The Annals of Statistics
2021, Vol. 49, No. 6, 3408–3433
https://doi.org/10.1214/21-AOS2089

OPTIMAL ADAPTIVITY OF SIGNED-POLYGON STATISTICS FOR NETWORK TESTING

BY JIASHUN JIN[1,*], ZHENG TRACY KE[2] AND SHENGMING LUO[1,†]

[1]*Department of Statistics & Data Science, Carnegie Mellon University,* [*]*jiashun@stat.cmu.edu;* [†]*shengmil@andrew.cmu.edu*
[2]*Department of Statistics, Harvard University, zke@fas.harvard.edu*

Given a symmetric social network, we are interested in testing whether it has only one community or multiple communities. The desired tests should (a) accommodate severe degree heterogeneity, (b) accommodate mixed memberships, (c) have a tractable null distribution and (d) adapt automatically to different levels of sparsity, and achieve the optimal phase diagram. How to find such a test is a challenging problem.

We propose the Signed Polygon as a class of new tests. Fixing $m \geq 3$, for each m-gon in the network, define a score using the centered adjacency matrix. The sum of such scores is then the mth order Signed Polygon statistic. The Signed Triangle (SgnT) and the Signed Quadrilateral (SgnQ) are special examples of the Signed Polygon.

We show that both the SgnT and SgnQ tests satisfy (a)–(d), and especially, they work well for both very sparse and less sparse networks. Our proposed tests compare favorably with existing tests. For example, the EZ and GC tests behave unsatisfactorily in the less sparse case and do not achieve the optimal phase diagram. Also, many existing tests do not allow for severe heterogeneity or mixed memberships, and they behave unsatisfactorily in our settings.

The analysis of the SgnT and SgnQ tests is delicate and extremely tedious, and the main reason is that we need a unified proof that covers a wide range of sparsity levels and a wide range of degree heterogeneity. For lower bound theory, we use a phase transition framework, which includes the standard minimax argument, but is more informative. The proof uses classical theorems on matrix scaling.

1. Introduction.

Given a symmetrical social network, we are interested in the *global testing problem* where we use the adjacency matrix of the network to test whether it has only one community or multiple communities. A good understanding of the problem is useful for discovering nonobvious social groups and patterns [5, 14], measuring diversity of individual nodes [15], determining stopping time in a recursive community detection scheme [33, 44]. It may also help understand other related problems such as membership estimation [43] and estimation of the number of communities [40, 42].

Natural networks have several characteristics that are ubiquitously found:

- *Severe degree heterogeneity*. The distribution of the node degrees usually has a power-law tail, implying severe degree heterogeneity.
- *Mixed memberships*. Communities are tightly woven clusters of nodes where we have more edges within than between [17, 39]. Communities are rarely nonoverlapping, and some nodes may belong to more than one community (and thus have mixed memberships).
- *Sparsity*. Many networks are sparse. The sparsity levels may range significantly from one network to another, and may also range significantly from one node to another (due to severe degree heterogeneity).

Received May 2019; revised May 2021.
MSC2020 subject classifications. Primary 62H30, 91C20; secondary 62P25.
Key words and phrases. Asymptotic normality, DCBM, DCMM, lower bound, matrix scaling, optimal phase diagram, phase transition, SBM, signed quadrilateral, signed triangle, Sinkhorn's theorem, sparsity.

Phase transition is a well-known optimality framework [13, 22, 34, 38]. It is related to the minimax framework but can be more informative in many cases. Conceptually, for the global testing problem, in the two-dimensional phase space with the two axes calibrating the "sparsity" and "signal strength," respectively, there is a "Region of Possibility" and a "Region of Impossibility." In the "Region of Possibility," any alternative is separable from the null. In the "Region of Impossibility," any alternative is inseparable from the null.

If a test is able to automatically adapt to different levels of sparsity and separate any given alternative in the "Region of Possibility" from the null, then we call it "optimally adaptive."

We are interested in finding tests that satisfy the following requirements:

(R1) Applicable to networks with severe degree heterogeneity.
(R2) Applicable to networks with mixed memberships.
(R3) The asymptotic null distribution is easy to track, so the rejection regions are easy to set.
(R4) Optimally adaptive: We desire a single test that is able to adapt to different levels of sparsity and is optimally adaptive.

1.1. *The DCMM model.* We adopt the *Degree Corrected Mixed Membership (DCMM)* model [24, 43]. Denote the adjacency matrix by A, where

$$A_{ij} = \begin{cases} 1, & \text{if node } i \text{ and node } j \text{ have an edge,} \\ 0, & \text{otherwise.} \end{cases} \tag{1.1}$$

Conventionally, self-edges are not allowed so all the diagonal entries of A are 0. In DCMM, we assume there are K perceivable communities $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_K$, and each node is associated with a mixed-membership weight vector $\pi_i = (\pi_i(1), \pi_i(2), \dots, \pi_i(K))'$ where for $1 \le k \le K$ and $1 \le i \le n$,

$$\pi_i(k) = \text{the weight node } i \text{ puts on community } k. \tag{1.2}$$

Moreover, for a $K \times K$ symmetric nonnegative matrix P, which models the community structure, and positive parameters $\theta_1, \theta_2, \dots, \theta_n$, which model the degree heterogeneity, we assume the upper triangular entries of A are independent Bernoulli variables satisfying

$$\mathbb{P}(A_{ij} = 1) = \theta_i \theta_j \cdot \pi_i' P \pi_j \equiv \Omega_{ij}, \quad 1 \le i < j \le n, \tag{1.3}$$

where Ω denotes the matrix $\Theta \Pi P \Pi' \Theta$, with Θ being the $n \times n$ diagonal matrix $\text{diag}(\theta_1, \dots, \theta_n)$ and Π being the $n \times K$ matrix $[\pi_1, \pi_2, \dots, \pi_n]'$. For identifiability (see [24] for more discussion), we assume

$$\text{all diagonal entries of } P \text{ are } 1. \tag{1.4}$$

When $K = 1$, (1.4) implies $P = 1$, and so $\Omega_{ij} = \theta_i \theta_j$, $1 \le i, j \le n$.

Write for short $\text{diag}(\Omega) = \text{diag}(\Omega_{11}, \Omega_{22}, \dots, \Omega_{nn})$, and let W be the matrix where for $1 \le i, j \le n$, $W_{ij} = A_{ij} - \Omega_{ij}$ if $i \ne j$ and $W_{ij} = 0$ otherwise. In matrix form, we have

$$A = \Omega - \text{diag}(\Omega) + W, \quad \text{where } \Omega = \Theta \Pi P \Pi' \Theta. \tag{1.5}$$

DCMM includes three models as special cases, each of which is well known and has been studied extensively recently.

- *Degree Corrected Block Model (DCBM)* [29]. If we do not allow mixed memberships (i.e., each weight vector π_i is degenerate with one entry being nonzero), then DCMM reduces to the DCBM.
- *Mixed-Membership Stochastic Block Model (MMSBM)* [1]. DCBM further reduces to MMSBM if $\theta_1 = \cdots = \theta_n (= \sqrt{\alpha_n})$. In this special case, $\Omega = \alpha_n \Pi P \Pi'$, and for identifiability, (1.4) is too strong, so we relax it to that the average of the diagonals of P is 1.

- *Stochastic Block Model (SBM)* [20]. MMSBM further reduces to the classical SBM if additionally we do not allow mixed memberships.

Under DCMM, the global testing problem is the problem of testing

$$H_0^{(n)} : K = 1 \quad \text{vs.} \quad H_1^{(n)} : K \geq 2. \tag{1.6}$$

The seeming simplicity of the two hypotheses is deceiving, as both of them are highly composite, consisting of many different parameter configurations.

1.2. *Phase transition*: *A preview of our main results*. Let $\lambda_1, \lambda_2, \ldots, \lambda_K$ be the first K eigenvalues of Ω, arranged in the descending order in magnitude. We can view (a) $\sqrt{\lambda_1}$ both as the sparsity level and the noise level [23] (i.e., spectral norm of the noise matrix W), (b) $|\lambda_2|$ as the signal strength, so that $|\lambda_2|/\sqrt{\lambda_1}$ is the Signal-to-Noise Ratio (SNR) and (c) $|\lambda_2|/\lambda_1$ as a measure of dissimilarity between different communities (Example 1 below illustrates why it measures "dissimilarity"). We note that [12, 19] also pointed out that $|\lambda_2|/\sqrt{\lambda_1}$ is a reasonable metric of SNR.

Now, in the two-dimensional *phase space* where the x-axis is $\sqrt{\lambda_1}$, which measures the sparsity level, and the y-axis is $|\lambda_2|/\lambda_1$, which measures the community dissimilarity, we have two regions.

- *Region of Possibility (*$1 \ll \sqrt{\lambda_1} \ll \sqrt{n}$, $|\lambda_2|/\sqrt{\lambda_1} \to \infty$*)*. For any alternative hypothesis in this region, it is possible to distinguish it from any null hypothesis, by the Signed Polygon tests to be introduced.
- *Region of Impossibility (*$1 \ll \sqrt{\lambda_1} \ll \sqrt{n}$, $|\lambda_2|/\sqrt{\lambda_1} \to 0$*)*. In this region, any alternative hypothesis is inseparable from the null hypothesis, provided with some mild conditions.

See Figure 1 (left panel). Also, see Sections 2 and 3 for our main theorems on *Possibility* and *Impossibility*, respectively. Note that the figure is only for illustration purposes, where the cases of $|\lambda_2| = c_0\sqrt{\lambda_1}$ for some constant $c_0 > 0$ are compressed in the separating the boundary of two regions (red curve). The Signed Polygon test satisfies all requirements (R1)–(R4) above. Since the test is able to separate all alternatives (ranging from very sparse to less sparse) in the Region of Possibility from the null, it is *optimally adaptive*.

REMARK 1. A stronger version of the phase transition is that for a constant $c_0 > 0$, the Region of Possibility and Region of Impossibility are given by $|\lambda_2|/\sqrt{\lambda_1} > c_0$ and $|\lambda_2|/\sqrt{\lambda_1} < c_0$, respectively. For the broad setting, we consider, this is an open problem, though for some special cases, there are some interesting works (e.g., [19]); see Remark 11.

It is instructive to consider a special DCMM model, which is a generalization of the symmetric SBM [37] to the case with degree heterogeneity.

EXAMPLE 1 (A special DCMM). Let $e_1, \ldots, e_K$ be the standard basis of $\mathbb{R}^K$. Fixing a positive vector $\theta \in \mathbb{R}^n$ and a scalar $b_n \in (0, 1)$, we assume

$$P = (1 - b_n)I_K + b_n 1_K 1_K', \qquad \pi_i \text{ are i.i.d. sampled from } e_1, \ldots, e_K. \tag{1.7}$$

In this model, $(1 - b_n)$ measures the "dissimilarity" between different communities (it quantifies how well we can tell whether two nodes i and j are from the same community or not; note that $b_n = 1$ corresponds to the null case where all communities are indistinguishable) and $\|\theta\|$ measures the sparsity level. In this model, $\lambda_1 \sim (1 + (K - 1)b_n)\|\theta\|^2$ and $\lambda_k \sim (1 - b_n)\|\theta\|^2$, $2 \leq k \leq K$. The sparsity level is $\sqrt{\lambda_1} \asymp \|\theta\|$, the community dissimilarity is characterized by $\lambda_2/\lambda_1 \asymp (1 - b_n)$, and the SNR is $|\lambda_2|/\sqrt{\lambda_1} \asymp \|\theta\|(1 - b_n)$. The Region of Possibility and Region of Impossibility are given by $\{1 \ll \|\theta\| \ll \sqrt{n}, \|\theta\|(1 - b_n) \to \infty\}$ and $\{1 \ll \|\theta\| \ll \sqrt{n}, \|\theta\|(1 - b_n) \to 0\}$, respectively. See Figure 1 (right panel).

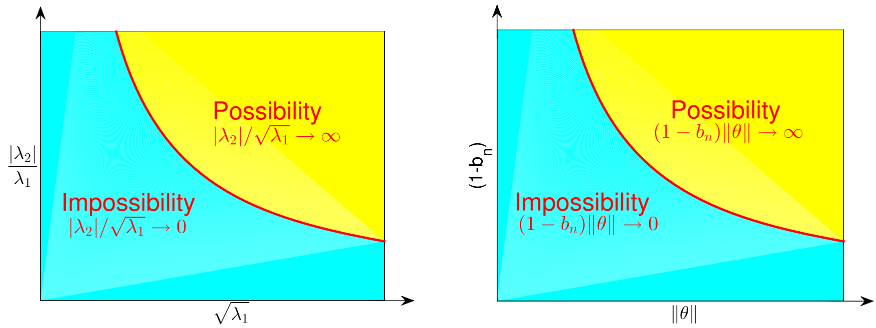

FIG. 1. *Left: Phase transition. In the Region of Impossibility, any alternative hypothesis is indistinguishable from a null hypothesis, provided that some mild conditions hold. In the Region of Possibility, the Signed Polygon test is able to separate any alternative hypothesis from a null hypothesis asymptotically. Right: Phase transition for the special DCMM model in Example* 1, *where* $\sqrt{\lambda_1} \asymp \|\theta\|$, $|\lambda_2|/\lambda_1 \asymp (1-b_n)$ *and* $|\lambda_2|/\sqrt{\lambda_1} \asymp (1-b_n)\|\theta\|$.

REMARK 2. As the phase transition is hinged on $\lambda_2/\sqrt{\lambda_1}$, one may think that the statistic $\hat{\lambda}_2/\sqrt{\hat{\lambda}_1}$ is optimally adaptive, where $\hat{\lambda}_k$ is the kth largest (in magnitude) eigenvalue of A. This is however not true, because the consistency of $\hat{\lambda}_2$ for estimating λ_2 cannot be guaranteed in our range of interest, unless with strong conditions on $\theta_{\max}$ [23].

1.3. *Literature review, the signed polygon and our contribution.* Recently, the global testing problem has attracted much attention and many interesting approaches have been proposed. To name a few, Mossel et al. [37] and Banerjee and Ma [3] (see also [4]) considered a special case of the testing problem, where they assume a simple null of Erdős–Renyi random graph model and a special alternative which is an SBM with two equal-sized communities. They provided the asymptotic distribution of the log-likelihood ratio within the contiguous regime. Since the likelihood ratio test statistic is NP-hard to compute, [3] introduced an approximation by linear spectral statistics. Lei [32] also considered the SBM model and studied the problem of testing whether $K = K_0$ or $K > K_0$, where K_0 is a prespecified integer. His approach is based on the Tracy–Widom law of extreme eigenvalues and requires delicate random matrix theory. Unfortunately, these works have been focused on the SBM (which allows neither severe degree heterogeneity nor mixed membership). Therefore, despite the elegant theory in these works, it remains unclear how to extend their ideas to our settings.

Along a different line, graphlet counts (GC) have been frequently used for hypothesis testing in nonparametric and parametric network models. This includes the EZ test [16] and GC test [25]. Other interesting works include [6, 7, 36]. In particular, [25] suggested a general recipe for constructing test statistics and showed that both GC and EZ tests have competitive power in a broad setting. Unfortunately, it turns out that in the less sparse case, the variance of the GC test statistic is much larger than expected, which largely hurts the power of the test. The underlying reason is that GC tests use *noncentered* cycle counts. If, however, we use *centered* cycle counts, we can largely reduce the variances and have a more powerful test. A similar phenomenon was discovered by Bubeck et al. [10] for the SBM setting.

This motivates a class of new tests, which we call *Signed Polygon*, including the Signed Triangle (SgnT) and the Signed Quadrilateral (SgnQ). The Signed Polygon statistics are related to the Signed Cycle statistics, first introduced by Bubeck et al. [10] and later generalized by Banerjee [2]. Both the Signed Polygon and Signed Cycle recognize that using centered-cycle counts may help reduce the variance, but there are some major differences. The study of the Signed Cycles has been focused on the SBM and similar models, where under the null, $\mathbb{P}(A_{ij} = 1) = \alpha$, $1 \le i \ne j \le n$, and α is the only unknown parameter. In this case, a natural approach to centering the adjacency matrix A is to first estimate α using the whole matrix

A (say, $\hat{\alpha}$), and then subtract all off-diagonal entries of A by $\hat{\alpha}$. However, under the null of our setting, $\mathbb{P}(A_{ij} = 1) = \theta_i\theta_j$, $1 \le i \ne j \le n$, and there are n different unknown parameters $\theta_1, \theta_2, \ldots, \theta_n$. In this case, how to center the matrix A is not only unclear but also *worrisome*, especially when the network is very sparse, because we have to use limited data to estimate a large number of unknown parameters. Also, for any approaches we may have, the analysis is seen to be much harder than that of the previous case. Note that the ways how two statistics are defined over the centered adjacency matrix are also different; see Section 1.4 and [2, 10].

In the Signed Polygon, we use a new approach to estimate $\theta_1, \theta_2, \ldots, \theta_n$ under the null, and use the estimates to center the matrix A. To our surprise, data limitation (though a challenge) does not ruin the idea: even for very sparse networks, the estimation errors of $\theta_1, \theta_2, \ldots, \theta_n$ only have a negligible effect. The main contributions of the paper are as follows:

- Discover the phase transition for global testing in the broad DCMM setting by identifying the Regions of Impossibility and Possibility.
- Propose the Signed Polygon as a class of new tests that are appropriate for networks with severe degree heterogeneity and mixed memberships.
- Prove that the Signed Triangle and Signed Quadrilateral tests satisfy all the requirements (R1)–(R4), and especially that they are optimally adaptive and perform well for all networks in the Region of Possibility, ranging from very sparse ones to the least sparse ones.

To show the success of the Signed Polygon test for the whole Region of Possibility is very subtle and extremely tedious. The main reason is that we hope to cover the *whole spectrum* of degree heterogeneity and sparsity levels. Crude bounds may work in one case but not another, and many seemingly negligible terms turn out to be nonnegligible (see Sections 1.4 and 4). The lower bound argument is also very subtle. Compared to work on SBM where there is only one unknown parameter under the null, our null has n unknown parameters. The difference provides a lot of freedom in constructing inseparable hypothesis pairs, and so the Region of Impossibility in our setting is much wider than that for SBM. Our construction of inseparable hypothesis pairs uses theorems on nonnegative matrix scaling, a mathematical area pioneered by Sinkhorn [41] and Olkin [35] among others (e.g., [9, 28]).

1.4. *The signed polygon statistic.* Recall that A is the adjacency matrix of the network. Introduce a vector $\hat{\eta}$ by ($\mathbf{1}_n$ denotes the vector of 1's)

$$\hat{\eta} = (1/\sqrt{V})\, A\, \mathbf{1}_n, \quad \text{where } V = \mathbf{1}_n' A \mathbf{1}_n. \tag{1.8}$$

Fixing $m \ge 3$, the order-m *Signed Polygon* statistic is defined by (notation: (*dist*) is short for "distinct," which means any two of $i_1, \ldots, i_m$ are unequal)

$$U_n^{(m)} = \sum_{i_1, i_2, \ldots, i_m (dist)} (A_{i_1 i_2} - \hat{\eta}_{i_1}\hat{\eta}_{i_2})(A_{i_2 i_3} - \hat{\eta}_{i_2}\hat{\eta}_{i_3}) \ldots (A_{i_m i_1} - \hat{\eta}_{i_m}\hat{\eta}_{i_1}). \tag{1.9}$$

When $m = 3$, we call it the Signed-Triangle (SgnT) statistic:

$$T_n = \sum_{i_1, i_2, i_3 (dist)} (A_{i_1 i_2} - \hat{\eta}_{i_1}\hat{\eta}_{i_2})(A_{i_2 i_3} - \hat{\eta}_{i_2}\hat{\eta}_{i_3})(A_{i_3 i_1} - \hat{\eta}_{i_3}\hat{\eta}_{i_1}). \tag{1.10}$$

When $m = 4$, we call it the Signed-Quadrilateral (SgnQ) statistic:

$$Q_n = \sum_{i_1, i_2, i_3, i_4 (dist)} (A_{i_1 i_2} - \hat{\eta}_{i_1}\hat{\eta}_{i_2})(A_{i_2 i_3} - \hat{\eta}_{i_2}\hat{\eta}_{i_3})(A_{i_3 i_4} - \hat{\eta}_{i_3}\hat{\eta}_{i_4})(A_{i_4 i_1} - \hat{\eta}_{i_4}\hat{\eta}_{i_1}). \tag{1.11}$$

For analysis, we focus on T_n and Q_n, but our main results are extendable to general m.

The key to understanding and analyzing the Signed Polygon is the *Ideal Signed Polygon*. Introduce a *nonstochastic counterpart* of $\hat{\eta}$ by

$$\eta^* = (1/\sqrt{v_0})\Omega \mathbf{1}_n, \quad \text{where } v_0 = \mathbf{1}_n' \Omega \mathbf{1}_n. \tag{1.12}$$

Define the order-m *Ideal Signed Polygon* statistic by

$$\widetilde{U}_n^{(m)} = \sum_{i_1,i_2,\dots,i_m (dist)} (A_{i_1 i_2} - \eta_{i_1}^* \eta_{i_2}^*)(A_{i_2 i_3} - \eta_{i_2}^* \eta_{i_3}^*) \dots (A_{i_m i_1} - \eta_{i_m}^* \eta_{i_1}^*). \tag{1.13}$$

We expect to see that $\hat{\eta} \approx \mathbb{E}[\hat{\eta}] \approx \eta^*$. We can view $\widetilde{U}_n^{(m)}$ as the oracle version of $U_n^{(m)}$, with η^* given. We can also view $U_n^{(m)}$ as the *plug-in* version of $\widetilde{U}_n^{(m)}$, where we replace η^* by $\hat{\eta}$.

For implementation, it is desirable to rewrite T_n and Q_n in matrix forms, which allows us to avoid using an for-loop and compute much faster (say, in MATLAB or R). For any two matrices $M, N \in \mathbb{R}^{n,n}$, let $\mathrm{tr}(M)$ be the trace of M, $\mathrm{diag}(M) = \mathrm{diag}(M_{11}, M_{22}, \dots, M_{nn})$, and $M \circ N$ be the Hadamard product of M and N (i.e., $M \circ N \in \mathbb{R}^{n,n}$, $(M \circ N)_{ij} = M_{ij} N_{ij}$). Denote $\widetilde{A} = A - \hat{\eta}\hat{\eta}'$. The following theorem is proved in the Supplementary Material [26].

THEOREM 1.1. *We have* $T_n = \mathrm{tr}(\widetilde{A}^3) - 3\,\mathrm{tr}(\widetilde{A} \circ \widetilde{A}^2) + 2\,\mathrm{tr}(\widetilde{A} \circ \widetilde{A} \circ \widetilde{A})$ *and* $Q_n = \mathrm{tr}(\widetilde{A}^4) - 4\,\mathrm{tr}(\widetilde{A} \circ \widetilde{A}^3) + 8\,\mathrm{tr}(\widetilde{A} \circ \widetilde{A} \circ \widetilde{A}^2) - 6\,\mathrm{tr}(\widetilde{A} \circ \widetilde{A} \circ \widetilde{A} \circ \widetilde{A}) - 2\,\mathrm{tr}(\widetilde{A}^2 \circ \widetilde{A}^2) + 2 \cdot 1_n'[\mathrm{diag}(\widetilde{A})(\widetilde{A} \circ \widetilde{A})\,\mathrm{diag}(\widetilde{A})]1_n + 1_n'[\widetilde{A} \circ \widetilde{A} \circ \widetilde{A} \circ \widetilde{A}]1_n$. *The complexity of computing both* T_n *and* Q_n *is* $O(n^2\bar{d})$, *where* $\bar{d}$ *is the average degree of the network.*

Compared to the EZ and GC tests [16, 25], the computational complexity of SgnT and SgnQ is of the same order.

REMARK 3. The computational complexity of $U_n^{(m)}$ remains as $O(n^2\bar{d})$ for larger m. Similarly as that in Theorem 1.1, the main complexity of $U_n^{(m)}$ comes from computing $\widetilde{A}^m$. Since we can compute $\widetilde{A}^m$ with $\widetilde{A}^m = \widetilde{A}^{m-1}\widetilde{A}$ and recursive matrix multiplications, each time with a complexity of $O(n^2\bar{d})$, the overall complexity is $O(n^2\bar{d})$.

REMARK 4 (Connection to the Signed Cycle). In the more idealized SBM or MMSBM model, we do not have degree heterogeneity, and $\Omega = \alpha_n \mathbf{1}_n \mathbf{1}_n'$ under the null, where α_n is the only unknown parameter. In this simple setting, it makes sense to estimate α_n by $\hat{\alpha}_n = \bar{d}/(n-1)$, where $\bar{d}$ is the average degree. This gives rise to the *Signed Cycle* statistics [2, 10]: $C_n^{(m)} = \sum_{i_1,i_2,\dots,i_m (dist)} (A_{i_1 i_2} - \hat{\alpha}_n)(A_{i_2 i_3} - \hat{\alpha}_n) \dots (A_{i_m i_1} - \hat{\alpha}_n)$. Bubeck et al. [10] first proposed $C_n^{(3)}$ for a global testing problem in a model similar to MMSBM. Although their test statistic is also called the Signed Triangle, it is different from our SgnT statistic (1.10), because their tests are only applicable to models without degree heterogeneity. The analysis of the Signed Polygon is also much more delicate than that of the Signed Cycle, as the error $(\hat{\alpha}_n - \alpha_n)$ is much smaller than the errors in $(\hat{\eta} - \eta^*)$.

It remains to understand (A) how the Signed Polygon manages to reduce variance, and (B) what are the analytical challenges.

Consider Question (A). We illustrate it with the Ideal Signed Polygon (1.13) and the null case. In this case, $\Omega = \theta\theta'$. It is seen $\eta^* = \theta$, $A_{ij} - \eta_i^*\eta_j^* = A_{ij} - \Omega_{ij} = W_{ij}$, for $i \neq j$ (see (1.5) for definition of W), and so $\widetilde{U}_n^{(m)} = \sum_{i_1,i_2,\dots,i_m (dist)} W_{i_1 i_2} W_{i_2 i_3} \dots W_{i_m i_1}$. Here, each term is an m-product of independent centered Bernoulli variables, and $W_{i_1 i_2} W_{i_2 i_3} \dots W_{i_m i_1}$ and $W_{i_1' i_2'} W_{i_2' i_3'} \dots W_{i_m' i_1'}$ are correlated only when $\{i_1, i_2, \dots, i_m\}$ and $\{i_1', i_2', \dots, i_m'\}$ are the vertices of the same polygon. Such a construction is known to be efficient in variance reduction (e.g., [10]).

In comparison, for an order-m GC statistic [25], $N_n^{(m)} = \sum_{i_1,i_2,\dots,i_m (dist)} A_{i_1 i_2} A_{i_2 i_3} \dots A_{i_m i_1}$ is the main term. Since here the Bernoulli variables are not centered, we can split $N_n^{(m)}$ into two uncorrelated terms: $N_n^{(m)} = \widetilde{U}_n^{(m)} + (N_n^{(m)} - \widetilde{U}_n^{(m)})$. Compared to the Signed Polygon, the additional variance comes from the second term, which is undesirably large in the less sparse case [30].

REMARK 5. The above also explains why the order-2 Signed Polygon does not work well. In fact, when $m = 2$, $\widetilde{U}_n^{(m)} = \sum_{i_1 \neq i_2} W_{i_1 i_2}^2$ under the null, which has an unsatisfactory variance due to the square of the W-terms.

Consider Question (B). We discuss with the SgnQ statistic. Recall that η^* is a nonstochastic proxy of $\hat{\eta}$. For any $1 \leq i, j \leq n$ and $i \neq j$, we decompose $\eta_i^* \eta_j^* - \hat{\eta}_i \hat{\eta}_j = \delta_{ij} + r_{ij}$, where δ_{ij} is the main term, which is a linear function of $\hat{\eta}_i$ and $\hat{\eta}_j$, and r_{ij} is the remainder term. Introduce

$$\widetilde{\Omega} = \Omega - \eta^* (\eta^*)'. \tag{1.14}$$

We have $A_{ij} - \hat{\eta}_i \hat{\eta}_j = \widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij} + r_{ij}$. After inserting this into Q_n, each 4-product is now the product of 4 bracketed terms, where each bracketed term is the sum of 4 terms. Expanding the brackets and reorganizing, Q_n splits into $4 \times 4 \times 4 \times 4 = 256$ *post-expansion* sums, each having the form $\sum_{i_1, i_2, i_3, i_4 (dist)} a_{i_1 i_2} b_{i_2 i_3} c_{i_3 i_4} d_{i_4 i_1}$, where a is a generic term, which can be equal to either of the four terms $\widetilde{\Omega}$, W, δ and r; same for b, c and d. While some of these terms may be equal to each other, the symmetry, we can exploit is limited, due to (a) degree heterogeneity, (b) mixed memberships and (c) the underlying polygon structure. As a result, we still have more than 50 post-expansion sums to analyze.

The analysis of a post-expansion sum with the presence of one or more r-term is the most tedious of all, where we need to further decompose each r-term into three different terms. This requires analysis of more than 100 additional post-expansion sums. We may think most of the post-expansion sums are easy to control via a crude bound (e.g., by the Cauchy–Schwarz inequality). Unfortunately, this is not the case, and many seemingly negligible terms turn out to be nonnegligible. Here are some of the reasons:

- We wish to cover most interesting cases. A crude bound may be enough for some cases but not for others.
- We desire to have a *single* test that achieves the phase transition for the whole range of interest. Alternatively, we may want to find several tests, each covering a subset of cases of interest, but this is less appealing.

As a result, we have to analyze a large number of post-expansion sums, where the analysis is subtle, extremely tedious and error-prone, involving delicate combinatorics, due to the underlying polygon structure. See Section 4.

REMARK 6. In Signed Polygon (1.9), we estimate Ω by $\hat{\eta}\hat{\eta}' = (\mathbf{1}_n' A \mathbf{1}_n)^{-1} A \mathbf{1}_n \mathbf{1}_n' A$ for the null. Alternatively, we may use a spectral approach and estimate Ω by $\hat{\lambda}_1 \hat{\xi}_1 \hat{\xi}_1'$, where $\hat{\lambda}_1$ and $\hat{\xi}_1$ are the first eigenvalue and eigenvector of A, respectively. Unfortunately, even in the more idealized SBM case, this estimate may be unsatisfactory for sparse networks (e.g., [11], Section 2.2). In fact, for our main results to hold, we need to have $|\hat{\lambda}_1 - \lambda_1| \leq C\|\theta\|$ with large probability, but the best concentration inequality we have is $|\hat{\lambda}_1 - \lambda_1| \leq C\sqrt{\theta_{\max}\|\theta\|_1}$ with large probability ([24], Lemma C.1). In the presence of severe degree heterogeneity, we often have $\sqrt{\theta_{\max}\|\theta\|_1} \gg \|\theta\|$. Also, unlike $\hat{\eta}\hat{\eta}'$ in our proposal, $\hat{\lambda}_1 \hat{\xi}_1 \hat{\xi}_1'$ is not an explicit function of A, so the alternative version of the Signed Polygon statistic is much harder to analyze.

1.5. *Organization of the paper.* Section 2 focuses on the Region of Possibility and contains the upper bound argument. Section 3 focuses on the Region of Impossibility and contains the lower bound argument. Section 4 presents the key proof ideas, with the proof of secondary lemmas deferred to the Supplementary Material. Section 5 presents the numerical study, and Section 6 discusses extensions and connections.

For any $q > 0$ and $\theta \in \mathbb{R}^n$, $\|\theta\|_q$ denotes the ℓ^q-norm of θ (when $q = 2$, we drop the subscript for simplicity). Also, $\theta_{\min}$ and $\theta_{\max}$ denote $\min\{\theta_1, \dots, \theta_n\}$ and $\max\{\theta_1, \dots, \theta_n\}$, respectively. For any $n > 1$, $\mathbf{1}_n \in \mathbb{R}^n$ denotes the vector of 1's. For two positive sequences $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$, we write $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$, and we write $a_n \asymp b_n$ if for sufficiently large n, there are two constants $c_2 > c_1 > 0$ such that $c_1 \le a_n/b_n \le c_2$. We use $\sum_{i_1,i_2,\dots,i_m(dist)}$ to denote the sum over all $(i_1, \dots, i_m)$ such that $1 \le i_k \le n$ and $i_k \ne i_\ell$ for $1 \le k \ne \ell \le m$. We use $C > 0$ as a generic constant that may vary from occurrence to occurrence. For constants that need to be more specific, we use c_0, c_1, etc.

2. The signed polygon test and the upper bound. For reasons aforementioned, we focus on the SgnT statistic T_n and SgnQ statistic Q_n, but the ideas are extendable to general Signed Polygon statistics. In this section, we study the upper bound. In detail, in Section 2.1, we establish the asymptotic normality of both test statistics. In Sections 2.2–2.3, we discuss the power of the two tests. We show that if $|\lambda_2|/\sqrt{\lambda_1} \to \infty$ and some mild regularity conditions hold, then for each of the two tests, the sum of Type I and Type II errors tends to 0 as $n \to \infty$. The lower bound is studied in Section 3, where we show that for an alternative hypothesis setting with $|\lambda_2|/\sqrt{\lambda_1} \to 0$, we can always pair it with a null setting so that two hypotheses are asymptotically inseparable.

In a DCMM model, $\Omega = \Theta\Pi P\Pi'\Theta$, where $\Theta = \mathrm{diag}(\theta_1, \dots, \theta_n)$, and Π is the $n \times K$ membership matrix $[\pi_1, \pi_2, \dots, \pi_n]'$. We assume as $n \to \infty$,

$$\|\theta\| \to \infty, \qquad \theta_{\max} \to 0, \quad \text{and} \quad (\|\theta\|^2/\|\theta\|_1)\sqrt{\log(\|\theta\|_1)} \to 0. \tag{2.1}$$

The first condition is necessary. In fact, if $\|\theta\| \to 0$, then the alternative is indistinguishable from the null, as suggested by lower bounds in Section 3. The second one is mild as we usually assume $\theta_{\max} \le C$. This is due to that under DCMM, P has unit diagonal entries and $\theta_i\theta_j(\pi_i' P\pi_j)$ is a probability for all $i \ne j$. The last one is weaker than that of $\theta_{\max}\sqrt{\log(n)} \to 0$, and is very mild. It is assumed mostly for technical reasons and is not required in many cases (e.g., the dense case where all $\theta_i = O(1)$). Moreover, introduce $G = \|\theta\|^{-2}\Pi'\Theta^2\Pi \in \mathbb{R}^{K\times K}$. This matrix is properly scaled and it can be shown that $\|G\| \le 1$ (Appendix E, Supplemental Material). When the null is true, $K = P = G = 1$, and we do not need any additional condition. When the alternative is true, we assume

$$\frac{\max_{1\le k\le K}\{\sum_{i=1}^n \theta_i\pi_i(k)\}}{\min_{1\le k\le K}\{\sum_{i=1}^n \theta_i\pi_i(k)\}} \le C, \qquad \|G^{-1}\| \le C, \qquad \|P\| \le C. \tag{2.2}$$

Here, $C > 0$ is a generic constant; see Section 1.5. The conditions are mild. Take the first two, for example. When there is no mixed membership, they only require the K classes to be relatively balanced.

2.1. *Asymptotic normality of the null.* Theorems 2.1–2.2 are proved in the supplement.

THEOREM 2.1 (Limiting null of the SgnT statistic). *Consider the testing problem* (1.6) *under the DCMM model* (1.1)–(1.4), *where the condition* (2.1) *is satisfied. Suppose the null hypothesis is true. As* $n \to \infty$, $\mathbb{E}[T_n] = o(\|\theta\|^3)$, $\mathrm{Var}(T_n) \sim 6\|\theta\|^6$ *and* $(T_n - \mathbb{E}[T_n])/\sqrt{\mathrm{Var}(T_n)} \longrightarrow N(0, 1)$ *in law.*

THEOREM 2.2 (Limiting null of the SgnQ statistic). *Consider the testing problem* (1.6) *under the DCMM model* (1.1)–(1.4), *where the condition* (2.1) *is satisfied. Suppose the null hypothesis is true. As* $n \to \infty$, $\mathbb{E}[Q_n] = (2 + o(1))\|\theta\|^4$, $\mathrm{Var}(Q_n) \sim 8\|\theta\|^8$ *and* $(Q_n - \mathbb{E}[Q_n])/\sqrt{\mathrm{Var}(Q_n)} \longrightarrow N(0, 1)$ *in law.*

Note that under the null, the limiting distributions of $T_n/\sqrt{\mathrm{Var}(T_n)}$ and $Q_n/\sqrt{\mathrm{Var}(Q_n)}$ are $N(0,1)$ and $N(1/\sqrt{2},1)$, respectively. To appreciate the difference, recall that the Signed Polygon can be viewed as a plug-in statistic, where we replace η^* in the Ideal Signed Polygon by $\hat{\eta}$. Under the null, the effect of the plug-in is negligible for SgnT but not for SgnQ, so the two limiting distributions are different. See Section 4 for details.

2.2. *The level-α SgnT and SgnQ tests.* By Theorems 2.1 and 2.2, the null variances of the two statistics depend on $\|\theta\|^2$. To use the two statistics as tests, we need to estimate $\|\theta\|^2$. For $\hat{\eta}$ and η^* defined in (1.8) and (1.12), respectively, we have $\hat{\eta}\approx\eta^*$ and $\eta^*=\theta$ under the null. A reasonable estimator for $\|\theta\|^2$ under the null is therefore $\|\hat{\eta}\|^2$. We propose to estimate $\|\theta\|^2$ with $(\|\hat{\eta}\|^2-1)$, which corrects the bias and is slightly more accurate than $\|\hat{\eta}\|^2$. The following lemma is proved in the Supplementary Material.

LEMMA 2.1 (Estimation of $\|\theta\|^2$). *Consider the testing problem* (1.6) *under the DCMM model* (1.1)–(1.4), *where the condition* (2.1) *holds when either hypothesis is true and condition* (2.2) *holds when the alternative is true. Then, under both hypotheses, as $n\to\infty$ $(\|\hat{\eta}\|^2-1)/\|\eta^*\|^2\to 1$ in probability, where $\|\eta^*\|^2=(\mathbf{1}_n'\Omega^2\mathbf{1}_n)/(\mathbf{1}_n'\Omega\mathbf{1}_n)$. Furthermore, $\|\eta^*\|^2=\|\theta\|^2$ under $H_0^{(n)}$ and $\|\eta^*\|^2\asymp\|\theta\|^2$ under $H_1^{(n)}$.*

Combining Lemma 2.1 with Theorem 2.1 gives

$$T_n/\sqrt{6(\|\hat{\eta}\|^2-1)^3}\longrightarrow N(0,1),\quad \text{in law.} \tag{2.3}$$

Fix $\alpha\in(0,1)$. We propose the following SgnT test, which is a two-sided test where we reject the null hypothesis if and only if

$$|T_n|\geq z_{\alpha/2}\sqrt{6}(\|\hat{\eta}\|^2-1)^{3/2},\quad z_{\alpha/2}\text{: upper }(\alpha/2)\text{-quantile of }N(0,1). \tag{2.4}$$

Similarly, combining Theorem 2.2 and Lemma 2.1, we have

$$[Q_n-2(\|\hat{\eta}\|^2-1)^2]/\sqrt{8(\|\hat{\eta}\|^2-1)^4}\longrightarrow N(0,1),\quad \text{in law.} \tag{2.5}$$

With the same α, we propose the following SgnQ test, which is a one-sided test where we reject the null hypothesis if and only if

$$Q_n\geq(2+z_\alpha\sqrt{8})(\|\hat{\eta}\|^2-1)^2,\quad z_\alpha\text{: upper }\alpha\text{-quantile of }N(0,1). \tag{2.6}$$

As a result, for both tests we just defined, the levels satisfy

$$\mathbb{P}_{H_0^{(n)}}(\text{Reject the null})\to\alpha,\quad \text{as } n\to\infty.$$

Figure 2 shows the histograms of $T_n/\sqrt{6(\|\hat{\eta}\|^2-1)^3}$ (left) and $(Q_n-2(\|\hat{\eta}\|^2-1)^2)/(\sqrt{8(\|\hat{\eta}\|^2-1)^4})$ (right) under a null and an alternative simulated from DCMM. Recall that in DCMM, $\Omega=\theta\theta'$ under the null and $\Omega=\Theta\Pi P\Pi\Theta$, where $\Theta=\mathrm{diag}(\theta_1,\ldots,\theta_n)$. For the null, we take $n=2000$ and draw θ_i from Pareto$(12,3/8)$ and scale θ to have an ℓ^2-norm of 8. For the alternative, we let $(n,K)=(2000,2)$, P be the matrix with 1 on the diagonal and 0.6 on the off-diagonal, rows of Π equal to $\{1,0\}$ and $\{0,1\}$ half by half, and with the same θ as in the null but (to make it harder to separate from the null) rescaled to have an ℓ^2-norm of 9. The results confirm the limiting null of $N(0,1)$ for both tests.

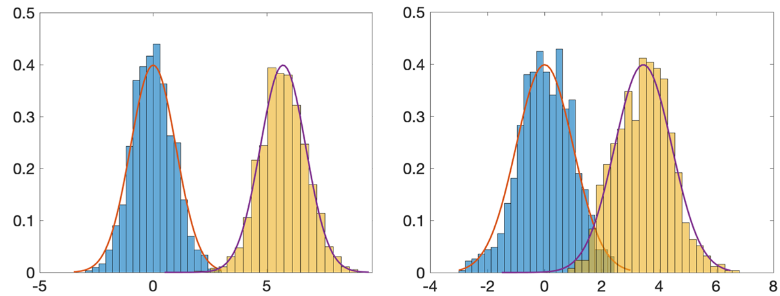

FIG. 2. *Left*: *histograms of the SgnT test statistics in* (2.3) *for the null* (*blue*) *and the alternative* (*yellow*). *Empirical mean and SD under the null*: 0.04 *and* 0.94. *Right*: *same but for SgnQ test statistic in* (2.5). *Empirical mean and SD under the null*: −0.02 *and* 0.92. *Repetition*: 1000 *times*. *See setting details in the main text*.

2.3. *Power analysis of the SgnT and SgnQ tests*. The matrices Ω and $\widetilde{\Omega}$ play a key role in power analysis. Recall that Ω is defined in (1.3) where $\text{rank}(\Omega) = K$, and $\widetilde{\Omega} = \Omega - \eta^*(\eta^*)'$ is defined in (1.14) with $\eta^* = \Omega \mathbf{1}_n / \sqrt{\mathbf{1}_n' \Omega \mathbf{1}_n}$ as in (1.12). Recall that $\lambda_1, \lambda_2, \ldots, \lambda_K$ are the K nonzero eigenvalues of Ω. Let $\xi_1, \xi_2, \ldots, \xi_K$ be the corresponding eigenvectors. The following theorems are proved in the Supplemental Material.

THEOREM 2.3 (Limiting behavior the SgnT statistic (alternative)). *Consider the testing problem* (1.6) *under the DCMM model* (1.1)–(1.4). *Suppose the alternative hypothesis is true*, *and the conditions* (2.1)–(2.2) *hold. As* $n \to \infty$, $\mathbb{E}[T_n] = \text{tr}(\widetilde{\Omega}^3) + o((|\lambda_2|/\lambda_1)^3 \|\theta\|^6) + o(\|\theta\|^3)$ *and* $\text{Var}(T_n) \le C[\|\theta\|^6 + (\lambda_2/\lambda_1)^4 \|\theta\|^4 \|\theta\|_3^6]$.

THEOREM 2.4 (Limiting behavior of the SgnQ statistic (alternative)). *Consider the testing problem* (1.6) *under the DCMM model* (1.1)–(1.4). *Suppose the alternative hypothesis is true and the conditions* (2.1)–(2.2) *hold. As* $n \to \infty$, $\mathbb{E}[Q_n] = \text{tr}(\widetilde{\Omega}^4) + o((\lambda_2/\lambda_1)^4 \|\theta\|^8) + o(\|\theta\|^4)$ *and* $\text{Var}(Q_n) \le C[\|\theta\|^8 + (\lambda_2/\lambda_1)^6 \|\theta\|^8 \|\theta\|_3^6]$.

We conjecture that both T_n and Q_n are asymptotically normal under the alternative. In fact, asymptotic normality is easy to establish for the Ideal SgnT and Ideal SgnQ. To establish results for the real SgnT and real SgnQ, we need very precise characterization of the plug-in effect. For reasons of space, we leave them to the future.

Consider the SgnT test (2.4) first. By Theorem 2.3 and Lemma 2.1, under the alternative,

$$\text{the mean and variance of } \frac{T_n}{\sqrt{6(\|\hat{\eta}\|^2 - 1)^3}} \text{ are } \frac{\text{tr}(\widetilde{\Omega}^3)}{\sqrt{6\|\eta^*\|^6}} \text{ and } \sigma_n^2, \text{ respectively,} \tag{2.7}$$

where σ_n^2 denotes the asymptotic variance, which satisfies that

$$\sigma_n^2 \le \begin{cases} C, & \text{if } |\lambda_2/\lambda_1| \ll \sqrt{\|\theta\|/\|\theta\|_3^3}, \\ C(\lambda_2/\lambda_1)^4 \cdot (\|\theta\|_3^6/\|\theta\|^2), & \text{if } |\lambda_2/\lambda_1| \gg \sqrt{\|\theta\|/\|\theta\|_3^3}. \end{cases} \tag{2.8}$$

If we fix the degree heterogeneity vector θ and let (λ_2/λ_1) range, there is a *phase change* in the variance. We shall call:

- the case of $|\lambda_2/\lambda_1| \le C\sqrt{\|\theta\|/\|\theta\|_3^3}$ as the *weak signal* case for SgnT.
- the case of $|\lambda_2/\lambda_1| \gg \sqrt{\|\theta\|/\|\theta\|_3^3}$ as the *strong signal* case for SgnT.

It remains to derive a more explicit formula for $\text{tr}(\widetilde{\Omega}^3)$. Recall that λ_k and ξ_k are the kth eigenvalue and eigenvector of Ω, $1 \le k \le K$, respectively. Define $\Lambda \in \mathbb{R}^{(K-1)\times(K-1)}$ and

$h \in \mathbb{R}^{K-1}$ by $\Lambda = \mathrm{diag}(\lambda_2, \lambda_3, \dots, \lambda_K)$ and $h_k = (\mathbf{1}_n'\xi_{k+1})/(\mathbf{1}_n'\xi_1)$, $1 \le k \le K-1$. It can be shown that $\mathbf{1}_n'\xi_1 \neq 0$ and $\|h\|_\infty \le C$ so the vector h is well defined. In the special case of $\|h\|_\infty = o(1)$ (this happens when the angle between $\mathbf{1}_n$ and ξ_1 is small):

- We can show that $\mathrm{tr}(\widetilde{\Omega}^3) \approx \sum_{k=2}^K \lambda_k^3$.
- Motivated by these, we say "signal cancellation" happens when $|\mathrm{tr}(\widetilde{\Omega}^3)| \ll \sum_{k=2}^K |\lambda_k|^3$.

Therefore, "signal cancellation" may happen if the $(K-1)$ eigenvalues $\lambda_2, \lambda_3, \dots, \lambda_K$ have different signs. In fact, in the extreme case, we can have $\sum_{k=2}^K \lambda_k^3 = 0$, though $\sum_{k=2}^K |\lambda_k|^3$ is very large (e.g., [25], Section 3.3). Normally, the "signal cancellation" is found for odd-order moment-based statistics (e.g., $3rd$, $5th$, …, moment), but not for even-order moment methods (in fact, the SgnQ test will not experience such "signal cancellation").

Fortunately, "signal cancellation" is only possible when $\lambda_2, \lambda_3, \dots, \lambda_K$ have different signs, and can be avoided in some special cases. We propose the following conditions.

CONDITION 2.1. (a) $\lambda_2, \lambda_3, \dots, \lambda_K$ *have the same signs*, (b) $K = 2$ *and* (c) $|\lambda_2|/\lambda_1 \to 0$, *and* $|\mathrm{tr}(\Lambda^3) + 3h'\Lambda^3 h + 3(h'\Lambda h)(h'\Lambda^2 h) + (h'\Lambda h)^3| \ge C\sum_{k=2}^K |\lambda_k|^3$.

In (a)–(b), $\lambda_2, \dots, \lambda_K$ have the same signs. Condition (c) is based on more delicate analysis; see the proof of Lemma 2.2 for details.

While the above discussion is motivated by the case of $\|h\|_\infty = o(1)$, the idea continues to be valid for more general cases. The following is proved in the Supplementary Material.

LEMMA 2.2 (Analysis of $\mathrm{tr}(\widetilde{\Omega}^3)$). *Suppose conditions of Theorem 2.3 hold. Under the alternative hypothesis*,

- *If* $|\lambda_2|/\lambda_1 \to 0$, *then* $\mathrm{tr}(\widetilde{\Omega}^3) = \mathrm{tr}(\Lambda^3) + 3h'\Lambda^3 h + 3(h'\Lambda h)(h'\Lambda^2 h) + (h'\Lambda h)^3 + o(|\lambda_2|^3)$.
- *If* $\lambda_2, \lambda_3, \dots, \lambda_K$ *have the same signs, then*

$$\big|\mathrm{tr}(\widetilde{\Omega}^3)\big| \ge \begin{cases} \sum_{k=2}^K |\lambda_k|^3 + o\big(|\lambda_2|^3\big), & \text{if } |\lambda_2/\lambda_1| \to 0, \\ C|\lambda_2|^3, & \text{if } |\lambda_2/\lambda_1| \ge C. \end{cases}$$

- *In the special case where* $K = 2$, *the vector* h *is a scalar, and*

$$\big|\mathrm{tr}(\widetilde{\Omega}^3)\big| \begin{cases} = \big[(h^2+1)^3 + o(1)\big] \cdot |\lambda_2|^3, & \text{if } |\lambda_2|/\lambda_1 \to 0, \\ \ge C|\lambda_2|^3, & \text{if } |\lambda_2/\lambda_1| \ge C. \end{cases}$$

As a result, when either one of (*a*)–(*c*) *holds*, $|\mathrm{tr}(\widetilde{\Omega}^3)| \ge C\sum_{k=2}^K |\lambda_k|^3$.

