{s.t.} \quad & \mathbf{Z}\mathbf{1} = \mathbf{1}, \mathbf{Z} \geq 0, \end{aligned} \tag{4}$$

where $\mathbf{A}$ is the anchor matrix, $\mathbf{Z}$ is the bipartite graph matrix, and λ is a hyper-parameter to balance the contribution of the regularizer that avoids the trivial solution.

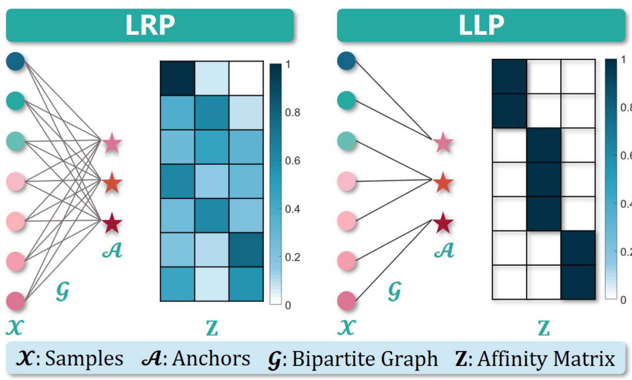


Fig. 1. Sketches of existing LRP and LLP paradigms.

Since LRP paradigm models each instance as a combination of all anchors linearly through probability/similarity, the resulting bipartite graph often shows a fuzzy representation, as shown in Fig. 1 (left).

Locally Linear Paradigm (LLP): LLP builds upon the manifold learning [45] that supposes that the original high-dimensional data actually reside on the low-dimensional manifold. This setting enables it to preserve locally linear structures. Formally, LLP is expressed by

$$\min_{\mathbf{Z}} \sum_{i=1}^{n}\sum_{j=1}^{m}\left(\|\mathbf{x}_i - \mathbf{a}_j\|_2^2 z_{ij} + \lambda z_{ij}^2\right), \quad \text{s.t.} \quad \mathbf{Z1} = \mathbf{1}, \mathbf{Z} \geq 0. \tag{5}$$

Typically, LLP measures similarity for anchor-instance pairs through Euclidean distance, where longer distances correspond to lower probabilities of being neighbors. As a result, the constructed bipartite graph is often sparse, as shown in Fig. 1 (right).

By reviewing existing BGC research, we find that most variants are built upon either LRP [24], [25], [26] or LLP [29], [30], [32], which focus on modeling linear or locally linear structures while failing to exploit global complementary structures, resulting in degraded discrimination of bipartite graphs.

III. Methodology

A. Probability Perspective for Bipartite Graph

To naturally motivate the subsequent model design, we begin by introducing a probabilistic perspective of bipartite graph [46] that how to recover instance-instance membership $w_{ij} = p(\mathbf{x}_i \mapsto \mathbf{x}_j)$ with instance-anchor membership $z_{ri} = p(\mathbf{x}_i \mapsto \mathbf{a}_r)$ and $z_{rj} = p(\mathbf{a}_r \mapsto \mathbf{x}_j)$.

Specifically, the one-step transition probability from the i-th instance ($\mathbf{x}_i$) to the r-th anchor ($\mathbf{a}_r$) is as follows

$$p^{(1)}(\mathbf{x}_i \mapsto \mathbf{a}_r) = \frac{z_{ri}}{\sum_{r=1}^{m} z_{ri}}, \quad p^{(1)}(\mathbf{a}_r \mapsto \mathbf{x}_j) = \frac{z_{rj}}{\sum_{j=1}^{n} z_{rj}}, \tag{6}$$

where $\mapsto$ denotes the transitioning.

The relationship between two instances can be viewed as a double-step transition process, and the transition probability from $\mathbf{x}_i$ to $\mathbf{x}_j$ is

$$p^{(2)}(\mathbf{x}_i \mapsto \mathbf{x}_j) = \sum_{r=1}^{m} p^{(1)}(\mathbf{x}_i \mapsto \mathbf{a}_r)\, p^{(1)}(\mathbf{a}_r \mapsto \mathbf{x}_j) = \sum_{r=1}^{m} \frac{z_{ri} z_{rj}}{\sum_{j=1}^{n} z_{rj}}. \tag{7}$$

Therefore, the instance-instance affinity matrix $\mathbf{S} \in \mathbb{R}^{n\times n}$ in conventional graph clustering can be approximated via the construction of bipartite graph $\mathbf{Z} \in \mathbb{R}^{n\times m}$, where $n \gg m$, which significantly reduces computational and memory costs. Anchors thus serve as representative points that capture the underlying structural relationships among instances.

B. Robust Latent Embedding Learning

From the probabilistic perspective of bipartite graph, the reliability of the double-step transition probability $p^{(2)}(\mathbf{x}_i \mapsto \mathbf{x}_j)$ depends on the quality of one-step transition probabilities $p^{(1)}(\mathbf{x}_i \mapsto \mathbf{a}_r)$. However, the high-dimensional raw features $\mathbf{X} \in \mathbb{R}^{d'\times n}$ may contain redundant or noisy features, which induce unreliable one-step transition, and further degrades double-step transition. To enhance robustness against noisy and redundant features, we propose mapping the raw features into a latent embedding space, built upon robust NMF [40] in (3). For the multi-view setting, we jointly optimize all views $\{\mathbf{X}_p \in \mathbb{R}^{d_p\times n}\}_{p=1}^{v}$ and fuse them into a unified embedding $\mathbf{V} \in \mathbb{R}^{d\times n}$. Our *Robust Latent Embedding Learning (RLEL)* is formulated as

$$\min_{\{\mathbf{U}_p\}_{p=1}^{v}, \mathbf{V}, \boldsymbol{\gamma}} \sum_{p=1}^{v} \gamma_p^2 \|\mathbf{X}_p - \mathbf{U}_p\mathbf{V}\|_{2,1}, \quad \text{s.t.} \begin{cases} \mathbf{V}\mathbf{V}^\top = \mathbf{I}; \\ \boldsymbol{\gamma}^\top \mathbf{1} = 1,\ \gamma_p \geq 0, \forall p \in \{1, 2, \ldots, v\}, \end{cases} \tag{8}$$

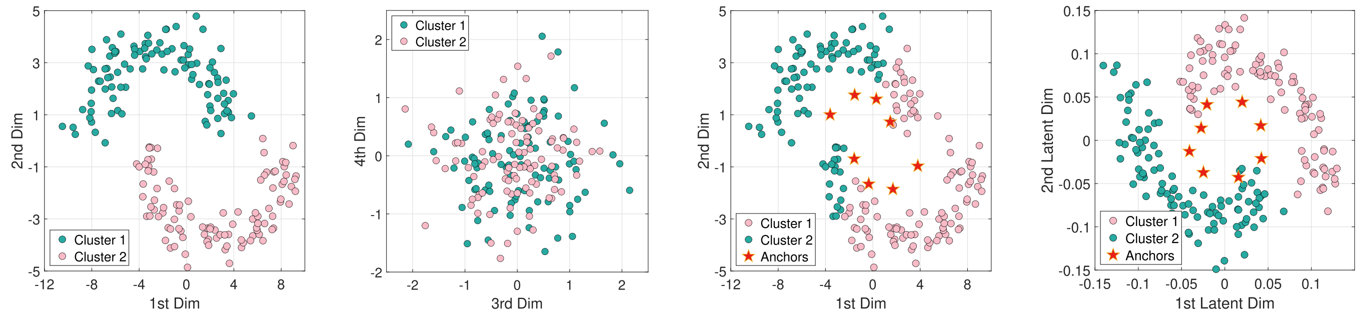
where $\{\mathbf{U}_p \in \mathbb{R}^{d_p\times d}\}_{p=1}^{v}$ are the view-related base matrices, $\boldsymbol{\gamma}$ measures the view importance.

Note that we introduce $\mathbf{V}$-orthogonal constraint to enhance the discrimination of the latent embedding. Moreover, we remove the non-negative constraints on $\mathbf{U}$ and $\mathbf{V}$, enabling it to handle input data with mixed signs, rather than being limited to non-negative data [47]. Empirical evidence supporting this relaxation is provided in the supplementary material (Section 4.1). Theorem 1 further bridges the connection between our RLEL and the relaxed multiple kernel k-means (MKKM) clustering. More details can be found in supplementary material (Section 1).

Theorem 1: A reformulation of our RLEL model under the Frobenius norm, i.e.,

$$\min_{\{\mathbf{U}_p\}_{p=1}^{v}, \mathbf{V}, \boldsymbol{\gamma}} \sum_{p=1}^{v} \gamma_p^2 \|\mathbf{X}_p - \mathbf{U}_p\mathbf{V}\|_{\mathrm{F}}^2, \quad \text{s.t.} \ \mathbf{V}\mathbf{V}^\top = \mathbf{I};\ \boldsymbol{\gamma}^\top \mathbf{1} = 1,\ \gamma_p \geq 0, \forall p \in \{1, 2, \ldots, v\}, \tag{9}$$

is equivalent to MKKM clustering with a linear kernel, under the condition that the latent feature dimension equals the cluster



(a) Input clean features $\mathbf{X}_{\text{clean}}$ (b) Input noisy features $\mathbf{X}_{\text{noisy}}$ (c) SFRF $\mathbf{X}$&$\mathbf{A}'$ (NMI: 21.51%) (d) JetBGC $\mathbf{V}$&$\mathbf{A}$ (NMI: 100%)

Fig. 2. Robustness of the proposed latent embedding module to noisy features, evaluated on a synthetic 4D single-view dataset (SData-1) with two clusters (100 instances each). The input $\mathbf{X} = (\mathbf{X}_{\text{clean}}^\top, \mathbf{X}_{\text{noisy}}^\top)^\top \in \mathbb{R}^{4\times 200}$ contains clean features in the first two dimensions and noisy features in the last two. (a) and (b) visualize the clean and noisy features, respectively. (c) shows the results of SFRF, a JetBGC variant without latent embedding module, where anchors $\mathbf{A}' \in \mathbb{R}^{4\times 8}$ are learned in the raw space. See Section III-H1 for details of SFRF. (d) shows the result of JetBGC, where features are projected into a 2D latent space $\mathbf{V} \in \mathbb{R}^{2\times 200}$, yielding anchors $\mathbf{A} \in \mathbb{R}^{2\times 8}$ that better capture the cluster structure.

number, i.e., $d = k$. Formally, this corresponds to:

$$\max_{\mathbf{H},\boldsymbol{\gamma}} \ \operatorname{Tr}\left(\mathbf{H}^\top \mathbf{K}_{\boldsymbol{\gamma}} \mathbf{H}\right),$$
$$\text{s.t.} \ \ \mathbf{H}^\top\mathbf{H} = \mathbf{I};\ \boldsymbol{\gamma}^\top \mathbf{1} = 1,\ \gamma_p \geq 0, \forall p \in \{1, 2, \ldots, v\}, \quad (10)$$

where $\mathbf{H}$ denotes the consensus kernel partition, $\mathbf{K}_{\boldsymbol{\gamma}} = \sum_{p=1}^{v} \gamma_p^2 \mathbf{K}_p$ represents a linear combination of multiple base kernels via kernel function $\kappa_{\boldsymbol{\gamma}}(\mathbf{x}_i, \mathbf{x}_j) = \psi_{\boldsymbol{\gamma}}(\mathbf{x}_i)^\top \psi_{\boldsymbol{\gamma}}(\mathbf{x}_j) = \sum_{p=1}^{v} \gamma_p^2 \kappa_p(\mathbf{x}_{p(i)}, \mathbf{x}_{p(j)})$, with $\psi(\cdot)$: $\mathbf{x} \in \mathcal{X} \mapsto \mathcal{H}$ denoting the mapping that transforms $\mathbf{x}$ into a kernel Hilbert space.

To demonstrate the robustness of RLEL module, we design a synthetic dataset (SData-1), shown in Fig. 2. For a fair comparison, all experimental settings for the compared SFRF (which excludes the RLEL module) and JetBGC are the same. For SFRF, the presence of noisy features lead to inaccurate decision boundaries and degrade the performance, resulting in a Normalized Mutual Information (NMI) of only 21.51%. In contrast, our strategy demonstrates robustness to noisy features and achieves competitive performance (NMI: 100%).

C. Flexible Anchor Learning

From the probabilistic perspective of bipartite graph, anchors act as intermediate "bridges" that connect the one-step transition probabilities $p^{(1)}(\mathbf{v}_i \mapsto \mathbf{a}_r)$ to the double-step transition probability $p^{(2)}(\mathbf{v}_i \mapsto \mathbf{v}_j)$. Therefore, anchor selection is critical.

