Modularity Regularized Locality Preserving Projection

Ibuki Masuda[1(✉)], Yasunori Futamura[1,2], Jianbo Lin[3], Anh Khoa Augustin Lu[4,5], Ryo Tamura[3,6], Tsuyoshi Miyazaki[2,5,7], and Tetsuya Sakurai[1,2]

[1] Department of Computer Science, University of Tsukuba, Tsukuba 305-8573, Japan
masuda.ibuki.tkb_eb@u.tsukuba.ac.jp
[2] Master's/Doctoral Program in Life Science Innovation, University of Tsukuba, Tsukuba 305-8577, Japan
[3] Center for Basic Research on Materials, National Institute for Materials Science, Tsukuba 305-0047, Japan
[4] Department of Materials Engineering, The University of Tokyo, Tokyo 113-8656, Japan
[5] Research Center for Materials Nanoarchitectonics (MANA), National Institute for Materials Science, Tsukuba 305-0044, Japan
[6] Graduate School of Frontier Sciences, The University of Tokyo, Chiba 277-8561, Japan
[7] Graduate School of Engineering, Nagoya university, Nagoya 464-8603, Japan

Abstract. Dimensionality reduction is a crucial preprocessing step for clustering. While nonlinear dimensionality reduction methods effectively capture complex data structures, linear dimensionality reduction methods are often preferred in applications requiring interpretability. However, existing linear approaches, while preserving local or global data characteristics, often struggle to effectively preserve cluster structures. In this paper, we propose a novel linear dimensionality reduction method based on modularity maximization, which is widely used in community detection. The proposed method is formulated as an extension of Locality Preserving Projection (LPP), controlled by a single parameter. More precisely, our method preserves local structural information based on the principle of LPP, while incorporating a distinctive regularization term designed to promote inter-cluster separation. We further show that this parameter serves to control the inter-cluster separation. Experimental results demonstrate that our method can robustly obtain a low-dimensional space that preserves cluster structures, particularly for datasets with high-variance clusters, compared to conventional methods.

Keywords: Linear Dimensionality Reduction · Modularity · Clustering.

1 Introduction

Dimensionality reduction is a technique that compresses high-dimensional data into lower dimensions while preserving essential information, playing a critical role in data visualization and enhancing computational efficiency. In recent

years, nonlinear dimensionality reduction methods have demonstrated a superior ability to capture complex data structures and provide effective visualization results. Prominent examples of nonlinear methods classified as graph embeddings include Laplacian Eigenmaps[2] and Unified Manifold Approximation and Projection[14]. These methods are designed based on the manifold hypothesis[3], which assumes that high-dimensional data lie on a low-dimensional manifold. Graph embeddings construct neighborhood graphs from data, assuming these graphs capture the geometric structure of the manifold. While nonlinear dimensionality reduction methods excel at capturing intrinsic data structures, they suffer from limited interpretability. In contrast, linear dimensionality reduction methods such as Principal Component Analysis (PCA) and Locality Preserving Projection (LPP)[11] offer greater interpretability due to their linear transformation constraints but are inherently limited in their capacity to represent complex nonlinear structures.

One important application of dimensionality reduction is clustering. Clustering is an unsupervised machine learning technique that groups data based on its underlying structure. Dimensionality reduction for clustering aims to eliminate redundancy in high-dimensional data while emphasizing cluster structures. In applications where interpretability is crucial, linear dimensionality reduction methods are preferred, but few existing linear methods are specifically designed with clustering in mind. Particularly in datasets with high-variance clusters, conventional methods like PCA and LPP tend to prioritize information from these high-variance clusters, potentially resulting in suboptimal low-dimensional representations from a clustering perspective.

In this research, we propose a linear dimensionality reduction method called Modularity Maximizing Projection (MMP) based on modularity[16], a widely used evaluation metric in community detection. Modularity is a measure that assigns high values when identifying clusters as sets of vertices with dense connections in a graph. Graph clustering methods that maximize modularity[4] are known to be more effective at capturing community structures in graphs compared to methods based on graph Laplacians, such as Normalized Cut[18].

