SCALABLE MULTI-VIEW CLUSTERING VIA LATENT TENSOR LOW-RANK REPRESENTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Existing TLRR(Tensor low-rank representation) methods typically use the observed tensor itself as the dictionary to learn robust representations. However, when the observed data are limited and contaminated by noise, such a self-representation strategy is unable to adequately capture the underlying tensor subspace structure, resulting in degraded performance. To overcome this limitation, we propose Latent Tensor Low-Rank Representation (LT-LRR), a general tensor representation framework that augments the observed tensor dictionary with an unobserved latent component. We show that the resulting latent tensor self-representation admits a closed-form structural characterization under the t-product framework, which naturally yields low-rank constraints on both the representation tensor and the latent correction tensor. Building upon this formulation, we further develop ETRL-SMVC, a scalable multi-view clustering method that applies LT-LRR to anchor graph learning by replacing computationally expensive sample-level self-representation with compact dictionary-based tensor factorization over anchor graphs. Extensive experiments on benchmark datasets demonstrate that the proposed method consistently outperforms state-of-the-art competitors while maintaining favorable scalability.

1 INTRODUCTION

Multi-view clustering aims to partition unlabeled samples by jointly exploiting complementary information from multiple views, and has been widely applied to pattern recognition, computer vision, and data mining (Fang et al., 2023; Li et al., 2022; Liu et al., 2021). By integrating heterogeneous feature descriptions of the same samples, multi-view clustering can usually recover more reliable clustering structures than single-view clustering. Among existing clustering paradigms, low-rank representation (LRR) has shown strong ability to discover latent subspace structures (Liu et al., 2010; 2012). It is built on the self-expressiveness assumption that samples from the same latent subspace can be linearly represented by each other, and therefore directly uses the observed data matrix itself as the representation dictionary (Elhamifar & Vidal, 2013; Liu et al., 2010; Liu & Yan, 2011). This idea has also been extended to multi-view clustering (Zhang et al., 2015; Xie et al., 2018), where the data from different views are jointly modeled to exploit their shared subspace structure.

When extended to the multi-view setting, multi-view data $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{n\times d\times V}$ can be modeled as a third-order tensor. The motivation is to represent the observed multi-view tensor using itself as the dictionary and impose low-rankness on the representation coefficients, thereby capturing shared subspace structure across views. A typical tensor low-rank representation (TLRR) model is formulated as

$$\min_{\boldsymbol{\mathcal{Z}}} \|\boldsymbol{\mathcal{Z}}\|_{\circledast}, \quad \text{s.t.} \quad \boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{B}} * \boldsymbol{\mathcal{Z}}, \tag{1}$$

where $*$ denotes the t-product defined below in Definition 2.1, $\|\cdot\|_{\circledast}$ denotes the tensor nuclear norm induced by t-SVD (Gao et al., 2020), and $\boldsymbol{\mathcal{Z}}$ is the tensor representation coefficient. This formulation provides an effective way to explore high-order correlations and complementary information among different views and recover the shared subspace structure under a unified tensor representation (Xie et al., 2018; Li et al., 2023; Gao et al., 2020).

However, by setting the observed tensor data as dictionary, the good performance of TLRR is achieved under the following two assumptions: first, number of samples is sufficient; second, the ob-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

served data contain a portion of clean data. This assumption may be restrictive in practical scenarios and the robustness of TLRR may deteriorate.

To overcome the limitations of small sample size and improve robustness under noise, we propose *Latent Tensor Low-rank Representation* (LT-LRR), a general tensor representation technique inspired by latent low-rank representation (Liu & Yan, 2011). Different from classical TLRR, which only uses the observed tensor as the dictionary, LT-LRR augments the observed tensor dictionary with an unobserved latent tensor component:

$$\min_{\boldsymbol{\mathcal{Z}}} \|\boldsymbol{\mathcal{Z}}\|_{\circledast}, \quad \text{s.t.} \quad \boldsymbol{\mathcal{B}}_o = [\boldsymbol{\mathcal{B}}_o, \boldsymbol{\mathcal{B}}_h] * \boldsymbol{\mathcal{Z}}, \tag{2}$$

where $\boldsymbol{\mathcal{B}}_o$ and $\boldsymbol{\mathcal{B}}_h$ denote the observed and latent tensors, respectively. Classical TLRR requires the observed tensor to be represented solely by a potentially incomplete observed dictionary, whereas model (2) introduces a latent dictionary component to supply the missing subspace information. This latent component is subsequently expressed as a learnable low-rank correction to the observed tensor. This formulation is not specifically tied to anchor graph clustering; rather, it provides a general mechanism to compensate for the insufficient representation ability of the observed tensor dictionary. We further derive a closed-form structural characterization of the optimal solution under the t-product, showing that LT-LRR naturally induces two low-rank components: the self-representation tensor and the latent correction operator. This gives a theoretical basis for improving the robustness of tensor representation learning under insufficient or noisy observations.

Based on LT-LRR, we then apply the proposed technique to anchor graph-based scalable multi-view clustering. In this setting, directly solving a large sample-level tensor self-representation is computationally expensive. Therefore, we develop *Efficient Tensor Representation Learning for Scalable Multi-view Clustering* (ETRL-SMVC), which replaces large-scale tensor self-representation with compact dictionary-based tensor factorization on anchor graphs:

$$\boldsymbol{\mathcal{S}} \approx \boldsymbol{\mathcal{G}} * \boldsymbol{\mathcal{H}}^{\top} + \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{S}}, \tag{3}$$

where $\boldsymbol{\mathcal{S}}$ is the anchor-graph tensor, $\boldsymbol{\mathcal{G}}$ and $\boldsymbol{\mathcal{H}}$ denote the anchor-label and sample-label tensors, respectively, and $\boldsymbol{\mathcal{L}}$ is the latent correction tensor derived from LT-LRR. The tensor Schatten-p norms of these tensors are minimized as surrogates for their ranks, thereby promoting compact and view-consistent representations. The resulting model can be efficiently optimized in the Fourier domain and is suitable for large-scale clustering.

The main contributions are summarized as follows:

- We propose LT-LRR, a general latent tensor low-rank representation technique that improves classical TLRR by introducing an unobserved latent tensor dictionary.
- We derive a closed-form structural characterization of LT-LRR, showing that both the self-representation tensor and the latent correction operator possess low-rank structures.
- We apply LT-LRR to anchor graph-based scalable multi-view clustering and develop ETRL-SMVC, which learns compact anchor-label and sample-label tensors with favorable computational efficiency.

2 NOTATIONS

Definition 2.1 (t-product, Kilmer & Martin (2011))**.** Let $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{n_1 \times m \times n_3}$ and $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{m \times n_2 \times n_3}$. The t-product $\boldsymbol{\mathcal{A}} * \boldsymbol{\mathcal{B}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined as

$$\boldsymbol{\mathcal{A}} * \boldsymbol{\mathcal{B}} = \text{ifft}\big(\text{bdiag}(\overline{\mathbf{A}}\,\overline{\mathbf{B}}), [\,], 3\big), \tag{4}$$

where $\overline{\boldsymbol{\mathcal{A}}} = \text{fft}(\boldsymbol{\mathcal{A}}, [\,], 3)$ and $\overline{\boldsymbol{\mathcal{B}}} = \text{fft}(\boldsymbol{\mathcal{B}}, [\,], 3)$. Here, $\overline{\mathbf{A}} = \text{bdiag}(\overline{\boldsymbol{\mathcal{A}}})$ and $\overline{\mathbf{B}} = \text{bdiag}(\overline{\boldsymbol{\mathcal{B}}})$ are block diagonal matrices whose diagonal blocks are the frontal slices of $\overline{\boldsymbol{\mathcal{A}}}$ and $\overline{\boldsymbol{\mathcal{B}}}$, respectively.

Definition 2.2 (Tensor Schatten p-norm, Gao et al. (2021))**.** Given $\boldsymbol{\mathcal{H}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $h = \min(n_1, n_2)$, the tensor Schatten p-norm of $\boldsymbol{\mathcal{H}}$ is defined as

$$\|\boldsymbol{\mathcal{H}}\|_{\mathcal{S}_p} = \Big(\sum_{i=1}^{n_3} \|\overline{\boldsymbol{\mathcal{H}}}^{(i)}\|_{\mathcal{S}_p}^{p}\Big)^{\frac{1}{p}} = \Big(\sum_{i=1}^{n_3}\sum_{j=1}^{h} \sigma_j^{p}(\overline{\boldsymbol{\mathcal{H}}}^{(i)})\Big)^{\frac{1}{p}}, \tag{5}$$

where $\overline{\boldsymbol{\mathcal{H}}} = \text{fft}(\boldsymbol{\mathcal{H}}, [\,], 3)$, $\overline{\boldsymbol{\mathcal{H}}}^{(i)}$ is its i-th frontal slice, and $\sigma_j(\overline{\boldsymbol{\mathcal{H}}}^{(i)})$ denotes the j-th singular value of $\overline{\boldsymbol{\mathcal{H}}}^{(i)}$. Compared with the convex tensor nuclear norm (Semerci et al., 2014), the tensor Schatten p-norm with $0 < p \leq 1$ provides a tighter nonconvex approximation to the tensor rank and imposes stronger shrinkage on insignificant singular values, thus better preserving dominant low-rank structures.

