Supplement of "Network Global Testing by Counting Graphlets"

Jiashun Jin [1] **Zheng Tracy Ke** [2] **Shengming Luo** [1]

Abstract

This is the supplementary material of (Jin et al., 2018). It contains Proposition A.1 and the proofs of Theorems 3.2-3.3, Corollary 3.1, and secondary lemmas.

A. An Alternative Expression of the GC Test Statistics

We rewrite the test statistic $\widehat{\chi}_{gc}$ (as well as $\widehat{L}_2$, $\widehat{L}_3$ and $\widehat{C}_4$) explicitly as a function of the adjacency matrix A. The following proposition is proved in Section D.4.

Proposition A.1 *The following are true:*

$$\widehat{L}_2 = \frac{1}{6\binom{n}{3}}\big[1'A^2 1 - \mathrm{tr}(A^2)\big],$$

$$\widehat{L}_3 = \frac{1}{24\binom{n}{4}}\big[1'A^3 1 - 2(1'A^2 1) + 1'A1 - \mathrm{tr}(A^3)\big],$$

and

$$\widehat{C}_4 = \frac{1}{24\binom{n}{4}}\big[\mathrm{tr}(A^4) - 2(1'A^2 1) + 1'A1\big].$$

Furthermore,

$$\widehat{\chi}_{gc} = \frac{\big[\mathrm{tr}(A^4) - 2(1'A^2 1) + 1'A1\big]}{n(n-1)(n-2)(n-3)} - \frac{1}{(n-3)^4}\left[\frac{1'A^3 1 - 2(1'A^2 1) + 1'A1 - \mathrm{tr}(A^3)}{1'A^2 1 - \mathrm{tr}(A^2)}\right]^4.$$

B. Proof of Theorem 3.2

We prove the case $m = 4$. The case of $m = 3$ is similar and thus omitted. From now on, we omit the superscripts "(4)" in all related quantities (e.g., we write $\delta_{gc}^{(4)}$ as δ_{gc}). Write

$$\frac{\sqrt{3\binom{n}{4}}}{\sqrt{\widehat{C}_4}}(\widehat{\chi}_{gc} - \chi_{gc}) = \sqrt{\frac{C_4}{\widehat{C}_4}} \cdot (I + II) \tag{1}$$

where

$$I = \frac{\sqrt{3\binom{n}{4}}}{\sqrt{C_4}}(\widehat{C}_4 - C_4), \quad II = -\frac{\sqrt{3\binom{n}{4}}}{\sqrt{C_4}}\Big[\Big(\frac{\widehat{L}_3}{\widehat{L}_2}\Big)^4 - \Big(\frac{L_3}{L_2}\Big)^4\Big].$$

Using the Slutsky's theorem, it suffices to show that

$$\widehat{C}_4/C_4 \xrightarrow{p} 1, \tag{2}$$

$$I \xrightarrow{d} N(0,1), \tag{3}$$

and

$$II \xrightarrow{p} 0, \tag{4}$$

The following lemma is useful, and its proof can be found in Section D.

Lemma B.1 *Under the assumptions of Theorem 3.2,*

$$C_4 \asymp n^{-4}\|\theta\|^8, \quad L_2 \asymp n^{-3}\|\theta\|_1^2\|\theta\|^2,$$
$$L_3 \asymp n^{-4}\|\theta\|_1^2\|\theta\|^4.$$

Moreover,

$$\mathrm{Var}(\widehat{C}_4) \le Cn^{-8}\|\theta\|^8, \quad \mathrm{Var}(\widehat{L}_2) \le Cn^{-6}\|\theta\|_1^3\|\theta\|_3^3,$$
$$\mathrm{Var}(\widehat{L}_3) \le Cn^{-8}\|\theta\|_1^4\|\theta\|_3^6.$$

We now show (2)-(4). The proof of (3) is relatively long, so we prove it in the end.

First, we prove (2). Recall that $C_4 = E[\widehat{C}_4]$. By Lemma B.1,

$$\mathbb{E}[(\widehat{C}_4/C_4 - 1)^2] = C_4^{-2}\mathrm{Var}(\widehat{C}_4) = O(\|\theta\|^{-8}),$$

where the right hand side $\to 0$ as $\|\theta\| \to \infty$. The claim follows by elementary probability theory.

Second, we prove (4). Define $\widehat{L}_2^* = (\|\theta\|^2/n)\widehat{L}_2$ and $L_2^* = (\|\theta\|^2/n)L_2$. Using Lemma B.1, it follows from direct calculations that

$$L_3/L_2^* = O(1). \tag{5}$$

---

*Equal contribution [1]Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, USA [2]Department of Statistics, University of Chicago, Chicago, USA. Correspondence to: Jiashun Jin <jiashun@stat.cmu.edu>.

Proceedings of the 35^{th} International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018.

With these notations, we have

$$|II| = \frac{\sqrt{3\binom{n}{4}}}{\sqrt{C_4}} \cdot \frac{\|\theta\|^8}{n^4} \left| \Big(\frac{\widehat{L}_3}{\widehat{L}_2^*}\Big)^4 - \Big(\frac{L_3}{L_2^*}\Big)^4 \right| \leq C\|\theta\|^4 \left| \Big(\frac{\widehat{L}_3}{\widehat{L}_2^*}\Big)^4 - \Big(\frac{L_3}{L_2^*}\Big)^4 \right|,$$

where we have used $C_4 \asymp n^{-4}\|\theta\|^8$ in the second equality; see Lemma B.1. Note that for any (x, y), $|x^4 - y^4| = |(x-y)(x^3 + x^2y + xy^3 + y^3)| \leq 3|x-y| \cdot (|x| + |y|)^3$. It follows that

$$|II| \leq C \cdot \|\theta\|^4 |Z| \cdot \left(\frac{L_3}{L_2^*} + |Z| \right)^3,$$

where for short we write

$$Z = \frac{\widehat{L}_3}{\widehat{L}_2^*} - \frac{L_3}{L_2^*}.$$

Recall that L_3/L_2^* is bounded. Therefore, to show (4), it suffices to show

$$\|\theta\|^4 \left(\frac{\widehat{L}_3}{\widehat{L}_2^*} - \frac{L_3}{L_2^*} \right) \xrightarrow{p} 0. \tag{6}$$

Below, we show (6). Write the term on the left by

$$\frac{\|\theta\|^4}{L_2^*}(\widehat{L}_3 - L_3) + \|\theta\|^4 \frac{\widehat{L}_3}{\widehat{L}_2^* L_2^*}(L_2^* - \widehat{L}_2^*) \equiv II_a + II_b.$$

To show (6), it suffices to show

$$II_a \xrightarrow{p} 0. \tag{7}$$

and

$$II_b \xrightarrow{p} 0. \tag{8}$$

Consider (7). Note that $L_3 = E[\widehat{L}_3]$. It follows from Lemma B.1 that

$$\mathrm{Var}(II_a) = \frac{\|\theta\|^8 \mathrm{Var}(\widehat{L}_3)}{(L_2^*)^2} \leq C \frac{\|\theta\|^8 \cdot n^{-8}\|\theta\|_1^4 \|\theta\|_3^6}{(n^{-4}\|\theta\|_1^2 \|\theta\|^4)^2} \leq C\|\theta\|_3^6,$$

where the last term $\to 0$ for $\|\theta\|_3 \to 0$ as $n \to \infty$; this is due to equation (7) of (Jin et al., 2018). By elementary probability, (7) follows.

Consider (8). To show the claim, we first show

$$\widehat{L}_2/L_2 \xrightarrow{p} 1, \qquad \widehat{L}_3/L_3 \xrightarrow{p} 1; \tag{9}$$

as the proofs are similar, we only show the first one. By Lemma B.1, $\mathrm{Var}(\widehat{L}_2) = O(n^{-6}\|\theta\|_1^3 \|\theta\|_3^3)$ and $L_2 \asymp n^{-3}\|\theta\|_1^2 \|\theta\|^2$. Using $E[\widehat{L}_2] = L_2$, $\mathbb{E}[(\widehat{L}_2/L_2 - 1)^2] = L_2^{-2}\mathrm{Var}(\widehat{L}_2) \leq C(\|\theta\|_3^3)/(\|\theta\|_1 \|\theta\|^4))$, which $\leq C/\|\theta\|^2$ since $\|\theta\|_3^3 \leq \|\theta\|_1 \|\theta\|^2$. This shows (9).

Using (9) and recalling $L_3/L_2^* \leq C$ (see (5)), to show (8), it is sufficient to show

$$\|\theta\|^4 \frac{1}{(L_2^*)}(\widehat{L}_2^* - L_2^*) \xrightarrow{p} 0,$$

and since $\widehat{L}_2^*/L_2^* = \widehat{L}_2/L_2$, it is equivalent to show

$$\|\theta\|^4 (\frac{\widehat{L}_2}{L_2} - 1) \xrightarrow{p} 0. \tag{10}$$

Last, we prove (3). We need some notations. Given 4 distinct nodes, there are 3 different possible cycles, denoted as $CC(i_1, i_2, i_3, i_4) = \{(i_1, i_2, i_3, i_4), (i_1, i_2, i_4, i_3), (i_1, i_3, i_2, i_4)\}$; moreover, for $B \subset \{1, 2, ..., n\}^4$, let $CC(B) = \cup_{(i_1, i_2, i_3, i_4) \in B} CC(i_1, i_2, i_3, i_4)$. For $1 \leq m \leq n$, let I_m be the collection of (i_1, i_2, i_3, i_4) such that $1 \leq i_1 < i_2 < i_3 < i_4 = m$. Write $\Omega_{ij}^* = \Omega_{ij}(1 - \Omega_{ij})$. Let $W = A - \Omega$. Define

$$S_{n,n} \equiv \frac{\sum_{CC(I_n)} W_{i_1 i_2} W_{i_2 i_3} W_{i_3 i_4} W_{i_4 i_1}}{\sqrt{\sum_{CC(I_n)} \Omega_{i_1 i_2}^* \Omega_{i_2 i_3}^* \Omega_{i_3 i_4}^* \Omega_{i_4 i_1}^*}}.$$

The following lemma is proved in Section D.

