Multiview Tensor Spectral Clustering via Co-Regularization

Hongmin Cai, *Senior Member, IEEE*, Yu Wang, Fei Qi, Zhuoyao Wang, and Yiu-ming Cheung, *Fellow, IEEE*

***Abstract*—Graph-based multi-view clustering encodes multi-view data into sample affinities to find consensus representation, effectively overcoming heterogeneity across different views. However, traditional affinity measures tend to collapse as the feature dimension expands, posing challenges in estimating a unified alignment that reveals both cross-view and inner relationships. To tackle this challenge, we propose to achieve multi-view uniform clustering via consensus representation co-regularization. First, the sample affinities are encoded by both popular dyadic affinity and recent high-order affinities to comprehensively characterize spatial distributions of the HDLSS data. Second, a fused consensus representation is learned through aligning the multi-view low-dimensional representation by co-regularization. The learning of the fused representation is modeled by a high-order eigenvalue problem within manifold space to preserve the intrinsic connections and complementary correlations of original data. A numerical scheme via manifold minimization is designed to solve the high-order eigenvalue problem efficaciously. Experiments on eight HDLSS datasets demonstrate the effectiveness of our proposed method in comparison with the recent thirteen benchmark methods.**

***Index Terms*—High-order affinity, clustering, fusing affinity, manifold optimization, tensor, spectral graph.**

Manuscript received 13 July 2023; revised 18 January 2024; accepted 6 April 2024. Date of publication 9 April 2024; date of current version 5 September 2024. This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFE0112200, in part by the Key-Area Research and Development of Guangdong Province under Grant 2022B1111050002, in part by the Key-Area Research and Development Program of Guangzhou City under Grant 202206030009, Grant 2023B01J1001, and Grant 2023B01J0022, in part by the National Natural Science Foundation of China under Grant 62325204, Grant U21A20520, and Grant 62172112, in part by the Guangdong Key Laboratory of Human Digital Twin Technology under Grant 2022B1212010004, in part by the NSFC / Research Grants Council (RGC) Joint Research Scheme under Grant N_HKBU214/21, in part by the General Research Fund of RGC under Grant 12201321, Grant 12202622, and Grant 12201323, and in part by the RGC Senior Research Fellow Scheme under Grant SRFS2324-2S02. Recommended for acceptance by X. Li. *(Corresponding author: Yiu-ming Cheung.)*

Hongmin Cai is with the School of Future Technology, South China University of Technology, Guangzhou 511442, China.

Yu Wang is with the School of Future Technology, South China University of Technology, Guangzhou 511442, China, and also with Peng Cheng Laboratory, Shenzhen 518055, China.

Fei Qi is with the School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China, and also with the School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang 550029, China.

Zhuoyao Wang is with the Peng Cheng Laboratory, Shenzhen 518055, China.

Yiu-ming Cheung is with the Department of Computer Science, Hong Kong Baptist University, Hong Kong, SAR, China (e-mail: ymc@comp.hkbu.edu.hk).

This article has supplementary downloadable material available at https://doi.org/10.1109/TPAMI.2024.3386828, provided by the authors.

Digital Object Identifier 10.1109/TPAMI.2024.3386828

I. Introduction

CLUSTERING is one of the crucial topics in unsupervised learning [1], [2], [3]. The goal of clustering is to partition unlabeled data into different subgroups. Traditional clustering methods have been extensively studied and have found wide applications in bioinformatics, computer vision, and other fields [4]. In real-world scenarios, the universal adoption of multi-view data is driven by the acquisition of data from different sources or multiple feature extractors [5], [6]. Consequently, the seamless integration of heterogeneous data has become a paramount focus in the realm of multi-view clustering [7], [8].

Recent studies have proposed different strategies to integrate complementary correlations from different views, with the aim of improving clustering performance [9], [10]. An intuitive approach begins by directly concatenating the data from different views as vectors. Subsequently, the conventional single-view methods are applied to the concatenated data [11]. However, this manner ignores the heterogeneity and difference of scale among multi-view data. Graph-based multi-view clustering methods are proposed to align the affinity graphs for the uniform representation [12]. Intuitively, the harmonization of heterogeneous multi-view data can be accomplished by employing affinities with same scale, thus reducing the disparities between different views. To achieve multi-view clustering, a joint analysis of the graphs across multiple views is required to extract consensus and complementary correlations. Accordingly, various approaches have been proposed to refine a consensus representation from multiple graphs [13]. For example, Kumar et al. [14] proposed to refine the affinity matrix via performing co-training of spectral results. However, there is a critical prerequisite in aforementioned methods: the data relationship can be accurately described by pairwise affinity. This can be challenging in real applications, especially for high-dimension m yet low-sample-size n (HDLSS) data when $n \ll m$ [15], [16]. The clustering performance of HDLSS data is hindered by the concentration effects, also known as the "curse of dimensionality" [17]. The collapse of pairwise distances in high-dimensional feature spaces presents a formidable challenge for clustering algorithms reliant on pairwise affinity, impeding accurate clustering [18].

For relieving the curse of dimensionality caused by HDLSS data, recent works have proposed to utilize high-order affinity to describe the spatial distribution of multiple samples. Employing high-order affinities is anticipated to mitigate the concentration effects on dyadic affinity and unveil crucial relationships within

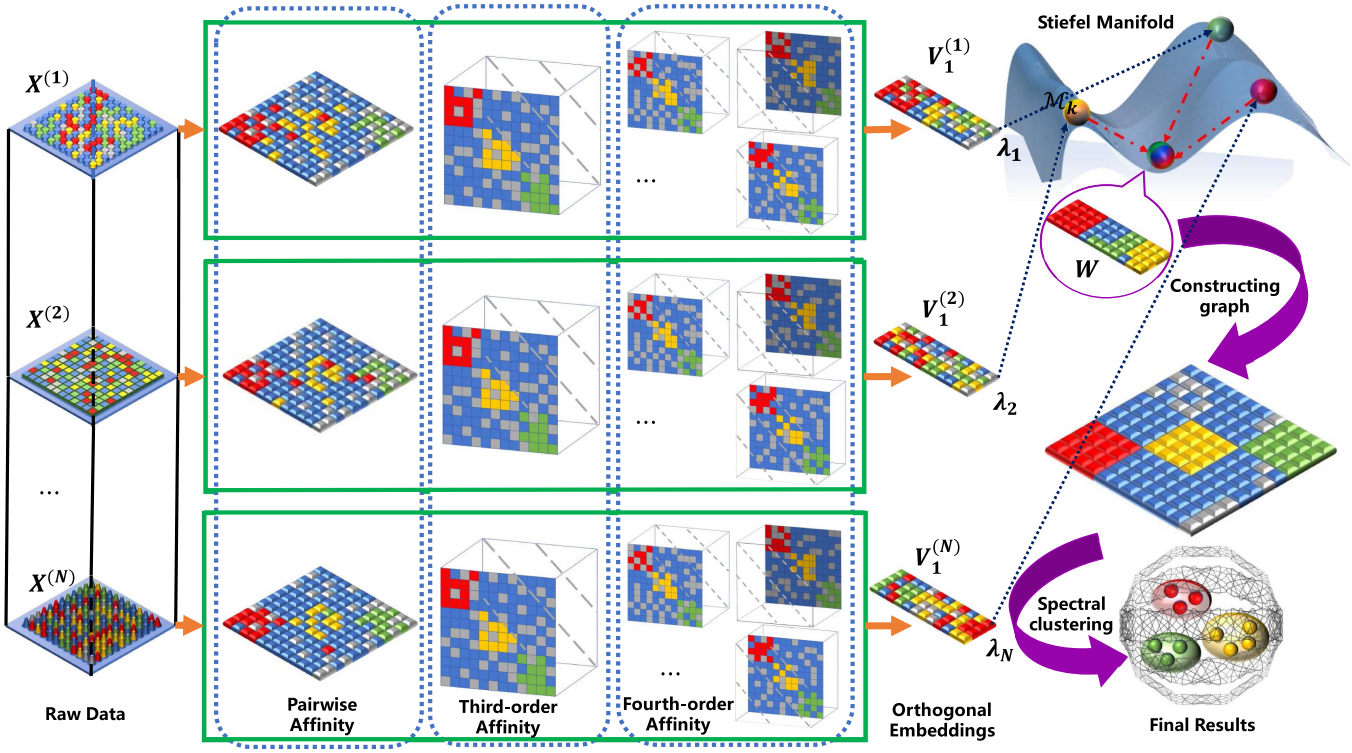


Fig. 1. Description of the CRMATS method. CRMATS demonstrates proficiency in accurately acquiring latent representations from multiple views on the manifold space. Specifically, given a multi-view dataset $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(N)}$, the corresponding pairwise affinities, third-order and fourth-order affinities are computed. Next, the local structure of the intrinsic subspace is encoded and low-dimensional representations $\boldsymbol{V}_1^{(1)}, \ldots, \boldsymbol{V}_1^{(N)}$ are obtained by employing the manifold constraint. To align the representations of each view, we incorporate co-regularized learning, resulting in a consensus representation $\boldsymbol{W}$. Additionally, a self-weighting fusion module is adopted to compute the corresponding weights $\lambda^{(1)}, \ldots, \lambda^{(N)}$ during the alignment process. Finally, we use spectral clustering on the consensus representation to obtain results.

the data. Along this line, Mei et al. [19] leveraged first and second-order affinities to mine the local structure of pairwise points, incorporate third-order similarity with a low-rank constraints for enhanced clustering performance and consensus correlations. Also, Ghoshdastidar [20] demonstrated a relationship between the relaxation of hypergraph clustering and the multilinear singular value decomposition with consideration of multiply affinities. Furthermore, IPS2 [21] validates that the higher-order affinities can be a complementary description of pairwise relations to enhance clustering performance. Following this vein, our previous work [22] unifies different order affinities to overcome the concentration effects, leading to remarkable estimating the spatial distribution of HDLSS data. The aforementioned methodology exhibits a deficiency to integrate multi-view data, thereby impeding the extraction of heterogeneous correlations.

To solve the key problems of extracting consensus correlations and accurately revealing relationships from multi-view data, especially for HDLSS data, we propose the **C**o-**R**egularized multi-view clustering via **M**anifold **A**lignment on **T**ensor **S**pectral embedding (*CRMATS*) method. Our method presents a unified multi-view clustering framework, which is based on the accurate description of intra-view sample relationships through the introduction of multi-order affinities. The low-dimensional representations from each view are co-regularized on the manifold space, aiming to minimize the geodesic distance and achieve alignment. This alignment process enhances the quality of the consensus representation, ultimately leading to improved partitioning performance as feedback. To accelerate the solution of the unified model, an efficient iterative strategy is designed to solve this model efficaciously. Extensive experiments on both synthetic and real-world datasets validate nice performance of our method. The framework of CRMATS is shown in Fig. 1, and the main contributions of this paper are summarized as follows:

- To precisely uncover the intra-view spatial correlations of HDLSS data, the incorporation of high-order affinities is employed, facilitating the capture of intricate sample interactions and effectively eliminating the concentration effects within each view.
- For effective integration multi-view data with heterogeneous correlations, co-regularized learning and manifold constraint are employed to align the respective low-dimensional representations, effectively leveraging cross-view spatial complementarity of HDLSS data.
- To improve computational efficiency, a singular value decomposition-based method is utilized to solve the quadratic problem on the manifold space.

The remainder of this paper is organized as follows. Section II introduces the background on multi-view and tensor spectral

clustering. Section III presents the proposed method and its optimization algorithm. In Section IV, we report the experimental results on the comparative datasets and methods. Finally, we draw a conclusion in Section V.

II. Background on Multiview Tensor Spectral Clustering

Notations: In this paper, we use bold calligraphy, uppercase letter, and lowercase letter to represent a tensor, matrix, and a vector, i.e., $(\boldsymbol{\mathcal{X}}, \boldsymbol{X}, \boldsymbol{x})$, respectively. For a matrix $\boldsymbol{X} \in \mathbb{R}^{n\times n}$, the j-th column is denoted as $\boldsymbol{X}_{:,j}$, the trace of $\boldsymbol{X}$ is represented as $Tr(\boldsymbol{X})$, and $\boldsymbol{I}$ represents the identity matrix. The square of the Frobenius norm of $\boldsymbol{X}$ is denoted as $||\boldsymbol{X}||_F^2$. The Khatri-Rao and Kronecker products are denoted by $*$ and $\otimes$, respectively. For a third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n\times n\times l}$, we denote the v-th frontal slices of $\boldsymbol{\mathcal{X}}$ as $\boldsymbol{X}^{(v)} \in \mathbb{R}^{n\times n}$.

A. Tensor Based Approaches for Multiview Clustering

Tensor-based clustering techniques harness the high-order representation of multi-view data through tensors, thereby elucidating the intricate inter-view relationships prior to performing cluster analysis [23]. In this process, most of the early works relied on tensor decomposition techniques. For example, Yu et al. [24] stacked the original data into a tensor and applied tensor-based factorization to obtain factor matrices that capture high-order relationships. Similarly, Nie et al. [13] proposed a co-clustering method via tensor factorization to learn a low-rank approximation for discovering high-order relationships. Following this approach, Guo et al. [25] utilized tensor logarithmic Schatten-p norm to obtain a more compact low-rank structure, which explores complementary information and characterizes the high-order correlations among multiple views. Meanwhile, Ji et al. [26] employed tensor decomposition to generate consistent and complementary tensors, while refining a tighter approximation of the tensor rank to explore the high-order consistency in consistent tensor. Similarly, Li et al. [27] stacked pairwise affinities into a tensor and employed a hypergraph-induced regularization for tensor factorization, enabling them to learn a consistent representation that preserved high-order correlations and improved performance. The aforementioned methods require the assurance that data relationships can be accurately described by pairwise affinities, which is difficult to hold in HDLSS data. Additionally, the utilization of tensors to represent high-order information between views does not leverage the high-order relationships between samples.

