Changes in the final manuscript

This file summarizes the technical changes in the final manuscript compared with the submission after the second-round revision. We list the changes by types and code the modifications by red color. See the color-coded manuscript "ieee_final_colorcode.pdf" for all detailed modifications. The minor wording changes including the corrections of typos and inconsistent notations are not highlighted.

Explanation and statement modifications.

1. We have revised the explanations for the boundedness constraints c_3, c_4 in Page 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 purely zero slide to avoid trivial non-identifiability of model (1); see Example 2 below. The upper bound c_4 is a technical constraint to avoid the slides with diverging norm as dimension grows. ..."

2. We have corrected the conclusions related to $\Delta^2_{\boldsymbol{X}}$ in Theorem 2:

   "... Further, we define the parameter space $\mathcal{P}'(\gamma') := \mathcal{P} \cap \{\Delta^2_{\boldsymbol{X}} = p^{\gamma'}\}$, where $\Delta^2_{\boldsymbol{X}}$ is the mean tensor minimal gap in (8). When $\gamma' < -(K-1)$, we have

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

Related discussion in Page 6 and the Proof of Theorem 2 in Appendix D are also revised correspondly.

Minor technical condition modifications.

1. We have added the ranges of the number of communities $r \geq 2$ (or $r \geq 1$), order $K \geq 2$, and dimension $p \to \infty$ in the statements of Theorems 1, 2, 3, 4, 5, Lemma 1, Corollary 1, Proposition 1 in the main text, and the Proofs of Theorems 1, 2, 3, 4, 5, Lemmas 8, 11, 12, 13 in the Appendices. We take the modification in Theorem 4 as a typical example here:

   "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 $p \to \infty$. ..."

2. We have clarified the technical assumptions in Lemma 1 and Theorems 3, 5.

   In Lemma 1, we have added the lower bound of degree $\boldsymbol{\theta}$ and removed the Assumption 1:

   "Consider the dTBM model (1) under the parameter space $\mathcal{P}$ in (2) with $r \geq 2$. Suppose $\boldsymbol{\theta}$ is balanced satisfying (6) 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 ..."

   In Theorems 3 and 5, we have clarified the linear local stability condition:

"... Assume that the locally linear stability of degree holds in the neighborhood $\mathcal{N}(z, \varepsilon)$ for all $\varepsilon \leq E_0$ and some $E_0 \gtrsim \log^{-1} p$. ..."

Proof modifications.

1. We have added discussions of extreme cases with $r = 1$ in the Proofs of Theorems 1, 4,and 5.

   In the Proof of Theorem 1, we have added following statements:

   > "... if the model (26) violates Assumption 2. Note that $\Delta_{\min}^2 = 1$ when there exists $k \in [K]$ 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 $\boldsymbol{M}_k$'s are 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; ..."

   In the Proofs of Theorems 4 and 5, we have added following statement:

   > "For the case $r = 1$, $\ell(z_k^{(t)}, z) = 0$ trivially for all $t \geq 0, k \in [k]$. Hence, we focus on the proof of the first mode clustering $z_1^{(t+1)}$ with $r \geq 2$; ..."

2. We have revised the Proofs of Lemmas 1 and 9 for better presentations.

   In Proof of Lemma 1, we showed the equivalence between mean tensor and core tensor minimal gaps via the cosine terms:

   > ..." 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 $\boldsymbol{\theta}$. ..."

   In Proof of Lemma 9, we used a more classical textbook result to upper bound maximal inner product between low-rank tensor and random noise tensor:

   > "... 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
   >
   > $$\begin{aligned} \sup_{\boldsymbol{T} \in \mathbb{R}^{m \times n}, \operatorname{rank}(\boldsymbol{T}) \leq 2r, \|\boldsymbol{T}\|_F = 1} \langle \boldsymbol{T}, \boldsymbol{Y} - \boldsymbol{X} \rangle &= \sup_{\boldsymbol{T} \in \mathbb{R}^{m \times n}, \operatorname{rank}(\boldsymbol{T}) \leq 2r, \|\boldsymbol{T}\|_F = 1} \langle \boldsymbol{U} \Sigma, \boldsymbol{E} \boldsymbol{V} \rangle \\ &= \sup_{\boldsymbol{v} \in \mathbb{R}^{2nr}} \boldsymbol{v}^T \boldsymbol{e} \leq C \sigma \sqrt{nr}, \end{aligned}$$
   >
   > with probability $1 - \exp(-C_2 nr)$, where C, C_2 are two positive constants, the vectorization $\boldsymbol{e} = \operatorname{Vec}(\boldsymbol{E}\boldsymbol{V}) \in \mathbb{R}^{2nr}$ has independent mean-zero sub-Gaussian entries with bounded variance σ^2 due to the orthogonality of $\boldsymbol{V}$, and the last inequality follows from Rigollet and Hütter (2015, Theorem 1.19). ..."