3 METHODOLOGY

3.1 LATENT TENSOR LOW-RANK REPRESENTATION

Classical tensor low-rank representation directly uses the observed tensor as the representation dictionary. This strategy is effective when the observed tensor contains sufficiently rich and clean samples. However, when the observed data are insufficient, noisy, or unable to fully span the underlying tensor subspaces, the learned representation may become unstable. To address this limitation, we first introduce a general *Latent Tensor Low-rank Representation* (LT-LRR) technique by augmenting the observed tensor dictionary with an unobserved latent tensor component.

Let $\boldsymbol{\mathcal{B}}_o \in \mathbb{R}^{n\times m\times V}$ be the observed tensor and $\boldsymbol{\mathcal{B}}_h \in \mathbb{R}^{n\times r_h\times V}$ be the latent tensor, where $m+r_h < n$. The augmented tensor dictionary is denoted by $[\boldsymbol{\mathcal{B}}_o, \boldsymbol{\mathcal{B}}_h] \in \mathbb{R}^{n\times(m+r_h)\times V}$. Different from the standard tensor self-representation model, LT-LRR augments the observed tensor dictionary with $\boldsymbol{\mathcal{B}}_h$ to compensate for the insufficient representation ability of $\boldsymbol{\mathcal{B}}_o$. The resulting latent representation is characterized by the following theorem. Specifically, consider the latent tensor self-representation problem

$$\min_{\boldsymbol{\mathcal{Z}}} \quad \|\boldsymbol{\mathcal{Z}}\|_{\circledast}, \quad \textit{s.t.} \quad \boldsymbol{\mathcal{B}}_o = [\boldsymbol{\mathcal{B}}_o, \boldsymbol{\mathcal{B}}_h] * \boldsymbol{\mathcal{Z}} \tag{6}$$

Theorem 3.1. *For problem*

$$\min_{\boldsymbol{\mathcal{Z}}} \quad \|\boldsymbol{\mathcal{Z}}\|_{\circledast}, \quad \text{s.t.} \qquad \boldsymbol{\mathcal{B}}_o = [\boldsymbol{\mathcal{B}}_o, \boldsymbol{\mathcal{B}}_h] * \boldsymbol{\mathcal{Z}} \tag{7}$$

, it is equivalent to

$$\min_{\boldsymbol{\mathcal{Z}}_o} \quad \|\boldsymbol{\mathcal{Z}}_o\|_{\circledast}, \quad \text{s.t.} \quad \boldsymbol{\mathcal{B}}_o = \boldsymbol{\mathcal{B}}_o * \boldsymbol{\mathcal{Z}}_o + \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{B}}_o. \tag{8}$$

where $\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{S}} * \boldsymbol{\mathcal{V}}_h^\top * \boldsymbol{\mathcal{V}}_h * \boldsymbol{\mathcal{S}}^{-1} * \boldsymbol{\mathcal{U}}^\top$. $[\boldsymbol{\mathcal{B}}_o, \boldsymbol{\mathcal{B}}_h] = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{S}} * \boldsymbol{\mathcal{V}}^\top$ is the compact t-SVD, where $\boldsymbol{\mathcal{U}}$ and $\boldsymbol{\mathcal{V}}$ are orthogonal tensors and $\boldsymbol{\mathcal{S}}$ is f-diagonal. The partitions $\boldsymbol{\mathcal{V}} = [\boldsymbol{\mathcal{V}}_o; \boldsymbol{\mathcal{V}}_h]$ and $\boldsymbol{\mathcal{Z}} = [\boldsymbol{\mathcal{Z}}_o; \boldsymbol{\mathcal{Z}}_h]$ correspond to the observed and latent dictionaries, respectively.

The proof is provided in Appendix A.1. This formulation shows that the hidden information introduced by the latent tensor can be equivalently modeled by a latent correction tensor $\boldsymbol{\mathcal{L}}$.

We further analyze the rank properties of $\boldsymbol{\mathcal{Z}}$ and $\boldsymbol{\mathcal{L}}$. Assume that $\boldsymbol{\mathcal{B}}_o$ and $\boldsymbol{\mathcal{B}}_h$ are drawn from t tensor subspaces, and the tensor rank of the i-th subspace is r_i for $i = 1, \ldots, t$. From the optimal solution derived in Appendix A.1 and the tensor-rank inequality, we have

$$\text{rank}(\boldsymbol{\mathcal{Z}}) \leq \text{rank}(\boldsymbol{\mathcal{V}}_o) \leq \text{rank}\left([\boldsymbol{\mathcal{B}}_o, \boldsymbol{\mathcal{B}}_h]\right) \leq \sum_{i=1}^{t} r_i = r. \tag{9}$$

Similarly, from the definition of $\boldsymbol{\mathcal{L}}$ in Theorem 3.1, we obtain

$$\text{rank}(\boldsymbol{\mathcal{L}}) \leq \text{rank}(\boldsymbol{\mathcal{V}}_h) \leq \text{rank}\left([\boldsymbol{\mathcal{B}}_o, \boldsymbol{\mathcal{B}}_h]\right) \leq r. \tag{10}$$

Thus, both the self-representation tensor $\boldsymbol{\mathcal{Z}}$ and the latent correction tensor $\boldsymbol{\mathcal{L}}$ have low-rank structures. This motivates the following latent tensor low-rank representation model:

$$\min_{\boldsymbol{\mathcal{Z}},\boldsymbol{\mathcal{L}}} \quad \text{rank}(\boldsymbol{\mathcal{Z}}) + \text{rank}(\boldsymbol{\mathcal{L}}), \quad \textit{s.t.} \qquad \boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{B}} * \boldsymbol{\mathcal{Z}} + \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{B}}. \tag{11}$$

Since direct tensor-rank minimization is difficult, we adopt the tensor Schatten p-norm as a nonconvex low-rank surrogate and obtain

$$\min_{\boldsymbol{\mathcal{Z}},\boldsymbol{\mathcal{L}}} \quad \|\boldsymbol{\mathcal{Z}}\|_{\circledS p}^p + \|\boldsymbol{\mathcal{L}}\|_{\circledS p}^p, \quad \textit{s.t.} \qquad \boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{B}} * \boldsymbol{\mathcal{Z}} + \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{B}}. \tag{12}$$

In many real-world scenarios, data inevitably contain noise, therefore, it is reasonable to assume that the model equation 12 still holds approximately when the data are corrupted. Thus, we revise the model equation 12 as the following general optimization problem.

$$\min_{\boldsymbol{\mathcal{Z}},\boldsymbol{\mathcal{L}},\boldsymbol{\mathcal{E}}} \quad \|\boldsymbol{\mathcal{Z}}\|_{\circledS p}^{p} + \|\boldsymbol{\mathcal{L}}\|_{\circledS p}^{p} + \|\boldsymbol{\mathcal{E}}\|_{F}^{2}, \quad s.t. \qquad \boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{B}} * \boldsymbol{\mathcal{Z}} + \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{B}} + \boldsymbol{\mathcal{E}}, \tag{13}$$

3.2 Application to Anchor Graph-based Multi-view Clustering

Anchor graph clustering based on tensor low-rank representation is an effective method for handling large-scale data and has become a current research topic. A key aspect of anchor graph clustering is that the number of anchors is small and the constructed anchor graph is inevitably corrupted. Therefore, anchor graph clustering is a representative application scenario of our proposed LT-LRR.

Given third-order tensor anchor graph $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{m\times n\times V}$, replacing $\boldsymbol{\mathcal{B}}$ in model equation 13, where m, n, and V denote the number of anchors, samples, and views, respectively. It can be found that it still requires learning an $n \times n$ sample-level coefficient tensor, which is computationally expensive for large-scale clustering. To improve scalability, we replace tensor self-representation with a compact dictionary-based tensor factorization:

$$\boldsymbol{\mathcal{S}} \approx \boldsymbol{\mathcal{G}} * \boldsymbol{\mathcal{H}}^{\top}, \tag{14}$$

where $\boldsymbol{\mathcal{G}} \in \mathbb{R}^{m\times c\times V}$ and $\boldsymbol{\mathcal{H}} \in \mathbb{R}^{n\times c\times V}$ denote the anchor-label tensor and sample-label tensor, respectively, and c is the number of clusters.