B. Revisiting Classical Spectral Clustering

Spectral clustering is a classical method that employs dyadic affinity to learn an optimal low-dimensional embedding from raw data for clustering purposes. Given a data matrix $\mathbf{X} \in \mathbb{R}^{n\times m}$, where n is the number of samples and m is the feature dimension, the objective of spectral clustering is to divide these n samples into c subgroups by reformulating the clustering problem as a minimum cost problem of graph-cut. The crucial step of this method is to construct a similarity graph by calculating the dyadic affinity matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. Specifically, the (i, j)-th element of $\boldsymbol{A}$ is calculated as $\boldsymbol{A}_{ij} = d(\boldsymbol{x}_i, \boldsymbol{x}_j)$, where $d(\cdot,\cdot)$ is a pairwise metric. The i-th diagonal element of the degree matrix $\boldsymbol{D} \in \mathbb{R}^{n\times n}$ is denoted as $\boldsymbol{D}_{ii} = \sum_{j=1}^{n} \boldsymbol{A}_{ij}$. The Laplacian matrix is then defined as $\boldsymbol{L} = \boldsymbol{I} - \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{D}^{-\frac{1}{2}}$. The spectral clustering seeks a low-dimensional embedding by minimizing the following objective model:

$$\min_{\boldsymbol{V}} Tr(\boldsymbol{V}^\top \boldsymbol{L}\boldsymbol{V}) \qquad \text{s.t.}\, \boldsymbol{V}^\top\boldsymbol{V} = \boldsymbol{I} \tag{1}$$

Equation (1) can be equivalently solved by seeking the dominant eigenvector of the Laplacian matrix, and thus degenerates to a standard eigenvalue problem. Moreover, this equation can be seen as a formulation that addresses the maximum partition problem in graph-cut, as described in [28]. Alternatively, one can define a normalized dyadic affinity matrix $\hat{\boldsymbol{L}} = \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{D}^{-\frac{1}{2}}$, other that the early ones to impose a strong positive definite Laplacian graph. The spectral clustering can also be popularly expressed as a maximization problem :

$$\max_{\boldsymbol{V}} Tr(\boldsymbol{V}^\top \hat{\boldsymbol{L}}\boldsymbol{V}) \qquad \text{s.t.}\, \boldsymbol{V}^\top\boldsymbol{V} = \boldsymbol{I} \tag{2}$$

Then, one can perform a clustering task like k-means on the obtained embedding.

C. Tensor Spectral Clustering

The core of model (2) is to maximize the intra-cluster affinity, thereby preserving the volume of each subgraph after graph cut. In our previous work [22], we introduced a normalized affinity entropy measurement, which effectively evaluates the volume of affinity using any number of samples.

Definition 1. The total normalized similarity: Let $\boldsymbol{C}$ be a group of samples belonging to the dataset $\boldsymbol{X}$, and $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{n_1\times n_2\cdots\times n_r}$ be an order-r similarity tensor. The total normalized similarity of the samples in $\boldsymbol{C}$ is defined as follows:

$$Sim(\boldsymbol{C}) = \sum_{\boldsymbol{x}_{i_1},\boldsymbol{x}_{i_2},\ldots,\boldsymbol{x}_{i_r}\in\boldsymbol{C}} \boldsymbol{\mathcal{L}}_{\boldsymbol{x}_{i_1},\boldsymbol{x}_{i_2},\ldots,\boldsymbol{x}_{i_r}} \tag{3}$$

where $\boldsymbol{\mathcal{L}}$ represents the normalized order-r affinity tensor. Let the samples $\boldsymbol{X}$ partitioned by k sets, i.e., $\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_k$, the normalized associativity (N_{Assoc}) of the resulting clustering is defined:

$$N_{Assoc}(\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_k) = \sum_{i=1}^{k} \frac{Sim(\boldsymbol{C}_i)}{|\boldsymbol{C}_i|^k} \tag{4}$$

where $|\boldsymbol{C}_i|$ denotes the cardinality of cluster $\boldsymbol{C}_i$.

Definition 2: mode-k product The mode-k product between an order-m tensor $\boldsymbol{\mathcal{T}} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_m}$ and a matrix $\boldsymbol{V} \in \mathbb{R}^{p\times n_k}$, denoted by $\boldsymbol{\mathcal{T}} \otimes_k \boldsymbol{V} \in \mathbb{R}^{n_1\times\cdots\times n_{k-1}\times p\times n_{k+1}\times\cdots\times n_m}$, with

$$(\boldsymbol{\mathcal{T}} \otimes_k \boldsymbol{V})_{i_1\ldots i_{k-1} j i_{k+1}\ldots i_m} = \sum_{i_k=1}^{n_k} \boldsymbol{\mathcal{T}}_{i_1\ldots i_{k-1} i_k i_{k+1}\ldots i_m} \boldsymbol{V}_{j i_k}. \tag{5}$$

Let $\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_k] \in \mathbb{R}^{n\times k}$ denote the latent representation, where $\boldsymbol{H}_{ik} = |\boldsymbol{C}_k|^{-1}$ if $\boldsymbol{x}_i \in \boldsymbol{C}_k$, and zero otherwise.

To obtain an optimal sample assignment $\boldsymbol{C}_1, \ldots, \boldsymbol{C}_k$, the maximization of the normalized associativity in (4) is pursued. This can be achieved through algebraic manipulation, which allows us to reformulate the problem as follows:

$$\max_{\boldsymbol{C}_1,\ldots,\boldsymbol{C}_k} \sum_{j=1}^{k} (\boldsymbol{\mathcal{L}} \otimes_r \boldsymbol{h}_j \otimes_{r-1} \boldsymbol{h}_j \cdots \otimes_1 \boldsymbol{h}_j) \tag{6}$$

Due to the NP-hardness of the maximum normalized associativity problem, it is necessary to use relaxation techniques to make the problem more tractable. One such technique involves relaxing the binary assignment matrix to an orthonormal matrix $\boldsymbol{V} \in \mathbb{R}^{n\times k}$, where $\boldsymbol{V}^\top \boldsymbol{V} = \boldsymbol{I}$. This relaxation reduces the strict binary assignment requirement and simplifies the problem as:

$$\max_{\boldsymbol{V}^\top \boldsymbol{V}=\boldsymbol{I}} \sum_{j=1}^{k} \boldsymbol{\mathcal{L}} \otimes_r \boldsymbol{v}_j \otimes_{r-1} \boldsymbol{v}_j \cdots \otimes_1 \boldsymbol{v}_j \tag{7}$$

where $\boldsymbol{v}_j$ represents the j-th column of $\boldsymbol{V}$.

D. Co-Regularized Multiview Clustering

In multi-view clustering, learning a consensus representation is the common approach to capture the local view-specific structure [29]. Based on this hypothesis, co-regularized learning is applied to align the low-dimensional representation of different views, reducing the impact of noise or errors and enhancing the quality of the consensus representation [30]. Given multi-view data $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n\times m\times l}$ with n samples, m features, and l views, one can construct the corresponding affinities $\boldsymbol{L}^{(i)} \in \mathbb{R}^{n\times n}$ of each view and project into the low-dimensional space to obtain representations $\boldsymbol{V}^{(i)} \in \mathbb{R}^{n\times k}$. The corresponding strategy is to align diverse low-dimensional representations and obtain a consensus representation $\boldsymbol{W} \in \mathbb{R}^{n\times k}$. This alignment essentially establishes a consensus representation by minimizing the geodesic distance on the manifold [31], and the measurement of geodesic distance can be:

$$\begin{aligned} dist^2_{proj}\left(\left\{\boldsymbol{V}^{(i)}\right\}_{i=1}^{l}, \boldsymbol{W}\right) &= \sum_{i=1}^{l} dist^2_{proj}\left(\left\{\boldsymbol{V}^{(i)}\right\}, \boldsymbol{W}\right) \\ &= \sum_{i=1}^{l} [k - Tr(\boldsymbol{V}^{(i)}\boldsymbol{V}^{(i)\top}\boldsymbol{W}\boldsymbol{W}^\top)] \\ &= kl - \sum_{i=1}^{l} Tr(\boldsymbol{V}^{(i)}\boldsymbol{V}^{(i)\top}\boldsymbol{W}\boldsymbol{W}^\top) \end{aligned} \tag{8}$$

where $\boldsymbol{V}^{(i)}$ is obtained through (2) to learn the low-dimensional representations within a single subspace, and the optimization of (8) aims to minimize geodesic distances and promote inter-subspace proximity [32]. Building upon these principles, the co-regularized learning process can be as:

$$\begin{aligned} \max_{\substack{\boldsymbol{V}^{(i)}\in\boldsymbol{\mathcal{M}}_k,\\ \boldsymbol{W},\lambda^{(i)}}} & \sum_{i=1}^{l} Tr(\boldsymbol{V}^{(i)\top}\hat{\boldsymbol{L}}_2^{(i)}\boldsymbol{V}^{(i)} + \boldsymbol{V}^{(i)}\boldsymbol{V}^{(i)\top}\boldsymbol{W}\boldsymbol{W}^\top) \\ \text{s.t. } & \boldsymbol{W}^\top\boldsymbol{W} = \boldsymbol{I} \end{aligned} \tag{9}$$

where $\hat{\boldsymbol{L}}_2^{(i)} \in \mathbb{R}^{n\times n}$ is the pairwise affinity of the i-th view. However, noise or errors of different views can adversely affect multi-view performance [33]. In order to tackle this challenge, $\lambda^{(i)}$ is introduced to control the influence of the deteriorating views, and the final co-regularized model with the manifold constraints can be as:

$$\begin{aligned} \max_{\substack{\boldsymbol{V}^{(i)}\in\boldsymbol{\mathcal{M}}_k,\\ \boldsymbol{W},\lambda^{(i)}}} & \sum_{i=1}^{l} Tr(\boldsymbol{V}^{(i)\top}\hat{\boldsymbol{L}}_2^{(i)}\boldsymbol{V}^{(i)} + \lambda^{(i)}\boldsymbol{V}^{(i)}\boldsymbol{V}^{(i)\top}\boldsymbol{W}\boldsymbol{W}^\top) \\ \text{s.t. } & \boldsymbol{W}^\top\boldsymbol{W} = \boldsymbol{I} \end{aligned} \tag{10}$$

where $\lambda^{(i)}$ is used to measure the weight of the i-th view.

III. Multi-View Tensor Spectral Clustering via Co-Regularization

Existing multi-view methods primarily rely on linear spaces, rendering them inadequate for analyzing multi-dimensional data characterized by intrinsic complex structures. Fortunately, manifolds can be conceptualized as the low-dimensional smooth surfaces embedded within higher-dimensional euclidean spaces, providing a framework that enables the effective capture and comprehension of the intricate structures in high-dimensional data [34], [35]. Follow this vein, Khan et al. [36] introduce manifold-based methods that effectively capture complex structures, leading to significant improvements in clustering performance. These techniques enable the representation of spatial structures, making them particularly advantageous for analyzing HDLSS data. Inspired by these manifold-based work, a unified tensor clustering model can be developed by combining high-order affinities, mitigating the concentration effects in HDLSS data. To enhance the discriminative power of the low-dimensional representation for HDLSS tasks, we impose a constraint that restricts the embedding to the manifold space. In this paper, to demo the effectiveness of our framework in handling both odd-order and even-order affinities, we introduce the third and fourth-order affinities. Thus, a model that incorporates high-order affinities and manifold constraints is proposed:

$$\begin{aligned} \min_{\boldsymbol{V}\in\boldsymbol{\mathcal{M}}_k} & \sum_{j=1}^{k} -\hat{\boldsymbol{L}}_2 \otimes_2 \boldsymbol{v}_j \otimes_1 \boldsymbol{v}_j - \boldsymbol{\mathcal{L}}_3 \otimes_3 \boldsymbol{v}_j \otimes_2 \boldsymbol{v}_j \otimes_1 \boldsymbol{v}_j \\ & - \boldsymbol{\mathcal{L}}_4 \otimes_4 \boldsymbol{v}_j \otimes_3 \boldsymbol{v}_j \otimes_2 \boldsymbol{v}_j \otimes_1 \boldsymbol{v}_j \end{aligned} \tag{11}$$

where $\boldsymbol{\mathcal{M}}_k = \{\boldsymbol{V} \in \mathbb{R}^{n\times k} | \boldsymbol{V}^\top\boldsymbol{V} = \boldsymbol{I}\}$ is the Stiefel manifold, $\boldsymbol{\mathcal{L}}_3$ and $\boldsymbol{\mathcal{L}}_4$ are triadic and tetradic affinity tensors for quantifying similarities among triplets and quadruplets data.

