SUPPLEMENT: PHASE TRANSITION FOR DETECTING A SMALL COMMUNITY IN A LARGE NETWORK

Jiashun Jin
Carnegie Mellon University
jiashun@stat.cmu.edu

Zheng Tracy Ke
Harvard University
zke@fas.harvard.edu

Paxton Turner
Harvard University
paxtonturner@fas.harvard.edu

Anru R. Zhang
Duke University
anru.zhang@duke.edu

CONTENTS

A Additional experiments **2**

A.1 Visualization of Carroll's network 2

A.2 SgnQ vs. Scan 2

A.3 χ_2 vs. SgnQ 3

B Proof of Lemma 2.1 (Identifiability) **3**

C Proof of Theorem 2.1 (Limiting null of the SgnQ statistic) **5**

D Proof of Lemma 2.2 (Properties of $\tilde{\Omega}$) **6**

E Proof of Theorem 2.2 (Power of the SgnQ test) and Corollary 2.1 **6**

E.1 Setup and results 6

E.2 Preliminary bounds 10

E.3 Mean and variance of SgnQ 12

E.3.1 Ideal SgnQ 13

E.3.2 Proxy SgnQ 13

E.3.3 Real SgnQ 15

E.4 Proofs of Lemmas E.7–E.12 15

E.4.1 Proof strategy 15

E.4.2 Proof of Lemma E.7 17

E.4.3 Proof of Lemma E.8 18

E.4.4 Proof of Lemma E.9 21

E.4.5 Proof of Lemma E.10 29

E.4.6 Proof of Lemma E.11 34

E.4.7 Proof of Lemma E.12 38

E.5 Calculations in the SBM setting 42

F Proof of Theorem 2.3 (Powerlessness of χ^2 test) 42

G Proof of Theorem 2.4 (Statistical lower bound) 47

G.1 Proof of Lemma G.1 51

H Proof of Theorem 2.5 (Tightness of the statistical lower bound) 51

I Proof of Theorem 2.6 (Computational lower bound) 61

J Proof of Theorem 2.7 (Power of EST) 63

A ADDITIONAL EXPERIMENTS

A.1 VISUALIZATION OF CARROLL'S NETWORK

In Figure 1, we display a subgraph of high-degree nodes of Raymond Carroll's personalized coauthorship network (figure borrowed with permission from Ji et al. (2022)). On the right of Figure 1 is shown the small community extracted by SCORE, and this cluster of size 17 is labeled by author names.

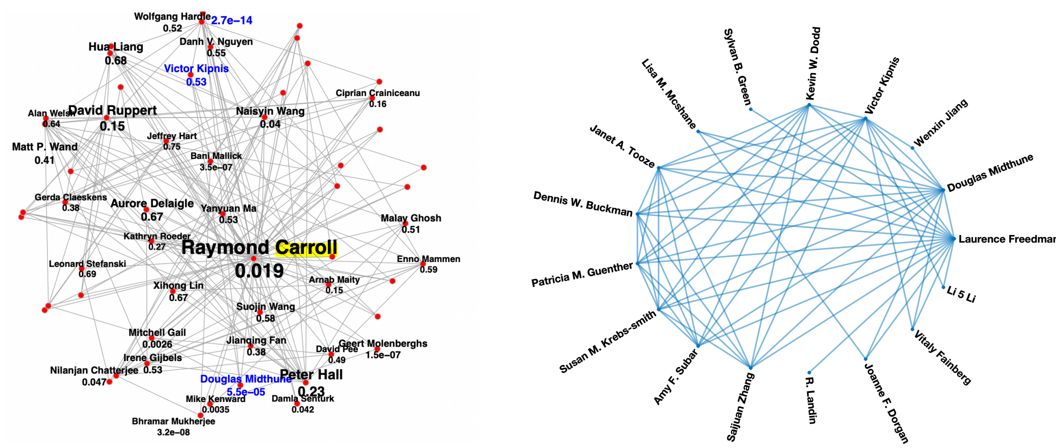

Figure 1: **Left:** Carroll's personalized network, figure taken from Ji et al. (2022). **Right:** A small community of 17 authors extracted by SCORE and whose SgnQ p-value is 0.6818.

A.2 SGNQ VS. SCAN

In this section we demonstrate evidence of a statistical-computational gap by means of numerical experiments.

We consider a SBM null and alternative model (as in Example 2 with $\theta \equiv 1$) with

$$P_0 = \begin{pmatrix} \alpha & \alpha \\ \alpha & \alpha \end{pmatrix}, \qquad P_1 = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

where $aN + b(n-N) = \alpha$. For this simple testing problem, we compare the power of SgnQ and the scan test. In our experiments, we set $\alpha = 0.2$ and allow the parameter a to vary from $a = \alpha$ to $a = a_{\max} \equiv an/N$. Once a and α are fixed, the parameters b and c are determined by

$$c = \frac{aN^2 + \alpha n^2 - 2\alpha n N}{(n-N)^2},$$

$$b = \frac{nc - (a+c)N}{n - 2N}.$$

In particular, $a_{\max}$ is the largest value of a such that $b \geq 0$.

Since the scan test ϕ_{sc} we defined is extremely computationally expensive, we study the power of an 'oracle' scan test $\tilde{\phi}_{sc}$ which knows the location of the true planted subset $\mathcal{C}_1$. The power of the oracle scan test is computed as follows. Let κ denote the desired level.

1. Using M_{cal} repetitions under the null, we calculate the (non-oracle) scan statistic $\phi_{sc}^{(1)}, \ldots, \phi_{sc}^{(M_{cal})}$ for each repetition. We set the threshold $\hat{\tau}$ to be the empirical $1 - \kappa$ quantile of $\phi_{sc}^{(1)}, \ldots, \phi_{sc}^{(M_{cal})}$.
2. Given a sample from the alternative model, we compute the power using M_{pow} repetitions, where we reject if
$$\tilde{\phi}_{sc} \equiv \mathbf{1}_{\mathcal{C}_1}(A - \hat{\eta}\hat{\eta}')\mathbf{1}_{\mathcal{C}_1} > \hat{\tau}.$$

In our experiments, we set $M_{cal} = 75$ and $M_{pow} = 200$.

Note that since $\tilde{\phi}_{sc} \leq \phi_{sc}$, the procedure above gives an underestimate of the power of the scan test (provide the threshold is correctly calibrated), which is helpful since this can be used to show evidence of a statistical-computational gap.

In our plots we also indicate the statistical (information-theoretic) and computational thresholds in addition to the power. Inspired by the sharp characterization of the statistical threshold in (Arias-Castro & Verzelen, 2014, Equation (10)) for planted dense subgraph, in all plots we draw a black vertical dashed line at the first value of a such that

$$(1/2)\sqrt{N}(a - c)/\sqrt{c(1 - c)} > 1.$$

We draw a blue vertical dashed line at the first value of a such that

$$N(a - c)/\sqrt{nc} > 1.$$

A.3 χ_2 vs. SgnQ

We also show additional experiments demonstrating the effect of degree-matching on the power of the χ^2 test. We compute the power with respect to the following alternative models (as in Example 2 with $\theta \equiv 1$) with

$$P^{(1)} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \qquad P^{(2)} = \begin{pmatrix} a & c \\ c & c \end{pmatrix}$$

where $b = \frac{cn-(a+c)N}{n-2N}$, c is fixed, and a ranges from c to $a'_{\max} = c(n - N)/N$ for the experiments with $P^{(1)}$. Similar to before, $a'_{\max}$ is the largest value of a such that $b \geq 0$. See Figure 3 for further details.

B Proof of Lemma 2.1 (Identifiability)

To prove identifiability, we make use of the following result from (Jin et al., 2021, Lemma 3.1), which is in line with Sinkhorn's work Sinkhorn (1974) on matrix scaling.

Lemma B.1 (Jin et al. (2021))**.** *Given a matrix $A \in \mathbb{R}^{K,K}$ with strictly positive diagonal entries and non-negative off-diagonal entries, and a strictly positive vector $h \in \mathbb{R}^K$, there exists a unique diagonal matrix $D = \mathrm{diag}(d_1, d_2, \ldots, d_K)$ such that $DADh = 1_K$ and $d_k > 0$, $1 \leq k \leq K$.*

We apply Lemma B.1 with $h = (h_1, \ldots, h_K)'$ and $A = P$ to construct a diagonal matrix $D = \mathrm{diag}(d_1, \ldots, d_K)$ satisfying $DADh = 1_K$. Note that P has positive diagonal entries since Ω does.

Define $P^* = DPD$ and $D^* = \mathrm{diag}(d_1^*, \ldots, d_n^*) \in \mathbb{R}^n$ where

$$d_i^* \equiv d_k \qquad \text{if } i \in \mathcal{C}_k$$

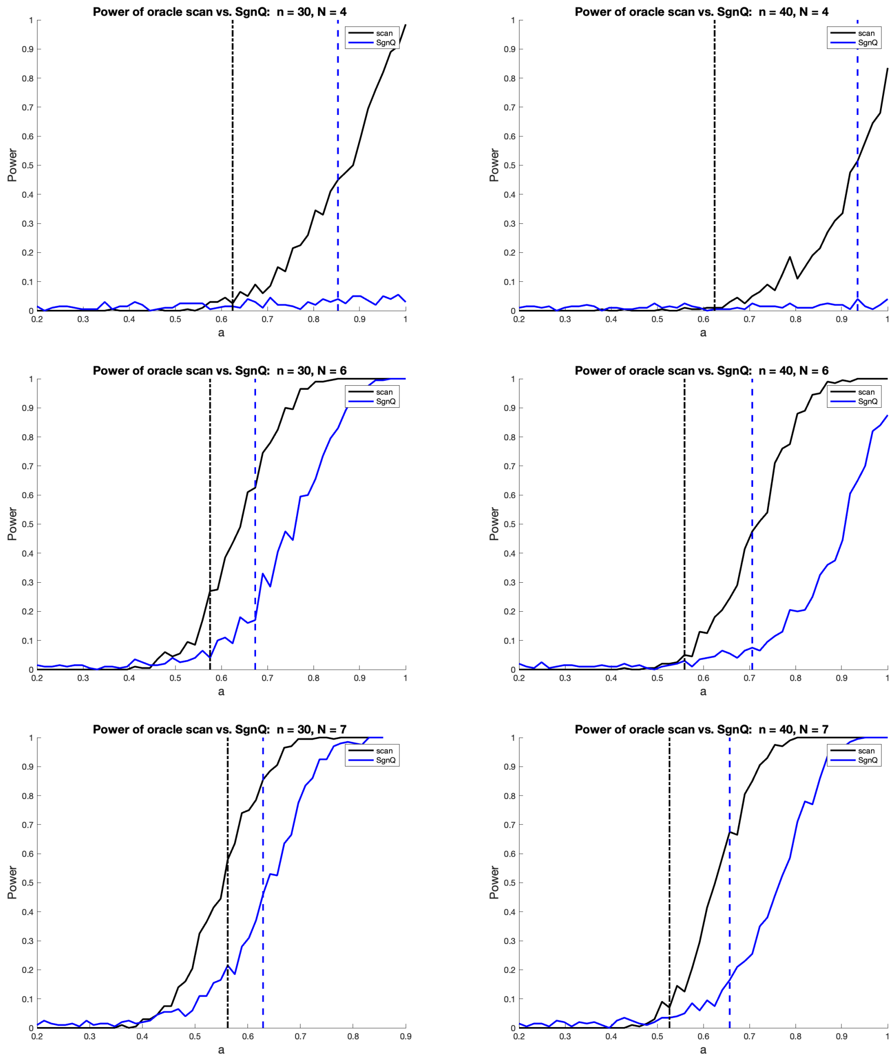

Figure 2: The power of SgnQ (blue curve) and oracle scan (black curve) for $n = 30, N \in \{4, 6, 7\}$ (left) and $n = 40, N \in \{4, 6, 7\}$ (right). The black dashed line indicates the theoretical statistical threshold, and the blue dashed line indicates the theoretical computational threshold.

Observe that

$$\Pi D^{-1} = (D^*)^{-1}\Pi.$$

Define $\Theta^* = \Theta(D^*)^{-1}$, and let $\theta^* = \text{diag}(\Theta^*)$. Next, let $\overline{\Theta} = \frac{n}{\|\theta^*\|_1} \cdot \Theta^*$, let $\overline{\theta} = \text{diag}(\overline{\Theta})$, and let $\overline{P} = \frac{\|\theta^*\|_1^2}{n^2} \cdot P^*$. Note that $\|\overline{\theta}\|_1 = n$ and $\overline{P}h \propto \mathbf{1}_K$.

Using the previous definitions and observations, we have

$$\Omega = \Theta\Pi D^{-1} DPDD^{-1}\Pi'\Theta = \Theta^*\Pi P^*\Pi'\Theta^* = \overline{\Theta}\Pi\overline{P}\Pi'\overline{\Theta}$$

which justifies existence.

To justify uniqueness, suppose that

$$\Omega = \Theta^{(1)}\Pi P^{(1)}\Pi'\Theta^{(1)} = \Theta^{(2)}\Pi P^{(2)}\Pi'\Theta^{(2)},$$

where $\theta^{(i)} = \text{diag}(\Theta^{(i)})$ satisfy $\|\theta^{(i)}\|_1 = n$ for $i = 1, 2$ and

$$P^{(1)}h \propto \mathbf{1}_K, \qquad P^{(2)}h \propto \mathbf{1}_K.$$

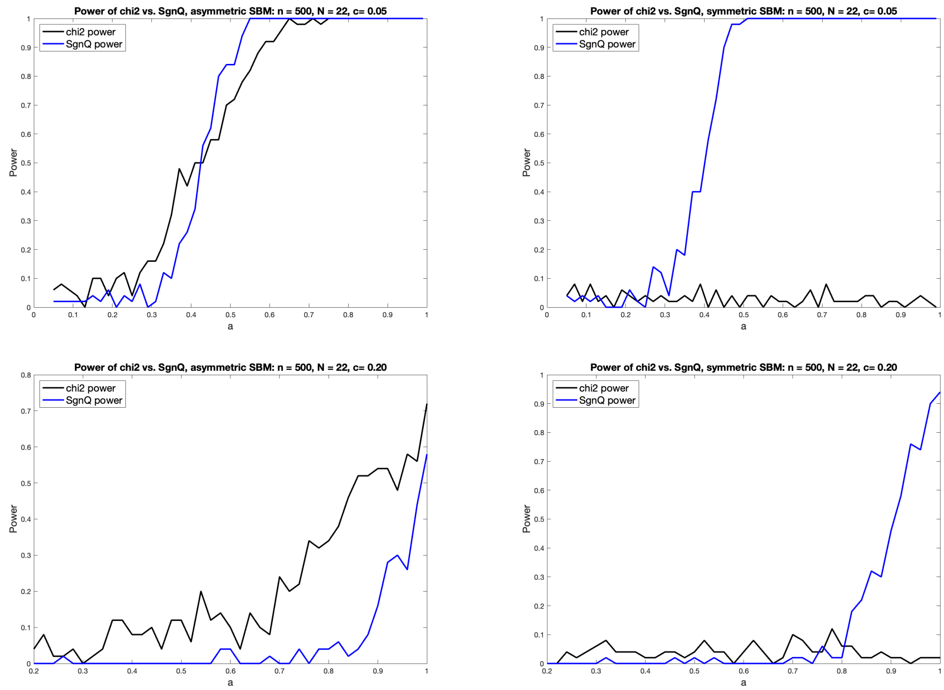

Figure 3: Power comparison of SgnQ and χ^2 ($n = 500$, $N = 22$, 50 repetitions). We consider a 2-community SBM with $P_{11} = a$, $P_{22} = c$, $P_{12} = c$ (left) and $P_{12} = \frac{an-(a+c)N}{n}$ (right plot, the case of degree matching) where $c = 0.05$ (top row) and $c = 0.20$ (bottom row).

Observe that

$$\Pi P^{(1)}\Pi'\mathbf{1}_n = \alpha^{(1)}n \cdot \mathbf{1}_n, \qquad \Pi P^{(2)}\Pi'\mathbf{1}_n = \alpha^{(2)}n \cdot \mathbf{1}_n.$$

for positive constants $\alpha^{(i)}, i \in \{1, 2\}$. Since Ω has nonnegative entries and positive diagonal elements, by Lemma B.1, there exists a unique diagonal matrix D such that

$$D\Omega D\mathbf{1}_n = \mathbf{1}_n.$$

We see that taking $D = \frac{1}{\sqrt{\alpha^{(i)}n}}(\Theta^{(i)})^{-1}$ satisfies this equation for $i = 1, 2$, and therefore by uniqueness,

$$\frac{1}{\sqrt{\alpha^{(1)}n}}(\Theta^{(1)})^{-1} = \frac{1}{\sqrt{\alpha^{(2)}n}}(\Theta^{(2)})^{-1}.$$

Since $\|\theta^{(1)}\|_1 = \|\theta^{(2)}\|_1 = n$, further we have $\alpha^{(1)} = \alpha^{(2)}$, and hence

$$\Theta^{(1)} = \Theta^{(2)}.$$

It follows that

$$\Pi P^{(1)}\Pi' = \Pi P^{(2)}\Pi',$$

which, since we assume $h_i > 0$ for $i = 1, \dots, K$, further implies that $P^{(1)} = P^{(2)}$. $\square$

C Proof of Theorem 2.1 (Limiting null of the SgnQ statistic)

Consider a null DCBM with $\Omega = \theta^*(\theta^*)'$. Note that this is a different choice of parameterization than the one we study in the main paper. In (Jin et al., 2021, Theorem 2.1) it is shown that the asymptotic distribution of ψ_n, the standardized version of SgnQ, is standard normal provided that

$$\|\theta^*\| \to \infty, \quad \theta^*_{max} \to 0, \quad \text{and} \quad (\|\theta^*\|^2/\|\theta^*\|_1)\sqrt{\log(\|\theta\|_1^*)} \to 0. \tag{C.1}$$

We verify that, in a DCBM with $\Omega = \alpha\theta\theta'$ and $\|\theta\|_1 = n$, these conditions are implied by the assumptions in (2.5), restated below:

$$n\alpha \to \infty, \qquad \text{and} \qquad \alpha\theta_{\max}^2 \log(n^2\alpha) \to 0 \tag{C.2}$$

In the parameterization of Jin et al. (2021), we have $\theta^* = \sqrt{\alpha}\theta$. First, $\|\theta^*\|^2 \to \infty$ because by (C.2),

$$\|\theta^*\|^2 \geq \frac{1}{n} \cdot \|\theta^*\|_1^2 = \alpha n \to \infty.$$

Next, $\theta^*_{\max} \to 0$ because by (C.2),

$$\theta_{\max} = \sqrt{\alpha}\theta_{\max} \to 0.$$

To show the last part of (C.1), note that

$$(\|\theta^*\|^2/\|\theta^*\|_1)\sqrt{\log(\|\theta\|_1^*)} \leq \sqrt{\alpha}\theta_{\max}\sqrt{\log(\sqrt{\alpha}n)} = \frac{1}{\sqrt{2}}\sqrt{\alpha}\theta_{\max}\sqrt{\log(\alpha n^2)} \to 0$$

by (C.2). Thus (C.1) holds, and ψ_n is asymptotically standard normal under the null. $\square$

D PROOF OF LEMMA 2.2 (PROPERTIES OF $\tilde{\Omega}$)

Lemma. *The rank and trace of the matrix $\widetilde{\Omega}$ are $(K-1)$ and $\|\theta\|^2\text{diag}(\tilde{P})'g$, respectively. When $K = 2$, $\tilde{\lambda}_1 = \text{trace}(\widetilde{\Omega}) = \|\theta\|^2(ac-b^2)(d_0^2 g_1 + d_1^2 g_0)/(ad_1^2 + 2bd_0d_1 + cd_0^2)$.*

Proof of Lemma 2.2. By basic algebra,

$$\widetilde{\Omega} = \Theta\Pi\widetilde{P}\Pi'\Theta, \qquad \text{where } \widetilde{P} = (P - (d'Pd)^{-1}Pdd'P).$$

It is seen $\widetilde{P}d = Pd - (d'Pd)^{-1}Pdd'Pd = 0$, so $\text{rank}(\widetilde{P}) \leq K-1$. At the same time, since for any matrix A and B of the same size, $\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$, it follows $\widetilde{P} \geq (K-1)$, as $\text{rank}(P) = K$ and $\text{rank}(Pdd'P) \leq 1$. This proves that $\text{rank}(\widetilde{P}) = K-1$.

At the same time, since for any matrices A and B, $\text{trace}(AB) = \text{trace}(BA)$,

$$\text{trace}(\widetilde{\Omega}) = \text{trace}(\widetilde{P}\Pi'\Theta^2\Pi) = \|\theta\|^2\text{trace}(\widetilde{P}G) = \|\theta\|^2\text{diag}(\tilde{P})'g.$$

This proves the second item of the lemma.

Last, when $K = 2$, $\widetilde{\Omega}$ is rank 1, and its eigenvalue is the same as its trace. First

$$(\tilde{P})_{11} = a - \frac{(ad_1 + bd_0)^2}{ad_1^2 + 2bd_0d_1 + cd_0^2} = (ac-b^2)\frac{d_0^2}{ad_1^2 + 2bd_0d_1 + cd_0^2}$$
$$(\tilde{P})_{22} = c - \frac{(bd_1 + cd_0)^2}{ad_1^2 + 2bd_0d_1 + cd_0^2} = (ac-b^2)\frac{d_1^2}{ad_1^2 + 2bd_0d_1 + cd_0^2}.$$

Thus

$$\tilde{\lambda}_1 = \|\theta\|^2\text{diag}(\tilde{P})'g = \|\theta\|^2(ac-b^2) \cdot \frac{d_0^2 g_1 + d_1^2 g_0}{ad_1^2 + 2bd_0d_1 + cd_0^2}$$

This proves the last item and completes the proof of the lemma.

$\square$

E PROOF OF THEOREM 2.2 (POWER OF THE SGNQ TEST) AND COROLLARY 2.1

E.1 SETUP AND RESULTS

Notation: Given sequences of real numbers $A = A_n$ and $B = B_n$, we write $A \lesssim B$ to signify that $A = O(B)$, $A \asymp B$ to signify that $A \lesssim B$ and $B \lesssim A$, and $A \sim B$ to signify that $A/B = 1 + o(1)$.

Throughout this section, we consider a DCBM with parameters (Θ, P) where $P \in \mathbb{R}^{2\times 2}$ has unit diagonals, and we analyze the behavior of SgnQ under the alternative. At the end of this subsection we explain how Theorem 2.2 and Corollary 2.1 follow from the results described next. Our results hinge on

$$\tilde{\lambda} \equiv \tilde{\lambda}_1 = \mathrm{tr}(\tilde{\Omega}).$$

Given a subset $U \subset [n]$, let $\theta_U \in \mathbb{R}^{|U|}$ denote the restriction of θ to the coordinates of U. For notational convenience, we let $S = \{i : \pi_i(1) = 1\}$, which was previously written as $\mathcal{C}_1$ in the main paper.

In a DCBM where P has unit diagonals, our main results hold under the following conditions.

$$\Omega_{ij} \lesssim \theta_i\theta_j \tag{E.1}$$

$$\|\theta\|_\infty = O(1), \text{ and} \tag{E.2}$$

$$\|\theta\|_2^2 \to \infty. \tag{E.3}$$

$$(\|\theta\|_2^2/\|\theta\|_1)\sqrt{\log(\|\theta\|_1)} \to 0. \tag{E.4}$$

First we justify that these assumptions are satisfied by an equivalent DCBM with the same Ω represented with the parameterization (2.1) and satisfying (2.7). Thus all results proved in this section transfer immediately to the main paper.

Lemma E.1. *Consider a DCBM with parameters* (Θ^*, P^*) *satisfying* (2.1) *and satisfying* (2.7)*. Define* $\Theta = \mathrm{diag}(\theta)$ *where*

$$\theta_i = \begin{cases} \sqrt{a}\theta_i^* & \text{if } i \in S \\ \sqrt{c}\theta_i^* & \text{if } i \in S^c, \end{cases}$$

and

$$P = \begin{pmatrix} 1 & \frac{b}{\sqrt{ac}} \\ \frac{b}{\sqrt{ac}} & 1 \end{pmatrix}.$$

Then

$$\Omega = \Theta\Pi P\Pi\Theta = \Theta^*\Pi P^*\Pi'\Theta^*$$

and (E.1)–(E.4) *are satisfied.*

Proof. The statement regarding Ω follows by basic algebra. (E.1) follows if we can show that

$$\frac{b}{\sqrt{ac}} \lesssim 1. \tag{E.5}$$

Since

$$b = \frac{cn - (a+c)N}{n - 2N} = c \cdot \frac{n-N}{n-2N} - a \cdot \frac{N}{n-2N},$$

we have $a \geq c \gtrsim b$, so (E.5) follows.

Next, (E.2) follows directly from $a\theta_{\max,1}^2 \lesssim 1$ since $c\theta_{\max,0}^2 = o(1)$ by (2.7).

For (E.3),

$$\|\theta\|_2^2 \geq \frac{1}{n} \cdot \|\theta\|_1^2 \geq cn \to \infty$$

by (2.7).

For the last part, note that

$$b = c \cdot \frac{n-N}{n-2N} - a \cdot \frac{N}{n-2N} \geq 0 \Rightarrow a\varepsilon \lesssim c.$$

Thus,

$$\begin{aligned} \frac{\|\theta\|_2^2}{\|\theta\|_1} &= \frac{a\|\theta_S^*\|_2^2 + c\|\theta_{S^c}^*\|_2^2}{\sqrt{a}\|\theta_S^*\|_1 + \sqrt{c}\|\theta_{S^c}^*\|_1} \lesssim \frac{a(N/n)\|\theta_{S^c}^*\|_2^2 + c\|\theta_{S^c}^*\|_2^2}{\sqrt{c}\|\theta_{S^c}^*\|_1} \\ &\lesssim \frac{c\|\theta_{S^c}^*\|_2^2}{\sqrt{c}\|\theta_{S^c}^*\|_1} \lesssim \sqrt{c}\theta_{\max,0} = o(\frac{1}{\sqrt{\log cn^2}}) = o(\frac{1}{\sqrt{\log(\|\theta\|_1)}}), \end{aligned}$$

which implies (E.4). Above we use that $a \geq c$ and $g_1 \asymp d_1 \asymp N/n$, by assumption. Precisely, in the first line, we used

$$a\|\theta^*_S\|_2^2 \asymp a \cdot (1 - N/n)^{-1}\frac{N}{n}\|\theta^*_{S^c}\|_2^2 \lesssim c\|\theta^*_{S^c}\|_2^2,$$

and in the second line we used

$$\|\theta\|_1 \geq \sqrt{c}\|\theta^*_{S^c}\|_1 \asymp \sqrt{c}(1 - N/n)^{-1}\|\theta^*\|_1 \asymp \sqrt{c}n.$$

$\square$

With Lemma E.1 in hand, we restrict in the remainder of this section to the setting where P has unit diagonals and (E.1)–(E.4) are satisfied.

Define $v_0 = \mathbf{1}'\Omega\mathbf{1}$, and let $\eta^* = (1/\sqrt{v_0})\Omega\mathbf{1}$. Recall $\tilde{\Omega} = \Omega - \eta^*\eta^{*\mathsf{T}}$, and $\tilde{\lambda} = \mathrm{tr}(\tilde{\Omega})$. Our main result concerning the alternative is the following.

Theorem E.1 (Limiting behavior of SgnQ test statistic)**.** *Suppose that the previous assumptions hold and that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$. Then under the null hypothesis, as $n \to \infty$, $\mathbb{E}[Q] \sim 2\|\theta\|_2^4$, $\mathrm{Var}(Q) \sim 8\|\theta\|_2^8$, and $(Q - \mathbb{E}Q)/\sqrt{\mathrm{Var}(Q)} \to N(0, 1)$ in law. Under the alternative hypothesis, as $n \to \infty$, $\mathbb{E}Q \sim \tilde{\lambda}^4$ and $\mathrm{Var}(Q) \lesssim |\tilde{\lambda}|^6 + |\tilde{\lambda}|^2\lambda_1^3 = o(\tilde{\lambda}^8)$.*

Following Jin et al. (2021), we introduce some notation:

$$\widetilde{\Omega} = \Omega - (\eta^*)(\eta^*)', \qquad \text{where} \quad \eta^* = \frac{1}{\sqrt{v_0}}\Omega\mathbf{1}_n, \;\; v_0 = \mathbf{1}_n'\Omega\mathbf{1}_n;$$

$$\delta_{ij} = \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_i - \tilde{\eta}_i), \qquad \text{where} \quad \eta = \frac{1}{\sqrt{v}}(\mathbb{E}A)\mathbf{1}_n, \;\; \tilde{\eta} = \frac{1}{\sqrt{V}}A\mathbf{1}_n, \;\; v = \mathbf{1}_n'(\mathbb{E}A)\mathbf{1}_n;$$

$$r_{ij} = (\eta_i^*\eta_j^* - \eta_i\eta_j) - (\eta_i - \tilde{\eta}_i)(\eta_j - \tilde{\eta}_j) + (1 - \frac{v}{V})\tilde{\eta}_i\tilde{\eta}_j, \qquad \text{where} \;\; V = \mathbf{1}_n'A\mathbf{1}_n.$$

The *ideal* and *proxy* SgnQ statistics, respectively, are defined as follows:

$$\widetilde{Q}_n = \sum_{i,j,k,\ell(dist)} (\widetilde{\Omega}_{ij} + W_{ij})(\widetilde{\Omega}_{jk} + W_{jk})(\widetilde{\Omega}_{k\ell} + W_{k\ell})(\widetilde{\Omega}_{\ell i} + W_{\ell i}) \tag{E.6}$$

$$Q_n^* = \sum_{i,j,k,\ell(dist)} (\widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij})(\widetilde{\Omega}_{jk} + W_{jk} + \delta_{jk})(\widetilde{\Omega}_{k\ell} + W_{k\ell} + \delta_{k\ell})(\widetilde{\Omega}_{\ell i} + W_{\ell i} + \delta_{\ell i}). \tag{E.7}$$

Moreover, we can express the original or *real* SgnQ as

$$Q_n = \sum_{i,j,k,\ell(dist)} \Big[(\widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij} + r_{ij})(\widetilde{\Omega}_{jk} + W_{jk} + \delta_{jk} + r_{jk}) (\widetilde{\Omega}_{k\ell} + W_{k\ell} + \delta_{k\ell} + r_{k\ell})(\widetilde{\Omega}_{\ell i} + W_{\ell i} + \delta_{\ell i} + r_{\ell i})\Big].$$

The next theorems handle the behavior of these statistics. Together the results imply Theorem E.1. Again, the analysis of the null carries over directly from Jin et al. (2021), so we only need to study the alternative. The claims regarding the alternative follow from Lemmas E.7–E.12 below.

Theorem E.2 (Ideal SgnQ test statistic)**.** *Suppose that the previous assumptions hold and that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$. Then under the null hypothesis, as $n \to \infty$, $\mathbb{E}[\tilde{Q}] = 0$ and $\mathrm{Var}(\tilde{Q}) = 8\|\theta\|_2^8 \cdot [1 + o(1)]$. Furthermore, under the alternative hypothesis, as $n \to \infty$, $\mathbb{E}[\tilde{Q}] \sim \tilde{\lambda}^4$ and $\mathrm{Var}(\tilde{Q}) \lesssim \lambda_1^4 + |\tilde{\lambda}|^6 = o(\tilde{\lambda}^8)$.*

Theorem E.3 (Proxy SgnQ test statistic)**.** *Suppose that the previous assumptions hold and that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$. Then under the null hypothesis, as $n \to \infty$, $|\mathbb{E}[\tilde{Q} - Q^*]| = o(\|\theta\|_2^4)$ and $\mathrm{Var}(\tilde{Q} - Q^*) = o(\|\theta\|_2^8)$. Furthermore, under the alternative hypothesis, as $n \to \infty$, $|\mathbb{E}[\tilde{Q} - Q^*]| \lesssim |\tilde{\lambda}|^2\lambda_1 = o(\tilde{\lambda}^4)$ and $\mathrm{Var}(\tilde{Q} - Q^*) \lesssim |\tilde{\lambda}|^2\lambda_1^3 + |\tilde{\lambda}|^6 = o(\tilde{\lambda}^8)$.*

Theorem E.4 (Real SgnQ test statistic)**.** *Suppose that the previous assumptions hold and that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$. Then under the null hypothesis, as $n \to \infty$, $|\mathbb{E}[Q - \tilde{Q}]| = o(\|\theta\|_2^4)$ and $\mathrm{Var}(Q - \tilde{Q}) = o(\|\theta\|_2^8)$. Furthermore, under the alternative hypothesis, as $n \to \infty$, $|\mathbb{E}[Q - Q^*]| \lesssim |\tilde{\lambda}|^2 \lambda_1 = o(\tilde{\lambda}^4)$ and $\mathrm{Var}(Q - Q^*) \lesssim |\tilde{\lambda}|^2 \lambda_1^3 = o(\tilde{\lambda}^8)$.*

The previous work Jin et al. (2021) establishes that under the assumptions above, if $\|\theta_S\|_1/\|\theta\|_1 \asymp 1$, then SgnQ distinguishes the null and alternative provided that $|\lambda_2|/\sqrt{\lambda_1} \to \infty$. To compare with the results above, note that $\lambda_2 \asymp \tilde{\lambda}$ if $\|\theta_S\|_1/\|\theta\|_1 \asymp 1$ (c.f. Lemma E.5 of Jin et al. (2021)). Thus when $K = 2$, our main result extends the upper bound of Jin et al. (2021) to the case when $\|\theta_S\|_1/\|\theta\|_1 = o(1)$. We note that $|\tilde{\lambda}| \gtrsim |\lambda_2|$ in general (see Lemma E.3 and Corollary E.1).

The theorems above apply to the symmetric SBM. Recall that in this model,

$$\Omega_{ij} = \begin{cases} a & \text{if } i, j \in S \\ c & \text{if } i, j \notin S \\ \tilde{b} = \frac{nc - (a+c)N}{n - 2N} & \text{otherwise.} \end{cases}$$

where $N = |S|$ and $a, b, c \in (0, 1)$. To obtain this model from our DCBM, set

$$P = \begin{pmatrix} 1 & \tilde{b}/\sqrt{ac} \\ \tilde{b}/\sqrt{ac} & 1 \end{pmatrix}, \tag{E.8}$$

and

$$\theta = \sqrt{a}\mathbf{1}_S + \sqrt{c}\mathbf{1}_{S^c}. \tag{E.9}$$

The assumption (E.1) implies that $\tilde{b} \lesssim \sqrt{ac}$, which is automatically satisfied since we assume $a \geq c$.

In SBM, it holds that $\lambda_2 = \tilde{\lambda}$ (see Lemma E.3). Furthermore, explicit calculations in Section E.5 reveal that

$$\begin{aligned} &\lambda_1 \sim nc, \text{ and} \\ &\lambda_2 = \tilde{\lambda} \sim N(a - c). \end{aligned} \tag{E.10}$$

In addition, with $P, a, \tilde{b}, c$ as above, if we have

$$\theta_i = \begin{cases} \rho_i \sqrt{a} & \text{if } i \in S \\ \rho_i \sqrt{c} & \text{if } i \notin S \end{cases}$$

for $\rho > 0$ with $\rho_{\min} \gtrsim \rho_{\max}$ in the DCBM setting, a very similar calculation, which we omit, reveals that

$$\begin{aligned} \lambda_1 &\asymp nc, \text{ and} \\ \tilde{\lambda} &\asymp N(a - c). \end{aligned} \tag{E.11}$$

With the previous results of this subsection in hand (which are proved in the remaining subsections) we justify Theorem 2.2 and Corollary 2.1.

Proof of Theorem 2.2. The SgnQ test has level κ by Theorem 2.1, so it remains to study the type II error. Using Theorem E.1 and Lemma E.1, the fact that the type II error tends to 0 directly follows from Chebyshev's inequality and the fact that $\|\hat{\eta}\|_2^2 - 1 \approx \|\theta\|_2^2$ with high probability. In particular, note that since $|\tilde{\lambda}| \gg \sqrt{\lambda_1}$, the expectation of SgnQ under the alternative is much larger than its standard deviation, under the null or alternative. We omit the details as they are very similar to the proof of Theorem 2.6 in (Jin et al., 2021, Supplement,pgs. 5–6). □

Proof of Corollary 2.1. This result follows immediately from (E.11) and Theorem 2.2. □

E.2 Preliminary bounds

Define $v_0 = \mathbf{1}^\mathsf{T}\Omega\mathbf{1}$, and let $\eta^* = 1/\sqrt{v_0} \cdot \Omega\mathbf{1}$. For the analysis of SgnQ, it is important is to understand $\tilde{\Omega} = \Omega - \eta^*\eta^{*\mathsf{T}}$. The next lemma establishes that $\tilde{\Omega}$ is rank one and has a simple expression when $K = 2$.