The proposed MMP is based on eigenvalue calculation of the modularity matrix and has deep relationships with PCA and LPP. MMP's objective function is formulated as the sum of an attractive term and a repulsive term. We introduce a parameter to control the strength of repulsion and show that MMP becomes equivalent to LPP as the repulsion strength approaches zero. We propose this parameterized linear dimensionality reduction framework as Modularity Regularized Locality Preserving Projection (MR-LPP). Within this framework, selecting an appropriate parameter enables the acquisition of low-dimensional spaces that robustly preserve cluster structures. Additionally, since MR-LPP can be formulated as an extension of LPP, it inherits LPP's theoretical background.

This paper is organized as follows. In Section 2, we explain related work, including LPP and methods related to modularity maximization. In Section 3, we propose MMP as a linear dimensionality reduction method based on modularity maximization. In Section 4, we propose MR-LPP, a generalized method

that encompasses both MMP and LPP. In Section 5, we demonstrate through experiments that the proposed method provides low-dimensional representations suitable for clustering in datasets containing high-variance clusters. Finally, in Section 6, we present conclusions and future prospects.

2 Related Work

2.1 Laplacian Eigenmaps

Laplacian Eigenmaps (LE)[2] is an unsupervised nonlinear dimensionality reduction method based on eigenvalue decomposition of the graph Laplacian. LE captures the essential nonlinear structure of data by preserving local neighborhood structures during dimensionality reduction. This method pioneered the application of spectral theory from graph cut minimization problems[10] to dimensionality reduction.

LE transforms a high-dimensional data matrix $X \in \mathbb{R}^{n\times m}$ into a low-dimensional data matrix $Y \in \mathbb{R}^{n\times d}$. Here, if we denote the transpose vector of the i-th row vector of X as $\mathbf{x}_i \in \mathbb{R}^m$, then X can be represented as $X = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]^\top$.

First, a similarity matrix $W \in \mathbb{R}^{n\times n}$ is constructed from the high-dimensional data matrix X using an appropriate method. For example, a similarity matrix can be constructed using k-nearest neighbors and a Gaussian kernel.

Next, from this similarity matrix, the graph Laplacian $L = D - W$ is defined, where D is the degree matrix satisfying $D_{ii} = \sum_{j=1}^{n} W_{ij}$. LE performs dimensionality reduction by solving the following optimization problem:

$$\min_{Y\in\mathbb{R}^{n\times d}} \operatorname{tr}(Y^\top L Y) \quad \text{s.t.} \quad Y^\top D Y = I, \tag{1}$$

where I is the identity matrix. This optimization problem reduces to the following generalized eigenvalue problem:

$$L\phi = \lambda D\phi, \tag{2}$$

where λ represents eigenvalues and ϕ represents eigenvectors. It is known that the graph Laplacian L always has zero as its minimum eigenvalue. Therefore, the low-dimensional representation $Y = [\phi^{(2)}, \phi^{(3)}, \cdots, \phi^{(d+1)}]$ is obtained using the eigenvectors corresponding to the second smallest eigenvalue $\phi^{(2)}$ through the $(d+1)$-th eigenvector $\phi^{(d+1)}$.

2.2 Locality Preserving Projection

Locality Preserving Projection (LPP)[11] is an unsupervised linear dimensionality reduction method proposed as a linear approximation of LE. LPP preserves the local structures inherent in LE while offering high interpretability through linear transformation. Furthermore, its ability to project new data points significantly expanded its practical applications.

For high-dimensional data $X \in \mathbb{R}^{n \times m}$, LPP seeks a low-dimensional representation $Z \in \mathbb{R}^{n \times d}$ and a projection matrix $F \in \mathbb{R}^{m \times d}$ satisfying $Z = XF$. The optimization problem for LPP is formulated as follows:

$$\min_{F \in \mathbb{R}^{m \times d}} \mathrm{tr}(F^\top X^\top L X F) \quad \text{s.t.} \quad F^\top X^\top D X F = I. \tag{3}$$

This reduces to the following generalized eigenvalue problem:

$$X^\top L X \psi = \lambda X^\top D X \psi, \tag{4}$$

where λ represents eigenvalues and ψ represents eigenvectors. The projection matrix $F = [\psi^{(1)}, \psi^{(2)}, \cdots, \psi^{(d)}]$ is obtained using the eigenvectors corresponding to the smallest eigenvalue $\psi^{(1)}$ through the d-th eigenvector $\psi^{(d)}$. The low-dimensional representation is calculated by $Z = XF$.