Lemma B.2 *Under the conditions of Theorem 3.2,*

$$\frac{\sqrt{3\binom{n}{4}}}{\sqrt{C_4}}(\widehat{C}_4 - C_4) - S_{n,n} \xrightarrow{p} 0.$$

By Lemma B.2, to show (3), it suffices to show that

$$S_{n,n} \xrightarrow{d} N(0, 1). \tag{11}$$

Below, we prove (11). For $1 \leq m \leq n$, define the σ-algebra $\mathcal{F}_{n,m} = \sigma(\{A_{ij}\}_{1 \leq i < j \leq m})$ and

$$X_{n,m} = S_{n,m} - S_{n,m-1},$$

where $S_{n,0} = 0$ and

$$S_{n,m} = \frac{\sum_{CC(I_m)} W_{i_1 i_2} W_{i_2 i_3} W_{i_3 i_4} W_{i_4 i_1}}{\sqrt{\sum_{CC(I_n)} \Omega_{i_1 i_2}^* \Omega_{i_2 i_3}^* \Omega_{i_3 i_4}^* \Omega_{i_4 i_1}^*}}, \quad 1 \leq m \leq n.$$

It is easy to see that $\mathbb{E}[S_{n,m} | \mathcal{F}_{n,m-1}] = S_{n,m-1}$. Hence, $\{X_{n,m}\}_{m=1}^n$ is a martingale difference sequence relative to the filtration $\{\mathcal{F}_{n,m}\}_{m=1}^n$, and $S_{n,n} = \sum_{m=1}^n X_{n,m}$. To show (11), we apply the martingale central limit theorem in (Hall & Heyde, 2014) and check:

(a) $\sum_{m=1}^n \mathbb{E}(X_{n,m}^2 | \mathcal{F}_{n,m-1}) \xrightarrow{p} 1$.

(b) $\sum_{m=1}^{n} \mathbb{E}(X_{n,m}^2 1_{\{|X_{n,m}|>\epsilon\}}|\mathcal{F}_{n,m-1}) \xrightarrow{p} 0$, for any $\epsilon > 0$.

Note that once we have checked that both conditions (a) and (b) are satisfied, then by the martingale central limit theorem, $S_{n,n} \xrightarrow{d} N(0,1)$. Combining it with Lemma B.2, we have proved (3).

It remains to check (a)-(b). For preparation, we first derive an alternative expression of $\mathbb{E}(X_{n,m}|\mathcal{F}_{n,m-1})$ as (14) below. By definition,

$$X_{n,m} = \frac{1}{\sqrt{M_n}} \sum_{\sum_{CC(I_m)\setminus CC(I_{m-1})}} W_{i_1i_2}W_{i_2i_3}W_{i_3i_4}W_{i_4i_1},$$

where $M_n \equiv \sum_{CC(I_n)} \Omega^*_{i_1i_2}\Omega^*_{i_2i_3}\Omega^*_{i_3i_4}\Omega^*_{i_4i_1}$ and the summation is over all 4-cycles in $CC(I_m) \setminus CC(I_{m-1})$. Note that a cycle in $CC(I_m)\setminus CC(I_{m-1})$ has to include the node m. Hence, we can use the following way to get all such cycles: First, select 2 indices (i,j) from $\{1,2,...,m-1\}$ and use them as the two neighboring nodes of m; second, select an index $k \in \{1,2,...,m-1\}\setminus\{i,j\}$ as the last node in the cycle. This allows us to write

$$X_{n,m} = \frac{1}{\sqrt{M_n}} \sum_{1\leq i<j\leq m-1} W_{mi}W_{mj} \cdot Y_{(m-1)ij}, \quad (12)$$

where

$$Y_{(m-1)ij} = \sum_{1\leq k\leq m-1, k\notin\{i,j\}} W_{ki}W_{kj}. \quad (13)$$

Conditioning on $\mathcal{F}_{n,m-1}$, $\{W_{mi}W_{mj}\}_{1\leq i<j\leq m-1}$ are mutually uncorrelated and $Y_{(m-1)ij}$ is a constant. Hence, it follows from (12)-(13) that

$$\mathbb{E}(X_{n,m}^2|\mathcal{F}_{n,m-1}) = \frac{1}{M_n} \sum_{1\leq i<j\leq m-1} Y^2_{(m-1)ij}\Omega^*_{mi}\Omega^*_{mj}. \quad (14)$$

We now check (a). It suffices to show that

$$\mathbb{E}\Big[\sum_{m=1}^{n} \mathbb{E}(X_{n,m}^2|\mathcal{F}_{n,m-1})\Big] = 1, \quad (15)$$

and

$$\mathrm{Var}\Big(\sum_{m=1}^{n} \mathbb{E}(X_{n,m}^2|\mathcal{F}_{n,m-1})\Big) \to 0. \quad (16)$$

Consider (15). In the definition (13), the terms in the sum are (unconditionally) mutually uncorrelated. As a result,

$$\mathbb{E}[Y^2_{(m-1)ij}] = \sum_{k<m, k\notin\{i,j\}} \Omega^*_{ki}\Omega^*_{kj}.$$

It follows that

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{m=1}^{n} \mathbb{E}(X_{n,m}^2|\mathcal{F}_{n,m-1})\Big] \\
=&\frac{1}{M_n}\sum_{m=1}^{n}\sum_{1\leq i<j\leq m-1}\sum_{1\leq k\leq m-1, k\notin\{i,j\}} \Omega^*_{ki}\Omega^*_{kj}\Omega^*_{mi}\Omega^*_{mj} \\
=&\frac{1}{M_n}\sum_{(m,i,j,k)\in CC(I_n)} \Omega^*_{mi}\Omega^*_{ik}\Omega^*_{kj}\Omega^*_{jm} = 1. \quad (17)
\end{aligned}$$

This proves (17).

Consider (16). We first decompose the random variable $\sum_{m=1}^{n} \mathbb{E}(X_{n,m}^2|\mathcal{F}_{n,m-1})$ into the sum of two parts, and then calculate its variance. By (13),

$$Y^2_{(m-1)ij} = \sum_k W^2_{ki}W^2_{kj} + \sum_{k\neq \ell} W_{ki}W_{kj}W_{\ell i}W_{\ell j},$$

where k and ℓ range in $\{1,2,...,m-1\}\setminus\{i,j\}$. Plugging it into (14), we have a decomposition

$$\sum_{m=1}^{n} \mathbb{E}(X_{n,m}^2|\mathcal{F}_{n,m-1}) = I_a + I_b, \quad (18)$$

where

$$I_a = \frac{1}{M_n}\sum_{m=1}^{n}\sum_{i<j\leq m-1}\sum_{\substack{k\leq m-1\\ k\notin\{i,j\}}} W^2_{ki}W^2_{kj}\Omega^*_{mi}\Omega^*_{mj},$$

$$I_b = \frac{1}{M_n}\sum_{m=1}^{n}\sum_{i<j\leq m-1}\sum_{\substack{k,\ell\leq m-1\\ k,\ell\notin\{i,j\}}} W_{ki}W_{kj}W_{\ell i}W_{\ell j}\Omega^*_{mi}\Omega^*_{mj}.$$

Then,

$$\begin{aligned}
&\mathrm{Var}\Big(\sum_{m=1}^{n} \mathbb{E}(X_{n,m}^2|\mathcal{F}_{n,m-1})\Big) \\
=&\mathrm{Var}(I_a) + \mathrm{Var}(I_b) + 2\mathrm{Cov}(I_a, I_b) \\
\leq&\Big(\sqrt{\mathrm{Var}(I_a)} + \sqrt{\mathrm{Var}(I_b)}\Big)^2 \quad (19)
\end{aligned}$$

It suffices to show that both $\mathrm{Var}(I_a) \to 0$ and $\mathrm{Var}(I_b) \to 0$.