Lemma E.2. *Let $f = (\|\theta_{S^c}\|_1, -\|\theta_S\|_1)^\mathsf{T}$ It holds that*

$$\tilde{\Omega} = \frac{(1-b^2)}{v_0} \cdot \Theta\Pi f f^\mathsf{T}\Pi^\mathsf{T}\Theta.$$

Proof. Let $\rho_0 = \|\theta_S\|_1$ and $\rho_1 = \|\theta_{S^c}\|_1$. Note that

$$(\Omega\mathbf{1})_i = \theta_i \sum_j \theta_j \pi_i^\mathsf{T} P \pi_j = \begin{cases} \theta_i(\rho_0 + b\rho_1) & \text{if } i \in S \\ \theta_i(b\rho_0 + \rho_1) & \text{if } i \notin S. \end{cases}$$

Hence

$$v_0 = \mathbf{1}^\mathsf{T}\Omega\mathbf{1} = \rho_0^2 + 2b\rho_0\rho_1 + \rho_1^2.$$

If $i, j \in S$, then

$$\tilde{\Omega}_{ij} = \theta_i\theta_j\big(1 - \frac{(\rho_0 + b\rho_1)^2}{v_0}\big) = \theta_i\theta_j \cdot \frac{(1-b^2)\rho_1^2}{v_0}$$

Similarly if $i \in S$ and $j \notin S$,

$$\tilde{\Omega}_{ij} = \theta_i\theta_j\big(b - \frac{(\rho_0 + b\rho_1)(b\rho_0 + \rho_1)}{v_0}\big) = -\theta_i\theta_j \cdot \frac{(1-b^2)\rho_0\rho_1}{v_0}$$

and

$$\tilde{\Omega}_{ij} = \theta_i\theta_j\big(1 - \frac{(b\rho_0 + \rho_1)^2}{v_0}\big) = \theta_i\theta_j \cdot \frac{(1-b^2)\rho_0^2}{v_0}$$

if $i, j \in S^c$. The claim follows. ☐

Let

$$w = \Theta\Pi f = \theta_S\|\theta_{S^c}\|_1 - \theta_{S^c}\|\theta_S\|_1 = \rho_1\theta_S - \rho_0\theta_{S^c}$$

Using the previous lemma, we have the rank one eigendecomposition

$$\tilde{\Omega} = \tilde{\lambda}\tilde{\xi}\tilde{\xi}^\mathsf{T}, \tag{E.12}$$

where we define

$$\tilde{\xi} = \frac{\rho_1\theta_S - \rho_0\theta_{S^c}}{\|\rho_1\theta_S - \rho_0\theta_{S^c}\|_2} = \frac{\rho_1\theta_S - \rho_0\theta_{S^c}}{\sqrt{\rho_1^2\|\theta_S\|_2^2 + \rho_0^2\|\theta_{S^c}\|_2^2}}, \text{ and} \tag{E.13}$$

$$\tilde{\lambda} = \frac{(1-b^2)}{v_0} \cdot \big(\rho_1^2\|\theta_S\|_2^2 + \rho_0^2\|\theta_{S^c}\|_2^2\big). \tag{E.14}$$

Lemma E.5 of Jin et al. (2021) implies that if $\|\theta_S\|_1/\|\theta\|_1 \asymp 1$, then $\lambda_2 \asymp \tilde{\lambda}_1$. If $\|\theta_S\|_1/\|\theta\|_1 = o(1)$, then this guarantee may not hold. Below, in the case $K = 2$, we express $\tilde{\lambda}$ in terms of the eigenvalues and eigenvectors of Ω. This allows us to compare λ_2 with $\tilde{\lambda}$ more generally, as in Corollary E.1.

Lemma E.3. *Let Ω have eigenvalues λ_1, λ_2 and eigenvectors ξ_1, ξ_2. Let $\tilde{\lambda}$ denote the eigenvalue of $\tilde{\Omega}$. Then*

$$\tilde{\lambda} = \frac{\lambda_1\lambda_2\big(\langle\xi_1, \mathbf{1}\rangle^2 + \langle\xi_2, \mathbf{1}\rangle^2\big)}{\lambda_1\langle\xi_1, \mathbf{1}\rangle^2 + \lambda_2\langle\xi_2, \mathbf{1}\rangle^2}. \tag{E.15}$$

Proof. By explicit computation,

$$\begin{aligned}
\tilde{\Omega} &= \Omega - \eta^* \eta^{*\mathsf{T}} \\
&= \lambda_1\big(1 - \frac{\lambda_1 \langle \xi_1, \mathbf{1}\rangle^2}{v_0}\big)\xi_1\xi_1^{\mathsf{T}} + \lambda_2\big(1 - \frac{\lambda_2 \langle \xi_2, \mathbf{1}\rangle^2}{v_0}\big)\xi_2\xi_2^{\mathsf{T}} - \frac{\lambda_1\lambda_2\langle \xi_1, \mathbf{1}\rangle\langle \xi_2, \mathbf{1}\rangle}{v_0}\big(\xi_1\xi_2^{\mathsf{T}}\xi_2 + \xi_1^{\mathsf{T}}\big) \\
&= \frac{\lambda_1\lambda_2}{v_0}\left(\langle \xi_2, \mathbf{1}\rangle\xi_1 + \langle \xi_1, \mathbf{1}\rangle\xi_2\right)\cdot\left(\langle \xi_2, \mathbf{1}\rangle\xi_1 + \langle \xi_1, \mathbf{1}\rangle\xi_2\right)^{\mathsf{T}}.
\end{aligned}$$

From (E.13) and (E.14), it follows that

$$\begin{aligned}
\tilde{\xi} &= \frac{\langle \xi_2, \mathbf{1}\rangle\xi_1 + \langle \xi_1, \mathbf{1}\rangle\xi_2}{\sqrt{\langle \xi_1, \mathbf{1}\rangle^2 + \langle \xi_2, \mathbf{1}\rangle^2}} \\
\tilde{\lambda} &= \frac{\lambda_1\lambda_2}{v_0}\big(\langle \xi_1, \mathbf{1}\rangle^2 + \langle \xi_2, \mathbf{1}\rangle^2\big).
\end{aligned}$$

□

Corollary E.1. *It holds that*

$$|\lambda_2| \lesssim |\tilde{\lambda}| \lesssim \lambda_1. \tag{E.16}$$

If $\lambda_2 \geq 0$*, then*

$$\lambda_2 \leq \tilde{\lambda} \leq \lambda_1 \tag{E.17}$$

Proof. Suppose that $\lambda_2 \geq 0$. Then

$$\lambda_2\big(\langle \xi_1, \mathbf{1}\rangle^2 + \langle \xi_2, \mathbf{1}\rangle^2\big) \leq \lambda_1\langle \xi_1, \mathbf{1}\rangle^2 + \lambda_2\langle \xi_2, \mathbf{1}\rangle^2 = v_0 \leq \lambda_1\big(\langle \xi_1, \mathbf{1}\rangle^2 + \langle \xi_2, \mathbf{1}\rangle^2\big),$$

implies (E.17).

Suppose that $\lambda_2 < 0$. Note that

$$\lambda_1\big(\langle \xi_1, \mathbf{1}\rangle^2 + \langle \xi_2, \mathbf{1}\rangle^2\big) \geq \lambda_1\langle \xi_1, \mathbf{1}\rangle^2 + \lambda_2\langle \xi_2, \mathbf{1}\rangle^2 = v_0 \geq 0,$$

which combined with (E.15) implies that $|\tilde{\lambda}| \geq |\lambda_2|$.

Next,

$$\lambda_2 \leq \tilde{\xi}^{\mathsf{T}}\Omega\tilde{\xi} = \tilde{\lambda} + \langle \tilde{\xi}, \eta^*\rangle^2,$$

which implies that

$$|\tilde{\lambda}| \leq |\lambda_2| + \langle \tilde{\xi}, \eta^*\rangle^2 \leq \lambda_1 + \|\eta^*\|_2^2 \lesssim \lambda_1 + \|\theta\|_1^2 \lesssim \lambda_1,$$

where the last inequality follows from Lemma E.5. □

The next results are frequently used in our analyis of SgnQ.

Lemma E.4. *Let* $v = \mathbf{1}^{\mathsf{T}}(\Omega - \mathrm{diag}(\Omega))\mathbf{1}$ *and* $v_0 = \mathbf{1}^{\mathsf{T}}\Omega\mathbf{1}$*. Then*

$$v_0 \sim v \sim \|\theta\|_1^2. \tag{E.18}$$

Proof. By (E.4), $\|\theta\|_2^2 = o(\|\theta\|_1)$. By (E.3), $\|\theta\|_1 \to \infty$. Hence

$$v = \mathbf{1}^{\mathsf{T}}(\Omega - \mathrm{diag}(\Omega))\mathbf{1} = \|\theta\|_1^2 - \|\theta\|_2^2 \sim \|\theta\|_1^2 \sim v_0 = \mathbf{1}^{\mathsf{T}}\Omega\mathbf{1}.$$

□

The next result is a direct corollary of Lemmas E.2 and E.4.

Corollary E.2. *Define* $\beta \in \mathbb{R}^n$ *by*

$$\beta = \sqrt{\frac{|1-b^2|}{v_0}}\cdot\big(\|\theta_{S^c}\|_1\mathbf{1}_S + \|\theta_S\|_1\mathbf{1}_{S^c}\big) \tag{E.19}$$

Then

$$|\tilde{\Omega}_{ij}| \lesssim \beta_i\theta_i\beta_j\theta_j. \tag{E.20}$$

Lemma E.5. *Let λ_1 denote the largest eigenvalue of Ω. Then*
$$\lambda_1 \gtrsim \|\theta\|_2^2. \tag{E.21}$$

Proof. Using the universal inequality $a^2 + b^2 \geq \frac{1}{2}(a+b)^2$, we have
$$\begin{aligned}\lambda_1 &\geq \frac{\theta^\mathsf{T}\Omega\theta}{\|\theta\|_2^2} \geq \frac{1}{\|\theta\|_2^2} \cdot \sum_{i,j} \theta_i\theta_j\Omega_{ij} \geq \frac{1}{\|\theta\|_2^2} \cdot \big(\sum_{i,j \in S} \theta_i^2\theta_j^2 + \sum_{i,j \notin S} \theta_i^2\theta_j^2 \big) \\ &\geq \frac{\|\theta_S\|_2^4 + \|\theta_{S^c}\|_2^4}{\|\theta\|_2^2} \gtrsim \|\theta\|_2^2.\end{aligned}$$
□

Lemma E.6. *Define $\eta = \frac{1}{\sqrt{v}}(\Omega - \mathrm{diag}(\Omega))\mathbf{1}$. Then*
$$\eta_i \lesssim \eta_i^* \lesssim \theta_i \tag{E.22}$$

Proof. The left-hand side is immediate, so we prove that $\eta_i^* \lesssim \theta_i$. We have
$$(\Omega\mathbf{1})_i = \begin{cases} \theta_i(\|\theta_S\|_1 + b\|\theta_{S^c}\|_1) & \text{if } i \in S \\ \theta_i(b\|\theta_S\|_1 + \|\theta_{S^c}\|_1) & \text{if } i \notin S \end{cases}$$
Since $\Omega_{ii} = \theta_i^2$,
$$\sqrt{v_0} \cdot \eta_i = \begin{cases} \theta_i(\|\theta_S\|_1 + b\|\theta_{S^c}\|_1) - \theta_i^2 & \text{if } i \in S \\ \theta_i(b\|\theta_S\|_1 + \|\theta_{S^c}\|_1) - \theta_i^2 & \text{if } i \notin S. \end{cases}$$
Since $b = O(1)$, $\theta_i = O(1)$, and $v_0 \gtrsim \|\theta\|_1^2$ (c.f. Lemma E.4),
$$\eta_i^* \lesssim \frac{\theta_i\|\theta\|_1}{\sqrt{\|\theta\|_1^2}} = \theta_i,$$
as desired. □

We use the bounds (E.18) – (E.22) throughout. We also use repeatedly that
$$\|\theta\|_p^p \lesssim \|\theta\|_q^q, \text{ if } p \geq q, \tag{E.23}$$
which holds by (E.2), and
$$\begin{aligned} \|\beta \circ \theta\|_2^2 &= |\tilde{\lambda}| \\ |\beta_i| &\lesssim 1 \\ \|\beta \circ \theta^{\circ 2}\|_1 &\leq \|\beta \circ \theta\|_2 \|\theta\|_2 \lesssim \|\theta\|_2^2, \end{aligned} \tag{E.24}$$
where the second line holds by Cauchy–Schwarz.

E.3 Mean and variance of SgnQ

The previous work Jin et al. (2021) decomposes $\tilde{Q}$ and $\tilde{Q} - Q^*$ into a finite number of terms. For each term an exact expression for its mean and variance is derived in Jin et al. (2021) that depends on θ, η, v, and $\tilde{\Omega}$. These expression are then bounded using the inequalities (E.2), (E.3), (E.18), (E.21)–(E.23), as well as an inequality of the form
$$|\tilde{\Omega}_{ij}| \lesssim \alpha\theta_i\theta_j.$$

In our case, an inequality of this form is still valid, but it does not attain sharp results because it does not properly capture the signal $|\tilde{\lambda}|$ from the smaller community. Instead, we use the inequality (E.20), followed by the bounds in (E.24) to handle terms involving $\tilde{\Omega}$.

Therefore, for terms of $\tilde{Q}$ and $\tilde{Q} - Q^*$ that do not depend on $\tilde{\Omega}$, the bounds in Jin et al. (2021) carry over immediately. In particular, their analysis of the null hypothesis carries over directly. Hence we can focus solely on the alternative hypothesis.

Furthermore, any terms with zero mean in Jin et al. (2021) also have zero mean in our setting : for every term that is mean zero, it is simply the sum of mean zero subterms, and each mean zero subterm is a product of independent, centered random variables (eg, X_1 below).

E.3.1 IDEAL SGNQ

The previous work Jin et al. (2021) shows that $\tilde{Q} = X_1 + 4X_2 + 4X_3 + 2X_4 + 4X_5 + X_6$, where $X_1, \dots, X_6$ are defined in their Section G.1. For convenience, we state explicitly the definitions of these terms.

$$X_1 = \sum_{i,j,k,\ell(dist)} W_{ij}W_{jk}W_{k\ell}W_{\ell i}, \qquad X_2 = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}W_{jk}W_{k\ell}W_{\ell i},$$
$$X_3 = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}W_{k\ell}W_{\ell i}, \qquad X_4 = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}W_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i},$$
$$X_5 = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i}, \qquad X_6 = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}.$$

Since X_1 does not depend on $\tilde{\Omega}$, the bounds for X_1 below are directly quoted from Lemma G.3 of Jin et al. (2021). Also note that X_6 is a non-stochastic term.

Lemma E.7. *Under the alternative hypothesis, we have*

$$\begin{aligned}
\mathbb{E}[X_k] &= 0 \textit{ for } 1 \le k \le 5, \\
\mathrm{Var}(X_1) &\lesssim \|\theta\|_2^8 \lesssim \lambda_1^4 \\
\mathrm{Var}(X_2) &\lesssim \|\beta \circ \theta\|_2^4 \, \|\theta\|_2^4 \lesssim |\tilde{\lambda}|^2 \lambda_1^2 \\
\mathrm{Var}(X_3) &\lesssim \|\beta \circ \theta\|_2^8 \, \|\theta\|_2^2 \lesssim |\tilde{\lambda}|^4 \lambda_1 \\
\mathrm{Var}(X_4) &\lesssim \|\beta \circ \theta\|_2^8 \le |\tilde{\lambda}|^4 \\
\mathrm{Var}(X_5) &\lesssim \|\beta \circ \theta\|_2^{12} \lesssim |\tilde{\lambda}|^6, \textit{ and} \\
\mathbb{E}[X_6] &= X_6 \sim |\tilde{\lambda}^4|
\end{aligned}$$

Since we assume $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$ under the alternative hypothesis, it holds that

$$\mathrm{Var}(\tilde{Q}) \lesssim \lambda_1^4 + |\tilde{\lambda}|^6.$$

Theorem E.2 follows directly from this bound and that $\mathbb{E}X_6 = \mathbb{E}\tilde{Q} \sim \tilde{\lambda}^4$.

E.3.2 PROXY SGNQ

The previous work Jin et al. (2021) shows that

$$\tilde{Q} - Q^* = U_a + U_b + U_c,$$

where

$$\begin{aligned}
U_a &= 4Y_1 + 8Y_2 + 4Y_3 + 8Y_4 + 4Y_5 + 4Y_6 \\
U_b &= 4Z_1 + 2Z_2 + 8Z_3 + 4Z_4 + 4Z_5 + 2Z_6 \\
U_c &= 4T_1 + 4T_2 + F.
\end{aligned}$$

These terms are defined in Section G.2 of Jin et al. (2021), and for convenience, we define them explicitly below. The previous equations are obtained by expanding carefully $\tilde{Q}$ and Q^* as defined in (E.6) and (E.7). Thus, the terms on the right-hand-side above are referred as *post-expansion* terms, and we can analyze each one individually. Now we proceed to their definitions.

First $Y_1, \dots, Y_6$ are defined as follows.

$$Y_1 = \sum_{i,j,k,\ell(dist)} \delta_{ij}W_{jk}W_{k\ell}W_{\ell i}, \qquad Y_2 = \sum_{i,j,k,\ell(dist)} \delta_{ij}\widetilde{\Omega}_{jk}W_{k\ell}W_{\ell i},$$
$$Y_3 = \sum_{i,j,k,\ell(dist)} \delta_{ij}W_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i}, \qquad Y_4 = \sum_{i,j,k,\ell(dist)} \delta_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i},$$

$$Y_5 = \sum_{i,j,k,\ell(dist)} \delta_{ij}\widetilde{\Omega}_{jk}W_{k\ell}\widetilde{\Omega}_{\ell i}, \qquad Y_6 = \sum_{i,j,k,\ell(dist)} \delta_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}.$$

Next, $Z_1, \ldots, Z_6$ are defined as follows.

$$\begin{aligned}
Z_1 &= \sum_{i,j,k,\ell(dist)} \delta_{ij}\delta_{jk}W_{k\ell}W_{\ell i}, & Z_2 &= \sum_{i,j,k,\ell(dist)} \delta_{ij}W_{jk}\delta_{k\ell}W_{\ell i}, \\
Z_3 &= \sum_{i,j,k,\ell(dist)} \delta_{ij}\delta_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i}, & Z_4 &= \sum_{i,j,k,\ell(dist)} \delta_{ij}\widetilde{\Omega}_{jk}\delta_{k\ell}W_{\ell i}, \\
Z_5 &= \sum_{i,j,k,\ell(dist)} \delta_{ij}\delta_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}, & Z_6 &= \sum_{i,j,k,\ell(dist)} \delta_{ij}\widetilde{\Omega}_{jk}\delta_{k\ell}\widetilde{\Omega}_{\ell i}.
\end{aligned}$$

Last, we have the definitions of T_1, T_2, and F.

$$\begin{aligned}
T_1 &= \sum_{i,j,k,\ell(dist)} \delta_{ij}\delta_{jk}\delta_{k\ell}W_{\ell i}, & T_2 &= \sum_{i,j,k,\ell(dist)} \delta_{ij}\delta_{jk}\delta_{k\ell}\widetilde{\Omega}_{\ell i}, \\
F &= \sum_{i,j,k,\ell(dist)} \delta_{ij}\delta_{jk}\delta_{k\ell}\delta_{\ell i}.
\end{aligned}$$

The following post-expansion terms below appear in Lemma G.5 of Jin et al. (2021). The term Y_1 does not depend on $\widetilde{\Omega}$, so we may directly quote the result.

Lemma E.8. *Under the alternative hypothesis, it holds that*

$$\begin{aligned}
&|\mathbb{E}Y_1| = 0, && \mathrm{Var}(Y_1) \lesssim \|\theta\|_2^2\| \, \|\theta\|_3^6 \lesssim \lambda_1^4 \\
&|\mathbb{E}Y_2| = 0, && \mathrm{Var}(Y_2) \lesssim \|\beta \circ \theta\|_2^2 \, \|\theta\|_2^6 \lesssim |\tilde{\lambda}|\lambda_1^3 \\
&|\mathbb{E}Y_3| = 0, && \mathrm{Var}(Y_3) \lesssim \|\beta \circ \theta\|_2^4 \, \|\theta\|_2^4 \lesssim |\tilde{\lambda}|^2\lambda_1^2 \\
&|\mathbb{E}Y_4| \lesssim \|\beta \circ \theta\|_2^4\|\theta\|_2^2 \lesssim |\tilde{\lambda}|^2\lambda_1, && \mathrm{Var}(Y_4) \lesssim \frac{\|\beta \circ \theta\|_2^6 \, \|\theta\|_2^6}{\|\theta\|_1} \lesssim |\tilde{\lambda}|^3\lambda_1^2 \\
&|\mathbb{E}Y_5| = 0, && \mathrm{Var}(Y_5) \lesssim \frac{\|\beta \circ \theta\|_2^6 \, \|\theta\|_2^4}{\|\theta\|_1} \lesssim |\tilde{\lambda}|^3\lambda_1 \\
&|\mathbb{E}Y_6| = 0, && \mathrm{Var}(Y_6) \lesssim \frac{\|\beta \circ \theta\|_2^{12}\|\theta\|_2^2}{\|\theta\|_1} \lesssim |\tilde{\lambda}|^6.
\end{aligned}$$

As a result,

$$|\mathbb{E}U_a| \lesssim |\tilde{\lambda}|^2\lambda_1 = o(\tilde{\lambda}^4). \tag{E.25}$$

Also using Corollary E.1 and that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$, we have

$$\mathrm{Var}(U_a) \lesssim \lambda_1^4 + |\tilde{\lambda}|^3\lambda_1^2 + |\tilde{\lambda}|^6. \tag{E.26}$$

The terms below appear in Lemma G.7 of Jin et al. (2021). The bounds on Z_1 and Z_2 are quoted directly from Jin et al. (2021).

Lemma E.9. *Under the alternative hypothesis, it holds that*

$$\begin{aligned}
&|\mathbb{E}Z_1| \lesssim \|\theta\|_2^4 \lesssim \lambda_1^2, && \mathrm{Var}(Z_1) \lesssim \|\theta\|_2^2 \, \|\theta\|_3^6 \lesssim \lambda_1^4 \\
&|\mathbb{E}Z_2| \lesssim \|\theta\|_2^4 \lesssim \lambda_1^2, && \mathrm{Var}(Z_2) \lesssim \frac{\|\theta\|_2^6 \, \|\theta\|_3^3}{\|\theta\|_1} \lesssim \lambda_1^3 \\
&|\mathbb{E}Z_3| = 0, && \mathrm{Var}(Z_3) \lesssim \|\beta \circ \theta\|_2^4 \, \|\theta\|_2^6 \lesssim |\tilde{\lambda}|^2\lambda_1^3 \\
&|\mathbb{E}Z_4| \lesssim \|\beta \circ \theta\|_2^2 \, \|\theta\|_2^2 \lesssim |\tilde{\lambda}|\lambda_1, && \mathrm{Var}(Z_4) \lesssim \frac{\|\beta \circ \theta\|_2^4 \, \|\theta\|_2^6}{\|\theta\|_1} \lesssim |\tilde{\lambda}|^2\lambda_1^2
\end{aligned}$$

$$|\mathbb{E}Z_5| \lesssim \|\beta\circ\theta\|_2^4\,\|\theta\|_2^2 \lesssim |\tilde{\lambda}|^2\lambda_1, \quad \mathrm{Var}(Z_5) \lesssim \frac{\|\beta\circ\theta\|_2^8\,\|\theta\|_2^6}{\|\theta\|_1^2} \lesssim |\tilde{\lambda}|^4\lambda_1$$
$$|\mathbb{E}Z_6| \lesssim \frac{\|\beta\circ\theta\|_2^4\,\|\theta\|_2^4}{\|\theta\|_1^2} \lesssim |\tilde{\lambda}|^2, \qquad \mathrm{Var}(Z_6) \lesssim \frac{\|\beta\circ\theta\|_2^8\,\|\theta\|_2^4}{\|\theta\|_1^2} \lesssim |\tilde{\lambda}|^4.$$

Using Corollary E.1 and the fact that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$ under the alternative hypothesis, we have

$$|\mathbb{E}U_b| \lesssim |\tilde{\lambda}|^2\lambda_1, \tag{E.27}$$

and

$$\mathrm{Var}(U_b) \lesssim |\tilde{\lambda}|^2\lambda_1^3. \tag{E.28}$$

The terms below appear in Lemma G.9 of Jin et al. (2021). The bounds on T_1 and F are quoted directly from Jin et al. (2021) since they do not depend on $\tilde{O}mega$.

Lemma E.10. *Under the alternative hypothesis, it holds that*

$$|\mathbb{E}T_1| \le \frac{\|\theta\|_2^6}{\|\theta\|_1^2} \lesssim \lambda_1, \qquad \mathrm{Var}(T_1) \lesssim \frac{\|\theta\|_2^6\,\|\theta\|_3^3}{\|\theta\|_1} \lesssim \lambda_1^3$$
$$|\mathbb{E}T_2| \le \frac{\|\beta\circ\theta\|_2^2\,\|\theta\|_2^4}{\|\theta\|_1^2} \lesssim |\tilde{\lambda}|, \quad \mathrm{Var}(T_2) \lesssim \frac{\|\beta\circ\theta\|_2^4\,\|\theta\|_2^8}{\|\theta\|_1^2} \lesssim |\tilde{\lambda}|^2\lambda_1^2$$
$$|\mathbb{E}F| \lesssim \|\theta\|_2^4 \lesssim \lambda_1^2, \qquad \mathrm{Var}(F) \lesssim \frac{\|\theta\|_2^{10}}{\|\theta\|_1^2} \lesssim \lambda_1^3$$

Using Corollary E.1 and the fact that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$ under the alternative hypothesis, we have

$$|\mathbb{E}U_c| \lesssim \lambda_1^2, \tag{E.29}$$

and

$$\mathrm{Var}(U_c) \lesssim |\tilde{\lambda}|^2\lambda_1^2. \tag{E.30}$$

Using Corollary E.1 and that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$, the inequalities (E.25)–(E.30) imply Theorem E.3.

E.3.3 Real SgnQ

Our first lemma regarding real SgnQ plays the part of Lemma G.11 from Jin et al. (2021).

Lemma E.11. *Under the previous assumptions, as $n \to \infty$,*

- *Under the null hypothesis, $|\mathbb{E}[Q^* - \tilde{Q}^*]| = o(\|\theta\|_2^4)$ and $\mathrm{Var}(Q^* - \tilde{Q}^*) = o(\|\theta\|_2^8)$.*
- *Under the alternative hypothesis, if $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$, then $|\mathbb{E}[Q^* - \tilde{Q}^*]| \lesssim |\tilde{\lambda}|^2\lambda_1$ and $\mathrm{Var}(Q^* - \tilde{Q}^*) \lesssim |\tilde{\lambda}|^2\lambda_1^3$.*

The following lemma plays the part of Lemma G.12 from Jin et al. (2021).

Lemma E.12. *Under the previous assumptions, as $n \to \infty$,*

- *Under the null hypothesis, $|\mathbb{E}[Q - \tilde{Q}^*]| = o(\|\theta\|_2^4)$ and $\mathrm{Var}(Q - \tilde{Q}^*) = o(\|\theta\|_2^8)$.*
- *Under the alternative hypothesis, if $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$, then $|\mathbb{E}[Q - \tilde{Q}^*]| \lesssim \lambda_1^2 + |\tilde{\lambda}|^3$ and $\mathrm{Var}(Q - \tilde{Q}^*) \lesssim \lambda_1^4$.*

E.4 Proofs of Lemmas E.7–E.12

E.4.1 Proof strategy

First we describe our method of proof for Lemmas E.7–E.10. We borrow the following strategy from Jin et al. (2021). Let T denote a term appearing in one of the Lemmas E.7–E.10, which takes the general form

$$T = \sum_{i_1,\dots,i_m \in \mathcal{R}} c_{i_1,\dots,i_m} G_{i_1,\dots,i_m}$$

where

- $m = O(1)$,
- $\mathcal{R}$ is a subset of $[n]^m$,
- $c_{i_1,\dots,i_m} = \prod_{(s,s')\in A} \Gamma^{(s,s')}_{i_s,i_{s'}}$ is a nonstochastic coefficient where $A \subset [m]\times[m]$ and $\Gamma^{(s,s')} \in \{\tilde{\Omega}, \eta^*\mathbf{1}^\mathsf{T}, \eta\mathbf{1}^\mathsf{T}, \mathbf{1}\mathbf{1}^\mathsf{T}\}$, and
- $G_{i_1,\dots,i_m} = \prod_{(s,s')\in B} W_{i_s,i_{s'}}$ where $B \subset [m]\times[m]$.

Since we are studying signed quadrilateral, one can simply take $m = 4$ above, though we wish to state the lemma in a general way.

Define a *canonical upper bound* $\overline{\Gamma^{(s,s')}_{i_s,i_{s'}}}$ (up to constant factor) on $\Gamma^{(s,s')}_{i_s,i_{s'}}$ as follows:

$$\overline{\Gamma^{(s,s')}_{i_s,i_{s'}}} = \begin{cases} \beta_{i_s}\theta_{i_s}\beta_{i_{s'}}\theta_{i_{s'}} & \text{if } \Gamma^{(s,s')} = \tilde{\Omega}, \\ \theta_{i_s} & \text{if } \Gamma^{(s,s')} \in \{\eta^*\mathbf{1}^\mathsf{T}, \eta\mathbf{1}^\mathsf{T}\} \\ 1 & \text{otherwise.} \end{cases} \tag{E.31}$$

Define

$$\overline{c_{i_1,\dots,i_m}} = \prod_{(s,s')\in A} \overline{\Gamma^{(s,s')}_{i_s,i_{s'}}}. \tag{E.32}$$

By Corollary E.1 and Lemma E.6,

$$|c_{i_1,\dots i_m}| \lesssim \overline{c_{i_1,\dots,i_m}}.$$

In Jin et al. (2021), each term T is decomposed into a sum of $L = O(1)$ terms:

$$T = \sum_{\ell=1}^{L} T^{(L)} = \sum_{\ell=1}^{L} \sum_{i_1,\dots,i_m\in\mathcal{R}^{(\ell)}} c_{i_1,\dots,i_m} G_{i_1,\dots,i_m}. \tag{E.33}$$

In our analysis below and that of Jin et al. (2021), an upper bound $\overline{\mathbb{E}T}$ on $|\mathbb{E}T|$ is obtained by

$$\begin{aligned} |\mathbb{E}T| &\le \sum_{\ell=1}^{L} |\mathbb{E}T^{(\ell)}| \le \sum_{\ell=1}^{L} \sum_{i_1,\dots,i_m\in\mathcal{R}^{(\ell)}} |c_{i_1,\dots,i_m}| \cdot |\mathbb{E}G_{i_1,\dots,i_m}| \\ &\le \sum_{\ell=1}^{L} \sum_{i_1,\dots,i_m\in\mathcal{R}^{(\ell)}} \overline{c_{i_1,\dots,i_m}} \cdot |\mathbb{E}G_{i_1,\dots,i_m}| \\ &=: \overline{\mathbb{E}T}. \end{aligned} \tag{E.34}$$

Also an upper bound $\overline{\mathrm{Var}T}$ on $\mathrm{Var}T$ is obtained by

$$\begin{aligned} \mathrm{Var}T &\le L\sum_{\ell=1}^{L} \mathrm{Var}(T^{(\ell)}) \\ &\le L\sum_{\ell=1}^{L} \sum_{\substack{i_1,\dots,i_m\in\mathcal{R}^{(\ell)} \\ i'_1,\dots,i'_m\in\mathcal{R}^{(\ell)}}} |c_{i_1,\dots,i_m} c_{i'_1,\dots,i'_m}| \cdot \big|\mathrm{Cov}\big(G_{i_1,\dots,i_m}, G_{i'_1,\dots,i'_m}\big)\big| \\ &\le L\sum_{\ell=1}^{L} \sum_{\substack{i_1,\dots,i_m\in\mathcal{R}^{(\ell)} \\ i'_1,\dots,i'_m\in\mathcal{R}^{(\ell)}}} \overline{c_{i_1,\dots,i_m}} \cdot \overline{c_{i'_1,\dots,i'_m}} \cdot \big|\mathrm{Cov}\big(G_{i_1,\dots,i_m}, G_{i'_1,\dots,i'_m}\big)\big| \\ &=: \overline{\mathrm{Var}T}. \end{aligned} \tag{E.35}$$

In Lemmas E.7–E.10, all stated upper bounds are obtained in this manner and are therefore upper bounds on $\overline{\mathbb{E}T}$ and $\overline{\mathrm{Var}T}$.

Note that the definition of $\overline{\mathbb{E}T}$ and $\overline{\mathrm{Var}T}$ depends on the specific decomposition (E.33) of T given in Jin et al. (2021). Refer to the proofs below for details including the explicit decomposition. Again we remark that the difference between our setting and Jin et al. (2021) is that the canonical upper bound on $|\tilde{\Omega}_{ij}|$ used in Jin et al. (2021) is of the form $\alpha\theta_i\theta_j$ rather than the inequality $\beta_i\theta_i\beta_j\theta_j$ which is required for our purposes.

The formalism above immediately yields the following useful fact that allows us to transfer bounds between terms that have similar structures.

Lemma E.13. *Suppose that*

$$T = \sum_{i_1,\dots,i_m \in \mathcal{R}} c_{i_1,\dots,i_m} G_{i_1,\dots,i_m},$$
$$T^* = \sum_{i_1,\dots,i_m \in \mathcal{R}} c^*_{i_1,\dots,i_m} G_{i_1,\dots,i_m},$$

where

$$|c_{i_1,\dots,i_m}| \lesssim \overline{c^*_{i_1,\dots,i_m}}$$

Then

$$|\mathbb{E}T| \lesssim \overline{\mathbb{E}[T^*]}$$

and

$$\mathrm{Var}\, T \lesssim \overline{\mathrm{Var}\, T^*}.$$

In the second part of our analysis, we show that Lemmas E.11 and E.12 follow from Lemmas E.7–E.10 and repeated applications of Lemma E.13.

E.4.2 PROOF OF LEMMA E.7

The bounds for X_1 follow immediately from Jin et al. (2021).

In (Jin et al., 2021, Supplement, pg.37) it is shown that $\mathbb{E}X_2 = 0$, and

$$\mathrm{Var}(X_2) = 2 \sum_{i,j,k,\ell (dist.)} \tilde{\Omega}_{ij}^2 \cdot \mathrm{Var}(W_{jk}W_{k\ell}W_{\ell i}).$$

Thus by (E.1) and (E.2),

$$\mathrm{Var}(X_2) \lesssim \sum_{i,j,k,\ell (dist.)} \tilde{\Omega}_{ij}^2 \cdot \mathrm{Var}(W_{jk}W_{k\ell}W_{\ell i}) \lesssim \sum_{i,j,k,\ell} \beta_i^2\theta_i^2\beta_j^2\theta_j^2 \cdot \Omega_{jk}\Omega_{k\ell}\Omega_{\ell i}$$
$$\lesssim \sum_{i,j,k,\ell} \beta_i^2\theta_i^2\beta_j^2\theta_j^2 \cdot \theta_j\theta_k^2\theta_\ell^2\theta_i = \|\beta\circ\theta\|_2^4\, \|\theta\|_2^4$$

In (Jin et al., 2021, Supplement, pg. 38) it is shown that $\mathbb{E}X_3 = 0$ and

$$\mathrm{Var}(X_3) \lesssim \sum_{i,k,\ell (dist)} \Big(\sum_{j \notin \{i,k,\ell\}} \tilde{\Omega}_{ij}\tilde{\Omega}_{jk} \Big)^2 \cdot \mathrm{Var}(W_{k\ell}W_{\ell i}).$$

By (E.20) and (E.24),

$$\Big(\sum_{j \notin \{i,k,\ell\}} \tilde{\Omega}_{ij}\tilde{\Omega}_{jk} \Big)^2 \le \beta_i^2\theta_i^2\, \beta_k^2\theta_k^2\, \|\beta\circ\theta\|_2^4$$

Thus by (E.1) and (E.2),

$$\mathrm{Var}(X_3) \lesssim \sum_{i,k,\ell} \beta_i^2\theta_i^2\, \beta_k^2\theta_k^2\, \|\beta\circ\theta\|_2^4 \cdot \Omega_{k\ell}\Omega_{\ell i} \lesssim \sum_{i,k,\ell} \beta_i^2\theta_i^3\, \beta_k^2\theta_k^3\, \|\beta\circ\theta\|_2^4 \cdot \theta_\ell^2 \quad \lesssim \|\beta\circ\theta\|_2^8\, \|\theta\|_2^2.$$

In (Jin et al., 2021, Supplement, pg. 38) it is shown that $\mathbb{E}X_4 = 0$ and

$$\mathrm{Var}(X_4) \lesssim \sum_{i,j,k,\ell (dist.)} \tilde{\Omega}_{ij}^2\tilde{\Omega}_{k\ell}^2 \cdot \mathrm{Var}(W_{jk}W_{\ell i}).$$

By (E.1) and (E.20),

$$\mathrm{Var}(X_4) \lesssim \sum_{i,j,k,\ell} \beta_i^2\theta_i^2\beta_j^2\theta_j^2\beta_k^2\theta_k^2\beta_\ell^2\theta_\ell^2 \cdot \theta_j\theta_k\theta_\ell\theta_i \lesssim \|\beta\circ\theta\|_2^8.$$

In (Jin et al., 2021, Supplement, pg. 39) it is shown that $\mathbb{E}X_5 = 0$ and

$$\mathrm{Var}(X_5) = 2\sum_{i<\ell}\Big(\sum_{\substack{j,k\notin\{i,\ell\}\\ j\neq k}} \tilde{\Omega}_{ij}\tilde{\Omega}_{jk}\tilde{\Omega}_{k\ell}\Big)^2 \cdot \mathrm{Var}(W_{\ell i}).$$

We have

$$\Big|\sum_{\substack{j,k\notin\{i,\ell\}\\ j\neq k}} \tilde{\Omega}_{ij}\tilde{\Omega}_{jk}\tilde{\Omega}_{k\ell}\Big| \lesssim \beta_i\theta_i\|\beta\circ\theta\|_2^4\beta_\ell\theta_\ell.$$

Thus by (E.1) and (E.2),

$$\mathrm{Var}(X_5) \lesssim \sum_{i,\ell}\big(\beta_i\theta_i\|\beta\circ\theta\|_2^4\beta_\ell\theta_\ell\big)^2 \cdot \theta_\ell\theta_i \lesssim \|\beta\circ\theta\|_2^{12}.$$

Note that X_6 is a nonstochastic term. Mimicking (Jin et al., 2021, Supplement, pg. 39), we have by (E.24),

$$|X_6 - \tilde{\lambda}^4| \lesssim \sum_{i,j,k,\ell (not\, dist.)} \beta_i^2\theta_i^2\beta_j^2\theta_j^2\beta_k^2\theta_k^2\beta_\ell^2\theta_\ell^2 \lesssim \sum_{i,j,k} \beta_i^2\theta_i^2\beta_j^2\theta_j^2\beta_k^4\theta_k^4 \lesssim \|\beta\circ\theta\|_2^6 \lesssim |\tilde{\lambda}|^3.$$

This completes the proof. □

E.4.3 Proof of Lemma E.8

The bounds on Y_1 carry over directly from (Jin et al., 2021, Lemma G.5).