2.3 Modularity

Modularity[16] is a measure to evaluate community structures in graphs. This measure is innovative as it evaluates the statistical significance of community structures by comparing the given graph with a random graph model, and it has been widely adopted in complex network analysis.

Let $A \in \mathbb{R}^{n \times n}$ be the adjacency matrix of graph G, $\mathrm{vol}(G)$ be the total degree of graph G, δ be the Kronecker delta, k_i be the degree of vertex i, and $\mathbf{u} \in \mathbb{R}^n$ be a vector representing community partitioning. Then, modularity Q is formulated as follows:

$$Q = \frac{1}{\mathrm{vol}(G)} \sum_{ij}^{n} \left(A_{ij} - \frac{k_i k_j}{\mathrm{vol}(G)} \right) \delta(u_i, u_j). \tag{5}$$

Here, the modularity matrix is defined as $B_{ij} = A_{ij} - \frac{k_i k_j}{\mathrm{vol}(G)}$.

The graph partitioning method that directly maximizes modularity is called modularity maximization. Since modularity maximization is an NP-hard problem, approximate methods[15] are typically used in practice. In spectral methods, the problem is addressed by relaxing the discrete cluster assignments to continuous values:

$$\max_{\mathbf{u} \in \mathbb{R}^n} \mathbf{u}^\top B \mathbf{u} \quad \text{s.t.} \quad \mathbf{u}^\top \mathbf{u} = 1. \tag{6}$$

This reduces to the following eigenvalue problem:

$$B\mathbf{u} = \lambda \mathbf{u}. \tag{7}$$

The eigenvector $\mathbf{u}^{(1)}$ corresponding to the maximum eigenvalue of the modularity matrix gives the optimal segmentation.

Generalized Modularity One of the challenges of modularity maximization is the resolution limit[8]. This issue arises when the graph scale becomes large, making it difficult to detect smaller cluster structures. To address the resolution limit, a resolution parameter[17] was introduced, leading to the proposal of

generalized modularity[7]. This extension enhances the versatility of modularity maximization, allowing for the detection of cluster structures at various scales.

Resolution parameter-based generalized modularity is formulated as follows:

$$Q_\gamma = \frac{1}{\mathrm{vol}(G)} \sum_{ij}^{n} \left(A_{ij} - \gamma \frac{k_i k_j}{\mathrm{vol}(G)} \right) \delta(u_i, u_j). \tag{8}$$

Here, γ is called the resolution parameter, and the value of γ can be adjusted to control the size of the detected clusters. The larger γ is, the easier it is to detect small clusters.

2.4 Modularity Embedding

Modularity Embedding (ME)[12] is an unsupervised nonlinear dimensionality reduction method based on eigenvalue decomposition of the modularity matrix. ME applies graph partitioning theory to dimensionality reduction, enabling the acquisition of entirely new low-dimensional representations. Specifically, ME utilizes the modularity matrix to obtain low-dimensional representations where cluster centers are positioned far apart.

For a low-dimensional data matrix $Y \in \mathbb{R}^{n \times d}$, the optimization problem for ME is formulated as follows:

$$\max_{Y \in \mathbb{R}^{n \times d}} \mathrm{tr}(Y^\top B Y) \quad \text{s.t.} \quad Y^\top Y = I. \tag{9}$$

This reduces to an eigenvalue problem, and the low-dimensional data matrix Y is formed by selecting the d eigenvectors corresponding to the largest eigenvalues of the modularity matrix B.

3 Modularity Maximizing Projection

In this section, we propose Modularity Maximizing Projection (MMP), a linear dimensionality reduction method based on the modularity matrix. MMP is presented as a linear approximation of Modularity Embedding, achieving low-dimensional representations suitable for clustering via linear transformation.