According to the rank inequality of tensor products, the rank of the factorized representation is bounded by

$$\operatorname{rank}\left(\boldsymbol{\mathcal{G}} * \boldsymbol{\mathcal{H}}^{\top}\right) \leq \min\left\{\operatorname{rank}(\boldsymbol{\mathcal{G}}), \operatorname{rank}(\boldsymbol{\mathcal{H}})\right\}. \tag{15}$$

Since the ranks of $\boldsymbol{\mathcal{G}}$ and $\boldsymbol{\mathcal{H}}$ are bounded by the cluster number c, the factorized representation naturally has a low-rank structure. Therefore, the explicit self-representation tensor $\boldsymbol{\mathcal{Z}}$ in equation 13 can be avoided, leading to

$$\min_{\boldsymbol{\mathcal{G}},\boldsymbol{\mathcal{H}},\boldsymbol{\mathcal{L}},\boldsymbol{\mathcal{E}}} \quad \|\boldsymbol{\mathcal{L}}\|_{\circledS p}^{p} + \|\boldsymbol{\mathcal{E}}\|_{F}^{2} + \boldsymbol{\mathcal{R}}(\boldsymbol{\mathcal{G}}, \boldsymbol{\mathcal{H}}), \quad s.t. \qquad \boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{G}} * \boldsymbol{\mathcal{H}}^{\top} + \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{S}} + \boldsymbol{\mathcal{E}}. \tag{16}$$

To enhance interpretability and avoid post-processing, we exploit the probabilistic property of anchor graphs and represent the anchor-graph tensor through the product of the anchor-label tensor and the sample-label tensor (Xue et al., 2025; Wang et al., 2025). Meanwhile, tensor Schatten p-norm regularization is imposed on $\boldsymbol{\mathcal{G}}$ and $\boldsymbol{\mathcal{H}}$ to capture high-order cross-view consistency and obtain compact clustering representations. Finally, the proposed objective is formulated as

$$\begin{aligned} \min_{\boldsymbol{\mathcal{H}},\boldsymbol{\mathcal{G}},\boldsymbol{\mathcal{L}}} \quad & \left\|\boldsymbol{\mathcal{S}} - \boldsymbol{\mathcal{G}} * \boldsymbol{\mathcal{H}}^{\top} - \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{S}}\right\|_{F}^{2} + \lambda_1 \|\boldsymbol{\mathcal{H}}\|_{\circledS p}^{p} + \lambda_2 \|\boldsymbol{\mathcal{G}}\|_{\circledS p}^{p} + \lambda_3 \|\boldsymbol{\mathcal{L}}\|_{\circledS p}^{p} \\ s.t. \quad & \boldsymbol{\mathcal{H}} \geq 0, \quad \boldsymbol{\mathcal{H}}^{\top} * \boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{I}}, \quad \mathbf{G}^{(v)} \geq 0, \quad \mathbf{G}^{(v)}\mathbf{1} = \mathbf{1} \end{aligned} \tag{17}$$

where $0 < p \leq 1$, and λ_1, λ_2, and λ_3 are trade-off parameters.

3.3 Optimization

Utilizing the Augmented Lagrange Multiplier (ALM) method, we introduce four auxiliary variables $\boldsymbol{\mathcal{Q}}$, $\boldsymbol{\mathcal{J}}$, $\boldsymbol{\mathcal{F}}$, and $\boldsymbol{\mathcal{M}}$, setting $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{Q}}$, $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{J}}$, $\boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{F}}$, and $\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{M}}$ respectively, with the condition $\boldsymbol{\mathcal{Q}} \geqslant 0$. Consequently, the objective function (17) can be reformulated as follows:

$$\begin{aligned} \min \ & \left\|\boldsymbol{\mathcal{S}} - \boldsymbol{\mathcal{G}} * \boldsymbol{\mathcal{H}}^{\top} - \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{S}}\right\|_{F}^{2} + \lambda_1 \|\boldsymbol{\mathcal{J}}\|_{\circledS p}^{p} + \lambda_2 \|\boldsymbol{\mathcal{F}}\|_{\circledS p}^{p} + \lambda_3 \|\boldsymbol{\mathcal{M}}\|_{\circledS p}^{p} + \frac{\mu}{2}\left\|\boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{Q}} + \frac{\boldsymbol{\mathcal{Y}_1}}{\mu}\right\|_{F}^{2} \\ & + \frac{\rho}{2}\left\|\boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{J}} + \frac{\boldsymbol{\mathcal{Y}_2}}{\rho}\right\|_{F}^{2} + \frac{\sigma}{2}\left\|\boldsymbol{\mathcal{G}} - \boldsymbol{\mathcal{F}} + \frac{\boldsymbol{\mathcal{Y}_3}}{\sigma}\right\|_{F}^{2} + \frac{\beta}{2}\left\|\boldsymbol{\mathcal{L}} - \boldsymbol{\mathcal{M}} + \frac{\boldsymbol{\mathcal{Y}_4}}{\beta}\right\|_{F}^{2} \\ s.t. \ & \boldsymbol{\mathcal{Q}} \geqslant 0, \boldsymbol{\mathcal{H}}^{\top} * \boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{I}}, \mathbf{G}^{(v)} \geqslant 0, \mathbf{G}^{(v)} \cdot \mathbf{1} = \mathbf{1}, \end{aligned} \tag{18}$$

Here, $\boldsymbol{\mathcal{Y}_1}, \boldsymbol{\mathcal{Y}_2}, \boldsymbol{\mathcal{Y}_3}$ and $\boldsymbol{\mathcal{Y}_4}$ denote the Lagrange multipliers, while μ, ρ, σ and β represent the penalty parameters. To address this optimization challenge (18), we employ an alternating optimization approach.

•Solving for $\mathcal{G}$ with fixed $\mathcal{L}, \mathcal{M}, \mathcal{Q}, \mathcal{H}, \mathcal{J}$, and $\mathcal{F}$. The objective function (18) transforms into:

$$\begin{aligned} \min \ & \left\| \boldsymbol{\mathcal{S}} - \boldsymbol{\mathcal{G}} * \boldsymbol{\mathcal{H}}^\top - \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{S}} \right\|_F^2 + \frac{\sigma}{2} \left\| \boldsymbol{\mathcal{G}} - \boldsymbol{\mathcal{F}} + \frac{\boldsymbol{\mathcal{Y}}_3}{\sigma} \right\|_F^2, \\ & \textit{s.t.} \quad \mathbf{G}^{(v)} \geqslant 0, \mathbf{G}^{(v)} \cdot \mathbf{1} = \mathbf{1} \end{aligned} \tag{19}$$

Furthermore, (19) can be transformed into the Fourier domain and simplified as:

$$\begin{aligned} & \min (1 + \frac{\sigma}{2}) \{ \mathrm{tr}[(\overline{\boldsymbol{\mathcal{G}}}^{(v)})^\top \overline{\boldsymbol{\mathcal{G}}}^{(v)}] - 2\mathrm{tr}[(\overline{\boldsymbol{\mathcal{G}}}^{(v)})^\top \overline{\boldsymbol{\mathcal{B}}}_1^{(v)}] \}, \\ & \textit{s.t.} \quad \mathbf{G}^{(v)} \geqslant 0, \mathbf{G}^{(v)} \cdot \mathbf{1} = \mathbf{1}, \end{aligned} \tag{20}$$

where $\overline{\boldsymbol{\mathcal{B}}}_1^{(v)} = \{2[\overline{\boldsymbol{\mathcal{S}}}^{(v)} - \overline{\boldsymbol{\mathcal{L}}}^{(v)} \overline{\boldsymbol{\mathcal{S}}}^{(v)}]^{\mathrm{H}} \overline{\boldsymbol{\mathcal{H}}}^{(v)} + \sigma(\overline{\boldsymbol{\mathcal{F}}}^{(v)} - \overline{\boldsymbol{\mathcal{Y}}}_3^{(v)})/\sigma\}/(2+\sigma)$.

By returning (20) to the time domain and simplifying it, we have:

$$\min_{\mathbf{G}^{(v)} \geqslant 0, \mathbf{G}^{(v)} \cdot \mathbf{1} = \mathbf{1}} (1 + \frac{\sigma}{2}) \left\| \boldsymbol{\mathcal{G}} - \boldsymbol{\mathcal{B}}_1 \right\|_F^2 \tag{21}$$

We can derive the solution of (21) as $\mathbf{G}_i^{(v)} = (\mathbf{B_1}_i^{(v)} + \gamma \mathbf{1})$ according to Nie et al. (2016), where $\mathbf{G}_i^{(v)}$ represents pulling the matrix $\mathbf{G}^{(v)}$ into a vector and γ is a Lagrange multiplier.