In (Jin et al., 2021, Supplement, pg. 43) it is shown that $\mathbb{E}Y_2 = 0$. To study $\mathrm{Var}(Y_2)$, we write $Y = Y_{2a} + Y_{2b} + Y_{2c}$ where as in (Jin et al., 2021, Supplement, pg. 43), we define

$$\begin{aligned} Y_2 \;&=\; -\frac{1}{\sqrt{v}} \sum_{\substack{i,j,k,\ell(dist)\\ s\neq j}} \eta_i\widetilde{\Omega}_{jk}W_{js}W_{k\ell}W_{\ell i} \\ &\quad -\frac{1}{\sqrt{v}} \sum_{i,k,\ell(dist)}\Big(\sum_{j\notin\{i,k,\ell\}} \eta_j\widetilde{\Omega}_{jk}\Big)W_{i\ell}^2W_{k\ell} \\ &\quad -\frac{1}{\sqrt{v}} \sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\}}}\Big(\sum_{j\notin\{i,k,\ell\}} \eta_j\widetilde{\Omega}_{jk}\Big)W_{is}W_{k\ell}W_{\ell i} \\ &\equiv\; Y_{2a} + Y_{2b} + Y_{2c}. \end{aligned} \tag{E.36}$$

There it is shown that

$$\mathrm{Var}(Y_{2a}) \lesssim \frac{1}{v}\sum_{ijk\ell s}\big|\eta_i\tilde{\Omega}_{jk} + \eta_i\tilde{\Omega}_{sk} + \eta_k\tilde{\Omega}_{ji} + \eta_k\tilde{\Omega}_{si}\big|^2 \cdot \mathrm{Var}(W_{js}W_{k\ell}W_{\ell i}).$$

We have by (E.22)

$$\big|\eta_i\tilde{\Omega}_{jk} + \eta_i\tilde{\Omega}_{sk} + \eta_k\tilde{\Omega}_{ji} + \eta_k\tilde{\Omega}_{si}\big| \lesssim \theta_i\beta_j\theta_j\beta_k\theta_k + \theta_i\beta_s\theta_s\beta_k\theta_k + \theta_k\beta_j\theta_j\beta_i\theta_i + \theta_k\beta_s\theta_s\beta_i\theta_i.$$

Hence by (E.1), (E.2), and (E.18),

$$\begin{aligned} \mathrm{Var}(Y_{2a}) &\lesssim \frac{1}{v}\sum_{ijk\ell s}\big(\theta_i\beta_j\theta_j\beta_k\theta_k + \theta_i\beta_s\theta_s\beta_k\theta_k + \theta_k\beta_j\theta_j\beta_i\theta_i + \theta_k\beta_s\theta_s\beta_i\theta_i\big)^2 \cdot \theta_j\theta_s\theta_k\theta_\ell^2\theta_i \\ &\lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^4}{\|\theta\|_1} \end{aligned}$$

Next, in (Jin et al., 2021, Supplement, pg. 43), it is shown that

$$\mathrm{Var}(Y_{2b}) \lesssim \frac{1}{v} \sum_{\substack{ik\ell(dist) \\ i'k'\ell'(dist)}} |\alpha_{ik\ell}\alpha_{i'k'\ell'}| \cdot \mathbb{E}[W_{i\ell}^2 W_{k\ell}, W_{i'\ell'}^2 W_{k'\ell'}]$$

where $\alpha_{ik\ell} = \sum_{j \notin \{i,k,\ell\}} \eta_j \tilde{\Omega}_{jk}$. By (E.24),

$$|\alpha_{ik\ell}| \lesssim \|\beta \circ \theta\|_2 \|\theta\|_2 \, \theta_k.$$

By (E.1), (E.18), the inequalities above, and the casework in (Jin et al., 2021, Supplement, pg.44) on $E[W_{i\ell}^2 W_{k\ell}, W_{i'\ell'}^2 W_{k'\ell'}]$,

$$\begin{aligned}
\mathrm{Var}(Y_{2b}) &\lesssim \frac{1}{v} \sum_{\substack{ik\ell(dist) \\ i'k'\ell'(dist)}} \|\beta \circ \theta\|_2^2 \|\theta\|_2^2 \theta_k \theta_{k'} \mathbb{E}[W_{i\ell}^2 W_{k\ell}, W_{i'\ell'}^2 W_{k'\ell'}] \\
&\lesssim \frac{\|\beta \circ \theta\|_2^2 \|\theta\|_2^2}{v} \big(\sum_{ik\ell} \theta_i \theta_k^3 \theta_\ell^2 + \sum_{ik\ell i'} \theta_i \theta_k^3 \theta_\ell^3 \theta_{i'} + \sum_{ik\ell} \theta_i^2 \theta_k^2 \theta_\ell^2 \big) \\
&\lesssim \|\beta \circ \theta\|_2^2 \|\theta\|_2^6.
\end{aligned}$$

Next, in (Jin et al., 2021, Supplement, pg.44) it is shown that

$$\mathrm{Var}(Y_{2c}) \lesssim \frac{1}{v} \sum_{\substack{ik\ell(dist) \\ s \notin \{i,\ell\}}} \beta_{ik\ell}^2 \mathrm{Var}(W_{is} W_{k\ell} W_{\ell i})$$

where $\alpha_{ik\ell}$ is defined the same as with Y_{2b}. Thus

$$\mathrm{Var}(Y_{2c}) \lesssim \frac{1}{v} \sum_{\substack{ik\ell(dist) \\ s \notin \{i,\ell\}}} \|\beta \circ \theta\|_2^2 \|\theta\|_2^2 \theta_k^2 \cdot \theta_k \theta_\ell^2 \theta_s \lesssim \frac{\|\beta \circ \theta\|_2^2 \|\theta\|_2^8}{\|\theta\|_1}.$$

Combining the results for Y_{2a}, Y_{2b}, Y_{2c} gives the claim for $\mathrm{Var}(Y_2)$.

In (Jin et al., 2021, Supplement, pg.45) it is shown that $\mathbb{E}Y_3 = 0$ and the decomposition

$$\begin{aligned}
Y_3 &= -\frac{2}{\sqrt{v}} \sum_{i,j,k,\ell(dist)} \eta_i \widetilde{\Omega}_{k\ell} W_{jk}^2 W_{\ell i} - \frac{2}{\sqrt{v}} \sum_{\substack{i,j,k,\ell(dist) \\ s \notin \{j,k\}}} \eta_i \widetilde{\Omega}_{k\ell} W_{js} W_{jk} W_{\ell i} \\
&\equiv Y_{3a} + Y_{3b},
\end{aligned} \tag{E.37}$$

is introduced. There it is shown that

$$\mathrm{Var}(Y_{3a}) = \frac{4}{v} \sum_{\substack{i,j,k,\ell(dist) \\ i',j',k',\ell'(dist)}} (\eta_i \widetilde{\Omega}_{k\ell} \eta_{i'} \widetilde{\Omega}_{k'\ell'}) \cdot \mathbb{E}[W_{jk}^2 W_{\ell i} W_{j'k'}^2 W_{\ell' i'}].$$

Using (E.1), (E.2) (E.24) and the casework in (Jin et al., 2021, Supplement, pg.45),

$$\begin{aligned}
\mathrm{Var}(Y_{3a}) &\lesssim \frac{1}{\|\theta\|_1^2} \left(\sum_{ijk\ell} [\beta_k^2 \beta_\ell^2 + \beta_i \beta_j \beta_k \beta_\ell] \theta_i^2 \theta_j^2 \theta_k^2 \theta_\ell^2 + \sum_{ijk\ell j'k'} \beta_k \beta_\ell^2 \beta_{k'} \theta_i^3 \theta_j \theta_k^2 \theta_\ell^3 \theta_{j'} \theta_{k'}^2 \right) \\
&\lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^4}{\|\theta\|_1^2} + \|\beta \circ \theta\|_2^4 \|\theta\|_2^4 \lesssim \|\beta \circ \theta\|_2^4 \|\theta\|_2^4
\end{aligned}$$

Similar to the study of Y_{2a} we have

$$\mathrm{Var}(Y_{3b}) \lesssim \frac{1}{v} \sum_{ijk\ell s} \big(\theta_i \beta_k \theta_k \beta_\ell \theta_\ell + \theta_\ell \beta_k \theta_k \beta_i \theta_i + \theta_i \beta_s \theta_s \beta_\ell \theta_\ell + \theta_\ell \beta_s \theta_s \beta_i \theta_i \big)^2 \cdot \mathrm{Var}(W_{sj} W_{jk} W_{\ell i})$$

$$
\begin{aligned}
&\lesssim \frac{1}{v}\sum_{ijk\ell s}\big(\theta_i\beta_k\theta_k\beta_\ell\theta_\ell + \theta_\ell\beta_k\theta_k\beta_i\theta_i + \theta_i\beta_s\theta_s\beta_\ell\theta_\ell + \theta_\ell\beta_s\theta_s\beta_i\theta_i\big)^2 \cdot \theta_s\theta_j^2\theta_k\theta_\ell\theta_i \\
&\lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^4}{\|\theta\|_1}.
\end{aligned}
$$

Combining the bounds on $\mathrm{Var}(Y_{3a})$ and $\mathrm{Var}(Y_{3b})$ yields the desired bound on $\mathrm{Var}(Y_3)$.

Following (Jin et al., 2021, Supplement, pg.46) we obtain the decomposition

$$
\begin{aligned}
Y_4 &= -\frac{1}{\sqrt{v}}\sum_{\substack{i,j,\ell(dist)\\ s\neq j}}\Big(\sum_{k\notin\{i,j,\ell\}}\eta_i\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\Big)W_{js}W_{\ell i} - \frac{1}{\sqrt{v}}\sum_{\substack{i,\ell(dist)\\ s\neq i}}\Big(\sum_{j,k\notin\{i,\ell\}}\eta_j\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\Big)W_{is}W_{\ell i} \\
&\equiv Y_{4a} + Y_{4b}.
\end{aligned}
$$

First we study Y_{4a}, which is shown in Jin et al. (2021) to have zero mean and satisfy the following:

$$
\mathrm{Var}(Y_{4a}) \lesssim \frac{1}{v}\sum_{\substack{ij\ell(dist)\\ s\neq j}}\alpha_{ij\ell}^2\mathrm{Var}(W_{js}W_{\ell i})
$$

where $\alpha_{ij\ell} = \sum_{k\notin\{i,j,\ell\}}\eta_i\tilde{\Omega}_{jk}\tilde{\Omega}_{k\ell}$. Simlar to previous arguments, we have

$$
\begin{aligned}
\mathrm{Var}(Y_{4a}) &\lesssim \frac{1}{\|\theta\|_1^2}\sum_{ij\ell s}\theta_i^2(\beta_j\theta_j)^2(\beta_\ell\theta_\ell)^2\|\beta\circ\theta\|_2^4\cdot\theta_i\theta_j\theta_\ell\theta_s \\
&\lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^2}{\|\theta\|_1}.
\end{aligned}
$$

Next we study Y_{4b} using the decomposition

$$
Y_{4b} = -\frac{1}{\sqrt{v}}\sum_{i,\ell(dist)}\beta_{i\ell}W_{\ell i}^2 - \frac{1}{\sqrt{v}}\sum_{\substack{i,\ell(dist)\\ s\notin\{i,\ell\}}}\beta_{i\ell}W_{is}W_{\ell i} \equiv \widetilde{Y}_{4b} + Y_{4b}^*.
$$

from (Jin et al., 2021, Supplement,pg.47). There it is shown that only $\mathbb{E}\tilde{Y}_{4b}$ is nonzero and

$$
|\mathbb{E}\tilde{Y}_{4b}| \lesssim \frac{1}{\|\theta\|_1}\sum_{i,\ell}|\alpha_{i\ell}|\theta_i\theta_\ell.
$$

where $\alpha_{i,\ell} = \sum_{j,k\notin\{i,\ell\}}\eta_j\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}$. In our case, we derive from (E.24),

$$
|\alpha_{i\ell}| \lesssim \beta_\ell\theta_\ell\|\beta\circ\theta\|_2^3\|\theta\|_2.
$$

Using similar arguments from before,

$$
|\mathbb{E}\tilde{Y}_{4b}| \lesssim \frac{1}{\|\theta\|_1}\sum_{i\ell}\beta_\ell\theta_\ell\|\beta\circ\theta\|_2^3\|\theta\|_2\cdot\theta_i\theta_\ell \lesssim \|\beta\circ\theta\|_2^4\|\theta\|_2^2.
$$

Now we study $\mathrm{Var}(Y_{4b})$. Using the bound above on $|\alpha_{i\ell}|$ and direct calculations,

$$
\begin{aligned}
\mathrm{Var}(\widetilde{Y}_{4b}) &= \frac{2}{v}\sum_{i,\ell(dist)}\alpha_{i\ell}^2\cdot\mathrm{Var}(W_{i\ell}^2) \lesssim \frac{1}{\|\theta\|_1^2}\sum_{i,\ell}\beta_\ell^2\theta_\ell^2\|\beta\circ\theta\|_2^6\|\theta\|_2^2\cdot\theta_i\theta_\ell \lesssim \frac{\|\beta\circ\theta\|_2^8\|\theta\|_2^2}{\|\theta\|_1}, \\
\mathrm{Var}(Y_{4b}^*) &\leq \frac{1}{v}\sum_{\substack{i,\ell(dist)\\ s\notin\{i,\ell\}}}\alpha_{i\ell}^2\cdot\mathrm{Var}(W_{is}W_{\ell i}) \leq \frac{1}{\|\theta\|_1^2}\sum_{i,\ell,s}\beta_\ell^2\theta_\ell^2\|\beta\circ\theta\|_2^6\|\theta\|_2^2\cdot\theta_i^2\theta_\ell\theta_s \leq \frac{\|\beta\circ\theta\|_2^8\|\theta\|_2^4}{\|\theta\|_1}.
\end{aligned}
$$

Combining the results above yields the required bounds on $\mathbb{E}Y_{4b}$ and $\mathrm{Var}(Y_{4b})$.

In (Jin et al., 2021, Supplement, pg.48) it is shown that $\mathbb{E}Y_5 = 0$ and

$$\mathrm{Var}(Y_5) \lesssim \frac{1}{v} \sum_{\substack{j,k,\ell(dist)\\ s\neq j}} \alpha_{jk\ell}^2 \cdot \mathrm{Var}(W_{js}W_{k\ell})$$

where

$$\alpha_{jk\ell} \equiv \sum_{i\notin\{j,k,\ell\}} \eta_i \widetilde{\Omega}_{jk}\widetilde{\Omega}_{\ell i}.$$

We have using (E.20), (E.24) and the triangle inequality,

$$|\alpha_{jk\ell}| \lesssim \|\theta\|_2^2(\beta_j\theta_j)(\beta_k\theta_k)(\beta_\ell\theta_\ell).$$

Thus, by similar arguments to before,

$$\mathrm{Var}(Y_5) \lesssim \frac{1}{\|\theta\|_1^2} \sum_{jk\ell} \big(\|\theta\|_2^4(\beta_j\theta_j)^2(\beta_k\theta_k)^2(\beta_\ell\theta_\ell)^2\big)\theta_j\theta_s\theta_k\theta_\ell \lesssim \frac{\|\theta\|_2^4\|\beta\circ\theta\|_2^6}{\|\theta\|_1}.$$

Next, in (Jin et al., 2021, Supplement, pg.49) it is shown that $\mathbb{E}Y_6 = 0$ and

$$\mathrm{Var}(Y_6) = \frac{8}{v} \sum_{j,s(dist)} \Big(\sum_{i,k,\ell(dist)\notin\{j\}} \eta_i \widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\Big)^2 \cdot \mathrm{Var}(W_{js}).$$

We have using (E.20), (E.24) and the triangle inequality,

$$\Big| \sum_{i,k,\ell(dist)\notin\{j\}} \eta_i \widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\Big| \lesssim \beta_j\theta_j\|\beta\circ\theta\|_2^5\|\theta\|_2.$$

Thus

$$\mathrm{Var}(Y_6) \lesssim \frac{1}{\|\theta\|_1^2} \sum_{j,s} \big(\beta_j^2\theta_j^2\|\beta\circ\theta\|_2^{10}\|\theta\|_2^2\big)\theta_j\theta_s \lesssim \frac{\|\beta\circ\theta\|_2^{12}\|\theta\|_2^2}{\|\theta\|_1}.$$

This completes the proof. □

E.4.4 PROOF OF LEMMA E.9

The bounds on Z_1 and Z_2 carry over directly from (Jin et al., 2021, Lemma G.7) since neither term depends on $\tilde{\Omega}$.

We consider Z_3. In (Jin et al., 2021, Supplement, pg.61), the decomposition

$$\begin{aligned} Z_3 &= \sum_{\substack{i,j,k,\ell\\(dist)}} \eta_i(\eta_j-\tilde{\eta}_j)\eta_j(\eta_k-\tilde{\eta}_k)\widetilde{\Omega}_{k\ell}W_{\ell i} + \sum_{\substack{i,j,k,\ell\\(dist)}} \eta_i(\eta_j-\tilde{\eta}_j)^2\eta_k\widetilde{\Omega}_{k\ell}W_{\ell i} \\ &+ \sum_{\substack{i,j,k,\ell\\(dist)}} (\eta_i-\tilde{\eta}_i)\eta_j^2(\eta_k-\tilde{\eta}_k)\widetilde{\Omega}_{k\ell}W_{\ell i} + \sum_{\substack{i,j,k,\ell\\(dist)}} (\eta_i-\tilde{\eta}_i)\eta_j(\eta_j-\tilde{\eta}_j)\eta_k\widetilde{\Omega}_{k\ell}W_{\ell i} \\ &\equiv Z_{3a} + Z_{3b} + Z_{3c} + Z_{3d}. \end{aligned} \tag{E.38}$$

is introduced. We study each term separately.

In (Jin et al., 2021, Supplement, pg.61) it is shown that $\mathbb{E}Z_{3a} = 0$ and the decomposition

$$\begin{aligned} Z_{3a} &= \frac{1}{v}\sum_{i,j,k,\ell(dist)} \alpha_{ijk\ell}W_{jk}^2W_{\ell i} + \frac{1}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq j, t\neq k, (s,t)\neq(k,j)}} \alpha_{ijk\ell}W_{js}W_{kt}W_{\ell i} \\ &\equiv \widetilde{Z}_{3a} + Z_{3a}^*. \end{aligned}$$

is introduced, where $\alpha_{ijk\ell} \equiv \eta_i \eta_j \tilde{\Omega}_{k\ell}$. Then

$$\mathrm{Var}(\tilde{Z}_{3a}) \lesssim \sum_{\substack{ijk\ell(dist)\\ i'j'k'\ell'(dist)}} |\alpha_{ijk\ell}||\alpha_{i'k'j'\ell'}| \cdot |\mathrm{Cov}(W_{jk}^2 W_{\ell i}, W_{j'k'}^2 W_{\ell' i'})|.$$

Using the casework in (Jin et al., 2021, Supplement, pg.62), (E.1), (E.2), and (E.24), we obtain

$$\begin{aligned}\mathrm{Var}(\tilde{Z}_{3a}) &\lesssim \frac{1}{v^2}\big(\sum_{ijk\ell}[\beta_k^2\beta_\ell^2 + \beta_k\beta_\ell\beta_i\beta_j]\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3 + \sum_{ijk\ell j'k'} \beta_k\beta_\ell^2\beta_{k'}\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3\theta_{j'}^2\theta_{k'}^2\big)\\ &\lesssim \frac{1}{\|\theta\|_1^4}\big(\|\beta\circ\theta\|_2^4\|\theta\|_2^2 + \|\beta\circ\theta\|_2^4\|\theta\|_2^4 + \|\beta\circ\theta\|_2^4\|\theta\|_2^8\big) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^8}{\|\theta\|_1^4}.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{Var}(Z_{3a}^*) &\lesssim \frac{1}{v^2}\left(\sum_{ijk\ell st}\beta_k^2\beta_\ell^2\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3\theta_s\theta_t + \sum_{ijk\ell st}[\beta_k^2\beta_\ell\beta_j + \beta_k\beta_\ell^2\beta_j]\theta_i^2\theta_j^3\theta_k^3\theta_\ell^3\theta_s^2\theta_t\right)\\ &\lesssim \frac{1}{\|\theta\|_1^4}\big(\|\beta\circ\theta\|_2^4\|\theta\|_2^4\|\theta\|_1^2 + \|\beta\circ\theta\|_2^4\|\theta\|_2^6\|\theta\|_1\big) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^4}{\|\theta\|_1^2}.\end{aligned}$$

It follows that

$$\mathrm{Var}(Z_{3a}) \lesssim \|\beta\circ\theta\|_2^4.$$

Next, in (Jin et al., 2021, Supplement, pg.63), it is shown that $\mathbb{E}Z_{3b}] = 0$ and the decomposition

$$Z_{3b} = \frac{1}{v}\sum_{\substack{i,j,\ell(dist)\\ s\neq j}} \beta_{ij\ell}W_{js}^2 W_{\ell i} + \frac{1}{v}\sum_{\substack{i,j,\ell(dist)\\ s,t(dist)\notin\{j\}}} \beta_{ij\ell}W_{js}W_{jt}W_{\ell i} \equiv \widetilde{Z}_{3b} + Z_{3b}^*.$$

is given. Using (Jin et al., 2021, Supplement, pg.63) we have

$$\mathrm{Var}(\tilde{Z}_{3b}) \lesssim \sum_{\substack{i,j,\ell,s,t\\ i',j',\ell',s',t'}} |\alpha_{ij\ell}||\alpha_{i'j'\ell'}||\mathrm{Cov}(W_{js}^2 W_{\ell_i}, W_{j's}'^2 W_{\ell' i'})|.$$

where

$$\alpha_{ij\ell} = \sum_{k\notin\{i,j,\ell\}} \eta_i\eta_k\widetilde{\Omega}_{k\ell}.$$

Using (E.24), (E.18), and similar arguments to before,

$$|\alpha_{ij\ell}| \lesssim \theta_i(\beta_\ell\theta_\ell)\|\theta\|_2^2.$$

By the casework in (Jin et al., 2021, Supplement, pg.63), (E.1), and (E.2),

$$\begin{aligned}\mathrm{Var}(\tilde{Z}_{3b}) &\lesssim \frac{1}{v^2}\left(\sum_{ij\ell s}\beta_\ell^2\|\theta\|_2^4\theta_i^3\theta_j\theta_\ell^3\theta_s + \sum_{ij\ell sj's'}\beta_\ell^2\|\theta\|_2^4\theta_i^3\theta_j\theta_\ell^3\theta_s\theta_{j'}\theta_{s'} + \sum_{ij\ell s}\beta_\ell\beta_j\|\theta\|_2^4\theta_i^2\theta_j^2\theta_\ell^2\theta_s^2\right)\\ &\lesssim \frac{1}{\|\theta\|_1^4}\big(\|\beta\circ\theta\|_2^2\|\theta\|_2^6\|\theta\|_1 + \|\beta\circ\theta\|_2^2\|\theta\|_2^6\|\theta\|_1^3 + \|\beta\circ\theta\|_2^2\|\theta\|_2^8\big) \lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^6}{\|\theta\|_1}.\end{aligned}$$

By a similar argument,

$$\mathrm{Var}(Z_{3b}^*) \lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^8}{\|\theta\|_1}.$$

Hence by (E.2),

$$\mathrm{Var}(Z_{3b}) \lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^8}{\|\theta\|_1} \lesssim \|\beta\circ\theta\|_2^2\|\theta\|_2^6.$$

For Z_{3c}, in (Jin et al., 2021, Supplement, pg.64), it is shown that $\mathbb{E}Z_{3c} = 0$ and the decomposition

$$Z_{3c} = \frac{1}{v} \sum_{\substack{i,k,\ell(dist)\\ t\neq k}} \alpha_{ik\ell} W_{i\ell}^2 W_{kt} + \frac{1}{v} \sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\}, t\neq k}} \alpha_{ik\ell} W_{is} W_{kt} W_{\ell i} \equiv \widetilde{Z}_{3c} + Z_{3c}^*.$$

is given. We have

$$|\alpha_{ik\ell}| = |\sum_{j\notin\{i,k,\ell\}} \eta_j^2 \widetilde{\Omega}_{k\ell}| \lesssim (\beta_k\theta_k)(\beta_\ell\theta_\ell)\|\theta\|_2^2.$$

By the casework in (Jin et al., 2021, Supplement, pg.65)

$$\begin{aligned}
\mathrm{Var}(\tilde{Z}_{3c}) \lesssim & \sum_{\substack{ik\ell(dist)\\ s\notin\{i,\ell\}, t\neq k}} \sum_{\substack{i'k'\ell'(dist)\\ s'\notin\{i',\ell'\}, t'\neq k'}} |\alpha_{ik\ell}\alpha_{i'k'\ell'}||\mathbb{E}W_{i\ell}^2 W_{kt} W_{i'\ell'}^2 W_{k't'}| \\
& \lesssim \frac{\|\theta\|_2^4}{\|\theta\|_1^4} \sum_{ik\ell t} \Big[\beta_k^2\beta_\ell^2\theta_i\theta_k^3\theta_\ell^3\theta_t + \beta_k^2\beta_\ell\beta_i\theta_i^2\theta_k^3\theta_\ell^2\theta_t + \beta_k\beta_\ell^2\beta_t\theta_i^1\theta_k^2\theta_\ell^3\theta_t^2 \\
& \quad + \beta_k\beta_\ell\beta_t\beta_i\theta_i^2\theta_k^2\theta_\ell^2\theta_t^2 + \beta_k^2\beta_\ell\beta_i\theta_i^2\theta_k^3\theta_\ell^2\theta_t^1 + \beta_k\beta_\ell^2\beta_t\theta_i\theta_k^2\theta_\ell^3\theta_t^2 + \beta_k^2\beta_\ell^2\theta_i\theta_k^3\theta_\ell^3\theta_t\Big] \\
& \quad + \sum_{ik\ell t i'\ell'} \Big[\beta_k^2\beta_\ell\beta_{\ell'}\theta_i\theta_k^3\theta_\ell^2\theta_t\theta_{i'}\theta_{\ell'}^2 + \beta_k\beta_\ell\beta_{\ell'}\beta_t\theta_i\theta_k^2\theta_\ell^2\theta_t^2\theta_{i'}\theta_{\ell'}^2\Big]
\end{aligned}$$

We have by (E.2) and (E.24) that

$$\begin{aligned}
\sum_{ik\ell t} \Big[&\beta_k^2\beta_\ell^2\theta_i\theta_k^3\theta_\ell^3\theta_t + \beta_k^2\beta_\ell\beta_i\theta_i^2\theta_k^3\theta_\ell^2\theta_t + \beta_k\beta_\ell^2\beta_t\theta_i^1\theta_k^2\theta_\ell^3\theta_t^2 + \beta_k\beta_\ell\beta_t\beta_i\theta_i^2\theta_k^2\theta_\ell^2\theta_t^2 \\
& + \beta_k^2\beta_\ell\beta_i\theta_i^2\theta_k^3\theta_\ell^2\theta_t^1 + \beta_k\beta_\ell^2\beta_t\theta_i\theta_k^2\theta_\ell^3\theta_t^2 + \beta_k^2\beta_\ell^2\theta_i\theta_k^3\theta_\ell^3\theta_t\Big] \\
& \lesssim \|\beta\circ\theta\|_2^4\|\theta\|_1^2 + \|\beta\circ\theta\|_2^2\|\theta\|_2^4\|\theta\|_1 + \|\beta\circ\theta\|_2^4\|\theta\|_2^4 + \|\beta\circ\theta\|_2^4\|\theta\|_1^2
\end{aligned}$$

and

$$\sum_{ik\ell t i'\ell'} \Big[\beta_k^2\beta_\ell\beta_{\ell'}\theta_i\theta_k^3\theta_\ell^2\theta_t\theta_{i'}\theta_{\ell'}^2 + \beta_k\beta_\ell\beta_{\ell'}\beta_t\theta_i\theta_k^2\theta_\ell^2\theta_t^2\theta_{i'}\theta_{\ell'}^2\Big] \lesssim \|\beta\circ\theta\|_2^2\|\theta\|_2^4\|\theta\|_1^3 + \|\beta\circ\theta\|_2^4\|\theta\|_2^4\|\theta\|_1^2$$

Thus

$$\mathrm{Var}(\tilde{Z}_{3c}) \lesssim \frac{\|\theta\|_2^4}{\|\theta\|_1^4}\big(\|\beta\circ\theta\|_2^4\|\theta\|_1^2 + \|\beta\circ\theta\|_2^2\|\theta\|_2^4\|\theta\|_1^3 + \|\beta\circ\theta\|_2^4\|\theta\|_2^4\|\theta\|_1^2\big) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^8}{\|\theta\|_1}$$

To study Z_{3c}^*, in (Jin et al., 2021, Supplement, pg.65) the decomposition

$$Z_{3c}^* = \frac{1}{v}\sum_{i,k,\ell(dist)} \alpha_{ik\ell} W_{ik}^2 W_{\ell i} + \frac{1}{v} \sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\}, t\neq k, (s,t)\neq(k,i)}} \alpha_{ik\ell} W_{is} W_{kt} W_{\ell i} \equiv Z_{3c,1}^* + Z_{3c,2}^*$$

is used, where recall $\alpha_{ik\ell} = \sum_{j\notin\{i,k,\ell\}} \eta_j^2 \widetilde{\Omega}_{k\ell}$. Using a similar argument as before, we have

$$\begin{aligned}
\mathrm{Var}(Z_{3c,1}^*) &\lesssim \frac{\|\theta\|_2^4}{\|\theta\|_1^4}\Big(\sum_{ik\ell}\beta_k^2\beta_\ell^2\theta_i^2\theta_k^3\theta_\ell^3 + \sum_{ik\ell k'}[\beta_k\beta_{k'}\beta_\ell^2 + \beta_k\beta_{k'}\beta_i\beta_\ell]\theta_i^3\theta_k^2\theta_\ell^2\theta_{k'}^2\Big) \\
&\lesssim \frac{\|\theta\|_2^4}{\|\theta\|_1^4}\big(\|\beta\circ\theta\|_2^4\|\theta\|_2^2 + \|\theta\|_2^4\|\beta\circ\theta\|_2^2\|\theta\|_3^3 + \|\beta\circ\theta\|_2^4\|\theta\|_2^4\big) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^{10}}{\|\theta\|_1^4}.
\end{aligned}$$

We omit the argument for $Z_{3c,2}^*$ as it is similar and simply state the bound:

$$\mathrm{Var}(Z_{3c,2}^*) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^6}{\|\theta\|_1^2}.$$

Combining the results for $\tilde{Z}_{3c}$ and Z_{3c}^*, we have

$$\mathrm{Var}(Z_{3c}) \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^8}{\|\theta\|_1} \lesssim \|\beta \circ \theta\|_2^4 \|\theta\|_2^6.$$

Next we study Z_{3d}, which is defined as

$$Z_{3d} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_k \eta_j \tilde{\Omega}_{j\ell})(\eta_i - \tilde{\eta}_i)(\eta_k - \tilde{\eta}_k) W_{\ell i} = \sum_{\substack{i,k,\ell(dist) \\ s \neq i, t \neq k}} \alpha_{ik\ell} W_{is} W_{kt} W_{\ell i}$$

where $\alpha_{ik\ell} = \sum_{j \notin \{i,k,\ell\}} \eta_k \eta_j \tilde{\Omega}_{j\ell}$. We see that $\mathbb{E} Z_{3d} = 0$. To study the variance, we use a similar decomposition to that of Z_{3c}. Write

$$Z_{3d} = \frac{1}{v} \sum_{\substack{i,k,\ell(dist) \\ t \neq k}} \alpha_{ik\ell} W_{i\ell}^2 W_{kt} + \frac{1}{v} \sum_{\substack{i,k,\ell(dist) \\ s \notin \{i,\ell\}, t \neq k}} \alpha_{ik\ell} W_{is} W_{kt} W_{\ell i} \equiv \widetilde{Z}_{3d} + Z_{3d}^*.$$

Mimicking the arguments for $\widetilde{Z}_{3c}$ and Z_{3c}^* we obtain

$$\mathrm{Var}(\tilde{Z}_{3d}) \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^6}{\|\theta\|_1},$$

and

$$\mathrm{Var}(Z_{3d}^*) \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^{10}}{\|\theta\|_1^4}.$$

Hence

$$\mathrm{Var}(Z_{3d}) \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^6}{\|\theta\|_1}.$$

Combining the results for $Z_{3a}, \ldots, Z_{3d}$, we have

$$\mathbb{E} Z_3 = 0, \quad \mathrm{Var}(Z_3) \lesssim \|\beta \circ \theta\|_2^4 \|\theta\|_2^6.$$

We proceed to study Z_4. In (Jin et al., 2021, Supplement,pg.67) the following decomposition is given:

$$\begin{aligned} Z_4 &= 2 \sum_{i,j,k,\ell(dist)} \eta_i (\eta_j - \tilde{\eta}_j) \widetilde{\Omega}_{jk} \eta_k (\eta_\ell - \tilde{\eta}_\ell) W_{\ell i} \\ &+ \sum_{i,j,k,\ell(dist)} \eta_i (\eta_j - \tilde{\eta}_j) \widetilde{\Omega}_{jk} (\eta_k - \tilde{\eta}_k) \eta_\ell W_{\ell i} \\ &+ \sum_{i,j,k,\ell(dist)} (\eta_i - \tilde{\eta}_i) \eta_j \widetilde{\Omega}_{jk} \eta_k (\eta_\ell - \tilde{\eta}_\ell) W_{\ell i} \\ &\equiv Z_{4a} + Z_{4b} + Z_{4c}. \end{aligned} \tag{E.39}$$

There it is shown that $\mathbb{E} Z_{4a} = 0$. To study $\mathrm{Var}(Z_{4a})$, we note that Z_{4a} and Z_{3c} have similar structure. In particular we have the decomposition

$$Z_{4a} = \frac{1}{v} \sum_{\substack{i,k,\ell(dist) \\ t \neq k}} \alpha_{ik\ell} W_{i\ell}^2 W_{kt} + \frac{1}{v} \sum_{\substack{i,k,\ell(dist) \\ s \notin \{i,\ell\}, t \neq k}} \alpha_{ik\ell} W_{is} W_{kt} W_{\ell i} \equiv \widetilde{Z}_{4a} + Z_{4a}^*.$$

where $\alpha_{ik\ell} = \sum_{j \notin \{i,k,\ell\}} \eta_j \eta_\ell \tilde{\Omega}_{k\ell}$. Mimicking the argument for $\tilde{Z}_{3c}$ we have

$$\begin{aligned} \mathrm{Var}(\tilde{Z}_{4a}) \lesssim \frac{\|\beta \circ \theta\|_2^2 \|\theta\|_2^2}{\|\theta\|_1^4} \Big(&\sum_{ik\ell t} \big[\beta_k^2 (\theta_i \theta_k^2 \theta_\ell^2 \theta_t + \theta_i^2 \theta_k^2 \theta_\ell^2 \theta_t) + \beta_k \beta_t \theta_i \theta_k^2 \theta_\ell^2 \theta_t^2 + \beta_k \beta_t \theta_i^2 \theta_k^2 \theta_\ell^2 \theta_t^2 \\ &+ \beta_k^2 \theta_i^2 \theta_k^2 \theta_\ell^2 \theta_t + \beta_k \beta_t \theta_i \theta_k^2 \theta_\ell^2 \theta_t^2 \big] + \sum_{ik\ell t i' \ell'} \big[\beta_k^2 \theta_i \theta_k^2 \theta_\ell^2 \theta_t \theta_{i'} \theta_{\ell'}^2 + \beta_k \beta_t \theta_i \theta_k^2 \theta_\ell^2 \theta_t^2 \theta_{i'} \theta_{\ell'}^2 \big] \Big) \end{aligned}$$

$$\lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^2}{\|\theta\|_1^4}\big(\|\beta\circ\theta\|_2^2\|\theta\|_2^2\|\theta\|_1 + \|\beta\circ\theta\|_2^2\|\theta\|_2^4\|\theta\|_1 + \|\beta\circ\theta\|_2^2\|\theta\|_2^4\|\theta\|_1^3 +$$
$$\|\beta\circ\theta\|_2^2\|\theta\|_2^4\|\theta\|_1^2\big) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^6}{\|\theta\|_1}.$$