It can be shown that $\|\eta^*\| \asymp \sqrt{\lambda_1} \asymp \|\theta\|$. We combine Lemma 2.2 with (2.7)–(2.8). In the *weak signal* case, $\frac{\mathbb{E}[T_n]}{\sqrt{\mathrm{Var}(T_n)}} \ge \frac{C(\sum_{k=2}^K |\lambda_k|^3)}{\|\theta\|^3} \ge C(\lambda_1^{-\frac{3}{2}} \sum_{k=2}^K |\lambda_k|^3)$. In the *strong signal* case, since $(\lambda_2/\lambda_1)^2 \le \lambda_1^{-2}(\sum_{k=2}^K |\lambda_k|^3)^{\frac{2}{3}}$, we have $\frac{\mathbb{E}[T_n]}{\sqrt{\mathrm{Var}(T_n)}} \ge \frac{C(\sum_{k=2}^K |\lambda_k|^3)}{\lambda_1^{-2}(\sum_{k=2}^K |\lambda_k|^3)^{\frac{2}{3}} \|\theta\|_3^3 \|\theta\|^2} \ge \frac{C\|\theta\|^3}{\|\theta\|_3^3}(\lambda_1^{-\frac{3}{2}} \sum_{k=2}^K |\lambda_k|^3)^{\frac{1}{3}}$, where it should be noted that in our setting, $\|\theta\|^3/\|\theta\|_3^3 \to \infty$. As a result, in both cases, the power of the SgnT test $\to 1$ as long as $\lambda_1^{-3/2}\sum_{k=2}^K |\lambda_k|^3 \to \infty$. This is validated in the following theorem, which is proved in the Supplemental Material.

THEOREM 2.5 (Power of the SgnT test). *Under the conditions of Theorem 2.3, for any fixed* $\alpha \in (0, 1)$, *consider the SgnT test in* (2.4). *Suppose one of the cases in Condition 2.1 holds. As* $n \to \infty$, *if* $\lambda_1^{-1/2}(\sum_{k=2}^K |\lambda_k|^3)^{1/3} \to \infty$, *then the Type I error* $\to \alpha$, *and the Type II error* $\to 0$.

Next, consider the SgnQ test (2.6). By Theorem 2.4 and Lemma 2.1, under the alternative, the mean and variance of $[Q_n - 2(\|\hat{\eta}\|^2-1)^2]/\sqrt{8(\|\hat{\eta}\|^2-1)^4}$ are $\mathrm{tr}(\widetilde{\Omega}^4)/\sqrt{8\|\eta^*\|^8}$ and σ_n^2, respectively, where σ_n^2 denotes the asymptotic variance and satisfies

$$\sigma_n^2 \le \begin{cases} C, & \text{if } |\lambda_2/\lambda_1| \ll \|\theta\|_3^{-1}, \\ C(\lambda_2/\lambda_1)^6 \cdot \|\theta\|_3^6, & \text{if } |\lambda_2/\lambda_1| \gg \|\theta\|_3^{-1}. \end{cases}$$

Similar to the SgnT test, if we fix the degree heterogeneity vector θ and let (λ_2/λ_1) range, there is a *phase change* in the variance. We shall call:

- the case of $|\lambda_2/\lambda_1| \le C\|\theta\|_3^{-1}$ as the *weak signal* case for SgnQ.
- the case of $|\lambda_2/\lambda_1| \gg \|\theta\|_3^{-1}$ as the *strong signal* case for SgnQ.

We now analyze $\mathrm{tr}(\widetilde{\Omega}^4)$. The following lemma is proved in the Supplementary Material.

LEMMA 2.3 (Analysis of $\mathrm{tr}(\widetilde{\Omega}^4)$). *Suppose the conditions of Theorem 2.4 hold. Under the alternative hypothesis*,

- *If* $|\lambda_2|/\lambda_1 \to 0$, *then* $\mathrm{tr}(\widetilde{\Omega}^4) = \mathrm{tr}(\Lambda^4) + (q'\Lambda q)^4 + 2(h'\Lambda^2 h)^2 + 4(h'\Lambda h)^2(h'\Lambda^2 h) + 4h'\Lambda^4 h + 4(h'\Lambda h)(h'\Lambda^3 h) + o(\lambda_2^4) \gtrsim \sum_{k=2}^4 \lambda_k^4$.
- *If* $|\lambda_2|/\lambda_1 \ge C$, *then* $\mathrm{tr}(\widetilde{\Omega}^4) \ge C\sum_{k=2}^K \lambda_k^4$.
- *In the special case of* $K = 2$, h *is a scalar and* $\mathrm{tr}(\widetilde{\Omega}^4) = [(h^2+1)^4 + o(1)] \cdot \lambda_2^4$.

As a result, the SgnQ test has no issue of "signal cancellation," and it always holds that $\mathrm{tr}(\widetilde{\Omega}^4) \ge C\sum_{k=2}^K \lambda_k^4$. Then, in the *weak signal* case, we have $\frac{\mathbb{E}[Q_n]}{\sqrt{\mathrm{Var}(Q_n)}} \ge \frac{C(\sum_{k=2}^K \lambda_k^4)}{\|\theta\|^4} \ge C(\lambda_1^{-2}\sum_{k=2}^K \lambda_k^4)$. In the *strong signal* case, since $(\lambda_2/\lambda_1)^3 \le \lambda_1^{-3}(\sum_{k=2}^K \lambda_k^4)^{\frac{3}{4}}$, we have $\frac{\mathbb{E}[Q_n]}{\sqrt{\mathrm{Var}(Q_n)}} \ge \frac{C(\sum_{k=2}^K \lambda_k^4)}{\lambda_1^{-3}(\sum_{k=2}^K \lambda_k^4)^{\frac{3}{4}}\|\theta\|_3^3\|\theta\|^4} \ge \frac{C\|\theta\|^3}{\|\theta\|_3^3}(\lambda_1^{-2}\sum_{k=2}^K \lambda_k^4)^{\frac{1}{4}}$, where $\|\theta\|^3/\|\theta\|_3^3 \to \infty$. So, in both cases, the power of the SgnQ test goes to 1 if $\lambda_1^{-2}\sum_{k=2}^K \lambda_k^4 \to \infty$. This is validated in Theorem 2.6, which is proved in the Supplemental Material.

THEOREM 2.6 (Power of the SgnQ test). *Under the conditions of Theorem* 2.4, *for any fixed* $\alpha \in (0,1)$, *consider the SgnQ test in* (2.6). *As* $n \to \infty$, *if* $\lambda_1^{-1/2}(\sum_{k=2}^K \lambda_k^4)^{1/4} \to \infty$, *then the Type I error* $\to \alpha$, *and the Type II error* $\to 0$.

In summary, Theorem 2.5 and Theorem 2.6 imply that as long as

$$|\lambda_2|/\sqrt{\lambda_1} \to \infty, \tag{2.9}$$

the levels of SgnT and SgnQ tests tend to α as expected, and their powers tend to 1. The SgnT test requires mild conditions to avoid "signal cancellation," but the SgnQ test has no such issue (such an advantage of SgnQ test is confirmed by numerical study in Section 5).

REMARK 7. Practically, we prefer to fix α, say, $\alpha = 5\%$. If we allow the level α to change with n, then when (2.9) holds, there is a sequence of α_n that tends to 0 slowly enough such that $|\lambda_2|/(z_{\alpha_n/2}\cdot\sqrt{\lambda_1}) \to \infty$. As a result, for either of the two tests, the Type I error $\to 0$ and the power $\to 1$, so the sum of Type I and Type II errors $\to 0$.

EXAMPLE 1 (contd). For this example, $\lambda_1 \sim (1+(K-1)b_n)\|\theta\|^2$, and $\lambda_k \sim (1-b_n)\|\theta\|^2$, $k = 2, 3, \dots, K$. The condition (2.9) of $|\lambda_2|/\sqrt{\lambda_1} \to \infty$ translates to $(1-b_n)\|\theta\| \to \infty$. See Section 1.2 and also Section 3 for more discussion.

3. Optimal adaptivity, lower bound and region of impossibility. We now focus on the region of impossibility, where $|\lambda_2|/\sqrt{\lambda_1} \to 0$. We first present a standard minimax lower bound, from which we can conclude that there is a sequence of hypothesis pairs (one alternative and one null) that are asymptotically indistinguishable. However, this does not answer the question whether *all alternatives* in the region of impossibility are indistinguishable from the null. To answer this question, we need much more sophisticated study; see Section 3.2.

3.1. *Minimax lower bound.* Given an integer $K \geq 1$, a constant $c_0 > 0$, and two positive sequences $\{\alpha_n\}_{n=1}^{\infty}$ and $\{\beta_n\}_{n=1}^{\infty}$, we define a class of parameters for DCMM (recall that $\Omega = \Theta\Pi P\Pi'\Theta$, $G = \|\theta\|^{-2}\Pi'\Theta^2\Pi$ and is properly scaled, and λ_k is the kth largest eigenvalue of Ω in magnitude):

$$\begin{aligned}&\mathcal{M}_n(K, c_0, \alpha_n, \beta_n)\\&\quad = \left\{\begin{array}{c}(\theta, \Pi, P) : \theta_{\max} \leq \beta_n, \|\theta\|^{-1} \leq \beta_n, \|\theta\|^2\|\theta\|_1^{-1}\sqrt{\log(\|\theta\|_1)} \leq \beta_n,\\ \dfrac{\max_k\{\sum_{i=1}^n \theta_i\pi_i(k)\}}{\min_k\{\sum_{i=1}^n \theta_i\pi_i(k)\}} \leq c_0, \|G^{-1}\| \leq c_0, |\lambda_2|/\sqrt{\lambda_1} \geq \alpha_n\end{array}\right\}.\end{aligned}$$

For the null case, $K = P = \pi_i = 1$, and the above defines a class of θ, which we write for short by $\mathcal{M}_n(1, c_0, \alpha_n, \beta_n) = \mathcal{M}_n^*(\beta_n)$.

THEOREM 3.1 (Minimax lower bound). *Fix $K \geq 2$, a constant $c_0 > 0$ and any sequences $\{\alpha_n\}_{n=1}^{\infty}$ and $\{\beta_n\}_{n=1}^{\infty}$ such that $\alpha_n \to 0$ and $\beta_n \to 0$ as $n \to \infty$. Then, as $n \to \infty$,*

$$\inf_{\psi}\Big\{\sup_{\theta \in \mathcal{M}_n^*(\beta_n)} \mathbb{P}(\psi = 1) + \sup_{(\theta,\Pi,P)\in\mathcal{M}_n(K,c_0,\alpha_n,\beta_n)} \mathbb{P}(\psi = 0)\Big\} \to 1,$$

where the infimum is taken over all possible tests ψ.

Theorem 3.1 says that in the region of impossibility, *there exists a sequence of alternatives* that are inseparable from the null. This does not show what we desire, that is *any sequence in the region of impossibility* is inseparable from the null. This is discussed in the next section.

3.2. *Region of impossibility.* Recall that under DCMM, $\Omega = \Theta\Pi P\Pi'\Theta$ and $\Pi = [\pi_1, \pi_2, \ldots, \pi_n]'$. Since our model is a mixed-membership latent variable model, in order to characterize the *least favorable configuration*, it is conventional to use a *random mixed-membership (RMM) model* for the matrix Π, while (Θ, P) are still nonstochastic. In detail,

- Let $V = \{x \in \mathbb{R}^K, x_k \geq 0, \sum_{k=1}^K x_k = 1\}$.
- Let $V_0 = \{e_1, e_2, \ldots, e_K\}$, where e_k is the kth Euclidean basis vector.

In DCMM–RMM, we fix a distribution F defined over V and assume $\pi_i \overset{\text{i.i.d.}}{\sim} F$ where $h \equiv \mathbb{E}[\pi_i]$. If we further restrict that F is defined over V_0, then the network has no mixed membership, and DCMM–RMM reduces to DCBM–RMM.

The desired result is to show that, for any given P and F, there is a sequence of hypothesis pairs (a null and an alternative)

$$H_0^{(n)} : \Omega = \theta\theta', \quad \text{and} \quad H_1^{(n)} : \Omega = \widetilde{\Theta}\Pi P\Pi'\widetilde{\Theta}, \tag{3.1}$$

where $\widetilde{\Theta} = \mathrm{diag}(\widetilde{\theta}_1, \widetilde{\theta}_1, \ldots, \widetilde{\theta}_n)$ and $\widetilde{\theta}_i$ can be different from θ_i, such that the two hypotheses within each pair are asymptotically indistinguishable from each other, provided that under the alternative $|\lambda_2|/\sqrt{\lambda_1} \to 0$.

Here, since Ω depends on π_i, λ_k is random, and it is more convenient to translate the condition of $|\lambda_2|/\sqrt{\lambda_1} \to 0$ to the condition of

$$\|\theta\| \cdot \left|\mu_2(P)\right| \to 0, \tag{3.2}$$

where $\mu_k(P)$ is the kth largest eigenvalue of P in magnitude. The equivalence of two conditions are justified in Section F.1 of the Supplementary Material. Condition (2.2) can also be ensured with high probability, by assuming that all entries of $\mathbb{E}[\pi_i]$ are at the order of $O(1)$.

Under the DCBM, the desired result can be proved satisfactorily. The key is the following lemma, which is in the line of Sinkhorn's beautiful work on scalable matrices [41] (see also [9, 28, 35]) and is proved in the Supplementary Material.

LEMMA 3.1. *Fix a matrix $A \in \mathbb{R}^{K,K}$ with strictly positive diagonal entries and nonnegative off-diagonal entries, and a strictly positive vector $h \in \mathbb{R}^K$, there exists a diagonal matrix $D = \mathrm{diag}(d_1, d_2, \dots, d_K)$ such that $DADh = 1_K$ and $d_k > 0$, $1 \le k \le K$.*

In detail, consider a DCBM–RMM setting where $\pi_i \overset{\text{i.i.d.}}{\sim} F$ and F is supported over V_0 (with possibly unequal probabilities on the K points). Recall $h = \mathbb{E}[\pi_i]$. By Lemma 3.1, there is a unique diagonal matrix D such that $DPDh = 1_K$. Let

$$\widetilde{\theta}_i = d_k \cdot \theta_i, \quad \text{if } \pi_i = e_k, 1 \le i \le n, 1 \le k \le K. \tag{3.3}$$

The following theorem is proved in the Supplementary Material.

THEOREM 3.2 (Region of impossibility (DCBM)). *Fix $K > 1$ and a distribution F defined over V_0. Consider a sequence of DCBM model pairs indexed by n:*

$$H_0^{(n)} : \Omega = \theta\theta' \quad \text{and} \quad H_1^{(n)} : \Omega = \widetilde{\Theta}\Pi P \Pi'\widetilde{\Theta},$$

where $\pi_i \overset{\text{i.i.d.}}{\sim} F$ and $\widetilde{\Theta} = \mathrm{diag}(\widetilde{\theta}_1, \widetilde{\theta}_2, \dots, \widetilde{\theta}_n)$ with $\widetilde{\theta}_i$ defined as in (3.3). If $\theta_{\max} \le c_0$ for a constant $c_0 < 1$, $\min_{1\le k\le K}\{h_k\} \ge C$, and $\|\theta\| \cdot |\mu_2(P)| \to 0$, then for each pair of two hypotheses, the χ^2-distance between the two joint distributions tends to 0, as $n \to \infty$.

To generalize this to RMM–DCMM, we fix a distribution F defined over V. Given a set of (Θ, P, Π) with $\Theta = \mathrm{diag}(\theta_1, \theta_2, \dots, \theta_n)$ and $\pi_i \overset{\text{i.i.d.}}{\sim} F$, let $\widetilde{h}_D = \mathbb{E}[D^{-1}\pi_i/\|D^{-1}\pi_i\|_1]$ for any diagonal matrix $D \in \mathbb{R}^{K\times K}$ with positive diagonals. We assume that there is a D such that

$$DPD\widetilde{h}_D = 1_K, \qquad \min_{1\le k\le K}\{\widetilde{h}_{D,k}\} \ge C. \tag{3.4}$$

When such a D exists, we let

$$\widetilde{\theta}_i = \theta_i/\left\|D^{-1}\pi_i\right\|_1, \quad 1 \le i \le n. \tag{3.5}$$

When the support of F is restricted to V_0, all realizations of π_i are degenerate (i.e., one entry is 1, and other entries are 0), so $\widetilde{h}_D = h$, $\widetilde{\theta}_i$ is the same as that in (3.3), and (3.4) holds by Lemma 3.1. Under DCMM–RMM, π_i's are not degenerate. We conjecture that (3.4) continues to hold generally (we can show it for the cases of $K = 2, 3$; the proof is elementary so is omitted). The following theorem is proved in the Supplementary Material.

THEOREM 3.3 (Region of Impossibility (DCMM)). *Fix $K > 1$ and a distribution F defined over V. Consider a sequence of DCMM model pairs indexed by n:*

$$H_0^{(n)} : \Omega = \theta\theta' \quad \text{and} \quad H_1^{(n)} : \Omega = \widetilde{\Theta}\Pi P \Pi'\widetilde{\Theta},$$

where $\pi_i \overset{\text{iid}}{\sim} F$ *and* $\widetilde{\Theta} = \text{diag}(\widetilde{\theta}_1, \widetilde{\theta}_2, \dots, \widetilde{\theta}_n)$ *with* $\widetilde{\theta}_i$ *defined as in* (3.5). *If* (3.4) *holds*, $\theta_{\max} \leq c_0$ *for a constant* $c_0 < 1$, *and* $\|\theta\| \cdot |\mu_2(P)| \to 0$, *then for each pair of two hypotheses*, *the* χ^2*-distance between the two joint distributions tends to* 0, *as* $n \to \infty$.

One of the main strengths of Theorems 3.2–3.3 is that this lower bound is valid for an arbitrary choice of $\theta \in \mathbb{R}^n_+$. This is stronger than the standard minimax lower bound.

In Theorem 3.3, we try to be as general as we can so Π is given (and we are not allowed to change it in our construction). For any P and F, by Lemma 3.1, there is a unique positive diagonal matrix D such that $DPDh = 1_K$ where $h = \mathbb{E}[\pi_i]$. We now consider a special case where we allow Π to depend on D in our construction. In this case, Condition (3.4) can be removed. Let $\widetilde{\Pi} = [\widetilde{\pi}_1, \widetilde{\pi}_2, \dots, \widetilde{\pi}_n]'$ and $\widetilde{\Theta} = \text{diag}(\widetilde{\theta}_1, \widetilde{\theta}_2, \dots, \widetilde{\theta}_n)$, with

$$\widetilde{\pi}_i = D\pi_i / \|D\pi_i\|_1, \qquad \widetilde{\theta}_i = \|D\pi_i\|_1 \cdot \theta_i. \tag{3.6}$$

THEOREM 3.4 (Region of impossibility (DCMM with flexible Π)). *Fix* $K > 1$ *and a distribution* F *defined over* V. *Consider a sequence of DCMM model pairs indexed by* n: $H_0^{(n)} : \Omega = \theta\theta'$ *and* $H_1^{(n)} : \Omega = \widetilde{\Theta}\widetilde{\Pi}P\widetilde{\Pi}'\widetilde{\Theta}$, *where* $\widetilde{\Pi}$ *and* $\widetilde{\Theta}$ *are defined as in* (3.6). *If* $\theta_{\max} \leq c_0$ *for a constant* $c_0 < 1$, $\min_{1 \leq k \leq K}\{h_k\} \geq C$, *and* $\|\theta\| \cdot |\mu_2(P)| \to 0$, *then for each pair of two hypotheses*, *the* χ^2*-distance between the two joint distributions tends to* 0, *as* $n \to \infty$.

Finally, we consider the case where we require that the null and the alternative have perfectly matching Θ matrix (up to an overall scaling). This is especially of interest when we consider SBM or MMSBM models where we have little freedom in choosing the Θ matrix. In this case, in order that the two hypotheses are indistinguishable, the expected node degrees under the alternative have to match those under the null. For each $1 \leq i \leq n$, conditional on π_i and neglecting the effect of no self-edges, the expected degree of node i equals to $\|\theta\|_1 \cdot \theta_i$ and $\|\theta\|_1 \cdot (\pi_i' Ph) \cdot \theta_i$ under the null and under the alternative, respectively, where $\{\pi_j\}_{j \neq i} \overset{\text{iid}}{\sim} F$ and $h = \mathbb{E}[\pi_j]$. For the expected degrees to match under any realized π_i, it is necessary that

$$Ph = q_n 1_K, \quad \text{for some scaling parameter } q_n > 0. \tag{3.7}$$

THEOREM 3.5 (Region of impossibility (DCMM with matching Θ)). *Fix* $K > 1$ *and a distribution* F *defined over* V. *Consider a sequence of DCMM model pairs indexed by* n: $H_0^{(n)} : \Omega = q_n \cdot \theta\theta'$ *and* $H_1^{(n)} : \Omega = \Theta\Pi P\Pi'\Theta$, *where* $\Theta = \text{diag}(\theta_1, \theta_2, \dots, \theta_n)$, $\pi_i \overset{\text{iid}}{\sim} F$, *and* (P, h, q_n) *satisfy* (3.7). *If* $\theta_{\max} \leq c_0$ *for a constant* $c_0 < 1$, $\min_{1 \leq k \leq K}\{h_k\} \geq C$ *and* $\|\theta\| \cdot |\mu_2(P)| \to 0$, *then for each pair of two hypotheses*, *the* χ^2*-distance between the two joint distributions tends to* 0, *as* $n \to \infty$.

Theorems 3.4–3.5 are proved in the Supplementary Material.

EXAMPLE 1 (continued). In Example 1, π_i is drawn from $e_1, e_2, \dots, e_K$ with equal probabilities, and $P = (1 - b_n)I_K + b_n 1_K 1_K'$. Therefore, $h = \mathbb{E}[\pi_i] = (1/K)1_K$. In this case, all conditions of Theorem 3.5 hold. Note $q_n = (1/K) + (K-1)b_n/K$ and $\mu_2(P) = (1 - b_n)$.

REMARK 8 (Least favorable configuration of LDA-DCMM). The Dirichlet model is often used for mixed memberships [1]. Consider the model pairs $H_0^{(n)} : \Omega = q_n\theta\theta'$ and $H_1^{(n)} : \Omega = \Theta\Pi P\Pi'\Theta$ and where $\pi_i \overset{\text{iid}}{\sim} \text{Dir}(\alpha)$ ($\text{Dir}(\alpha)$: Dirichlet distribution with parameters $\alpha = (\alpha_1, \dots, \alpha_K)'$). By Theorem 3.5, as long as $P\alpha \propto 1_K$, the null and alternative hypotheses are asymptotically indistinguishable if $(1 - q_n)\|\theta\| \to 0$. One can easily construct P such that $P\alpha \propto 1_K$. For example, $P = (1 - q_n)MM' + q_n 1_K 1_K'$, where $M \in \mathbb{R}^{K \times (K-1)}$ is a matrix whose columns are from $\text{Span}^{\perp}(\alpha)$ and satisfy $\text{diag}(MM') = I_K$.

3.3. *Optimal adaptivity.* Recall that $\sqrt{\lambda_1}$, $|\lambda_2|/\lambda_1$, and $|\lambda_2|/\sqrt{\lambda_1}$ can be viewed as a measure for the sparsity, community dissimilarity and SNR, respectively. Combining Theorems 2.1–2.4, Theorems 3.2–3.5 and Remark 7 in Section 2.3, in the two-dimensional phase space where the x-axis is $\sqrt{\lambda_1}$ and the y-axis is the $|\lambda_2|/\lambda_1$, we have a partition to two regions, the region of possibility and the region of impossibility.

- *Region of impossibility* ($1 \ll \sqrt{\lambda_1} \ll \sqrt{n}$, $|\lambda_2|/\sqrt{\lambda_1} = o(1)$). In this region, any DCBM alternative is asymptotically inseparable from the null, and up to a mild condition, any DCMM alternative is also asymptotically inseparable from the null.
- *Region of possibility* ($1 \ll \sqrt{\lambda_1} \ll \sqrt{n}$, $|\lambda_2|/\sqrt{\lambda_1} \to \infty$). In this region, asymptotically, any alternative is completely separable form any null.

The SgnQ test is optimally adaptive: for any alternative in the region of possibility, the test is able to separate it from the null with a sum of Type I and Type II errors tending to 0. The SgnT test is also optimally adaptive, provided that some mild conditions hold to avoid signal cancellation. To the best of our knowledge, the Signed Polygon is the only known test that is both applicable to general DCMM (where we allow severe degree heterogeneity and arbitrary mixed memberships) and optimally adaptive. The EZ and GC tests are the only other tests we know that are applicable to general DCMM, but their variances are unsatisfactorily large for the less sparse case, so they are not optimally adaptive. See [30] for details.

REMARK 9. Most existing lower bound results [2, 16, 37] are within the standard minimax framework, where they focus on a particular sequence of alternative (e.g., the off-diagonals of P are equal). In our case, the standard minimax theorem only implies that in the region of impossibility, there is a sequence of alternative that are inseparable from the null. Our results (Theorems 3.2–3.5) shed new light on the region of impossibility, saying that for each alternative, we can pair it with a null such that two hypotheses are asymptotically inseparable.

REMARK 10. Existing minimax lower bounds [2, 4, 37] are largely focused on the SBM. Though a least favorable scenario for SBM is least favorable for DCMM, the former does not provide much insight on how the least favorable configurations and the phase transition depend on the degree heterogeneity and mixed memberships. Moreover, our results (see also [19]) suggest that $\|\theta\|$, not $\|\theta\|_1$, determines the separating boundary. In the SBM case, $\theta_1 = \cdots = \theta_n$ and $\|\theta\|_1 = \sqrt{n}\|\theta\|$, so it is hard to tell which of the two norms decides the boundary. In DCMM, there is no simple relationship between $\|\theta\|_1$ and $\|\theta\|$, and we can tell this clearly.

REMARK 11. A sharper version of the phase transition is that there exists a constant $c_0 > 0$ such that the region of possibility and region of impossibility are given by $|\lambda_2|/\sqrt{\lambda_1} > c_0$ and $|\lambda_2|/\sqrt{\lambda_1} < c_0$, respectively. In some special cases, these kinds of results exist for community detection (a related but different problem). For example, [19] considered a setting where (i) there is no mixed membership, (ii) for some constants $a, b > 0$, $P(k, \ell) = a$ if $k = \ell$ and b otherwise, (iii) the communities have equal size and (iv) for a constant $\phi > 0$, $\{\sqrt{n}\theta_i\}_{i=1}^n$ are i.i.d. drawn from a fixed distribution supported in $[\phi, \infty)$. They showed that, when $(a - b)^2\mathbb{E}\|\theta\|^2 < K(a + b)$, it is impossible to reconstruct the community label matrix Π. Moreover, in the special case of $K = 2$, [18] (also, see [12]) showed that when $(a - b)^2\mathbb{E}\|\theta\|^2 > 2(a + b)$, it is possible to construct an estimate of Π that is positively correlated with the true community labels. By connecting $(a, b, \mathbb{E}\|\theta\|^2)$ with eigenvalues, it is seen that these results give a sharp phase transition at $c_0 = 1$, in the special case where $K = 2$ and (i)–(iv) hold. For more general settings, whether such a sharp phase transition exists is unclear: a slight change in conditions (i)–(iv) may affect the lower bounds, and the optimal tests (for the sharp phase transition) are hard to find as they usually need to adapt to specific features of the model. Also, technically, allowing for mixed memberships makes the lower bound much harder to study, and allowing for unequal community sizes and unequal off-diagonal entries in P requires an application of DAD theorem in lower bound construction (which is not needed in [19]). Moreover, [12, 18, 19] are for community detection and our paper is on global testing. For general DCMM settings, it is unclear whether the phase transitions for two problems are the same.

4. The behavior of the SgnQ test statistics. In this section, we study the SgnQ test statistic Q_n and explain how to prove Theorems 2.2, 2.4 and 2.6. We introduce a proxy SgnQ test statistic Q_n^* and an Ideal SgnQ test statistic $\widetilde{Q}_n$. Writing $Q_n = \widetilde{Q}_n + (Q_n^* - \widetilde{Q}_n) + (Q_n - Q_n^*)$, we study the three terms on the RHS in Sections 4.1–4.3, respectively. Given these results, the proofs of Theorems 2.2, 2.4 and 2.6 are straightforward and contained in Section B of the Supplementary Material. The study of the SgnT test statistic T_n is similar and contained in Section A of the Supplementary Material, where we also prove Theorems 2.1, 2.3 and 2.5.

Recall that the SgnQ statistic Q_n is defined as

$$Q_n = \sum_{i_1,i_2,i_3,i_4(dist)} (A_{i_1i_2} - \hat{\eta}_{i_1}\hat{\eta}_{i_2})(A_{i_2i_3} - \hat{\eta}_{i_2}\hat{\eta}_{i_3})(A_{i_3i_4} - \hat{\eta}_{i_3}\hat{\eta}_{i_4})(A_{i_4i_1} - \hat{\eta}_{i_4}\hat{\eta}_{i_1}),$$

where $\hat{\eta} = A\mathbf{1}_n/\sqrt{V}$, with $V = \mathbf{1}_n' A\mathbf{1}_n$. In Section 1.4, we have introduced the following nonstochastic proxy of $\hat{\eta}$: $\eta^* = \Omega\mathbf{1}_n/\sqrt{v_0}$, where $v_0 = \mathbf{1}_n'\Omega\mathbf{1}_n$. We now introduce another (stochastic) proxy $\tilde{\eta}$ by

$$\tilde{\eta} = A\mathbf{1}_n/\sqrt{v}, \quad \text{where } v = \mathbb{E}[\mathbf{1}_n' A\mathbf{1}_n] = \mathbf{1}_n'(\Omega - \text{diag}(\Omega))\mathbf{1}_n. \tag{4.1}$$

Denoting the mean of $\tilde{\eta}$ by η, it is seen that

$$\eta = ([\Omega - \text{diag}(\Omega)]\mathbf{1}_n)/\sqrt{\mathbf{1}_n'(\Omega - \text{diag}(\Omega))\mathbf{1}_n}. \tag{4.2}$$

Here, η and η^* are close to each other but η^* has a more explicit form. For example, under the null hypothesis, $\Omega = \theta\theta'$, and it is seen that $\eta^* = \theta$. Recall that $A = \Omega - \text{diag}(\Omega) + W$ and $\widetilde{\Omega} = \Omega - \eta^*(\eta^*)'$. Fix $1 \le i, j \le n$ and $i \ne j$. First, we write

$$A_{ij} - \hat{\eta}_i\hat{\eta}_j = (A_{ij} - \eta_i^*\eta_j^*) + (\eta_i^*\eta_j^* - \hat{\eta}_i\hat{\eta}_j) = \widetilde{\Omega}_{ij} + W_{ij} + (\eta_i^*\eta_j^* - \hat{\eta}_i\hat{\eta}_j).$$

Second, we write $\eta_i^*\eta_j^* - \hat{\eta}_i\hat{\eta}_j = \delta_{ij} + r_{ij}$, where

$$\delta_{ij} = \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_i - \tilde{\eta}_i) \tag{4.3}$$

is the linear approximation term of $(\eta_i^*\eta_j^* - \hat{\eta}_i\hat{\eta}_j)$ and $r_{ij} \equiv (\eta_i^*\eta_j^* - \hat{\eta}_i\hat{\eta}_j) - \delta_{ij}$ is the remainder term. By definition and elementary algebra,

$$r_{ij} = (\eta_i^*\eta_j^* - \eta_i\eta_j) - (\eta_i - \tilde{\eta}_i)(\eta_j - \tilde{\eta}_j) + \left(1 - \frac{v}{V}\right)\tilde{\eta}_i\tilde{\eta}_j. \tag{4.4}$$

It is shown that r_{ij} is of a smaller order than that of δ_{ij}. The remainder term can be shown to have a negligible effect over T_n and Q_n, in terms of the variances of T_n and Q_n, respectively; see Theorem 4.3.

Let X be the symmetric matrix where all diagonal entries are 0 and for $1 \le i, j \le n$ but $i \ne j$, $X_{ij} = A_{ij} - \hat{\eta}_i\hat{\eta}_j$, or equivalently,

$$X_{ij} = \widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij} + r_{ij}. \tag{4.5}$$

If we omit the remainder term, then we have a proxy of X, denoted by X^*, where all diagonal entries of X^* are 0, and for $1 \le i, j \le n$ but $i \ne j$,

$$X_{ij}^* = \widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij}. \tag{4.6}$$

If we further omit the δ term, then we have another proxy of X, denoted by $\widetilde{X}$, where all diagonal entries of $\widetilde{X}$ are 0, and for $1 \le i, j \le n$ but $i \neq j$,

$$\widetilde{X}_{ij} = \widetilde{\Omega}_{ij} + W_{ij}. \tag{4.7}$$

With the above notation, we can rewrite Q_n as $Q_n = \sum_{i_1,i_2,i_3,i_4(dist)} X_{i_1 i_2} X_{i_2 i_3} X_{i_3 i_4} X_{i_4 i_1}$. We introduce the *Proxy SgnQ test statistic* and *Ideal SgnQ test statistic* by

$$Q_n^* = \sum_{i_1,i_2,i_3,i_4(dist)} X^*_{i_1 i_2} X^*_{i_2 i_3} X^*_{i_3 i_4} X^*_{i_4 i_1}, \qquad \widetilde{Q}_n = \sum_{i_1,i_2,i_3,i_4(dist)} \widetilde{X}_{i_1 i_2} \widetilde{X}_{i_2 i_3} \widetilde{X}_{i_3 i_4} \widetilde{X}_{i_4 i_1}.$$

The Ideal SgnQ test statistic $\widetilde{Q}_n$ is the same as that defined in (1.13). Using these notation, we partition Q_n as $Q_n = \widetilde{Q}_n + (Q_n^* - \widetilde{Q}_n) + (Q_n - Q_n^*)$. In Sections 4.1–4.3, we study the three terms on the right-hand side, respectively.

4.1. *The behavior of the ideal SgnQ test statistics.* In view of (4.7), the Ideal SgnQ test statistic $\widetilde{Q}_n$ is written as

$$\widetilde{Q}_n = \sum_{i_1,i_2,i_3,i_4(dist)} (\widetilde{\Omega}_{i_1 i_2} + W_{i_1 i_2})(\widetilde{\Omega}_{i_2 i_3} + W_{i_2 i_3})(\widetilde{\Omega}_{i_3 1 i_4} + W_{i_3 i_4})(\widetilde{\Omega}_{i_4 i_1} + W_{i_4 i_1}). \tag{4.8}$$

Under the null, $\Omega = \theta\theta'$ and $\eta^* = \theta$. By definition, $\widetilde{\Omega}_{ij} = 0$, and the statistic reduces to $\widetilde{Q}_n = \sum_{i_1,i_2,i_3,i_4(dist)} W_{i_1 i_2} W_{i_2 i_3} W_{i_3 i_4} W_{i_4 i_1}$. The right-hand side is the sum of a large number of uncorrelated terms, with each term being a 4-product of independent centered-Bernoulli variables. It can be shown that the statistic is asymptotically normal, with $\mathbb{E}[\widetilde{Q}_n] = 0$ and $\mathrm{Var}(\widetilde{Q}_n) \sim 8\|\theta\|^8$.

Consider the alternative hypothesis. In the right-hand side of (4.8), expanding the bracket and rearranging, we have $2 \times 2 \times 2 \times 2 = 16$ post-expansion sums, each having the form of $\sum_{i_1,i_2,i_3,i_4(dist)} a_{i_1 i_2} b_{i_2 i_3} c_{i_3 i_4} d_{i_4 i_1}$, where a is a generic notation which may either equal to $\widetilde{\Omega}$ or W; same for b, c and (d). For example, $\sum_{i_1,i_2,i_3,i_4(dist)} W_{i_1 i_2} \widetilde{\Omega}_{i_2 i_3} W_{i_3 i_4} W_{i_4 i_1}$ is one of the 16 post-expansion sums, corresponding to $b = \widetilde{\Omega}$, and $a = c = d = W$. Note that each of 16 post-expansion sums is the sum of many 4-product, where the number of the $\widetilde{\Omega}$ factors in each product is the same; denote this number (which can be 0, 1, 2, 3 or 4) by $N_{\widetilde{\Omega}}$. Similarly, the number of the W factors in each product are also the same. Denote it by N_W, we have $N_{\widetilde{\Omega}} + N_W = 4$. For the example above, $(N_{\widetilde{\Omega}}, N_W) = (1, 3)$.

According to $(N_{\widetilde{\Omega}}, N_W)$, we can group the 16 post-expansion sums into 6 different types. Table 1 presents the mean and variance of each type (Recall that $\lambda_1, \ldots, \lambda_K$ are the K eigenvalues of Ω, arranged in descending order in magnitude. In Table 1, $\alpha = |\lambda_2|/\lambda_1$. In the alternative, we assume $|\lambda_2|/\sqrt{\lambda_1} \to \infty$, which translates to $\alpha\|\theta\| \to \infty$ since $\sqrt{\lambda_1} \asymp \|\theta\|$).

TABLE 1
The 6 different types of the 16 post-expansion sums of $\widetilde{Q}_n$ ($\|\theta\|_q$ is the ℓ^q-norm of θ (the subscript is dropped when $q = 2$). In our setting, $\alpha\|\theta\| \to \infty$, and $\|\theta\|_4^4 \ll \|\theta\|_3^3 \ll \|\theta\|^2 \ll \|\theta\|_1$

Type	#	$(N_{\widetilde{\Omega}}, N_W)$	Examples	Mean	Variance
I	1	(0, 4)	$\sum_{i,j,k,\ell(dist)} W_{ij} W_{jk} W_{k\ell} W_{\ell i}$	0	$\asymp \|\theta\|^8$
II	4	(1, 3)	$\sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij} W_{jk} W_{k\ell} W_{\ell i}$	0	$\le C\alpha^2 \|\theta\|^4 \|\theta\|_3^6 = o(\|\theta\|^8)$
IIIa	4	(2, 2)	$\sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij} \widetilde{\Omega}_{jk} W_{k\ell} W_{\ell i}$	0	$\le C\alpha^4 \|\theta\|^6 \|\theta\|_3^6 = o(\alpha^6 \|\theta\|^8 \|\theta\|_3^6)$
IIIb	2	(2, 2)	$\sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij} W_{jk} \widetilde{\Omega}_{k\ell} W_{\ell i}$	0	$\le C\alpha^4 \|\theta\|_3^{12} = o(\|\theta\|^8)$
IV	4	(3, 1)	$\sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij} \widetilde{\Omega}_{jk} \widetilde{\Omega}_{k\ell} W_{\ell i}$	0	$\le \alpha^6 \|\theta\|^8 \|\theta\|_3^6$
V	1	(4, 0)	$\sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij} \widetilde{\Omega}_{jk} \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i}$	$\sim \mathrm{tr}(\widetilde{\Omega}^4)$	0

From the table, among all 16 post-expansion sums, the total mean is $\sim \mathrm{tr}(\widetilde{\Omega}^4)$, and the total variance $\le C\|\theta\|^8 + C(|\lambda_2|/\lambda_1)^6\|\theta\|^8\|\theta\|_3^6$, with Type I sum and Type IV sum being the major contributors. The following theorem is proved in the Supplementary Material.

THEOREM 4.1 (Ideal SgnQ test statistic). *Consider the testing problem* (1.6) *under the DCMM model* (1.1)–(1.4), *where the condition* (2.2) *is satisfied under the alternative hypothesis. Suppose* $\theta_{\max} \to 0$ *and* $\|\theta\| \to \infty$ *as* $n \to \infty$, *and suppose* $|\lambda_2|/\sqrt{\lambda_1} \to \infty$ *under the alternative hypothesis. Then, under the null hypothesis, as* $n \to \infty$, $\mathbb{E}[\widetilde{Q}_n] = 0$, $\mathrm{Var}(\widetilde{Q}_n) = 8\|\theta\|^8 \cdot [1 + o(1)]$, *and* $(\widetilde{Q}_n - \mathbb{E}[\widetilde{Q}_n])/\sqrt{\mathrm{Var}(\widetilde{Q}_n)} \longrightarrow N(0,1)$ *in law. Furthermore, under the alternative hypothesis, as* $n \to \infty$, $\mathbb{E}[\widetilde{Q}_n] = \mathrm{tr}(\widetilde{\Omega}^4) + o(\|\theta\|^4)$ *and* $\mathrm{Var}(\widetilde{T}_n) \le C[\|\theta\|^8 + (|\lambda_2|/\lambda_1)^6\|\theta\|^8\|\theta\|_3^6]$.

4.2. *The behavior of* $(Q_n^* - \widetilde{Q}_n)$. The Proxy SgnQ test statistic is defined as $Q_n^* = \sum_{i_1,i_2,i_3,i_4(dist)} X^*_{i_1i_2}X^*_{i_2i_3}X^*_{i_3i_4}X^*_{i_4i_1}$. Inserting $X^*_{ij} = \widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij}$ and expanding every bracket, we similarly obtain $3 \times 3 \times 3 \times 3 = 81$ different post-expansion sums, where 15 of them do not involve any δ term. The sum of the remaining 65 terms is $(Q_n^* - \widetilde{Q}_n)$. For each of these 65 post-expansion sums, we are summing over many 4-products, where each of them has the same number of $\widetilde{\Omega}$ factors, W factors, and δ factors, which we denote by $N_{\widetilde{\Omega}}$, N_W, and N_δ, respectively. According to $(N_{\widetilde{\Omega}}, N_W, N_\delta)$, we divide the 65 post-expansion sums into 10 different types. See Table 2, where we recall that $\alpha = |\lambda_2|/\lambda_1$.

We now analyze $Q_n^* - \widetilde{Q}_n$. Consider the null hypothesis first. Under the null, $\widetilde{\Omega}$ is a zero matrix, so the nonzero post-expansion sums only include Type Ia, Type IIa, Type IIIa and Type IV. It is seen that $|\mathbb{E}[Q_n^* - \widetilde{Q}_n]| \le C\|\theta\|^4$ and $\mathrm{Var}(Q_n^* - \widetilde{Q}_n) = o(\|\theta\|^8)$. Note that $\|\theta\|^8$ is the order of $\mathrm{Var}(\widetilde{Q}_n)$ under the null. The difference between the variance of Q_n^* and the variance of $\widetilde{Q}_n$ is negligible, but the difference between the mean of Q_n^* and the mean of $\widetilde{Q}_n$ is nonnegligible. With lengthy calculations (see the Supplementary Material), we can show that $\mathbb{E}[Q_n^* - \widetilde{Q}_n] \sim 2\|\theta\|^4$. Therefore, $(Q_n^* - 2\|\theta\|^4)$ and $\widetilde{Q}_n$ have a negligible difference under the null.

Consider the alternative hypothesis next. From Table 2, $|\mathbb{E}[Q_n^* - \widetilde{Q}_n]| \le C(|\lambda_2|/\lambda_1)^2\|\theta\|^6$, where the major contribution is from Type Ic and Type IIc post-expansion sums. Under our assumptions for the alternative, $|\lambda_2|/\sqrt{\lambda_1} \to \infty$ and $\lambda_1 \asymp \|\theta\|^4$. It is easy to see that $|\mathbb{E}[Q_n^* - \widetilde{Q}_n]| = o(\lambda_2^4)$, where λ_2^4 is the order of $\mathrm{tr}(\widetilde{\Omega}^4)$ and $\mathbb{E}[\widetilde{Q}_n]$; see Lemma 2.3 and Theorem 4.1. Additionally, $\|\theta\|^4 = O(\lambda_1^2) = o(\lambda_2^4)$, which is also of a smaller order of $\mathbb{E}[\widetilde{Q}_n]$. We conclude that $|\mathbb{E}[Q_n^* - \widetilde{Q}_n - 2\|\theta\|^4]| = o(\mathbb{E}[\widetilde{Q}_n])$. From the table, $\mathrm{Var}(Q_n^* - \widetilde{Q}_n) \le C(|\lambda_2|/\lambda_1)^6\|\theta\|^{12}\|\theta\|_3^3/\|\theta\|_1 + o(\|\theta\|^8)$, with the major contribution from Type Id. Here, the second term is smaller than $\mathrm{Var}(\widetilde{Q}_n)$, and the first term is upper bounded by $C(|\lambda_2|/\lambda_1)^6\|\theta\|^8\|\theta\|_3^6$ (using the universal inequality of $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$), which has a comparable order as $\mathrm{Var}(\widetilde{Q}_n)$. It follows that $\mathrm{Var}(Q_n^* - \widetilde{Q}_n - 2\|\theta\|^4) = \mathrm{Var}(Q_n^* - \widetilde{Q}_n) \le C\mathrm{Var}(\widetilde{Q}_n)$. Combining the above, we obtain that the SNR of $(Q_n^* - 2\|\theta\|^4)$ and $\widetilde{Q}_n$ are at the same order.

These results are summarized in Theorem 4.2 and proved in the Supplementary Material.

THEOREM 4.2 (Proxy SgnQ test statistic). *Consider the testing problem* (1.6) *under the DCMM model* (1.1)–(1.4), *where the condition* (2.2) *is satisfied under the alternative hypothesis. Suppose* $\theta_{\max} \to 0$ *and* $\|\theta\| \to \infty$ *as* $n \to \infty$, *and suppose* $|\lambda_2|/\sqrt{\lambda_1} \to \infty$ *under the alternative hypothesis. Then, under the null hypothesis, as* $n \to \infty$, $\mathbb{E}[(Q_n^* - 2\|\theta\|^4) - \widetilde{Q}_n] = o(\|\theta\|^4)$ *and* $\mathrm{Var}(Q_n^* - \widetilde{Q}_n) = o(\|\theta\|^8)$. *Furthermore, under the alternative hypothesis,* $\mathbb{E}[(Q_n^* - 2\|\theta\|^4) - \widetilde{Q}_n] = o((|\lambda_2|/\lambda_1)^4\|\theta\|^8)$ *and* $\mathrm{Var}(Q_n^* - \widetilde{Q}_n) \le C(|\lambda_2|/\lambda_1)^6\|\theta\|^8\|\theta\|_3^6 + o(\|\theta\|^8)$.

TABLE 2
The 10 *types of the post-expansion sums for* $(Q_n^* - \widetilde{Q}_n)$. *Notation*: *same as in Table* 1

Type	#	$(N_\delta, N_{\widetilde{\Omega}}, N_W)$	Examples	Abs. Mean	Variance
Ia	4	(1, 0, 3)	$\sum_{i,j,k,\ell (dist)} \delta_{ij} W_{jk} W_{k\ell} W_{\ell i}$	0	$\le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8)$
Ib	8	(1, 1, 2)	$\sum_{i,j,k,\ell (dist)} \delta_{ij} \widetilde{\Omega}_{jk} W_{k\ell} W_{\ell i}$	0	$\le C\alpha^2\|\theta\|^4\|\theta\|_3^6 = o(\|\theta\|^8)$
	4		$\sum_{i,j,k,\ell (dist)} \delta_{ij} W_{jk} \widetilde{\Omega}_{k\ell} W_{\ell i}$	0	$\le C\alpha^2\|\theta\|^4\|\theta\|_3^6 = o(\|\theta\|^8)$
Ic	8	(1, 2, 1)	$\sum_{i,j,k,\ell (dist)} \delta_{ij} \widetilde{\Omega}_{jk} \widetilde{\Omega}_{k\ell} W_{\ell i}$	$\le C\alpha^2\|\theta\|^6 = o(\alpha^4\|\theta\|^8)$	$\le \frac{C\alpha^4\|\theta\|^{10}\|\theta\|_3^3}{\|\theta\|_1} = o(\alpha^6\|\theta\|^8\|\theta\|_3^6)$
	4		$\sum_{i,j,k,\ell (dist)} \delta_{ij} \widetilde{\Omega}_{jk} W_{k\ell} \widetilde{\Omega}_{\ell i}$	0	$\le \frac{C\alpha^4\|\theta\|^4\|\theta\|_3^9}{\|\theta\|_1} = o(\|\theta\|^8)$
Id	4	(1, 3, 0)	$\sum_{i,j,k,\ell (dist)} \delta_{ij} \widetilde{\Omega}_{jk} \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i}$	0	$\le \frac{C\alpha^6\|\theta\|^{12}\|\theta\|_3^3}{\|\theta\|_1} = O(\alpha^6\|\theta\|^8\|\theta\|_3^6)$
IIa	4	(2, 0, 2)	$\sum_{i,j,k,\ell (dist)} \delta_{ij} \delta_{jk} W_{k\ell} W_{\ell i}$	$\le C\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$	$\le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8)$
	2		$\sum_{i,j,k,\ell (dist)} \delta_{ij} W_{jk} \delta_{k\ell} W_{\ell i}$	$\le C\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$	$\le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$
IIb	8	(2, 1, 1)	$\sum_{i,j,k,\ell (dist)} \delta_{ij} \delta_{jk} \widetilde{\Omega}_{k\ell} W_{\ell i}$	0	$\le C\alpha^2\|\theta\|^4\|\theta\|_3^6 = o(\|\theta\|^8)$
	4		$\sum_{i,j,k,\ell (dist)} \delta_{ij} \widetilde{\Omega}_{jk} \delta_{k\ell} W_{\ell i}$	$\le C\alpha\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$	$\le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$
IIc	4	(2, 2, 0)	$\sum_{i,j,k,\ell (dist)} \delta_{ij} \delta_{jk} \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i}$	$\le C\alpha^2\|\theta\|^6 = o(\alpha^4\|\theta\|^8)$	$\le \frac{C\alpha^4\|\theta\|^{14}}{\|\theta\|_1^2} = o(\alpha^6\|\theta\|^8\|\theta\|_3^6)$
	2		$\le \sum_{i,j,k,\ell (dist)} \delta_{ij} \widetilde{\Omega}_{jk} \delta_{k\ell} \widetilde{\Omega}_{\ell i}$	$\frac{C\alpha^2\|\theta\|^8}{\|\theta\|_1^2} = o(\alpha^4\|\theta\|^8)$	$\le \frac{C\alpha^4\|\theta\|^8\|\theta\|_3^6}{\|\theta\|_1^2} = o(\|\theta\|^8)$
IIIa	4	(3, 0, 1)	$\sum_{i,j,k,\ell (dist)} \delta_{ij} \delta_{jk} \delta_{k\ell} W_{\ell i}$	$\le C\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$	$\le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$
IIIb	4	(3, 1, 0)	$\le \sum_{i,j,k,\ell (dist)} \delta_{ij} \delta_{jk} \delta_{k\ell} \widetilde{\Omega}_{\ell i}$	$\le \frac{C\alpha\|\theta\|^6}{\|\theta\|_1^3} = o(\alpha^4\|\theta\|^8)$	$\le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$
IV	1	(4, 0, 0)	$\sum_{i,j,k,\ell (dist)} \delta_{ij} \delta_{jk} \delta_{k\ell} \delta_{\ell i}$	$\le C\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$	$\le \frac{C\|\theta\|^{10}}{\|\theta\|_1^2} = o(\|\theta\|^8)$

4.3. *The behavior of* $(Q_n - Q_n^*)$. Recall that $Q_n = \sum_{i_1,i_2,i_3,i_4 (dist)} X_{i_1 i_2} X_{i_2 i_3} X_{i_3 i_4} X_{i_4 i_1}$, where $X_{ij} = \widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij} + r_{ij}$ for any $i \neq j$. Similar to Sections 4.1–4.2, we first expand every bracket in the definitions and obtain $4 \times 4 \times 4 \times 4 = 256$. Out of the 256 post-expansion sums in Q_n, $3 \times 3 \times 3 \times 3 = 81$ of them do not involve any r term and are contained in Q_n^*; this leaves a total of $256 - 81 = 175$ different post-expansion sums in $(Q_n - Q_n^*)$. In the Supplementary Material, we investigate the order of mean and variance of each of the 175 post-expansion sums in $(Q_n - Q_n^*)$. The calculations are very tedious: although we expect these post-expansion sums to be of a smaller order than the post-expansion sums in Sections 4.1–4.2, it is impossible to prove this argument rigorously using only some crude bounds (such as the Cauchy–Schwarz inequality). Instead, we still need to do calculations for each post-expansion sum; details are in the Supplementary Material.