•Solve for $\mathcal{H}$ with fixed $\mathcal{L}, \mathcal{M}, \mathcal{Q}, \mathcal{G}, \mathcal{J}$ and $\mathcal{F}$. (18) then transforms into:

$$\min_{\boldsymbol{\mathcal{H}}^\top * \boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{I}}} \left\| \boldsymbol{\mathcal{S}} - \boldsymbol{\mathcal{G}} * \boldsymbol{\mathcal{H}}^\top - \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{S}} \right\|_F^2 + \frac{\mu}{2} \left\| \boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{Q}} + \frac{\boldsymbol{\mathcal{Y}}_1}{\mu} \right\|_F^2 + \frac{\rho}{2} \left\| \boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{J}} + \frac{\boldsymbol{\mathcal{Y}_2}}{\rho} \right\|_F^2, \tag{22}$$

By transforming (22) into the frequency domain and simplifying it, we have:

$$\max_{(\overline{\boldsymbol{\mathcal{H}}}^{(v)})^{\mathrm{H}} \overline{\boldsymbol{\mathcal{H}}}^{(v)} = \mathbf{I}} \mathrm{tr} \left((\overline{\boldsymbol{\mathcal{H}}}^{(v)})^{\mathrm{H}} \overline{\boldsymbol{\mathcal{B}}}_2^{(v)} \right) \tag{23}$$

where $\overline{\boldsymbol{\mathcal{B}}}_2^{(v)} = 2[(\overline{\boldsymbol{\mathcal{S}}}^{(v)})^{\mathrm{H}} - (\overline{\boldsymbol{\mathcal{S}}}^{(v)})^{\mathrm{H}} (\overline{\boldsymbol{\mathcal{L}}}^{(v)})^{\mathrm{H}}] \overline{\boldsymbol{\mathcal{G}}}^{(v)} + \mu \overline{\boldsymbol{\mathcal{W}}}_2^{(v)} + \rho \overline{\boldsymbol{\mathcal{W}}}_3^{(v)}$, $\overline{\boldsymbol{\mathcal{W}}}_2^{(v)} = \overline{\boldsymbol{\mathcal{Q}}}^{(v)} - \overline{\boldsymbol{\mathcal{Y}}}_1^{(v)}/\mu$, and $\overline{\boldsymbol{\mathcal{W}}}_3^{(v)} = \overline{\boldsymbol{\mathcal{J}}}^{(v)} - \overline{\boldsymbol{\mathcal{Y}}}_2^{(v)}/\rho$.

To address (23), we introduce the following theorems:

Theorem 3.2. *Given two matrices $\mathbf{P} \in \mathbb{R}^{n_1 \times n_2}$, $\mathbf{Q} \in \mathbb{R}^{n_1 \times n_2}$,we have*

$$\mathrm{tr}(\mathbf{P}^{\mathrm{H}} \mathbf{Q}) \leq \|\mathbf{P}\|_F \|\mathbf{Q}\|_F \tag{24}$$

Equality holds if and only if $\mathbf{Q}$ is a multiple of $\mathbf{P}$.

The proof is provided in Appendix A.2.

Theorem 3.3. *Given two matrices $\mathbf{P} \in \mathbb{R}^{n_1 \times n_2}$, $\mathbf{Q} \in \mathbb{R}^{n_1 \times n_2}$, then the optimal solution of*

$$\arg\max_{\mathbf{P}^{\mathrm{H}} \mathbf{P} = \mathbf{I}_{n_2}} \mathrm{tr}(\mathbf{P}^{\mathrm{H}} \mathbf{Q}) \tag{25}$$

is $\mathbf{P} = \mathbf{U}\mathbf{W}^{\mathrm{H}}$. Where, $\mathbf{Q} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^{\mathrm{H}}$ is the singular value decomposition (SVD) of $\mathbf{Q}$, $\mathbf{U} \in \mathbb{R}^{n_1 \times n_2}$, $\mathbf{W} \in \mathbb{R}^{n_2 \times n_2}$ is an orthogonal matrix, $\mathbf{\Sigma} \in \mathbf{R}^{n_2 \times n_2}$, $\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{I}_{n_2}$.

The proof is provided in Appendix A.3.

By Theorem 3.3, we can derive the solution of (23) as:

$$\overline{\boldsymbol{\mathcal{H}}}^{(v)} = \overline{\mathbf{\Lambda}}_2^{(v)} (\overline{\boldsymbol{V}}_2^{(v)})^{\mathrm{H}}, \tag{26}$$

where $\overline{\mathbf{\Lambda}}_2^{(v)}$ and $\overline{\boldsymbol{V}}_2^{(v)}$ are determined through SVD of $\overline{\boldsymbol{\mathcal{B}}}_2^{(v)}$.

•**Solve $\mathcal{Q}$ with fixed $\mathcal{L}, \mathcal{M}, \mathcal{G}, \mathcal{H}, \mathcal{J}$ and $\mathcal{F}$.** (18) becomes:

$$\min_{\mathcal{Q}\geqslant 0} \frac{\mu}{2}\left\|\mathcal{H}-\mathcal{Q}+\frac{\mathcal{Y}_1}{\mu}\right\|_F^2 \tag{27}$$

Transforming this into the frequency domain and rearranging the variables, we obtain:

$$\min_{\overline{\mathcal{Q}}^{(v)}\geqslant 0} \sum\nolimits_{v=1}^{V} \frac{\mu}{2}\|\overline{\mathcal{Q}}^{(v)}-(\overline{\mathcal{H}}^{(v)}+\frac{\overline{\mathcal{Y}}_1^{(v)}}{\mu})\|_F^2 \tag{28}$$

The solution for Equation (28) is given by:

$$\overline{\mathcal{Q}}^{(v)}=(\overline{\mathcal{H}}^{(v)}+\frac{\overline{\mathcal{Y}}_1^{(v)}}{\mu})_+ \tag{29}$$

• **Solve for $\mathcal{J}$ with fixed $\mathcal{L}, \mathcal{M}, \mathcal{Q}, \mathcal{H}, \mathcal{G}$, and $\mathcal{F}$.** The equation (18) is reformulated as:

$$\min \lambda_1\|\mathcal{J}\|_{\circledS p}^p+\frac{\rho}{2}\left\|\mathcal{H}-\mathcal{J}+\frac{\mathcal{Y}_2}{\rho}\right\|_F^2, \tag{30}$$

where, after completing the square with respect to $\mathcal{J}$, we obtain

$$\mathcal{J}^*=\arg\min \frac{1}{2}\left\|\mathcal{H}+\frac{\mathcal{Y}_2}{\rho}-\mathcal{J}\right\|_F^2+\frac{\lambda_1}{\rho}\|\mathcal{J}\|_{\circledS p}^p, \tag{31}$$

yielding a closed-form solution as outlined in Lemma A.2 (Gao et al., 2021) in Appendix A.4.

Thus, the solution for (31) is expressed as:

$$\mathcal{J}^*=\Gamma_{\frac{\lambda_1}{\rho}}(\mathcal{H}+\frac{\mathcal{Y}_2}{\rho}). \tag{32}$$

•**Solve for $\mathcal{F}$ with fixed $\mathcal{L}, \mathcal{M}, \mathcal{Q}, \mathcal{H}, \mathcal{J}$ and $\mathcal{G}$.** The objective function (18) simplifies to:

$$\min \lambda_2\|\mathcal{F}\|_{\circledS p}^p+\frac{\sigma}{2}\left\|\mathcal{G}-\mathcal{F}+\frac{\mathcal{Y}_3}{\sigma}\right\|_F^2, \tag{33}$$

Applying Lemma A.2, the solution to (33) is given by:

$$\mathcal{F}^*=\Gamma_{\frac{\lambda_2}{\sigma}}(\mathcal{G}+\frac{\mathcal{Y}_3}{\sigma}). \tag{34}$$

•**Solve for $\mathcal{M}$ with fixed $\mathcal{L}, \mathcal{Q}, \mathcal{H}, \mathcal{J}, \mathcal{G}$ and $\mathcal{F}$.** The objective function (18) simplifies to:

$$\min \lambda_3\|\mathcal{M}\|_{\circledS p}^p+\frac{\beta}{2}\left\|\mathcal{L}-\mathcal{M}+\frac{\mathcal{Y}_4}{\beta}\right\|_F^2, \tag{35}$$

