# Review for

## "Efficient estimation in tensor Ising models"

This paper compares the efficiency of the maximum pseudolikelihood estimator (MPLE) and the maximum likelihood estimator (MLE) for the hypothesis testing problem under tensor Curie-Weiss Ising model. The phase transition of the efficiency comparison under different regimes of null and alternative parameters are provided. Simulations are implemented to verify the theortical results.

This is a theory paper without discussions and experiments for the real data applications. The theoretical discrepancy of estimation efficiency under the matrix and tensor Ising models are interesting. However, here are several concerns make me on the fence to this paper.

1. (Matrix vs Tensor) Main Theorems 2 and 3 show different Bahadur asymptotic relative efficiency (ARE) behavior under the matrix and tensor Ising models: when $\beta_0 > \beta^*(p)$, the ARE is equal to 1 in matrix case ($p = 2$) and shows two-layer phase transition in strictly tensor case ($p \geq 3$). There are two counter-intuitive observations bring correctness concerns to the main theorems.

    - The matrix conclusion in Theorem 2 is not a natural reduction of the strictly tensor conclusion in Theorem 3.

      Since matrix is a special case of tensor, the conclusions for tensor models should be able to apply for the matrix models by taking $p = 2$. However, Theorems 2 and 3 disagree with this intuition. Specifically, the window $(\beta^*(p), \log 2)$ does not shrink to empty when $p = 2$. Discussions should be provided to explain this theoretical discontinuity between the matrix and tensor cases. Moreover, the critical threshold $\log 2$ is counter-intuitive due to its invariance to the order $p$. More discussions about the occurrence of the window $(\beta^*(p), \log 2)$ and the interpretation of $\log 2$ may help the understanding of Theorem 3.
    - In Theorem 3, a larger order $p$ does not lead to a larger behavior discrepancy between matrix and tensor models.

      Intuitively, people would expect a larger behavior difference between matrices and strictly tensors with a larger order $p$, because of the larger difference between the pairwise and higher-order information. However, Theorem 3 contradicts to this intuition. When $p \to \infty$, the window $(\beta^*(p), \log 2)$ shrinks to empty, and thus the two-layer phases for tensor estimation collapse to a single phase. Hence, Theorem 3 indicates that tensors with a larger order $p$ have a more similar behavior as matrices. Reasonable explanations are necessary to understand this counter-intuitive phenomenon.

2. (Hypergraph) Following questions bring correctness concerns to the Theorem 6:

- The adjacency tensor $\boldsymbol{A}$ has random entries, and thus the probability in equation (3.1) is random. Is (3.1) a valid probability distribution? Is the randomness in $\boldsymbol{A}$ reflected in properties of the estimators?
- The asymptotic minimum sample sizes for model (1.1) in Theorem 1 are independent with the mean $\alpha_n$ in the Bernoulli random variables. It is curious why $\alpha_n$ has no effect to the efficiency comparison between MPLE and MLE. By Lemma 1 and Corollary 1, assumptions on the order of $\alpha_n$ seems necessary to obtain Theorem 6.
- In reference Mukherjee et al. (2022), the critical value $\beta^*(p)$ for the hypergraph model is dependent to the mean $\alpha_n$. Are the relation between $\beta^*(p)$ and $\log 2$ and the phase transition in Section 2 still valid for hypergraph model without any conditions on $\alpha_n$?

3. (Practice) Several comments come from the perspective of practical usage.

   - Including the specific algorithms to compute the MLE and MPLE in simulations, time and computational resource comparisons, and the explicit computational complexities of the algorithms may better illustrate the computational efficiency of MPLE.
   - Adding real-data applications may better illustrate the practical usage of MPLE.
   - In the real scenario, how to choose the order $p$ if one want to fit the Ising model, considering the different performances of MPLE with different $p$?

## References

Mukherjee, S., Son, J., and Bhattacharya, B. B. (2022). Estimation in tensor ising models. *Information and Inference: A Journal of the IMA*, 11(4):1457–1500.