Cross-View Diversity Embedded Consensus Learning for Multi-View Clustering

Chong Peng[1] , **Kai Zhang**[1] , **Yongyong Chen**[2,*] , **Chenglizhao Chen**[3,*] , **Qiang Cheng**[4]
[1]College of Computer Science and Technology, Qingdao University, China
[2]School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), China
[3]College of Computer Science and Technology, China University of Petroleum (East China), China
[4]Department of Computer Science, University of Kentucky, USA
pchong1991@163.com, zk127585@163.com, yongyongchen.cn@gmail.com, cclz123@163.com, qiang.cheng@uky.edu

Abstract

Multi-view clustering (MVC) has garnered significant attention in recent studies. In this paper, we propose a novel MVC method, named CCL-MVC. The novel method constructs a cross-order neighbor tensor of multi-view data to recover a low-rank essential tensor, which preserves noise-free, comprehensive, and complementary cross-order relationships among the samples. Furthermore, it constructs a consensus representation matrix by fusing the low-rank essential tensor with auto-adjusted cross-view diversity embedding, fully exploiting both consensus and discriminative information of the data. An effective optimization algorithm is developed, which is theoretically guaranteed to converge. Extensive experimental results confirm the effectiveness of the proposed method.

1 Introduction

As one of the fundamental tasks of machine learning, clustering has achieved notable successes in diverse domains, such as image segmentation [Ng *et al.*, 2023], face clustering [Wu *et al.*, 2023], community detection [Park *et al.*, 2022], etc. Among the extensive clustering algorithms, the subspace clustering methods have drawn significant attention in recent years due to their effectiveness and elegant theory [Liu *et al.*, 2010; Peng *et al.*, 2017; Elhamifar and Vidal, 2009].

In real-world scenarios, the advancement of technology has led to an increase in the availability of data from multiple sources, leading to the natural emergence of multi-view data [Zhang *et al.*, 2020]. Consequently, multi-view data are prevalent and contain a wealth of useful, discriminative, and complementary information across multiple perspectives, making them essential for enhancing clustering capability [Xia *et al.*, 2021]. However, traditional clustering methods mainly focus on single-view data and are not suitable for analyzing multi-view data, as each view can provide a unique and valuable insight. Thus, multi-view clustering (MVC) has garnered significant attention in recent years due to its versatility and effectiveness in various domains [Xia *et al.*, 2022; Wen *et al.*, 2023; Brbić and Kopriva, 2018].

Due to the great success of single-view subspace clustering (SVSC) methods, they have been extensively developed for MVC [Fu *et al.*, 2023; Zhang *et al.*, 2020; Yang *et al.*, 2019]. Most of these methods follow a pipeline similar to the SVSC with a two-step strategy. In particular, they first construct an affinity matrix, on which the standard spectral clustering (SPC) [Ng *et al.*, 2001] is then performed. Since the SPC is standard, the step of affinity matrix construction has been the main focus of MVC [Chen *et al.*, 2021; Zhang *et al.*, 2020].

According to the way of affinity matrix construction, existing MVC methods can be mainly categorized into two types [Wu *et al.*, 2019]. The first type is developed from the SPC-based subspace clustering methods [Zhang *et al.*, 2015; Li *et al.*, 2019; Zhang *et al.*, 2020], while the second is developed from the graph-based clustering, which constructs affinity matrix using similarity matrix [Wang *et al.*, 2020; Xia *et al.*, 2014; Zhan *et al.*, 2018; Kumar *et al.*, 2011] and is the main focus of our paper. It has been revealed that the Markov random walk has a close connection with the SPC [Zhou *et al.*, 2005; Shi and Malik, 2000], based on which a number of methods have been developed [Zhou and Burges, 2007; Xia *et al.*, 2014; Wu *et al.*, 2019]. For example, the transductive inference approach constructs a Markov transition probability matrix (TPM) in each view, and combines the TPMs using the Markov mixture [Zhou and Burges, 2007]. The RMSC constructs the TPMs of different views and learns a common low-rank stochastic matrix to alleviate the noise effects [Xia *et al.*, 2014]. Different from them, the ETLMSC learns a low-rank essential tensor using the low-rank tensor recovery (LTR) approach from the TPMs [Wu *et al.*, 2019].

The LTR approach has been quite popular and successful in recent development of MVC [Xie *et al.*, 2018; Zhang *et al.*, 2015; Wu *et al.*, 2019], which has been extensively attempted in above mentioned methods [Zhou and Burges, 2007; Xia *et al.*, 2014; Wu *et al.*, 2019; Xie *et al.*, 2018]. However, these methods suffer from some key issues, which may severely degrade the learning performance. First, by adopting the tensor nuclear norm (TNN), the LTR approach may suffer from inaccurate approximation issue [Yu and Yang, 2023]. Second, the high-order neighbor information of the data is rarely considered, which has been revealed essential in graph learning [Tang *et al.*, 2015]. Third, simultaneous consensus and diversity learning is rarely considered on the low-rank essential tensor, which omits some essential properties. To this end, we develop a new method for MVC to address the above issues.

We summarize the key contributions of this paper as follows: 1) We recover a low-rank essential tensor from a cross-order neighbor graph tensor derived from multi-view data, which preserves comprehensive and complementary information of the data. 2) By embedding an automatically adjusted weighting vector with the learned noise-free cross-order neighbor graphs, our method explicitly preserves cross-view diversity information of the multi-view data to learn a consensus affinity matrix. 3) We design an efficient optimization algorithm, which is theoretically guaranteed to converge under some mild conditions. 4) Extensive experimental results confirm the effectiveness of the proposed method and its superiority to baselines.

2 Related Work

Given multi-view data $\{\mathbf{X}^{(v)}\}_{v=1}^{V}$, with superscript $(\cdot)^{(v)}$ denoting the v-th view of multi-view data or the v-th frontal slice of a tensor, $\mathbf{X}^{(v)} \in \mathbb{R}^{d_v \times n}$ being the v-th view samples, d_v and n being the number of features and samples in the v-th view, and V being the number of views, respectively, the ETLMSC recovers a low-rank essential tensor $\mathcal{Z} \in \mathbb{R}^{n\times n\times V}$ from a TPM-based tensor $\mathcal{P} \in \mathbb{R}^{n\times n\times V}$ constructed from multi-view data by [Wu *et al.*, 2019]:

$$\min_{\mathcal{Z},\mathcal{E}\in\mathbb{R}^{n\times n\times V}} \|\mathcal{Z}\|_{\circledast} + \lambda\|\mathcal{E}\|_{2,1} \quad s.t. \quad \mathcal{P} = \mathcal{Z} + \mathcal{E}, \tag{1}$$

where $\mathcal{E}$ denotes the noise, $\|\cdot\|_{\circledast}$ is the t-SVD-based TNN, and $\|\cdot\|_{2,1}$ denotes the tensor $\ell_{2,1}$ norm defined as the sum of ℓ_2-norm of mode-3 fibers. Here, the tensor $\mathcal{P}$ is constructed by slice as $\mathcal{P}^{(v)} = (\mathbb{D}_s(\mathcal{S}^{(v)}))^{-1}\mathcal{S}^{(v)}$, where $\mathbb{D}_s(\cdot)$ returns a diagonal matrix with diagonal elements being the sum of rows of the input matrix, and $\mathcal{S} \in \mathbb{R}^{n\times n\times V}$ is a similarity tensor constructed by slice with $\mathcal{S}^{(v)}$ being a pair-wise similarity matrix in the v-th view. Eq. (1) essentially follows the robust tensor PCA [Lu *et al.*, 2016], where the key difference between them is that they aim at different learning tasks with different types of data as input.

3 The Proposed Method

Due to the great success of low-rank essential recovery approach, in this paper, we follow the framework of Eq. (1) and develop a new method with some more desired properties for enhanced learning performance. As has been widely revealed in literature, the convex approach to rank or sparsity approximation is not accurate, and nonconvex approach may provide a more accurate approximation and enhance the structural learning capability of the model [Peng *et al.*, 2022a; Peng *et al.*, 2022b]. Therefore, in a way similar to the t-SVD-based TNN [Wu *et al.*, 2019] and tensor Schatten norm [Xia *et al.*, 2022], and inspired by [Peng *et al.*, 2022a; Peng *et al.*, 2020], we expand the log-based matrix rank and sparsity approximation to log-based tensor rank approximation (LTRA) and sparsity approximation (LTSA) as $\|\mathcal{Z}\|_{®} = \sum_{v=1}^{V}\sum_{i=1}^{n}\log(1+\boldsymbol{\sigma}_i(\hat{\mathcal{Z}}^{(v)}))$, and $\|\mathcal{E}\|_{Ⓢ} = \sum_{j=1}^{n}\log(1+(\sum_{i=1}^{n}\sum_{v=1}^{V}\mathcal{E}_{ijv}^2)^{1/2})$, respectively, where the LTRA is defined in the frequency domain, $\hat{\mathcal{Z}} = \mathbf{fft}(\mathcal{Z},[\,],3) \in \mathbb{R}^{n\times n\times V}$ with $\mathbf{fft}(\cdot)$ being the fast Fourier transform (FFT) along the third dimension, $\boldsymbol{\sigma}_i(\cdot)$ is the i-th largest singular value of the input matrix, and the LTSA is defined in a lateral slice-wise manner that enhances the cross-view sparsity and helps strengthen the connections among different views. Then, we may develop Eq. (1) into the following model:

$$\min_{\mathcal{Z},\mathcal{E}\in\mathbb{R}^{n\times n\times V}} \|\mathcal{Z}\|_{®} + \lambda\|\mathcal{E}\|_{Ⓢ} \quad s.t. \quad \mathcal{P} = \mathcal{Z} + \mathcal{E}, \tag{2}$$

in which the nonconvex approach provides more accurate approximations and helps preserve more useful information to enhance the learning capability [Peng *et al.*, 2022a].