For $\tilde{Z}_{4a}^*$ we adapt the decomposition used for $\tilde{Z}_{4c}^*$:

$$Z_{4a}^* = \frac{1}{v}\sum_{i,k,\ell(dist)} \alpha_{ik\ell} W_{ik}^2 W_{\ell i} + \frac{1}{v}\sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\}, t\neq k, (s,t)\neq(k,i)}} \alpha_{ik\ell} W_{is} W_{kt} W_{\ell i} =: Z_{4a,1}^* + Z_{4a,2}^*$$

Mimicking the argument for $Z_{3c,1}^*$ and $Z_{3c,2}^*$, we have

$$\mathrm{Var}(Z_{4a,1}^*) \lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^2}{\|\theta\|_1^4}\big(\sum_{ik\ell}\beta_k^2\theta_i^2\theta_k^2\theta_\ell^2 + \sum_{ik\ell k'}\beta_k\beta_{k'}\theta_i^2\theta_k^2\theta_\ell^2\theta_{k'}^2\big) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^8}{\|\theta\|_1^4},$$

and

$$\mathrm{Var}(Z_{4a,2}^*) \lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^2}{\|\theta\|_1^4}\sum_{ik\ell st}\big[\beta_k^2\theta_i^2\theta_k^2\theta_\ell^2\theta_s\theta_t + \beta_k\beta_t\theta_i^2\theta_k^2\theta_\ell^2\theta_s\theta_t^2 + \beta_k\beta_s\theta_i^2\theta_k^2\theta_\ell^2\theta_s^2\theta_t^2\big]$$
$$\lesssim \frac{\|\theta\|_2^4\|\theta\|_2^6}{\|\theta\|_1^2}.$$

It follows that

$$\mathrm{Var}(Z_{4a}) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^6}{\|\theta\|_1}.$$

Next we study

$$Z_{4b} = \sum_{\substack{i,j,k,\ell\\(dist)}} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell W_{\ell i} = \sum_{\substack{i,j,k,\ell\\(dist)}} \alpha_{ijk\ell}(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k) W_{\ell i}$$
$$= \frac{1}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq j, t\neq k}} \alpha_{ijk\ell} W_{js} W_{kt} W_{\ell i}$$

where $\alpha_{ijk\ell} = \eta_i\eta_\ell\widetilde{\Omega}_{jk}$. Mimicking the study of Z_{3a}, we have the decomposition

$$Z_{4b} = \frac{1}{v}\sum_{i,j,k,\ell(dist)} \alpha_{ijk\ell} W_{jk}^2 W_{\ell i} + \frac{1}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq j, t\neq k, (s,t)\neq(k,j)}} \alpha_{ijk\ell} W_{js} W_{kt} W_{\ell i}$$
$$\equiv \widetilde{Z}_{4b} + Z_{4b}^*.$$

Further we have, using (E.1), (E.2), (E.20), and (E.24), we have

$$\mathrm{Var}(\tilde{Z}_{4b}) \lesssim \frac{1}{\|\theta\|_1^4}\Bigg(\sum_{ijk\ell}\big[[\beta_j^2\beta_k^2 + \beta_j\beta_k\beta_\ell\beta_i]\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3\big] + \sum_{ijk\ell j'k'}\beta_j\beta_k\beta_{j'}\beta_{k'}\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3\theta_{j'}^2\theta_{k'}^2\Bigg)$$
$$\lesssim \frac{1}{\|\theta\|_1^4}\big(\|\beta\circ\theta\|_2^4\|\theta\|_2^4 + \|\beta\circ\theta\|_2^4\|\theta\|_2^8\big) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^8}{\|\theta\|_1^4}.$$

Similarly,

$$\mathrm{Var}(Z_{4b}^*) \lesssim \frac{1}{\|\theta\|_1^4}\Bigg(\sum_{ijk\ell st}\big[\beta_j^2\beta_k^2\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\theta_s\theta_t + \beta_k^2\beta_\ell\beta_j\theta_i^2\theta_j^3\theta_k^3\theta_\ell^3\theta_s^2\theta_t + \beta_j\beta_k^2\beta_\ell\theta_i^2\theta_j^3\theta_k^3\theta_\ell^3\theta_s^2\theta_t\Bigg)$$
$$\lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^4}{\|\theta\|_1^2}.$$

It follows that

$$\mathrm{Var}(Z_{4b}) \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^4}{\|\theta\|_1^2}.$$

We study Z_{4c} using the decomposition

$$\begin{aligned} Z_{4c} &= \frac{1}{v} \sum_{i,\ell(dist)} \beta_{i\ell} W_{\ell i}^3 + \frac{2}{v} \sum_{\substack{i,\ell(dist) \\ s \notin \{i,\ell\}}} \beta_{i\ell} W_{is} W_{\ell i}^2 + \frac{1}{v} \sum_{\substack{i,\ell(dist) \\ s \notin \{i,\ell\}, t \notin \{\ell, i\}}} \beta_{i\ell} W_{is} W_{\ell t} W_{\ell i} \\ &\equiv \widetilde{Z}_{4c} + Z_{4c}^* + Z_{4c}^\dagger. \end{aligned}$$

from (Jin et al., 2021, Supplement, pg.68). Only

$$\tilde{Z}_{4c} = \frac{1}{v} \sum_{i,\ell(dist)} \alpha_{i\ell} W_{\ell i}^3$$

has nonzero mean, where $\alpha_{i\ell} = \sum_{j,k(dist) \notin \{i,\ell\}} \eta_j \eta_k \widetilde{\Omega}_{jk}$. By (E.20)

$$|\alpha_{i\ell}| \lesssim \|\beta \circ \theta\|_2^2 \|\theta\|_2^2.$$

Hence

$$|\mathbb{E}\tilde{Z}_{4c}| \lesssim \frac{1}{\|\theta\|_1^2} \sum_{i\ell} \|\beta \circ \theta\|_2^2 \|\theta\|_2^2 \theta_i \theta_\ell \lesssim \|\beta \circ \theta\|_2^2 \|\theta\|_2^2.$$

Except for when $(i,\ell) = (\ell, i)$, the summands of $\tilde{Z}_{4c}$ are uncorrelated. Thus

$$\mathrm{Var}(\tilde{Z}_{4c}) \lesssim \frac{1}{\|\theta\|_1^4} \sum_{i\ell} \|\beta \circ \theta\|_2^4 \|\theta\|_2^4 \theta_i \theta_\ell \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^4}{\|\theta\|_1^2}.$$

Applying the casework from (Jin et al., 2021, Supplement, pg.68),

$$\begin{aligned} \mathrm{Var}(Z_{4c}^*) &\lesssim \sum_{\substack{i,\ell(dist) \\ s \notin \{i,\ell\}}} \sum_{\substack{i',\ell'(dist) \\ s' \notin \{i',\ell'\}}} |\alpha_{i\ell} \alpha_{i'\ell'}| \mathrm{Cov}(W_{is} W_{\ell i}^2, W_{i's'} W_{\ell' i'}^2)| \\ &\lesssim \frac{1}{\|\theta\|_1^4} \big(\sum_{i\ell s} \|\beta \circ \theta\|_2^4 \|\theta\|_2^4 \theta_i^2 \theta_\ell \theta_s + \sum_{i\ell s \ell'} \|\beta \circ \theta\|_2^4 \|\theta\|_2^4 \|\theta\|_2^4 \theta_i^3 \theta_\ell \theta_s \theta_{\ell'} \big) \\ &\lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^4}{\|\theta\|_1^4} \big(\|\theta\|_2^2 \|\theta\|_1^2 + \|\theta\|_2^2 \|\theta\|_1^3 \big) \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^6}{\|\theta\|_1}. \end{aligned}$$

Next, in (Jin et al., 2021, Supplement, pg.69) it is shown that

$$\mathrm{Var}(Z_{4c}^\dagger) \lesssim \frac{1}{v^2} \sum_{\substack{i,\ell(dist) \\ s \notin \{i,\ell\}, t \notin \{\ell, i\}}} \alpha_{i\ell}^2 \cdot \mathrm{Var}(W_{is} W_{\ell t} W_{\ell i})$$

Thus

$$\mathrm{Var}(Z_{4c}^\dagger) \lesssim \sum_{i\ell s} \|\beta \circ \theta\|_2^4 \|\theta\|_2^4 \theta_i^2 \theta_\ell^2 \theta_s \theta_t \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^8}{\|\theta\|_1^2}.$$

Combining the results for $\tilde{Z}_{4c}, Z_{4c}^*, Z_{4c}^\dagger$, we have

$$|\mathbb{E}Z_{4c}| \lesssim \|\beta \circ \theta\|_2^2 \|\theta\|_2^2, \quad \mathrm{Var}(Z_{4c}) \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^8}{\|\theta\|_1}.$$

Combining the results for Z_{4a}, Z_{4b}, and Z_{4c}, we have

$$|\mathbb{E}Z_4| \lesssim \|\beta \circ \theta\|_2^2 \|\theta\|_2^2, \quad \mathrm{Var}(Z_4) \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^6}{\|\theta\|_1}$$

To study Z_5, we use the decomposition

$$
\begin{aligned}
Z_5 = 2 &\sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\eta_j(\eta_k - \tilde{\eta}_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} + \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)^2\eta_k\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} \\
&+ \sum_{i,j,k,\ell(dist)} (\eta_i - \tilde{\eta}_i)\eta_j^2(\eta_k - \tilde{\eta}_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} \\
&\equiv Z_{5a} + Z_{5b} + Z_{5c}.
\end{aligned} \tag{E.40}
$$

from (Jin et al., 2021, Supplement, pg. 70). We further decompose Z_{5a} as in (Jin et al., 2021, Supplement, pg.70):

$$
Z_{5a} = \frac{2}{v}\sum_{j,k(dist)} \alpha_{jk}W_{jk}^2 + \frac{2}{v}\sum_{\substack{j,k(dist)\\ s\neq j, t\neq k,\\ (s,t)\neq(k,j)}} \alpha_{jk}W_{js}W_{kt} \equiv \widetilde{Z}_{5a} + Z_{5a}^*.
$$

where $\alpha_{jk} = \sum_{i,\ell(dist)\notin\{j,k\}} \eta_i\eta_j\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}$. Note that by (E.20) and (E.24),

$$
|\alpha_{jk}| \lesssim \sum_{i\ell} (\beta_k\theta_k)(\beta_\ell\theta_\ell)^2(\beta_i\theta_i) \lesssim \theta_j(\beta_k\theta_k)\|\beta\circ\theta\|_2^3\|\theta\|_2.
$$

Only $\tilde{Z}_{5a}$ has nonzero mean. By (E.1) and (E.2),

$$
|\mathbb{E}Z_{5a}| = |\mathbb{E}\tilde{Z}_{5a}| \lesssim \frac{1}{\|\theta\|_1^2}\sum_{jk} \theta_j(\beta_k\theta_k)\|\beta\circ\theta\|_2^3\|\theta\|_2 \cdot \theta_j\theta_k \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^4}{\|\theta\|_1^2}.
$$

Now we study the variance of Z_{5a}. In (Jin et al., 2021, Supplement, pg.70) it is shown that

$$
\begin{aligned}
\mathrm{Var}(\widetilde{Z}_{5a}) &\lesssim \frac{1}{v^2}\sum_{j,k(dist)} \alpha_{jk}^2\,\mathrm{Var}(W_{jk}^2) \\
\mathrm{Var}(Z_{5a}^*) &\lesssim \frac{1}{v^2}\sum_{\substack{j,k(dist)\\ s\neq j, t\neq k,\\ (s,t)\neq(k,j)}} \alpha_{jk}^2\,\mathrm{Var}(W_{js}W_{kt}).
\end{aligned}
$$

Thus by (E.2) and (E.24),

$$
\begin{aligned}
\mathrm{Var}(\widetilde{Z}_{5a}) &\lesssim \frac{\|\beta\circ\theta\|_2^6\|\theta\|_2^4}{\|\theta\|_1^4}\big(\sum_{jk}\theta_j^3\beta_k^2\theta_k^3\big) \lesssim \frac{\|\beta\circ\theta\|_2^8\|\theta\|_2^6}{\|\theta\|_1^4} \\
\mathrm{Var}(Z_{5a}^*) &\lesssim \frac{\|\beta\circ\theta\|_2^6\|\theta\|_2^4}{\|\theta\|_1^4}\big(\sum_{jk}\theta_j^2\beta_k^2\theta_k^2\cdot\theta_j\theta_s\theta_k\theta_t\big) \lesssim \frac{\|\beta\circ\theta\|_2^8\|\theta\|_2^6}{\|\theta\|_1^2}.
\end{aligned}
$$

We conclude that

$$
\mathrm{Var}(Z_{5a}) \lesssim \frac{\|\beta\circ\theta\|_2^8\|\theta\|_2^6}{\|\theta\|_1^2}.
$$

Next we study Z_{5b} using the decomposition

$$
Z_{5b} = \frac{1}{v}\sum_{j,s(dist)} \alpha_j W_{js}^2 + \frac{1}{v}\sum_{\substack{j\\ s,t(dist)\notin\{j\}}} \alpha_j W_{js}W_{jt} \equiv \widetilde{Z}_{5b} + Z_{5b}^*.
$$

from (Jin et al., 2021, Supplement, pg.71), where $\alpha_j = \sum_{i,k,\ell(dist)\notin\{j\}} \eta_i\eta_k\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}$. Note that by (E.2) and (E.20),

$$
|\alpha_j| \lesssim \sum_{ik\ell} \theta_i\theta_k(\beta_k\theta_k)(\beta_\ell\theta_\ell)^2(\beta_i\theta_i) \lesssim \|\beta\circ\theta\|_2^4\|\theta\|_2^2.
$$

Only $\widetilde{Z}_{5b}$ above has nonzero mean, and we have

$$|\mathbb{E}Z_{5b}| = |\mathbb{E}Z_{5b}| \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^2}{\|\theta\|_1^2} \sum_{j,s} \theta_j \theta_s \lesssim \|\beta \circ \theta\|_2^4 \|\theta\|_2^2.$$

Similarly for the variances,

$$\mathrm{Var}(\tilde{Z}_{5b}) \lesssim \frac{\|\beta \circ \theta\|_2^8 \|\theta\|_2^4}{\|\theta\|_1^4} \sum_{js} \theta_j \theta_s \lesssim \frac{\|\beta \circ \theta\|_2^8 \|\theta\|_2^4}{\|\theta\|_1^2}$$
$$\mathrm{Var}(Z_{5b}^*) \lesssim \frac{\|\beta \circ \theta\|_2^8 \|\theta\|_2^4}{\|\theta\|_1^4} \sum_{jst} \theta_j^2 \theta_s \theta_t \lesssim \frac{\|\beta \circ \theta\|_2^8 \|\theta\|_2^6}{\|\theta\|_1^2},$$

and it follows that

$$\mathrm{Var}(Z_{5b}) \lesssim \frac{\|\beta \circ \theta\|_2^8 \|\theta\|_2^6}{\|\theta\|_1^2}.$$

Next we study

$$\begin{aligned} Z_{5c} &= \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_j - \tilde{\eta}_j)\eta_i^2(\eta_k - \tilde{\eta}_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell j} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i^2 \widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell j})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k) \\ &= \frac{1}{v} \sum_{\substack{i,j,k,\ell(dist) \\ s \neq j, t \neq k}} (\eta_i^2 \widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell j}) W_{js} W_{kt} = \frac{1}{v} \sum_{\substack{j,k(dist) \\ s \neq j, t \neq k}} \alpha_{jk} W_{js} W_{kt} \end{aligned}$$

where $\alpha_{jk} = \sum_{\substack{i,\ell(dist) \\ i,\ell \notin \{j,k\}}} \eta_i^2 \widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell j}$. Note that by (E.20) and (E.18) ,

$$|\alpha_{jk}| \lesssim \sum_{i\ell} \theta_i^2 (\beta_k \theta_k)(\beta_\ell \theta_\ell)^2 (\beta_j \theta_j) \lesssim (\beta_j \theta_j)(\beta_k \theta_k) \|\theta\|_2^2 \|\beta \circ \theta\|_2^2. \tag{E.41}$$

We further decompose

$$Z_{5c} = \frac{1}{v} \sum_{\substack{j,k \\ (dist)}} \alpha_{jk} W_{jk}^2 + \frac{1}{v} \sum_{\substack{j,k(dist) \\ s,t \notin \{j,k\}}} \alpha_{jk} W_{js} W_{kt} \equiv \tilde{Z}_{5c} + Z_{5c}^*.$$

Only the first term has nonzero mean. It follows that

$$|\mathbb{E}Z_{5c}| = |\mathbb{E}\tilde{Z}_{5c}| \lesssim \frac{\|\theta\|_2^2 \|\beta \circ \theta\|_2^2}{\|\theta\|_1^2} \sum_{j,k,s,t} (\beta_j \theta_j)(\beta_k \theta_k) \cdot \theta_j \theta_k \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^4}{\|\theta\|_1^2}.$$

Note that Z_{5c} and Z_{5a} have the same form, but with a different setting of the coefficient α_{jk}. Mimicking the variance bounds for Z_{5a} we obtain the bound

$$\mathrm{Var}(Z_{5c}) \lesssim \frac{\|\beta \circ \theta\|_2^8 \|\theta\|_2^4}{\|\theta\|_1^2}.$$

Combining the previous bounds we obtain

$$|\mathbb{E}Z_5| \lesssim \|\beta \circ \theta\|_2^4 \|\theta\|_2^2, \quad \mathrm{Var}(Z_5) \lesssim \frac{\|\beta \circ \theta\|_2^8 \|\theta\|_2^6}{\|\theta\|_1^2}.$$

Next we study $Z_6 = Z_{6a} + Z_{6b}$ as defined in (Jin et al., 2021, Supplement, pg.72), where

$$Z_{6a} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i \eta_\ell \widetilde{\Omega}_{j\ell} \widetilde{\Omega}_{ki})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k) = \frac{1}{v} \sum_{\substack{j,k(dist) \\ s \neq j, t \neq k}} \alpha_{jk}^{(6a)} W_{js} W_{kt}$$

$$Z_{6b} = 2 \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i \eta_\ell \widetilde{\Omega}_{jk} \widetilde{\Omega}_{\ell i})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k) = \frac{1}{v} \sum_{\substack{j,k(dist) \\ s \neq j, t \neq k}} \alpha^{(6b)}_{jk} W_{js} W_{kt}$$

and

$$\alpha^{(6a)}_{jk} = \sum_{\substack{i,\ell(dist) \\ i,\ell \notin \{j,k\}}} \eta_i \eta_\ell \tilde{\Omega}_{jk} \tilde{\Omega}_{\ell i}$$
$$\alpha^{(6b)}_{jk} = \sum_{\substack{i,\ell(dist) \\ i,\ell \notin \{j,k\}}} \eta_i \eta_\ell \tilde{\Omega}_{j\ell} \tilde{\Omega}_{ki}.$$

Thus Z_{6a} and Z_{6b} take the same form as Z_{5c}, but with a different setting of α_{jk}. Note that by (E.24) and similar arguments from before,

$$\max(|\alpha^{(6a)}_{jk}|, |\alpha^{(6b)}_{jk}|) \lesssim (\beta_j \theta_j)(\beta_k \theta_k) \|\theta\|_2^2 \|\beta \circ \theta\|_2^2,$$

which is the same as the upper bound on $|\alpha_{jk}|$ associated to Z_{5c} given in (E.41). It follows that

$$|\mathbb{E} Z_6| \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^4}{\|\theta\|_1^2}, \quad \mathrm{Var}(Z_6) \lesssim \frac{\|\beta \circ \theta\|_2^8 \|\theta\|_2^4}{\|\theta\|_1^2}.$$

We have proved all claims in Lemma E.9. □

E.4.5 PROOF OF LEMMA E.10

The terms T_1 and F do not depend on $\tilde{\Omega}$, and thus the claimed bounds transfer directly from (Jin et al., 2021, Lemma G.9). Thus we focus on T_2. We use the decomposition $T_2 = 2(T_{2a} + T_{2b} + T_{2c} + T_{2d})$ from (Jin et al., 2021, Supplement, pg.73) where

$$T_{2a} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2} \eta_{i_3} \eta_{i_4} \big[(\eta_{i_1} - \tilde{\eta}_{i_1})(\eta_{i_2} - \tilde{\eta}_{i_2})(\eta_{i_3} - \tilde{\eta}_{i_3})\big] \cdot \widetilde{\Omega}_{i_4 i_1},$$
$$T_{2b} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2} \eta_{i_3}^2 \big[(\eta_{i_1} - \tilde{\eta}_{i_1})(\eta_{i_2} - \tilde{\eta}_{i_2})(\eta_{i_4} - \tilde{\eta}_{i_4})\big] \cdot \widetilde{\Omega}_{i_4 i_1},$$
$$T_{2c} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1} \eta_{i_3} \eta_{i_4} \big[(\eta_{i_2} - \tilde{\eta}_{i_2})^2(\eta_{i_3} - \tilde{\eta}_{i_3})\big] \cdot \widetilde{\Omega}_{i_4 i_1},$$
$$T_{2d} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1} \eta_{i_3}^2 \big[(\eta_{i_2} - \tilde{\eta}_{i_2})^2(\eta_{i_4} - \tilde{\eta}_{i_4})\big] \cdot \widetilde{\Omega}_{i_4 i_1}.$$

We study each term separately.

For T_{2a}, in (Jin et al., 2021, Supplement, pg.89), we have the decomposition $T_{2a} = X_{a1} + X_{a2} + X_{a3} + X_b$ where

$$X_{a1} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_3 \neq i_3} \eta_{i_2} \eta_{i_3} \eta_{i_4} W_{i_1 i_2}^2 W_{i_3 j_3} \widetilde{\Omega}_{i_1 i_4},$$
$$X_{a2} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2 \neq i_2} \eta_{i_2} \eta_{i_3} \eta_{i_4} W_{i_1 i_3}^2 W_{i_2 j_2} \widetilde{\Omega}_{i_1 i_4},$$
$$X_{a3} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_1 \neq i_1} \eta_{i_2} \eta_{i_3} \eta_{i_4} W_{i_2 i_3}^2 W_{i_1 j_1} \widetilde{\Omega}_{i_1 i_4},$$
$$X_b = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_1,j_2,j_3 \\ j_k \neq i_\ell, k,\ell = 1,2,3}} \eta_{i_2} \eta_{i_3} \eta_{i_4} W_{i_1 j_1} W_{i_2 j_2} W_{i_3 j_3} \widetilde{\Omega}_{i_1 i_4}.$$

There it is shown that $\mathbb{E} T_{2a} = 0$. Further it is argued that

$$\mathrm{Var}(X_{a1}) = \mathbb{E} X_{a1}^2$$

$$= \frac{1}{v^3} \sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ i'_1,i'_2,i'_3,i'_4(dist)}} \sum_{\substack{j_3,j'_3\\ j_3\neq i_3, j'_3\neq i'_3}} \eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i'_2}\eta_{i'_3}\eta_{i'_4}\mathbb{E}[W^2_{i_1i_2}W_{i_3j_3}W^2_{i'_1i'_2}W_{i'_3j'_3}]\widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i'_1i'_4} \tag{E.42}$$

$$\equiv V_A + V_B + V_C,$$

where the terms V_A, V_B, V_C correspond to the contributions from cases A, B, C, respectively, described in (Jin et al., 2021, Supplement, pg.89). Concretely, the nonzero terms of (E.42) fall into three cases:

Case A. $\{i_1, i_2\} = \{i'_3, j'_3\}$ and $\{i_3, j_3\} = \{i'_1, i'_2\}$

Case B. $\{i_3, j_3\} = \{i'_3, j'_3\}$ and $\{i_1, i_2\} = \{i'_1, i'_2\}$

Case C. $\{i_3, j_3\} = \{i'_3, j'_3\}$ and $\{i_1, i_2\} \neq \{i'_1, i'_2\}$.

Here V_A, V_B, and V_C are defined to be the contributions from each case.

Applying (E.2), (E.22), and (E.20),

$$\begin{aligned}|\eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i'_2}\eta_{i'_3}\eta_{i'_4}\widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i'_1i'_4}| &\lesssim \theta_{i_2}\theta_{i_3}\theta_{i_4}\theta_{i'_2}\theta_{i'_3}\theta_{i'_4}(\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4})(\beta_{i'_1}\theta_{i'_1})(\beta_{i'_4}\theta_{i'_4})\\ &\lesssim \theta_{i_2}\theta_{i_3}\theta_{i_4}\theta_{i'_2}\theta_{i'_3}\theta_{i'_4}(\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4})\theta_{i'_1}(\beta_{i'_4}\theta_{i'_4}).\end{aligned} \tag{E.43}$$

Note that using the last inequality reduces the required casework while still yielding a good enough bound. Mimicking the casework in Case A of (Jin et al., 2021, Supplement, pg.90) and applying (E.24), we have

$$\begin{aligned}V_A &\lesssim \frac{1}{\|\theta\|_1^6} \sum_{\substack{i_1,i_2,i_3\\ i_4,i'_4,j_3}} \sum_{\substack{b_1,b_2\\ (b_1+b_2=1)}} \beta_{i_1}\beta_{i_4}\beta_{i'_4}\theta_{i_1}^{2+b_1}\theta_{i_2}^{2+b_2}\theta_{i_3}^3\theta_{j_3}^2\theta_{i_4}^2\theta_{i'_4}^2\\ &\lesssim \frac{1}{\|\theta\|_1^6}\big(\|\beta\circ\theta\|_2^3\|\theta\|_2^3\|\theta\|_2^4\|\theta\|_3^3 + \|\beta\circ\theta\|_2^3\|\theta\|_2^3\|\theta\|_2^2\|\theta\|_3^6\big) \lesssim \frac{\|\beta\circ\theta\|_2^3\|\theta\|_2^9}{\|\theta\|_1^6}.\end{aligned}$$

Similarly, applying (E.43) along with (E.22), (E.20), and (E.24) yields

$$V_B \lesssim \frac{1}{\|\theta\|_1^5} \sum_{\substack{i_1,i_2,i_3\\ i_4,i'_4,j_3}} \sum_{\substack{c_1,c_2\\ (c_1+c_2=1)}} \beta_{i_1}\beta_{i_4}\beta_{i'_4}\theta_{i_1}^3\theta_{i_2}^3\theta_{i_3}^{2+c_1}\theta_{j_3}^{1+c_2}\theta_{i_2}^2\theta_{i'_4}^2 \lesssim \frac{\|\beta\circ\theta\|_2^3\|\theta\|_2^7}{\|\theta\|_1^5}.$$

and

$$V_C \lesssim \sum_{\substack{i_1,i_2,i_3,i_4\\ i'_1,i'_2,i'_4,j_3}} \sum_{\substack{c_1,c_2\\ (c_1+c_2=1)}} \beta_{i_1}\beta_{i_4}\beta_{i'_1}\beta_{i'_4}\theta_{i_1}^2\theta_{i_2}^2\theta_{i_3}^{2+c_1}\theta_{j_3}^{1+c_2}\theta_{i_4}^2\theta_{i'_1}^2\theta_{i'_2}^2\theta_{i'_4}^2 \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^{10}}{\|\theta\|_1^5}.$$

Thus

$$\mathrm{Var}(X_{a1}) \lesssim \|\beta\circ\theta\|_2^4.$$

The arguments for X_{a2} and X_{a3} are similar, and the corresponding V_A, V_B, V_C satisfy the same inequalities above. We simply state the bounds:

$$\mathbb{E}X_{a_2} = \mathbb{E}X_{a3} = 0, \quad \mathrm{Var}(X_{a2}) \lesssim \|\beta\circ\theta\|_2^4, \quad \mathrm{Var}(X_{a3}) \lesssim \|\beta\circ\theta\|_2^4.$$

Next we consider X_b as defined in (Jin et al., 2021, Supplement, pg.89). We have $\mathbb{E}X_b = 0$ and focus on the variance. In (Jin et al., 2021, Supplement, pg.91) it is shown that

$$\begin{aligned}\mathrm{Var}(X_b) &= \mathbb{E}[X_b^2]\\ &= v^{-3} \sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ i'_1,i'_2,i'_3,i'_4(dist)}} \sum_{\substack{j_3,j'_3\\ j_3\neq i_3, j'_3\neq i'_3}} \eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i'_2}\eta_{i'_3}\eta_{i'_4}\mathbb{E}[W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}W_{i'_1j'_1}W_{i'_2j'_2}W_{i'_3j'_3}]\widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i'_1i'_4},\end{aligned}$$

Note that

$$\mathbb{E}[W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}W_{i'_1j'_1}W_{i'_2j'_2}W_{i'_3j'_3}] \neq 0$$

if and only if the two sets of random variables $\{W_{i_1j_1}, W_{i_2j_2}, W_{i_3j_3}\}$ and $\{W_{i'_1j'_1}, W_{i'_2j'_2}, W_{i'_3j'_3}\}$ are identical. Applying (E.22) and (E.20),

$$\begin{aligned}|\eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i'_2}\eta_{i'_3}\eta_{i'_4}\widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i'_1i'_4}| &\lesssim \theta_{i_2}\theta_{i_3}\theta_{i_4}\theta_{i'_2}\theta_{i'_3}\theta_{i'_4}(\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4})\theta_{i'_1}(\beta_{i'_4}\theta_{i'_4})\\ &\lesssim \beta_{i_1}\beta_{i_4}\beta_{i'_4}\theta_{i_1}^{1+a_1}\theta_{j_1}^{a_2}\theta_{i_2}^{1+a_3}\theta_{j_2}^{a_4}\theta_{i_3}^{1+a_5}\theta_{j_3}^{a_6}\theta_{i_4}^2\theta_{i'_4}^2\end{aligned}$$

if $\mathbb{E}[W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}W_{i'_1j'_1}W_{i'_2j'_2}W_{i'_3j'_3}] \neq 0$, where $a_i \in \{0,1\}$ and $\sum_{i=1}^6 a_i = 3$. Thus by (E.1), (E.2), and (E.24),

$$\begin{aligned}\mathrm{Var}(X_b) &\lesssim \max_a \frac{1}{\|\theta\|_1^6}\sum_{\substack{i_1,i_2,i_3,i_4\\ i'_4,j_1,j_2,j_3}} \beta_{i_1}\beta_{i_4}\beta_{i'_4}\theta_{i_1}^{2+a_1}\theta_{j_1}^{1+a_2}\theta_{i_2}^{2+a_3}\theta_{j_2}^{1+a_4}\theta_{i_3}^{2+a_5}\theta_{j_3}^{1+a_6}\theta_{i_4}^2\theta_{i'_4}^2\\ &\lesssim \frac{1}{\|\theta\|_1^6}\sum_{\substack{i_1,i_2,i_3,i_4\\ i'_4,j_1,j_2,j_3}} \beta_{i_1}\beta_{i_4}\beta_{i'_4}\theta_{i_1}^2\theta_{j_1}^1\theta_{i_2}^2\theta_{j_2}^1\theta_{i_3}^2\theta_{j_3}^1\theta_{i_4}^2\theta_{i'_4}^2\\ &\lesssim \frac{\|\beta\circ\theta\|_2^3\|\theta\|_2^3\|\theta\|_2^4\|\theta\|_1^3}{\|\theta\|_1^6} \lesssim \frac{\|\beta\circ\theta\|_2^3\|\theta\|_2^7}{\|\theta\|_1^3} \lesssim \|\beta\circ\theta\|_2^3\|\theta\|_2.\end{aligned}$$

Combining the results for X_{a1}, X_{a2}, X_{a3} and X_b, we conclude that

$$\mathbb{E}T_{2a} = 0, \qquad \mathrm{Var}(T_{2a}) \lesssim \|\beta\circ\theta\|_2^4\|\theta\|_2.$$

The argument for T_{2b} is similar to the one for T_{2a}, so we simply state the results:

$$\mathbb{E}T_{2b} = 0, \qquad \mathrm{Var}(T_{2b}) \lesssim \|\beta\circ\theta\|_2^4\|\theta\|_2.$$

Next we study T_{2c}, providing full details for completeness. Using the definition of T_{2c} in (Jin et al., 2021, Supplement, pg.92), we have the following decomposition by careful casework.

$$\begin{aligned}Y_a &= -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2i_3}^3\widetilde{\Omega}_{i_1i_4},\\ Y_{b1} &= -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{\substack{(i_2,j_2)\neq(j_3,i_3)\\ j_2\neq i_2, j_3\neq i_3}} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2j_2}^2W_{i_3j_3}\widetilde{\Omega}_{i_1i_4},\\ Y_{b2} &= -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{\ell_2\notin\{i_3,i_2\}} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2i_3}^2W_{i_2\ell_2}\widetilde{\Omega}_{i_1i_4},\\ Y_{b3} &= -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{j_2\notin\{i_3,i_2\}} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2i_3}^2W_{i_2j_2}\widetilde{\Omega}_{i_1i_4},\\ Y_c &= -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{\substack{j_2,\ell_2,j_3\\ j_2\neq i_2,\ell_2\neq i_2,j_3\neq i_3\\ j_2\neq\ell_2,(i_2,j_2)\neq(j_3,i_3),(i_2,\ell_2)\neq(j_3,i_3)}} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}\widetilde{\Omega}_{i_1i_4}.\end{aligned}$$

Note that, by the change of variables $\ell_2 \to j_2$, it holds that $Y_{b2} = Y_{b3}$.