For high-dimensional data $X \in \mathbb{R}^{n \times m}$, MMP seeks a low-dimensional representation $Z \in \mathbb{R}^{n \times d}$ and a projection matrix $F \in \mathbb{R}^{m \times d}$ satisfying $Z = XF$. While LPP employs the graph Laplacian, MMP utilizes the modularity matrix B to obtain low-dimensional representations that effectively preserve cluster structures.

First, a similarity matrix $W \in \mathbb{R}^{n \times n}$ is constructed using k-nearest neighbors and a Gaussian kernel.

$$W_{ij} = \begin{cases} \exp\left(-\frac{||\mathbf{x}_i - \mathbf{x}_j||_2^2}{2\sigma^2} \right) & \text{if } \mathbf{x}_i \in \mathcal{N}_k(\mathbf{x}_j) \vee \mathbf{x}_j \in \mathcal{N}_k(\mathbf{x}_i) \\ 0 & \text{otherwise} \end{cases}. \tag{10}$$

Here, $\mathcal{N}_k(\mathbf{x}_i)$ denotes the set of k-nearest neighbors of point $\mathbf{x}_i$, and σ represents the bandwidth of the Gaussian kernel.

Next, the modularity matrix is calculated using the constructed weighted undirected graph W. The modularity matrix is defined as the difference between the given graph W and a random graph P. The Chung-Lu random-graph model[6] is commonly used in the context of modularity. Consequently, the random graph similarity matrix P is defined as $P_{ij} = \frac{k_i k_j}{\mathrm{vol}(G)}$, using the degree k_i of vertex i and the total degree $\mathrm{vol}(G)$ of the graph. Thus, the modularity matrix $B \in \mathbb{R}^{n \times n}$ is defined as follows:

$$B = W - \frac{\mathbf{k}\mathbf{k}^\top}{\mathrm{vol}(G)}. \tag{11}$$

Under linear constraints, the optimization problem for MMP is formulated as follows:

$$\max_{F \in \mathbb{R}^{m \times d}} \mathrm{tr}(F^\top X^\top B X F) \quad \text{s.t.} \quad F^\top X^\top X F = I. \tag{12}$$

This reduces to the following generalized eigenvalue problem:

$$X^\top B X \psi = \lambda X^\top X \psi, \tag{13}$$

where λ represents eigenvalues and ψ represents eigenvectors. The projection matrix $F = [\psi^{(1)}, \psi^{(2)}, \cdots, \psi^{(d)}]$ is formed by selecting the eigenvectors corresponding to the d largest eigenvalues, from $\psi^{(1)}$ (associated with the largest eigenvalue) to $\psi^{(d)}$. The low-dimensional representation is calculated by $Z = XF$.

We now analyze the characteristics of the low-dimensional space obtained by MMP. The modularity matrix can be expanded as $B = -(D - W) + (D - P)$ using the degree matrix $D = \mathrm{diag}(k_1, k_2, \cdots, k_n)$. This expansion reveals that the modularity matrix is defined as the difference between two graph Laplacians. Based on this expansion, the quantity maximized in MMP can be expressed as follows:

$$\mathrm{tr}(F^\top X^\top B X F) = -\mathrm{tr}(F^\top X^\top (D - W) X F) + \mathrm{tr}(F^\top X^\top (D - P) X F). \tag{14}$$

From equation (14), the first term is the trace of the graph Laplacian based on the similarity matrix, which is consistent with the quantity minimized in LPP. LPP exerts attractive forces, drawing data points with high similarity closer together. The second term corresponds to the trace of the graph Laplacian based on the random graph model. As the random graph model is defined by the product of degrees, the second term introduces repulsive forces between data points with high degrees. This effect facilitates the acquisition of a low-dimensional space that emphasizes inter-cluster separation.

This effect also explains why ME could obtain low-dimensional representations where cluster centers are positioned far apart.

4 Modularity Regularized Locality Preserving Projection

In this section, we propose Modularity Regularized Locality Preserving Projection (MR-LPP), a method that incorporates both MMP and LPP through the introduction of a parameter γ. MR-LPP enables the acquisition of low-dimensional spaces with well-separated clusters by adjusting the parameter γ.

