



# Align then Fusion: Generalized Large-scale Multi-view Clustering

## with Anchor Matching Correspondences

Siwei Wang, Xinwang Liu and En Zhu ({wangsiwei13,xinwangliu}@nudt.edu.cn)

<https://github.com/wangsiwei2010/NeurIPS22-FMVACC>



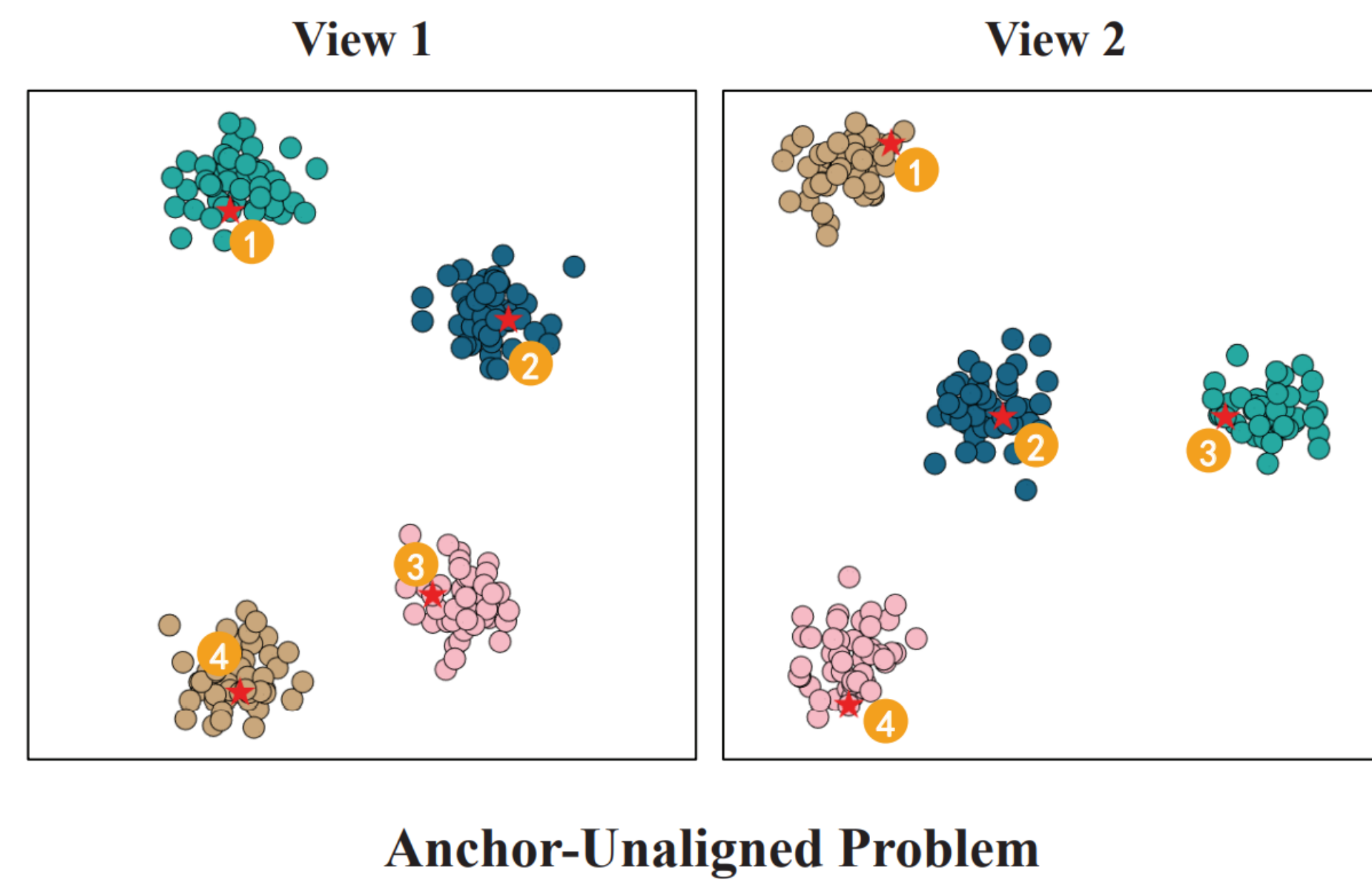
### Abstract

Existing MVGC approaches do not pay sufficient attention to establishing correct correspondences between the anchor sets across views. Such inaccurate anchor graph fusion may lead to wrong multi-view fusion and unsatisfactory performance. The AUP problem has been overlooked in existing literature and