Applying Lemma A.2, the solution to (35) is given by:

$$\mathcal{M}^*=\Gamma_{\frac{\lambda_3}{\beta}}(\mathcal{L}+\frac{\mathcal{Y}_4}{\beta}). \tag{36}$$

•**Solving for $\mathcal{L}$ with fixed $\mathcal{M}, \mathcal{G}, \mathcal{Q}, \mathcal{H}, \mathcal{J}$, and $\mathcal{F}$.** The objective function (18) transforms into:

$$\min \left\|\mathcal{S}-\mathcal{G}*\mathcal{H}^\top-\mathcal{L}*\mathcal{S}\right\|_F^2+\frac{\beta}{2}\left\|\mathcal{L}-\mathcal{M}+\frac{\mathcal{Y}_4}{\beta}\right\|_F^2 \tag{37}$$

Further, (37) can be reformulated in the frequency domain as:

$$\min \sum\nolimits_{v=1}^{V}\left\|\overline{\mathcal{S}}^{(v)}-\overline{\mathcal{G}}^{(v)}(\overline{\mathcal{H}}^{(v)})^{\mathrm{H}}-\overline{\mathcal{L}}^{(v)}\overline{\mathcal{S}}^{(v)}\right\|_F^2+\frac{\beta}{2}\left\|\overline{\mathcal{L}}^{(v)}-\overline{\mathcal{M}}^{(v)}+\frac{\overline{\mathcal{Y}}_4^{(v)}}{\beta}\right\|_F^2, \tag{38}$$

where $\overline{\mathcal{L}}=\mathrm{fft}(\mathcal{L},[\,],3)$, and other variables are transformed similarly. Taking the derivative of formula (38) with respect to $\overline{\mathcal{L}}^{(v)}$, and making the derivative zero, we get:

$$\overline{\mathcal{L}}^{(v)}=\{2[[\overline{\mathcal{S}}^{(v)}-\overline{\mathcal{G}}^{(v)}(\overline{\mathcal{H}}^{(v)})^{\mathrm{H}}](\overline{\mathcal{S}}^{(v)})^{\mathrm{H}}]+\beta\overline{\mathcal{M}}^{(v)}-\overline{\mathcal{Y}}_4^{(v)}\}[2\overline{\mathcal{S}}^{(v)}(\overline{\mathcal{S}}^{(v)})^{\mathrm{H}}+\beta\boldsymbol{I}]^{-1} \tag{39}$$

Algorithm 1 Efficient Tensor Representation Learning for Scalable Multi-view Clustering

Require: Data matrices $\{\mathbf{X}^{(v)}\}_{v=1}^{V} \in \mathbb{R}^{N\times d_v}$; anchors numbers m; cluster number K.
Ensure: Cluster labels $\mathbf{Y}$ of each data points.

1: **Initialize**: $\mu = \rho = \sigma = \beta = 10^{-5}$, $\eta = 1.2$, $\boldsymbol{\mathcal{Y}}_1 = 0$, $\boldsymbol{\mathcal{Y}}_2 = 0$, $\boldsymbol{\mathcal{Y}}_3 = 0$, $\boldsymbol{\mathcal{Y}}_4 = 0$ and $\overline{\boldsymbol{\mathcal{H}}}^{(v)}$ is identity matrix. Compute graph matrix $\mathbf{S}^{(v)}$ of each views;
2: **while** not condition **do**
3: Update $\boldsymbol{\mathcal{G}}$ by solving (21);
4: Update $\overline{\boldsymbol{\mathcal{H}}}^{(v)}$ by solving (26);
5: Update $\overline{\boldsymbol{\mathcal{Q}}}^{(v)}$ by solving (29);
6: Update $\boldsymbol{\mathcal{J}}$ by using (32);
7: Update $\boldsymbol{\mathcal{F}}$ by using (34);
8: Update $\boldsymbol{\mathcal{M}}$ by solving (36);
9: Update $\overline{\boldsymbol{\mathcal{L}}}^{(v)}$ by solving (39);
10: Update $\boldsymbol{\mathcal{Y}}_1$, $\boldsymbol{\mathcal{Y}}_2$, $\boldsymbol{\mathcal{Y}}_3$, $\boldsymbol{\mathcal{Y}}_4$, μ, ρ, σ and β.
11: **end while**
12: Calculate the K clusters by using $\mathbf{Q} = \sum_{v=1}^{V} \mathbf{Q}^{(v)}/V$;
13: **return** Clustering result.

Table 1: Experimental performance results. Optimal results are in bold.

Datasets	MSRC			NUS			AWA			Reuters		
Methods	ACC	NMI	Purity	ACC	NMI	Purity	ACC	NMI	Purity	ACC	NMI	Purity
LMVSC	0.814	0.717	0.814	0.133	0.110	0.167	0.075	0.086	0.089	0.589	0.335	0.615
PLCMF	0.634	0.567	0.660	0.120	0.098	0.208	0.082	0.085	0.097	0.397	0.138	0.422
FPMVS-CAG	0.786	0.686	0.786	0.174	0.128	0.210	0.091	0.109	0.096	0.526	0.323	0.603
FMCNOF	0.713	0.648	0.714	0.133	0.040	0.155	0.050	0.022	0.051	0.343	0.125	0.358
SFMC	0.810	0.721	0.810	0.179	0.191	0.255	0.039	0.003	0.040	0.602	0.354	0.604
AWMVC	0.843	0.727	0.843	0.126	0.122	0.229	0.094	0.113	0.115	0.536	0.303	0.576
FSMSC	0.876	0.753	0.876	0.143	0.131	0.237	0.097	0.113	0.118	0.517	0.307	0.558
TBGL	1.000	1.000	1.000	0.274	0.202	0.281	0.050	0.023	0.053	0.795	0.660	0.795
Orth-NTF	0.990	0.978	0.990	0.355	0.544	0.465	0.109	0.115	0.134	0.694	0.686	0.809
LMVC-MFCG	0.852	0.738	0.852	0.183	0.135	0.219	0.097	0.118	0.104	0.579	0.313	0.586
EMVC-NTLC	0.857	0.811	0.857	0.453	0.519	0.483	0.234	0.375	0.236	0.727	0.570	0.727
AGFS-OMVC	1.000	1.000	1.000	0.300	0.467	0.390	0.240	0.520	0.278	0.646	0.593	0.771
ALPC	0.881	0.775	0.881	0.143	0.149	0.256	0.081	0.086	0.103	0.481	0.268	0.564
Dual MoE MVKM	0.733	0.655	0.767	0.117	0.103	0.210	0.070	0.071	0.084	0.501	0.302	0.560
LMTC	0.881	0.808	0.881	0.122	0.101	0.212	0.109	0.120	0.126	0.459	0.268	0.528
Ours	**1.000**	**1.000**	**1.000**	**0.566**	**0.742**	**0.710**	**0.539**	**0.770**	**0.606**	**0.818**	**0.739**	**0.846**

3.4 Computational Complexity and Scalability Analysis

We analyze the computational complexity of the proposed method. Let $d^{(v)}$ denote the feature dimension of the v-th view, V the number of views, n the number of samples, m the number of anchors, and c the number of clusters. Let $h = \sum_{v=1}^{V} d^{(v)}$. The cost of constructing the anchor graph tensor $\boldsymbol{\mathcal{S}}$ is $\mathcal{O}(nmhV + nmV\log m)$. For the optimization procedure, the main computational costs come from updating $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{H}}$, which require $\mathcal{O}(nm^2 + nmc + m^3)$ and $\mathcal{O}(nVc^2 + mcV\log(mV))$, respectively. The remaining auxiliary variables are updated with lower costs. Since m, c, and V are usually much smaller than n, the overall complexity is approximately linear with respect to n. Therefore, by using a compact anchor graph tensor instead of an $n \times n$ sample-level graph, the proposed method avoids quadratic or cubic complexity in the number of samples and is suitable for large-scale multi-view clustering.

4 Experiments

4.1 Clustering Performance

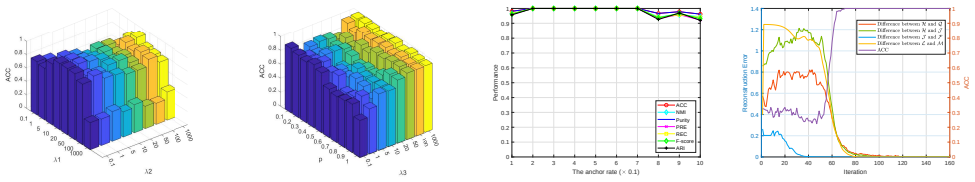
The detailed information on the datasets and compared algorithms is provided in Appendix A.5. The hyperparameter settings are provided in the Appendix. Table 1 presents the comparative result-

Table 2: Ablation experiments on Reuters, NUS, and AWA.