By multiplying the second term in Equation (14) by a parameter γ, we can control its repulsive effect. This parameter is also inspired by the resolution parameter in generalized modularity. We define the generalized modularity matrix B_γ, parameterized by γ, as follows:

$$B_\gamma = W - \gamma \frac{\mathbf{k}\mathbf{k}^\top}{\mathrm{vol}(G)}, \tag{15}$$

where W is the similarity matrix, $\mathbf{k} \in \mathbb{R}^n$ is the degree vector, and γ is a hyperparameter.

By replacing the modularity matrix B in the objective function of Equation (12) with B_γ, we define the objective function of MR-LPP as follows:

$$\mathrm{tr}(F^\top X^\top B_\gamma X F) = \mathrm{tr}(F^\top X^\top (D - W) X F) - \gamma \mathrm{tr}(F^\top X^\top (D - P) X F). \tag{16}$$

In Equation (16), when $\gamma = 0$, the objective function aligns with that of LPP, and when $\gamma = 1$, it aligns with that of MMP. Furthermore, by adjusting the parameter γ within the range $0 \leq \gamma \leq 1$, we can continuously change from LPP to MMP, positioning MMP as a generalization of LPP. At the same time, since the constraints of LPP and MMP are different, we also generalize the constraint using γ as follows:

$$\gamma F^\top X^\top X F + (1 - \gamma) F^\top X^\top D X F = I. \tag{17}$$

In practical scenarios, when a Gaussian kernel is used for weighting during graph construction, samples with a high density of neighboring data points in the high-dimensional space tend to have larger degrees. In such cases, as γ increases, samples with larger degrees are positioned further apart in the low-dimensional space. Conversely, samples with sparse surroundings in the high-dimensional space exhibit smaller degrees. In other words, samples distant from cluster centers or belonging to clusters with large variance are positioned relatively close to the origin. Therefore, by appropriately selecting γ within the range $0 \leq \gamma \leq 1$, it is possible to obtain a low-dimensional space that maintains a balanced overall cluster distribution by mitigating excessive spread.

5 Performance Evaluations

In this section, we evaluate the performance of the proposed MR-LPP method on multiple datasets.

The MR-LPP hyperparameters are set as follows. The number of neighbors for graph construction is fixed at 10. The Gaussian kernel bandwidth σ is set

such that the similarity between the closest points in the high-dimensional space is 1, while the similarity between the most distant points among the 10 nearest neighbors is 0.01.

For quantitative evaluation, we employ the Fisher Discriminant Ratio (FDR)[9]. FDR measures the ratio of between-class variance to within-class variance, thereby evaluating the degree of cluster structure separation. This metric aligns with the objective function of Linear Discriminant Analysis [9], where higher values indicate more clearly separated cluster structures. As FDR values are difficult to compare across different dimensions, we normalize them by the maximum value in each dimension to facilitate dimension-independent comparisons. In the quantitative evaluations, we observe the change in FDR values as the MR-LPP hyperparameter γ varies from 0 to 1 in increments of 0.05. Considering that in exploratory data analysis, it is desirable for essential information to be captured in the top dimensions, our evaluation focuses on the first few dimensions. For qualitative comparison, we utilize the existing unsupervised linear dimensionality reduction methods PCA and LPP.

In particular, this experiment demonstrates that the proposed method can effectively preserve cluster structures within datasets containing high-variance clusters.

5.1 Image Segmentation Data

First, we verify the effectiveness of the proposed method using the Image Segmentation dataset[20] from the UCI Machine Learning Repository. This dataset comprises pixel information from images classified into 7 classes, where each pixel is assigned a correct class label. Prior to analysis, the data was preprocessed using whitening Principal Component Analysis (PCA).