Conventional anchor selection strategies, such as random sampling [48] and k-means clustering [49], typically pre-select anchors before optimization, which lacks of flexibility and reduces the reliability of $p^{(1)}(\mathbf{v}_i \mapsto \mathbf{a}_r)$, and consequently undermines $p^{(2)}(\mathbf{v}_i \mapsto \mathbf{v}_j)$. Similar issues arise in constrained anchor learning [26], where predefined constraints limit the adaptability of anchors.

To address these limitations, we adopt a constraint-free anchor learning strategy that removes the reliance on heuristic methods or manual constraints. This enables the anchors to better adapt to the underlying data distribution and more reliable estimation of transition probabilities.

D. Structural Bipartite Graph Fusion

Given that LRP and LLP capture only linear or locally linear structures, this section analyzes their consistency and specificity.

By generalizing LRP into the embedding space (termed LRPE), and further expanding it mathematically, we have:

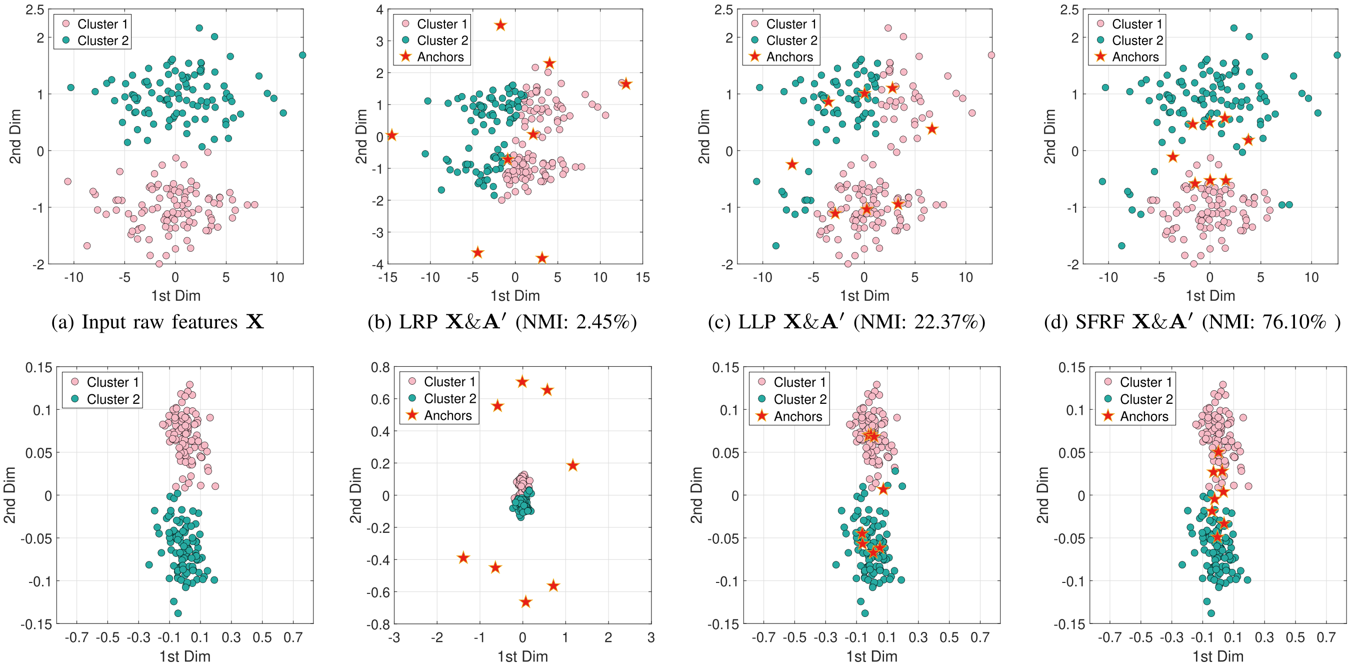
$$\begin{aligned}
&\left\|\mathbf{V} - \mathbf{A}\mathbf{Z}^\top\right\|_{\mathrm{F}}^2 + \lambda \left\|\mathbf{Z}\right\|_{\mathrm{F}}^2 \\
&= \sum_{i=1}^{n} \left\| \mathbf{v}_i - \sum_{j=1}^{m} \mathbf{a}_j z_{ij} \right\|_2^2 + \lambda \sum_{i=1}^{n}\sum_{j=1}^{m} z_{ij}^2, \\
&= \operatorname{Tr}\Bigg(\underbrace{\mathbf{V}\mathbf{V}^\top - 2\mathbf{V}\mathbf{Z}\mathbf{A}^\top}_{\text{Common Part}} + \underbrace{\mathbf{A}\mathbf{Z}^\top\mathbf{Z}\mathbf{A}^\top}_{\text{Linear Specific}} \Bigg) + \underbrace{\lambda \|\mathbf{Z}\|_{\mathrm{F}}^2}_{\text{Regularizer}}.
\end{aligned} \quad (11)$$

By generalizing LLP into the embedding space (termed LLPE), and further expanding it mathematically, we have:

$$\begin{aligned}
&\sum_{i=1}^{n}\sum_{j=1}^{m} \left(\|\mathbf{v}_i - \mathbf{a}_j\|_2^2\, z_{ij} + \lambda z_{ij}^2 \right) = \\
&\operatorname{Tr}\Bigg(\underbrace{\mathbf{V}\mathbf{V}^\top - 2\mathbf{V}\mathbf{Z}\mathbf{A}^\top}_{\text{Common Part}} + \underbrace{\mathbf{A}\mathbf{D}_m\mathbf{A}^\top}_{\text{Locally Linear Specific}} \Bigg) + \underbrace{\lambda \|\mathbf{Z}\|_{\mathrm{F}}^2}_{\text{Regularizer}},
\end{aligned} \quad (12)$$

where $\mathbf{D}_m = diag(\mathbf{Z}^\top \mathbf{1}) \in \mathbb{R}^{m\times m}$.

By reviewing (11) and (12), we find that they share a common part and a regularization term, while each retains a specific part. This motivates us to integrate them into a unified framework that captures global complementary structures. For simplicity, we combine their specific terms with equal weighting, along with the shared components, a unified *Structural Fusion on Latent*



(a) Input raw features $\mathbf{X}$ (b) LRP $\mathbf{X}\&\mathbf{A}'$ (NMI: 2.45%) (c) LLP $\mathbf{X}\&\mathbf{A}'$ (NMI: 22.37%) (d) SFRF $\mathbf{X}\&\mathbf{A}'$ (NMI: 76.10%)

(e) Latent embedding $\mathbf{V}$ (f) LRPE $\mathbf{V}\&\mathbf{A}$ (NMI: 68.38%) (g) LLPE $\mathbf{V}\&\mathbf{A}$ (NMI: 87.82%) (h) Proposed $\mathbf{V}\&\mathbf{A}$ (NMI: 100%)

Fig. 3. Comparison of two BGC paradigms on a synthetic 2D single-view dataset (SData-2). The input $\mathbf{X} \in \mathbb{R}^{2\times 200}$ contains two clusters, each with 100 instances. (a) Visualizes the input features. (b)-(d) show the clustering results of LRP, LLP, and SFRF in the raw space. Section III-H1 for details of SFRF. (e) Displays the learnt latent embedding, and (f)-(g) present the results of LRPE, LLPE, and the proposed JetBGC in the latent space.

Embedding (SFLE) can be formulated as,

$$\min_{\mathbf{A},\mathbf{Z}} \operatorname{Tr}\left(\underbrace{-2\mathbf{V}\mathbf{Z}\mathbf{A}^\top}_{\text{Common Part}} + \underbrace{\mathbf{A}\mathbf{Z}^\top\mathbf{Z}\mathbf{A}^\top}_{\text{Linear Specific}} + \underbrace{\mathbf{A}\mathbf{D}_m\mathbf{A}^\top}_{\text{Locally Linear Specific}}\right) + \underbrace{\lambda\|\mathbf{Z}\|_{\mathrm{F}}^2}_{\text{Regularizer}}, \text{ s.t. } \mathbf{Z}\mathbf{1} = \mathbf{1},\ \mathbf{Z} \geq 0. \tag{13}$$

Note that the constraint $\mathbf{V}\mathbf{V}^\top = \mathbf{I}$, thereby $\operatorname{Tr}(\mathbf{V}\mathbf{V}^\top)$ is a constant term and can be omitted from the optimization. Besides, we treat LRPE and LLPE as equally important, and add their specific parts with equal contribution.

Furthermore, we derive an alternative form of SFLE that aligns closely with LLPE, demonstrating that SFLE is a generalization of LLPE and establishing a connection between the proposed module and the prior methods.

Concretely, by introducing the constant term $\operatorname{Tr}(\mathbf{V}\mathbf{V}^\top)$, (13) can be reformulated as

$$\min_{\mathbf{Z},\mathbf{A}} \sum_{i=1}^{n}\sum_{j=1}^{m} \|\mathbf{v}_i - \mathbf{a}_j\|_2^2\, z_{ij} + \lambda \sum_{i=1}^{n} \|\mathbf{z}_i\|_\eta^2, \quad \text{s.t. } \mathbf{Z}\mathbf{1} = \mathbf{1}, \mathbf{Z} \geq 0, \tag{14}$$

where $\|\cdot\|_\eta$ denotes a newly defined vector norm in Theorem 2, termd η-norm, which is induced by a positive definite (PD) matrix $\mathbf{I} + \eta\mathbf{A}^\top\mathbf{A}$ with $\eta > 0$, namely

$$\|\mathbf{z}_i\|_\eta^2 = \langle \mathbf{z}_i, \mathbf{z}_i\rangle_\eta = \mathbf{z}_i\left(\mathbf{I} + \eta\mathbf{A}^\top\mathbf{A}\right)\mathbf{z}_i^\top. \tag{15}$$

Theorem 2: η-norm is a valid vector norm induced by a positive definite matrix $\mathbf{I} + \eta\mathbf{A}^\top\mathbf{A}$ with $\eta > 0$.

Detailed proof is provided in supplementary material (Section 2). In experiments, we set $\eta = \frac{1}{\lambda}$ to ensure that LRPE and LLPE contribute equally to the overall objective.

Remark 1: The introduced η-norm is a generalization of F-norm. In (15), we enforce $\eta > 0$ to hold the positive definiteness of $\mathbf{I} + \eta\mathbf{A}^\top\mathbf{A}$. Specifically, when $\eta = 0$, η-norm is simplified to F-norm. For LLP or LLPE variants, a F-norm based regularizer is typically incorporated to avoid trivial solutions. Beyond this functionality, η-norm also contributes to capturing linear structures by combining a LRPE-specific term. Moreover, η-norm is compatible with existing LLP or LLPE variants. Incorporating it as a regularizer enables these variants to extract both linear and locally linear structures.

To demonstrate the flexibility and complementarity of the proposed SFLE module, we design another synthetic dataset (SData-2), as shown in Fig. 3. In all cases, the anchors are learned via constraint-free optimization. The results show that: (a) For LRP, anchors are scattered separately, reflecting its linear reconstruction property. (b) For LLP, anchors are primarily located within clusters, capturing locally linear structures. (c) Fusing LRP and LLP combines their complementarity, resulting in improved performance. (d) LRPE and LLPE inherit the structural properties of their respective backbones, while benefiting from the robust latent embedding module. (e) JetBGC achieves the highest performance, indicating the importance of learning latent embedding and modeling both linear and locally linear structures.

E. The Proposed JetBGC Model

By integrating the above RLEL and SFLE modules, the proposed JetBGC model is as follows,

$$\min_{\substack{\{\mathbf{U}_p\}_{p=1}^{v},\mathbf{V},\\ \mathbf{A},\mathbf{Z},\boldsymbol{\gamma}}} \underbrace{\sum_{p=1}^{v}\gamma_p^2\|\mathbf{X}_p-\mathbf{U}_p\mathbf{V}\|_{2,1}}_{\text{Robust Embedding Learning}}$$
$$+\underbrace{\operatorname{Tr}\left(-2\mathbf{VZA}^\top+\eta\mathbf{AZ}^\top\mathbf{ZA}^\top+\mathbf{AD}_m\mathbf{A}^\top\right)+\lambda\|\mathbf{Z}\|_{\mathrm{F}}^2}_{\text{Structural Bipartite Graph Fusion}},$$
$$\text{s.t.}\begin{cases}\mathbf{Z1}=\mathbf{1},\,\mathbf{Z}\geq 0;\,\mathbf{VV}^\top=\mathbf{I};\\ \boldsymbol{\gamma}^\top\mathbf{1}=1,\,\gamma_p\geq 0,\,\forall p\in\{1,2,\ldots,v\}.\end{cases}\quad(16)$$

In summary, JetBGC integrates robust latent representation learning, flexible anchor optimization, and structural bipartite graph fusion, providing a unified solution for robustness, flexibility, and complementarity.

