# Changes in final file

This file lists the changes in the final manuscript compared with the submission after the second-round revision. We list the change by types and code the modification by red color.

**Corrected explanations and statements.**

1. Page 6, line 9. We revised the explanations for the boundedness constraints $c_3, c_4$.

   "... Third, the constant $c_3$ requires that all slides in $\mathcal{S}$ have non-degenerate norm. Particularly, the lower bound $c_3$ excludes the no purely zero slide case to avoid trivial non-identifiability of model (2); see Example 2 below. The upper bound $c_4$ is a technical constraint to avoid the slides with unbounded norm as dimension grows; in practice, the constraint $\max_{a:} \|\text{Mat}(\mathcal{S})_{a:}\| \leq c_4$ would likely never be active with a large $c_4 \geq \|\mathcal{Y}\|_F$. ..."

2. Page 8, before Remark 2. We corrected the conclusions related to $\Delta^2_{\boldsymbol{X}}$.

   "Based on our theory in later Sections, the dTBM is impossible to solve when $\Delta^2_{\boldsymbol{X}} \lesssim p^{-1}$; ..."

3. Theorem 2. We corrected the theoretical impossibility result related to $\Delta^2_{\boldsymbol{X}}$.

   "... Further, the impossibility can also be described by the minimal angle in mean tensor $\Delta^2_{\boldsymbol{X}}$ defined in (9). Let $\mathcal{P}'(\gamma') := \mathcal{P} \cap \{\Delta^2_{\boldsymbol{X}} = p^{\gamma'}\}$. We have

   $$\gamma' < -(K-1) \quad \Rightarrow \quad \liminf_{p\to\infty} \inf_{\hat{z}_{\text{stat}}} \sup_{(z,\boldsymbol{\theta},\mathcal{S})\in\mathcal{P}'(\gamma')} \mathbb{E}\left[p\ell(\hat{z}_{\text{stat}}, z)\right] \geq 1."$$

**Minor technical condition changes.**

Minor technical condition changes include adding ranges for the fixed number of communities $r$ and orders $K$, emphasizing the asymptotic results with "$p \to \infty$", and clarifying technical assumptions in the theorems.

1. Theorem 1. Add range of $K$.

   "Consider the dTBM with $r \geq 2$ and $K \geq 2$ ..."

2. Lemma 1. Add range of $r, p$, and clarify the assumption on $\boldsymbol{\theta}$.

   "Consider the dTBM model (2) under the parameter space $\mathcal{P}$ in (3) with $r \geq 2$. Suppose $\boldsymbol{\theta}$ is balanced satisfying (7) and $\min_{i\in[p]} \theta(i) \geq c$ from some constant $c > 0$. Then, as $p \to \infty$, for all $i, j$ such that $z(i) \neq z(j)$, we have ..."

3. Theorem 2. Add range of $r, K, p$.

   "**Impossibility.** Assume $p \to \infty$ and $2 \leq r \lesssim p^{1/3}$. ..."

"**MLE achievability.** Suppose the signal exponent satisfies $\gamma > -(K-1)+c_0$ for an arbitrary constant $c_0 > 0$. Furthermore, assume $\boldsymbol{\theta}$ is balanced and $\min_{i\in[p]} \theta(i) \geq c$ from some constant $c > 0$. Then, when $p \to \infty$, fixed $r \geq 1$, and $K \geq 3$, the MLE ..."

4. Theorem 3. Add range of $r, K, p$, and clarify the assumption for polynomial achievability.

   "**Impossibility.** Assume HPC conjecture holds and $r \geq 2$. ..."

   "**Polynomial-time algorithm achievability.**Suppose the parameter space satisfies $\gamma > -K/2 + c_0$ for an arbitrary constant $c_0 > 0$. Furthermore, assume fixed $r \geq 1, K \geq 2,$ the degree is balanced, lower bounded in that $\min_{i\in[p]} \theta_i \geq c$ for some constant $c > 0$, and satisfies the local linear stability in Definition 2 in the neighborhood $\mathcal{N}(z, \varepsilon)$ for all $\varepsilon \leq E_0$ and some $E_0 \geq \check{C} \log^{-1} p$ with some positive constant $\check{C}$. Then, as $p \to \infty$, there exists ... "

   The discussion following Theorem 3 also changes.

