Supplement of "Sharp Impossibility Results for Hyper-graph Global Testing"

In this supplement file, we first present the impossibility results for RMM-DCMM, which is omitted from the main text due to space limit. Then, we prove all the theorems and lemmas. Note that in this paper, C is a generic constant that may vary from occurrence to occurrence.

A The region of impossibility for RMM-DCMM

For RMM-DCMM models, we allow mixed-memberships. The discussion is quite similar, and the impossibility result in Section 2.2 continues to hold under a mild condition.

Similarly, consider a model pair, where we have a null DCMM model and an RMM-DCMM model with K communities as the alternative. Denote the Bernoulli probability tensors by $\mathcal{Q}$ and $\mathcal{Q}^*$, respectively. Similarly, for $1 \leq i_1, i_2, i_3 \leq n$, we assume

$$\mathcal{Q}_{i_1 i_2 i_3} = \theta_{i_1}\theta_{i_2}\theta_{i_3}, \tag{A.1}$$

$$\mathcal{Q}^*_{i_1 i_2 i_3} = \theta^*_{i_1}\theta^*_{i_2}\theta^*_{i_3} \cdot \pi'_{i_1}(\mathcal{P}\pi_{i_3})\pi_{i_2}, \tag{A.2}$$

where the community structure tensor $\mathcal{P}$ is as in (1.2), and π_i and $h = \mathbb{E}_F[\pi_i]$ are as in (1.5). Similarly, for any matrix $D = \mathrm{diag}(d_1, d_2, \ldots, d_K)$ with $d_k > 0$, $1 \leq k \leq K$, let $\mathcal{P}^D$ be the tensor with the same size of $\mathcal{P}$ satisfying $\mathcal{P}^D_{k_1k_2k_3} = d_{k_1}d_{k_2}d_{k_3}\mathcal{P}_{k_1k_2k_3}$. Also, let $h^D = \mathbb{E}[D^{-1}\pi_i/\|D^{-1}\pi_i\|_1]$ and $\tilde{a}^D = (\mathcal{P}^D h^D)h^D$. We assume that there is a matrix D such that

$$\tilde{a}^D = \mathbf{1}_K, \qquad \min_{1\leq k\leq K}\{h^D_k\} \geq C. \tag{A.3}$$

Recall that in Lemma 2.1, we have shown that such a matrix D always exists for DCBM. To see the point, note that if we do not allow mixed-memberships, then each realized π_i is degenerate (i.e., only one entry is 1, all other entries are 0). In this case, $h^D = \mathbb{E}_F[\pi_i] = h$, and $\tilde{a}^D = a^D$. Therefore, (A.3) always holds, by Lemma 2.1. For this reason, (A.3) is only a mild condition.

Suppose now (A.3) holds for a matrix $D = D_0$. Let $\mathcal{P}^*$ and $\tilde{a}^*$ be $\mathcal{P}^D$ and $\tilde{a}^D$ evaluated at $D = D_0$, respectively. By definitions, $\tilde{a}^* = \mathbf{1}_K$. For $1 \leq i \leq n$, let

$$\theta^*_i = \theta_i/\|D_0^{-1}\pi_i\|_1, \quad \pi^*_i = D_0^{-1}\pi_i/\|D_0^{-1}\pi_i\|_1. \tag{A.4}$$

Combining them with (A.2), for all $1 \leq i_1, i_2, i_3 \leq n$, we have $\mathcal{Q}^*_{i_1i_2i_3} = \theta^*_{i_1}\theta^*_{i_2}\theta^*_{i_3} \cdot \pi'_{i_1}(\mathcal{P}\pi_{i_3})\pi_{i_2} = \theta_{i_1}\theta_{i_2}\theta_{i_3}{\pi^*_{i_1}}'(\mathcal{P}^*\pi^*_{i_3})\pi^*_{i_2}$. By similar calculations, for $1 \leq i_1 \leq n$, the leading term of the expected degree of node i_1 under the alternative is $\theta_{i_1}\|\theta\|_1^2(\pi^*_{i_1})'\tilde{a}^* = \theta_{i_1}\|\theta\|_1^2$, where the right hand side is the leading term of the expected degree of node i_1 under the null. Therefore, we have the desired degree matching as before. The following theorem is proved in Section D.

Theorem A.1 (Impossibility for DCMM)**.** *Fix $K > 1$. Given $(\theta, \mathcal{P}, h, F)$, consider a pair of models, an alternative with K communities and a null, as in (A.2) and (A.1) respectively, where (A.3) holds and θ^* is given by (A.4). Suppose (2.6) hold and $\|\theta\|_1\|\theta\|^2\mu_2^2 = o(1)$. As $n \to \infty$,* *the χ^2-divergence between the pair tends to 0. Therefore, the two models are asymptotically indistinguishable in the sense that the sum of Type I and Type II errors of any test is no smaller than $1 + o(1)$.*

Similarly, in the parameter space $(\theta, \mathcal{P}, h, F)$ for DCMM, we call the region prescribed by $\|\theta\|_1\|\theta\|^2\mu_2^2 \to 0$ the *Region of Impossibility.* For any model in this region, we can pair it with a null so they are asymptotically inseparable.

We next generalize the result to non-uniform DCMM. Consider a DCMM null model with probability tensors $\mathcal{Q}[M] = \{\mathcal{Q}^{(2)}, \dots, \mathcal{Q}^{(M)}\}$ and an RMM-DCMM model with probability tensors $\mathcal{Q}^*[M] = \{\mathcal{Q}^{*(2)}, \dots, \mathcal{Q}^{*(M)}\}$, where for every $2 \le m \le M$ and $1 \le i_1, i_2, \dots, i_m \le n$,

$$\mathcal{Q}^{(m)}_{i_1,i_2,\dots,i_m} = \theta^{(m)}_{i_1}\theta^{(m)}_{i_2}\cdots\theta^{(m)}_{i_m}, \tag{A.5}$$

$$\mathcal{Q}^{*(m)}_{i_1,i_2,\dots,i_m} = \theta^{*(m)}_{i_1}\cdots\theta^{*(m)}_{i_m} \times [\mathcal{P}^{(m)}; \pi_{i_1}, \dots, \pi_{i_m}], \qquad \pi_i \overset{iid}{\sim} F. \tag{A.6}$$

For any matrix $D^{(m)} = \mathrm{diag}(d_1^{(m)}, d_2^{(m)}, \dots, d_K^{(m)})$ with $d_k^{(m)} > 0$, $1 \le k \le K$, let $\widetilde{\mathcal{P}}^{(m)}$ be the tensor with the same size of $\mathcal{P}^{(m)}$ satisfying $\widetilde{\mathcal{P}}^{(m)}_{k_1k_2\cdots k_m} = d^{(m)}_{k_1}d^{(m)}_{k_2}\cdots d^{(m)}_{k_m}\mathcal{P}^{(m)}_{k_1k_2\cdots k_m}$. Also, let $\widetilde{h}^{(m)} = \mathbb{E}[{D^{(m)}}^{-1}\pi_i/\|{D^{(m)}}^{-1}\pi_i\|_1]$ and $\widetilde{a}^{(m)} = \sum_{1\le i_2,\dots,i_m\le K} d^{(m)}_{i_1}\cdot\mathcal{P}^{(m)}_{i_1\cdots i_m}\cdot(d^{(m)}_{i_2}\widetilde{h}^{(m)}_{i_2})\cdots(d^{(m)}_{i_m}\widetilde{h}^{(m)}_{i_m})$, for every $1 \le i_1 \le K$. We assume that there are matrices $D^{(2)}, \dots, D^{(m)}$ such that for $m = 2, \dots, M$

$$\widetilde{a}^{(m)} = \mathbf{1}_K, \qquad \min_{1\le k\le K}\{\widetilde{h}^{(m)}_k\} \ge C. \tag{A.7}$$

Note that (A.7) always holds for non-uniform DCBM, by Lemma C.1 in Section C below. For this reason, (A.7) is only a mild condition.

Suppose now (A.7) holds for a matrix $D^{(m)} = D_0^{(m)}$, for $m = 2, \dots, M$. Let $\mathcal{P}^{*(m)}$ and $\widetilde{a}^{*(m)}$ be $\widetilde{\mathcal{P}}^{(m)}$ and $\widetilde{a}^{(m)}$ evaluated at $D^{(m)} = D_0^{(m)}$, respectively. By definitions, $\widetilde{a}^{*(m)} = \mathbf{1}_K$. For $1 \le i \le n$, $2 \le m \le M$, let

$$\theta^{*(m)}_i = \theta^{(m)}_i/\|{D_0^{(m)}}^{-1}\pi_i\|_1, \quad \pi^{*(m)}_i = {D_0^{(m)}}^{-1}\pi_i/\|{D_0^{(m)}}^{-1}\pi_i\|_1. \tag{A.8}$$

This is analogous to the degree matching strategy in (A.4), and it is conducted for each m separately. Let $\mu_2^{(m)}$ be the second singular value of $P^{(m)}$. For short, let $\ell_m = \|\theta^{(m)}\|_1^{m-2}\|\theta^{(m)}\|^2(\mu_2^{(m)})^2$. The following Theorem is for non-uniform DCMM.

Theorem A.2 (Impossibility for non-uniform RMM-DCMM)**.** *Fix $K > 1$ and $M \ge 2$. For any given (h, F) and $\{(\theta^{(m)}, \mathcal{P}^{(m)})\}_{2\le m\le M}$, consider a pair of models, a null as in (A.6) and an alternative with K communities as in (A.5), where (A.7) hold and $\{\theta^{*(m)}_i\}_{1\le i\le n, 2\le m\le M}$ are as in (A.8). Suppose $\|P^{(m)}\| \le C$ and $\max_{1\le i\le n}\theta^{(m)}_i \le C$. If $\max_{2\le m\le M}\{\ell_m\} = o(1)$, then as $n \to \infty$, the χ^2-divergence between the pair tends to 0.*

B Proof of Theorem 2.2

Fix an arbitrary $(\theta, \mathcal{P}, h, F)$ that satisfies the requirement of Theorem A.1. We consider a pair of models: a null model where $\mathcal{Q}_{i_1i_2i_3} = \theta_{i_1}\theta_{i_2}\theta_{i_3}$ and a K-community uniform RMM-DCMM model as in Theorem A.1. Let $\mathcal{P}_0^{(n)}$ and $\mathcal{P}_1^{(n)}$ denote the probability measures associated with these two models, respectively. We further modify $\mathcal{P}_1^{(n)}$ as follows. In this RMM-DCMM, the membership matrix Π is randomly generated. Let Π_0 be a non-random membership matrix such that $(\theta, \Pi_0, \mathcal{P}) \in \mathcal{M}_n(K, c_0, \alpha_n, \beta_n)$. We define

$$\widetilde{\Pi} = \begin{cases} \Pi, & \text{if } (\theta, \Pi, \mathcal{P}) \in \mathcal{M}_n(K, c_0, \alpha_n, \beta_n), \\ \Pi_0, & \text{otherwise.} \end{cases}, \qquad \text{where } \pi_i \overset{iid}{\sim} F. \tag{B.9}$$

We construct a similar RMM-DCMM by replacing Π with $\widetilde{\Pi}$ and denote $\widetilde{P}_1^{(n)}$ the probability measure associated with this new RMM-DCMM.

Consider a pair of hypotheses, where $\mathcal{A}$ is generated from $\mathcal{P}_0^{(n)}$ under the null hypothesis and it is generated from $\widetilde{P}_1^{(n)}$ under the alternative hypothesis. Given any test ψ, its sum of type I and type II errors is equal to

$$\begin{aligned} & \mathcal{P}_0^{(n)}(\psi = 1) + \widetilde{\mathcal{P}}_1^{(n)}(\psi = 0) \\ = \; & \mathbb{P}_0(\psi = 1) + \mathbb{E}_{\widetilde{\Pi}}\big[\mathbb{P}_1\big(\psi = 0 | \widetilde{\Pi}\big)\big] \\ \le \; & \sup_{\theta \in \mathcal{M}_n^*(\beta_n)} \mathbb{P}(\psi = 1) + \sup_{(\theta, \Pi, \mathcal{P}) \in \mathcal{M}_n(K, c_0, \alpha_n, \beta_n)} \mathbb{P}(\psi = 0). \end{aligned}$$

In the last inequality, we have used the fact that $(\theta, \widetilde{\Pi}, \mathcal{P}) \in \mathcal{M}_n(K, c_0, \alpha_n, \beta_n)$ for any realization of $\widetilde{\Pi}$ (this is guaranteed by the construction in (B.9)). At the same time, by Neyman-Pearson lemma,

$$\mathcal{P}_0^{(n)}(\psi = 1) + \widetilde{\mathcal{P}}_1^{(n)}(\psi = 0) \;\; \ge \;\; 1 - \|\mathcal{P}_0^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1,$$

where $\|\mathcal{P}_0^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1$ is the L_1-distance between two probability measures. Therefore, to show the claim, it suffices to show that

$$\|\mathcal{P}_0^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1 = o(1). \tag{B.10}$$

We now show (B.10). Recall that in Theorem A.1 we have seen that the χ^2-divergence between $\mathcal{P}_0^{(n)}$ and $\mathcal{P}_1^{(n)}$ tends to 0. Using the triangle inequality and the connection between L_1-distance and χ^2-divergence (e.g., equation (2.27) of [5]), we have

$$\begin{aligned} \|\mathcal{P}_0^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1 &\le \|\mathcal{P}_0^{(n)} - \mathcal{P}_1^{(n)}\|_1 + \|\mathcal{P}_1^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1 \\ &\le \sqrt{\chi^2(\mathcal{P}_0^{(n)}, \mathcal{P}_1^{(n)})} + \|\mathcal{P}_1^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1 \\ &\le o(1) + \|\mathcal{P}_1^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1. \end{aligned} \tag{B.11}$$

It suffices to show that $\|\mathcal{P}_1^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1 \to 0$. By (B.9), $\widetilde{\mathcal{P}}_1^{(n)}$ is obtained from $\mathcal{P}_1^{(n)}$ by modifying those realizations of Π where $(\theta, \Pi, \mathcal{P}) \notin \mathcal{M}_n(K, c_0, \alpha_n, \beta_n)$. By some elementary calculations, we have

$$\|\mathcal{P}_1^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1 \le 2\, \mathbb{P}\big((\theta, \Pi, \mathcal{P}) \notin \mathcal{M}_n(K, c_0, \alpha_n, \beta_n)\big),$$

where $\mathbb{P}$ is with respect to the randomness of Π. In the definition of $\mathcal{M}_n(K, c_0, \alpha_n, \beta_n)$, the only requirement involving Π is that $\max_{1 \le k \le K}\{g_k\} \le c_0^{-1} \min_{1 \le k \le K}\{g_k\}$. The following lemma is proved below:

Lemma B.1. *Fix a constant $c_0 \ge 1$. As $n \to \infty$, suppose $\|P\| \le c_0$, $\theta_{\max} \le c_0$, and $\|\theta\|_1 \to \infty$. Write $h = \mathbb{E}[D^{-1}\pi_i / \|D^{-1}\pi_i\|_1]$. If $\min_{1 \le k \le K}\{h_k\} \ge c_1$, for an appropriate constant $c_1 > 0$, then as $n \to \infty$, with probability $1 - o(1)$, the following condition is satisfied,*

$$\frac{\max_{1 \le k \le K}\{g_k\}}{\min_{1 \le k \le K}\{g_k\}} \le c_0^{-1}.$$

By Lemma B.1, the probability of $(\theta, \Pi, \mathcal{P}) \notin \mathcal{M}_n(K, c_0, \alpha_n, \beta_n)$ tends to 0 as $n \to \infty$. It follows that $\|\mathcal{P}_1^{(n)} - \widetilde{\mathcal{P}}_1^{(n)}\|_1 \to 0$. We plug it into (B.11) to get (B.10). This completes the proof.

B.1 Proof of Lemma B.1

Recall that $g_k = (1/\|\theta\|_1) \sum_{i=1}^n \theta_i \pi_i(k)$, for $1 \le k \le K$. Since $\max_k\{\sum_{i=1}^n \theta_i \pi_i(k)\} \le \|\theta\|_1$, it suffices to show that

$$\min_k \{\sum_{i=1}^n \theta_i \pi_i(k)\} \ge c_0 \, \|\theta\|_1. \tag{B.12}$$

Let c_1 be a constant such that $c_1 > c_0$. Our assumptions say that $\min_{1\le k\le K}\{h_k\} \ge c_1$, where $h = \mathbb{E}[D^{-1}\pi_i/\|D^{-1}\pi_i\|_1]$. Let $h^* = \mathbb{E}[\pi_i]$. We first show that $\min_{1\le k\le K}\{h_k\} \ge c_1$ implies $\min_{1\le k\le K}\{h_k^*\} \ge c_1 \cdot [1+o(1)]$. By Lemma E.5 in section E, we have

$$\max_{1\le i\le K}\{|d_i - 1|\} \le C\mu_2 \qquad \text{with } \mu_2 = o(1),$$

and so $d_i = 1 + o(1), 1 \le i \le K$. By definitions, it follows that

$$h_k \le \mathbb{E}[(\min_{1\le k\le K}\{d_k\})^{-1}\pi_i(k)/(\max_{1\le k\le K}\{d_k\})^{-1})] \le h_k^* \cdot [1+o(1)].$$

Combining this with $\min_{1\le k\le K}\{h_k\} \ge c_1$, we have $\min_{1\le k\le K}\{h_k^*\} \ge c_1 \cdot [1+o(1)]$.

Now we are going to show (B.12). Note that $X = \sum_{i=1}^n \theta_i(\pi_i(k) - h_k^*)$ is a sum of independent mean-zero random variables, where $\theta_i(\pi_i(k) - h_k^*) \le C\theta_{\max}$ and $\sum_{i=1}^n \mathrm{Var}(\theta_i(\pi_i(k) - h_k^*)) \le C\|\theta\|^2$. By Bernstein's inequality,

$$\mathbb{P}(|X| > t) \le \exp\Big(-\frac{t^2}{C\|\theta\|^2 + C\theta_{\max}t}\Big), \qquad \text{for any } t > 0.$$

Taking $t = C\|\theta\|\sqrt{\log(\|\theta\|_1)} + C\theta_{\max}\log(\|\theta\|_1)$, it follows that, with probability at least $1 - \|\theta\|_1^{-1}$,

$$|\sum_i \theta_i(\pi_i(k) - h_k^*\|\theta\|_1| = |X| \le C\|\theta\|\sqrt{\log(\|\theta\|_1)} + C\theta_{\max}\log(\|\theta\|_1),$$

where by $\|\theta\|^2 \le \|\theta\|_1$, the RHS is $o(\|\theta\|_1)$. Combining this with $\min_k\{h_k^*\} \ge c_1 \cdot [1+o(1)]$,

$$\sum_i \theta_i\pi_i(k) = h_k^*\|\theta\|_1 \cdot [1+o(1)] \ge c_1\|\theta\|_1 \cdot [1+o(1)],$$

where c_1 is a constant strictly larger than c_0. This proves (B.12). The claim follows.

C Proof of Lemma 2.1

We prove a version of this lemma for $m-$uniform hypergraph below where the desired result is by letting $m = 3$.

Lemma C.1 (Lemma 2.1 for $m-$uniform hypergraph)**.** *Fix $K > 1$ and $m > 1$. Let $\mathcal{P}$ be a nonnegative m-uniform tensor of dimension K and h be a vector in $\mathbb{R}^K$, where we assume $\mathcal{P}_{i\ldots i} = 1$, for $i = 1, \ldots, K$ and $\min\{h_1, h_2, \ldots, h_K\} \ge C$. There exists an unique diagonal matrix $D = \mathrm{diag}(d_1, d_2, \ldots, d_K)$ such that*

$$\sum_{i_2,\ldots,i_m=1}^K d_{i_1}\mathcal{P}_{i_1\cdots i_m} \cdot (d_{i_2}h_{i_2})\cdots(d_{i_m}h_{i_m}) = 1, \qquad \textit{for all } i_1 = 1, \ldots, K. \tag{C.13}$$

To begin with, we transform the problem (C.13) into an equivalent form (C.14).

Multiplying h_{i_1} on both sides of (C.13) and let $\widetilde{d}_i = d_ih_i$ for $i = 1, \ldots, K$. It is equivalent to find an unique diagonal matrix $\widetilde{D} = diag(\widetilde{d}_1, \ldots, \widetilde{d}_K)$ with strictly positive entries such that

$$\sum_{i_2,\ldots,i_m=1}^K \widetilde{d}_{i_1}\mathcal{P}_{i_1\cdots i_m}\widetilde{d}_{i_2}\cdots\widetilde{d}_{i_m} = h_{i_1}, \qquad \text{for all } i_1 = 1, \ldots, K. \tag{C.14}$$

Now, by the Theorem 6 in [1], for a nonnegative order-m tensor $\mathcal{P}$ of dimension K (not necessarily symmetric) such that $\mathcal{P}_{i\ldots i} > 0$, $i = 1, \ldots, K$, and K positive numbers $h_1, \ldots, h_K$, there exist positive numbers $x_1, \ldots, x_K$ such that

$$\sum_{i_2,\ldots,i_m=1}^K x_{i_1}\mathcal{P}_{i_1\cdots i_m}x_{i_2}\cdots x_{i_m} = h_{i_1}, \qquad \text{for all } i_1 = 1, \ldots, K. \tag{C.15}$$

which gives the existence of such $\widetilde{D}$ satisfying (C.14).

The uniqueness of such $\widetilde{D}$ is given by the Theorem 1.1 in [2] which states that there is an unique tensor $\mathcal{A}$ that is defined by $\mathcal{A}_{i_1\cdots i_m} = \widetilde{d}_{i_1}\mathcal{P}_{i_1\cdots i_m}\widetilde{d}_{i_2}\cdots\widetilde{d}_{i_m}$ for $i_1,\ldots,i_m = 1,\ldots,K$ and satisfies

$$\sum_{i_2,\ldots,i_m=1}^{K} \mathcal{A}_{i_1\cdots i_m} = h_{i_1}, \qquad \text{for all } i_1 = 1,\ldots,K. \tag{C.16}$$

Therefore, $\widetilde{D}$ is unique since $\mathcal{A}$ is unique and one-to-one correpondence with $\widetilde{D}$. This completes the proof.

D Proof of Theorem 2.1, Theorem 2.3 and Theorem A.1-A.2

Theorem 2.1 and Theorem 2.3 are the special cases of Theorem 3.1, which do not need separate proofs. Furthermore, in the proof of Theorem 3.1 below, we actually consider the more general setting of non-uniform DCMM where θ_i^* is constructed as $\theta_i^* = \theta_i/\|D^{-1}\pi_i\|_1$ (note that when π_i is degenerate, this reduces to the construction of $\theta_i^* = \theta_i d_k$ for DCBM). Therefore, the proof of Theorem 3.1 (for non-uniform DCMM) already includes the proof of Theorem A.1 (for 3-uniform DCMM) and Theorem A.2 (for non-uniform DCMM). It remains to prove Theorem 3.1, which is contained in Section E.

E Proof of Theorem 3.1

We first state the preliminary lemmas, Lemmas E.1-E.5, needed for the proof of Theorem 3.1. Next, we prove this theorem. Finally, we prove all the preliminary lemmas.

E.1 Preliminary lemmas

The following lemmas are used in the main proof and proved after the main proof.