F. Optimization

This section designs an ADMM solver. To separate constraints and simplify our model, we introduce v auxiliary variables $\{\mathbf{E}_p=\mathbf{X}_p-\mathbf{U}_p\mathbf{V}\}_{p=1}^{v}$. The Augmented Lagrange Multiplier (ALM) function of (16) is expressed by

$$\min_{\substack{\{\mathbf{U}_p,\mathbf{E}_p,\boldsymbol{\Lambda}_p\}_{p=1}^{v}\\ \mathbf{V},\mathbf{Z},\mathbf{A},\boldsymbol{\gamma}}} \sum_{p=1}^{v}\gamma_p^2\,\|\mathbf{E}_p\|_{2,1}+$$
$$\operatorname{Tr}\left(-2\mathbf{VZA}^\top+\mathbf{AZ}^\top\mathbf{ZA}^\top+\mathbf{AD}_m\mathbf{A}^\top\right)+\lambda\|\mathbf{Z}\|_{\mathrm{F}}^2$$
$$+\sum_{p=1}^{v}\left(\langle\boldsymbol{\Lambda}_p,\mathbf{X}_p-\mathbf{U}_p\mathbf{V}-\mathbf{E}_p\rangle+\frac{\mu}{2}\|\mathbf{X}_p-\mathbf{U}_p\mathbf{V}-\mathbf{E}_p\|_{\mathrm{F}}^2\right)$$
$$\text{s.t.}\begin{cases}\mathbf{Z1}=\mathbf{1},\,\mathbf{Z}\geq 0;\,\mathbf{VV}^\top=\mathbf{I};\\ \boldsymbol{\gamma}^\top\mathbf{1}=1,\,\gamma_p\geq 0,\,\forall p\in\{1,2,\ldots,v\},\end{cases}\quad(17)$$

where $\boldsymbol{\Lambda}_p$ denotes the ALM multiplier to penalize the gap between the original target and the auxiliary variables, and μ is the ALM parameter. Equation (17) can be solved by block-coordinate descent method.

1) Update $\mathbf{U}_p$*:* Each $\mathbf{U}_p$ is independently updated by

$$\min_{\mathbf{U}_p}\left\|\mathbf{X}_p-\mathbf{U}_p\mathbf{V}-\mathbf{E}_p+\frac{1}{\mu}\boldsymbol{\Lambda}_p\right\|_{\mathrm{F}}^2.\quad(18)$$

Since $\mathbf{VV}^\top=\mathbf{I}$, the solution of $\mathbf{U}_p$ is

$$\mathbf{U}_p=\left(\mathbf{X}_p-\mathbf{E}_p+\frac{1}{\mu}\boldsymbol{\Lambda}_p\right)\mathbf{V}^\top.\quad(19)$$

2) Update $\mathbf{V}$*:* $\mathbf{V}$ is updated by

$$\max_{\mathbf{V}}\operatorname{Tr}(\mathbf{V}\boldsymbol{\Delta}),\ \text{s.t. }\mathbf{VV}^\top=\mathbf{I},\quad(20)$$

where $\boldsymbol{\Delta}=2\mathbf{ZA}^\top+\mu\sum_{p=1}^{v}\boldsymbol{\Pi}_p^\top\mathbf{U}_p$ and $\boldsymbol{\Pi}_p=\mathbf{X}_p-\mathbf{E}_p+\frac{1}{\mu}\boldsymbol{\Lambda}_p$. The problem can be solved by singular value decomposition (SVD) [27].

3) Update $\mathbf{A}$*:* $\mathbf{A}$ is updated by

$$\min_{\mathbf{A}}\operatorname{Tr}\left(\mathbf{A}\left(\mathbf{Z}^\top\mathbf{Z}+\mathbf{D}_m\right)\mathbf{A}^\top-2\mathbf{VZA}^\top\right).\quad(21)$$

By enforcing the partial derivative $\frac{\partial(\cdot)}{\partial\mathbf{A}}=0$, we have

$$\mathbf{A}=\mathbf{VZ}(\mathbf{Z}^\top\mathbf{Z}+\mathbf{D}_m)^{-1}.\quad(22)$$

Remark 2: Let $\boldsymbol{\Omega}=\mathbf{Z}^\top\mathbf{Z}+\mathbf{D}_m$, the inverse $\boldsymbol{\Omega}^{-1}$ exists if and only if $\boldsymbol{\Omega}$ is positive definite (PD) matrix, i.e., all eigenvalues $\{\omega_\iota>0\}_{\iota=1}^{m}$. It is easy to verify that $\mathbf{Z}^\top\mathbf{Z}$ is a positive semi-definite (PSD) matrix and thus its eigenvalues are greater than 0. For the diagonal matrix $\mathbf{D}_m=\operatorname{diag}(\mathbf{Z}^\top\mathbf{1})$, its eigenvalues correspond to its diagonal elements $\{\delta_\iota\}_{\iota=1}^{m}$. Ideally, if every anchor connects to at least one instance, then $\mathbf{Z}^\top\mathbf{1}>0$, and $\mathbf{D}_m=diag(\mathbf{Z}^\top\mathbf{1})$ is PD matrix. In this case, $\boldsymbol{\Omega}$ is the sum of a PSD and a PD matrix, ensuring $\boldsymbol{\Omega}$ is PD and thus invertible.

An undesirable case may arise when an anchor is not connected to any instance, i.e., $\mathbf{a}_j$ is an isolated anchor without building correlations with all samples. In this case, the corresponding diagonal entry δ_j of $\mathbf{D}_m=\operatorname{diag}(\mathbf{Z}^\top\mathbf{1})$ becomes zero, inducing $\mathbf{D}_m$ is not a diagonal matrix. However, such cases are not observed in experiments. Therefore, we empirically assume that $\boldsymbol{\Omega}^{-1}$ exists.

4) Update $\mathbf{Z}$*:* Each $\mathbf{Z}_{[i,:]}$ can be independently updated by quadratic programming (QP) problem,

$$\min_{\mathbf{Z}_{[i,:]}}\frac{1}{2}\mathbf{Z}_{[i,:]}\mathbf{G}\mathbf{Z}_{[i,:]}^\top+\mathbf{r}^\top\mathbf{Z}_{[i,:]}^\top,$$
$$\text{s.t. }\mathbf{Z}_{[i,:]}\mathbf{1}=1,\,\mathbf{Z}_{[i,:]}\geq 0,\quad(23)$$

where $\mathbf{G}=2(\mathbf{A}^\top\mathbf{A}+\lambda\mathbf{I})$ and $\mathbf{r}^\top=diag(\mathbf{A}^\top\mathbf{A})^\top-(2\mathbf{V}^\top\mathbf{A})_{[i,:]}$.

5) Update $\mathbf{E}_p$*:* Each $\mathbf{E}_p$ is independently updated by

$$\max_{\mathbf{E}_p}\frac{\gamma_p^2}{\mu}\|\mathbf{E}_p\|_{2,1}+\frac{1}{2}\|\mathbf{E}_p-\mathbf{Q}_p\|_{\mathrm{F}}^2,\quad(24)$$

where $\mathbf{Q}_p=\mathbf{X}_p-\mathbf{U}_p\mathbf{V}+\frac{1}{\mu}\boldsymbol{\Lambda}_p$. According to [50], the solution is

$$\mathbf{e}_p^i=\begin{cases}\left(1-\frac{\gamma_p}{\mu\|\mathbf{q}_p^i\|_2}\right)\mathbf{q}_p^i, & if\ \frac{\gamma_p}{\mu}<\left\|\mathbf{q}_p^i\right\|_2,\\ 0, & otherwise.\end{cases}\quad(25)$$

6) Update $\boldsymbol{\gamma}$*:* Each γ_p is independently updated by

$$\min_{\gamma_p}\sum_{p=1}^{v}\gamma_p^2\xi_p,\ \text{s.t. }\boldsymbol{\gamma}^\top\mathbf{1}=1,\,\gamma_p\geq 0,\quad(26)$$

where $\xi_p=\|\mathbf{X}_p-\mathbf{U}_p\mathbf{V}\|_{2,1}$. According to Cauchy-Schwarz inequality, we have $\gamma_p=\frac{1/\xi_p}{\sum_{p=1}^{v}1/\xi_p}$.

7) Update $\boldsymbol{\Lambda}$ *and* μ*:* ALM multiplier $\boldsymbol{\Lambda}_p$ and μ are updated by

$$\boldsymbol{\Lambda}_p=\boldsymbol{\Lambda}_p+\mu\left(\mathbf{X}_p-\mathbf{U}_p\mathbf{V}-\mathbf{E}_p\right),$$
$$\mu=\sigma\mu,\quad(27)$$

where μ is the ALM penalty parameter used to update the Lagrange multipliers, while σ is a scaling factor.

G. Initiation of Parameters

Initiation of $\mathbf{Z}$ *and* λ*:* Following a widely used strategy in previous MVBGC works [28], [51], [52] that initializes $\mathbf{Z}$ by

Algorithm 1: JetBGC.

1: **Input**: $\{\mathbf{X}_p\}_{p=1}^{v}$, k, m, d, maximal iteration Γ.
2: **Initialize** $\{\mathbf{U}_p\}_{p=1}^{v}$, $\mathbf{A}$, $\mathbf{Z}$, $\mathbf{V}$, $\boldsymbol{\gamma}$, $\{\boldsymbol{\Lambda}_p\}_{p=1}^{v}$.
3: **while** not converged and iteration less than Γ **do**
4: Update $\{\mathbf{E}_p\}_{p=1}^{v}$ by solving (24).
5: Update $\{\mathbf{U}_p\}_{p=1}^{v}$ by solving (18).
6: Update $\mathbf{A}$ by solving (21).
7: Update $\mathbf{Z}$ by solving (23).
8: Update $\mathbf{V}$ by solving (20).
9: Update $\boldsymbol{\gamma}$ by solving (26).
10: Update $\{\boldsymbol{\Lambda}_p\}_{p=1}^{v}$ by solving (27).
11: **end while**
12: Perform spectral decomposition on $\mathbf{Z}$ to get partition.
13: **Output**: The predicted labels.

retaining the top-ϵ nearest anchors for each instance (measured by Euclidean distance), while setting all other connections to zero. This method has been shown to (i) preserve sparsity of $\mathbf{Z}$ and (ii) avoid the need for manually tuning the parameter λ [51], [52], [53].

For brevity, we only present the solution, the detailed derivations can refer to [52], [53]. In LLPE setting, (23) can be reformulated as

$$\min_{\mathbf{Z}_{[i,:]}} \frac{1}{2}\left\|\mathbf{Z}_{[i,:]} + \frac{1}{2\lambda}\mathbf{r}^{\top}\right\|_2^2, \text{ s.t. } \mathbf{Z}_{[i,:]}\mathbf{1} = 1,\ \mathbf{Z}_{[i,:]} \geq 0. \quad (28)$$

According to [27], the closed-form solution to (28) is $\mathbf{Z}_{[i,:]} = \max\{-\frac{1}{2\lambda}\mathbf{r}^{\top} + \tau_i \mathbf{1}_n, 0\}$, where τ_i can be solved by Newton's method.

Furthermore, by using the Lagrange multiplier technique in [53], λ can be pre-determined by

$$\lambda = \frac{\epsilon}{2}\mathbf{r}_{i,\epsilon+1}^{\top} - \frac{1}{2}\sum_{j=1}^{\epsilon}\mathbf{r}_{i,j}^{\top}, \quad (29)$$

where ϵ denotes the number of neighbors assigned to each instance in initialization, we empirically set $\epsilon = 3$ in experiments, which demonstrates favorable performance in most cases. Further analysis on the sensitivity of ϵ is available in the supplementary material (Section 4.2).

Initiation of other variables: $\mathbf{U}_p$ is initialized to zero due to unconstrained property; $\mathbf{A}$ is initialized as the centroids obtained by applying k-means to the left singular vectors of the concatenated multi-view data; $\mathbf{V}$ is initialized with the left singular vectors of the concatenated multi-view data to satisfy orthogonality constraint; $\boldsymbol{\Lambda_p}$ is initialized to zero; the penalty parameter μ and the scaling parameter σ are initialized to 10 and 2, respectively. Further analysis on the sensitivity of μ and σ are in supplementary material (Section 4.2–4.3).