A. Co-Regularized Multi-View Clustering via Manifold Alignment on Tensor Spectral Embedding

Equation (11) achieves the low-dimensional representation on the manifold space by integrating multi-order affinities. Building upon this, a co-regularized learning step with the manifold constraints is introduced for handling the concentration effects and extracting heterogeneous correlations in multi-view HDLSS clustering. The goal of the proposed method is to effectively leverage cross-view correlations in HDLSS data by combining

(11) and (10). Consequently, a **C**o-**R**egularized multi-view clustering via **M**anifold **A**lignment on **T**ensor **S**pectral embedding (*CRMATS*) method is proposed as follows:

$$\begin{aligned}\min_{\substack{\boldsymbol{V}_1^{(i)}\in\boldsymbol{\mathcal{M}}_k,\\ \boldsymbol{W},\lambda^{(i)}}} & -\sum_{i=1}^{l} Tr\Big(\boldsymbol{V}_1^{(i)^\top}\hat{\boldsymbol{L}}_2^{(i)}\boldsymbol{V}_1^{(i)}+\Big(\boldsymbol{V}_1^{(i)}*\boldsymbol{V}_1^{(i)}\Big)^\top\hat{\boldsymbol{L}}_3^{(i)}\boldsymbol{V}_1^{(i)}\\ &+\Big(\boldsymbol{V}_1^{(i)}*\boldsymbol{V}_1^{(i)}\Big)^\top\hat{\boldsymbol{L}}_4^{(i)}\Big(\boldsymbol{V}_1^{(i)}*\boldsymbol{V}_1^{(i)}\Big)\\ &+\lambda^{(i)}\Big(\boldsymbol{V}_1^{(i)}\boldsymbol{V}_1^{(i)^\top}\boldsymbol{W}\boldsymbol{W}^\top\Big)\Big)\\ &\text{s.t.}\quad \boldsymbol{W}^\top\boldsymbol{W}=\boldsymbol{I}\end{aligned}\tag{12}$$

where $\hat{\boldsymbol{L}}_3$ and $\hat{\boldsymbol{L}}_4$ are the normalized unfold form of tensor affinities, calculated by:

$$\hat{\boldsymbol{L}}_3=\hat{\boldsymbol{D}}_{3_1}^{-\frac{1}{2}}\hat{\boldsymbol{T}}_3\hat{\boldsymbol{D}}_{3_2}^{-\frac{1}{2}},\hat{\boldsymbol{L}}_4=\hat{\boldsymbol{D}}_4^{-\frac{1}{2}}\hat{\boldsymbol{T}}_4\hat{\boldsymbol{D}}_4^{-\frac{1}{2}}\tag{13}$$

where $\hat{\boldsymbol{D}}_{3_1}$ and $\hat{\boldsymbol{D}}_{3_2}$ are the diagonal matrices with diagonal elements obtained by taking the square root of the Khatri-Rao product of the column-wise sums and the column-wise sums of $\boldsymbol{\mathcal{T}}_3$, respectively. $\boldsymbol{D}_4$ is a diagonal matrix in which the elements on the diagonal are computed as the reciprocal square root of the sum of each row of $\boldsymbol{\mathcal{T}}_4$. The matrices $\hat{\boldsymbol{T}}_3$ and $\hat{\boldsymbol{T}}_4$ are the unfold form of third and fourth order affinity tensors. Tensor affinities $\boldsymbol{\mathcal{T}}_3$ and $\boldsymbol{\mathcal{T}}_4$ can be defined as:

$$\boldsymbol{\mathcal{T}}_{3_{ijk}}=1-\frac{<(\boldsymbol{x}_i-\boldsymbol{x}_j),(\boldsymbol{x}_k-\boldsymbol{x}_j)>}{\boldsymbol{d}_{ij}\boldsymbol{d}_{jk}}\tag{14}$$

$$\boldsymbol{\mathcal{T}}_{4_{ijkl}}=exp\left(-\sigma\frac{\boldsymbol{d}_{ij}+\boldsymbol{d}_{kl}}{\boldsymbol{d}_{ik}+\boldsymbol{d}_{jl}+\varepsilon}\right)\tag{15}$$

for $i,j,k,l\in n$, where $\boldsymbol{d}_{ij}$ denotes the distance between samples $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, and σ is an scaling constant. However, the computation in (12) involves high-order polynomial function that may lead to numerical difficulties. To circumvent this, one can introduce a slack variable $\boldsymbol{V}_2^{(i)}$ to approximate the term $\boldsymbol{V}_1^{(i)}*\boldsymbol{V}_1^{(i)}$ in (12). Furthermore, we introduce γ_1,γ_2 to trade off the affinities term and co-regularization term. The final model can be:

$$\begin{aligned}\min_{\substack{\boldsymbol{V}_1^{(i)}\in\boldsymbol{\mathcal{M}}_k,\\ \boldsymbol{V}_2^{(i)},\boldsymbol{W},\lambda^{(i)}}} & -\gamma_1\sum_{i=1}^{l}Tr\Big(\boldsymbol{V}_1^{(i)^\top}\hat{\boldsymbol{L}}_2^{(i)}\boldsymbol{V}_1^{(i)}+\boldsymbol{V}_2^{(i)^\top}\hat{\boldsymbol{L}}_3^{(i)}\boldsymbol{V}_1^{(i)}\\ &+\boldsymbol{V}_2^{(i)^\top}\hat{\boldsymbol{L}}_4^{(i)}\boldsymbol{V}_2^{(i)}\Big)-\gamma_2\sum_{i=1}^{l}\lambda^{(i)}Tr\left(\boldsymbol{V}_1^{(i)}\boldsymbol{V}_1^{(i)^\top}\boldsymbol{W}\boldsymbol{W}^\top\right)\\ &\text{s.t. } \boldsymbol{V}_1^{(\mathrm{i})}*\boldsymbol{V}_1^{(\mathrm{i})}=\boldsymbol{V}_2^{(\mathrm{i})},\mathrm{i}=1,\ldots,\mathrm{N};\quad \boldsymbol{W}^\top\boldsymbol{W}=\boldsymbol{I}\end{aligned}\tag{16}$$

Equation (16) aims to fuse multiple affinities to produce a consistent representation that is robust against noise and concentration effects. The clustering task is then accomplished by applying spectral clustering, leading to the final group assignments.

B. Numerical Scheme to Solve CRMATS

An efficient alternating direction minimization strategy is employed to solve CRMATS. Using the Augmented Lagrange formulation methodology, the corresponding function of (16), defined as $\mathbb{L}$, is obtained by:

$$\begin{aligned}&\mathbb{L}\left(\boldsymbol{V}_1^{(i)},\boldsymbol{V}_2^{(i)},\boldsymbol{Y}_i,\boldsymbol{W},\lambda^{(i)},\boldsymbol{Z},\mu_1^{(i)},\mu_2\right)\\ &=-\gamma_1 Tr\sum_{i=1}^{l}\Big(\boldsymbol{V}_1^{(i)^\top}\hat{\boldsymbol{L}}_2^{(i)}\boldsymbol{V}_1^{(i)}+\boldsymbol{V}_2^{(i)^\top}\hat{\boldsymbol{L}}_3^{(i)}\boldsymbol{V}_1^{(i)}\\ &+\boldsymbol{V}_2^{(i)^\top}\hat{\boldsymbol{L}}_4^{(i)}\boldsymbol{V}_2^{(i)}\Big)-\gamma_2\sum_{i=1}^{l}\lambda^{(i)}Tr(\boldsymbol{V}_1^{(i)}\boldsymbol{V}_1^{(i)^\top}\boldsymbol{W}\boldsymbol{W}^\top)\\ &+<\boldsymbol{Z},\boldsymbol{W}^\top\boldsymbol{W}-\boldsymbol{I}>+\sum_{i=1}^{l}\frac{\mu_1^{(i)}}{2}\left\|\boldsymbol{V}_1^{(i)}*\boldsymbol{V}_1^{(i)}-\boldsymbol{V}_2^{(i)}\right\|_F^2\\ &+\sum_{i=1}^{l}<\boldsymbol{Y}_i,\boldsymbol{V}_1^{(i)}*\boldsymbol{V}_1^{(i)}-\boldsymbol{V}_2^{(i)}>+\frac{\mu_2}{2}\|\boldsymbol{W}^\top\boldsymbol{W}-\boldsymbol{I}\|_F^2\end{aligned}\tag{17}$$

where $\boldsymbol{Y}_i$ and $\boldsymbol{Z}$ are Lagrange multipliers, $\mu_1^{(i)}$ and $\mu_2>0$ are penalty parameters. Our goal is to minimize $\mathbb{L}$ by dividing it into several subproblems. We accomplish this by considering the following variables alternatively and solving each variable while keeping the others fixed.

Step 1): Solving the subproblem with respect to the variable $\boldsymbol{V}_1^{(i)}$

When keeping the related items, (17) can be:

$$\begin{aligned}\min_{\boldsymbol{V}_1^{(i)}\in\boldsymbol{\mathcal{M}}_k}\ &\gamma_1\sum_{i=1}^{l}Tr\Big(\boldsymbol{V}_1^{(i)^\top}\left(-\left(\hat{\boldsymbol{L}}_2^{(i)}+\frac{\gamma_2}{\gamma_1}\lambda^{(i)}\boldsymbol{W}\boldsymbol{W}^\top\right)\right)\boldsymbol{V}_1^{(i)}\\ &-2\boldsymbol{V}_1^{(i)^\top}\left(\tfrac{1}{2}\hat{\boldsymbol{L}}_3^{(i)^\top}\boldsymbol{V}_2^{(i)}\right)\Big)+\frac{\mu_1^{(i)}}{2}\left\|\boldsymbol{V}_1^{(i)}*\boldsymbol{V}_1^{(i)}-\boldsymbol{V}_2^{(i)}+\frac{\boldsymbol{Y}_i}{\mu_1^{(i)}}\right\|_F^2\end{aligned}\tag{18}$$

In (18), we encounter a quadratic term $-(\hat{\boldsymbol{L}}_2^{(i)}+\frac{\gamma_2}{\gamma_1}\lambda^{(i)}\boldsymbol{W}\boldsymbol{W}^\top)$ and a first-order term $\frac{1}{2}\hat{\boldsymbol{L}}_3^{(i)^\top}\boldsymbol{V}_2^{(i)}$. By incorporating the manifold constraint, (18) can be simplified as a quadratic optimization problem within each view. To solve this problem, we introduce a universal Stiefel manifold $\boldsymbol{\mathcal{M}}_p=\{\boldsymbol{V}\in\mathbb{R}^{n\times k}|\boldsymbol{V}^\top\boldsymbol{V}=\boldsymbol{I}\}$. The quadratic and first-order terms for each view are denoted as $\boldsymbol{A}$ and $\boldsymbol{B}$.

$$\min_{\boldsymbol{V}\in\boldsymbol{\mathcal{M}}_p}Tr(\boldsymbol{V}^\top\boldsymbol{A}\boldsymbol{V}-2\boldsymbol{V}^\top\boldsymbol{B})\tag{19}$$

where $\boldsymbol{A}\in\mathbb{R}^{n\times n}$ is a symmetric matrix. In order to solve the problem in (19), it can be relaxed into:

$$\max_{\boldsymbol{V}\in\boldsymbol{\mathcal{M}}_p}Tr(\boldsymbol{V}^\top\hat{\boldsymbol{A}}\boldsymbol{V}+2\boldsymbol{V}^\top\boldsymbol{B})\tag{20}$$

where $\hat{\boldsymbol{A}}=\alpha\boldsymbol{I}-\boldsymbol{A}\in\mathbb{R}^{n\times n}$. The parameter α is an arbitrary constant such that $\hat{\boldsymbol{A}}$ is a positive definite matrix. A closed-form solution to (20) can be achieved through the corresponding

Algorithm 1: Generalized Power Iteration Method (GPI).

Input: The matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, matrix $\boldsymbol{B} \in \mathbb{R}^{n\times k}$ and α
Output: The orthonormal base $\boldsymbol{V}$ of Stiefel manifold $\mathcal{M}_p = \{\boldsymbol{V} \in \mathbb{R}^{n\times k} | \boldsymbol{V}^\top \boldsymbol{V} = \boldsymbol{I}\}$.

1: Initialize a random orthonormal base of Stiefel manifold, definite matrix $\hat{\boldsymbol{A}} = \alpha \boldsymbol{I} - \boldsymbol{A}$ via α.
2: **while** Not Converged **do**
3: Compute the gradient, $\boldsymbol{M} \in \mathbb{R}^{n\times k} \leftarrow 2\hat{\boldsymbol{A}}\boldsymbol{V} + 2\boldsymbol{B}$.
4: Calculate $\boldsymbol{M} = \boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{R}^\top$ via the compact SVD of $\boldsymbol{M}$.
5: Update $\boldsymbol{V} \in \mathbb{R}^{n\times k} \leftarrow \boldsymbol{F}\boldsymbol{R}^\top$.
6: **end while**
7: **return**The orthonormal base $\boldsymbol{V}$.

partial derivative $\boldsymbol{M}$ [37]. Suppose that $\boldsymbol{M} \leftarrow \boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{R}^\top$ represents the details of the singular value decomposition, in which $\boldsymbol{F} \in \mathbb{R}^{n\times n}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times k}$ $\boldsymbol{R} \in \mathbb{R}^{k\times k}$. Thus, (20) can be solved equally:

$$\max_{\boldsymbol{V}\in\mathcal{M}_k} Tr(\boldsymbol{V}^\top \boldsymbol{M}) \tag{21}$$

Then we have the following equations:

$$\begin{aligned} Tr(\boldsymbol{V}^\top \boldsymbol{M}) &= Tr(\boldsymbol{V}^\top \boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{R}^\top) = Tr(\boldsymbol{\Sigma}\boldsymbol{R}^\top \boldsymbol{V}^\top \boldsymbol{F}) \\ &= Tr(\boldsymbol{\Sigma}\boldsymbol{Z}) = \sum_{i=1}^{k} \sigma_{ii}\boldsymbol{z}_{ii} \end{aligned} \tag{22}$$

where $\boldsymbol{Z} = \boldsymbol{R}^\top \boldsymbol{V}^\top \boldsymbol{F} \in \mathbb{R}^{k\times n}$. Apparently, we conduct that $\boldsymbol{z}_{ii} \leq 1$ since $\boldsymbol{Z}\boldsymbol{Z}^\top = \boldsymbol{I}$. In this way, $Tr(\boldsymbol{V}^\top \boldsymbol{M})$ reaches maximum when the matrix $\boldsymbol{Z} = [\boldsymbol{I}, \boldsymbol{0}] \in \mathbb{R}^{k\times n}$. Thus, the optimal solution of $\boldsymbol{V}$ can be:

$$\boldsymbol{V} = \boldsymbol{F}\boldsymbol{Z}^\top \boldsymbol{R}^\top = \boldsymbol{F}[\boldsymbol{I}; \boldsymbol{0}]\boldsymbol{R}^\top \tag{23}$$

Algorithm 1 provides a concise description of the quadratic optimization algorithm on the manifold. For a comprehensive analysis of its convergence, we refer readers to Section A of the supplementary materials, available online. Once we have derived the objective function for the subproblem, we can utilize Algorithm 1 to obtain the solution for $\boldsymbol{V}^{(i)}$. As we are imposing the orthogonal constraint of manifold, the gradient of a differentiable function $f : \boldsymbol{V}_{k,m} \to \mathbb{R}$ can be given as $\nabla_{\boldsymbol{X}} f = f_{\boldsymbol{X}} - \boldsymbol{X} f_{\boldsymbol{X}}^\top \boldsymbol{X}$ with $f_{\boldsymbol{X}} = \frac{\partial f}{\partial \boldsymbol{X}}$ [38]. Thus, the gradient of $\boldsymbol{V}_1^{(i)}$ in (18) is:

$$\begin{aligned} \frac{\partial \mathbb{L}}{\partial \boldsymbol{V}_1^{(i)}} = \gamma_1 \Bigg(& -2\left(\hat{\boldsymbol{L}}_2^{(i)} + \frac{\gamma_2}{\gamma_1}\lambda^{(i)}\boldsymbol{W}\boldsymbol{W}^\top\right)\boldsymbol{V}_1^{(i)} \\ & + \boldsymbol{V}_1^{(i)}\boldsymbol{V}_1^{(i)\top}\left(\hat{\boldsymbol{L}}_2^{(i)} + \frac{\gamma_2}{\gamma_1}\lambda^{(i)}\boldsymbol{W}\boldsymbol{W}^\top\right)^\top \boldsymbol{V}_1^{(i)} \\ & - \hat{\boldsymbol{L}}_3^{(i)\top}\boldsymbol{V}_2^{(i)} + \boldsymbol{V}_1^{(i)}\boldsymbol{V}_2^{(i)\top}\hat{\boldsymbol{L}}_3^{(i)}\boldsymbol{V}_1^{(i)}\Bigg) \\ & + \mu_1^{(i)}\Bigg(\sum_{j=1}^{k}\left(\boldsymbol{V}_{1:j}^{(i)} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{V}_{1:j}^{(i)}\right)^\top \\ & \left(\boldsymbol{V}_{1:j}^{(i)} * \boldsymbol{V}_{1:j}^{(i)} - \boldsymbol{V}_{2:j}^{(i)} + \boldsymbol{Y}_{i:j}/\mu_1^{(i)}\right) \\ & + \boldsymbol{V}_1^{(i)}\Bigg(\sum_{j=1}^{k}\left(\boldsymbol{V}_{1:j}^{(i)} * \boldsymbol{V}_{1:j}^{(i)} - \boldsymbol{V}_{2:j}^{(i)} + \boldsymbol{Y}_{i:j}/\mu_1^{(i)}\right)^\top \\ & \left(\boldsymbol{V}_{1:j}^{(i)} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{V}_{1:j}^{(i)}\right)\boldsymbol{V}_1^{(i)}\Bigg) \end{aligned} \tag{24}$$

where $\boldsymbol{V}_1^{(i)}$ can be updated iteratively through (24) and Algorithm 1.