Lemma E.1. *Let $\mathcal{P}$ be a $m-$way symmetric K dimensional tensor, $\mathcal{P}_0$ be the tensor with the same size as $\mathcal{P}$ where all entries are 1, and introduce a tensor $\mathcal{M}$ by $\mathcal{M} = \mathcal{P} - \mathcal{P}_0$. Let h,π_i be weight vectors in $\mathbb{R}^K$ and $y_i = \pi_i - h$, for $1 \le i \le n$. Then*

$$[\mathcal{P};\pi_1,\ldots,\pi_m] = 1 + x^{(m)} + z^{(m)}, \qquad \textit{holds for any } m > 1,$$

where

$$x^{(m)} = [\mathcal{M}; h,\ldots,h] + \sum_{s=1}^{m}[\mathcal{M}; \underbrace{h,\ldots,h}_{s-1}, y_s, \underbrace{h,\ldots,h}_{m-s}],$$

$$z^{(m)} = \sum_{s_1=1}^{m-1}\sum_{s_2=s_1+1}^{m}[\mathcal{M}; \underbrace{h,\ldots,h}_{s_1-1}, y_{s_1}, \underbrace{h,\ldots,h}_{s_2-s_1-1}, y_{s_2}, \underbrace{\pi_{s_2+1}\ldots,\pi_m}_{m-s_2}].$$

Lemma E.2. *With the same notations as in Section E.2, let $\{w_i^{(j)} : 1 \le i \le n, 1 \le j \le m\}$ be a set of weight vectors in $\mathbb{R}^K$ and $\{\widetilde{w}_i^{(j)}\}$ be an independent copy of $\{w_i^{(j)}\}$. Assume that for distinct $i_1,\ldots,i_m$, vectors $y_{i_1}, y_{i_2}, w_{i_3}^{(3)},\ldots,w_{i_m}^{(m)}$ are mutually independent and that $\|\mathcal{M}_{::k_3\cdots k_m}\| \le C\mu$, for $1 \le k_3,\ldots,k_m \le K$. Denote*

$$S = \sum_{i_1,\ldots,i_m(dist)} \frac{(\theta_{i_1}\cdots\theta_{i_m})^t}{a_t}[\mathcal{M}; y_{i_1}, y_{i_2}, w_{i_3}^{(3)},\ldots,w_{i_m}^{(m)}][\mathcal{M}; \widetilde{y}_{i_1}, \widetilde{y}_{i_2}, \widetilde{w}_{i_3}^{(3)},\ldots,\widetilde{w}_{i_m}^{(m)}].$$

Then, for any constant c independent of n,

$$\mathbb{E}\Big[\exp(cS)\Big] \le \mathbb{E}\Big[\exp\Big(C\mu^2\|\theta\|_t^{t(m-2)}|T|/a_t\Big)\Big]\cdot\exp(C\mu^2\|\theta\|_t^{t(m-2)}\|\theta\|_{2t}^{2t}/a_t),$$

where T is a random variable satisfying $\mathbb{P}(|T|>x)\le 2\exp(-x/(2K^2\|\theta\|_{2t}^{2t}))$, for $x>0$.

Lemma E.3. *With the same setting in Lemma E.2, denote*

$$S=\sum_{i_1,\ldots,i_m(dist)}\frac{(\theta_{i_1}\cdots\theta_{i_m})^t}{a_t}[\mathcal{M};y_{i_1},y_{i_2},w_{i_3}^{(3)},\ldots,w_{i_m}^{(m)}][\mathcal{M};\widetilde{y}_{i_1},h,\widetilde{y}_{i_3},\widetilde{w}_{i_4}^{(4)},\ldots,\widetilde{w}_{i_m}^{(m)}].$$

Then, for any constant c independent of n,

$$\mathbb{E}\Big[\exp(cS)\Big] \le \mathbb{E}\Big[\exp\Big(C\mu^2\|\theta\|_t^{t(m-2)}|T|/a_t\Big)\Big]\cdot\exp(C\mu^2\|\theta\|_t^{t(m-2)}\|\theta\|_{2t}^{2t}/a_t),$$

where T is a random variable satisfying $\mathbb{P}(|T|>x)\le 4\exp(-x/(2K^2\|\theta\|_{2t}^{2t}))$, for $x>0$.

Lemma E.4. *With the same setting in Lemma E.2, denote*

$$S=\sum_{i_1,\ldots,i_m(dist)}\frac{(\theta_{i_1}\cdots\theta_{i_m})^t}{a_t}[\mathcal{M};y_{i_1},y_{i_2},w_{i_3}^{(3)},\ldots,w_{i_m}^{(m)}][\mathcal{M};h,h,\widetilde{y}_{i_3},\widetilde{y}_{i_4},\widetilde{w}_{i_5}^{(5)},\ldots,\widetilde{w}_{i_m}^{(m)}].$$

Then, for any constant c independent of n,

$$\mathbb{E}\Big[\exp(cS)\Big] \le \mathbb{E}\Big[\exp\Big(C\mu^2\|\theta\|_t^{t(m-2)}|T|/a_t\Big)\Big]\cdot\exp(C\mu^2\|\theta\|_t^{t(m-2)}\|\theta\|_{2t}^{2t}/a_t),$$

where T is a random variable satisfying $\mathbb{P}(|T|>x)\le 4\exp(-x/(2K\|\theta\|_{2t}^{2t}))$, for $x>0$.

Lemma E.5. *Under the conditions of Theorem 3.1, for $m=2,\ldots,M$ we have*

$$\max_{1\le k_3,\ldots,k_m\le K}\|\mathcal{M}^{(m)}_{::k_3\cdots k_m}\|\le C|\mu_2^{(m)}|,\qquad \max_{1\le i\le K}|d_i^{(m)}-1|\le C|\mu_2^{(m)}|,$$

where $\mathcal{M}^{(m)}$ is a $m-$way symmetric tensor defined by $\mathcal{M}^{(m)}_{k_1\cdots k_m}=(\mathcal{P}^{(m)}_{k_1\cdots k_m}-1)d^{(m)}_{k_1}\cdots d^{(m)}_{k_m}$, $1\le k_1,\ldots,k_m\le K$.

E.2 Proof of Theorem 3.1

Let $P_0^{(n)}$ and $P_1^{(n)}$ denote the probability measures associated with the null and alternative hypotheses, respectively, and let $\chi^2(P_0^{(n)},P_1^{(n)})$ be the χ^2 divergence between the two probability measures. By definitions,

$$\chi^2(P_0^{(n)},P_1^{(n)})=\int_{\mathcal{A}}\left[\frac{dP_1^{(n)}(\mathcal{A})}{dP_0^{(n)}(\mathcal{A})}\right]^2 dP_0^{(n)}(\mathcal{A})-1.$$

To show the claim, it suffices to show that when $(\mu_2^{(m)})^2\|\theta^{(m)}\|_1^{m-2}\|\theta^{(m)}\|_2^2\to 0$, $m=1,\ldots,M$, we have

$$\int_{\mathcal{A}}\left[\frac{dP_1^{(n)}(\mathcal{A})}{dP_0^{(n)}(\mathcal{A})}\right]^2 dP_0^{(n)}(\mathcal{A})=1+o(1). \tag{E.17}$$

By definitions,

$$dP_0^{(n)}(\mathcal{A})=\prod_{m=2}^{M}\prod_{i_1<\cdots<i_m}dP_0^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}),$$

$$dP_1^{(n)}(\mathcal{A})=\mathbb{E}_\Pi\Big[\prod_{m=2}^{M}\prod_{i_1<\cdots<i_m}dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\Pi)\Big],$$

Let $\widetilde{\Pi}$ be an independent copy of Π. Putting the above two equations into (E.17) gives

$$\int_{\mathcal{A}} \left[\frac{dP_1^{(n)}(\mathcal{A})}{dP_0^{(n)}(\mathcal{A})}\right]^2 dP_0^{(n)}(\mathcal{A}) = \int_{\mathcal{A}} \frac{\mathbb{E}_{\Pi,\widetilde{\Pi}}\Big[\prod_{m=2}^M \prod_{i_1<\cdots<i_m} dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\Pi) dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\widetilde{\Pi})\Big]}{\prod_{m=2}^M \prod_{i_1<\cdots<i_m} dP_0^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m})}$$
$$= \int_{\mathcal{A}} \mathbb{E}_{\Pi,\widetilde{\Pi}}\Big[\prod_{m=2}^M \prod_{i_1<\cdots<i_m} \frac{dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\Pi) dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\widetilde{\Pi})}{dP_0^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m})}\Big].$$

Exchanging the order of integral and expectation in the last equation and by elementary probability,

$$\int_{\mathcal{A}} \left[\frac{dP_1^{(n)}(\mathcal{A})}{dP_0^{(n)}(\mathcal{A})}\right]^2 dP_0^{(n)}(\mathcal{A}) = \mathbb{E}_{\Pi,\widetilde{\Pi}}\Big[\int_{\mathcal{A}} \prod_{m=2}^M \prod_{i_1<\cdots<i_m} \frac{dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\Pi) dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\widetilde{\Pi})}{dP_0^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m})}\Big]$$
$$= \mathbb{E}_{\Pi,\widetilde{\Pi}}\Big[\prod_{m=2}^M \prod_{i_1<\cdots<i_m} \int_{\mathcal{A}^{(m)}_{i_1\cdots i_m}} \frac{dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\Pi) dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\widetilde{\Pi})}{dP_0^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m})}\Big].$$

Let $\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi})$ denote $\int_{\mathcal{A}^{(m)}_{i_1\cdots i_m}} dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\Pi) dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\widetilde{\Pi}) / dP_0^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}) - 1$. Hence

$$\int_{\mathcal{A}} \left[\frac{dP_1^{(n)}(\mathcal{A})}{dP_0^{(n)}(\mathcal{A})}\right]^2 dP_0^{(n)}(\mathcal{A}) = \mathbb{E}_{\Pi,\widetilde{\Pi}}\Big[\prod_{m=2}^M \prod_{i_1<\cdots<i_m} (\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi}) + 1)\Big]. \tag{E.18}$$

Note that by inequality $\prod_{i=1}^n (1+x_i) \le \exp(\sum_{i=1}^n x_i)$, for all x_i such that $1 + x_i \ge 0$, we have

$$\prod_{m=2}^M \prod_{i_1<\cdots<i_m} (\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi}) + 1) \le \exp\Bigg(\sum_{m=2}^M \sum_{i_1<\cdots<i_m} \chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi})\Bigg), \tag{E.19}$$

Further, by Jensen's inequality, $\exp(\sum_{i=2}^M x_i) \le \frac{1}{M-1}\sum_{i=2}^M \exp(x_i)$. It follows that

$$\exp\Bigg(\sum_{m=2}^M \sum_{i_1<\cdots<i_m} \chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi})\Bigg) \le \sum_{m=2}^M \frac{1}{M-1} \exp\Bigg((M-1) \sum_{i_1<\cdots<i_m} \chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi})\Bigg). \tag{E.20}$$

Combining (E.18)-(E.20) gives

$$\int_{\mathcal{A}} \left[\frac{dP_1^{(n)}(\mathcal{A})}{dP_0^{(n)}(\mathcal{A})}\right]^2 dP_0^{(n)}(\mathcal{A}) \le \sum_{m=2}^M \frac{1}{M-1} \mathbb{E}_{\Pi,\widetilde{\Pi}}\Bigg[\exp\Bigg((M-1) \sum_{i_1<\cdots<i_m} \chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi})\Bigg)\Bigg].$$

Therefore, to show (E.17), it is sufficient to show that when the conditions hold, for each $m = 2,\dots M$ we have

$$\mathbb{E}_{\Pi,\widetilde{\Pi}}\Bigg[\exp\Bigg((M-1) \sum_{i_1<\cdots<i_m} \chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi})\Bigg)\Bigg] = 1 + o_n(1). \tag{E.21}$$

Fix m, recall that

$$\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi}) = \int_{\mathcal{A}} \frac{dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\Pi) dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\widetilde{\Pi})}{dP_0^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m})} - 1. \tag{E.22}$$

By definitions,

$$dP_0^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}) = (\mathcal{Q}^{(m)}_{i_1\cdots i_m})^{\mathcal{A}^{(m)}_{i_1\cdots i_m}} (1 - \mathcal{Q}^{(m)}_{i_1\cdots i_m})^{1-\mathcal{A}^{(m)}_{i_1\cdots i_m}},$$
$$dP_1^{(n,m)}(\mathcal{A}^{(m)}_{i_1\cdots i_m}|\Pi) = (\mathcal{Q}^{*(m)}_{i_1\cdots i_m}(\Pi))^{\mathcal{A}^{(m)}_{i_1\cdots i_m}} (1 - \mathcal{Q}^{*(m)}_{i_1\cdots i_m}(\Pi))^{1-\mathcal{A}^{(m)}_{i_1\cdots i_m}}.$$

Putting the above two equations into (E.22) gives

$$
\begin{aligned}
\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi}) =& \frac{\mathcal{Q}^{*(m)}_{i_1\cdots i_m}(\Pi)\mathcal{Q}^{*(m)}_{i_1\cdots i_m}(\widetilde{\Pi})}{\mathcal{Q}^{(m)}_{i_1\cdots i_m}} + \frac{(1-\mathcal{Q}^{*(m)}_{i_1\cdots i_m}(\Pi))(1-\mathcal{Q}^{*(m)}_{i_1\cdots i_m}(\widetilde{\Pi}))}{1-\mathcal{Q}^{(m)}_{i_1\cdots i_m}} - 1 \\
=& \frac{\Big(\mathcal{Q}^{*(m)}_{i_1\cdots i_m}(\Pi) - \mathcal{Q}^{(m)}_{i_1\cdots i_m}\Big)\Big(\mathcal{Q}^{*(m)}_{i_1\cdots i_m}(\widetilde{\Pi}) - \mathcal{Q}^{(m)}_{i_1\cdots i_m}\Big)}{\mathcal{Q}^{(m)}_{i_1\cdots i_m}(1-\mathcal{Q}^{(m)}_{i_1\cdots i_m})}.
\end{aligned} \tag{E.23}
$$

Based on the expression of $\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi})$, it is seen that the LHS of (E.21) only relates to the variables in $m-$uniform tensor DCMM (e.g., $\mathcal{A}^{(m)}, \mathcal{Q}^{(m)}, \mathcal{P}^{(m)}, \theta^{(m)}$), for ease of notations, we remove the superscript (m) whenever it is clear from the context.

Next we continue to simplify $\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi})$. According to the constructions of our model,

$$
\mathcal{Q}_{i_1\cdots i_m} = \theta_{i_1}\cdots\theta_{i_m} \qquad \text{and} \qquad \mathcal{Q}^*_{i_1\cdots i_m} = \theta_{i_1}\cdots\theta_{i_m}[\mathcal{P}^*;\pi^*_{i_1},\ldots,\pi^*_{i_m}],
$$

where we recall that $\mathcal{P}^*$ is the m-uniform tensor defined by $\mathcal{P}^*_{k_1\cdots k_m} = d_{k_1}\cdots d_{k_m}\mathcal{P}_{k_1\cdots k_m}$, $1\le k_1,\ldots,k_m\le K$, $\pi^*_i = D^{-1}\pi_i/\|D^{-1}\pi_i\|_1$, $1\le i\le n$ and $D=\mathrm{diag}(d_1,d_2,\ldots,d_K)$ is the scaling matrix given by degree matching.

Let $\mathcal{P}_0$ the tensor with the same size as $\mathcal{P}^*$ and where all entries are 1, and introduce a tensor $\mathcal{M}$ by $\mathcal{M}=\mathcal{P}^*-\mathcal{P}_0$. Let $h=\mathbb{E}_F[\pi^*_i]$, and $y_i=\pi^*_i-h$, $1\le i\le n$. By Lemma E.1, we can write the Bernoulli probability tensor for the alternative $\mathcal{Q}^*$ by

$$
\mathcal{Q}^*_{i_1\cdots i_m} = \theta_{i_1}\cdots\theta_{i_m}(1+x_{i_1\cdots i_m}+z_{i_1\cdots i_m}), \qquad 1\le i_1,\ldots,i_m\le n, \tag{E.24}
$$

where

$$
\begin{aligned}
x_{i_1\cdots i_m} =& [\mathcal{M};h,\ldots,h] + \sum_{s=1}^{m}[\mathcal{M};\underbrace{h,\ldots,h}_{s-1},y_{i_s},\underbrace{h,\ldots,h}_{m-s}],\\
z_{i_1\cdots i_m} =& \sum_{s_1=1}^{m-1}\sum_{s_2=s_1+1}^{m}[\mathcal{M};\underbrace{h,\ldots,h}_{s_1-1},y_{i_{s_1}},\underbrace{h,\ldots,h}_{s_2-s_1-1},y_{i_{s_2}},\underbrace{\pi^*_{i_{s_2+1}}\ldots,\pi^*_{i_m}}_{m-s_2}].
\end{aligned}
$$

Let e_{i_1} be the i_1-th standard basis vector of the Euclidean space $\mathbb{R}^K$, $1\le i_1\le K$. Note that by definitions and symmetry,

$$
\begin{aligned}
[\mathcal{M};h,\ldots,h,e_{i_1},h,\ldots,h] =& \sum_{i_2,\ldots,i_m=1}^{K}(\mathcal{P}^*_{i_1\cdots i_m}-1)\cdot h_{i_2}\cdots h_{i_m}\\
=& \sum_{i_2,\ldots,i_m=1}^{K}\mathcal{P}^*_{i_1\cdots i_m}\cdot h_{i_2}\cdots h_{i_m}-1\\
(\text{By degree matching}) =& 0
\end{aligned}
$$

This indicates that any linear combination of elements in $\{[\mathcal{M};h,\ldots,h,e_i,h,\ldots,h]:,1\le i\le K\}$ equals to 0. It follows that the term $x_{i_1\cdots i_m}$ in the RHS of (E.24) equals to 0.

Write for short $z_{i_1\cdots i_m}(s_1,s_2)=[\mathcal{M};h,\ldots,h,y_{i_{s_1}},h,\ldots,h,y_{i_{s_2}},\pi^*_{i_{s_2+1}}\ldots,\pi^*_{i_m}]$, we get

$$
\mathcal{Q}^*_{i_1\cdots i_m} = \theta_{i_1}\cdots\theta_{i_m}\Big(1+\sum_{s_1=1}^{m-1}\sum_{s_2=s_1+1}^{m}z_{i_1\cdots i_m}(s_1,s_2)\Big), \tag{E.25}
$$

Let $\widetilde{z}_{i_1\cdots i_m}(s_1,s_2)$ be $z_{i_1\cdots i_m}(s_1,s_2)$ evaluated at $\widetilde{\Pi}$. Inserting (E.25) into (E.23) gives

$$
\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi}) = \frac{\theta_{i_1}\cdots\theta_{i_m}}{1-\theta_{i_1}\cdots\theta_{i_m}}\sum_{\substack{s_1=1,\\ \widetilde{s}_1=1}}^{m-1}\sum_{\substack{s_2=s_1+1\\ \widetilde{s}_2=\widetilde{s}_1+1}}^{m} z_{i_1\cdots i_m}(s_1,s_2)\widetilde{z}_{i_1\cdots i_m}(\widetilde{s}_1,\widetilde{s}_2).
$$

Note that $\frac{x}{1-x} = \sum_{i=1}^{\infty} x^i$ for any $x \in [0,1)$, we have $\frac{\theta_{i_1}\cdots\theta_{i_m}}{1-\theta_{i_1}\cdots\theta_{i_m}} = \sum_{i=1}^{\infty}(\theta_{i_1}\cdots\theta_{i_m})^t$ and so

$$\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi}) = \sum_{t=1}^{\infty}(\theta_{i_1}\cdots\theta_{i_m})^t \sum_{\substack{s_1=1,\\ \widetilde{s}_1=1}}^{m-1}\sum_{\substack{s_2=s_1+1\\ \widetilde{s}_2=\widetilde{s}_1+1}}^{m} z_{i_1\cdots i_m}(s_1,s_2)\widetilde{z}_{i_1\cdots i_m}(\widetilde{s}_1,\widetilde{s}_2).$$

Introduce

$$\begin{aligned} a_t =& \theta_{\max}^{m(t-1)}(1-\theta_{\max}^m), \\ S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2) =& (M-1)4^m \sum_{i_1<\cdots<i_m} \frac{(\theta_{i_1}\cdots\theta_{i_m})^t}{a^t} z_{i_1\cdots i_m}(s_1,s_2)\widetilde{z}_{i_1\cdots i_m}(\widetilde{s}_1,\widetilde{s}_2). \end{aligned} \tag{E.26}$$

Exchanging the order of summation, we then can write

$$(M-1)\sum_{i_1<\cdots<i_m}\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi}) = \sum_{t=1}^{\infty}\sum_{\substack{s_1=1,\\ \widetilde{s}_1=1}}^{m-1}\sum_{\substack{s_2=s_1+1\\ \widetilde{s}_2=\widetilde{s}_1+1}}^{m} \frac{a_t}{4^m} S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2).$$

Note that $\sum_{t=1}^{\infty}\sum_{s_1,\widetilde{s}_1=1}^{m-1}\sum_{s_2=s_1+1,\widetilde{s}_2=\widetilde{s}_1+1}^{m} a_t/4^m = 1$ and $\exp(\cdot)$ is convex, by Jensen's inequality

$$\exp\Big((M-1)\sum_{i_1<\cdots<i_m}\chi^2_{i_1\cdots i_m}(\Pi,\widetilde{\Pi})\Big) \le \sum_{t=1}^{\infty}\sum_{\substack{s_1=1,\\ \widetilde{s}_1=1}}^{m-1}\sum_{\substack{s_2=s_1+1\\ \widetilde{s}_2=\widetilde{s}_1+1}}^{m} \frac{a_t}{4^m}\exp\Big(S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2)\Big).$$

Therefore, to prove (E.21), it is sufficient to show that

$$\max_{t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2}\Big\{\mathbb{E}\Big[\exp\Big(S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2)\Big)\Big]\Big\} \le 1 + o_n(1). \tag{E.27}$$

Fix $t, s_1, s_2, \widetilde{s}_1, \widetilde{s}_2$, we are going to bound $\mathbb{E}\Big[\exp\Big(S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2)\Big)\Big]$. Recall that by construction, $s_1 < s_2$ and $\widetilde{s}_1 < \widetilde{s}_2$. By symmetry, without loss of generality, assume $s_2 \le \widetilde{s}_2$. Now, we can separate the situations into three cases. *Case 1:* $s_1 = \widetilde{s}_1, s_2 = \widetilde{s}_2$; *Case 2:* Only one of $\{s_1, s_2\}$ matches any one of $\{\widetilde{s}_1, \widetilde{s}_2\}$ (e.g., $\widetilde{s}_1 = s_1 < s_2 < \widetilde{s}_2$ or $s_1 < s_2 = \widetilde{s}_1 < \widetilde{s}_2$ or $s_1 \neq \widetilde{s}_1, s_2 = \widetilde{s}_2$); *Case 3:* None of $\{s_1, s_2\}$ matches one of $\{\widetilde{s}_1, \widetilde{s}_2\}$.

Remark: Case 2 only exists for $m \ge 3$ and *Case 3* only exists for $m \ge 4$. They require much tricky and delicate analysis to resolve extra random effects induced by Π. This indicates one of the differences on the calculations of the χ^2-divergence between hypergraph and network.