H. Analysis and Discussion

1) Structural Bipartite Graph Construction in Raw Feature Space: If the bipartite graph is constructed in the input space, JetBGC is reduced to the following form:

$$\begin{aligned} \min_{\{\mathbf{A}_p\}_{p=1}^{v},\mathbf{Z},\boldsymbol{\gamma}} \quad & \sum_{p=1}^{v}\gamma_p^2 \mathrm{Tr}\big(-2\mathbf{X}_p\mathbf{Z}\mathbf{A}_p^{\top} + \eta\mathbf{A}_p\mathbf{Z}^{\top}\mathbf{Z}\mathbf{A}_p^{\top} \\ & + \mathbf{A}_p\mathbf{D}_m\mathbf{A}_p^{\top}\big) + \lambda\|\mathbf{Z}\|_{\mathrm{F}}^2, \\ \text{s.t.} \quad & \begin{cases} \mathbf{Z}\mathbf{1} = \mathbf{1},\ \mathbf{Z} \geq 0; \\ \boldsymbol{\gamma}^{\top}\mathbf{1} = 1,\ \gamma_p \geq 0,\ \forall p \in \{1,\ldots,v\}. \end{cases} \end{aligned} \quad (30)$$

We refer to the above model as *Structural Fusion on Raw Features (SFRF)*. The detailed optimization procedure is provided in the supplementary material (Section 3). Since it excludes the RLEL module, SFRF shows less robustness to noisy features.

2) Convergence: To solve the proposed optimization model in (16), we develop an ADMM-based solver that adopts block coordinate descent strategy [54]. The original objective is decomposed into six subproblems, each of which has a closed-form solution. As discussed in [55], [56], the scaling parameter σ controls the update of ALM penalty μ. A larger σ typically corresponds to fewer iterations to reach the convergence criterion, but may also induces precision loss of the final objective value. Moreover, with increasing μ, the last term in (17) approaches zero, thereby the ALM objective asymptotically converges to the original function, which is bounded by 0. According to previous convergence analyses of ALM [39], [41], [57] and block coordinate descent [58], the original function decreases monotonically during iteration and converges to a local optimum. In experiments, the stopping criterion is as follow,

$$\begin{aligned} &\text{if } (iter > 9) \text{ and } \left(\frac{|obj(iter-1) - obj(iter)|}{obj(iter-1)} < 10^{-3}\right. \\ &\left.\quad \text{or } iter > 30 \text{ or } obj(iter) < 10^{-10}\right), \end{aligned} \quad (31)$$

where $iter$ is the iteration index, and $obj(iter)$ denotes the corresponding objective value.

3) Complexity Analysis: This section analyses the complexities. For simplicity, we set $\zeta = \sum_{p=1}^{v} d_p$.

Time Complexity: The time complexity consists of nine parts. Updating $\{\mathbf{E}_p\}_{p=1}^{v}$ requires $\mathcal{O}(n\zeta d)$ time. Updating $\{\mathbf{U}_p\}_{p=1}^{v}$ requires $\mathcal{O}(n\zeta d)$ time. Updating $\mathbf{A}$ requires $\mathcal{O}(nm(d+m))$ time. Updating $\mathbf{V}$ requires $\mathcal{O}(n(\zeta + dm + d^2 + d + m))$ time. Updating $\mathbf{Z}$ requires $\mathcal{O}(nm(dm+d))$ time. Updating $\boldsymbol{\gamma}$ requires $\mathcal{O}(n\zeta d)$ time. Updating $\{\boldsymbol{\Lambda}_p\}_{p=1}^{v}$ requires $\mathcal{O}(n\zeta d)$ time. The total time complexity is $\mathcal{O}(n(m^2 d + d^2 + \zeta d))$.

Space Complexity: Space complexity is mainly caused by storing matrices, i.e., $\{\mathbf{X}_p, \mathbf{E}_p, \boldsymbol{\Lambda}_p\}_{p=1}^{v} \in \mathbb{R}^{d_p \times n}$, $\{\mathbf{U}_p\}_{p=1}^{v} \in \mathbb{R}^{d_p \times d}$, $\mathbf{V} \in \mathbb{R}^{d \times n}$, $\mathbf{A} \in \mathbb{R}^{d \times m}$, $\mathbf{Z} \in \mathbb{R}^{n \times m}$. The total space complexity is $\mathcal{O}(n(\zeta + d + m) + \zeta d + dm)$.

Therefore, the complexities are linear with n, making it can scale to large-scale datasets with $n \geq 100,000$.

TABLE II
MVC DATASETS

Datasets	View	Instance	Features	Cluster
WebKB_cor	2	195	195/1,703	5
MSRCV1	6	210	1,302/48/512/100/256/210	7
Dermatology	2	358	12/22	6
ORLRnSp	2	400	1,024/288	40
ORL_4Views	4	400	256/256/256/256	40
Movies	2	617	1,878/1,398	17
Flower17	7	1,360	5,376/512/5,376/5,376/1,239/5,376/5,376	17
BDGP	3	2,500	1,000/500/250	5
VGGFace2_50	4	16,936	944/576/512/640	50
YouTubeFace10	4	38,654	944/576/512/640	10
EMNIST Digits	4	280,000	944/576/512/640	10

IV. EXPERIMENT

A. Synthetic Datasets

To visualize the effectiveness of JetBGC intuitively, we design two single-view synthetic datasets.

SData-1: a 4D dataset shown in Fig. 2, consisting of two clusters, each containing 100 samples, i.e., $\mathbf{X} \in \mathbb{R}^{4\times 200}$. The first two dimensionnal features (1st and 2nd dimensions) exhibit a two-moons shape, while the last two dimensions (the 3rd and 4th dimensions) are noisy features that follow Gaussian distributions. Specifically, the pink and green clusters are distributed as $\mathbf{C}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ and $\mathbf{C}_2 \sim \mathcal{C}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$, respectively, where $\boldsymbol{\mu}_1 = (0,0)^\top$, $\boldsymbol{\mu}_2 = (0,0)^\top$, $\boldsymbol{\Sigma} = \mathrm{diag}(0.5, 0.5)$, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the mean vectors, and $\boldsymbol{\Sigma}$ represents the variance.

SData-2: a 2D dataset consist of two clusters, with each cluster containing 100 samples, i.e., $\mathbf{X} \in \mathbb{R}^{2\times 200}$. Both clusters follow Gaussian distributions. Specifically, the 1st cluster (green) is distributed as $\mathbf{C}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$, while the 2nd cluster (pink) follows $\mathbf{C}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$. The mean vectors for the two clusters are $\boldsymbol{\mu}_1 = (0,1)^\top$, $\boldsymbol{\mu}_2 = (0,-1)^\top$. The covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(s_1^2, s_2^2)$, where $s_1 = 3.5$ and $s_2 = 0.4$.

B. Real-World Datasets

Table II summarizes 11 real-world MVC datasets, with the number of instances ranging from 195 to 286,000, the number of clusters ranging from 5 to 50, and the feature dimensions varying from 48 to 5,376. WebKB_cor is a sub-network of WebKB[1] dataset which consists of web pages and hyperlinks, including course, faculty, student, project, and staff categories. MSRCV1[2] is a scene dataset which comprises 210 images from 7 categories, including CENTRIST, CMT, GIST, HOG, LBP, and DoG-SIFT. ORLRnSp and ORL_4views[3] are face datasets containing 400 images from 40 categories but with different views. Movies[4] involves 617 movies drawn from 17 categories, characterized by 2 views (actors and keywords). Flower17[5] is a flower dataset with 17 categories and each one contains 80 images. BDGP[6] contains 2,500 images of drosophila embryos from 5 classes. Each image is described by a 1000D-lateral visual vector, a 500D dorsal visual vector, and a 250D texture vector. VGGFace2_50 is extracted from large-scale face recognition dataset VGGFace2.[7] YouTubeFace10 [59] is a face video datasets collected from YouTube. EMNIST Digits[8] is a subset of handwritten character digits extracted from the NIST Special Database-19,[9] containing 280,000 characters with 10 balanced categories.

C. Compared Baselines

Thirteen state-of-the-art models are compared as baselines, where MCLES [60], PMSC [61], FMR [62], AMGL [63], and RMKM [64] are MVC methods with $\mathcal{O}(n^3)$ computational complexity and $\mathcal{O}(n^2)$ space complexity. BMVC [65], LMVSC [24], SMVSC [25], FPMVS-CAG [26], SFMC [32], FMCNOF [66], SDAFG [33], and MGSL [67] are BGC models with $\mathcal{O}(n)$ time and space complexities. Source codes are collected from public websites or authors' homepage. The hyper-parameters are tuned according to authors' recommendations and we report the best metrics.

D. Experimental Setup

Following common experimental settings in clustering, the cluster number k is known in advance [68], [69], [70], [71]. For baselines requiring k-means as a post-processing to generate discrete clustering labels, we execute k-means 50 times repeatedly to reduce randomness caused by stochastic centroid initialization, and then report mean $\pm$ std. For JetBGC, the anchor number m varies in $[k, 2k, 3k]$, the latent feature dimension d varies in $[k, 2k, 3k, 4k]$, and $d \leq \min\{d_p\}_{p=1}^{v}$ should be satisfied.

Five widely used metrics, namely Accuracy (ACC), Normalized Mutual Information (NMI), F-score, Adjusted Rand Index (ARI), and Purity, are used to measure clustering performance [72], [73], [74]. Experiments are obtained from a server with 12 Core Intel(R) i9 10900 K CPUs @3.6 GHZ, 64 GB RAM, and Matlab 2020b.

E. Effectiveness

Table III reports clustering metrics, due to the space limitation, the results of MSGL are available in supplementary material (Section 4.7). From the results, we find that:

1) JetBGC achieves competitive clustering performance, whose ACC outperforms the second best with large margins of 12.32%, 10.10%, 10.65%, 0.79%, 4.05%, 2.86%, 15.38%, 8.96%, 2.99%, 4.91%, and 10.42% on eleven datasets, respectively. On average, our model outperforms competitors with 7.59%, 7.95%, 5.22%, 7.69%, and 9.55% improvements of ACC, NMI, F-score, ARI, and Purity, respectively, fully validating our effectiveness.
2) MCLES, PMSC, FMR, AMGL, and RMKM are MVGC models with $\mathcal{O}(n^3)$ time complexity and $\mathcal{O}(n^2)$ space complexity, these baselines cannot scale to large-scale

[1] https://starling.utdallas.edu/datasets/webkb/
[2] https://www.microsoft.com/en-us/research/project/image-understanding/downloads/
[3] https://cam-orl.co.uk/facedatabase.html
[4] https://lig-membres.imag.fr/grimal/data.html
[5] https://www.robots.ox.ac.uk/vgg/data/flowers/17/
[6] https://www.fruitfly.org/
[7] https://www.robots.ox.ac.uk/vgg/data/vgg_face2/
[8] https://www.nist.gov/itl/products-and-services/emnist-dataset
[9] https://www.nist.gov/srd/nist-special-database-19