   "... and the second part shows the existence of such algorithm when $\gamma > -K/2 + c_0$ for an arbitrary constant $c_0 > 0$ under extra technical assumptions. ..."

5. Lemma 2. Add the parameter space.

   "Consider the dTBM under the parameter space $\mathcal{P}$. Suppose ..."

6. Theorem 4. Add range of $r, K, p$.

   "Consider the general sub-Gaussian dTBM with fixed $r \geq 1$, $K \geq 2$, i.i.d. noise ..."

   "...With probability going to 1 as $p \to \infty$, we have ..."

   "...We have ... with probability going to 1 as $p \to \infty$. ..."

7. Theorem 5. Add range of $r, K, p$, and clarify the assumption of local linearity.

   "Consider the general sub-Gaussian dTBM with fixed $r \geq 1$, $K \geq 2$, independent noise under the parameter space $\mathcal{P}$, and Assumption 1. Assume the local linear stability of degree holds in the neighborhood $\mathcal{N}(z, \varepsilon)$ for all $\varepsilon \leq E_0$ and some $E_0 \geq \check{C} \log^{-1} p$ with some positive constant $\check{C}$. Let $\{z_k^{(0)}\}_{k=1}^K$ be the initialization for Sub-Algorithm 2 and $z_k^{(t)}$ be the $t$-th iteration output on $k$-th mode. Suppose $\min_{i\in[p]} \theta(i) \geq c$ for some constant $c > 0$, ..."

   "... With probability going to 1 as $p \to \infty$, there exists ..."

8. Corollary 1. Add range of $p$, and clarify the statement.

   "... Combining all parameter assumptions and the results in Theorems 4 and 5, with probability going to 1 as $p \to \infty$, ..."

9. Proposition 1. Add range of $r, K, p$.

   "Consider the Bernoulli dTBM in the parameter space $\mathcal{P}$ with fixed $r \geq 1, K \geq 2$ and Assumption 1 holds. ... With probability going to 1 as $p \to \infty$, we have ..."

*Following minor changes occur in Appendices.*

10. Proof of Theorem 2in Appendix D. Add range of $r, K, p$.

    "Consider the general asymmetric dTBM (27) in the special case that $p_k = p$ and $r_k = r$ for all $k \in [K]$ with $K \geq 1$, $2 \leq r \lesssim p^{1/3}$ as $p \to \infty$. For simplicity, ..."

11. Proof of Theorem 3 in Appendix E. Add range of $r$.

    "... show the computational lower bound for a special class of degree-corrected tensor clustering model with $K \geq 2$ and $r \geq 2$. We construct ..."

12. Proof of Theorem 4 in Appendix F. Add range of $r$, and add discussion of extreme case with $r = 1$.

    "We prove Theorem 4 under the dTBM (2) with symmetric mean tensor, parameters $(z, \mathcal{S}, \boldsymbol{\theta})$, fixed $r \geq 1, K \geq 2$, and i.i.d. noise. For the case $r = 1$, we have $L(z^{(0)}, z) = 0, \ell(z^{(0)}, z) = 0$ trivially. Hence, we focus on the proof of the first mode clustering $z_1^{(0)}$ with $r \geq 2$; the proofs for the other modes ..."

13. Lemma 6. Clarify the constant assumption.

    " Under the parameter space (3) and assumption that $\min_{i \in [p]} \theta(i) \geq c$ for some constant $c > 0$, the singular values ..."

14. Lemma 8. Add range of $r$.

    "Let $z : [p] \mapsto [r]$ be a cluster assignment such that $|z^{-1}(a)| \asymp p/r$ for all $a \in [r]$ with $r \geq 2$. ..."

15. Lemma 9. Add condition for dimension $n, m$.

    " where $n > m$ and $\boldsymbol{E}$ contains independent ..."

16. Proof of Theorem 5 in Appendix G. Add range of $r$ and add discussion of extreme case with $r = 1$.

    "We consider dTBM (2) with symmetric mean tensor, parameters $(z, \mathcal{S}, \boldsymbol{\theta})$, fixed $r \geq 1, K \geq 2$, and i.i.d. noise. Let $(\hat{z}, \hat{\boldsymbol{\theta}}, \hat{\mathcal{S}})$ denote the MLE in (10), and $(z_k^{(0)}, \boldsymbol{\theta}^{(0)}, \mathcal{S}^{(0)})$ denote parameters related to the initialization. For the case $r = 1$, we have $\ell(z^{(t)}, z) = 0$ trivially for all $t \geq 0$. Hence, we focus on the proof of the first mode clustering $z_1^{(t+1)}$ with $r \geq 2$; the extension for other modes..."