THEOREM 4.3 (Real SgnQ test statistic). *Consider the testing problem* (1.6) *under the DCMM model* (1.1)–(1.4), *where the condition* (2.2) *is satisfied under the alternative hypothesis*. *Suppose* $\theta_{\max} \to 0$ *and* $\|\theta\| \to \infty$ *as* $n \to \infty$, *and suppose* $|\lambda_2|/\sqrt{\lambda_1} \to \infty$ *under the alternative hypothesis*. *Then*, *under the null hypothesis*, *as* $n \to \infty$, $|\mathbb{E}[Q_n - Q_n^*]| = o(\|\theta\|^4)$ *and* $\mathrm{Var}(Q_n - Q_n^*) = o(|\theta\|^8)$. *Under the alternative hypothesis*, *as* $n \to \infty$, $|\mathbb{E}[Q_n - Q_n^*]| = o((|\lambda_2|/\lambda_1)^4\|\theta\|^8)$ *and* $\mathrm{Var}(Q_n - Q_n^*) = o((|\lambda_2|/\lambda_1)^6\|\theta\|^8|\theta\|_3^6) + o(\|\theta\|^8)$.

5. Simulations. We investigate the numerical performance of two Signed Polygon tests, the SgnT test (2.4) and the SgnQ test (2.6). We also include the EZ test [16] and the GC test [25] for comparison. For reasons mentioned in [25], we use a two-sided rejection region for EZ and a one-sided rejection region for GC.

Given (n, K), a scalar $\beta_n > 0$ that controls $\|\theta\|$, a symmetric nonnegative matrix $P \in \mathbb{R}^{K\times K}$, a distribution $f(\theta)$ on $\mathbb{R}_+$, and a distribution $g(\pi)$ on the standard simplex of $\mathbb{R}^K$, we generate two network adjacency matrices A^{null} and A^{alt}, under the null and the alternative, respectively, as follows: (i) Generate $\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_n$ *i.i.d.* from $f(\theta)$. Let $\theta_i = \beta_n \cdot \tilde{\theta}_i/\|\tilde{\theta}\|$, $1 \le i \le n$. (ii) Generate $\pi_1, \pi_2, \dots, \pi_n$ iid from $g(\pi)$. (iii) Let $\Omega^{\text{alt}} = \Theta\Pi P\Pi'\Theta'$, where $\Theta = \text{diag}(\theta_1, \dots, \theta_n)$ and $\Pi = [\pi_1, \pi_2, \dots, \pi_n]'$. Generate A^{alt} from Ω^{alt} according to Model (1.1). (iv) Let $\Omega^{\text{null}} = (a'Pa)\cdot\theta\theta'$, where $a = \mathbb{E}_g\pi \in \mathbb{R}^K$ is the mean vector of $g(\pi)$. Generate A^{null} from Ω^{null} according to Model (1.1). The pair $(\Omega^{\text{null}}, \Omega^{\text{alt}})$ is constructed in a way such that the corresponding networks have approximately the same expected average degree. This is the most subtle case for distinguishing two hypotheses (see Section 3).

It is of interest to explore different sparsity levels and also focus on the parameter settings where the SNR is neither too large nor too small. Therefore, for most experiments, we let $\beta_n = \|\theta\|$ range but fix the SNR at more or less the same level. See details below. For each parameter setting, we generate 200 networks under the null hypothesis and 200 networks under the alternative hypothesis, run all the four tests with a target level $\alpha = 5\%$ and then record the sum of percent of type I errors and percent of type II errors. For space limit, we do not report separately the percent of each type of errors but relegate these results to the Supplementary Material.

5.1. *Experiment* 1. We study the role of degree heterogeneity. Fix $(n, K) = (2000, 2)$. Let P be a 2×2 matrix with unit diagonal entries and all off-diagonal entries equal to b_n. Let $g(\pi)$ be the uniform distribution on $\{(0, 1), (1, 0)\}$. We consider three subexperiments, Exp 1a–1c, where respectively we take $f(\theta)$ to be the following: (a) Uniform(2, 3), (b) two-point distribution $0.95\delta_1 + 0.05\delta_3$, where δ_a is a point mass at a and (c) Pareto(10, 0.375), where 10 is the shape parameter and 0.375 is the scale parameter. The degree heterogeneity is moderate in Exp 1a–1b, but more severe in Exp 1c. In such a setting, SNR is at the order of $\|\theta\|(1 - b_n)$. Therefore, for each subexperiment, we let $\beta_n = \|\theta\|$ vary while fixing the SNR to be $\|\theta\|(1 - b_n) = 3.2$. The sum of Type I and Type II errors are displayed in Figure 3.

First, both the SgnQ test and the GC test are based on the counts of 4 cycles, but the *GC* test counts *noncentered* cycles and the SgnQ test counts *centered* cycles. As we pointed out in Section 1, counting *centered* cycles may have much smaller variances than counting *noncentered* cycles, especially in the less sparse case, and thus improves the testing power. This is confirmed by numerical results here, where the SgnQ test is consistently better than the GC test, significantly so in the less sparse case. Similarly, both the SgnT test and the

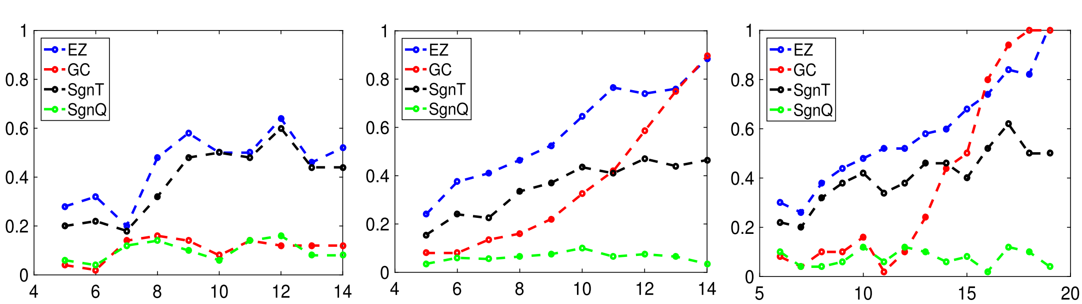

FIG. 3. *From left to right*: *Experiment* 1*a*, 1*b and* 1*c*. *The* y*-axis are the sum of Type I and Type II errors* (*testing level is fixed at* 5%). *The* x*-axis are* $\|\theta\|$ *or sparsity levels. Results are based on* 200 *repetitions.*

EZ test are based on the counts of 3 cycles, but the EZ test counts *noncentered* cycles and the SgnT test counts *centered* cycles, and we expect that SgnT significantly improves EZ, especially in the less sparse case. This is also confirmed in the experiment.

Second, SgnQ and GC are order-4 graphlet counting statistics, and SgnT and EZ are order-3 graphlet counting statistics. In comparison, SgnQ significantly outperforms SgnT, and GC significantly outperforms EZ (in the more sparse case; see discussion below for the less sparse case). A possible explanation is that higher-order graphlet counting statistics have larger SNR. Investigation toward this direction is interesting, and we leave it to future study. Note that SgnQ is the best among all four tests.

Last, GC outperforms EZ in the more sparse case but underperforms EZ in the less sparse case. The reason for the latter is as follows. The biases of both tests are negligible in the more sparse case, but are nonnegligible in the less sparse case, with that of GC much larger. In [30], we propose a bias correction method, where the performance of GC is significantly improved. However, GC continues to underperform SgnQ, because even with the bias corrected, it still has a variance that is unsatisfactorily large.

5.2. *Experiment* 2. We study the cases with larger K and a more complicated matrix of P. For some $b_n \in (0,1)$, let $\epsilon_n = \frac{1}{6}\min(1-b_n, b_n)$, and let P be the matrix with 1 on the diagonal and the off-diagonal entries i.i.d. drawn from $\mathrm{Unif}(b_n - \epsilon_n, b_n + \epsilon_n)$; once the P matrix is drawn, it is fixed throughout different repetitions. We consider two subexperiments, Exp 2a and 2b. In Exp 2a, we take $(n, K) = (1000, 5)$, $f(\theta)$ to be $\mathrm{Pareto}(10, 0.375)$, and $g(\pi)$ to be the uniform distribution on $\{e_1, \dots, e_K\}$ (the standard basis vectors of $\mathbb{R}^K$). We let β_n range but fix $\|\theta\|(1-b_n)$ at 4.5, so the SNR will not change drastically. In Exp 2b, we take $(n, K) = (3000, 10)$, $f(\theta)$ to be $0.95\delta_1 + 0.05\delta_3$, and $g(\pi) = 0.1\sum_{k=1}^{2}\delta_{e_k} + 0.15\sum_{k=3}^{6}\delta_{e_k} + 0.05\sum_{k=7}^{10}\delta_{e_k}$ (so to have unbalanced community sizes). Similarly, we let β_n range but fix $\|\theta\|(1-b_n) = 5.2$. The sum of Type I and II errors are shown in Figure 4.

In these examples, EZ and GC underperform SgnT and SgnQ, especially in the less sparse case, and the performances of SgnT and SgnQ are more similar to each other, compared to those in Experiment 1. In these examples, we have larger K, more complicated P and unbalanced community sizes, and the performance of SgnT and SgnQ test statistics suggest that they are relatively robust.

5.3. *Experiment* 3. We investigate the role of mixed membership. We have three subexperiments, Exp 3a–3c. where the memberships are not mixed, lightly mixed and significantly mixed, respectively. For all subexperiments, we take $(n, K) = (2000, 3)$ and $f(\theta)$ to be $\mathrm{Unif}(2, 3)$. For Exp 3a, we let $g_1(\pi) = 0.4\delta_{e_1} + 0.3\delta_{e_2} + 0.3\delta_{e_3}$. In Exp 3b, we let $g_2(\pi) = 0.3\sum_{k=1}^{3}\delta_{e_k} + 0.1\cdot\mathrm{Dirichlet}$, and in Exp 3c, we let $g_3(\pi) = 0.25\sum_{k=1}^{3}\delta_{e_k} + 0.25\cdot\mathrm{Dirichlet}$,

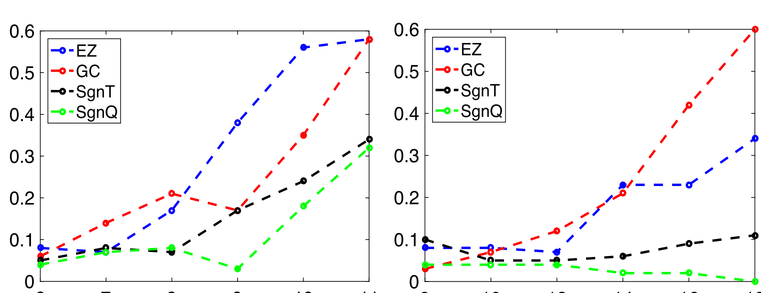

FIG. 4. *From left to right*: *Experiment* 2a *and* 2b. *The y-axis are the sum of Type I and Type II errors* (*testing level is fixed at* 5%). *The x-axis are $\|\theta\|$ or sparsity levels. Results are based on* 200 *repetitions.*

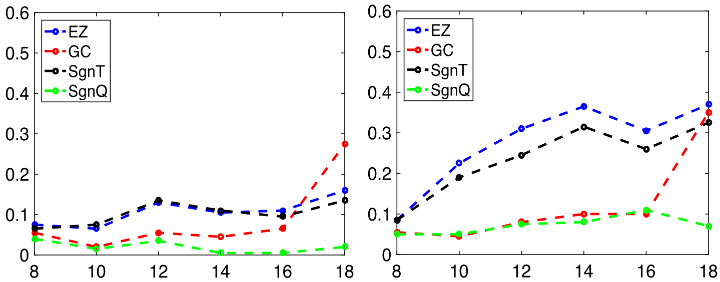

FIG. 5. *From left to right: Experiment* 3a, 3b *and* 3c. *The* y-*axis are the sum of Type I and Type II errors* (*testing level is fixed at* 5%). *The* x-*axis are* $\|\theta\|$ *or sparsity levels. Results are based on* 200 *repetitions.*

where Dirichlet represents the symmetric K-dimensional Dirichlet distribution. In Exp 3a–3b, we let β_n range while $(1-b_n)\|\theta\|$ is fixed at 4.2 so the SNR's are roughly the same. In Exp 3c, we also let β_n range but $(1-b_n)\|\theta\| = 4.5$ (the SNR's need to be slightly larger to counter the effect of mixed membership, which makes the testing problem harder).

The sum of Type I and Type II errors are presented in Figure 5. First, the results confirm that mixed memberships make the testing problem harder. For example, the value of $\|\theta\|(1-b_n)$ in Exp 3c is higher than that of Exp 3a–3b, but the testing errors are higher, due to that the memberships in Exp 3c are more mixed. Second, SgnQ consistently outperforms EZ and SgnT. Third, GC is comparable with SgnQ in the more sparse case, but performs unsatisfactorily in the less sparse case, for reasons explained before. Last, in these settings, SgnT is uniformly better than EZ, and more so when the memberships become more mixed.

5.4. *Experiment* 4. We vary the size of network and study its impact on testing errors. We fix $K = 2$ and let P be a 2×2 matrix with unit diagonals and off-diagonals equal to b_n. Let $g(\pi)$ be the uniform distribution on $\{(0,1),(1,0)\}$ and let $f(\theta)$ be Pareto$(8, 0.375)$. We let n ranges from $\{100, 300, 1000, 3000\}$. Note that in our data generating process, $\beta_n = \|\theta\|$ controls the sparsity level and $(1-b_n)\|\theta\|$ is the SNR. As n varies, we fix $\beta_n = 4$ and change b_n accordingly so that the SNR is fixed at 3. The results are in Table 3. This is a sparse setting, therefore, the biases in EZ and GC are negligible and they both control the Type I error well. The SgnT and SgnQ tests also control the Type I error well. In terms of the Type II errors, GC and SgnQ are better than EZ and SgnT. The results are relatively stable as n varies.

6. Discussions. A closely related idea is to use $\|A - \hat{\eta}\hat{\eta}'\|$ as the test statistics. To see why this is a reasonable choice, consider the proxy test statistic $\|A - \eta^*(\eta^*)'\|$, where we recall that $\eta^* = \theta$ under the null; see (1.12). Therefore, $A - \eta^*(\eta^*)'$ is equal to W and $(\Omega - (\eta^*(\eta^*)') + W$, under the null and the alternative, respectively. The test has reasonable power, as $\|A - \eta^*(\eta^*)'\|$ is expected to be bigger in the alternative than in the null. Another related idea is to extend the Signed Polygon to address the problem of testing whether

TABLE 3
Experiment 4. *Numbers in each cell are Type I error, Type II error and their sum*

n	100	300	1000	3000
EZ	(0.025, 0.22, 0.245)	(0.055, 0.26, 0.315)	(0.05, 0.27, 0.32)	(0.06, 0.275, 0.335)
GC	(0.02, 0.02, 0.04)	(0.06, 0.02, 0.08)	(0.04, 0.005, 0.045)	(0.04, 0.005, 0.045)
SgnT	(0.01, 0.15, 0.16)	(0.04, 0.14, 0.18)	(0.065, 0.175, 0.24)	(0.06, 0.14, 0.2)
SgnQ	(0.05, 0.015, 0.02)	(0.04, 0.005, 0.045)	(0.04, 0, 0.04)	(0.02, 0.005, 0.025)

$K = k_0$ versus $K > k_0$, where $k_0 > 1$ is a prescribed integer. Let $\hat{\Omega} = \sum_{k=1}^{k_0} \hat{\lambda}_k \hat{\xi}_k \hat{\xi}_k'$, where $\hat{\lambda}_k$ are the kth eigenvalue of A, arranged in the descending order in magnitude and $\hat{\xi}_k$ is the corresponding eigenvector. The Signed Polygon test statistic can then be extended to $U_{n,k_0}^{(m)} = \sum_{i_1,i_2,\dots,i_m (dist)} (A_{i_1 i_2} - \hat{\Omega}_{i_1 i_2})(A_{i_2 i_3} - \hat{\Omega}_{i_2 i_3}) \dots (A_{i_m i_1} - \hat{\Omega}_{i_m i_1})$. See [27] for more discussion. It remains unclear whether these test statistics are optimally adaptive, and we leave the study to the future.

Another testing idea would be using the first eigenvalue of $\widetilde{A} = \hat{\theta}^{-1} A \hat{\theta}^{-1} - \hat{b} \mathbf{1}_n \mathbf{1}_n'$, for a reasonable estimate $\hat{\theta}$ for θ and a proper $\hat{b}$. Unfortunately, even if $\hat{\theta} = \theta$, the distribution of the test is unknown for general cases. In fact, this is essentially the approaches proposed in [8, 32]). Both papers showed that in the dense case of $\theta_1 = \theta_2 = \cdots = \theta_n = O(1)$, the largest eigenvalue of $\widetilde{A}$ (when standardized) converges to the Tracy–Widom law. Unfortunately, the approaches have been focused on the more idealized SBM model and the less sparse case where $\theta_1 = \theta_2 = \cdots = \theta_n = \sqrt{\alpha_n} \geq O(n^{-1/6})$, and the limiting distribution remains unknown for other cases.

The testing problem is also closely related to the problem of estimating K. In fact, we can cast the estimation problem as a sequential testing problem where we test $K = k_0$ vs. $K > k_0$ for $k_0 = 1, 2, 3, \dots$, and estimate K to be the smallest k_0 where we accept the null.

Note also the lower bound argument for the global testing problem sheds useful insight for many other problems (e.g., estimating K, community detection, mixed membership). Take the problem of estimating K, for example. Given an alternative setting, if we cannot distinguish it from some null setting, then the underlying parameter K is not estimable.

In a high level, these ideas, together with the Signed Polygon, are related to the ideas in [21] on testing $K = k_0$ versus $K > k_0$, in [32] on goodness-of-fit, and in [31] on estimating K. However, the focus of these works are on the more idealized model where we do not have degree heterogeneity, and how to extend their ideas to the current setting remains unclear.

Acknowledgments. The authors would like to thank the anonymous Associate Editor and referees for helpful comments. ZK would like to thank Sebastien Bubeck, Fang Han, and Elchanan Mossel for helpful pointers and comments.

Funding. JJ and SL are supported in part by NSF Grant DMS-2015469. ZK is supported in part by NSF Grant DMS-1925845 and NSF CAREER Grant DMS-1943902.

SUPPLEMENTARY MATERIAL

Additional results and technical proofs (DOI: 10.1214/21-AOS2089SUPP; .pdf). The supplemental material contains the results not reported in the main article due to space limit and the proofs of all theorems and lemmas.

REFERENCES

[1] AIROLDI, E., BLEI, D., FIENBERG, S. and XING, E. (2008). Mixed membership stochastic blockmodels. *J. Mach. Learn. Res.* **9** 1981–2014.

[2] BANERJEE, D. (2018). Contiguity and non-reconstruction results for planted partition models: The dense case. *Electron. J. Probab.* **23** Paper No. 18, 28. MR3771755 https://doi.org/10.1214/17-EJP128

[3] BANERJEE, D. and MA, Z. (2017). Optimal hypothesis testing for stochastic block models with growing degrees. Available at arXiv:1705.05305.

[4] BANKS, J., MOORE, C., NEEMAN, J. and NETRAPALLI, P. (2016). Information-theoretic thresholds for community detection in sparse networks. In *COLT* 383–416.

[5] BÉJAR, J., ÁLVAREZ, S., GARCÍA, D., GÓMEZ, I., OLIVA, L., TEJEDA, A. and VÁZQUEZ-SALCEDA, J. (2016). Discovery of spatio-temporal patterns from location-based social networks. *J. Exp. Theor. Artif. Intell.* **28** 313–329.

[6] BHATTACHARYYA, S. and BICKEL, P. J. (2015). Subsampling bootstrap of count features of networks. *Ann. Statist.* **43** 2384–2411. MR3405598 https://doi.org/10.1214/15-AOS1338

[7] BICKEL, P. J., CHEN, A. and LEVINA, E. (2011). The method of moments and degree distributions for network models. *Ann. Statist.* **39** 2280–2301. MR2906868 https://doi.org/10.1214/11-AOS904

[8] BICKEL, P. J. and SARKAR, P. (2016). Hypothesis testing for automated community detection in networks. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 253–273. MR3453655 https://doi.org/10.1111/rssb.12117

[9] BRUALDI, R. A. (1974). The DAD theorem for arbitrary row sums. *Proc. Amer. Math. Soc.* **45** 189–194. MR0354737 https://doi.org/10.2307/2040060

[10] BUBECK, S., DING, J., ELDAN, R. and RÁCZ, M. Z. (2016). Testing for high-dimensional geometry in random graphs. *Random Structures Algorithms* **49** 503–532. MR3545825 https://doi.org/10.1002/rsa.20633

[11] CHATTERJEE, S. (2015). Matrix estimation by universal singular value thresholding. *Ann. Statist.* **43** 177–214. MR3285604 https://doi.org/10.1214/14-AOS1272

[12] DALL'AMICO, L., COUILLET, R. and TREMBLAY, N. (2019). Revisiting the Bethe-Hessian: Improved community detection in sparse heterogeneous graphs. In *Adv. Neural. Inf. Process. Syst.* 4037–4047. Curran Associates, Red Hook.

[13] DONOHO, D. and JIN, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.* **32** 962–994. MR2065195 https://doi.org/10.1214/009053604000000265

[14] DU, H. and YANG, S. J. (2011). Discovering collaborative cyber attack patterns using social network analysis. In *Proceedings of the 4th International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction. SBP'*11 129–136. Springer, Berlin.

[15] FU, Y.-H., HUANG, C.-Y. and SUN, C.-T. (2015). Using global diversity and local topology features to identify influential network spreaders. *Phys. A* **433** 344–355. https://doi.org/10.1016/j.physa.2015.03.042

[16] GAO, C. and LAFFERTY, J. (2017). Testing for global network structure using small subgraph statistics. Available at arXiv:1710.00862.

[17] GIRVAN, M. and NEWMAN, M. E. J. (2002). Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **99** 7821–7826. MR1908073 https://doi.org/10.1073/pnas.122653799

[18] GULIKERS, L., LELARGE, M. and MASSOULIÉ, L. (2017). Non-backtracking spectrum of degree-corrected stochastic block models. In *8th Innovations in Theoretical Computer Science Conference. LIPIcs. Leibniz Int. Proc. Inform.* **67** Art. No. 44, 27. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern. MR3754968

[19] GULIKERS, L., LELARGE, M. and MASSOULIÉ, L. (2018). An impossibility result for reconstruction in the degree-corrected stochastic block model. *Ann. Appl. Probab.* **28** 3002–3027. MR3847979 https://doi.org/10.1214/18-AAP1381

[20] HOLLAND, P. W., LASKEY, K. B. and LEINHARDT, S. (1983). Stochastic blockmodels: First steps. *Soc. Netw.* **5** 109–137. MR0718088 https://doi.org/10.1016/0378-8733(83)90021-7

[21] HU, J., QIN, H., YAN, T., ZHANG, J. and ZHU, J. (2017). Using maximum entry-wise deviation to test the goodness-of-fit for stochastic block models. Available at arXiv:1703.06558.

[22] INGSTER, Y. I., TSYBAKOV, A. B. and VERZELEN, N. (2010). Detection boundary in sparse regression. *Electron. J. Stat.* **4** 1476–1526. MR2747131 https://doi.org/10.1214/10-EJS589

[23] JIN, J. (2015). Fast community detection by SCORE. *Ann. Statist.* **43** 57–89. MR3285600 https://doi.org/10.1214/14-AOS1265

[24] JIN, J., KE, Z. T. and LUO, S. (2017). Estimating network memberships by simplex vertices hunting. Available at arXiv:1708.07852.

[25] JIN, J., KE, Z. T. and LUO, S. (2018). Network global testing by counting graphlets. In *Proceedings of the 35th International Conference on Machine Learning, ICML* 2018, *Stockholmsmässan, Stockholm, Sweden, July* 10–15, 2018 2338–2346.

[26] JIN, J., KE, Z. T. and LUO, S. (2021). Supplement to "Optimal adaptivity of signed-polygon statistics for network testing." https://doi.org/10.1214/21-AOS2089SUPP

[27] JIN, J., LI, W. and WANG, T. (2019). A spectral approach for network global testing. Manuscript.

[28] JOHNSON, C. R. and REAMS, R. (2009). Scaling of symmetric matrices by positive diagonal congruence. *Linear Multilinear Algebra* **57** 123–140. MR2492099 https://doi.org/10.1080/03081080600872327

[29] KARRER, B. and NEWMAN, M. E. J. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E* (3) **83** 016107, 10. MR2788206 https://doi.org/10.1103/PhysRevE.83.016107

[30] KE, Z. T. (2019). Comparison of different network testing procedures. Manuscript.

[31] LE, C. M. and LEVINA, E. (2015). Estimating the number of communities in networks by spectral methods. Available at arXiv:1507.00827.

[32] LEI, J. (2016). A goodness-of-fit test for stochastic block models. *Ann. Statist.* **44** 401–424. MR3449773 https://doi.org/10.1214/15-AOS1370

[33] LI, T., LEI, L., BHATTACHARYYA, S., VAN DEN BERGE, K. SARKAR, P., BICKEL, P. J. and LEVINA, E. (2020). Hierarchical community detection by recursive partitioning. *J. Amer. Statist. Assoc.* https://doi.org/10.1080/01621459.2020.1833888

[34] MA, Z. and WU, Y. (2015). Computational barriers in minimax submatrix detection. *Ann. Statist.* **43** 1089–1116. MR3346698 https://doi.org/10.1214/14-AOS1300

[35] MARSHALL, A. W. and OLKIN, I. (1968). Scaling of matrices to achieve specified row and column sums. *Numer. Math.* **12** 83–90. MR0238875 https://doi.org/10.1007/BF02170999

[36] MAUGIS, P.-A. G., PRIEBE, C. E., OLHEDE, S. C. and WOLFE, P. J. (2017). Statistical inference for network samples using subgraph counts. Available at arXiv:1701.00505.

[37] MOSSEL, E., NEEMAN, J. and SLY, A. (2015). Reconstruction and estimation in the planted partition model. *Probab. Theory Related Fields* **162** 431–461. MR3383334 https://doi.org/10.1007/s00440-014-0576-6

[38] PAUL, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** 1617–1642. MR2399865

[39] RADICCHI, F., CASTELLANO, C., CECCONI, F., LORETO, V. and PARISI, D. (2004). Defining and identifying communities in networks. *Proc. Natl. Acad. Sci. USA* **101** 2658–2663.

[40] SALDAÑA, D. F., YU, Y. and FENG, Y. (2017). How many communities are there? *J. Comput. Graph. Statist.* **26** 171–181. MR3610418 https://doi.org/10.1080/10618600.2015.1096790

[41] SINKHORN, R. (1974). Diagonal equivalence to matrices with prescribed row and column sums. II. *Proc. Amer. Math. Soc.* **45** 195–198. MR0357434 https://doi.org/10.2307/2040061

[42] WANG, Y. X. R. and BICKEL, P. J. (2017). Likelihood-based model selection for stochastic block models. *Ann. Statist.* **45** 500–528. MR3650391 https://doi.org/10.1214/16-AOS1457

[43] ZHANG, Y., LEVINA, E. and ZHU, J. (2020). Detecting overlapping communities in networks using spectral methods. *SIAM J. Math. Data Sci.* **2** 265–283. MR4081803 https://doi.org/10.1137/19M1272238

[44] ZHAO, Y., LEVINA, E. and ZHU, J. (2011). Community extraction for social networks. *Proc. Natl. Acad. Sci. USA* **108** 7321–7326.

Submitted to the Annals of Statistics

SUPPLEMENT OF "OPTIMAL ADAPTIVITY OF SIGNED-POLYGON STATISTICS FOR NETWORK TESTING"

BY JIASHUN JIN[1], ZHENG TRACY KE[2] AND SHENGMING LUO[1]

[1]*Department of Statistics & Data Science, Carnegie Mellon University, jiashun@stat.cmu.edu; shengmil@andrew.cmu.edu*

[2]*Department of Statistics, Harvard University, zke@fas.harvard.edu*

This supplement contains additional results and technical proofs for the main article [4]. Appendix A studies the behavior of the SgnT test statistic and proves Theorems 2.1, 2.3, and 2.5. Appendix B is about the properties of the SgnQ test statistic and proves Theorems 2.2, 2.4, and 2.6. Appendix C derives the matrix forms of signed-polygon statistics and proves Theorem 1.1. Appendix D studies the estimation error of $\|\theta\|^2$ and proves Lemma 2.1. Appendix E contains spectral analysis for Ω and $\widetilde{\Omega}$ and proves Lemmas 2.2-2.3. Appendix F analyzes the region of impossibility and proves Lemma 3.1 and Theorems 3.1-3.5. Appendix G calculates the mean and variance of signed-polygon statistics and proves the results in Tables 1-2, Tables A.1-A.2, Theorems 4.1-4.3, and Theorems A.1-A.3. Appendix H contains additional simulation results.

CONTENTS

APPENDIX A: THE BEHAVIOR OF THE SGNT TEST STATISTIC

We now discuss the behavior of the SgnT test statistic and prove Theorems 2.1, 2.3, and 2.5. The discussion is similar to that of SgnQ in Section 4, and so we keep it brief.

Recall that the SgnT test statistic is defined by

$$T_n = \sum_{i_1,i_2,i_3(dist)} (A_{i_1 i_2} - \hat{\eta}_{i_1}\hat{\eta}_{i_2})(A_{i_2 i_3} - \eta_{i_2}\hat{\eta}_{i_3})(A_{i_3 i_1} - \hat{\eta}_{i_3}\hat{\eta}_{i_1}).$$

Similarly, define the Ideal SgnT test statistic $\widetilde{T}_n$ and the Proxy SgnT test statistic and T_n^*, and write

$$T_n = \widetilde{Q}_n + (Q_n^* - \widetilde{Q}_n) + (Q_n - Q_n^*). \tag{1}$$

We have the following observations.

- $\widetilde{Q}_n$ is the sum of 8 different post-expansion sums, divided into 4 types. See Table A.1.
- $Q_n^* - \widetilde{Q}_n$ is the sum of 19 different post-expansion sums, divided into 6 different types. See Table A.2.
- $Q_n - Q_n^*$ is the sum of 37 different post-expansion sums.

The following lemmas are proved in the supplementary material.

TABLE A.1

The 4 types of the 8 post-expansion sums for $\widetilde{T}_n$ ($\|\theta\|_q$ is the ℓ^q-norm of θ (the subscript is dropped when $q = 2$). In our setting, $\alpha\|\theta\| \to \infty$, and $\|\theta\|_4^4 \ll \|\theta\|_3^3 \ll \|\theta\|^2 \ll \|\theta\|_1$.

Type	#	$(N_{\widetilde{\Omega}}, N_W)$	Examples	Mean	Variance
I	1	(0, 3)	$\sum_{i,j,k(dist)} W_{ij}W_{jk}W_{ki}$	0	$\asymp \|\theta\|^6$
II	3	(1, 2)	$\sum_{i,j,k(dist)} \widetilde{\Omega}_{ij}W_{jk}W_{ki}$	0	$\leq C\alpha^2\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^6)$
III	3	(2, 1)	$\sum_{i,j,k(dist)} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}W_{ki}$	0	$\leq C\alpha^4\|\theta\|^4\|\theta\|_3^6$
IV	1	(3, 0)	$\sum_{i,j,k(dist)} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{ki}$	$\sim \mathrm{tr}(\widetilde{\Omega}^3)$	0

THEOREM A.1 (Ideal SgnT test statistic). *Consider the testing problem (1.6) under the DCMM model (1.1)-(1.4), where the condition (2.2) is satisfied under the alternative hypothesis. Suppose $\theta_{\max} \to 0$ and $\|\theta\| \to \infty$ as $n \to \infty$, and suppose $|\lambda_2|/\sqrt{\lambda_1} \to \infty$ under the alternative hypothesis. Then, under the null hypothesis, as $n \to \infty$,*

$$\mathbb{E}[\widetilde{T}_n] = 0, \qquad \mathrm{Var}(\widetilde{T}_n) = 6\|\theta\|^6 \cdot [1 + o(1)],$$

and

$$\frac{\widetilde{T}_n - \mathbb{E}[\widetilde{T}_n]}{\sqrt{\mathrm{Var}(\widetilde{T}_n)}} \longrightarrow N(0,1), \qquad \textit{in law}.$$

Furthermore, under the alternative hypothesis, as $n \to \infty$,

$$\mathbb{E}[\widetilde{T}_n] = \mathrm{tr}(\widetilde{\Omega}^3) + o(\|\theta\|^3), \qquad \mathrm{Var}(\widetilde{T}_n) \leq C\|\theta\|^6 + C(|\lambda_2|/\lambda_1)^4\|\theta\|^4\|\theta\|_3^6.$$

THEOREM A.2 (Proxy SgnT test statistic). *Consider the testing problem (1.6) under the DCMM model (1.1)-(1.4), where the condition (2.2) is satisfied under the alternative hypothesis. Suppose $\theta_{\max} \to 0$ and $\|\theta\| \to \infty$ as $n \to \infty$, and suppose $|\lambda_2|/\sqrt{\lambda_1} \to \infty$ under the alternative hypothesis. Then, under the null hypothesis, as $n \to \infty$,*

$$\mathbb{E}[T_n^* - \widetilde{T}_n] = o(\|\theta\|^3), \qquad \mathrm{Var}(T_n^* - \widetilde{T}_n) = o(\|\theta\|^6).$$

TABLE A.2
The 6 types of the 19 post-expansion sums for $(T_n^* - \widetilde{T}_n)$. Notations: same as Table A.1.

Type	#	$(N_\delta, N_{\widetilde{\Omega}}, N_W)$	Examples	Abs. Mean	Variance
Ia	3	(1, 0, 2)	$\sum_{i,j,k\,(dist)} \delta_{ij} W_{jk} W_{ki}$	0	$\leq C\frac{\|\theta\|^4\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^6)$
Ib	6	(1, 1, 1)	$\sum_{i,j,k\,(dist)} \delta_{ij} \widetilde{\Omega}_{jk} W_{ki}$	$\leq C\alpha\|\theta\|^4 = o(\alpha^3\|\theta\|^6)$	$\leq \frac{C\alpha^2\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^6)$
Ic	3	(1, 2, 0)	$\sum_{i,j,k\,(dist)} \delta_{ij} \widetilde{\Omega}_{jk} \widetilde{\Omega}_{ki}$	0	$\leq \frac{C\alpha^4\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1} = O(\alpha^4\|\theta\|^4\|\theta\|_3^6)$
IIa	3	(2, 0, 1)	$\sum_{i,j,k\,(dist)} \delta_{ij} \delta_{jk} W_{ki}$	$\leq C\|\theta\|^2 = o(\|\theta\|^3)$	$\leq C\|\theta\|_3^6 = o(\|\theta\|^6)$
IIb	3	(2, 1, 0)	$\sum_{i,j,k\,(dist)} \delta_{ij} \delta_{jk} \widetilde{\Omega}_{ki}$	$\leq \frac{C\alpha\|\theta\|^6}{\|\theta\|_1^2} = o(\|\theta\|^3)$	$\leq \frac{C\alpha^2\|\theta\|^{10}}{\|\theta\|_1^2} = o(\|\theta\|^6)$
III	1	(3, 0, 0)	$\sum_{i,j,k\,(dist)} \delta_{ij} \delta_{jk} \delta_{ki}$	$\leq \frac{C\|\theta\|^4}{\|\theta\|_1^2} = o(\|\theta\|^3)$	$\leq \frac{C\|\theta\|^4\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^6)$

Furthermore, under the alternative hypothesis,

$$\mathbb{E}[T_n^* - \widetilde{T}_n] = o((|\lambda_2|/\lambda_1)^3\|\theta\|^6),$$
$$\mathrm{Var}(T_n^* - \widetilde{T}_n) \leq C(|\lambda_2|/\lambda_1)^4\|\theta\|^4\|\theta\|_3^6 + o(\|\theta\|^6).$$

THEOREM A.3 (Real SgnT test statistic). *Consider the testing problem (1.6) under the DCMM model (1.1)-(1.4), where the condition (2.2) is satisfied under the alternative hypothesis. Suppose $\theta_{\max} \to 0$ and $\|\theta\| \to \infty$ as $n \to \infty$, and suppose $|\lambda_2|/\sqrt{\lambda_1} \to \infty$ under the alternative hypothesis. Then, under the null hypothesis, as $n \to \infty$,*

$$|\mathbb{E}[T_n - T_n^*]| = o(\|\theta\|^3), \qquad \textit{and} \qquad \mathrm{Var}(T_n - T_n^*) = o(|\theta\|^6).$$

Under the alternative hypothesis, as $n \to \infty$,

$$|\mathbb{E}[T_n - T_n^*]| = o((|\lambda_2|/\lambda_1)^3\|\theta\|^6),$$
$$\mathrm{Var}(T_n - T_n^*) = o((|\lambda_2|/\lambda_1)^4\|\theta\|^4|\theta\|_3^6) + o(\|\theta\|^6).$$

Combining Theorems A.1, A.2, and A.3, Theorems 2.1, 2.3, and 2.5 follow by similar arguments as in Appendix B.

APPENDIX B: THE BEHAVIOR OF THE SGNQ TEST STATISTIC

We prove Theorems 2.2, 2.4, and 2.6. We use the same notations as those in Section 4 of the main article, and the proof here relies on Theorems 4.1-4.3 in the main article.

Consider Theorem 2.2. In this theorem, we assume the null is true. First, by Theorems 4.2 and 4.3 and elementary statistics, $\mathbb{E}[Q_n^* - \widetilde{Q}_n] \sim 2\|\theta\|^4$, $|\mathbb{E}[Q_n - Q_n^*]| = o(\|\theta\|^4)$, $\mathrm{Var}(Q_n^* - \widetilde{Q}_n) = o(\|\theta\|^8)$, and $\mathrm{Var}(Q_n - Q_n^*) = o(\|\theta\|^8)$. It follows that

$$\mathbb{E}[Q_n] - \mathbb{E}[\widetilde{Q}_n] = (2 + o(1))\|\theta\|^4, \qquad \mathrm{Var}(Q_n - \widetilde{Q}_n) = o(\|\theta\|^8). \tag{2}$$

By Theorem 4.1.

$$\mathbb{E}[\widetilde{Q}_n] = o(\|\theta\|^4), \quad \mathrm{Var}(\widetilde{Q}_n) \sim 8\|\theta\|^8, \quad \frac{\widetilde{Q}_n - \mathbb{E}[\widetilde{Q}_n]}{\sqrt{\mathrm{Var}(\widetilde{Q}_n)}} \to N(0,1). \tag{3}$$

Since for any random variables X and Y, $\mathrm{Var}(X+Y) \leq (1+a_n)\mathrm{Var}(X) + (1+\frac{1}{a_n})\mathrm{Var}(Y)$ for any number $a_n > 0$, combining the above and letting a_n tend to 0 appropriately slow,

$$\mathbb{E}[Q_n] \sim 2\|\theta\|^4, \qquad \mathrm{Var}(Q_n) \sim 8\|\theta\|^8. \tag{4}$$

Moreover, write

$$\frac{Q_n - \mathbb{E}[Q_n]}{\sqrt{\mathrm{Var}(Q_n)}} = \sqrt{\frac{\mathrm{Var}(\widetilde{Q}_n)}{\mathrm{Var}(Q_n)}} \cdot \left[\frac{(Q_n - \widetilde{Q}_n)}{\sqrt{\mathrm{Var}(\widetilde{Q}_n)}} + \frac{\widetilde{Q}_n - \mathbb{E}[\widetilde{Q}_n]}{\sqrt{\mathrm{Var}(\widetilde{Q}_n)}} + \frac{\mathbb{E}[\widetilde{Q}_n] - \mathbb{E}[Q_n]}{\sqrt{\mathrm{Var}(\widetilde{Q}_n)}}\right].$$

On the right hand side, by (2)-(4), as $n \to \infty$, the term outside the bracket $\to 1$, and for the three terms in the bracket, the first one has a mean and variance that tend to 0 so it tends to 0 in probability, the second one weakly converges to $N(0, 1)$, and the last one $\to 0$. Combining these,

$$\frac{Q_n - \mathbb{E}[Q_n]}{\sqrt{\mathrm{Var}(Q_n)}} \to N(0, 1), \qquad \text{in law.} \tag{5}$$

Combining (4) and (5) proves Theorem 2.2.

Next, we consider Theorem 2.4, where we assume the alternative is true. First, similarly, by Theorems 4.2 and 4.3,

$$\mathbb{E}[Q_n^* - \widetilde{Q}_n] = (2 + o(1))\|\theta\|^4 + o((|\lambda_2|/\lambda_1)^4\|\theta\|^8),$$
$$\mathrm{Var}(Q_n - \widetilde{Q}_n) \le C(\lambda_2/\lambda_1)^6\|\theta\|^8\|\theta\|_3^6 + o(\|\theta\|^8).$$

Second, by Theorem 4.1,

$$\mathbb{E}[\widetilde{Q}_n] = \mathrm{tr}(\widetilde{\Omega}^4) + o(\|\theta\|^4), \qquad \mathrm{Var}(\widetilde{Q}_n) \le C[\|\theta\|^8 + (\lambda_2/\lambda_1)^6\|\theta\|^8\|\theta\|_3^6].$$

Combining these proves Theorem 2.4.

Last, we consider Theorems 2.5-2.6. Since the proofs are similar, we only show Theorem 2.6. First, by Theorem 2.2 and Lemma 2.1, under the null, $\frac{Q_n - 2(\|\hat{\eta}\|^2-1)^2}{\sqrt{8(\|\hat{\eta}\|^2-1)^4}} \to N(0, 1)$, so the Type I error is

$$\mathbb{P}_{H_0^{(n)}}\Big(Q_n \ge (2 + z_\alpha\sqrt{8})(|\hat{\eta}\|^2 - 1)^2\Big) = P\bigg(\frac{Q_n - 2(\|\hat{\eta}\|^2 - 1)^2}{\sqrt{8(\|\hat{\eta}\|^2 - 1)^4}} \ge z_\alpha\bigg) = \alpha + o(1).$$

Second, fixing $0 < \epsilon < 1$, let A_ϵ be the event $\{(\|\hat{\eta}\|^2 - 1) \le (1 + \epsilon)\|\eta^*\|^2\}$. By Lemma 2.1 and definitions, on one hand, over the event A_ϵ, $(\|\hat{\eta}\|^2 - 1) \le (1 + \epsilon)\|\eta^*\|^2 \le C\|\theta\|^2$, and on the other hand, $\mathbb{P}(A_\epsilon^c) = o(1)$. Therefore, the Type II error

$$\begin{aligned}
&\mathbb{P}_{H_1^{(n)}}\Big(Q_n \le (2 + z_\alpha\sqrt{8})(\|\hat{\eta}\|^2 - 1)^2\Big) \\
\le&\mathbb{P}_{H_1^{(n)}}\Big(Q_n \le (2 + z_\alpha\sqrt{8})(\|\hat{\eta}\|^2 - 1)^2, A_\epsilon\Big) + \mathbb{P}(A_\epsilon^c) \\
\le&\mathbb{P}_{H_1^{(n)}}\Big(Q_n \le C(2 + z_\alpha\sqrt{8})\|\theta\|^4\Big) + o(1),
\end{aligned}$$

where by Chebyshev's inequality, the first term in the last line

$$\le [\mathbb{E}(Q_n) - C(2 + z_\alpha\sqrt{8})\|\theta\|^4]^{-2} \cdot \mathrm{Var}(Q_n). \tag{6}$$

By Lemma D.2 of the supplementary material and our assumptions, $\lambda_1 \asymp \|\theta\|^2$, $|\lambda_2|/\sqrt{\lambda_1} \to \infty$, and $\|\theta\| \to \infty$. Using Lemma 2.3 $\mathbb{E}[Q_n] \ge C\lambda_2^4 \gg \lambda_1^2$, and it follows that $\mathbb{E}(Q_n) \gg C(2 + z_\alpha\sqrt{8})\|\theta\|^4$, so for sufficiently large n,

$$\mathbb{E}(Q_n) - C(2 + z_\alpha\sqrt{8})\|\theta\|^4 \ge \frac{1}{2}\mathbb{E}[Q_n] \ge C\lambda_2^4.$$

At the same time, by Theorem 2.4,

$$\mathrm{Var}(Q_n) \le C(\|\theta\|^8 + (\lambda_2/\lambda_1)^6 \|\theta\|^8 \|\theta\|_3^6).$$

Combining these, the right hand side of (6) does not exceed

$$C\frac{\|\theta\|^8 + (\lambda_2/\lambda_1)^6 \|\theta\|^8 \|\theta\|_3^6}{\lambda_2^8} = (I) + (II), \tag{7}$$

where $(I) = C\lambda_2^{-8}\|\theta\|^8$ and $(II) = C\lambda_2^{-8}(\lambda_2/\lambda_1)^6\|\theta\|^8\|\theta\|_3^6$. Now, first, since $\lambda_1 \asymp \|\theta\|^2$ and $|\lambda_2|/\sqrt{\lambda_1} \to 0$, $(I) \le C(\lambda_2/\sqrt{\lambda_1})^{-8} \to 0$. Second, since $\lambda_1 \asymp \|\theta\|^2$ and $\|\theta\|_3^6 \le \|\theta\|^4$, $(II) = C\lambda_2^{-2}\lambda_1^{-6}\|\theta\|^8\|\theta\|_3^6 \le C\lambda_2^{-2}$. As $|\lambda_2|/\sqrt{\lambda_1} \to \infty$, $\sqrt{\lambda_1} \asymp \|\theta\|$ with $\|\theta\| \to \infty$, $|\lambda_2| \to \infty$ and $(II) \to 0$. Inserting these into (7), the Type II error $\to 0$ and the claim follows. □

APPENDIX C: MATRIX FORMS OF SIGNED-POLYGON STATISTICS

We prove Theorem 1.1. Recall that $\widetilde{A} = A - \hat{\eta}\hat{\eta}$. By definition,

$$T_n = \mathrm{tr}(\widetilde{A}^3) - \sum_{\substack{\text{at least two of} \\ i,j,k \text{ are equal}}} \widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{ki},$$

$$Q_n = \mathrm{tr}(\widetilde{A}^4) - \sum_{\substack{\text{at least two of} \\ i,j,k,\ell \text{ are equal}}} \widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{k\ell}\widetilde{A}_{\ell i}.$$

First, we derive the matrix form of T_n. If at least two of $\{i,j,k\}$ are equal, there are four cases: (a) $i=j$, $k\neq i$, (b) $j=k$, $i \neq j$, (c) $k=i$, $j\neq k$, (d) $i=j=k$. The first three cases are similar. It follows that

$$\begin{aligned}
T_n &= \mathrm{tr}(\widetilde{A}^3) - 3\sum_{i,k(dist)} \widetilde{A}_{ii}\widetilde{A}_{ik}^2 - \sum_i \widetilde{A}_{ii}^3 \\
&= \mathrm{tr}(\widetilde{A}^3) - 3\Big(\sum_{i,k} \widetilde{A}_{ii}\widetilde{A}_{ik}^2 - \sum_i \widetilde{A}_{ii}^3\Big) - \sum_i \widetilde{A}_{ii}^3 \\
&= \mathrm{tr}(\widetilde{A}^3) - 3\mathrm{tr}(\widetilde{A}\circ\widetilde{A}^2) + 2\mathrm{tr}(\widetilde{A}\circ\widetilde{A}\circ\widetilde{A}).
\end{aligned}$$

This gives the desired expression of T_n.

Next, we derive the matrix form of Q_n. When at least two of $\{i,j,k,\ell\}$ are equal, depending on how many indices are equal, we have four patterns: $\{i,i,i,i\}$, $\{i,i,i,j\}$, $\{i,i,j,j\}$, $\{i,i,j,k\}$, where (i,j,k) are distinct. For each pattern, depending on the appearing locations of the next distinct indices, there are a few variations. Take the pattern $\{i,i,j,k\}$ for example: (a) when a new distinct index appears at location 2 and at location 3, the variations are (i,j,k,i), (i,j,k,j), (i,j,k,k); (b) when a new distinct index appears at location 2 and at location 4, the variations are (i,j,i,k), (i,j,j,k); (c) when a new distinct index appears at location 3 and location 4, the variation is (i,i,j,k). Using similar arguments, we can find all variations of each pattern. They are summarized in Table C.3. Define

$$S_1 = \sum_{i,j,k(dist)} \widetilde{A}_{ii}\widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{ki}, \qquad S_2 = \sum_{i,j,k(dist)} \widetilde{A}_{ij}^2\widetilde{A}_{ik}^2,$$

$$S_3 = \sum_{i,j(dist)} \widetilde{A}_{ii}^2\widetilde{A}_{ij}^2, \qquad S_4 = \sum_{i,j(dist)} \widetilde{A}_{ij}^4,$$

$$S_5 = \sum_{i,j(dist)} \widetilde{A}_{ii}\widetilde{A}_{ij}^2\widetilde{A}_{jj}, \qquad S_6 = \sum_i \widetilde{A}_{ii}^4.$$

TABLE C.3
Decomposition of $\mathrm{tr}(\widetilde{A}^4)$. *We note that the last column sums to* n^4.