Consider the variance of I_a. In the sum of I_a, all 4-cycles (k,i,m,j) involved are selected in this way: We first select m, then select a pair (i,j) from $\{1,2,\ldots,m-1\}$ and connect both i and j to m, and finally select k to close the cycle. In fact, these 4-cycles can be selected in an alternative way: First, select a V-shape (i,k,j) with k being the middle point. Second, select $m > \max\{i,k,j\}$ to make the V-shape a cycle. Hence, we can rewrite

$$I_a = \frac{1}{M_n}\sum_{k=1}^{n}\sum_{\substack{1\leq i<j\leq n\\ i\neq k, j\neq k}} W^2_{ki}W^2_{kj} \underbrace{\sum_{m>\max\{i,j,k\}} \Omega^*_{mi}\Omega^*_{mj}}_{\equiv b_{kij}}$$

The terms $W_{ki}^2 W_{kj}^2$ corresponding to different k are independent of each other. We now fix k and calculate the covariance between $W_{ki}^2 W_{kj}^2$ and $W_{ki'}^2 W_{kj'}^2$ for $(i,j) \neq (i',j')$. There are three cases. Case (i): $(i,j) = (i',j')$. In this case, $\mathrm{Var}(W_{ki}^2 W_{kj}^2) \leq \mathbb{E}[W_{ki}^4 W_{kj}^4] \leq \mathbb{E}[W_{ki}^2 W_{kj}^2] \leq \Omega_{ki}^* \Omega_{kj}^*$. Case (ii): $i = i'$ but $j \neq j'$. In this case, we have $\mathrm{Cov}(W_{ki}^2 W_{kj}^2, W_{ki}^2 W_{kj'}^2) = \mathrm{Var}(W_{ki}^2) \cdot \mathbb{E}[W_{kj}^2]\mathbb{E}[W_{kj'}^2] \leq \Omega_{ki}^* \Omega_{kj}^* \Omega_{kj'}^*$. Case (iii): $(i,j) \cap (i',j') = \emptyset$. The two terms are independent, and their covariance is zero. Combining the above gives

$$\mathrm{Var}(I_a) \leq \frac{1}{M_n} \sum_{k=1}^n \Bigg(\sum_{\substack{1 \leq i < j \leq n \\ i \neq k, j \neq k}} b_{kij}^2 \Omega_{ki}^* \Omega_{kj}^* + \sum_{\substack{i,j,j' \in \{1,\ldots,n\} \setminus \{k\} \\ i,j,j' \text{ are distinct}}} b_{kij} b_{kij'} \Omega_{ki}^* \Omega_{kj}^* \Omega_{kj'}^* \Bigg).$$

We now bound the right hand side. By condition (9), $\Omega_{ij}^* \leq C\theta_i\theta_j$. Hence, $b_{kij} \leq C \sum_{m>k} \theta_m^2 \theta_i \theta_j \leq C\|\theta\|^2 \theta_i\theta_j$. As a result,

$$\begin{aligned} \mathrm{Var}(I_a) &\leq \frac{C}{M_n^2} \Big[\sum_{k,i,j} \|\theta\|^4 \theta_k^2 \theta_i^3 \theta_j^3 + \sum_{k,i,j,j'} \|\theta\|^4 \theta_k^3 \theta_i^3 \theta_j^2 \theta_{j'}^2 \Big] \\ &\leq \frac{C}{M_n^2} (\|\theta\|^6 \|\theta\|_3^6 + \|\theta\|^8 \|\theta\|_3^6). \end{aligned}$$

By (7), $\|\theta\| \to \infty$, so the second term dominates. Moreover, since $\Omega_{ij}^* = \Omega_{ij}(1-\Omega_{ij}) \geq c\Omega_{ij}$ (in our setting, all Ω_{ij}'s are bounded away from 1). As a result, we have $M_n \geq c \sum_{CC(I_n)} \Omega_{i_1 i_2} \Omega_{i_2 i_3} \Omega_{i_3 i_4} \Omega_{i_4 i_1} \geq C^{-1} n^4 C_4$. By Lemma B.1, $n^4 C_4 \asymp \|\theta\|^8$. Combining the above gives $\mathrm{Var} = O(\|\theta\|_3^6 / \|\theta\|^8)$, i.e.,

$$\sqrt{\mathrm{Var}(I_a)} \leq \frac{C \sum_i \theta_i^3}{(\sum_i \theta_i^2)^2} \leq \frac{C\theta_{\max}}{\sum_i \theta_i^2} = o(1). \tag{20}$$

Consider the variance of I_b. Rewrite

$$I_b = \frac{1}{M_n} \sum_{k,\ell,i,j \text{ are distinct}} c_{k\ell ij} G_{k\ell ij},$$

where

$$G_{k\ell ij} \equiv W_{ki} W_{kj} W_{\ell i} W_{\ell j}, \quad c_{k\ell ij} = \sum_{\substack{m> \\ \max\{k,\ell,i,j\}}} \Omega_{mi}^* \Omega_{mj}^*.$$

Since I_b has a mean zero, $\mathrm{Var}(I_b) = \mathbb{E}(I_b^2)$. Additionally, for 2 cycles (k,ℓ,i,j) and (k',ℓ',i',j'), only when they are exactly equal, we have $\mathbb{E}[G_{k\ell ij} G_{k'\ell' i' j'}] \neq 0$. As a result,

$$\begin{aligned} \mathrm{Var}(I_b) &= \frac{1}{M_n} \sum_{k,\ell,i,j \text{ are distinct}} c_{k\ell ij}^2 \mathbb{E}[G_{k\ell ij}^2] \\ &= \frac{1}{M_n} \sum_{k,\ell,i,j \text{ are distinct}} c_{k\ell ij}^2 \Omega_{ki}^* \Omega_{kj}^* \Omega_{\ell i}^* \Omega_{\ell j}^*. \end{aligned}$$

Similarly to how we get the bound for b_{kij}, we can derive that $c_{k\ell ij} \leq C\|\theta\|^2 \theta_i\theta_j$. Moreover, $\Omega_{ki}^* \Omega_{kj}^* \Omega_{\ell i}^* \Omega_{\ell j}^* \leq C\theta_i^2 \theta_j^2 \theta_k^2 \theta_\ell^2$. Hence,

$$\mathrm{Var}(I_b) \leq \frac{C}{\|\theta\|^{16}} \sum_{k,\ell,i,j} \|\theta\|^4 \theta_k^2 \theta_\ell^2 \theta_i^4 \theta_j^4 \leq \frac{C\|\theta\|_4^8}{\|\theta\|^8}.$$

As a result,

$$\sqrt{\mathrm{Var}(I_b)} \leq \frac{C \sum_i \theta_i^4}{(\sum_i \theta_i^2)^2} \leq \frac{C\theta_{\max}^2}{\sum_i \theta_i^2} = o(1). \tag{21}$$

Plugging (20)-(21) into (19) gives (16). Combining (15) and (16), we have proved (a).

We now check (b). By the Cauchy-Schwarz inequality and the Chebyshev's inequality,

$$\begin{aligned} &\sum_{m=1}^n \mathbb{E}(X_{n,m}^2 1_{\{|X_{n,m}|>\epsilon\}} | \mathcal{F}_{n,m-1}) \\ \leq &\sum_{m=1}^n \sqrt{\mathbb{E}(X_{n,m}^4 | \mathcal{F}_{n,m-1})} \sqrt{\mathbb{P}\big(|X_{n,m}| \geq \epsilon | \mathcal{F}_{n,m-1}\big)} \\ \leq &\epsilon^{-2} \sum_{m=1}^n \mathbb{E}(X_{n,m}^4 | \mathcal{F}_{n,m-1}). \end{aligned}$$

Therefore, it suffices to show that the right hand side converges to zero in probability. Then, it suffices to show that its L^1-norm converges to zero. Since the right hand is a nonnegative random variable, we only need to prove that its expectation converges to zero, i.e.,

$$\mathbb{E}\Big[\sum_{m=1}^n X_{n,m}^4\Big] = o(1). \tag{22}$$

We now prove (22). We use the expression of $X_{n,m}$ in (12). Conditioning on $\mathcal{F}_{n,m-1}$, the $Y_{(m-1)ij}$'s are non-random. It follows that

$$\begin{aligned} \mathbb{E}[X_{n,m}^4 | \mathcal{F}_{n,m-1}] &= \frac{1}{M_n^4} \sum_{\substack{i,j=1 \\ i \neq j}}^{m-1} Y_{(m-1)ij}^2 \mathbb{E}[W_{mi}^4 W_{mj}^4] \\ &+ \frac{1}{M_n^4} \sum_{i=1}^{m-1} \sum_{\substack{j,j'=1 \\ j,j' \notin \{i\}}}^{m-1} Y_{(m-1)ij} Y_{(m-1)ij'} \mathbb{E}[W_{mi}^4 W_{mj}^2 W_{mj'}^2] \\ &+ \frac{1}{M_n^4} \sum_{\substack{i,j,i',j'=1 \\ \text{distinct}}}^{m-1} Y_{(m-1)ij} Y_{(m-1)i'j'} \mathbb{E}[W_{mi}^2 W_{mj}^2 W_{mi'}^2 W_{mj'}^2]. \end{aligned}$$

First, we shall use the independence across entries of W and the fact that $\mathbb{E}[W_{ij}^4] \leq \mathbb{E}[W_{ij}^2] \leq \Omega_{ij} \leq C\theta_i\theta_j$. Second, in proving (17), we have seen that $\mathbb{E}[Y_{(m-1)ij}^2] = \sum_{k<m, k \notin \{i,j\}} \Omega_{ki}^* \Omega_{kj}^* \leq C \sum_k \theta_k^2 \theta_i \theta_j \leq C\|\theta\|^2 \theta_i\theta_j$.