The only term with nonzero mean is Y_a. We have by (E.18), (E.20), (E.22), and (E.24) that

$$\begin{aligned}|\mathbb{E}Y_a| &\lesssim \frac{1}{\|\theta\|_1^3}\sum_{i_1,i_2,i_3,i_4}\theta_{i_1}\theta_{i_3}\theta_{i_4}(\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4})\cdot|\mathbb{E}W_{i_2i_3}^3| \lesssim \frac{1}{\|\theta\|_1^3}\sum_{i_1,i_2,i_3,i_4}\beta_{i_1}\beta_{i_4}\theta_{i_1}^2\theta_{i_2}\theta_{i_3}^2\theta_{i_4}^2\\ &\lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^4}{\|\theta\|_1^2}.\end{aligned}$$

For the variance, by independence of $\{W_{ij}\}_{i>j}$, (E.2), (E.20), and (E.24), we have

$$\mathrm{Var}(Y_a) \lesssim \frac{1}{\|\theta\|_1^6}\sum_{i_2,i_3}\big(\sum_{i_1,i_4}\theta_{i_1}\theta_{i_3}\theta_{i_4}(\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4})\big)^2\theta_{i_2}\theta_{i_3} \lesssim \frac{1}{\|\theta\|_1^6}\sum_{i_2,i_3}\|\beta\circ\theta\|_2^4\|\theta\|_2^4\theta_{i_2}\theta_{i_3}^2$$

$$\lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^6}{\|\theta\|_1^5}.$$

For Y_{b1}, Y_{b2}, Y_{b3} we make note of the identity

$$W_{ij}^2 = (1 - 2\Omega_{ij})W_{ij} + \Omega_{ij}(1 - \Omega_{ij}) \equiv A_{ij}W_{ij} + B_{ij}. \tag{E.44}$$

Write

$$Y_{b1} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{(i_2,j_2)\neq(j_3,i_3) \\ j_2\neq i_2, j_3\neq i_3}} \eta_{i_1}\eta_{i_3}\eta_{i_4} A_{i_2j_2} W_{i_2j_2} W_{i_3j_3} \widetilde{\Omega}_{i_1i_4}$$
$$- \frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{(i_2,j_2)\neq(j_3,i_3) \\ j_2\neq i_2, j_3\neq i_3}} \eta_{i_1}\eta_{i_3}\eta_{i_4} B_{i_2j_2} W_{i_3j_3} \widetilde{\Omega}_{i_1i_4} \equiv Y_{b1,A} + Y_{b1,B}.$$

By similar arguments from before, and noting that $|A_{i_2,j_2}| \lesssim 1$,

$$\mathrm{Var}(Y_{b1,A}) \lesssim \frac{1}{\|\theta\|_1^6} \sum_{\substack{(i_2,j_2)\neq(j_3,i_3) \\ j_2\neq i_2, j_3\neq i_3}} \Big(\sum_{i_1,i_4} \eta_{i_1}\eta_{i_3}\eta_{i_4} (\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4}) \Big)^2 |\mathbb{E} W_{i_2j_2} W_{i_3j_3}|$$
$$\lesssim \frac{1}{\|\theta\|_1^6} \sum_{i_2,j_2,i_3,j_3} \Big(\sum_{i_1,i_4} \eta_{i_1}\eta_{i_3}\eta_{i_4} (\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4}) \Big)^2 \cdot \theta_{i_2}\theta_{j_2}\theta_{i_3}\theta_{j_3}$$
$$\lesssim \frac{1}{\|\theta\|_1^6} \sum_{i_2,j_2,i_3,j_3} \|\beta \circ \theta\|_2^4 \|\theta\|_2^4 \theta_{i_2}\theta_{j_2}\theta_{i_3}^3\theta_{j_3} \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^6}{\|\theta\|_1^3}.$$

Similarly, using $|B_{ij}| \lesssim \Omega_{ij} \lesssim \theta_i\theta_j$,

$$\mathrm{Var}(Y_{b1,B}) \lesssim \frac{1}{\|\theta\|_1^6} \sum_{i_3,j_3(dist)} \Big(\sum_{i_1,i_2,i_4,j_2} \eta_{i_1}\eta_{i_3}\eta_{i_4} \theta_{i_2}\theta_{j_2} (\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4}) \Big)^2 \cdot |\mathbb{E} W_{i_3,j_3}|$$
$$\lesssim \frac{1}{\|\theta\|_1^6} \sum_{i_3,j_3} \|\beta \circ \theta\|_2^4 \|\theta\|_2^4 \|\theta\|_1^2 \theta_{i_3}^3\theta_{j_3} \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^6}{\|\theta\|_1^3}.$$

It follows that

$$\mathrm{Var}(Y_{b1}) \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^6}{\|\theta\|_1^3}$$

To control $\mathrm{Var}(Y_{b2})$, again we invoke the identity (E.44) to write

$$Y_{b2} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\ell_2 \notin \{i_3,i_2\}} \eta_{i_1}\eta_{i_3}\eta_{i_4} A_{i_2i_3} W_{i_2i_3} W_{i_2\ell_2} \widetilde{\Omega}_{i_1i_4}$$
$$- \frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\ell_2 \notin \{i_3,i_2\}} \eta_{i_1}\eta_{i_3}\eta_{i_4} B_{i_2i_3} W_{i_2\ell_2} \widetilde{\Omega}_{i_1i_4} \equiv Y_{b2,A} + Y_{b2,B}.$$

Using similar arguments from before, we have

$$\mathrm{Var}(Y_{b2,A}) \lesssim \frac{1}{\|\theta\|_1^6} \sum_{i_2i_3\ell_2} \Big(\sum_{i_1i_4} \theta_{i_1}\theta_{i_3}\theta_{i_4} (\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4}) \Big)^2 \theta_{i_2}^2\theta_{i_3}\theta_{\ell_2}$$
$$\lesssim \frac{1}{\|\theta\|_1^6} \sum_{i_2i_3\ell_2} \|\beta \circ \theta\|_2^4 \|\theta\|_2^4 \theta_{i_2}^2\theta_{i_3}^3\theta_{\ell_2} \lesssim \frac{\|\beta \circ \theta\|_2^4 \|\theta\|_2^8}{\|\theta\|_1^5}.$$

Furthermore,

$$\mathrm{Var}(Y_{b2,B}) \lesssim \frac{1}{\|\theta\|_1^6} \sum_{i_2,\ell_2} \Big(\sum_{i_1,i_3,i_4} \theta_{i_1}\theta_{i_3}\theta_{i_4} (\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4}) \theta_{i_2}\theta_{i_3} \Big)^2 \theta_{i_2}\theta_{\ell_2}$$

$$\lesssim \frac{1}{\|\theta\|_1^6}\sum_{i_2,\ell_2}\|\beta\circ\theta\|_2^4\|\theta\|_2^8\theta_{i_2}^3\theta_{\ell_2} \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^{10}}{\|\theta\|_1^5}.$$

Since $Y_{b2} = Y_{b3}$, we have

$$\mathrm{Var}(Y_{b2}) = \mathrm{Var}(Y_{b3}) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^{10}}{\|\theta\|_1^5}.$$

Next we study the variance of Y_{2c}. For notational brevity, let

$$\mathcal{R}_{i_1,i_2,i_3} = \left\{(j_2,\ell_2,j_3)\middle| j_2\neq i_2, \ell_2\neq i_2, j_3\neq i_3 j_2\neq \ell_2, (i_2,j_2)\neq(j_3,i_3), (i_2,\ell_2)\neq(j_3,i_3)\right\}.$$

We have

$$\mathrm{Var}(Y_c) = \frac{1}{v^3}\sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ i_1',i_2',i_3',i_4'(dist)}}\sum_{\substack{(j_2,\ell_2,j_3)\in\mathcal{R}_{i_1,i_2,i_3}\\ (j_2',\ell_2',j_3')\in\mathcal{R}_{i_1',i_2',i_3'}}}\eta_{i_1}\eta_{i_3}\eta_{i_4}\widetilde{\Omega}_{i_1i_4}\eta_{i_1'}\eta_{i_3'}\eta_{i_4'}\widetilde{\Omega}_{i_1'i_4'}\mathbb{E}[W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}W_{i_2'j_2'}W_{i_2'\ell_2'}W_{i_3'j_3'}] \tag{E.45}$$

Note that $W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}$ and $W_{i_2'j_2'}W_{i_2'\ell_2'}W_{i_3'j_3'}$ above are uncorrelated unless

$$\left\{\{i_2,j_2\},\{i_2,\ell_2\},\{i_3,j_3\}\right\} = \left\{\{i_2',j_2'\},\{i_2',\ell_2'\},\{i_3',j_3'\}\right\}.$$

In particular, $i_3'\in\{i_2,j_2,\ell_2,i_3,j_3\}$ when the above holds. Hence for some choice of $a_i\in\{0,1\}$ with $\sum_{i=1}^5 a_i = 1$,

$$\begin{aligned}
\mathrm{Var}(Y_c) &\lesssim \frac{1}{v^3}\sum_{\substack{i_1,i_2,i_3,i_4\\ i_1',i_4',j_2,\ell_2,j_3}}\theta_{i_2}^{a_1}\theta_{j_2}^{a_2}\theta_{\ell_2}^{a_3}\theta_{i_3}^{a_4}\theta_{j_3}^{a_5}\cdot\theta_{i_1}\theta_{i_3}\theta_{i_4}(\beta_{i_1}\theta_{i_1})(\beta_{i_4}\theta_{i_4})\theta_{i_1'}\theta_{i_4'}(\beta_{i_1'}\theta_{i_1'})(\beta_{i_4'}\theta_{i_4'})\cdot\theta_{i_2}^2\theta_{j_2}\theta_{\ell_2}\theta_{i_3}\theta_{j_3}\\
&\lesssim \frac{1}{v^3}\sum_{\substack{i_1,i_2,i_3,i_4\\ i_1',i_4',j_2,\ell_2,j_3}}\beta_{i_1}\beta_{i_1'}\beta_{i_4}\beta_{i_4'}\theta_{i_1}^2\theta_{i_2}^{2+a_1}\theta_{i_3}^{2+a_4}\theta_{i_4}^2\theta_{i_1'}^2\theta_{i_4'}^2\theta_{j_2}^{1+a_2}\theta_{\ell_2}^{1+a_3}\theta_{j_3}^{1+a_5}\\
&\lesssim \frac{1}{v^3}\sum_{\substack{i_1,i_2,i_3,i_4\\ i_1',i_4',j_2,\ell_2,j_3}}\beta_{i_1}\beta_{i_1'}\beta_{i_4}\beta_{i_4'}\theta_{i_1}^2\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}^2\theta_{i_1'}^2\theta_{i_4'}^2\theta_{j_2}^1\theta_{\ell_2}^1\theta_{j_3}^1 \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^8}{\|\theta\|_1^3},
\end{aligned}$$

where in the last line we apply (E.2) followed by (E.24). Combining our results above we have

$$|\mathbb{E}T_{2c}| \lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^4}{\|\theta\|_1^2}, \qquad \mathrm{Var}(T_{2c}) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^6}{\|\theta\|_1^2}.$$

The argument for T_{2d} is omitted since it is similar to the one for T_{2c} (note that the two terms have similar structure). The results are stated below.

$$|\mathbb{E}T_{2d}| \lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^4}{\|\theta\|_1^2}, \qquad \mathrm{Var}(T_{2d}) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^8}{\|\theta\|_1^3}.$$

Combining the results for $T_{2a},\ldots,T_{2d}$ yields

$$|\mathbb{E}T_2| \lesssim \frac{\|\beta\circ\theta\|_2^2\|\theta\|_2^4}{\|\theta\|_1^2}, \qquad \mathrm{Var}(T_2) \lesssim \frac{\|\beta\circ\theta\|_2^4\|\theta\|_2^8}{\|\theta\|_1^2},$$

as desired. $\square$

E.4.6 Proof of Lemma E.11

As before, we only need to analyze the alternative hypothesis. In (Jin et al., 2021, Supplement,pg.103) it is shown that $\tilde{Q}^* - Q^*$ is a sum of $O(1)$ terms of the form

$$Y = \left(\frac{v}{V}\right)^{N_{\tilde{r}}} \sum_{i,j,k,\ell(dist)} a_{ij} b_{jk} c_{k\ell} d_{\ell i}, \tag{E.46}$$

where $a, b, c, d \in \{\widetilde{\Omega}, W, \delta, -(\tilde{\eta} - \eta)(\tilde{\eta} - \eta)^\mathsf{T}\}$, and $N_{\tilde{r}}$ denotes the number of a, b, c, d that are equal to $-(\tilde{\eta} - \eta)(\tilde{\eta} - \eta)^\mathsf{T}$.

Similarly, let N_W denote the number of a, b, c, d that are equal to W, and $N_{\tilde{\Omega}}$ and N_δ are similarly defined. Write

$$Y = \left(\frac{v}{V}\right)^m X, \qquad \text{where} \quad X = \sum_{i,j,k,\ell(dist)} a_{ij} b_{jk} c_{k\ell} d_{\ell i}. \tag{E.47}$$

Note that for this proof, we do not need the explicit decomposition: we only will use the fact that $\tilde{Q}^* - Q^*$ is a sum of $O(1)$ terms. At times, we refer to these terms of the form Y composing $\tilde{Q}^* - Q^*$ as *post-expansion sums*.

In Jin et al. (2021) it is shown that $4 \geq N_{\tilde{r}} \geq 1$ for every post-expansion sum (note that the upper bound of 4 is trivial). It turns out that this is the *only* constraint on the post-expansion sums; so we need to analyze every single possible combination of nonnegative integers $(N_{\tilde{\Omega}}, N_W, N_\delta, N_{\tilde{r}})$ where their sum is 4 and $N_{\tilde{r}} \geq 1$ and then arrange $a, b, c, d \in \{\tilde{\Omega}, W, \delta, -(\tilde{\eta} - \eta)(\tilde{\eta} - \eta)^\mathsf{T}\}$ in all possible ways according to (E.46). This leads to a total of 34 possibilities, all of which are shown in Table 1 reproduced from Jin et al. (2021).

In (Jin et al., 2021, Supplement,pg.103) it is shown that

$$\begin{aligned} |\mathbb{E}[Y - X]| &\leq o(\|\theta\|_2^{-2})\sqrt{\mathbb{E}[X^2]} + o(1), \text{ and} \\ \mathrm{Var}(Y) &\leq 2\mathrm{Var}(X) + o(\|\theta\|_2^{-4})\mathbb{E}[X^2] + o(1). \end{aligned} \tag{E.48}$$

The proof of (E.48) in Jin et al. (2021) only requires the heterogeneity assumptions (E.2)–(E.4) and the following two conditions. First, we must have the tail inequality

$$\mathbb{P}(|V - v| > t) \leq \begin{cases} 2\exp\left(-\frac{C_1}{\|\theta\|_1^2} t^2\right), & \text{when } x_n\|\theta\|_1 \leq t \leq \|\theta\|_1^2, \\ 2\exp\left(-C_2 t\right), & \text{when } t > \|\theta\|_1^2. \end{cases} \tag{E.49}$$

Second, it must hold that $|Y - X|$ is dominated by a polynomial in V. See (Jin et al., 2021, Lemma G.10 and G.11) for further details. Both conditions are satisfied in our setting, so indeed (E.48) applies.

Let N_W and N_δ denote the number of a, b, c, d that are equal to W and δ, respectively. As in Jin et al. (2021), we define

$$N_W^* = N_W + N_\delta + 2N_{\tilde{r}} \tag{E.50}$$

and divide our analysis into parts based on this parameter.

Analysis of terms with $N_W^* \leq 4$ For convenience, we reproduce Table G.5 from Jin et al. (2021) in Table 2. The left column of Table 2 lists all of the terms with $N_W^* \leq 4$, where note that factors of $(\frac{v}{V})^{N_{\tilde{r}}}$ are removed. In the right column terms are listed that have similar structure to those on the left. Precisely, a term in the left column has the form

$$X = \sum_{i_1,\ldots,i_m \in \mathcal{R}} c_{i_1,\ldots,i_m} G_{i_1,\ldots,i_m},$$

and its adjacent term on the right column has the form

$$X^* = \sum_{i_1,\ldots,i_m \in \mathcal{R}} c^*_{i_1,\ldots,i_m} G_{i_1,\ldots,i_m},$$

Table 1: *Note: This table and caption reproduced from Table G.4 of Jin et al. (2021).*The 34 types of the 175 post-expansion sums for $(\widetilde{Q}_n^* - Q_n^*)$.

Notation	#	$N_{\tilde{r}}$	$(N_\delta, N_{\widetilde{\Omega}}, N_W)$	Examples	N_W^*
R_1	4	1	(0, 0, 3)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} W_{jk} W_{k\ell} W_{\ell i}$	5
R_2	8	1	(0, 1, 2)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \widetilde{\Omega}_{jk} W_{k\ell} W_{\ell i}$	4
R_3	4			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} W_{jk} \widetilde{\Omega}_{k\ell} W_{\ell i}$	4
R_4	8	1	(0, 2, 1)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \widetilde{\Omega}_{jk} \widetilde{\Omega}_{k\ell} W_{\ell i}$	3
R_5	4			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \widetilde{\Omega}_{jk} W_{k\ell} \widetilde{\Omega}_{\ell i}$	3
R_6	4	1	(0, 3, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \widetilde{\Omega}_{jk} \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i}$	2
R_7	8	1	(1, 0, 2)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} W_{k\ell} W_{\ell i}$	5
R_8	4			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} W_{jk} \delta_{k\ell} W_{\ell i}$	5
R_9	8	1	(1, 1, 1)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} \widetilde{\Omega}_{k\ell} W_{\ell i}$	4
R_{10}	8			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \widetilde{\Omega}_{jk} W_{k\ell} \delta_{\ell i}$	4
R_{11}	8			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} W_{jk} \delta_{k\ell} \widetilde{\Omega}_{\ell i}$	4
R_{12}	8	1	(1, 2, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i}$	3
R_{13}	4			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \widetilde{\Omega}_{jk} \delta_{k\ell} \widetilde{\Omega}_{\ell i}$	3
R_{14}	8	1	(2, 0, 1)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} \delta_{k\ell} W_{\ell i}$	5
R_{15}	4			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} W_{k\ell} \delta_{\ell i}$	5
R_{16}	8	1	(2, 1, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} \delta_{k\ell} \widetilde{\Omega}_{\ell i}$	4
R_{17}	4			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} \widetilde{\Omega}_{k\ell} \delta_{\ell i}$	4
R_{18}	4	1	(3, 0, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} \delta_{k\ell} \delta_{\ell i}$	5
R_{19}	4	2	(0, 0, 2)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} W_{k\ell} W_{\ell i}$	6
R_{20}	2			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} W_{jk} \tilde{r}_{k\ell} W_{\ell i}$	6
R_{21}	4	2	(0, 2, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i}$	4
R_{22}	2			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \widetilde{\Omega}_{jk} \tilde{r}_{k\ell} \widetilde{\Omega}_{\ell i}$	4
R_{23}	4	2	(2, 0, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} \delta_{k\ell} \delta_{\ell i}$	6
R_{24}	2			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} \tilde{r}_{k\ell} \delta_{\ell i}$	6
R_{25}	8	2	(0, 1, 1)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} \widetilde{\Omega}_{k\ell} W_{\ell i}$	5
R_{26}	4			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \widetilde{\Omega}_{jk} \tilde{r}_{k\ell} W_{\ell i}$	5
R_{27}	8	2	(1, 1, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} \delta_{k\ell} \widetilde{\Omega}_{\ell i}$	5
R_{28}	4			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} \tilde{r}_{k\ell} \widetilde{\Omega}_{\ell i}$	5
R_{29}	8	2	(1, 0, 1)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} \delta_{k\ell} W_{\ell i}$	6
R_{30}	4			$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \delta_{jk} \tilde{r}_{k\ell} W_{\ell i}$	6
R_{31}	4	3	(0, 0, 1)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} \tilde{r}_{k\ell} W_{\ell i}$	7
R_{32}	4	3	(0, 1, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} \tilde{r}_{k\ell} \widetilde{\Omega}_{\ell i}$	6
R_{33}	4	3	(1, 0, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} \tilde{r}_{k\ell} \delta_{\ell i}$	7
R_{34}	1	4	(0, 0, 0)	$\sum_{i,j,k,\ell(dist)} \tilde{r}_{ij} \tilde{r}_{jk} \tilde{r}_{k\ell} \tilde{r}_{\ell i}$	8

analogous to T and T^* from Lemma E.13. By inspection, we see that for each term in the left column, the canonical upper bounds $\overline{c_{i_1,\ldots,i_m}}$ and $\overline{c^*_{i_1,\ldots,i_m}}$ on the coefficients $c_{i_1,\ldots,i_m}$ and $c^*_{i_1,\ldots,i_m}$ satisfy

$$\overline{c_{i_1,\ldots,i_m}} \lesssim \overline{c^*_{i_1,\ldots,i_m}}.$$

Recall that these canonical upper bounds were defined in Section E.4.1. Thus the conclusion of Lemma E.13 applies, and we have for each term X in the left column of Table 2,

$$|\mathbb{E}X| \lesssim \overline{\mathbb{E}X^*}, \qquad \mathrm{Var}(X) \lesssim \overline{\mathrm{Var}(X^*)}.$$

As discussed in Section E.4.1, the upper bounds on the means and variances in Lemmas E.7–E.10 are in fact upper bounds on $\overline{\mathbb{E}X^*}$ and $\overline{\mathrm{Var}(X^*)}$. By (E.48) and Lemmas E.7–E.10, for every post-expansion sum Y with $N_W^* \leq 4$ we have

$$|\mathbb{E}Y| \leq |\mathbb{E}X| + o(\|\theta\|_2^{-2})\sqrt{\mathbb{E}[X^2]} = |\mathbb{E}X| + o(\|\theta\|_2^{-2})\sqrt{\mathbb{E}[X]^2 + \mathrm{Var}(X)}$$

Table 2: *For clarity, this table and caption are borrowed from Table G.5 of Jin et al. (2021).*The 14 types of post-expansion sums with $N_W^* \leq 4$. The right column displays the post-expansion sums defined before which have similar forms as the post-expansion sums in the left column. For some terms in the right column, we permute (i, j, k, ℓ) in the original definition for ease of comparison with the left column. (In all expressions, the subscript "$i, j, k, \ell(dist)$" is omitted.)

	Expression		Expression
R_2	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}W_{k\ell}W_{\ell i}$	Z_{1b}	$\sum(\tilde{\eta}_i - \eta_i)\eta_j(\tilde{\eta}_j - \eta_j)\eta_k W_{k\ell}W_{\ell i}$
R_3	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)W_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i}$	Z_{2a}	$\sum \eta_\ell(\tilde{\eta}_j - \eta_j)W_{jk}\eta_k(\tilde{\eta}_i - \eta_i)W_{i\ell}$
R_4	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i}$	Z_{3d}	$\sum(\tilde{\eta}_i - \eta_i)\eta_j(\tilde{\eta}_j - \eta_j)\eta_k\widetilde{\Omega}_{k\ell}W_{\ell i}$
R_5	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}W_{k\ell}\widetilde{\Omega}_{\ell i}$	Z_{4b}	$\sum \widetilde{\Omega}_{ij}(\tilde{\eta}_j - \eta_j)\eta_k W_{k\ell}\eta_\ell(\tilde{\eta}_i - \eta_i)$
R_6	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}$	Z_{5a}	$\sum \eta_i(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\eta_\ell(\tilde{\eta}_i - \eta_i)$
R_9	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)^2\eta_k\widetilde{\Omega}_{k\ell}W_{\ell i}$	T_{1d}	$\sum \eta_\ell(\tilde{\eta}_j - \eta_j)^2\eta_k^2(\tilde{\eta}_i - \eta_i)W_{i\ell}$
	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)\widetilde{\Omega}_{k\ell}W_{\ell i}$	T_{1a}	$\sum \eta_\ell(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)\eta_k(\tilde{\eta}_i - \eta_i)W_{i\ell}$
R_{10}	$\sum(\tilde{\eta}_i - \eta_i)^2(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}W_{k\ell}\eta_\ell$	T_{1c}	$\sum(\tilde{\eta}_j - \eta_j)\eta_k W_{k\ell}\eta_\ell(\tilde{\eta}_i - \eta_i)^2\eta_j$
	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}W_{k\ell}(\tilde{\eta}_\ell - \eta_\ell)\eta_i$	T_{1a}	$\sum(\tilde{\eta}_j - \eta_j)\eta_k W_{k\ell}(\tilde{\eta}_\ell - \eta_\ell)\eta_i(\tilde{\eta}_i - \eta_i)\eta_j$
R_{11}	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)W_{jk}\eta_k(\tilde{\eta}_\ell - \eta_\ell)\widetilde{\Omega}_{\ell i}$	T_{1a}	$\sum(\tilde{\eta}_i - \eta_i)\eta_k W_{kj}(\tilde{\eta}_j - \eta_j)\eta_\ell(\tilde{\eta}_\ell - \eta_\ell)\eta_i$
	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)W_{jk}(\tilde{\eta}_k - \eta_k)\eta_\ell\widetilde{\Omega}_{\ell i}$	T_{1b}	$\sum \eta_i(\tilde{\eta}_k - \eta_k)W_{kj}(\tilde{\eta}_j - \eta_j)\eta_\ell^2(\tilde{\eta}_i - \eta_i)$
R_{12}	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)^2\eta_k\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}$	T_{2c}	$\sum \eta_i(\tilde{\eta}_j - \eta_j)^2\eta_k\widetilde{\Omega}_{k\ell}\eta_\ell(\tilde{\eta}_i - \eta_i)$
	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}$	T_{2a}	$\sum \eta_i(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)\widetilde{\Omega}_{k\ell}\eta_\ell(\tilde{\eta}_i - \eta_i)$
R_{13}	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}(\tilde{\eta}_k - \eta_k)\eta_\ell\widetilde{\Omega}_{\ell i}$	T_{2b}	$\sum \eta_i(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}(\tilde{\eta}_k - \eta_k)\eta_\ell^2(\tilde{\eta}_i - \eta_i)$
R_{16}	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)^2\eta_k(\tilde{\eta}_k - \eta_k)\eta_\ell\widetilde{\Omega}_{\ell i}$	F_b	$\sum \eta_i(\tilde{\eta}_j - \eta_j)^2\eta_k(\tilde{\eta}_k - \eta_k)\eta_\ell^2(\tilde{\eta}_i - \eta_i)$
	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)^2\eta_k^2(\tilde{\eta}_\ell - \eta_\ell)\widetilde{\Omega}_{\ell i}$	F_b	$\sum \eta_i(\tilde{\eta}_j - \eta_j)^2\eta_k^2(\tilde{\eta}_\ell - \eta_\ell)\eta_\ell(\tilde{\eta}_i - \eta_i)$
	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)^2\eta_\ell\widetilde{\Omega}_{\ell i}$	F_b	$\sum \eta_i(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)^2\eta_\ell^2(\tilde{\eta}_i - \eta_i)$
	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)\eta_k(\tilde{\eta}_\ell - \eta_\ell)\widetilde{\Omega}_{\ell i}$	F_a	$\sum \eta_i(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)\eta_k(\tilde{\eta}_\ell - \eta_\ell)\eta_\ell(\tilde{\eta}_i - \eta_i)$
R_{17}	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)\widetilde{\Omega}_{k\ell}(\tilde{\eta}_\ell - \eta_\ell)\eta_i$	F_a	$\sum \eta_i(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)\eta_k(\tilde{\eta}_\ell - \eta_\ell)\eta_\ell(\tilde{\eta}_i - \eta_i)$
	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)^2\eta_k\widetilde{\Omega}_{k\ell}(\tilde{\eta}_\ell - \eta_\ell)\eta_i$	F_b	$\sum \eta_i(\tilde{\eta}_j - \eta_j)^2\eta_k^2(\tilde{\eta}_\ell - \eta_\ell)\eta_\ell(\tilde{\eta}_i - \eta_i)$
	$\sum(\tilde{\eta}_i - \eta_i)^2(\tilde{\eta}_j - \eta_j)^2\eta_k\widetilde{\Omega}_{k\ell}\eta_\ell$	F_c	$\sum \eta_\ell(\tilde{\eta}_i - \eta_i)^2\eta_k^2(\tilde{\eta}_j - \eta_j)^2\eta_\ell$
R_{21}	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)^2(\tilde{\eta}_k - \eta_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}$	F_b	$\sum \eta_i(\tilde{\eta}_j - \eta_j)^2\eta_k(\tilde{\eta}_k - \eta_k)\eta_\ell^2(\tilde{\eta}_i - \eta_i)$
R_{22}	$\sum(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\widetilde{\Omega}_{jk}(\tilde{\eta}_k - \eta_k)(\tilde{\eta}_\ell - \eta_\ell)\widetilde{\Omega}_{\ell i}$	F_a	$\sum \eta_i(\tilde{\eta}_j - \eta_j)\eta_j(\tilde{\eta}_k - \eta_k)\eta_k(\tilde{\eta}_\ell - \eta_\ell)\eta_\ell(\tilde{\eta}_i - \eta_i)$

$$
\begin{aligned}
&\lesssim \tilde{\lambda}^2\lambda_1 + o(\|\theta\|_2^{-2}) \cdot \sqrt{\tilde{\lambda}^4\lambda_1^2 + \lambda_1^4 + \tilde{\lambda}^6 + \tilde{\lambda}^2\lambda_1^3} \\
&\lesssim \tilde{\lambda}^2\lambda_1 + \lambda_1^2 + \tilde{\lambda}^3 + |\tilde{\lambda}|\lambda_1^{3/2} = o(\tilde{\lambda}^4)
\end{aligned}
$$

by the assumption that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$. Similarly,

$$
\begin{aligned}
\mathrm{Var}(Y) &\lesssim \mathrm{Var}(X) + o(\|\theta\|_2^{-4})\mathbb{E}[X^2] = \mathrm{Var}(X) + o(\|\theta\|_2^{-4})(\mathbb{E}[X]^2 + \mathrm{Var}(X)) \\
&\lesssim \lambda_1^4 + \tilde{\lambda}^6 + \tilde{\lambda}^2\lambda_1^3 + o(\|\theta\|_2^{-4}) \cdot \left(\tilde{\lambda}^4\lambda_1^2 + \lambda_1^4 + \tilde{\lambda}^6 + \tilde{\lambda}^2\lambda_1^3\right) \lesssim o(\tilde{\lambda}^8).
\end{aligned}
$$

Analysis of terms with $N_W^* > 4$ Recall that

$$
\eta = \frac{1}{\sqrt{v}}(\mathbb{E}A)\mathbf{1}_n, \quad \tilde{\eta} = \frac{1}{\sqrt{v}}A\mathbf{1}_n, \quad v = \mathbf{1}_n'(\mathbb{E}A)\mathbf{1}_n
$$

.

Define

$$
G_i = \tilde{\eta}_i - \eta_i. \tag{E.51}
$$

Among the post-expansion sums in Table (1) satisfying $N_W^* = 5$, only R_7, R_8, and R_{25}–R_{28} depend on $\tilde{\Omega}$. Each of these terms falls into one of the types

$$
\begin{aligned}
J_5' &= \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{jk}(G_iG_jG_kG_\ell W_{\ell i}), \\
J_6' &= \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{k\ell}(G_iG_j^2G_kW_{\ell i})
\end{aligned}
$$

$$J_9 = \sum_{i,j,k,\ell(dist)} \eta_k \widetilde{\Omega}_{\ell i}(G_i G_j^2 G_k G_\ell)$$
$$J_{10} = \sum_{i,j,k,\ell(dist)} \eta_\ell \widetilde{\Omega}_{\ell i}(G_i G_j^2 G_k^2).$$

See (Jin et al., 2021, Supplement, Section G.4.10.2) for more details.

To handle J_5' and J_6', we compare them to

$$J_5 = \sum_{i,j,k,\ell(dist)} \eta_j \eta_k (G_i G_j G_k G_\ell W_{\ell i})$$
$$J_6 = \sum_{i,j,k,\ell(dist)} \eta_k \eta_\ell (G_i G_j^2 G_k W_{\ell i}),$$

both of which are considered in (Jin et al., 2021, Supplement, Section G.4.10.2). Note that neither J_5 nor J_5 depends on $\tilde{\Omega}$. Setting $T = J_5'$ and $T^* = J_5$ in Lemma E.13 and noting that $|\tilde{\Omega}_{jk}| \lesssim \theta_j \theta_k$ by (E.24), we see that the hypotheses of Lemma E.13 are satisfied. In (Jin et al., 2021, Supplement, Section G.4.10.2), it is shown that

$$\mathbb{E}[J_5^2] \leq \overline{\mathbb{E}[J_5]}^2 + \overline{\mathrm{Var}(J_5)} = o(\|\theta\|_2^8).$$

Applying Lemma E.13, we conclude that

$$\mathbb{E}[J_5^{'2}] = o(\|\theta\|_2^8).$$

Similarly, it is shown in (Jin et al., 2021, Supplement, Section G.4.10.2) that

$$\mathbb{E}[J_6^2] \leq \overline{\mathbb{E}[J_6]}^2 + \overline{\mathrm{Var}(J_6)} = o(\|\theta\|_2^8).$$

Setting $T = J_6'$ and $T^* = J_6$, the hypotheses of Lemma E.13 are satisfied because $|\tilde{\Omega}_{k\ell}| \lesssim \theta_k \theta_\ell$. We conclude that

$$\mathbb{E}[J_6^{'2}] = o(\|\theta\|_2^8).$$

The terms J_9 and J_{10} can be analyzed explicitly using the strategy described in Section E.4.1. We omit the full details and instead give a simplified proof in the case where $\|\theta\|_2 \gg [\log(n)]^{5/2}$. The event

$$E = \cap_{i=1}^n E_i, \qquad \text{where} \quad E_i = \left\{\sqrt{v}|G_i| \leq C_0 \sqrt{\theta_i \|\theta\|_1 \log(n)}\right\}. \tag{E.52}$$

is introduced in (Jin et al., 2021, Supplement,pg.110). By applying Bernstein's inequality and the union bound, it is shown that E holds with probability at least $1 - n^{-C_0/2.01}$. Applying the crude bound $|G_i| \leq n$ and triangle inequality, we see that $|J_9| \lesssim n^9$ with high probability, and thus for C_0 sufficiently large,

$$\mathbb{E}[|J_9|^2 \cdot \mathbf{1}_{E^c}] = o(1).$$

Under the event E, we have by (E.20),

$$\begin{aligned}
|J_9| &\leq \sum_{i,j,k,\ell} |\eta_k \widetilde{\Omega}_{\ell i}||G_i G_j^2 G_k G_\ell| \\
&\lesssim \sum_{i,j,k,\ell} (\theta_i \theta_k \theta_\ell) \frac{\sqrt{\theta_i \theta_j^2 \theta_k \theta_\ell \|\theta\|_1^5 [\log(n)]^5}}{\sqrt{v^5}} \\
&\lesssim \frac{[\log(n)]^{5/2}}{\sqrt{\|\theta\|_1^5}} \Big(\sum_i \theta_i^{3/2}\Big)\Big(\sum_j \theta_j\Big)\Big(\sum_k \theta_k^{3/2}\Big)\Big(\sum_\ell \theta_\ell^{3/2}\Big) \\
&\lesssim \frac{[\log(n)]^{5/2}}{\sqrt{\|\theta\|_1^3}} \Big(\sum_i \theta_i^{3/2}\Big)^3 \\
&\lesssim \frac{[\log(n)]^{5/2}}{\sqrt{\|\theta\|_1^3}} \Big(\sum_i \theta_i^2\Big)^{3/2} \Big(\sum_i \theta_i\Big)^{3/2}
\end{aligned}$$

$$\lesssim [\log(n)]^{5/2}\|\theta\|^3.$$

It follows that

$$\mathbb{E}[J_9^2] = \mathrm{Var}(J_9) + \mathbb{E}[J_9]^2 = o(\|\theta\|_2^8).$$

We give a similar, simplified argument for J_{10} assuming that $\|\theta\|_2 \gg [\log(n)]^{5/2}$. Under the event E, we have

$$\begin{aligned}
|J_{10}| &\leq \sum_{i,j,k,\ell} |\eta_\ell \widetilde{\Omega}_{\ell i}||G_i G_j^2 G_k^2| \\
&\lesssim \sum_{i,j,k,\ell} (\theta_i \theta_\ell^2) \frac{\sqrt{\theta_i \theta_j^2 \theta_k^2 \|\theta\|_1^5 [\log(n)]^5}}{\sqrt{v^5}} \\
&\lesssim \frac{[\log(n)]^{5/2}}{\sqrt{\|\theta\|_1^5}} \Big(\sum_i \theta_i^{3/2}\Big)\Big(\sum_j \theta_j\Big)\Big(\sum_k \theta_k\Big)\Big(\sum_\ell \theta_\ell^2\Big) \\
&\lesssim \frac{[\log(n)]^{5/2}}{\sqrt{\|\theta\|_1^5}} \big(\|\theta\|\sqrt{\|\theta\|_1}\big)\|\theta\|_1^2\|\theta\|^2 \\
&\lesssim [\log(n)]^{5/2}\|\theta\|^3;
\end{aligned}$$

Hence

$$\mathbb{E}[J_{10}^2] = \mathrm{Var}(J_{10}) + \mathbb{E}[J_{10}]^2 = o(\|\theta\|_2^8).$$

Next we consider the terms with $N_W^* = 6$. The only term that depends on $\tilde{\Omega}$ is R_{32}, which has the form

$$K_5' = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ik} G_i G_j^2 G_k G_\ell^2.$$

The variance of K_5' can be analyzed explicitly using the strategy described in Section E.4.1. To save space, we give a simplified argument when $\|\theta\|_2 \gg [\log(n)]^{3/2}$. Again let E denote the event (E.52). Under this event we have

$$\begin{aligned}
|K_5'| &\lesssim \sum_{i,j,k,\ell} (\theta_i \theta_k) \frac{\sqrt{\theta_i \theta_j^2 \theta_k \theta_\ell^2}\|\theta\|_1^3 [\log(n)]^3}{v^3} \\
&\lesssim \frac{[\log(n)]^3}{\|\theta\|_1^3} \Big(\sum_i \theta_i^{3/2}\Big)\Big(\sum_j \theta_j\Big)\Big(\sum_k \theta_k^{3/2}\Big)\Big(\sum_\ell \theta_\ell\Big) \\
&\lesssim \frac{[\log(n)]^3}{\|\theta\|_1^3} \big(\|\theta\|\sqrt{\|\theta\|_1}\big)^2\|\theta\|_1^2 \\
&\lesssim [\log(n)]^3\|\theta\|^2,
\end{aligned}$$

Above we apply (E.20) and (E.24) as well as Cauchy–Schwarz. It follows that

$$\mathbb{E}[K_5^{'2}] = \mathrm{Var}(K_5') + \mathbb{E}[K_5']^2 = o(\|\theta\|_2^8).$$

Finally, all terms with $N_W^* \geq 7$ have no dependence on $\tilde{\Omega}$, and thus the bounds carry over immediately (see (Jin et al., 2021, Supplement, Section G.4.10.4) for details). This completes the proof of the lemma. □

E.4.7 Proof of Lemma E.12

Define

$$\epsilon_{ij}^{(1)} = \eta_i^* \eta_j^* - \eta_i \eta_j, \quad \epsilon_{ij}^{(2)} = (1 - \frac{v}{V})\eta_i \eta_j, \quad \epsilon_{ij}^{(3)} = -(1 - \frac{v}{V})\delta_{ij}.$$

Note that $\epsilon_{ij}^{(1)}$ is a nonstochastic term. As shown in (Jin et al., 2021, Supplement, pg. 119), we have

$$|\epsilon_{ij}^{(1)}| \lesssim \frac{\|\theta\|_\infty}{\|\theta\|_1} \cdot \theta_i \theta_j,$$

which implies that

$$|\epsilon_{ij}^{(1)}| \lesssim \frac{1}{\|\theta\|_2^2} \cdot \theta_i \theta_j \tag{E.53}$$

by (E.2).

As discussed in (Jin et al., 2021, Supplement, Section G.3), $Q - Q^*$ is a finite sum of terms of the form

$$\sum_{i,j,k,\ell(dist)} a_{ij} b_{jk} c_{k\ell} d_{\ell i}, \qquad \text{where} \quad a, b, c, d \in \{\widetilde{\Omega}, W, \delta, \tilde{r}, \epsilon^{(1)}, \epsilon^{(2)}, \epsilon^{(3)}\}. \tag{E.54}$$

Let Y denote an arbitrary term of the form above, and given $X \in \{\widetilde{\Omega}, W, \delta, \tilde{r}, \epsilon^{(1)}, \epsilon^{(2)}, \epsilon^{(3)}\}$, let N_X denote the total number of a, b, c, d that are equal to X. It holds that

$$Y = \big(\frac{v}{V}\big)^{N_{\tilde{r}}} (-1)^{N_\epsilon^{(3)}} \Big(1 - \frac{v}{V}\Big)^{N_\epsilon^{(2)} + N_\epsilon^{(3)}} X, \qquad X \equiv \sum_{i,j,k,\ell(dist)} a_{ij} b_{jk} c_{k\ell} d_{\ell i}.$$

where

$$\begin{cases} a, b, c, d \in \{\widetilde{\Omega}, W, \delta, (V/v)\tilde{r}, \epsilon^{(1)}, \eta\eta^\mathsf{T}\}, \\ \text{number of } \eta_i \eta_j \text{ in the product is } N_\epsilon^{(2)}, \\ \text{number of } \delta_{ij} \text{ in the product is } N_\delta + N_\epsilon^{(3)}, \\ \text{number of any other term in the product is same as before.} \end{cases} \tag{E.55}$$

Let x_n denote a sequence of real numbers such that $\sqrt{\log(\|\theta\|_1)} \ll x_n \ll \|\theta\|_1$. Mimicking the argument in (Jin et al., 2021, Supplement,pg.121), it holds that

$$\mathbb{E}[Y^2] \lesssim \Big(\frac{x_n^2}{\|\theta\|_1^2}\Big)^{N_\epsilon^{(2)} + N_\epsilon^{(3)}} \cdot \mathbb{E}[X^2] + o(1),$$

By (E.4), there exists a sequence $\log(\|\theta\|_1) \ll x_n \ll \|\theta\|_1 / \|\theta\|_2^2$. Hence,

$$\mathbb{E}[Y^2] \lesssim \Big(\frac{1}{\|\theta\|_2^4}\Big)^{N_\epsilon^{(2)} + N_\epsilon^{(3)}} \cdot \mathbb{E}[X^2] + o(1), \tag{E.56}$$

Thus we focus on controlling $\mathbb{E}[X^2]$.