Method	Reuters			NUS			AWA		
	ACC	NMI	Purity	ACC	NMI	Purity	ACC	NMI	Purity
w/o $\|\mathcal{H}\|^p_{\circledS p} + \|\mathcal{G}\|^p_{\circledS p} + \|\mathcal{L}\|^p_{\circledS p}$	0.412	0.158	0.489	0.113	0.092	0.120	0.066	0.067	0.082
w/o $\|\mathcal{H}\|^p_{\circledS p} + \|\mathcal{L}\|^p_{\circledS p}$	0.436	0.128	0.482	0.131	0.077	0.192	0.066	0.069	0.081
w/o $\|\mathcal{G}\|^p_{\circledS p} + \|\mathcal{L}\|^p_{\circledS p}$	0.715	0.666	0.793	0.447	0.532	0.572	0.455	0.669	0.495
w/o $\|\mathcal{L}\|^p_{\circledS p}$	0.796	0.731	0.796	0.442	0.614	0.580	0.510	0.704	0.560
Full Model	**0.818**	**0.739**	**0.846**	**0.566**	**0.742**	**0.710**	**0.539**	**0.770**	**0.606**

s on four datasets. Compared with Nonnegative Matrix Factorization (NMF)-based methods, our dictionary-based tensor factorization method achieves superior performance. NMF-based methods usually decompose each view separately and fuse the clustering indicators only at the decision level, which may ignore high-order cross-view correlations. In contrast, our method learns a compact tensor dictionary representation under the t-product framework, thereby capturing correlations within the anchor graph tensor and producing more consistent clustering results.

Compared with traditional graph denoising methods, such as TBGL, our method further improves performance by introducing latent tensor representation to compensate for missing structural information in the observed anchor graph. Since limited anchors and view-specific construction errors often make anchor graphs noisy and incomplete, directly performing low-rank self-representation may be insufficient. By introducing a latent correction term and coupling it with low-rank regularization, our method can better exploit cross-view complementary information and recover a more reliable consensus clustering structure.



(a) Parameter analysis of λ_1 and λ_2. (b) Parameter analysis of λ_3 and p. (c) Parameter analysis of m. (d) Convergence curves.

Figure 1: Parameter analysis and convergence curves on the MSRC dataset.

4.2 Parameter Analysis

Impact of the anchor rate m: Experimental results in Figure 1 show that the clustering performance is largely insensitive to changes in the anchor rate. This is because, within a certain range, the selected anchors can sufficiently cover the structure of the sample space without significantly affecting the overall similarity distribution. Therefore, a moderate number of anchors is sufficient to ensure stable clustering performance, and there is no need to rely heavily on a specific value.

Impact of p: Results in Figure 1b show that Smaller values of p help the model more effectively exploit the low-rank structure. This is because smaller p values impose stronger penalties on low-rank patterns, allowing the model to more accurately capture shared information across views and thereby improve clustering performance.

Impact of λ_1 and λ_2: Results in Figure 1a show that when λ_1 or λ_2 is too large or too small, clustering performance deteriorates. This is because these parameters control the strength of the tensor Schatten-p norm regularization on the low-rank structure: excessively large values may over-smooth the representation and suppress the expression of true structures, while excessively small values fail to fully exploit the low-rank prior, weakening the integration of complementary information across views.

Impact of λ_3: Results in Figure 1bshow that appropriate settings of λ_3 can significantly improve clustering performance, effectively validating the role of the tensor low-rank self-representation term in optimizing the anchor graph structure and capturing high-order hidden effects.

4.3 Convergence Analysis

To optimize the objective function (17), we introduce four auxiliary variables: $\mathcal{Q}$, $\mathcal{J}$, $\mathcal{F}$, and $\mathcal{M}$. The convergence of the model is evaluated by analyzing the differences between the original variables and their corresponding auxiliary variables: $\mathcal{H}$ and $\mathcal{Q}$, $\mathcal{H}$ and $\mathcal{J}$, $\mathcal{G}$ and $\mathcal{F}$, as well as $\mathcal{L}$ and $\mathcal{M}$. Figure 1d illustrates the variations between these variable pairs. Typically, around the 80th iteration, the differences stabilize, indicating that the model has reached a steady state. Meanwhile, clustering metrics such as ACC also converge, empirically validating the effectiveness of our method.

4.4 Ablation Experiments

As shown in Table 2, we conducted an ablation study on the tensor $\mathcal{L}$ regularization term. The results indicate that incorporating the low-rank self-representation constraint on the anchor graph can substantially improve its quality. This is because constructing the anchor graph inevitably introduces noise, and the low-rank self-representation constraint effectively suppresses such noise, thereby further enhancing clustering performance.

4.5 Denoising Visualization

Taking synthetic data as an example, we visualize the anchor graphs learned by the proposed method. In Fig. 2, the first row shows the anchor graphs with added random noise, the second row shows the anchor graphs obtained without the low-rank self-representation constraint, and the third row shows the anchor graphs learned using our method. The results indicate that by incorporating the proposed low-rank tensor self-representation, the learned anchor graph exhibits 7 clearly connected components. This demonstrates that our method can effectively uncover the underlying cluster structure in noisy multi-view data while simultaneously performing denoising.

5 Conclusion

In this work, we propose Latent Tensor Low-rank Representation, a general tensor representation technique that augments the observed tensor dictionary with an unobserved latent component and provides a closed-form structural characterization for latent correction. Based on LT-LRR, we further develop Efficient Tensor Representation Learning for Scalable Multi-view Clustering (ETRL-SMVC), which applies the proposed technique to anchor graph-based multi-view clustering. ETRL-SMVC replaces tensor self-representation with dictionary-based tensor factorization, thereby avoiding large-scale self-representation tensors and directly learning anchor labels and sample labels. Extensive experiments validate the effectiveness and superiority of the proposed method.

AI use statement

We used generative AI tools solely for language polishing to improve the grammar, clarity, and readability of this manuscript. All AI-assisted edits were reviewed and verified by the authors. We take full responsibility for the final content of this work.

Reproducibility statement

To support the reproducibility of this work, Sections 2 and 3 present the model formulations, constraints, and block-wise optimization updates of LT-LRR and ETRL-SMVC. Algorithm 1 summarizes the overall optimization procedure. Section 4 introduces the benchmark datasets and evaluation metrics used in the experiments. The hyperparameter settings are additionally provided in the appendix.

REFERENCES

Yawei Chen, Huibing Wang, Jinjia Peng, and Yang Wang. Anchor learning with potential cluster constraints for multi-view clustering. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(15):15939–15947, Apr. 2025.

Zhe Chen, Xiaojun Wu, Tianyang Xu, and Josef Kittler. Fast self-guided multi-view subspace clustering. *IEEE Trans. Image Process.*, 32:6514–6525, 2023.

Liang Du, Yukai Shi, Yan Chen, Feijiang Li, Peng Zhou, Lei Duan, and Yuhua Qian. A dual mixture-of-experts framework for multi-view k-means clustering with view balance and regional sparsity. *Information Fusion*, 125:103449, 2026. ISSN 1566-2535.

Ehsan Elhamifar and René Vidal. Sparse subspace clustering: Algorithm, theory, and applications. *IEEE transactions on pattern analysis and machine intelligence*, 35(11):2765–2781, 2013.

Uno Fang, Man Li, Jianxin Li, Longxiang Gao, Tao Jia, and Yanchun Zhang. A comprehensive survey on multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 35 (12):12350–12368, 2023.

Quanxue Gao, Wei Xia, Zhizhen Wan, Deyan Xie, and Pu Zhang. Tensor-svd based graph learning for multi-view subspace clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 3930–3937, 2020.

Quanxue Gao, Pu Zhang, Wei Xia, Deyan Xie, Xinbo Gao, and Dacheng Tao. Enhanced tensor rpca and its application. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(6): 2133–2140, 2021.

Zhao Kang, Wangtao Zhou, Zhitong Zhao, Junming Shao, Meng Han, and Zenglin Xu. Large-scale multi-view subspace clustering in linear time. In *Proceedings of the AAAI conference on Artificial Intelligence*, pp. 4412–4419, 2020.

Misha E Kilmer and Carla D Martin. Factorization strategies for third-order tensors. *Linear Algebra and its Applications*, 435(3):641–658, 2011.

Jing Li, Quanxue Gao, QIANQIAN WANG, Ming Yang, and Wei Xia. Orthogonal non-negative tensor factorization based multi-view clustering. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Jing Li, Qianqian Wang, Ming Yang, Quanxue Gao, and Xinbo Gao. Efficient anchor graph factorization for multi-view clustering. *IEEE Trans. Multim.*, 26:5834–5845, 2024.