By symmetry of $\mathcal{M}$, we summerized the derivation of the bounds on $\mathbb{E}\Big[\exp\Big(S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2)\Big)\Big]$ for *Case 1,2,3* into Lemma E.2, E.3, E.4, respectively. Take *Case* 1 for example,

Case 1 ($s_1 = \widetilde{s}_1, s_2 = \widetilde{s}_2$): By definitions and symmetry of $\mathcal{M}$, we can rewrite

$$\begin{aligned} S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2) :=& 4^m(M-1)\sum_{i_1<\cdots<i_m}\frac{(\theta_{i_1}\cdots\theta_{i_m})^t}{a^t}[\mathcal{M}; h,\ldots,h,y_{i_{s_1}},h,\ldots,h,y_{i_{s_2}},\pi^*_{i_{s_2+1}}\ldots,\pi^*_{i_m}] \\ &\cdot[\mathcal{M}; h,\ldots,h,\widetilde{y}_{i_{s_1}},h,\ldots,h,\widetilde{y}_{i_{s_2}},\widetilde{\pi}^*_{i_{s_2+1}}\ldots,\widetilde{\pi}^*_{i_m}]. \\ =& \frac{4^m(M-1)}{m!}\sum_{i_1,\ldots,i_m(dist)}\frac{(\theta_{i_1}\cdots\theta_{i_m})^t}{a^t}[\mathcal{M}; y_{i_1},y_{i_2},h\ldots,h,\pi^*_{i_{s_2+1}}\ldots,\pi^*_{i_m}] \\ &\cdot[\mathcal{M}; \widetilde{y}_{i_1},\widetilde{y}_{i_2},h\ldots,h,\widetilde{\pi}^*_{i_{s_2+1}}\ldots,\widetilde{\pi}^*_{i_m}]. \end{aligned}$$

which is implied by the standard forms discussed in Lemma E.2. Similarly, *Case 2* is implied by Lemma E.3 and *Case 3* is implied by Lemma E.4.

Combining Lemmas E.2-E.4 with Lemma E.5, we have

$$\mathbb{E}\Big[\exp\Big(S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2)\Big)\Big] \le \mathbb{E}\Big[\exp\Big(C\frac{\mu_2^2\|\theta\|_t^{t(m-2)}}{a_t}|T|\Big)\Big]\cdot\exp\Big(C\frac{\mu_2^2\|\theta\|_t^{t(m-2)}\|\theta\|_{2t}^{2t}}{a_t}\Big), \quad \text{(E.28)}$$

where μ_2 is the second singular value of the matricization of the tensor $\mathcal{P}^{(m)}$ and T is a random variable satisfying $\mathbb{P}(|T|>x)\le 4\exp(-x/(2K^2\|\theta\|_{2t}^{2t}))$, for any $x>0$.

Now, we are ready to calculate a bound for $\mathbb{E}\Big[\exp\Big(S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2)\Big)\Big]$. By direct calculations,

$$\begin{aligned}\mathbb{E}\Big[\exp\Big(C\frac{\|\theta\|_t^{t(m-2)}}{a_t}\mu_2^2|T|\Big)\Big] &= \Big(1+\int_0^\infty e^x\cdot\mathbb{P}(C\frac{\|\theta\|_t^{t(m-2)}}{a_t}\mu_2^2|T|>x)dx\Big)\\ &\le \Big(1+\int_0^\infty e^x\cdot 4\exp(-\frac{a_t x}{2CK^2\mu_2^2\|\theta\|_t^{t(m-2)}\|\theta\|_{2t}^{2t}})dx\Big)\end{aligned} \quad \text{(E.29)}$$

By $\theta_{\max}\le c_0$, $\|\theta\|_t^t\le\|\theta\|_1\theta_{\max}^{t-1}$ and $\|\theta\|_{2t}^{2t}\le\|\theta\|^2\theta_{\max}^{t-2}$, we have

$$\frac{a_t}{\mu_2^2\|\theta\|_t^{t(m-2)}\|\theta\|_{2t}^{2t}} = \frac{\theta_{\max}^{m(t-1)}(1-\theta_{\max}^m)}{\mu_2^2\|\theta\|_t^{t(m-2)}\|\theta\|_{2t}^{2t}} \ge \frac{1-c_0^m}{\mu_2^2\|\theta\|_1^{m-2}\|\theta\|_2^2}$$

Combining this with (E.28)-(E.29), we get

$$\begin{aligned}\mathbb{E}\Big[\exp\Big(S(t,s_1,s_2,\widetilde{s}_1,\widetilde{s}_2)\Big)\Big] \le&\Big(1+\int_0^\infty e^x\cdot 4\exp(-\frac{(1-c_0^m)x}{2CK^2\mu_2^2\|\theta\|_1^{(m-2)}\|\theta\|_2^2})dx\Big)e^{\frac{C}{1-c_0^m}\mu_2^2\|\theta\|_1^{m-2}\|\theta\|^2}\\ =&e^{\frac{C}{1-c_0^m}\mu_2^2\|\theta\|_1^{m-2}\|\theta\|^2}\Big(1+4\Big(\frac{(1-c_0^m)}{2CK^2\mu_2^2\|\theta\|_1^{m-2}\|\theta\|_2^2}-1\Big)^{-1}\Big),\end{aligned}$$

where the RHS on the last inequality goes 1 as $\mu_2^2\|\theta\|_1^{m-2}\|\theta\|_2^2\to 0$. This proves (E.27) and finishes the proof.

E.3 Proof of Lemma E.1

Recall the definition of $[\mathcal{P};\pi_1,\ldots,\pi_m]$

$$[\mathcal{P};\pi_1,\ldots,\pi_m] := \sum_{k_1,\ldots,k_m=1}^K \mathcal{P}_{k_1\ldots k_m}\pi_1(k_1)\cdots\pi_m(k_m).$$

Note that $\mathcal{P}=\mathcal{M}+\mathcal{P}_0$ and $\sum_{k=1}^K\pi_i(k)=1$, for $1\le i\le n$. By direct calculations

$$\begin{aligned}[\mathcal{P};\pi_1,\ldots,\pi_m] =& \sum_{k_1,\ldots,k_m=1}^K \mathcal{M}_{k_1\ldots k_m}\pi_1(k_1)\cdots\pi_m(k_m)+\sum_{k_1,\ldots,k_m=1}^K 1\cdot\pi_1(k_1)\cdots\pi_m(k_m)\\ =&[\mathcal{M};\pi_1,\ldots,\pi_m]+1.\end{aligned}$$

Therefore, we are left to show for $m>1$

$$[\mathcal{M};\pi_1,\ldots,\pi_m] = x^{(m)}+z^{(m)}. \quad \text{(E.30)}$$

We prove it by induction. When $m=2$, $\mathcal{M}\in\mathbb{R}^{K\times K}$. By definitions and elementary algebra,

$$\begin{aligned}[\mathcal{M};\pi_1,\pi_2] =&\pi_1'\mathcal{M}\pi_2\\ =&h'\mathcal{M}h+y_1'\mathcal{M}h+h'\mathcal{M}y_2+y_1'\mathcal{M}y_2\\ =&\underbrace{[\mathcal{M};h,h]+[\mathcal{M};y_1,h]+[\mathcal{M};h,y_2]}_{x^{(2)}}+\underbrace{[\mathcal{M};y_1,y_2]}_{z^{(2)}}.\end{aligned}$$

Hence, the claim holds for $m = 2$.

Assume that for $m = r$, the claim holds. Note that for each $k_{r+1} \in \{1, \ldots, K\}$, $\{\mathcal{M}_{k_1 \ldots k_r k_{r+1}} : 1 \le k_1, \ldots, k_r \le K\}$ forms a r-way symmetric tensor of K dimensions. It follows that

$$\begin{aligned}[\mathcal{M}; \pi_1, \ldots, \pi_{r+1}] =& [\mathcal{M}; h, \ldots, h, \pi_{r+1}] + \sum_{s=1}^{r} [\mathcal{M}; h, \ldots, h, y_s, h, \ldots, h, \pi_{r+1}] \\ &+ \sum_{s_1=1}^{r-1} \sum_{s_2=s_1+1}^{r} [\mathcal{M}; h, \ldots, h, y_{s_1}, h, \ldots, h, y_{s_2}, \pi_{s_2+1} \ldots, \pi_{r+1}].\end{aligned}$$

Further, decompose π_{r+1} into $h + y_{r+1}$. By direct calculations

$$\begin{aligned}[\mathcal{M}; \pi_1, \ldots, \pi_r, \pi_{r+1}] =& \Big([\mathcal{M}; h, \ldots, h, h] + [\mathcal{M}; h, \ldots, h, y_{r+1}] \Big) \\ &+ \Big(\sum_{s=1}^{r} [\mathcal{M}; h, \ldots, h, y_s, h, \ldots, h, h] + \sum_{s=1}^{r} [\mathcal{M}; h, \ldots, h, y_s, h, \ldots, h, y_{r+1}] \Big) \\ &+ \sum_{s_1=1}^{m-1} \sum_{s_2=s_1+1}^{m} [\mathcal{M}; h, \ldots, h, y_{s_1}, h, \ldots, h, y_{s_2}, \pi_{s_2+1} \ldots, \pi_{r+1}] \\ =& [\mathcal{M}; h, \ldots, h] + \sum_{s=1}^{r+1} [\mathcal{M}; h, \ldots, h, y_s, h, \ldots, h] \\ &+ \sum_{s_1=1}^{r} \sum_{s_2=s_1+1}^{r+1} [\mathcal{M}; h, \ldots, h, y_{s_1}, h, \ldots, h, y_{s_2}, \pi_{s_2+1} \ldots, \pi_{r+1}], \\ =& x^{r+1} + z^{r+1},\end{aligned}$$

which suggests that the claim also holds for $m = r + 1$. By induction, (E.30) is proved.

E.4 Proof of Lemma E.2

Introduce $N_\theta = \sum_{i_3, \ldots, i_m (dist)} (\theta_{i_3} \cdots \theta_{i_m})^t$ and $I^{(i)}$ be the shorthand notation for set $\{1, \ldots, n\} \setminus \{i_3, \ldots, i_m\}$. Here, for convenience, we misuse the superscript (i) to indicate that this element depends on the choice of $(i_3, \ldots, i_m)$ whenever it is clear from the context.

By definitions and elementary algebra,

$$\begin{aligned}S = & \sum_{i_3, \ldots, i_m (dist)} \frac{(\theta_{i_1} \cdots \theta_{i_m})^t}{N_\theta} \sum_{\substack{k_3, \ldots, k_m = 1 \\ k'_3, \ldots, k'_m = 1}}^{K} \prod_{s=3}^{m} w_{i_s}^{(s)}(k_s) \widetilde{w}_{i_s}^{(s)}(k_s) \frac{N_\theta}{a_t} \\ & \cdot \Big[\sum_{i_1, i_2 (dist) \in I^{(i)}} (\theta_{i_1} \theta_{i_2})^t (y'_{i_1} \mathcal{M}_{::k_3 \cdots k_m} y_{i_2}) (\widetilde{y}'_{i_1} \mathcal{M}_{::k'_3 \cdots k'_m} \widetilde{y}_{i_2}) \Big],\end{aligned} \tag{E.31}$$

Let $\mathcal{M}_{::k_3 \cdots k_m} = \sum_{j=1}^{K} b_j^{(k)} b_j^{(k)'} \delta_j^{(k)}$, and $\mathcal{M}_{::k'_3 \cdots k'_m} = \sum_{j'=1}^{K} b_{j'}^{(k')} b_{j'}^{(k')'} \delta_j^{(k')}$ be the eigen-decomposition of the matrices $\mathcal{M}_{::k_3 \cdots k_m}$ and $\mathcal{M}_{::k'_3 \cdots k'_m}$, respectively. Introduce

$$X(i, j, j', k, k') = \sum_{i_1, i_2 (dist) \in I^{(i)}} (\theta_{i_1} \theta_{i_2})^t \delta_j^{(k)} \delta_{j'}^{(k')} (y'_{i_1} b_j^{(k)}) (y'_{i_2} b_j^{(k)}) (\widetilde{y}'_{i_1} b_{j'}^{(k')}) (\widetilde{y}'_{i_2} b_{j'}^{(k')}).$$

Then we can write

$$\sum_{i_1, i_2 (dist) \in I^{(i)}} (\theta_{i_1} \theta_{i_2})^t (y'_{i_1} \mathcal{M}_{::k_3 \cdots k_m} y_{i_2}) (\widetilde{y}'_{i_1} \mathcal{M}_{::k'_3 \cdots k'_m} \widetilde{y}_{i_2}) = \sum_{j, j'=1}^{K} X(i, j, j', k, k').$$

Inserting this into (E.31) gives

$$S = \sum_{i_3,\ldots,i_m(dist)} \frac{(\theta_{i_1}\cdots\theta_{i_m})^t}{N_\theta} \sum_{\substack{k_3,\ldots,k_m=1\\ k'_3,\ldots,k'_m=1}}^{K} \prod_{s=3}^{m} w_{i_s}^{(s)}(k_s)\widetilde{w}_{i_s}^{(s)}(k_s) \sum_{j,j'=1}^{K} \frac{1}{K^2}\Big(\frac{K^2 N_\theta}{a_t} X(i,j,j',k,k')\Big).$$

Note that $\sum_{i_3,\ldots,i_m(dist)} \frac{(\theta_{i_1}\cdots\theta_{i_m})^t}{N_\theta} \sum_{k_3,k'_3,\ldots,k_m,k'_m=1}^{K} \prod_{s=3}^{m} w_{i_s}^{(s)}(k_s)\widetilde{w}_{i_s}^{(s)}(k_s) \sum_{j,j'=1}^{K} \frac{1}{K^2} = 1$ and that $\exp(\cdot)$ is convex. By Jensen's inequality,

$$\exp(cS) \leq \sum_{\substack{i_3,\ldots,i_m\\ (dist)}} \frac{(\theta_{i_3}\cdots\theta_{i_m})^t}{N_\theta} \sum_{\substack{k_3,\ldots,k_m=1\\ k'_3,\ldots,k'_m=1}}^{K} \prod_{s=3}^{m} w_{i_s}^{(s)}(k_s)\widetilde{w}_{i_s}^{(s)}(k_s) \sum_{j,j'=1}^{K} \frac{1}{K^2} \exp\Big(\frac{cK^2 N_\theta}{a_t} X(i,j,j',k,k')\Big)$$

By assumptions $w_{i_s}^{(s)}, \widetilde{w}_{i_s}^{(s)}$ are independent of $y_{i_1}, y_{i_2}, \widetilde{y}_{i_1}, \widetilde{y}_{i_2}$, $3 \leq s \leq m$. Taking expectation on both sides gives

$$\begin{aligned}
\mathbb{E}[\exp(cS)] \leq & \sum_{\substack{i_3,\ldots,i_m\\ (dist)}} \frac{(\theta_{i_3}\cdots\theta_{i_m})^t}{N_\theta} \sum_{\substack{k_3,\ldots,k_m=1\\ k'_3,\ldots,k'_m=1}}^{K} \prod_{s=3}^{m} \mathbb{E}[w_{i_s}^{(s)}(k_s)]\mathbb{E}[\widetilde{w}_{i_s}^{(s)}(k_s)] \sum_{j,j'=1}^{K} \frac{1}{K^2} \\
& \cdot \mathbb{E}\Big[\exp\Big(\frac{cK^2 N_\theta}{a_t} X(i,j,j',k,k')\Big)\Big] \\
\leq & \max_{i,j,j',k,k'} \mathbb{E}\Big[\exp\Big(\frac{cK^2 N_\theta}{a_t} X(i,j,j',k,k')\Big)\Big].
\end{aligned}$$

Now, to show the claim, note that $N_\theta := \sum_{i_3,\ldots,i_m(dist)} (\theta_{i_3}\cdots\theta_{i_m})^t \leq \|\theta\|_t^{t(m-2)}$, we are sufficient to show that

$$X(i,j,j',k,k') \leq C\mu^2|T| + C\mu^2\|\theta\|_{2t}^{2t}, \tag{E.32}$$

where T is a random variable satisfying $\mathbb{P}(|T| > x) \leq 2\exp(-x/(2K^2\|\theta\|_{2t}^{2t}))$, for $x > 0$.

To see this, we rewrite

$$\begin{aligned}
X(i,j,j',k,k') := & \sum_{i_1,i_2\in I^{(i)}} (1-\mathbb{I}_{\{i_1=i_2\}})(\theta_{i_1}\theta_{i_2})^t \delta_j^{(k)}\delta_{j'}^{(k')} (y'_{i_1}b_j^{(k)})(y'_{i_2}b_j^{(k)})(\widetilde{y}'_{i_1}b_{j'}^{(k')})(\widetilde{y}'_{i_2}b_{j'}^{(k')}) \\
= & \delta_j^{(k)}\delta_{j'}^{(k')}(T_1 - T_2),
\end{aligned}$$

where

$$T_1 = \Big(\sum_{i_1\in I^{(i)}} \theta_{i_1}^t (y'_{i_1}b_j^{(k)})(\widetilde{y}'_{i_1}b_{j'}^{(k')})\Big)^2, \qquad T_2 = \sum_{i_1\in I^{(i)}} \Big(\theta_{i_1}^t (y'_{i_1}b_j^{(k)})(\widetilde{y}'_{i_1}b_{j'}^{(k')})\Big)^2.$$

Consider T_2 first. Note that $\max_{i_1}\{\|y_{i_1}\|, \|\widetilde{y}_{i_1}\|\} \leq \sqrt{K}$ and that $\|b_j^{(k)}\| = \|b_{j'}^{(k')}\| = 1$, $\forall$ j, j', k, k'. By direct calculations

$$|T_2| \leq (K)^2 \sum_{i_1} \theta_{i_1}^{2t} \leq C\|\theta\|_{2t}^{2t}.$$

Next, consider T_1. Let $Z = \sum_{i_1\in I^{(i)}} \theta_{i_1}^t (y'_{i_1}b_j^{(k)})(\widetilde{y}'_{i_1}b_{j'}^{(k')})$. Note that Z is a sum of $n-(m-2)$ independent random variables with $|\theta_{i_1}^t (y'_{i_1}b_j^{(k)})(\widetilde{y}'_{i_1}b_{j'}^{(k')})| \leq \sqrt{K}^2\theta_{i_1}^t$. By Hoeffding's inequality

$$\mathbb{P}(|Z| > x) \leq 2\exp\Big(-2x^2/(\sum_{i_1\in I^{(i)}} (2\sqrt{K}^2\theta_{i_1}^t)^2)\Big), \qquad \text{for } x > 0.$$

Combining this with $\sum_{i_1 \in I^{(i)}} (2\sqrt{K}^2 \theta_{i_1}^t)^2 \le 4K^2 \|\theta\|_{2t}^{2t}$ and $T_1 = Z^2$, it follows that

$$\mathbb{P}(|T_1| > x) \le 2\exp(-x/(2K^2\|\theta\|_{2t}^{2t})), \qquad \text{for } x > 0. \tag{E.33}$$

At the same time, recall that $\delta_j^{(k)}$,$\delta_{j'}^{(k')}$ are the eigenvalues of the matrices $\mathcal{M}_{::k_3\cdots k_m}$ and $\mathcal{M}_{::k_3'\cdots k_m'}$. By the assumption $\|\mathcal{M}_{::k_3\cdots k_m}\| \le C\mu$, for $1 \le k_3, \ldots, k_m \le K$, $\max_{j,k}\{|\delta_j^{(k)}|\} \le C\mu$. It is seen that

$$X(i,j,j',k,k') := \delta_j^{(k)}\delta_{j'}^{(k')}(T_1 - T_2) \le C\mu^2|T_1| + C\mu^2\|\theta\|_{2t}^{2t}, \qquad \text{with } T_1 \text{ satisfying (E.33)}.$$

This shows (E.32) and finishes the proof.

E.5 Proof of Lemma E.3

Similarly, let $N_\theta = \sum_{i_3,\ldots,i_m(dist)}(\theta_{i_3}\cdots\theta_{i_m})^t$ and $I^{(i)}$ be the shorthand notation for set $\{1,\ldots,n\}\setminus\{i_3,\ldots,i_m\}$. Here, for convenience, we misuse the superscript (i) to indicate that this element depends on the choice of $(i_3,\ldots,i_m)$ whenever it is clear from the context. Let $\mathcal{M}_{::k_3\cdots k_m} = \sum_{j=1}^K b_j^{(k)} b_j^{(k)'} \delta_j^{(k)}$, and $\mathcal{M}_{:k_2':k_4'\cdots k_m'} = \sum_{j'=1}^K b_{j'}^{(k')} b_{j'}^{(k')'} \delta_j^{(k')}$ be the eigen-decomposition of the matrices $\mathcal{M}_{::k_3\cdots k_m}$ and $\mathcal{M}_{:k_2':k_4'\cdots k_m'}$, respectively. Following the procedures in the proof of Lemma E.2, we can obtain

$$\begin{aligned}\exp(cS) \le \sum_{\substack{i_3,\ldots,i_m \\ (dist)}} \frac{(\theta_{i_3}\cdots\theta_{i_m})^t}{N_\theta} \sum_{\substack{k_3,\ldots,k_m=1 \\ k_2',k_4',\ldots,k_m'=1}}^{K} h(k_2')\widetilde{w}_{i_3}^{(3)}(k_3)\prod_{s=4}^m w_{i_s}^{(s)}(k_s)\widetilde{w}_{i_s}^{(s)}(k_s)\sum_{j,j'=1}^K \frac{1}{K^2} \\ \cdot \exp\Big(\frac{cK^2N_\theta}{a_t}X(i,j,j',k,k')\Big),\end{aligned} \tag{E.34}$$

where

$$X(i,j,j',k,k') = \sum_{i_1,i_2(dist)\in I^{(i)}} (\theta_{i_1}\theta_{i_2})^t \delta_j^{(k)}\delta_{j'}^{(k')}(y_{i_1}'b_j^{(k)})(y_{i_2}'b_j^{(k)})(\widetilde{y}_{i_1}'b_{j'}^{(k')})(\widetilde{y}_{i_3}'b_{j'}^{(k')}).$$

Note that $\widetilde{w}_{i_3}^{(3)}$ may not be independent of $\widetilde{y}_{i_3}$ which exists in $X(i,j,j',k,k')$. Consequently, we can not directly take expectation on both sides of (E.34) like that in Lemma E.2 to eliminate weight vectors $\{w_{i_j}^{(j)}\}$ by a maximum bound. To resolve this, we first derive a bound on $X(i,j,j',k,k')$ to eliminate $\widetilde{y}_{i_3}$.We rewrite

$$\begin{aligned}X(i,j,j',k,k') :=& \sum_{i_1,i_2\in I^{(i)}} (1-\mathbb{I}_{\{i_1=i_2\}})(\theta_{i_1}\theta_{i_2})^t \delta_j^{(k)}\delta_{j'}^{(k')}(y_{i_1}'b_j^{(k)})(y_{i_2}'b_j^{(k)})(\widetilde{y}_{i_1}'b_{j'}^{(k')})(\widetilde{y}_{i_3}'b_{j'}^{(k')}) \\ =&\delta_j^{(k)}\delta_{j'}^{(k')}(T_1 - T_2)(\widetilde{y}_{i_3}'b_{j'}^{(k')}),\end{aligned}$$

where

$$T_1 = \Big(\sum_{i_1\in I^{(i)}} \theta_{i_1}^t(y_{i_1}'b_j^{(k)})(\widetilde{y}_{i_1}'b_{j'}^{(k')})\Big)\Big(\sum_{i_2\in I^{(i)}} \theta_{i_2}^t(y_{i_2}'b_j^{(k)})\Big), \qquad T_2 = \sum_{i_1\in I^{(i)}}\Big(\theta_{i_1}^t(y_{i_1}'b_j^{(k)})\Big)^2(\widetilde{y}_{i_1}'b_{j'}^{(k')}).$$