In Eq. (2), $\mathcal{P}$ reveals the probability of a one-step random walk from one example to another in each view, which provides soft neighbor relationships of the samples. In practice, such relationships may be insufficient to explicitly measure the structure of the data, because samples may have latent higher-order neighbor relationships that are not directly preserved in a first-order graph [Kang *et al.*, 2022]. Intuitively, a higher-order relationship can be measured by the probability of a multi-step random walk from one example to another [Tang *et al.*, 2015]. Therefore, we may incorporate the higher-order neighbor relationships into LTR by defining a high-order TPM for multi-step random walks in each view as: $(\mathcal{P}^{(v)})^k = \mathcal{P}^{(v)}\cdot\mathcal{P}^{(v)}\cdot\;\cdots\;\cdot\mathcal{P}^{(v)}$, with k being the step of random walks or the order of neighbor relationships. To fully exploit the cross-order neighbor relationships, we define the following fine-grained probability tensor $\mathcal{P}_K \in \mathbb{R}^{n\times n\times V}$ by slice as: $\mathcal{P}_K^{(v)} = \sum_{k=1}^{K}\frac{(\mathcal{P}^{(v)})^k + ((\mathcal{P}^{(v)})^k)^T}{2}$, where the transpose ensures the symmetry of the neighbor relationships that is natural and essentially desired for neighbor relationships. Moreover, to exploit local structure of the data and filter the redundancy of neighbor relationships, we construct a local similarity tensor $\mathcal{S}$, where in each view we keep N neighbors in the neighbor graph while setting the others to zero for each sample. Consequently, we construct $\mathcal{P}_K$ with local structure of the data and our model becomes

$$\min_{\mathcal{Z},\mathcal{E}} \|\mathcal{Z}\|_{®} + \lambda\|\mathcal{E}\|_{Ⓢ} \quad s.t. \quad \mathcal{P}_K = \mathcal{Z} + \mathcal{E}. \tag{3}$$

In particular, Eq. (3) is treated as a K-th order model. In literature, an intuitive and common way of clustering is to fuse $\mathcal{Z}$ across different views to obtain a representation matrix by $\mathrm{Z}_0 = \frac{1}{V}\sum_{v=1}^{V}\mathcal{Z}^{(v)}$, on which the final clustering algorithm is performed for the final clustering result [Wu *et al.*, 2019]. However, the two-step strategy omits the close connection between the learning and fusion tasks, and the straight fusion omits the discriminative information embedded in different views, which might be insufficient to fully exploit structural information of the data [Pan *et al.*, 2023].

To build a close connection between learning and fusion, we embed the fusion task into our model, leading to:

$$\begin{aligned} &\min_{\mathcal{Z},\mathcal{E},w} \|\mathcal{Z}\|_{®} + \lambda\|\mathcal{E}\|_{Ⓢ} + \alpha\sum\nolimits_v \|\mathrm{Z}_0 - w_v\mathcal{Z}^{(v)}\|_F^2 \\ &s.t.\ \mathcal{P}_K = \mathcal{Z} + \mathcal{E}, w_v \geq 0, \sum\nolimits_v w_v = 1, \end{aligned} \tag{4}$$

where Z_0 is a consensus affinity matrix that fuses cross-view neighbor graphs, and $w \in \mathbb{R}^V$ is a weighting vector that is automatically adjusted to balance the discrinimative information embedded across different views. To further enhance the cross-view diversity embedded in the weighting vector w, and thus in the consensus affinity matrix Z_0, it is desired to prevent the weights from being simultaneously large if the TPMs of their corresponding views are highly similar [Kang *et al.*, 2023], which leads to the diversity embedding defined as $\min_{w\geq 0,\sum_i w_i=1}\sum_{i,j} w_i w_j \mathrm{Tr}((\mathcal{P}_K^{(i)})^T\mathcal{P}_K^{(j)})$. Although the embedding is straightforward and easy to solve, $\mathcal{P}_K$ may be prone to noise effects and thus the embedding might be inaccurate. As an alternative, we may embed the diversity of w with $\mathcal{Z}$, which efficiently alleviates the noise effects while is more challenging to solve. Therefore, to embed the cross-view noise-free diversity, we finally develop Eq. (4) as:

$$\begin{aligned}&\min_{\mathcal{P}_K=\mathcal{Z}+\mathcal{E},w\geq 0,\sum_i w_i=1}\|\mathcal{Z}\|_{\circledR}+\lambda\|\mathcal{E}\|_{\circledS}\\&+\alpha\sum_{v=1}^{V}\|\mathrm{Z}_0-w_v\mathcal{Z}^{(v)}\|_F^2+\beta\sum_{i,j=1}^{V}w_iw_j\mathrm{Tr}((\mathcal{Z}^{(i)})^T\mathcal{Z}^{(j)}),\end{aligned}\tag{5}$$

where $\beta > 0$ is a balancing parameter. We name the above model the Cross-view diversity embedded Consensus Learning for Multi-View Clustering (CCL-MVC). We will develop an effective optimization algorithm in the next section.

4 Optimization

In this section, we develop an effective optimization algorithm using the augmented Lagrange multiplier method (ALM). We introduce an auxiliary variable $\mathcal{Q}\in\mathbb{R}^{n\times n\times V}$ and obtain the following augmented Lagrange function:

$$\begin{aligned}&L(\mathcal{Z},\mathcal{E},\mathcal{Q},\mathrm{Z}_0,w,\mathcal{Y}_1,\mathcal{Y}_2,\rho)=\|\mathcal{Q}\|_{\circledR}+\lambda\|\mathcal{E}\|_{\circledS}\\&+\alpha\sum_{v}\|\mathrm{Z}_0-w_v\mathcal{Z}^{(v)}\|_F^2+\beta\sum_{i,j}w_iw_j\mathrm{Tr}((\mathcal{Z}^{(i)})^T\mathcal{Z}^{(j)})\\&+\frac{\rho}{2}\|\mathcal{Q}-\mathcal{Z}+\mathcal{Y}_1/\rho\|_F^2+\frac{\rho}{2}\|\mathcal{P}_K-\mathcal{Z}-\mathcal{E}+\mathcal{Y}_2/\rho\|_F^2,\end{aligned}\tag{6}$$

where $\mathcal{Y}_1,\mathcal{Y}_2\in\mathbb{R}^{n\times n\times V}$ are Lagrange multipliers and ρ is the penalty parameter. Then, we alternately optimize each variable with details presented as follows.

4.1 Optimization of $\mathcal{Z}$

The subproblem associated with $\mathcal{Z}$ is to minimize

$$\begin{aligned}&\min_{\mathcal{Z}}\alpha\sum_{v}\|\mathrm{Z}_0-w_v\mathcal{Z}^{(v)}\|_F^2+\beta\sum_{i,j}w_iw_j\mathrm{Tr}((\mathcal{Z}^{(i)})^T\mathcal{Z}^{(j)})\\&+\frac{\rho}{2}\|\mathcal{Q}-\mathcal{Z}+\mathcal{Y}_1/\rho\|_F^2+\frac{\rho}{2}\|\mathcal{P}_K-\mathcal{Z}-\mathcal{E}+\mathcal{Y}_2/\rho\|_F^2.\end{aligned}\tag{7}$$

An intuitive way is to optimize each $\mathcal{Z}^{(v)}$ while keeping the others fixed. However, this strategy may suffer from the following issues. First, all $\mathcal{Z}^{(v)}$'s depend on each other and we cannot obtain the global solution for all $\mathcal{Z}^{(v)}$'s simultaneously. Second, we may need to iteratively update all $\mathcal{Z}^{(v)}$'s within an inner loop until convergence to obtain the optimal $\mathcal{Z}$, which lacks efficiency. To address these issues, we design an efficient optimization strategy as follows. First, we define the following augmented variables:

$$\bar{\mathrm{Z}}=\begin{bmatrix}\mathcal{Z}^{(1)}\\\vdots\\\mathcal{Z}^{(V)}\end{bmatrix},\bar{\mathrm{Z}}_0=\begin{bmatrix}\mathrm{Z}_0\\\vdots\\\mathrm{Z}_0\end{bmatrix},\bar{\mathrm{I}}_1=\begin{bmatrix}w_1\mathrm{I}&&\\&\ddots&\\&&w_V\mathrm{I}\end{bmatrix},\bar{\mathrm{I}}_2=\begin{bmatrix}w_1\mathrm{I}\\\vdots\\w_V\mathrm{I}\end{bmatrix}^T.$$

Then, with these notations, we derive an equavilent form of Eq. (7) to facilitate the optimization. For the first term of Eq. (7), we have $\alpha\sum_v\|\mathrm{Z}_0-w_v\mathcal{Z}^{(v)}\|_F^2=\alpha\|\bar{\mathrm{Z}}_0-\bar{\mathrm{I}}_1\bar{\mathrm{Z}}\|_F^2$. For the second term, we have $\beta\sum_{ij}w_iw_j\mathrm{Tr}((\mathcal{Z}^{(i)})^T\mathcal{Z}^{(j)})=\beta\sum_{ij}\mathrm{Tr}((w_i\mathcal{Z}^{(i)})^T(w_j\mathcal{Z}^{(j)}))=\beta\|\sum_i w_i\mathcal{Z}^{(i)}\|_F^2=\beta\|\bar{\mathrm{I}}_2\bar{\mathrm{Z}}\|_F^2$. For the last two terms, we may combine them and obtain $\frac{\rho}{2}\|\mathcal{Q}-\mathcal{Z}+\mathcal{Y}_1/\rho\|_F^2+\frac{\rho}{2}\|\mathcal{P}_K-\mathcal{Z}-\mathcal{E}+\mathcal{Y}_2/\rho\|_F^2=\rho\|\bar{\mathrm{Z}}-\bar{\mathrm{H}}\|_F^2$, where $\bar{\mathrm{H}}=[\frac{(\mathcal{Q}^{(1)}+\mathcal{P}_K^{(1)}-\mathcal{E}^{(1)})^T}{2}+\frac{(\mathcal{Y}_1^{(1)}+\mathcal{Y}_2^{(1)})^T}{2\rho},\cdots,\frac{(\mathcal{Q}^{(V)}+\mathcal{P}_K^{(V)}-\mathcal{E}^{(V)})^T}{2}+\frac{(\mathcal{Y}_1^{(V)}+\mathcal{Y}_2^{(V)})^T}{2\rho}]^T\in\mathbb{R}^{nV\times n}$. Thus, we may convert Eq. (7) to the following quadratic problem:

$$\min_{\bar{Z}}\alpha\|\bar{\mathrm{Z}}_0-\bar{\mathrm{I}}_1\bar{\mathrm{Z}}\|_F^2+\beta\|\bar{\mathrm{I}}_2\bar{\mathrm{Z}}\|_F^2+\rho\|\bar{\mathrm{Z}}-\bar{\mathrm{H}}\|_F^2.\tag{8}$$

The first-order optimality condition of Eq. (10) is:

$$2\alpha\bar{\mathrm{I}}_1^T\bar{\mathrm{I}}_1\bar{\mathrm{Z}}+2\beta\bar{\mathrm{I}}_2^T\bar{\mathrm{I}}_2\bar{\mathrm{Z}}-2\alpha\bar{\mathrm{I}}_1^T\bar{\mathrm{Z}}_0+2\rho\bar{\mathrm{Z}}-2\rho\bar{\mathrm{H}}=0,\tag{9}$$

which leads to the following closed-form solution:

$$\bar{\mathrm{Z}}=(\alpha\bar{\mathrm{I}}_1^T\bar{\mathrm{I}}_1+\beta\bar{\mathrm{I}}_2^T\bar{\mathrm{I}}_2+\rho\mathrm{I})^{-1}(\alpha\bar{\mathrm{I}}_1^T\bar{\mathrm{Z}}_0+\rho\bar{\mathrm{H}}).\tag{10}$$

