On Learning from Low Rank Tensor Data: A Random Tensor Theory Perspective

1. **Summary and Contributions: Briefly summarize the paper and its contributions**

   This work studies the theoretical property of supervised and unsupervised discriminant problems with tensor observations. The clustering performance guarantees for vectorization based and low-rank tensor decomposition based methods are provided. The theorems indicate the benefits of exploiting low-rank tensor structure when the observations are generated from the (low-rank tensor signal) + (noise model).

2. **Strengths: Please describe the strengths of the work according (but not limited) to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, and relevance to the AISTATS community.**

   All simulations and figures are informative and well reveal the difference between vectorization and tensor based methods. The phase transition of tensor discriminant problem in Figure 6 is interesting.

3. **Weaknesses: Please describe the limitations of this work according (but not limited) to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, and relevance to the AISTATS community.**

   - The key conclusion "fewer training sample achieves better performance with low-rank tensor" is not super surprising. As the observations are assumed to have underlying low-rank structure (equation 2), it is natural that low-rank tensor based method has a better performance and vectorization method has a worse performance due to model misspecification. This conclusion also is also not applicable for general applications since we have no idea about the low-rank structure in real life.

   - The phase transition in Figure 6 seems not solid for me. Propositions 3.6 and 3.7 do not show the minimax rates but show the performance for a paritcular method. Then, it seems not sound to state the "impossibility" and "NP-hard" for the problem and all possible estimator.

4. **Correctness: Are the method and claims correct? Is the empirical methodology correct?**

   The conclusions seem sound.

5. **Clarity: Is the paper well written? Does it clearly state its contributions, notation and results?**

   Most parts of the manuscript are well-written.

- Adding more explanations to the complex results, for example the interpretation of function Q and equation (8), may benefit the understanding. - Under the Assumption 2.2, the theoretical results may be more concise if we replace p by $O(kn)$ and replace d by $O(n^k)$. - Organizing the algorithms step by step under the "algorithm" environment may be a better way.

6. **Relation to prior work: Is it clearly discussed how this work differs from or relates to prior work in the literature?**

   As far as I am concerned, there are rich and increasing literature that process tensor data with low-rank structure and consider the exact (or optimal) estimation. For example, [1] [2] study the exact recovery of block structured tensor observations and rigorously show the statistical-computational gap similar with the phase transition in Figure 6; [3] [4] study the optimal estimation with low-rank tensors; and so on for other tensor problems including tensor discriminate, regression, and completion. So, I personally disagree with the statement that "...few works in the literature were focused on the exact estimation of the performance of ML methods when processing tensor data with low-rank structure."

   [1] Han, Rungang, Yuetian Luo, Miaoyan Wang, and Anru R. Zhang. "Exact clustering in tensor block model: Statistical optimality and computational limit." arXiv preprint arXiv:2012.09996 (2020).

   [2] Hu, Jiaxin, and Miaoyan Wang. "Multiway Spherical Clustering via Degree-Corrected Tensor Block Models." In International Conference on Artificial Intelligence and Statistics, pp. 1078-1119. PMLR, 2022.

   [3] Zhang, Anru, and Dong Xia. "Tensor SVD: Statistical and computational limits." IEEE Transactions on Information Theory 64, no. 11 (2018): 7311-7338.

   [4] Xia, Dong, Anru R. Zhang, and Yuchen Zhou. "Inference for low-rank tensors—no need to debias." The Annals of Statistics 50, no. 2 (2022): 1220-1245.

7. **Additional Comments: Add your additional comments, feedback and suggestions for improvement, as well as any further questions for the authors.**

   - Adding real data analysis may better support the conclusions. - Highlighting the specific random tensor theory techniques (e.g., random tensor concentration, spectral property of random tensor) used in the analysis may better reflect the title.

References