Recall that $\delta_j^{(k)}$,$\delta_{j'}^{(k')}$ are the eigenvalues of the matrices $\mathcal{M}_{::k_3\cdots k_m}$ and $\mathcal{M}_{:k_2':k_4'\cdots k_m'}$. By the assumption $\|\mathcal{M}_{::k_3\cdots k_m}\| \le C\mu$, for $1 \le k_3,\ldots,k_m \le K$, $\max_{j,k}\{|\delta_j^{(k)}|\} \le C\mu$. Combining this with $\|b_{j'}^{(k')}\| = 1$ and $\|y_{i_3}\| \le \sqrt{K}$, we have

$$X(i,j,j',k,k') := \delta_j^{(k)}\delta_{j'}^{(k')}(T_1 - T_2)(\widetilde{y}_{i_3}'b_{j'}^{(k')}) \le C\mu^2(|T_1| + |T_2|).$$

Note that T_1, T_2 (and so the bound) are independent of $w_{i_s}^{(s)}, \widetilde{w}_{i_s}^{(s)}$, $3 \le s \le m$. Applying this inequality to the RHS of (E.34) and taking expectation on both sides give

$$\begin{aligned}\mathbb{E}[\exp(cS)] \le & \sum_{\substack{i_3,\dots,i_m \\ (dist)}} \frac{(\theta_{i_3}\cdots\theta_{i_m})^t}{N_\theta} \sum_{\substack{k_3,\dots,k_m=1 \\ k_2',k_4',\dots,k_m'=1}}^{K} h(k_2')\mathbb{E}[\widetilde{w}_{i_3}^{(3)}(k_3)] \prod_{j=4}^{m} \mathbb{E}[w_{i_s}^{(s)}(k_s)]\mathbb{E}[\widetilde{w}_{i_s}^{(s)}(k_s)] \sum_{j,j'=1}^{K} \frac{1}{K^2} \\ & \cdot \mathbb{E}\Big[\exp\Big(\frac{CN_\theta}{a_t}\mu^2(|T_1|+|T_2|)\Big)\Big] \\ \le & \max_{i,j,j',k,k'} \mathbb{E}\Big[\exp\Big(\frac{CN_\theta}{a_t}\mu^2(|T_1|+|T_2|)\Big)\Big].\end{aligned}$$

Now, to show the claim, note that $N_\theta := \sum_{i_3,\dots,i_m(dist)}(\theta_{i_3}\cdots\theta_{i_m})^t \le \|\theta\|_t^{t(m-2)}$, it is then sufficient to show that

$$(I): \mathbb{P}(|T_1| > x) \le 4\exp(-x/(2K^2\|\theta\|_{2t}^{2t})), \quad \forall x > 0, \qquad (II): |T_2| \le C\|\theta\|_{2t}^{2t}. \tag{E.35}$$

Consider (I) first. Let $Z_1 = \sum_{i_1 \in I^{(i)}} \theta_{i_1}^t (y_{i_1}' b_j^{(k)})(\widetilde{y}_{i_1}' b_{j'}^{(k')})$, $Z_2 = \sum_{i_2 \in I^{(i)}} \theta_{i_2}^t (y_{i_2}' b_j^{(k)})$ and so $T_1 = Z_1 \cdot Z_2$. Note that Z_1 and Z_2 are the sum of $n-(m-2)$ independent random variables. Similarly, by Hoeffding's inequality, for any $x > 0$

$$\mathbb{P}(|Z_1| > x) \le 2\exp(-2x/((2K)^2\|\theta\|_{2t}^{2t})), \qquad \mathbb{P}(|Z_2| > x) \le 2\exp(-2x/((2\sqrt{K})^2\|\theta\|_{2t}^{2t})).$$

Combining this with $|T_1| = |Z_1|\cdot|Z_2|$ and union bound $\mathbb{P}(|Z_1||Z_2| > x) \le \mathbb{P}(|Z_1| > \sqrt{x}) + \mathbb{P}(|Z_1||Z_2| > \sqrt{x})$,

$$\mathbb{P}(|T_1| > x) \le 2\exp(-x/(2K^2\|\theta\|_{2t}^{2t})) + 2\exp(-x/(2K\|\theta\|_{2t}^{2t})) \le 4\exp(-x/(2K^2\|\theta\|_{2t}^{2t})),$$

which proves the first claim in (E.35).

Next, consider (II) in (E.35). By $\max_{i_1}\{\|y_{i_1}\|, \|\widetilde{y}_{i_1}\|\} \le \sqrt{K}$, $\|b_j^{(k)}\| = \|b_{j'}^{(k')}\| = 1, \forall\, j, j', k, k'$

$$|T_2| := \sum_{i_1 \in I^{(i)}} \Big(\theta_{i_1}^t (y_{i_1}' b_j^{(k)})\Big)^2 (\widetilde{y}_{i_1}' b_{j'}^{(k')}) \le \sum_{i_1} \theta_{i_1}^{2t} (\sqrt{K})^2 \sqrt{K} \le C\|\theta\|_{2t}^{2t},$$

which proves (II) and finishes the whole proof.

E.6 Proof of Lemma E.4

The proof is similar to that in Lemma E.3. Similarly, let $N_\theta = \sum_{i_3,\dots,i_m(dist)}(\theta_{i_3}\cdots\theta_{i_m})^t$ and $I^{(i)}$ be the shorthand notation for set $\{1,\dots,n\} \setminus \{i_3,\dots,i_m\}$. Here, for convenience, we misuse the superscript (i) to indicate that this element depends on the choice of $(i_3,\dots,i_m)$ whenever it is clear from the context. Let $\mathcal{M}_{::k_3\cdots k_m} = \sum_{j=1}^K b_j^{(k)} b_j^{(k)\prime} \delta_j^{(k)}$, and $\mathcal{M}_{k_1'k_2'::k_5'\cdots k_m'} = \sum_{j'=1}^K b_{j'}^{(k')} b_{j'}^{(k')\prime} \delta_j^{(k')}$ be the eigen-decomposition of the matrices $\mathcal{M}_{::k_3\cdots k_m}$ and $\mathcal{M}_{k_1'k_2'::k_5'\cdots k_m'}$, respectively. Following the procedures in the proof of Lemma E.2, we can obtain

$$\begin{aligned}\exp(cS) \le & \sum_{\substack{i_3,\dots,i_m \\ (dist)}} \frac{(\theta_{i_3}\cdots\theta_{i_m})^t}{N_\theta} \sum_{\substack{k_3,\dots,k_m=1 \\ k_1'k_2',k_5',\dots,k_m'=1}}^{K} h(k_1')h(k_2') \prod_{s=5}^{m} w_{i_s}^{(s)}(k_s)\widetilde{w}_{i_s}^{(s)}(k_s) \\ & \cdot \widetilde{w}_{i_3}^{(3)}(k_3)\widetilde{w}_{i_4}^{(4)}(k_4) \sum_{j,j'=1}^{K} \frac{1}{K^2} \cdot \exp\Big(\frac{cK^2 N_\theta}{a_t} X(i,j,j',k,k')\Big),\end{aligned} \tag{E.36}$$

where

$$X(i,j,j',k,k') = \sum_{i_1,i_2(dist)\in I^{(i)}} (\theta_{i_1}\theta_{i_2})^t \delta_j^{(k)}\delta_{j'}^{(k')}(y'_{i_1}b_j^{(k)})(y'_{i_2}b_j^{(k)})(\widetilde{y}'_{i_3}b_{j'}^{(k')})(\widetilde{y}'_{i_4}b_{j'}^{(k')}).$$

Note that $\widetilde{w}_{i_3}^{(3)}$ and $\widetilde{w}_{i_4}^{(4)}$ may not be independent of $\widetilde{y}_{i_3}$ and $\widetilde{y}_{i_4}$ which exist in $X(i,j,j',k,k')$. Similar to the proof of Lemma E.3, we rewrite

$$\begin{aligned} X(i,j,j',k,k') &:= \sum_{i_1,i_2\in I^{(i)}} (1-\mathbb{I}_{\{i_1=i_2\}})(\theta_{i_1}\theta_{i_2})^t \delta_j^{(k)}\delta_{j'}^{(k')}(y'_{i_1}b_j^{(k)})(y'_{i_2}b_j^{(k)})(\widetilde{y}'_{i_3}b_{j'}^{(k')})(\widetilde{y}'_{i_4}b_{j'}^{(k')}) \\ &= \delta_j^{(k)}\delta_{j'}^{(k')}(T_1-T_2)(\widetilde{y}'_{i_3}b_{j'}^{(k')})(\widetilde{y}'_{i_4}b_{j'}^{(k')}), \end{aligned}$$

where

$$T_1 = \Big(\sum_{i_1\in I^{(i)}} \theta_{i_1}^t (y'_{i_1}b_j^{(k)})\Big)^2, \qquad T_2 = \sum_{i_1\in I^{(i)}} \Big(\theta_{i_1}^t (y'_{i_1}b_j^{(k)})\Big)^2.$$

Recall that $\delta_j^{(k)}$,$\delta_{j'}^{(k')}$ are the eigenvalues of the matrices $\mathcal{M}_{::k_3\cdots k_m}$ and $\mathcal{M}_{:k'_2:k'_4\cdots k'_m}$. By the assumption $\|\mathcal{M}_{::k_3\cdots k_m}\| \leq C\mu$, for $1 \leq k_3,\dots,k_m \leq K$, $\max_{j,k}\{|\delta_j^{(k)}|\} \leq C\mu$. Combining this with $\|b_{j'}^{(k')}\| = 1$ and $\|y_{i_3}\| \leq \sqrt{K}$, we have

$$X(i,j,j',k,k') := \delta_j^{(k)}\delta_{j'}^{(k')}(T_1-T_2)(\widetilde{y}'_{i_3}b_{j'}^{(k')})(\widetilde{y}'_{i_4}b_{j'}^{(k')}) \leq C\mu^2(|T_1|+|T_2|).$$

Note that T_1, T_2 (and so the bound) are independent of $w_{i_s}^{(s)}, \widetilde{w}_{i_s}^{(s)}$, $3 \leq s \leq m$. Applying this inequality to the RHS of (E.36) and taking expectation on both sides give

$$\mathbb{E}[\exp(cS)] \leq \max_{i,j,j',k,k'} \mathbb{E}\Big[\exp\Big(\frac{CN_\theta}{a_t}\mu^2(|T_1|+|T_2|)\Big)\Big].$$

Now, to show the claim, note that $N_\theta := \sum_{i_3,\dots,i_m(dist)}(\theta_{i_3}\cdots\theta_{i_m})^t \leq \|\theta\|_t^{t(m-2)}$, it is then sufficient to show that

$$(I): \mathbb{P}(|T_1|>x) \leq 2\exp(-x/(2K\|\theta\|_{2t}^{2t})), \quad \forall x>0, \qquad (II): |T_2| \leq C\|\theta\|_{2t}^{2t}.$$

The procedures to show them are the same as that in the proof of Lemma E.2. So we omit them.

E.7 Proof of Lemma E.5

The following lemma is used in this proof and we prove it below.

Lemma E.6 (Each element of community structure tensor is close to one)**.** *Using the same notations of Theorem 3.1, for each $m \in \{2,\dots,M\}$,*

$$\max_{1\leq i_1,\dots,i_m\leq K}\{|\mathcal{P}_{i_1\cdots i_m}^{(m)}-1|\} \asymp |\mu_2^{(m)}|. \tag{E.37}$$

Fix m, for simplicity of notation, we remove the superscript (m) whenever it is clear from the context. Recall that $D = \text{diag}(d_1,\cdots,d_K)$ and $h = \mathbb{E}[D^{-1}\pi_i/\|D^{-1}\pi_i\|_1]$. Write for short $s = \sum_{k=1}^K d_k h_k$ and $v = (d_1,\dots,d_K)'$. With these notations and direct calculations, for $1 \leq k_3,\dots,k_m \leq K$

$$\mathcal{M}_{::k_3\cdots k_m} = D(\mathcal{P}_{::k_3\cdots k_m} - \mathbf{1}_K\mathbf{1}'_K)D\prod_{j=3}^m d_{k_j} + (\prod_{j=3}^m d_{k_j} - s^{m-2})vv' + (s^{m-2}vv' - \mathbf{1}_K\mathbf{1}'_K).$$

Therefore, to prove the first claim of this lemma, by elementary algebra, it is sufficient to show that

$$(a): \max_{1\leq k_1,\ldots,k_m\leq K}\{|\mathcal{P}_{k_1\cdots k_m}-1|\}\leq C|\mu_2|,$$
$$(b): \max_{1\leq k\leq K}\{d_k\}\leq C,$$
$$(c): \max_{1\leq i,j\leq K}\{|(s^{m-2}vv'-\mathbf{1}_K\mathbf{1}_K')_{ij}|\}\leq C|\mu_2|,$$
$$(d): \max_{1\leq k\leq K}\{|d_k-s|\}\leq C|\mu_2|,$$

where we note that (a) is implied by Lemma E.6.

Consider (b). Recall that by degree matching

$$\sum_{k_2,\ldots,k_m=1}^{K} D\mathcal{P}_{:k_2\cdots k_m}\prod_{j=2}^{m}(d_{k_j}h_{k_j})=\mathbf{1}_K. \tag{E.38}$$

Note that each element of $\mathcal{P}$ is non-negative and $\mathcal{P}_{k_1\cdots k_1}=1$ for $1\leq k_1\leq K$. It follows that

$$d_{k_1}(d_{k_1}h_{k_1})^{m-1}\leq \sum_{k_2,\ldots,k_m=1}^{K} d_{k_1}\mathcal{P}_{k_1\cdots k_m}\prod_{j=2}^{m}(d_{k_j}h_{k_j})=1, \qquad 1\leq k_1\leq K.$$

Combining this with our assumption $\min_{1\leq k\leq K}\{h_k\}\geq C$,

$$d_k\leq h_k^{-(m-1)/m}\leq C, \qquad 1\leq k\leq K, \tag{E.39}$$

which proves (b).

Next consider (c). Let $\mathcal{H}$ be a tensor defined by $\mathcal{H}_{k_1\cdots k_m}=\mathcal{P}_{k_1\cdots k_m}-1$, for all $1\leq k_1,\ldots,k_m\leq K$ and introduce w as the vector $\sum_{k_2\cdots k_m=1}^{K} D\mathcal{H}_{:k_2\cdots k_m}\prod_{j=2}^{m}(d_{k_j}h_{k_j})$. Recall that $s=\sum_{k=1}^{K} d_k h_k$. By definitions and calculations, (E.38) can be written as

$$w+s^{m-1}v=\mathbf{1}_K. \tag{E.40}$$

Note that $h'v=s$. Left multiplying h' on both sides gives

$$h'w+s^m=1. \tag{E.41}$$

At the same time, inserting (E.40) into $s^{m-2}vv'-\mathbf{1}_K\mathbf{1}_K'$ through $\mathbf{1}_K$ gives

$$\begin{aligned} s^{m-2}vv'-\mathbf{1}_K\mathbf{1}_K' =& s^{m-2}vv'-(w+s^{m-1}v)(w+s^{m-1}v)' \\ =& s^{m-2}(1-s^m)vv'-s^{m-1}wv'-s^{m-1}vw'-ww'. \end{aligned}$$

Note that by (E.41), $1-s^m=h'w$. It follows that

$$s^{m-2}vv'-\mathbf{1}_K\mathbf{1}_K' = s^{m-2}h'wvv'-s^{m-1}wv'-s^{m-1}vw'-ww'.$$

By (E.39), $\max_{1\leq k\leq K}\{h_k\}\leq 1$ and elementary algebra

$$\max_{1\leq i,j\leq K}\{|(s^{m-2}vv'-\mathbf{1}_K\mathbf{1}_K')_{ij}|\}\leq C\|h\|_{\max}\cdot\|v\|_{\max}\cdot\|w\|_{\max}\leq C\|\mathcal{H}\|_{\max},$$

where $\|\cdot\|_{\max}$ is the element-wise maximum norm and $\|\mathcal{H}\|_{\max}:=\max_{k_1,\ldots,k_m}\{|\mathcal{P}_{k_1,\ldots,k_m}-1|\}\leq C|\mu_2|$. This proves (c).

On the other hand, by elementary algebra, $|(s^{m-2}vv'-\mathbf{1}_K\mathbf{1}_K')_{ii}|\leq\|s^{m-2}vv'-\mathbf{1}_K\mathbf{1}_K'\|$, for all $1\leq i\leq K$ and so

$$|s^{m-2}d_i d_i-1|\leq C|\mu_2|.$$

Transforming the above formula gives,

$$d_i = s^{-(m-2)/2} + O(|\mu_2|). \tag{E.42}$$

Summing up with weight h_i in terms of i on two sides and noting that $\sum_i h_i = 1$, it gives

$$s = s^{-(m-2)/2} + O(|\mu_2|). \tag{E.43}$$

Combining this with (E.42) gives (d).

Next we consider the second claim of this lemma i.e. $\max_{1\le i\le K}\{|d_i - 1|\} \le C|\mu_2|$. By elementary algebra, (E.43) can be rewritten as

$$s = 1 + \frac{\sqrt{s}^{m-1} + \sqrt{s}^{m-2}}{\sum_{j=0}^{m-1}\sqrt{s}^j}\cdot O(|\mu_2|),$$

where we note that $\frac{\sqrt{s}^{m-1}+\sqrt{s}^{m-2}}{\sum_{j=0}^{m-1}\sqrt{s}^j} \le 1$. Combining this with (E.42) proves the second claim.

E.8 Proof of Lemma E.6

Since the claim is argued for each m-uniform tensor $\mathcal{P}^{(m)}$ separately, fixing m, we remove the superscript (m) whenever it is clear from the context.

Let the $K \times K^{m-1}$ matrix P denote the matricization of $\mathcal{P}^{(m)}$. Let $U\Sigma V'$ be the SVD of P, where $U = (u_1, \ldots, u_K)$, $V = (v_1, \ldots, v_{K^{m-1}})$ and $\Sigma = (\mathrm{diag}(\mu_1, \ldots, \mu_K), \mathbf{0}_{K\times(K^{m-1}-K)})$.

To show that claim, it is sufficient to show that

$$(I): |\mu_2| \le C \max_{1\le i_1,\ldots,i_m\le K}\{|\mathcal{P}_{i_1\cdots i_m} - 1|\}, \qquad (II): \max_{1\le i_1,\ldots,i_m\le K}\{|\mathcal{P}_{i_1\cdots i_m} - 1|\} \le C|\mu_2|.$$

Consider (I) first. Let P_0 be the $K \times K^{m-1}$ matrix of ones. Recall that μ_2 is the second singular value of P, and note that the second singular value of P_0 is 0. By [4, Corollary 7.3.5, Page 451],

$$|\mu_2| \le \|P - P_0\|.$$

At the same time, by elementary algebra, $\|P - P_0\| \le C\max_{1\le i_1,\ldots,i_m\le K}\{|\mathcal{P}_{i_1\cdots i_m} - 1|\}$. Combining these proves (I).

Next we consider (II).

By our assumption $\|P\| \le C$ and elemantary algebra,

$$\max_{1\le i_1,\ldots,i_m\le K}\{|\mathcal{P}_{i_1\cdots i_m}|\} = \|P\|_{\max} \le \|P\| \le C,$$

where $\|\cdot\|_{\max}$ is the element-wise maximum norm. Therefore, (II) directly holds for the case that $|\mu_2| \ge \epsilon$ for some positive constants $\epsilon < 1$. It is then sufficient to consider the case when $|\mu_2| < \epsilon$.

By definitions,

$$(PP')_{ii} \ge \mathcal{P}^2_{i\cdots i} = 1, \qquad (PP')_{ij} \ge 0, \qquad 1 \le i,j \le K.$$

Therefore, by Perron's theorem [4], the first eigenvalue (in magnitude) and each entry of the first eigenvector of PP' are positive. Note that $PP' = U\Sigma^2U'$, it follows that

$$\mu_1 > 0, \qquad u_1(i) > 0, \qquad 1 \le i \le K.$$

Let $a = u_1\mu_1^{\frac{1}{m}}$ and $b = v_1\mu_1^{\frac{m-1}{m}}$ be the scaled version of u_1 and v_1, where $a_i > 0$ since $u_1(i) > 0$, $1 \le i \le K$. Introduce $\widetilde{P} = ab'$. For simplicity, we misuse the notation $b_{i_2\cdots i_m}$

for $b_{i_2+\sum_{s=3}^m K^{s-2}(i_s-1)}$. To show (II), by triangle inequality, it is sufficient to show that for $1 \le i_1, \ldots, i_m \le K$,

$$(IIa): |\mathcal{P}_{i_1\cdots i_m} - a_{i_1} b_{i_2\cdots i_m}| \le C|\mu_2|, \qquad (IIb): |a_{i_1} b_{i_2\cdots i_m} - 1| \le C|\mu_2|.$$

Note that by elementary algebra

$$|\mathcal{P}_{i_1\cdots i_m} - a_{i_1} b_{i_2\cdots i_m}| \le \|P - \widetilde{P}\|_{\max} \le \|P - \widetilde{P}\| = |\mu_2|, \tag{E.44}$$

This proves (IIa).