Pattern	Variations	Summand	Sum	#Summands
$\{i,j,k,\ell\}$	(i,j,k,ℓ)	$\widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{k\ell}\widetilde{A}_{\ell i}$	Q_n	$n(n-1)(n-2)(n-3)$
$\{i,i,j,k\}$	(i,j,k,i)	$\widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{ki}\widetilde{A}_{ii}$	S_1	$6n(n-1)(n-2)$
	(i,j,k,j)	$\widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{kj}\widetilde{A}_{ji}$	S_2	
	(i,j,k,k)	$\widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{kk}\widetilde{A}_{ki}$	S_1	
	(i,j,i,k)	$\widetilde{A}_{ij}\widetilde{A}_{ji}\widetilde{A}_{ik}\widetilde{A}_{ki}$	S_2	
	(i,j,j,k)	$\widetilde{A}_{ij}\widetilde{A}_{jj}\widetilde{A}_{jk}\widetilde{A}_{ki}$	S_1	
	(i,i,j,k)	$\widetilde{A}_{ii}\widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{ki}$	S_1	
$\{i,i,i,j\}$	(i,j,i,i)	$\widetilde{A}_{ij}\widetilde{A}_{ji}\widetilde{A}_{ii}\widetilde{A}_{ii}$	S_3	$4n(n-1)$
	(i,j,j,j)	$\widetilde{A}_{ij}\widetilde{A}_{jj}\widetilde{A}_{jj}\widetilde{A}_{ji}$	S_3	
	(i,i,j,i)	$\widetilde{A}_{ii}\widetilde{A}_{ij}\widetilde{A}_{ji}\widetilde{A}_{ii}$	S_3	
	(i,i,i,j)	$\widetilde{A}_{ii}\widetilde{A}_{ii}\widetilde{A}_{ij}\widetilde{A}_{ji}$	S_3	
$\{i,i,j,j\}$	(i,j,i,j)	$\widetilde{A}_{ij}\widetilde{A}_{ji}\widetilde{A}_{ij}\widetilde{A}_{ji}$	S_4	$3n(n-1)$
	(i,j,j,i)	$\widetilde{A}_{ij}\widetilde{A}_{jj}\widetilde{A}_{ji}\widetilde{A}_{ii}$	S_5	
	(i,i,j,j)	$\widetilde{A}_{ii}\widetilde{A}_{ij}\widetilde{A}_{jj}\widetilde{A}_{ji}$	S_5	
$\{i,i,i,i\}$	(i,i,i,i)	$\widetilde{A}_{ii}\widetilde{A}_{ii}\widetilde{A}_{ii}\widetilde{A}_{ii}$	S_6	n

It follows from Table C.3 that

$$Q_n = \mathrm{tr}(\widetilde{A}^4) - 4S_1 - 2S_2 - 4S_3 - S_4 - 2S_5 - S_6. \tag{8}$$

What remains is to derive the matrix form of S_1-S_6. By direct calculations,

$$\begin{aligned}
S_1 &= \sum_i \widetilde{A}_{ii}\Big[\sum_{j\neq i,k\neq i}\widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{ki} - \sum_{j\neq i}\widetilde{A}_{ij}\widetilde{A}_{jj}\widetilde{A}_{ji}\Big]\\
&= \sum_i \widetilde{A}_{ii}\Big[\Big(\sum_{j,k}\widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{ki} - 2\sum_j \widetilde{A}_{ij}^2\widetilde{A}_{ii} + \widetilde{A}_{ii}^3\Big) - \Big(\sum_j \widetilde{A}_{ij}^2\widetilde{A}_{jj} - \widetilde{A}_{ii}^3\Big)\Big]\\
&= \sum_{i,j,k}\widetilde{A}_{ii}\widetilde{A}_{ij}\widetilde{A}_{jk}\widetilde{A}_{ki} - 2\sum_{i,j}\widetilde{A}_{ii}^2\widetilde{A}_{ij}^2 - \sum_{i,j}\widetilde{A}_{ii}\widetilde{A}_{ij}^2\widetilde{A}_{jj} + 2\sum_i \widetilde{A}_{ii}^4\\
&= \mathrm{tr}(\widetilde{A}\circ\widetilde{A}^3) - 2\mathrm{tr}(\widetilde{A}\circ\widetilde{A}\circ\widetilde{A}^2) - 1_n'[\mathrm{diag}(\widetilde{A})(\widetilde{A}\circ\widetilde{A})\mathrm{diag}(\widetilde{A})]1_n + 2S_6.
\end{aligned}$$

Moreover, we can derive that

$$\begin{aligned}
S_2 &= \sum_i \Big[\sum_{j\neq i,k\neq i}\widetilde{A}_{ij}^2\widetilde{A}_{ik}^2 - \sum_{j\neq i}\widetilde{A}_{ij}^4\Big]\\
&= \sum_i \Big[\Big(\sum_{j,k}\widetilde{A}_{ij}^2\widetilde{A}_{ik}^2 - 2\sum_j \widetilde{A}_{ij}^2\widetilde{A}_{ii}^2 + \widetilde{A}_{ii}^4\Big) - \Big(\sum_j \widetilde{A}_{ij}^4 - \widetilde{A}_{ii}^4\Big)\Big]\\
&= \sum_{i,j,k}\widetilde{A}_{ij}^2\widetilde{A}_{ik}^2 - 2\sum_{i,j}\widetilde{A}_{ij}^2\widetilde{A}_{ii}^2 - \sum_{i,j}\widetilde{A}_{ij}^4 + 2\sum_i \widetilde{A}_{ii}^4\\
&= \mathrm{tr}(\widetilde{A}^2\circ\widetilde{A}^2) - 2\mathrm{tr}(\widetilde{A}\circ\widetilde{A}\circ\widetilde{A}^2) - 1_n'[\widetilde{A}\circ\widetilde{A}\circ\widetilde{A}\circ\widetilde{A}]1_n + 2S_6.
\end{aligned}$$

It is also easy to see that

$$S_3 = \sum_{i,j}\widetilde{A}_{ii}^2\widetilde{A}_{ij}^2 - \sum_i \widetilde{A}_{ii}^4 = \mathrm{tr}(\widetilde{A}\circ\widetilde{A}\circ\widetilde{A}^2) - S_6,$$

$$S_4 = \sum_{i,j} \widetilde{A}_{ij}^4 - \sum_i \widetilde{A}_{ii}^4 = 1_n'[\widetilde{A} \circ \widetilde{A} \circ \widetilde{A} \circ \widetilde{A}]1_n - S_6,$$

$$S_5 = \sum_{i,j} \widetilde{A}_{ii}\widetilde{A}_{ij}^2\widetilde{A}_{jj} - S_6 = 1_n'[\mathrm{diag}(\widetilde{A})(\widetilde{A} \circ \widetilde{A})\mathrm{diag}(\widetilde{A})]1_n - S_6,$$

$$S_6 = \mathrm{tr}(\widetilde{A} \circ \widetilde{A} \circ \widetilde{A} \circ \widetilde{A}).$$

Plugging the matrix forms of S_1-S_6 into (8), we obtain

$$\begin{aligned} Q_n =& \mathrm{tr}(\widetilde{A}^4) - 4\mathrm{tr}(\widetilde{A} \circ \widetilde{A}^3) - 2\mathrm{tr}(\widetilde{A}^2 \circ \widetilde{A}^2) + 8\mathrm{tr}(\widetilde{A} \circ \widetilde{A} \circ \widetilde{A}^2) - 6\mathrm{tr}(\widetilde{A} \circ \widetilde{A} \circ \widetilde{A} \circ \widetilde{A}) \\ &+ 2 \cdot 1_n'[\mathrm{diag}(\widetilde{A})(\widetilde{A} \circ \widetilde{A})\mathrm{diag}(\widetilde{A})]1_n + 1_n'[\widetilde{A} \circ \widetilde{A} \circ \widetilde{A} \circ \widetilde{A}]1_n. \end{aligned}$$

This gives the desired expression of Q_n.

Last, we discuss the complexity of computing T_n and Q_n. It involves the following operations:

- Compute the matrix $\widetilde{A} = A - \hat{\eta}\hat{\eta}'$.
- Compute the Hadamard product of finitely many matrices.
- Compute the trace of a matrix.
- Compute the matrix DMD for a matrix M and a diagonal matrix D.
- Compute $1_n'M1_n$ for a matrix M.
- Compute the matrices $\widetilde{A}^k$, for $k = 2, 3, 4$.

Excluding the last operation, the complexity is $O(n^2)$. For the last operation, since we can compute $\widetilde{A}^k$ recursively from $\widetilde{A}^k = \widetilde{A}^{k-1}\widetilde{A}$, it suffices to consider the complexity of computing $B\widetilde{A}$, for an arbitrary $n \times n$ matrix B. Write

$$B\widetilde{A} = BA - B\hat{\eta}(\hat{\eta})'.$$

Consider computing BA. The (i,j)-th entry of BA is $\sum_{\ell: A_{\ell j} \neq 0} B_{i\ell}A_{\ell j}$, where the total number of nonzero $A_{\ell j}$ equals to d_j, the degree of node j. Hence, the complexity of computing the (i,j)-th entry of BA is $O(d_j)$. It follows that the complexity of computing BA is $O(\sum_{i,j=1}^n d_j) = O(n^2\bar{d})$. Consider computing $B\hat{\eta}(\hat{\eta})'$. We first compute the vector $v = B\hat{\eta}$ and then compute $v(\hat{\eta})'$, where the complexity of both steps is $O(n^2)$. Combining the above, the complexity of computing $B\widetilde{A}$ is $O(n^2\bar{d})$. We have seen that this is the dominating step in computing T_n and Q_n, so the complexity of the latter is also $O(n^2\bar{d})$.

APPENDIX D: ESTIMATION OF $\|\theta\|$

We prove Lemma 2.1. First, we show that

$$\|\eta^*\|^2 \begin{cases} = \|\theta\|^2, & \text{under the null,} \\ \asymp \|\theta\|^2, & \text{under the alternative.} \end{cases}$$

Recall that $\eta^* = (1/\sqrt{1_n'\Omega 1_n})\Omega 1_n$. Hence,

$$\|\eta^*\|^2 = (1_n'\Omega^2 1_n)/(1_n'\Omega 1_n). \tag{9}$$

Under the null, $\Omega = \theta\theta'$, and the claim follows by direct calculations. Under the alternative, $\Omega = \sum_{k=1}^K \lambda_k \xi_k \xi_k'$, so

$$1_n'\Omega 1_n = \sum_{k=1}^K \lambda_k (1_n'\xi_k)^2, \qquad 1_n'\Omega^2 1_n = \sum_{k=1}^K \lambda_k^2 (1_n'\xi_k)^2.$$

By Lemma E.2, $\lambda_1 \asymp \|\theta\|^2$. By Lemma E.3, $1_n'\xi_1 \asymp \|\theta\|^{-1}\|\theta\|_1$ and $|1_n'\xi_k| = O(\|\theta\|^{-1}\|\theta\|_1)$. It follows that $1_n'\Omega^2 1_n \geq \lambda_1^2(1_n'\xi_1)^2 \geq C\|\theta\|_1^2\|\theta\|^2$ and $1_n'\Omega^2 1_n \leq \lambda_1^2\sum_{k=1}^K(1_n'\xi_k)^2 \leq C\|\theta\|_1^2\|\theta\|^2$. We conclude that

$$1_n'\Omega^2 1_n \asymp \|\theta\|_1^2\|\theta\|^2. \tag{10}$$

Moreover, $1_n'\Omega 1_n \leq |\lambda_1|\sum_{k=1}^K(1_n'\xi_k)^2 \leq C\|\theta\|_1^2$, and by Lemma E.4, $1_n'\Omega 1_n \geq C\|\theta\|_1^2$. It follows that

$$1_n'\Omega 1_n \asymp \|\theta\|_1^2. \tag{11}$$

Plugging (10)-(11) into (9) gives the claim.

Next, we show $(\|\hat{\eta}\|^2 - 1)/\|\eta^*\|^2 \to 1$ in probability. Since $\|\eta^*\| \asymp \|\theta\| \to \infty$ as $n \to \infty$, it suffices to show $\|\hat{\eta}\|^2/\|\eta^*\|^2 \to 1$ in probability. By definition,

$$\|\hat{\eta}\|^2 = \frac{1_n'A^2 1_n}{1_n'A1_n}.$$

Compare this with (9), all we need to show is that in probability,

$$\frac{1_n'A1_n}{1_n'\Omega 1_n} \to 1, \qquad \text{and} \qquad \frac{1_n'A^2 1_n}{1_n'\Omega^2 1_n} \to 1. \tag{12}$$

Since the proofs are similar, we only show the second one. By elementary probability, it is sufficient to show that as $n \to \infty$,

$$\frac{\mathbb{E}[1_n'A^2 1_n]}{1_n'\Omega^2 1_n} \to 1, \qquad \frac{\mathrm{Var}(1_n'A^2 1_n)}{(1_n'\Omega^2 1_n)^2} \to 0. \tag{13}$$

We now prove (13). Consider the first claim. Write

$$1_n'A^2 1_n = \sum_{i,j,k} A_{ij}A_{jk} = \sum_{i\neq j} A_{ij}^2 + \sum_{i,j,k(dist)} A_{ij}A_{jk}. \tag{14}$$

It follows that

$$\mathbb{E}[1_n'A^2 1_n] = \sum_{i\neq j}\Omega_{ij} + \sum_{i,j,k(dist)}\Omega_{ij}\Omega_{jk}.$$

Since $\Omega_{ij} \leq \theta_i\theta_j$ under both hypotheses, we have

$$\begin{aligned}
\left|\mathbb{E}[1_n'A^2 1_n] - 1_n'\Omega 1_n - 1_n'\Omega^2 1_n\right| &\leq \Big|\sum_i \Omega_{ii} + \sum_{\substack{(i,j,k)\text{ are}\\ \text{not distinct}}}\Omega_{ij}\Omega_{jk}\Big| \\
&\leq \sum_i \theta_i^2 + C\sum_{i,j}\theta_i^2\theta_j^2 + C\sum_{i,k}\theta_i^3\theta_k \\
&\leq C\|\theta\|^2 + C\|\theta\|^4 + C\|\theta\|_3^3\|\theta\|_1 \\
&\leq C\|\theta\|_3^3\|\theta\|_1,
\end{aligned}$$

where we have used the universal inequality $\|\theta\|^4 \leq \|\theta\|_3^3\|\theta\|_1$. Since $\|\theta\|_3^3 \leq \theta_{\max}^2\|\theta\|_1 = o(\|\theta\|_1)$, the right hand side is $o(\|\theta\|_1^2) = o(1_n'\Omega 1_n)$. So,

$$\mathbb{E}[1_n'A^2 1_n] = 1_n'\Omega^2 1_n + 1_n'\Omega 1_n + o(1_n'\Omega 1_n). \tag{15}$$

Combining this with (10)-(11) gives

$$\left|\frac{\mathbb{E}[1_n'A^2 1_n]}{1_n'\Omega^2 1_n} - 1\right| \lesssim \frac{1_n'\Omega 1_n}{1_n'\Omega^2 1_n} \asymp \frac{1}{\|\theta\|^2},$$

and the claim follows by $\|\theta\| \to \infty$.

Consider the second claim. By (14),

$$\text{Var}(1_n' A^2 1_n) \leq 2\text{Var}\Big(\sum_{i\neq j} A_{ij}^2\Big) + 2\text{Var}\Big(\sum_{i,j,k(dist)} A_{ij}A_{jk}\Big). \tag{16}$$

We re-write $\sum_{i\neq j} A_{ij}^2 = \sum_{i\neq j} A_{ij} = 2\sum_{i<j} A_{ij}$. The variables $\{A_{ij}\}_{1\leq i<j\leq n}$ are mutually independent. It follows that

$$\text{Var}\Big(\sum_{i\neq j} A_{ij}^2\Big) = 4\sum_{i<j}\text{Var}(A_{ij}) \leq C\sum_{i,j}\Omega_{ij} \leq C\|\theta\|_1^2. \tag{17}$$

Moreover, since $A_{ij}A_{jk} = (\Omega_{ij} + W_{ij})(\Omega_{jk} + W_{jk})$, we have

$$\begin{aligned}\sum_{i,j,k(dist)} A_{ij}A_{jk} &= \sum_{i,j,k(dist)} \Omega_{ij}\Omega_{jk} + 2\sum_{i,j,k(dist)} \Omega_{ij}W_{jk} + \sum_{i,j,k(dist)} W_{ij}W_{jk} \\ &\equiv \sum_{i,j,k(dist)} \Omega_{ij}\Omega_{jk} + X_1 + X_2.\end{aligned}$$

By elementary probability,

$$\text{Var}\Big(\sum_{i,j,k(dist)} A_{ij}A_{jk}\Big) \leq 2\text{Var}(X_1) + 2\text{Var}(X_2).$$

To compute the variance of X_1, we note that

$$X_1 = 4\sum_{j<k}\beta_{jk}W_{jk}, \qquad \beta_{jk} = \sum_{i\notin\{j,k\}}\Omega_{ij}.$$

The variables $\{W_{jk}\}_{1\leq j<k\neq n}$ are mutually independent, and $|\beta_{jk}| \leq C\sum_i \theta_i\theta_j \leq C\|\theta\|_1\theta_j$. It follows that

$$\text{Var}(X_1) \leq C\sum_{j,k}(\|\theta\|_1\theta_j)^2(\theta_j\theta_k) \leq C\|\theta\|_1^3\|\theta\|_3^3.$$

To compute the variance of X_2, we note that

$$\text{Var}(X_2) = \sum_{i,j,k(dist)}\sum_{i',j',k'(dist)}\mathbb{E}[W_{ij}W_{jk}W_{i'j'}W_{j'k'}].$$

The summand is nonzero only when the two variables $\{W_{i'j'}, W_{j'k'}\}$ are the same as the two variables $\{W_{ij}, W_{jk}\}$. This can only happen if $(i,j,k) = (i',j',k')$ or $(i,j,k) = (k',j',i')$, where in either case the summand equals to $\mathbb{E}[W_{ij}^2W_{jk}^2]$. It follows that

$$\text{Var}(X_2) = \sum_{i,j,k(dist)} 2\mathbb{E}[W_{ij}^2W_{jk}^2] \leq C\sum_{i,j,k}\theta_i\theta_j^2\theta_k \leq C\|\theta\|^2\|\theta\|_1^2.$$

Combining the above gives

$$\text{Var}\Big(\sum_{i,j,k(dist)} A_{ij}A_{jk}\Big) \leq C\|\theta\|_1^3\|\theta\|_3^3 + C\|\theta\|^2\|\theta\|_1^2 \leq C\|\theta\|_1^3\|\theta\|_3^3, \tag{18}$$

where we have used the fact that $\|\theta\|_1\|\theta\|_3^3 \geq \|\theta\|^4$ (Cauchy-Schwarz inequality) and $\|\theta\| \to \infty$. Plugging (17)-(18) into (16) gives

$$\text{Var}(1_n' A^2 1_n) \leq C\|\theta\|_1^3\|\theta\|_3^3. \tag{19}$$

Comparing this with (10) and using $\|\theta\|_3^3 \leq \theta_{\max}^2\|\theta\|_1$, we obtain

$$\frac{\text{Var}(1_n' A^2 1_n)}{(1_n'\Omega^2 1_n)^2} \leq \frac{C\|\theta\|_1^3\|\theta\|_3^3}{\|\theta\|_1^4\|\theta\|^4} \leq \frac{C\theta_{\max}^2}{\|\theta\|^4},$$

and the claim follows by $\|\theta\| \to \infty$.

APPENDIX E: SPECTRAL ANALYSIS FOR Ω AND $\widetilde{\Omega}$

We state and prove some useful results about eigenvalues and eigenvectors of Ω and $\widetilde{\Omega}$. In Section E.4, we prove Lemma 2.2 and 2.3 of the main file.

For $1 \leq k \leq K$, let λ_k be the k-th largest (in absolute value) eigenvalue of Ω and let $\xi_k \in \mathbb{R}^n$ be the corresponding unit-norm eigenvector. We write

$$\Xi = [\xi_1, \xi_2, \ldots, \xi_K] = [u_1, u_2, \ldots, u_n]',$$

so that u_i is the i-th row of Ξ. Recall that G is the $K \times K$ matrix $\|\theta\|^{-2}(\Pi'\Theta^2\Pi)$.

E.1. Spectral analysis of Ω. The following lemma relates λ_k and ξ_k to the eigenvalues and eigenvectors of the $K \times K$ matrix $G^{\frac{1}{2}}PG^{\frac{1}{2}}$.

LEMMA E.1. *Consider the DCMM model. Let d_k be the k-th largest (in absolute value) eigenvalue of $G^{\frac{1}{2}}PG^{\frac{1}{2}}$ and let $\beta_k \in \mathbb{R}^K$ be the associated eigenvector, $1 \leq k \leq K$. Then under the null,*

$$\lambda_1 = \|\theta\|^2, \qquad \xi_1 = \pm\theta/\|\theta\|.$$

Under the alternative, for $1 \leq k \leq K$,

$$\lambda_k = d_k\|\theta\|^2, \qquad \xi_k = \|\theta\|^{-1}[\theta \circ (\Pi G^{-\frac{1}{2}}\beta_k)].$$

Under the alternative hypothesis, we further have the following lemma:

LEMMA E.2. *Under the DCMM model, as $n \to \infty$, suppose (2.2) holds. As $n \to \infty$, under the alternative hypothesis,*

$$\lambda_1 \asymp \|\theta\|^2, \qquad \|u_i\| \leq C\|\theta\|^{-1}\theta_i, \quad \textit{for all } 1 \leq i \leq n.$$

The quantities $(1_n'\xi_k)$ play key roles in the analysis of the Signed Polygon tests. By Lemma E.1,

$$\xi_1 = (\|\theta\|)^{-1}\Theta\Pi G^{-1/2}\beta_1,$$

where β_1 is the first eigenvector of $G^{1/2}PG^{1/2}$, corresponding to the largest eigenvalue of $G^{1/2}PG^{1/2}$. It is seen $G^{-1/2}\beta_1$ is the eigenvector of the matrix PG associated with the largest eigenvalue of GP, which is the same as the largest eigenvalue of $G^{1/2}PG^{1/2}$. Since PG is a non-negative matrix, by Perron's theorem, we can assume all entries of $G^{-1/2}\beta_1$ are non-negative. As a result, all entries of ξ_1 are non-negative, and

$$1_n'\xi_1 > 0.$$

The following lemma is proved in Section E.3.

LEMMA E.3. *Under the DCMM model, as $n \to \infty$, suppose (2.2) holds. As $n \to \infty$,*

$$\max_{1 \leq k \leq K} |1_n'\xi_k| \leq C\|\theta\|^{-1}\|\theta\|_1, \qquad 1_n'\xi_1 \geq C\|\theta\|^{-1}\|\theta\|_1.$$

and so for any $2 \leq k \leq K$,

$$|1_n'\xi_k| \leq C|1_n'\xi_1|$$

We also have a lower bound for $1_n'\Omega 1_n$. The following lemma is proved in Section E.3.

LEMMA E.4. *Under the DCMM model, as $n \to \infty$, suppose (2.2) holds. As $n \to \infty$, both under the null hypothesis and the alternative hypothesis,*

$$1_n'\Omega 1_n \geq C\|\theta\|_1^2.$$

E.2. Spectral analysis of $\tilde{\Omega}$. Recall that

$$\widetilde{\Omega} = \Omega - (\eta^*)(\eta^*)', \qquad \text{where } \eta^* = (1/\sqrt{1_n'\Omega 1_n})\Omega 1_n,$$

and $\lambda_1, \ldots, \lambda_K$ are the K nonzero eigenvalues of Ω, arranged in the descending order in magnitude, and $\xi_1, \ldots, \xi_K$ are the corresponding unit-norm eigenvectors of Ω The following lemma is proved in Section E.3.

LEMMA E.5. *Under the DCMM model, as $n \to \infty$, suppose (2.2) holds. Then,*

$$|\lambda_2| \le \|\widetilde{\Omega}\| \le C|\lambda_2|.$$

Moreover, for any fixed integer $m \ge 1$,

$$|(\widetilde{\Omega}^m)_{ij}| \le C|\lambda_2|^m \cdot \|\theta\|^{-2}\theta_i\theta_j, \quad \textit{for all } 1 \le i, j \le n.$$

Recall that $d_1, \ldots, d_K$ are the nonzero eigenvalues of $G^{\frac{1}{2}}PG^{\frac{1}{2}}$. Introduce

$$D = \mathrm{diag}(d_1, d_2, \ldots, d_K), \qquad \widetilde{D} = \mathrm{diag}(d_2, d_3, \ldots, d_K),$$

and

$$h = \Big(\frac{1_n'\xi_2}{1_n'\xi_1}, \frac{1_n'\xi_3}{1_n'\xi_1}, \ldots, \frac{1_n'\xi_K}{1_n'\xi_1}\Big)', \qquad u_0 = \sum_{k=2}^{K} \frac{d_k(1_n'\xi_k)^2}{d_1(1_n'\xi_1)^2}.$$

By Lemma E.3, $1_n'\xi_1 > 0$, so h and u_0 are both well-defined. Write $\Xi = [\xi_1, \xi_2, \ldots, \xi_K]$. The following lemma gives an alternative expression of $\widetilde{\Omega}$.

LEMMA E.6. *Under the DCMM model,*

$$\widetilde{\Omega} = \|\theta\|^2 \cdot \Xi M \Xi',$$

where M is a $K \times K$ matrix satisfying

$$M = \begin{bmatrix} (1+u_0)^{-1}h'\widetilde{D}h & -(1+u_0)^{-1}h'\widetilde{D} \\ -(1+u_0)^{-1}\widetilde{D}h & \widetilde{D} - (d_1(1+u_0))^{-1}\widetilde{D}hh'\widetilde{D} \end{bmatrix}.$$

If additionally $|\lambda_2|/\lambda_1 \to 0$, then for the matrix $\widetilde{M} \in \mathbb{R}^{K,K}$,

$$\widetilde{M} = \|\theta\|^2 \cdot \begin{bmatrix} h'\widetilde{D}h & -h'\widetilde{D} \\ -\widetilde{D}h & \widetilde{D} \end{bmatrix},$$

we have

$$|M_{ij} - \widetilde{M}_{ij}| \le C\lambda_2^2/\lambda_1, \qquad \textit{for all } 1 \le i, j \le K.$$

We now study $\mathrm{tr}(\widetilde{\Omega}^3)$ and $\mathrm{tr}(\widetilde{\Omega}^4)$. They are related to the power of the SgnT test and SgnQ test, respectively. We discuss the two cases $|\lambda_2|/\lambda_1 \to 0$ and $|\lambda_2|/\lambda_1 \ge c_0$ separately. Consider the case of $|\lambda_2|/\lambda_1 = o(1)$. Since $\widetilde{\Omega} = \Xi M \Xi'$, where $\Xi'\Xi = I_K$, we have

$$\mathrm{tr}(\widetilde{\Omega}^3) = \mathrm{tr}(M^3), \qquad \text{and} \qquad \mathrm{tr}(\widetilde{\Omega}^4) = \mathrm{tr}(M^4).$$

The following lemma is proved in Section E.3.

LEMMA E.7. *Consider the DCMM model, where (2.2) holds. As $n \to \infty$, if $|\lambda_2|/\lambda_1 \to 0$, then*

$$|\mathrm{tr}(\widetilde{\Omega}^3) - \mathrm{tr}(\widetilde{M}^3)| \leq o(|\lambda_2|^3), \qquad |\mathrm{tr}(\widetilde{\Omega}^4) - \mathrm{tr}(\widetilde{M}^4)| \leq o(|\lambda_2|^3), \tag{20}$$

Moreover,

$$\mathrm{tr}(\widetilde{M}^3) = \mathrm{tr}(\widetilde{D}^3) + 3h'\widetilde{D}^3h + 3(h'\widetilde{D}h)(h'\widetilde{D}^2h) + (h'\widetilde{D}h)^3,$$

and

$$\begin{aligned}\mathrm{tr}(\widetilde{M}^4) &= \mathrm{tr}(\widetilde{D}^4) + (h'\widetilde{D}h)^4 + 4(h'\widetilde{D}^2h)^2 + 4(h'\widetilde{D}h)^2(h'\widetilde{D}^2h) + 4h'\widetilde{D}^4h + 4(h'\widetilde{D}h)(h'\widetilde{D}^3h)\\ &\geq \mathrm{tr}(\widetilde{D}^4) + (h'\widetilde{D}h)^4 + 2[(h'\widetilde{D}^2h)^2 + (h'\widetilde{D}h)^2(h'\widetilde{D}^2h) + h'\widetilde{D}^4h]\\ &\geq \mathrm{tr}(\widetilde{D}^4).\end{aligned}$$

- *In the special case where $\lambda_2, \lambda_3, \ldots, \lambda_K$ have the same signs,*

$$|\mathrm{tr}(\widetilde{M}^3)| \geq |\sum_{k=2}^{K} \lambda_k^3| = \sum_{k=2}^{K} |\lambda_k|^3,$$

  and so

$$|\mathrm{tr}(\widetilde{\Omega}^3)| \geq \sum_{k=2}^{K} |\lambda_k|^3 + o(|\lambda_2|^3).$$

- *In the special case where $K = 2$, the vector h is a scalar, and*

$$\mathrm{tr}(\widetilde{M}^3) = (1+h^2)^3\lambda_2^3, \qquad \mathrm{tr}(\widetilde{M}^4) = (1+h^2)^4\lambda_2^4,$$

  and so

$$\mathrm{tr}(\widetilde{\Omega}^3) = [(1+h^2)^3 + o(1)]\lambda_2^3, \qquad \mathrm{tr}(\widetilde{\Omega}^4) = [(1+h^2)^4 + o(1)]\lambda_2^4.$$

We now consider the case $|\lambda_2/\lambda_1| \geq c_0$. In this case, $\widetilde{M}$ is not a good proxy for M any more, so we can not derive a simple formula for $\mathrm{tr}(\widetilde{\Omega}^3)$ or $\mathrm{tr}(\widetilde{\Omega}^4)$ as above. However, for $\mathrm{tr}(\widetilde{\Omega}^4)$, since

$$\mathrm{tr}(\widetilde{\Omega}^4) \geq \|\widetilde{\Omega}\|^4,$$

by Lemma E.5, we immediately have

$$\mathrm{tr}(\widetilde{\Omega}^4) \geq C\lambda_2^4 \geq C(\sum_{k=2}^{K} \lambda_k^4)/(K-1) \geq C\sum_{k=2}^{K} \lambda_k^4. \tag{21}$$

E.3. Proof of Lemmas E.1-E.7.

E.3.1. *Proof of Lemma E.1.* The proof for the null case is straightforward, so we only prove the lemma for the alternative case. Consider the spectral decomposition

$$G^{1/2}PG^{1/2} = BDB'.$$

where

$$D = \mathrm{diag}(d_1, \ldots, d_K) \qquad \text{and} \qquad B = [\beta_1, \ldots, \beta_K].$$

Combining this with $\Omega = \Theta\Pi P\Pi'\Theta$ gives

$$\begin{aligned}\Omega &= \Theta\Pi G^{-\frac{1}{2}}(G^{\frac{1}{2}}PG^{\frac{1}{2}})G^{-\frac{1}{2}}\Pi'\Theta\\ &= \Theta\Pi G^{-\frac{1}{2}}(BDB')G^{-\frac{1}{2}}\Pi'\Theta\\ &= (\|\theta\|^{-1}\Theta\Pi G^{-\frac{1}{2}}B)(\|\theta\|^2 D)(\|\theta\|^{-1}\Theta\Pi G^{-\frac{1}{2}}B)'\\ &= H(\|\theta\|^2 D)H',\end{aligned}$$

where

$$H = \|\theta\|^{-1}\Theta\Pi G^{-\frac{1}{2}}B.$$

Recalling that $G = (\|\theta\|^2)^{-1}\cdot\Pi'\Theta^2\Pi$, it is seen

(22) $$H'H = \|\theta\|^{-2}B'G^{-\frac{1}{2}}(\Pi'\Theta^2\Pi)G^{-\frac{1}{2}}B = B'B = I_K,$$

Therefore,

$$\Omega = H(\|\theta\|^2 D)H'$$

is the spectral decomposition of Ω. Since $(\widetilde{D}_k, \xi_k)$ are the k-th eigenvalue of Ω and unit-norm eigenvector respectively, we have

$$\xi_k = \pm 1 \cdot \text{the } k\text{-th column of H} = \pm(\|\theta\|)^{-1}\Theta\Pi G^{-1/2}\beta_k.$$

This proves the claim. □

E.3.2. *Proof of Lemma E.2.* Consider the first claim. By Lemma E.1, $\lambda_1 = d_1\|\theta\|^2$, where d_1 is the maximum eigenvalue of $G^{\frac{1}{2}}PG^{\frac{1}{2}}$. It suffices to show that $d_1 \asymp 1$. Since all entries of P are upper bounded by constants, we have

$$\|P\| \le C.$$

Additionally, since G is a nonnegative symmetric matrix,

(23) $$\|G\| \le \|G\|_{\max} = \max_{1\le k\le K}\sum_{\ell=1}^{K} G(k,\ell) = \|\theta\|^{-2}\max_{1\le k\le K}\sum_{\ell=1}^{K}\sum_{i=1}^{n}\pi_i(k)\pi_i(\ell)\theta_i^2 \le 1.$$

It follows that

(24) $$d_1 \le \|G\|\|P\| \le C.$$

At the same time, for any unit-norm non-negative vector $x \in \mathbb{R}^K$, since all entries of P are non-negative and all diagonal entries of P are 1,

$$x'Px \ge x'x = 1.$$

It follows that

$$d_1 = \|G^{\frac{1}{2}}PG^{\frac{1}{2}}\| \ge \frac{(G^{-\frac{1}{2}}x)'(G^{\frac{1}{2}}PG^{\frac{1}{2}})(G^{-\frac{1}{2}}x)}{\|(G^{-\frac{1}{2}}x)\|^2} = \frac{x'Px}{x'G^{-1}x} \ge \frac{1}{\|G^{-1}\|}.$$

Combining it with the assumption (2.2) gives

(25) $$d_1 \ge C.$$

where we note C denotes a generic constant which may vary from occurrence to occurrence. Combining (24)-(25) gives the claim.

Consider the second claim. Let $B = [\beta_1, \beta_2, \ldots, \beta_K]$ and $D = \mathrm{diag}(d_1, d_2, \ldots, d_K)$ as in the proof of Lemma E.1, where we note B is orthonormal. By Lemma E.1 and definitions,

$$u_i' = \|\theta\|^{-1}\theta_i \pi_i' G^{-\frac{1}{2}} B.$$

It follows that

$$\|u_i\| \leq \|\theta\|^{-1}\theta_i \cdot \|\pi_i\| \|G^{-\frac{1}{2}}\| \|B\| \leq (\|\theta\|)^{-1}\theta_i \|G^{-1/2}\|,$$

where we have used $\|B\| = 1$ and $\|\pi_i\| = [\sum_{k=1}^K \pi_i(k)^2]^{1/2} \leq 1$. Finally, by the assumption (2.2), $\|G^{-1}\| \leq C$ and so $\|G^{-1/2}\| \leq C$. Combining these gives the claim. □

E.3.3. *Proof of Lemma E.3*. It is sufficient to show the first two claims. Consider the first claim. By Lemma E.2, for all $1 \leq k \leq K$ and $1 \leq i \leq n$,

$$|\xi_k(i)| \leq C\|\theta\|^{-1}\theta_i.$$

It follows that

$$|1_n' \xi_k| \leq C \sum_{i=1}^n \|\theta\|^{-1}\theta_i \leq C\|\theta\|^{-1}\|\theta\|_1, \qquad \text{for all } 1 \leq k \leq K, \tag{26}$$

and the claim follows.

Consider the second claim. By Lemma E.1,

$$\xi_1 = \|\theta\|^{-1}\Theta\Pi(G^{-\frac{1}{2}}\beta_1), \tag{27}$$

where β_1 is the (unit-norm) eigenvector of $G^{\frac{1}{2}}PG^{\frac{1}{2}}$ associated with λ_1, which is the largest eigenvalue of $G^{1/2}PG^{1/2}$. By basic algebra, λ_1 is also the largest eigenvalue of the matrix PG, with $G^{-1/2}\beta_1$ being the corresponding eigenvector. Since PG is a nonnegative matrix, $G^{-\frac{1}{2}}\beta_1$ is a nonnegative vector (e.g., [2, Theorem 8.3.1]). Denote for short by

$$h = G^{-1/2}\beta_1.$$

It follows from (27) that

$$1_n'\xi_1 = (\|\theta\|)^{-1} \cdot 1_n'\Theta\Pi h = \|\theta\|^{-1} \cdot \sum_{k=1}^K \Big(\sum_{i=1}^n \pi_i(k)\theta_i\Big) h_k. \tag{28}$$

We note that $\sum_{k=1}^K \big(\sum_{i=1}^n \pi_i(k)\theta_i\big) = \|\theta\|_1$. Combining it with the assumption (2.2) yields

$$\min_{1 \leq k \leq K} \Big\{\sum_{i=1}^n \pi_i(k)\theta_i\Big\} \geq C\|\theta\|_1.$$

Inserting this into (28) gives

$$1_n'\xi_1 \geq C(\|\theta\|)^{-1}\|\theta\|_1 \cdot \|h\|_1. \tag{29}$$

We claim that $\|h\| \geq 1$. Otherwise, if $\|h\| < 1$, then every entry of h is no greater than 1 in magnitude, and so

$$\|h\|_1 \geq \|h\|^2 = \|G^{-1}\beta_1\|^2.$$

Since $\|G^{-1}\| = \|G\|^{-1} \geq 1$ (see (23)) and $\|\beta_1\| = 1$,

$$\|G^{-\frac{1}{2}}\beta_1\| \geq 1.$$

and so it follows $\|h\| \geq 1$. The contradiction show that $\|h\| \geq 1$. The claim follows by combining this with (29). □

E.3.4. *Proof of Lemma E.4.* For $1 \leq k \leq K$, let

$$c = (\|\theta\|_1)^{-1}\Pi'\Theta 1_n = (\|\theta\|_1)^{-1}(1_n'\Theta\Pi)'.$$

Since $\Omega = \Theta\Pi P\Pi'\Theta$ and all entries of P are non-negative,

$$1_n'\Omega 1_n = \|\theta\|_1^2(c'Pc) \geq \|\theta\|^2\Big(\sum_{k=1}^K c_k^2\Big). \tag{30}$$

Note that, first, $c_k \geq 0$, and second, $\|\theta\|_1 \sum_{k=1}^K c_k = 1_n'\Pi\Theta 1_n = 1_n'\Theta 1_n$, where the last term is $\|\theta\|_1$, and so

$$\sum_{k=1}^K c_k = 1.$$

Together with the Cauchy-Schwartz inequality, we have

$$\sum_{k=1}^K c_k^2 \geq (\sum_{k=1}^K c_k)^2/K = 1/K.$$

Combining this with (30) gives the claim. □

E.3.5. *Proof of Lemma E.5.* Consider the first claim. We first derive a lower bound for $\|\widetilde{\Omega}\|$. By Lemma E.6,

$$\|\widetilde{\Omega}\| = \|\theta\|^2 \cdot \|M\|, \tag{31}$$

where with the same notations as in the proof of Lemma E.6, $M = D - (1+u_0)^{-1}vv'$. Let M_0 be the top left 2×2 block of M. Let $D_0 = \mathrm{diag}(d_1, d_2)$, and let v_0 be the sub-vector of v in (36) restricted to the first two coordinates. By (36),

$$M_0 = D_0 - (1+u_0)^{-1}v_0v_0' = D_0^{\frac{1}{2}}\Big(I_2 - (1+u_0)^{-1}D_0^{-1/2}v_0v_0'D_0^{-\frac{1}{2}}\Big)D_0^{\frac{1}{2}},$$

and so by $\|D_0^{-1/2}\| = |d_2|^{-1/2}$ we have

$$\|\Big(I_2 - (1+u_0)^{-1}D_0^{-1/2}v_0v_0'D_0^{-\frac{1}{2}}\Big)\| \leq \|D_0^{-1/2}M_0D_0^{-1/2}\| \leq |d_2|^{-1}\cdot\|M_0\|. \tag{32}$$

At the same time, since $(1+u_0)^{-1}D_0^{-1/2}v_0v_0'D_0^{-1/2}$ is a rank-1 matrix, there is an orthonormal matrix and a number b such that

$$Q(1+u_0)^{-1}D_0^{-1/2}v_0v_0'D_0^{-1/2}Q' = \mathrm{diag}(b, 0).$$

It follows

$$\|\Big(I_2 - (1+u_0)^{-1}D_0^{-1/2}v_0v_0'D_0^{-\frac{1}{2}}\Big)\| = \|I_2 - \mathrm{diag}(b,0)\| = \max\{|1-b|, 1\} \geq 1.$$

Inserting this into (32) gives

$$\|M_0\| \geq |d_2|,$$

Note that $\|M\| \geq \|M_0\|$. Combining this with (31) gives

$$\|\widetilde{\Omega}\| \geq |d_2|\|\theta\|^2.$$

Next, we derive an upper bound for $\|\widetilde{\Omega}\|$. By Lemma E.3,

$$\max_{1\leq k\leq K}|1_n'\xi_k| \leq C\|\theta\|^{-1}\|\theta\|_1, \qquad 1_n'\xi_1 \geq C\|\theta\|^{-1}\|\theta\|_1. \tag{33}$$

By (33), all the entries of M are upper bounded by $C|\lambda_2|$, which implies $\|M\| \leq C|d_2|$. Plugging it into (31) gives

$$\|\widetilde{\Omega}\| \leq \frac{C}{|1+u_0|}|d_2|\|\theta\|^2, \tag{34}$$

and all remains to show is

$$1 + u_0 \geq C > 0.$$

Now, recalling that $\Omega = \sum_{k=1}^{K} \lambda_k \xi_k \xi_k'$ and $\lambda_k = d_k\|\theta\|^2$, by definitions,

$$d_1(1_n'\xi_1)^2(1+u_0) = \sum_{k=1}^{K} d_k(1_n'\xi_k)^2 = \|\theta\|^{-2}1_n'\Omega 1_n.$$

By Lemma E.4 which gives $1_n'\Omega 1_n \geq C\|\theta\|_1^2$. It follows that

$$1 + u_0 \geq \frac{\|\theta\|^{-2}1_n'\Omega 1_n}{d_1(1_n'\xi_1)^2} \geq C\frac{\|\theta\|^{-2}\cdot\|\theta\|_1^2}{\|\theta\|^{-2}\cdot\|\theta\|_1^2} \geq C,$$

where in the second inequality we have used (33) and $d_1 = (\|\theta\|)^{-2}\cdot\lambda_1 \leq 1$ (see Lemma E.2). Inserting this into (34) gives the claim.

Consider the second claim. By Lemma E.6,

$$\widetilde{\Omega} = \Xi M \Xi',$$

where Ξ and M are the same there. Write

$$\Xi = [\xi_1, \xi_2, \ldots, \xi_K] = [u_1, u_2, \ldots, u_n]'.$$

Note that $\widetilde{\Omega}$ and M have the same spectral norm. It follows that

$$\widetilde{\Omega}^m = \Xi M^m \Xi',$$

and

$$|(\widetilde{\Omega}^m)_{ij}| = |u_i' M^m u_j| \leq \|u_i\|\|M\|^m\|u_j\|.$$

By Lemma E.2, $\|u_i\|\|u_j\| \leq C\|\theta\|^{-2}\theta_i\theta_j$, and by the first part of the current lemma,

$$\|M\| = \|\tilde{\Omega}\| \leq C|d_2|\|\theta\|^2.$$

It follows that

$$|(\widetilde{\Omega}^m)_{ij}| \leq C|d_2|^m\|\theta\|^{2m-2}\theta_i\theta_j.$$

This proves the claim. □

E.3.6. *Proof of Lemma E.6.* Consider the first claim. By definitions,

$$\widetilde{\Omega} = \Omega - (\eta^*)(\eta^*)', \qquad \text{where } \eta^* = \frac{1}{\sqrt{1_n'\Omega 1_n}}\Omega 1_n. \tag{35}$$

Recalling $\widetilde{D}_k = d_k\|\theta\|^2$ and $\Xi = [\xi_1, \xi_2, \ldots, \xi_K]$, we have

$$\Omega = \sum_{k=1}^{K} \widetilde{D}_k \xi_k \xi_k' = \|\theta\|^2 \cdot \Xi D \Xi'.$$

It follows that

$$1_n'\Omega 1_n = \|\theta\|^2 \sum_{k=1}^K d_k(1_n'\xi_k)^2,$$

and

$$\eta^* = \frac{\|\theta\|}{\sqrt{\sum_{s=1}^K d_s(1_n'\xi_s)^2}} \sum_{k=1}^K d_k(1_n'\xi_k)\xi_k = \frac{\|\theta\|}{\sqrt{(1+u_0)}} \Big[\sqrt{d_1}\,\xi_1 + \sum_{k=2}^K \frac{d_k(1_n'\xi_k)}{\sqrt{d_1}(1_n'\xi_1)}\xi_k\Big],$$

where the vector in the big bracket on the right is Ξv, if we let

$$v = (\sqrt{d_1}, \frac{d_2(1_n'\xi_2)}{\sqrt{d_1}(1_n'\xi_1)}, \ldots, \frac{d_K(1_n'\xi_K)}{\sqrt{d_1}(1_n'\xi_1)})'.$$

Combining these gives

$$\widetilde{\Omega} = \|\theta\|^2 \Xi D \Xi' - \frac{\|\theta\|^2}{1+u_0} \Xi v v' \Xi.$$

Plugging it into (35) gives

$$\widetilde{\Omega} = \|\theta\|^2 \Xi D \Xi' - \frac{\|\theta\|^2}{1+u_0} \Xi v v' \Xi = \|\theta\|^2 \Xi (D - (1+u_0)^{-1} v v') \Xi'. \tag{36}$$

By definitions,

$$D = \mathrm{diag}(d_1, d_2, \ldots, d_K), \qquad \text{and} \qquad v = d_1^{-1/2} \cdot (d_1, h'\widetilde{D})'.$$

It follows

$$D - (1+u_0)^{-1} v v' = \begin{bmatrix} (1+u_0)^{-1} d_1 u_0 & -(1+u_0)^{-1} h'\widetilde{D} \\ -(1+u_0)^{-1}\widetilde{D}h & \widetilde{D} - (d_1(1+u_0))^{-1}\widetilde{D}hh'\widetilde{D} \end{bmatrix},$$

where we note that

$$d_1 u_0 = \sum_{s=2}^K d_s \frac{(1_n'\xi_s)^2}{(1_n'\xi_1)^2} = h'\widetilde{D}h,$$

Combining these gives the claim.

Consider the second claim. By definitions,

$$M - \widetilde{M} = \|\theta\|^2 \cdot \begin{bmatrix} [(1+u_0)^{-1} - 1] d_1 u_0 & (1 - (1+u_0)^{-1}) h'\widetilde{D} \\ (1 - (1+u_0)^{-1})\widetilde{D}h & -(d_1(1+u_0))^{-1}\widetilde{D}hh'\widetilde{D} \end{bmatrix}.$$

Note that

$$|1 - (1+u_0)^{-1}| \le C|u_0| \le C|\widetilde{D}_2|/\widetilde{D}_1,$$

and that by Lemma E.3,

$$|(1_n'\xi_k)| \le C 1_n'\xi_1,$$

and so each entry of $\widetilde{D}h$ does not exceed $C|d_2|$. It follows that for all $2 \le i, j \le K$,

$$|M_{1i} - \widetilde{M}_{1i}| \le C\|\theta\|^2 (|\widetilde{D}_2|/\widetilde{D}_1) d_2^2 \le C\widetilde{D}_2^2/\widetilde{D}_1,$$

and

$$|M_{ij} - \widetilde{M}_{ij}| \le C\|\theta\|^2 d_1^{-1} d_2^2 \le C\widetilde{D}_2^2/\widetilde{D}_1.$$

Finally,

$$d_1 u_0^2 = d_1^{-1} (\sum_{s=2} d_2 \frac{(1_n' \xi_s)^2}{(1_n' \xi_1)^2})^2 \leq C d_2^2 / d_1,$$

so

$$|M_{11} - \widetilde{M}_{11}| \leq C \|\theta\|^2 d_2^2 / d_1 \leq C \widetilde{D}_2^2 / \widetilde{D}_1.$$

Combining these gives the claim. □

E.3.7. *Proof of Lemma E.7.* It is sufficient to show (20). In fact, once (20) is proved, other claims follow by direct calculations, except for the first inequality regarding $\mathrm{tr}(\widetilde{\Omega}^4)$, we have used

$$|(h'\widetilde{D}h)(h'\widetilde{D}^3 h)| \leq |h'\widetilde{D}h| \sqrt{(h'\widetilde{D}^2 h)(h'\widetilde{D}^4 h)} \leq \frac{1}{2}\Big[(h'\widetilde{D}h)^2 (h'\widetilde{D}^2 h) + h'\widetilde{D}^4 h\Big].$$

We now show (20). Since $\mathrm{tr}(\widetilde{\Omega}^m) = \mathrm{tr}(\widetilde{M}^m)$, for $m = 3, 4$, it is sufficient to show

$$|\mathrm{tr}(M^3) - \mathrm{tr}(\widetilde{M}^3)| \leq C\lambda_2^4/\lambda_1), \qquad |\mathrm{tr}(M^4) - \mathrm{tr}(\widetilde{M}^4)| \leq C|\lambda_2|^5/\lambda_1. \tag{37}$$

Since the proofs are similar, we only show the first one. By basic algebra,

$$\mathrm{tr}(M^3 - \widetilde{M}^3) = \mathrm{tr}((M - \widetilde{M})^3) + 3\mathrm{tr}(\widetilde{M}(M - \widetilde{M})^2) + 3\mathrm{tr}(\widetilde{M}^2(M - \widetilde{M})).$$

By Lemma E.6, for all $1 \leq i, j \leq K$,

$$|M_{ij} - \widetilde{M}_{ij}| \leq C\lambda_2^2/\lambda_1.$$

Also, by Lemma E.3, all entries of h are bounded, so for all $1 \leq i, j \leq K$,

$$|\widetilde{M}_{ij}| \leq |\lambda_2|.$$

It follows

$$|\mathrm{tr}((M - \widetilde{M})^3| \leq C(\lambda_2^2/\lambda_1)^3,$$

$$|\mathrm{tr}(\widetilde{M}(M - \widetilde{M})^2)| \leq C|\lambda_2|(\lambda^2/\lambda_1)^2 \leq C|\lambda_2|^5/\lambda_1^2,$$

and

$$|\mathrm{tr}(\widetilde{M}^2(M - \widetilde{M})| \leq C\lambda_2^2(\lambda^2/\lambda_1) \leq C\lambda_2^4/\lambda_1.$$

where we note that $\lambda_2/\lambda_1 = o(1)$. Combining these gives the claim.