Figure 1 shows the 2D visualization results for each method. The parameter γ for MR-LPP was set to 0.7, based on Figure 2 (a). PCA exhibits insufficient separation between classes and fails to capture the underlying cluster structure. LPP prioritizes the identification of small cluster structures along the second dimension, thereby obscuring the major cluster structures concentrated in the lower left. In contrast, MR-LPP($\gamma = 0.7$) and MMP robustly capture the structure of densely populated clusters, yielding a low-dimensional space where clusters are easily distinguishable.

Figure 2 (a) illustrates the change in FDR values plotted against the MR-LPP hyperparameter γ. Note that when $\gamma = 0$, MR-LPP coincides with the LPP method, and when $\gamma = 1$, it coincides with MMP. In the 2D case, the highest FDR value is achieved at $\gamma = 0.7$, while in the 3D case, it is achieved at $\gamma = 1$. These results align well with the qualitative observations.

5.2 Materials Informatics Data

To evaluate the performance of our proposed method in a relevant application domain, we utilized a materials informatics dataset. This dataset contains local structural information for silicon (Si) atoms, derived from large-scale Density

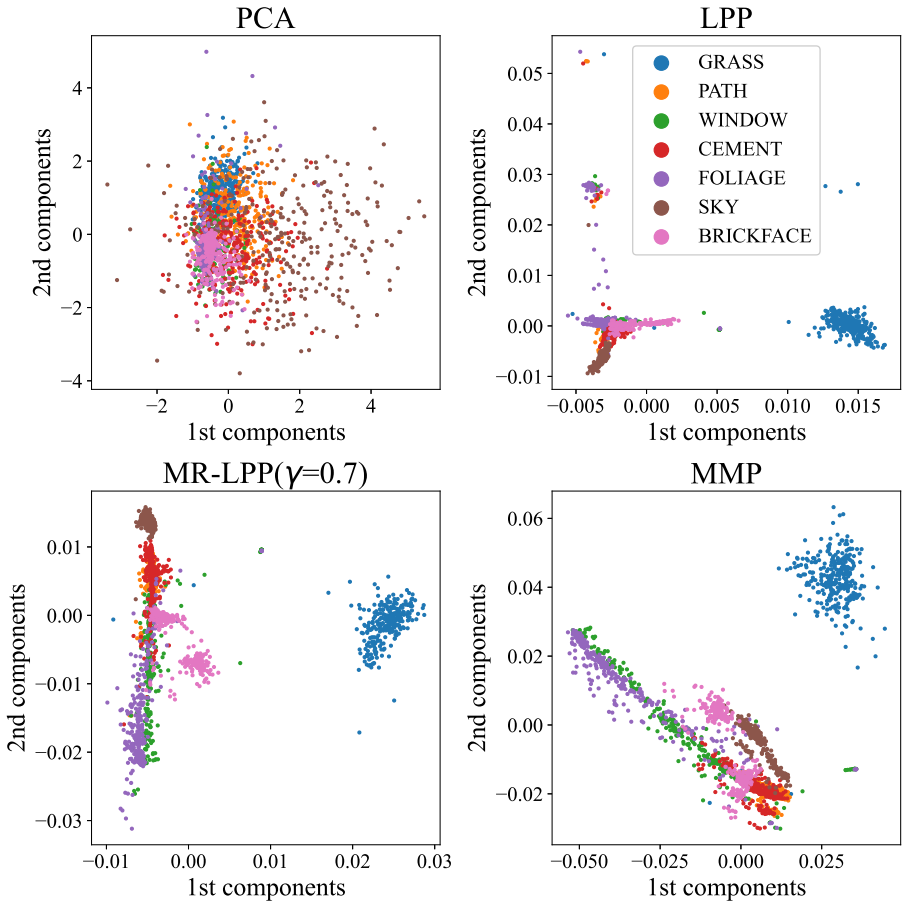
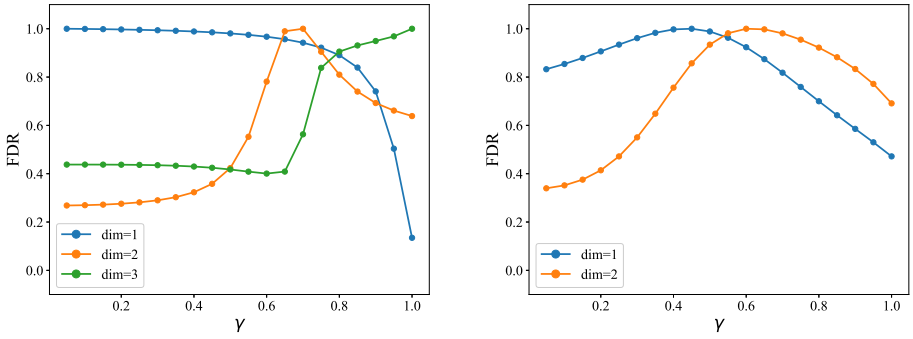


