Review for

"Clustering of Diverse Multiplex Networks"

Jiaxin Hu

This work focuses on the clustering of diverse multiplex (DIMPLE) network, in which each single layer network follows a stochastic block model (SBM) and layers are partitioned into groups. The layers clustered in one group share the same network community structure while the probability matrices for all layers are allowed to be different. The clustering aims to recover the layer partition and the network communities in each group of layers. Main contributions of this work lie in the proposal of the general DIMPLE network model and the novel singular-space-based clustering techniques for layer partition.

The DIMPLE network model and the idea of singular-space-based clustering algorithm are novel and interesting to the field. Though, several concerns arise in the theoretical correctness and experiments, and there exists a room to improve the manuscript quality.

Major concerns:

1. Conditions for identification and signal-to-noise ratio (SNR) seem missing in the between-layer clustering. As Section 3.1 mentions, the success of between-layer clustering relies on the difference among the singular-spaces of different groups of networks, which is supposed to serve as the SNR in the context. Intuitively, the layer partition is identifiable if and only if SNR is larger than 0, and the algorithm should perform better as SNR increases. However, there is no rigorous definition or assumption on the gap among different group singular-spaces. Specifically, assumption A2, the only assumption related to multiple singular-spaces, can be satisfied with the equal singular-space case (i.e., $\boldsymbol{U}_z^{(1)} = \cdots = \boldsymbol{U}_z^{(M)}$) without any restriction on the rank r for $\bar{\boldsymbol{Z}}$. The lacks of identification and SNR condition lead to the doubts for all following analyses. For example, if we have $r^2 \leq M$, the conclusion $\text{rank}(\boldsymbol{F}) = M$ in Lemma 1 can not hold by any means. Exposing and highlighting the underlying conditions for identification and SNR is critical to improve the evaluation for the theoretical results.
2. More explanations are necessary for the result in Theorem 1. It is quite surprising that the guarantee for between-layer clustering in Theorem 1 is independent with the number of layers L and the gap among different singular-spaces (i.e., SNR). Intuitively, the clustering performance should increase when L increases with fixed number of groups M and when SNR increases. The simulations in Figure 7 and 8 also indicate that there is a general decreasing trend in between-layer clustering error with larger L. Adding discussions for this unexpected theoretical result helps to verify the correctness of the conclusions and improve readers' understanding to the novel clustering techniques.
3. More numerical experiments should be implemented for comparison and real application. Current methods including tensor block model (Wang and Zeng, 2019; Han et al., 2020; Hu

and Wang, 2022) and mixture multi-layer SBM (Lei et al., 2020; Fan et al., 2021; Jing et al., 2021) are able to recover the membership of layers and the network community structures. Numerical comparisons, for particular aspects (e.g., layer partition, model misspecification), with related methods help to emphasize advantage of the proposed method. In addition, real data analysis will strengthen the motivation and practical usage of the proposed method.

Minor concerns:

There are several typos and imprecise presentations.

1. Page 6, second paragraph, "diagonal of a matrix $\boldsymbol{A}$" $\to$ "diagonal of a square matrix $\boldsymbol{A}$" ;
2. Page 6, third equation, no definition for $\Pi_{\boldsymbol{U}}$;
3. Page 6, third paragraph, "$\boldsymbol{U} \in \boldsymbol{O}_{n_1,K}$" $\to$ "$\boldsymbol{U} \in \mathcal{O}_{n_1,K}$";
4. Page 7, first equation, "$\|\boldsymbol{U}^T\tilde{\boldsymbol{U}})\|_F^2$ " $\to$ "$\|\boldsymbol{U}^T\tilde{\boldsymbol{U}}\|_F^2$ ";
5. Page 8, equation (14), "$\tilde{\mathcal{V}}\tilde{\Lambda}\tilde{\mathcal{W}}$" $\to$ "$\tilde{\mathcal{V}}\tilde{\Lambda}\tilde{\mathcal{W}}^T$", as far as I am concerned;
6. Page 11, equation (22), it is unclear which term is minimized over $\aleph(M)$;
7. Page 16, Figure 3 and Figure 4 have exactly the same caption.

References

Fan, X., Pensky, M., Yu, F., and Zhang, T. (2021). Alma: Alternating minimization algorithm for clustering mixture multilayer network. *arXiv preprint arXiv:2102.10226*.

Han, R., Luo, Y., Wang, M., and Zhang, A. R. (2020). Exact clustering in tensor block model: Statistical optimality and computational limit. *arXiv preprint arXiv:2012.09996*.

Hu, J. and Wang, M. (2022). Multiway spherical clustering via degree-corrected tensor block models. In *International Conference on Artificial Intelligence and Statistics*, pages 1078–1119. PMLR.

Jing, B.-Y., Li, T., Lyu, Z., and Xia, D. (2021). Community detection on mixture multilayer networks via regularized tensor decomposition. *The Annals of Statistics*, 49(6):3181–3205.

Lei, J., Chen, K., and Lynch, B. (2020). Consistent community detection in multi-layer network data. *Biometrika*, 107(1):61–73.

Wang, M. and Zeng, Y. (2019). Multiway clustering via tensor block models. *Advances in neural information processing systems*, 32.