It is seen that the above solution involves inversion of an $nV\times nV$ matrix, which generally has $O(n^3V^3)$ complexity. Fortunately, the above solution admits a special structure, which can be further simplified by the Sherman-Morrison-Woodbury formula. First, we have

$$\begin{aligned}&(\alpha\bar{\mathrm{I}}_1^T\bar{\mathrm{I}}_1+\beta\bar{\mathrm{I}}_2^T\bar{\mathrm{I}}_2+\rho\mathrm{I})^{-1}=(\alpha\bar{\mathrm{I}}_1^T\bar{\mathrm{I}}_1+\rho\mathrm{I}+\beta\bar{\mathrm{I}}_2^T\bar{\mathrm{I}}_2)^{-1}\\&=\mathrm{A}^{-1}-\beta\mathrm{A}^{-1}\bar{\mathrm{I}}_2^T(\mathrm{I}+\beta\bar{\mathrm{I}}_2\mathrm{A}^{-1}\bar{\mathrm{I}}_2^T)^{-1}\bar{\mathrm{I}}_2\mathrm{A}^{-1}\\&=\mathrm{A}^{-1}-\beta\mathrm{A}^{-1}\bar{\mathrm{I}}_2^T\left(\mathrm{I}+\left(\beta\sum_{i=1}^{V}\frac{w_i^2}{\rho+\alpha w_i^2}\right)\mathrm{I}\right)^{-1}\bar{\mathrm{I}}_2\mathrm{A}^{-1}\\&=\mathrm{A}^{-1}-\frac{\beta}{1+\sum_{i=1}^{V}\frac{\beta w_i^2}{\rho+\alpha w_i^2}}\mathrm{A}^{-1}\bar{\mathrm{I}}_2^T\bar{\mathrm{I}}_2\mathrm{A}^{-1}\\&=\mathrm{A}^{-1}-\beta\frac{\begin{bmatrix}\frac{w_1}{\rho+\alpha w_1^2}\mathrm{I}\\\vdots\\\frac{w_V}{\rho+\alpha w_V^2}\mathrm{I}\end{bmatrix}\begin{bmatrix}\frac{w_1}{\rho+\alpha w_1^2}\mathrm{I}&\cdots&\frac{w_V}{\rho+\alpha w_V^2}\mathrm{I}\end{bmatrix}}{1+\sum_{i=1}^{V}\frac{\beta w_i^2}{\rho+\alpha w_i^2}}\\&=\begin{bmatrix}\frac{1}{\xi_1}\mathrm{I}&\cdots&0\\\vdots&\ddots&\vdots\\0&\cdots&\frac{1}{\xi_V}\mathrm{I}\end{bmatrix}-\frac{\beta}{1+\sum_i\frac{\beta w_i^2}{\xi_i}}\begin{bmatrix}\xi'_{11}\mathrm{I}&\cdots&\xi'_{1V}\mathrm{I}\\\vdots&\ddots&\vdots\\\xi'_{V1}\mathrm{I}&\cdots&\xi'_{VV}\mathrm{I}\end{bmatrix},\end{aligned}\tag{11}$$

with $\xi_i=\rho+\alpha w_i^2$ and $\xi'_{ij}=\frac{w_iw_j}{\xi_i\xi_j}$ for $i,j=1,\cdots,V$, and $\mathrm{A}=\alpha\bar{\mathrm{I}}_1^T\bar{\mathrm{I}}_1+\rho\mathrm{I}$. Define $\mathcal{B}\in\mathbb{R}^{n\times n\times V}$ with $\mathcal{B}^{(v)}=\alpha w_v\mathrm{Z}_0+\frac{\rho(\mathcal{Q}^{(v)}+\mathcal{P}_K^{(v)}-\mathcal{E}^{(v)})+\mathcal{Y}_1^{(v)}+\mathcal{Y}_2^{(v)}}{2}$ for $v=1,\cdots,V$, then, by plugging Eq. (11) and $\mathcal{B}$ into Eq. (10), we have

$$\bar{\mathrm{Z}}=\begin{bmatrix}\frac{\mathcal{B}^{(1)}}{\rho+\alpha w_1^2}-\frac{1}{\frac{1}{\beta}+\sum_i\frac{w_i^2}{\rho+\alpha w_i^2}}\sum_j\frac{w_1w_j\mathcal{B}^{(j)}}{(\rho+\alpha w_1^2)(\rho+\alpha w_j^2)}\\\vdots\\\frac{\mathcal{B}^{(V)}}{\rho+\alpha w_V^2}-\frac{1}{\frac{1}{\beta}+\sum_i\frac{w_i^2}{\rho+\alpha w_i^2}}\sum_j\frac{w_Vw_j\mathcal{B}^{(j)}}{(\rho+\alpha w_V^2)(\rho+\alpha w_j^2)}\end{bmatrix},\tag{12}$$

whose complexity surprisingly mainly comes from scatter-matrix product. Then, $\mathcal{Z}$ is obtained by tensorizing $\bar{\mathrm{Z}}$.

4.2 Optimization of $\mathcal{Q}$

The subproblem of $\mathcal{Q}$ is $\min_{\mathcal{Q}} \|\mathcal{Q}\|_{\circledR} + \frac{\rho}{2}\|\mathcal{Q} - \mathcal{Z} + \mathcal{Y}_1/\rho\|_F^2$. This problem is not straightforward to solve since the LTRA is defined in the frequency domain. In a way similar to the shrinkage problems in [Gao *et al.*, 2021; Pan *et al.*, 2023; Wu *et al.*, 2019], the above problem can be converted to the frequency domain as

$$\min_{\hat{\mathcal{Q}}}\sum_{v=1}^{V}\Big\{\frac{1}{2}\|\hat{\mathcal{K}}^{(v)}-\hat{\mathcal{Q}}^{(v)}\|_F^2+\frac{1}{\rho}\sum_{i=1}^{n}\log(1+\boldsymbol{\sigma}_i(\hat{\mathcal{Q}}^{(v)}))\Big\}, \quad (13)$$

where $\hat{\mathcal{K}} = \mathbf{fft}(\mathcal{Z} - \mathcal{Y}_1/\rho, [\,], 3)$. Then, the above problem can be divided into V independent subproblems of $\mathcal{Q}^{(v)}$ with $v = 1, \cdots, V$, which are the standard log-determinant regularized shrinkage problems described in [Peng *et al.*, 2022a]. We denote $\mathbb{U}(\cdot)$ and $\mathbb{V}(\cdot)$ as operators that return the left and right singular vectors of the input, and $\mathbb{D}_e(\cdot)$ as an operator that returns a diagonal matrix based on the input elements, then according to [Peng *et al.*, 2022a; Peng *et al.*, 2015], $\hat{\mathcal{Q}}^{(v)}$ is obtained by

$$\hat{\mathcal{Q}}^{(v)} = \mathbb{U}(\hat{\mathcal{K}}^{(v)})\mathbb{D}_e\{\hat{\sigma}_i^{(v)}, \cdots, \hat{\sigma}_n^{(v)}\}(\mathbb{V}(\hat{\mathcal{K}}^{(v)}))^T,$$

with $\hat{\sigma}_i^{(v)} = \arg\min_{x\geq 0} \frac{1}{2}(x - \sigma_i^{(v)})^2 + \frac{1}{\rho}\log(1 + x)$ for $i = 1, \cdots, n$. With straightforward algebra, we may obtain $\hat{\sigma}_i^{(v)} = \mathbb{I}_{\{\hat{\sigma}_i^{(v)}\in\mathbb{S}_{i,v}\}} \cdot \frac{\sigma_i^{(v)}-1+\sqrt{(1+\sigma_i^{(v)})^2-\frac{4}{\rho}}}{2}$, with $\mathbb{I}_{\{\cdot\}}$ being an indicator function that returns 1 if the conditions in subscript hold and 0 otherwise, and $\mathbb{S}_{i,v} = \Big\{x|x\in\mathbb{R}, f:\mathbb{R}\to\mathbb{R}, f(x) = \frac{1}{2}(x - \sigma_i^{(v)})^2 + \frac{1}{\rho}\log(1 + x), \xi = \frac{\sigma_i^{(v)}-1+\sqrt{(1+\sigma_i^{(v)})^2-\frac{4}{\rho}}}{2} > 0, f(\xi) \leq \frac{(\sigma_i^{(v)})^2}{2}, (1 + \sigma_i^{(v)})^2 > \frac{4}{\rho}\Big\}$. Thus, we may obtain $\mathcal{Q} = \mathbf{ifft}(\hat{\mathcal{Q}}, [\,], 3)$, with $\mathbf{ifft}(\cdot)$ being the inverse FFT of the input tensor along the third dimension. For ease of notation, we summarize the above procedure as

$$\mathcal{Q} = \mathbb{LTRA}_{1/\rho}(\mathcal{Z} - \mathcal{Y}_1/\rho). \quad (14)$$

4.3 Optimization of $\mathcal{E}$

The LTSA-regularized shrinkage problem of $\mathcal{E}$ is:

$$\min_{\mathcal{E}} \lambda\|\mathcal{E}\|_{\circledS} + \frac{\rho}{2}\|\mathcal{P}_K - \mathcal{Z} - \mathcal{E} + \mathcal{Y}_2/\rho\|_F^2, \quad (15)$$

which is a LTSA-regularized shrinkage problem. It is noted that by reforming the lateral slices of the tensors in Eq. (15) into vectors, the problem is a direct extension of the $\ell_{2,\log}$-norm regularized shrinkage problem in [Peng *et al.*, 2022a]. Then, by performing a vectorization-matrixization procedure to the $\ell_{2,\log}$-norm shrinkage problem in [Peng *et al.*, 2022a], we may obtain the above solution. Thus, it is straightforward that Eq. (15) admits the following closed-form solution by lateral slice as $\mathcal{E}_{L,j} = \frac{\delta_j-1+\sqrt{(1+\delta_j)^2-\frac{4\lambda}{\rho}}}{2\delta_j} \cdot \bar{\mathcal{K}}_{L,j} \cdot \mathbb{I}_{\{\delta_j\in\mathbb{S}_j\}}$, where $(\cdot)_{L,j}$ denotes the j-th lateral slice of the input tensor, $\bar{\mathcal{K}} = \mathcal{P}_K - \mathcal{Z} + \mathcal{Y}_2/\rho$, $\delta_j = \|(\mathcal{P}_K - \mathcal{Z} + \mathcal{Y}_2/\rho)_{L,j}\|_F$, and $\mathbb{S}_j = \Big\{x|x \in \mathbb{R}, f : \mathbb{R} \to \mathbb{R}, f(x)=\frac{1}{2}(x-\delta_j)^2+\frac{\lambda}{\rho}\log(1+x), \xi=\frac{\delta_j-1+\sqrt{(1+\delta_j)^2-\frac{4\lambda}{\rho}}}{2} > 0, f(\xi) \leq \frac{\delta_j^2}{2}, (1 + \delta_j)^2 > \frac{4\lambda}{\rho}\Big\}$. For ease of notation, we summarize the above procedure as

$$\mathcal{E} = \mathbb{LTSA}_{\lambda/\rho}(\mathcal{P}_K - \mathcal{Z} + \mathcal{Y}_2/\rho). \quad (16)$$

4.4 Optimization of Z_0

The subproblem of Z_0 is $\alpha\sum_v \|\mathrm{Z}_0 - w_v\mathcal{Z}^{(v)}\|_F^2$, which is convex and quadratic, and admits a closed-form solution by the first-order optimality condition: $\mathrm{Z}_0 = \frac{1}{V}\sum_{v=1}^{V} w_v\mathcal{Z}^{(v)}$.