Step 2): Solving the subproblem with respect to the variable $\boldsymbol{V}_2^{(i)}$

Discarding the items irrelevant to $\boldsymbol{V}_2^{(i)}$, the augmented Lagrange function can be simplified as:

$$\begin{aligned} \min_{\boldsymbol{V}_2^{(i)}} & -\gamma_1 \sum_{i=1}^{l} Tr\left(\boldsymbol{V}_2^{(i)\top}\hat{\boldsymbol{L}}_3^{(i)}\boldsymbol{V}_1^{(i)} + \boldsymbol{V}_2^{(i)\top}\hat{\boldsymbol{L}}_4^{(i)}\boldsymbol{V}_2^{(i)}\right) \\ & + \frac{\mu_1^{(i)}}{2}\left\|\boldsymbol{V}_1^{(i)} * \boldsymbol{V}_1^{(i)} - \boldsymbol{V}_2^{(i)} + \frac{\boldsymbol{Y}_i}{\mu_1^{(i)}}\right\|_F^2 \end{aligned} \tag{25}$$

The gradient of the objective function is:

$$\begin{aligned} \frac{\partial \mathbb{L}}{\partial \boldsymbol{V}_2^{(i)}} = & -\gamma_1\left(\hat{\boldsymbol{L}}_3^{(i)}\boldsymbol{V}_1^{(i)} + 2\hat{\boldsymbol{L}}_4^{(i)}\boldsymbol{V}_2^{(i)}\right) \\ & + \mu_1^{(i)}\left(\boldsymbol{V}_2^{(i)} - \boldsymbol{V}_1^{(i)} * \boldsymbol{V}_1^{(i)} - \frac{\boldsymbol{Y}_i}{\mu_1^{(i)}}\right) \end{aligned} \tag{26}$$

By setting the gradient to zero, $\hat{\boldsymbol{L}}_4^{(i)}$ is symmetric and its diagonal elements are not yet zero, one can obtain the implicit solution as:

$$\begin{aligned} \boldsymbol{V}_2^{(i)^*} = & \left(\mu_1^{(i)}\boldsymbol{I} - 2\gamma_1\hat{\boldsymbol{L}}_4^{(i)}\right)^{-1} \\ & \left(\mu_1^{(i)}\boldsymbol{V}_1^{(i)} * \boldsymbol{V}_1^{(i)} + \gamma_1\hat{\boldsymbol{L}}_3^{(i)}\boldsymbol{V}_1^{(i)} + \boldsymbol{Y}_i\right) \end{aligned} \tag{27}$$

Step 3): Solving the subproblem with respect to the latent representation $\boldsymbol{W}$.

When the terms associated with $\boldsymbol{W}$ are kept, the following subproblem is obtained:

$$\max_{\boldsymbol{W}} \gamma_2 Tr\left(\boldsymbol{W}^\top\left(\sum_{i=1}^{l}\lambda^{(i)}\boldsymbol{V}_1^{(i)}\boldsymbol{V}_1^{(i)\top}\right)\boldsymbol{W}\right) \quad s.t. \boldsymbol{W}^\top\boldsymbol{W} = \boldsymbol{I} \tag{28}$$

where (28) is actually a standard k-means problems with a specific kernel $\sum_{i=1}^{l}\lambda^{(i)}\boldsymbol{V}_1^{(i)}\boldsymbol{V}_1^{(i)\top}$. We can obtain $\boldsymbol{W}$ via eigenvalue decomposition on this specific matrix.

Step 4): Solving the subproblem with respect to $\lambda^{(i)}$.

$$\max_{\lambda^{(i)}} \gamma_2 Tr\left(\sum_{i=1}^{l}\lambda^{(i)}\boldsymbol{V}_1^{(i)}\boldsymbol{V}^{(i)\top}\boldsymbol{W}\boldsymbol{W}^\top\right) \tag{29}$$

By defining $Tr(\boldsymbol{V}_1^{(i)}\boldsymbol{V}^{(i)\top}\boldsymbol{W}\boldsymbol{W}^\top) = d^{(i)}$, and combining the Cauchy-Schwarz inequality [39], the optimal solution for $\lambda^{(i)}$ can be obtained as:

$$\lambda^{(i)} = \frac{d^{(i)}}{\sqrt{\sum_{i=1}^{l} d^{(i)^2}}}, \tag{30}$$

Algorithm 2: **C**o-**R**egularized Multi-View Clustering via **M**anifold **A**lignment on **T**ensor **S**pectral Embedding (CR-MATS).

Input: Multi-view dataset $\mathcal{X} \in \mathbb{R}^{n\times m\times l}$, Cluster number c.
Output: The consensus matrix $\boldsymbol{W}$.
1: Construct the affinity $\hat{\boldsymbol{L}}_2^{(i)} \in \mathbb{R}^{n\times n}, \hat{\boldsymbol{L}}_3^{(i)} \in \mathbb{R}^{n^2\times n}, \hat{\boldsymbol{L}}_4^{(i)} \in \mathbb{R}^{n^2\times n^2}$. Set $\boldsymbol{Y}_i = \boldsymbol{0}, \boldsymbol{W} = \boldsymbol{0}, \boldsymbol{Z} = \boldsymbol{0}, \mu_1^{(i)} = \mu_2 = 10^{-1}, \mu^{max} = 10^6, \rho = 1.1\, and\, \epsilon = 10^{-6}$.
2: Initialize $\boldsymbol{V}_1^{(i)}, \boldsymbol{V}_2^{(i)}$.
3: **while** Not Converged **do**
4: **for** i $= 1, \ldots, l$ **do**
4: Update $\boldsymbol{V}_1^{(i)}$ via (24) as well as Algorithm 1.
4: Computer $\boldsymbol{V}_2^{(i)}$ by (27).
4: Update $\boldsymbol{Y}_i$ and $\lambda^{(i)}$ by (31) and (30).
4: $\mu_1^{(i)} = \min(\rho\mu_1^{(i)}, \mu^{\max})$.
5: **end for**
6: $\mu_2 = \min(\rho\mu_2, \mu^{\max})$.
7: Update $\boldsymbol{Z}$ and $\boldsymbol{W}$ via (32) as well as (28).
8: Check the convergence conditions.
9: **end while**
10: **return**Consensus representation $\boldsymbol{W}$.
11: Perform spectral clustering on $\boldsymbol{W}$ to have sample assignment $\boldsymbol{Y}_{pred}$.

Step 5): Updating the multipliers $\boldsymbol{Y}_i$ *and* $\boldsymbol{Z}$, their formulations are follows:

$$\boldsymbol{Y}_i^{(t+1)} = \boldsymbol{Y}_i^{(t)} + \mu_1^{(i)}(\boldsymbol{V}_1^{(i)(t)} * \boldsymbol{V}_1^{(i)(t)} - \boldsymbol{V}_2^{(i)(t)}) \tag{31}$$

$$\boldsymbol{Z}^{(t+1)} = \boldsymbol{Z}^{(t)} + \mu_2(\boldsymbol{W}^{\top(t)}\boldsymbol{W}^{(t)} - \boldsymbol{I}) \tag{32}$$

where t is current number of iterations.

The five steps are iteratively updated until convergence or until a stopping criterion is met: $\max(\|\boldsymbol{V}_1^{(i)(t+1)} - \boldsymbol{V}_1^{(i)(t)}\|_\infty, \|\boldsymbol{V}_2^{(i)(t+1)} - \boldsymbol{V}_2^{(i)(t)}\|_\infty, \|\boldsymbol{V}_1^{(i)(t+1)} \otimes \boldsymbol{V}_1^{(i)(t+1)} - \boldsymbol{V}_2^{(i)(t+1)}\|_\infty, \|\boldsymbol{Y}_1^{(i)(t+1)} - \boldsymbol{Y}_1^{(i)(t)}\|_\infty, \|\boldsymbol{Z}^{(t+1)} - \boldsymbol{Z}^{(t)}\|_\infty) < \epsilon$. Algorithm 2 presents a comprehensive outline of the solving process in CRMATS, serving as a valuable reference for a detailed understanding of the methodology. The convergence proof of CRMATS is elaborated in Section A of the supplementary materials, available online.

IV. Experiments

In this section, a comprehensive experimental study is conducted on eight HDLSS datasets to showcase the effectiveness of CRMATS. All of the experiments are implemented in Matlab 2020a on 64-bit Windows OS PC with an Intel 2.30-GHz CPU.

A. Comparative Datasets and Methods

A total of eight datasets are utilized to validate the effectiveness of our method in the experiment, comprising six real datasets and two HDLSS synthetic datasets: Syndata1 and Syndata2. Syndata1 consists of 120 samples divided into two categories, with each category containing 60 samples. Each sample is described from three views. To verify the robustness of our methods on HDLSS data, we extend the dimensions and the number of views, and reduce the sample size on Syndata1 to obtain Syndata2. Specifically, Syndata2 comprises 90 samples divided into 3 categories, with each sample being described from four views. Each subcategory of synthetic data is generated from independent and identically distributed normal distributions with the mean values of 2 and the standard deviations of 0.5. In addition, we evaluate the effectiveness of CRMATS in six public benchmark datasets, including Coil-20 [40], MSRC_v1 [41], Yale [42], BBCSport [43], 3Sources [44] and Reuters [44]. To demonstrate the effectiveness of our method on HDLSS datasets, we randomly select samples from these datasets for experiments. More details about the datasets are provided in Table I, and our method is compared with other multi-view clustering methods. The details of the comparative methods are shown as follows:

TABLE I
Statistics on Tested Datasets

Dataset	Instances	Views	Features	Classes	Types
Syndata1	120	3	2352,5097,9998	2	Synthetic
Syndata2	90	4	2341,3988,7236,16996	3	Synthetic
Coil-20	100	3	1024, 3304, 6750	5	Image
Yale	165	3	4096, 3304, 6750	15	Image
MSRC_v1	90	5	24,576,512,256,254	2	Image
BBCSport	78	2	3183, 3203	3	Text
3Sources	169	3	3560, 3631, 3068	6	Text
Reuters	120	5	21526,24892,34121,15487,11539	6	Text

I. *Scalable Multi-view Subspace Clustering (SMSC)* [45] constructs latent graph after anchor learning.
II. *Pure graph-guided multi-view subspace clustering (PGSC)* [46] learns consensus graph by leveraging the sparsity and connectivity of each affinity graph.
III. *Robust Multi-View Spectral Clustering (RMSC)* [12] considers low rankness and sparsity of matrix to learn a common graph after decomposition.
IV. *Low-rank Tensor Based Proximity Learning (LTBPL)* [42] performs probability affinity to recover the low rankness and high-order correlations.
V. *Multiview Subspace Clustering via Low-Rank Symmetric Affinity Graph (LSGMC)* [44] pursues a consistent low-rank structure across views.
VI. *Measuring Diversity in Graph Learning: A Unified Framework for Structured Multi-View Clustering (CDMGC)* [47] leverages the multi-view consistency and the diversity in a unified framework.
VII. *Co-regularized kernel k-means for multi-view clustering (Co-reg)* [39] combines similarities of different view and latent representation for clustering.
VIII. *Multiview Clustering via Co-Training Robust Representation (CoMSC)* [48] finds a consensus matrix and complementary information.
IX. *Efficient Multi-view Graph Clustering (EMGC)* [49] finds a consistent cluster indicator matrix with a Super Nodes Similarity Minimization module.

X. *Multi-View Clustering on Topological Manifold (MVCT)* [50] integrates multiple affinity graphs into a consensus one with the topological relevance.
XI. *A Tensor Approach for Uncoupled Multiview Clustering (T-UMC)* [51] couples the representation matrix to explore high-order relationship.
XII. *Multiview Subspace Clustering by an Enhanced Tensor Nuclear Norm (WTSNM)* [52] studies the Schatten p-norm to solve the minimization problem.
XIII. *High-order Complementarity Multi-View Clustering with Enhanced Tensor Rank (HCETR)* [26] adopts Tensor Rank to find high-order consistency.

B. Clustering Performance

The clustering performance is evaluated using several commonly employed metrics, including Accuracy (ACC), Normalized Mutual Information (NMI), Purity, and Fscore [32]. A larger value denotes superior performance. Consequently, the best results are highlighted in bold. Considering that the clustering problem does not include the number of groups present in the data, we further use the Calinski Harabasz index (CHI) to further assess the quality of the clustering results. CHI measures the compactness and separation of clustering outcomes by evaluating the ratio of inter-class variance to intra-class variance [53]. A higher CHI value indicates better clustering results. In our evaluation, a comprehensive assessment is provided using five metrics, with each metric representing a specific property of the clustering outcomes. The clustering performance of our method and comparison methods on eight benchmark datasets is presented in Tables II and III. Moreover, we report the CHI values between the ground truth labels and the features for each dataset. To ensure the robustness, each algorithm is repeated 20 times to obtain the mean value. We then use Student's t-test to test the statistical significance of the results, with the p-value represented in parentheses.