Consider a new random variable X^* defined to be

$$X^* \equiv \sum_{i,j,k,\ell(dist)} a_{ij}^* b_{jk}^* c_{k\ell}^* d_{\ell i}^*$$

where

$$a^* = \begin{cases} \frac{1}{\|\theta\|_2^2} \cdot \theta\theta^\mathsf{T} & \text{if } a = \epsilon^{(1)} \\ \theta\theta^\mathsf{T} & \text{if } a \in \{\tilde{\Omega}, \eta\eta^\mathsf{T}\} \\ a & \text{otherwise} \end{cases}$$

$$b^* = \begin{cases} \frac{1}{\|\theta\|_2^2} \cdot \theta\theta^\mathsf{T} & \text{if } b = \epsilon^{(1)} \\ \theta\theta^\mathsf{T} & \text{if } b \in \{\tilde{\Omega}, \eta\eta^\mathsf{T}\} \\ b & \text{otherwise} \end{cases}$$

$$c^* = \begin{cases} \frac{1}{\|\theta\|_2^2} \cdot \theta\theta^\mathsf{T} & \text{if } c = \epsilon^{(1)} \\ \theta\theta^\mathsf{T} & \text{if } c \in \{\tilde{\Omega}, \eta\eta^\mathsf{T}\} \\ c & \text{otherwise} \end{cases}$$

$$d^* = \begin{cases} \frac{1}{\|\theta\|_2^2} \cdot \theta\theta^\mathsf{T} & \text{if } d = \epsilon^{(1)} \\ \theta\theta^\mathsf{T} & \text{if } \in \{\tilde{\Omega}, \eta\eta^\mathsf{T}\} \\ d & \text{otherwise .} \end{cases}$$

Also define

$$\tilde{X} = \sum_{ijk\ell(dist)} \tilde{a}_{ij}\tilde{b}_{jk}\tilde{c}_{k\ell}\tilde{d}_{\ell i}$$

where

$$\tilde{a} = \begin{cases} \theta\theta^\mathsf{T} & \text{if } a \in \{\epsilon^{(1)}, \tilde{\Omega}, \eta\eta^\mathsf{T}\} \\ a & \text{otherwise} \end{cases}$$

$$\tilde{b} = \begin{cases} \theta\theta^\mathsf{T} & \text{if } b \in \{\epsilon^{(1)}, \tilde{\Omega}, \eta\eta^\mathsf{T}\} \\ b & \text{otherwise} \end{cases}$$

$$\tilde{c} = \begin{cases} \theta\theta^\mathsf{T} & \text{if } c \in \{\epsilon^{(1)}, \tilde{\Omega}, \eta\eta^\mathsf{T}\} \\ c & \text{otherwise} \end{cases}$$

$$\tilde{d} = \begin{cases} \theta\theta^\mathsf{T} & \text{if } d \in \{\epsilon^{(1)}, \tilde{\Omega}, \eta\eta^\mathsf{T}\} \\ d & \text{otherwise .} \end{cases}$$

Note that $X^* = \big(\frac{1}{\|\theta\|_2^2}\big)^{N_\epsilon^{(1)}} \tilde{X}$ and $\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d} \in \{\theta\theta^\mathsf{T}, W, \delta, (V/v)\tilde{r}\}$. Later we show that

$$\mathbb{E}[X^2] \lesssim \mathbb{E}[X^{*2}] \tag{E.57}$$

First we bound $\mathbb{E}[\tilde{X}^2]$ in the case when $N_W + N_\delta + N_{\tilde{r}} = 0$. Note that for all such terms in $Q - Q^*$, we have $N_\epsilon^{(1)} + N_\epsilon^{(2)} + N_\epsilon^{(3)} + N_{\tilde{\Omega}} = 4$ and $N_{\tilde{\Omega}} < 4$. In particular, $\tilde{X}$ and X^* are nonstochastic. If $N_{\tilde{\Omega}} = 3$, then by (E.22) and (E.24),

$$|\tilde{X}| = \Big|\sum_{ijk\ell(dist)} \tilde{\Omega}_{ij}\tilde{\Omega}_{jk}\tilde{\Omega}_{k\ell}\theta_i\theta_\ell\Big| \lesssim \frac{1}{\|\theta\|_2^2}\sum_{ijk\ell} \beta_i\theta_i^2\beta_j^2\theta_j^2\beta_k^2\theta_k^2\beta_\ell\theta_\ell^2 \lesssim \|\beta\circ\theta\|_2^6\|\theta\|_2^2$$

If $N_{\tilde{\Omega}} = 2$, there are two cases. First,

$$|\tilde{X}| = \Big|\sum_{ijk\ell(dist)} \tilde{\Omega}_{ij}\tilde{\Omega}_{jk}\theta_k\theta_\ell\theta_\ell\theta_i\Big| \lesssim \sum_{ijk\ell} \beta_i\theta_i\beta_j^2\theta_j^2\beta_k\theta_k^2\theta_\ell^2\theta_i \lesssim \|\beta\circ\theta\|_2^4\|\theta\|_2^4,$$

and second

$$|\tilde{X}| = \Big|\sum_{ijk\ell(dist)} \tilde{\Omega}_{ij}\theta_j\theta_k\tilde{\Omega}_{k\ell}\theta_\ell\theta_i\Big| \lesssim \sum_{ijk\ell} \beta_i\theta_i^2\beta_j\theta_j^2\beta_k\theta_k^2\beta_\ell\theta_\ell^2 \lesssim \|\beta\circ\theta\|_2^4\|\theta\|_2^4$$

Finally if $N_{\tilde{\Omega}} = 1$,

$$|\tilde{X}| = \Big|\sum_{ijk\ell(dist)} \tilde{\Omega}_{ij}\theta_j\theta_k^2\theta_\ell^2\theta_i\Big| \lesssim \sum_{ijk\ell} \tilde{\beta}_i\theta_i^2\beta_j\theta_j^2\theta_k^2\theta_\ell^2 \lesssim \|\beta\circ\theta\|_2^2\|\theta\|_2^6.$$

Note that when $N_W + N_\delta + N_{\tilde{r}} = 0$

$$|X| \lesssim |X^*|$$

by (E.22), (E.20), and (E.53). By the bounds above, we conclude that

$$|Y| \lesssim \Big(\frac{1}{\|\theta\|_2^2}\Big)^{N_\varepsilon^{(1)}+N_\varepsilon^{(2)}+N_\varepsilon^{(3)}}|\tilde{X}| \lesssim \max_{1\le k\le 3}\|\beta\circ\theta\|_2^{2k}\|\theta\|_2^{2(4-k)} \lesssim |\tilde{\lambda}|^3. \tag{E.58}$$

Next we bound $\mathbb{E}[\tilde{X}^2]$ in the case when $N_W + N_\delta + N_{\tilde{r}} > 0$. By Lemma E.2 and the definition of $f \in \mathbb{R}^2$ there, we have $\tilde{\Omega}_{ij} = \alpha_i\alpha_j\theta_i\theta_j$ where $\alpha = \Pi f$. Observe that in Lemmas E.7–E.11, we bound the mean and variance of all terms of the form

$$Z \equiv \sum_{i,j,k,\ell(dist)} a_{ij}b_{jk}c_{k\ell}d_{\ell i}, \qquad \text{where} \quad a, b, c, d \in \{\widetilde{\Omega}, W, \delta, (V/v)\tilde{r}\}.$$

As a result, the proofs of Lemmas E.7–E.11 produce a function F such that

$$\mathbb{E}[Z^2] \le F(\theta, \beta; N_{\tilde{\Omega}}, N_W, N_\delta, N_{\tilde{r}}),$$

where recall that $|\alpha_i| \le \beta_i$.

Note that in what follows, we use $'$ to denote a new variable rather than the transpose. As a direct corollary to the proofs of Lemmas E.7–E.11, if we define a new matrix $\tilde{\Omega}' = \alpha'_i \alpha'_j \theta_i \theta_j$ where α' is a vector with a coordinate-wise bound of the form $|\alpha'_i| \le \beta'_i$, then

$$Z' \equiv \sum_{i,j,k,\ell (dist)} a_{ij} b_{jk} c_{k\ell} d_{\ell i}, \qquad \text{where} \quad a, b, c, d \in \{\widetilde{\Omega}', W, \delta, (V/v)\tilde{r}\}$$

satisfies

$$\mathbb{E}[Z'^2] \le F(\theta, \beta'; N'_{\tilde{\Omega}'}, N'_W, N'_\delta, N'_{\tilde{r}}),$$

where, for example, N'_δ counts the number of appearances of δ in Z'. This can be verified by tracing each calculation in Lemmas E.7–E.11 line by line, replacing all occurences of $\tilde{\Omega}$ with $\tilde{\Omega}'$, and replacing every usage of the bound $|\alpha_i| \le \beta_i$ with $|\alpha'_i| \le \beta'_i$ instead. In other words, our proofs make no use of the specific value of $\alpha = \Pi f$.

In particular, if $\alpha = \mathbf{1}$, then $\tilde{\Omega}' = \theta\theta^\mathsf{T}$. In this case we may set $\beta = \mathbf{1}$. Observe that $\tilde{X}$ has the form of Z' with this choice of $\tilde{\Omega}'$. Hence,

$$\mathbb{E}[\tilde{X}^2] \le F(\theta, \mathbf{1}; \tilde{N}_{\tilde{\Omega}'}, \tilde{N}_W, \tilde{N}_\delta, \tilde{N}_{\tilde{r}}). \tag{E.59}$$

By careful inspection of the bounds in Lemmas E.7–E.11, we see that

$$F(\theta, \mathbf{1}; N_{\tilde{\Omega}'}, N_W, N_\delta, N_{\tilde{r}}) \lesssim \|\theta\|_2^{12}. \tag{E.60}$$

In (Jin et al., 2021, Supplement, Section G.3) it is shown that all terms in the decomposition of $Q - Q^*$ satisfy $N_\epsilon^{(1)} + N_\epsilon^{(2)} + N_\epsilon^{(3)} > 0$. Using this fact along with (E.56), (E.57), (E.59) and (E.60),

$$\mathbb{E}[Y^2] \lesssim \Big(\frac{1}{\|\theta\|_2^4}\Big)^{N_\epsilon^{(2)} + N_\epsilon^{(3)}} \cdot \big(\frac{1}{\|\theta\|_2^2}\big)^{2N_\epsilon^{(1)}} \cdot \mathbb{E}[\tilde{X}^2] + o(1) \lesssim \|\theta\|_2^8. \tag{E.61}$$

Observe that (E.58) and (E.61) recover the bounds in Lemma E.12 under the alternative hypothesis, and the bounds under the null hypothesis transfer directly from (Jin et al., 2021, Lemma G.12). Thus it only remains to justify (E.57) when $N_W + N_\delta + N_{\tilde{r}} > 0$. Let us write

$$X = \sum_{i_1, \dots, i_m} c_{i_1, \dots, i_m} G_{i_1, \dots, i_m}$$
$$X^* = \sum_{i_1, \dots, i_m} c^*_{i_1, \dots, i_m} G_{i_1, \dots, i_m}$$

in the form described in Section E.4.1, where now

- $c_{i_1, \dots, i_m} = \prod_{(s,s') \in A} \Gamma^{(s,s')}_{i_s, i_{s'}}$ is a nonstochastic term where $A \subset [m] \times [m]$ and
$$\Gamma^{(s,s')} \in \{\tilde{\Omega}, \eta^* \mathbf{1}^\mathsf{T}, \eta \mathbf{1}^\mathsf{T}, \mathbf{1}\mathbf{1}^\mathsf{T}, \epsilon^{(1)}, \eta\eta^\mathsf{T}\}$$
- $c^*_{i_1, \dots, i_m} = \prod_{(s,s') \in A} \Gamma^{(s,s')}_{i_s, i_{s'}}$ is a nonstochastic term where $A \subset [m] \times [m]$ and
$$\Gamma^{(s,s')} \in \{\eta^* \mathbf{1}^\mathsf{T}, \eta \mathbf{1}^\mathsf{T}, \mathbf{1}\mathbf{1}^\mathsf{T}, \theta\theta^\mathsf{T}/\|\theta\|_2^2, \theta\theta^\mathsf{T}\}$$
- $G_{i_1, \dots, i_m} = \prod_{(s,s') \in B} W_{i_s, i_{s'}}$ where $B \subset [m] \times [m]$.

If $\Gamma^{(s,s')} \in \{\theta\theta^\mathsf{T}, \theta\theta^\mathsf{T}/\|\theta\|_2^2\}$, we simply let $\overline{\Gamma^{(s,s')}} = \Gamma^{(s,s')}$ and define

$$\overline{c^*_{i_1, \dots, i_m}} = \prod_{(s,s') \in A} \overline{\Gamma^{(s,s')}_{i_s, i_{s'}}}$$

as in Section E.4.1. We also define the canonical upper bound $\overline{\mathbb{E}X^*}$ on $|\mathbb{E}X^*|$ and the canonical upper bound $\overline{\mathrm{Var}(X^*)}$ on $\mathrm{Var}(X^*)$ similarly to Section E.4.1. By the discussion above and (E.59),

$$\overline{\mathbb{E}[X^*]} \equiv \big(\frac{1}{\|\theta\|_2^2}\big)^{N_\epsilon^{(1)}} \sqrt{F(\theta, \mathbf{1}; \tilde{N}_{\tilde{\Omega}'}, \tilde{N}_W, \tilde{N}_\delta, \tilde{N}_{\tilde{r}})},$$

and

$$\overline{\mathrm{Var}(X^*)} \equiv \big(\frac{1}{\|\theta\|_2^2}\big)^{2N_\epsilon^{(1)}} F(\theta, \mathbf{1}; \tilde{N}_{\tilde{\Omega}'}, \tilde{N}_W, \tilde{N}_\delta, \tilde{N}_{\tilde{r}}).$$

Next observe that

$$|c_{i_1,\ldots,i_m}| \lesssim |c^*_{i_1,\ldots,i_m}| \lesssim |\overline{c^*_{i_1,\ldots,i_m}}|.$$

By a mild extension of Lemma E.13 it follows that

$$|\mathbb{E}X| \lesssim \overline{\mathbb{E}X^*}$$
$$\mathrm{Var}(X) \lesssim \overline{\mathrm{Var}(X^*)},$$

which verifies (E.57) and completes the proof. □

E.5 CALCULATIONS IN THE SBM SETTING

We compute the order of λ_1 and $\tilde{\lambda}_1 = \lambda_2$ in the SBM setting (which are the two nonzero eigenvalues of Ω). By basic algebra, λ_1, λ_2 are also the two nonzero eigenvalues of the following matrix

$$\begin{bmatrix} N & 0 \\ 0 & n-N \end{bmatrix}^{1/2} \times \begin{bmatrix} a & b \\ b & c \end{bmatrix} \times \begin{bmatrix} N & 0 \\ 0 & n-N \end{bmatrix}^{1/2} = \begin{bmatrix} aN & \sqrt{N(n-N)}b \\ \sqrt{N(n-N)}b & (n-N)c \end{bmatrix},$$

where b is given by (H.1). By direct calculations and pluging the definitions of b,

$$\begin{aligned}\lambda_1 =& \frac{aN + (n-N)c + \sqrt{(aN-(n-N)c)^2 + 4N(n-N)b^2}}{2} \\ =& \frac{aN + (n-N)c + |(n-N)c - aN|\frac{n}{n-2N}}{2}.\end{aligned}$$

Recall that

$$b = \frac{nc - N(a+c)}{n-2N}.$$

It is required that $b \geq 0$. Therefore,

$$nc - (a+c)N \geq 0, \qquad \text{and so} \qquad (n-N)c \geq aN. \tag{E.62}$$

By direct calculations, it follows that

$$\lambda_1 = \frac{(n-N)^2c - aN^2}{n-2N} = \frac{(n-N)c((n-N) - \frac{aN}{(n-N)c}N)}{n-2N} \sim \frac{(n-N)c(n-N)}{n-2N} \sim nc$$

where in the last two $\asymp$, we have used $(n-N)c \geq aN$ and $N = o(n)$. Similarly,

$$\lambda_2 = \frac{aN + (n-N)c - \sqrt{(aN-(n-N)c)^2 + 4N(n-N)b^2}}{2} = \frac{(a-c)N(n-N)}{n-2N} \sim N(a-c).$$

F PROOF OF THEOREM 2.3 (POWERLESSNESS OF χ^2 TEST)

We compare the SgnQ test with the χ^2 test. Recall we assume $\theta_i = \mathbf{1}_n$. The χ^2 test statistic is defined to be

$$X_n = \frac{1}{\hat{\alpha}(1-\hat{\alpha})(n-1)} \sum_{i=1}^n \big((A\mathbf{1}_n)_i - \hat{\alpha}n\big)^2, \qquad \text{where } \hat{\alpha} = \frac{1}{n(n-1)} \sum_{i \neq j} A_{ij}.$$

We also define an idealized χ^2 test statistic by

$$\tilde{X}_n = \frac{1}{\alpha(1-\alpha)(n-1)} \sum_{i=1}^{n} \big((A\mathbf{1}_n)_i - \alpha n\big)^2, \qquad \text{where } \alpha = \frac{1}{n(n-1)} \sum_{i \neq j} \Omega_{ij}.$$

The χ^2 test is defined to be

$$\chi_n^2 = \mathbf{1}\left[\frac{|X_n - n|}{\sqrt{2n}} > z_{\gamma/2}\right],$$

where z_γ is such that $\mathbb{P}[|N(0,1)| \geq z_\gamma] = \gamma$. Similarly, the idealized χ^2 test is defined by

$$\tilde{\chi}_n^2 = \mathbf{1}\left[\frac{|\tilde{X}_n - n|}{\sqrt{2n}} > z_{\gamma/2}\right],$$

In certain degree-homogeneous settings, the χ^2 test is known to have full power Arias-Castro & Verzelen (2014); Cammarata & Ke (2022).

We prove the following, which directly implies Theorem 2.3.

Theorem F.1. *Suppose that* (2.7) *holds and that* $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$*, and recall that under these conditions, the power of the SgnQ test goes to* 1*. Next suppose that the following regularity conditions hold under the null and alternative:*

(i) $\theta = \mathbf{1}_n$

(ii) $\alpha \to 0$

(iii) $\alpha^2 n \to \infty$

(iv) $\sum_{ij} (\Omega_{ij} - \alpha)^2 = o(\alpha n^{3/2})$.

Then the power of both the χ^2*-test and idealized* χ^2*-test goes to* γ *(which is the prescribed level of the test).*

Note that the previous theorem implies Theorem 2.3. By Theorem 2.2, SgnQ has full power even without the extra regularity conditions (i)–(iv). On the other hand, for any fixed alternative DCBM satisfying (i)–(iv), Theorem F.1 implies that χ^2 has power κ.

Proof of Theorem F.1. Theorem 2.2 confirms that SgnQ has full power provided that (2.7) holds and that $|\tilde{\lambda}|/\sqrt{\lambda_1} \to \infty$. It remains to justify the powerlessness of the χ^2 test.

Consider an SBM in the alternative such that $\Omega\mathbf{1} = (\alpha n)\mathbf{1}$ and $|\tilde{\lambda}|/\sqrt{\lambda_1} \asymp N(a-c)/\sqrt{nc} \to \infty$. To do this we select an integer $N > 0$ to be the size of the smaller community and set $b = \frac{cn-(a+c)N}{n-2N}$. The remaining regularity conditions are satisfied if $c \to 0$ and $cn \ll N(a-c)^2 \ll cn^{3/2}$. We show that both X_n and $\tilde{X}_n$ are asymptotically normal under the specified alternative, which is enough to imply Theorem F.1.

In Cammarata & Ke (2022) it is shown that

$$\hat{T}_n \equiv [(n-1)\hat{\alpha}(1-\hat{\alpha})](X_n - n) = \sum_{i,j,k \text{ (dist.)}} (A_{ik} - \hat{\alpha})(A_{jk} - \hat{\alpha}). \tag{F.1}$$

We introduce an idealized version T_n of $\hat{T}_n$, which is

$$T_n = \sum_{i,j,k \text{ (dist.)}} (A_{ik} - \alpha)(A_{jk} - \alpha),$$

Following Cammarata & Ke (2022), we have

$$\frac{X_n - n}{\sqrt{2n}} = \left(\frac{n-2}{n-1}\right)^{1/2} U_n V_n Z_n. \tag{F.2}$$

where

$$U_n = \frac{\alpha_n(1-\alpha_n)}{\hat{\alpha}_n(1-\hat{\alpha}_n)}, \qquad V_n = \frac{\hat{T}_n}{T_n}, \qquad Z_n = \frac{\frac{T_n}{(n-1)\alpha_n(1-\alpha_n)}}{\sqrt{\frac{2n(n-2)}{(n-1)}}}.$$

Since the terms of $\hat{\alpha}$ are bounded, the law of large numbers implies that $U_n \xrightarrow{\mathbb{P}} 1$. Furthermore, since $\alpha n \to \infty$ by assumption that $\alpha^2 n \to \infty$, a straightforward application of the Berry-Esseen theorem implies that

$$\sqrt{\frac{n(n-1)}{2}} \frac{\hat{\alpha}_n - \alpha_n}{\sqrt{\alpha_n(1-\alpha_n)}} \Rightarrow \mathcal{N}(\mu, 1).$$

With the previous fact, mimicking the argument in (Cammarata & Ke, 2022, pg.32), it also follows that

$$V_n \xrightarrow{\mathbb{P}} 1,$$

provided we can show that $Z_n \Rightarrow N(0,1)$. We omit the details since the argument is very similar.

Thus it suffices to study Z_n. We first analyze T_n, which we decompose as

$$\begin{aligned} T_n = & \sum_{i,j,k \text{ (dist.)}} (A_{ik} - \Omega_{ik})(A_{jk} - \Omega_{jk}) + 2 \sum_{ijk(dist)} (\Omega_{ik} - \alpha)(A_{jk} - \Omega_{jk}) \\ & + \sum_{ijk(dist)} (\Omega_{ik} - \alpha)(\Omega_{jk} - \alpha) \equiv T_{n1} + T_{n2} + T_{n3}. \end{aligned}$$

Observe that T_{n3} is non-stochastic. The second and third term are negligible compared to T_{n1}. Define $\overline{\Omega} = \Omega - \alpha \mathbf{1}\mathbf{1}'$. By direct calculations,

$$\mathbb{E} T_{n2} = 0,$$

and

$$\operatorname{Var}(T_{n2}) = 8 \sum_{j<k(dist)} \big(\sum_{i \notin \{j,k\}} \overline{\Omega}_{ik} \big)^2 \Omega_{jk}(1-\Omega_{jk}) = 8 \sum_{j<k(dist)} \big(\overline{\Omega}_{jk} + \overline{\Omega}_{kk}\big)^2 \Omega_{jk}(1-\Omega_{jk}) \lesssim \alpha n^2.$$

Next,

$$\begin{aligned} |T_{n3}| &= \big| \sum_{ijk} \overline{\Omega}_{ik}\overline{\Omega}_{jk} - \sum_{ijk(not\, dist.)} \overline{\Omega}_{ik}\overline{\Omega}_{jk} \big| = \big| \sum_{ijk(not\, dist.)} \overline{\Omega}_{ik}\overline{\Omega}_{jk} \big| \\ &\lesssim \big| \sum_{ij} \overline{\Omega}_{ii}\overline{\Omega}_{ji} \big| + \big| \sum_{ik} \overline{\Omega}_{ik}^2 \big| + \big| \sum_{i} \overline{\Omega}_{ii}^2 \big| = 0 + o(\alpha n^{3/2}) + n = o(\alpha n^{3/2}), \end{aligned}$$

where we apply the third regularity condition.

Now we focus on T_{n1}. By direct calculations

$$\mathbb{E} T_{n1} = 0,$$

and

$$\begin{aligned} \operatorname{Var} T_{n1} &= 2 \sum_{i,j,k(dist)} \Omega_{ik}(1-\Omega_{ik})\Omega_{jk}(1-\Omega_{jk}) \\ &= 2 \sum_{i,j,k} \Omega_{ik}(1-\Omega_{ik})\Omega_{jk}(1-\Omega_{jk}) - 2 \sum_{i,j,k(not\, dist.)} \Omega_{ik}(1-\Omega_{ik})\Omega_{jk}(1-\Omega_{jk}) \\ &= 2\mathbf{1}'\Omega^2\mathbf{1} - 2 \sum_{i,j,k(not\, dist.)} \Omega_{ik}(1-\Omega_{ik})\Omega_{jk}(1-\Omega_{jk}) \end{aligned}$$

Note that

$$2\mathbf{1}'\Omega^2\mathbf{1} \sim 2n(n-1)(n-2)\alpha^2$$

since $\alpha \to 0$. Moreover, with some simple casework we can show

$$\sum_{i,j,k(not\,dist.)} \Omega_{ik}(1-\Omega_{ik})\Omega_{jk}(1-\Omega_{jk}) \lesssim \alpha n^2 = o(\alpha^2 n^3),$$

where we use that $\alpha n \to \infty$ (because $\alpha^2 n \to \infty$). Hence

$$\operatorname{Var} T_{n1} \sim 2n(n-1)(n-2)\alpha^2(1-\alpha)^2 \sim 2n(n-1)(n-2)\alpha^2(1-\alpha)^2.$$

To study T_{n1} we apply the martingale central limit theorem using a similar argument to Cammarata & Ke (2022)). Define $W_{ij} = A_{ij} - \Omega_{ij}$ and

$$T_{n,m} = \sum_{(i,j,k)\in I_m} W_{ik}W_{jk}, \qquad \text{and} \qquad T_{n,0} = 0,$$

$$Z_{n,m} = \sqrt{\frac{n-1}{2n(n-2)}}\frac{T_{n,m}}{(n-1)\alpha_n(1-\alpha_n)}, \qquad \text{and} \qquad Z_{n,0} = 0.$$

where

$$I_m = \{(i,j,k) \in [m]^3 \text{ s.t. } i,j,k \text{ are distinct}\},$$

and $m \le n$. Define a filtration $\{\mathcal{F}_{n,m}\}$ where $\mathcal{F}_{n,m} = \sigma\{W_{ij}, (i,j) \in [m]^2\}$ for all $m \in [n]$, and let $\mathcal{F}_{n,0}$ be the trivial σ-field. It is straightforward to verify that $T_{n,m}$ and $Z_{n,m}$ are martingales with respect to this filtration. We further define a martingale difference sequence

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

for all $m \in [n]$.

If we can show that the following conditions hold

$$\text{(a) } \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 | \mathcal{F}_{n,m-1}] \xrightarrow{\mathbb{P}} 1, \tag{F.3}$$

$$\text{(b) } \forall \epsilon > 0, \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 \mathbf{1}\{|X_{n,m} > \epsilon|\} | \mathcal{F}_{n,m-1}] \xrightarrow{\mathbb{P}} 0, \tag{F.4}$$

then the Martingale Central Limit Theorem implies that $Z_n \Rightarrow \mathcal{N}(0,1)$.

Our argument follows closely Cammarata & Ke (2022). First consider (F.3). It suffices to show that

$$\mathbb{E}\left[\sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 | \mathcal{F}_{n,m-1}]\right] \xrightarrow{n\to\infty} 1, \tag{F.5}$$

and

$$\operatorname{Var}\left(\sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 | \mathcal{F}_{n,m-1}]\right) \xrightarrow{n\to\infty} 0. \tag{F.6}$$

For notational brevity, define

$$C_n := (n-1)\alpha_n(1-\alpha_n)\sqrt{\frac{2n(n-2)}{n-1}}.$$

Mimicking the argument in (Cammarata & Ke, 2022, pgs.33-34) shows the following. Note that all sums below are indexed up to $m-1$.

$$\begin{aligned} \mathbb{E}[C_n^2 X_{n,m}^2 | \mathcal{F}_{n,m-1}] &= 4\sum_{k\neq j;\; i\neq l} W_{jk}W_{il}\mathbb{E}\left[W_{mk}W_{mi}\right] + 4\sum_{k\neq j;\; i\neq l} W_{jk}\mathbb{E}\left[W_{im}W_{km}W_{lm}\right] \\ &+ \sum_{i\neq j;\; k\neq l} \mathbb{E}\left[W_{im}W_{jm}W_{km}W_{lm}\right]. \end{aligned} \tag{F.7}$$

Continuing, we have

$$\mathbb{E}[C_n^2 X_{n,m}^2 | \mathcal{F}_{n,m-1}] = 4\sum_{i}\sum_{j\neq i, l\neq i} W_{ij}W_{il}\Omega_{mi}(1-\Omega_{mi}) + 2\sum_{i,j(dist)} \Omega_{im}(1-\Omega_{im})\Omega_{jm}(1-\Omega_{jm})$$

$$
\begin{aligned}
&= 4 \sum_{ij\ell(dist)} W_{ij} W_{il} \Omega_{mi}(1-\Omega_{mi}) + 4 \sum_{i,j(dist)} W_{ij}^2 \Omega_{mi}(1-\Omega_{mi}) \\
&\quad + 2 \sum_{i,j(dist)} \Omega_{im}(1-\Omega_{im})\Omega_{jm}(1-\Omega_{jm}).
\end{aligned} \tag{F.8}
$$

Computing expectations,

$$
\begin{aligned}
&\mathbb{E}[\mathbb{E}[C_n^2 X_{n,m}^2 | \mathcal{F}_{n,m-1}]] \\
&\qquad = 4 \sum_{i,j(dist)} \Omega_{ij}(1-\Omega_{ij})\Omega_{mi}(1-\Omega_{mi}) + 2 \sum_{i,j(dist)} \Omega_{im}(1-\Omega_{im})\Omega_{jm}(1-\Omega_{jm})
\end{aligned}
$$

Summing over m and a simple combinatorial argument yields

$$
C_n^2 \mathbb{E}\big[\sum_{m=1}^n \mathbb{E}[X_{n,m}^2 | \mathcal{F}_{n,m-1}]\big] = 2 \sum_{i,j,k(dist)} \Omega_{ik}(1-\Omega_{ik})\Omega_{jk}(1-\Omega_{jk}) \sim C_n^2.
$$

Using the identity

$$
W_{ij}^2 = (1-2\Omega_{ij})W_{ij} + \Omega_{ij}(1-\Omega_{ij}),
$$

we have

$$
\begin{aligned}
\mathbb{E}[C_n^2 X_{n,m}^2 | \mathcal{F}_{n,m-1}] &= 4 \sum_{ij\ell(dist)} W_{ij} W_{il} \Omega_{mi}(1-\Omega_{mi}) + 4 \sum_{i,j(dist)} W_{ij}^2 \Omega_{mi}(1-\Omega_{mi}) \\
&= 24 \sum_{i<j<\ell} W_{ij} W_{il} \Omega_{mi}(1-\Omega_{mi}) + 8 \sum_{i<j} W_{ij}(1-2\Omega_{ij})\Omega_{mi}(1-\Omega_{mi}) \\
&\quad + 4 \sum_{i<j} \Omega_{ij}(1-\Omega_{ij})\Omega_{mi}(1-\Omega_{mi}).
\end{aligned}
$$

Thus

$$
\begin{aligned}
\sum_{m=1}^n \mathbb{E}[C_n^2 X_{n,m}^2 | \mathcal{F}_{n,m-1}] &= 24 \sum_{i<j<\ell} \big(\sum_{m>\max(i,j,\ell)} \Omega_{mi}(1-\Omega_{mi}) \big) W_{ij} W_{i\ell} \\
&\quad + 8 \sum_{i<j} \big(\sum_{m>\max(i,j,\ell)} \Omega_{mi}(1-\Omega_{mi}) \big) (1-2\Omega_{ij}) W_{ij}.
\end{aligned}
$$

All terms above are uncorrelated. Hence,

$$
\begin{aligned}
\mathrm{Var}\left(\sum_{m=1}^n \mathbb{E}[C_n^2 X_{n,m}^2 | \mathcal{F}_{n,m-1}] \right) &= 24^2 \sum_{i<j<\ell} \big(\sum_{m>\max(i,j,\ell)} \Omega_{mi}(1-\Omega_{mi}) \big)^2 \Omega_{ij}(1-\Omega_{ij})\Omega_{i\ell}(1-\Omega_{i\ell}) \\
&\quad + 64 \sum_{i<j} \big(\sum_{m>\max(i,j,\ell)} \Omega_{mi}(1-\Omega_{mi}) \big)^2 (1-2\Omega_{ij})^2 \Omega_{ij}(1-\Omega_{ij}) \\
&\lesssim n^2 \cdot C_n^2,
\end{aligned}
$$

whence,

$$
\mathrm{Var}\left(\sum_{m=1}^n \mathbb{E}[X_{n,m}^2 | \mathcal{F}_{n,m-1}] \right) \lesssim \frac{n^2}{C_n^2} \asymp \frac{n^2}{\alpha^2 n^3} \to 0
$$

since $\alpha^2 n \to \infty$. Thus we have shown (F.5) and (F.6), which together prove (F.3).

Next we prove (F.4), again following the argument in Cammarata & Ke (2022). In (Cammarata & Ke, 2022, pg.36) it is shown that it suffices to prove

$$
\sum_{m=1}^n \mathbb{E}[X_{n,m}^4] \xrightarrow{n\to\infty} 0. \tag{F.9}
$$

Further in (Cammarata & Ke, 2022, pg.37), it is shown that

$$\begin{aligned}\mathbb{E}[C_n^4 X_{n,m}^4] =16\Bigg[&\sum_{i<j}\mathbb{E}[W_{jm}^4]\mathbb{E}[(W_{ij}+W_{im})^4]\\ &+3\sum_{\substack{i<j,u<v\\ i\neq u,j\neq v}}\mathbb{E}[W_{jm}^2]\,\mathbb{E}[(W_{ij}+W_{im})^2]\,\mathbb{E}[W_{vm}^2]\,\mathbb{E}[(W_{uv}+W_{um})^2]\\ &+3\sum_{\substack{i<j,v\\ j\neq v}}\mathbb{E}[W_{jm}^2]\,\mathbb{E}[W_{vm}^2]\,\mathbb{E}[(W_{ij}+W_{im})^2(W_{iv}+W_{im})^2]\\ &+3\sum_{\substack{i,u<j\\ i\neq u}}\mathbb{E}[(W_{ij}+W_{im})^2]\,\mathbb{E}[(W_{uj}+W_{um})^2]\,\mathbb{E}[W_{jm}^4]\Bigg].\end{aligned}$$

Going through term by term, we have for n sufficiently large

$$\sum_{i<j}\mathbb{E}[W_{jm}^4]\mathbb{E}[(W_{ij}+W_{im})^4]\lesssim\sum_{i,j}\Omega_{jm}\big(\Omega_{ij}+\Omega_{im}\big)\lesssim\alpha^2n^2$$

Next

$$\begin{aligned}\sum_{\substack{i<j,u<v\\ i\neq u,j\neq v}}\mathbb{E}[W_{jm}^2]\,\mathbb{E}[(W_{ij}+W_{im})^2]\,\mathbb{E}[W_{vm}^2]\,\mathbb{E}[(W_{ij}+W_{im})^2]&\lesssim\sum_{ijuv}\Omega_{jm}(\Omega_{ij}+\Omega_{jm})\Omega_{vm}(\Omega_{uv}+\Omega_{um})\\ &=\sum_{ijuv}\Omega_{jm}\Omega_{ij}\Omega_{vm}\Omega_{uv}+\sum_{ijuv}\Omega_{jm}\Omega_{ij}\Omega_{vm}\Omega_{um}+\sum_{ijuv}\Omega_{jm}^2\Omega_{vm}\Omega_{uv}\\ &\quad+\sum_{ijuv}\Omega_{jm}^2\Omega_{vm}\Omega_{um}\\ &\lesssim\alpha^4n^4+\alpha^3n^3\end{aligned}$$

With a similar argument, we also have, for n sufficiently large,

$$\sum_{\substack{i<j,v\\ j\neq v}}\mathbb{E}[W_{jm}^2]\,\mathbb{E}[W_{vm}^2]\,\mathbb{E}[(W_{ij}+W_{im})^2(W_{iv}+W_{im})^2]\lesssim\alpha^2n^2+\alpha^3n^3$$

$$\sum_{\substack{i,u<j\\ i\neq u}}\mathbb{E}[(W_{ij}+W_{im})^2]\,\mathbb{E}[(W_{uj}+W_{um})^2]\,\mathbb{E}[W_{jm}^4]\Bigg]\lesssim\alpha^3n^3+\alpha^2n^2.$$

Thus

$$\sum_{m=1}^{n}\mathbb{E}[X_{n,m}^4]\lesssim\frac{\alpha^4n^5}{C_n^4}\sim\frac{\alpha^4n^5}{\alpha^4n^6}\to 0,$$

which verifies (F.9). Since (F.9) implies (F.4), this completes the proof. □

G Proof of Theorem 2.4 (Statistical lower bound)

Let $f_0(A)$ be the density under the null hypothesis. Let $\mu(\Pi)$ be the density of Π, and let $f_1(A|\Pi)$ be the conditional density of A given Π. The L_1 distance between two hypotheses is

$$\ell^*\equiv\frac{1}{2}\mathbb{E}_{A\sim f_0}\big|\mathbb{E}_{\Pi\sim\mu}L(A,\Pi)-1\big|,\qquad L(A,\Pi)=f_1(A|\Pi)/f_0(A).$$

Define

$$\mathcal{M}=\big\{\Pi:\Pi\text{ is an eligible membership matrix and }\textstyle\sum_i\pi_i(1)\leq 2n\epsilon\big\}.\tag{G.1}$$

Write $L^{\mathcal{M}}(A,\Pi)=L(A,\Pi)\cdot 1\{\Pi\in\mathcal{M}\}$ and define $L^{\mathcal{M}^c}(A,\Pi)$ similarly. By direct calculations, we have

$$\ell^*=\frac{1}{2}\mathbb{E}_{A\sim f_0}\big|\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}}(A,\Pi)-1+\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}^c}(A,\Pi)\big|$$

$$
\begin{aligned}
&\leq \frac{1}{2}\mathbb{E}_{A\sim f_0}\big|\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}}(A,\Pi)-1\big| + \frac{1}{2}\mathbb{E}_{A\sim f_0}\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}^c}(A,\Pi) \\
&\equiv \frac{1}{2}\ell_0 + \frac{1}{2}\ell_1. \qquad \text{(G.2)}
\end{aligned}
$$

Note that $\mathbb{E}_{A\sim f_0}\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}^c}(A,\Pi) = \int_{\Pi\in\mathcal{M}^c} f_1(A|\Pi)\mu(\Pi)d\Pi dA = \int_{\Pi\in\mathcal{M}^c}\mu(\Pi)d\Pi = \mu(\mathcal{M}^c)$. We bound the probability of $\mu \in \mathcal{M}^c$. Note that $\pi_i(1)$ are independent Bernoulli variables with mean ϵ, where $\epsilon \asymp n^{-1}N$. It follows by Bernstein inequality that if $t = 100\sqrt{N\log N}$, the we have conservatively,

$$
\mathbb{P}\Big(\Big|\sum_i \pi_i(1) - N\Big| > t\Big) \leq 2\exp\Big(-\frac{t^2/2}{n\varepsilon + t/3}\Big) \leq 2\exp\Big(-\frac{100^2 N(\log N)/2}{200N}\Big) \lesssim N^{-c} = o(1) \qquad \text{(G.3)}
$$

for some $c > 0$. It follows that

$$
\ell_1 = \mu(\mathcal{M}^c) = o(1). \qquad \text{(G.4)}
$$

By Cauchy-Schwarz inequality,

$$
\begin{aligned}
\ell_0^2 &\leq \mathbb{E}_{A\sim f_0}\big|\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}}(A,\Pi)-1\big|^2 \\
&= \mathbb{E}_{A\sim f_0}\big(\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}}(A,\Pi)\big)^2 - 2\mathbb{E}_{A\sim f_0}\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}}(A,\Pi) + 1 \\
&= \mathbb{E}_{A\sim f_0}\big(\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}}(A,\Pi)\big)^2 - 2\big[1 - \mathbb{E}_{A\sim f_0}\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}^c}(A,\Pi)\big] + 1 \\
&\leq \mathbb{E}_{A\sim f_0}\big(\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}}(A,\Pi)\big)^2 - 1 + o(1),
\end{aligned}
$$

where the third line is from $\mathbb{E}_{A\sim f_0}\mathbb{E}_{\Pi\sim\mu}L(A,\Pi) = 1$ and the last line is from (G.4). We plug it into (G.2) to get

$$
\ell^* \leq \sqrt{\ell_2 - 1} + o(1), \qquad \text{where} \quad \ell_2 \equiv \mathbb{E}_{A\sim f_0}\big(\mathbb{E}_{\Pi\sim\mu}L^{\mathcal{M}}(A,\Pi)\big)^2. \qquad \text{(G.5)}
$$

It suffices to prove that $\ell_2 \leq 1 + o(1)$.