Xuelong Li, Han Zhang, Rong Wang, and Feiping Nie. Multiview clustering: A scalable and parameter-free bipartite graph fusion method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1):330–344, 2022.

Guangcan Liu and Shuicheng Yan. Latent low-rank representation for subspace segmentation and feature extraction. In *2011 international conference on computer vision*, pp. 1615–1622. IEEE, 2011.

Guangcan Liu, Zhouchen Lin, and Yong Yu. Robust subspace segmentation by low-rank representation. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 663–670, 2010.

Guangcan Liu, Zhouchen Lin, Shuicheng Yan, Ju Sun, Yong Yu, and Yi Ma. Robust recovery of subspace structures by low-rank representation. *IEEE transactions on pattern analysis and machine intelligence*, 35(1):171–184, 2012.

Jiyuan Liu, Xinwang Liu, Yuexiang Yang, Li Liu, Siqi Wang, Weixuan Liang, and Jiangyong Shi. One-pass multi-view clustering for large-scale data. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 12324–12333, 2021.

Jiyuan Liu, Xinwang Liu, Chuankun Li, Xinhang Wan, Hao Tan, Yi Zhang, Weixuan Liang, Qian Qu, Yu Feng, Renxiang Guan, and Ke Liang. Large-scale multi-view tensor clustering with implicit linear kernels. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20727–20736, 2025.

Feiping Nie, Xiaoqian Wang, Michael I. Jordan, and Heng Huang. The constrained laplacian rank algorithm for graph-based clustering. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, AAAI'16, pp. 19691976. AAAI Press, 2016.

Oguz Semerci, Ning Hao, Misha E Kilmer, and Eric L Miller. Tensor-based formulation and nuclear norm regularization for multienergy computed tomography. *IEEE Transactions on Image Processing*, 23(4):1678–1693, 2014.

Xinhang Wan, Xinwang Liu, Jiyuan Liu, Siwei Wang, Yi Wen, Weixuan Liang, En Zhu, Zhe Liu, and Lu Zhou. Auto-weighted multi-view clustering for large-scale data. In *AAAI Conference on Artificial Intelligence*, pp. 10078–10086, 2023.

Di Wang, Songwei Han, Quan Wang, Lihuo He, Yu min Tian, and Xinbo Gao. Pseudo-label guided collective matrix factorization for multiview clustering. *IEEE Transactions on Cybernetics*, 52: 8681–8691, 2021a.

Rui Wang, Yuxuan Liu, Guangyu Yang, Quanxue Gao, and Cheng Deng. Bi-orthogonal nonnegative tensor tri-factorization for tensorized label learning. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pp. 2304–2312, 2025.

Siwei Wang, Xinwang Liu, Xinzhong Zhu, Pei Zhang, Yi Zhang, Feng Gao, and En Zhu. Fast parameter-free multi-view subspace clustering with consensus anchor guidance. *IEEE Transactions on Image Processing*, 31:556–568, 2021b.

Wei Xia, Quanxue Gao, Qianqian Wang, Xinbo Gao, Chris Ding, and Dacheng Tao. Tensorized bipartite graph learning for multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):5187–5202, 2023.

Yuan Xie, Dacheng Tao, Wensheng Zhang, Yan Liu, Lei Zhang, and Yanyun Qu. On unifying multiview self-representations for clustering by tensor multi-rank minimization. *Int. J. Comput. Vis.*, 126(11):1157–1179, 2018.

Xingyu Xue, Wenhui Zhao, Quanxue Gao, Ming Yang, and Cheng Deng. Image clustering with transition probabilities learning. *IEEE Transactions on Image Processing*, 2025.

Ben Yang, Xuetao Zhang, Feiping Nie, Fei Wang, Weizhong Yu, and Rong Wang. Fast multi-view clustering via nonnegative and orthogonal factorization. *IEEE transactions on image processing*, 30:2575–2586, 2021.

Zengbiao Yang, Yihua Tan, and Tao Yang. Large-scale multi-view clustering via matrix factorization of consensus graph. *Pattern Recognition*, 155:110716, 2024. ISSN 0031-3203.

Changqing Zhang, Huazhu Fu, Si Liu, Guangcan Liu, and Xiaochun Cao. Low-rank tensor constrained multiview subspace clustering. In *Proceedings of the IEEE international conference on computer vision*, pp. 1582–1590, 2015.

Wenhui Zhao, Qin Li, Huafu Xu, Quanxue Gao, Qianqian Wang, and Xinbo Gao. Anchor graph-based feature selection for one-step multi-view clustering. *IEEE Trans. Multim.*, 26:7413–7425, 2024.

A Appendix

A.1 Proof of Theorem 3.1

Proof.

Lemma A.1. *Suppose the tensor linear equation $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{A}} * \boldsymbol{\mathcal{Z}}$, where $\boldsymbol{\mathcal{Z}}$ is the unknown variable, admits a unique feasible solution. Then this solution can be defined as*

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{A}}^{-1} * \boldsymbol{\mathcal{X}}, \tag{40}$$

where $\boldsymbol{\mathcal{A}}^{-1}$ is defined as the pseudoinverse of $\boldsymbol{\mathcal{A}}$.

According to Lemma A.1, we can reformulate equation 6 as follows:

$$\boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{B}} * \boldsymbol{\mathcal{V}}_o^\top = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{B}} * \boldsymbol{\mathcal{V}}^\top * \boldsymbol{\mathcal{Z}}. \tag{41}$$

By the orthogonality of $\boldsymbol{\mathcal{U}}$ and the invertibility of $\boldsymbol{\mathcal{B}}$, we get

$$\boldsymbol{\mathcal{V}}_o^\top = \boldsymbol{\mathcal{V}}^\top * \boldsymbol{\mathcal{Z}}. \tag{42}$$

Consequently, we obtain an equivalent formulation of problem equation 6

$$\begin{aligned} &\min_{\boldsymbol{\mathcal{Z}}} \|\boldsymbol{\mathcal{Z}}\|_* \\ &\text{s.t. } \boldsymbol{\mathcal{V}}_o^\top = \boldsymbol{\mathcal{V}}^\top * \boldsymbol{\mathcal{Z}}. \end{aligned} \tag{43}$$

Based on Lemma equation A.1, the above problem admits the following closed-form solution:

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{V}} * \boldsymbol{\mathcal{V}}_o^\top = [\boldsymbol{\mathcal{V}}_o * \boldsymbol{\mathcal{V}}_o^\top; \boldsymbol{\mathcal{V}}_h * \boldsymbol{\mathcal{V}}_o^\top]. \tag{44}$$

To obtain the equivalent correction form, write the t-SVD and its conformable partition as

$$[\boldsymbol{\mathcal{B}}_o, \boldsymbol{\mathcal{B}}_h] = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{S}} * \boldsymbol{\mathcal{V}}^\top, \qquad \boldsymbol{\mathcal{V}} = [\boldsymbol{\mathcal{V}}_o; \boldsymbol{\mathcal{V}}_h],$$

which gives

$$\boldsymbol{\mathcal{B}}_o = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{S}} * \boldsymbol{\mathcal{V}}_o^\top, \qquad \boldsymbol{\mathcal{B}}_h = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{S}} * \boldsymbol{\mathcal{V}}_h^\top.$$

Partitioning $\boldsymbol{\mathcal{Z}}^*$ accordingly yields

$$\boldsymbol{\mathcal{B}}_o = \boldsymbol{\mathcal{B}}_o * \boldsymbol{\mathcal{Z}}_o^* + \boldsymbol{\mathcal{B}}_h * \boldsymbol{\mathcal{Z}}_h^*.$$

Since

$$\boldsymbol{\mathcal{Z}}_h^* = \boldsymbol{\mathcal{V}}_h * \boldsymbol{\mathcal{V}}_o^\top, \qquad \boldsymbol{\mathcal{V}}_o^\top = \boldsymbol{\mathcal{S}}^{-1} * \boldsymbol{\mathcal{U}}^\top * \boldsymbol{\mathcal{B}}_o,$$

the latent term becomes

$$\boldsymbol{\mathcal{B}}_h * \boldsymbol{\mathcal{Z}}_h^* = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{S}} * \boldsymbol{\mathcal{V}}_h^\top * \boldsymbol{\mathcal{V}}_h * \boldsymbol{\mathcal{S}}^{-1} * \boldsymbol{\mathcal{U}}^\top * \boldsymbol{\mathcal{B}}_o = \boldsymbol{\mathcal{L}} * \boldsymbol{\mathcal{B}}_o.$$