4.5 Optimization of w

The subproblem of w is $\min_{w\geq 0,\sum_i w_i=1}\{\alpha\sum_i \|\mathrm{Z}_0 - w_i\mathcal{Z}^{(i)}\|_F^2 + \beta\sum_{ij} w_iw_j\mathrm{Tr}((\mathcal{Z}^{(i)})^T\mathcal{Z}^{(j)})\}$. Let $\mathbf{z}_i$ be vectorized $\mathcal{Z}^{(i)}$ for $i = 1, \cdots, V$, $\mathrm{M}_1 \in \mathbb{R}^{V\times V}$ be a diagonal matrix with $(\mathrm{M}_1)_{ii} = \mathbf{z}_i^T\mathbf{z}_i$, $\mathrm{M}_2 \in \mathbb{R}^{V\times V}$ with $(\mathrm{M}_2)_{ij} = \mathbf{z}_i^T\mathbf{z}_j$, and $h \in \mathbb{R}^V$ with $h_i = -\alpha\mathrm{Tr}(\mathrm{Z}_0^T\mathcal{Z}^{(i)})$, then the subproblem of w is equavilent to

$$\min_{w} w^T(\alpha\mathrm{M}_1+\beta\mathrm{M}_2)w+2h^Tw, s.t.\ w \geq 0, \sum_i w_i = 1. \quad (17)$$

It is easy to verify that $\forall z \in \mathbb{R}^V$ and $z \neq 0$, we have $z^T(\alpha\mathrm{M}_1 + \beta\mathrm{M}_2)z = \alpha\sum_i z_i^2\mathbf{z}_i^T\mathbf{z}_i + \beta\sum_{i,j} z_iz_j\mathbf{z}_i^T\mathbf{z}_j = \alpha\sum_i(z_i\mathbf{z}_i)^T(z_i\mathbf{z}_i)+\beta\sum_i\sum_j(z_i\mathbf{z}_i)^T(z_j\mathbf{z}_j)=\alpha\sum_i \|z_i\mathbf{z}_i\|_2^2 +\beta\|[\mathbf{z}_1, \cdots, \mathbf{z}_V]\mathbb{D}_e(z)\|_F^2 \geq 0$, with $\mathbf{z}_i$ and $\mathbf{z}_j$ being the vectorized $\mathcal{Z}^{(i)}$ and $\mathcal{Z}^{(j)}$, and z_i and z_j denoting the i-th and j-th elements of z, respectively. Thus, $\alpha\mathrm{M}_1 + \beta\mathrm{M}_2 \succeq 0$, and Eq. (17) is convex. Then, w can be efficiently solved by standard convex programming technique, which is denoted as

$$w = \mathrm{quadprog}(\alpha\mathrm{M}_1 + \beta\mathrm{M}_2, h). \quad (18)$$

4.6 Updating of $\mathcal{Y}_1$, $\mathcal{Y}_2$, and ρ

The Lagrange multipliers and the penalty parameter are updated in a standard way as follows:

$$\mathcal{Y}_1=\mathcal{Y}_1+\rho(\mathcal{Q}-\mathcal{Z}), \mathcal{Y}_2=\mathcal{Y}_2+\rho(\mathcal{P}_K-\mathcal{Z}-\mathcal{E}), \rho=\rho\kappa, \quad (19)$$

with $\kappa > 1$ being a parameter that keeps ρ increasing. In our paper, we update the variables in the order of $\mathcal{Z}$, $\mathcal{E}$, $\mathcal{Q}$, Z_0, w, $\mathcal{Y}_1$, $\mathcal{Y}_2$, and ρ, which is essentially important for the convergence analysis in Section 5. After we obtain the solution, the standard SPC is applied to Z_0 for the final clustering result.

5 Convergence Analysis

In this section, we analyze the convergent property of the CCL-MVC. The main results are presented in the following.

Theorem 1. Let t in the superscript denote the iteration number. Under assumptions that $\sum \frac{1}{\rho^t} < \infty$ and $\sum \frac{\rho^{t+1}}{(\rho^t)^2} < \infty$, and given a bounded initialization of the variables, the variable sequences $\{\mathcal{Z}^t\}$, $\{\mathcal{E}^t\}$, $\{\mathcal{Q}^t\}$, $\{\mathrm{Z}_0^t\}$, $\{w^t\}$, $\{\mathcal{Y}_1^t\}$, and $\{\mathcal{Y}_2^t\}$ generated by our optimization algorithm are bounded.

Proof: According to the constraints of w, it is easy to verify that $\{w^t\}$ is bounded. At the $(t + 1)$-th iteration, according to the first-order optimality condition of $\mathcal{Q}$, we have

$$\partial_{\mathcal{Q}}L(\mathcal{Z}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Q}, \mathrm{Z}_0^t, w^t, \mathcal{Y}_1^t, \mathcal{Y}_2^t, \rho^t)|_{\mathcal{Q}^{t+1}} \quad (20)$$

$$=\partial\|\mathcal{Q}\|_{\circledR}/\partial\mathcal{Q}|_{\mathcal{Q}^{t+1}}+\rho^t(\mathcal{Q}^{t+1}-\mathcal{Z}^{t+1})+\mathcal{Y}_1^t$$
$$=\partial\|\mathcal{Q}\|_{\circledR}/\partial\mathcal{Q}|_{\mathcal{Q}^{t+1}}+\mathcal{Y}_1^{t+1}=0.$$

By definition of the LTRA, and according to [Wu *et al.*, 2019; Peng *et al.*, 2022a], it is clear that $\|\partial\|\mathcal{Q}\|_{\circledR}/\partial\mathcal{Q}|_{\mathcal{Q}^{t+1}}\|_F^2<\infty$ and thus $\mathcal{Y}_1^{t+1}$ is bounded.

Similarly, according to the first-order optimality condition of $\mathcal{E}$, at the $(t+1)$-th iteration, we have

$$\begin{aligned}&\partial_{\mathcal{E}}L(\mathcal{Z}^{t+1},\mathcal{E},\mathcal{Q}^t,\mathrm{Z}_0^t,w^t,\mathcal{Y}_1^t,\mathcal{Y}_2^t,\rho^t)|_{\mathcal{E}^{t+1}}\\&=\partial(\alpha\|\mathcal{E}\|_{\circledS})/\partial\mathcal{E}|_{\mathcal{E}^{t+1}}-\rho^t(\mathcal{P}_K-\mathcal{Z}^{t+1}-\mathcal{E}^{t+1})-\mathcal{Y}_2^t\\&=\alpha\cdot\partial\|\mathcal{E}\|_{\circledS}/\partial\mathcal{E}|_{\mathcal{E}^{t+1}}-\mathcal{Y}_2^{t+1}=0.\end{aligned}\tag{21}$$

By definition of the LTSA, it is clear that $\|\partial\|\mathcal{E}\|_{\circledS}/\partial\mathcal{E}|_{\mathcal{E}^{t+1}}\|_F<\infty$ and thus $\mathcal{Y}_2^{t+1}$ is bounded. Then, according to the updating rules of $\mathcal{Y}_1$ and $\mathcal{Y}_2$, we have

$$\begin{aligned}&L(\mathcal{Z}^{t+1},\mathcal{E}^{t+1},\mathcal{Q}^{t+1},\mathrm{Z}_0^{t+1},w^{t+1},\mathcal{Y}_1^{t+1},\mathcal{Y}_2^{t+1},\rho^{t+1})\\&=L(\mathcal{Z}^{t+1},\mathcal{E}^{t+1},\mathcal{Q}^{t+1},\mathrm{Z}_0^{t+1},w^{t+1},\mathcal{Y}_1^t,\mathcal{Y}_2^t,\rho^t)\\&\quad+\frac{\rho^{t+1}-\rho^t}{2}\|\mathcal{Q}^{t+1}-\mathcal{Z}^{t+1}\|_F^2+\langle\mathcal{Q}^{t+1}-\mathcal{Z}^{t+1},\\&\quad\mathcal{Y}_1^{t+1}-\mathcal{Y}_1^t\rangle+\frac{1}{2\rho^{t+1}}\|\mathcal{Y}_1^{t+1}\|_F^2-\frac{1}{2\rho^t}\|\mathcal{Y}_1^t\|_F^2\\&\quad+\frac{\rho^{t+1}-\rho^t}{2}\|\mathcal{P}_K-\mathcal{Z}^{t+1}-\mathcal{E}^{t+1}\|_F^2+\langle\mathcal{Y}_2^{t+1}-\mathcal{Y}_2^t,\\&\quad\mathcal{P}_K-\mathcal{Z}^{t+1}-\mathcal{E}^{t+1}\rangle+\frac{1}{2\rho^{t+1}}\|\mathcal{Y}_2^{t+1}\|_F^2-\frac{1}{2\rho^t}\|\mathcal{Y}_2^t\|_F^2\\&\le L(\mathcal{Z}^t,\mathcal{E}^t,\mathcal{Q}^t,\mathrm{Z}_0^t,w^t,\mathcal{Y}_1^t,\mathcal{Y}_2^t,\rho^t)\\&\quad+\frac{\rho^{t+1}+\rho^t}{2(\rho^t)^2}\|\mathcal{Y}_1^{t+1}-\mathcal{Y}_1^t\|_F^2+\frac{1}{2\rho^{t+1}}\|\mathcal{Y}_1^{t+1}\|_F^2\\&\quad+\frac{\rho^{t+1}+\rho^t}{2(\rho^t)^2}\|\mathcal{Y}_2^{t+1}-\mathcal{Y}_2^t\|_F^2+\frac{1}{2\rho^{t+1}}\|\mathcal{Y}_2^{t+1}\|_F^2,\end{aligned}\tag{22}$$