### Introduction

Despite being widely applied in large-scale applications, one vital factor of successful multi-view anchor graph clustering is to build correct correspondences for anchor sets. The selected anchor sets generated by different views may mismatch without guarantees of correspondences, since k-means or sampling is performed on each view separately. Such an unaligned issue would cause inappropriate graph fusion and degrade the clustering performance in return. One pioneer work, SFMC provides an intuitive way to select samples with same indexes to implicitly avoid wrong correspondences. However, this way could destroy the flexibility of anchors. To the best of our knowledge, no generalized framework for flexible multi-view anchor correspondences has been proposed so far.

### Methodology



### Anchor-unaligned Problem (AUP)

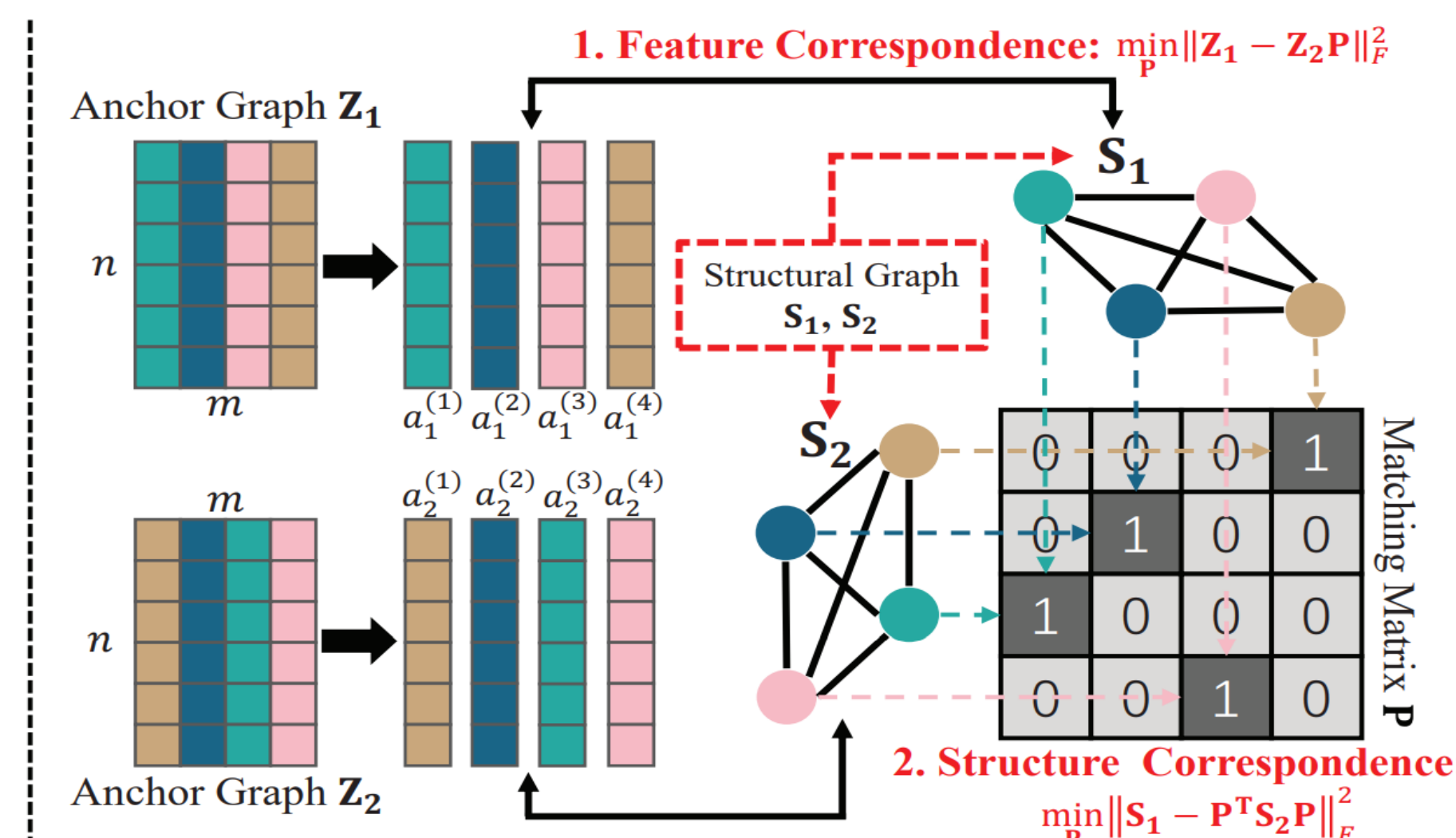
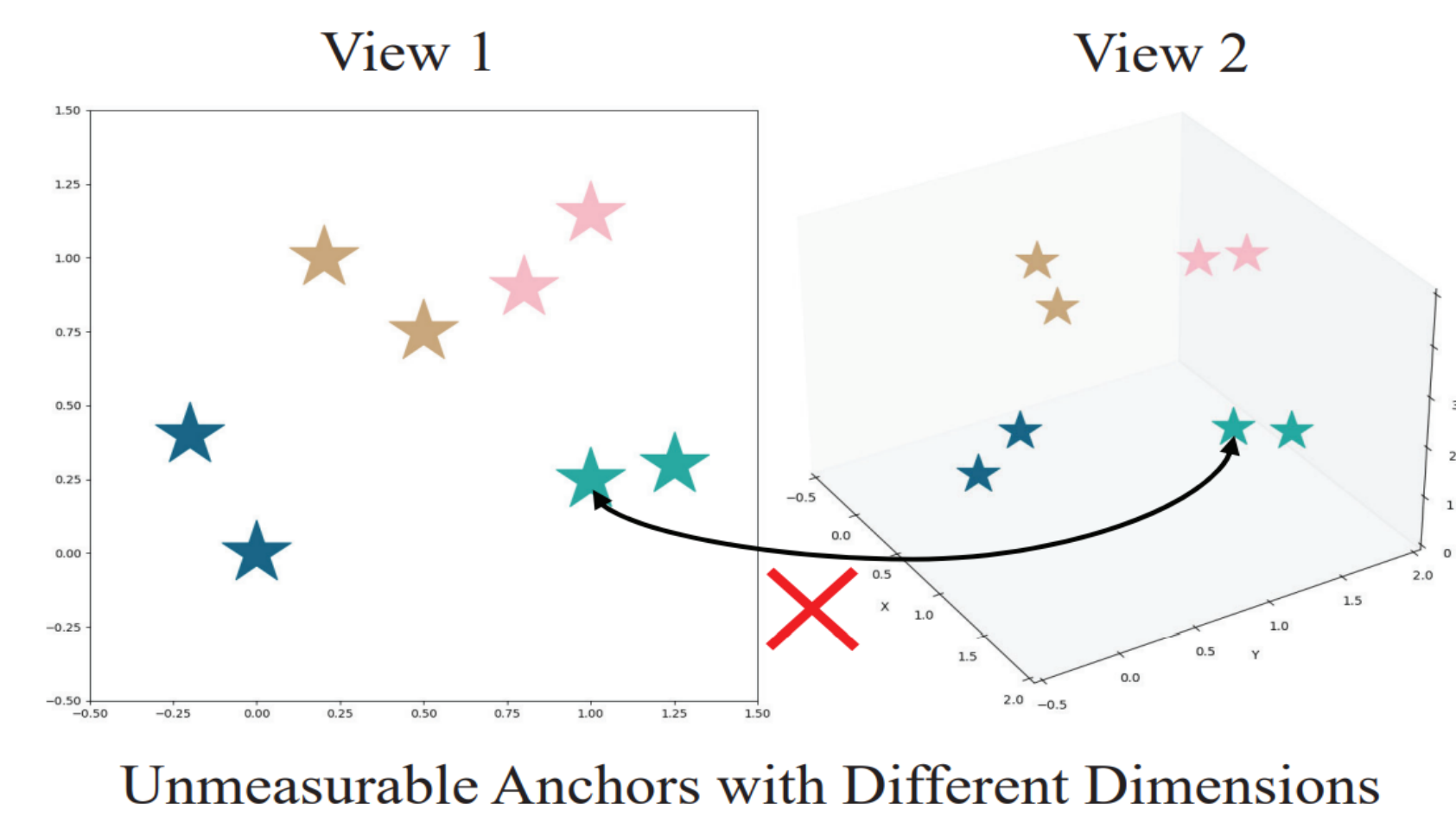
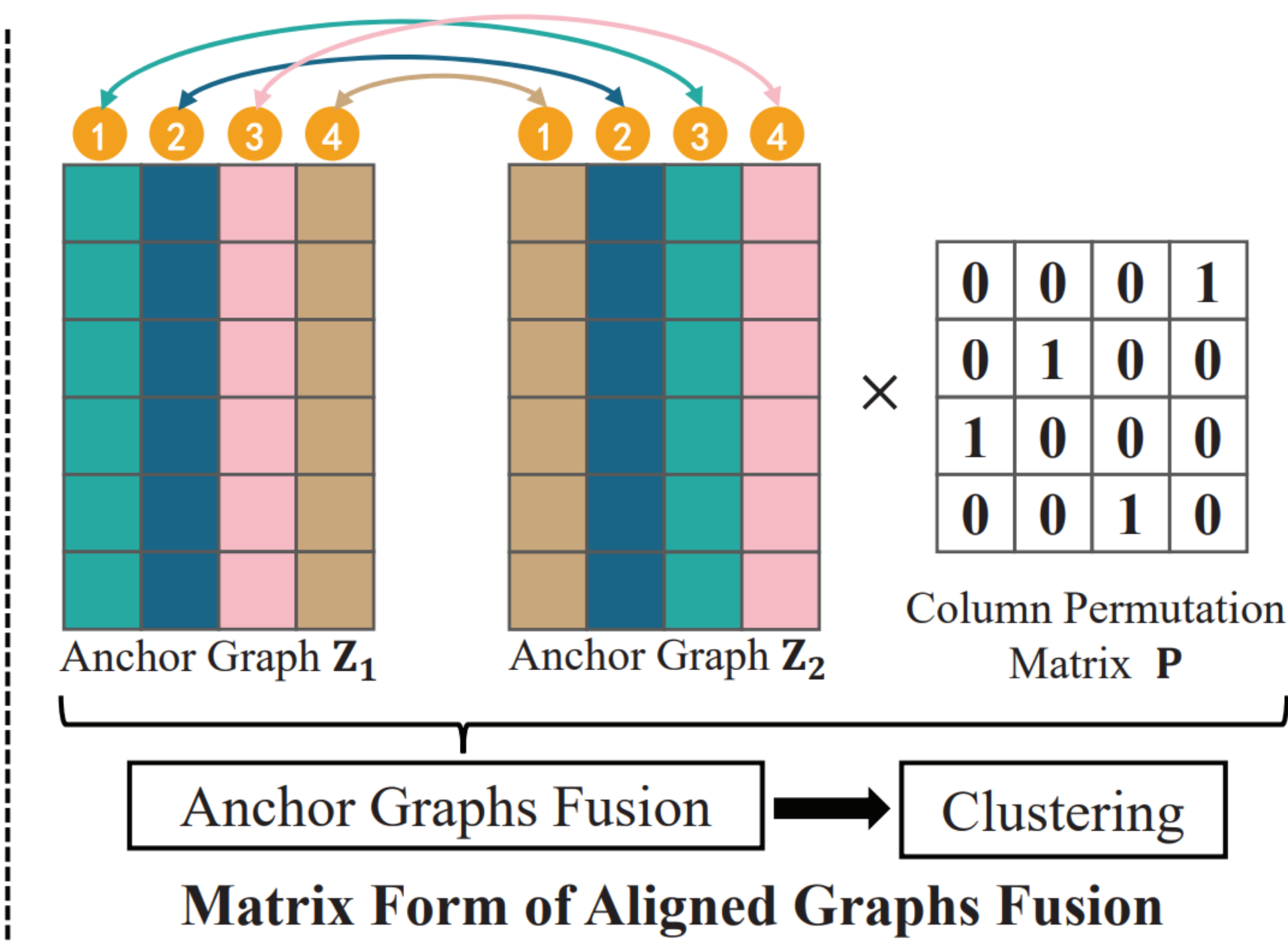