Substituting this identity into the partitioned representation establishes equation 8. This completes the proof. □

A.2 Proof of Theorem 3.2

Proof. $vec(\mathbf{P})$ and $vec(\mathbf{Q})$ denote the column vectors that concatenate all the columns of matrix $\mathbf{P}$ and $\mathbf{Q}$, respectively. Then,

$$\mathrm{tr}(\mathbf{P}^{\mathrm{H}}\mathbf{Q}) = (vec(\mathbf{P}))^{\mathrm{H}} vec(\mathbf{Q}). \tag{45}$$

Leveraging CauchySchwarz inequality, we have

$$(vec(\mathbf{P}))^{\mathrm{H}} vec(\mathbf{Q}) \leq \|vec(\mathbf{P})\|_2 \|vec(\mathbf{Q})\|_2 = \|\mathbf{P}\|_F \|\mathbf{Q}\|_F. \tag{46}$$

Comparing inequalities (45) and (46) yields

$$\mathrm{tr}(\mathbf{P}^{\mathrm{H}}\mathbf{Q}) \leq \|\mathbf{P}\|_F \|\mathbf{Q}\|_F \tag{47}$$

□

A.3 Proof of Theorem 3.3

Proof.

$$\mathrm{tr}(\mathbf{P}^{\mathrm{H}}\mathbf{Q}) = \mathrm{tr}(\mathbf{P}^{\mathrm{H}}\mathbf{U}\boldsymbol{\Sigma}\mathbf{W}^{\mathrm{H}}) = \mathrm{tr}(\mathbf{U}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}\mathbf{W}^{\mathrm{H}}\mathbf{P}^{\mathrm{H}}). \tag{48}$$

Using Theorem 3.2, we get

$$\mathrm{tr}(\mathbf{P}^{\mathrm{H}}\mathbf{Q}) \leq \|\mathbf{U}\boldsymbol{\Sigma}^{1/2}\|_F \|\boldsymbol{\Sigma}^{1/2}\mathbf{W}^{\mathrm{H}}\mathbf{P}^{\mathrm{H}}\|_F = \|\boldsymbol{\Sigma}^{1/2}\|_F \|\boldsymbol{\Sigma}^{1/2}\|_F. \tag{49}$$

Equality holds only when

$$\boldsymbol{\Sigma}^{1/2}\mathbf{U}^{\mathrm{H}} = \boldsymbol{\Sigma}^{1/2}\mathbf{W}^{\mathrm{H}}\mathbf{P}^{\mathrm{H}} \tag{50}$$

Thus, $\mathbf{P} = \mathbf{U}\mathbf{W}^{\mathrm{H}}$. □

Table 3: Multi-view datasets used in our experiments.

Dataset	# Samples	# Views	# Classes	#Feature
MSRC	210	5	7	24, 576, 512, 256, 254
Reuters	18758	5	6	21531, 24892, 34251, 15506, 11547
NUS	30000	5	31	64, 225, 144, 73, 128
AWA	30475	6	50	2688, 2000, 252, 2000, 2000, 2000

A.4 LEMMA A.2

Lemma A.2. *Given the input $\mathcal{Z} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with a t-SVD $\mathcal{Z} = \mathcal{U} * \mathcal{S} * \mathcal{V}^\top$, the optimal solution for*

$$\min_{\mathcal{X}} \tfrac{1}{2} \|\mathcal{X} - \mathcal{Z}\|_F^2 + \tau \|\mathcal{X}\|_{\circledS p}^p \tag{51}$$

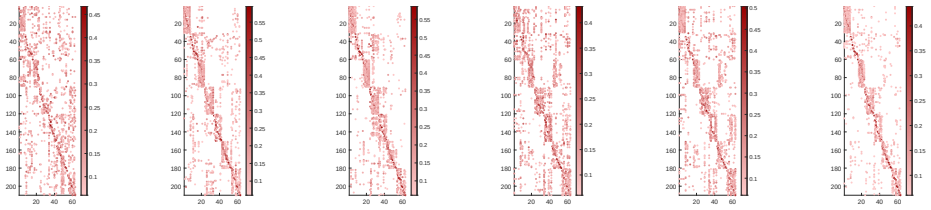
is $\mathcal{X}^* = \Gamma_\tau(\mathcal{Z}) = \mathcal{U} * \text{ifft}(P_\tau(\overline{\mathcal{Z}})) * \mathcal{V}^\top$, where τ is the input parameter, $\tau > 0$, $P_\tau(\overline{\mathcal{Z}}) = diag(z_1, z_2, ..., z_l)$ is an f-diagonal 3rd-order tensor. Its diagonal elements $z_1, z_2, \ldots, z_l$ ($l = \min(n_1, n_2)$) can be determined using the GST algorithm described by Gao et al. (2021).

A.5 DATASETS AND COMPARED METHODS

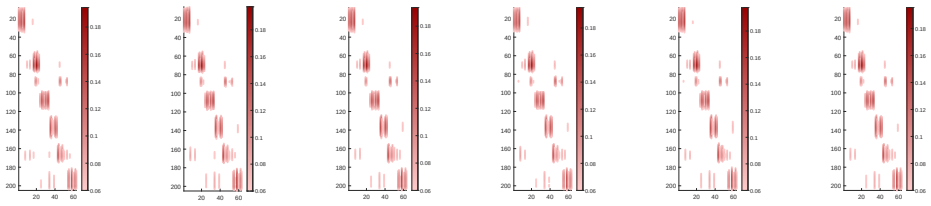
In this section, we conduct comprehensive comparative experiments to evaluate the proposed methods. The effectiveness of clustering is measured using three key metrics: Accuracy (ACC), Normalized Mutual Information (NMI), and Purity. Higher values for these metrics correspond to better clustering performance. Datasets employed in our experiments are detailed in Table 3. We compare our approach against the following state-of-the-art methods: LMVSC (Kang et al., 2020), PLCM-F (Wang et al., 2021a), FPMVS-CAG (Wang et al., 2021b), FMCNOF (Yang et al., 2021), SFMC (Li et al., 2022), AWMVC (Wan et al., 2023), FSMSC (Chen et al., 2023), TBGL (Xia et al., 2023), Orth-NTF (Li et al., 2023), LMVC-MFCG (Yang et al., 2024), EMVC-NTLC (Li et al., 2024), AGFS-OMVC (Zhao et al., 2024), ALPC (Chen et al., 2025), Dual MoE MVKM (Du et al., 2026), and LMTC (Liu et al., 2025). Experiments using the MSRC dataset was conducted on a laptop equipped with an Intel Core i5-8300H CPU and 16 GB RAM, utilizing Matlab R2021b. In contrast, NUS, AWA and Reuters datasets were processed on a standard Windows 10 server, featuring dual Intel(R) Xeon(R) Gold 6230 CPUs at 2.10 GHz and 128 GB RAM, with MATLAB R2020b.

A.6 HYPERPARAMETER SETTINGS

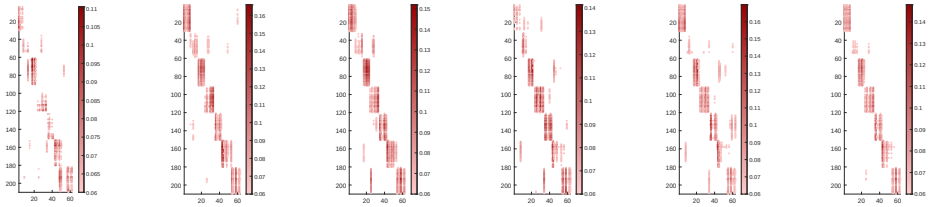
The hyperparameters are set separately for different datasets. Specifically, for Reuters, we set $m = 0.001$, $p = 0.7$, $\lambda_1 = 6$, $\lambda_2 = 3$, and $\lambda_3 = 90$. For AWA, we set $m = 0.002$, $p = 0.1$, $\lambda_1 = 1$, $\lambda_2 = 0.1$, and $\lambda_3 = 1000$. For MSRC, we set $m = 0.3$, $p = 1$, $\lambda_1 = 10$, $\lambda_2 = 0.001$, and $\lambda_3 = 1000$. For NUS, we set $m = 0.005$, $p = 0.1$, $\lambda_1 = 10$, $\lambda_2 = 0.001$, and $\lambda_3 = 0.1$.



(a) Input noisy anchor graphs from five views and average anchor graph



(b) The learned five anchor graphs without $\|\mathcal{L}\|_{\textcircled{Sp}}^{p}$ and fused anchor graph



(c) The obtained five anchor graphs by our method and fused view

Figure 2: Denoising visualization on synthetic dataset.