Fig. 1: 2D visualization results of PCA, LPP, MR-LPP($\gamma = 0.7$), and MMP on the Image Segmentation dataset.

Functional Theory molecular dynamics simulations of Si-Ge systems performed using the CONQUEST code[1][5]. The primary objective is the identification of distinct material phases as clusters within this dataset[13]. Each Si atom is characterized by a 40-dimensional feature vector, computed using the Locality Averaged Atomic Fingerprint (LAAF) method[19]. The LAAF hyperparameters were fixed at $R_c = 4.0$ Å and $R_a = 4.0$ Å. The dataset comprises 3500 samples, categorized into three distinct phases: crystalline, amorphous, and liquid. The liquid phase, which is present at high temperatures, tends to exhibit larger intra-cluster variance due to inherent thermodynamic fluctuations. Prior to analysis, the data was preprocessed using whitening Principal Component Analysis (PCA).



(a) Image Segmentation dataset (b) Materials Informatics dataset

Fig. 2: Graph comparing FDR values against γ for each dataset.

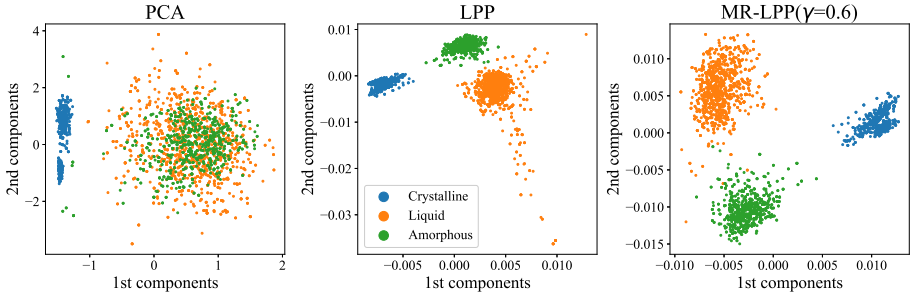


Fig. 3: 2D visualization results of PCA, LPP, and MR-LPP($\gamma = 0.6$) on the Materials Informatics dataset.

Figure 3 displays the 2D visualization results for each method. The MR-LPP parameter γ was set to 0.6, based on Figure 2 (b). PCA fails to distinguish between the liquid and amorphous cluster structures. In the LPP results, the liquid phase exhibits large variance and spreads excessively across the entire low-dimensional space. In contrast, MR-LPP mitigates the spread of the liquid cluster, resulting in a more balanced representation of each phase structure.

Figure 2 (b) illustrates the trend of FDR values with respect to the MR-LPP hyperparameter γ. By determining the target dimension and appropriately setting γ, higher FDR values can be achieved compared to the conventional LPP method (i.e., MR-LPP with $\gamma = 0$).

6 Conclusions

In this paper, we have proposed MR-LPP, an unsupervised linear dimensionality reduction method designed with clustering applications in mind. The main con-

tributions of this work are threefold: First, we have introduced MMP, a linear dimensionality reduction method based on the eigenvalue decomposition of the modularity matrix. MMP incorporates a repulsive effect based on the degrees of the graph constructed from high-dimensional data, enabling robust capture of cluster structures. Second, by introducing a parameter to control repulsion in MMP, we have developed MR-LPP, a method that extends LPP and continuously relates to MMP. Third, we have experimentally demonstrated that MR-LPP achieves more effective cluster separation than conventional methods, particularly in datasets with high-variance clusters. MR-LPP can adapt to various data characteristics through hyperparameter adjustment and can effectively capture cluster structures within the subspace spanned by the top eigenvectors. Owing to these characteristics, MR-LPP is expected to perform well as a preprocessing step for clustering, positioning it as a potentially valuable tool for exploratory data analysis.