TABLE III
COMPARISON OF CLUSTERING METRICS

Datasets	MCLES	PMSC	FMR	AMGL	RMKM	BMVC	LMVSC	SMVSC	FPMVS-CAG	SFMC	FMCNOF	SDAFG	Proposed
ACC (%)													
WebKB_cor	37.92±2.52	*45.51±3.96*	42.28±2.07	27.13±1.75	43.08±0.00	43.59±0.00	44.50±1.77	37.58±4.55	44.71±1.80	44.10±0.00	42.05±0.00	44.10±0.00	**57.83±1.88**
MSRCV1	60.86±3.45	47.45±4.23	77.48±6.40	76.44±6.30	71.43±0.00	26.67±0.00	*83.73±7.20*	70.51±4.98	71.95±5.36	60.48±0.00	47.14±0.00	70.95±0.00	**93.84±7.55**
Dermatology	49.67±4.72	80.75±4.46	81.72±5.66	22.57±0.59	74.86±0.00	63.97±0.00	79.02±6.63	78.64±5.41	*82.96±7.44*	49.44±0.00	62.01±0.00	56.70±0.00	**93.61±5.87**
ORLRnSp	31.64±1.72	31.36±1.60	47.62±2.58	63.55±3.11	54.75±0.00	46.25±0.00	60.37±2.28	69.60±2.82	*70.88±1.47*	37.25±0.00	19.50±0.00	61.50±0.00	**71.67±2.69**
ORL_4Views	29.18±1.99	21.48±1.02	25.21±1.10	59.73±2.78	47.00±0.00	43.25±0.00	*61.50±2.96*	47.76±2.36	54.63±1.49	37.00±0.00	21.50±0.00	57.75±0.00	**65.55±2.64**
Movies	*29.65±1.34*	21.70±0.82	22.49±1.04	11.66±0.54	17.99±0.00	20.58±0.00	26.45±1.59	25.57±1.01	25.76±0.05	11.51±0.00	17.02±0.00	9.08±0.00	**32.50±1.37**
Flower17	OM	20.82±0.74	33.43±1.75	9.70±1.53	23.24±0.00	26.99±0.00	*37.12±1.86*	27.13±0.84	25.99±1.83	7.57±0.00	17.43±0.00	8.68±0.00	**52.50±1.94**
BDGP	OM	27.86±0.00	41.84±0.03	34.48±2.23	41.36±0.00	39.76±0.00	*49.52±2.39*	37.22±2.03	32.62±1.17	14.32±0.00	31.08±0.00	20.16±0.00	**58.48±0.03**
VGGFace2_50	OM	OM	OM	2.95±0.35	8.23±0.00	10.30±0.00	10.56±0.26	*13.36±0.60*	12.06±0.36	3.64±0.00	5.51±0.00	3.58±0.00	**16.35±0.41**
YouTubeFace10	OM	OM	OM	OM	*74.88±0.00*	58.58±0.00	74.48±5.30	72.93±3.96	67.09±5.70	55.80±0.00	43.42±0.00	64.48±0.00	**79.79±5.51**
EMNIST Digits	OM	OM	OM	OM	OM	*68.99±0.00*	61.75±4.05	-	62.24±4.08	-	36.50±0.00	61.54±0.00	**79.40±2.62**
Avg Rank	10.6	9.2	7.0	8.5	6.4	7.0	3.4	6.2	4.2	10.1	9.6	7.9	1.0
NMI (%)													
WebKB_cor	4.94±0.62	20.13±1.04	*22.39±1.50*	6.78±1.59	14.84±0.00	6.59±0.00	16.61±1.04	9.52±2.38	14.49±1.94	5.08±0.00	13.36±0.00	5.08±0.00	**38.88±1.58**
MSRCV1	51.72±2.37	34.29±2.81	69.48±3.31	77.65±3.23	63.03±0.00	8.29±0.00	*78.93±4.60*	62.01±2.61	65.69±3.27	60.23±0.00	38.42±0.00	76.23±0.00	**92.62±4.94**
Dermatology	41.67±3.43	*85.11±1.91*	79.97±3.67	3.20±0.57	71.10±0.00	60.79±0.00	70.17±3.94	66.62±2.66	71.90±5.05	38.68±0.00	54.24±0.00	51.61±0.00	**89.69±2.62**
ORLRnSp	57.91±1.60	56.34±1.30	68.86±1.52	83.03±1.44	74.35±0.00	66.94±0.00	78.93±0.93	*84.84±0.79*	84.75±0.41	76.42±0.00	39.50±0.00	80.51±0.00	**86.57±1.01**
ORL_4Views	54.03±1.67	43.87±0.84	48.33±0.81	*80.28±1.37*	71.83±0.00	65.32±0.00	79.39±1.15	72.73±1.13	77.41±0.54	76.30±0.00	43.32±0.00	76.89±0.00	**81.77±1.02**
Movies	*29.76±1.02*	20.03±0.68	19.58±0.98	12.14±0.66	14.92±0.00	18.19±0.00	25.94±1.10	23.21±1.10	25.07±0.19	28.72±0.00	12.42±0.00	5.72±0.00	**31.91±0.96**
Flower17	OM	19.13±0.48	30.65±0.91	10.25±4.01	22.07±0.00	25.62±0.00	*35.37±1.10*	25.78±0.76	25.81±1.59	7.87±0.00	14.68±0.00	5.64±0.00	**49.57±0.94**
BDGP	OM	4.52±0.00	12.57±0.05	17.21±3.46	16.98±0.00	15.74±0.00	25.85±1.92	9.85±0.71	10.02±1.33	*26.23±0.00*	10.29±0.00	0.32±0.00	**35.70±0.05**
VGGFace2_50	OM	OM	OM	2.04±0.50	9.66±0.00	13.48±0.00	12.64±0.28	*16.21±0.49*	14.74±0.55	1.63±0.00	4.74±0.00	2.01±0.00	**19.25±0.35**
YouTubeFace10	OM	OM	OM	OM	*78.83±0.00*	54.66±0.00	77.74±2.03	78.57±2.80	76.11±3.06	77.46±0.00	39.15±0.00	71.97±0.00	**84.48±2.38**
EMNIST Digits	OM	OM	OM	OM	OM	70.08±0.00	61.87±2.47	-	53.47±2.44	-	28.36±0.00	*72.73±0.00*	**72.77±0.99**
Avg Rank	10.9	9.2	7.2	7.5	6.4	7.8	3.9	6.4	5.0	8.1	9.4	8.3	1.0
F-score (%)													
WebKB_cor	37.14±2.83	36.45±2.61	34.38±1.30	28.81±2.08	33.82±0.00	40.32±0.00	36.10±1.37	31.11±3.30	36.93±0.97	*42.89±0.00*	35.94±0.00	*42.89±0.00*	**49.76±1.79**
MSRCV1	48.53±2.10	34.05±2.34	66.76±4.50	70.28±4.42	59.98±0.00	16.01±0.00	*77.43±6.43*	59.31±2.82	61.55±3.54	52.43±0.00	33.85±0.00	63.58±0.00	**91.64±7.36**
Dermatology	42.79±2.67	*83.50±4.24*	77.59±5.15	18.46±0.78	74.82±0.00	56.62±0.00	70.50±4.17	70.06±3.60	77.37±6.23	42.90±0.00	57.89±0.00	53.89±0.00	**91.93±5.04**
ORLRnSp	20.64±1.50	16.83±1.38	33.01±2.50	40.03±6.27	39.54±0.00	28.86±0.00	49.26±2.30	*56.80±2.18*	56.61±1.22	19.44±0.00	8.57±0.00	27.98±0.00	**62.31±2.85**
ORL_4Views	17.63±1.35	6.48±0.54	9.27±0.83	35.12±4.54	33.66±0.00	24.84±0.00	*50.10±2.86*	32.37±1.71	42.87±1.47	23.74±0.00	12.09±0.00	23.11±0.00	**53.20±2.63**
Movies	*16.85±0.70*	13.27±0.40	11.12±0.47	10.48±0.08	9.81±0.00	9.93±0.00	15.06±0.90	15.65±0.57	16.24±0.07	10.92±0.00	13.94±0.00	11.48±0.00	**20.48±1.17**
Flower17	OM	12.33±0.36	20.09±0.81	11.49±0.55	14.35±0.00	16.61±0.00	*23.99±0.95*	17.53±0.21	17.29±0.39	10.94±0.00	13.93±0.00	11.02±0.00	**36.03±1.05**
BDGP	OM	28.59±0.00	29.21±0.04	34.79±1.04	30.25±0.00	32.40±0.00	*38.51±1.04*	28.81±0.38	28.79±0.58	25.09±0.00	28.89±0.00	33.27±0.00	**46.29±0.03**
VGGFace2_50	OM	OM	OM	3.91±0.02	3.69±0.00	5.10±0.00	5.09±0.15	*6.35±0.18*	6.10±0.07	4.16±0.00	4.36±0.00	4.15±0.00	**8.37±0.20**
YouTubeFace10	OM	OM	OM	OM	66.93±0.00	52.53±0.00	*68.93±3.19*	68.34±5.78	66.10±5.06	61.25±0.00	32.88±0.00	52.40±0.00	**78.70±4.53**
EMNIST Digits	OM	OM	OM	OM	OM	*61.38±0.00*	54.07±3.51	-	49.29±3.01	-	25.64±0.00	57.93±0.00	**70.30±1.41**
Avg Rank	10.2	9.6	7.8	8.1	7.4	7.0	3.8	6.3	4.8	9.2	8.7	7.1	1.0
ARI (%)													
WebKB_cor	3.03±2.41	*14.72±1.53*	12.39±1.30	0.59±0.58	10.27±0.00	7.78±0.00	13.55±1.67	6.51±4.67	12.25±2.08	1.35±0.00	11.21±0.00	1.35±0.00	**33.95±2.35**
MSRCV1	38.66±2.82	22.74±2.83	61.19±5.41	64.81±5.50	53.33±0.00	2.26±0.00	*73.60±7.66*	52.28±3.58	54.55±4.63	42.25±0.00	21.60±0.00	55.70±0.00	**90.24±8.65**
Dermatology	21.54±4.29	*79.17±5.33*	72.16±6.44	0.20±0.17	68.33±0.00	45.02±0.00	63.25±5.63	62.79±4.82	71.54±8.52	17.22±0.00	43.75±0.00	40.43±0.00	**89.97±6.18**
ORLRnSp	17.83±1.60	14.44±1.45	31.38±2.57	38.02±6.61	37.92±0.00	27.06±0.00	47.96±2.38	*55.61±2.27*	55.43±1.26	17.06±0.00	4.91±0.00	25.27±0.00	**61.32±2.94**
ORL_4Views	14.66±1.46	3.81±0.55	7.06±0.84	32.88±4.83	31.74±0.00	22.88±0.00	*48.82±2.96*	30.13±1.81	41.20±1.54	21.86±0.00	8.60±0.00	20.14±0.00	**51.96±2.73**
Movies	*10.81±0.85*	5.07±0.57	4.99±0.51	0.11±0.05	1.98±0.00	4.00±0.00	9.09±1.05	9.20±0.84	9.26±0.10	-0.01±0.00	4.31±0.00	0.01±0.00	**15.20±1.24**
Flower17	OM	6.70±0.36	15.09±0.86	0.72±0.71	8.83±0.00	11.06±0.00	*19.18±1.02*	10.57±0.26	9.86±0.53	0.01±0.00	5.26±0.00	0.26±0.00	**31.85±1.11**
BDGP	OM	3.26±0.00	11.26±0.05	7.90±2.26	11.61±0.00	12.59±0.00	*22.82±1.85*	6.65±0.37	5.69±0.85	0.49±0.00	6.01±0.00	0.00±0.00	**32.62±0.04**
VGGFace2_50	OM	OM	OM	0.05±0.05	1.59±0.00	2.98±0.00	3.06±0.16	*3.77±0.22*	3.10±0.09	0.01±0.00	0.65±0.00	0.01±0.00	**6.39±0.21**
YouTubeFace10	OM	OM	OM	OM	62.80±0.00	46.36±0.00	*64.86±3.75*	63.93±7.01	61.13±6.13	55.95±0.00	22.54±0.00	44.44±0.00	**76.07±5.13**
EMNIST Digits	OM	OM	OM	OM	OM	*56.85±0.00*	48.70±3.97	-	43.13±3.77	-	16.13±0.00	51.83±0.00	**66.93±1.59**
Avg Rank	10.5	8.9	6.9	8.3	6.2	6.7	3.3	5.9	4.7	10.5	8.9	9.3	1.0
Purity (%)													
WebKB_cor	43.15±0.36	*54.05±1.59*	51.33±1.48	27.59±1.91	49.23±0.00	43.59±0.00	49.26±1.39	47.71±3.76	49.42±1.43	44.62±0.00	49.23±0.00	44.62±0.00	**68.62±1.57**
MSRCV1	61.57±2.91	49.91±3.78	79.01±4.16	80.45±4.29	74.76±0.00	27.14±0.00	*85.25±5.56*	71.51±4.02	72.33±5.01	62.86±0.00	50.48±0.00	70.95±0.00	**94.52±6.17**
Dermatology	54.59±3.51	*85.37±1.89*	84.79±2.87	23.12±0.50	75.70±0.00	65.08±0.00	80.97±4.29	80.35±3.73	83.55±6.84	50.00±0.00	62.85±0.00	66.48±0.00	**94.89±2.15**
ORLRnSp	33.78±1.64	34.39±1.58	50.97±2.45	70.35±2.44	57.75±0.00	49.25±0.00	63.96±1.89	73.29±2.42	74.60±1.42	**80.25±0.00**	21.25±0.00	67.50±0.00	*75.03±2.35*
ORL_4Views	31.96±1.78	23.90±1.08	26.48±1.14	66.92±2.07	53.00±0.00	47.50±0.00	65.68±2.53	51.91±2.16	58.97±1.39	**78.00±0.00**	21.75±0.00	63.00±0.00	*68.95±2.06*
Movies	*32.92±1.05*	24.69±1.24	24.84±1.14	12.07±0.52	19.29±0.00	23.01±0.00	29.36±1.44	27.90±1.42	28.83±0.10	28.53±0.00	17.83±0.00	10.53±0.00	**35.39±1.03**
Flower17	OM	22.20±0.62	34.74±1.38	10.76±1.53	24.49±0.00	29.41±0.00	*38.81±1.54*	27.88±0.78	26.38±1.85	10.29±0.00	17.57±0.00	9.19±0.00	**54.13±1.76**
BDGP	OM	30.31±0.00	42.36±0.06	35.75±2.67	42.12±0.00	40.52±0.00	49.64±1.90	37.80±1.22	34.82±1.23	*56.60±0.00*	33.32±0.00	20.16±0.00	**60.65±0.03**
VGGFace2_50	OM	OM	OM	3.17±0.37	9.28±0.00	11.44±0.00	11.41±0.25	*13.90±0.58*	12.27±0.36	3.86±0.00	5.67±0.00	3.74±0.00	**17.23±0.35**
YouTubeFace10	OM	OM	OM	OM	*79.70±0.00*	63.87±0.00	78.39±3.43	77.35±4.61	69.43±5.88	74.10±0.00	46.53±0.00	69.50±0.00	**85.33±3.30**
EMNIST Digits	OM	OM	OM	OM	OM	*71.38±0.00*	65.08±2.93	-	62.30±4.04	-	36.90±0.00	66.12±0.00	**79.99±1.80**
Avg Rank	10.9	9.1	6.8	8.6	6.4	7.7	3.9	6.3	5.2	6.9	9.6	8.4	1.2

The best metrics are marked in bold, the second-best ones are italicized and underlined, "OOM" is out-of-memory error, and "-" means unavailable results caused by extremely long running time.