where the last inequality holds by $\mathcal{Y}_1^{t+1}=\mathcal{Y}_1^t+\rho^t(\mathcal{Q}^{t+1}-\mathcal{Z}^{t+1})$ and $\mathcal{Y}_2^{t+1}=\mathcal{Y}_2^t+\rho^t(\mathcal{P}_K-\mathcal{Z}^{t+1}-\mathcal{E}^{t+1})$. Thus, by repeating the above inequality $t+1$ times, it is clear that

$$\begin{aligned}&L(\mathcal{Z}^{t+1},\mathcal{E}^{t+1},\mathcal{Q}^{t+1},\mathrm{Z}_0^{t+1},w^{t+1},\mathcal{Y}_1^{t+1},\mathcal{Y}_2^{t+1},\rho^{t+1})\\&\le L(\mathcal{Z}^0,\mathcal{E}^0,\mathcal{Q}^0,\mathrm{Z}_0^0,w^0,\mathcal{Y}_1^0,\mathcal{Y}_2^0,\rho^0)\\&\quad+\sum_{s=0}^{t}\frac{\rho^{s+1}+\rho^s}{2(\rho^s)^2}\|\mathcal{Y}_1^{s+1}-\mathcal{Y}_1^s\|_F^2+\sum_{s=0}^{t}\frac{\|\mathcal{Y}_1^{s+1}\|_F^2}{2\rho^{s+1}}\\&\quad+\sum_{s=0}^{t}\frac{\rho^{s+1}+\rho^s}{2(\rho^s)^2}\|\mathcal{Y}_2^{s+1}-\mathcal{Y}_2^s\|_F^2+\sum_{s=0}^{t}\frac{\|\mathcal{Y}_2^{s+1}\|_F^2}{2\rho^{s+1}}\\&\le L(\mathcal{Z}^0,\mathcal{E}^0,\mathcal{Q}^0,\mathrm{Z}_0^0,w^0,\mathcal{Y}_1^0,\mathcal{Y}_2^0,\rho^0)\\&\quad+2(\Gamma_1+\Gamma_2)\sum_{s=0}^{t}\frac{\rho^{s+1}}{(\rho^s)^2}+\frac{5(\Gamma_1+\Gamma_2)}{2}\sum_{s=0}^{t+1}\frac{1}{\rho^s},\end{aligned}\tag{23}$$

with $\Gamma_1=\max_{s=0}^{t+1}\{\|\mathcal{Y}_1^s\|_F^2\}$ and $\Gamma_2=\max_{s=0}^{t+1}\{\|\mathcal{Y}_2^s\|_F^2\}$. Thus, with bounded initialization, Eq. (23) must be bounded.

By definition of $L(\mathcal{Z},\mathcal{E},\mathcal{Q},\mathrm{Z}_0,w,\mathcal{Y}_1,\mathcal{Y}_2,\rho)$ in Eq. (6), it is clear that all of its terms are nonnegative and are thus bounded. Then, according to the definitions of $\|\mathcal{Q}^{t+1}\|_{\circledR}$ and $\|\mathcal{E}^{t+1}\|_{\circledS}$, the boundedness directly implies that $\mathcal{Q}^{t+1}$ and $\mathcal{E}^{t+1}$ are bounded. Because $\mathcal{Q}^{t+1}$, $\mathcal{Y}_1^{t+1}$, and $\rho^{t+1}\|\mathcal{Q}^{t+1}-\mathcal{Z}^{t+1}+\mathcal{Y}_1^{t+1}/\rho^{t+1}\|_F^2$ are all bounded, it is clear that $\mathcal{Z}^{t+1}$ is also bounded. Then, $\mathrm{Z}_0^{t+1}=\frac{1}{V}\sum_{v=1}^{V}w_i^t(\mathcal{Z}^{t+1})^{(v)}$ is also bounded. Now, we may conclude that the sequences of $\{\mathcal{Z}^t\}$, $\{\mathcal{E}^t\}$, $\{\mathcal{Q}^t\}$, $\{\mathrm{Z}_0^t\}$, $\{w^t\}$, $\{\mathcal{Y}_1^t\}$, and $\{\mathcal{Y}_2^t\}$ are all bounded. □

Theorem 2. Let $\{\mathcal{Z}^t,\mathcal{E}^t,\mathcal{Q}^t,\mathrm{Z}_0^t,w^t,\mathcal{Y}_1^t,\mathcal{Y}_2^t\}$ be a sequence generated by our algorithm. Under assumptions that $\sum\frac{1}{\rho^t}<\infty$, $\sum\frac{\rho^{t+1}}{(\rho^t)^2}<\infty$, $\rho^t(\mathcal{Q}^{t+1}-\mathcal{Q}^t)\to 0$, and $\rho^t(\mathcal{E}^{t+1}-\mathcal{E}^t)\to 0$, the sequence $\{\mathcal{Z}^t,\mathcal{E}^t,\mathcal{Q}^t,\mathrm{Z}_0^t,w^t,\mathcal{Y}_1^t,\mathcal{Y}_2^t\}$ has at least one accumulation point. For any accumulation point, denoted as $\{\mathcal{Z}^*,\mathcal{E}^*,\mathcal{Q}^*,\mathrm{Z}_0^*,w^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*\}$, $\{\mathcal{Z}^*,\mathcal{E}^*,\mathcal{Q}^*,\mathrm{Z}_0^*,w^*\}$ is a stationary point of the optimization problem in Eq. (5).

Proof: By Theorem 1, we know $\{\mathcal{Z}^t,\mathcal{E}^t,\mathcal{Q}^t,\mathrm{Z}_0^t,w^t,\mathcal{Y}_1^t,\mathcal{Y}_2^t\}$ is bounded. Then, according to the Bolzano-Weierstrass theorem, the sequence has at least one accumulation point, denoted as $\{\mathcal{Z}^*,\mathcal{E}^*,\mathcal{Q}^*,\mathrm{Z}_0^*,w^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*\}$. Then, there exists a sub-sequence of $\{\mathcal{Z}^t,\mathcal{E}^t,\mathcal{Q}^t,\mathrm{Z}_0^t,w^t,\mathcal{Y}_1^t,\mathcal{Y}_2^t\}$ that converges to $\{\mathcal{Z}^*,\mathcal{E}^*,\mathcal{Q}^*,\mathrm{Z}_0^*,w^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*\}$. Without loss of generality, we assume that $\{\mathcal{Z}^t,\mathcal{E}^t,\mathcal{Q}^t,\mathrm{Z}_0^t,w^t,\mathcal{Y}_1^t,\mathcal{Y}_2^t\}$ itself converges to $\{\mathcal{Z}^*,\mathcal{E}^*,\mathcal{Q}^*,\mathrm{Z}_0^*,w^*,\mathcal{Y}_1^*,\mathcal{Y}_2^*\}$. Next, we will show that $\{\mathcal{Z}^*,\mathcal{E}^*,\mathcal{Q}^*,\mathrm{Z}_0^*,w^*\}$ is a stationary point of Eq. (5).

By the assumption that $\sum\frac{1}{\rho^t}<\infty$, it is clear that $\rho^t\to\infty$. Then, by the boundedness of $\{\mathcal{Y}_1^t\}$ and $\{\mathcal{Y}_2^t\}$, we have

$$\begin{aligned}&\mathcal{Q}^*-\mathcal{Z}^*=\lim_{t\to\infty}\mathcal{Q}^{t+1}-\mathcal{Z}^{t+1}=\lim_{t\to\infty}(\mathcal{Y}_1^{t+1}-\mathcal{Y}_1^t)/\rho^t=0,\\&\mathcal{P}_K-\mathcal{Z}^*-\mathcal{E}^*=\lim_{t\to\infty}\mathcal{P}_K-\mathcal{Z}^{t+1}-\mathcal{E}^{t+1}\\&\qquad=\lim_{t\to\infty}(\mathcal{Y}_2^{t+1}-\mathcal{Y}_2^t)/\rho^t=0.\end{aligned}\tag{24}$$

At the $(t+1)$-th iteration, according to the optimality conditions of $\mathcal{Q}$ and $\mathcal{E}$, we have the following equalities:

$$\begin{aligned}&\lim_{t\to\infty}\partial_{\mathcal{Q}}L(\mathcal{Z}^{t+1},\mathcal{E}^{t+1},\mathcal{Q},\mathrm{Z}_0^t,w^t,\mathcal{Y}_1^t,\mathcal{Y}_2^t,\rho^t)|_{\mathcal{Q}^{t+1}}\\&=\lim_{t\to\infty}\{\partial\|\mathcal{Q}\|_{\circledR}/\partial\mathcal{Q}|_{\mathcal{Q}^{t+1}}+\rho^t(\mathcal{Q}^{t+1}-\mathcal{Z}^{t+1})+\mathcal{Y}_1^t\}\\&=\partial\|\mathcal{Q}\|_{\circledR}/\partial\mathcal{Q}|_{\mathcal{Q}^*}+\mathcal{Y}_1^*=0,\end{aligned}\tag{25}$$

and

$$\begin{aligned}&\lim_{t\to\infty}\partial_{\mathcal{E}}L(\mathcal{Z}^{t+1},\mathcal{E},\mathcal{Q}^t,\mathrm{Z}_0^t,w^t,\mathcal{Y}_1^t,\mathcal{Y}_2^t,\rho^t)|_{\mathcal{E}^{t+1}}\\&=\lim_{t\to\infty}\{\partial(\alpha\|\mathcal{E}\|_{\circledS})/\partial\mathcal{E}|_{\mathcal{E}^{t+1}}-\rho^t(\mathcal{P}_K-\mathcal{Z}^{t+1}-\mathcal{E}^{t+1})-\mathcal{Y}_2^t\}\\&=\alpha\cdot\partial\|\mathcal{E}\|_{\circledS}/\partial\mathcal{E}|_{\mathcal{E}^*}-\mathcal{Y}_2^*=0.\end{aligned}\tag{26}$$