Future work includes, first, investigating alternative graph construction algorithms. The choice of graph construction algorithm significantly influences how cluster structures are captured, as it markedly alters the degree distribution of the graph. Second, establishing guidelines for selecting the hyperparameter γ and determining the target dimensionality also warrants future research. Furthermore, a potential future direction involves the development of supervised dimensionality reduction methods based on modularity.

Acknowledgments. This study was supported by in part by JSPS KAKENHI Grant Number 21H01008, 23K20816, 22K12046 and 23H03411.

References

1. CONQUEST: Linear Scaling DFT. https://ordern.github.io/, accessed: May 2025
2. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. **15**(6), 1373–1396 (Jun 2003)
3. Bengio, Y., Courville, A., Vincent, P.: Representation learning: a review and new perspectives. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1798–1828 (Aug 2013)
4. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. J. Stat. Mech. **2008**(10), P10008 (Oct 2008)
5. Bowler, D.R., Choudhury, R., Gillan, M.J., Miyazaki, T.: Recent progress with large-scale *ab initio* calculations: the CONQUEST code. Phys. Status Solidi B Basic Res. **243**(5), 989–1000 (Apr 2006)
6. Chung, F., Lu, L.: Connected components in random graphs with given expected degree sequences. Annals of Combinatorics **6**(2), 125–145 (Nov 2002)
7. Fasino, D., Tudisco, F.: Generalized modularity matrices. Linear Algebra Appl. **502**, 327–345 (Aug 2016)
8. Fortunato, S., Barthélemy, M.: Resolution limit in community detection. Proc. Natl. Acad. Sci. U. S. A. **104**(1), 36–41 (Jan 2007)
9. Fukunaga, K.: Introduction to Statistical Pattern Recognition. Academic Press, San Diego, CA, 2nd edn. (1990), iSBN: 978-0122692628
10. Hagen, L., Kahng, A.: New spectral methods for ratio cut partitioning and clustering. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **11**(9), 1074–1085 (1992). https://doi.org/10.1109/43.159993

11. He, X., Niyogi, P.: Locality preserving projections. Adv. Neural Inf. Process. Syst. pp. 153–160 (Dec 2003)
12. Li, W.: Modularity embedding. In: Neural Information Processing. pp. 92–99. Springer Berlin Heidelberg (2013)
13. Lin, J., Tamura, R., Futamura, Y., Sakurai, T., Miyazaki, T.: Determination of hyper-parameters in the atomic descriptors for efficient and robust molecular dynamics simulations with machine learning forces. Phys. Chem. Chem. Phys. **25**(27), 17978–17986 (Jul 2023)
14. McInnes, L., Healy, J., Melville, J.: UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv [stat.ML] (Feb 2018)
15. Newman, M.E.J.: Finding community structure in networks using the eigenvectors of matrices. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. **74**(3 Pt 2), 036104 (Sep 2006)
16. Newman, M.E.J.: Modularity and community structure in networks. Proc. Natl. Acad. Sci. U. S. A. **103**(23), 8577–8582 (Jun 2006)
17. Reichardt, J., Bornholdt, S.: Statistical mechanics of community detection. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. **74**(1 Pt 2), 016110 (Jul 2006)
18. Shi, J., Malik, J., Member: Normalized cuts and image segmentation. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE **22**(8) (2000)
19. Tamura, R., Matsuda, M., Lin, J., Futamura, Y., Sakurai, T., Miyazaki, T.: Structural analysis based on unsupervised learning: Search for a characteristic low-dimensional space by local structures in atomistic simulations. Phys. Rev. B Condens. Matter **105**(7), 075107 (Feb 2022)
20. UCI Machine Learning Repository: Image Segmentation dataset. https://archive.ics.uci.edu/ml/datasets/Image+Segmentation (1990), accessed: May 2025