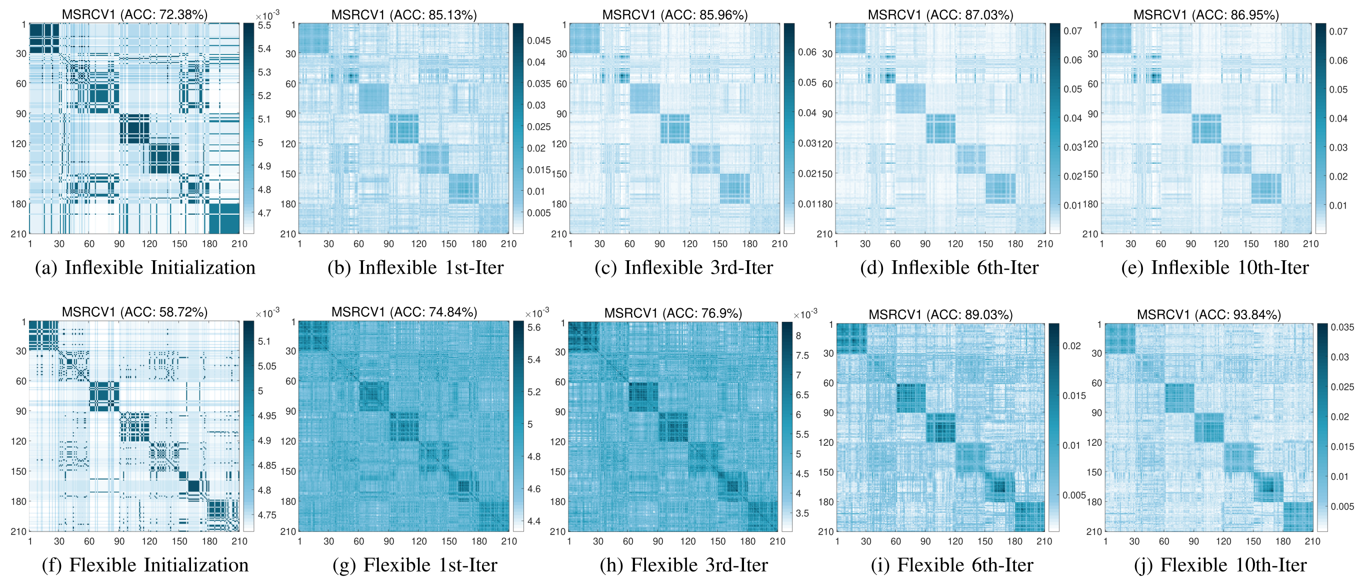
datasets ($n \geq 100,000$) and easily incur "OOM" error, greatly limiting their applications. Two MVBGC baselines, SFMC and SMVSC, also suffer unavailable results "-" on EMNIST Digits, due to unacceptable extremely long running time caused by complex optimization.

3) Most compared baselines construct graphs directly from the input data rather than latent embedding, making them sensitive to noisy or redundant features, thereby degrading the quality of graph. In addition, they are typically built upon either the LRP or LLP paradigm. These limitations can explain their degraded performance.

F. Comparison With Inflexible Anchor Selection

This section validates the "flexibility" of constrained-free anchor optimization. We denote constrained-free anchor selection strategy as "Flexible", while the baseline that select anchors via k-means is denoted as "Inflexible". For a fair comparison, all experimental settings are kept consistent.

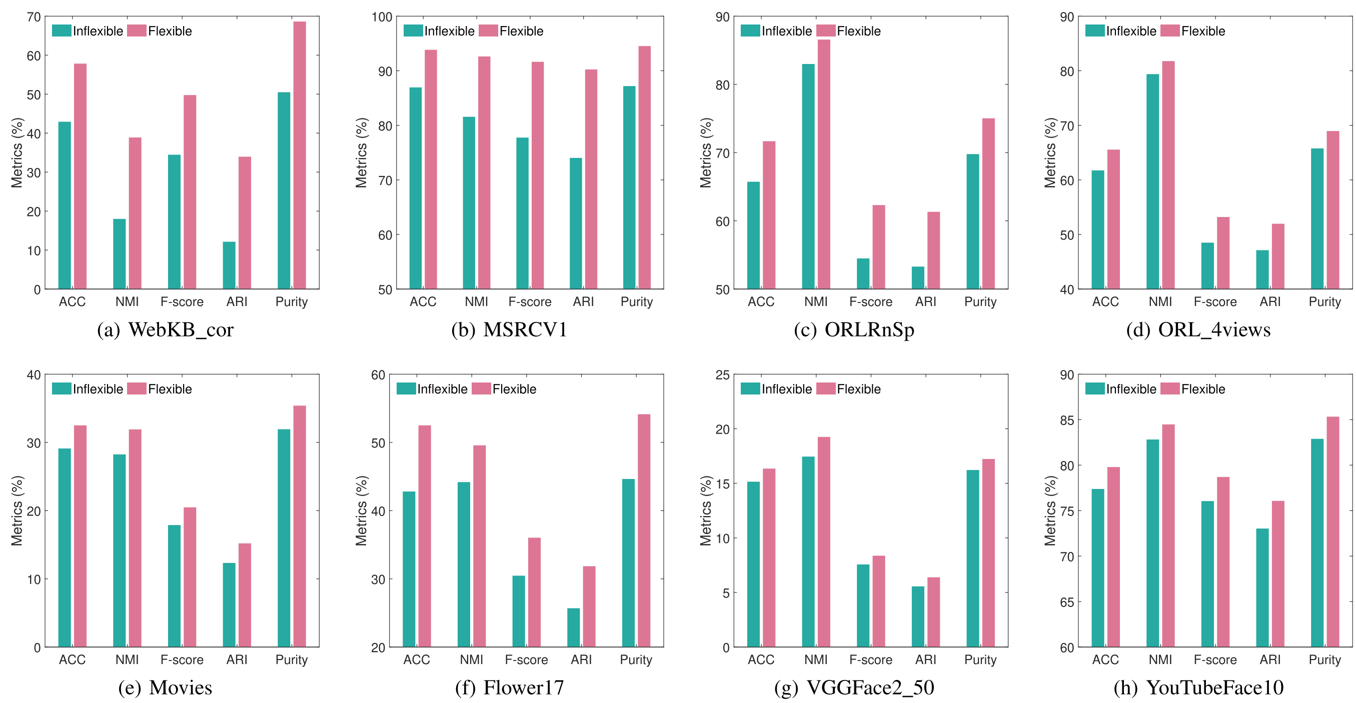
Fig. 4 visualizes the evolution of the normalized affinity graph ($\mathbf{Z}\mathbf{D}_m^{-1}\mathbf{Z}^\top$) on MSRCV1. It is worth noting that the "Inflexible" method achieves its best performance with $m = 1k$ and $d = 2k$, while the "Flexible" method performs best with $m = 2k$ and $d = 3k$, thereby their initialized graphs are different. Moreover,



(a) Inflexible Initialization (b) Inflexible 1st-Iter (c) Inflexible 3rd-Iter (d) Inflexible 6th-Iter (e) Inflexible 10th-Iter

(f) Flexible Initialization (g) Flexible 1st-Iter (h) Flexible 3rd-Iter (i) Flexible 6th-Iter (j) Flexible 10th-Iter

Fig. 4. Evolution of the normalized affinity matrix over iterations on MSRCV1.



(a) WebKB_cor (b) MSRCV1 (c) ORLRnSp (d) ORL_4views

(e) Movies (f) Flower17 (g) VGGFace2_50 (h) YouTubeFace10

Fig. 5. Clustering performance: flexible versus inflexible anchor selection.

the "Inflexible" method achieves a 14.57% improvement in ACC through optimization, with most of the gains occurring in the first iteration. However, in subsequent iterations, the reinforcement of the graph is inconspicuous. Differently, the "Flexible" method progressively improves the performance over iterations. It gradually reduces inter-cluster noisy similarities, refines block-diagonal structures, and achieves a 35.12% improvement in ACC. The visualization results demonstrate the effectiveness of our unconstrained anchor optimization strategy.

Fig. 5 further quantifies clustering metrics. The "Flexible" method consistently outperforms the "Inflexible" manner by large margins with an average of 4.79%, 5.29%, 5.41%, 6.47%, and 5.22% improvements respecting ACC, NMI, F-score, ARI, and Purity, respectively. The improvements demonstrate the superiority of the flexible learnable anchor strategy.

G. Ablation Study

This section validates the "complementarity" by comparing JetBGC with LRPE, and LLPE backbones. For comparison, we also report the results of SFRF and "Inflexible" baseline. Table IV gives the experimental settings.

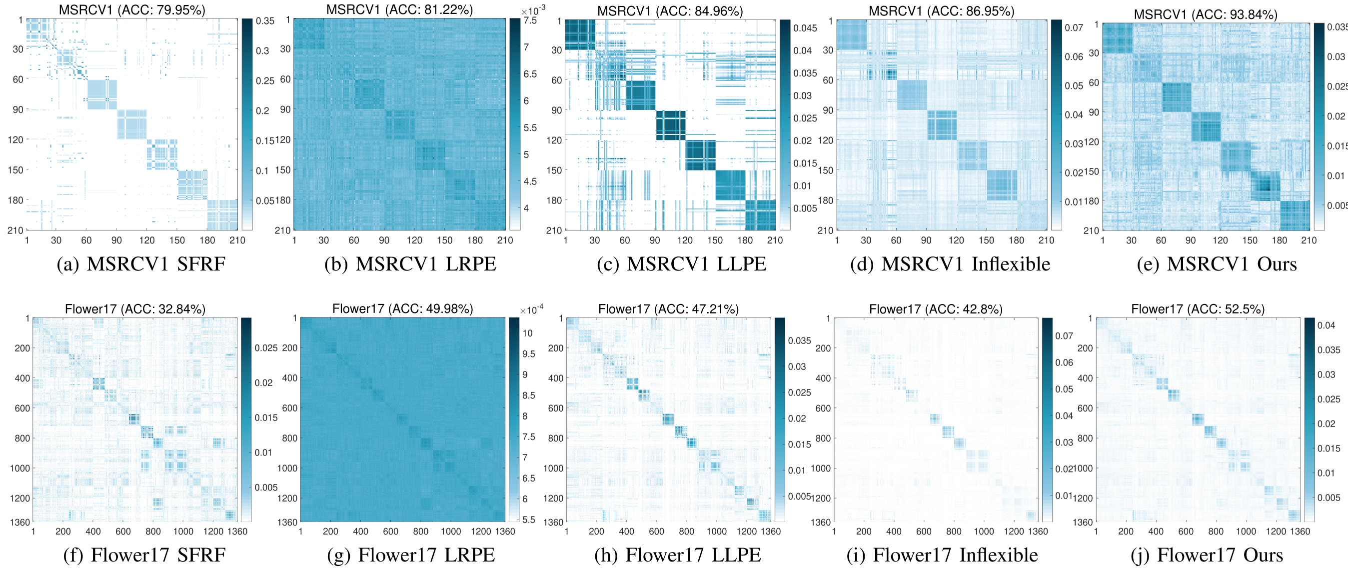


Fig. 6. Visualization of normalized affinity graph of SFRF, LRPE, LLPE, and our SFLE methods.