Figure 2: Overview of the proposed FMVACC consisting of two parts: feature and structure correspondence. (Left) Anchor sets are represented by different dimensions, which makes them **unmeasurable** in various metric spaces; (Right) In FMVACC, each column of the graph is taken as a new  $n$ -dimensional feature of each anchor, and the optimal matching matrix  $\mathbf{P}$  is obtained by minimizing the both the feature and structure correspondences.

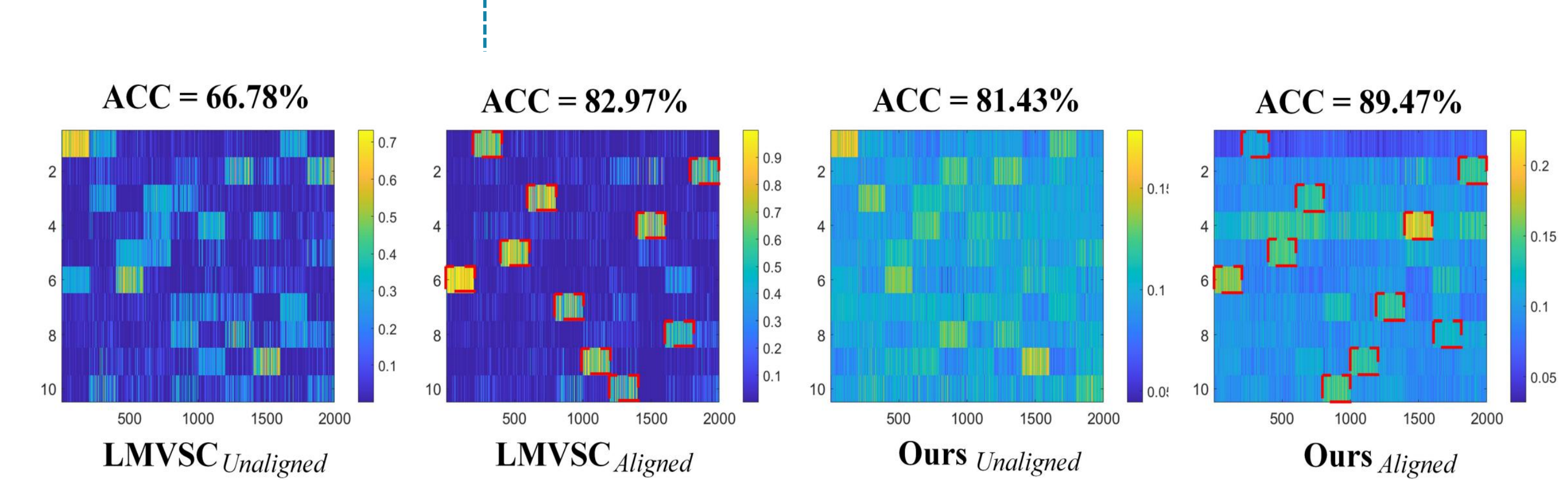


Figure 4: Visualization of the aligned and unaligned anchor graphs (LMVSC and Ours) on UCI-Digit.

### Conclusion

- We study a new paradigm for large-scale multi-view anchor graph clustering. The selected anchor sets in multi-view data are not aligned, which may lead to inaccurate graph fusion and degrade the clustering performance.
- We propose a flexible anchor graph fusion framework termed FMVACC to tackle the AUP problem. After generating flexible anchor candidates, an anchor alignment module is proposed to solve AUP with both feature and structure information. To the best of our knowledge, it is the first study of the flexible anchor correspondence fusion framework

### Contact

If you have any problem, please don't hesitate to contact Siwei Wang (wangsiwei13@nudt.edu.cn).