Third, from (13), it is easy to see that when (i, j, i', j') are distinct, $\mathbb{E}[Y_{(m-1)ij}Y_{(m-1)i'j'}] = 0$; moreover, for $j \neq j'$, $\mathbb{E}[Y_{(m-1)ij}Y_{(m-1)ij'}] = \sum_k \mathbb{E}[W_{ki}^2]\mathbb{E}[W_{kj}W_{kj'}] = 0$. Last, in proving (20), we have seen that $M_n \geq c\|\theta\|^8$. Combining the above, we find that

$$\begin{aligned}\mathbb{E}[X_{n,m}^4] &= \frac{1}{M_n^2}\sum_{\substack{i,j=1\\ i\neq j}}^{m-1}\mathbb{E}[Y_{(m-1)ij}^2]\mathbb{E}[W_{mi}^4 W_{mj}^4]\\ &\leq \frac{C}{\|\theta\|^{16}}\sum_{i,j=1}^{m-1}(\|\theta\|^2\theta_i\theta_j)(\theta_m\theta_i)(\theta_m\theta_j)\\ &\leq C\theta_m^2/\|\theta\|^{10}.\end{aligned}$$

As a result,

$$\sum_{n=1}^{n}\mathbb{E}[X_{n,m}^4] \leq C\|\theta\|^{-8} = o(1).$$

This gives (22) and (b) follows. □

C. Proof of Theorem 3.3 and Corollary 3.1

Consider Theorem 3.3 first. For short, let

$$Z_n^{(m)} = \sqrt{\frac{B_{n,m}}{2m}}\widehat{C}_m^{-1/2}\widehat{\chi}_{gc}^{(m)}, \quad x_0^* = \mathbb{P}(Z_n^{(m)} \geq z_\alpha).$$

It suffices to show that under the null and alternative,

$$|x_0^* - \Phi(\delta_{gc}^{(m)} - z_\alpha)| \leq o(1). \tag{23}$$

Denote $a_n = (C_m/\widehat{C}_m)^{1/2}$ for short. It is seen that

$$a_n \xrightarrow{p} 1, \tag{24}$$

and

$$\frac{1}{a_n}Z_n^{(m)} = \sqrt{\frac{B_{n,m}}{2m}}C_m^{-1/2}\widehat{\chi}_{gc}^{(m)}. \tag{25}$$

Combining Theorem 3.1 and the proof of Theorem 3.2, we have shown that

$$\sqrt{\frac{B_{n,m}}{2m}}C_m^{-1/2}\big[\widehat{\chi}_{gc}^{(m)} - \chi_{gc,0}^{(m)}\big] \xrightarrow{p} N(0,1), \tag{26}$$

where by definitions,

$$\sqrt{\frac{B_{n,m}}{2m}}C_m^{-1/2}\chi_{gc,0}^{(m)} = \delta_{gc}^{(m)}. \tag{27}$$

Combining (25)-(27) gives

$$\frac{1}{a_n}Z_n^{(m)} - \delta_{gc}^{(m)} \xrightarrow{d} N(0,1). \tag{28}$$

Denote the CDF of $\frac{1}{a_n}Z_n^{(m)} - \delta_{gc}^{(m)}$ by F_n. Recall that Φ denotes the CDF of $N(0,1)$. It follows from (28) that

$$\sup_x |F_n(x) - \Phi(x)| \to 0. \tag{29}$$

We now rewrite

$$x_0^* = \mathbb{P}\bigg(\frac{1}{a_n}Z_n^{(m)} - \delta_{gc}^{(m)} \geq \frac{1}{a_n}z_\alpha - \delta_{gc}^{(m)}\bigg),$$

and introduce a proxy by

$$x_0 = \mathbb{P}\bigg(\frac{1}{a_n}Z_n^{(m)} - \delta_{gc}^{(m)} \geq z_\alpha - \delta_{gc}^{(m)}\bigg).$$

By triangle inequality,

$$|x_0^* - \Phi(\delta_{gc}^{(m)} - z_\alpha)| \leq |x_0^* - x_0| + |x_0 - \Phi(\delta_{gc}^{(m)} - z_\alpha)|. \tag{30}$$

where by (29),

$$|x_0 - \Phi(\delta_{gc}^{(m)} - z_\alpha)| \to 0. \tag{31}$$

Moreover, for any fixed $\epsilon > 0$, it is seen that

$$|x_0^* - x_0| \leq I + II,$$

where

$$I = \mathbb{P}(|a_n - 1| \geq \epsilon),$$

and

$$II = \mathbb{P}(\tfrac{1}{a_n}Z_n^{(m)} - \delta_{gc}^{(m)} \text{ falls between } (1 \pm \epsilon)z_\alpha - \delta_{gc}^{(m)}),$$

which by (29) does not exceed

$$\mathbb{P}(N(0,1) \text{ falls between } (1 \pm \epsilon)z_\alpha - \delta_{gc}^{(m)}) + o(1);$$

note the first term does not exceed $(2/\sqrt{2\pi})z_\alpha\epsilon$. Combining these gives that for any $\epsilon > 0$,

$$|x_0^* - \Phi(\delta_{gc}^{(m)} - z_\alpha)| \leq (2/\sqrt{2\pi})z_\alpha\epsilon + \mathbb{P}(|a_n - 1| \geq \epsilon) + o(1).$$

Recall that $a_n \xrightarrow{p} 1$, the claim follows.

Next, consider Corollary 3.1. It is seen that $\delta_{gc}^{(m)} = 0$ under the null and that under the alternative,

$$\delta_{gc}^{(m)} \geq \sum_{k=2}^{K}\lambda_k^m/[\sum_{k=1}^{K}\lambda_k^m]^{1/2}.$$

When $m = 4$, by Lemma 6.1, $\delta_{gc}^{(4)} \geq c_4\|\theta\|^4$ for some constant $c_4 > 0$. When $m = 3$ and P is positive definite, λ_k are the eigenvalues of $\Theta\Pi P\Pi'\Theta$, so for $1 \leq k \leq K$, $\lambda_k \geq 0$. Using Lemma 6.1, $\delta_{gc}^{(3)} \geq c_3\|\theta\|^3$ for some constant c_3. Combining these with Theorem 3.3 gives the claim.

D. Proof of Secondary Lemmas

D.1. Proof of Lemma 6.1

We first consider the claim about λ_k's. Recall that λ_k's are the eigenvalues of the matrix $G^{1/2}PG^{1/2}$. First, we have $\|G\| \leq \sum_{k,\ell} G(k,\ell) = \sum_{k,\ell}\sum_i \theta_i^2 \pi_i(k)\pi_i(\ell) = \sum_i \theta_i^2 \sum_{k,\ell}\pi_i(k)\pi_i(\ell) = \|\theta\|^2$. Second, let $g_k = \sum_{i\in\mathcal{N}_k}\theta_i^2$ for $1 \leq k \leq K$, and write $\Theta = \Theta_1 + \Theta_2$, where $\Theta_1(i,i) = \theta_i \cdot 1\{i \in \cup_{k=1}^K \mathcal{N}_k\}$ and $\Theta_2 \equiv \Theta - \Theta_1$. It yields that $G = \Pi'\Theta_1^2\Pi + \Pi'\Theta_2^2\Pi = \mathrm{diag}(g_1, \cdots, g_K) + \Pi'\Theta_2^2\Pi$. Hence, $\lambda_{\min}(G) \geq \min_{1\leq k\leq K} g_k \geq c_2\|\theta\|^2$, where the last inequality is from condition (8). Combining the above gives

$$c_2\|\theta\|^2 \leq \lambda_{\min}(G) \leq \|\theta\|^2. \qquad (32)$$

Using condition (9), we find that $|\lambda_k| \leq \|PG\| \leq C\|G\| = O(\|\theta\|^2)$. Additionally, since $|\lambda_k|^2$ is an eigenvalue of $(G^{1/2}PG^{1/2})^2 = G^{1/2}PGPG^{1/2}$, we then have $|\lambda_k|^2 \geq \lambda_{\min}(G)\cdot\lambda_{\min}(PGP) \geq \lambda_{\min}^2(G)\cdot s_{\min}^2(P) \geq c_1^2c_2^2\|\theta\|^4$. It gives

$$C^{-1}\|\theta\|^2 \leq |\lambda_k| \leq C\|\theta\|^2, \qquad 1 \leq k \leq K.$$

We then consider the claim about η. Since $\max_k |\eta'\xi_k|^2$ is upper bounded by $\sum_k |\eta'\xi_k|^2$ and lower bounded by $K^{-1}\sum_k |\eta'\xi_k|^2$, it suffices to show that

$$C^{-1}\|\theta\|_1^2 \leq \sum_{1\leq k\leq K} |\eta'\xi_k|^2 \leq C\|\theta\|_1^2. \qquad (33)$$

Since $\xi_1, \ldots, \xi_K$ form an orthonormal basis,

$$\sum_{1\leq k\leq K} |\eta'\xi_k|^2 = \|\eta\|^2 = 1_n'\Theta\Pi G^{-1}\Pi'\Theta 1_n.$$

It follows from (32) that the right hand side has the same order as $\|\theta\|^{-2}\|\Pi'\Theta 1_n\|^2$. Write $v = \Pi'\Theta 1_n$. For $1 \leq k \leq K$, $v(k) = \sum_i \pi_i(k)\theta_i$. It follows that $v(k) \leq \|\theta\|_1$. At the same time, $\sum_{k=1}^K v^2(k) \geq \frac{\left(\sum_{k=1}^K v(k)\right)^2}{K} = \frac{\|\theta\|_1^2}{K}$, where we've used Cauchy-Schwarz inequality.