For ease of notation, we denote $\mathcal{G}^{t+1}=\partial\{\alpha\sum_{v=1}^{V}\|\mathrm{Z}_0-w_v\mathcal{Z}^{(v)}\|_F^2+\beta\sum_{i,j=1}^{V}w_iw_j\mathrm{Tr}((\mathcal{Z}^{(i)})^T\mathcal{Z}^{(j)})\}/\partial\mathcal{Z}|_{\mathcal{Z}^{t+1}}$, which may be obtained by tensorizing the first three terms of Eq. (9). Then, at the $(t+1)$-th iteration, according to the optimality condition of $\mathcal{Z}$, and under the assumptions that $\rho^t(\mathcal{Q}^{t+1}-\mathcal{Q}^t)\to 0$ and $\rho^t(\mathcal{E}^{t+1}-\mathcal{E}^t)\to 0$, it is clear that

$$\begin{aligned}&\lim_{t\to\infty}\{\mathcal{G}^{t+1}+\rho^t(\mathcal{Z}^{t+1}-\mathcal{Q}^{t+1})+\rho^t(\mathcal{Z}^{t+1}+\mathcal{E}^{t+1}-\mathcal{P}_K)\\&\qquad-\mathcal{Y}_1^t-\mathcal{Y}_2^t+\rho^t(\mathcal{Q}^t-\mathcal{Q}^{t+1})+\rho^t(\mathcal{E}^t-\mathcal{E}^{t+1})\}\\&=\lim_{t\to\infty}\{\mathcal{G}^{t+1}-\mathcal{Y}_1^{t+1}-\mathcal{Y}_2^{t+1}\}=\mathcal{G}^*-\mathcal{Y}_1^*-\mathcal{Y}_2^*=0,\end{aligned}\tag{27}$$

where $\mathcal{G}^*=\lim_{t\to\infty}\mathcal{G}^t$ must exist according to Eq. (9) and the boundedness of $\{\mathcal{Z}^t\}$ and $\{w^t\}$. Besides, since w and Z_0

Datasets	BBC-4view				BBC-Sport				Flowers			
Methods	ACC	NMI	AR	F-Score	ACC	NMI	AR	F-Score	ACC	NMI	AR	F-Score
AWP	0.904±0.000	0.760±0.000	0.797±0.000	0.845±0.000	0.809±0.000	0.723±0.000	0.726±0.000	0.796±0.000	0.435±0.000	0.430±0.000	0.246±0.000	0.292±0.000
MvDSCN	0.495±0.019	0.247±0.022	0.224±0.023	0.437±0.015	0.931±0.001	0.935±0.000	0.909±0.001	0.860±0.000	0.276±0.012	0.285±0.012	0.108±0.010	0.182±0.008
OMVFC-LICAG	0.718±0.000	0.586±0.000	0.584±0.000	0.688±0.000	0.724±0.000	0.571±0.000	0.397±0.000	0.627±0.000	0.397±0.000	0.404±0.000	0.208±0.000	0.259±0.000
MLAN	0.853±0.007	0.698±0.010	0.716±0.005	0.783±0.004	0.721±0.000	0.779±0.000	0.591±0.000	0.714±0.000	0.501±0.008	0.532±0.003	0.331±0.010	0.373±0.009
GMC	0.693±0.000	0.563±0.000	0.479±0.000	0.633±0.000	0.807±0.000	0.760±0.000	0.722±0.000	0.794±0.000	0.177±0.000	0.247±0.000	0.020±0.000	0.125±0.000
UOMvSC	0.391±0.000	0.170±0.000	0.105±0.000	0.267±0.000	0.529±0.000	0.302±0.000	0.209±0.000	0.415±0.000	0.507±0.000	0.508±0.000	0.313±0.000	0.355±0.000
EOMSC-CA	0.425±0.000	0.090±0.000	0.090±0.000	0.365±0.000	0.384±0.000	0.083±0.000	0.058±0.000	0.375±0.000	0.374±0.000	0.418±0.000	0.221±0.000	0.281±0.000
LMSC	0.883±0.000	0.699±0.000	0.746±0.000	0.806±0.000	0.847±0.003	0.739±0.001	0.749±0.001	0.810±0.001	0.442±0.009	0.444±0.009	0.275±0.007	0.318±0.012
MCLES	0.819±0.000	0.637±0.000	0.662±0.000	0.742±0.000	0.921±0.000	0.802±0.000	0.795±0.000	0.945±0.000	0.469±0.000	0.516±0.000	0.337±0.000	0.390±0.000
FPMVS-CAG	0.323±0.000	0.030±0.000	0.037±0.000	0.276±0.000	0.406±0.000	0.106±0.000	0.103±0.000	0.304±0.000	0.272±0.000	0.356±0.000	0.182±0.000	0.248±0.000
CLR-MVP	0.626±0.000	0.476±0.000	0.437±0.000	0.597±0.000	0.925±0.000	0.815±0.000	0.851±0.000	0.886±0.000	0.520±0.000	0.533±0.001	0.365±0.001	0.403±0.001
t-SVD-MSC	0.858±0.001	0.685±0.002	0.725±0.002	0.789±0.001	0.879±0.000	0.765±0.000	0.784±0.000	0.834±0.000	**0.836±0.005**	**0.852±0.002**	**0.766±0.002**	**0.780±0.002**
SM^{2}SC	**0.934±0.008**	0.812±0.001	**0.853±0.003**	**0.887±0.006**	**0.982±0.000**	**0.937±0.000**	**0.952±0.000**	**0.963±0.000**	0.442±0.008	0.453±0.005	0.276±0.007	0.319±0.007
ETLMSC	0.872±0.094	**0.826±0.028**	0.811±0.082	0.855±0.063	0.959±0.086	**0.972±0.058**	**0.949±0.107**	**0.961±0.081**	**0.811±0.066**	**0.874±0.025**	**0.763±0.057**	**0.778±0.054**
LMVSC	0.480±0.000	0.242±0.000	0.403±0.000	0.380±0.000	0.517±0.000	0.382±0.000	0.151±0.000	0.394±0.000	0.360±0.000	0.385±0.000	0.198±0.000	0.246±0.000
E^{2}OMVC	0.849±0.000	0.707±0.000	0.713±0.000	0.783±0.000	0.971±0.000	0.903±0.000	0.920±0.000	0.940±0.000	0.490±0.000	0.477±0.000	0.256±0.000	0.306±0.000
RMSL	**0.943±0.009**	**0.831±0.005**	**0.862±0.004**	**0.894±0.002**	**0.972±0.002**	0.905±0.005	0.931±0.002	0.947±0.004	0.511±0.006	0.490±0.007	0.332±0.010	0.372±0.005
CCL-MVC	**0.984±0.000**	**0.951±0.000**	**0.962±0.000**	**0.971±0.000**	**0.996±0.000**	**0.986±0.000**	**0.991±0.000**	**0.993±0.000**	**0.890±0.053**	**0.892±0.019**	**0.832±0.052**	**0.852±0.048**
Datasets	UCI-3view				StillDB				MITindoor			
Methods	ACC	NMI	AR	F-Score	ACC	NMI	AR	F-Score	ACC	NMI	AR	F-Score
AWP	0.806±0.000	0.842±0.000	0.759±0.000	0.785±0.000	0.306±0.000	0.093±0.000	0.058±0.000	0.223±0.000	0.499±0.000	0.629±0.000	0.317±0.000	0.329±0.000
MvDSCN	0.308±0.011	0.299±0.013	0.158±0.009	0.281±0.006	0.377±0.023	0.245±0.020	0.169±0.003	0.320±0.015	0.084±0.003	0.182±0.004	0.014±0.002	0.037±0.001
OMVFC-LICAG	0.833±0.000	0.811±0.000	0.731±0.000	0.759±0.000	0.376±0.000	0.129±0.000	0.087±0.000	0.273±0.000	0.319±0.000	0.453±0.000	0.157±0.000	0.171±0.000
MLAN	0.874±0.000	0.910±0.000	0.847±0.000	0.864±0.000	0.349±0.000	0.138±0.000	0.098±0.000	0.272±0.000	0.232±0.010	0.408±0.012	0.012±0.009	0.041±0.003
GMC	0.736±0.000	0.815±0.000	0.678±0.000	0.713±0.000	0.251±0.000	0.078±0.000	0.005±0.000	0.278±0.000	0.099±0.000	0.204±0.000	0.003±0.000	0.032±0.000
UOMvSC	**0.981±0.000**	**0.956±0.000**	**0.958±0.000**	**0.962±0.000**	0.328±0.000	0.131±0.000	0.084±0.000	0.246±0.000	0.344±0.000	0.506±0.000	0.064±0.000	0.088±0.000
EOMSC-CA	0.545±0.000	0.673±0.000	0.459±0.000	0.533±0.000	0.308±0.000	0.127±0.000	0.085±0.000	0.245±0.000	0.147±0.000	0.298±0.000	0.042±0.000	0.066±0.000
LMSC	0.893±0.000	0.815±0.000	0.783±0.000	0.805±0.000	0.327±0.003	0.136±0.003	0.084±0.011	0.269±0.005	0.384±0.006	0.506±0.005	0.243±0.005	0.254±0.004
MCLES	0.941±0.004	0.891±0.008	0.877±0.009	0.889±0.008	0.338±0.000	0.153±0.000	0.098±0.000	0.264±0.000	—	—	—	—
FPMVS-CAG	0.722±0.000	0.744±0.000	0.645±0.000	0.683±0.000	0.328±0.000	0.124±0.000	0.089±0.000	0.251±0.000	0.204±0.000	0.390±0.000	0.085±0.000	0.108±0.000
CLR-MVP	0.965±0.000	0.920±0.001	0.924±0.001	0.932±0.001	0.337±0.002	0.127±0.003	0.095±0.001	0.273±0.002	—	—	—	—
t-SVD-MSC	0.830±0.000	0.884±0.005	0.786±0.003	0.800±0.004	0.347±0.010	0.130±0.004	0.088±0.003	0.255±0.004	**0.684±0.005**	**0.750±0.007**	**0.555±0.005**	**0.562±0.008**
SM^{2}SC	0.961±0.001	0.914±0.001	0.914±0.001	0.923±0.001	**0.452±0.002**	**0.336±0.002**	**0.312±0.002**	**0.370±0.002**	0.477±0.009	0.583±0.006	0.321±0.006	0.332±0.006
ETLMSC	0.958±0.078	**0.977±0.028**	**0.953±0.069**	**0.958±0.062**	**0.604±0.043**	**0.520±0.015**	**0.423±0.029**	**0.523±0.024**	**0.775±0.040**	**0.899±0.011**	**0.729±0.037**	**0.733±0.036**
LMVSC	0.790±0.000	0.756±0.000	0.643±0.000	0.681±0.000	0.317±0.000	0.189±0.000	0.066±0.000	0.239±0.000	0.371±0.000	0.522±0.000	0.112±0.000	0.132±0.000
E^{2}OMVC	**0.974±0.000**	0.941±0.000	0.943±0.000	0.949±0.000	0.321±0.000	0.130±0.000	0.089±0.000	0.265±0.000	0.404±0.000	0.550±0.000	0.173±0.000	0.191±0.000
RMSL	0.578±0.013	0.511±0.014	0.407±0.017	0.474±0.007	0.356±0.003	0.131±0.001	0.090±0.002	0.243±0.001	0.279±0.004	0.372±0.003	0.125±0.005	0.139±0.002
CCL-MVC	**0.988±0.037**	**0.996±0.013**	**0.988±0.037**	**0.990±0.033**	**0.752±0.077**	**0.682±0.042**	**0.597±0.079**	**0.666±0.066**	**0.860±0.038**	**0.946±0.011**	**0.842±0.038**	**0.844±0.037**

Table 1: Comparison of Clustering Performance of Different Methods on Benchmark Data Sets

are optimized using the standard convex programming techniques, the corresponding Karush-Kuhn-Tucker (KKT) conditions are satisfied by w^* and Z_0^*.