E.4. Proof of Lemmas 2.2 and 2.3. Lemma 2.2 follows directly from Lemma E.7 of this appendix. Consider Lemma 2.3. The second bullet point is a direct result of (21), and the other two bullet points follow directly from Lemma E.7 of this appendix.

APPENDIX F: LOWER BOUNDS, REGION OF IMPOSSIBILITY

We study the Region of Impossibility by considering a DCMM with random mixed memberships. First, in Section F.1, we establish the equivalence between regularity conditions for a DCMM with non-random mixed memberships and those for a DCMM with random mixed memberships. Next, we prove Lemma 3.1, which is key to the construction of inseparable hypothesis pairs. Last, we prove Theorems 3.1-3.5 in the main article.

F.1. Equivalence of regularity conditions. Let $\mu_1, \mu_2, \dots, \mu_K$ be the eigenvalues of P, arranged in the descending order in magnitude. Recall that $\lambda_1, \lambda_2, \dots, \lambda_K$ are the eigenvalues of Ω. The following lemma is proved in Section F.5.

LEMMA F.1 (Equivalent definition of Region of Impossibility). *Consider the DCMM model (1.1)-(1.4), where the alternative is true and the condition (2.2) holds. Suppose $\theta_{\max} \to 0$ and $\|\theta\| \to \infty$ as $n \to \infty$. Then, as $n \to \infty$,*

$$\mu_1 \asymp 1, \qquad \frac{|\mu_2|}{\mu_1} \asymp \frac{|\lambda_2|}{\lambda_1}, \qquad \max_{1\le i,j\le K} |P_{ij} - 1| \le C(|\lambda_2|/\lambda_1).$$

As a result, $|\lambda_2|/\sqrt{\lambda_1} \to 0$ if and only if $\|\theta\| \cdot |\mu_2(P)| \to 0$.

We now consider DCMM with random mixed memberships: Given (Θ, P) and a distribution F over V (the standard simplex in $\mathbb{R}^K$), let

$$\Omega = \Theta\Pi P \Pi'\Theta, \qquad \Pi = [\pi_1, \pi_2, \dots, \pi_n]', \qquad \pi_i \overset{iid}{\sim} F. \tag{38}$$

We notice that the conclusion of Lemma F.1 holds provided that the regularity condition (2.2) is satisfied. The next lemma shows that (2.2) holds with high probability. It is proved in Section F.5.

LEMMA F.2 (Equivalence of regularity conditions). *Consider the model (38). Let $h = \mathbb{E}[\pi_i]$ and $\Sigma = \mathbb{E}[\pi_i\pi_i']$. Suppose $\|P\| \le C$, $\min_{1\le k\le K}\{h_k\} \ge C$ and $\|\Sigma^{-1}\| \le C$. Suppose $\theta_{\max} \to 0$, $\|\theta\| \to \infty$, and $(\|\theta\|^2/\|\theta\|_1)\sqrt{\log(\|\theta\|_1)} \to 0$, as $n \to \infty$. Then, as $n \to \infty$, with probability $1 - o(1)$, the condition (2.2) is satisfied, i.e.,*

$$\frac{\max_{1\le k\le K}\{\sum_{i=1}^n \theta_i \pi_i(k)\}}{\min_{1\le k\le K}\{\sum_{i=1}^n \theta_i \pi_i(k)\}} \le C_0, \qquad \|G^{-1}\| \le C_0,$$

for a constant $C_0 > 0$ and $G = \|\theta\|^{-2}(\Pi'\Theta^2\Pi)$.

F.2. Proof of Lemma 3.1. Let $M = \mathrm{diag}(\mu_1, \mu_2, \dots, \mu_K)$. It is seen $\mu = M1_K$ and so the desired result is to find a D such that

$$DADM1_K = 1_K \iff MDADM1_K = M1_K = \mu \iff D(MAM)D1_K = \mu.$$

Since MAM has strictly positive entries, it is sufficient to show that for any matrix A (MAM in our case; a slight misuse notation here) with strictly positive entries, there is a unique diagonal matrix D with strictly positive diagonal entries such that

$$DAD1_k = \mu_K. \tag{39}$$

We now show the existence and uniqueness separately.

For existence, we follow the proof in [6]. Consider $d'Ad$ for a vector $d \in \mathbb{R}^K$ with strictly positive entries. It is shown there that $d'Ad$ can be minimized using Lagrange multiplier:

$$\frac{1}{2}d'Ad - \lambda\sum_{k=1}^K \mu_k \log(d_k).$$

Differentiating with respect to d and set the derivative to 0 gives

$$Ad = \lambda\sum_{k=1}^K \mu_k/d_k, \tag{40}$$

where $\lambda = d'Ad/(\sum_{k=1}^{K}\mu_k) > 0$. Letting $D = \lambda^{-1/2}\text{diag}(d_1, d_2, \ldots, d_K)$. It is seen that (40) can be rewritten as

$$DAD1_K = \mu,$$

and the claim follows.

For uniqueness, we adapt the proof in [5] to our case. Suppose there are two different eligible diagonal matrices D_1 and D_2 satisfying (39). Let $d_1 = D_1 1_K$ and $d_2 = D_2 1_K$, and let $M = \text{diag}(\mu_1, \mu_2, \ldots, \mu_K)$. It follows that

$$D_2 D_1 A d_1 = D_2 D_1 A D_1 1_K = D_2 \mu = M d_2,$$

and so

$$M^{-1} D_2 D_1 A d_1 = d_2.$$

Now, for a diagonal matrix S with strictly positive diagonal entries to be determined, we have

$$S^{-1} M^{-1} D_2 D_1 A S S^{-1} d_1 = S^{-1} d_2.$$

We pick S such that

$$S^{-1} M^{-1} D_2 D_1 = S,$$

and denote such an S by S_0. It follows

$$S_0 A S_0 (S_0^{-1} d_1) = S_0^{-1} d_2.$$

or equivalently, if we let $\tilde{d}_1 = S_0^{-1} d_1$ and $\tilde{d}_2 = S_0^{-1} d_2$,

$$S_0 A S_0 \tilde{d}_1 = \tilde{d}_2. \tag{41}$$

Similarly, by switching the places of D_1 and D_2, we have

$$S_0 A S_0 \tilde{d}_2 = \tilde{d}_1. \tag{42}$$

Combining (41) and (42) gives

$$S_0 A S_0 (\tilde{d}_1 + \tilde{d}_2) = (\tilde{d}_1 + \tilde{d}_2), \qquad \text{and} \qquad S_0 A S_0 (\tilde{d}_1 - \tilde{d}_2) = -(\tilde{d}_1 - \tilde{d}_2).$$

This implies that 1 and -1 are the two eigenvalues of $S_0 A S_0$, with $\tilde{d}_1 + \tilde{d}_2$ and $\tilde{d}_1 - \tilde{d}_2$ being the corresponding eigenvectors, respectively, where we note that especially, $\tilde{d}_1 + \tilde{d}_2$ has all strictly positive entries. By Perron's theorem [2], since $S_0 A S_0$ have all strictly positive entries, the eigenvector corresponding to the largest eigenvalue (i.e., the Perron root) have all strictly positive entries. As for any symmetric matrix, we can only have one eigenvector that has all strictly positive entries, so 1 must be the Perron root of $S_0 A S_0$. Using Perron's Theorem again, all eigenvalues of $S_0 A S_0$ except the Perron root itself should be strictly smaller than 1 in magnitude. This contradicts with the fact that -1 is an eigenvalue of $S_0 A S_0$. The contradiction proves the uniqueness. □

F.3. Proof of Theorem 3.1. This theorem follows easily from Theorem 3.2 and Theorems 3.3-3.5. Fix (Θ, P, F) such that $\theta \in \mathcal{M}_n^*(\beta_n/2)$ and $\|\theta\| \cdot |\mu_2(P)| \geq 2\alpha_n$. Consider a sequence of hypotheses indexed by n, where $\Omega = \theta\theta'$ under $H_0^{(n)}$, and Ω follows the construction in any of Theorem 3.2 and Theorems 3.3-3.5 under $H_1^{(n)}$. Let $P_0^{(n)}$ and $P_1^{(n)}$ be the probability measures associated with two hypotheses, respectively. By those theorems, the χ^2-distance satisfy

$$\mathcal{D}(P_0^{(n)}, P_1^{(n)}) = o(1), \qquad \text{as } n \to \infty.$$

By connection between L^1-distance and χ^2-distance, it follows that

$$\|P_0^{(n)} - P_1^{(n)}\|_1 = o(1), \qquad \text{as } n \to \infty.$$

We now slightly modify the alternative hypothesis. Let Π_0 be a non-random membership matrix such that $(\theta, \Pi_0, P) \in \mathcal{M}_n(K, c_0, \alpha_n, \beta_n)$. In the modified alternative hypothesis $\widetilde{H}_1^{(n)}$,

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

Let $\widetilde{P}_1^{(n)}$ be the probability measure associated with $\widetilde{H}_1^{(n)}$. By Lemmas F.1-F.2, $\Pi = \widetilde{\Pi}$, except for a vanishing probability. It follows that

$$\|P_1^{(n)} - \widetilde{P}_1^{(n)}\|_1 = o(1), \qquad \text{as } n \to \infty.$$

Under $\widetilde{H}_1^{(n)}$, all realizations (θ, Π, P) are in the class $\mathcal{M}_n(K, c_0, \alpha_n, \beta_n)$. By Neyman-Pearson lemma and elementary inequalities,

$$\begin{aligned} &\inf_{\psi} \Big\{ \sup_{\theta \in \mathcal{M}_n^*(\beta_n)} \mathbb{P}(\psi = 1) + \sup_{(\theta, \Pi, P) \in \mathcal{M}_n(K, c_0, \alpha_n, \beta_n)} \mathbb{P}(\psi = 0) \Big\} \\ &\geq \inf_{\psi} \Big\{ \mathcal{P}_0^{(n)}(\psi = 1) + \widetilde{\mathcal{P}}_1^{(n)}(\psi = 0) \Big\} \\ &\geq 1 - \|P_0^{(n)} - \widetilde{P}_1^{(n)}\|_1 \\ &\geq 1 - \|P_0^{(n)} - P_1^{(n)}\|_1 - \|P_1^{(n)} - \widetilde{P}_1^{(n)}\|_1 \\ &\geq 1 - o(1), \end{aligned}$$

where the second line is because all realizations in $\widetilde{\mathcal{P}}_1^{(n)}$ are in the class $\mathcal{M}_n(K, c_0, \alpha_n, \beta_n)$, and the third line follows from the Neyman-Pearson lemma. □

F.4. Proof of Theorems 3.2-3.5. We note that Theorem 3.2, Theorem 3.4 and Theorem 3.5 can be deduced from Theorem 3.3. To see this, recall that Theorem 3.3 assumes there exists a positive diagonal matrix D such that

$$DPD\widetilde{h}_D = 1_K, \qquad \min_{1 \leq k \leq K} \{\widetilde{h}_{D,k}\} \geq C, \tag{43}$$

where $\widetilde{h}_D = \mathbb{E}[D^{-1}\pi_i / \|D^{-1}\pi_i\|_1]$. We show that the condition (43) is implied by conditions of other theorems. Theorem 3.2 assumes $\pi_i \in \{e_1, e_2, \ldots, e_K\}$. It follows that $D^{-1}\pi_i / \|D^{-1}\pi_i\|_1 = \pi_i$, and so $\widetilde{h}_D = h$. By Lemma 3.1, there exists D such that $DPDh = 1_K$, hence, (43) is satisfied. Theorem 3.4 constructs the alternative hypothesis using $\widetilde{\pi}_i = D\pi_i / \|D\pi_i\|_1$. Equivalently, $D^{-1}\widetilde{\pi}_i / \|D^{-1}\widetilde{\pi}_i\|_1 = \pi_i$, and so $\widetilde{h}_D$ becomes h. Since $DPDh = 1_K$, condition (43) holds. Theorem 3.5 assumes $Ph = q_n 1_K$. Let $D = q_n^{-1/2} I_K$. Then, $\widetilde{h}_D = h$ and $DPDh = q_n^{-1} Ph = 1_K$. Again, (43) is satisfied.

We only need to prove Theorem 3.3. Let $P_0^{(n)}$ and $P_1^{(n)}$ be the probability measure associated with $H_0^{(n)}$ and $H_1^{(n)}$, respectively. Let $\mathcal{D}(P_0^{(n)}, P_1^{(n)})$ be the chi-square distance between two probability measures. By elementary probability,

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

It suffices to show that, when $\|\theta\| \cdot \mu_2(P) \to 0$,

$$\int \left[\frac{dP_1^{(n)}}{dP_0^{(n)}}\right]^2 dP_0^{(n)} = 1 + o(1). \tag{44}$$

Let p_{ij} and $q_{ij}(\Pi)$ be the corresponding Ω_{ij} under the null and the alternative, respectively. It is seen that

$$dP_0^{(n)} = \prod_{i<j} p_{ij}^{A_{ij}} (1-p_{ij})^{1-A_{ij}}, \qquad dP_1^{(n)} = \mathbb{E}_\Pi \Big[\prod_{i<j} [q_{ij}(\Pi)]^{A_{ij}} [1 - q_{ij}(\Pi)]^{1-A_{ij}}\Big].$$

Let $\tilde{\Pi}$ be an independent copy of Π. Then,

$$\begin{aligned}\left[\frac{dP_1^{(n)}}{dP_0^{(n)}}\right]^2 &= \mathbb{E}_\Pi \Big[\prod_{i<j} \Big(\frac{q_{ij}(\Pi)}{p_{ij}}\Big)^{A_{ij}} \Big(\frac{1-q_{ij}(\Pi)}{1-p_{ij}}\Big)^{1-A_{ij}}\Big] \cdot \mathbb{E}_{\tilde{\Pi}} \Big[\prod_{i<j} \Big(\frac{q_{ij}(\tilde{\Pi})}{p_{ij}}\Big)^{A_{ij}} \Big(\frac{1-q_{ij}(\tilde{\Pi})}{1-p_{ij}}\Big)^{1-A_{ij}}\Big] \\ &= \mathbb{E}_{\Pi,\tilde{\Pi}} \underbrace{\Big[\prod_{i<j} \Big(\frac{q_{ij}(\Pi) q_{ij}(\tilde{\Pi})}{p_{ij}^2}\Big)^{A_{ij}} \Big(\frac{[1-q_{ij}(\Pi)][1-q_{ij}(\tilde{\Pi})]}{[1-p_{ij}]^2}\Big)^{1-A_{ij}}\Big]}_{S(A,\Pi,\tilde{\Pi})}.\end{aligned}$$

It follows that

$$\begin{aligned}\int \left[\frac{dP_1^{(n)}}{dP_0^{(n)}}\right]^2 dP_0^{(n)} &= \mathbb{E}_A \left[\frac{dP_1^{(n)}}{dP_0^{(n)}}\right]^2 \\ &= \mathbb{E}_{A,\Pi,\tilde{\Pi}}[S(A,\Pi,\tilde{\Pi})] \\ &= \mathbb{E}_{\Pi,\tilde{\Pi}} \big\{\mathbb{E}_A \big[S(A,\Pi,\tilde{\Pi})|\Pi,\tilde{\Pi}\big]\big\},\end{aligned}$$

where the distribution of $A|(\Pi,\tilde{\Pi})$ is under the null hypothesis. Under the null hypothesis, A is independent of $(\Pi,\tilde{\Pi})$, the upper triangular entries of A are independent of each other, and $A_{ij} \sim \text{Bernoulli}(p_{ij})$. It follows that

$$\begin{aligned}\mathbb{E}_A \big[S(A,\Pi,\tilde{\Pi})|\Pi,\tilde{\Pi}\big] &= \prod_{i<j} \mathbb{E}_A \bigg[\Big(\frac{q_{ij}(\Pi) q_{ij}(\tilde{\Pi})}{p_{ij}^2}\Big)^{A_{ij}} \Big(\frac{[1-q_{ij}(\Pi)][1-q_{ij}(\tilde{\Pi})]}{[1-p_{ij}]^2}\Big)^{1-A_{ij}} \bigg| \Pi,\tilde{\Pi}\bigg] \\ &= \prod_{i<j} \bigg\{p_{ij} \frac{q_{ij}(\Pi) q_{ij}(\tilde{\Pi})}{p_{ij}^2} + (1-p_{ij}) \frac{[1-q_{ij}(\Pi)][1-q_{ij}(\tilde{\Pi})]}{[1-p_{ij}]^2}\bigg\} \\ &= \prod_{i<j} \bigg\{\frac{q_{ij}(\Pi) q_{ij}(\tilde{\Pi})}{p_{ij}} + \frac{[1-q_{ij}(\Pi)][1-q_{ij}(\tilde{\Pi})]}{1-p_{ij}}\bigg\}.\end{aligned}$$

Let $\Delta_{ij} = q_{ij}(\Pi) - p_{ij}$ and $\tilde{\Delta}_{ij} = q_{ij}(\tilde{\Pi}) - p_{ij}$. By direct calculations,

$$\frac{q_{ij}(\Pi) q_{ij}(\tilde{\Pi})}{p_{ij}} + \frac{[1-q_{ij}(\Pi)][1-q_{ij}(\tilde{\Pi})]}{1-p_{ij}} = 1 + \frac{\Delta_{ij}\tilde{\Delta}_{ij}}{p_{ij}(1-p_{ij})}.$$

Combining the above gives

$$\int \left[\frac{dP_1^{(n)}}{dP_0^{(n)}}\right]^2 dP_0^{(n)} = \mathbb{E}_{\Pi,\tilde{\Pi}} \Big[\prod_{i<j} \Big(1 + \frac{\Delta_{ij}\tilde{\Delta}_{ij}}{p_{ij}(1-p_{ij})}\Big)\Big]. \tag{45}$$

We then plug in the expressions of Δ_{ij} and $\tilde{\Delta}_{ij}$ from the model. Let D be the matrix in (43). Introduce $M = DPD - \mathbf{1}_K\mathbf{1}_K'$. We re-write

$$DPD = \mathbf{1}_K\mathbf{1}_K' + M.$$

It is seen that $M\widetilde{h}_D = \mathbf{0}_K$. The following lemma is proved in Section F.5.

LEMMA F.3. *Under the conditions of Theorem 3.3,* $\|M\| \leq C|\mu_2(P)|$.

Write for short $\pi_i^D = \frac{1}{\|D^{-1}\pi_i\|_1}D^{-1}\pi_i$ and $y_i = \pi_i^D - \mathbb{E}[\pi_i^D] = \pi_i^D - \widetilde{h}_D$. Under the alternative hypothesis,

$$\begin{aligned}
q_{ij}(\Pi) &= \theta_i\theta_j\|D^{-1}\pi_i\|_1\|D^{-1}\pi_j\|_1 \cdot \pi_i'P\pi_j \\
&= \theta_i\theta_j \cdot (\pi_i^D)'(DPD)(\pi_j^D) \\
&= \theta_i\theta_j \cdot (\pi_i^D)'(\mathbf{1}_K\mathbf{1}_K' + M)(\pi_j^D) \\
&= \theta_i\theta_j \cdot \left[1 + (\pi_i^D)'M(\pi_j^D)\right] \\
&= \theta_i\theta_j \cdot \left[1 + (\widetilde{h}_D + y_i)'M(\widetilde{h}_D + y_j)\right] \\
&= \theta_i\theta_j \cdot (1 + y_i'My_j).
\end{aligned}$$

Here, the fourth line is due to $\mathbf{1}_K'\pi_i = 1$ and the last line is due to $M\widetilde{h}_D = \mathbf{0}_K$. Under the null hypothesis, $p_{ij} = \theta_i\theta_j$. As a result,

$$\Delta_{ij} = \theta_i\theta_j \cdot y_i'My_j, \qquad y_i \equiv \pi_i^D - \mathbb{E}[\pi_i^D].$$

Similarly, $\tilde{\Delta}_{ij} = \theta_i\theta_j \cdot \tilde{y}_i'M\tilde{y}_j$, with $\tilde{y}_i = \tilde{\pi}_i^D - \mathbb{E}[\tilde{\pi}_i^D]$. We plug them into (45) and use $p_{ij} = \theta_i\theta_j$. It gives

$$\int\left[\frac{dP_1^{(n)}}{dP_0^{(n)}}\right]^2 dP_0^{(n)} = \mathbb{E}\left[\prod_{i<j}\left(1 + \frac{\theta_i\theta_j}{1-\theta_i\theta_j}(y_i'My_j)(\tilde{y}_i'M\tilde{y}_j)\right)\right], \tag{46}$$

where $\{y_i, \tilde{y}_i\}_{i=1}^n$ are iid random vectors with $\mathbb{E}[y_i] = \mathbf{0}_K$.

We bound the right hand side of (46). Since $1 + x \leq e^x$ for all $x \in \mathbb{R}$,

$$\mathcal{D}(P_0^{(n)}, P_1^{(n)}) \leq \mathbb{E}[\exp(S)], \qquad \text{where} \quad S \equiv \sum_{i<j}\frac{\theta_i\theta_j}{1-\theta_i\theta_j}(y_i'My_j)(\tilde{y}_i'M\tilde{y}_j).$$

Let $M = \sum_{k=1}^K \delta_k b_k b_k'$ be the eigen-decomposition of M. Then,

$$(y_i'My_j)(\tilde{y}_i'M\tilde{y}_j) = \sum_{1\leq k,\ell\leq K}\delta_k\delta_\ell(b_k'y_i)(b_k'y_j)(b_\ell'\tilde{y}_i)(b_\ell'\tilde{y}_j).$$

This allows us to decompose

$$S = \frac{1}{K^2}\sum_{1\leq k,\ell\leq K}S_{k\ell}, \qquad \text{where } S_{k\ell} = K^2\delta_k\delta_\ell\sum_{i<j}\frac{\theta_i\theta_j}{1-\theta_i\theta_j}(b_k'y_i)(b_k'y_j)(b_\ell'\tilde{y}_i)(b_\ell'\tilde{y}_j).$$

By Jensen's inequality, $\exp(\frac{1}{K^2}\sum_{k,\ell}S_{k\ell}) \leq \frac{1}{K^2}\sum_{k,\ell}\exp(S_{k\ell})$. It follows that

$$\int\left[\frac{dP_1^{(n)}}{dP_0^{(n)}}\right]^2 dP_0^{(n)} \leq \mathbb{E}[\exp(S)] \leq \max_{1\leq k,\ell\leq K}\mathbb{E}[\exp(S_{k\ell})]. \tag{47}$$

We now fix (k,ℓ) and derive a bound for $\mathbb{E}[\exp(S_{k\ell})]$. For n large enough, $\theta_{\max} \leq 1/2$ and $K^4\|M\|^2\|\theta\|^2 \leq 1/9$. By Taylor expansion of $(1-\theta_i\theta_j)^{-1}$,

$$S_{k\ell} = K^2\delta_k\delta_\ell \sum_{i<j}\sum_{m=1}^{\infty} \theta_i^m\theta_j^m (b_k' y_i)(b_k' y_j)(b_\ell' \tilde{y}_i)(b_\ell' \tilde{y}_j)$$

$$\equiv \sum_{m=1}^{\infty} X_m, \qquad \text{where} \quad X_m \equiv K^2\delta_k\delta_\ell \sum_{i<j} \theta_i^m\theta_j^m (b_k' y_i)(b_k' y_j)(b_\ell' \tilde{y}_i)(b_\ell' \tilde{y}_j).$$

Since $|X_m| \leq C\|M\|^2\|\theta\|_m^{2m} \leq C\|M\|\|\theta\|_1^2\theta_{\max}^{2(m-1)}$, where $\sum_{m=1}^\infty \theta_{\max}^{2(m-1)} < \infty$, the random variable $\sum_{m=1}^\infty X_m$ is always well-defined. For $m \geq 1$, let $a_m = \theta_{\max}^{2(m-1)}(1-\theta_{\max}^2)$. Then, $\sum_{m=1}^\infty a_m = 1$. By Jenson's inequality,

$$\exp\Big(\sum_{m=1}^{\infty} X_m\Big) = \exp\Big(\sum_{m=1}^{\infty} a_m \cdot a_m^{-1}|X_m|\Big) \leq \sum_{m=1}^{\infty} a_m \cdot \exp(a_m^{-1}X_m).$$

Using Fatou's lemma, we have

$$\mathbb{E}[\exp(S_{k\ell})] \leq \sum_{m=1}^{\infty} a_m \cdot \mathbb{E}\big[\exp(a_m^{-1}X_m)\big]. \tag{48}$$

By definition of X_m,

$$X_m = K^2\delta_k\delta_\ell \Big\{ \Big[\sum_i \theta_i^m (b_k' y_i)(b_\ell' \tilde{y}_i)\Big]^2 - \sum_i \theta_i^{2m}(b_k' y_i)^2(b_\ell' \tilde{y}_i)^2 \Big\}.$$

Note that $\max_i\{\|y_i\|, \|\tilde{y}_i\|\} \leq \sqrt{K}$ and $\max_k |\delta_k| = \|M\|$. Therefore,

$$|X_m| \leq K^2\|M\|^2 \Big[\sum_i \theta_i^m (b_k' y_i)(b_\ell' \tilde{y}_i)\Big]^2 + K^4\|M\|^2\|\theta\|_{2m}^{2m}.$$

Write $Y = \sum_i \theta_i^m (b_k' y_i)(b_\ell' \tilde{y}_i)$. We see that Y is sum of independent, mean-zero random variables. Since $|(b_k' y_i)(b_\ell' \tilde{y}_i)| \leq K$, by Hoeffding's inequality,

$$\mathbb{P}(|Y| > t) \leq 2\exp\Big(-\frac{t^2}{4K^2\|\theta\|_{2m}^{2m}}\Big), \qquad \text{for any } t > 0.$$

Since $\|\theta\|_{2m}^{2m} \leq \|\theta\|^2\theta_{\max}^{2(m-1)} \leq 2a_m\|\theta\|^2$, we have $a_m^{-1}K^4\|M\|^2\|\theta\|_{2m}^{2m} \leq 2K^4\|M\|^2\|\theta\|^2$. Note that $K^4\|M\|^2\|\theta\|^2 \leq 1/9$. By direct calculations,

$$\begin{aligned}
\mathbb{E}\big[\exp(a_m^{-1}|X_m|)\big] &\leq e^{a_m^{-1}K^4\|M\|^2\|\theta\|_{2m}^{2m}} \cdot \mathbb{E}\big[e^{a_m^{-1}K^2\|M\|^2Y^2}\big] \\
&\leq e^{2K^4\|M\|^2\|\theta\|^2} \cdot \mathbb{E}\big[e^{a_m^{-1}K^2\|M\|^2Y^2}\big] \\
&= e^{2K^4\|M\|^2\|\theta\|^2}\Big[1 + \int_0^\infty e^t \cdot \mathbb{P}\big(a_m^{-1}K^2\|M\|^2Y^2 > t\big)dt\Big] \\
&\leq e^{2K^4\|M\|^2\|\theta\|^2}\Big[1 + \int_0^\infty e^t \cdot e^{-\frac{t}{8K^4\|M\|^2\|\theta\|^2}}dt\Big] \\
&\leq e^{K^4\|M\|^2\|\theta\|^2} \cdot (1 + 72K^4\|M\|^2\|\theta\|^2).
\end{aligned}$$

We plug it into (48) and notice that $\sum_{m=1}^\infty a_m = 1$. It gives

$$\mathbb{E}[\exp(S_{k\ell})] \leq e^{K^4\|M\|^2\|\theta\|^2} \cdot (1 + 72K^4\|M\|^2\|\theta\|^2). \tag{49}$$

Combining (47) and (49) gives

$$\int \left[\frac{dP_1^{(n)}}{dP_0^{(n)}}\right]^2 dP_0^{(n)} \leq e^{K^4\|M\|^2\|\theta\|^2} \cdot (1 + 72K^4\|M\|^2\|\theta\|^2).$$

We recall that $\|\theta\| \cdot \|M\| \leq C\|\theta\| \cdot |\mu_2(P)| \to 0$. Hence, the right hand side is $1 + o(1)$. This proves (44).

F.5. Proof of Lemmas F.1-F.3.

F.5.1. *Proof of Lemma F.1.* The first claim follows by our assumptions on P, so we omit the proof. Consider the second claim. Recall that $G = \|\theta\|^{-2}\Pi'\Theta^2\Pi$ and $d_1, d_2, \ldots, d_K$ are the eigenvalues of $G^{1/2}PG^{1/2}$, arranged in the descending order in magnitude. By Lemmas D.1 and D.2, $\lambda_k = \|\theta\|^2 d_k$, $1 \leq k \leq K$, and $d_1 \asymp 1$. Combining these, it suffices to show

$$|\mu_2| \asymp |d_2|.$$

We now prove for the cases where P is non-singular and singular, separately. Consider the first case. Since $1/d_k$ and $1/\mu_K$ are the largest eigenvalue of $G^{-1/2}P^{-1/2}G^{-1/2}$ and P^{-1} in magnitude, respectively, and $\|G\| \leq C$ and $\|G^{-1}\| \leq C$, it is seen that $|\mu_K| \asymp |d_K|$. To show the claim, it sufficient to show that for any $m \geq 2$, if $|\mu_k| \asymp |d_k|$ for $k = m+1, \ldots, K$, then $|\mu_m| \asymp |d_m|$.

We now fix $m \geq 2$, and assume $|\mu_k| \asymp |d_k|$ for $k = m+1, \ldots, K$. The goal is to show $|\mu_m| \asymp |d_m|$. By symmetry, it is sufficient to show that

$$|d_m| \leq C|\mu_m|. \tag{50}$$

Let $P = V\text{diag}(d_1, d_2, \ldots, d_K)V'$ be the SVD of P, where $V \in \mathbb{R}^{K,K}$ is orthonormal, and let V_m be the sub-matrix of V consisting the first m columns of V. Introduce

$$\widetilde{P}_m = V_m D_m V_m', \qquad \text{where } D_m = \text{diag}(d_1, d_2, \ldots, d_m).$$

Let $\mu_1^*, \mu_2^*, \ldots, \mu_m^*$ and $d_1^*, d_2^*, \ldots, d_m^*$ be the first m eigenvalues of $\widetilde{P}_m$ and $G^{1/2}P_mG^{1/2}$, respectively, arranged in the descending order in magnitude. Since $\|G\| \leq C$, we have

$$\|P - P_m\| \leq C|\mu_{m+1}|, \qquad \|G^{1/2}(P - P_m)G^{1/2}\| \leq C|\mu_{m+1}|.$$

By Theorem [1, Theorem A.46],

$$|\mu_m - \mu_m^*| \leq C\|P - P_m\| \leq |\lambda_{m+1}|, \tag{51}$$

and

$$|d_m - d_m^*| \leq \|G^{1/2}(P - P_m)G^{1/2}\| \leq C|\mu_{m+1}|. \tag{52}$$

At the same time, note that the nonzero eigenvalues of $G^{1/2}P_mG^{1/2}$ are the same as the nonzero eigenvalues of $D_mV_m'GV_m$, and also the same as those of $(V_m'GV_m)^{1/2}D_m(V_m'GV_m)^{1/2}$. Since $\|G\| \leq C$ and $\|G^{-1}\| \leq C$, it is seen $\|V_m'GV_m\| \leq C$ and $\|V_m'GV_m)^{-1}\| \leq C$. Therefore, by similar arguments,

$$|\mu_m^*| \asymp |d_m^*|. \tag{53}$$

Combining (51), (52), and (53) gives

$$\begin{aligned} |\mu_m| \leq |\mu_m^*| + |\mu_m - \mu_m^*| &\leq C(|d_m^*| + |d_{m+1}|) \\ \leq C[(|d_m| + |d_m - d_m^*|) + |d_{m+1}|] &\leq C|d_m|. \end{aligned}$$

This proves (50) and the claim follows.

We now consider the case where P is singular, say, $rank(P) = r < K$, and the nonzero eigenvalues are $\mu_1, \mu_2, \ldots, \mu_r$. Let $P = UDU'$ be the SVD, where $U \in \mathbb{R}^{n,r}$ and $D = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_r)$. By similar argument, the nonzero eigenvalues of $G^{1/2}PG^{1/2}$ are the same as $(U'GU)^{1/2}D(U'GU)^{1/2}$, where $\|U'GU\| \leq C$ and $\|(U'GU)^{-1}\| \leq C$. The remaining part of the proof is similar so is omitted.

Consider the last claim. Let $\widetilde{P} = \eta\eta'$, where η is the first eigenvector of P, scaled to have a ℓ^2-norm of $\sqrt{\mu_1}$. Write

$$|P_{ij} - 1| = |P_{ij} - \eta_i\eta_j| + |\eta_i\eta_j - 1|. \tag{54}$$

Now, first, by definitions and elementary algebra, for $1 \leq i, j \leq K$,

$$|P_{ij} - \eta_i\eta_j| \leq |P_{ij} - \widetilde{P}_{ij}| \leq \|P - \widetilde{P}\| \leq \mu_2, \tag{55}$$

where by the second claim, $\mu_2 = o(1)$. Note that for $1 \leq i, j \leq K$, $P_{ii} = 1$ and $P_{ij} \geq 0$. It is seen that $|\eta_i| = 1 + o(1)$ and all η_i must have the positive sign. It follows $|\eta_i - 1| = (1+\eta_i)^{-1}(1 - \eta_i^2) \leq \mu_2$, and so

$$|1 - \eta_i\eta_j| \leq |(1-\eta_i)(1-\eta_j)| + |1 - \eta_i| + |1 - \eta_j| \leq C\mu_2. \tag{56}$$

Combining (54)-(56) gives the claim. □

F.5.2. *Proof of Lemma F.2.* Consider the first claim about $\sum_i \theta_i\pi_i(k)$. Write $X = \sum_{i=1}^n \theta_i(\pi_i(k) - h_k)$. It is seen that X is sum of independent mean-zero random variables, where $\theta_i|\pi_i(k) - h_k| \leq C\theta_{\max}$ and $\sum_{i=1}^n \mathrm{Var}(\theta_i(\pi_i(k) - h_k)) \leq C\|\theta\|^2$. By Bernstein's inequality, for any $t > 0$,

$$\mathbb{P}(|X| > t) \leq \exp\Big(-\frac{t^2}{C\|\theta\|^2 + C\theta_{\max}t}\Big).$$

It follows that, with probability $1 - \|\theta\|_1^{-1}$,

$$\Big|\sum_i \theta_i\pi_i(k) - h_k\|\theta\|_1\Big| = |X| \leq C\|\theta\|\sqrt{\log(\|\theta\|_1)} + C\theta_{\max}\log(\|\theta\|_1).$$

Since $\|\theta\| \to \infty$, $\theta_{\max} \to 0$, and $(\|\theta\|^2/\|\theta\|_1)\sqrt{\log(\|\theta\|_1)} \to 0$, the right hand side is $o(\|\theta\|_1)$. Combining it with the assumption of $\min_k\{h_k\} \geq C$, we have

$$\sum_i \theta_i\pi_i(k) \geq C\|\theta\|_1, \qquad \text{with probability } 1 - \|\theta\|^{-1} = 1 - o(1).$$

Additionally, since $\pi_i(k) \leq 1$, $\sum_i \theta_i\pi_i(k) \leq \|\theta\|_1$. Therefore, with probability $1 - o(1)$, each $\sum_i \theta_i\pi_i(k)$ is at the order of $\|\theta\|_1$. This proves the first claim.

Consider the second claim about G. Let $y_i = \pi_i - h$. Then, $\pi_i\pi_i' = hh' + hy_i' + y_ih' + y_iy_i'$ and $\Sigma = \mathbb{E}[\pi_i\pi_i'] = hh' + \mathbb{E}[y_iy_i']$. It follows that

$$\begin{aligned}
\|\theta\|^2 G = \sum_{i=1}^n \theta_i^2\pi_i\pi_i' &= \sum_{i=1}^n \theta_i^2\big(\Sigma + hy_i' + y_ih' + y_iy_i' - \mathbb{E}[y_iy_i']\big) \\
&= \|\theta\|^2\Sigma + \sum_{i=1}^n \theta_i^2(y_iy_i' - \mathbb{E}[y_iy_i']) + \sum_{i=1}^n \theta_i^2 hy_i' + \sum_{i=1}^n \theta_i^2 y_ih' \\
&\equiv \|\theta\|^2\Sigma + Z_0 + Z_1 + Z_2.
\end{aligned}$$

Here, Z_0 is the sum of independent, mean-zero random matrices. We apply the matrix Hoeffding inequality [7] to bound its operator norm. Since $\theta_i^2\|y_iy_i' - \mathbb{E}[y_iy_i']\| \leq C\theta_i^2$, the matrix

Hoeffding inequality implies that $\mathbb{P}\big(\|Z_0\| > t\big) \le \exp\big(-\frac{t^2}{C^*\|\theta\|_4^4}\big)$ for all $t > 0$, where $C^* > 0$ is a constant. Let ζ_n be a sequence such that $\zeta_n \to \infty$. With $t = \|\theta\|_4^2\sqrt{C^*\log(\zeta_n)}$, we have

$$\|Z_0\| \le C\|\theta\|_4^2\sqrt{\log(\zeta_n)}, \qquad \text{with probability } 1-\zeta_n.$$

Similarly, we can apply the matrix Hoeffding inequality to Z_1 and Z_2. It gives

$$\|Z_1 + Z_2\| \le C\|\theta\|_4^2\sqrt{\log(\zeta_n)}, \qquad \text{with probability } 1-\zeta_n.$$

Since $\|\theta\|_4^2 \le \theta_{\max}\|\theta\| \ll \|\theta\|^2$, we can choose ζ_n so that $\|\theta\|_4^2\sqrt{\log(\zeta_n)} = o(\|\theta\|^2)$. It follows that, with probability $1-o(1)$,

$$\|Z_0 + Z_1 + Z_2\| = o(\|\theta\|^2).$$

At the same time, $\lambda_{\min}(\|\theta\|^2\Sigma) = \|\theta\|^2\,\|\Sigma^{-1}\|^{-1} \ge C\|\theta\|^2$. Therefore, with probability $1-o(1)$,

$$\lambda_{\min}(\|\theta\|^2 G) \ge \lambda_{\min}(\|\theta\|^2\Sigma) - \|Z_0 + Z_1 + Z_2\| \ge C\|\theta\|^2.$$

This guarantees $\|G^{-1}\| \le C$. □

F.5.3. *Proof of Lemma F.3*. Let $Q = P - 1_K 1_K'$, and introduce $d \in \mathbb{R}^K$ such that $D = \mathrm{diag}(d)$. By Lemma F.1, $\|Q\| \le C|\mu_2|$. With these notations,

$$DPD - 1_K 1_K' = dd' + DQD - 1_K 1_K'. \tag{57}$$

Using the same notations, the assumption $DPD\widetilde{h}_D = 1_K$ can be written as $D(1_K 1_K' + Q)D\widetilde{h}_D = 1_K$. It implies

$$1_K = (d'\widetilde{h}_D)d + DQD\widetilde{h}_D. \tag{58}$$

We multiply $\widetilde{h}_D'$ on both sides and notice that $1_K'\widetilde{h}_D = 1$. It gives

$$(d'\widetilde{h}_D)^2 = 1 - \widetilde{h}_D' DQD\widetilde{h}_D. \tag{59}$$

Combining (58)-(59) gives

$$\begin{aligned}dd' - 1_K 1_K' &= [1-(d'\widetilde{h}_D)^2]dd' - (d'\widetilde{h}_D)(DQD\widetilde{h}_D d + d\widetilde{h}_D DQD) - DQD\widetilde{h}_D\widetilde{h}_D' DQD \\ &= (\widetilde{h}_D' DQD\widetilde{h}_D)\cdot dd' - (d'\widetilde{h}_D)(DQD\widetilde{h}_D d + d\widetilde{h}_D DQD) - DQD\widetilde{h}_D\widetilde{h}_D' DQD.\end{aligned}$$

Since $\|\widetilde{h}_D\| \le C$ and $\|d\| \le C$, we immediately have

$$\|dd' - 1_K 1_K'\| \le C\|Q\| \le C|\mu_2|.$$

Plugging it into (57) gives

$$\|DPD - 1_K 1_K'\| \le C\|Q\| \le C|\mu_2|.$$

□

APPENDIX G: PROPERTIES OF SIGNED POLYGON STATISTICS

We prove Tables A.1-2 and Theorem A.1-4.3. The analysis of T_n and Q_n is very similar. To save space, we only present the proof for results of Q_n. The proof for results of T_n (Tables A.1, A.2, and Theorems A.1, A.2, A.3) is omitted.

We recall the following notations:

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

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

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

Then, the Ideal SgnQ statistic equals to

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

the Proxy SgnQ statistic equals to

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

and the SgnQ statistic equals to

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

As explained in Section 4, each of $\widetilde{Q}_n, Q_n^*, Q_n$ is the sum of a finite number of post-expansion sums, each having the form

$$\sum_{i,j,k,\ell(dist)} a_{ij}b_{jk}c_{k\ell}d_{\ell i}, \tag{60}$$

where a_{ij} equals to one of $\{\widetilde{\Omega}_{ij}, W_{ij}, \delta_{ij}, r_{ij}\}$; same for b_{ij}, c_{ij} and d_{ij}. Let $N_{\widetilde{\Omega}}$ be the (common) number of $\widetilde{\Omega}$ terms in each product; similarly, we define N_W, N_δ, N_r. These numbers satisfy $N_{\widetilde{\Omega}} + N_W + N_\delta + N_r = 4$. For example, for the post-expansion sum $\sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}W_{jk}W_{k\ell}W_{\ell i}$, $(N_{\widetilde{\Omega}}, N_W, N_\delta, N_r) = (1,3,0,0)$. In Section G.1, we study $\widetilde{Q}_n$, and it involves these post-expansion sums such that

$$N_\delta = N_r = 0,$$

In Section G.2, we study $(Q_n^* - \widetilde{Q}_n)$, which involves post-expansion sums such that

$$N_\delta > 0, \text{ and } N_r = 0,$$

In Section G.3, we study $(Q_n - Q_n^*)$, which is related to the sums such that

$$N_r > 0.$$

G.1. Analysis of Table 1, proof of Theorem 4.1. Define

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

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

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

We first consider the null hypothesis. Since $\widetilde{\Omega}$ is a zero matrix, it is not hard to see that

$$\widetilde{Q}_n = X_1.$$

The following lemmas are proved in Section G.4.

LEMMA G.1. *Suppose the conditions of Theorem 4.1 hold. Under the null hypothesis, as $n \to \infty$, $\mathbb{E}[\widetilde{Q}_n] = 0$ and $\mathrm{Var}(\widetilde{Q}_n) = 8\|\theta\|^8 \cdot [1 + o(1)]$.*

LEMMA G.2. *Suppose the conditions of Theorem 4.1 hold. Under the null hypothesis, as $n \to \infty$,*

$$\frac{\widetilde{Q}_n - E[\widetilde{Q}_n]}{\sqrt{\mathrm{Var}(\widetilde{Q}_n)}} \longrightarrow N(0,1), \qquad \textit{in law}.$$

We then consider the alternative hypothesis. By elementary algebra,

$$\widetilde{Q}_n = X_1 + 4X_2 + 4X_3 + 2X_4 + 4X_5 + X_6.$$

The following lemma characterizes the asymptotic mean and variance of X_1-X_6 under the alternative hypothesis. It gives rise to Columns 5-6 of Table 1.

LEMMA G.3 (Table 1). *Suppose conditions of Theorem 4.1 hold. Write $\alpha = |\lambda_2|/\lambda_1$. Under the alternative hypothesis, as $n \to \infty$,*

- $\mathbb{E}[X_k] = 0$ *for* $1 \leq k \leq 5$, *and* $\mathbb{E}[X_6] = \mathrm{tr}(\widetilde{\Omega}^4) \cdot [1 + o(1)]$.
- $C^{-1}\|\theta\|^8 \leq \mathrm{Var}(X_1) \leq C\|\theta\|^8$.
- $\mathrm{Var}(X_2) \leq C\alpha^2\|\theta\|^4\|\theta\|_3^6 = o(\|\theta\|^8)$.
- $\mathrm{Var}(X_3) \leq C\alpha^4\|\theta\|^6\|\theta\|_3^6 = o(\alpha^6\|\theta\|^8\|\theta\|_3^6)$.
- $\mathrm{Var}(X_4) \leq C\alpha^4\|\theta\|_3^{12} = o(\|\theta\|^8)$.
- $\mathrm{Var}(X_5) \leq C\alpha^6\|\theta\|^8\|\theta\|_3^6$.

As a result, $\mathbb{E}[\widetilde{Q}_n] \sim \mathrm{tr}(\widetilde{\Omega}^4)$ *and* $\mathrm{Var}(\widetilde{Q}_n) \leq C(\|\theta\|^8 + \alpha^6\|\theta\|^8\|\theta\|_3^6)$.

Theorem 4.1 follows directly from Lemmas G.1-G.3.

G.2. Analysis of Table 2, proof of Theorem 4.2. We introduce U_a, U_b and U_c such that

$$Q_n^* - \widetilde{Q}_n = U_a + U_b + U_c,$$

where U_a, U_b and U_c contain post-expansion sums (60) with $N_\delta = 1$, $N_\delta = 2$, and $N_\delta \geq 3$, respectively.

First, we consider the post-expansion sums with $N_\delta = 1$. Define

$$U_a = 4Y_1 + 8Y_2 + 4Y_3 + 8Y_4 + 4Y_5 + 4Y_6, \tag{61}$$

where

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

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

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

Under the null hypothesis, only Y_1 is nonzero, and

$$U_a = 4Y_1.$$

LEMMA G.4. *Suppose the conditions of Theorem 4.1 hold. Under the null hypothesis, as* $n \to \infty$, $\mathbb{E}[U_a] = 0$ *and* $\mathrm{Var}(U_a) \le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8)$.

Under the alternative hypothesis, the following lemma characterizes the asymptotic means and variances of Y_1-Y_6. It gives rise to Rows 1-6 of Table 2 and is proved in Section G.4.

LEMMA G.5 (Table 2, Rows 1-6). *Suppose the conditions of Theorem 4.1 hold. Let* $\alpha = |\lambda_2|/\lambda_1$. *Under the alternative hypothesis, as* $n \to \infty$,

- $\mathbb{E}[Y_k] = 0$ *for* $k \in \{1,2,3,5,6\}$, *and* $|\mathbb{E}[Y_4]| \le C\alpha^2\|\theta\|^6 = o(\alpha^4\|\theta\|^8)$.
- $\mathrm{Var}(Y_1) \le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8)$.
- $\mathrm{Var}(Y_2) \le C\alpha^2\|\theta\|^4\|\theta\|_3^6 = o(\|\theta\|^8)$.
- $\mathrm{Var}(Y_3) \le C\alpha^2\|\theta\|^4\|\theta\|_3^6 = o(\|\theta\|^8)$.
- $\mathrm{Var}(Y_4) \le \frac{C\alpha^4\|\theta\|^{10}\|\theta\|_3^3}{\|\theta\|_1} = o(\alpha^6\|\theta\|^8\|\theta\|_3^6)$.
- $\mathrm{Var}(Y_5) \le \frac{C\alpha^4\|\theta\|^4\|\theta\|_3^9}{\|\theta\|_1} = o(\|\theta\|^8)$.
- $\mathrm{Var}(Y_6) \le \frac{C\alpha^6\|\theta\|^{12}\|\theta\|_3^3}{\|\theta\|_1} = O(\alpha^6\|\theta\|^8\|\theta\|_3^6)$.

As a result, $\mathbb{E}[U_a] = o(\alpha^4\|\theta\|^8)$ *and* $\mathrm{Var}(U_a) \le C\alpha^6\|\theta\|^8\|\theta\|_3^6 + o(\|\theta\|^8)$.

Next, we consider the post-expansion sums with $N_\delta = 2$. Define

$$U_b = 4Z_1 + 2Z_2 + 8Z_3 + 4Z_4 + 4Z_5 + 2Z_6, \tag{62}$$

where

$$Z_1 = \sum_{i,j,k,\ell (dist)} \delta_{ij}\delta_{jk}W_{k\ell}W_{\ell i}, \qquad Z_2 = \sum_{i,j,k,\ell (dist)} \delta_{ij}W_{jk}\delta_{k\ell}W_{\ell i},$$

$$Z_3 = \sum_{i,j,k,\ell (dist)} \delta_{ij}\delta_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i}, \qquad Z_4 = \sum_{i,j,k,\ell (dist)} \delta_{ij}\widetilde{\Omega}_{jk}\delta_{k\ell}W_{\ell i},$$

$$Z_5 = \sum_{i,j,k,\ell (dist)} \delta_{ij}\delta_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}, \qquad Z_6 = \sum_{i,j,k,\ell (dist)} \delta_{ij}\widetilde{\Omega}_{jk}\delta_{k\ell}\widetilde{\Omega}_{\ell i}.$$

Under the null hypothesis, only Z_1 and Z_2 are nonzero, and

$$U_b = 4Z_1 + 2Z_2.$$

LEMMA G.6. *Suppose the conditions of Theorem 4.1 hold. Under the null hypothesis, as* $n \to \infty$,

- $\mathbb{E}[Z_1] = \|\theta\|^4 \cdot [1 + o(1)]$, *and* $\mathrm{Var}(Z_1) \le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8)$.
- $\mathbb{E}[Z_2] = 2\|\theta\|^4 \cdot [1 + o(1)]$, *and* $\mathrm{Var}(Z_2) \le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$.