It follows that

$$C^{-1}\|\theta\|_1^2 \leq \|\Pi'\Theta 1_n\|^2 \leq C\|\theta\|_1^2.$$

Hence, (33) follows. □

D.2. Proof of Lemma B.1

Consider the first item. By (16) of (Jin et al., 2018),

$$C_4 = \frac{1}{B_{n,4}}\left[\sum_{k=1}^K \lambda_k^4 + O(\|\theta\|_4^4\|\theta\|^4)\right],$$

where we note $B_{n,4} \sim n^{-4}$. First, by Lemma 6.1 of (Jin et al., 2018),

$$\sum_{k=1}^K \lambda_k^4 \asymp \|\theta\|^8,$$

Second, by (7) of (Jin et al., 2018), $\theta_{max} \leq \|\theta\|_3 \to 0$, so it is seem $\|\theta\|_4^4 \leq \theta_{max}\|\theta\|_3^3 \leq o(1)$, and so $\|\theta\|_4^4\|\theta\|^4 \leq o(\theta\|^4)$. Combining these give the claim.

Consider the second item. By (17) of (Jin et al., 2018),

$$L_2 = \frac{1}{B_{n,3}}\left[\sum_{k=1}^K \lambda_k^2(\eta,\xi_k)^2 + O(\|\theta\|_1^2\|\theta\|_4^4\|\theta\|^{-2})\right],$$

where by Lemma 6.1 of (Jin et al., 2018),

$$\sum_{k=1}^K \lambda_k^2(\eta,\xi_k)^2 \asymp \|\theta\|^2\|\theta\|_1^2.$$

By similar argument, $\|\theta\|_1^2\|\theta\|_4^4\|\theta\|^{-2} \leq o(\|\theta\|^2\|\theta\|_1^2)$, so the claim follows by noting $B_{n,3} \sim n^{-3}$.

Consider the third item. By similar argument, it is seen that

$$\begin{aligned} L_3 &= \frac{1}{B_{n,4}}\left[\sum_{k=1}^K \lambda_k^3(\eta,\xi_k)^2 + O(\|\theta\|_1^2\|\theta\|_4^4)\right] \\ &\leq Cn^{-4}\|\theta\|_1^2\|\theta\|^4. \end{aligned}$$

For the lower bound, we use a different proof as λ_k may be negative. By $L_3 = E[\widehat{L}_3]$ and $E[A_{ij}] = \Omega_{ij}$ when $i \neq j$,

$$L_3 = \frac{1}{B_{n,4}} \sum_{\substack{1\leq i_1,i_2,i_3,i_4\leq n \\ \text{are distinct}}} \Omega_{i_1i_2}\Omega_{i_2i_3}\Omega_{i_3i_4}.$$

As before, let $\mathcal{N}_1$ denote the set of pure nodes in community 1. It is not hard to see that

$$L_3 \geq \sum_{k=1}^K \sum_{\substack{i_1,i_2,i_3,i_4\in\mathcal{N}_k \\ \text{are distinct}}} \Omega_{i_1i_2}\Omega_{i_2i_3}\Omega_{i_3i_4}.$$

In our model, all diagonal entries of P are 1, so for any $i, j \in \mathcal{N}_1$, $\Omega_{ij} = \theta_i\theta_j$. Therefore,

$$L_3 \geq \sum_{k=1}^K \sum_{\substack{i_1,i_2,i_3,i_4\in\mathcal{N}_k \\ \text{are distinct}}} \theta_{i_1}\theta_{i_4}\theta_{i_2}^2\theta_{i_3}^2. \qquad (34)$$

Now, we can lower bound the right hand side of (34) by

$$I - II - III - IV,$$

where

$$I = \sum_{k=1}^K \sum_{i_1,i_2,i_3,i_4\in\mathcal{N}_k} \theta_{i_1}\theta_{i_4}\theta_{i_2}^2\theta_{i_3}^2,$$

$$II = \sum_{k=1}^K \sum_{\substack{i_1,i_2,i_3,i_4\in\mathcal{N}_k \\ i_1=i_4}} \theta_{i_1}\theta_{i_4}\theta_{i_2}^2\theta_{i_3}^2,$$

$$III = \sum_{\substack{i_1,i_2,i_3,i_4\in\mathcal{N}_1 \\ i_2=i_3}} \theta_{i_1}\theta_{i_4}\theta_{i_2}^2\theta_{i_3}^2,$$

and

$$IV = 4\sum_{k=1}^{K} \sum_{\substack{i_1,i_2,i_3,i_4 \in \mathcal{N}_k \\ i_1 = i_2}} \theta_{i_1}\theta_{i_4}\theta_{i_2}^2\theta_{i_3}^2.$$

First, by (8) of (Jin et al., 2018),

$$\begin{aligned} I &= \sum_{k=1}^{K} \big(\sum_{i\in\mathcal{N}_k} \theta_i\big)^2 \big(\sum_{i\in\mathcal{N}_k} \theta_i^2\big)^2 \\ &\geq C\|\theta\|^4 \sum_{k=1}^{K} \big(\sum_{i\in\mathcal{N}_k} \theta_i\big)^2 \\ &\geq C\|\theta\|_1^2\|\theta\|^4, \end{aligned}$$

where the last inequality is Cauchy-Schwarz inequality.

Second, by (7) of (Jin et al., 2018) that $\theta_{\max} \leq \|\theta\|_3 = o(1)$, we obtain $\|\theta\|^2 \leq o(1)\cdot\|\theta\|_1$ (note $\|\theta\| \to \infty$), and

$$\begin{aligned} II &\leq C\sum_{k=1}^{K} \sum_{i_1,i_2,i_3\in\mathcal{N}_k} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2 \\ &\leq C\|\theta\|_1\|\theta\|^4 \\ &= o\big(\|\theta\|_1^2\|\theta\|^4\big). \end{aligned}$$

Similarly, we have $\|\theta\|_3^3 \leq o(1)\cdot\|\theta\|_2^2$ and $\|\theta\|_4^4 \leq o(1)\cdot \|\theta\|_2^2$, which implies

$$III = o\big(\|\theta\|_1^2\|\theta\|^4\big), \quad \text{and} \quad IV = o\big(\|\theta\|_1^2\|\theta\|^4\big).$$

Combining these gives $L_3 \geq c\|\theta\|_1^2\|\theta\|^4$, and the claim follows.

We now prove the next three items (on the variances). In the Proof of Lemma B.2, we've already shown that $\frac{1}{B_{n,4}}\sum_{i_1,\cdots,i_4 \text{ distinct}} G_{i_1i_2i_3i_4}(W)$ is the dominating term of $(\widehat{C}_4 - C_4)$, and that

$$\begin{aligned} &\mathrm{Var}\Big(\frac{1}{B_{n,4}}\sum_{\substack{i_1,\cdots,i_4 \\ \text{distinct}}} G_{i_1i_2i_3i_4}(W)\Big) \\ \leq& Cn^{-4}\sum_{\substack{i_1,\cdots,i_4 \\ \text{distinct}}} G_{i_1i_2i_3i_4}(\Omega) \\ \leq& Cn^{-4}\sum_{i_1,\cdots,i_4} \theta_{i_1}^2\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}^2 = Cn^{-4}\|\theta\|^8. \end{aligned}$$

Combining it with $C_4 \asymp n^{-4}\|\theta\|^8$, we get $\mathrm{Var}(\widehat{C}_4) = O(n^{-4}C_4)$.

Consider $\mathrm{Var}(\widehat{L}_2)$. By definitions and that $B_{n,m} \asymp n^m$, we bound

$$\mathbb{E}(\widehat{L}_2 - L_2)^2 \leq Cn^{-6}\mathbb{E}\Big[\sum_{i_1<i_2<i_3} (A_{i_1i_2}A_{i_2i_3} - \Omega_{i_1i_2}\Omega_{i_2i_3})\Big]^2. \quad (35)$$

Recall that when $i \neq j$, $A_{ij} = \Omega_{ij} + W_{ij}$. Since for any numbers x, y, a, b, $(a+x)(b+y) - ab = xy + ay + bx$, we can write

$$\sum_{i_1<i_2<i_3} (A_{i_1i_2}A_{i_2i_3} - \Omega_{i_1i_2}\Omega_{i_2i_3}) = I + II + III,$$

where

$$I = \sum_{i_1<i_2<i_3} W_{i_1i_2}W_{i_2i_3},$$

$$II = \sum_{i_1<i_2<i_3} \Omega_{i_1i_2}W_{i_2i_3},$$

and

$$III = \sum_{i_1<i_2<i_3} \Omega_{i_2i_3}W_{i_1i_2}.$$

Inserting this into (35) and using Cauchy-Schwarz inequality,

$$\mathbb{E}(\widehat{L}_2 - L_2)^2 \leq Cn^{-6}(\mathbb{E}[(I)^2] + \mathbb{E}[(II)^2] + \mathbb{E}[(III)^2]).$$

It then suffices to show

$$\mathbb{E}[(I)^2] \lesssim \|\theta\|_1^3\|\theta\|_3^3, \quad (36)$$

$$\mathbb{E}[(II)^2] \lesssim \|\theta\|_1^3\|\theta\|_3^3, \quad (37)$$

and

$$\mathbb{E}[(III)^2] \lesssim \|\theta\|_1^3\|\theta\|_3^3. \quad (38)$$

We now show (36)-(38) separately.