1) Experiments on Synthetic Datasets: Experiments are conducted to assess the stability of our method on Syndata1 and Syndata2. The results are summarized in Tables II and III. Notably, our method demonstrates substantial improvements in terms of NMI compared to other methods on Syndata1. Specifically, we observe improvements of 37.95%, 41.34%, 58.90%, 67.53%, 23.37%, 42.26%, 47.04% and 21.67% when compared to Co-reg, CoMSC, LTBPL, LSGMC, MVCT, T-UMC, WTSNM, and HCETR, respectively. In relation to Table III, the results of CRMATS align with the ground truth labels, exhibiting a consistent value of 2.13. Regarding Syndata2, our method outperforms Co-reg, CoMSC, LTBPL, LSGMC, MVCT, T-UMC, WTSNM, and HCETR by 73.42%, 66.89%, 22.23%, 63.25% , 18.63%, 33.54%, 52.09% and 36.19%. At this point, the evaluation result in Table III is 1.86, surpassing the results of the other comparative methods.

To assess the discriminative power of the consensus representation, we utilize t-SNE for visualizing the differences by projecting the latent representation onto a two-dimensional space. Syndata2, which has the highest dimensionality, is specifically selected to represent the synthetic scenarios. When visualizing the raw data from their respective views using t-SNE (Fig. 2(a)–(d)), it is evident that most of the samples in Syndata2 appear intermingled and lack clear separation. In contrast, the consensus representation achieved by our method successfully separates subcategories without any overlap (Fig. 2(e)). For the purpose of further validating the effectiveness of CRMATS, an analysis of the heatmaps of similarities is performed to evaluate the distinctions between groups. In Fig. 2(f)–(i), the affinity heatmap of the raw samples lacks clear boundaries and block structures. However, Fig. 2(j) shows that the affinity obtained from the low-dimensional embedding after applying CRMATS exhibits distinct boundaries, indicating the method's ability to mitigate potential biases. The results indicate that the fusion of different order affinities outperforms the traditional pairwise affinity, low-rankness, and tensor-based methods in fully capturing the data structure on synthetic data.

2) Experiments on Real Datasets: To further validate the efficacy of CRMATS in real-world scenarios, we evaluate its performance on several benchmark datasets, including Coil-20, Yale, MSRC_v1, BBCSport, 3Sources, and Reuters. Detailed information and corresponding results can be found in Tables I, II and III, respectively. On Coil-20, CRMATS demonstrates superior performance compared to LTBPL, LSGMC, MVCT, CoMSC, Co-reg, T-UMC, WTSNM, and HCETR, with NMI clustering results exceeding them by 29.12%, 21.27%, 82.64%, 63.11%, 8.94%, 25.56%, 58.76%, and 39.59%, respectively. Likewise, on Yale, CRMATS outperforms Co-reg, CoMSC, LTBPL, LSGMC, MVCT, T-UMC, WTSNM, and HCETR with performance improvements of 15.99%, 39.47%, 22.20%, 19.96%, 28.60%, 12.23%, 2.30%, and 20.11%. Comparing against alternative methods such as SMSC, RMSC, PGSC, CDMGC, and EMGC2F, CRMATS achieves significantly higher performance, surpassing them by 32.39%, 1.58%, 77.91%, 18.66%, and 17.31%. Furthermore, CRMATS exhibits superior performance on BBCSport, MSRC_v1, 3Sources , and Reuters. Notably, the NMI result of CRMATS on MSRC_v1 reaches 93.90%, making it the second-best performer. Additionally, CRMATS achieves a perfect NMI result of 100.00% on BBCSport, surpassing the second-best performance of 82.81% achieved by T-UMC. In the case of the 3Sources dataset, CRMATS achieves an NMI result of 87.46%, which is over 16.18% higher than the second-best method, T-UMC. Similarly, on the Reuters dataset, CRMATS achieves an NMI result of 77.38%, surpassing the second-best method, EMGC2F, by 11.39%. Table III presents the CHI results of CRMATS alongside those of the comparison methods in the dataset. Notably, in the BBCSport dataset, CRMATS achieves results that perfectly align with the ground truth labels. Across the Coil-20, Yale, MSRC_v1, 3Sources, and Reuters datasets, the respective CHI results are 349.00, 58.37, 12.87, 9.52, and 4.40. These findings provide compelling evidence of the outstanding performance of CRMATS in the remaining comparison datasets, further corroborating the trends observed in Table II.

TABLE II
COMPARISON RESULTS (%): THE MEAN AND p VALUE MEASURED BY DIFFERENT CLUSTERING METHODS ON ALL THE CORRESPONDING DATASETS

Dataset	Method	ACC	NMI	Purity	Fscore	Dataset	ACC	NMI	Purity	Fscore
Syndata1	SMSC	$83.30_{(0.00)}$	$46.91_{(0.00)}$	$83.32_{(0.00)}$	$73.32_{(0.00)}$	Syndata2	$56.72_{(0.00)}$	$30.61_{(0.09)}$	$62.21_{(0.00)}$	$52.21_{(0.00)}$
	RMSC	$51.72_{(0.00)}$	$10.91_{(0.00)}$	$51.74_{(0.00)}$	$49.51_{(0.01)}$		$56.11_{(0.08)}$	$17.91_{(0.00)}$	$55.60_{(0.00)}$	$43.39_{(0.00)}$
	PGSC	$50.82_{(0.00)}$	$0.81_{(0.00)}$	$59.17_{(0.01)}$	$65.90_{(0.00)}$		$2.31_{(0.01)}$	$2.30_{(0.00)}$	$67.81_{(0.00)}$	$48.61_{(0.00)}$
	LTBPL	$3.33_{(0.00)}$	$41.10_{(0.00)}$	$53.12_{(0.00)}$	$41.11_{(0.01)}$		$52.19_{(0.02)}$	$73.30_{(0.01)}$	$54.37_{(0.00)}$	$49.21_{(0.00)}$
	LSGMC	$70.00_{(0.00)}$	$32.47_{(0.00)}$	$68.88_{(0.00)}$	$53.52_{(0.00)}$		$53.33_{(0.00)}$	$32.28_{(0.00)}$	$54.44_{(0.00)}$	$49.49_{(0.00)}$
	CDMGC	$51.67_{(0.00)}$	$61.69_{(0.00)}$	$51.67_{(0.00)}$	$65.57_{(0.00)}$		$40.00_{(0.00)}$	$4.25_{(0.00)}$	$41.12_{(0.00)}$	$37.73_{(0.00)}$
	Co-reg	$58.31_{(0.00)}$	$62.05_{(0.00)}$	$58.28_{(0.01)}$	$50.60_{(0.00)}$		$38.91_{(0.00)}$	$22.11_{(0.00)}$	$38.94_{(0.00)}$	$31.17_{(0.00)}$
	CoMSC	$74.44_{(0.00)}$	$58.66_{(0.00)}$	$74.44_{(0.01)}$	$58.66_{(0.00)}$		$76.67_{(0.00)}$	$28.64_{(0.00)}$	$76.67_{(0.00)}$	$28.64_{(0.00)}$
	EMGC2F	$55.10_{(0.00)}$	$0.71_{(0.00)}$	$55.03_{(0.00)}$	$49.47_{(0.00)}$		$60.01_{(0.00)}$	$25.21_{(0.01)}$	$60.03_{(0.00)}$	$45.91_{(0.00)}$
	MVCT	$51.71_{(0.00)}$	$76.63_{(0.03)}$	$91.72_{(0.04)}$	$62.53_{(0.03)}$		$35.63_{(0.01)}$	$76.90_{(0.01)}$	$65.61_{(0.00)}$	$38.61_{(0.00)}$
	T-UMC	$53.83_{(0.00)}$	$57.74_{(0.00)}$	$53.83_{(0.00)}$	$52.51_{(0.01)}$		$77.33_{(0.02)}$	$61.99_{(0.00)}$	$77.33_{(0.00)}$	$67.19_{(0.00)}$
	WTSNM	$60.00_{(0.00)}$	$52.96_{(0.00)}$	$60.00_{(0.00)}$	$51.62_{(0.00)}$		$56.67_{(0.00)}$	$43.44_{(0.00)}$	$56.67_{(0.00)}$	$51.81_{(0.00)}$
	HCETR	$78.33_{(0.01)}$	$64.93_{(0.00)}$	$78.33_{(0.01)}$	$65.65_{(0.00)}$		$80.00_{(0.00)}$	$59.34_{(0.00)}$	$80.00_{(0.00)}$	$69.72_{(0.00)}$
	CRMATS	$\mathbf{100.00}_{(0.00)}$	$\mathbf{100.00}_{(0.00)}$	$\mathbf{100.00}_{(0.00)}$	$\mathbf{100.00}_{(0.00)}$		$\mathbf{98.92}_{(0.01)}$	$\mathbf{95.53}_{(0.05)}$	$\mathbf{97.71}_{(0.02)}$	$\mathbf{98.91}_{(0.01)}$
Coil-20	SMSC	$94.01_{(0.00)}$	$89.10_{(0.02)}$	$94.04_{(0.00)}$	$88.41_{(0.00)}$	Yale	$50.91_{(0.00)}$	$55.73_{(0.09)}$	$52.71_{(0.00)}$	$36.41_{(0.00)}$
	RMSC	$48.01_{(0.00)}$	$69.42_{(0.00)}$	$60.03_{(0.00)}$	$55.22_{(0.01)}$		$73.25_{(0.00)}$	$86.54_{(0.01)}$	$73.91_{(0.00)}$	$50.11_{(0.00)}$
	PGSC	$42.00_{(0.00)}$	$4.15_{(0.00)}$	$86.00_{(0.00)}$	$41.53_{(0.00)}$		$7.91_{(0.00)}$	$10.21_{(0.00)}$	$61.51_{(0.00)}$	$12.11_{(0.00)}$
	LTBPL	$68.51_{(0.21)}$	$67.53_{(0.00)}$	$78.51_{(0.00)}$	$73.93_{(0.00)}$		$60.04_{(0.00)}$	$62.92_{(0.00)}$	$60.02_{(0.00)}$	$36.41_{(0.00)}$
	LSGMC	$84.00_{(0.00)}$	$75.38_{(0.00)}$	$84.12_{(0.00)}$	$74.86_{(0.00)}$		$64.85_{(0.00)}$	$68.16_{(0.00)}$	$62.23_{(0.00)}$	$45.84_{(0.00)}$
	CDMGC	$80.00_{(0.00)}$	$77.52_{(0.00)}$	$80.00_{(0.00)}$	$75.41_{(0.00)}$		$66.67_{(0.00)}$	$69.46_{(0.00)}$	$67.27_{(0.00)}$	$48.63_{(0.00)}$
	Co-reg	$87.10_{(0.00)}$	$87.71_{(0.02)}$	$87.02_{(0.00)}$	$80.72_{(0.00)}$		$67.91_{(0.00)}$	$72.13_{(0.00)}$	$67.93_{(0.00)}$	$53.92_{(0.00)}$
	CoMSC	$36.33_{(0.00)}$	$33.54_{(0.00)}$	$44.33_{(0.00)}$	$36.67_{(0.00)}$		$76.07_{(0.00)}$	$48.65_{(0.00)}$	$76.07_{(0.00)}$	$56.17_{(0.00)}$
	EMGC2F	$24.21_{(0.00)}$	$56.73_{(0.00)}$	$31.03_{(0.00)}$	$31.62_{(0.00)}$		$66.69_{(0.00)}$	$70.81_{(0.00)}$	$67.32_{(0.00)}$	$50.21_{(0.00)}$
	MVCT	$25.03_{(0.00)}$	$14.01_{(0.00)}$	$26.01_{(0.00)}$	$30.43_{(0.00)}$		$54.53_{(0.00)}$	$59.52_{(0.00)}$	$65.41_{(0.01)}$	$35.71_{(0.00)}$
	T-UMC	$65.51_{(0.00)}$	$71.09_{(0.00)}$	$65.51_{(0.00)}$	$70.03_{(0.00)}$		$74.62(0.01)$	$75.89_{(0.00)}$	$74.67_{(0.00)}$	$64.84_{(0.00)}$
	WTSNM	$43.51_{(0.00)}$	$37.89_{(0.00)}$	$50.00_{(0.00)}$	$39.72_{(0.00)}$		$81.52_{(0.00)}$	$85.82_{(0.00)}$	$82.73_{(0.00)}$	$80.88_{(0.00)}$
	HCETR	$70.00_{(0.00)}$	$57.06_{(0.00)}$	$70.00_{(0.00)}$	$59.05_{(0.00)}$		$59.39_{(0.00)}$	$68.01_{(0.00)}$	$63.03_{(0.00)}$	$51.27_{(0.00)}$
	CRMATS	$\mathbf{97.02}_{(0.13)}$	$\mathbf{96.65}_{(0.11)}$	$\mathbf{98.64}_{(0.03)}$	$\mathbf{97.13}_{(0.21)}$		$\mathbf{91.12}_{(0.01)}$	$\mathbf{88.12}_{(0.01)}$	$\mathbf{91.12}_{(0.01)}$	$\mathbf{83.33}_{(0.00)}$
MSRC_v1	SMSC	$78.90_{(0.00)}$	$47.73_{(0.02)}$	$78.90_{(0.00)}$	$66.11_{(0.00)}$	BBCSport	$66.67_{(0.00)}$	$56.51_{(0.15)}$	$66.67_{(0.00)}$	$66.22_{(0.00)}$
	RMSC	$54.41_{(0.00)}$	$55.43_{(0.00)}$	$66.70_{(0.00)}$	$62.93_{(0.00)}$		$88.51_{(0.00)}$	$69.41_{(0.00)}$	$88.54_{(0.01)}$	$79.10_{(0.00)}$
	PGSC	$33.44_{(0.00)}$	$2.31_{(0.00)}$	$67.82_{(0.00)}$	$48.56_{(0.00)}$		$64.64_{(0.00)}$	$2.60_{(0.00)}$	$67.44_{(0.00)}$	$48.36_{(0.00)}$
	LTBPL	$67.81_{(0.00)}$	$64.83_{(0.06)}$	$67.79_{(0.00)}$	$71.61_{(0.07)}$		$74.34_{(0.00)}$	$52.71_{(0.00)}$	$74.44_{(0.00)}$	$60.62_{(0.00)}$
	LSGMC	$95.55_{(0.00)}$	$83.59_{(0.00)}$	$92.22_{(0.00)}$	$91.10_{(0.00)}$		$35.90_{(0.00)}$	$4.74_{(0.00)}$	$36.96_{(0.00)}$	$11.76_{(0.00)}$
	CDMGC	$71.11_{(0.00)}$	$50.88_{(0.00)}$	$71.11_{(0.00)}$	$66.73_{(0.00)}$		$51.28_{(0.00)}$	$23.60_{(0.00)}$	$53.85_{(0.00)}$	$50.88_{(0.00)}$
	Co-reg	$94.43_{(0.00)}$	$81.21_{(0.02)}$	$94.39_{(0.00)}$	$89.21_{(0.01)}$		$73.32_{(0.00)}$	$57.32_{(0.00)}$	$74.44_{(0.00)}$	$66.51_{(0.00)}$
	CoMSC	$44.81_{(0.00)}$	$9.25_{(0.00)}$	$45.19_{(0.00)}$	$10.26_{(0.00)}$		$58.55_{(0.00)}$	$20.76_{(0.00)}$	$59.83_{(0.00)}$	$21.20_{(0.00)}$
	EMGC2F	$\mathbf{99.44}_{(0.05)}$	$\mathbf{95.91}_{(0.06)}$	$\mathbf{99.43}_{(0.03)}$	$\mathbf{96.89}_{(0.07)}$		$91.01_{(0.00)}$	$81.81_{(0.00)}$	$91.32_{(0.00)}$	$84.23_{(0.00)}$
	MVCT	$40.00_{(0.00)}$	$7.77_{(0.00)}$	$56.73_{(0.00)}$	$39.29_{(0.00)}$		$51.31_{(0.00)}$	$29.81_{(0.00)}$	$79.54_{(0.01)}$	$50.81_{(0.00)}$
	T-UMC	$78.11_{(0.00)}$	$67.73_{(0.00)}$	$78.11_{(0.00)}$	$68.16_{(0.00)}$		$94.62_{(0.00)}$	$82.81_{(0.00)}$	$94.62_{(0.01)}$	$90.13_{(0.00)}$
	WTSNM	$79.05_{(0.00)}$	$66.31_{(0.00)}$	$79.05_{(0.00)}$	$64.07_{(0.00)}$		$74.36_{(0.00)}$	$68.99_{(0.00)}$	$74.36_{(0.00)}$	$60.64_{(0.00)}$
	HCETR	$84.29_{(0.00)}$	$74.01_{(0.00)}$	$84.29_{(0.00)}$	$72.61_{(0.00)}$		$82.05_{(0.00)}$	$57.72_{(0.00)}$	$82.05_{(0.00)}$	$70.20_{(0.00)}$
	CRMATS	$98.13_{(0.02)}$	$93.90_{(0.01)}$	$96.28_{(0.01)}$	$95.12_{(0.02)}$		$\mathbf{100.00}_{(0.00)}$	$\mathbf{100.00}_{(0.00)}$	$\mathbf{100.00}_{(0.00)}$	$\mathbf{100.00}_{(0.00)}$
3Sources	SMSC	$26.04_{(0.00)}$	$7.88_{(0.15)}$	$43.21_{(0.00)}$	$21.77_{(0.00)}$	Reuters	$53.50_{(0.00)}$	$43.76_{(0.00)}$	$53.50_{(0.00)}$	$50.83_{(0.00)}$
	RMSC	$47.34_{(0.00)}$	$38.65_{(0.00)}$	$61.54_{(0.00)}$	$40.17_{(0.00)}$		$51.50_{(0.00)}$	$51.06_{(0.00)}$	$51.50_{(0.00)}$	$64.88_{(0.00)}$
	PGSC	$34.91_{(0.00)}$	$3.13_{(0.00)}$	$77.44_{(0.00)}$	$37.46_{(0.00)}$		$51.05_{(0.00)}$	$50.15_{(0.00)}$	$71.05_{(0.00)}$	$66.23_{(0.00)}$
	LTBPL	$63.31_{(0.00)}$	$53.46_{(0.00)}$	$68.86_{(0.00)}$	$55.12_{(0.00)}$		$51.00_{(0.00)}$	$55.17_{(0.00)}$	$51.00_{(0.00)}$	$66.06_{(0.00)}$
	LSGMC	$75.10_{(0.00)}$	$65.17_{(0.00)}$	$76.16_{(0.00)}$	$69.96_{(0.00)}$		$66.73_{(0.01)}$	$61.09_{(0.00)}$	$66.73_{(0.01)}$	$74.12_{(0.00)}$
	CDMGC	$69.11_{(0.00)}$	$56.17_{(0.00)}$	$68.26_{(0.00)}$	$60.95_{(0.00)}$		$64.63_{(0.00)}$	$59.93_{(0.00)}$	$64.63_{(0.00)}$	$77.12_{(0.00)}$
	Co-reg	$43.20_{(0.00)}$	$43.52_{(0.00)}$	$64.54_{(0.00)}$	$38.64_{(0.00)}$		$52.00_{(0.00)}$	$58.49_{(0.00)}$	$52.00_{(0.00)}$	$56.48_{(0.00)}$
	CoMSC	$76.10_{(0.00)}$	$60.17_{(0.00)}$	$79.83_{(0.00)}$	$62.20_{(0.00)}$		$71.11_{(0.00)}$	$65.66_{(0.00)}$	$71.11_{(0.00)}$	$77.13_{(0.00)}$
	EMGC2F	$63.31_{(0.00)}$	$63.12_{(0.00)}$	$76.92_{(0.00)}$	$59.35_{(0.00)}$		$70.31_{(0.01)}$	$65.99_{(0.00)}$	$70.31_{(0.01)}$	$73.42_{(0.00)}$
	MVCT	$60.36_{(0.00)}$	$54.28_{(0.00)}$	$83.43_{(0.00)}$	$56.63_{(0.00)}$		$51.00_{(0.00)}$	$61.73_{(0.00)}$	$51.15_{(0.00)}$	$69.93_{(0.00)}$
	T-UMC	$76.91_{(0.00)}$	$71.28_{(0.00)}$	$83.61_{(0.00)}$	$72.35_{(0.00)}$		$75.67_{(0.00)}$	$64.35_{(0.00)}$	$75.67_{(0.00)}$	$66.73_{(0.00)}$
	WTSNM	$68.46_{(0.00)}$	$57.12_{(0.00)}$	$63.85_{(0.00)}$	$70.75_{(0.00)}$		$60.61_{(0.00)}$	$61.36_{(0.00)}$	$62.17_{(0.00)}$	$75.35_{(0.00)}$
	HCETR	$66.86_{(0.00)}$	$48.50_{(0.00)}$	$66.86_{(0.00)}$	$50.50_{(0.00)}$		$85.00_{(0.00)}$	$64.22_{(0.00)}$	$85.00_{(0.00)}$	$74.75_{(0.00)}$
	CRMATS	$\mathbf{95.19}_{(0.00)}$	$\mathbf{87.46}_{(0.00)}$	$\mathbf{93.95}_{(0.00)}$	$\mathbf{89.64}_{(0.00)}$		$\mathbf{87.68}_{(0.00)}$	$\mathbf{77.38}_{(0.00)}$	$\mathbf{87.68}_{(0.00)}$	$\mathbf{79.95}_{(0.00)}$