It is left to show (IIb). We start by showing that a is a vector with elements are almost the same. By equality $x^m - y^m = (x-y)\sum_{j=0}^{m-1} x^{m-1-j} y^j$, we have,

$$|a_{i_1} - a_{i_2}| = \frac{|a_{i_1}^m - a_{i_2}^m|}{\sum_{j=0}^{m-1} a_{i_1}^{m-j-1} a_{i_2}^j} = \frac{|a_{i_1}/a_{i_2} - (a_{i_2}/a_{i_1})^{m-1}|}{\sum_{j=0}^{m-1} a_{i_1}^{-j} a_{i_2}^{j-1}}, \qquad 1 \le i_1, i_2 \le K.$$

Combining this with triangle's inequality $|a_{i_1}/a_{i_2} - (a_{i_2}/a_{i_1})^{m-1}| \le |a_{i_1} b_{i_2\cdots i_2} - a_{i_1}/a_{i_2}| + |a_{i_1} b_{i_2\cdots i_2} - (a_{i_2}/a_{i_1})^{m-1}|$,

$$|a_{i_1} - a_{i_2}| \le \frac{|a_{i_1} b_{i_2\cdots i_2} - a_{i_1}/a_{i_2}| + |a_{i_1} b_{i_2\cdots i_2} - (a_{i_2}/a_{i_1})^{m-1}|}{\sum_{j=0}^{m-1} a_{i_1}^{-j} a_{i_2}^{j-1}}, \qquad 1 \le i_1, i_2 \le K. \tag{E.45}$$

We claim that for $1 \le k \le m$ the following holds and prove it later.

$$\Big| a_{i_1} b_{i_2\cdots i_k i_1 \cdots i_1} - \frac{\prod_{j=1}^k a_{i_j}}{a_{i_1}^k} \Big| \le \Big(2\sum_{s=2}^k \frac{\prod_{j=s+1}^k a_{i_j}}{a_{i_1}^{k-s}} + \frac{\prod_{j=1}^k a_{i_j}}{a_{i_1}^k} \Big) |\mu_2|, \quad 1 \le i_1, \ldots, i_m \le K. \tag{E.46}$$

By setting $k = m; i_3, \ldots, i_m = i_2$ and $k = 1, i_1 = i_2$ separately in the above inequality, we obtain

$$\Big| a_{i_1} b_{i_2\cdots i_2} - \frac{a_{i_2}^{m-1}}{a_{i_1}^{m-1}} \Big| \le \Big(2\sum_{s=2}^m \frac{a_{i_2}^{m-s}}{a_{i_1}^{m-s}} + \frac{a_{i_2}^{m-1}}{a_{i_1}^{m-1}} \Big) |\mu_2|, \qquad \Big| a_{i_1} b_{i_2\cdots i_2} - \frac{a_{i_1}}{a_{i_2}} \Big| \le \frac{a_{i_1}}{a_{i_2}} |\mu_2|.$$

Inserting the above into the RHS of (E.45) and by direct calculations

$$|a_{i_1} - a_{i_2}| \le \frac{1}{\sum_{j=0}^{m-1} a_{i_1}^{-j} a_{i_2}^{j-1}} \Big(2\sum_{s=2}^m \frac{a_{i_2}^{m-s}}{a_{i_1}^{m-s}} |\mu_2| + \frac{a_{i_2}^{m-1}}{a_{i_1}^{m-1}} |\mu_2| + \frac{a_{i_1}}{a_{i_2}} |\mu_2| \Big) = (a_{i_1} + a_{i_2}) |\mu_2|.$$

Combining this inequality with $\sum_{j=1}^K (a_i - |a_i - a_j|) \le \sum_{j=1}^K a_j \le \sum_{j=1}^K (a_i + |a_i - a_j|)$ give

$$\sum_{i_2=1}^K \Big(a_{i_1} - (a_{i_1} + a_{i_2}) |\mu_2| \Big) \le \sum_{i_2=1}^K a_{i_2} \le \sum_{i_2=1}^K \Big(a_{i_1} + (a_{i_1} + a_{i_2}) |\mu_2| \Big).$$

By $\sum_{i_2=1}^K a_{i_2} = \|a\|_1$, we can rewrite it as

$$\frac{\|a\|_1}{K} \frac{1 - |\mu_2|}{1 + |\mu_2|} \le a_{i_1} \le \frac{\|a\|_1}{K} \frac{1 + |\mu_2|}{1 - |\mu_2|}.$$

Note that $|\mu_2| < \epsilon < 1$, it is seen that

$$a_{i_1} = \frac{\|a\|_1}{K} (1 + O(|\mu_2|)), \qquad 1 \le i_1 \le K. \tag{E.47}$$

Now we are ready to show (IIb). By triangle inequality

$$|a_{i_1} b_{i_2\cdots i_m} - 1| \le |a_{i_1} b_{i_2\cdots i_m} - \frac{\prod_{j=1}^m a_{i_j}}{a_{i_1}^m}| + |\frac{\prod_{j=1}^m a_{i_j}}{a_{i_1}^m} - 1|. \tag{E.48}$$

Note that setting $k=m$ in (E.46) gives

$$\Big|a_{i_1}b_{i_2\cdots i_m}-\frac{\prod_{j=1}^m a_{i_j}}{a_{i_1}^m}\Big|\le\Big(2\sum_{s=2}^m\frac{\prod_{j=s+1}^m a_{i_j}}{a_{i_1}^{m-s}}+\frac{\prod_{j=1}^m a_{i_j}}{a_{i_1}^m}\Big)|\mu_2|.$$

Inserting this into (E.48). By direct calculations and (E.47)

$$|a_{i_1}b_{i_2\cdots i_m}-1|\le\Big(2\sum_{s=2}^m\frac{\prod_{j=s+1}^m a_{i_j}}{a_{i_1}^{m-s}}+\frac{\prod_{j=1}^m a_{i_j}}{a_{i_1}^m}\Big)|\mu_2|+|\frac{\prod_{j=1}^m a_{i_j}}{a_{i_1}^m}-1|=O(|\mu_2|).$$

which holds proves (IIb) and finishes the main proof of this lemma.

Lastly, we prove the claim (E.46), which is done by induction. Consider $k=1$, the goal is to show

$$|a_{i_1}b_{i_1\cdots i_1}-1|\le|\mu_2|,\qquad 1\le i_1\le K \tag{E.49}$$

Since $\mathcal{P}_{i_1\cdots i_1}=1$, for $1\le i_1\le K$. By (E.44), we have

$$|a_{i_1}b_{i_1\cdots i_1}-1|\le|\mu_2|,$$

which is exactly (E.49) and so the claim (E.46) holds for $k=1$.

Now, assume that the claim holds for $k=k_0$ and the goal is to show that this implies that the claim holds for $k=k_0+1$. By triangle's inequality,

$$\begin{aligned}&\Big|a_{i_1}b_{i_2\cdots i_{k_0+1}i_1\cdots i_1}-\frac{\prod_{j=1}^{k_0+1}a_{i_j}}{a_{i_1}^{k_0+1}}\Big|\\ \le&\Big|a_{i_1}b_{i_2\cdots i_{k_0+1}i_1\cdots i_1}-\mathcal{P}_{i_1\cdots i_k i_{k_0+1}i_1\cdots i_1}\Big|+\Big|\mathcal{P}_{i_1\cdots i_k i_{k_0+1}i_1\cdots i_1}-\mathcal{P}_{i_{k_0+1}i_1\cdots i_{k_0}i_1\cdots i_1}\Big|\\ &+\Big|\mathcal{P}_{i_{k_0+1}i_1\cdots i_{k_0}i_1\cdots i_1}-a_{i_{k_0+1}}b_{i_2\cdots i_{k_0}i_1\cdots i_1}\Big|+\Big|a_{i_{k_0+1}}b_{i_2\cdots i_{k_0}i_1\cdots i_1}-\frac{\prod_{j=1}^{k_0+1}a_{i_j}}{a_{i_1}^{k_0+1}}\Big|\end{aligned}$$

By (E.44), the first term and the third is bounded by $|\mu_2|$. Also, by symmetry of $\mathcal{P}$, the second term is 0. Moving a factor $a_{i_{k_0+1}}/a_{i_1}$ from the last term, it follows that

$$\begin{aligned}\Big|a_{i_1}b_{i_2\cdots i_{k_0+1}i_1\cdots i_1}-\frac{\prod_{j=1}^{k_0+1}a_{i_j}}{a_{i_1}^{k_0+1}}\Big|&\le 2|\mu_2|+\frac{a_{i_{k_0+1}}}{a_{i_1}}\Big|a_{i_1}b_{i_2\cdots i_{k_0}i_1\cdots i_1}-\frac{\prod_{j=1}^{k_0}a_{i_j}}{a_{i_1}^{k_0}}\Big|\\ \text{(By the assumption for } k=k_0\text{)}\quad&\le 2|\mu_2|+\frac{a_{i_{k_0+1}}}{a_{i_1}}\Big(2\sum_{s=2}^{k_0}\frac{\prod_{j=s+1}^{k}a_{i_j}}{a_{i_1}^{k_0-s}}+\frac{\prod_{j=1}^{k}a_{i_j}}{a_{i_1}^{k_0}}\Big)|\mu_2|\\ &=\Big(2\sum_{s=2}^{k_0+1}\frac{\prod_{j=s+1}^{k_0+1}a_{i_j}}{a_{i_1}^{k_0+1-s}}+\frac{\prod_{j=1}^{k_0+1}a_{i_j}}{a_{i_1}^{k_0+1}}\Big)|\mu_2|,\end{aligned}$$

which shows (E.46) also holds for $k=k_0+1$. Hence, by induction, (E.46) holds for $1\le k\le m$.

F Proof of Lemma 2.2

We have the following lemma which is used in the proof of Lemma 2.2 and prove it below.

Lemma F.1. *Under the conditions of Lemma 2.2, as $n\to\infty$, with probability at least $1-O(1/n)$,*

- *(a) Under both the null and under the alternative, $|\hat{\alpha}_n-\widetilde{\alpha}_n|\le C\log(n)(\widetilde{\alpha}_n/n^3)^{1/2}$.*
- *(b) Under the alternative, $\widetilde{\alpha}_n\le\max_{1\le k_1,k_2,k_3\le K}\{\mathcal{P}_{k_1k_2k_3}\}\le C\widetilde{\alpha}_n$ and $\widetilde{\alpha}_n=h'(\mathcal{P}h)h+O(\frac{\widetilde{\alpha}_n}{n})$.*

To show the claims of Lemma 2.2, it is sufficient to show as $n \to \infty$, for any positive constant M,

$$\psi_n \to N(0,1) \text{ under the null}, \qquad \text{and} \qquad \mathbb{P}(|\psi_n| \le M) \to 0 \text{ under the alternative.} \tag{F.50}$$

Recall that

$$\widetilde{\alpha}_n = \mathbb{E}[\hat{\alpha}_n],$$

Let $\mathcal{A}^*$ and $\widetilde{\mathcal{A}}$ be two tensors with the same size as $\mathcal{A}$, where $\mathcal{A}^*_{i_1 i_2 i_3} = \mathcal{A}_{i_1 i_2 i_3} - \hat{\alpha}_n$ and $\widetilde{\mathcal{A}}_{i_1 i_2 i_3} = \mathcal{A}_{i_1 i_2 i_3} - \widetilde{\alpha}_n$ if i_1, i_2, i_3 are distinct, and $\mathcal{A}^*_{i_1 i_2 i_3} = \widetilde{\mathcal{A}}_{i_1 i_2 i_3} = 0$ otherwise. By definitions,

$$\sqrt{2n}\psi_n = \frac{\sum_{1 \le i \le n} \left(\sum_{j<k} \mathcal{A}^*_{ijk} \right)^2 - n\binom{n-1}{2}\hat{\alpha}_n(1-\hat{\alpha}_n)}{\binom{n-1}{2}\hat{\alpha}_n(1-\hat{\alpha}_n)}. \tag{F.51}$$

Let $S_0 = \{(i_1, i_2, i_3, i_4, i_5) : 1 \le i_1, i_2, i_3, i_4, i_5 \le n; i_1 < i_2; i_4 < i_5; i_1, i_2, i_4, i_5 \ne i_3\}$, and write for short $x = (i_1, i_2, i_3, i_4, i_5)$. Introduce a subset of S_0 by $S = \{x \in S_0 : (i_1, i_2) \ne (i_4, i_5)\}$. Note that for any $x \in S_0 \setminus S$, $(i_1, i_2) = (i_4, i_5)$. It is seen that the numerator on the RHS of (F.51) is

$$\begin{aligned} &\sum_{x \in S_0} \mathcal{A}^*_{i_1 i_2 i_3}\mathcal{A}^*_{i_3 i_4 i_5} - n\binom{n-1}{2}\hat{\alpha}_n(1-\hat{\alpha}_n) \\ =&\sum_{x \in S} \mathcal{A}^*_{i_1 i_2 i_3}\mathcal{A}^*_{i_3 i_4 i_5} + \sum_{x \in S_0 \setminus S} \mathcal{A}^*_{i_1 i_2 i_3}\mathcal{A}^*_{i_3 i_4 i_5} - n\binom{n-1}{2}\hat{\alpha}_n(1-\hat{\alpha}_n) \\ =&(I) + (II), \end{aligned} \tag{F.52}$$

where

$$(I) = \sum_{x \in S} \mathcal{A}^*_{i_1 i_2 i_3}\mathcal{A}^*_{i_3 i_4 i_5}, \qquad (II) = \sum_{x \in S_0 \setminus S} \mathcal{A}^*_{i_1 i_2 i_3}\mathcal{A}^*_{i_3 i_4 i_5} - n\binom{n-1}{2}\hat{\alpha}_n(1-\hat{\alpha}_n).$$

Consider (I) first. Write

$$(I) = (Ia) + (Ib), \tag{F.53}$$

where

$$(Ia) = \sum_{x \in S} \widetilde{\mathcal{A}}_{i_1 i_2 i_3}\widetilde{\mathcal{A}}_{i_3 i_4 i_5}, \qquad (Ib) = \sum_{x \in S} (\mathcal{A}^*_{i_1 i_2 i_3}\mathcal{A}^*_{i_3 i_4 i_5} - \widetilde{\mathcal{A}}_{i_1 i_2 i_3}\widetilde{\mathcal{A}}_{i_3 i_4 i_5}).$$

Now, by direct calculations,

$$(Ib) = (\widetilde{\alpha}_n - \hat{\alpha}_n) \sum_{x \in S} (\mathcal{A}_{i_1 i_2 i_3} + \mathcal{A}_{i_3 i_4 i_5} - \hat{\alpha}_n - \widetilde{\alpha}_n). \tag{F.54}$$

Note that for each tuple (i_1, i_2, i_3), there are $\binom{n-1}{2} - 1$ different $x = (i_1, i_2, i_3, i_4, i_5)$ in S with the same (i_1, i_2, i_3). It follows

$$\sum_{x \in S} \mathcal{A}_{i_1 i_2 i_3} = \left(\binom{n-1}{2} - 1 \right) \sum_{\substack{i_1, i_2, i_3 (dist) \\ i_1 < i_2}} \mathcal{A}_{i_1 i_2 i_3} = \frac{n^2(n-1)(n-2)(n-3)}{4}\hat{\alpha}_n. \tag{F.55}$$

Similarly, we have

$$\sum_{x \in S} \mathcal{A}_{i_3 i_4 i_5} = \frac{n^2(n-1)(n-2)(n-3)}{4}\hat{\alpha}_n. \tag{F.56}$$

Inserting (F.55)-(F.56) into (F.54) gives

$$(Ib) = -\frac{n^2(n-1)(n-2)(n-3)}{4}(\widetilde{\alpha}_n - \hat{\alpha}_n)^2.$$

Combining this with (F.53) gives

$$(I) = (Ia) - \frac{n^2(n-1)(n-2)(n-3)}{4}(\widetilde{\alpha}_n - \hat{\alpha}_n)^2. \tag{F.57}$$

Next consider (II). Note that for any $x \in S_0 \setminus S$, $i_1 < i_2$ and $(i_1, i_2) = (i_4, i_5)$. By direct calculations

$$\sum_{x \in S_0 \setminus S} \mathcal{A}^*_{i_1 i_2 i_3} \mathcal{A}^*_{i_3 i_4 i_5} = \frac{1}{2} \sum_{i_1, i_2, i_3 (dist)} (\mathcal{A}^*_{i_1 i_2 i_3})^2 = \frac{1}{2} \sum_{i_1, i_2, i_3 (dist)} (\mathcal{A}^2_{i_1 i_2 i_3} - 2\hat{\alpha}_n \mathcal{A}_{i_1 i_2 i_3} + \hat{\alpha}_n^2). \tag{F.58}$$

Since $\mathcal{A}_{i_1 i_2 i_3} \in \{0, 1\}$, we have $\mathcal{A}^2_{i_1 i_2 i_3} = \mathcal{A}_{i_1 i_2 i_3}$. Combining this with definitions, the RHS of (F.58) reduces to

$$\frac{n(n-1)(n-2)}{2}\hat{\alpha}_n(1 - \hat{\alpha}_n).$$

It follows that

$$(II) = 0. \tag{F.59}$$

Combining (F.52), (F.57), and (F.59), it follows from (F.51) that

$$\psi_n = \frac{(Ia) - (1/4)n^2(n-1)(n-2)(n-3)(\widetilde{\alpha}_n - \hat{\alpha}_n)^2}{\sqrt{2n}\binom{n-1}{2}\hat{\alpha}_n(1 - \hat{\alpha}_n)}.$$

Now, by Lemma F.1, $|\hat{\alpha}_n - \widetilde{\alpha}_n| \le C \log(n)(\widetilde{\alpha}_n / n^3)^{1/2}$ except for a probability of $1 - O(1/n)$. It is seen that except for a probability of $1 - O(1/n)$

$$\left|\frac{\hat{\alpha}_n}{\widetilde{\alpha}_n} - 1\right| \le C\frac{\log(n)}{\sqrt{n^3 \widetilde{\alpha}_n}}, \qquad \left|\frac{(1/4)n^2(n-1)(n-2)(n-3)(\widetilde{\alpha}_n - \hat{\alpha}_n)^2}{\sqrt{2n}\binom{n-1}{2}\hat{\alpha}_n(1 - \hat{\alpha}_n)}\right| \le C\frac{\log^2(n)}{n^{1/2}}.$$

By $n^2 \widetilde{\alpha}_n \to \infty$, we have that in probability,

$$\frac{\hat{\alpha}_n}{\widetilde{\alpha}_n} \to 1, \qquad \frac{(1/4)n^2(n-1)(n-2)(n-3)(\widetilde{\alpha}_n - \hat{\alpha}_n)^2}{\sqrt{2n}\binom{n-1}{2}\hat{\alpha}_n(1 - \hat{\alpha}_n)} \to 0.$$

Let

$$Z_n = \frac{(Ia)}{\sqrt{2n}\binom{n-1}{2}\widetilde{\alpha}_n(1 - \widetilde{\alpha}_n)}.$$

To show (F.50), it is sufficient to show that as $n \to \infty$,

$$Z_n \to N(0, 1), \qquad \text{under the null}, \tag{F.60}$$

and

$$\mathbb{P}(|Z_n| > M) \to 1 \text{ for any } M > 0, \qquad \text{under the alternative}. \tag{F.61}$$

We now show (F.60)-(F.61). We consider (F.61) first since the proof is shorter. The following lemma is proved below.

Lemma F.2. *Under the conditions of Lemma 2.2, if the alternative hypothesis is true, then as $n \to \infty$*

$$\mathbb{E}[Z_n] \ge C n^{2.5} \widetilde{\alpha}_n \delta_n^2, \qquad Var(Z_n) \le C n^2 \widetilde{\alpha}_n.$$

Now, suppose the alternative hypothesis is true. Note that by triangle inequality

$$\mathbb{P}(|Z_n| \le M) \le \mathbb{P}\Big(\big|\mathbb{E}[Z_n]\big| - \big|Z_n - \mathbb{E}[Z_n]\big| \le M\Big) = \mathbb{P}(\big|Z_n - \mathbb{E}[Z_n]\big| \ge \big|\mathbb{E}[Z_n]\big| - M),$$

where by Chebyshev's inequality,

$$\mathbb{P}(\left|Z_n - \mathbb{E}[Z_n]\right| \geq \left|\mathbb{E}[Z_n]\right| - M) \leq \frac{\mathrm{Var}(Z_n)}{(\mathbb{E}[Z_n] - M)^2}.$$

At the same time, by Lemma F.2 and our assumptions of $n^2\widetilde{\alpha}_n \to \infty$ and $n^{3/2}\widetilde{\alpha}_n^{1/2}\delta_n^2 \to \infty$,

$$\frac{\mathrm{Var}(Z_n)}{(\mathbb{E}[Z_n] - M)^2} \leq \frac{Cn^2\widetilde{\alpha}_n}{(Cn^{2.5}\widetilde{\alpha}_n\delta_n^2 - M)^2} \leq \frac{1}{C(n^{3/2}\widetilde{\alpha}_n^{1/2}\delta_n^2)^2} \to 0.$$

Combining these proves (F.61).

We now consider (F.60). For $1 \leq m \leq n$, introduce a subset of S by

$$S^{(m)} = \{x = (i_1, i_2, i_3, i_4, i_5) \in S : \max\{i_1, i_2, i_3, i_4, i_5\} \leq m\}.$$

Introduce

$$\widetilde{T}_{n,m} = \sum_{x \in S^{(m)}} \widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}, \qquad Z_{n,m} = \frac{\widetilde{T}_{n,m}}{\sqrt{2n\binom{n-1}{2}\widetilde{\alpha}_n(1-\widetilde{\alpha}_n)}}, \qquad (\widetilde{T}_{n,0} = Z_{n,0} = 0),$$

and

$$X_{n,m} = Z_{n,m} - Z_{n,m-1}.$$

It is seen that

$$(Ia) = \widetilde{T}_{n,n}, \qquad \text{and} \qquad Z_n = Z_{n,n} = \sum_{m=1}^{n} X_{n,m}. \tag{F.62}$$

Consider the filtration $\{\mathcal{F}_{n,m}\}_{1 \leq m \leq n}$ with $\mathcal{F}_{n,m} = \sigma\Big(\{\widetilde{\mathcal{A}}_{i_1i_2i_3} : 1 \leq i_1, i_2, i_3 \leq m\}\Big)$. It is seen that for all $1 \leq m \leq n$,

$$\mathbb{E}[X_{n,m}|\mathcal{F}_{n,m-1}] = \mathbb{E}[Z_{n,m}|\mathcal{F}_{n,m-1}] - Z_{n,m-1} = 0,$$

so $\{X_{n,m}\}_{m=1}^n$ is a martingale difference sequence with respect to $\{\mathcal{F}_{n,m}\}_{1 \leq m \leq n}$. We have the following lemma which is proved below.

Lemma F.3. *Under the conditions of Lemma 2.2, if the null hypothesis is true, then as $n \to \infty$,*

$$(a) \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2|\mathcal{F}_{n,m-1}] \to 1, \qquad \text{in probability },$$

$$(b) \forall \epsilon > 0, \quad \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 \mathbb{I}\{|X_{n,m}| > \epsilon\}|\mathcal{F}_{n,m-1}] \to 0, \qquad \text{in probability }.$$

By Lemma F.3 and [3, Corollary 3.1], it follows from (F.62) that under the null,

$$Z_n = Z_{n,n} \to N(0,1).$$

This proves (F.60).

F.1 Proof of Lemma F.1

We first prove the claim (b). By definitions

$$\widetilde{\alpha}_n = \mathbb{E}[\hat{\alpha}_n] = \frac{\sum_{i_1,i_2,i_3(dist)} \mathcal{Q}_{i_1i_2i_3}}{n(n-1)(n-2)}.$$

Recall that under alternative

$$\mathcal{Q}_{i_1i_2i_3} = \sum_{1\le k_1,k_2,k_3\le K} \pi_{i_1}(k_1)\pi_{i_2}(k_2)\pi_{i_3}(k_3)\mathcal{P}_{k_1k_2k_3}, \qquad 1\le i_1,i_2,i_3\le n.$$

It is seen that $\mathcal{Q}_{i_1i_2i_3} \le \max_{1\le k_1,k_2,k_3\le K}\{\mathcal{P}_{k_1k_2k_3}\}$, $1\le i_1,i_2,i_3\le n$ and so

$$\widetilde{\alpha}_n \le \max_{1\le k_1,k_2,k_3\le K}\{\mathcal{P}_{k_1k_2k_3}\}, \qquad \widetilde{\alpha}_n = h'(\mathcal{P}h)h + O(\frac{\max_{1\le k_1,k_2,k_3\le K}\{\mathcal{P}_{k_1k_2k_3}\}}{n}).$$

At the same time, by our assumption $\min_{k=1}^K\{h_k\}\ge c_0$ and elementary calculations

$$\max_{1\le k_1,k_2,k_3\le K}\{\mathcal{P}_{k_1k_2k_3}\} \le C\sum_{1\le k_1,k_2,k_3\le K} h_{k_1}h_{k_2}h_{k_3}\mathcal{P}_{k_1k_2k_3} \le C\widetilde{\alpha}_n.$$

These prove the claims in (b). Now we show the claim (a).

Note that, $\hat{\alpha}_n$ is the average of $\binom{n}{3}$ independent Bernoulli random variables with parameter bounded by $C\widetilde{\alpha}_n$ under both null and alternative hypothesis. By Bernstein's inequality,

$$\mathbb{P}\Big(\big(\tbinom{n}{3}\big)|\hat{\alpha}_n - \widetilde{\alpha}_n| \ge t\Big) \le 2\exp\big(-\frac{t^2}{\binom{n}{3}C\widetilde{\alpha}_n(1-C\widetilde{\alpha}_n) + \frac{t}{3}}\big).$$

Let $t = C\binom{n}{3}\frac{\log(n)\widetilde{\alpha}_n^{1/2}}{n^{3/2}}$, by elementary calculations, we get

$$\mathbb{P}\Big(|\hat{\alpha}_n - \widetilde{\alpha}_n| \ge C\frac{\log(n)\widetilde{\alpha}_n^{1/2}}{n^{3/2}}\Big) \le O(1/n). \tag{F.63}$$

This is equivalent to the claim in (a).