Consider (36). Note that for two sets of indices (i_1, i_2, i_3) and (j_1, j_2, j_3) such that $i_1 < i_2 < i_3$, $j_1 < j_2 < j_3$, by basic statistics, we have that when $(i_1, i_2, i_3) \neq (j_1, j_2, j_3)$,

$$\mathbb{E}[W_{i_1i_2}W_{i_2i_3}W_{j_1j_2}W_{j_2j_3}] = 0.$$

and when $(i_1, i_2, i_3) = (j_1, j_2, j_3)$,

$$\begin{aligned} \mathbb{E}[W_{i_1i_2}W_{i_2i_3}W_{j_1j_2}W_{j_2j_3}] &= E[W_{i_1i_2}^2W_{i_2i_3}^2] \\ &= \Omega_{i_1i_2}(1-\Omega_{i_1i_2})\Omega_{i_2i_3}(1-\Omega_{i_2i_3}). \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbb{E}[(I)]^2 &= \sum_{i_1<i_2<i_3}\sum_{j_1<j_2<j_3} E[W_{i_1i_2}W_{i_2i_3}W_{j_1j_2}W_{j_2j_3}] \\ &\leq \sum_{i_1<i_2<i_3} \mathbb{E}[W_{i_1i_2}^2W_{i_2i_3}^2] \\ &\leq \sum_{i_1<i_2<i_3} \Omega_{i_1i_2}(1-\Omega_{i_1i_2})\Omega_{i_2i_3}(1-\Omega_{i_2i_3}). \end{aligned}$$

Recall that for any $i < j$,

$$\Omega_{ij}(1-\Omega_{ij}) \leq \Omega_{ij} \leq \theta_i\theta_j,$$

it follows that

$$\mathbb{E}[(I)]^2 \leq \sum_{i_1<i_2<i_3} \theta_{i_1}\theta_{i_2}\theta_{i_2}\theta_{i_3} \leq \|\theta\|_1^2\|\theta\|_2^2,$$

and the claim follows by Cauchy Schwarz inequality that $\|\theta\|_1^2\|\theta\|_2^2 \leq \|\theta\|_1^2(\|\theta\|_1\|\theta\|_3^3) = \|\theta\|_1^3\|\theta\|_3^3$.

Consider (37)-(38). Since the proofs are similar, we only show (37).

Note that for two sets of indices (i_1, i_2, i_3) and (j_1, j_2, j_3) such that $i_1 < i_2 < i_3$, $j_1 < j_2 < j_3$, by basic statistics, we have that when $(i_2, i_3) \neq (j_2, j_3)$,

$$\mathbb{E}[W_{i_2i_3}W_{j_2j_3}] = 0.$$

and when $(i_2, i_3) = (j_2, j_3)$,

$$\mathbb{E}[W_{i_2i_3}W_{j_2j_3}] = \mathbb{E}[W_{i_2i_3}^2] = \Omega_{i_2i_3}(1 - \Omega_{i_2i_3}).$$

Therefore,

$$\begin{aligned}
\mathbb{E}[(II)^2] &= \sum_{i_1<i_2<i_3} \sum_{j_1<j_2<j_3} \mathbb{E}[\Omega_{i_1i_2}\Omega_{j_1j_2}W_{i_2i_3}W_{j_2j_3}] \\
&= \sum_{i_1<i_2<i_3} \mathbb{E}[\Omega_{i_1i_2}W_{i_2i_3} \sum_{j_1<j_2<j_3} (\Omega_{j_1j_2}W_{j_2j_3})] \\
&= \sum_{i_1<i_2<i_3} \mathbb{E}[\Omega_{i_1i_2}W_{i_2i_3}^2 (\sum_{j_1<i_2} \Omega_{j_1i_2})] \\
&= \sum_{i_1<i_2<i_3} \mathbb{E}[\Omega_{i_1i_2}\Omega_{i_2i_3}(1-\Omega_{i_2i_3})(\sum_{j_1<i_2} \Omega_{j_1i_2})]
\end{aligned}$$

Again by $\Omega_{ij}(1-\Omega_{ij}) \leq \Omega_{ij} \leq \theta_i\theta_j$ for any $i < j$, we find

$$\begin{aligned}
\mathbb{E}[(II)^2] &\leq \sum_{i_1<i_2<i_3} \mathbb{E}[\Omega_{i_1i_2}\Omega_{i_2i_3}(\sum_{j_1<i_2} \Omega_{j_1i_2})] \\
&\leq \sum_{i_1<i_2<i_3} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}(\sum_{j_1<i_2} \theta_{j_1}\theta_{i_2}) \\
&\leq \|\theta\|_1^3\|\theta\|_3^3.
\end{aligned}$$

Last, we prove the claim on $\mathrm{Var}(\widehat{L}_3)$. It suffices to control the covariance between $(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4})$ and $(A_{j_1j_2}A_{j_2j_3}A_{j_3j_4})$. To be more specific, define the set

$$\mathcal{J} = \big\{(i_1, i_2), (i_2, i_3), (i_3, i_4), (j_1, j_2), (j_2, j_3), (j_3, j_4)\big\},$$

whose elements are pairs of unordered integers, i.e. we treat (i_1, i_2) and (i_2, i_1) as the same element.

Let $|\mathcal{J}|$ be the number of distinct elements of $\mathcal{J}$, where $3 \leq |\mathcal{J}| \leq 6$ under the condition that $i_1 < i_2 < i_3 < i_4$ and $j_1 < j_2 < j_3 < j_4$. To control the variance of $\widehat{L}_3$, it suffices to bound the following quantity

$$\sum_{s=3}^{6} \sum_{|\mathcal{J}|=s} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}).$$

Furthermore, it suffices to show for $3 \leq s \leq 6$,

$$\sum_{|\mathcal{J}|=s} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \lesssim \|\theta\|_1^4\|\theta\|_3^6. \tag{39}$$

When $|\mathcal{J}| = 6$, it's not hard to see $(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4})$ and $(A_{j_1j_2}A_{j_2j_3}A_{j_3j_4})$ are independent because the six elements in $\mathcal{J}$ are all distinct, which indicates

$$\sum_{|\mathcal{J}|=6} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) = 0.$$

The following basic property is frequently used in the discussion of remaining cases. For non-negative random variables X and Y, we have

$$\mathrm{Cov}(X, Y) \leq \mathbb{E}[XY]. \tag{40}$$

Consider the case where $|\mathcal{J}| = 5$. By symmetry, it's enough to consider three situations where $(i_1, i_2) = (j_1, j_2)$, $(i_1, i_2) = (j_2, j_3)$ and $(i_2, i_3) = (j_2, j_3)$, separately.

If $(i_1, i_2) = (j_1, j_2)$, we have

$$\begin{aligned}
&\sum_{(i_1,i_2)=(j_1,j_2)} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \\
\leq &\sum_{(i_1,i_2)=(j_1,j_2)} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}A_{j_2j_3}A_{j_3j_4}\big] \\
\leq C &\sum_{(i_1,i_2)=(j_1,j_2)} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}\theta_{j_2}\theta_{j_3}^2\theta_{j_4} \\
\leq C &\sum \theta_{i_1}\theta_{i_2}^3\theta_{i_3}^2\theta_{i_4}\theta_{j_3}^2\theta_{j_4} \\
= C &\|\theta\|_1^3\|\theta\|^4\|\theta\|_3^3 \leq C\|\theta\|_1^4\|\theta\|_3^6,
\end{aligned}$$

where the last inequality is due to $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$ by Cauchy-Schwarz inequality.

If $(i_1, i_2) = (j_2, j_3)$, we have

$$\begin{aligned}
&\sum_{(i_1,i_2)=(j_2,j_3)} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \\
\leq &\sum_{(i_1,i_2)=(j_2,j_3)} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}A_{j_1j_2}A_{j_3j_4}\big] \\
\leq C &\sum_{(i_1,i_2)=(j_2,j_3)} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}\theta_{j_1}\theta_{j_2}\theta_{j_3}\theta_{j_4} \\
= C &\sum_{i_1,\cdots,i_4,j_1,j_4} \theta_{i_1}^2\theta_{i_2}^3\theta_{i_3}^2\theta_{i_4}\theta_{j_1}\theta_{j_4} \\
= C &\|\theta\|_1^3\|\theta\|^4\|\theta\|_3^3 \leq C\|\theta\|_1^4\|\theta\|_3^6,
\end{aligned}$$

where the last inequality is due to $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$.

If $(i_2, i_3) = (j_2, j_3)$, we have

$$\begin{aligned}
&\sum_{(i_2,i_3)=(j_2,j_3)} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \\
\leq &\sum_{(i_2,i_3)=(j_2,j_3)} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}A_{j_1j_2}A_{j_3j_4}\big] \\
\leq &C \sum_{(i_2,i_3)=(j_2,j_3)} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}\theta_{j_1}\theta_{j_2}\theta_{j_3}\theta_{j_4} \\
= &C \sum_{i_1,\cdots,i_4,j_1,j_4} \theta_{i_1}\theta_{i_2}^3\theta_{i_3}^3\theta_{i_4}\theta_{j_1}\theta_{j_4} = C\|\theta\|_1^4\|\theta\|_3^6.
\end{aligned}$$

Combining above three inequalities, we derive

$$\sum_{|\mathcal{J}|=5} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \leq C\|\theta\|_1^4\|\theta\|_3^6.$$

Consider the case where $|\mathcal{J}| = 4$. By symmetry, $\mathcal{J}$ either equals to $\mathcal{J}_1 = \big\{(i_1, i_2), (i_2, i_3), (i_3, i_4), (j_1, j_2)\big\}$ or $\mathcal{J}_2 = \big\{(i_1, i_2), (i_2, i_3), (i_3, i_4), (j_2, j_3)\big\}$.