The best scores are highlighted in bold.

To visually demonstrate the effectiveness of our method on Coil-20, we present the spatial distributions of the raw data and consensus representation obtained using t-SNE in Fig. 3. Additionally, Fig. 4 showcases the corresponding heatmaps and consensus representations learned by our method. These visualizations provide compelling evidence supporting the superiority of our approach. In the pairwise affinity-based visualization (Fig. 3(a)–(c)), most samples appear mixed together, making it challenging to accurately distinguish between different subgroups . However, our method generates a consensus representation that exhibits improved separation and reduced overlap, leading to enhanced clustering performance. The effectiveness

TABLE III
CHI RESULTS: THE MEAN AND p VALUE MEASURED BY DIFFERENT CLUSTERING METHODS ON BENCHMARK DATASETS

Method	Dataset	CHI	Dataset	CHI	Dataset	CHI	Dataset	CHI
SMSC	Syndata1	$1.38_{(0.00)}$	Syndata2	$1.61_{(0.00)}$	Coil-20	$335.85_{(0.00)}$	Yale	$36.86_{(0.00)}$
RMSC		$1.38_{(0.00)}$		$1.13_{(0.00)}$		$134.91_{(0.00)}$		$48.81_{(0.00)}$
PGSC		$0.86_{(0.00)}$		$1.12_{(0.00)}$		$108.65_{(0.00)}$		$13.87_{(0.00)}$
LTBPL		$1.34_{(0.00)}$		$0.41_{(0.00)}$		$224.46_{(0.00)}$		$44.74_{(0.00)}$
LSGMC		$1.35_{(0.00)}$		$1.41_{(0.00)}$		$292.13_{(0.00)}$		$44.31_{(0.00)}$
CGMGC		$1.22_{(0.00)}$		$1.13_{(0.00)}$		$274.65_{(0.00)}$		$45.29_{(0.00)}$
Co-reg		$1.21_{(0.00)}$		$1.23_{(0.00)}$		$305.67_{(0.00)}$		$45.95_{(0.00)}$
CoMSC		$1.67_{(0.00)}$		$1.47_{(0.00)}$		$83.89_{(0.00)}$		$50.31_{(0.00)}$
EMGC2F		$1.41_{(0.00)}$		$1.18_{(0.00)}$		$30.84_{(0.00)}$		$45.30_{(0.00)}$
MVCT		$1.17_{(0.00)}$		$1.13_{(0.00)}$		$34.52_{(0.00)}$		$38.80_{(0.00)}$
T-UMC		$1.68_{(0.00)}$		$1.24_{(0.00)}$		$211.35_{(0.00)}$		$49.54_{(0.00)}$
WTSNM		$1.38_{(0.00)}$		$1.26_{(0.00)}$		$128.18_{(0.00)}$		$53.23_{(0.00)}$
HCETR		$1.74_{(0.00)}$		$1.53_{(0.00)}$		$229.77_{(0.00)}$		$41.40_{(0.00)}$
CRMATS		$\mathbf{2.13_{(0.00)}}$		$\mathbf{1.86_{(0.00)}}$		$\mathbf{349.00_{(0.00)}}$		$\mathbf{58.37_{(0.00)}}$
Truth		2.13		1.90		362.02		63.12
SMSC	MSRC_v1	$12.29_{(0.00)}$	BBCSport	$56.47_{(0.00)}$	3Scources	$8.82_{(0.00)}$	Reuters	$4.27_{(0.00)}$
RMSC		$11.57_{(0.00)}$		$104.62_{(0.00)}$		$9.04_{(0.00)}$		$4.28_{(0.00)}$
PGSC		$10.94_{(0.00)}$		$51.99_{(0.00)}$		$8.91_{(0.00)}$		$4.27_{(0.00)}$
LTBPL		$11.96_{(0.00)}$		$73.38_{(0.00)}$		$9.20_{(0.00)}$		$4.27_{(0.00)}$
LSGMC		$12.79_{(0.00)}$		$11.39_{(0.00)}$		$9.32_{(0.00)}$		$4.32_{(0.00)}$
CGMGC		$12.06_{(0.00)}$		$22.53_{(0.00)}$		$9.26_{(0.00)}$		$4.31_{(0.00)}$
Co-reg		$12.76_{(0.00)}$		$71.13_{(0.00)}$		$8.99_{(0.00)}$		$4.27_{(0.00)}$
CoMSC		$11.28_{(0.00)}$		$38.56_{(0.00)}$		$9.33_{(0.00)}$		$4.34_{(0.00)}$
EMGC2F		$\mathbf{12.90_{(0.00)}}$		$110.14_{(0.00)}$		$9.20_{(0.00)}$		$4.33_{(0.00)}$
MVCT		$11.13_{(0.00)}$		$22.59_{(0.00)}$		$9.17_{(0.00)}$		$4.27_{(0.00)}$
T-UMC		$12.27_{(0.00)}$		$118.10_{(0.00)}$		$9.33_{(0.00)}$		$4.35_{(0.00)}$
WTSNM		$12.30_{(0.00)}$		$73.42_{(0.00)}$		$9.25_{(0.00)}$		$4.30_{(0.00)}$
HCETR		$12.45_{(0.00)}$		$90.38_{(0.00)}$		$9.25_{(0.00)}$		$4.38_{(0.00)}$
CRMATS		$12.87_{(0.00)}$		$\mathbf{129.97_{(0.00)}}$		$\mathbf{9.52_{(0.00)}}$		$\mathbf{4.40_{(0.00)}}$
Truth		12.93		129.97		9.58		4.64

The best scores are highlighted in bold.