As a result, $\mathbb{E}[U_b] \sim 8\|\theta\|^4$ *and* $\mathrm{Var}(U_b) = o(\|\theta\|^8)$.

Under the alternative hypothesis, the following lemma provides the asymptotic means and variances of Z_1-Z_6. It gives rise to Rows 7-12 of Table 2:

LEMMA G.7 (Table 2, Rows 7-12). *Suppose conditions of Theorem 4.1 hold. Write* $\alpha = |\lambda_2|/\lambda_1$. *Under the alternative hypothesis, as* $n \to \infty$,

- $|\mathbb{E}[Z_1]| \le C\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$, *and* $\mathrm{Var}(Z_1) \le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8)$.
- $|\mathbb{E}[Z_2]| \le C\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$, *and* $\mathrm{Var}(Z_2) \le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$.
- $\mathbb{E}Z_3 = 0$, *and* $\mathrm{Var}(Z_3) \le C\alpha^2\|\theta\|^4\|\theta\|_3^6 = o(\|\theta\|^8)$.
- $|\mathbb{E}[Z_4]| \le C\alpha\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$, *and* $\mathrm{Var}(Z_4) \le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$.
- $|\mathbb{E}[Z_5]| \le C\alpha^2\|\theta\|^6 = o(\alpha^4\|\theta\|^8)$, *and* $\mathrm{Var}(Z_5) \le \frac{C\alpha^4\|\theta\|^{14}}{\|\theta\|_1^2} = o(\alpha^6\|\theta\|^8\|\theta\|_3^6)$.
- $|\mathbb{E}[Z_6]| \le \frac{C\alpha^2\|\theta\|^8}{\|\theta\|_1^2} = o(\alpha^4\|\theta\|^8)$, *and* $\mathrm{Var}(Z_6) \le \frac{C\alpha^4\|\theta\|^8\|\theta\|_3^6}{\|\theta\|_1^2} = o(\|\theta\|^8)$.

As a result, $\mathbb{E}[U_b] = o(\alpha^4\|\theta\|^8)$ *and* $\mathrm{Var}(U_b) = o(\|\theta\|^8 + \alpha^6\|\theta\|^8\|\theta\|_3^6)$.

Last, we consider the post-expansion sums with $N_\delta \ge 3$. Define

$$U_c = 4T_1 + 4T_2 + F, \tag{63}$$

where

$$T_1 = \sum_{i,j,k,\ell (dist)} \delta_{ij}\delta_{jk}\delta_{k\ell}W_{\ell i}, \qquad T_2 = \sum_{i,j,k,\ell (dist)} \delta_{ij}\delta_{jk}\delta_{k\ell}\widetilde{\Omega}_{\ell i},$$

$$F = \sum_{i,j,k,\ell (dist)} \delta_{ij}\delta_{jk}\delta_{k\ell}\delta_{\ell i}.$$

Under the null hypothesis, only T_1 and F are nonzero, and

$$U_b = 4T_1 + F.$$

LEMMA G.8. *Suppose the conditions of Theorem 4.1 hold. Under the null hypothesis, as* $n \to \infty$,

- $\mathbb{E}[T_1] = -2\|\theta\|^4 \cdot [1 + o(1)]$, *and* $\mathrm{Var}(T_1) \le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$.
- $|\mathbb{E}[F]| = 2\|\theta\|^4 \cdot [1 + o(1)]$, *and* $\mathrm{Var}(F) \le \frac{C\|\theta\|^{10}}{\|\theta\|_1^2} = o(\|\theta\|^8)$.

As a result, $\mathbb{E}[U_c] \sim -6\|\theta\|^4$ *and* $\mathrm{Var}(U_c) = o(\|\theta\|^8)$.

Under the alternative hypothesis, the next lemma studies the asymptotic means and variances of T_1, T_2 and F. It gives rise to Rows 13-15 of Table 2:

LEMMA G.9 (Table 2, Rows 13-15). *Suppose conditions of Theorem 4.1 hold. Write* $\alpha = |\lambda_2|/\lambda_1$. *Under the alternative hypothesis, as* $n \to \infty$,

- $|\mathbb{E}[T_1]| \le C\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$, *and* $\mathrm{Var}(T_1) \le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$.
- $|\mathbb{E}[T_2]| \le \frac{C\alpha\|\theta\|^6}{\|\theta\|_1^3} = o(\alpha^4\|\theta\|^8)$, *and* $\mathrm{Var}(T_2) \le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8)$.
- $|\mathbb{E}[F]| \le C\|\theta\|^4 = o(\alpha^4\|\theta\|^8)$, *and* $\mathrm{Var}(F) \le \frac{C\|\theta\|^{10}}{\|\theta\|_1^2} = o(\|\theta\|^8)$.

As a result, $\mathbb{E}|U_c| = o(\alpha^4\|\theta\|^8)$ *and* $\mathrm{Var}(U_c) = o(\|\theta\|^8)$.

We now prove Theorem 4.2. Since $Q_n^* - \widetilde{Q}_n = U_a + U_b + U_c$, we have

$$\mathbb{E}[Q_n^* - \widetilde{Q}_n] = \mathbb{E}[U_a] + \mathbb{E}[U_b] + \mathbb{E}[U_c],$$
$$\mathrm{Var}(Q_n^* - \widetilde{Q}_n) \le 3\mathrm{Var}(U_a) + 3\mathrm{Var}(U_b) + 3\mathrm{Var}(U_c).$$

Consider the null hypothesis. By Lemmas G.4, G.6, G.8,

$$\mathbb{E}[Q_n^* - \widetilde{Q}_n] = 0 + 8\|\theta\|^4 - 6\|\theta\|^4 + o(\|\theta\|^4) \sim 2\|\theta\|^4,$$

and

$$\mathrm{Var}(Q_n^* - \widetilde{Q}_n) \le C\|\theta\|^2\|\theta\|_3^6 + \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} + \frac{C\|\theta\|^{10}}{\|\theta\|_1^2}.$$

Using the universal inequality $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$, we further have

$$\mathrm{Var}(Q_n^* - \widetilde{Q}_n) \le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8),$$

where $\|\theta\|_3^3 = o(\|\theta\|^2)$ and $\|\theta\| \to \infty$ in our range of interest. This proves claims for the null hypothesis. Consider the alternative hypothesis. By Lemmas G.5, G.7, G.9,

$$\left|\mathbb{E}[Q_n^* - \widetilde{Q}_n]\right| \le C\alpha^2\|\theta\|^6,$$

where the main contributors are Y_4 and Z_5. Since $\alpha\|\theta\| \to \infty$ in our range of interest, the above is $o(\alpha^4\|\theta\|^8)$. By Lemmas G.5, G.7, G.9,

$$\mathrm{Var}(Q_n^* - \widetilde{Q}_n) \le \frac{C\alpha^6\|\theta\|^{12}\|\theta\|_3^3}{\|\theta\|_1},$$

where the main contributor is Y_6. Using the universal inequality of $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$, the above is $O(\alpha^6\|\theta\|^8\|\theta\|_3^6)$. This proves claims for the alternative hypothesis.

G.3. Analysis of $(Q_n - Q_n^*)$, proof of Theorem 4.3. By definition, $(Q_n - Q_n^*)$ expands to the sum of 175 post-expansion sums, where each has the form (60) and satisfies $N_r > 0$. Recall that

$$r_{ij} = (\eta_i^*\eta_j^* - \eta_i\eta_j) - (\eta_i - \tilde{\eta}_i)(\eta_j - \tilde{\eta}_j) + (1 - \frac{v}{V})\tilde{\eta}_i\tilde{\eta}_j.$$

Since $\delta_{ij} = \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_i - \tilde{\eta}_i)$, we have $\tilde{\eta}_i\tilde{\eta}_j = \eta_i\eta_j - \delta_{ij} + (\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)$. Inserting it into the definition of r_{ij} gives

$$r_{ij} = (\eta_i^*\eta_j^* - \eta_i\eta_j) + (1 - \frac{v}{V})\eta_i\eta_j - (1 - \frac{v}{V})\delta_{ij} - \frac{v}{V}(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j). \tag{64}$$

Define

$$\tilde{r}_{ij} = -\frac{v}{V}(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j), \qquad \epsilon_{ij} = (\eta_i^*\eta_j^* - \eta_i\eta_j) + (1 - \frac{v}{V})\eta_i\eta_j - (1 - \frac{v}{V})\delta_{ij}.$$

Then, we can write

$$r_{ij} = \tilde{r}_{ij} + \epsilon_{ij}. \tag{65}$$

Using this notation, we re-write

$$Q_n = \sum_{i,j,k,\ell (dist)} M_{ij}M_{jk}M_{k\ell}M_{\ell i}, \qquad \text{where } M_{ij} = \widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij} + \tilde{r}_{ij} + \epsilon_{ij},$$

and

$$Q_n^* = \sum_{i,j,k,\ell (dist)} M_{ij}^*M_{jk}^*M_{k\ell}^*M_{\ell i}^*, \qquad \text{where } M_{ij}^* \equiv \widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij}.$$

We then introduce an intermediate variable:

$$\widetilde{Q}_n^* = \sum_{i,j,k,\ell(dist)} \widetilde{M}_{ij}^* \widetilde{M}_{jk}^* \widetilde{M}_{k\ell}^* \widetilde{M}_{\ell i}^*, \quad \text{where } \widetilde{M}_{ij}^* = \widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij} + \tilde{r}_{ij}. \tag{66}$$

As a result, $(Q_n - Q_n^*)$ decomposes into

$$Q_n - Q_n^* = (\widetilde{Q}_n^* - Q_n^*) + (Q_n - \widetilde{Q}_n^*). \tag{67}$$

We note that Q_n can be expanded to the sum of $5^4 = 625$ post-expansion sums, each with the form

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

where each of $a_{ij}, b_{ij}, c_{ij}, d_{ij}$ takes values in $\{\widetilde{\Omega}_{ij}, W_{ij}, \delta_{ij}, \tilde{r}_{ij}, \epsilon_{ij}\}$. Let $N_{\widetilde{\Omega}}$ be the (common) number of $\widetilde{\Omega}$ terms in each product and define $N_W, N_\delta, N_{\tilde{r}}, N_\epsilon$ similarly. Among the 625 post-expansion sums,

- $3^4 = 81$ of them are contained in Q_n^*,
- $4^4 - 3^4 = 175$ of them are contained in $(\widetilde{Q}_n^* - Q_n^*)$,
- and $5^4 - 4^4 = 369$ of them are contained in $(Q_n - \widetilde{Q}_n^*)$.

We shall study $(\widetilde{Q}_n^* - Q_n^*)$ and $(Q_n - \widetilde{Q}_n^*)$, separately.

In our analysis, one challenge is to deal with the random variable V that appears in the denominator in the expression of r_{ij}. The following lemma is useful and proved in Section G.4.

LEMMA G.10. *Suppose conditions of Theorem 4.3 hold. As $n \to \infty$, for any sequence x_n such that $\sqrt{\log(\|\theta\|_1)} \ll x_n \ll \|\theta\|_1$,*

$$\mathbb{E}\big[(\widetilde{Q}_n - Q_n)^2 \cdot I\{|V - v| > \|\theta\|_1 x_n\}\big] \to 0.$$

The next two lemmas are proved in Section G.4.

LEMMA G.11. *Suppose conditions of Theorem 4.3 hold. Write $\alpha = |\lambda_2|/\lambda_1$. As $n \to \infty$,*

- *Under the null hypothesis, $|\mathbb{E}[\widetilde{Q}_n^* - Q_n^*]| = o(\|\theta\|^4)$ and $\mathrm{Var}(\widetilde{Q}_n^* - Q_n^*) = o(\|\theta\|^8)$.*
- *Under the alternative hypothesis, $|\mathbb{E}[\widetilde{Q}_n^* - Q_n^*]| = o(\alpha^4\|\theta\|^8)$ and $\mathrm{Var}(\widetilde{Q}_n^* - Q_n^*) = o(\|\theta\|^8 + \alpha^6\|\theta\|^8\|\theta\|_3^6)$.*

LEMMA G.12. *Suppose conditions of Theorem 4.3 hold. Write $\alpha = |\lambda_2|/\lambda_1$. As $n \to \infty$,*

- *Under the null hypothesis, $|\mathbb{E}[Q_n - \widetilde{Q}_n^*]| = o(\|\theta\|^4)$ and $\mathrm{Var}(Q_n - \widetilde{Q}_n^*) = o(\|\theta\|^8)$.*
- *Under the alternative hypothesis, $|\mathbb{E}[Q_n - \widetilde{Q}_n^*]| = o(\alpha^4\|\theta\|^8)$ and $\mathrm{Var}(\widetilde{Q}_n^* - Q_n^*) = O(\|\theta\|^8 + \alpha^6\|\theta\|^8\|\theta\|_3^6)$.*

Theorem 4.3 follows directly from (67) and Lemmas G.11-G.12.

G.4. Proof of Lemmas G.1-G.12.

G.4.1. *Proof of Lemma G.1.* Under the null hypothesis,

$$\widetilde{Q}_n = X_1 = \sum_{i,j,k,\ell (dist)} W_{ij} W_{jk} W_{k\ell} W_{\ell i}.$$

For mutually distinct indices (i,j,k,ℓ), $(W_{ij}, W_{jk}, W_{k\ell}, W_{\ell i})$ are independent of each other, each with mean zero. So $\mathbb{E}[W_{ij} W_{jk} W_{k\ell} W_{\ell i}] = 0$. It follows that

$$\mathbb{E}[\widetilde{Q}_n] = 0.$$

We now calculate the variance of $\widetilde{Q}_n$. Under the null hypothesis, $\Omega_{ij} = \theta_i \theta_j$; hence, $\mathrm{Var}(W_{ij}) = \Omega_{ij}(1-\Omega_{ij}) = \theta_i\theta_j - \theta_i^2\theta_j^2 = \theta_i\theta_j[1 + O(\theta_{\max}^2)]$. It follows that

$$\begin{aligned} \mathrm{Var}(W_{ij} W_{jk} W_{k\ell} W_{\ell i}) &= \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2 \cdot [1 + O(\theta_{\max}^2)]^4 \\ &= \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2 \cdot [1 + O(\theta_{\max}^2)]. \end{aligned} \tag{68}$$

Note that each (i,j,k,ℓ) corresponds to a 4-cycle in a complete graph of n nodes. For (i,j,k,ℓ) and (i',j',k',ℓ'), we can write $W_{ij} W_{jk} W_{k\ell} W_{\ell i} \cdot W_{i'j'} W_{j'k'} W_{k'\ell'} W_{\ell' i'}$ in the form of $\prod_t (W_{i_t j_t})^{m_t}$, where $\{W_{i_t j_t}\}$ are mutually distinct with each other and m_t is the number of times that $W_{i_t j_t}$ appears in this product. If the two 4-cycles corresponding to (i,j,k,ℓ) and (i',j',k',ℓ') are not exactly overlapping, then at least two of m_t equals to 1. As a result, the mean of $\prod_t (W_{i_t j_t})^{m_t}$ is zero. In other words, we have argued that

(69) $\mathrm{Cov}(W_{ij} W_{jk} W_{k\ell} W_{\ell i},\ W_{i'j'} W_{j'k'} W_{k'\ell'} W_{\ell' i'}) = 0$ if the two cycles corresponding to (i,j,k,ℓ) and (i',j',k',ℓ') are not exactly overlapping.

In the sum over all distinct (i,j,k,ℓ), each 4-cycle is repeatedly counted by 8 times

$$\begin{aligned} &(i,j,k,\ell), (j,k,\ell,i), (k,\ell,i,j), (\ell,i,j,k), \\ &(\ell,k,j,i), (k,j,i,\ell), (j,i,\ell,k), (i,\ell,k,j). \end{aligned}$$

It follows that

$$\begin{aligned} \mathrm{Var}(\widetilde{Q}_n) &= \mathrm{Var}\Bigg(8 \sum_{\substack{\text{unique} \\ \text{4-cycles}}} W_{ij} W_{jk} W_{k\ell} W_{\ell i} \Bigg) \\ &= 64 \cdot \mathrm{Var}\Bigg(\sum_{\substack{\text{unique} \\ \text{4-cycles}}} W_{ij} W_{jk} W_{k\ell} W_{\ell i} \Bigg) \\ &= 64 \sum_{\substack{\text{unique} \\ \text{4-cycles}}} \mathrm{Var}\big(W_{ij} W_{jk} W_{k\ell} W_{\ell i}\big) \\ &= 8 \sum_{i,j,k,\ell (dist)} \mathrm{Var}\big(W_{ij} W_{jk} W_{k\ell} W_{\ell i}\big) \\ &= [1 + O(\theta_{\max}^2)] \cdot 8 \sum_{i,j,k,\ell (dist)} \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2, \end{aligned} \tag{70}$$

where the third line is from (69) and the last line is from (68). We then compute the right hand side of (70). Note that

$$\sum_{i,j,k,\ell (dist)} \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2 = \|\theta\|^8 - \sum_{i,j,k,\ell (not\ dist)} \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2,$$

where

$$\sum_{i,j,k,\ell (not\ dist)} \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2 \le \binom{4}{2}\sum_{i,j,k}\theta_i^2\theta_j^2\theta_k^4 \le C\|\theta\|^4\|\theta\|_4^4 = \|\theta\|^8\cdot O\Big(\frac{\|\theta\|_4^4}{\|\theta\|^4}\Big).$$

Combining the above gives

$$\sum_{i,j,k,\ell(dist)} \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2 = \|\theta\|^8\cdot\Big[1+O\Big(\frac{\|\theta\|_4^4}{\|\theta\|^4}\Big)\Big]. \tag{71}$$

We combine (70)-(71) and note that $\theta_{\max}=o(1)$ and $\|\theta\|_4^4/\|\theta\|^4 \le (\|\theta\|^2\theta_{\max}^2)/\|\theta\|^4 = o(1)$. So,

$$\mathrm{Var}(\widetilde{Q}_n) = 8\|\theta\|^8\cdot[1+o(1)].$$

This completes the proof.

G.4.2. *Proof of Lemma G.2.* Under the null hypothesis,

$$\widetilde{Q}_n = X_1 = \sum_{i,j,k,\ell(dist)} W_{ij}W_{jk}W_{k\ell}W_{\ell i}.$$

In the proof of Theorem 3.2 of [3], it has been shown that $X_1/\sqrt{\mathrm{Var}(X_1)}\to N(0,1)$ in law (in the proof there, $X_1/\sqrt{\mathrm{Var}(X_1)}$ is denoted as $S_{n,n}$). Since $\mathbb{E}[X_1]=0$, we can directly quote their results to get the desired claim.

G.4.3. *Proof of Lemma G.3.* We shall study the mean and variance of each of X_1-X_6 and then combine those results.

Consider X_1. We have analyzed this term under the null hypothesis. Under the alternative hypothesis, the difference is that we no longer have $\Omega_{ij}=\theta_i\theta_j$. Instead, we have an upper bound $\Omega_{ij}=\theta_i\theta_j(\pi_i' P\pi_j)\le C\theta_i\theta_j$. Using similar proof as that for the null hypothesis, we can derive that

$$\mathbb{E}[X_1]=0, \qquad \mathrm{Var}(X_1)\le C\|\theta\|^8. \tag{72}$$

To get a lower bound for $\mathrm{Var}(X_1)$, we notice that $\mathrm{Var}(W_{ij})=\Omega_{ij}(1-\Omega_{ij})\ge\Omega_{ij}[1-O(\theta_{\max}^2)]\ge\Omega_{ij}/2$; this inequality is true even when $\Omega_{ij}=0$. It follows that

$$\mathrm{Var}(W_{ij}W_{jk}W_{k\ell}W_{\ell i}) \ge \frac{1}{16}\Omega_{ij}\Omega_{jk}\Omega_{k\ell}\Omega_{\ell i}.$$

Note that the second last line of (70) is still true. As a result,

$$\begin{aligned}
\mathrm{Var}(X_1) &= 8\sum_{i,j,k,\ell(dist)}\mathrm{Var}\big(W_{ij}W_{jk}W_{k\ell}W_{\ell i}\big)\\
&\ge \frac{1}{2}\sum_{i,j,k,\ell(dist)}\Omega_{ij}\Omega_{jk}\Omega_{k\ell}\Omega_{\ell i}\\
&= \frac{1}{2}\mathrm{tr}(\Omega^4) - \frac{1}{2}\sum_{i,j,k,\ell(not\ dist)}\Omega_{ij}\Omega_{jk}\Omega_{k\ell}\Omega_{\ell i}\\
&\ge \frac{1}{2}\mathrm{tr}(\Omega^4) - C\sum_{i,j,k,\ell(not\ dist)}\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\\
&\ge \frac{1}{2}\mathrm{tr}(\Omega^4) - o(\|\theta\|^8),
\end{aligned}$$

where the last inequality is due to (71). Recall that $\lambda_1, \ldots, \lambda_K$ denote the K nonzero eigenvalues of Ω. By Lemma E.2, $\lambda_1 \geq C^{-1}\|\theta\|^2$. It follows that

$$\mathrm{tr}(\Omega^4) = \sum_{k=1}^{K} \lambda_k^4 \geq \lambda_1^4 \geq C^{-1}\|\theta\|^8.$$

Combining the above gives

$$\mathrm{Var}(X_1) \geq C^{-1}\|\theta\|^8. \tag{73}$$

So far, we have proved all claims about X_1.

Consider X_2. Recall that

$$X_2 = \sum_{i,j,k,\ell (dist)} \widetilde{\Omega}_{ij} W_{jk} W_{k\ell} W_{\ell i}.$$

It is easy to see that $\mathbb{E}[X_2] = 0$. Below, we bound its variance. Each index choice (i, j, k, ℓ) defines a undirected path j-k-ℓ-i in the complete graph of n nodes. If the two paths j-k-ℓ-i and j'-k'-ℓ'-i' are not exactly overlapping, then $W_{jk}W_{k\ell}W_{\ell i} \cdot W_{j'k'}W_{k'\ell'}W_{\ell' i'}$ have mean zero. In the sum above, each unique path j-k-ℓ-i is counted twice as (i, j, k, ℓ) and (j, i, ℓ, k). Mimicking the argument in (70), we immediately have

$$\begin{aligned}
\mathrm{Var}(X_2) &= 2 \sum_{i,j,k,\ell (dist)} \mathrm{Var}\big(\widetilde{\Omega}_{ij} W_{jk} W_{k\ell} W_{\ell i}\big) \\
&= 2 \sum_{i,j,k,\ell (dist)} \widetilde{\Omega}_{ij}^2 \cdot \mathrm{Var}\big(W_{jk} W_{k\ell} W_{\ell i}\big).
\end{aligned}$$

By Lemma E.5, $|\widetilde{\Omega}_{ij}| \leq |\lambda_2|\|\theta\|^{-2}\theta_i\theta_j$. In our notations, $\alpha = |\lambda_2|/\lambda_1$; additionally, by Lemma E.2, $\lambda_1 \leq C\|\theta\|^2$. Combining them gives

$$|\widetilde{\Omega}_{ij}| \leq C\alpha\theta_i\theta_j. \tag{74}$$

Moreover, $\mathrm{Var}(W_{jk}W_{k\ell}W_{\ell i}) \leq \Omega_{jk}\Omega_{k\ell}\Omega_{\ell i} \leq C\theta_j\theta_k^2\theta_\ell^2\theta_i$. It follows that

$$\begin{aligned}
\mathrm{Var}(X_2) &\leq C \sum_{i,j,k,\ell (dist)} (\alpha\theta_i\theta_j)^2 \cdot \theta_j\theta_k^2\theta_\ell^2\theta_i \\
&\leq C\alpha^2 \sum_{i,j,k,\ell} \theta_i^3\theta_j^3\theta_k^2\theta_\ell^2 \\
&\leq C\alpha^2\|\theta\|^4\|\theta\|_3^6.
\end{aligned}$$

Since $\|\theta\|_3^3 \leq \theta_{\max}\sum_i \theta_i^2 = \theta_{\max}\|\theta\|^2$, the right hand side is $\leq C\alpha^2\|\theta\|^8\theta_{\max}^2$. Note that $\alpha \leq 1$ and $\theta_{\max} \to 0$. So, this term is $o(\|\theta\|^8)$. We have proved all claims about X_2.

Consider X_3. Recall that

$$X_3 = \sum_{i,j,k,\ell (dist)} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk} W_{k\ell} W_{\ell i} = \sum_{i,k,\ell (dist)} \Big(\sum_{j \notin \{i,k,\ell\}} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk} \Big) W_{k\ell} W_{\ell i}.$$

It is easy to see that $\mathbb{E}[X_3] = 0$. We then study its variance. We note that for $W_{k\ell}W_{\ell i}$ and $W_{k'\ell'}W_{\ell' i'}$ to be correlated, we must have that $(k', \ell', i') = (k, \ell, i)$ or $(k', \ell', i') = (i, \ell, k)$; in other words, the two underlying paths k-ℓ-i and k'-ℓ'-i' have to be equal. Mimicking the argument in (70), we have

$$\mathrm{Var}(X_3) \le C \sum_{i,k,\ell (dist)} \mathrm{Var}\Big[\Big(\sum_{j\notin\{i,k,\ell\}} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\Big) W_{k\ell}W_{\ell i}\Big]$$
$$\le C \sum_{i,k,\ell (dist)} \Big(\sum_{j\notin\{i,k,\ell\}} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\Big)^2 \cdot \mathrm{Var}(W_{k\ell}W_{\ell i}).$$

By (74),

$$\Big|\sum_{j\notin\{i,k,\ell\}} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\Big| \le C\sum_j \alpha^2\theta_i\theta_j^2\theta_k \le C\alpha^2\|\theta\|^2\cdot\theta_i\theta_k.$$

Combining the above gives

$$\mathrm{Var}(X_3) \le C\sum_{i,k,\ell}(\alpha^2\|\theta\|^2\theta_i\theta_k)^2\cdot\theta_k\theta_\ell^2\theta_i$$
$$\le C\alpha^4\|\theta\|^4\sum_{i,k,\ell}\theta_i^3\theta_k^3\theta_\ell^2$$
$$\le C\alpha^4\|\theta\|^6\|\theta\|_3^6.$$

Since $\|\theta\|\to\infty$, the right hand side is $o(\alpha^4\|\theta\|^8\|\theta\|_3^6)$. We have proved all claims about X_3.

Consider X_4. Recall that

$$X_4 = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}W_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i} = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{k\ell}W_{jk}W_{\ell i}.$$

It is easy to see that $\mathbb{E}[X_4]=0$. To calculate its variance, note that $W_{jk}W_{\ell i}$ and $W_{j'k'}W_{\ell' i'}$ are uncorrelated unless (i) $\{j',k'\}=\{j,k\}$ and $\{\ell',i'\}=\{\ell,i\}$ or (ii) $\{j',k'\}=\{\ell,i\}$ and $\{\ell',i'\}=\{j,k\}$. Mimicking the argument in (70), we immediately have

$$\mathrm{Var}(X_4) \le C\sum_{i,j,k,\ell(dist)} \mathrm{Var}\big(\widetilde{\Omega}_{ij}\widetilde{\Omega}_{k\ell}W_{jk}W_{\ell i}\big)$$
$$\le C\sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}^2\widetilde{\Omega}_{k\ell}^2\cdot\mathrm{Var}(W_{jk}W_{\ell i})$$
$$\le C\sum_{i,j,k,\ell}(\alpha\theta_i\theta_j)^2(\alpha\theta_k\theta_\ell)^2\cdot\theta_j\theta_k\theta_\ell\theta_i$$
$$\le C\alpha^4\sum_{i,j,k,\ell}\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3$$
$$\le C\alpha^4\|\theta\|_3^{12}.$$

Since $\|\theta\|_3^3\le\theta_{\max}\|\theta\|^2=o(\|\theta\|^2)$, the right hand side is $o(\|\theta\|^8)$. This proves the claims of X_4.

Consider X_5. Recall that

$$X_5 = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i} = 2\sum_{i<\ell}\Big(\sum_{\substack{j,k\notin\{i,\ell\}\\ j\neq k}} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\Big)W_{\ell i}.$$

It is easily seen that $\mathbb{E}[X_5]=0$. Furthermore, we have

$$\mathrm{Var}(X_5) = 2\sum_{i<\ell}\Big(\sum_{\substack{j,k\notin\{i,\ell\}\\ j\neq k}} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\Big)^2\cdot\mathrm{Var}(W_{\ell i}). \tag{75}$$

By (74),

$$\Big|\sum_{\substack{j,k\notin\{i,\ell\}\\ j\neq k}} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\Big| \le C\sum_{j,k}\alpha^3\theta_i\theta_j^2\theta_k^2\theta_\ell \le C\alpha^3\|\theta\|^4\cdot\theta_i\theta_\ell$$

We plug it into (75) and use $\mathrm{Var}(W_{\ell i}) \le \Omega_{\ell i} \le C\theta_\ell\theta_i$. It yields that

$$\begin{aligned}\mathrm{Var}(X_5) &\le C\sum_{\ell,i(dist)}(\alpha^3\|\theta\|^4\theta_i\theta_\ell)^2\cdot\theta_\ell\theta_i\\ &\le C\alpha^6\|\theta\|^8\sum_{\ell,i}\theta_i^3\theta_\ell^3\\ &\le C\alpha^6\|\theta\|^8\|\theta\|_3^6. \end{aligned} \tag{76}$$

This proves the claims of X_5.

Consider X_6. Recall that

$$X_6 = \sum_{i,j,k,\ell(dist)}\widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} = \mathrm{tr}(\widetilde{\Omega}^4) - \sum_{i,j,k,\ell(not\,dist)}\widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}.$$

This is a non-stochastic number, so its variance is zero and its mean is X_6 itself. By Lemma E.5, $|\lambda_2| \le \|\widetilde{\Omega}\| \le C|\lambda_2|$. Since $\|\widetilde{\Omega}\|^4 \le \mathrm{tr}(\widetilde{\Omega}^4) \le K\|\widetilde{\Omega}\|^4$, we immediately have $\mathrm{tr}(\widetilde{\Omega}^4) \asymp \|\widetilde{\Omega}\|^4 \asymp |\lambda_2|^4$. Additionally, $|\lambda_2| = \alpha\lambda_1$ in our notation, and $\lambda_1 \asymp \|\theta\|^2$ by Lemma E.2. It follows that

$$\mathrm{tr}(\widetilde{\Omega}^4) \asymp |\lambda_2|^4 \asymp \alpha^4\|\theta\|^8.$$

At the same time, by (74), $|\widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}| \le C\alpha^4\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2$. We thus have

$$\begin{aligned}|X_6 - \mathrm{tr}(\widetilde{\Omega}^4)| &\le C\alpha^4\sum_{i,j,k,\ell(not\,dist)}\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\\ &\le C\alpha^4\sum_{i,j,k}\theta_i^2\theta_j^2\theta_k^4\\ &\le C\alpha^4\|\theta\|^4\|\theta\|_4^4 = o(\alpha^4\|\theta\|^8),\end{aligned}$$

where the last equality is due to $\|\theta\|_4^4 \le \theta_{\max}^2\|\theta\|^2 = o(\|\theta\|^4)$. Combining the above gives

$$X_6 = \mathrm{tr}(\widetilde{\Omega}^4)\cdot[1+o(1)].$$

This proves the claims of X_6.

Last, we combine the results for X_1-X_6 to study $\widetilde{Q}_n$. Note that

$$\widetilde{Q}_n = X_1 + 4X_2 + 4X_3 + 2X_4 + 4X_5 + X_6.$$

Only X_6 has a nonzero mean. So,

$$\mathbb{E}[\widetilde{Q}_n] = \mathbb{E}[X_6] = \mathrm{tr}(\widetilde{\Omega}^4)\cdot[1+o(1)].$$

At the same time, given random variables $Z_1, Z_2, \ldots, Z_m$, $\mathrm{Var}(\sum_{k=1}^m Z_k) = \sum_k \mathrm{Var}(Z_k) + \sum_{k\neq\ell}\mathrm{Cov}(Z_k, Z_\ell) \le \sum_k \mathrm{Var}(Z_k) + \sum_{k\neq\ell}\sqrt{\mathrm{Var}(Z_k)\mathrm{Var}(Z_\ell)} \le m^2\max_k\{\mathrm{Var}(Z_k)\}$. We thus have

$$\mathrm{Var}(\widetilde{Q}_n) \le C\max_{1\le k\le 6}\mathrm{Var}(X_k) \le C\big(\|\theta\|^8 + \alpha^6\|\theta\|^8\|\theta\|_3^6\big).$$

The proof of this lemma is now complete.

G.4.4. *Proof of Lemma G.4*. Recall that $U_a = 4Y_1 = 4\sum_{i,j,k,\ell(dist)} \delta_{ij} W_{jk}W_{k\ell}W_{\ell i}$. By definition, $\delta_{ij} = \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_i - \tilde{\eta}_i)$. It follows that

$$U_a = 4 \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)W_{jk}W_{k\ell}W_{\ell i} + 4 \sum_{i,j,k,\ell(dist)} \eta_j(\eta_i - \tilde{\eta}_i)W_{jk}W_{k\ell}W_{\ell i}.$$

In the second sum, if we relabel $(i,j,k,\ell) = (j', i', \ell', k')$, it becomes

$$4 \sum_{i',j',k',\ell'(dist)} \eta_{i'}(\eta_{j'} - \tilde{\eta}_{j'})W_{i'\ell'}W_{\ell'k'}W_{k'j'} = 4 \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)W_{i\ell}W_{\ell k}W_{kj},$$

which is the same as the first term. It follows that

$$U_a = 8 \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)W_{jk}W_{k\ell}W_{\ell i}.$$

By definition, $\eta_j = \frac{1}{\sqrt{v}}\sum_{s\neq j}\mathbb{E}A_{js}$ and $\tilde{\eta}_j = \frac{1}{\sqrt{v}}\sum_{s\neq j} A_{js}$. Hence,

$$\tilde{\eta}_j - \eta_j = \frac{1}{\sqrt{v}}\sum_{s\neq j} W_{js}. \tag{77}$$

We then re-write

$$\begin{aligned} U_a &= -8 \sum_{i,j,k,\ell(dist)} \eta_i\Big(\frac{1}{\sqrt{v}}\sum_{s\neq j} W_{js}\Big)W_{jk}W_{k\ell}W_{\ell i} \\ &= -\frac{8}{\sqrt{v}} \sum_{\substack{i,j,k,\ell(dist)\\ s\neq j}} \eta_i W_{js}W_{jk}W_{k\ell}W_{\ell i}. \end{aligned}$$

In the summand, (i,j,k,ℓ) are distinct, but s is only required to be distinct from j. We consider two different cases: (a) the case of $s = k$, where the summand becomes $W_{jk}^2 W_{k\ell}W_{\ell i}$, and (b) the case of $s \neq k$. Correspondingly, we write

$$\begin{aligned} U_a &= -\frac{8}{\sqrt{v}} \sum_{i,j,k,\ell(dist)} \eta_i W_{jk}^2 W_{k\ell}W_{\ell i} - \frac{8}{\sqrt{v}} \sum_{\substack{i,j,k,\ell(dist)\\ s\notin\{j,k\}}} \eta_i W_{js}W_{jk}W_{k\ell}W_{\ell i} \\ &\equiv U_{a1} + U_{a2}. \end{aligned} \tag{78}$$

It is easy to see that the summands in both sums have mean zero. Therefore,

$$\mathbb{E}[U_a] = 0.$$

Next, we bound the variance of U_a. Since $\mathrm{Var}(U_a) \leq 2\mathrm{Var}(U_{a1}) + 2\mathrm{Var}(U_{a2})$, it suffices to bound the variances of U_{a1} and U_{a2}. Consider U_{a1}. Note that

$$\mathrm{Var}(U_{a1}) = \frac{64}{v} \sum_{\substack{i,j,k,\ell(dist)\\ i',j',k',\ell'(dist)}} \eta_i\eta_{i'} \cdot \mathbb{E}[W_{jk}^2 W_{k\ell}W_{\ell i}W_{j'k'}^2 W_{k'\ell'}W_{\ell' i'}]. \tag{79}$$

By definition, $v = 1_n'(\mathbb{E}A)1_n = 1_n'\Omega 1_n - \sum_i \Omega_{ii}$. Since $\Omega_{ii} \leq \theta_i^2$, it implies $v = 1_n'\Omega 1_n - O(\|\theta\|^2) = 1_n'\Omega 1_n + o(\|\theta\|_1^2)$. Moreover, we note that $1_n'\Omega 1_n \leq C\sum_{i,j}\theta_i\theta_j \leq C\|\theta\|_1^2$, and by Lemma E.4, $1_n'\Omega 1_n \geq C^{-1}\|\theta\|_1^2$. Combining these results gives

$$C^{-1}\|\theta\|_1^2 \leq v \leq C\|\theta\|_1^2. \tag{80}$$

Moreover, $\eta_i = \frac{1}{\sqrt{v}}\sum_{s\neq i}\Omega_{is} \leq \frac{C}{\|\theta\|_1}\sum_s \theta_i\theta_s$. This gives

$$0 \leq \eta_i \leq C\theta_i, \qquad \text{for all } 1 \leq i \leq n. \tag{81}$$

We plug (80)-(81) into (79) and find out that

$$\mathrm{Var}(U_{a1}) \leq \frac{C}{\|\theta\|_1^2} \sum_{\substack{i,j,k,\ell(dist)\\ i',j',k',\ell'(dist)}} \theta_i\theta_{i'} \cdot \mathbb{E}[W_{jk}^2 W_{k\ell} W_{\ell i} W_{j'k'}^2 W_{k'\ell'} W_{\ell' i'}].$$

In order for the summand to be nonzero, all W terms have to be perfectly paired. By elementary calculations,

$$\theta_i\theta_{i'}\mathbb{E}[W_{jk}^2 W_{k\ell} W_{\ell i} W_{j'k'}^2 W_{k'\ell'} W_{\ell' i'}] = \begin{cases} \theta_i^2\,\mathbb{E}[W_{jk}^2 W_{k\ell}^2 W_{\ell i}^2 W_{j'k}^2], & \text{if } (\ell',k',i')=(\ell,k,i);\\ \theta_i\theta_k\,\mathbb{E}[W_{jk}^2 W_{k\ell}^2 W_{\ell i}^2 W_{j'i}^2], & \text{if } (\ell',k',i')=(\ell,i,k);\\ \theta_i\theta_j\,\mathbb{E}[W_{jk}^3 W_{k\ell}^2 W_{\ell i}^3], & \text{if } (j',k')=(i,\ell),\ (i',\ell')=(j,k);\\ 0, & \text{otherwise.}\end{cases}$$

Here, (i,j,k,ℓ) are distinct. In the second case above, $(W_{jk}^2, W_{k\ell}^2, W_{\ell i}^2, W_{j'i}^2)$ are independent of each other, no matter $j=j'$ or $j\neq j'$ (we remark that $j'\neq \ell$, because $j'\notin\{i',k',\ell'\}=\{i,k,\ell\}$). It follows that $\mathbb{E}[W_{jk}^2 W_{k\ell}^2 W_{\ell i}^2 W_{j'i}^2] \leq \Omega_{jk}\Omega_{k\ell}\Omega_{\ell i}\Omega_{j'i} \leq C\theta_i^2\theta_j\theta_k^2\theta_\ell^2\theta_{j'}$. In the first case, when $j\neq j'$, $\mathbb{E}[W_{jk}^2 W_{k\ell}^2 W_{\ell i}^2 W_{j'k}^2] \leq \Omega_{jk}\Omega_{k\ell}\Omega_{\ell i}\Omega_{j'k} \leq C\theta_i\theta_j\theta_k^3\theta_\ell^2\theta_{j'}$; when $j=j'$, it holds that $\mathbb{E}[W_{jk}^2 W_{k\ell}^2 W_{\ell i}^2 W_{j'k}^2] = \mathbb{E}[W_{jk}^4 W_{k\ell}^2 W_{\ell i}^2] \leq C\theta_i\theta_j\theta_k^2\theta_\ell^2$. In the third case, $(W_{jk}^3, W_{k\ell}^2, W_{\ell i}^3)$ are mutually independent, so $\mathbb{E}[W_{jk}^2 W_{k\ell}^2 W_{\ell i}^2] \leq \Omega_{jk}\Omega_{k\ell}\Omega_{\ell i} \leq C\theta_i\theta_j\theta_k^2\theta_\ell^2$. We then have

$$\theta_i\theta_{i'}\mathbb{E}[W_{jk}^2 W_{k\ell} W_{\ell i} W_{j'k'}^2 W_{k'\ell'} W_{\ell' i'}] \leq \begin{cases} C\theta_i^3\theta_j\theta_k^2\theta_\ell^2, & \text{if } (\ell',k',i')=(\ell,k,i),\ j'=j;\\ C\theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_{j'}, & \text{if } (\ell',k',i')=(\ell,k,i),\ j'\neq j;\\ C\theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_{j'}, & \text{if } (\ell',k',i')=(\ell,i,k);\\ C\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2, & \text{if } (j',k')=(i,\ell),\ (i',\ell')=(j,k);\\ 0, & \text{otherwise.}\end{cases}$$

It follows that

$$\begin{aligned}\mathrm{Var}(U_{a1}) &\leq \frac{C}{\|\theta\|_1^2}\Big(\sum_{i,j,k,\ell}\theta_i^3\theta_j\theta_k^2\theta_\ell^2 + \sum_{i,j,k,\ell,j'}\theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_{j'} + \sum_{i,j,k,\ell}\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\Big)\\ &\leq \frac{C}{\|\theta\|_1^2}\big(\|\theta\|^4\|\theta\|_3^3\|\theta\|_1 + \|\theta\|^2\|\theta\|_3^6\|\theta\|_1^2 + \|\theta\|^8\big)\\ &\leq C\|\theta\|^2\|\theta\|_3^6, \qquad (82)\end{aligned}$$

where we obtain the last inequality as follows: By Cauchy-Schwarz inequality, $\|\theta\|^4 = (\sum_i \theta_i^{1/2}\cdot\theta^{3/2})^2 \leq (\sum_i\theta_i)(\sum_i\theta_i^3) \leq \|\theta\|_1\|\theta\|_3^3$; therefore, $\|\theta\|^8 \leq \|\theta\|^4\|\theta\|_3^3\|\theta\|_1 \leq \|\theta\|_3^6\|\theta\|_1^2$. We then consider U_{a2}. Define

$$\mathcal{P}_5^* = \left\{\begin{array}{l}\text{path } i\text{-}\ell\text{-}k\text{-}j\text{-}s \text{ in a complete}\\ \text{graph with } n \text{ nodes}\end{array} : \begin{array}{l}\text{nodes } i,j,k,\ell \text{ are distinct,}\\ \text{and node } s \text{ is different from } j,\, k\end{array}\right\}.$$

Fix a path i-ℓ-k-j-s in $\mathcal{P}_5^*$. If $s\notin\{i,\ell\}$, then this path is counted twice in the definition of U_{a2}, as i-ℓ-k-j-s and s-j-k-ℓ-i, respectively. If $s\in\{i,\ell\}$, then it is counted only once in the definition of U_{a2}. Hence, we can re-write

$$U_{a2} = -\frac{8}{\sqrt{v}}\sum_{\substack{\text{path in } \mathcal{P}_5^*\\ s\notin\{i,\ell\}}}(\eta_i+\eta_s)W_{sj}W_{jk}W_{k\ell}W_{\ell i} - \frac{8}{\sqrt{v}}\sum_{\substack{\text{path in } \mathcal{P}_5^*\\ s\in\{i,\ell\}}}\eta_i W_{sj}W_{jk}W_{k\ell}W_{\ell i}.$$

For two distinct paths in $\mathcal{P}_5^*$, the corresponding summands are uncorrelated with each other. It follows that

$$\begin{aligned}
\mathrm{Var}(U_{a2}) &= \frac{64}{v} \sum_{\substack{\text{path in } \mathcal{P}_5^* \\ s\notin\{i,\ell\}}} (\eta_i+\eta_s)^2\, \mathrm{Var}(W_{sj}W_{jk}W_{k\ell}W_{\ell i}) \\
&\quad + \frac{64}{v} \sum_{\substack{\text{path in } \mathcal{P}_5^* \\ s\in\{i,\ell\}}} \eta_i^2\, \mathrm{Var}(W_{sj}W_{jk}W_{k\ell}W_{\ell i}) \\
&\le \frac{C}{v} \sum_{i,j,k,\ell,s} (\eta_i^2+\eta_s^2)\cdot \theta_i\theta_j^2\theta_k^2\theta_\ell^2\theta_s \\
&\le \frac{C}{\|\theta\|_1^2} \sum_{i,j,k,\ell,s} (\theta_i^3\theta_j^2\theta_k^2\theta_\ell^2\theta_s + \theta_i\theta_j^2\theta_k^2\theta_\ell^2\theta_s^3) \\
&\le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1}. \qquad (83)
\end{aligned}$$

By Cauchy-Schwarz inequality, $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$, so the right hand side of (83) is $\le C\|\theta\|^2\|\theta\|_3^6$. Combining it with (82) gives

$$\mathrm{Var}(U_a) \le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8).$$

This proves the claim.