F.2 Proof of Lemma F.2

Recall that

$$Z_n = (2n)^{-1/2}\frac{(Ia)}{\binom{n-1}{2}\widetilde{\alpha}_n(1-\widetilde{\alpha}_n)}, \qquad \text{with } (Ia) = \textstyle\sum_{x\in S}(\mathcal{A}_{i_1i_2i_3} - \widetilde{\alpha}_n)(\mathcal{A}_{i_3i_4i_5} - \widetilde{\alpha}_n).$$

Therefore, to show the claims, it is sufficient to show that as $n\to\infty$

$$\mathbb{E}[(Ia)] \ge Cn^5\widetilde{\alpha}_n^2\delta_n^2, \tag{F.64}$$

and

$$\mathrm{Var}((Ia)) \le Cn^7\widetilde{\alpha}_n^3. \tag{F.65}$$

Consider (F.64) first. Since for each $x = (i_1,i_2,i_3,i_4,i_5)\in S$, $\mathcal{A}_{i_1i_2i_3}$ is independent of $\mathcal{A}_{i_3i_4i_5}$, by direct calculations,

$$\mathbb{E}[(Ia)] = \sum_{x\in S}(\mathcal{Q}_{i_1i_2i_3} - \widetilde{\alpha}_n)(\mathcal{Q}_{i_3i_4i_5} - \widetilde{\alpha}_n).$$

Let $\widetilde{Q}_{i_1i_2i_3} = \mathcal{Q}_{i_1i_2i_3} - \widetilde{\alpha}_n$, by definitions,

$$\mathbb{E}[(Ia)] = \frac{1}{4}\big(\sum_{x\in S_0'}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5} - \sum_{x\in(S_0'\setminus S_1')}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}\big),$$

where

$$\begin{aligned} S_0' &= \{x : 1\le i_1,i_2,i_3,i_4,i_5\le n\} \\ S_1' &= \{x\in S_0' : i_1,i_2,i_3(dist); i_3,i_4,i_5(dist); (i_1,i_2)\ne(i_4,i_5)\}. \end{aligned}$$

To show (F.64), it is sufficient to show that

$$\sum_{x\in S_0'} \widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5} \geq Cn^5\widetilde{\alpha}_n^2\delta_n^2, \qquad \text{and} \qquad \sum_{x\in(S_0'\setminus S_1')} \widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5} = o(\sum_{x\in S_0'} \widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}). \tag{F.66}$$

Consider the first claim in (F.66). Recall that

$$\widetilde{Q}_{i_1i_2i_3} = \mathcal{Q}_{i_1i_2i_3} - \widetilde{\alpha}_n = \sum_{k_1,k_2,k_3} \pi_{i_1}(k_1)\pi_{i_2}(k_2)\pi_{i_3}(k_3)\mathcal{P}_{k_1k_2k_3} - \widetilde{\alpha}_n, \qquad \text{and} \qquad h = \sum_{i=1}^n \pi_i/n.$$

By direct calculations and elementary algebra,

$$\sum_{x\in S_0'} \widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5} = n^4\|\Pi(\mathcal{P}h)h - \widetilde{\alpha}_n\mathbf{1}_n\|^2.$$

By triangle inequality, we have $\|\Pi(\mathcal{P}h)h - \widetilde{\alpha}_n\mathbf{1}_n\| \geq \big|\|\Pi(\mathcal{P}h)h - h'(\mathcal{P}h)h\mathbf{1}_n\| - \|(h'(\mathcal{P}h)h - \widetilde{\alpha}_n)\mathbf{1}_n\|\big|$. It follows that

$$\sum_{x\in S_0'} \widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5} \geq n^4(\|\Pi(\mathcal{P}h)h - h'(\mathcal{P}h)h\mathbf{1}_n\| - \|(h'(\mathcal{P}h)h - \widetilde{\alpha}_n)\mathbf{1}_n\|)^2. \tag{F.67}$$

Recall that $\Sigma = \Pi'\Pi/n - hh'$ and note that $\Sigma\mathbf{1}_K = 0$. Also, recall that $H_K = K^{-1}\mathbf{1}_K\mathbf{1}_K'$ and note that $I_K - H_K$ is a projection matrix. By elementary algebra,

$$\Sigma = (I_K - H_K)\Sigma(I_K - H_K).$$

First, by elementary algebra,

$$\|\Pi(\mathcal{P}h)h - h'(\mathcal{P}h)h\mathbf{1}_n\|^2 = n\big(h'(\mathcal{P}h)\frac{\Pi'\Pi}{n}(\mathcal{P}h)h - h'(\mathcal{P}h)hh'(\mathcal{P}h)h\big) = n((\mathcal{P}h)h)'\Sigma((\mathcal{P}h)h), \tag{F.68}$$

where the RHS equals to

$$n((\mathcal{P}h)h)'(I_K - H_K)\Sigma(I_K - H_K)(\mathcal{P}h)h. \tag{F.69}$$

By our assumption $\lambda_{K-1}(\Sigma) = \min_{\|v\|=1, v\perp\mathbf{1}_K} v'\Sigma v \geq c_0$, it is seen that

$$n((\mathcal{P}h)h)'(I_K - H_K)\Sigma(I_K - H_K)(\mathcal{P}h)h \geq c_0n\widetilde{\alpha}_n^2\|\widetilde{\alpha}_n^{-1}(I_K - H_K)(\mathcal{P}h)h\|^2. \tag{F.70}$$

Recall that $\delta_n = \|\widetilde{\alpha}_n^{-1}(I_K - H_K)(\mathcal{P}h)h\|$, combining with (F.68)-(F.70), we get

$$\|\Pi(\mathcal{P}h)h - h'(\mathcal{P}h)h\mathbf{1}_n\|^2 \geq c_0n\widetilde{\alpha}_n^2\delta_n^2. \tag{F.71}$$

At the same time, by Lemma F.1,

$$\widetilde{\alpha}_n = h'(\mathcal{P}h)h + O(\frac{\widetilde{\alpha}_n}{n}). \tag{F.72}$$

By direct calculations,

$$\|(h'(\mathcal{P}h)h - \widetilde{\alpha}_n)\mathbf{1}_n\|^2 = n(h'(\mathcal{P}h)h - \widetilde{\alpha}_n)^2 = O(\frac{\widetilde{\alpha}_n^2}{n}), \tag{F.73}$$

where by $\widetilde{\alpha}_n \leq \max_{1\leq i_1,i_2,i_3\leq n}\{\mathcal{P}_{i_1i_2i_3}\} \leq c_0$ and our condition $n^{3/2}\widetilde{\alpha}_n^{1/2}\delta_n^2 \to \infty$,

$$\frac{\widetilde{\alpha}_n^2}{n} = o(1)\cdot(n\widetilde{\alpha}_n^2\delta_n^2). \tag{F.74}$$

Combining (F.72)-(F.74),

$$\|(h'(\mathcal{P}h)h - \widetilde{\alpha}_n)\mathbf{1}_n\|^2 = o(n\widetilde{\alpha}_n^2\delta_n^2). \tag{F.75}$$

Inserting (F.71) and (F.75) into (F.67) proves the first claim in (F.66).

Next, we consider the second claim in (F.66). Notice that by symmetry, the two leading terms of $\sum_{x\in(S_0'\setminus S_1')}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}$ are the following:

$$O(\sum_{\substack{1\leq i_1,i_2,i_3,i_4,i_5\leq n\\ i_3=i_4}}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}), \qquad \text{and} \qquad O(\sum_{\substack{1\leq i_1,i_2,i_3,i_4,i_5\leq n\\ i_4=i_5}}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}). \tag{F.76}$$

The other terms are $O(n^3\widetilde{\alpha}_n^2) = o(n^5\widetilde{\alpha}_n^2\delta_n^2)$ and thus are negligible. It is therefore adequate to consider the two terms in (F.76).

Consider the first term in (F.76). By Cauchy-Schwarz inequality,

$$|\sum_{\substack{1\leq i_1,i_2,i_3,i_4,i_5\leq n\\ i_3=i_4}}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}| \leq \sqrt{\sum_{1\leq i_3\leq n}(\sum_{1\leq i_5\leq n}\widetilde{Q}_{i_3i_3i_5})^2}\sqrt{\sum_{1\leq i_3\leq n}(\sum_{1\leq i_1,i_2\leq n}\widetilde{Q}_{i_1i_2i_3})^2}. \tag{F.77}$$

Note that by definitions and Lemma F.1, $|\widetilde{Q}_{i_3i_3i_5}| \leq C\widetilde{\alpha}_n$. It is seen that

$$\sum_{1\leq i_3\leq n}(\sum_{1\leq i_5\leq n}\widetilde{Q}_{i_3i_3i_5})^2 \leq Cn^3\widetilde{\alpha}_n^2. \tag{F.78}$$

By our condition $n^{3/2}\widetilde{\alpha}_n^{1/2}\delta_n^2 \to \infty$, we have $n^2\delta_n^2 \to \infty$. Comparing the RHS of (F.78) with the first claim of (F.66), the RHS is at a smaller order of $\sum_{x\in S_0'}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}$. At the same time,

$$\sum_{1\leq i_3\leq n}(\sum_{1\leq i_1,i_2\leq n}\widetilde{Q}_{i_1i_2i_3})^2 = \sum_{x\in S_0'}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}. \tag{F.79}$$

Inserting (F.78)-(F.79) into (F.77), we have

$$|\sum_{\substack{1\leq i_1,i_2,i_3,i_4,i_5\leq n\\ i_3=i_4}}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}| = o(\sum_{x\in S_0'}\widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}).$$

For the second term in (F.76), the analysis is similar, so we omit the details. These prove the second claim of (F.66), and so complete the proof of (F.64).

Next we consider (F.65). Let $\mathcal{W}$ be the tensor with the same size as $\mathcal{A}$, where $\mathcal{W}_{i_1i_2i_3} = \mathcal{A}_{i_1i_2i_3} - \mathcal{Q}_{i_1i_2i_3}$ if i_1, i_2, i_3 are distinct, and $\mathcal{W}_{i_1i_2i_3} = 0$ otherwise. By symmetry and definitions,

$$(Ia) = \sum_{x\in S}(\mathcal{W}_{i_1i_2i_3} - \widetilde{Q}_{i_1i_2i_3})(\mathcal{W}_{i_3i_4i_5} - \widetilde{Q}_{i_3i_4i_5}) = \sum_{x\in S}(\mathcal{W}_{i_1i_2i_3}\mathcal{W}_{i_3i_4i_5} - 2\widetilde{Q}_{i_1i_2i_3}\mathcal{W}_{i_3i_4i_5} + \widetilde{Q}_{i_1i_2i_3}\widetilde{Q}_{i_3i_4i_5}). \tag{F.80}$$

Since for any random variables X and Y, $\mathrm{Var}(X+Y) \leq 2\mathrm{Var}(X) + 2\mathrm{Var}(Y)$, we have

$$\mathrm{Var}((Ia)) \leq 2\mathrm{Var}(\sum_{x\in S}\mathcal{W}_{i_1i_2i_3}\mathcal{W}_{i_3i_4i_5}) + 2\mathrm{Var}(\sum_{x\in S}2\widetilde{Q}_{i_1i_2i_3}\mathcal{W}_{i_3i_4i_5}).$$

Here, we note that $\widetilde{Q}$ is non-random, so the variance of the last term in (F.80) is 0. By direct calculations,

$$\mathrm{Var}(\sum_{x\in S}\mathcal{W}_{i_1i_2i_3}\mathcal{W}_{i_3i_4i_5}) = \sum_{x\in S}\mathrm{Var}(\mathcal{W}_{i_1i_2i_3}\mathcal{W}_{i_3i_4i_5}) = O(n^5\widetilde{\alpha}_n^2),$$

$$\mathrm{Var}(\sum_{x\in S}2\widetilde{Q}_{i_1i_2i_3}\mathcal{W}_{i_3i_4i_5}) = \frac{1}{4}\sum_{i_3i_4i_5(dist)}(\sum_{\substack{i_1i_2(dist)\\ \{i_1,i_2\}\neq\{i_4,i_5\}\\ i_1,i_2\neq i_3}}\mathcal{Q}_{i_1i_2i_3})^2\mathrm{Var}(W_{i_3i_4i_5}) = O(n^7\widetilde{\alpha}_n^3).$$

By our assumptions, $n^2\widetilde{\alpha}_n \to \infty$, and so $n^5\widetilde{\alpha}_n = o(1)\cdot n^7\widetilde{\alpha}_n^3$. Combining these gives that

$$\mathrm{Var}((Ia)) \leq Cn^7\widetilde{\alpha}_n^3.$$

This proves (F.65).

F.3 Proof of Lemma F.3

We first show claim (a). By Chebyshev's inequality, it is sufficient to show that

$$\mathbb{E}\Big[\sum_{m=1}^n \mathbb{E}[X_{n,m}^2|\mathcal{F}_{n,m-1}]\Big] \to 1, \qquad \mathrm{Var}(\sum_{m=1}^n \mathbb{E}[X_{n,m}^2|\mathcal{F}_{n,m-1}]) \to 0. \tag{F.81}$$

Introduce

$$T^{(m)} = \mathbb{E}[(\sum_{x\in S^{(m)}\backslash S^{(m-1)}} \widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5})^2|\mathcal{F}_{n,m-1}].$$

By definitions,

$$\mathbb{E}[X_{n,m}^2|\mathcal{F}_{n,m-1}] = \frac{\mathbb{E}[(\sum_{x\in S^{(m)}\backslash S^{(m-1)}} \widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5})^2|\mathcal{F}_{n,m-1}]}{(\sqrt{2n}\binom{n-1}{2}\alpha_n(1-\alpha_n))^2} = \frac{T^{(m)}}{(\sqrt{2n}\binom{n-1}{2}\alpha_n(1-\alpha_n))^2}.$$

To show (F.81), it is sufficient to show that

$$\mathbb{E}[\sum_{m=1}^n T^{(m)}] = \frac{n^5\alpha_n^2(1-\alpha_n)^2}{2}(1+o(1)), \tag{F.82}$$

and that

$$\mathrm{Var}(\sum_{m=1}^n T^{(m)}) = o(n^{10}\alpha_n^4). \tag{F.83}$$

Consider (F.82) first. Recall that $S^{(m)} = \{x = (i_1,i_2,i_3,i_4,i_5) \in S : \max\{i_1,i_2,i_3,i_4,i_5\} \leq m\}$ and $x = (i_1,i_2,i_3,i_4,i_5)$ for short. Similarly, for short, we write $x' = (i_1',i_2',i_3',i_4',i_5')$ and let

$$(S^{(m)}\backslash S^{(m-1)})^2 = \{(x,x') : x \in S^{(m)}\backslash S^{(m-1)}, x' \in S^{(m)}\backslash S^{(m-1)}\}.$$

Let

$$\begin{aligned} SS_1^{(m)} =& \{(x,x') \in (S^{(m)}\backslash S^{(m-1)})^2 : i_3 = i_3', \{i_1,i_2,i_4,i_5\} = \{i_1',i_2',i_4',i_5'\}\}, \\ SS_2^{(m)} =& (S^{(m)}\backslash S^{(m-1)})^2\backslash SS_1^{(m)}. \end{aligned}$$

It is seen that the LHS of (F.82) equals to

$$(I) + (II),$$

where

$$(I) = \mathbb{E}\Big[\sum_{m=1}^n \mathbb{E}[\sum_{(x,x')\in SS_1^{(m)}} \widetilde{\mathcal{A}}_{i_1i_2i_3}^2\widetilde{\mathcal{A}}_{i_3i_4i_5}^2|\mathcal{F}_{n,m-1}]\Big],$$

and

$$(II) = \mathbb{E}\Big[\sum_{m=1}^n \mathbb{E}[\sum_{(x,x')\in SS_2^{(m)}} \widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'i_5'}|\mathcal{F}_{n,m-1}]\Big].$$

Notice that for any $(x,x') \in SS_2^m$, each $\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'i_5'}$ is a mean-zero random variable. It follows that

$$(II) = 0.$$

At the same time, note that for any $(x,x') \in SS_1^{(m)}$ (where $x = (i_1,i_2,i_3,i_4,i_5)$ and $x' = (i_1',i_2',i_3',i_4',i_5')$), there are two possibilities: $(i_1,i_2,i_4,i_5) = (i_1',i_2',i_4',i_5')$ and $(i_1,i_2,i_4,i_5) = (i_4',i_5',i_1',i_2')$. By symmetry,

$$(I) = 2\sum_{m=1}^{n}\sum_{x\in S^{(m)}\backslash S^{(m-1)}}\mathbb{E}\Big[\widetilde{\mathcal{A}}_{i_1i_2i_3}^2\widetilde{\mathcal{A}}_{i_3i_4i_5}^2\Big] = 2\sum_{x\in S}\alpha_n^2(1-\alpha_n)^2 = 12n\binom{n}{4}\alpha_n^2(1-\alpha_n^2).$$

Combining these gives (F.82).

Next, consider (F.83). In $S^{(m)}\backslash S^{(m-1)}$, we have $i_3 = m$ or $i_2 = m$ or $i_5 = m$. Let

$$\begin{aligned}S_1^{(m)} =&\big\{x\in S^{(m)}\backslash S^{(m-1)} : \text{either } i_2 = m, i_5 < m \text{ or } i_5 = m, i_2 < m\big\},\\ S_2^{(m)} =&(S^{(m)}\backslash S^{(m-1)})\backslash S_1^{(m)}.\end{aligned}$$

Write

$$T^{(m)} = T_1^{(m)} + 2T_2^{(m)} + T_3^{(m)},$$

where

$$\begin{aligned}T_1^{(m)} =&\mathbb{E}[\sum_{x,x'\in S_1^{(m)}}\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'i_5'}|\mathcal{F}_{n,m-1}],\\ T_2^{(m)} =&\mathbb{E}[\sum_{x\in S_1^{(m)},x'\in S_2^{(m)}}\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'i_5'}|\mathcal{F}_{n,m-1}],\\ T_3^{(m)} =&\mathbb{E}[\sum_{x,x'\in S_2^{(m)}}\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'i_5'}|\mathcal{F}_{n,m-1}].\end{aligned}$$

Notice that for $x\in S_1^{(m)}, x'\in S_2^{(m)}$, $\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'i_5'}$ is mean-zero conditional on $\mathcal{F}_{n,m-1}$. It follows directly that

$$T_2^{(m)} = 0.$$

Also, by definitions, for each $x\in S_2^{(m)}$, we must have $i_3 = m$ or $i_2 = i_5 = m$. Let $E_m = \big\{(x,x')\in S_2^{(m)}\times S_2^{(m)} : \{i_1,i_2,i_3,i_4,i_5\} = \{i_1',i_2',i_3',i_4',i_5'\}\big\}$, by direct calculations

$$T_3^{(m)} = |E_m|\alpha_n^2(1-\alpha_n)^2.$$

It is seen that $T_3^{(m)}$ is non-random. Therefore,

$$T^{(m)} = T_1^{(m)} + |E_m|\alpha_n^2(1-\alpha_n)^2, \qquad \text{and} \qquad \mathrm{Var}(\sum_{m=1}^{n}T^{(m)}) = \mathrm{Var}(\sum_{m=1}^{n}T_1^{(m)}),$$

and to show (F.83), it is sufficient to show that

$$\mathrm{Var}(\sum_{m=1}^{n}T_1^{(m)}) = o(n^{10}\alpha_n^4). \tag{F.84}$$

By definitions and symmetry

$$T_1^{(m)} = \mathbb{E}[4\sum_{\substack{1\le i_1,i_2,i_3,i_4,i_1',i_2',i_3',i_4'\le m-1\\ i_1<i_2;i_1'<i_2'\\ i_1,i_2,i_4\neq i_3;i_1',i_2',i_4'\neq i_3'}}\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4m}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'm}|\mathcal{F}_{n,m-1}].$$

If $\{i_3,i_4\}\neq\{i_3',i_4'\}$, then $\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4m}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'm}$ has a conditional mean of zero. Therefore, we have

$$T_1^{(m)} = T_{11}^{(m)} + T_{12}^{(m)},$$

where

$$T_{11}^{(m)} = \mathbb{E}[4 \sum_{\substack{1 \le i_1, i_2, i_3, i_4, i_1', i_2' \le m-1 \\ i_1 < i_2; i_1' < i_2' \\ i_1, i_2, i_1', i_2', i_4 \neq i_3}} \widetilde{\mathcal{A}}_{i_1 i_2 i_3} \widetilde{\mathcal{A}}_{i_3 i_4 m}^2 \widetilde{\mathcal{A}}_{i_1' i_2' i_3} | \mathcal{F}_{n,m-1}],$$

$$T_{12}^{(m)} = \mathbb{E}[4 \sum_{\substack{1 \le i_1, i_2, i_3, i_4, i_1', i_2' \le m-1 \\ i_1 < i_2; i_1' < i_2' \\ i_1, i_2, i_4 \neq i_3; i_1' i_2' \neq i_4}} \widetilde{\mathcal{A}}_{i_1 i_2 i_3} \widetilde{\mathcal{A}}_{i_3 i_4 m}^2 \widetilde{\mathcal{A}}_{i_1' i_2' i_4} | \mathcal{F}_{n,m-1}].$$

Since for any random variables X and Y, $\mathrm{Var}(X+Y) \le 2\mathrm{Var}(X) + 2\mathrm{Var}(Y)$, to show (F.84), it is sufficient to show that

$$\mathrm{Var}(\sum_{m=1}^{n} T_{11}^{(m)}) = o(n^{10}\alpha_n^4), \qquad \text{and} \qquad \mathrm{Var}(\sum_{m=1}^{n} T_{12}^{(m)}) = o(n^{10}\alpha_n^4). \tag{F.85}$$

Consider the first claim in (F.85). Recall that

$$\widetilde{T}_{n,m} = \sum_{x \in S^{(m)}} \widetilde{\mathcal{A}}_{i_1 i_2 i_3} \widetilde{\mathcal{A}}_{i_3 i_4 i_5} = \sum_{\substack{1 \le i_1, \cdots, i_5 \le m \\ i_1 < i_2; i_4 < i_5 \\ i_1, i_2, i_4, i_5 \neq i_3 \\ (i_1, i_2) \neq (i_4, i_5)}} \widetilde{\mathcal{A}}_{i_1 i_2 i_3} \widetilde{\mathcal{A}}_{i_3 i_4 i_5}.$$

By elementary calculations

$$T_{11}^{(m)} = 4(m-2)\alpha_n(1-\alpha_n)\widetilde{T}_{n,m-1} + 4(m-2)\alpha_n(1-\alpha_n) \sum_{\substack{1 \le i_1, i_2, i_3 \le m-1 \\ i_1 < i_2 \\ i_1, i_2 \neq i_3}} \widetilde{\mathcal{A}}_{i_1 i_2 i_3}^2.$$

By inequality $\mathrm{Var}(X+Y) \le 2\mathrm{Var}(X) + 2\mathrm{Var}(Y)$, to show the first claim in (F.85), it is sufficient to show that

$$\mathrm{Var}(\sum_{m=1}^{n} 4(m-2)\alpha_n(1-\alpha_n)\widetilde{T}_{n,m-1}) = o(n^{10}\alpha_n^4), \tag{F.86}$$

and

$$\mathrm{Var}(\sum_{m=1}^{n} 4(m-2)\alpha_n(1-\alpha_n) \sum_{\substack{1 \le i_1, i_2, i_3 \le m-1 \\ i_1 < i_2 \\ i_1, i_2 \neq i_3}} \widetilde{\mathcal{A}}_{i_1 i_2 i_3}^2) = o(n^{10}\alpha_n^4). \tag{F.87}$$

Consider the LHS of (F.86), by definitions,

$$\mathrm{Var}(\sum_{m=1}^{n} 4(m-2)\alpha_n(1-\alpha_n)\widetilde{T}_{n,m-1}) = \sum_{m,m'=1}^{n} 16(m-2)(m'-2)\alpha_n^2(1-\alpha_n)^2 \mathrm{Cov}(\widetilde{T}_{n,m-1}, \widetilde{T}_{n,m'-1}). \tag{F.88}$$

Notice that

$$\mathrm{Cov}(\widetilde{T}_{n,m-1}, \widetilde{T}_{n,m'-1}) = \sum_{\substack{1 \le i_1, \cdots, i_5 \le m \\ i_1 < i_2; i_4 < i_5 \\ i_1, i_2, i_4, i_5 \neq i_3 \\ (i_1, i_2) \neq (i_4, i_5)}} \sum_{\substack{1 \le i_1', \cdots, i_5' \le m \\ i_1' < i_2'; i_4' < i_5' \\ i_1', i_2', i_4', i_5' \neq i_3' \\ (i_1', i_2') \neq (i_4', i_5')}} \mathbb{E}[\widetilde{\mathcal{A}}_{i_1 i_2 i_3} \widetilde{\mathcal{A}}_{i_3 i_4 i_5} \widetilde{\mathcal{A}}_{i_1' i_2' i_3'} \widetilde{\mathcal{A}}_{i_3' i_4' i_5'}].$$

Only if $\{i_1, i_2, i_3, i_4, i_5\} = \{i_1', i_2', i_3', i_4', i_5'\}$, $\mathbb{E}[\widetilde{\mathcal{A}}_{i_1 i_2 i_3} \widetilde{\mathcal{A}}_{i_3 i_4 i_5} \widetilde{\mathcal{A}}_{i_1' i_2' i_3'} \widetilde{\mathcal{A}}_{i_3' i_4' i_5'}]$ will be non-zero. Since there are only a bounded number of ways to pair the indexes, by direct calculations

$$\mathrm{Cov}(\widetilde{T}_{n,m-1}, \widetilde{T}_{n,m'-1}) = O(\sum_{\substack{1 \le i_1, \cdots, i_5 \le m \\ i_1 < i_2; i_4 < i_5 \\ i_1, i_2, i_4, i_5 \neq i_3 \\ (i_1, i_2) \neq (i_4, i_5)}} \mathbb{E}[(\widetilde{\mathcal{A}}_{i_1 i_2 i_3}^2 \widetilde{\mathcal{A}}_{i_3 i_4 i_5}^2]) = O(n^5\alpha_n^2).$$

Combining this with (F.88), it is seen that

$$\text{Var}(\sum_{m=1}^{n} 4(m-2)\alpha_n(1-\alpha_n)\widetilde{T}_{n,m-1}) = O(n^4 n^5 \alpha_n^4) = o(n^{10}\alpha_n^4).$$

This proves (F.86).