Therefore, we decompose and bound

$$\begin{aligned}
&\sum_{|\mathcal{J}|=4} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \\
\lesssim &\sum_{\mathcal{J}_1} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \\
&+ \sum_{\mathcal{J}_2} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \\
\leq &\sum_{\mathcal{J}_1} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}A_{j_1j_2}\big] + \sum_{\mathcal{J}_2} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}A_{j_2j_3}\big]
\end{aligned}$$

It then suffices to show

$$\sum_{\mathcal{J}_1} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}A_{j_1j_2}\big] \leq C\|\theta\|_1^4\|\theta\|_3^6, \qquad (41)$$

and

$$\sum_{\mathcal{J}_2} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}A_{j_2j_3}\big] \leq C\|\theta\|_1^4\|\theta\|_3^6, \qquad (42)$$

For (41), j_2 must equal to one of $i_1, \cdots, i_4$ since (j_2, j_3) equals to some (i_s, i_{s+1}) by definition of $\mathcal{J}_1$. By symmetry, we only need to consider $j_2 = i_1$ and $j_2 = i_2$. Again by $\Omega_{ij} \leq \theta_i\theta_j$, we obtain

$$\begin{aligned}
&\sum_{\mathcal{J}_1} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}A_{j_1j_2}\big] \\
\leq &\sum_{j_2=i_1} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}\theta_{j_1}\theta_{j_2} + \sum_{j_2=i_2} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}\theta_{j_1}\theta_{j_2} \\
\leq &\sum \theta_{i_1}^2\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}\theta_{j_1} + \sum \theta_{i_1}\theta_{i_2}^3\theta_{i_3}^2\theta_{i_4}\theta_{j_1} \\
= &\|\theta\|_1^2\|\theta\|^6 + C\|\theta\|_1^3\|\theta\|^2\|\theta\|_3^3 \\
\leq &\|\theta\|_1^4\|\theta\|_3^6.
\end{aligned}$$

Here we explain the last inequality. By Cauchy-Schwartz inequality, $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$. Combining with (7) that $\|\theta\| \to \infty$, $\|\theta\|_1^2\|\theta\|^6 \lesssim \|\theta\|_1^2\|\theta\|^8 \leq \|\theta\|_1^4\|\theta\|_3^6$. Moreover, $\|\theta\|_1^3\|\theta\|^2\|\theta\|_3^3 \leq \|\theta\|_1^3\|\theta\|_3^3(\|\theta\|^4) \leq \|\theta\|_1^4\|\theta\|_3^6$.

For (42), we similarly found j_2, j_3 must equal to some $i_1, \cdots, i_4$. By (7), $\theta_{j_3} \leq C$. Thus we only need to discuss the cases where $j_2 = i_1$ or $j_2 = i_2$.

$$\begin{aligned}
&\sum_{\mathcal{J}_2} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}A_{j_2j_3}\big] \\
\leq &\sum_{j_2=i_1} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}\theta_{j_2} + \sum_{j_2=i_2} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}\theta_{j_2} \\
\leq &\sum \theta_{i_1}^2\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}\theta_{j_2} + \sum \theta_{i_1}\theta_{i_2}^3\theta_{i_3}^2\theta_{i_4}\theta_{j_2} \\
= &C\|\theta\|_1^2\|\theta\|^6 + C\|\theta\|_1^3\|\theta\|^2\|\theta\|^3 \leq C\|\theta\|_1^4\|\theta\|_3^6,
\end{aligned}$$

where the last inequality has been explained in the proof of (41).

Combining (41) and (42), we bound

$$\sum_{|\mathcal{J}|=4} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \leq C\|\theta\|_1^4\|\theta\|_3^6.$$

Finally, consider the case where $|\mathcal{J}| = 3$. In this case, the covariance is in fact variance. Therefore,

$$\begin{aligned}
&\sum_{|\mathcal{J}|=3} \mathrm{Cov}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}, A_{j_1j_2}A_{j_2j_3}A_{j_3j_4}) \\
= &\sum_{i_1,\cdots,i_4} \mathrm{Var}(A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}) \\
\leq &\sum_{i_1,\cdots,i_4} \mathbb{E}\big[A_{i_1i_2}A_{i_2i_3}A_{i_3i_4}\big] \\
\leq &\sum_{i_1,\cdots,i_4} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4} \\
= &C\|\theta\|_1^2\|\theta\|^2 \lesssim C\|\theta\|_1^2\|\theta\|^6 \leq C\|\theta\|_1^4\|\theta\|_3^6,
\end{aligned}$$

where the second last inequality is by (7) that $\|\theta\| \to \infty$ and last inequality is Cauchy-Schwarz inequality.

This proves (39). $\square$

D.3. Proof of Lemma B.2

Write for short $T_n = \frac{\sqrt{B_{n,4}}}{\sqrt{C_4}}(\widehat{C}_4 - C_4)$. We introduce some useful notations. For any $n \times n$ matrix M and distinct indices (i_1, i_2, i_3, i_4), define

$$G_{i_1i_2i_3i_4}(M) = M_{i_1i_2}M_{i_2i_3}M_{i_3i_4}M_{i_4i_1},$$
$$G(M) = \sum_{(i_1,i_2,i_3,i_4)\in CC(I_n)} G_{i_1i_2i_3i_4}(M).$$

Additionally, let $W = A - \Omega$ and let Ω^* be the matrix where $\Omega^*_{ij} = \Omega_{ij}(1 - \Omega_{ij})$ for all $1 \leq i, j \leq n$. We now rewrite

$$T_n = \frac{G(A) - G(\Omega)}{\sqrt{G(\Omega)}}, \qquad S_{n,n} = \frac{G(W)}{\sqrt{G(\Omega^*)}}. \qquad (43)$$

Therefore,

$$T_n - S_{n,n} = \frac{G(A) - G(\Omega) - G(W)}{\sqrt{G(\Omega)}} + S_{n,n}\left[\frac{\sqrt{G(\Omega^*)}}{\sqrt{G(\Omega)}} - 1\right] \equiv J_1 + S_{n,n} \cdot J_2.$$

In the proof of Theorem 3.2, we have shown $S_{n,n} \xrightarrow{d} N(0,1)$. Hence, to show $(T_n - S_{n,n}) \xrightarrow{p} 0$, it suffices to show that

$$J_1 \xrightarrow{p} 0 \tag{44}$$

and

$$J_2 \to 0. \tag{45}$$

First, we prove (44). We can decompose $G_{i_1i_2i_3i_4}(A) - G_{i_1i_2i_3i_4}(\Omega) - G_{i_1i_2i_3i_4}(W)$ as the sum of three terms

$$\begin{aligned}
\Delta^{(1)}_{i_1i_2i_3i_4} &= W_{i_1i_2}\Omega_{i_2i_3}\Omega_{i_3i_4}\Omega_{i_4i_1} + \Omega_{i_1i_2}W_{i_2i_3}\Omega_{i_3i_4}\Omega_{i_4i_1} \\
&\quad + \Omega_{i_1i_2}\Omega_{i_2i_3}W_{i_3i_4}\Omega_{i_4i_1} + \Omega_{i_1i_2}\Omega_{i_2i_3}\Omega_{i_3i_4}W_{i_4i_1}, \\
\Delta^{(2)}_{i_1i_2i_3i_4} &= W_{i_1i_2}W_{i_2i_3}\Omega_{i_3i_4}\Omega_{i_4i_1} + W_{i_1i_2}\Omega_{i_2i_3}W_{i_3i_4}\Omega_{i_4i_1} \\
&\quad + W_{i_1i_2}\Omega_{i_2i_3}\Omega_{i_3i_4}W_{i_4i_1} + \Omega_{i_1i_2}W_{i_2i_3}W_{i_3i_4}\Omega_{i_4i_1} \\
&\quad + \Omega_{i_1i_2}W_{i_2i_3}\Omega_{i_3i_4}W_{i_4i_1} + \Omega_{i_1i_2}\Omega_{i_2i_3}W_{i_3i_4}W_{i_4i_1}, \\
\Delta^{(3)}_{i_1i_2i_3i_4} &= \Omega_{i_1i_2}W_{i_2i_3}W_{i_3i_4}W_{i_4i_1} + W_{i_1i_2}\Omega_{i_2i_3}W_{i_3i_4}W_{i_4i_1} \\
&\quad + W_{i_1i_2}W_{i_2i_3}\Omega_{i_3i_4}W_{i_4i_1} + W_{i_1i_2}W_{i_2i_3}W_{i_3i_4}\Omega_{i_4i_1}.
\end{aligned}$$

It is easy to see that

$$\mathbb{E}\left[\sum_{CC(I_n)} \Delta^{(1)}_{i_1i_2i_3i_4}\right] = 0. \tag{46}$$