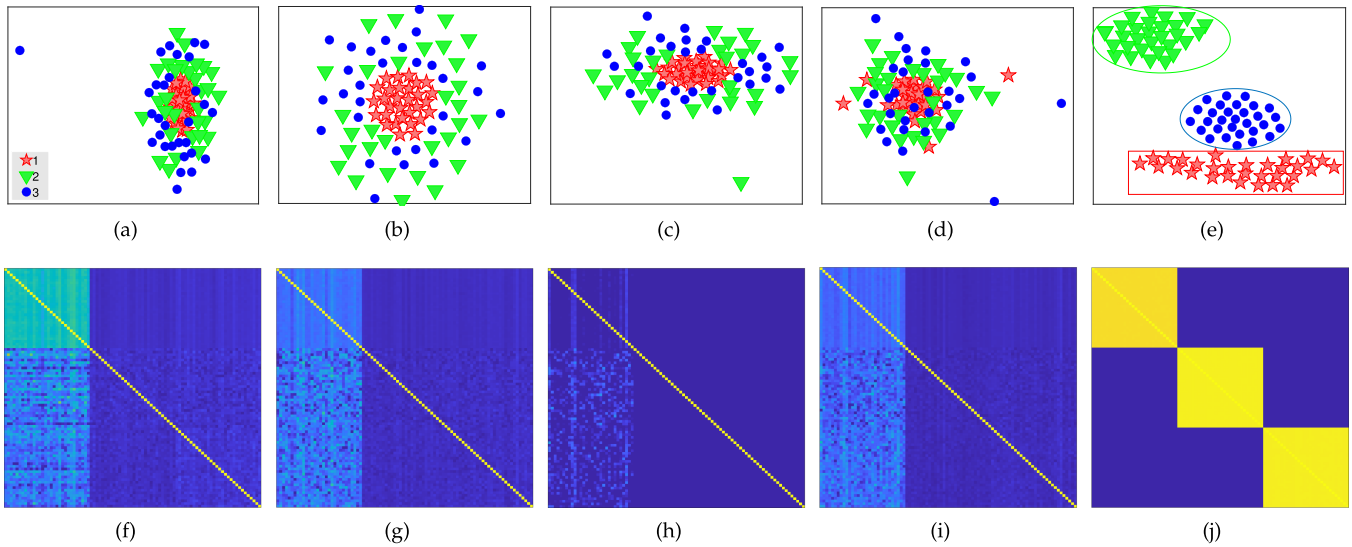


Fig. 2. Visualizations of the consensus representation, raw data with t-SNE, and the heatmap on Syndata2. Subfigure (a)–(d) depict the spatial distribution of Syndata2 from different views using t-SNE. Subfigure (e) is the visualization of the obtained consensus representations $\boldsymbol{W}$ using t-SNE. Subfigure (f)–(i) show the heatmap of Syndata2 from different views. The affinity heatmap on raw samples has blurred boundaries and no apparent block structures. In contrast, Subfigure (j) shows the affinity from consensus representations has clear-cut boundaries.

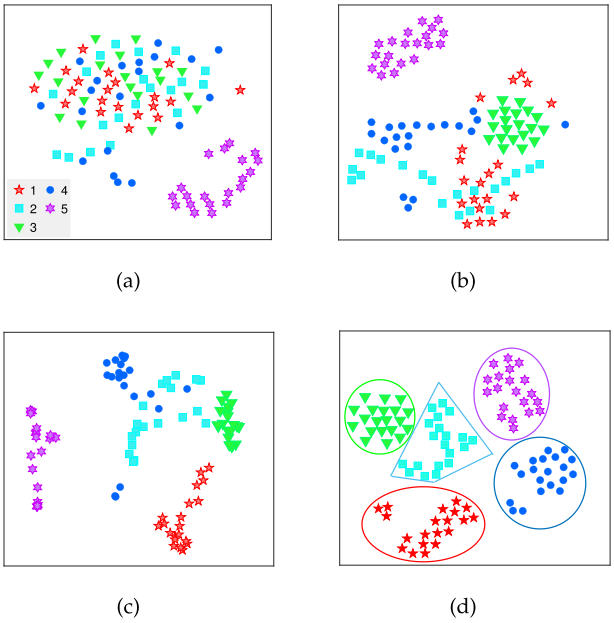


Fig. 3. Visualization of the consensus representations and raw data. Subfigure (a)–(c) depict the visualization of Coil-20 using t-SNE. Subfigure (d) is visualization of the consensus representations $\boldsymbol{W}$.

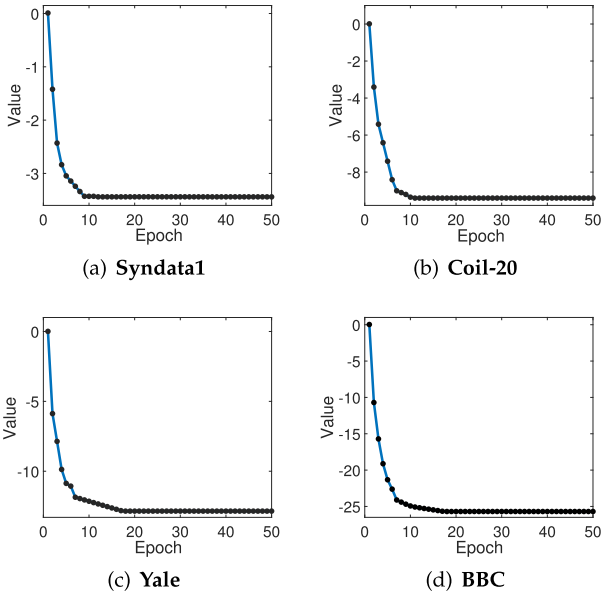


Fig. 5. The convergence curve of the proposed CRMATS method, on (a) Syndata1; (b) Coil-20; (c) Yale; and (d) BBC. The objective value decreases consistently with respect to the iteration number.

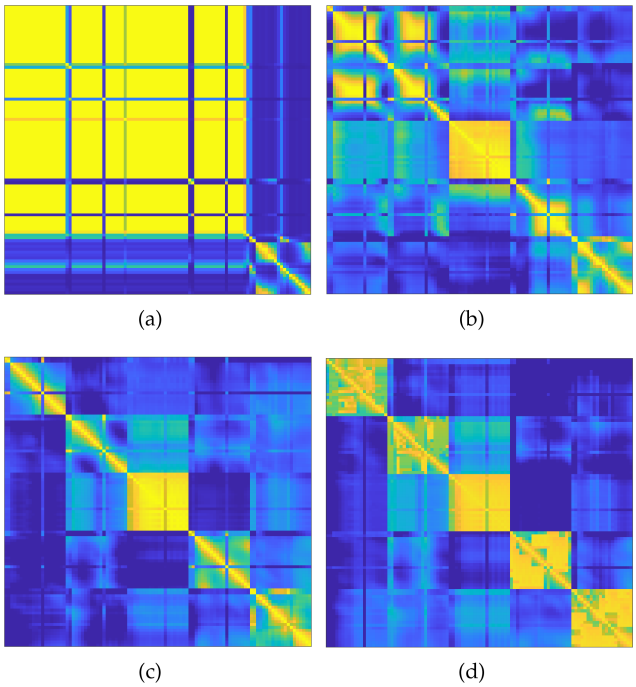


Fig. 4. Visualization of heatmap on Coil-20. Subfigure (a)–(c) are the heatmap from different views. The affinity heatmap on raw samples has blurred boundaries. In contrast, Subfigure (d) shows the affinity from consensus representation has clear-cut boundaries.

of our method is demonstrated through the comparison of affinity heatmaps. In Coil-20, the affinity heatmaps of each view on the raw samples (Fig. 4(a)–(c)) exhibit blurred boundaries and lack a clear block structure. However, after applying our method and generating the affinity matrix from the consensus representation, the boundaries become well-defined. This is supported by the t-SNE visualization of the consensus representation, where most samples are clustered with their corresponding partners, and distinct blocks are observed on the diagonal of the heatmap (Figs. 3(d) and 4(d)). Overall, CRMATS yields superior clustering performance on all real-world datasets, as shown in Tables II and III. Furthermore, we conducted an analysis on the randomness of views and resource consumption in CRMATS, which can be found in Section B and Section C of the supplementary materials, available online.

Based on the experimental results, CRMATS demonstrates several advantages in handling HDLSS data. First, our method effectively utilizes high and low-order affinities to comprehensively capture the spatial structure of HDLSS data. Second, our co-regularization approach aligns the different low-dimensional representations to seek a consensus graph and incorporate cross-view correlations, thereby avoiding suboptimal clustering results. Lastly, by learning a consensus representation on the manifold space, we consider the complex connections among samples of the original data, leading to improved clustering performance.

C. Convergence Analysis

An alternate minimizing algorithm is developed to solve the optimization problem. As for the convergence of CRMATS, we have provided the corresponding theoretical proof in Section A of the supplementary materials, available online. In this subsection, we compare objective value of the benchmark datasets with diverse backgrounds to eliminate randomness and then illustrate clearly in Fig. 5. We show the objective value with 50 epochs. The corresponding objective value of each benchmark dataset decreases sharply in the first 5 iterations and then stays steady with more iterations, implying that CRMATS

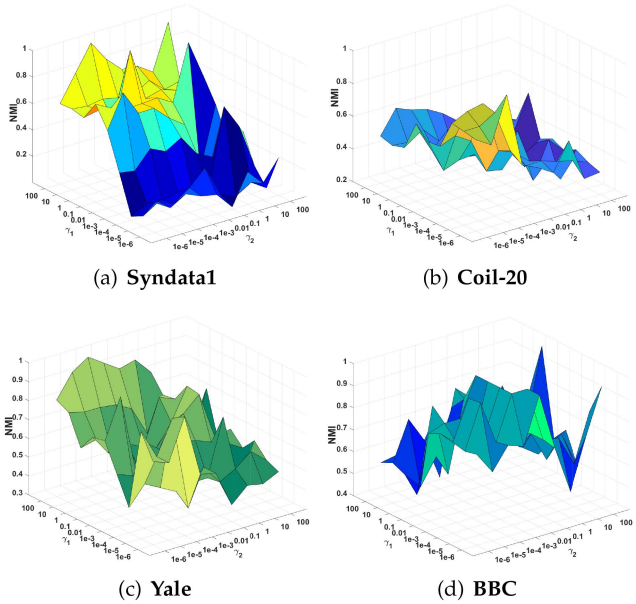


Fig. 6. Hyperparameter tuning (γ_1 and γ_2) in terms of NMI, on (a) Syndata1; (b) Coil-20; (c) Yale; and (d) BBC.

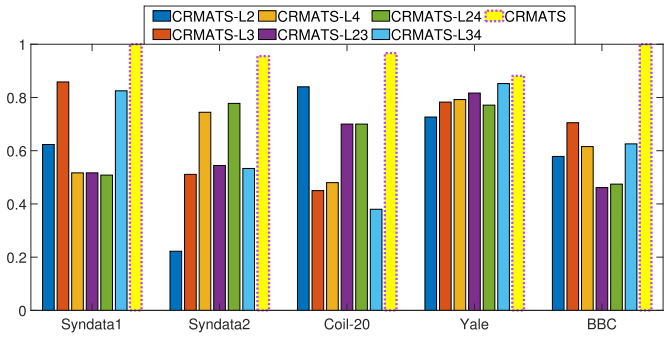


Fig. 7. Ablation analysis experiments of the CRMATS: 1) CRMATS-L2 which only uses the pairwise affinity matrix $\hat{\boldsymbol{L}}_2$; 2) CRMATS-L3, which only uses the third-order affinity matrix $\hat{\boldsymbol{L}}_3$; 3) CRMATS-L4, which only uses the fourth-order affinity matrix $\hat{\boldsymbol{L}}_4$; 4) CRMATS-L23, which combines $\hat{\boldsymbol{L}}_2$ and $\hat{\boldsymbol{L}}_3$; 5) CRMATS-L24 combines $\hat{\boldsymbol{L}}_2$ and $\hat{\boldsymbol{L}}_4$; 6) CRMATS-L34 combines $\hat{\boldsymbol{L}}_3$ and $\hat{\boldsymbol{L}}_4$; 7) CRMATS is our model.

converges steadily after just a few iterations. The fast convergence of CRMATS is due to the alternate minimizing algorithm, which updates each variable separately. Moreover, by integrating co-regularization techniques and leveraging high-order affinities on the manifold space, our proposed method adeptly captures the intrinsic structure of the data. Consequently, our method exhibits rapid convergence, highlighting its effectiveness as a solution for optimizing the problem in multi-view HDLSS clustering.

D. Hyperparameter Sensitivity Analysis

The influence of hyperparameter sets on the clustering performance in terms of NMI is investigated, and the results are represented in Fig. 6. Notice that (16) consists of two parts, the affinities term and co-regularization term. We set hyperparameters for each part, denoted as γ_1 and γ_2, respectively. We use a combination of hyperparameters [1e-6, 1e-5, 1e-4, 1e-3, 0.01, 0.1, 1, 10, 100] for both γ_1 and γ_2, and evaluate the performance on four benchmark datasets.

According to Fig. 6, the clustering performance is influenced by γ_1 in certain instances. For example, the NMI of Syndata1 and Yale (Fig. 6(a) and (c)) decreases with lower γ_1. In addition, the NMI changes little while γ_2 increases. Although the performance of CRMATS changes with different combinations of γ_1 and γ_2, as shown in Fig. 6(b) and (d), CRMATS still outperforms its comparative methods on benchmark datasets, demonstrating the stability of our model. Furthermore, the results depicted in Fig. 6 provide empirical evidence of the robustness of CRMATS to the selection of hyperparameters. Our method consistently outperforms the comparative methods on multiple datasets under different settings, indicating its insensitivity to specific hyperparameter choices and its ability to achieve favorable performance across a wide range of hyperparameter values. The observed effectiveness and robustness of CRMATS in multi-view clustering tasks, as demonstrated in the results of Fig. 6, underscore the practicality and versatility of our method.

E. Ablation Analysis

In this subsection, an ablation study of our proposed CRMATS method is conducted to investigate the roles played by different order of affinities and their combinations. In order to assess the significance of multi-order affinities, an ablation study is conducted as follows. First, the removal of $\hat{\boldsymbol{L}}_3$ and $\hat{\boldsymbol{L}}_4$ results in the method degrading into the traditional Co-reg approach. This variant is referred to as CRMATS-L2. Subsequently, each high-order affinity is individually applied in isolation, yielding the methods CRMATS-L3 and CRMATS-L4, respectively. In the multiple affinities situation, we combine $\hat{\boldsymbol{L}}_2$, $\hat{\boldsymbol{L}}_3$, and $\hat{\boldsymbol{L}}_4$ in pairs. For instance, CRMATS-L23 is the combination of $\hat{\boldsymbol{L}}_2$ and $\hat{\boldsymbol{L}}_3$. The clustering performance of each ablation method is evaluated using the same benchmark datasets and metrics as employed in the previous experiments.

The experimental results presented in Fig. 7 demonstrate the importance and effectiveness of incorporating multi-order affinities in multi-view clustering tasks. First, our method achieves the best performance on the following datasets, indicating its effectiveness in capturing the underlying structure of the data in multi-view clustering tasks. Second, singular affinity experiments demonstrate that high-order affinities (CRMATS-L3, CRMATS-L4) supplement the inherent information of traditional pairwise affinity. For example, the performance of CRMATS-L3 is better than CRMATS-L2 on Syndata1 and BBC. Finally, the fusion of the second, third, and fourth order affinities exhibits superior performance in comparison to any two-order fusion (e.g., CRMATS-L23, CRMATS-L24, CRMATS-L34). This observation suggests that the inclusion of each order of fusion contributes to the enhancement of internal information, resulting in incremental improvements.

