We thank all reviewers for helpful comments and feedback! Please see our responses to each reviewer's critical concerns below.

Reviewer #2.

- *"Standard deviation in the experimental part":* please see Figure 1. Results are shown with 95% confidence intervals. The CIs can hardly be seen in the right figure. We will update the paper with these figures.

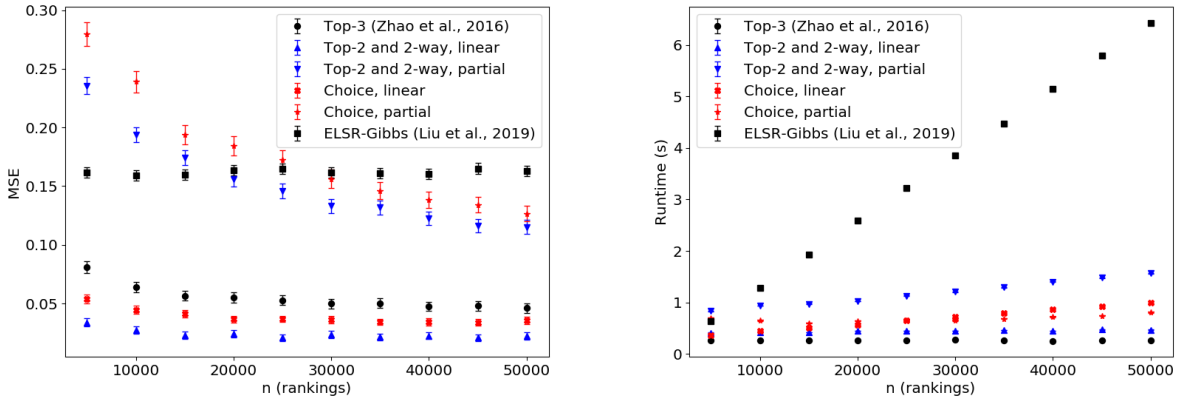

Figure 1: MSE and runtime with 95% confidence intervals.

- *"Real data study":* we feel that large-scale experiments on real-world data may not serve the purposes of this (mostly theoretical) paper, because no ground truth parameter is given in such datasets. We certainly believe that large-scale experiments on real-world datasets are important and would greatly appreciate the reviewer's suggestions on how to conduct such experiments.

Reviewer #3.

- *"Identifiability theorems should cover broader range of structured partial orders":* As a first step, our theorems have already significantly extended some previous work in non-trivial ways: for example, Theorem 2 covers a much broader range of partial orders than previous works ([7] and [33], as discussed in L107 and L118 of the submission). We believe that the proofs for our theorems are highly nontrivial and similar techniques do not seem to work for more general cases. Identifiability for other structures and $k \geq 3$ are certainly important (and highly challenging) future directions as the reviewer suggested.
- *"Experiments on real data and beyond mixture of 2 components"* Please see our response to Reviewer #2 on real data experiments. Identifiability of mixtures of three or more components is still an open question. Therefore, we felt that it is a little premature to run experiments on such cases. Extending the algorithm is not hard, but even the consistency of the algorithm is not known and hard to prove.
- *"Combinatorial explosion in the number of components":* Our proposed algorithms are only designed for $k = 2$. GMM algorithms for $k \geq 3$ cases need to be carefully designed w.r.t. identifiability, which is still an open question. Existing EM heuristics (as done in [16]) can be applied but their theoretical guarantee, e.g. consistency, is unclear.
- *"How does one extend the algorithm beyond just the data structures stated in Theorem 2":* We believe that new structures can be conveniently integrated into the proposed framework by introducing a parameter to represent the probability of generating the new structure. Nonetheless, it can be highly non-trivial to prove the identifiability of the new model and design consistent algorithms as we did for the most commonly-studied structures. We will add more discussions.

Reviewer #4. Thank you for your very encouraging comments!