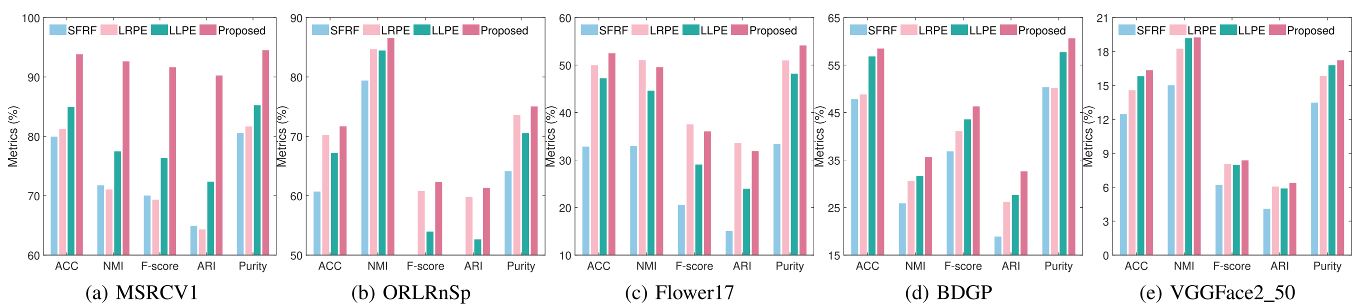


Fig. 7. Performance comparison of LRPE, LLPE, SFRF, and the proposed model.

TABLE IV
EXPERIMENTAL SETTINGS OF ABLATION ANALYSIS

Model	Embedding	Linear	Locally linear
SFRF	–	✓	✓
LRPE	✓	–	✓
LLPE	✓	✓	–
Proposed	✓	✓	✓

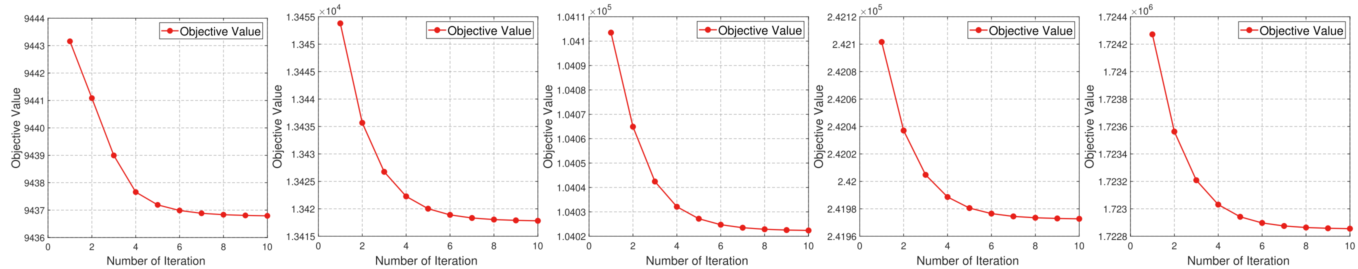
Fig. 6 visualizes the normalized affinity graph on MSRCV1 and Flower17, and Fig. 7 presents a comparison of clustering metrics. We observe that:

1) LRPE, derived from self-expressive subspace clustering, constructs correlations between each instance and all anchors. As a result, the graph shows a fuzzy representation, with many noisy inter-cluster similarities, which degrade the block-diagonal structure and the quality of clustering.
2) LLPE, grounded in manifold learning, emphasizes locality by connecting each instance to a few neighbor anchors. Therefore, the graphs is more sparser and contains fewer noisy connections. However, several dominant noisy similarities may mislead clustering partition.
3) SFLE integrates the properties of LRPE and LLPE, achieving a more discriminative graph with promising metrics. Compared to LRPE, JetBGC exploits clearer block-diagonal structures, while compared to LLPE, JetBGC reduces the noisy similarities.
4) Compared to SFRF, which constructs graphs from raw features, JetBGC introduces a robust embedding module for feature extraction, which reduces the negative impact of the noisy features.
5) JetBGC outperforms the compared baselines with competitive performance on almost all datasets.

These results are convincing evidence to verify the overall superiority of JetBGC.

H. Convergence

Fig. 8 empirically validates the "convergence" of JetBGC. As discussed in our theoretical analysis, with the increase of ALM parameter μ, the ALM objective function gradually converges to the original function, which is bounded by 0. In experiments, we find that the objective decreases and stabilizes within ten



(a) Flower17 (b) BDGP (c) VGGFace2_50 (d) YouTubeFace10 (e) EMNIST Digits

Fig. 8. Experimental validation of the convergence.

iterations, verifying convergence. More convergence results are available in supplementary material (Section 10).

V. Conclusion

This paper proposes a novel bipartite graph clustering model, JetBGC, which focuses on three aspects: robustness, flexibility, and complementarity. To improve robustness, we derive a new feature extractor that learns robust latent embedding, which reduces the adverse impact of noisy features. To achieve flexibility, we design a constraint-free anchor optimization strategy instead of following the existing fixed anchor or constrained learnable methods. To enhance complementarity, we bride the connection of two popular BGC paradigms, and design a novel structural bipartite graph fusion strategy from a unified perspective, to integrate global complementary structures. Overall, JetBGC integrates robust embedding learning, constraint-free anchor optimization, and structural bipartite graph fusion into a unified framework. This paper provides new insights into enhancing bipartite graph clustering that will inspire more variants in the BGC community. One limitation of JetBGC is its reliance on post-processing to generate the discrete clustering labels, which may introduce variance in the performance. Alternative strategies, such as Laplacian rank constraint [53] or one-pass clustering method [75], provide potential solutions for directly generating labels, which will be our future research. Another limitation is its assumption of complete multi-view data. However, in many scenarios, missing data is common due to sensor failures or data corruption. Tackling incomplete multi-view data remains a challenging yet practical problem, and we leave this in subsequent research.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[2] S. Huang, I. W. Tsang, Z. Xu, and J. Lv, "Latent representation guided multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 7, pp. 7082–7087, Jul. 2023.
[3] J. Xu et al., "Self-supervised discriminative feature learning for deep multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 7, pp. 7470–7482, Jul. 2023.
[4] C. Liu, J. Wen, Y. Xu, B. Zhang, L. Nie, and M. Zhang, "Reliable representation learning for incomplete multi-view missing multi-label classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 47, no. 6, pp. 4940–4956, Jun. 2025.
[5] Y. Liu et al., "Dink-net: Neural clustering on large graphs," in *Proc. 40th Int. Conf. Mach. Learn.*, 2023, pp. 21794–21812.
[6] W. Tu et al., "WAGE: Weight-sharing attribute-missing graph autoencoder," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 47, no. 7, pp. 5760–5777, Jul. 2025.
[7] Y. Wang, Y. Tong, C. Long, P. Xu, K. Xu, and W. Lv, "Adaptive dynamic bipartite graph matching: A reinforcement learning approach," in *Proc. 35th IEEE Int. Conf. Data. Mining*, 2019, pp. 1478–1489.
[8] Y. Gao, T. Zhang, L. Qiu, Q. Linghu, and G. Chen, "Time-respecting flow graph pattern matching on temporal graphs," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 10, pp. 3453–3467, Oct. 2021.
[9] Y. Liu et al., "Deep graph clustering via dual correlation reduction," in *Proc. 36th AAAI Conf. Artif. Intell.*, 2022, pp. 7603–7611.
[10] H. Li, Y. Feng, C. Xia, and J. Cao, "Overlapping graph clustering in attributed networks via generalized cluster potential game," *ACM Trans. Knowl. Discov. Data*, vol. 18, no. 1, pp. 27:1–27:26, 2024.
[11] H. Zhong et al., "Graph contrastive clustering," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 9224–9233.
[12] Y. Liu et al., "Improved dual correlation reduction network with affinity recovery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 4, pp. 6159–6173, Apr. 2025.
[13] E. Pan and Z. Kang, "Multi-view contrastive graph clustering," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2021, pp. 2148–2159.
[14] Z. Lin, Z. Kang, L. Zhang, and L. Tian, "Multi-view attributed graph clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 2, pp. 1872–1880, Feb. 2023.
[15] C. Liu, J. Wen, Z. Wu, X. Luo, C. Huang, and Y. Xu, "Information recovery-driven deep incomplete multiview clustering network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 11, pp. 15442–15452, Nov. 2024.
[16] U. Fang, M. Li, J. Li, L. Gao, T. Jia, and Y. Zhang, "A comprehensive survey on multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12350–12368, Dec. 2023.
[17] Y. Liu et al., "End-to-end learnable clustering for intent learning in recommendation," in *Proc. 38th Int. Conf. Neural Inf. Process. Syst.*, 2024, pp. 5913–5949.
[18] H. Xiao, Y. Chen, and X. Shi, "Knowledge graph embedding based on multi-view clustering framework," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 2, pp. 585–596, Feb. 2021.
[19] W. Zhang, L. Jiao, F. Liu, S. Yang, and J. Liu, "Adaptive contourlet fusion clustering for SAR image change detection," *IEEE Trans. Image Process.*, vol. 31, pp. 2295–2308, 2022.
[20] S. Huang, I. Tsang, Z. Xu, and J. C. Lv, "Measuring diversity in graph learning: A unified framework for structured multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 12, pp. 5869–5883, Dec. 2022.
[21] F. Nie, W. Chang, Z. Hu, and X. Li, "Robust subspace clustering with low-rank structure constraint," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 3, pp. 1404–1415, Mar. 2022.
[22] S. Shi, F. Nie, R. Wang, and X. Li, "Fast multi-view clustering via prototype graph," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 1, pp. 443–455, Jan. 2023.
[23] L. Li et al., "BGAE: Auto-encoding multi-view bipartite graph clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 8, pp. 3682–3696, Aug. 2024.
[24] Z. Kang, W. Zhou, Z. Zhao, J. Shao, M. Han, and Z. Xu, "Large-scale multi-view subspace clustering in linear time," in *Proc. 34th Conf. Artif. Intell.*, 2020, pp. 4412–4419.
[25] M. Sun et al., "Scalable multi-view subspace clustering with unified anchors," in *Proc. 29th ACM Int. Conf. Multimedia*, 2021, pp. 3528–3536.
[26] S. Wang et al., "Fast parameter-free multi-view subspace clustering with consensus anchor guidance," *IEEE Trans. Image Process.*, vol. 31, pp. 556–568, 2022.

[27] L. Li, J. Zhang, S. Wang, X. Liu, K. Li, and K. Li, "Multi-view bipartite graph clustering with coupled noisy feature filter," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12842–12854, Dec. 2023.

[28] Z. Wang, L. Zhang, R. Wang, F. Nie, and X. Li, "Semi-supervised learning via bipartite graph construction with adaptive neighbors," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 5, pp. 5257–5268, May 2023.

[29] L. Li and H. He, "Bipartite graph based multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 7, pp. 3111–3125, Jul. 2022.

[30] F. Nie, X. Dong, L. Tian, R. Wang, and X. Li, "Unsupervised feature selection with constrained $\ell_{2,0}$-norm and optimized graph," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 4, pp. 1702–1713, Apr. 2022.

[31] H. Chen, F. Nie, R. Wang, and X. Li, "Fast unsupervised feature selection with bipartite graph and $\ell_{2,0}$-norm constraint," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 5, pp. 4781–4793, May 2023.

[32] X. Li, H. Zhang, R. Wang, and F. Nie, "Multiview clustering: A scalable and parameter-free bipartite graph fusion method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 1, pp. 330–344, Jan. 2022.

[33] X. Lu and S. Feng, "Structure diversity-induced anchor graph fusion for multi-view clustering," *ACM Trans. Knowl. Discov. Data*, vol. 17, no. 2, pp. 17:1–17:18, 2023.

[34] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[35] C. Ding, X. He, and H. D. Simon, "On the equivalence of nonnegative matrix factorization and spectral clustering," in *Proc. SIAM Int. Conf. Data. Mining*, 2005, pp. 606–610.

[36] D. Cai, X. He, J. Han, and T. S. Huang, "Graph regularized nonnegative matrix factorization for data representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1548–1560, Aug. 2011.

[37] D. Kuang, C. Ding, and H. Park, "Symmetric nonnegative matrix factorization for graph clustering," in *Proc. SIAM Int. Conf. Data. Mining*, 2012, pp. 106–117.