V. Conclusion

In this paper, we have presented a unified multi-view clustering framework, which aligns the latent representations of views on the manifold space based on the accurate description of intra-view sample relationships through the introduction of multi-order affinities. The efficiency of CRMATS is improved by employing an alternating minimization strategy and singular value decomposition. Furthermore, a novel set of evaluation metrics is devised to comprehensively assess the performance of CRMATS in capturing the underlying structure of the data, followed by taking into account the similarity within clusters and the dissimilarity between clusters in clustering tasks. Experimental results on eight HDLSS datasets have demonstrated the effectiveness of the proposed method in comparison with the other popular approaches.

Although our method effectively addresses the concentration effects in high-dimensional data clustering and outperforms several baseline methods on the benchmark dataset, there are still potential directions for improvement. First, our strategy of improving the HDLSS clustering results through high-order affinity requires more time and memory costs for computation. Additionally, existing high-order affinity methods have limitations when dealing with the complex graph data. To address these limitations, we consider incorporating deep graph neural network for high-dimensional learning. The deep neural network can leverage GPU computing units to reduce memory costs within each mini-batch. Furthermore, the graph neural network enables us to extract more potential high-order correlations from high-dimensional data.

References

[1] T. Hofmann and J. M. Buhmann, "Pairwise data clustering by deterministic annealing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 1, pp. 1–14, Jan. 1997.

[2] L. Law and Y. Cheung, "Color image segmentation using rival penalized controlled competitive learning," in *Proc. Int. Joint Conf. Neural Netw.*, 2003, pp. 108–112.

[3] Y. Cheung and L. Law, "Rival-model penalized self-organizing map," *IEEE Trans. Neural Netw.*, vol. 18, no. 1, pp. 289–295, Jan. 2007.

[4] Y. Cheung, "Maximum weighted likelihood via rival penalized EM for density mixture clustering with automatic model selection," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 750–761, Jun. 2005.

[5] D. Huang, C. Wang, J. Wu, J. Lai, and C. Kwoh, "Ultra-scalable spectral clustering and ensemble clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 6, pp. 1212–1226, Jun. 2020.

[6] T. Debnath and M. Song, "Fast optimal circular clustering and applications on round genomes," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 6, pp. 2061–2071, Nov./Dec. 2021.

[7] L. Fu, P. Lin, A. V. Vasilakos, and S. Wang, "An overview of recent multi-view clustering," *Neurocomputing*, vol. 402, pp. 148–161, 2020.

[8] H. Peng, H. Wang, Y. Hu, W. Zhou, and H. Cai, "Multi-dimensional clustering through fusion of high-order similarities," *Pattern Recognit.*, vol. 121, 2022, Art. no. 108108.

[9] M. Liu, Y. Wang, V. Palade, and Z. Ji, "Multi-view subspace clustering network with block diagonal and diverse representation," *Inf. Sci.*, vol. 626, pp. 149–165, 2023.

[10] M. Brbić and I. Kopriva, "Multi-view low-rank sparse subspace clustering," *Pattern Recognit.*, vol. 73, pp. 247–258, 2018.

[11] Y. Yang and H. Wang, "Multi-view clustering: A survey," *Big Data Mining Analytics*, vol. 1, no. 2, pp. 83–107, 2018.

[12] R. Xia, Y. Pan, L. Du, and J. Yin, "Robust multi-view spectral clustering via low-rank and sparse decomposition," in *Proc. 28th AAAI Conf. Artif. Intell.*, 2014, pp. 2149–2155.

[13] F. Nie, S. Shi, and X. Li, "Auto-weighted multi-view co-clustering via fast matrix factorization," *Pattern Recognit.*, vol. 102, 2020, Art. no. 107207.

[14] A. Kumar and H. Daume III, "A co-training approach for multi-view spectral clustering," in *Proc. 28th Int. Conf. Mach. Learn.*, 2011, pp. 393–400.

[15] W. Yang, C. Hui, D. Sun, X. Sun, and Q. Liao, "Clustering through probability distribution analysis along eigenpaths," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 2, pp. 875–884, Feb. 2021.

[16] Y. Chen et al., "KNN-BLOCK DBSCAN: Fast clustering for large-scale data," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 6, pp. 3939–3953, Jun. 2021.

[17] D. François, V. Wertz, and M. Verleysen, "The concentration of fractional distances," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 7, pp. 873–886, Jul. 2007.

[18] S. Sarkar and A. K. Ghosh, "On perfect clustering of high dimension, low sample size data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 9, pp. 2257–2272, Sep. 2020.

[19] Y. Mei, Z. Ren, B. Wu, T. Yang, and Y. Shao, "Multi-order similarity learning for multi-view spectral clustering," *Pattern Recognit.*, vol. 137, 2023, Art. no. 109264.

[20] D. Ghoshdastidar and A. Dukkipati, "Spectral clustering using multilinear SVD: Analysis, approximations and applications," in *Proc. 29th AAAI Conf. Artif. Intell.*, 2015, pp. 2610–2616.

[21] H. Peng, Y. Hu, J. Chen, H. Wang, Y. Li, and H. Cai, "Integrating tensor similarity to enhance clustering performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 5, pp. 2582–2593, May 2022.

[22] H. Cai et al., "Uniform tensor clustering by jointly exploring sample affinities of various orders," 2023, *arXiv:2302.01569*.

[23] J. Wu, Z. Lin, and H. Zha, "Essential tensor learning for multi-view spectral clustering," *IEEE Trans. Image Process.*, vol. 28, no. 12, pp. 5910–5922, Dec. 2019.

[24] J. Yu, G. Zhou, C. Li, Q. Zhao, and S. Xie, "Low tensor-ring rank completion by parallel matrix factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 7, pp. 3020–3033, Jul. 2021.

[25] J. Guo, Y. Sun, J. Gao, Y. Hu, and B. Yin, "Logarithmic Schatten-p norm minimization for tensorial multi-view subspace clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 3, pp. 3396–3410, Mar. 2023.

[26] J. Ji and S. Feng, "High-order complementarity induced fast multi-view clustering with enhanced tensor rank minimization," in *Proc. 31st ACM Int. Conf. Multimedia*, 2023, pp. 328–336.

[27] Z. Li, C. Tang, X. Zheng, X. Liu, W. Zhang, and E. Zhu, "High-order correlation preserved incomplete multi-view subspace clustering," *IEEE Trans. Image Process.*, vol. 31, pp. 2067–2080, 2022.

[28] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proc. 14th Int. Conf. Neural Inf. Process. Syst.*, 2001, pp. 849–856.

[29] A. Kumar, P. Rai, and H. Daumé, "Co-regularized multi-view spectral clustering," in *Proc. 24th Int. Conf. Neural Inf. Process. Syst.*, 2011, pp. 1413–1421.

[30] J. Yu, Q. Duan, H. Huang, S. He, and T. Zou, "Effective incomplete multi-view clustering via low-rank graph tensor completion," *Mathematics*, vol. 11, no. 3, pp. 1–18, 2023.

[31] Y. Chikuse, *Statistics on Special Manifolds*. Berlin, Germany: Springer, 2003.

[32] H. Wang, G. Han, B. Zhang, G. Tao, and H. Cai, "Multi-view learning a decomposable affinity matrix via tensorself-representation on Grassmann manifold," *IEEE Trans. Image Process.*, vol. 30, pp. 8396–8409, 2021.

[33] Y. Hu, E. Guo, Z. Xie, X. Liu, and H. Cai, "Robust multi-view clustering through partition integration on Stiefel manifold," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 10, pp. 10397–10410, Oct. 2023.

[34] B. Wang, Y. Hu, J. Gao, Y. Sun, F. Ju, and B. Yin, "Adaptive fusion of heterogeneous manifolds for subspace clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 8, pp. 3484–3497, Aug. 2021.

[35] Y. Hu and H. Cai, "Multi-view clustering through hypergraphs integration on Stiefel manifold," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2022, pp. 01–06.

[36] A. Khan and P. Maji, "Multi-manifold optimization for multi-view subspace clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 8, pp. 3895–3907, Aug. 2022.

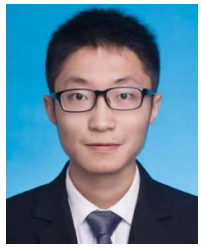
[37] F. Nie, J. Yuan, and H. Huang, "Optimal mean robust principal component analysis," in *Proc. 31st Int. Conf. Mach. Learn.*, 2014, pp. 1062–1070.

[38] A. Chaudhry, N. Khan, P. K. Dokania, and P. H. S. Torr, "Continual learning in low-rank orthogonal subspaces," in *Proc. 34th Int. Conf. Neural Inf. Process. Syst.*, 2020, pp. 2654–2666.

[39] Y. Ye, X. Liu, J. Yin, and E. Zhu, "Co-regularized kernel k-means for multi-view clustering," in *Proc. 23rd Int. Conf. Pattern Recognit.*, 2016, pp. 1583–1588.
[40] H. Wang, J. Chen, B. Zhang, and H. Cai, "Accurate multi-view clustering by exploiting within-view high-order affinities through tensorself-representation," in *Proc. IEEE Int. Conf. Bioinf. Biomed.*, 2022, pp. 595–600.
[41] B. Cai, G. Lu, L. Yao, and H. Li, "High-order manifold regularized multi-view subspace clustering with robust affinity matrices and weighted TNN," *Pattern Recognit.*, vol. 134, 2023, Art. no. 109067.
[42] M. Chen, C. Wang, and J. Lai, "Low-rank tensor based proximity learning for multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 5, pp. 5076–5090, May 2023.
[43] Y. Li, D. Tao, W. Liu, and Y. Wang, "Co-regularization for classification," in *Proc. IEEE Int. Conf. Secur. Pattern Anal. Cybern.*, 2014, pp. 218–222.
[44] W. Lan et al., "Multiview subspace clustering via low-rank symmetric affinity graph," *IEEE Trans. Neural Netw. Learn. Syst.*, pp. 1–14, 2023.
[45] M. Sun et al., "Scalable multi-view subspace clustering with unified anchors," in *Proc. 29th ACM Int. Conf. Multimedia*, 2021, pp. 3528–3536.
[46] H. Wu, S. Huang, C. Tang, Y. Zhang, and J. Lv, "Pure graph-guided multi-view subspace clustering," *Pattern Recognit.*, vol. 136, 2023, Art. no. 109187.
[47] S. Huang, I. W. Tsang, Z. Xu, and J. Lv, "Measuring diversity in graph learning: A unified framework for structured multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 12, pp. 5869–5883, Dec. 2022.
[48] J. Liu, X. Liu, Y. Yang, X. Guo, M. Kloft, and L. He, "Multiview subspace clustering via co-training robust data representation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 10, pp. 5177–5189, Oct. 2022.
[49] D. Wu, J. Lu, F. Nie, R. Wang, and Y. Yuan, "EMGC^2;F: Efficient multi-view graph clustering with comprehensive fusion," in *Proc. 31st Int. Joint Conf. Artif. Intell.*, 2022, pp. 3566–3572.
[50] S. Huang, I. Tsang, Z. Xu, J. Lv, and Q.-H. Liu, "Multi-view clustering on topological manifold," in *Proc. 36th AAAI Conf. Artif. Intell.*, 2022, pp. 6944–6951.
[51] J.-Q. Lin, M.-S. Chen, C.-D. Wang, and H. Zhang, "A tensor approach for uncoupled multiview clustering," *IEEE Trans. Cybern.*, vol. 54, no. 2, pp. 1236–1249, Feb. 2024.
[52] W. Xia, X. Zhang, Q. Gao, X. Shu, J. Han, and X. Gao, "Multiview subspace clustering by an enhanced tensor nuclear norm," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 8962–8975, Sep. 2022.
[53] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 12, pp. 1650–1654, Dec. 2002.

Hongmin Cai (Senior Member, IEEE) received the BS and MS degrees in mathematics from the Harbin Institute of Technology, Harbin, China, in 2001 and 2003, respectively, and the PhD degree in applied mathematics from Hong Kong University, in 2007. He is a professor with the School of Computer Science and Engineering, South China University of Technology, Guangzhou, China. From 2005 to 2006, he was a research assistant with the Center of Bioinformatics, Harvard University, and the Section for Biomedical Image Analysis, University of Pennsylvania. His current research interests include bioinformatics, machine learning, and medical image analysis.

Yu Wang received the BS and MS degrees from the Hunan University of Technology and Business, Changsha, Hunan, China, in 2019 and 2022, respectively. He is currently working toward the PhD degree with the School of Future Technology, South China University of Technology, Guangzhou, Guangdong, China. His research interests include machine learning and image processing.



Fei Qi received the BS and MS degrees from Xiamen University, Xiamen, Fujian, China, in 2013 and 2016, respectively. He is currently working toward the PhD degree in computer science and engineering with the South China University of Technology, Guangzhou, Guangdong, China. His research interests include machine learning and image processing.

Zhuoyao Wang received the PhD degree in electrical and computer engineering with a minor in mathematics from the University of New Mexico, Albuquerque, NM, USA, in 2016. He is currently a research scientist with Peng Cheng Laboratory, Shenzhen, China. He also spent several years working in industry at large Blue Chip organizations (Tencent, ZTE) and hi-tech startups. His research interests include risk analysis of cyber-physical systems, cloud computing, and industrial internet. He is currently a co-chair of the CCSA TC13 WG4.

Yiu-ming Cheung (Fellow, IEEE) received the PhD degree from the Department of Computer Science and Engineering, Chinese University of Hong Kong, Hong Kong. He is currently a chair professor with the Department of Computer Science, Hong Kong Baptist University, Hong Kong. His current research interests include machine learning and visual computing, as well as their applications in data science, pattern recognition, multi-objective optimization, and information security. He is the editor-in-chief of *IEEE Transactions on Emerging Topics in Computational Intelligence*. Also, he serves as an associate editor of *IEEE Transactions on Cybernetics*, *IEEE Transactions on Cognitive and Developmental Systems*, *Pattern Recognition*, *Knowledge and Information Systems*, and *Neurocomputing*, just to name a few. He is a fellow of the AAAS, IET, BCS, and AAIA.