Below, we study ℓ_2. Let $\widetilde{\Pi}$ be an independent copy of Π. Define

$$
S(A,\Pi,\widetilde{\Pi}) = L(A,\Pi)\cdot L(\widetilde{\Pi},A).
$$

It is easy to see that

$$
\ell_2 = \mathbb{E}_{A\sim f_0,\Pi,\tilde{\Pi}\sim\mu}\big[S(A,\Pi,\widetilde{\Pi})\cdot 1\{\Pi\in\mathcal{M},\widetilde{\Pi}\in\mathcal{M}\}\big]. \qquad \text{(G.6)}
$$

Denote by p_{ij} and $q_{ij}(\Pi)$ the values of Ω_{ij} under the null and the alternative, respectively. Write $\delta_{ij}(\Pi) = (q_{ij}(\Pi) - p_{ij})/p_{ij}$. By definition,

$$
S(A,\Pi,\widetilde{\Pi}) = \prod_{i<j}\left[\frac{q_{ij}(\Pi)q_{ij}(\widetilde{\Pi})}{p_{ij}^2}\right]^{A_{ij}}\left[\frac{(1-q_{ij}(\Pi))(1-q_{ij}(\widetilde{\Pi}))}{(1-p_{ij})^2}\right]^{1-A_{ij}}.
$$

Write for short $q_{ij}(\Pi) = q_{ij}$, $q_{ij}(\widetilde{\Pi}) = \tilde{q}_{ij}$, $\delta_{ij}(\Pi) = \delta_{ij}$ and $\delta_{ij}(\widetilde{\Pi}) = \tilde{\delta}_{ij}$. By straightforward calculations, we have the following claims:

$$
\mathbb{E}_{A\sim f_0}[S(A,\Pi,\widetilde{\Pi})] = \prod_{i<j}\Big(1 + \frac{p_{ij}\delta_{ij}\tilde{\delta}_{ij}}{1-p_{ij}}\Big), \qquad \text{(G.7)}
$$

and

$$
\begin{aligned}
\ln S(A,\Pi,\widetilde{\Pi}) = \sum_{i<j} A_{ij}\ln&\Big[\frac{(1+\delta_{ij})(1+\tilde{\delta}_{ij})}{(1-\frac{p_{ij}}{1-p_{ij}}\delta_{ij})(1-\frac{p_{ij}}{1-p_{ij}}\tilde{\delta}_{ij})}\Big] \\
&+ \ln\Big[\Big(1 - \frac{p_{ij}}{1-p_{ij}}\delta_{ij}\Big)\Big(1 - \frac{p_{ij}}{1-p_{ij}}\tilde{\delta}_{ij}\Big)\Big]. \qquad \text{(G.8)}
\end{aligned}
$$

The expression (G.8) may be useful for the case of $Nc \to 0$. In the current case of $Nc \to \infty$, we use (G.7). It follows from (G.6) that

$$
\ell_2 = \mathbb{E}_{\Pi,\tilde{\Pi}\sim\mu}\Big[\prod_{i<j}\Big(1 + \frac{p_{ij}\delta_{ij}\tilde{\delta}_{ij}}{1-p_{ij}}\Big)\cdot 1\{\Pi\in\mathcal{M},\widetilde{\Pi}\in\mathcal{M}\}\Big]
$$

$$
\begin{aligned}
&= \mathbb{E}_{\Pi,\tilde{\Pi}\sim\mu}\Big[\exp\Big(\sum_{i<j}\ln\Big(1+\frac{p_{ij}\delta_{ij}\tilde{\delta}_{ij}}{1-p_{ij}}\Big)\Big)\cdot 1\{\Pi\in\mathcal{M},\widetilde{\Pi}\in\mathcal{M}\}\Big] \\
&\leq \mathbb{E}_{\Pi,\tilde{\Pi}\sim\mu}\Big[\exp(X)\cdot 1\{\Pi\in\mathcal{M},\widetilde{\Pi}\in\mathcal{M}\}\Big], \quad \text{with } X\equiv\sum_{i<j}\frac{p_{ij}\delta_{ij}\tilde{\delta}_{ij}}{1-p_{ij}}. \qquad \text{(G.9)}
\end{aligned}
$$

where the last line is from the universal inequality of $\ln(1+t)\leq t$.

We further work out the explicit expressions of p_{ij}, δ_{ij} and $\tilde{\delta}_{ij}$. Let $h=(\epsilon,1-\epsilon)'$, and recall that $\alpha_0=a\epsilon+b(1-\epsilon)$. The condition of b in (H.1) guarantees that

$$
Ph=\alpha_0\mathbf{1}_2, \qquad \alpha_0=a\epsilon+b(1-\epsilon).
$$

By direct calculations,

$$
\alpha_0=\frac{c(1-\epsilon)^2-a\epsilon^2}{1-2\epsilon}. \qquad \text{(G.10)}
$$

It follows that

$$
P=\alpha_0\mathbf{1}_2\mathbf{1}_2'+M, \qquad \text{where} \quad M=\frac{a-c}{1-2\epsilon}\xi\xi', \quad \xi=(1-\epsilon,-\epsilon)'. \qquad \text{(G.11)}
$$

Write $z_i=\pi_i-h$. Since $Ph=\alpha_0\mathbf{1}_2$ and $z_i'\mathbf{1}_2=0$, we have

$$
\begin{aligned}
\Omega_{ij} &= \theta_j\theta_j(h+z_i)'P(h+z_i) \\
&= \theta_i\theta_j(h'Ph+z_i'Pz_j) \\
&= \theta_i\theta_j(\alpha_0+z_i'Pz_j) \\
&= \theta_i\theta_j(\alpha_0+z_i'Mz_j) \\
&= \theta_i\theta_j\Big[\alpha_0+\frac{a-c}{1-2\epsilon}(\xi'z_i)(\xi'z_j)\Big].
\end{aligned}
$$

Let t_i be the indicator that node i belongs to the first community and write $u_i=t_i-\frac{N}{n}$. Then, $\pi_i=(t_i,1-t_i)$ and $z_i=u_i(1,-1)'$. It follows that $\xi'z_i=u_i$. Therefore,

$$
\Omega_{ij}=\theta_i\theta_j\Big[\alpha_0+\frac{a-c}{1-2\epsilon}u_iu_j\Big], \qquad \text{where} \quad u_i\overset{iid}{\sim}\text{Bernoulli}(\epsilon)-\epsilon. \qquad \text{(G.12)}
$$

Consequently,

$$
p_{ij}=\alpha_0\theta_i\theta_j, \qquad \delta_{ij}(\Pi)=\frac{a-c}{(1-2\epsilon)\alpha_0}u_iu_j.
$$

We plug it into (G.9) to obtain

$$
X=\sum_{i<j}\frac{\theta_i\theta_j}{1-\alpha_0\theta_i\theta_j}\frac{(a-c)^2}{(1-2\epsilon)^2\alpha_0}u_iu_j\tilde{u}_i\tilde{u}_j. \qquad \text{(G.13)}
$$

Below, we use (G.13) to bound ℓ^2. Since $\alpha_0\theta_{\max}^2=O(c\theta_{\max}^2)=o(1)$, by Taylor expansion of $(1-\alpha_0\theta_i\theta_j)^{-1}$, we have

$$
X=\frac{(a-c)^2}{(1-2\epsilon)^2\alpha_0}\sum_{i<j}\sum_{s=1}^{\infty}\alpha_0^{s-1}\theta_i^s\theta_j^su_iu_j\tilde{u}_i\tilde{u}_j.
$$

Let $b_i=\theta_i\theta_{\max}^{-1}<1$. We re-write X as

$$
X=\gamma\sum_{s=1}^{\infty}w_sX_s,
$$

where

$$
\gamma=\frac{\theta_{\max}^2(a-c)^2}{(1-\alpha_0\theta_{\max}^2)(1-2\epsilon)^2\alpha_0}, \quad w_s=(1-\alpha_0\theta_{\max}^2)\alpha_0^{s-1}\theta_{\max}^{2s-2}, \quad \text{and } X_s=\sum_{i<j}b_i^sb_j^su_iu_j\tilde{u}_i\tilde{u}_j. \qquad \text{(G.14)}
$$

Let $\check{\mathbb{E}}$ be the conditional expectation by conditioning on the event of $\{\Pi \in \mathcal{M}, \widetilde{\Pi} \in \mathcal{M}\}$. It follows from (G.9) that

$$\begin{aligned}
\ell_2 &= \mathbb{P}(\Pi \in \mathcal{M}, \widetilde{\Pi} \in \mathcal{M}) \cdot \check{\mathbb{E}}[\exp(X)] \\
&= \mathbb{P}(\Pi \in \mathcal{M}, \widetilde{\Pi} \in \mathcal{M}) \cdot \check{\mathbb{E}}\Big[\exp\Big(\gamma \sum_{s=1}^{\infty} w_s X_s\Big)\Big] \\
&\leq \mathbb{P}(\Pi \in \mathcal{M}, \widetilde{\Pi} \in \mathcal{M}) \cdot \sum_{s=1}^{\infty} w_s \check{\mathbb{E}}[\exp(\gamma X_s)] \\
&= \sum_{s=1}^{\infty} w_s \, \mathbb{E}\big[\exp(\gamma X_s) \cdot 1\{\Pi \in \mathcal{M}, \widetilde{\Pi} \in \mathcal{M}\}\big]. \qquad \text{(G.15)}
\end{aligned}$$

The third line follows using Jensen's inequality and that $\sum_{s \geq 1} w_s = 1$.

It suffices to bound the term in (G.15) for each $s \geq 1$. Note that

$$X_s \leq Y_s^2, \qquad Y_s = \sum_i b_i^s u_i \tilde{u}_i. \qquad \text{(G.16)}$$

We recall that $u_i = t_i - \epsilon$, where $t_i = \pi_i(1) \in \{0, 1\}$. The event $\{\Pi \in \mathcal{M}, \widetilde{\Pi} \in \mathcal{M}\}$ translates to $\max\{\sum_i t_i, \ \sum_i \tilde{t}_i\} \leq 2n\epsilon$. Note that

$$u_i \tilde{u}_i = \begin{cases} (1-\epsilon)^2, & \text{when } t_i + \tilde{t}_i = 2, \\ -\epsilon(1-\epsilon), & \text{when } t_i + \tilde{t}_i = 1, \\ \epsilon^2, & \text{where } t_i + \tilde{t}_i = 0. \end{cases}$$

It follows that $|u_i \tilde{u}_i| \leq (t_i + \tilde{t}_i)/2 + \epsilon^2$. Note that $\epsilon = O(N/n)$. Therefore, on this event,

$$|Y_s| \leq \sum_i [(t_i + \tilde{t}_i)/2 + \epsilon^2] \leq 2n\epsilon + n\epsilon^2 \leq 3N.$$

We immediately have

$$\mathbb{E}\big[\exp(\gamma X_s) \cdot 1\{\Pi \in \mathcal{M}, \widetilde{\Pi} \in \mathcal{M}\}\big] \leq \mathbb{E}\Bigg[\exp(\gamma Y_s^2) \cdot 1\{|Y_s| \leq 3N\}\Bigg]. \qquad \text{(G.17)}$$

The following lemma is useful.

Lemma G.1. *Let Z be a random variable satisfying that*

$$\mathbb{P}(|Z| > t) \leq 2\exp\Big(-\frac{t^2/2}{\sigma^2 + bt}\Big), \qquad \textit{for all } t > 0.$$

Then, for any $\gamma > 0$ and $B > 0$ such that $\gamma(\sigma^2 + bB) < 1/2$, we have

$$\mathbb{E}\big[\exp(\gamma Z^2) 1\{|Z| \leq B\}\big] \leq 1 + \frac{4\gamma(\sigma^2 + bB)}{1 - 2\gamma(\sigma^2 + bB)}.$$

Note that $Y_s = \sum_i b_i^s u_i \tilde{u}_i$ is a sum of independent, mean-zero variables, where $|b_i^s u_i \tilde{u}_i| \leq 2$ and $\sum_i \mathrm{Var}(b_i^s u_i \tilde{u}_i) \leq \sum_i b_i^{2s} 2\epsilon^2 \leq 2n\epsilon^2$. It follows from Bernstein's inequality that

$$\mathbb{P}(|Y_s| > t) \leq \exp\Bigg(-\frac{t^2/2}{2n\epsilon^2 + 2t}\Bigg), \qquad \text{for all } t > 0.$$

To apply Lemma G.1, we set

$$b = 2, \qquad \sigma^2 = 2n\epsilon^2 \leq 2n^{-1}N^2, \qquad Z = Y_s, \qquad B = 3N,$$

and γ as in (G.14). The choice of B is in light of (G.17). Furthermore, by (G.10), we have $\alpha_0 \asymp c$. Also we have $\theta_{\max}^2 \alpha_0 \to 0$. Hence,

$$\gamma = \frac{\theta_{\max}^2 (a-c)^2}{(1 - \alpha_0 \theta_{\max}^2)(1 - 2\epsilon)^2 \alpha_0} \leq C \cdot \big(\frac{\theta_{\max}^2 (a-c)^2}{c}\big).$$

Thus by the hypothesis $\frac{\theta_{\max}^2 N(a-c)^2}{c} \to 0$, it holds that $\gamma(\sigma^2 + bB) < 1/2$ for n sufficiently large. Applying Lemma G.1, we obtain

$$\begin{aligned}\mathbb{E}\big[\exp(\gamma X_s)\cdot 1\{\Pi \in \mathcal{M}, \widetilde{\Pi} \in \mathcal{M}\}\big] &\le 1 + C(\gamma(\sigma^2 + bB)) \\ &\le 1 + C\cdot\big(\frac{\theta_{\max}^2 N(a-c)^2}{c}\big)\end{aligned}$$

We further plug it into (G.15) to get

$$\ell_2 \le \sum_{s=1}^{\infty} w_s\Big[1 + C\cdot\big(\frac{\theta_{\max}^2 N(a-c)^2}{c}\big)\Big] \le 1 + \big(\frac{\theta_{\max}^2 N(a-c)^2}{c}\big),$$

where we use that $\sum w_s = 1$.

It follows immediately that

$$\ell_2 \le 1 + o(1), \qquad \text{if} \quad \theta_{\max}\frac{\sqrt{N}(a-c)}{\sqrt{c}} \to 0.$$

This proves the claim. □

G.1 Proof of Lemma G.1

Let X denote a nonnegative random variable, and define $\overline{F}(x) = \mathbb{P}_X[X \ge x]$. For any positive number $\beta > 0$, we have

$$\begin{aligned}\mathbb{E}[\exp(\gamma X)1\{X < \beta\}] &= \int_0^\beta e^{\gamma x}\,\mathrm{d}\mathbb{P}_X(x) \\ &= -e^{\gamma x}\bar{F}(x)\Big|_0^\beta + \int_0^\beta \gamma e^{\gamma x}\bar{F}(x)dx \\ &= 1 - e^{\gamma\beta}\bar{F}(\beta) + \int_0^\beta \gamma e^{\gamma x}\bar{F}(x)dx \\ &\le 1 + \int_0^\beta \gamma e^{\gamma x}\bar{F}(x)dx.\end{aligned}$$

We apply it to $X = Z^2$ and $\beta = B^2$ to get

$$\begin{aligned}\mathbb{E}\big[\exp(\gamma Z^2)1\{|Z| \le B\}\big] &\le 1 + \int_0^{B^2} \gamma\exp(\gamma x)\mathbb{P}(|Z| > \sqrt{x})dx \\ &\le 1 + 2\gamma\int_0^{B^2} \exp(\gamma x)\exp\Big\{-\frac{x}{2(\sigma^2 + b\sqrt{x})}\Big\}dx \\ &\le 1 + 2\gamma\int_0^{B^2} \exp(\gamma x)\exp\Big\{-\frac{x}{2(\sigma^2 + bB)}\Big\}dx \\ &\le 1 + 2\gamma\int_0^{\infty} \exp\Big\{-\frac{1 - 2\gamma(\sigma^2 + bB)}{2(\sigma^2 + bB)}x\Big\}dx \\ &\le 1 + \frac{4\gamma(\sigma^2 + bB)}{1 - 2\gamma(\sigma^2 + bB)}.\end{aligned}$$

This proves the claim. □

H Proof of Theorem 2.5 (Tightness of the statistical lower bound)

Let $\rho \in \mathbb{R}^n$. We consider the global testing problem in the DCBM model where

A) $P = \begin{pmatrix} 1 & b \\ b & 1 \end{pmatrix}$

B) $b = \tilde{b}/\sqrt{ac}$,

C) $\theta_i = \rho_i \sqrt{a}$ for $i \in S$,

D) $\theta_i = \rho_i \sqrt{c}$ for $i \notin S$, and

E) $aN_0 + \tilde{b}(n - N_0) = \tilde{b}N_0 + c(n - N_0)$,

Recall that $h = (N_0/n, 1 - N_0/n)^\mathsf{T}$, and N_0 is the size of the smaller community in the alternative. Observe that the null model $K = 1$ is parameterized by setting $a = c = \tilde{b} = 1$.

Recall that $\varepsilon = N/n$. We define

$$\alpha_0 \equiv \frac{aN_0 + \tilde{b}(n - N_0)}{n}.$$

Note that by Assumption (*E*),

$$\tilde{b} = \frac{nc - (a + c)N_0}{n - 2N_0} \tag{H.1}$$

$$a\epsilon = O(c), \text{ and} \tag{H.2}$$

$$c \sim \tilde{b} \sim \alpha_0. \tag{H.3}$$

Our assumptions in this section are the following:

a) There exists an absolute constant $C_\rho > 0$ such that $\rho_{\max} \le C_\rho\, \rho_{\min}$

b) $\frac{\rho_{\max}^2 \alpha_0 n}{\sqrt{\log n}} \to \infty$

c) An integer N is known such that $N_0 = N[1 + o(1)]$.

Note that since we tolerate a small error in the clique size by Assumption (*c*), our setting indeed matches that of the statistical lower bound, by (G.3).

Define the signed scan statistic

$$\phi_{sc} = \max_{D \subset [n]: |D| = N} \mathbf{1}_D' \big(A - \hat{\eta}\hat{\eta}^\mathsf{T}\big) \mathbf{1}_D. \tag{H.4}$$

For notational brevity, define $n^{(2)} = \binom{n}{2}$. Let

$$\hat{\gamma} = \frac{1}{n^{(2)}} \sum_{i,j} A_{ij}.$$

The estimator $\hat{\gamma}$ provides a constant factor approximation of the edge density of the least-favorable null model. See Lemma H.1 for further details.

Next let

$$h(u) = (1 + u)\log(1 + u) - u, \tag{H.5}$$

and note that this function is strictly increasing on $\mathbb{R}_{\ge 0}$. Define a random threshold $\hat{\tau}$ to be

$$\hat{\tau} = C^* \hat{\gamma} N^2 h^{-1}\left(\frac{C^* N \log(\frac{ne}{N})}{\hat{\gamma} N^2}\right) \tag{H.6}$$

Let $C^* > 0$ denote a sufficiently large constant, to be determined, that depends only on C_ρ from Assumption (*a*). Finally define the scan test to be

$$\varphi_{sc} = \mathbf{1}\big[|\phi_{sc}| > \hat{\tau}\big]$$

Note that, if we assume $a \ge c$, as in the main text, then $b < 1$. In this case, we can simply take

$$\varphi_{sc} = \mathbf{1}\big[\phi_{sc} > \hat{\tau}\big],$$

and the same guarantees hold. On the other hand, if $b > 1$, then the scan test skews negative, as our proof shows.

Theorem H.1. *If*

$$h\Big(\frac{\|\theta_S\|_1^2|1-b^2|}{\rho_{\max}^2\alpha_0 N_0^2}\Big) \gg \frac{\log\frac{ne}{N_0}}{\rho_{\max}^2\alpha_0 N_0}, \tag{H.7}$$

then the type 1 and 2 error of φ_{sc} tend to 0 as $n\to\infty$.

We interpret the previous result in the following concrete settings.

Corollary H.1. *If*

$$\frac{\rho_{\max}^2\alpha_0 N_0}{\log\frac{ne}{N_0}} \to 0,$$

then φ_{sc} has type 1 and 2 errors tending to 0 as $n\to\infty$, provided that

$$\frac{\rho_{\max}^2 N_0(a-c)}{\log\frac{ne}{N_0}} \gg 1.$$

If

$$\frac{\rho_{\max}^2\alpha_0 N_0}{\log\frac{ne}{N_0}} \to \infty,$$

then φ_{sc} has type 1 and 2 errors tending to 0 as $n\to\infty$, provided that

$$\frac{\rho_{\max}^2 N_0(a-c)}{\sqrt{\rho_{\max}^2 N_0\alpha_0\log\frac{ne}{N_0}}} \gg 1.$$

Proof. Note that

$$\|\theta_S\|_1^2|1-b^2| = \rho_{\max}^2 N_0^2(a-\tilde{b}^2/\sqrt{c}) \sim \rho_{\max}^2 N_0^2(a-c).$$

In the first case,

$$h\Big(\frac{\|\theta_S\|_1^2|1-b^2|}{\rho_{\max}^2\alpha_0 N_0^2}\Big) \gg h\Big(\frac{\log\frac{ne}{N_0}}{\rho_{\max}^2\alpha_0 N_0}\Big) \gtrsim \frac{\log\frac{ne}{N_0}}{\rho_{\max}^2\alpha_0 N_0}.$$

We use the fact that $h(u)\gtrsim u$ for $u\geq 1$.

In the second case,

$$h\Big(\frac{\|\theta_S\|_1^2|1-b^2|}{\rho_{\max}^2\alpha_0 N_0^2}\Big) \gg h\Big(\frac{N_0\cdot\sqrt{\rho_{\max}^2 N_0\alpha_0\log\frac{ne}{N_0}}}{\rho_{\max}^2\alpha_0 N_0^2}\Big) = h\Big(\sqrt{\frac{\log\frac{ne}{N_0}}{\rho_{\max}^2\alpha_0 N_0}}\Big) \gtrsim \frac{\log\frac{ne}{N_0}}{\rho_{\max}^2\alpha_0 N_0}.$$

□

The upper bounds in the second part of Corollary H.1 is the best possible up to logarithmic factors. For example, suppose that $\theta_{\max}\lesssim\theta_{\min}$ in Theorem 2.4. Then the upper bound for the second case of Corollary H.1 matches the lower bound of Theorem 2.4 up to logarithmic factors.

To prove Theorem 2.5, first we establish concentration of $\hat{\gamma}$.

Lemma H.1. *Recall*

$$\hat{\gamma} = \frac{1}{n^{(2)}}\sum_{i,j(dist)} A_{ij}.$$

There exists an absolute constant $C>0$ such that for all $\delta>0$, it holds that

$$|\hat{\gamma}-\mathbb{E}\hat{\gamma}| \leq \frac{C\sqrt{\rho_{\max}^2\alpha_0\log(1/\delta)}}{n}$$

with probability at least $1-\delta$.

Proof. As a preliminary, we claim that

$$(\Omega \mathbf{1})_i \asymp \rho_{\max}^2 \alpha_0 n. \tag{H.8}$$

To see this, note that if $i \in S$, then by (*E*)

$$\begin{aligned}(\Omega \mathbf{1})_i &= \sum_j \Omega_{ij} = \theta_i(\|\theta_S\|_1 + b\|\theta_{S^c}\|_1) \\ &\asymp \rho_{\max}\sqrt{a} \cdot \big(\sqrt{a} N \rho_{\max} + \frac{\tilde{b}}{\sqrt{ac}} \cdot \sqrt{c}\rho_{\max}\big) = \rho_{\max}^2 \alpha_0 n.\end{aligned}$$

The claim for $i \notin S$ follows by a similar argument applying (*E*). It follows that

$$v_0 = \mathbf{1}^\mathsf{T} \Omega \mathbf{1} \asymp \rho_{\max}^2 \alpha_0 n^2$$

The expectation is

$$\mathbb{E}\hat{\gamma} = \frac{1}{n^{(2)}} \sum_{i,j(dist)} \Omega_{ij},$$

and the variance is

$$\mathrm{Var}(\hat{\gamma}) = \frac{1}{(n^{(2)})^2} \sum_{i,j(dist)} \Omega_{ij}(1 - \Omega_{ij}).$$

By Bernstein's inequality,

$$\mathbb{P}\big[n^{(2)}\big|\hat{\gamma} - \mathbb{E}\hat{\gamma}\big| > t\big] \le 2\exp\bigg(-\frac{ct^2}{\sum_{i,j(dist)} \Omega_{ij} + t}\bigg). \tag{H.9}$$

By Assumptions (*a*) and (*b*),

$$\sum_{i,j(dist)} \Omega_{ij} \asymp \rho_{\max}^2 \alpha_0 n^2 \gg n.$$

Setting

$$t = \tau \equiv C\sqrt{\rho_{\max}^2 \alpha_0 n^2 \log(1/\delta)}$$

for a large enough absolute constant $C > 0$, (H.9) implies that

$$|\hat{\gamma} - \mathbb{E}\hat{\gamma}| \le \frac{\tau}{n^2} \asymp \frac{\sqrt{\rho_{\max}^2 \alpha_0 \log(1/\delta)}}{n}$$

with probability at least $1 - \delta$. □

Next we control the error arising from the plug-in effect of approximating η^* by $\hat{\eta}$.

Lemma H.2. *Given $D \subset [n]$, define*

$$L_D \equiv \mathbf{1}_D^\mathsf{T}(\eta^* \eta^{*\mathsf{T}} - \hat{\eta}\hat{\eta}^\mathsf{T})\mathbf{1}_D.$$

Then under the null and alternative hypothesis,

$$\max_{|D|=N} |L_D| \lesssim \sqrt{N_0^3 \rho_{\max}^2 \alpha_0 \log(\frac{ne}{N_0})}$$

with probability at least $1 - \binom{n}{N}^{-1} - 2v_0^{-c_1}$, for an absolute constant $c_1 > 0$.

Proof. In this proof, $c > 0$ is an absolute constant that may vary from line to line.

Given $D \subset [n]$, let

$$L_D \equiv \mathbf{1}_D^\mathsf{T}(\eta^* \eta^{*\mathsf{T}} - \hat{\eta}\hat{\eta}^\mathsf{T})\mathbf{1}_D = \mathbf{1}_D^\mathsf{T}\eta^*(\eta^* - \hat{\eta})^\mathsf{T}\mathbf{1}_D + \mathbf{1}_D^\mathsf{T}(\eta^* - \hat{\eta})\hat{\eta}^\mathsf{T}\mathbf{1}_D \tag{H.10}$$

Our first goal is to control

$$\big|\mathbf{1}_D^{\mathsf{T}}(\hat{\eta}-\eta^*)\big|.$$

Define $\overline{\Omega} = \Omega - \mathrm{diag}(\Omega)$. Note that

$$\hat{\eta}-\eta^* = \frac{A\mathbf{1}}{\sqrt{V}} - \frac{\Omega\mathbf{1}}{\sqrt{v_0}} = \big(\frac{A\mathbf{1}}{\sqrt{V}} - \frac{A\mathbf{1}}{\sqrt{v_0}}\big) + \big(\frac{A\mathbf{1}}{\sqrt{v_0}} - \frac{\overline{\Omega}\mathbf{1}}{\sqrt{v_0}}\big) + \big(\frac{\overline{\Omega}\mathbf{1}}{\sqrt{v_0}} - \frac{\Omega\mathbf{1}}{\sqrt{v_0}}\big) \tag{H.11}$$

We study each term of (H.11). First note that

$$(\overline{\Omega}\mathbf{1})_i = (\Omega\mathbf{1})_i - \Omega_{ii} = \rho_{\max}^2\alpha_0 n + O(1),$$

and thus

$$v_0 = \sum_i(\Omega\mathbf{1})_i \sim \sum_i(\overline{\Omega}\mathbf{1})_i = v, \text{ and}$$
$$|v_0 - v| \lesssim 1 \tag{H.12}$$

Next note that

$$\mathrm{Var}\big(\mathbf{1}_D^{\mathsf{T}}\big(A\mathbf{1} - \overline{\Omega}\mathbf{1}\big)\big) \lesssim \sum_{\substack{i\in[n],j\in D\\ i\neq j}} \Omega_{ij} \lesssim |D|\rho_{\max}^2\alpha_0 n.$$

By Bernstein's inequality,

$$\mathbb{P}\big[\big|\mathbf{1}_D^{\mathsf{T}}\big(A\mathbf{1} - \overline{\Omega}\mathbf{1}\big)\big| \geq t\big] \leq 2\exp\left(-\frac{ct^2}{|D|\rho_{\max}^2\alpha_0 n + t}\right) \tag{H.13}$$

for all $t > 0$. Setting

$$t = \tau \equiv \sqrt{4/c}\cdot\sqrt{|D|\rho_{\max}^2\alpha_0 n\log(1/\delta)},$$

we have

$$\frac{1}{\sqrt{v_0}}\big|\mathbf{1}_D^{\mathsf{T}}\big(A\mathbf{1} - \overline{\Omega}\mathbf{1}\big)\big| \lesssim \frac{\sqrt{|D|\rho_{\max}^2\alpha_0 n\log(1/\delta)}}{\sqrt{\rho_{\max}^2\alpha_0 n^2}} = \sqrt{(|D|/n)\cdot\log(1/\delta)} \tag{H.14}$$

with probability at least $1-\delta$.

Next, it is shown in (Jin et al., 2021, Supplement, pg.100) that for $\sqrt{\log\|\theta\|_1} \ll x_n \ll \|\theta\|_1$,

$$\mathbb{P}\big[|V - v| > x_n\|\theta\|_1\big] = \mathbb{P}\left[|\sqrt{V} - \sqrt{v}| > \frac{x_n\|\theta\|_1}{\sqrt{V}+\sqrt{v}}\right] \leq 2\exp(-cx_n^2).$$

Hence

$$\mathbb{P}\left[|\sqrt{V} - \sqrt{v}| > \frac{x_n\|\theta\|_1}{\sqrt{v}}\right] \leq 2\exp(-cx_n^2),$$

Note that by (H.2) and (H.3),

$$\frac{\|\theta\|_1}{\sqrt{v}} \asymp \frac{N_0\rho_{\max}\sqrt{a} + (n-N_0)\rho_{\max}\sqrt{c}}{\rho_{\max}\sqrt{\alpha_0}n} \asymp 1.$$

By (H.12), we have

$$\mathbb{P}\left[|\sqrt{V} - \sqrt{v_0}| > \frac{x_n\|\theta\|_1}{\sqrt{v}}\right] \leq 2\exp(-cx_n^2). \tag{H.15}$$

Hence with probability at least $1 - 2\exp(-cx_n^2)$,

$$V \gtrsim v_0.$$

It follows that

$$\mathbb{P}\left[\big|\frac{1}{\sqrt{V}} - \frac{1}{\sqrt{v_0}}\big| \geq \frac{x_n\|\theta\|_1}{v_0\sqrt{v}}\right] = \mathbb{P}\left[\frac{|\sqrt{V}-\sqrt{v_0}|}{\sqrt{V\cdot v_0}} \geq \frac{x_n\|\theta\|_1}{v_0\sqrt{v}}\right] \leq 2\exp(-cx_n^2).$$

Hence with probability at least $1 - \delta - 2\exp(-cx_n^2)$,

$$\left|\mathbf{1}_D^\mathsf{T}(\frac{A\mathbf{1}}{\sqrt{V}} - \frac{A\mathbf{1}}{\sqrt{v_0}})\right| \leq \frac{x_n \cdot \big(|D|\rho_{\max}^2\alpha_0 n + \sqrt{|D|\rho_{\max}^2\alpha_0 n\log(1/\delta)}\big)}{v_0}$$
$$\asymp \frac{x_n \cdot \big(|D|\rho_{\max}^2\alpha_0 n + \sqrt{|D|\rho_{\max}^2\alpha_0 n\log(1/\delta)}\big)}{\rho_{\max}^2\alpha_0 n^2}. \tag{H.16}$$

For the last term of (H.11),

$$\mathbf{1}_D^\mathsf{T}(\frac{\overline{\Omega}\mathbf{1}}{\sqrt{v_0}} - \frac{\Omega\mathbf{1}}{\sqrt{v_0}}) = \frac{\sum_{i\in D}\Omega_{ii}}{\sqrt{v_0}} \asymp \frac{\rho_{\max}^2 a|D\cap S| + \rho_{\max}^2 c|D\cap S^c|}{\sqrt{\rho_{\max}^2\alpha_0 n^2}}$$
$$\lesssim \rho_{\max} a\epsilon/\sqrt{\alpha_0} \lesssim \rho_{\max}\sqrt{c} \lesssim 1. \tag{H.17}$$

Next we control $\mathbf{1}_D^\mathsf{T}\hat{\eta}$. By (H.13) and (H.15),

$$|\mathbf{1}_D^\mathsf{T}\hat{\eta}| = \frac{|\mathbf{1}_D^\mathsf{T}A\mathbf{1}|}{\sqrt{V}} \lesssim \frac{|D|\rho_{\max}^2\alpha_0 n + \sqrt{|D|\rho_{\max}^2\alpha_0 n\log(1/\delta)}}{\sqrt{v_0} - cx_n} \tag{H.18}$$

with probability at least $1 - \delta - 2\exp(-cx_n^2)$. It also holds that

$$|\mathbf{1}_D^\mathsf{T}\eta^*| = \frac{|\mathbf{1}_D^\mathsf{T}\Omega\mathbf{1}|}{\sqrt{v_0}} = \frac{|D|\rho_{\max}^2\alpha_0 n}{\rho_{\max}\sqrt{\alpha_0}n} = |D|\rho_{\max}\sqrt{\alpha_0}. \tag{H.19}$$

Next we set $x_n = \sqrt{\log\|\theta\|_1} \asymp \sqrt{\log v_0}$. Then from (H.16) and (H.18),

$$\left|\mathbf{1}_D^\mathsf{T}(\frac{A\mathbf{1}}{\sqrt{V}} - \frac{A\mathbf{1}}{\sqrt{v_0}})\right| \asymp \frac{\sqrt{\log v_0} \cdot \big(|D|\rho_{\max}^2\alpha_0 n + \sqrt{|D|\rho_{\max}^2\alpha_0 n\log(1/\delta)}\big)}{\rho_{\max}^2\alpha_0 n^2}$$
$$\asymp \sqrt{\log v_0} \cdot \big((|D|/n) + \frac{\sqrt{(|D|/n)\log(1/\delta)}}{\rho_{\max}\sqrt{\alpha_0}n}\big), \tag{H.20}$$

and

$$|\mathbf{1}_D^\mathsf{T}\hat{\eta}| \lesssim \frac{|D|\rho_{\max}^2\alpha_0 n + \sqrt{|D|\rho_{\max}^2\alpha_0 n\log(1/\delta)}}{\sqrt{v_0}}$$
$$\asymp \frac{|D|\rho_{\max}^2\alpha_0 n + \sqrt{|D|\rho_{\max}^2\alpha_0 n\log(1/\delta)}}{\rho_{\max}\sqrt{\alpha_0}n}$$
$$\asymp |D|\rho_{\max}\sqrt{\alpha_0} + \sqrt{(|D|/n)\cdot\log(1/\delta)} \tag{H.21}$$

with probability at least $1 - \delta - 2v_0^{-c_1}$.

