Third-round review for
"Clustering of Diverse Multiplex Networks"

Jiaxin Hu

We thank authors for the reply. The revised introduction and discussion make the model comparison with related works much clearer. However, the technical contribution is not well-conveyed, and there are still correctness and sharpness concerns to the theoretical results based on the reply.

- (Between-layer clustering) I agree that DIMPLE has more parameters than MMLSBM (Fan et al., 2022) due to the layer-specific connection matrices. But, when come to between-layer clustering, the membership matrix $\boldsymbol{C} \in \{0,1\}^{L\times M}$ is the only parameter of our estimation interest. Hence, I am confused why we consider the total number of parameters, including the within-layer estimation parameters $LK^2 + Mn\log K$, for the between-layer clustering. It is also curious how the extra complexity in connection matrices affects the between-layer clustering.

  Theoretically, for the between-layer clustering problem, both DIMPLE and MMLSBM have the same target parameter $\boldsymbol{C}$ and thus share the same number of parameters $L\log M$ with the same number of observations n^2L. For fixed K, M and no sparsity, the optimal between-layer clustering rate should have order $1/n^2$, which is the same as R_{BL}^{ALMA} reported in Fan et al. (2022). However, current version of Theorem 1 reports a between-layer clustering rate $1/n$, which seems sub-optimal. Is this sub-optimiality led by the extra complexity in connection matrices?

  In Algorithm 1, the proposed between-layer clustering relies only on the estimations of subspaces $\hat{\boldsymbol{U}}_l$'s. The subspace estimation should be robust to the core connection matrix. Because the subspaces spanned by $\boldsymbol{U}_l$'s do not change with layer-specific cores $\boldsymbol{B}^{(l)}$'s or group-specific cores $\boldsymbol{B}^{(m)}$'s, and the estimation errors of $\boldsymbol{U}_l$'s via SVD do not inflate with different $\boldsymbol{B}^{(l)}$'s under some spectral assumptions. Hence, the performance of Algorithm 1 seems invariant to the modelling of connection matrices. Then, the algorithm accuracy should be comparable with MMLSBM between-layer clustering performance or other network clustering algorithms.

  In addition, a notable part in Theorem 1 is the right hand side of inequality (38), the lower bound $1 - Ln^{-\tau}$, led by the usage of union bound over layers. When L becomes larger, we have a smaller probability for $R_{BL} < \mathcal{O}(K^2/n)$. This indicates that we have less confidence for the error bound when we have more samples in the clustering task. Is this observation a desirable or expected consequence of Theorem 1?

- (Within-layer clustering) I agree that the provided within-layer estimation error in Theorem 2 should involve the extra between-layer clustering error, and the DIMPLE with $M = 1$ is equivalent to model in Lei and Lin (2022). However, in the theoretical comparison of Corollary 1, the error bound (44) seems not a direct consequence of (43) after plugging $M = 1$. Adding more explanations would be helpful. Also, it is curious whether the "high probability" in (44) is still equal to the right hand side of (43). If so, the probability $1 - \mathcal{O}(Ln^{1-\tau})$ seems sub-optimal compared with the "high probability" $1 - \mathcal{O}(1/(L+n))$ in Lei and Lin (2022).

- (Contribution) I agree that DIMPLE imposes a generalized model from previous multi-layer SBM (Lei and Lin, 2022) and MMLSBM (Fan et al., 2022). The key concern is that whether the generalization leads to a significant technical innovation. In particular, one would wonder whether any DIMPLE-specific techniques are required in the algorithms and theoretical analysis. Though previous algorithms may not be directly applicable for DIMPLE model, the proposed algorithms seem not adapt innovative algorithmic strategies. Both between-layer and within-layer clustering use basic spectral clustering techniques to avoid the complexity in connection matrices. Pointing out the DIMPLE-specific techniques and adding more discussions on the technical innovations would better convey the novelty of the work.

References

Fan, X., Pensky, M., Yu, F., and Zhang, T. (2022). Alma: Alternating minimization algorithm for clustering mixture multilayer network. *Journal of Machine Learning Research*, 23(330):1–46.

Lei, J. and Lin, K. Z. (2022). Bias-adjusted spectral clustering in multi-layer stochastic block models. *Journal of the American Statistical Association*, pages 1–13.