17. Lemma 11. Add range of $r, p$.

    "Under the Condition 1 and the setup of Theorem 5 with fixed $r \geq 2$, assume ... As $p \to \infty$, we have ..."

18. Lemma 12. Add range of $r, p$.

    " Under the Condition 1 and the setup of Theorem 5 with fixed $r \geq 2$, as $p \to \infty$, ..."

19. Lemma 13. Add range of $K$, and the assumptions on $\boldsymbol{\theta}$.

" Let $(\hat{z}, \hat{\boldsymbol{\theta}}, \hat{\mathcal{S}})$ denote the MLE in (10) with fixed $K \geq 2$, and $\hat{\mathcal{X}}$ denote the mean tensor consisting of parameter $(\hat{z}, \hat{\boldsymbol{\theta}}, \hat{\mathcal{S}})$. With high probability going to 1 as $p \to \infty$, we have ... When SNR $\gtrsim p^{-(K-1)} \log p$, $\boldsymbol{\theta}$ is balanced and $\min_{i \in [p]} \theta(i) \geq c$ for some constant $c$, the MLE satisfies ..."

**Proof modifications.**

1. Proof of Theorem 1. Add discussions for extreme cases.

   "... if the model (27) violates Assumption 2. Note that $\Delta^2_{\min} = 1$ when there exists a $k \in [K]$ such that $r_k = 1$. Hence, we consider the case that $r_k \geq 2$ for all $k \in [K]$. Without loss of generality,..."

   "... First, we show the uniqueness of $\boldsymbol{M}_k$ for all $k \in [K]$. When $r_k = 1$, all possible $\boldsymbol{M}_k$ is equal to the vector $\mathbf{1}_{p_k}$, and the uniqueness holds trivially. Hence, we consider the case that $r_k \geq 2$. Without loss of generality, we consider $k = 1$ with $r_1 \geq 2$ and show the uniqueness of the first mode membership matrix; i.e., $\boldsymbol{M}_1' = \boldsymbol{M}_1 \boldsymbol{P}_1$ where $\boldsymbol{P}_1$ is a permutation matrix. The conclusion for $k \geq 2$ can be showed similarly and thus omitted. ..."

2. Proof of Lemma 1. Revise the last part for better and more concise presentation.

   ..." Then, we have

   $$\cos(\boldsymbol{S}_{z_1(i):}, \boldsymbol{S}_{z_1(j):}) = \frac{\langle \boldsymbol{S}_{z_1(i):}, \boldsymbol{S}_{z_1(j):} \rangle}{\|\boldsymbol{S}_{z_1(i):}\| \|\boldsymbol{S}_{z_1(j):}\|} = (1+o(1)) \frac{\langle \boldsymbol{X}_{i:}, \boldsymbol{X}_{j:} \rangle}{\|\boldsymbol{X}_{i:}\| \|\boldsymbol{X}_{j:}\|} = (1+o(1)) \cos(\boldsymbol{X}_{i:}, \boldsymbol{X}_{j:}),$$

   where the second inequality follows by the balance assumption on $\boldsymbol{\theta}$.

   Further, notice that $\|\boldsymbol{v}_1^s - \boldsymbol{v}_2^s\|^2 = 2(1 - \cos(\boldsymbol{v}_1, \boldsymbol{v}_2))$. For all $i, j$ such that $z(i) \neq z(j)$, when $p \to \infty$, we have

   $$\|\boldsymbol{X}_{i:}^s - \boldsymbol{X}_{j:}^s\| \asymp \|\boldsymbol{S}_{z_1(i):}^s - \boldsymbol{S}_{z_1(j):}^s\| \gtrsim \Delta_{\min}."$$