G.4.5. *Proof of Lemma G.5.* It suffices to prove the claims for each of Y_1-Y_6. Consider Y_1. We have analyzed this term under the null hypothesis. Using similar proof, we can easily derive that

$$\mathbb{E}[Y_1] = 0, \qquad \mathrm{Var}(Y_1) \le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8).$$

Consider Y_2. Using the definition of Y_2 and the expression of $\tilde{\eta}_i$ in (77), we have

$$\begin{aligned}
Y_2 &= \sum_{i,j,k,\ell (dist)} \delta_{ij}\widetilde{\Omega}_{jk}W_{k\ell}W_{\ell i} \\
&= \sum_{i,j,k,\ell (dist)} \eta_i(\eta_j-\tilde{\eta}_j)\widetilde{\Omega}_{jk}W_{k\ell}W_{\ell i} + \sum_{i,j,k,\ell (dist)} \eta_j(\eta_i-\tilde{\eta}_i)\widetilde{\Omega}_{jk}W_{k\ell}W_{\ell i} \\
&= \frac{1}{\sqrt{v}} \sum_{i,j,k,\ell (dist)} \eta_i\Big(-\sum_{s\neq j} W_{js}\Big)\widetilde{\Omega}_{jk}W_{k\ell}W_{\ell i} + \frac{1}{\sqrt{v}} \sum_{i,j,k,\ell (dist)} \eta_j\Big(-\sum_{s\neq i} W_{is}\Big)\widetilde{\Omega}_{jk}W_{k\ell}W_{\ell i} \\
&= -\frac{1}{\sqrt{v}} \sum_{\substack{i,j,k,\ell (dist) \\ s\neq j}} \eta_i\widetilde{\Omega}_{jk}W_{js}W_{k\ell}W_{\ell i} - \frac{1}{\sqrt{v}} \sum_{\substack{i,k,\ell (dist) \\ s\neq i}} \Big(\sum_{j\notin\{i,k,\ell\}} \eta_j\widetilde{\Omega}_{jk}\Big) W_{is}W_{k\ell}W_{\ell i}.
\end{aligned}$$

In the second sum above, we further separate two cases, $s=\ell$ and $s\neq\ell$. It then gives rise to three terms:

$$\begin{aligned}
Y_2 = &-\frac{1}{\sqrt{v}} \sum_{\substack{i,j,k,\ell (dist) \\ s\neq j}} \eta_i\widetilde{\Omega}_{jk}W_{js}W_{k\ell}W_{\ell i} \\
&-\frac{1}{\sqrt{v}} \sum_{i,k,\ell (dist)} \Big(\sum_{j\notin\{i,k,\ell\}} \eta_j\widetilde{\Omega}_{jk}\Big) W_{i\ell}^2 W_{k\ell}
\end{aligned}$$

$$-\frac{1}{\sqrt{v}}\sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\}}}\Big(\sum_{j\notin\{i,k,\ell\}}\eta_j\widetilde{\Omega}_{jk}\Big)W_{is}W_{k\ell}W_{\ell i}$$

(84)
$$\equiv Y_{2a}+Y_{2b}+Y_{2c}.$$

Since (i,j,k,ℓ) are distinct, it is easy to see that all three terms have mean zero. We thus have

$$\mathbb{E}[Y_2]=0.$$

Below, we calculate the variances. First, we bound the variance of Y_{2a}. Each (i,j,k,ℓ,s) is associated with a length-3 path i-k-ℓ and an edge j-s in the complete graph. For (i,j,k,ℓ,s) and (i',j',k',ℓ',s'), if the associated path and edge are the same, then we group them together. Given a length-3 path i-k-ℓ and an edge j-s (such that the edge is not in the path), they are counted four times in the definition of Y_{2a}, as (i) i-k-ℓ and j-s, (ii) i-k-ℓ and s-j, (iii) ℓ-k-i and j-s, (iv) ℓ-k-i and s-j, so we group these four summands together. After grouping the summands, we re-write

$$Y_{2a}=-\frac{1}{\sqrt{v}}\sum_{\substack{\text{length-3}\\ \text{path}}}\sum_{\substack{\text{edge not}\\ \text{in the path}}}\big(\eta_i\widetilde{\Omega}_{jk}+\eta_i\widetilde{\Omega}_{sk}+\eta_k\widetilde{\Omega}_{ji}+\eta_k\widetilde{\Omega}_{si}\big)W_{js}W_{k\ell}W_{\ell i}.$$

In this new expression of Y_{2a}, two summands are correlated only when the underlying path&edge pairs are exactly the same. Additionally, by (74) and (81),

$$\big|\eta_i\widetilde{\Omega}_{jk}+\eta_i\widetilde{\Omega}_{sk}+\eta_k\widetilde{\Omega}_{ji}+\eta_k\widetilde{\Omega}_{si}\big|\le C\alpha(\theta_j+\theta_s)\theta_i\theta_k.$$

It follows that

$$\begin{aligned}\mathrm{Var}(Y_{2a})&\le\frac{C}{v}\sum_{i,j,k,\ell,s}\alpha^2(\theta_j+\theta_s)^2\theta_i^2\theta_k^2\cdot\mathrm{Var}(W_{js}W_{k\ell}W_{\ell i})\\&\le\frac{C}{\|\theta\|_1^2}\sum_{i,j,k,\ell,s}\alpha^2(\theta_j+\theta_s)^2\theta_i^2\theta_k^2\cdot\theta_i\theta_j\theta_k\theta_\ell^2\theta_s\\&\le\frac{C\alpha^2}{\|\theta\|_1^2}\sum_{i,j,k,\ell,s}(\theta_i^3\theta_j^3\theta_k^3\theta_\ell^2\theta_s+\theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_s^3)\end{aligned}$$

(85)
$$\le\frac{C\alpha^2\|\theta\|^2\|\theta\|_3^9}{\|\theta\|_1}.$$

Second, we bound the variance of Y_{2b}. Write $\beta_{ik\ell}=\sum_{j\notin\{i,k,\ell\}}\eta_j\widetilde{\Omega}_{jk}$. By (74) and (81), $|\beta_{ik\ell}|\le C\sum_j\theta_j\cdot\alpha\theta_j\theta_k\le C\alpha\|\theta\|^2\theta_k$. Using this notation,

$$Y_{2b}=\frac{1}{v}\sum_{i,j,k,\ell(dist)}\beta_{ik\ell}W_{i\ell}^2W_{k\ell},\qquad\text{where}\quad|\beta_{ik\ell}|\le C\alpha\|\theta\|^2\theta_k.$$

It follows that

$$\begin{aligned}\mathrm{Var}(Y_{2b})=\mathbb{E}[Y_{2b}^2]&\le\frac{C}{v}\sum_{\substack{i,k,\ell(dist)\\ i',k',\ell'(dist)}}\beta_{ik\ell}\beta_{i'k'\ell'}\cdot\mathbb{E}[W_{i\ell}^2W_{k\ell}W_{i'\ell'}^2W_{k'\ell'}]\\&\le\frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^2}\sum_{\substack{i,k,\ell(dist)\\ i',k',\ell'(dist)}}\theta_k\theta_{k'}\cdot\mathbb{E}[W_{i\ell}^2W_{k\ell}W_{i'\ell'}^2W_{k'\ell'}].\end{aligned}$$

The summand is nonzero only when the two variables $W_{k\ell}$ and $W_{k'\ell'}$ equal to each other or when each of them equals to some other squared variables. By elementary calculations,

$$\theta_k\theta_{k'}\cdot\mathbb{E}[W_{i\ell}^2W_{k\ell}W_{i'\ell'}^2W_{k'\ell'}]$$
$$=\begin{cases}\theta_k^2\,\mathbb{E}[W_{i\ell}^4W_{k\ell}^2]\le C\theta_i\theta_k^3\theta_\ell^2, & \text{if } (k',\ell')=(k,\ell),\ i'=i;\\ \theta_k^2\,\mathbb{E}[W_{i\ell}^2W_{k\ell}^2W_{i'\ell}^2]\le C\theta_i\theta_k^3\theta_\ell^3\theta_{i'}, & \text{if } (k',\ell')=(k,\ell),\ i'\neq i;\\ \theta_k\theta_\ell\,\mathbb{E}[W_{i\ell}^2W_{k\ell}^2W_{i'k}^2]\le C\theta_i\theta_k^3\theta_\ell^3\theta_{i'}, & \text{if } (k',\ell')=(\ell,k);\\ \theta_k^2\,\mathbb{E}[W_{i\ell}^3W_{k\ell}^3]\le C\theta_i\theta_k^3\theta_\ell^2, & \text{if } \ell'=\ell,\ (i',k')=(i,k);\\ \theta_k\theta_i\,\mathbb{E}[W_{i\ell}^3W_{k\ell}^3]\le C\theta_i^2\theta_k^2\theta_\ell^2, & \text{if } \ell'=\ell,\ (i',k')=(k,i);\\ 0, & \text{otherwise.}\end{cases}$$

As a result,

$$\begin{aligned}\mathrm{Var}(Y_{2b}) &\le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^2}\Big(\sum_{i,k,\ell}\theta_i\theta_k^3\theta_\ell^2+\sum_{i,k,\ell,i'}\theta_i\theta_k^3\theta_\ell^3\theta_{i'}+\sum_{i,k,\ell}\theta_i^2\theta_k^2\theta_\ell^2\Big)\\ &\le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^2}\big(\|\theta\|_3^3\|\theta\|^2\|\theta\|_1+\|\theta\|_3^6\|\theta\|_1^2+\|\theta\|^6\big)\\ &\le C\alpha^2\|\theta\|^4\|\theta\|_3^6, \end{aligned}\tag{86}$$

where to get the last inequality we have used $\|\theta\|^6\ll\|\theta\|^8\le(\|\theta\|_1\|\theta\|_3^3)^2$ and $\|\theta\|_3^3\|\theta\|^2\|\theta\|_1\ll\|\theta\|_3^3\|\theta\|^4\|\theta\|_1\le(\|\theta\|_1\|\theta\|_3^3)^2$. Last, we bound the variance of Y_{2c}. Let $\beta_{ik\ell}=\sum_{j\notin\{i,k,\ell\}}\eta_j\widetilde{\Omega}_{jk}$ be the same as above. We write

$$Y_{2c}=\frac{1}{\sqrt{v}}\sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\}}}\beta_{ik\ell}W_{is}W_{k\ell}W_{\ell i},\qquad \text{where}\quad |\beta_{ik\ell}|\le C\alpha\|\theta\|^2\theta_k.$$

For $\mathbb{E}[W_{is}W_{k\ell}W_{\ell i}\cdot W_{i's'}W_{k'\ell'}W_{\ell'i'}]$ to be nonzero, it has to be the case that $(W_{is},W_{k\ell},W_{\ell i})$ and $(W_{i's'},W_{k'\ell'},W_{\ell'i'})$ are the same set of variables, up to an order permutation. For each fixed (i,k,ℓ,s), there are only a constant number of (i',k',ℓ',s') such that the above is satisfied. As we have argued many times before (e.g., see (70)), it is true that

$$\begin{aligned}\mathrm{Var}(Y_{2c}) &\le \frac{C}{v}\sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\}}}\beta_{ik\ell}^2\cdot\mathrm{Var}(W_{is}W_{k\ell}W_{\ell i})\\ &\le \frac{C}{\|\theta\|_1^2}\sum_{i,k,\ell,s}(\alpha\|\theta\|^2\theta_k)^2\cdot\theta_i^2\theta_k\theta_\ell^2\theta_s\\ &\le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1}. \end{aligned}\tag{87}$$

We now combine the variances of Y_{2a}-Y_{2c}. Since $\|\theta\|_3^3\le\theta_{\max}^2\|\theta\|_1\ll\|\theta\|_1$, the right hand side is (85) is $o(\alpha^2\|\theta\|^2\|\theta\|_3^6)=o(\alpha^2\|\theta\|^4\|\theta\|_3^6)$. Since $\|\theta\|^4\le\|\theta\|_1\|\theta\|_3^3$, the right hand side is (87) is $\le C\alpha^2\|\theta\|^4\|\theta\|_3^6$. It follows that

$$\mathrm{Var}(Y_2)\le C\alpha^2\|\theta\|^4\|\theta\|_3^6=o(\|\theta\|^8).$$

This proves the claims of Y_2.

Consider Y_3. By definition,

$$Y_3=\sum_{i,j,k,\ell(dist)}\eta_i(\eta_j-\tilde{\eta}_j)W_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i}+\sum_{i,j,k,\ell(dist)}\eta_j(\eta_i-\tilde{\eta}_i)W_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i}.$$

In the second sum, if we relabel $(i,j,k,\ell)=(j',i',\ell',k')$, it can be written as $\sum_{i',j',k',\ell'(dist)} \eta_{i'}(\eta_{j'} - \tilde{\eta}_{j'})W_{i'\ell'}\widetilde{\Omega}_{\ell'k'}W_{k'j'}$. This shows that the second sum is indeed equal to the first sum. As a result,

$$
\begin{aligned}
Y_3 &= 2\sum_{i,j,k,\ell(dist)} \eta_i(\eta_j-\tilde{\eta}_j)W_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i} \\
&= 2\sum_{i,j,k,\ell(dist)} \eta_i\Big(-\frac{1}{\sqrt{v}}\sum_{s\neq j}W_{js}\Big)W_{jk}\widetilde{\Omega}_{k\ell}W_{\ell i} \\
&= -\frac{2}{\sqrt{v}}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq j}} \eta_i\widetilde{\Omega}_{k\ell}W_{js}W_{jk}W_{\ell i} \\
&= -\frac{2}{\sqrt{v}}\sum_{i,j,k,\ell(dist)} \eta_i\widetilde{\Omega}_{k\ell}W_{jk}^2W_{\ell i} - \frac{2}{\sqrt{v}}\sum_{\substack{i,j,k,\ell(dist)\\ s\notin\{j,k\}}} \eta_i\widetilde{\Omega}_{k\ell}W_{js}W_{jk}W_{\ell i} \\
&\equiv Y_{3a}+Y_{3b}, \qquad (88)
\end{aligned}
$$

where the second line is from (77) and the second last line is from dividing all summands into two cases of $s=k$ and $s\neq k$. Both terms have mean zero, so

$$\mathbb{E}[Y_3]=0.$$

Below, first, we calculate the variance of Y_{3a}.

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

The summand is nonzero only if either the two variables $W_{\ell i}$ and $W_{\ell' i'}$ are the same, or each of the two variables $W_{\ell i}$ and $W_{\ell' i'}$ equals to another squared W term. By (74), (81), and elementary calculations,

$$
\begin{aligned}
&(\eta_i\widetilde{\Omega}_{k\ell}\eta_{i'}\widetilde{\Omega}_{k'\ell'})\cdot\mathbb{E}[W_{jk}^2W_{\ell i}W_{j'k'}^2W_{\ell'i'}] \\
&\leq C\alpha^2\theta_i\theta_k\theta_\ell\theta_{i'}\theta_{k'}\theta_{\ell'}\cdot\mathbb{E}[W_{jk}^2W_{\ell i}W_{j'k'}^2W_{\ell'i'}] \\
&= \begin{cases}
C\alpha^2\theta_i^2\theta_\ell^2\theta_k^2\,\mathbb{E}[W_{jk}^4W_{\ell i}^2] \leq C\alpha^2\theta_i^3\theta_j\theta_k^3\theta_\ell^3, & \text{if } \{\ell',i'\}=\{\ell,i\},\ (j',k')=(j,k); \\
C\alpha^2\theta_i^2\theta_\ell^2\theta_k\theta_j\,\mathbb{E}[W_{jk}^4W_{\ell i}^2] \leq C\alpha^2\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3, & \text{if } \{\ell',i'\}=\{\ell,i\},\ (j',k')=(k,j); \\
C\alpha^2\theta_i^2\theta_\ell^2\theta_k\theta_{k'}\,\mathbb{E}[W_{jk}^2W_{\ell i}^2W_{j'k'}^2] \leq C\alpha^2\theta_i^3\theta_j\theta_k^2\theta_\ell^3\theta_{j'}\theta_{k'}^2, & \text{if } \{\ell',i'\}=\{\ell,i\},\ \{j',k'\}\neq\{j,k\}; \\
C\alpha^2\theta_i^2\theta_\ell\theta_j\theta_k^2\,\mathbb{E}[W_{jk}^3W_{\ell i}^3] \leq C\alpha^2\theta_i^3\theta_j^2\theta_k^3\theta_\ell^2, & \text{if } \{\ell',i'\}=\{j,k\},\ (j',k')=(\ell,i); \\
C\alpha^2\theta_i\theta_\ell^2\theta_j\theta_k^2\,\mathbb{E}[W_{jk}^3W_{\ell i}^3] \leq C\alpha^2\theta_i^2\theta_j^2\theta_k^3\theta_\ell^3, & \text{if } \{\ell',i'\}=\{j,k\},\ (j',k')=(i,\ell); \\
0, & \text{otherwise.}
\end{cases}
\end{aligned}
$$

There are only three different cases in the bounds. It follows that

$$
\begin{aligned}
\mathrm{Var}(Y_{3a}) &\leq \frac{C\alpha^2}{\|\theta\|_1^2}\Big(\sum_{i,j,k,\ell}\theta_i^3\theta_j\theta_k^3\theta_\ell^3 + \sum_{i,j,k,\ell}\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3 + \sum_{i,j,k,\ell,j',k'}\theta_i^3\theta_j\theta_k^2\theta_\ell^3\theta_{j'}\theta_{k'}^2\Big) \\
&\leq \frac{C\alpha^2}{\|\theta\|_1^2}\big(\|\theta\|_1\|\theta\|_3^9 + \|\theta\|^4\|\theta\|_3^6 + \|\theta\|^4\|\theta\|_1^2\|\theta\|_3^6\big) \\
&\leq C\alpha^2\|\theta\|^4\|\theta\|_3^6, \qquad (89)
\end{aligned}
$$

where in the last line we have used $\|\theta\|_3^9 \le \|\theta\|_3^6(\theta_{\max}\|\theta\|^2) = o(\|\theta\|^2\|\theta\|_3^6)$ and $\|\theta\|_1 \ge \theta_{\max}^{-1}\|\theta\|^2 \to \infty$. Next, we calculate the variance of Y_{3b}. We mimic the argument in (85) and group summands according to the underlying path s-j-k and edge ℓ-i in a complete graph. It yields

$$Y_{3b} = -\frac{2}{\sqrt{v}} \sum_{\substack{\text{length-3} \\ \text{path}}} \sum_{\substack{\text{edge not} \\ \text{in the path}}} \big(\eta_i\widetilde{\Omega}_{k\ell} + \eta_\ell\widetilde{\Omega}_{ki} + \eta_i\widetilde{\Omega}_{s\ell} + \eta_\ell\widetilde{\Omega}_{si}\big)W_{sj}W_{jk}W_{\ell i},$$

where

$$\big|\eta_i\widetilde{\Omega}_{k\ell} + \eta_\ell\widetilde{\Omega}_{ki} + \eta_i\widetilde{\Omega}_{s\ell} + \eta_\ell\widetilde{\Omega}_{si}\big| \le C\alpha(\theta_k+\theta_s)\theta_i\theta_\ell.$$

It follows that

$$\begin{aligned}
\mathrm{Var}(Y_{3b}) &\le \frac{C}{v}\sum_{i,j,k,\ell,s} \alpha^2(\theta_k+\theta_s)^2\theta_i^2\theta_\ell^2 \cdot \mathrm{Var}(W_{sj}W_{jk}W_{\ell i}) \\
&\le \frac{C\alpha^2}{\|\theta\|_1^2}\sum_{i,j,k,\ell,s}(\theta_i^3\theta_j^2\theta_k^3\theta_\ell^3\theta_s + \theta_i^3\theta_j^2\theta_k\theta_\ell^3\theta_s^3) \\
&\le \frac{C\alpha^2\|\theta\|^2\|\theta\|_3^9}{\|\theta\|_1}. \qquad (90)
\end{aligned}$$

Since $\|\theta\|_3^9 \le \|\theta\|_3^6(\theta_{\max}\|\theta\|_1) = o(\|\theta\|_1\|\theta\|_3^6)$, so the right hand side of (90) is much smaller than the right hand side of (89). Together, we have

$$\mathrm{Var}(Y_3) \le C\alpha^2\|\theta\|^4\|\theta\|_3^6 = o(\|\theta\|^8).$$

This proves the claims of Y_3.

Consider Y_4. We plug in $\delta_{ij} = \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_i - \tilde{\eta}_i)$ and the expression (77). It gives

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

First, we analyze Y_{4a}. When (i,j,ℓ) are distinct, $W_{js}W_{\ell i}$ has a mean zero. Therefore,

$$\mathbb{E}[Y_{4a}] = 0.$$

To calculate the variance, we rewrite

$$Y_{4a} = -\frac{1}{\sqrt{v}}\sum_{\substack{i,j,\ell(dist)\\ s\neq j}} \beta_{ij\ell}W_{js}W_{\ell i}, \qquad \text{where} \quad \beta_{ij\ell} = \sum_{k\notin\{i,j,\ell\}} \eta_i\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}$$

By (74) and (81), $|\beta_{ij\ell}| \le C\sum_k \alpha^2\theta_i\theta_j\theta_k^2\theta_\ell \le C\alpha^2\|\theta\|^2\theta_i\theta_j\theta_\ell$. Also, for $W_{js}W_{\ell i}$ and $W_{j's'}W_{\ell' i'}$ to be correlated, there are only two cases: $(W_{js}, W_{\ell i}) = (W_{j's'}, W_{\ell' i'})$ or

$(W_{js}, W_{\ell i}) = (W_{\ell' i'}, W_{j's'})$. Mimicking the argument in (85) or (90), we can easily obtain that

$$\mathrm{Var}(Y_{4a}) \leq \frac{C}{v} \sum_{\substack{i,j,\ell (dist)\\ s\neq j}} \beta_{ij\ell}^2 \cdot \mathrm{Var}(W_{js}W_{\ell i})$$

$$\leq \frac{C}{\|\theta\|_1^2} \sum_{i,j,\ell,s} (\alpha^2 \|\theta\|^2 \theta_i\theta_j\theta_\ell)^2 \cdot \theta_i\theta_j\theta_\ell\theta_s$$

$$\leq \frac{C\alpha^4\|\theta\|^4\|\theta\|_3^9}{\|\theta\|_1}. \tag{91}$$

Next, we analyze Y_{4b}. We re-write

$$Y_{4b} = -\frac{1}{\sqrt{v}} \sum_{\substack{i,\ell (dist)\\ s\neq i}} \beta_{i\ell} W_{is} W_{\ell i}, \qquad \text{where} \quad \beta_{i\ell} = \sum_{j,k\notin\{i,\ell\}} \eta_j \widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}.$$

By separating the case of $s = \ell$ from the case of $s \neq \ell$, we have

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

Only $\widetilde{Y}_{4b}$ has a nonzero mean. By (74) and (81),

$$|\beta_{i\ell}| \leq C \sum_{j,k} \alpha^2 \theta_j^2 \theta_k^2 \theta_\ell \leq C\alpha^2 \|\theta\|^4 \theta_\ell.$$

It follows that

$$|\mathbb{E}[Y_{4b}]| = |\mathbb{E}[\widetilde{Y}_{4b}]| \leq \frac{C}{\|\theta\|_1} \sum_{i,\ell} (\alpha^2\|\theta\|^4\theta_\ell)\theta_i\theta_\ell \leq C\alpha^2\|\theta\|^6. \tag{92}$$

We now bound the variances of $\widetilde{Y}_{4b}$ and Y_{4b}^*. By direct calculations,

$$\mathrm{Var}(\widetilde{Y}_{4b}) = \frac{2}{v} \sum_{i,\ell (dist)} \beta_{i\ell}^2 \cdot \mathrm{Var}(W_{i\ell}^2) \leq \frac{C}{\|\theta\|_1^2} \sum_{i,\ell} (\alpha^2\|\theta\|^4\theta_\ell)^2 \cdot \theta_i\theta_\ell \leq \frac{C\alpha^4\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1},$$

$$\mathrm{Var}(Y_{4b}^*) \leq \frac{C}{v} \sum_{\substack{i,\ell (dist)\\ s\notin\{i,\ell\}}} \beta_{i\ell}^2 \cdot \mathrm{Var}(W_{is}W_{\ell i}) \leq \frac{C}{\|\theta\|_1^2} \sum_{i,\ell,s} (\alpha^2\|\theta\|^4\theta_\ell)^2 \cdot \theta_i^2\theta_\ell\theta_s \leq \frac{C\alpha^4\|\theta\|^{10}\|\theta\|_3^3}{\|\theta\|_1}.$$

Together, we have

$$\mathrm{Var}(Y_{4b}) \leq 2\mathrm{Var}(\widetilde{Y}_{4b}) + 2\mathrm{Var}(Y_{4b}^*) \leq \frac{C\alpha^4\|\theta\|^{10}\|\theta\|_3^3}{\|\theta\|_1}. \tag{93}$$

We combine the results of Y_{4a} and Y_{4b}. Since $\|\theta\|_3^6 \leq (\theta_{\max}\|\theta\|^2)^2 = o(\|\theta\|^4)$, the right hand side of (92) dominates the right hand side of (91). It follows that

$$|\mathbb{E}[Y_4]| \leq C\alpha^2\|\theta\|^6 = o(\alpha^4\|\theta\|^8), \quad \mathrm{Var}(Y_4) \leq \frac{C\alpha^4\|\theta\|^{10}\|\theta\|_3^3}{\|\theta\|_1} = o(\alpha^6\|\theta\|^8\|\theta\|_3^6).$$

Here, we explain the equalities. The first one is due to $\alpha^2\|\theta\|^2 \to \infty$. To get the second equality, we compare $\mathrm{Var}(Y_4)$ with the order of $\alpha^6\|\theta\|^8\|\theta\|_3^6$. Note that $\frac{\|\theta\|^{10}\|\theta\|_3^3}{\|\theta\|_1} = \frac{\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1}\|\theta\|^4 \leq$

$\frac{\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1}\|\theta\|_1\|\theta\|_3^3 \leq \|\theta\|^6\|\theta\|_3^6$. It follows that $\mathrm{Var}(Y_4) \leq C\alpha^4\|\theta\|^6\|\theta\|_3^6 \ll C\alpha^6\|\theta\|^8\|\theta\|_3^6$, where the last inequality is due to $\alpha^2\|\theta\|^2 \to \infty$. So far, we have proved all claims about Y_4.

Consider Y_5. Recall that

$$Y_5 = \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}W_{k\ell}\widetilde{\Omega}_{\ell i} + \sum_{i,j,k,\ell(dist)} (\eta_i - \tilde{\eta}_i)\eta_j\widetilde{\Omega}_{jk}W_{k\ell}\widetilde{\Omega}_{\ell i}.$$

With relabeling of $(i,j,k,\ell) = (j',i',\ell',k')$, the second sum can be written as $\sum_{i',j',k',\ell'(dist)}(\eta_{j'} - \tilde{\eta}_{j'})\eta_{i'}\widetilde{\Omega}_{i'\ell'}W_{\ell'k'}\widetilde{\Omega}_{k'j'}$. This suggests that it is actually equal to the first sum above. Hence,

$$\begin{aligned}
Y_5 &= 2\sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}W_{k\ell}\widetilde{\Omega}_{\ell i} \\
&= \sum_{i,j,k,\ell(dist)} \eta_i\Big(-\frac{2}{\sqrt{v}}\sum_{s\neq j} W_{js}\Big)\widetilde{\Omega}_{jk}W_{k\ell}\widetilde{\Omega}_{\ell i} \\
&= -\frac{2}{\sqrt{v}}\sum_{\substack{j,k,\ell(dist)\\ s\neq j}}\Big(\sum_{i\notin\{j,k,\ell\}}\eta_i\widetilde{\Omega}_{jk}\widetilde{\Omega}_{\ell i}\Big)W_{js}W_{k\ell} \\
&\equiv -\frac{2}{\sqrt{v}}\sum_{\substack{j,k,\ell(dist)\\ s\neq j}}\beta_{jk\ell}W_{js}W_{k\ell}, \qquad \text{where} \quad \beta_{jk\ell} \equiv \sum_{i\notin\{j,k,\ell\}}\eta_i\widetilde{\Omega}_{jk}\widetilde{\Omega}_{\ell i}.
\end{aligned}$$

It is easy to see that $\mathbb{E}[W_{js}W_{k\ell}] = 0$ when (j,k,ℓ) are distinct. Hence,

$$\mathbb{E}[Y_5] = 0.$$

By (74) and (81), $|\beta_{jk\ell}| \leq C\sum_i \theta_i \cdot \alpha^2\theta_j\theta_k\theta_\ell\theta_i \leq C\alpha^2\|\theta\|^2\theta_j\theta_k\theta_\ell$. Similar to the argument in (85) or (90), we can show that

$$\begin{aligned}
\mathrm{Var}(Y_5) &\leq \frac{C}{v}\sum_{\substack{j,k,\ell(dist)\\ s\neq j}}\beta_{jk\ell}^2 \cdot \mathrm{Var}(W_{js}W_{k\ell}) \\
&\leq \frac{C}{\|\theta\|_1^2}\sum_{j,k,\ell,s}(\alpha^2\|\theta\|^2\theta_j\theta_k\theta_\ell)^2\theta_j\theta_s\theta_k\theta_\ell \\
&\leq \frac{C\alpha^4\|\theta\|^4\|\theta\|_3^9}{\|\theta\|_1}.
\end{aligned}$$

Since $\|\theta\|_3^9 = (\|\theta\|_3^3)^2\|\theta\|_3^3 \leq (\theta_{\max}\|\theta\|^2)^2(\theta_{\max}^2\|\theta\|_1) = o(\|\theta\|^4\|\theta\|_1)$, the right hand side is $o(\|\theta\|^8)$. This proves the claims of Y_5.

Consider Y_6. By definition and elementary calculations,

$$\begin{aligned}
Y_6 &= \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} + \sum_{i,j,k,\ell(dist)} \eta_j(\eta_i - \tilde{\eta}_i)\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} \\
&= 2\sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} \\
&= 2\sum_{i,j,k,\ell(dist)} \eta_i\Big(-\frac{1}{\sqrt{v}}\sum_{s\neq j}W_{js}\Big)\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} \\
&= -\frac{2}{\sqrt{v}}\sum_{j,s(dist)}\Big(\sum_{i,k,\ell(dist)\notin\{j\}}\eta_i\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\Big)W_{js}.
\end{aligned}$$

Here, to get the second line above, we relabeled $(i,j,k,\ell)=(j',i',\ell',k')$ in the second sum and found out the two sums are equal; the third line is from (77). We immediately see that

$$\mathrm{E}[Y_6]=0.$$

By (74) and (81),

$$\Big|\sum_{i,k,\ell(dist)\notin\{j\}}\eta_i\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\Big|\leq\sum_{i,k,\ell}C\theta_i\cdot\alpha^3\theta_j\theta_k^2\theta_\ell^2\theta_i\leq C\alpha^3\|\theta\|^6\theta_j.$$

It follows that

$$\begin{aligned}\mathrm{Var}(Y_6)&=\frac{8}{v}\sum_{j,s(dist)}\Big(\sum_{i,k,\ell(dist)\notin\{j\}}\eta_i\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\Big)^2\cdot\mathrm{Var}(W_{js})\\&\leq\frac{C}{\|\theta\|_1^2}\sum_{j,s}(\alpha^3\|\theta\|^6\theta_j)^2\theta_j\theta_s\\&\leq\frac{C\alpha^6\|\theta\|^{12}\|\theta\|_3^3}{\|\theta\|_1}.\end{aligned}$$

Since $\|\theta\|^4\leq\|\theta\|_1\|\theta\|_3^3$, the variance is bounded by $C\alpha^6\|\theta\|^8\|\theta\|_3^6$. This proves the claims of Y_6.

G.4.6. *Proof of Lemma G.6.* It suffices to prove the claims for each of Z_1 and Z_2; then, the claims of U_b follow immediately.

We first analyze Z_1. Plugging $\delta_{ij}=\eta_i(\eta_j-\tilde{\eta}_j)+\eta_j(\eta_i-\tilde{\eta}_i)$ into the definition of Z_1 gives

$$\begin{aligned}Z_1=&\sum_{i,j,k,\ell(dist)}\eta_i(\eta_j-\tilde{\eta}_j)\eta_j(\eta_k-\tilde{\eta}_k)W_{k\ell}W_{\ell i}+\sum_{i,j,k,\ell(dist)}\eta_i(\eta_j-\tilde{\eta}_j)^2\eta_kW_{k\ell}W_{\ell i}\\&+\sum_{i,j,k,\ell(dist)}(\eta_i-\tilde{\eta}_i)\eta_j^2(\eta_k-\tilde{\eta}_k)W_{k\ell}W_{\ell i}+\sum_{i,j,k,\ell(dist)}(\eta_i-\tilde{\eta}_i)\eta_j(\eta_j-\tilde{\eta}_j)\eta_kW_{k\ell}W_{\ell i}.\end{aligned}$$

In the last term above, if we relabel $(i,j,k,\ell)=(k',j',i',\ell')$, it becomes $\sum_{i',j',k',\ell'(dist)}(\eta_{k'}-\tilde{\eta}_{k'})\eta_{j'}(\eta_{j'}-\tilde{\eta}_{j'})\eta_{i'}W_{i'\ell'}W_{\ell'k'}$. This shows that the last sum equals to the first sum. Therefore,

$$\begin{aligned}Z_1=&\sum_{i,j,k,\ell(dist)}\eta_i(\eta_j-\tilde{\eta}_j)^2\eta_kW_{k\ell}W_{\ell i}\\&+2\sum_{i,j,k,\ell(dist)}\eta_i(\eta_j-\tilde{\eta}_j)\eta_j(\eta_k-\tilde{\eta}_k)W_{k\ell}W_{\ell i}\\&+\sum_{i,j,k,\ell(dist)}(\tilde{\eta}_i-\eta_i)\eta_j^2(\tilde{\eta}_k-\eta_k)W_{k\ell}W_{\ell i}\\\equiv&\, Z_{1a}+Z_{1b}+Z_{1c}.\end{aligned}\tag{94}$$

Below, we compute the means and variances of Z_{1a}-Z_{1c}.

First, we study Z_{1a}. When (i,j,k,ℓ) are distinct, $W_{k\ell}W_{\ell i}$ has a mean zero and is independent of $(\tilde{\eta}_j-\eta_j)^2$, so $\mathbb{E}[(\eta_j-\tilde{\eta}_j)^2W_{k\ell}W_{\ell i}]=0$. It follows that

$$\mathbb{E}[Z_{1a}]=0.$$

To bound the variance of Z_{1a}, we use (77) to re-write

$$Z_{1a}=\sum_{i,j,k,\ell(dist)}\eta_i\Big(-\frac{1}{\sqrt{v}}\sum_{s\neq j}W_{js}\Big)\Big(-\frac{1}{\sqrt{v}}\sum_{t\neq j}W_{jt}\Big)\eta_kW_{k\ell}W_{\ell i}$$

$$
\begin{aligned}
&= \frac{1}{v} \sum_{\substack{i,j,k,\ell(dist)\\ s,t\notin\{j\}}} \eta_i\eta_k W_{js}W_{jt}W_{k\ell}W_{\ell i} \\
&= \frac{1}{v} \sum_{\substack{i,j,k,\ell(dist)\\ s\notin\{j\}}} \eta_i\eta_k W_{js}^2 W_{k\ell}W_{\ell i} + \frac{1}{v} \sum_{\substack{i,j,k,\ell(dist)\\ s,t(dist)\notin\{j\}}} \eta_i\eta_k W_{js}W_{jt}W_{k\ell}W_{\ell i} \\
&\equiv \widetilde{Z}_{1a} + Z_{1a}^*.
\end{aligned}
$$

We first bound the variance of $\widetilde{Z}_{1a}$. It is seen that

$$
\mathrm{Var}(\widetilde{Z}_{1a}) = \frac{1}{v^2} \sum_{\substack{i,j,k,\ell(dist),s\notin\{j\}\\ i',j',k',\ell'(dist),s'\notin\{j'\}}} \eta_i\eta_k\eta_{i'}\eta_{k'} \cdot \mathbb{E}[W_{js}^2 W_{k\ell}W_{\ell i} \cdot W_{j's'}^2 W_{k'\ell'}W_{\ell' i'}].
$$

The summand is nonzero only if $\ell' = \ell$ and $\{k', i'\} = \{k, i\}$. We also note that, if we switch i' and k', the summand remains unchanged. So, it suffices to consider the case of $\ell' = \ell$ and $(k', i') = (k, i)$. By (81) and elementary calculations,

$$
\begin{aligned}
&\eta_i\eta_k\eta_{i'}\eta_{k'} \cdot \mathbb{E}[W_{js}^2 W_{k\ell}W_{\ell i} \cdot W_{j's'}^2 W_{k'\ell'}W_{\ell' i'}] \\
&= \begin{cases}
\eta_i^2\eta_k^2\, \mathbb{E}[W_{js}^4 W_{k\ell}^2 W_{\ell i}^2] \le C\theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_s, & \text{if } (\ell', k', i') = (\ell, k, i),\ \{j', s'\} = \{j, s\};\\
\eta_i^2\eta_k^2\, \mathbb{E}[W_{js}^2 W_{k\ell}^2 W_{\ell i}^2 W_{j's'}^2] \le C\theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_s\theta_{j'}\theta_{s'}, & \text{if } (\ell', k', i') = (\ell, k, i),\ \{j', s'\} \ne \{j, s\};\\
0, & \text{otherwise.}
\end{cases}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\mathrm{Var}(\widetilde{Z}_{1a}) &\le \frac{C}{\|\theta\|_1^4} \Big(\sum_{i,j,k,\ell,s} \theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_s + \sum_{i,j,k,\ell,s,j',s'} \theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_s\theta_{j'}\theta_{s'} \Big) \\
&\le \frac{C}{\|\theta\|_1^4} \big(\|\theta\|^2\|\theta\|_3^6\|\theta\|_1^2 + \|\theta\|^2\|\theta\|_3^6\|\theta\|_1^4 \big) \\
&\le C\|\theta\|^2\|\theta\|_3^6.
\end{aligned}
$$

We then bound the variance of Z_{1a}^*. Note that

$$
\begin{aligned}
&\eta_i\eta_k\eta_{i'}\eta_{k'} \cdot \mathbb{E}[W_{js}W_{jt}W_{k\ell}W_{\ell i} \cdot W_{j's'}W_{j't'}W_{k'\ell'}W_{\ell' i'}] \\
&= \begin{cases}
\eta_i^2\eta_k^2\, \mathbb{E}[W_{js}^2 W_{jt}^2 W_{k\ell}^2 W_{\ell i}^2] \le C\theta_i^3\theta_j^2\theta_k^3\theta_\ell^2\theta_s\theta_t, & \text{if } (j', \ell') = (j, \ell), \{s', t'\} = \{s, t\}, \{k', i'\} = \{k, i\};\\
\eta_i\eta_k\eta_s\eta_t\, \mathbb{E}[W_{js}^2 W_{jt}^2 W_{k\ell}^2 W_{\ell i}^2] \le C\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\theta_s^2\theta_t^2, & \text{if } (j', \ell') = (\ell, j), \{s', t'\} = \{k, i\}, \{k', i'\} = \{s, t\};\\
0, & \text{otherwise.}
\end{cases}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\mathrm{Var}(Z_{1a}^*) &\le \frac{C}{\|\theta\|_1^4} \Big(\sum_{i,j,k,\ell,s,t} \theta_i^3\theta_j^2\theta_k^3\theta_\ell^2\theta_s\theta_t + \sum_{i,j,k,\ell,s,t} \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\theta_s^2\theta_t^2 \Big) \\
&\le \frac{C}{\|\theta\|_1^4} \big(\|\theta\|^4\|\theta\|_3^6\|\theta\|_1^2 + \|\theta\|^{12} \big) \\
&\le \frac{C\|\theta\|^4\|\theta\|_3^6}{\|\theta\|_1^2},
\end{aligned}
$$

where the last inequality is because of $\|\theta\|^{12} = \|\theta\|^4(\|\theta\|^4)^2 \le \|\theta\|^4(\|\theta\|_1\|\theta\|_3^3)^2 = \|\theta\|^4\|\theta\|_3^6\|\theta\|_1^2$. Combining the above gives

$$
\mathrm{Var}(Z_{1a}) \le 2\mathrm{Var}(\widetilde{Z}_{1a}) + 2\mathrm{Var}(Z_{1a}^*) \le C\|\theta\|^2\|\theta\|_3^6. \tag{95}
$$

Second, we study Z_{1b}. Since $(\eta_j - \tilde{\eta}_j)$, $(\eta_k - \tilde{\eta}_k)W_{k\ell}$ and $W_{\ell i}$ are independent of each other, each summand in Z_{1b} has a zero mean. It follows that

$$\mathbb{E}[Z_{1b}] = 0.$$

We now compute its variance. By direct calculations,

$$\begin{aligned}
Z_{1b} &= 2\sum_{i,j,k,\ell(dist)} \eta_i\Big(-\frac{1}{\sqrt{v}}\sum_{s\neq j} W_{js}\Big)\eta_j\Big(-\frac{1}{\sqrt{v}}\sum_{t\neq k} W_{kt}\Big)W_{k\ell}W_{\ell i} \\
&= \frac{2}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq j, t\neq k}} \eta_i\eta_j W_{js}W_{kt}W_{k\ell}W_{\ell i} \\
&= \frac{2}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq j}} \eta_i\eta_j W_{js}W_{k\ell}^2 W_{\ell i} + \frac{2}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq j, t\notin\{k,\ell\}}} \eta_i\eta_j W_{js}W_{kt}W_{k\ell}W_{\ell i} \\
&\equiv \widetilde{Z}_{1b} + Z_{1b}^*.
\end{aligned}$$

We first bound the variance of $\widetilde{Z}_{1b}$. Note that

$$\mathrm{Var}(\widetilde{Z}_{1b}) = \frac{4}{v}\sum_{\substack{i,j,k,\ell(dist), s\neq j\\ i',j',k',\ell'(dist), s'\neq j'}} \eta_i\eta_j\eta_{i'}\eta_{j'}\cdot\mathbb{E}[W_{js}W_{k\ell}^2W_{\ell i}\cdot W_{j's'}W_{k'\ell'}^2W_{\ell' i'}].$$

For this summand to be nonzero, there are only two cases. In the first case, $(W_{js}, W_{\ell i})$ are paired with $(W_{j's'}, W_{\ell' i'})$. It follows that

$$\eta_i\eta_j\eta_{i'}\eta_{j'}\cdot\mathbb{E}[W_{js}W_{k\ell}^2W_{\ell i}W_{j's'}W_{k'\ell'}^2W_{\ell' i'}] = \eta_i\eta_j\eta_{i'}\eta_{j'}\cdot\mathbb{E}[W_{js}^2W_{k\ell}^2W_{\ell i}^2W_{k'\ell'}^2].$$

This happens only if (i) $\{j', s'\} = \{j, s\}$ and $\{\ell', i'\} = \{\ell, i\}$, or (ii) $\{j', s'\} = \{\ell, i\}$ and $\{\ell', i'\} = \{j, s\}$. By (81) and elementary calculations,

$$\begin{aligned}
&\eta_i\eta_j\eta_{i'}\eta_{j'}\cdot\mathbb{E}[W_{js}W_{k\ell}^2W_{\ell i}\cdot W_{j's'}W_{k'\ell'}^2W_{\ell' i'}] \\
&= \begin{cases}
\eta_i^2\eta_j^2\cdot\mathbb{E}[W_{js}^2W_{\ell i}^2W_{k\ell}^2W_{k'\ell}^2] \leq C\theta_i^3\theta_j^3\theta_k\theta_\ell^3\theta_s\theta_{k'}, & \text{if } (j',s') = (j,s), (\ell',i') = (\ell,i); \\
\eta_i\eta_j^2\eta_\ell\cdot\mathbb{E}[W_{js}^2W_{\ell i}^2W_{k\ell}^2W_{k'i}^2] \leq C\theta_i^3\theta_j^3\theta_k\theta_\ell^3\theta_s\theta_{k'}, & \text{if } (j',s') = (j,s), (\ell',i') = (i,\ell); \\
\eta_i^2\eta_j\eta_s\cdot\mathbb{E}[W_{js}^2W_{\ell i}^2W_{k\ell}^2W_{k'\ell}^2] \leq C\theta_i^3\theta_j^2\theta_k\theta_\ell^3\theta_s^2\theta_{k'}, & \text{if } (j',s') = (s,j), (\ell',i') = (\ell,i); \\
\eta_i\eta_j\eta_\ell\eta_s\cdot\mathbb{E}[W_{js}^2W_{\ell i}^2W_{k\ell}^2W_{k'i}^2] \leq C\theta_i^3\theta_j^2\theta_k\theta_\ell^3\theta_s^2\theta_{k'}, & \text{if } (j',s') = (s,j), (\ell',i') = (i,\ell); \\
\eta_i\eta_j\eta_\ell\eta_s\cdot\mathbb{E}[W_{js}^2W_{\ell i}^2W_{k\ell}^2W_{k'j}^2] \leq C\theta_i^2\theta_j^3\theta_k\theta_\ell^3\theta_s^2\theta_{k'}, & \text{if } (j',s') = (\ell,i), (\ell',i') = (j,s); \\
\eta_i\eta_j^2\eta_\ell\cdot\mathbb{E}[W_{js}^2W_{\ell i}^2W_{k\ell}^2W_{k's}^2] \leq C\theta_i^2\theta_j^3\theta_k\theta_\ell^3\theta_s^2\theta_{k'}, & \text{if } (j',s') = (\ell,i), (\ell',i') = (s,j); \\
\eta_i^2\eta_j\eta_s\cdot\mathbb{E}[W_{js}^2W_{\ell i}^2W_{k\ell}^2W_{k'j}^2] \leq C\theta_i^3\theta_j^3\theta_k\theta_\ell^2\theta_s^2\theta_{k'}, & \text{if } (j',s') = (i,\ell), (\ell',i') = (j,s); \\
\eta_i^2\eta_j^2\cdot\mathbb{E}[W_{js}^2W_{\ell i}^2W_{k\ell}^2W_{k's}^2] \leq C\theta_i^3\theta_j^3\theta_k\theta_\ell^2\theta_s^2\theta_{k'}, & \text{if } (j',s') = (i,\ell), (\ell',i') = (s,j); \\
0, & \text{otherwise.}
\end{cases}
\end{aligned}$$

The upper bound on the right hand side only has two types $C\theta_i^3\theta_j^3\theta_k\theta_\ell^3\theta_s\theta_{k'}$ and $C\theta_i^3\theta_j^2\theta_k\theta_\ell^3\theta_s^2\theta_{k'}$. The contribution of this case to $\mathrm{Var}(\widetilde{Z}_{1b})$ is

$$\begin{aligned}
&\leq \frac{C}{v^2}\Big(\sum_{i,j,k,\ell,s,k'} \theta_i^3\theta_j^3\theta_k\theta_\ell^3\theta_s\theta_{k'} + \sum_{i,j,k,\ell,s,k'} \theta_i^3\theta_j^2\theta_k\theta_\ell^3\theta_s^2\theta_{k'}\Big) \\
&\leq \frac{C}{\|\theta\|_1^4}\big(\|\theta\|_3^9\|\theta\|_1^3 + \|\theta\|^4\|\theta\|_3^6\|\theta\|_1^2\big)
\end{aligned}$$

$$\leq \frac{C\|\theta\|_3^9}{\|\theta\|_1}.$$

In the second case, $\{W_{js}, W_{k\ell}, W_{\ell i}\}$ and $\{W_{j's'}, W_{k'\ell'}, W_{\ell' i'}\}$ are two sets of same variables. Then,

$$\eta_i\eta_j\eta_{i'}\eta_{j'} \cdot \mathbb{E}[W_{js}W_{k\ell}^2 W_{\ell i}W_{j's'}W_{k'\ell'}^2 W_{\ell' i'}] = \eta_i\eta_j\eta_{i'}\eta_{j'} \cdot \mathbb{E}[W_{js}^3 W_{k\ell}^3 W_{\ell i}^3].$$

This can only happen if $\ell' = \ell$, $\{i', k'\} = \{i, k\}$, and $\{j', s'\} = \{j, s\}$. By (81) and elementary calculations,

$$\eta_i\eta_j\eta_{i'}\eta_{j'} \cdot \mathbb{E}[W_{js}W_{k\ell}^2 W_{\ell i} \cdot W_{j's'}W_{k'\ell'}^2 W_{\ell' i'}]$$

$$= \begin{cases} \eta_i^2\eta_j^2 \cdot \mathbb{E}[W_{js}^3 W_{\ell i}^3 W_{k\ell}^3] \leq C\theta_i^3\theta_j^3\theta_k\theta_\ell^2\theta_s, & \text{if } \ell' = \ell,\ (i', k') = (i, k),\ (j', s') = (j, s); \\ \eta_i^2\eta_j\eta_s \cdot \mathbb{E}[W_{js}^3 W_{\ell i}^3 W_{k\ell}^3] \leq C\theta_i^3\theta_j^2\theta_k\theta_\ell^2\theta_s^2, & \text{if } \ell' = \ell,\ (i', k') = (i, k),\ (j', s') = (s, j); \\ \eta_i\eta_k\eta_j^2 \cdot \mathbb{E}[W_{js}^3 W_{\ell i}^3 W_{k\ell}^3] \leq C\theta_i^2\theta_j^3\theta_k^2\theta_\ell^2\theta_s, & \text{if } \ell' = \ell,\ (i', k') = (k, i),\ (j', s') = (j, s); \\ \eta_i\eta_k\eta_j\eta_s \cdot \mathbb{E}[W_{js}^3 W_{\ell i}^3 W_{k\ell}^3] \leq C\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\theta_s^2, & \text{if } \ell' = \ell,\ (i', k') = (i, k),\ (j', s') = (s, j); \\ 0, & \text{otherwise.} \end{cases}$$

The upper bound on the right hand side has three types, and the contribution of this case to $\mathrm{Var}(\widetilde{Z}_{1b})$ is

$$\leq \frac{C}{v^2}\Big(\sum_{i,j,k,\ell,s} \theta_i^3\theta_j^3\theta_k\theta_\ell^2\theta_s + \sum_{i,j,k,\ell,s} \theta_i^3\theta_j^2\theta_k\theta_\ell^2\theta_s^2 + \sum_{i,j,k,\ell,s} \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\theta_s^2\Big)$$

$$\leq \frac{C}{\|\theta\|_1^4}\big(\|\theta\|^2\|\theta\|_3^6\|\theta\|_1^2 + \|\theta\|^6\|\theta\|_3^3\|\theta\|_1 + \|\theta\|^{10}\big)$$

$$\leq \frac{C\|\theta\|^2\|\theta\|_3^6}{\|\theta\|_1^2},$$

where we use $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$ (Cauchy-Schwarz) in the last line. It is seen that the contribution of the first case is dominating, and so

$$\mathrm{Var}(\widetilde{Z}_{1b}) \leq \frac{C\|\theta\|_3^9}{\|\theta\|_1}.$$

We then bound the variance of Z_{1b}^*. Note that

$$\mathrm{Var}(Z_{1b}^*) = \frac{4}{v^2} \sum_{\substack{i,j,k,\ell(dist), s\neq j, t\notin\{k,\ell\} \\ i',j',k',\ell'(dist), s'\neq j', t'\notin\{k',\ell'\}}} \eta_i\eta_j\eta_{i'}\eta_{j'} \cdot \mathbb{E}[W_{js}W_{kt}W_{k\ell}W_{\ell i} \cdot W_{j's'}W_{k't'}W_{k'\ell'}W_{\ell' i'}].$$

For the summand to be nonzero, all W terms have to be perfectly matched, so that the expectation in the summand becomes

$$\mathbb{E}[W_{js}W_{kt}W_{k\ell}W_{\ell i} \cdot W_{j's'}W_{k't'}W_{k'\ell'}W_{\ell' i'}] = \mathbb{E}[W_{js}^2 W_{kt}^2 W_{k\ell}^2 W_{\ell i}^2] \leq C\theta_i\theta_j\theta_k^2\theta_\ell^2\theta_s\theta_t.$$

For this perfect match to happen, we need $(t', k', \ell', i') = (t, k, \ell, i)$ or $(t', k', \ell', i') = (i, \ell, k, t)$, as well as $\{j', s'\} = \{j, s\}$. This implies that, i' can only take values in $\{i, t\}$ and j' can only take values in $\{j, s\}$. It follows that $\eta_i\eta_j\eta_{i'}\eta_{j'}$ belongs to one of the following cases:

$$\eta_i\eta_j(\eta_i\eta_j) \leq C\theta_i^2\theta_j^2, \qquad \eta_i\eta_j(\eta_i\eta_s) = C\theta_i^2\theta_j\theta_s,$$

$$\eta_i\eta_j(\eta_t\eta_j) \leq C\theta_i\theta_j^2\theta_t, \qquad \eta_i\eta_j(\eta_t\eta_s) \leq C\theta_i\theta_j\theta_t\theta_s.$$

Combining the above gives

$$
\begin{aligned}
\mathrm{Var}(Z_{1b}^*) &\le \frac{C}{v^2} \sum_{i,j,k,\ell,s,t} (\theta_i^2\theta_j^2 + \theta_i^2\theta_j\theta_s + \theta_i\theta_j^2\theta_t + \theta_i\theta_j\theta_t\theta_s) \cdot \theta_i\theta_j\theta_k^2\theta_\ell^2\theta_s\theta_t \\
&\le \frac{C}{\|\theta\|_1^4} \big(\|\theta\|^4\|\theta\|_3^6\|\theta\|_1^2 + 2\|\theta\|^8\|\theta\|_3^3\|\theta\|_1 + \|\theta\|^{12} \big) \\
&\le \frac{C\|\theta\|^4\|\theta\|_3^6}{\|\theta\|_1^2}.
\end{aligned}
$$

We combine the variances of $\widetilde{Z}_{1b}$ and Z_{1b}^*. Since $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$, the variance of $\widetilde{Z}_{1b}$ dominates. It follows that

$$
\mathrm{Var}(Z_{1b}) \le 2\mathrm{Var}(\widetilde{Z}_{1b}) + 2\mathrm{Var}(Z_{1b}^*) \le \frac{C\|\theta\|_3^9}{\|\theta\|_1}. \tag{96}
$$

Third, we study Z_{1c}. It is seen that

$$
\begin{aligned}
Z_{1c} &= \sum_{i,j,k,\ell (dist)} \Big(-\frac{1}{\sqrt{v}} \sum_{s\neq i} W_{is} \Big) \eta_j^2 \Big(-\frac{1}{\sqrt{v}} \sum_{t\neq k} W_{kt} \Big) W_{k\ell} W_{\ell i} \\
&= \frac{1}{v} \sum_{\substack{i,k,\ell(dist)\\ s\neq i, t\neq k}} \Big(\sum_{j\notin\{i,k,\ell\}} \eta_j^2 \Big) W_{is}W_{kt}W_{k\ell}W_{\ell i} \\
&\equiv \frac{1}{v} \sum_{\substack{i,k,\ell(dist)\\ s\neq i, t\neq k}} \beta_{ik\ell} W_{is}W_{kt}W_{k\ell}W_{\ell i},
\end{aligned}
$$

where

$$
\beta_{ik\ell} \equiv \sum_{j\notin\{i,k,\ell\}} \eta_j^2 \le C \sum_j \theta_j^2 \le C\|\theta\|^2. \tag{97}
$$