[38] C. H. Q. Ding, T. Li, W. Peng, and H. Park, "Orthogonal nonnegative matrix t-factorizations for clustering," in *Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2006, pp. 126–135.

[39] D. Kong, C. H. Q. Ding, and H. Huang, "Robust nonnegative matrix factorization using $\ell_{2,1}$-norm," in *Proc. 20th ACM Int. Conf. Knowl. Manage.*, 2011, pp. 673–682.

[40] C. H. Q. Ding, D. Zhou, X. He, and H. Zha, " R_1-PCA: Rotational invariant ℓ_1-norm principal component analysis for robust subspace factorization," in *Proc. 23th Int. Conf. Mach. Learn.*, 2006, pp. 281–288.

[41] J. Huang, F. Nie, H. Huang, and C. Ding, "Robust manifold nonnegative matrix factorization," *ACM Trans. Knowl. Discov. Data*, vol. 8, no. 3, pp. 1–21, 2014.

[42] X. Li, M. Chen, and Q. Wang, "Discrimination-aware projected matrix factorization," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 4, pp. 809–814, Apr. 2020.

[43] T. Li and C.-C. Ding, "Nonnegative matrix factorizations for clustering: A survey," in *Data Clustering*. London, U.K./Boca Raton, FL, USA: Chapman and Hall/CRC Press, 2018, pp. 149–176.

[44] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, Nov. 2013.

[45] L. K. Saul and S. T. Roweis, "Think globally, fit locally: Unsupervised learning of low dimensional manifolds," *J. Mach. Learn. Res.*, vol. 4, no. Jun., pp. 119–155, 2003.

[46] H. Zhang, J. Shi, R. Zhang, and X. Li, "Non-graph data clustering via $\mathcal{O}(n)$ bipartite graph convolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 7, pp. 8729–8742, Jul. 2023.

[47] C. H. Ding, T. Li, and M. I. Jordan, "Convex and semi-nonnegative matrix factorizations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 1, pp. 45–55, Jan. 2010.

[48] D. Cai and X. Chen, "Large scale spectral clustering via landmark-based sparse representation," *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1669–1680, Aug. 2015.

[49] Y. Li, F. Nie, H. Huang, and J. Huang, "Large-scale multi-view spectral clustering via bipartite graph," in *Proc. 39th AAAI Conf. Artif. Intell.*, 2015, pp. 2750–2756.

[50] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. Roy. Statist. Soc. Ser. Statist. Methodol.*, vol. 68, no. 1, pp. 49–67, 2006.

[51] X. Yang, G. Lin, Y. Liu, F. Nie, and L. Lin, "Fast spectral embedded clustering based on structured graph learning for large-scale hyperspectral image," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 5501705.

[52] J. Wang, L. Wang, F. Nie, and X. Li, "Fast unsupervised projection for large-scale data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 8, pp. 3634–3644, Aug. 2022.

[53] F. Nie, X. Wang, M. I. Jordan, and H. Huang, "The constrained Laplacian rank algorithm for graph-based clustering," in *Proc. 30th AAAI Conf. Artif. Intell.*, 2016, pp. 1969–1976.

[54] S. J. Wright, "Coordinate descent algorithms," *Math. Program.*, vol. 151, no. 1, pp. 3–34, 2015.

[55] S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[56] T. Goldstein, B. O'Donoghue, S. Setzer, and R. Baraniuk, "Fast alternating direction optimization methods," *SIAM J. Imag. Sci.*, vol. 7, no. 3, pp. 1588–1623, 2014.

[57] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," 2010, *arXiv:1009.5055*.

[58] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory. Appl.*, vol. 109, no. 3, 2001, Art. no. 475.

[59] D. Huang, C. Wang, and J. Lai, "Fast multi-view clustering via ensembles: Towards scalability, superiority, and simplicity," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 11, pp. 11388–11402, Nov. 2023.

[60] M. Chen, L. Huang, C. Wang, and D. Huang, "Multi-view clustering in latent embedding space," in *Proc. 34th AAAI Conf. Artif. Intell.*, 2020, pp. 3513–3520.

[61] Z. Kang et al., "Partition level multiview subspace clustering," *Neural Netw.*, vol. 122, pp. 279–288, 2020.

[62] R. Li, C. Zhang, Q. Hu, P. Zhu, and Z. Wang, "Flexible multi-view representation learning for subspace clustering," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, 2019, pp. 2916–2922.

[63] F. Nie, J. Li, and X. Li, "Parameter-free auto-weighted multiple graph learning: A framework for multiview clustering and semi-supervised classification," in *Proc. 25th Int. Joint Conf. Artif. Intell.*, 2016, pp. 1881–1887.

[64] X. Cai, F. Nie, and H. Huang, "Multi-view k-means clustering on big data," in *Proc. 23th Int. Joint Conf. Artif. Intell.*, 2013, pp. 2598–2604.

[65] Z. Zhang, L. Liu, F. Shen, H. T. Shen, and L. Shao, "Binary multi-view clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 7, pp. 1774–1782, Jul. 2019.

[66] B. Yang, X. Zhang, F. Nie, F. Wang, W. Yu, and R. Wang, "Fast multi-view clustering via nonnegative and orthogonal factorization," *IEEE Trans. Image Process.*, vol. 30, pp. 2575–2586, 2021.

[67] Z. Kang, Z. Lin, X. Zhu, and W. Xu, "Structured graph learning for scalable subspace clustering: From single view to multiview," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 8976–8986, Sep. 2022.

[68] Y. Liu et al., "Simple contrastive graph clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 10, pp. 13789–13800, Oct. 2024.

[69] C. Liu, Z. Wu, J. Wen, Y. Xu, and C. Huang, "Localized sparse incomplete multi-view clustering," *IEEE Trans. Multimedia*, vol. 25, pp. 5539–5551, 2023.

[70] Y. Liu et al., "Hard sample aware network for contrastive deep graph clustering," in *Proc. 37th AAAI Conf. Artif. Intell.*, 2023, pp. 8914–8922.

[71] W. Tu et al., "Attribute-missing graph clustering network," in *Proc. 38th AAAI Conf. Artif. Intell.*, 2024, pp. 15392–15401.

[72] J. Zhang et al., "TFMKC: Tuning-free multiple kernel clustering coupled with diverse partition fusion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 5, pp. 9592–9605, May 2025.

[73] J. Zhang et al., "Multiple kernel clustering with dual noise minimization," in *Proc. 30th ACM Int. Conf. Multimedia*, 2022, pp. 3440–3450.

[74] L. Li et al., "Local sample-weighted multiple kernel clustering with consensus discriminative graph," *IEEE Trans. Neural Netw.*, vol. 35, no. 2, pp. 1721–1734, Feb. 2024.

[75] X. Liu et al., "One pass late fusion multi-view clustering," in *Proc. 38th Int. Conf. Mach. Learn.*, 2021, pp. 6850–6859.

Liang Li (Graduate Student Member, IEEE) received the B.S. degree from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2018 and the Ph.D. degree in computer science from the National University of Defense Technology (NUDT), Changsha, China, in 2025, where he also received the M.S. degree in 2020. He is a visiting Ph.D. student at A∗STAR Centre for Frontier AI Research, Singapore from 2023 to 2025. His research interests include graph learning and AI4Science.

Yuangang Pan received the PhD degree in computer science from the University of Technology Sydney (UTS), Ultimo, NSW, Australia, in 2020. He is working as a research scientist with the A∗STAR Centre for Frontier AI Research, Singapore. He has authored or coauthored articles in various top journals, such as the *Journal of Machine Learning Research*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *IEEE Transactions on Neural Networks and Learning Systems*, *IEEE Transactions on Knowledge and Data Engineering*, and *ACM Transactions on Information Systems*. His research interests include deep clustering, deep generative learning, and robust ranking aggregation.

Junpu Zhang received the bachelor's degree from the Ocean University of China, Qingdao, China, in 2020. He is currently working toward the master degree with the National University of Defense Technology, Changsha, China. His current research interests include kernel learning, ensemble learning, and multi-view clustering.

Pei Zhang received the bachelor's degree from Yunnan University, China, in 2018, and the master's degree from the National University of Defense Technology (NUDT), China, in 2020, where she is currently working toward the PhD degree. Her current research interests include incomplete multi-view clustering, and deep clustering.

Jie Liu received the PhD degrees in computer science from the National University of Defense Technology (NUDT), Changsha, China. He is a professor with the College of Computer Science and Technology, National University of Defense Technology. His research interests include high performance computing and machine learning.

Xinwang Liu (Senior Member, IEEE) received the PhD degree from the National University of Defense Technology (NUDT), Changsha, China, in 2013. He is currently a full professor with the College of Computer Science and Technology, NUDT. His current research interests include kernel learning and unsupervised feature learning. He has published more than 100 peer-reviewed papers, such as *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *IEEE Transactions on Knowledge and Data Engineering*, *IEEE Transactions on Image Processing*, *IEEE Transactions on Neural Networks and Learning Systems*, *IEEE Transactions on Multimedia*, *IEEE Transactions on Information Forensics and Security*, ICML, NeurIPS, ICCV, CVPR, AAAI, and IJCAI. He serves as an associated editor of *IEEE Transactions on Neural Networks and Learning Systems* and *IEEE Transactions on Cybernetics*.

Kenli Li (Senior Member, IEEE) received the PhD degree from the Huazhong University of Science and Technology (HUST), China, in 2003. He is currently a full professor of computer science and technology with Hunan University and director with the National Supercomputing Center in Changsha. His research interests include parallel and distributed computing, high-performance computing, AI, cloud computing, and Big Data. He has published more than 600 research papers, such as *IEEE Transactions on Computers*, *IEEE Transactions on Parallel and Distributed Systems*, *IEEE Transactions on Cloud Computing*, *IEEE Transactions on Knowledge and Data Engineering*, DAC, AAAI, ICPP, etc.

Ivor W. Tsang (Fellow, IEEE) is the director with A∗STAR Centre for Frontier AI Research, Singapore. He is a professor of artificial intelligence with the University of Technology Sydney (UTS), Australia, and the research director with the Australian Artificial Intelligence Institute (AAII). His research interests include transfer learning, deep generative models, learning with weakly supervision, Big Data analytics for data with extremely high dimensions in features, samples and labels. He was the recipient of the ARC Future Fellowship for his outstanding research on Big Data analytics and large-scale machine learning, in 2013. In 2019, his JMLR article toward ultrahigh dimensional feature selection for Big Data was the recipient of the International Consortium of Chinese Mathematicians Best Paper Award. In 2020, he was recognized as the AI 2000 AAAI/IJCAI Most Influential Scholar in Australia for his outstanding contributions to the field between 2009 and 2019. He serves as the editorial board of the *Journal of Machine Learning Research*, *Machine Learning Journal*, *Journal of Artificial Intelligence Research*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *IEEE Transactions on Artificial Intelligence*, *IEEE Transactions on Big Data*, and *IEEE Transactions on Emerging Topics in Computational Intelligence*. He serves/served as an AC or senior AC for NeurIPS, ICML, AAAI, and IJCAI, and the steering committee of ACML.

Keqin Li (Fellow, IEEE) received the BS degree in computer science from Tsinghua University, in 1985, and the PhD degree in computer science from the University of Houston, in 1990. He is currently a SUNY distinguished professor with the State University of New York and a National distinguished professor with Hunan University (China). He has authored or co-authored more than 950 journal articles, book chapters, and refereed conference papers. He received several best paper awards from international conferences including PDPTA-1996, NAECON-1997, IPDPS-2000, ISPA-2016, NPC-2019, ISPA-2019, and CPSCom-2022. He holds nearly 70 patents announced or authorized by the Chinese National Intellectual Property Administration. He is among the world's top five most influential scientists in parallel and distributed computing in terms of single-year and career-long impacts based on a composite indicator of the Scopus citation database. He was a 2017 recipient of the Albert Nelson Marquis Lifetime Achievement Award for being listed in Marquis Who's Who in Science and Engineering, Who's Who in America, Who's Who in the World, and Who's Who in American Education for more than twenty consecutive years. He received the Distinguished Alumnus Award from the Computer Science Department at the University of Houston in 2018. He received the IEEE TCCLD Research Impact Award from the IEEE CS Technical Committee on Cloud Computing in 2022 and the IEEE TCSVC Research Innovation Award from the IEEE CS Technical Community on Services Computing in 2023. He is a member of the SUNY Distinguished Academy. He is an AAAS fellow, and an AAIA fellow. He is a member of Academia Europaea (Academician of the Academy of Europe).