Therefore, $\{\mathcal{Z}^*, \mathcal{E}^*, \mathcal{Q}^*, \mathrm{Z}_0^*, w^*, \mathcal{Y}_1^*, \mathcal{Y}_2^*\}$ satisfies KKT conditions of Eq. (6) and thus $\{\mathcal{Z}^*, \mathcal{E}^*, \mathcal{Q}^*, \mathrm{Z}_0^*, w^*\}$ is a stationary point of the original problem Eq. (5).

□

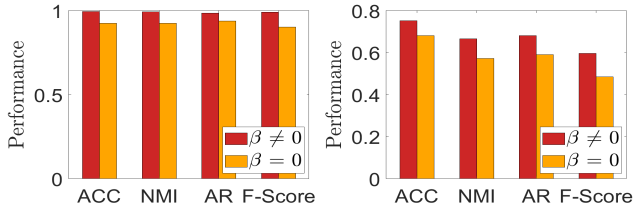
6 Experiments

In this section, we conduct extensive experiments to evaluate the proposed method. In particular, we use six benchmark data sets, including the BBC-4view, BBC-Sport, Flowers, UCI-3view, StillDB, and MITindoor, and four evaluation metrics, including the clustering accuracy (ACC), normalized mutual information (NMI), adjusted rand index (AR), and F-Score, of which the detailed descriptions can be found in [Wu *et al.*, 2019; Larson, 2019], respectively. Seventeen state-of-the-art methods are adopted as baselines for comparison, including the AWP [Nie *et al.*, 2018b], MvDSCN [Zhu *et al.*, 2019], MLAN [Nie *et al.*, 2018a], UOMvSC [Tang *et al.*, 2023], EOMSC-CA [Liu *et al.*, 2022], LMSC [Zhang *et al.*, 2020], GMC [Wang *et al.*, 2020], MCLES [Chen *et al.*, 2020], FPMVS-CAG [Wang *et al.*, 2022], CLR-MVP [Kang *et al.*, 2023], t-SVD-MSC [Xie *et al.*, 2018], SM^{2}SC [Yang *et al.*, 2019], ETLMSC [Wu *et al.*, 2019], LMVSC [Kang *et al.*, 2020], E^{2}OMVC [Wang *et al.*, 2023], RMSL [Li *et al.*, 2019], and OMVFC-LICAG [Zhang *et al.*, 2024], among which six are developed within the last two years.

For the baseline methods, we follow the parameters in the original papers. For the proposed method, we set the parameters in the following way. For all balancing parameters, we tune them within the set {0.001,0.01,0.1,1,10,100,1000}. For ρ, κ, and N, we fix them to 0.001, 1.5, and 5 throughout the paper. If not otherwise clarified, we use a third-order CCL-MVC in the experiment. For all methods, the final clustering step is repeated 10 times and we report the averaged results with parameters tuned to the best.

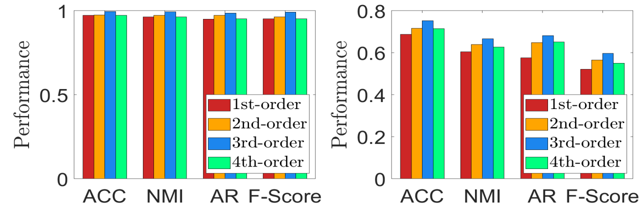
6.1 Clustering Performance

We compare the CCL-MVC with the baseline methods and report the clustering results in Section 5. In general, the CCL-MVC has the best performance among all methods, where it obtains the best results in all 24 cases. Among the baseline methods, the RMSL, SM^{2}SC, ETLMSC and t-SVD-MSC are among the most competitive ones, which obtain the top three results in 5, 11, 19, and 8 out of a total number of 24 cases, respectively. Compared with the baseline methods, the CCL-MVC has significantly improved performance, where it improves the performance by about 0.04-0.12, 0.01-0.04, 0.02-0.07, 0.01-0.03, 0.14-0.17, and 0.05-0.11 in different metrics on the BBC-4view, BBC-Sport, Flowers, UCI-3view, StillDB, and MITindoor data sets, respectively. Among these data sets, the StillDB data set is considered as a "difficult" one, on which the baseline methods rarely obtain results higher than 0.4. On this data set, the CCL-MVC improves the performance by about 0.14-0.17 and 0.29-0.35 in different



(a) BBC-Sport (b) StillDB

Figure 1: Comparison of the CCL-MVC models to show the significance of cross-view diversity embedding.



(a) BBC-Sport (b) StillDB

Figure 2: Comparison of the CCL-MVC models to show the significance of cross-order neighbor relationship preservation.

metrics, compared with the top second and third methods, i.e., the ETLMSC and SM^{2}SC, respectively. Such improvements are indeed significant. Moreover, the CCL-MVC has superior performance in terms of its stability. In particular, the CCL-MVC has the best performance on all data sets, while none of the baseline methods have the top three performance on all data sets. For example, as the most competitive method among the baselines, the ETLMSC is not among the top three on the BBC-4view data set; although the RMSL has the top second performance on the BBC-4view data set, it fails on several other data sets. These observations confirm the effectiveness and stability of the CCL-MVC on these data sets.

6.2 Ablation Study

We conduct ablation study from two perspectives, using the BBC-Sport and StillDB data sets without loss of generality. First, we validate the significance of adopting cross-view diversity and present the results in Fig. 1. In particular, we test how the CCL-MVC performs when the cross-view diversity embedding is eliminated from the model, for which we set $\beta = 0$ and tune the other parameters in the same way as described in Section 6.1. From Fig. 1, we may observe that the performance of the CCL-MVC is significantly degraded when the diversity is not considered, which confirms the significance of cross-view diversity embedding.

Then, we verify the significance of adopting cross-order neighbor information and show the results in Fig. 2. Compared with the first-order CCL-MVC, the second and third-order models have significantly improved performance, which confirms the significance of adopting cross-order neighbor information in our model. Moreover, when the order is higher than 3, the CCL-MVC cannot be further improved, or may even degrade. This may be explained by the fact that a very high order may introduce redundancy or noise to the relationships. Given the difficulty in finding a suitable K, and considering the significance of incorporating the third-order information, it is reasonably convincing to recommend utilizing a third-order CCL-MVC in practical applications.

6.3 Convergence Study

Besides the theoretical results about convergence provided in Section 5, we further show some empirical results to illustrate the convergent behavior of the CCL-MVC. Without loss of generality, we show the results on the BBC-4view and StillDB data sets in Fig. 3. Due to space limit, rather than ploting the curves for each variable, we plot the curve of error sequence, where at the t-th iteration the error is defined as $\max\{\|\mathcal{Z}^t - \mathcal{Z}^{t-1}\|_F, \|\mathcal{E}^t - \mathcal{E}^{t-1}\|_F, \|\mathcal{Q}^t - \mathcal{Q}^{t-1}\|_F, \|\mathrm{Z}_0^t - \mathrm{Z}_0^{t-1}\|_F, \|w^t - w^{t-1}\|_2, \|\mathcal{P}_K - \mathcal{Z}^t - \mathcal{E}^t\|_F, \|\mathcal{Z}^t - \mathcal{Q}^t\|_F\}$. From the results, it is seen that the curves converge to zero within about 50 iterations, which is quite efficient. Similar observations can be found on other data sets as well, which is convincing to claim the convergent behavior of the CCL-MVC.

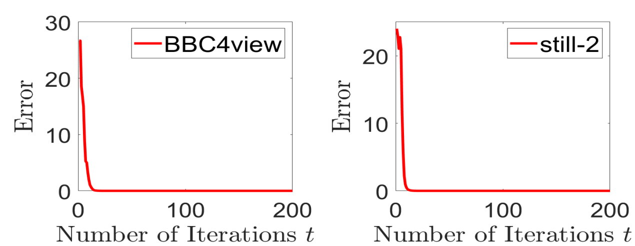


Figure 3: Error curves of the CCL-MVC on BBC-4view and StillDB data sets, with parameters fixed to $\lambda = 0.01$, $\alpha = 0.001$, and $\beta = 1$.

7 Conclusion

In this paper, we propose a novel CCL-MVC method for multi-view clustering. The CCL-MVC incorporates cross-view diversity to learn a consensus affinity matrix by fusing a low-rank essential tensor recovered from a fine-grained neighbor tensor that encompasses comprehensive and complementary cross-order information of multi-view data. We develop an effective optimization algorithm for the CCL-MVC, which is proved to converge with theoretical guarantee. Extensive experimental results show the superiority of the CCL-MVC.

Acknowledgements

Y.C. and C.C. are corresponding authors. This work is supported in part by the National Natural Science Foundation of China (NSFC) under Grants 62276147, 62172246, and 62106063; in part by the Shandong Province Colleges and Universities Youth Innovation Technology Plan Innovation Team Project under Grants 2022KJ149, 2021KJ062, and 2020KJN011; and in part by the Guangdong Major Project of Basic and Applied Basic Research under Grant 2023B0303000010.

References

[Brbić and Kopriva, 2018] Maria Brbić and Ivica Kopriva. Multi-view low-rank sparse subspace clustering. *Pattern Recognition*, 73:247–258, 2018.

[Chen *et al.*, 2020] Man-Sheng Chen, Ling Huang, Chang-Dong Wang, and Dong Huang. Multi-view clustering in latent embedding space. *Proceedings of the AAAI Conference on Artificial Intelligence*, page 3513–3520, Jun 2020.