Next consider the LHS of (F.87), by direct calculations,

$$\begin{aligned}\text{Var}(\sum_{m=1}^{n} 4(m-2)\alpha_n(1-\alpha_n) \sum_{\substack{1\le i_1,i_2,i_3\le m-1\\ i_1<i_2\\ i_1,i_2\neq i_3}} \widetilde{\mathcal{A}}_{i_1i_2i_3}^2) \le & 16n^4\alpha_n^2(1-\alpha_n)^2 \text{Var}(\sum_{\substack{1\le i_1,i_2,i_3\le n\\ i_1<i_2\\ i_1,i_2\neq i_3}} \widetilde{\mathcal{A}}_{i_1i_2i_3}^2)\\ =& 16n^4\alpha_n^2(1-\alpha_n)^2\cdot \sum_{\substack{1\le i_1,i_2,i_3\le n\\ i_1<i_2\\ i_1,i_2\neq i_3}} 3\cdot \text{Var}(\widetilde{\mathcal{A}}_{i_1i_2i_3}^2)\\ =& O(n^7\alpha_n^3).\end{aligned}$$

By our assumption $n^2\widetilde{\alpha}_n \to \infty$ (i.e., $n^2\alpha_n \to \infty$), the RHS of the above inequality is $o(n^{10}\alpha_n^4)$. This proves (F.87) and completes the first claim of (F.85).

Next consider the second claim in (F.85), by definitions,

$$\text{Var}(\sum_{m=1}^{n} T_{12}^{(m)}) = \sum_{m,m'=1}^{n} 16\alpha_n^2(1-\alpha_n)^2 \sum_{\substack{1\le i_1,\cdots,i_6\le m\\ i_1<i_2;i_4<i_5\\ i_1,i_2\neq i_3;i_4,i_5\neq i_6\\ i_3\neq i_6}} \sum_{\substack{1\le i_1',\cdots,i_6'\le m\\ i_1'<i_2';i_4'<i_5'\\ i_1',i_2'\neq i_3';i_4',i_5'\neq i_6'\\ i_3'\neq i_6'}} \mathbb{E}[\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_4i_5i_6}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_4'i_5'i_6'}].$$

Similarly, it is sufficient to consider terms that satisfy $\{i_1,\cdots,i_6\} = \{i_1',\cdots,i_6'\}$, hence

$$\text{Var}(\sum_{m=1}^{n} T_{12}^{(m)}) = O(\sum_{m,m'=1}^{n} 16\alpha_n^2(1-\alpha_n)^2 \sum_{\substack{1\le i_1,\cdots,i_6\le m\\ i_1<i_2;i_4<i_5\\ i_1,i_2\neq i_3;i_4,i_5\neq i_6\\ i_3\neq i_6}} \mathbb{E}[\widetilde{\mathcal{A}}_{i_1i_2i_3}^2\widetilde{\mathcal{A}}_{i_4i_5i_6}^2]) = O(n^8\alpha_n^4).$$

Note that the RHS above is $o(n^{10}\alpha_n^4)$. This proves the second claim in (F.85) and completes the proof of claim (a) of (F.81).

Now we consider the claim (b), where the goal is to show that

$$\forall\epsilon>0, \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2\mathbb{I}\{|X_{n,m}|>\epsilon\}|\mathcal{F}_{n,m-1}] \to 0, \qquad \text{in probability.} \tag{F.89}$$

By Cauchy-Schwarz inequality

$$\big|\sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2\mathbb{I}\{|X_{n,m}|>\epsilon\}|\mathcal{F}_{n,m-1}]\big| \le \sum_{m=1}^{n} \sqrt{\mathbb{E}[X_{n,m}^4|\mathcal{F}_{n,m-1}]}\sqrt{\mathbb{P}(|X_{n,m}|>\epsilon|\mathcal{F}_{n,m-1})}. \tag{F.90}$$

At the same time, by Markov's inequality,

$$\sqrt{\mathbb{P}(|X_{n,m}|>\epsilon|\mathcal{F}_{n,m-1})} \le \sqrt{\mathbb{E}[X_{n,m}^4|\mathcal{F}_{n,m-1}]/\epsilon^4}. \tag{F.91}$$

Combining (F.90) and (F.91) gives

$$\big|\sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2\mathbb{I}\{|X_{n,m}|>\epsilon\}|\mathcal{F}_{n,m-1}]\big| \le \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^4|\mathcal{F}_{n,m-1}]/\epsilon^2.$$

To show (F.89), by Markov's inequality, it is sufficient to show that

$$\mathbb{E}\Big[\sum_{m=1}^{n}\mathbb{E}[X_{n,m}^4|\mathcal{F}_{n,m-1}]\Big]\to 0. \tag{F.92}$$

Recall that

$$X_{n,m}=\frac{\sum_{x\in S^{(m)}\backslash S^{(m-1)}}\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}}{\sqrt{2n}\binom{n-1}{2}\widetilde{\alpha}_n(1-\widetilde{\alpha}_n)}.$$

Write for short $y=(i_1,i_2,i_3,i_4,i_5,j_1,j_2,j_3,j_4,j_5)$, similarly, $y'=(i_1',i_2',i_3',i_4',i_5',j_1',j_2',j_3',j_4',j_5')$. To show (F.92), it is sufficient to show that

$$\mathbb{E}[\sum_{m=1}^{n}\sum_{y,y'\in(S^{(m)}\backslash S^{(m-1)})^2}\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}\widetilde{\mathcal{A}}_{j_1j_2j_3}\widetilde{\mathcal{A}}_{j_3j_4j_5}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'i_5'}\widetilde{\mathcal{A}}_{j_1'j_2'j_3'}\widetilde{\mathcal{A}}_{j_3'j_4'j_5'}]=o(n^{10}\alpha_n^4).$$

Similarly, to have non-zero expected value, $\widetilde{\mathcal{A}}_{i_1i_2i_3}\widetilde{\mathcal{A}}_{i_3i_4i_5}\widetilde{\mathcal{A}}_{j_1j_2j_3}\widetilde{\mathcal{A}}_{j_3j_4j_5}\widetilde{\mathcal{A}}_{i_1'i_2'i_3'}\widetilde{\mathcal{A}}_{i_3'i_4'i_5'}\widetilde{\mathcal{A}}_{j_1'j_2'j_3'}\widetilde{\mathcal{A}}_{j_3'j_4'j_5'}$ must be in quadratic form. Since there are only a bounded number of ways to pair them into quadratic forms, it is sufficient to show that

$$\sum_{m=1}^{n}\sum_{y\in(S^{(m)}\backslash S^{(m-1)})^2}\mathbb{E}[\widetilde{\mathcal{A}}_{i_1i_2i_3}^2\widetilde{\mathcal{A}}_{i_3i_4i_5}^2\widetilde{\mathcal{A}}_{j_1j_2j_3}^2\widetilde{\mathcal{A}}_{j_3j_4j_5}^2]=o(n^{10}\alpha_n^4).$$

Recall that for each $x\in S^{(m)}\backslash S^{(m-1)}$, there are at least one index of (i_1,i_2,i_3,i_4,i_5) is m. It is seen that

$$\sum_{m=1}^{n}\sum_{y\in(S^{(m)}\backslash S^{(m-1)})^2}\mathbb{E}[\widetilde{\mathcal{A}}_{i_1i_2i_3}^2\widetilde{\mathcal{A}}_{i_3i_4i_5}^2\widetilde{\mathcal{A}}_{j_1j_2j_3}^2\widetilde{\mathcal{A}}_{j_3j_4j_5}^2]\le\sum_{m=1}^{n}n^{10-2}\Big(\alpha_n(1-\alpha_n)\Big)^4=o(n^{10}\alpha_n^4).$$

This finishes the proof.

G Proof of Theorem 3.2

Recall that $\phi_n=\max_{2\le m\le M}\{\phi_n^{(m)}\}$. To prove this theorem, it is sufficient to show that if there is a $m\in\{2,\dots,M\}$ such that $\|\theta^{(m)}\|_1^{m-2}\|\theta^{(m)}\|^2(\mu_2^{(m)})^2\gg\log(n)$, we will have

$$\phi_n^{(m)}\to 0\text{ under }H_0,\qquad\text{and }\phi_n^{(m)}\to\infty\text{ under }H_1.$$

Fix m. For simplicity, we remove the superscript (m) whenever it is clear from the context.

Let

$$\widetilde{\alpha}_n=\mathbb{E}[\hat{\alpha}_n],\qquad\beta=\sum_{k_2,\dots,k_m=1}^{K}\mathcal{P}_{:k_2\cdots k_m}g_{k_2}\cdots g_{k_m}/([\mathcal{P};g,\dots,g])^{(m-1)/m}.,$$

where $g\in\mathbb{R}^K$ is defined by $g_k=(1/\|\theta\|_1)\sum_{i=1}^n\theta_i\pi_i(k)$, $1\le k\le K$.

Introduce ideal counterparts of V_n and η by

$$\widetilde{V}_n=\binom{n}{m}\widetilde{\alpha}_n(1-\widetilde{\alpha}_n)\qquad\text{and}\qquad\eta^*=\Theta\Pi\beta,\qquad\text{respectively.} \tag{G.93}$$

The following lemma is used in this proof and we prove it after the main proof.

Lemma G.1. *With the conditions of Theorem 3.2, as $n\to\infty$,*

- *(a) Under both the null and alternative, $\widetilde{V}_n/V_n\to 1$ in probability.*

- *(b) Under the null, with a probability at least* $1 - O(1/n)$*,* $\max_{1\le i\le n}\{|\eta_i/\eta_i^* - 1|\} \le C(n^{m-1}\theta_{\max}^m/\log(n))^{-1/2}$.
- *(c) Under the alternative, with a probability at least* $1 - O(1/n)$*,* $\max_{1\le i\le n}\{|\eta_i/\eta_i^* - 1|\} \le C(n^{m-1}\theta_{\max}^m/\log(n))^{-1/2} + C\gamma_n/n$ *and* $n^m\theta_{\max}^m\gamma_n/(n^{m+1}\theta_{\max}^m\log(n))^{1/2} \to \infty$*, where* $\gamma_n = \max_{1\le k_1,\ldots,k_m\le K}\{|\mathcal{P}_{k_1\cdots k_m} - \beta_{k_1}\cdots\beta_{k_m}|\}$.

G.1 Main Proof of Theorem 3.2

Recall that $\phi_n^{(m)} = Q_n/\sqrt{n\log(n)^{1.1}V_n}$. The goal is to show that with probability $1 - o(1)$

$$Q_n \le (n\log(n)^{1.1}V_n)^{1/2} \text{ under } H_0^{(n)}, \qquad Q_n \ge (n\log(n)^{1.1}V_n)^{1/2} \text{ under } H_1^{(n)}, \tag{G.94}$$

By (a) in Lemma G.1, $\widetilde{V}_n/V_n \to 1$ in probability. Hence to show (G.94), it is sufficient to show that with probability $1 - o(1)$

$$Q_n \le 0.5(n\log(n)^{1.1}\widetilde{V}_n)^{1/2} \text{ under } H_0^{(n)}, \qquad Q_n \ge 1.5(n\log(n)^{1.1}\widetilde{V}_n)^{1/2} \text{ under } H_1^{(n)}. \tag{G.95}$$

Recall that

$$Q_n = \max_{S=(S_1,\ldots,S_{m+1})\in B}\ \max_{1\le k_1,\ldots,k_m\le m+1}\{|X_{S,k_1\cdots k_m}|\},$$

where

$$X_{S,k_1\cdots k_m} = \sum_{\substack{i_1\in S_{k_1},\ldots,i_m\in S_{k_m}\\ i_1,\ldots,i_m(dist)}}(\mathcal{A}_{i_1\cdots i_m} - \eta_{i_1}\cdots\eta_{i_m}).$$

Also, recall that η^* is the ideal counterparts of η, defined in (G.93). Introduce a counterpart of $X_{S,k_1\cdots k_m}$ by replacing η with η^*

$$\widetilde{X}_{S,k_1\cdots k_m} = \sum_{\substack{i_1\in S_{k_1},\ldots,i_m\in S_{k_m}\\ i_1,\ldots,i_m(dist)}}(\mathcal{A}_{i_1\cdots i_m} - \eta_{i_1}^*\cdots\eta_{i_m}^*).$$

Let

$$\widetilde{Q}_n = \max_{S=(S_1,\ldots,S_{m+1})\in B}\ \max_{1\le k_1,\ldots,k_m\le m+1}\{|\widetilde{X}_{S,k_1\cdots k_m}|\}.$$

Note that for any number $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$,

$$|\max\{x_1, x_2, \ldots, x_n\} - \max\{y_1, y_2, \ldots, y_n\}| \le \max\{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|\},$$

It is seen that

$$|Q_n - \widetilde{Q}_n| \le \max_S\ \max_{1\le k_1,\ldots,k_m\le m+1}\{|X_{S,k_1\cdots k_m} - \widetilde{X}_{S,k_1\cdots k_m}|\}. \tag{G.96}$$

At the same time, by definitions and direct calculations, for all $S = (S_1, \ldots, S_{m+1}) \in B$ and $1 \le k_1, \ldots, k_m \le m+1$

$$|X_{S,k_1\cdots k_m} - \widetilde{X}_{S,k_1\cdots k_m}| \le |S_{k_1}|\cdots|S_{k_m}|\max_{1\le i_1,\ldots,i_m\le n}|\eta_{i_1}\cdots\eta_{i_m} - \eta_{i_1}^*\cdots\eta_{i_m}^*|, \tag{G.97}$$

where by (b) and (c) in Lemma G.1, except for a probability $O(1/n)$

$$\max_{1\le i\le n}\Big\{|\frac{\eta_i}{\eta_i^*} - 1|\Big\} \le C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} \qquad \text{under } H_0, \tag{G.98}$$

and

$$\max_{1\le i\le n}\Big\{|\frac{\eta_i}{\eta_i^*} - 1|\Big\} \le C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C\gamma_n}{n} \qquad \text{under } H_1. \tag{G.99}$$

Here γ_n denotes $\max_{1\leq k_1,\ldots,k_m\leq K}\{|\mathcal{P}_{k_1\cdots k_m}-\beta_{k_1}\cdots\beta_{k_m}|\}$ under H_1. Note that by our regular conditions and elementary calculations, $\log(n)/(n^{m-1}\theta_{\max}^m)=o(1)$ and $\gamma_n/n=O(1/n)$. Therefore, $\max_{1\leq i\leq n}\Big\{|\frac{\eta_i}{\eta_i^*}-1|\Big\}=o(1)$ under both hypotheses. By Taylor's expansion, for $1\leq i_1,\ldots,i_m\leq n$

$$|\eta_{i_1}\cdots\eta_{i_m}-\eta_{i_1}^*\cdots\eta_{i_m}^*|\leq C\eta_{i_1}^*\cdots\eta_{i_m}^*\max_{1\leq i\leq n}\Big\{|\frac{\eta_i}{\eta_i^*}-1|\Big\}. \tag{G.100}$$

Combining (G.96)-(G.100) and observe that $\eta_i^*\leq C\theta_{\max}$ and $|S_{k_j}|\leq n$, $1\leq j\leq m$, with probability $1-o(1)$

$$|Q_n-\widetilde{Q}_n|\leq C\Big(\log(n)n^{m+1}\theta_{\max}^m\Big)^{1/2}\qquad\text{under } H_0, \tag{G.101}$$

and

$$|Q_n-\widetilde{Q}_n|\leq C\Big(\log(n)n^{m+1}\theta_{\max}^m\Big)^{1/2}+C\gamma_n n^{m-1}\theta_{\max}^m\qquad\text{under } H_1 \tag{G.102}$$

Note that by direct calculations, we have $\widetilde{V}_n\asymp n^m\theta_{\max}^m$. Therefore, to show (G.95), it is sufficient to show that with probability $1-o(1)$

$$(I):\widetilde{Q}_n\leq 0.5(n\log(n)^{1.1}\widetilde{V}_n)^{1/2}\qquad\text{under } H_0^{(n)},$$
$$(II):\widetilde{Q}_n\geq 2(n\log(n)^{1.1}\widetilde{V}_n)^{1/2}+C\gamma_n n^{m-1}\theta_{\max}^m\qquad\text{under } H_1^{(n)}.$$

Consider (I) first. Recall that

$$\widetilde{Q}_n=\max_{S=(S_1,\ldots,S_{m+1})\in B}\max_{1\leq k_1,\ldots,k_m\leq m+1}\{|\widetilde{X}_{S,k_1\cdots k_m}|\},$$

where the RHS is the maximum of

$$\leq m^n m^m=m^{n+m}$$

random variables. By union bound, it is sufficient to show that for every $S=(S_1,\ldots,S_{m+1})\in B$ and $1\leq k_1,\ldots,k_m\leq m+1$, except for a probability of $O(m^{-(n+m)}n^{-1})$

$$\big|\widetilde{X}_{S,k_1\cdots k_m}\big|\leq 0.5(n\log(n)^{1.1}\widetilde{V}_n)^{1/2}. \tag{G.103}$$

Now we are going to prove (G.103). Note that under null hypothesis, $\eta^*=\theta$. By definitions

$$\widetilde{X}_{S,k_1\cdots k_m}=\sum_{\substack{i_1\in S_{k_1},\ldots,i_m\in S_{k_m}\\ i_1,\ldots,i_m(dist)}}(\mathcal{A}_{i_1\cdots i_m}-\theta_{i_1}\cdots\theta_{i_m}),$$

where by symmetry the RHS is a sum of no more than $\binom{n}{m}$ unique independent random variables, each of which has mean 0 and variance $\leq(m!)^2\theta_{i_1}\cdots\theta_{i_m}(1-\theta_{i_1}\cdots\theta_{i_m})$. By Bernstein's inequality, for any $t>0$,

$$\mathbb{P}\big(\big|\widetilde{X}_{S,k_1\cdots k_m}\big|\geq t\big)\leq 2\exp\Big(-\frac{t^2}{\sum_{\substack{i_1\in S_{k_1},\ldots,i_m\in S_{k_m}\\ i_1,\ldots,i_m(unique)}}(m!)^2\theta_{i_1}\cdots\theta_{i_m}(1-\theta_{i_1}\cdots\theta_{i_m})+t/3}\Big).$$

Since $\sum_{\substack{i_1\in S_{k_1},\ldots,i_m\in S_{k_m}\\ i_1,\ldots,i_m(unique)}}(m!)^2\theta_{i_1}\cdots\theta_{i_m}(1-\theta_{i_1}\cdots\theta_{i_m})\leq Cn^m\theta_{\max}^m$, it follows that

$$\mathbb{P}\big(\big|\widetilde{X}_{S,k_1\cdots k_m}\big|\geq t\big)\leq 2\exp\Big(-\frac{t^2}{Cn^m\theta_{\max}^m+t/3}\Big). \tag{G.104}$$

Taking $t = (n\log(n)\widetilde{V}_n)^{1/2}$ and noting that $(1/C)\sqrt{n^{m+1}\log(n)\theta_{\max}^m} \le t \le \sqrt{n^{m+1}\log(n)\theta_{\max}^m}$,

$$\exp\Big(-\frac{t^2}{Cn^m\theta_{\max}^m + t/3}\Big) \le \exp\Big(-\frac{\Big((1/C)\sqrt{n^{m+1}\log(n)\theta_{\max}^m}\Big)^2}{Cn^m\theta_{\max}^m + \sqrt{n^{m+1}\log(n)\theta_{\max}^m}/3}\Big).$$

Combining this with our assumption $\|\theta\|_1^{m-2}\|\theta\|^2/\log(n) \to \infty$ and $\theta_{\max} \le C\theta_{\min}$, by elementary calculations, the RHS of (G.104) is $O(\exp(-Cn\log(n)))$. This proves (G.103).

Next, consider (II) for the alternative case. Let S_k^* denote the true partition set $\{1 \le i \le n : \text{node } i \text{ is in community } k\}$, $1 \le k \le K$. Also, recall that

$$\gamma_n = \max_{1\le k_1,\ldots,k_m\le K}\{|\mathcal{P}_{k_1\cdots k_m} - \beta_{k_1}\cdots\beta_{k_m}|\}.$$

Suppose the maximum on the right hand side is assumed at $(k_1,\ldots,k_m) = (k_1^*,\ldots,k_m^*)$ and so

$$\gamma_n = |\mathcal{P}_{k_1^*\cdots k_m^*} - \beta_{k_1^*}\cdots\beta_{k_m^*}|.$$

Without loss of generality, assume $k_1^*,\ldots,k_m^*$ are distinct. The proofs for the cases that $k_1^*,\ldots,k_m^*$ are not distinct are similar, so we omit them.