We then study the variance of this term. Note that the four terms in $\Delta^{(1)}_{i_1i_2i_3i_4}$ are independent of each other. Let (j,s,m,ℓ) be any cycle on the four nodes $\{i_1,i_2,i_3,i_4\}$. Then, the variance of $W_{js}\Omega_{sm}\Omega_{m\ell}\Omega_{\ell j}$ is bounded by $\Omega_{js}\Omega^2_{sm}\Omega^2_{m\ell}\Omega^2_{\ell j} = O(\theta_j^3\theta_s^3\theta_m^4\theta_\ell^4)$. Hence,

$$\begin{aligned}
\sum_{CC(I_n)} \mathrm{Var}(\Delta^{(1)}_{i_1i_2i_3i_4}) &\le C\sum_{j,s,m,\ell} \theta_j^3\theta_s^3\theta_m^4\theta_\ell^4 \\
&\le C\|\theta\|_3^6\|\theta\|_4^8 = o(\|\theta\|_3^6\|\theta\|^8),
\end{aligned}$$

where the last inequality is from the condition (7) and the fact that $\|\theta\|_4^4 = (\sum_i \theta_i^4) \le \theta^2_{\max}(\sum_i \theta_i^2) = O(\|\theta\|^2) = o(\|\theta\|^4)$. We then look at the covariance between $\Delta^{(1)}_{i_1i_2i_3i_4}$ and $\Delta^{(1)}_{i_1'i_2'i_3'i_4'}$. Let (j,s,m,ℓ) be any cycle on the four nodes $\{i_1,i_2,i_3,i_4\}$, and let (j',s',m',ℓ') be any cycle on the four nodes $\{i_1',i_2',i_3',i_4'\}$. As long as $\{j,s\} \neq \{j',s'\}$, the two terms $W_{js}\Omega_{sm}\Omega_{m\ell}\Omega_{\ell j}$ and $W_{j's'}\Omega_{s'm'}\Omega_{m'\ell'}\Omega_{\ell' j'}$ are independent, hence, their covariance is zero. If $\{j,s\} = \{j',s'\}$, their covariance is bounded by $\Omega_{js} \cdot \Omega_{sm}\Omega_{m\ell}\Omega_{\ell j}\Omega_{s'm'}\Omega_{m'\ell'}\Omega_{\ell' j} = O(\theta_j^3\theta_s^3\theta_m^2\theta_\ell^2\theta_{m'}^2\theta_{\ell'}^2)$. As a result,

$$\begin{aligned}
&\sum_{CC(I_n)\times CC(I_n)} \mathrm{Cov}(\Delta^{(1)}_{i_1i_2i_3i_4}, \Delta^{(1)}_{i_1'i_2'i_3'i_4'}) \\
\le\, & C\sum_{j,s,m,\ell,m',\ell'} \theta_j^3\theta_s^3\theta_m^2\theta_\ell^2\theta_{m'}^2\theta_{\ell'}^2 \le C\|\theta\|_3^6\|\theta\|^8.
\end{aligned}$$

Note that $G(\Omega) \asymp n^4C_4 \asymp \|\theta\|^8$ by Lemma B.1. Additionally, from the condition (7), $\|\theta\|_3 = o(1)$. Hence, the above imply

$$\mathrm{Var}\left(\sum_{CC(I_n)} \Delta^{(1)}_{i_1i_2i_3i_4}\right) \ll G(\Omega). \tag{47}$$

Combining (46)-(47) gives

$$\frac{1}{\sqrt{G(\Omega)}}\sum_{CC(I_n)} \Delta^{(1)}_{i_1i_2i_3i_4} \xrightarrow{p} 0. \tag{48}$$

We can consider other terms similarly. By direct calculations,

$$\begin{aligned}
\mathrm{Var}\left(\sum_{CC(I_n)} \Delta^{(2)}_{i_1i_2i_3i_4}\right) &\le \sum_{j,s,m,\ell} \Omega_{js}\Omega_{sm}\Omega^2_{m\ell}\Omega^2_{\ell j} \\
&\quad + \sum_{\substack{j,s,m\\ \ell,\ell'}} \Omega_{js}\Omega_{sm}\Omega_{m\ell}\Omega_{m,\ell'}\Omega_{\ell j}\Omega_{\ell' j} \\
&\le C\sum_{j,s,m,\ell} \theta_j^3\theta_s^2\theta_m^3\theta_\ell^4 + C\sum_{\substack{j,s,m\\ \ell,\ell'}} \theta_j^3\theta_s^2\theta_m^3\theta_\ell^2\theta_{\ell'}^2 \\
&\le C\|\theta\|_3^6\|\theta\|^2\|\theta\|_4^4 + C\|\theta\|_3^6\|\theta\|^6 = o(\|\theta\|^8),
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}\left(\sum_{CC(I_n)} \Delta^{(3)}_{i_1i_2i_3i_4}\right) &\le \sum_{j,s,m,\ell} \Omega^2_{js}\Omega_{sm}\Omega_{m\ell}\Omega_{\ell j} \\
&\le C\sum_{j,s,m,\ell} \theta_j^3\theta_s^3\theta_m^2\theta_\ell^2 \\
&\le C\|\theta\|_3^6\|\theta\|^4 = o(\|\theta\|^8).
\end{aligned}$$

Hence, for the terms related to $\Delta^{(2)}_{i_1i_2i_3i_4}$ and $\Delta^{(3)}_{i_1i_2i_3i_4}$, we also have a similar convergence as that of (48). These together imply $J_1 \xrightarrow{p} 0$. Hence, (44) is true.

Next, we prove (45). It is seen that

$$\begin{aligned}
0 \le G(\Omega) - G(\Omega^*) &\le C\sum_{j,s,m,\ell} \Omega^2_{js}\Omega_{sm}\Omega_{m\ell}\Omega_{mj} \\
&\le C\sum_{j,s,m,\ell} \theta_j^3\theta_s^3\theta_m^2\theta_\ell^2 \le C\|\theta\|_3^6\|\theta\|^4 = o(\|\theta\|^8).
\end{aligned}$$

As a result, $|G(\Omega^*)/G(\Omega) - 1| = o(1)$. This proves (45). □

D.4. Proof of Proposition A.1

The last item follows once the first three items are proved, so we only consider the first three items.

Consider the first item. Write

$$1'A^2 1 = \sum_{1 \le i_1, i_2, i_3 \le n} A_{i_1 i_2} A_{i_2 i_3}.$$

Recall that all diagonal entries of A are 0, we can exclude the case $i_1 = i_2$ or $i_2 = i_3$ from the summation. Therefore, we only need to sum over either the cases where i_1, i_2, i_3 are distinct and the cases $i_1 = i_3$ but $i_1 \neq i_2$. It follows

$$1'A^2 1 = \Bigg(\sum_{\substack{i_1, i_2, i_3 \\ \text{are distinct}}} + \sum_{\substack{i_1, i_2, i_3 \\ i_1 = i_3, i_1 \neq i_2}} \Bigg) A_{i_1 i_2} A_{i_2 i_3} = I + II.$$

Now, first, by definition,

$$I = B_{n,3} \widehat{L}_2, \qquad \text{where } B_{n,3} = 6\tbinom{n}{3},$$

and second (recall all diagonal entries of A are 0),

$$II = \sum_{i_1, i_2} A^2_{i_1 i_2} = \mathrm{tr}(A^2).$$

Combining these gives

$$\widehat{L}_2 = \frac{1}{6\binom{n}{3}} (1'A^2 1 - \mathrm{tr}(A^2)),$$

and the claim follows.

Consider the second item. Using similar arguments, we decompose

$$1'A^3 1 = \sum_{1 \le i_1, i_2, i_3, i_4 \le n} A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} = I + II + III + IV,$$

where

$$I = \sum_{\substack{i_1, i_2, i_3, i_4 \\ \text{are distinct}}} A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} = B_{n,4} \widehat{L}_3,$$

with $B_{n,4} = 24\binom{n}{4}$,

$$II = \Bigg(\sum_{\substack{i_1, i_2, i_3, i_4 \\ i_1 = i_3}} + \sum_{\substack{i_1, i_2, i_3, i_4 \\ i_2 = i_4}} \Bigg) A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} = 2 \cdot (1'A^2 1),$$

$$III = \sum_{\substack{i_1, i_2, i_3, i_4 \\ i_1 = i_4}} A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} = 1'A^3 1,$$

and

$$IV = - \sum_{\substack{i_1, i_2, i_3, i_4 \\ i_1 = i_3, i_2 = i_4}} A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} = -1'A1.$$

Combining these gives

$$\widehat{L}_3 = \frac{1}{24\binom{n}{4}} [1'A^3 1 - 2 \cdot 1'A^2 1 + 1'A1 - \mathrm{tr}(A^3)],$$

and the claim follows.

Consider the third item. Note first

$$\mathrm{tr}(A^4) = \sum_{1 \le i_1, i_2, i_3, i_4 \le n} A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} A_{i_4 i_1}.$$

Similarly, we have

$$\mathrm{tr}(A^4) = \sum_{i_1, i_2, i_3, i_4} A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} A_{i_4 i_1} = I + II + III,$$

where

$$I = \sum_{\substack{i_1, i_2, i_3, i_4 \\ \text{are distinct}}} A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} A_{i_4 i_1} = 24 \binom{n}{4} \widehat{C}_4,$$

$$II = \Bigg(\sum_{\substack{i_1, i_2, i_3, i_4 \\ i_1 = i_3}} + \sum_{\substack{i_1, i_2, i_3, i_4 \\ i_2 = i_4}} \Bigg) A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} A_{i_4 i_1} = 2 \cdot (1'A^2 1)$$

and

$$III = - \sum_{\substack{i_1, i_2, i_3, i_4 \\ i_1 = i_3, i_2 = i_4}} A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_4} A_{i_4 i_1} = -1'A1.$$

Combining these gives

$$\widehat{C}_4 = \frac{1}{24\binom{n}{4}} \Big(tr(A^4) - 2 \cdot 1'A^2 1 + 1'A1 \Big),$$

and the claim follows. □

References

Hall, Peter and Heyde, Christopher C. *Martingale limit theory and its application*. Academic press, 2014.

Jin, Jiashun, Ke, Zheng Tracy, and Luo, Shengming. Network global testing by counting graphlets. *Manuscript*, 2018.