3. Proof of Theorem 2. Add discussions related to the conclusions of $\Delta^2_{\boldsymbol{X}}$.

   "... Last, with constructed $z_k^*, \boldsymbol{\theta}_k^*$ satisfying properties (i) and (ii) and $\gamma' < -(K-1)$, we construct a core tensor $\mathcal{S}^*$ such that $\Delta^2_{\boldsymbol{X}^*} \leq p^{-(K-1)}$. Based on the property (ii) and the boundedness constraint of $\mathcal{S}^*$ in $\mathcal{P}$, we still have $\|\theta^* \boldsymbol{x}_1^* - \theta^* \boldsymbol{x}_2^*\|_F \leq 2c_4\sigma$. Hence, we obtain the desired result

   $$\liminf_{p \to \infty} \inf_{\hat{z}_1} \sup_{(z, \boldsymbol{\theta}, \mathcal{S}) \in \mathcal{P}'(\gamma')} \mathbb{E}\left[p\ell(\hat{z}_1, z_1)\right] \geq \liminf_{p \to \infty} \inf_{\hat{z}_{\text{stat}}} \mathbb{E}\left[p\ell(\hat{z}_1, z_1^*) | z_k^*, \boldsymbol{\theta}_k^*, \mathcal{S}^*\right] \geq 1."$$

4. Proof of Lemma 9. Use a more classical prior result in the proof.

   *Proof of Lemma 9.* Note that $\|\hat{\boldsymbol{X}} - \boldsymbol{Y}\|_F^2 \leq \|\boldsymbol{X} - \boldsymbol{Y}\|_F^2$ by the definition of least square estimator.

We have

$$\|\hat{\boldsymbol{X}} - \boldsymbol{X}\|_F^2 \le 2\left\langle \hat{\boldsymbol{X}} - \boldsymbol{X}, \boldsymbol{Y} - \boldsymbol{X} \right\rangle$$
$$\le 2\|\hat{\boldsymbol{X}} - \boldsymbol{X}\|_F \sup_{\boldsymbol{T}\in\mathbb{R}^{m\times n}, \mathrm{rank}(\boldsymbol{T})\le 2r, \|\boldsymbol{T}\|_F=1} \langle \boldsymbol{T}, \boldsymbol{Y} - \boldsymbol{X}\rangle \tag{1}$$

with probability at least $1 - \exp(-C_2 nr)$, where the second inequality follows by re-arrangement.

Consider the SVD for matrix $\boldsymbol{T} = \boldsymbol{U}\Sigma\boldsymbol{V}^T$ with orthogonal matrices $\boldsymbol{U} \in \mathbb{R}^{m\times 2r}, \boldsymbol{V} \in \mathbb{R}^{n\times 2r}$ and diagonal matrix $\Sigma \in \mathbb{R}^{2r\times 2r}$. We have

$$\sup_{\boldsymbol{T}\in\mathbb{R}^{m\times n}, \mathrm{rank}(\boldsymbol{T})\le 2r, \|\boldsymbol{T}\|_F=1} \langle \boldsymbol{T}, \boldsymbol{Y} - \boldsymbol{X}\rangle = \sup_{\boldsymbol{T}\in\mathbb{R}^{m\times n}, \mathrm{rank}(\boldsymbol{T})\le 2r, \|\boldsymbol{T}\|_F=1} \langle \boldsymbol{U}\Sigma, \boldsymbol{E}\boldsymbol{V}\rangle$$
$$= \sup_{\boldsymbol{v}\in\mathbb{R}^{2nr}} \boldsymbol{v}^T\boldsymbol{e} \le C\sigma\sqrt{nr}, \tag{2}$$

with probability $1 - \exp(-C_2 nr)$, where $C, C_2$ are two positive constants, the vectorization $\boldsymbol{e} = \mathrm{Vec}(\boldsymbol{E}\boldsymbol{V}) \in \mathbb{R}^{2nr}$ has independent mean-zero sub-Gaussian entries with bounded variance $\sigma^2$ due to the orthogonality of $\boldsymbol{V}$, and the last inequality follows from Rigollet and Hütter (2015, Theorem 1.19).

Combining inequalities (1) and (2), we obtain the desired conclusion.

□

5. Page 58, Proof of Lemma 12. Add sentences for better explanations.

   "Note that we have $\ell^{(t)} \le \frac{L^{(t)}}{\Delta_{\min}^2} \le \frac{\bar{C}}{\tilde{C}} r \log^{-1}(p)$ by Condition 1 and Lemma 2. Then, with the locally linear stability assumption, the $\boldsymbol{\theta}$ is $\ell^{(t)}$-locally linearly stable; i.e., ..."