Now let $S^* = \Big(S_{k_1^*},\ldots,S_{k_m^*},\{1,\cdots,n\}\backslash(S_{k_1^*}\cup\cdots\cup S_{k_m^*})\Big)$. It follows that $S^* \in B$. By definitions,

$$\widetilde{Q}_n \ge |\widetilde{X}_{S^*,k_1^*\cdots k_m^*}|.$$

Therefore, to show (II), it is sufficient to show that except for a probability of $1 - O(1/n)$,

$$|\widetilde{X}_{S^*,k_1^*\cdots k_m^*}| \ge C(n\log(n)^{1.1}\widetilde{V}_n)^{1/2} + C\gamma_n n^{m-1}\theta_{\max}^m. \tag{G.105}$$

Write

$$\widetilde{X}_{S^*,k_1^*\cdots k_m^*} := \sum_{i_1\in S_{k_1^*},\ldots,i_m\in S_{k_m^*}}(\mathcal{A}_{i_1\cdots i_m} - \eta_{i_1}^*\cdots\eta_{i_m}^*) = (I) + (II), \tag{G.106}$$

where

$$(I) = \sum_{i_1\in S_{k_1^*},\ldots,i_m\in S_{k_m^*}}(\theta_{i_1}\cdots\theta_{i_m}\mathcal{P}_{k_1^*\ldots k_m^*} - \eta_{i_1}^*\cdots\eta_{i_m}^*),$$

and

$$(II) = \sum_{i_1\in S_{k_1^*},\ldots,i_m\in S_{k_m^*}}(\mathcal{A}_{i_1\cdots i_m} - \theta_{i_1}\cdots\theta_{i_m}\mathcal{P}_{k_1^*\cdots k_m^*}).$$

By definitions, $\eta_{i_1}^*\cdots\eta_{i_m}^* = \theta_{i_1}\cdots\theta_{i_m}\beta_{k_1^*}\cdots\beta_{k_m^*}$, for $i_1 \in S_{k_1^*},\ldots,i_m \in S_{k_m^*}$. It is seen that

$$|(I)| = \|\theta\|_1^m g_{k_1^*}\cdots g_{k_m^*}\gamma_n.$$

By our assumption $\max_{k=1}^K\{h_k\} \le C\min_{k=1}^K\{h_k\}$ and $\theta_{\max} \le C\theta_{\min}$,

$$\|\theta\|_1^m g_{k_1^*}\cdots g_{k_m^*} \ge Cn^m\theta_{\max}^m,$$

and so

$$|(I)| \ge Cn^m\theta_{\max}^m\gamma_n. \tag{G.107}$$

Write for short

$$N = |S_{k_1^*}^*|\cdots|S_{k_m^*}^*|.$$

Note that (II) is a sum of no more than N independent random variables, each with a mean of 0 and a variance less than $C\theta_{\max}^m$. By Bernstein's Lemma, for any $t > 0$,

$$\mathbb{P}(|(II)| \ge t) \le \exp(-\frac{t^2}{NC\theta_{\max}^m + t/3}). \tag{G.108}$$

Taking $t = (\log(n)\widetilde{V}_n)^{1/2}$. Note that $t \asymp (\log(n)n^m\theta_{\max}^m)^{1/2}$ and $N \leq n^m$, by direct calculations

$$\exp(-\frac{t^2}{NC\theta_{\max}^m + t/3}) = O(1/n).$$

Putting this into (G.108), gives except for a probability of $O(1/n)$,

$$|(II)| \leq (\log(n)\widetilde{V}_n)^{1/2}. \tag{G.109}$$

Inserting (G.107)-(G.109) into (G.106) gives that except for a probability of $O(1/n)$,

$$|\widetilde{X}_{S^*,k_1^*\cdots k_m^*}| \geq Cn^m\theta_{\max}^m\gamma_n - (\log(n)\widetilde{V}_n)^{1/2}, \tag{G.110}$$

where we note that by Lemma G.1, $n^m\theta_{\max}^m\gamma_n/(n\log(n)^{1.1}\widetilde{V}_n)^{1/2} \to \infty$. This proves (G.105) and finishes the proof.

G.2 Proof of Lemma G.1

Consider the claim (a). By definitions

$$\frac{V_n}{\widetilde{V}_n} - 1 = \frac{(\hat{\alpha}_n - \widetilde{\alpha}_n)(1 - \hat{\alpha}_n - \widetilde{\alpha}_n)}{\widetilde{\alpha}_n(1 - \widetilde{\alpha}_n)}. \tag{G.111}$$

Note that $\hat{\alpha}_n$ is the average of $\binom{n}{m}$ independent Bernoulli random variables with parameters bounded by $C\theta_{\max}^m$ under both null and alternative hypothesis. By Bernstein's inequality,

$$\mathbb{P}((\tbinom{n}{m})|\hat{\alpha}_n - \widetilde{\alpha}_n| \geq t\Big) \leq 2\exp(-\frac{t^2}{C\binom{n}{m}\theta_{\max}^m + \frac{t}{3}}).$$

Let $t = C\log(n)(\binom{n}{m}\theta_{\max}^m)^{1/2}$, by elementary calculations, we get

$$\mathbb{P}\Big(|\hat{\alpha}_n - \widetilde{\alpha}_n| \geq C\log(n)(\theta_{\max}^m/\tbinom{n}{m})^{1/2}\Big) \leq o(1/n).$$

Combining this with (G.111) and $\widetilde{\alpha}_n \leq C\theta_{\max}^m \leq Cc_0^m < 1$, by elementary calculations,

$$\Big|\frac{V_n}{\widetilde{V}_n} - 1\Big| \leq C\log(n)(\tbinom{n}{m}\theta_{\max}^m)^{-1/2}, \qquad \text{except for a probability of } O(1/n),$$

where by our conditions $n^{m-1}\theta_{\max}^m/\log(n) \to \infty$ (implied by $\|\theta\|_1^{m-2}\|\theta\|^2\mu_2^2/\log(n) \to \infty$), the RHS is $o(1)$. Therefore $V_n/\widetilde{V}_n \to 1$ in probability.

Combining this with Slutsky's Lemma, we get $\widetilde{V}_n/V_n \to 1$ in probability and finish the proof of (a).

Next we consider the claim (b) and the first claim in (c). Our goal is to show that except for a probability $O(1/n)$

$$\max_{1\leq i\leq n}\Big\{|\frac{\eta_i}{\eta_i^*} - 1|\Big\} \leq C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2}, \qquad \text{under } H_0 \tag{G.112}$$

and

$$\max_{1\leq i\leq n}\Big\{|\frac{\eta_i}{\eta_i^*} - 1|\Big\} \leq C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C\gamma_n}{n}, \qquad \text{under } H_1. \tag{G.113}$$

Recall that

$$\eta = u^{(\lceil\frac{m-1}{2}\rceil)} \qquad \text{and} \qquad u^{(k)} = g(u^{(k-1)}), \qquad 1 \leq k \leq m,$$

where for $1 \le i \le n$

$$L_{i_1}(u) = \frac{\sum_{i_2,\dots,i_m(\text{distinct})} \mathcal{A}_{i_1\cdots i_m} + \sum_{i_2,\dots,i_m(\text{non-distinct})} u_{i_1}\cdots u_{i_m}}{\big(\sum_{i_1,\dots,i_m(\text{distinct})} \mathcal{A}_{i_1\cdots i_m} + \sum_{i_1,\dots,i_m(\text{non-distinct})} u_{i_1}\cdots u_{i_m}\big)^{(m-1)/m}}.$$

Let $I^{(i_1)}$ denote $\{1,\dots,n\}\setminus\{i_1\}$. We claim that if the following events

$$\begin{aligned} E_1: & \max_{1\le i_1\le n}\Big\{\Big|\sum_{\substack{i_2,\dots,i_m\in I^{(i_1)}\\ (dist)}} (\mathcal{A}_{i_1\cdots i_m} - \mathcal{Q}_{i_1\cdots i_m})\Big|\Big\} \le (n^{m-1}\theta_{\max}^m \log(n))^{1/2}, \\ E_2: & \quad \Big|\sum_{\substack{i_1,\dots,i_m\\ (dist)}} (\mathcal{A}_{i_1\cdots i_m} - \mathcal{Q}_{i_1\cdots i_m})\Big| \le (n^m\theta_{\max}^m)^{1/2} \end{aligned} \tag{G.114}$$

hold then for $1 \le k \le m$

$$\max_{1\le i\le n}\{|\frac{L_i(u^{(k)})}{\eta_i^*} - 1|\} \le C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C}{n}\max_{1\le i\le n}\{|\frac{u_i^{(k)}}{\eta_i^*} - 1|\} + C\frac{\gamma_n}{n}, \tag{G.115}$$

where by definitions γ_n is 0 under H_0.

Note that inequality (G.115) implies the claims (G.112)-(G.113). To see this, recall that $u^{(k)} = g(u^{(k-1)})$. If inequality (G.115) holds, then

$$\begin{aligned} \max_{1\le i\le n}\{|\frac{u_i^{(k)}}{\eta_i^*} - 1|\} &= \max_{1\le i\le n}\{|\frac{L_i(u^{(k-1)})}{\eta_i^*} - 1|\} \\ &\le C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C}{n}\max_{1\le i\le n}\{|\frac{u_i^{(k-1)}}{\eta_i^*} - 1|\} + \frac{C\gamma_n}{n} \\ &\overset{\cdots}{\le} C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2}(1+o(1)) + \frac{C}{n^k}\max_{1\le i\le n}\{|\frac{u_i^{(0)}}{\eta_i^*} - 1|\} + \frac{C\gamma_n}{n}(1+o(1)) \\ \text{(Note that } u^{(0)} = 0) \quad &\le C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C}{n^k} + \frac{C\gamma_n}{n}. \end{aligned}$$

Combining this with $\eta = u^{(\lceil \frac{m-1}{2}\rceil)}$, it follows that n^{-k} ($k = \lceil \frac{m-1}{2}\rceil$) is a minor term and so $\max_{1\le i\le n}\{|\eta_i/\eta_i^* - 1|\} \le C(\log(n)/n^{m-1}\theta_{\max}^m)^{1/2} + C\gamma_n/n$ (i.e., the claims (G.112)-(G.113)).

Therefore, it is sufficient to show that events (G.114) hold except for a probability $O(1/n)$ and that inequality (G.115) holds for $1 \le k \le m$ given these events.

First, we show that the events E_1 and E_2 hold with a probability of $1 - O(1/n)$.

Consider event E_1 first. For $1 \le i_1 \le n$, note that by symmetry,

$$\sum_{\substack{i_2,\dots,i_m\in I^{(i_1)}\\ (dist)}} (\mathcal{A}_{i_1\cdots i_m} - \mathcal{Q}_{i_1\cdots i_m}) = \sum_{i_2<\cdots<i_m\in I^{(i_1)}} (m-1)!(\mathcal{A}_{i_1\cdots i_m} - \mathcal{Q}_{i_1\cdots i_m}),$$

where the RHS is a sum of $\binom{n-1}{m-1}$ independent centered Bernoulli random variables with parameters bounded by $C\theta_{\max}^m$. By Bernstein's inequality, for any $t_1 > 0$

$$\mathbb{P}\Big(\sum_{\substack{i_2,\dots,i_m\in I^{(i_1)}\\ (dist)}} (\mathcal{A}_{i_1\cdots i_m} - \mathcal{Q}_{i_1\cdots i_m}) > t_1\Big) \le \exp(-\frac{t_1^2}{Cn^{m-1}\theta_{\max}^m + t_1/3}).$$

Similarly, for event E_2, we have for any $t_2 > 0$

$$\mathbb{P}\Big(\sum_{\substack{i_1,\dots,i_m\\ (dist)}} (\mathcal{A}_{i_1\cdots i_m} - \mathcal{Q}_{i_1\cdots i_m}) > t_2\Big) \le \exp(-\frac{t_2^2}{Cn^m\theta_{\max}^m + t_2/3}).$$

Letting $t_1 = \sqrt{2C}(n^{m-1}\theta_{\max}^m \log(n))^{1/2}$ and $t_2 = (n^m\theta_{\max}^m)^{1/2}$ and by direct calculations

$$\mathbb{P}\Big(\sum_{\substack{i_2,\ldots,i_m\in I^{(i_1)}\\(dist)}} (\mathcal{A}_{i_1\cdots i_m} - \mathcal{Q}_{i_1\cdots i_m}) > \sqrt{2C}(n^{m-1}\theta_{\max}^m \log(n))^{1/2}\Big) \le \exp(-2\log(n)) = O(1/n^2).$$

and

$$\mathbb{P}\Big(\sum_{i_1,\ldots,i_m} (\mathcal{A}_{i_1\cdots i_m} - \mathcal{Q}_{i_1\cdots i_m}) > (n^m\theta_{\max}^m)^{1/2}\Big) \le \exp(-n/C) = o(1/n^2).$$

Combining these with union bound over $1 \le i_1 \le n$, we see that events E_1 and E_2 hold except for a probability $O(1/n)$.

Next, we show inequality (G.115) when (G.114) is given.

By definitions (G.93) and elementary algebra, η^* can be written as

$$\eta^* = \frac{\sum_{i_2,\ldots,i_m=1}^n \mathcal{Q}_{i_1\cdots i_m}}{(\sum_{i_1,\ldots,i_m=1}^n \mathcal{Q}_{i_1\cdots i_m})^{\frac{m-1}{m}}}.$$

For $1 \le i_1 \le n$ and $0 \le k \le m$, we can then write

$$\frac{L_{i_1}(u^{(k)})}{\eta_{i_1}^*} = (I^{(k)})_{i_1}(II^{(k)})_{i_1}^{-\frac{m-1}{m}},$$

where

$$(I^{(k)})_{i_1} = \frac{\sum_{i_2,\ldots,i_m(\text{distinct})}\mathcal{A}_{i_1\cdots i_m} + \sum_{i_2,\ldots,i_m(\text{non-distinct})} u_{i_1}^{(k)}\cdots u_{i_m}^{(k)}}{\sum_{i_2,\ldots,i_m=1}^n \mathcal{Q}_{i_1\cdots i_m}}$$

and

$$(II^{(k)})_{i_1} = \frac{\sum_{i_1,\ldots,i_m(\text{distinct})}\mathcal{A}_{i_1\cdots i_m} + \sum_{i_1,\ldots,i_m(\text{non-distinct})} u_{i_1}^{(k)}\cdots u_{i_m}^{(k)}}{\sum_{i_1,\ldots,i_m=1}^n \mathcal{Q}_{i_1\cdots i_m}}.$$

Therefore to show (G.115), by Taylor's expansion, it is sufficient to show that

$$\max_{1\le i\le n}\{|(I^{(k)})_i - 1|\} = o(1), \qquad \max_{1\le i\le n}\{|(II^{(k)})_i - 1|\} = o(1), \tag{G.116}$$

$$\max_{1\le i\le n}\{|(I^{(k)})_i - 1|\} \le C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C}{n}\max_{1\le i\le n}\{|\frac{u_i^{(k)}}{\eta_i^*} - 1|\} + C\frac{\gamma_n}{n} \tag{G.117}$$

and that

$$\max_{1\le i\le n}\{|(II^{(k)})_i - 1|\} \le C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C}{n}\max_{1\le i\le n}\{|\frac{u_i^{(k)}}{\eta_i^*} - 1|\} + C\frac{\gamma_n}{n}. \tag{G.118}$$

Note that by triangle's inequality,

$$\begin{aligned}|(I^{(k)})_{i_1} - 1| \le &\Big|\frac{\sum_{\substack{i_2,\ldots,i_m\in I^{(i_1)}\\(\text{distinct})}}(\mathcal{A}_{i_1\cdots i_m} - \mathcal{Q}_{i_1\cdots i_m})}{\sum_{i_2,\ldots,i_m=1}^n \mathcal{Q}_{i_1\cdots i_m}}\Big| \\ &+ \Big|\frac{\sum_{\substack{i_2,\ldots,i_m\\(\text{non-distinct})}}(u_{i_1}^{(k)}\cdots u_{i_m}^{(k)} - \eta_{i_1}^*\cdots\eta_{i_m}^*)}{\sum_{i_2,\ldots,i_m=1}^n \mathcal{Q}_{i_1\cdots i_m}}\Big| \\ &+ \Big|\frac{\sum_{\substack{i_2,\ldots,i_m\\(\text{non-distinct})}}(\eta_{i_1}^*\cdots\eta_{i_m}^* - \mathcal{Q}_{i_1\cdots i_m})}{\sum_{i_2,\ldots,i_m=1}^n \mathcal{Q}_{i_1\cdots i_m}}\Big|.\end{aligned}$$

By event E_1 and $\mathcal{Q}_{i_1\cdots i_m} \asymp \theta_{\max}^m$, the first term on the RHS is $\leq C(n^{m-1}\theta_{\max}^m/\log(n))^{-1/2}$. At the same time, by definitions and elementary algebra, $|\eta_{i_1}^* \cdots \eta_{i_m}^* - \mathcal{Q}_{i_1\cdots i_m}| \leq \theta_{i_1}\cdots\theta_{i_m}\gamma_n$. It follows that

$$|(I^{(k)})_{i_1} - 1| \leq C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C}{n}\max_{1\leq i_1,\ldots,i_m\leq n}\Big\{\Big|\frac{u_{i_1}^{(k)}\cdots u_{i_m}^{(k)}}{\eta_{i_1}^*\cdots\eta_{i_m}^*} - 1\Big|\Big\} + C\frac{\gamma_n}{n}. \tag{G.119}$$

Similarly, by event E_2 and elementary calculations, we have

$$\begin{aligned}|(II^{(k)})_{i_1} - 1| \leq & C\Big(\frac{1}{n^m\theta_{\max}^m}\Big)^{1/2} + \frac{C}{n}\max_{1\leq i_1,\ldots,i_m\leq n}\Big\{\Big|\frac{u_{i_1}^{(k)}\cdots u_{i_m}^{(k)}}{\eta_{i_1}^*\cdots\eta_{i_m}^*} - 1\Big|\Big\} + C\frac{\gamma_n}{n} \\ \leq & C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C}{n}\max_{1\leq i_1,\ldots,i_m\leq n}\Big\{\Big|\frac{u_{i_1}^{(k)}\cdots u_{i_m}^{(k)}}{\eta_{i_1}^*\cdots\eta_{i_m}^*} - 1\Big|\Big\} + C\frac{\gamma_n}{n}.\end{aligned} \tag{G.120}$$

Therefore, using Taylor's expansion on $u_{i_1}^{(k)}\cdots u_{i_m}^{(k)}/(\eta_{i_1}^*\cdots\eta_{i_m}^*)$, to show (G.116)-(G.118), it is sufficient to show that

$$\max_{1\leq i\leq n}\{|\frac{u_i^{(k)}}{\eta_i^*} - 1|\} = o(1), \qquad 1\leq k\leq K,$$

where we recall that our original goal is to show

$$\max_{1\leq i\leq n}\{|\frac{L_i(u^{(k)})}{\eta_i^*} - 1|\} \leq C\Big(\frac{\log(n)}{n^{m-1}\theta_{\max}^m}\Big)^{1/2} + \frac{C}{n}\max_{1\leq i\leq n}\{|\frac{u_i^{(k)}}{\eta_i^*} - 1|\} + C\frac{\gamma_n}{n},$$

Noting that $u^{(k)} = g(u^{(k-1)})$. Using induction, we only need to verify that $\max_{1\leq i\leq n}\{|L_i(u^{(0)})/\eta_i^* - 1|\} = o(1)$. To see this, by $u^{(0)} = 0$, we have

$$\max_{1\leq i_1,\ldots,i_m\leq n}\Big|\frac{u_{i_1}^{(0)}\cdots u_{i_m}^{(0)}}{\eta_{i_1}^*\cdots\eta_{i_m}^*} - 1\Big| = 1 = \max_{1\leq i\leq n}\Big\{\Big|\frac{u_i^{(0)}}{\eta_i^*} - 1\Big|\Big\}.$$

Combining this with (G.119)-(G.120), we get (G.116)-(G.118) hold for $k = 0$. It follows that

$$\max_{1\leq i\leq n}\{|\frac{L_i(u^{(0)})}{\eta_i^*} - 1|\} \leq C\max_{1\leq i\leq n}\{|(I^{(0)})_i - 1|\} + C\max_{1\leq i\leq n}\{|(II^{(0)})_i - 1|\} = o(1).$$

This finishes the proof of the claim (b) and the first claim in (c).

Lastly, consider the second claim of (c). Let $\mathcal{G}$ be a $m-$way symmetric tensor of dimension K defined by

$$\mathcal{G}_{k_1\cdots k_m} = \beta_{k_1}\cdots\beta_{k_m}, \qquad 1\leq k_1,\ldots,k_m\leq K,$$

and G be the matricization of $\mathcal{G}$. By [4, Corollary 7.3.5, Page 451],

$$|\sigma_2(P) - \sigma_2(G)| \leq \|P - G\|, \tag{G.121}$$

where $\sigma_2(B)$ denotes the second largest singular value of matrix B. Note that by definitions, the $k_2 + \sum_{j=3}^m K^{k_j-1}(k_j-1)$-th column of the matrix G can be written as the following form

$$G_{:,k_2+\sum_{j=3}^m K^{k_j-1}(k_j-1)} = \beta\cdot(\beta_{k_2}\cdots\beta_{k_m}), \qquad 1\leq k_2,\ldots,k_m\leq K.$$

It is seen that G is a rank-one matrix and so $\sigma_2(G) = 0$. Also, by the definition $\sigma_2(P) = |\mu_2|$. Combining these with (G.121) and noting that $\|P - G\| \leq C\max_{1\leq k_1,\ldots,k_m\leq K}\{|\mathcal{P}_{k_1\cdots k_m} - \beta_{k_1}\cdots\beta_{k_m}|\} = C\gamma_n$, we obtain

$$|\mu_2| \leq \|P - G\| \leq C\gamma_n.$$

By our assumption $\|\theta\|_1^{m-2}\|\theta\|^2\mu_2^2/\log(n)^{1.1} \to \infty$ and $\theta_{\max} \leq C\theta_{\min}$, the above inequality implies $n^{m-1}\theta_{\max}^m\gamma_n^2/\log(n)^{1.1} \to \infty$. It follows that

$$n^m\theta_{\max}^m\gamma_n/(n^{m+1}\theta_{\max}^m\log(n)^{1.1})^{1/2} = C(n^{m-1}\theta_{\max}^m\gamma_n^2/\log(n)^{1.1})^{1/2} \to \infty.$$

This proves the last claim in (c).

References

[1] Ravindra Bapat. D1ad2 theorems for multidimensional matrices. *Linear Algebra and its Applications*, 48:437–442, 1982.

[2] Shmuel Friedland. Positive diagonal scaling of a nonnegative tensor to one with prescribed slice sums. *Linear algebra and its applications*, 434(7):1615–1619, 2011.

[3] Peter Hall and Christopher C Heyde. *Martingale limit theory and its application*. Academic press, 1980.

[4] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[5] Alexandre B Tsybakov. *Introduction to nonparametric estimation*. Springer Science & Business Media, 2008.