By (H.14),(H.17), (H.19), (H.20), and (H.21)

$$|L_D| \leq \big|\mathbf{1}_D^\mathsf{T}\eta^*(\eta^* - \hat{\eta})^\mathsf{T}\mathbf{1}_D\big| + \big|\mathbf{1}_D^\mathsf{T}(\eta^* - \hat{\eta})\hat{\eta}^\mathsf{T}\mathbf{1}_D\big|$$
$$\lesssim \big(|D|\rho_{\max}\sqrt{\alpha_0} + \sqrt{(|D|/n)\cdot\log(1/\delta)}\big) \cdot \big(\sqrt{\log v_0}(|D|/n) + \sqrt{(|D|/n)\log(1/\delta)} + 1\big).$$

with probability at least $1 - \delta - 2v_0^{-c_1}$.

It follows that, setting $\delta = 1/\binom{n}{N}^2$ above and applying the union bound,

$$\max_{|D|=N}|L_D| \lesssim \big(N\rho_{\max}\sqrt{\alpha_0} + \sqrt{N\epsilon\cdot\log(\frac{ne}{N})}\,\big) \cdot \big(\epsilon\sqrt{\log v_0} + \sqrt{N\epsilon\cdot\log(\frac{ne}{N})} + 1\big)$$

with probability at least $1 - \binom{n}{N}^{-1} - 2v_0^{-c_1} \to 1$. Note that

$$\frac{n\log\frac{ne}{N}}{\log v_0} \asymp \frac{n\log\frac{ne}{N}}{\log(\rho_{\max}^2\alpha_0 n^2)} \gtrsim 1 \Rightarrow$$

$$\frac{N^2}{n}\log\frac{ne}{N} \gtrsim \frac{N^2}{n^2}\log(\rho_{\max}^2\alpha_0 n^2) \Rightarrow$$
$$\sqrt{N\epsilon\cdot\log(\frac{ne}{N})} \gtrsim \epsilon\sqrt{\log v_0}.$$

Further, since $(N/n)\log\frac{ne}{N} \ll 1$ and $\rho_{\max}^2\alpha_0 n \to \infty$ by Assumption (*b*),

$$N\log\frac{ne}{N} \lesssim \rho_{\max}^2\alpha_0 n^2 \Rightarrow$$
$$\frac{N}{n}\sqrt{\log\frac{ne}{N}} \lesssim \sqrt{N\rho_{\max}^2\alpha_0} \Rightarrow$$
$$N\epsilon\log\frac{ne}{N} \lesssim \sqrt{N^3\rho_{\max}^2\alpha_0\log\frac{ne}{N}}.$$

Hence

$$\max_{|D|=N}|L_D| \lesssim \sqrt{N^3\rho_{\max}^2\alpha_0\log(\frac{ne}{N})} + N\epsilon\log(\frac{ne}{N}) \lesssim \sqrt{N^3\rho_{\max}^2\alpha_0\log(\frac{ne}{N})}$$

with probability at least $1-\binom{n}{N}^{-1}-2v_0^{-c_1}$. Recalling that $N=N_0[1+o(1)]$ yields the statement of the lemma.

□

Next we study an ideal version of ϕ_{sc}.

Lemma H.3. *Define the ideal scan statistic*

$$\tilde{\phi}_{sc} = \max_{|D|=N}\mathbf{1}_D^{\mathsf{T}}(A-\eta^*\eta^{*\mathsf{T}})\mathbf{1}_D,$$

and corresponding test

$$\tilde{\varphi}_{sc} = \mathbf{1}\Big[\tilde{\phi}_{sc} > \tilde{\tau}\Big],$$

where

$$\tilde{\tau} \equiv \tilde{C}\hat{\gamma}N^2 h^{-1}\bigg(\frac{\tilde{C}N\log(\frac{ne}{N})}{\hat{\gamma}N^2}\bigg),$$

and $\tilde{C}>0$ is a sufficiently large absolute constant that depends only on C_ρ from Assumption (a). Then under the null hypothesis,

$$\mathbb{P}\big[|\tilde{\phi}_{sc}| > \tilde{\tau}\big] \le n^{-c_0} + \exp\big(-N\log\frac{ne}{N}\big)$$

and under the alternative hypothesis,

$$\mathbb{P}\big[|\tilde{\phi}_{sc}| \le \tilde{\tau}\big] \le n^{-c_0} + \big(\frac{N}{ne}\big)^{10}$$

for n sufficiently large, where c_0 is an absolute constant.

Proof. In this proof, $c>0$ is an absolute constant that may vary form line to line.

Define the ideal scan statistic

$$\tilde{\phi}_{sc} = \max_{|D|=N}\mathbf{1}_D^{\mathsf{T}}(A-\eta^*\eta^{*\mathsf{T}})\mathbf{1}_D.$$

Also define

$$Z_D \equiv \sum_{i,j\in D(dist)}(A_{ij}-\Omega_{ij})$$

First consider the type 1 error. Under the null hypothesis, we have $\eta^*=\theta=\rho$ and $\alpha_0=1$. Observe that

$$\sigma_D^2 \equiv \mathrm{Var}(Z_D) = \mathrm{Var}\Big(\sum_{i,j\in D(dist)}(A_{ij}-\theta_i\theta_j)\Big) \lesssim \|\theta_D\|_1^2 \asymp \rho_{\max}^2 N^2 \sim \rho_{\max}^2 N_0^2$$

By the Bennett inequality, (Vershynin, 2018, Theorem 2.9.2),

$$\mathbb{P}\Big[\sum_{i,j\in D}(A_{ij}-\theta_i\theta_j)>t\Big]\le\exp\bigg(-\sigma_D^2\,h\Big(\frac{t}{\sigma_D^2}\Big)\bigg), \tag{H.22}$$

where $h(u)=(1+u)\log(1+u)-u$.

Next, by Lemma H.1,

$$|\hat\gamma-\mathbb{E}\hat\gamma|\lesssim\frac{\sqrt{\log n}}{n}$$

with probability n^{-c_0}. Also recall that

$$\mathbb{E}\,\hat\gamma=\frac{1}{n^{(2)}}\sum_{i,j(dist)}\Omega_{ij}\asymp\rho_{\max}^2\alpha_0=\rho_{\max}^2\gg\frac{\sqrt{\log n}}{n}$$

by Assumptions (*a*) and (*b*). It follows that there exist absolute constants $c_0,c_\gamma,C_\gamma>0$ such that

$$c_\gamma\rho_{\max}^2<\hat\gamma<C_\gamma\rho_{\max}^2 \tag{H.23}$$

with probability n^{-c_0}. Let $\mathcal{E}$ denote this event. Under $\mathcal{E}$, we have that for $\tilde C$ sufficiently large,

$$\tilde C\hat\gamma N^2h^{-1}\bigg(\frac{\tilde CN\log(\frac{ne}{N})}{\hat\gamma N^2}\bigg)\ge\sigma_D^2h^{-1}\bigg(\frac{2N\log\frac{ne}{N}}{\sigma_D^2}\bigg)$$

It follows from this, the union bound, and the Bennett inequality,

$$\begin{aligned}\mathbb{P}\bigg[|\tilde\phi_{sc}|>\tilde C\hat\gamma N^2h^{-1}\bigg(\frac{\tilde CN\log(\frac{ne}{N})}{\hat\gamma N^2}\bigg)\bigg]&\le\mathbb{P}[\mathcal{E}^c]+\mathbb{P}\bigg[|\tilde\phi_{sc}|>\tilde C\hat\gamma N^2h^{-1}\bigg(\frac{\tilde CN\log(\frac{ne}{N})}{\hat\gamma N^2}\bigg),\ \mathcal{E}\bigg]\\&\le n^{-c_0}+\sum_{|D|=N}\mathbb{P}\bigg[|Z_D|>\tilde C\hat\gamma N^2h^{-1}\bigg(\frac{\tilde CN\log(\frac{ne}{N})}{\hat\gamma N^2}\bigg)\bigg]\\&\le n^{-c_0}+\sum_{|D|=N}\mathbb{P}\bigg[|Z_D|>\sigma_D^2h^{-1}\bigg(\frac{2N\log\frac{ne}{N}}{\sigma_D^2}\bigg)\bigg]\\&\le n^{-c_0}+\big(\frac{ne}{N}\big)^N\exp\big(-2N\log\frac{ne}{N}\big).\end{aligned}$$

This shows that the type 1 error for the ideal scan statistic is $o(1)$.

Next consider the type 2 error. We have by Lemma (E.2),

$$\mathbf{1}_S^{\mathsf T}(A-\eta^*\eta^{*\mathsf T})\mathbf{1}_S=\sum_{i,j\in S(dist)}(A_{ij}-\Omega_{ij})+\mathbf{1}_S^{\mathsf T}\tilde\Omega\mathbf{1}_S=Z_S+\|\theta_S\|_1^2(1-b^2)\cdot\frac{\|\theta_{S^c}\|_1^2}{v_0}.$$

Note that by (H.12)

$$\|\theta_S\|_1^2(1-b^2)\cdot\frac{\|\theta_{S^c}\|_1^2}{v_0}\sim\|\theta_S\|_1^2(1-b^2).$$

Next,

$$\mathrm{Var}(Z_S)=\sum_{i,j\in S(dist)}\Omega_{ij}(1-\Omega_{ij})\lesssim\|\theta_S\|_1^2\asymp\rho_{\max}^2Na\sim\rho_{\max}^2N_0a$$

By Bernstein's inequality,

$$|Z_S|\lesssim\sqrt{\|\theta_S\|_1^2\log(1/\delta)}\vee\log(1/\delta)\le\|\theta_S\|_1\log(1/\delta)$$

with probability at least $1-\delta$. Setting $\delta=(\frac{N}{ne})^{10}$, we have

$$|Z_S|\lesssim\|\theta_S\|_1\log\big(\frac{ne}{N}\big)$$

with probability at least $1-(\frac{N}{ne})^{10}$.

Next we show that

$$\|\theta_S\|_1|1-b^2| \gtrsim \log\frac{ne}{N} \tag{H.24}$$

using (H.7), which we rewrite as

$$\|\theta_S\|_1^2|1-b^2| \gg \gamma N_0^2 h^{-1}\left(\frac{\log\frac{ne}{N_0}}{\gamma N_0}\right) \sim \gamma N^2 h^{-1}\left(\frac{\log\frac{ne}{N}}{\gamma N}\right) \tag{H.25}$$

where $\gamma=\rho_{\max}^2\alpha_0$. Recall that $\alpha_0=1$ under the null, and $\alpha_0\sim c$ under the alternative. Let

$$u=\frac{\log\frac{ne}{N}}{\gamma N}.$$

Consider two cases: (i) $u\le 0.01$, and (ii) $u\ge 0.01$. For $u'\le h^{-1}(0.01)$, we have $h(u')\asymp(u')^2$, and therefore $h^{-1}(u)\asymp u^2$ for $u\le 0.01$. In this case (H.25) implies

$$\|\theta_S\|_1^2|1-b^2| \gg \gamma N^2\sqrt{\frac{\log\frac{ne}{N}}{\gamma N}}=\sqrt{\gamma N^3\log\frac{ne}{N}}.$$

In addition,

$$\|\theta_S\|_1=N\sqrt{a}\rho_{\max},$$

so that

$$\|\theta_S\|_1(1-b^2)\gg\sqrt{\frac{\gamma N\log\frac{ne}{N}}{a\rho_{\max}^2}}\gtrsim\log\frac{ne}{N}$$

since $u\le 0.01$ and $a\rho_{\max}^2\lesssim 1$. Thus in case (i), (H.24) is satisfied for n sufficiently large.

Now consider case (ii) where $u\ge 0.01$. Note that $h(u)\le(u+1)\log(u+1)$, and thus

$$\frac{1}{2}(u+1)\le u\le h^{-1}((u+1)\log(u+1)).$$

Let $\varphi\equiv(u+1)\log(u+1)\ge u$ and observe that

$$u+1=\frac{\varphi}{\log(u+1)}\ge\frac{\varphi}{\log\varphi}.$$

Hence

$$h^{-1}((u+1)\log(u+1))\ge\frac{1}{2}\cdot\frac{(u+1)\log(u+1)}{\log\left[(u+1)\log(u+1)\right]}.$$

Applying (H.25),

$$\|\theta_S\|_1^2|1-b^2|\gg\gamma N^2\cdot\frac{(\frac{\log\frac{ne}{N}}{\gamma N}+1)\log(\frac{\log\frac{ne}{N}}{\gamma N}+1)}{\log\left[(\frac{\log\frac{ne}{N}}{\gamma N}+1)\log(\frac{\log\frac{ne}{N}}{\gamma N}+1)\right]}\gtrsim N\log\frac{ne}{N}.$$

Hence

$$\|\theta_S\|_1|1-b^2|\gg\frac{\log\frac{ne}{N}}{\sqrt{a}\rho_{\max}}\gtrsim\log\frac{ne}{N}.$$

Thus in case (ii), (H.24) is also satisfied.

Next we have,

$$\begin{aligned}\mathbb{P}&\left[|\tilde{\phi}_{sc}|\le\tilde{C}\hat{\gamma}N^2h^{-1}\left(\frac{\tilde{C}N\log(\frac{ne}{N})}{\hat{\gamma}N^2}\right)\right]\\&\qquad\le n^{-c_0}+\mathbb{P}\left[|\tilde{\phi}_{sc}|\le\tilde{C}\hat{\gamma}N^2h^{-1}\left(\frac{\tilde{C}N\log(\frac{ne}{N})}{\hat{\gamma}N^2}\right),\,\mathcal{E}\right]\end{aligned}$$

$$\leq n^{-c_0} + \mathbb{P}\Big[\,\Big|\|\theta_S\|_1^2(1-b^2) + Z_S\Big| \leq C\gamma N^2 h^{-1}\Big(\frac{CN\log(\frac{ne}{N})}{\gamma N^2}\Big)\,\Big]$$
$$\leq n^{-c_0} + \mathbb{P}\Big[|Z_S| \geq \big|\|\theta_S\|_1^2(1-b^2)\big| - C\gamma N^2 h^{-1}\Big(\frac{CN\log(\frac{ne}{N})}{\gamma N^2}\Big)\Big],$$

where $C > 0$ is a sufficiently large absolute constant. In the second line and third lines we use the event $\mathcal{E}$ from (H.23), and in the last line we use the triangle inequality. By (H.7), we have conservatively that

$$\big|\,\|\theta_S\|_1^2(1-b^2)\big| - C\gamma N^2 h^{-1}\Big(\frac{CN\log(\frac{ne}{N})}{\gamma N^2}\Big) \geq \frac{1}{2}\big|\|\theta_S\|_1^2(1-b^2)\big| \gg \|\theta_S\|_1 \log\frac{ne}{N}$$

for n sufficiently large. Thus for n sufficiently large,

$$\mathbb{P}\Big[|\tilde{\phi}_{sc}| \leq \tilde{C}\hat{\gamma}N^2 h^{-1}\Big(\frac{\tilde{C}N\log(\frac{ne}{N})}{\hat{\gamma}N^2}\Big)\Big] \leq n^{-c_0} + \mathbb{P}\Big[|Z_S| \geq \frac{1}{2}\big|\,\|\theta_S\|_1^2(1-b^2)\big|\,\Big]$$
$$\leq n^{-c_0} + \big(\frac{N}{ne}\big)^{10}.$$

Therefore the type 2 error for the ideal scan statistic is also $o(1)$. $\square$

Lemma H.4. *Let ϕ_{sc} denote the scan statistic defined in* (H.4)*, and let $\hat{\tau}$ denote the random threshold defined in* (H.6)*. Then under the null hypothesis,*

$$\mathbb{P}\big[|\phi_{sc}| > \hat{\tau}\big] \leq \binom{n}{N}^{-1} + v_0^{-c_1} + n^{-c_0} + \exp\big(-N\log\frac{ne}{N}\big),$$

and under the alternative hypothesis,for n sufficiently large we have

$$\mathbb{P}\big[|\phi_{sc}| < \hat{\tau}\big] \leq \binom{n}{N}^{-1} + v_0^{-c_1} + n^{-c_0} + \big(\frac{N}{ne}\big)^{10}.$$

Proof. We show that the plug-in effect is negligible compared to the threshold and signal-strength.

By Lemma H.2,

$$\max_{|D|=N}|L_D| \lesssim \sqrt{N_0^3\gamma\log(\frac{ne}{N_0})}$$

with high probability. Since $h(u) \leq u^2$ for $u \geq 0$, it follows that

$$h\Big(\frac{\sqrt{N_0^3\gamma\log(\frac{ne}{N_0})}}{\gamma N_0^2}\Big) \leq \frac{N_0^3\gamma\log(\frac{ne}{N_0})}{\gamma^2 N_0^4} = \frac{\log\frac{ne}{N_0}}{\gamma N_0} \Rightarrow$$
$$\sqrt{N_0^3\gamma\log(\frac{ne}{N_0})} \leq \gamma N_0^2 h^{-1}\Big(\frac{\log\frac{ne}{N_0}}{\gamma N_0}\Big) \Rightarrow$$
$$\sqrt{N^3\gamma\log(\frac{ne}{N})} \leq [1+o(1)]\gamma N^2 h^{-1}\Big(\frac{\log\frac{ne}{N}}{\gamma N}\Big).$$

Under the null, we have by Lemma H.3 that

$$\mathbb{P}\big[|\phi_{sc}| \geq \hat{\tau}\big] \leq \mathbb{P}\big[\,|\tilde{\phi}_{sc}| \geq \hat{\tau} - \max_{|D|=N}|L_D|\,\big]$$
$$\leq \binom{n}{N}^{-1} + v_0^{-c_1} + \mathbb{P}\Big[|\tilde{\phi}_{sc}| \geq C^*\hat{\gamma}N^2 h^{-1}\Big(\frac{C^* N\log(\frac{ne}{N})}{\hat{\gamma}N^2}\Big) - \gamma N^2 h^{-1}\Big(\frac{\log\frac{ne}{N}}{\gamma N}\Big)\Big]$$
$$\leq \binom{n}{N}^{-1} + v_0^{-c_1} + n^{-c_0} + \exp\big(-N\log\frac{ne}{N}\big)$$

for $C^* > 0$ a sufficiently large absolute constant. It suffices to take $C^* \geq 2\tilde{C}$.

Under the alternative hypothesis, we have by Lemma H.3 that

$$\begin{aligned}
\mathbb{P}\big[|\phi_{sc}| \le \hat{\tau}\big] &\le \mathbb{P}\big[\,|\tilde{\phi}_{sc}| \le \hat{\tau} + \max_{|D|=N} |L_D|\,\big] \\
&\le \binom{n}{N}^{-1} + v_0^{-c_1} + \mathbb{P}\bigg[|\tilde{\phi}_{sc}| \le C^*\hat{\gamma}N^2h^{-1}\Big(\frac{C^*N\log(\frac{ne}{N})}{\hat{\gamma}N^2}\Big) + \gamma N^2 h^{-1}\Big(\frac{\log\frac{ne}{N}}{\gamma N}\Big)\bigg] \\
&\le \binom{n}{N}^{-1} + v_0^{-c_1} + \mathbb{P}\bigg[|\tilde{\phi}_{sc}| \le 2C^*\hat{\gamma}N^2h^{-1}\Big(\frac{C^*N\log(\frac{ne}{N})}{\hat{\gamma}N^2}\Big)\bigg] \\
&\le \binom{n}{N}^{-1} + v_0^{-c_1} + n^{-c_0} + \big(\frac{N}{ne}\big)^{10}
\end{aligned}$$

for n sufficiently large.

□

Observe that Theorem 2.5 follows directly from Lemma H.4.

I Proof of Theorem 2.6 (Computational lower bound)

In this section, we provide the proof of Theorem 2.6. For convenience, we denote $b = \frac{nc-(a+c)N}{n-2N}$, $d = \frac{c(n-N)^2-aN^2}{n(n-2N)}$. Under H_0, all upper triangular entries A are i.i.d. Bernoulli distributed with probability d. Then an orthonormal basis of the adjacency matrix of graph D is

$$f_\Gamma(A) = \prod_{i<j:(i,j)\in\Gamma} \frac{A_{ij}-d}{\sqrt{d(1-d)}}.$$

Here, $\Gamma \subseteq \{(i,j) : 1 \le i < j \le n\}$ takes all subsets of all upper triagonal entries of A. Denote $|\Gamma|$ as the cardinality of Γ and $B(D) = \{\Gamma \subseteq \{\text{unordered pairs } (i,j) : i \ne j, i,j \in [n]\}, \Gamma \ne \emptyset, |\Gamma| \le D\}$ as all subsets of off-diagonal entries of A of cardinality at most D. By Proposition I.1 and the property of the orthonormal basis function of A,

$$\begin{aligned}
&\sup_{\substack{f \text{ is polynomial; degree}(f)\le D \\ \mathbb{E}_{H_0}f(A)=0;\mathrm{Var}_{H_0}(A)=1}} \mathbb{E}_{H_1} f(A) = \|LR^{\le D} - 1\| \\
&= \left\{\sum_{\Gamma\in B(D)} \big(\mathbb{E}_{H_0} f_\Gamma(A)(LR^{\le D}(A)-1)\big)^2\right\}^{1/2} \overset{(*)}{=} \left\{\sum_{\Gamma\in B(D)} \big(\mathbb{E}_{H_0} f_\Gamma(A) LR(A)\big)^2\right\}^{1/2} \\
&= \left\{\sum_{\Gamma\in B(D)} \mathbb{E}_{H_1}\left(f_\Gamma(A)\right)^2\right\}^{1/2} = \left\{\sum_{\Gamma\in B(D)} \left(\mathbb{E}_{H_1} \prod_{(i,j)\in\Gamma} \frac{A_{ij}-d}{\sqrt{d(1-d)}}\right)^2\right\}^{1/2}.
\end{aligned}$$

Here, $(*)$ is due to $\mathbb{E}_{H_0} f_\Gamma LR^{\le D} = \mathbb{E}_{H_0} f_\Gamma LR$ by the property of projection and $\mathbb{E}_{H_0} f_\Gamma(A) = 0$ for any $\Gamma \in B(D)$. Therefore, to establish the desired computational lower bound, we only need to prove

$$\sum_{\Gamma\in B(D)} \left(\mathbb{E}_{H_1} \prod_{(i,j)\in\Gamma} \frac{A_{ij}-d}{\sqrt{d(1-d)}}\right)^2 = o(1)$$

under the described asymptotic regime. For convenience, we denote

$$p_1 = \frac{a-d}{\sqrt{d(1-d)}}, \quad p_2 = \frac{b-d}{\sqrt{d(1-d)}}, \quad p_3 = \frac{c-d}{\sqrt{d(1-d)}}.$$

We can calculate that

$$a-d = \frac{(n-N)^2(a-c)}{n(n-2N)}, \quad b-d = -\frac{(n-N)N(a-c)}{n(n-2N)}, \quad c-d = \frac{N^2(a-c)}{n(2-2N)}.$$

and

$$c - d = -\frac{N}{n-N}(b-d) = \left(\frac{N}{n-N}\right)^2 (a-d). \tag{I.1}$$

Since $b = \frac{c(n-N)-aN}{n-2N} \geq 0$ and $N \leq n/3$, we know $a \leq c(n-N)/N$ and

$$c \geq d = \frac{c(n-N)^2 - aN^2}{n(n-2N)} \geq \frac{c(n-N)^2 - N(n-N)c}{n(n-2N)} \geq (n-N)/n \cdot c \geq 2/3 \cdot c.$$

Under the asymptotic regime of this theorem, we have $d = \frac{c(n-N)^2 - aN^2}{n(n-2N)}$ and

$$p_1 = \frac{(n-N)^2(a-c)}{n(n-2N)\sqrt{d(1-d)}} \asymp \frac{a-c}{\sqrt{c}}, \tag{I.2}$$

i.e., there exists constant $\delta > 1$ such that $\delta^{-1}c \leq p_1 \leq \delta c$. By (I.1), we have $p_3 = -N/(n-N)p_2 = N^2/(n-N)^2 p_1$. For any fixed $\Gamma \subseteq \{(i,j) : 1 \leq i < j \leq n\}$,

$$\begin{aligned}
&\mathbb{E}_{H_1} \prod_{(i,j)\in\Gamma} \frac{A_{ij} - d}{\sqrt{d(1-d)}} = \mathbb{E}_{\Pi} \left\{ \mathbb{E}\left\{ \prod_{(i,j)\in\Gamma} \frac{A_{ij}-d}{\sqrt{d(1-d)}} \middle| \ A \text{ has two communities assigned by } \Pi \right\}\right\} \\
=&\mathbb{E}_{\Pi} p_1^{|\Gamma \cap K\otimes K|} \cdot p_2^{|\Gamma\cap K\otimes K^c|} \cdot p_3^{|\Gamma \cap K^c \otimes K^c|} = \mathbb{E}_{\Pi} \prod_{(i,j)\in\Gamma} \left\{ p_1 \cdot (-N/(n-N))^{\pi_i + \pi_j - 2} \right\} \\
=&p_1^{|\Gamma|} \cdot \left(\frac{-N}{n-N}\right)^{\sum_{(i,j)\in\Gamma}(\pi_i+\pi_j-2)} = p_1^{|\Gamma|} \cdot \left(\frac{-N}{n-N}\right)^{\sum_{(i,j)\in\Gamma}(\pi_i+\pi_j-2)} \\
=&p_1^{|\Gamma|} \cdot \prod_{i=1}^n \left(\frac{-N}{n-N}\right)^{(\pi_i - 1)\cdot|\{j':(i,j')\in\Gamma\}|} \overset{(a)}{=} p_1^{|\Gamma|} \cdot \prod_{i=1}^n \left\{ \left(\frac{N}{n}\right) + \frac{n-N}{n}\left(\frac{-N}{n-N}\right)^{|\{j':(i,j')\in\Gamma\}|} \right\}.
\end{aligned}$$

Here, (a) is because $\mathbb{P}(\pi_i = 1) = N/n$; $\mathbb{P}(\pi_i = 2) = (n-N)/n$. Thus, the following fact holds: if there exists a node i that appears exactly one time in Γ, i.e., $|\{j' : (i,j') \in \Gamma\}| = 1$, $\mathbb{E}_{H_1} \prod_{(i,j)\in\Gamma} \frac{A_{ij}-d}{\sqrt{d(1-d)}} = 0$. On the other hand, for all Γ that each node appear zero times or at least two times, we have

$$\begin{aligned}
&\mathbb{E}_{H_1} \prod_{(i,j)\in\Gamma} \frac{A_{ij}-d}{\sqrt{d(1-d)}} \leq p_1^{|\Gamma|} \cdot \left\{ \frac{N}{n} + \frac{n-N}{n}\left(\frac{-N}{n-N}\right)^2 \right\}^{|\{i: i \text{ appears at least 2 times in } \Gamma\}|} \\
\leq & p_1^{|\Gamma|} \cdot \left(\frac{2N}{n}\right)^{|\{i: i \text{ appears at least 2 times in } \Gamma\}|}.
\end{aligned}$$

Finally, we denote

$$B_0(D) = \{\Gamma \in B(D) : \text{ each node in } [n] \text{ appears zero time or at least 2 times}\},$$

$$m(\Gamma) = |\{i : i \text{ appears in some pair of } \Gamma\}|.$$

For any $\Gamma \in B_0(D)$, we must have $m(\Gamma) \le |\Gamma| \le m(\Gamma)(m(\Gamma)-1)/2$. Then,

$$
\begin{aligned}
&\sum_{\Gamma \in B(D)} \left(\mathbb{E}_{H_1} \prod_{(i,j)\in\Gamma} \frac{A_{ij}-d}{\sqrt{d(1-d)}} \right)^2 = \sum_{\Gamma \in B_0(D)} \left(\mathbb{E}_{H_1} \prod_{(i,j)\in\Gamma} \frac{A_{ij}-d}{\sqrt{d(1-d)}} \right)^2 \\
=& \sum_{\Gamma \in B_0(D)} p_1^{2|\Gamma|} \cdot \left(\frac{2N}{n}\right)^{2|\{i: i \text{ appears at least 2 times in } \Gamma\}|} \le \sum_{\Gamma \in B_0(D)} p_1^{2|\Gamma|} \cdot \left(\frac{2N}{n}\right)^{2m(\Gamma)} \\
=& \sum_{m=2}^{D} \sum_{g=m}^{D\wedge m(m-1)/2} \sum_{\substack{\Gamma \in B_0(D)\\ m(\Gamma)=m \\ |\Gamma|=g}} p_1^{2g} \left(\frac{2N}{n}\right)^{2m} \overset{(a)}{\le} \sum_{m=2}^{D} \sum_{g=m}^{D\wedge \frac{m(m-1)}{2}} \binom{n}{m} m^g p_1^g \left(\frac{2N}{n}\right)^{m} \\
\le& \sum_{m=2}^{D} \sum_{g=m}^{D\wedge \frac{m(m-1)}{2}} \frac{m^g p_1^{2g} (2N)^{2m}}{m! \cdot n^m} \le \sum_{m=2}^{D} \frac{D \max\left\{(mp_1^2)^m, (mp_1^2)^{D\wedge m(m-1)/2}\right\} \cdot (2N)^{2m}}{n^m} \\
=& D \sum_{m=2}^{D} \left(\frac{\max\{mp_1^2, (mp_1^2)^M\} \cdot (2N)^2}{n} \right)^m \overset{(b)}{=} o(1)
\end{aligned}
$$

Here, $M = \max_{m\ge 1} \frac{D\wedge m(m-1)/2}{m} \le \sqrt{D/2-1}$; (a) is because the number of $\Gamma \in B_0(D)$ with $m(\Gamma) = m$ and $|\Gamma| = g$ is at most $\binom{n}{m} \cdot m^g$; (b) is due to the asymptotic assumption and (I.2), which leads to

$$\frac{N}{\sqrt{n}} \left(p_1 \vee p_1^M\right) \le n^{-\varepsilon}.$$

We have thus finished the proof of this theorem. $\square$

Proposition I.1 (Proposition 1.15 of Kunisky et al. (2019))**.** *Given data A, consider the simple hypothesis testing problem: H_0 versus H_1. Let the likelihood ratio function be $LR(A) = \frac{p_{H_1}(A)}{p_{H_0}(A)}$. Define $\|f\| = \sqrt{\mathbb{E}_{H_0} f^2(A)}$ and $f^{\le D}$ as the projection of any function f to the subspace of polynomials of degree at most D, i.e., $f^{\le D} = \text{argmin}_{\substack{g \text{ is polynomial}\\ \text{degree}(g)\le D}} \|f-g\|$. Then for any positive integer D, we have*

$$\|LR^{\le D}(A) - 1\| = \max_{\substack{f:\text{degree}(f)\le D\\ \mathbb{E}_{H_0} f^2(A)=1 \\ \mathbb{E}_{H_0} f(A)=0}} \mathbb{E}_{H_1} f(A);$$

$$\frac{LR^{\le D}(A)-1}{\|LR^{\le D}(A)-1\|} = \text{argmax}_{\substack{f:\text{degree}(f)\le D\\ \mathbb{E}_{H_0} f^2(A)=1 \\ \mathbb{E}_{H_0} f(A)=0}} \mathbb{E}_{H_1} f(A).$$

J Proof of Theorem 2.7 (Power of EST)

The EST statistic is defined to be

$$\phi_{EST}^{(v)} \equiv \sup_{|S|\le v} \sum_{i,j\in S} A_{ij},$$

and the EST is defined to be

$$\varphi_{EST} = \mathbf{1}\left[\phi_{EST}^{(r)} \ge e\right],$$

where v, e are relatively prime and satisfy

$$\frac{\omega}{1-\beta} < \frac{v}{e} < \delta.$$

Such v and e exist because

$$\frac{\omega}{1-\beta} < \delta,$$

by assumption. Furthermore, we have

$$v < e$$

since $\omega, \delta \in (0, 1)$.

To prove the statement, we require some preliminaries. Let $G(n, p)$ denote an Erdős-Rényi graph with parameter p. A graph H with v vertices and e edges is said to be *balanced* if for all (not necessarily induced) subgraphs $H' \subset H$ with v' vertices and e' edges, it holds that

$$e/v > e'/v'.$$

Next, the power of EST hinges on two well-known facts from probabilistic combinatorics. The first concerns the appearance of an arbitrary graph H in $G(n, p)$.

Theorem J.1 (Adapted from Theorem 4.4.2. of Alon & Spencer (2016))**.** *Let H denote a graph with v vertices and e edges. Then if $p \ll n^{-v/e}$, the random graph $G(n, p)$ does not have H as a subgraph, with high probability as $n \to \infty$.*

On the other hand, if H is balanced and $p \gg n^{-v/e}$, the random graph $G(n, p)$ contains H as a subgraph, with high probability as $n \to \infty$.

Theorem J.2 (Ruciński & Vince (1986); Catlin et al. (1988))**.** *There exists a balanced graph with v vertices and e edges if and only if $1 \le v - 1 \le e \le \binom{v}{2}$.*

Now we continue the proof. Recall that v and e are integers chosen such that $\frac{\omega}{1-\beta} < v/e < \delta$.

Type 1 error: Observe that

$$b = \frac{cn - (a + c)N}{n - 2N} = c \cdot \frac{n - N}{n - 2N} - a \cdot \frac{N}{n - 2N},$$

and thus

$$\begin{aligned}\alpha &= a\varepsilon + b(1 - \varepsilon) = a\varepsilon + (1 - \varepsilon)\big(c \cdot \frac{n - N}{n - 2N} - a \cdot \frac{N}{n - 2N}\big) \\ &= a\left(\frac{N}{n} - (1 - \varepsilon)\frac{N}{n - 2N}\right) + (1 - \varepsilon) \cdot \frac{n - N}{n - 2N} \cdot c = -a \cdot \frac{N^2}{n(n - 2N)} + (1 - \varepsilon) \cdot \frac{n - N}{n - 2N} \cdot c \sim c.,\end{aligned}$$

where above we use that $a\varepsilon \le c$.

Thus under the alternative, A is distributed as Erdős-Rényi with parameter

$$\alpha \sim c = n^{-\delta} \ll n^{-v/e},$$

by our choice of v and e. By the first part of Theorem J.1, no subset of size v of A contains more than e edges, with high probability as $n \to \infty$.

To be more precise, there are a finite number of graphs $H_1, \ldots, H_L$ with v vertices and at least e edges, where L is a constant depending only on v. For each graph H_i, Theorem J.1 contains H_i as a subgraph with probability tending 0 as $n \to \infty$. The type 1 error of EST thus vanishes by the union bound.

Type 2 error: Let H denote a balanced graph on v vertices and e edges, whose existence is guaranteed by Theorem J.2. Consider the induced subgraph on $\mathcal{C}_1$, the smaller community, which is an Erdős-Rényi random graph on N vertices with parameter $a = n^{-\omega}$. By our choice of v and e, we have

$$a = n^{-\omega} = N^{-\frac{\omega}{1-\beta}} \gg N^{-v/e}.$$

By Theorem J.1, $\mathcal{C}_1$ contains a copy of H with high probability. Since H has e edges, we conclude that $\phi_{EST}^{(v)} \ge e$, and thus the null is rejected with high probability as $n \to \infty$.

References

Noga Alon and Joel H Spencer. *The probabilistic method.* John Wiley & Sons, 2016.

Ery Arias-Castro and Nicolas Verzelen. Community detection in dense random networks. *The Annals of Statistics*, 42(3):940–969, 2014.

Louis Cammarata and Zheng Tracy Ke. Power enhancement and phase transitions for global testing of the mixed membership stochastic block model. *arXiv preprint arXiv:2204.11109*, 2022.

Paul A Catlin, Jerrold W Grossman, and Arthur M Hobbs. Graphs with uniform density. *Congr. Numer*, 65:281–285, 1988.

Pengsheng Ji, Jiashun Jin, Zheng Tracy Ke, and Wanshan Li. Co-citation and co-authorship networks for statisticians (with discussions). *J. Bus. Econ. Statist.*, 40(2), 2022.

Jiashun Jin, Zheng Tracy Ke, and Shengming Luo. Optimal adaptivity of signed-polygon statistics for network testing. *The Annals of Statistics*, 49(6):3408–3433, 2021.

Dmitriy Kunisky, Alexander S Wein, and Afonso S Bandeira. Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. *arXiv preprint arXiv:1907.11636*, 2019.

Andrzej Ruciński and Andrew Vince. Strongly balanced graphs and random graphs. *Journal of graph theory*, 10(2):251–264, 1986.

Richard Sinkhorn. Diagonal equivalence to matrices with prescribed row and column sums. *Proc. Amer. Math. Soc.*, (2):195–198, 1974.

Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.