[Chen *et al.*, 2021] Yongyong Chen, Shuqin Wang, Chong Peng, Zhongyun Hua, and Yicong Zhou. Generalized non-convex low-rank tensor approximation for multi-view subspace clustering. *IEEE Transactions on Image Processing*, 30:4022–4035, 2021.

[Elhamifar and Vidal, 2009] Ehsan Elhamifar and René Vidal. Sparse subspace clustering. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2790–2797. IEEE, 2009.

[Fu *et al.*, 2023] Lele Fu, Zhaoliang Chen, Yongyong Chen, and Shiping Wang. Unified low-rank tensor learning and spectral embedding for multi-view subspace clustering. *IEEE Transactions on Multimedia*, 25:4972–4985, 2023.

[Gao *et al.*, 2021] 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.

[Kang *et al.*, 2020] Zhao Kang, Wangtao Zhou, Zhitong Zhao, Junming Shao, Meng Han, and Zenglin Xu. Large-scale multi-view subspace clustering in linear time. *Proceedings of the AAAI Conference on Artificial Intelligence*, page 4412–4419, Jun 2020.

[Kang *et al.*, 2022] Zhao Kang, Zhanyu Liu, Shirui Pan, and Ling Tian. Fine-grained attributed graph clustering. In *Proceedings of the 2022 SIAM International Conference on Data Mining (SDM)*, pages 370–378, 2022.

[Kang *et al.*, 2023] Kehan Kang, Chenglizhao Chen, and Chong Peng. Consensus low-rank multi-view subspace clustering with cross-view diversity preserving. *IEEE Signal Processing Letters*, 30:1512–1516, 2023.

[Kumar *et al.*, 2011] Abhishek Kumar, Piyush Rai, and Hal Daumé. Co-regularized multi-view spectral clustering. In *Proceedings of the 24th International Conference on Neural Information Processing Systems*, NIPS'11, pages 1413–1421, Red Hook, NY, USA, 2011. Curran Associates Inc.

[Larson, 2019] Ray R. Larson. Introduction to information retrieval. *J. Assoc. Inf. Sci. Technol.*, 61:852–853, 2019.

[Li *et al.*, 2019] Ruihuang Li, Changqing Zhang, Huazhu Fu, Xi Peng, Tianyi Zhou, and Qinghua Hu. Reciprocal multi-layer subspace learning for multi-view clustering. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 8172–8180, 2019.

[Liu *et al.*, 2010] Guangcan Liu, Zhouchen Lin, and Yong Yu. Robust subspace segmentation by low-rank representation. In *Proceedings of the 27th International Conference on Machine Learning*, pages 663–670, 2010.

[Liu *et al.*, 2022] Suyuan Liu, Siwei Wang, Pei Zhang, Kai Xu, Xinwang Liu, Changwang Zhang, and Feng Gao. Efficient one-pass multi-view subspace clustering with consensus anchors. *Proceedings of the AAAI Conference on Artificial Intelligence*, page 7576–7584, Jul 2022.

[Lu *et al.*, 2016] Canyi Lu, Jiashi Feng, Yudong Chen, Wei Liu, Zhouchen Lin, and Shuicheng Yan. Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5249–5257, 2016.

[Ng *et al.*, 2001] Andrew Ng, Michael Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, volume 14, pages 849–856, 2001.

[Ng *et al.*, 2023] Tsz Ching Ng, Siu Kai Choy, Shu Yan Lam, and Kwok Wai Yu. Fuzzy superpixel-based image segmentation. *Pattern Recognition*, 134:109045, 2023.

[Nie *et al.*, 2018a] Feiping Nie, Guohao Cai, Jing Li, and Xuelong Li. Auto-weighted multi-view learning for image clustering and semi-supervised classification. *IEEE Transactions on Image Processing*, page 1501–1511, Mar 2018.

[Nie *et al.*, 2018b] Feiping Nie, Lai Tian, and Xuelong Li. Multiview clustering via adaptively weighted procrustes. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2022–2030, 2018.

[Pan *et al.*, 2023] Baicheng Pan, Chuandong Li, and Hangjun Che. Nonconvex low-rank tensor approximation with graph and consistent regularizations for multi-view subspace learning. *Neural Networks*, 161:638–658, 2023.

[Park *et al.*, 2022] Namyong Park, Ryan Rossi, Eunyee Koh, Iftikhar Ahamath Burhanuddin, Sungchul Kim, Fan Du, Nesreen Ahmed, and Christos Faloutsos. Cgc: Contrastive graph clustering forcommunity detection and tracking. In *The Web Conference (WWW)*, pages 1115–1126, 2022.

[Peng *et al.*, 2015] Chong Peng, Zhao Kang, Huiqing Li, and Qiang Cheng. Subspace clustering using log-determinant rank approximation. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 925–934. ACM, 2015.

[Peng *et al.*, 2017] Chong Peng, Zhao Kang, and Qiang Cheng. Subspace clustering via variance regularized ridge regression. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2931–2940, 2017.

[Peng *et al.*, 2020] Chong Peng, Yongyong Chen, Zhao Kang, Chenglizhao Chen, and Qiang Cheng. Robust principal component analysis: A factorization-based approach with linear complexity. *Information Sciences*, 513:581–599, 2020.

[Peng *et al.*, 2022a] Chong Peng, Yang Liu, Kehan Kang, Yongyong Chen, Xinxing Wu, Andrew Cheng, Zhao

Kang, Chenglizhao Chen, and Qiang Cheng. Hyperspectral image denoising using nonconvex local low-rank and sparse separation with spatial-spectral total variation regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–17, 2022.

[Peng *et al.*, 2022b] Chong Peng, Yiqun Zhang, Yongyong Chen, Zhao Kang, Chenglizhao Chen, and Qiang Cheng. Log-based sparse nonnegative matrix factorization for data representation. *Knowledge-Based Systems*, page 109127, 2022.

[Shi and Malik, 2000] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[Tang *et al.*, 2015] Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. Line: Large-scale information network embedding. In *Proceedings of the 24th international conference on world wide web*, pages 1067–1077, 2015.

[Tang *et al.*, 2023] Chang Tang, Zhenglai Li, Jun Wang, Xinwang Liu, Wei Zhang, and En Zhu. Unified one-step multi-view spectral clustering. *IEEE Transactions on Knowledge and Data Engineering*, 35(6):6449–6460, 2023.

[Wang *et al.*, 2020] Hao Wang, Yan Yang, and Bing Liu. Gmc: Graph-based multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, page 1116–1129, Jun 2020.

[Wang *et al.*, 2022] 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, 2022.

[Wang *et al.*, 2023] Jun Wang, Chang Tang, Zhiguo Wan, Wei Zhang, Kun Sun, and Albert Y. Zomaya. Efficient and effective one-step multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–12, 2023.

[Wen *et al.*, 2023] Jie Wen, Zheng Zhang, Lunke Fei, Bob Zhang, Yong Xu, Zhao Zhang, and Jinxing Li. A survey on incomplete multiview clustering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(2):1136–1149, 2023.

[Wu *et al.*, 2019] Jianlong Wu, Zhouchen Lin, and Hongbin Zha. Essential tensor learning for multi-view spectral clustering. *IEEE Transactions on Image Processing*, 28(12):5910–5922, 2019.

[Wu *et al.*, 2023] Yang Wu, Zhiwei Ge, Yuhao Luo, Lin Liu, and Sulong Xu. Face clustering via graph convolutional networks with confidence edges. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 20990–20999, 2023.

[Xia *et al.*, 2014] Rongkai Xia, Yan Pan, Lei Du, and Jian Yin. Robust multi-view spectral clustering via low-rank and sparse decomposition. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence,*, pages 2149–2155. AAAI Press, 2014.

[Xia *et al.*, 2021] Wei Xia, Xiangdong Zhang, Quanxue Gao, Xiaochuang Shu, Jungong Han, and Xinbo Gao. Multi-view subspace clustering by an enhanced tensor nuclear norm. *IEEE Transactions on Cybernetics*, 52(9):1–14, 2021.

[Xia *et al.*, 2022] Wei Xia, Quanxue Gao, Qianqian Wang, and Xinbo Gao. Tensor completion-based incomplete multiview clustering. *IEEE Transactions on Cybernetics*, 52(12):13635–13644, 2022.

[Xie *et al.*, 2018] Yuan Xie, Dacheng Tao, Wensheng Zhang, Yan Liu, Lei Zhang, and Yanyun Qu. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *International Journal of Computer Vision*, page 1157–1179, Nov 2018.

[Yang *et al.*, 2019] Zhiyong Yang, Qianqian Xu, Weigang Zhang, Xiaochun Cao, and Qingming Huang. Split multiplicative multi-view subspace clustering. *IEEE Transactions on Image Processing*, 28(10):5147–5160, 2019.

[Yu and Yang, 2023] Quan Yu and Ming Yang. Low-rank tensor recovery via non-convex regularization, structured factorization and spatio-temporal characteristics. *Pattern Recognition*, 137:109343, 2023.

[Zhan *et al.*, 2018] Kun Zhan, Changqing Zhang, Junpeng Guan, and Junsheng Wang. Graph learning for multiview clustering. *IEEE Transactions on Cybernetics*, 48(10):2887–2895, 2018.

[Zhang *et al.*, 2015] Changqing Zhang, Huazhu Fu, Si Liu, Guangcan Liu, and Xiaochun Cao. Low-rank tensor constrained multiview subspace clustering. In *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*, pages 1582–1590. IEEE Computer Society, 2015.

[Zhang *et al.*, 2020] Changqing Zhang, Huazhu Fu, Qinghua Hu, Xiaochun Cao, Yuan Xie, Dacheng Tao, and Dong Xu. Generalized latent multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, page 86–99, Jan 2020.

[Zhang *et al.*, 2024] Chuanbin Zhang, Long Chen, Zhaoyin Shi, and Weiping Ding. Latent information-guided one-step multi-view fuzzy clustering based on cross-view anchor graph. *Information Fusion*, 102:102025, 2024.

[Zhou and Burges, 2007] Dengyong Zhou and Christopher JC Burges. Spectral clustering and transductive learning with multiple views. In *Proceedings of the 24th international conference on Machine learning*, pages 1159–1166, 2007.

[Zhou *et al.*, 2005] Dengyong Zhou, Jiayuan Huang, and Bernhard Schölkopf. Learning from labeled and unlabeled data on a directed graph. In *Proceedings of the 22nd international conference on Machine learning*, pages 1036–1043, 2005.

[Zhu *et al.*, 2019] Pengfei Zhu, Binyuan Hui, Changqing Zhang, Dawei Du, Longyin Wen, and Qinghua Hu. Multi-view deep subspace clustering networks. *arXiv preprint arXiv:1908.01978*, 2019.