We divide all summands into four groups: (i) $s = t = \ell$; (ii) $s = \ell$, $t \neq \ell$; (iii) $s \neq \ell$, $t = \ell$; (iv) $s \neq \ell$, $t \neq \ell$. It yields that

$$
\begin{aligned}
Z_{1c} = {} & \frac{1}{v} \sum_{i,k,\ell(dist)} \beta_{ik\ell} W_{k\ell}^2 W_{\ell i}^2 + \frac{1}{v} \sum_{\substack{i,k,\ell(dist)\\ t\neq\{k,\ell\}}} \beta_{ik\ell} W_{kt} W_{k\ell} W_{\ell i}^2 \\
& + \frac{1}{v} \sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\}}} \beta_{ik\ell} W_{is} W_{k\ell}^2 W_{\ell i} + \frac{1}{v} \sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\}, t\notin\{k,\ell\}}} \beta_{ik\ell} W_{is}W_{kt}W_{k\ell}W_{\ell i}.
\end{aligned}
$$

In the third sum, if we relabel $(i,k,\ell,s) = (k',i',\ell',t')$, it has the form $\sum_{i',k',\ell'(dist), t'\notin\{k',\ell'\}} \beta_{k'i'\ell'} W_{k't'} W_{i'\ell'}^2 W_{\ell' k'}$. This shows that this sum equals to the second sum. We thus have

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

Among all three terms, only $\widetilde{Z}_{1c}$ has a nonzero mean. It follows that

$$\mathbb{E}[Z_{1c}] = \mathbb{E}[\widetilde{Z}_{1c}] = \frac{1}{v}\sum_{i,k,\ell(dist)} \beta_{ik\ell}\Omega_{k\ell}(1-\Omega_{k\ell})\Omega_{\ell i}(1-\Omega_{\ell i})$$
$$= \frac{1}{v}\sum_{i,k,\ell(dist)} \beta_{ik\ell}\Omega_{k\ell}\Omega_{\ell i}[1+O(\theta_{\max}^2)].$$

Under the null hypothesis, $\Omega_{ij}=\theta_i\theta_j$. It follows that $\eta_j = \frac{\theta_j}{\sqrt{v}}\sum_{i:i\neq j}\theta_i = [1+o(1)]\frac{\theta_j\|\theta\|_1}{\sqrt{v}}$ and that $\beta_{ik\ell}=\sum_{j\notin\{i,k,\ell\}}\eta_j^2 = [1+o(1)]\frac{\|\theta\|_1^2}{v}\sum_{j\notin\{i,k,\ell\}}\theta_j^2 = [1+o(1)]\frac{\|\theta\|_1^2\|\theta\|^2}{v}$. Additionally, $v=\sum_{i\neq j}\theta_i\theta_j = \|\theta\|_1^2\cdot[1+o(1)]$. As a result,

$$\mathbb{E}[Z_{1c}] = \frac{1}{v}\sum_{i,k,\ell(dist)}[1+o(1)]\frac{\|\theta\|_1^2\|\theta\|^2}{v}\cdot\theta_k\theta_\ell^2\theta_i$$
$$= [1+o(1)]\cdot\frac{\|\theta\|_1^2\|\theta\|^2}{v^2}\sum_{i,k,\ell(dist)}\theta_k\theta_\ell^2\theta_i$$
$$= [1+o(1)]\cdot\frac{\|\theta\|_1^2\|\theta\|^2}{\|\theta\|_1^4}\big[\|\theta\|_1^2\|\theta\|^2 - O(\|\theta\|^4+\|\theta\|_1\|\theta\|_3^3)\big]$$
$$= [1+o(1)]\cdot\|\theta\|^4, \tag{98}$$

where in the last line we have used $\|\theta\|^2=o(\|\theta\|_1)$, $\|\theta\|_3^3=o(\|\theta\|_1)$ and $\|\theta\|_1\to\infty$. We then bound the variance of Z_{1c} by studying the variance of each of the three variables, $\widetilde{Z}_{1c}$, Z_{1c}^* and $Z_{1c}^\dagger$. Consider $\widetilde{Z}_{1c}$ first. For $W_{k\ell}^2W_{\ell i}^2$ and $W_{k'\ell'}^2W_{\ell' i'}^2$ to be correlated, it has to be the case of either $\{k',\ell'\}=\{k,\ell\}$ or $\{i',\ell'\}=\{i,\ell\}$. By symmetry between k and i in the expression, it suffices to consider $\{k',\ell'\}=\{k,\ell\}$. Direct calculations show that

$$\mathrm{Cov}(W_{k\ell}^2W_{\ell i}^2,\, W_{k'\ell'}^2W_{\ell' i'}^2) \le \begin{cases} \mathbb{E}[W_{k\ell}^4W_{\ell i}^4] \le C\theta_k\theta_\ell^2\theta_i, & \text{if } (k',\ell')=(k,\ell),\ i'=i;\\ \mathbb{E}[W_{k\ell}^4W_{\ell i}^2W_{\ell i'}^2] \le C\theta_k\theta_\ell^3\theta_i\theta_{i'}, & \text{if } (k',\ell')=(k,\ell),\ i'\neq i;\\ \mathbb{E}[W_{k\ell}^4W_{\ell i}^2W_{ki}^2] \le C\theta_k^2\theta_\ell^2\theta_i^2, & \text{if } (k',\ell')=(\ell,k),\ i'=i;\\ \mathbb{E}[W_{k\ell}^4W_{\ell i}^2W_{ki'}^2] \le C\theta_k^2\theta_\ell^2\theta_i\theta_{i'}, & \text{if } (k',\ell')=(\ell,k),\ i'\neq i;\\ 0, & \text{otherwise.}\end{cases}$$

Combining it with (97) and the fact of $v\ge C^{-1}\|\theta\|_1^2$, we have

$$\mathrm{Var}(\widetilde{Z}_{1c}) \le \frac{C\|\theta\|^4}{\|\theta\|_1^4}\Big(\sum_{i,k,\ell}\theta_k\theta_\ell^2\theta_i + \sum_{i,k,\ell,i'}\theta_k\theta_\ell^3\theta_i\theta_{i'} + \sum_{i,k,\ell}\theta_k^2\theta_\ell^2\theta_i^2 + \sum_{i,k,\ell,i'}\theta_k^2\theta_\ell^2\theta_i\theta_{i'}\Big)$$
$$\le \frac{C\|\theta\|^4}{\|\theta\|_1^4}\big(\|\theta\|^2\|\theta\|_1^2 + \|\theta\|_3^3\|\theta\|_1^3 + \|\theta\|^6 + \|\theta\|^4\|\theta\|_1^2\big)$$
$$\le \frac{C\|\theta\|^4\|\theta\|_3^3}{\|\theta\|_1}.$$

Consider Z_{1c}^*. By direct calculations,

$$\mathbb{E}[W_{kt}W_{k\ell}W_{\ell i}^2W_{k't'}W_{k'\ell'}W_{\ell' i'}^2]$$

$$
= \begin{cases} \mathbb{E}[W_{kt}^2 W_{k\ell}^2 W_{\ell i}^4] \leq C\theta_i\theta_k^2\theta_\ell^2\theta_t, & \text{if } (k',t',\ell') = (k,t,\ell),\ i = i'; \\ \mathbb{E}[W_{kt}^2 W_{k\ell}^2 W_{\ell i}^2 W_{\ell i'}^2] \leq C\theta_i\theta_k^2\theta_\ell^3\theta_t\theta_{i'}, & \text{if } (k',t',\ell') = (k,t,\ell),\ i \neq i'; \\ \mathbb{E}[W_{kt}^2 W_{k\ell}^2 W_{\ell i}^2 W_{ti'}^2] \leq C\theta_i\theta_k^2\theta_\ell^2\theta_t^2\theta_{i'}, & \text{if } (k',t',\ell') = (k,\ell,t); \\ \mathbb{E}[W_{kt}^3 W_{k\ell}^2 W_{\ell i}^3] \leq C\theta_i\theta_k^2\theta_\ell^2\theta_t, & \text{if } (k',t',\ell',i') = (\ell,i,k,t); \\ 0, & \text{otherwise.} \end{cases}
$$

We combine it with (97) and find that

$$
\begin{aligned}
\mathrm{Var}(Z_{1c}^*) &= \frac{4}{v^2} \sum_{\substack{i,k,\ell(dist),t\neq\{k,\ell\} \\ i',k',\ell'(dist),t'\neq\{k',\ell'\}}} \beta_{ik\ell}\beta_{i'k'\ell'} \cdot \mathbb{E}[W_{kt}W_{k\ell}W_{\ell i}^2 W_{k't'}W_{k'\ell'}W_{\ell' i'}^2] \\
&\leq \frac{C\|\theta\|^4}{\|\theta\|_1^4}\Big(\sum_{i,k,\ell,t} \theta_i\theta_k^2\theta_\ell^2\theta_t + \sum_{i,k,\ell,t,i'} \theta_i\theta_k^2\theta_\ell^3\theta_t\theta_{i'} + \sum_{i,k,\ell,t,i'} \theta_i\theta_k^2\theta_\ell^2\theta_t^2\theta_{i'} \Big) \\
&\leq \frac{C\|\theta\|^4}{\|\theta\|_1^4}\big(\|\theta\|^4\|\theta\|_1^2 + \|\theta\|^2\|\theta\|_3^3\|\theta\|_1^3 + \|\theta\|^6\|\theta\|_1^2 \big) \\
&\leq \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1}.
\end{aligned}
$$

Consider $Z_{1c}^\dagger$. Re-write

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

Regarding the first term, by direct calculations,

$$
\begin{aligned}
&\mathbb{E}[W_{ik}^2 W_{k\ell}W_{\ell i} \cdot W_{i'k'}^2 W_{k'\ell'}W_{\ell' i'}] \\
&= \begin{cases} \mathbb{E}[W_{ik}^4 W_{k\ell}^2 W_{\ell i}^2] \leq C\theta_i^2\theta_k^2\theta_\ell^2, & \text{if } \ell' = \ell,\ \{i',k'\} = \{i,k\}; \\ \mathbb{E}[W_{ik}^3 W_{k\ell}^2 W_{\ell i}^3] \leq C\theta_i^2\theta_k^2\theta_\ell^2, & \text{if } (\ell',k') = (k,\ell),\ i' = i; \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}
$$

Combining it with (97) gives

$$
\mathrm{Var}\Big(\frac{1}{v}\sum_{i,k,\ell(dist)} \beta_{ik\ell}W_{ik}^2 W_{k\ell}W_{\ell i}\Big) \leq \frac{C\|\theta\|^4}{\|\theta\|_1^4}\sum_{i,j,k,\ell} \theta_i^2\theta_k^2\theta_\ell^2 \leq \frac{C\|\theta\|^{10}}{\|\theta\|_1^4}.
$$

Regarding the second term, for $W_{is}W_{kt}W_{k\ell}W_{\ell i}$ and $W_{i's'}W_{k't'}W_{k'\ell'}W_{\ell' i'}$ to be correlated, all W terms have to be perfectly matched. For each fixed (i,k,ℓ,s,t), there are only a constant number of (i',k',ℓ',s',t') so that the above is satisfied. Mimicking the argument in (70), we have

$$
\begin{aligned}
\mathrm{Var}\Big(\frac{1}{v}\sum_{\substack{i,k,\ell(dist) \\ s\notin\{i,\ell\},t\notin\{k,\ell\} \\ (s,t)\neq(k,i)}} \beta_{ik\ell}W_{is}W_{kt}W_{k\ell}W_{\ell i}\Big) &\leq \frac{C}{v^2}\sum_{\substack{i,k,\ell(dist) \\ s\notin\{i,\ell\},t\notin\{k,\ell\} \\ (s,t)\neq(k,i)}} \beta_{ik\ell}^2 \cdot \mathrm{Var}(W_{is}W_{kt}W_{k\ell}W_{\ell i}) \\
&\leq \frac{C}{\|\theta\|_1^4}\sum_{i,k,\ell,s,t} \|\theta\|^4 \cdot \theta_i^2\theta_k^2\theta_\ell^2\theta_s\theta_t \leq \frac{C\|\theta\|^{10}}{\|\theta\|_1^2}.
\end{aligned}
$$

It follows that

$$\mathrm{Var}(Z_{1c}^{\dagger}) \leq \frac{C\|\theta\|^{10}}{\|\theta\|_1^2}.$$

Combining the above results and noticing that $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$, we immediately have

$$\mathrm{Var}(Z_{1c}) \leq 3\mathrm{Var}(\widetilde{Z}_{1c}) + 3\mathrm{Var}(Z_{1c}^{*}) + 3\mathrm{Var}(Z_{1c}^{\dagger}) \leq \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1}. \tag{99}$$

We now combine (95), (96), (98), and (99). Since $Z_1 = Z_{1a} + Z_{1b} + Z_{1c}$, it follows that

$$\mathbb{E}[Z_1] = \|\theta\|^4 \cdot [1 + o(1)], \qquad \mathrm{Var}(Z_1) \leq C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8).$$

This proves the claims of Z_1.

Next, we analyze Z_2. Since $\delta_{ij} = \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_i - \tilde{\eta}_i)$, by direct calculations,

$$\begin{aligned} Z_2 = &\sum_{i,j,k,\ell (dist)} \eta_i(\eta_j - \tilde{\eta}_j)W_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)W_{\ell i} + \sum_{i,j,k,\ell (dist)} \eta_i(\eta_j - \tilde{\eta}_j)W_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell W_{\ell i} \\ &+ \sum_{i,j,k,\ell (dist)} (\eta_i - \tilde{\eta}_i)\eta_j W_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)W_{\ell i} + \sum_{i,j,k,\ell (dist)} (\eta_i - \tilde{\eta}_i)\eta_j W_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell W_{\ell i}. \end{aligned}$$

By relabeling the indices, we find out that the first and last sums are equal and that the second and third sums are equal. It follows that

$$\begin{aligned} Z_2 &= 2\sum_{i,j,k,\ell (dist)} \eta_i(\eta_j - \tilde{\eta}_j)W_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)W_{\ell i} \\ &\quad + 2\sum_{i,j,k,\ell (dist)} \eta_i(\eta_j - \tilde{\eta}_j)W_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell W_{\ell i} \\ &\equiv Z_{2a} + Z_{2b}. \end{aligned} \tag{100}$$

First, we study Z_{2a}. It is seen that

$$\begin{aligned} Z_{2a} &= 2\sum_{i,j,k,\ell (dist)} \eta_i\Big(-\frac{1}{\sqrt{v}}\sum_{s\neq j} W_{js}\Big)W_{jk}\eta_k\Big(-\frac{1}{\sqrt{v}}\sum_{t\neq \ell} W_{\ell t}\Big)W_{\ell i} \\ &= \frac{2}{v}\sum_{\substack{i,j,k,\ell (dist) \\ s\neq j, t\neq \ell}} \eta_i\eta_k W_{js}W_{jk}W_{\ell t}W_{\ell i}. \end{aligned}$$

We divide summands into four groups: (i) $s = k$ and $t = i$, (ii) $s = k$ and $t \neq i$, (iii) $s \neq k$ and $t = i$, (iv) $s \neq k$ and $t \neq i$. By symmetry between (j, k, s) and (ℓ, i, t), the sum of group (ii) and group (iii) are equal. We end up with

$$\begin{aligned} Z_{2a} &= \frac{2}{v}\sum_{i,j,k,\ell (dist)} \eta_i\eta_k W_{jk}^2 W_{\ell i}^2 + \frac{4}{v}\sum_{\substack{i,j,k,\ell (dist) \\ s\notin\{j,k\}}} \eta_i\eta_k W_{js}W_{jk}W_{\ell i}^2 \\ &\quad + \frac{2}{v}\sum_{\substack{i,j,k,\ell (dist) \\ s\notin\{j,k\}, t\notin\{\ell,i\}}} \eta_i\eta_k W_{js}W_{jk}W_{\ell t}W_{\ell i} \\ &\equiv \widetilde{Z}_{2a} + Z_{2a}^{*} + Z_{2a}^{\dagger}, \end{aligned}$$

Only $\widetilde{Z}_{2a}$ has a nonzero mean. It follows that

$$\mathbb{E}[Z_{2a}] = \mathbb{E}[\widetilde{Z}_{2a}] = \frac{2}{v}\sum_{i,j,k,\ell (dist)} \eta_i\eta_k\Omega_{jk}(1 - \Omega_{jk})\Omega_{\ell i}(1 - \Omega_{\ell i}).$$

Under the null hypothesis, $\Omega_{ij} = \theta_i\theta_j$. Hence, $\Omega_{jk}(1-\Omega_{jk})\Omega_{\ell i}(1-\Omega_{\ell i}) = \theta_j\theta_k\theta_\ell\theta_i \cdot [1 + O(\theta_{\max}^2)]$. Additionally, in the proof of (98), we have seen that $v = [1+o(1)]\cdot\|\theta\|_1^2$ and $\eta_j = [1+o(1)]\cdot\theta_j$. Combining these results gives

$$
\begin{aligned}
\mathbb{E}[Z_{2a}] &= \frac{2[1+o(1)]}{\|\theta\|_1^2}\sum_{i,j,k,\ell (dist)}(\theta_i\theta_k)(\theta_j\theta_k\theta_\ell\theta_i) \\
&= \frac{2[1+o(1)]}{\|\theta\|_1^2}\Big[\sum_{i,j,k,\ell}\theta_i^2\theta_j\theta_k^2\theta_\ell - \sum_{\substack{i,j,k,\ell \\ \text{(not dist)}}}\theta_i^2\theta_j\theta_k^2\theta_\ell\Big] \\
&= \frac{2[1+o(1)]}{\|\theta\|_1^2}\Big[\|\theta\|^4\|\theta\|_1^2 - O\big(\|\theta\|_4^4\|\theta\|_1^2 + \|\theta\|_3^3\|\theta\|^2\|\theta\|_1 + \|\theta\|^6\big)\Big] \\
&= \frac{2[1+o(1)]}{\|\theta\|_1^2}\cdot\|\theta\|^4\|\theta\|_1^2[1+o(1)] \\
&= [1+o(1)]\cdot 2\|\theta\|^4. \qquad (101)
\end{aligned}
$$

We then bound the variance of Z_a. Consider $\widetilde{Z}_{2a}$ first. Note that $W_{jk}^2W_{\ell i}^2$ and $W_{j'k'}^2W_{\ell' i'}^2$ are correlated only if either $\{j',k'\} = \{j,k\}$ or $\{j',k'\} = \{\ell, i\}$. By symmetry, it suffices to consider $\{j',k'\} = \{j,k\}$. Direct calculations show that

$$
\begin{aligned}
&\mathrm{Cov}(\eta_i\eta_k W_{jk}^2W_{\ell i}^2,\, \eta_{i'}\eta_{k'}W_{j'k'}^2W_{\ell' i'}^2) \\
&\le \begin{cases}
\eta_k^2\eta_i^2\mathbb{E}[W_{jk}^4W_{\ell i}^4] \le C\theta_i^3\theta_j\theta_k^3\theta_\ell, & \text{if } (j',k') = (j,k),\ i = i',\ \ell = \ell'; \\
\eta_k^2\eta_i^2\mathbb{E}[W_{jk}^4W_{\ell i}^2W_{\ell' i}^2] \le C\theta_i^4\theta_j\theta_k^3\theta_\ell\theta_{\ell'}, & \text{if } (j',k') = (j,k),\ i = i',\ \ell \neq \ell'; \\
\eta_k^2\eta_i\eta_{i'}\mathbb{E}[W_{jk}^4W_{\ell i}^2W_{\ell' i'}^2] \le C\theta_i^2\theta_j\theta_k^3\theta_\ell\theta_{i'}^2\theta_{\ell'}, & \text{if } (j',k') = (j,k),\ i \neq i'; \\
\eta_j\eta_k\eta_i^2\mathbb{E}[W_{jk}^4W_{\ell i}^4] \le C\theta_i^3\theta_j^2\theta_k^2\theta_\ell, & \text{if } (j',k') = (k,j),\ i = i',\ \ell = \ell'; \\
\eta_j\eta_k\eta_i^2\mathbb{E}[W_{jk}^4W_{\ell i}^2W_{\ell' i}^2] \le C\theta_i^4\theta_j^2\theta_k^2\theta_\ell\theta_{\ell'}, & \text{if } (j',k') = (k,j),\ i = i',\ \ell \neq \ell'; \\
\eta_j\eta_k\eta_i\eta_{i'}\mathbb{E}[W_{jk}^4W_{\ell i}^2W_{\ell' i'}^2] \le C\theta_i^2\theta_j^2\theta_k^2\theta_\ell\theta_{i'}^2\theta_{\ell'}, & \text{if } (j',k') = (k,j),\ i \neq i'; \\
0, & \text{otherwise.}
\end{cases}
\end{aligned}
$$

As a result,

$$
\begin{aligned}
\mathrm{Var}(\widetilde{Z}_{2a}) &= \frac{4}{v^2}\sum_{\substack{i,j,k,\ell (dist) \\ i',j',k',\ell' (dist)}}\mathrm{Cov}(\eta_i\eta_k W_{jk}^2W_{\ell i}^2,\, \eta_{i'}\eta_{k'}W_{j'k'}^2W_{\ell' i'}^2) \\
&\le \frac{C}{\|\theta\|_1^4}\big(\|\theta\|_3^6\|\theta\|_1^2 + \|\theta\|_4^4\|\theta\|_3^3\|\theta\|_1^3 + \|\theta\|_3^3\|\theta\|^4\|\theta\|_1^3 \\
&\qquad + \|\theta\|_3^3\|\theta\|^4\|\theta\|_1 + \|\theta\|_4^4\|\theta\|^4\|\theta\|_1^2 + \|\theta\|^8\|\theta\|_1^2\big) \\
&\le \frac{C\|\theta\|^4\|\theta\|_3^3}{\|\theta\|_1},
\end{aligned}
$$

where the last line is obtained as follows: There are six terms in the brackets; since $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$, the last three terms are dominated by the first three terms; for the first three terms, since $\|\theta\|_3^3 \le \theta_{\max}^2\|\theta\|_1 = o(\|\theta\|_1)$ and $\|\theta\|_4^4 \le \theta_{\max}^2\|\theta\|^2 = o(\|\theta\|^2)$, the third term dominates. Consider Z_{2a}^* next. We note that for

$$\mathbb{E}[W_{js}W_{jk}W_{\ell i}^2 \cdot W_{j's'}W_{j'k'}W_{\ell' i'}^2]$$

to be nonzero, it has to be the case of either $(W_{j's'}, W_{j'k'}) = (W_{js}, W_{jk})$ or $(W_{j's'}, W_{j'k'}) = (W_{jk}, W_{js})$. This can only happen if $(j', s', k') = (j, s, k)$ or $(j', s', k') = (j, k, s)$. By elementary calculations,

$$\eta_i\eta_k\eta_{i'}\eta_{k'}\cdot\mathbb{E}[W_{js}W_{jk}W_{\ell i}^2 \cdot W_{j's'}W_{j'k'}W_{\ell' i'}^2]$$

$$
= \begin{cases} \eta_i^2\eta_k^2\,\mathbb{E}[W_{js}^2W_{jk}^2W_{\ell i}^4] \leq C\theta_i^3\theta_j^2\theta_k^3\theta_\ell\theta_s, & \text{if } (j',s',k')=(j,s,k),\ i'=i,\ \ell'=\ell; \\ \eta_i^2\eta_k^2\,\mathbb{E}[W_{js}^2W_{jk}^2W_{\ell i}^2W_{\ell' i}^2] \leq C\theta_i^4\theta_j^2\theta_k^3\theta_\ell\theta_s\theta_{\ell'}, & \text{if } (j',s',k')=(j,s,k),\ i'=i,\ \ell'\neq\ell; \\ \eta_i\eta_{i'}\eta_k^2\,\mathbb{E}[W_{js}^2W_{jk}^2W_{\ell i}^2W_{\ell' i'}^2] \leq C\theta_i^2\theta_j^2\theta_k^3\theta_\ell\theta_s\theta_{i'}^2\theta_{\ell'}, & \text{if } (j',s',k')=(j,s,k),\ i\neq i'; \\ \eta_i^2\eta_k\eta_s\,\mathbb{E}[W_{js}^2W_{jk}^2W_{\ell i}^4] \leq C\theta_i^3\theta_j^2\theta_k^2\theta_\ell\theta_s^2, & \text{if } (j',s',k')=(j,k,s),\ i'=i,\ \ell'=\ell; \\ \eta_i^2\eta_k\eta_s\,\mathbb{E}[W_{js}^2W_{jk}^2W_{\ell i}^2W_{\ell' i}^2] \leq C\theta_i^4\theta_j^2\theta_k^2\theta_\ell\theta_s^2\theta_{\ell'}, & \text{if } (j',s',k')=(j,k,s),\ i'=i,\ \ell'\neq\ell; \\ \eta_i\eta_{i'}\eta_k\eta_s\,\mathbb{E}[W_{js}^2W_{jk}^2W_{\ell i}^2W_{\ell' i'}^2] \leq C\theta_i^2\theta_j^2\theta_k^2\theta_\ell\theta_s^2\theta_{i'}^2\theta_{\ell'}, & \text{if } (j',s',k')=(j,k,s),\ i\neq i'; \\ 0, & \text{otherwise.} \end{cases}
$$

It follows that

$$
\begin{aligned}
\mathrm{Var}(Z_{2a}^*) &= \frac{16}{v^2}\sum_{\substack{i,j,k,\ell(dist)\\ i',j',k',\ell'(dist)}} \eta_i\eta_k\eta_{i'}\eta_{k'}\cdot\mathbb{E}[W_{js}W_{jk}W_{\ell i}^2\cdot W_{j's'}W_{j'k'}W_{\ell' i'}^2] \\
&\leq \frac{C}{\|\theta\|_1^4}\big(\|\theta\|_3^6\|\theta\|^2\|\theta\|_1^2 + \|\theta\|_4^4\|\theta\|_3^3\|\theta\|^2\|\theta\|_1^3 + \|\theta\|_3^3\|\theta\|^6\|\theta\|_1^3 \\
&\quad + \|\theta\|_3^3\|\theta\|^6\|\theta\|_1 + \|\theta\|_4^4\|\theta\|^6\|\theta\|_1^2 + \|\theta\|^{10}\|\theta\|_1^2\big) \\
&\leq \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1},
\end{aligned}
$$

where the last inequality is obtained similarly as in the calculation of $\mathrm{Var}(\widetilde{Z}_{2a})$. Last, consider $Z_{2a}^\dagger$. Write

$$
Z_{2a}^\dagger = \frac{2}{v}\sum_{i,j,k,\ell(dist)} \eta_i\eta_k W_{j\ell}^2W_{jk}W_{\ell i} + \frac{2}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\notin\{j,k\},t\notin\{\ell,i\}\\ (s,t)\neq(\ell,j)}} \eta_i\eta_k W_{js}W_{jk}W_{\ell t}W_{\ell i} \tag{102}
$$

Regarding the first term, we note that

$$
\begin{aligned}
&\eta_i\eta_k\eta_{i'}\eta_{k'}\cdot\mathbb{E}[W_{j\ell}^2W_{jk}W_{\ell i}\cdot W_{j'\ell'}^2W_{j'k'}W_{\ell' i'}] \\
&= \begin{cases} \eta_i^2\eta_k^2\,\mathbb{E}[W_{jk}^2W_{\ell i}^2W_{j\ell}^4] \leq C\theta_i^3\theta_j^2\theta_k^3\theta_\ell^2, & \text{if } (j',k')=(j,k),\ (i',\ell')=(i,\ell); \\ \eta_i\eta_k^2\eta_\ell\,\mathbb{E}[W_{jk}^2W_{\ell i}^2W_{j\ell}^2W_{ji}^2] \leq C\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3, & \text{if } (j',k')=(j,k),\ (i',\ell')=(\ell,i); \\ \eta_i^2\eta_k\eta_\ell\,\mathbb{E}[W_{jk}^2W_{\ell i}^2W_{j\ell}^2W_{k\ell}^2] \leq C\theta_i^3\theta_j^2\theta_k^3\theta_\ell^4, & \text{if } (j',k')=(k,j),\ (i',\ell')=(i,\ell); \\ \eta_i\eta_k\eta_\ell\eta_j\,\mathbb{E}[W_{jk}^2W_{\ell i}^2W_{j\ell}^2W_{ki}^2] \leq C\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3, & \text{if } (j',k')=(k,j),\ (i',\ell')=(\ell,i); \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\mathrm{Var}\Big(\frac{2}{v}\sum_{i,j,k,\ell(dist)} \eta_i\eta_k W_{j\ell}^2W_{jk}W_{\ell i}\Big) \\
\leq &\frac{C}{\|\theta\|_1^4}\sum_{i,j,k,\ell}(\theta_i^3\theta_j^2\theta_k^3\theta_\ell^2 + \theta_i^3\theta_j^3\theta_k^3\theta_\ell^3 + \theta_i^3\theta_j^2\theta_k^3\theta_\ell^4) \\
\leq &\frac{C}{\|\theta\|_1^4}\big(\|\theta\|_3^6\|\theta\|^4 + \|\theta\|_3^{12} + \|\theta\|_4^4\|\theta\|_3^6\|\theta\|^2\big) \\
\leq &\frac{C\|\theta\|_3^6\|\theta\|^4}{\|\theta\|_1^4}.
\end{aligned}
$$

Regarding the second term in (102). We note that, for $\eta_i\eta_k W_{js}W_{jk}W_{\ell t}W_{\ell i}$ and $\eta_{i'}\eta_{k'}W_{j's'}W_{j'k'}W_{\ell' t'}W_{\ell' i'}$ to be correlated, all the W terms have to be perfectly paired. It turns out that

$$\mathbb{E}[W_{js}W_{jk}W_{\ell t}W_{\ell i}\cdot W_{j's'}W_{j'k'}W_{\ell' t'}W_{\ell' i'}] = \mathbb{E}[W_{js}^2 W_{jk}^2 W_{\ell t}^2 W_{\ell i}^2].$$

To perfectly pair the W terms, there are two possible cases: (i) $(j',\ell') = (j,\ell)$, $\{s',k'\} = \{s,k\}$, $\{\ell',i'\} = \{\ell,i\}$. (ii) $(j',\ell') = (\ell,j)$, $\{s',k'\} = \{\ell,i\}$, $\{\ell',i'\} = \{s,k\}$. As a result, $\eta_i\eta_k\eta_{i'}\eta_{k'}$ only has the following possibilities:

$$\begin{aligned}&\eta_i\eta_k(\eta_i\eta_k) = \eta_i^2\eta_k^2,\ \eta_i\eta_k(\eta_i\eta_s) = \eta_i^2\eta_k\eta_s,\ \eta_i\eta_k(\eta_\ell\eta_k) = \eta_i\eta_k^2\eta_\ell,\ \eta_i\eta_k(\eta_\ell\eta_s) = \eta_i\eta_k\eta_\ell\eta_s,\\ &\eta_i\eta_k(\eta_k\eta_i) = \eta_i^2\eta_k^2,\ \eta_i\eta_k(\eta_k\eta_\ell) = \eta_i\eta_k^2\eta_\ell,\ \eta_i\eta_k(\eta_s\eta_i) = \eta_i^2\eta_k\eta_s,\ \eta_i\eta_k(\eta_s\eta_\ell) = \eta_i\eta_k\eta_\ell\eta_s.\end{aligned}$$

By symmetry, there are only three different types: $\eta_i^2\eta_k^2$, $\eta_i^2\eta_k\eta_s$, and $\eta_i\eta_k\eta_\ell\eta_s$. It follows that

$$\begin{aligned}&\mathrm{Var}\Big(\frac{2}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\notin\{j,k\},t\notin\{\ell,i\},(s,t)\neq(\ell,j)}}\eta_i\eta_k W_{js}W_{jk}W_{\ell t}W_{\ell i}\Big)\\ \leq&\frac{C}{\|\theta\|_1^4}\sum_{i,j,k,\ell,s,t}(\theta_i^2\theta_k^2 + \theta_i^2\theta_k\theta_s + \theta_i\theta_k\theta_\ell\theta_s)\cdot\theta_j^2\theta_s\theta_k\theta_\ell^2\theta_t\theta_i\\ \leq&\frac{C}{\|\theta\|_1^4}\sum_{i,j,k,\ell,s,t}(\theta_i^3\theta_j^2\theta_k^3\theta_\ell^2\theta_s\theta_t + \theta_i^3\theta_j^2\theta_k^2\theta_\ell^2\theta_s^2\theta_t + \theta_i^2\theta_j^2\theta_k^2\theta_\ell^3\theta_s^2\theta_t)\\ \leq&\frac{C}{\|\theta\|_1^4}\big(\|\theta\|_3^6\|\theta\|^4\|\theta\|_1^2 + \|\theta\|_3^3\|\theta\|^8\|\theta\|_1\big) \leq \frac{C\|\theta\|^4\|\theta\|_3^6}{\|\theta\|_1^2}.\end{aligned}$$

It follows that

$$\mathrm{Var}(Z_{2a}^\dagger) \leq \frac{C\|\theta\|^4\|\theta\|_3^6}{\|\theta\|_1^2}.$$

Comparing the variances of $\widetilde{Z}_{2a}$, Z_{2a}^* and $Z_{2a}^\dagger$, we find out that the variance of Z_{2a}^* dominates. As a result,

$$\mathrm{Var}(Z_{2a}) \leq 3\mathrm{Var}(\widetilde{Z}_{2a}) + 3\mathrm{Var}(Z_{2a}^*) + 3\mathrm{Var}(Z_{2a}^\dagger) \leq \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1}. \tag{103}$$

Second, we study Z_{2b}. It is seen that

$$\begin{aligned}Z_{2b} =& 2\sum_{i,j,k,\ell(dist)}\eta_i\Big(-\frac{1}{\sqrt{v}}\sum_{s\neq j}W_{js}\Big)W_{jk}\Big(-\frac{1}{\sqrt{v}}\sum_{t\neq k}W_{kt}\Big)\eta_\ell W_{\ell i}\\ =&\frac{2}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq j,t\neq k}}\eta_i\eta_\ell W_{js}W_{jk}W_{kt}W_{\ell i}.\end{aligned}$$

We divide summands into four groups: (i) $s = k$ and $t = j$, (ii) $s = k$ and $t \neq j$, (iii) $s \neq k$ and $t = j$, (iv) $s \neq k$ and $t \neq j$. By index symmetry, the sums of group (ii) and group (iii) are equal. We end up with

$$\begin{aligned}Z_{2b} =& \frac{2}{v}\sum_{i,j,k,\ell(dist)}\eta_i\eta_\ell W_{jk}^3 W_{\ell i} + \frac{4}{v}\sum_{i,j,k,\ell(dist),t\notin\{k,j\}}\eta_i\eta_\ell W_{jk}^2 W_{kt}W_{\ell i}\\ &+\frac{2}{v}\sum_{i,j,k,\ell(dist),s\neq\{j,k\},t\neq\{j,k\}}\eta_i\eta_\ell W_{js}W_{jk}W_{kt}W_{\ell i}\\ \equiv&\widetilde{Z}_{2b} + Z_{2b}^* + Z_{2b}^\dagger.\end{aligned}$$

It is easy to see that all three terms have mean zero. Therefore,

$$\mathbb{E}[Z_{2b}] = 0. \tag{104}$$

We then bound the variances. Consider $\widetilde{Z}_{2b}$ first. By direct calculations,

$$\eta_i\eta_\ell\eta_{i'}\eta_{\ell'} \cdot \mathbb{E}[W_{jk}^3 W_{\ell i} \cdot W_{j'k'}^3 W_{\ell' i'}]$$
$$= \begin{cases} \eta_i^2\eta_\ell^2 \cdot \mathbb{E}[W_{jk}^6 W_{\ell i}^2] \leq C\theta_i^3\theta_j\theta_k\theta_\ell^3, & \text{if } \{j',k'\} = \{j,k\},\ \{\ell',i'\} = \{\ell,i\}; \\ \eta_i\eta_\ell\eta_j\eta_k \cdot \mathbb{E}[W_{jk}^4 W_{\ell i}^4] \leq C\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2, & \text{if } \{j',k'\} = \{\ell,i\},\ \{\ell',i'\} = \{j,k\}; \\ 0, & \text{otherwise.} \end{cases}$$

It follows that

$$\begin{aligned} \mathrm{Var}(\widetilde{Z}_{2b}) &\leq \frac{C}{\|\theta\|_1^4}\Big(\sum_{i,j,k,\ell} \theta_i^3\theta_j\theta_k\theta_\ell^3 + \sum_{i,j,k,\ell} \theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\Big) \\ &\leq \frac{C}{\|\theta\|_1^4}\big(\|\theta\|_3^6\|\theta\|_1^2 + \|\theta\|^8\big) \\ &\leq \frac{C\|\theta\|_3^6}{\|\theta\|_1^2}. \end{aligned}$$

Consider Z_{2b}^* next. By direct calculations,

$$\eta_i\eta_\ell\eta_{i'}\eta_{\ell'} \cdot \mathbb{E}[W_{jk}^2 W_{kt} W_{\ell i} \cdot W_{j'k'}^2 W_{k't'} W_{\ell' i'}]$$
$$= \begin{cases} \eta_i^2\eta_\ell^2\,\mathbb{E}[W_{jk}^4 W_{kt}^2 W_{\ell i}^2] \leq C\theta_i^3\theta_j\theta_k^2\theta_\ell^3\theta_t, & \text{if } (k',t') = (k,t),\ \{\ell',i'\} = \{\ell,i\},\ j' = j; \\ \eta_i^2\eta_\ell^2\,\mathbb{E}[W_{jk}^2 W_{kt}^2 W_{\ell i}^2 W_{j'k}^2] \leq C\theta_i^3\theta_j\theta_k^3\theta_\ell^3\theta_t\theta_{j'}, & \text{if } (k',t') = (k,t),\ \{\ell',i'\} = \{\ell,i\},\ j' \neq j; \\ \eta_i^2\eta_\ell^2\,\mathbb{E}[W_{jk}^2 W_{kt}^2 W_{\ell i}^2 W_{j't}^2] \leq C\theta_i^3\theta_j\theta_k^2\theta_\ell^3\theta_t^2\theta_{j'}, & \text{if } (k',t') = (t,k),\ \{\ell',i'\} = \{\ell,i\}; \\ \eta_i\eta_\ell\eta_k\eta_t\,\mathbb{E}[W_{jk}^2 W_{kt}^2 W_{\ell i}^4] \leq C\theta_i^2\theta_j\theta_k^3\theta_\ell^2\theta_t^2, & \text{if } (k',t') = (\ell,i),\ \{\ell',i'\} = \{k,t\},\ j' = i; \\ \eta_i\eta_\ell\eta_k\eta_t\,\mathbb{E}[W_{jk}^2 W_{kt}^2 W_{\ell i}^2 W_{j'\ell}^2] \leq C\theta_i^2\theta_j\theta_k^3\theta_\ell^3\theta_t^2\theta_{j'}, & \text{if } (k',t') = (\ell,i),\ \{\ell',i'\} = \{k,t\},\ j' \neq i; \\ \eta_i\eta_\ell\eta_k\eta_t\,\mathbb{E}[W_{jk}^2 W_{kt}^2 W_{\ell i}^4] \leq C\theta_i^2\theta_j\theta_k^3\theta_\ell^2\theta_t^2, & \text{if } (k',t') = (i,\ell),\ \{\ell',i'\} = \{k,t\},\ j' = \ell; \\ \eta_i\eta_\ell\eta_k\eta_t\,\mathbb{E}[W_{jk}^2 W_{kt}^2 W_{\ell i}^2 W_{j'i}^2] \leq C\theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_t^2\theta_{j'}, & \text{if } (k',t') = (i,\ell),\ \{\ell',i'\} = \{k,t\},\ j' \neq \ell; \\ \eta_i^2\eta_\ell^2\,\mathbb{E}[W_{jk}^3 W_{kt}^3 W_{\ell i}^2] \leq C\theta_i^3\theta_j\theta_k^2\theta_\ell^3\theta_t, & \text{if } (k',t',j') = (k,j,t),\ \{i',\ell'\} = \{i,\ell\}; \\ 0, & \text{otherwise.} \end{cases}$$

There are only two four types on the right hand side. It follows that

$$\begin{aligned} \mathrm{Var}(Z_{2b}^*) &\leq \frac{C}{\|\theta\|_1^4}\Big(\sum_{i,j,k,\ell,t,j'} \theta_i^3\theta_j\theta_k^3\theta_\ell^3\theta_t\theta_{j'} + \sum_{i,j,k,\ell,t,j'} \theta_i^3\theta_j\theta_k^2\theta_\ell^3\theta_t^2\theta_{j'} \\ &\qquad + \sum_{i,j,k,\ell,t} \theta_i^3\theta_j\theta_k^2\theta_\ell^3\theta_t + \sum_{i,j,k,\ell,t} \theta_i^2\theta_j\theta_k^3\theta_\ell^2\theta_t^2\Big) \\ &\leq \frac{C}{\|\theta\|_1^4}\big(\|\theta\|_3^9\|\theta\|_1^3 + \|\theta\|_3^6\|\theta\|^4\|\theta\|_1^2 + \|\theta\|_3^6\|\theta\|^2\|\theta\|_1^2 + \|\theta\|_3^3\|\theta\|^6\|\theta\|_1\big) \\ &\leq \frac{C\|\theta\|_3^9}{\|\theta\|_1}. \end{aligned}$$

Last, consider $Z_{2b}^\dagger$. By direct calculations,

$$\eta_i\eta_\ell\eta_{i'}\eta_{\ell'} \cdot \mathbb{E}[W_{js}W_{jk}W_{kt}W_{\ell i} \cdot W_{j's'}W_{j'k'}W_{k't'}W_{\ell' i'}]$$

$$= \begin{cases} \eta_i^2\eta_\ell^2\,\mathbb{E}[W_{js}^2W_{jk}^2W_{kt}^2W_{\ell i}^2] \le C\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3\theta_s\theta_t, & \text{if } (j',s')=(j,s),\ (k',t')=(k,t),\ \{\ell',i'\}=\{\ell,i\};\\ \eta_i^2\eta_\ell^2\,\mathbb{E}[W_{js}^2W_{jk}^2W_{kt}^2W_{\ell i}^2] \le C\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3\theta_s\theta_t, & \text{if } (j',s')=(k,t),\ (k',t')=(j,s),\ \{\ell',i'\}=\{\ell,i\};\\ 0, & \text{otherwise.}\end{cases}$$

It follows that

$$\mathrm{Var}(Z_{2b}^\dagger) \le \frac{C}{\|\theta\|_1^4}\sum_{i,j,k,\ell,s,t}\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3\theta_s\theta_t \le \frac{C\|\theta\|^4\|\theta\|_3^6}{\|\theta\|_1^2}.$$

Since $\|\theta\|_1\|\theta\|_3^3 \ge \|\theta\|^4 \to \infty$, the variance of Z_{2b}^* dominates the variances of $\widetilde{Z}_{2b}$ and $Z_{2b}^\dagger$. We thus have

$$\mathrm{Var}(Z_{2b}) \le 3\mathrm{Var}(\widetilde{Z}_{2b}) + 3\mathrm{Var}(Z_{2b}^*) + 3\mathrm{Var}(Z_{2b}^\dagger) \le \frac{C\|\theta\|_3^9}{\|\theta\|_1}. \tag{105}$$

We now combine (101), (103), (104), and (105). Since $\|\theta\|_3^6 \le \theta_{\max}^2\|\theta\|^4 \ll \|\theta\|^6$, the right hand side of (105) is much smaller than the right hand side of (103). It yields that

$$\mathbb{E}[Z_2] = 2\|\theta\|^4\cdot[1+o(1)], \qquad \mathrm{Var}(Z_2) \le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8).$$

This proves the claims of Z_2.

G.4.7. *Proof of Lemma G.7.* It suffices to prove the claims for each of Z_1-Z_6. We have analyzed Z_1-Z_2 under the null hypothesis. The proof for the alternative hypothesis is similar and omitted. We obtain that

$$\big|\mathbb{E}[Z_1]\big| \le C\|\theta\|^4, \qquad \mathrm{Var}(Z_1) \le C\|\theta\|^2\|\theta\|_3^6 = o(\|\theta\|^8),$$

$$\big|\mathbb{E}[Z_2]\big| \le C\|\theta\|^4, \qquad \mathrm{Var}(Z_2) \le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8).$$

First, we analyze Z_3. Since $\delta_{ij} = \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_i - \tilde{\eta}_i)$, we have

$$\begin{aligned} Z_3 = &\sum_{\substack{i,j,k,\ell\\(dist)}}\eta_i(\eta_j-\tilde{\eta}_j)\eta_j(\eta_k-\tilde{\eta}_k)\widetilde{\Omega}_{k\ell}W_{\ell i} + \sum_{\substack{i,j,k,\ell\\(dist)}}\eta_i(\eta_j-\tilde{\eta}_j)^2\eta_k\widetilde{\Omega}_{k\ell}W_{\ell i}\\ &+ \sum_{\substack{i,j,k,\ell\\(dist)}}(\eta_i-\tilde{\eta}_i)\eta_j^2(\eta_k-\tilde{\eta}_k)\widetilde{\Omega}_{k\ell}W_{\ell i} + \sum_{\substack{i,j,k,\ell\\(dist)}}(\eta_i-\tilde{\eta}_i)\eta_j(\eta_j-\tilde{\eta}_j)\eta_k\widetilde{\Omega}_{k\ell}W_{\ell i}\\ \equiv\ & Z_{3a}+Z_{3b}+Z_{3c}+Z_{3d}. \end{aligned} \tag{106}$$

First, we study Z_{3a}. By direct calculations,

$$\begin{aligned} Z_{3a} &= \sum_{i,j,k,\ell(dist)}\eta_i\Big(-\frac{1}{\sqrt{v}}\sum_{s\neq j}W_{js}\Big)\eta_j\Big(-\frac{1}{\sqrt{v}}\sum_{t\neq k}W_{kt}\Big)\widetilde{\Omega}_{k\ell}W_{\ell i}\\ &= \frac{1}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq j,t\neq k}}\beta_{ijk\ell}W_{js}W_{kt}W_{\ell i}, \qquad \text{where } \beta_{ijk\ell} = \eta_i\eta_j\widetilde{\Omega}_{k\ell}. \end{aligned}$$

Since (i,j,k,ℓ) are distinct, all summands have mean zero. Hence,

$$\mathbb{E}[Z_{3a}] = 0. \tag{107}$$

To bound its variance, re-write

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

We note that $|\beta_{ijk\ell}| \leq C\alpha\theta_i\theta_j\theta_k\theta_\ell$ by (74) and (81). Consider the variance of $\widetilde{Z}_{3a}$. By direct calculations,

$$
\begin{aligned}
&\beta_{ijk\ell}\beta_{i'j'k'\ell'} \cdot \mathrm{Cov}(W_{jk}^2 W_{\ell i}, W_{j'k'}^2 W_{\ell' i'})\\
&= \begin{cases}
C\alpha^2\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\, \mathbb{E}[W_{jk}^4 W_{\ell i}^2] \leq C\alpha^2\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3, & \text{if } \{\ell', i'\} = \{\ell, i\},\ \{j',k'\} = \{j,k\};\\
C\alpha^2\theta_i^2\theta_j\theta_k\theta_\ell^2\theta_{j'}\theta_{k'}\, \mathbb{E}[W_{jk}^2 W_{j'k'}^2 W_{\ell i}^2] \leq C\alpha^2\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3\theta_{j'}^2\theta_{k'}^2, & \text{if } \{\ell', i'\} = \{\ell, i\},\ \{j',k'\} \neq \{j,k\};\\
C\alpha^2\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\, \mathbb{E}[W_{jk}^3 W_{\ell i}^3] \leq C\alpha^2\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3, & \text{if } \{j', k'\} = \{\ell, i\},\ \{\ell',i'\} = \{j,k\};\\
0, & \text{otherwise.}
\end{cases}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\mathrm{Var}(\widetilde{Z}_{3a}) &\leq \frac{C\alpha^2}{\|\theta\|_1^4}\Big(\sum_{i,j,k,\ell}\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3 + \sum_{i,j,k,\ell,j',k'}\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3\theta_{j'}^2\theta_{k'}^2\Big)\\
&\leq \frac{C\alpha^2}{\|\theta\|_1^4}\big(\|\theta\|_3^{12} + \|\theta\|^8\|\theta\|_3^6\big)\\
&\leq \frac{C\alpha^2\|\theta\|_3^{12}}{\|\theta\|_1^2}.
\end{aligned}
$$

Consider the variance of Z_{3a}^*. For $W_{js}W_{kt}W_{\ell i}$ and $W_{j's'}W_{k't'}W_{\ell' i'}$ to be correlated, all W terms have to be perfectly paired. By symmetry across indices, it reduces to three cases: (i) $(\ell', i') = (\ell, i)$, $(j', s') = (j, s)$, $(k', t') = (k, t)$; (ii) $(\ell', i') = (j, s)$, $(j', s') = (\ell, i)$, $(k', t') = (k, t)$; (iii) $(\ell', i') = (j, s)$, $(j', s') = (k, t)$, $(k', t') = (\ell, i)$. It follows that

$$
\begin{aligned}
&\beta_{ijk\ell}\beta_{i'j'k'\ell'} \cdot \mathbb{E}[W_{js}W_{kt}W_{\ell i} \cdot W_{j's'}W_{k't'}W_{\ell' i'}]\\
&\leq C\alpha^2(\theta_i\theta_j\theta_k\theta_\ell)(\theta_{i'}\theta_{j'}\theta_{k'}\theta_{\ell'}) \cdot \mathbb{E}[W_{js}^2 W_{kt}^2 W_{\ell i}^2]\\
&\leq \begin{cases}
C\alpha^2\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\mathbb{E}[W_{js}^2 W_{kt}^2 W_{\ell i}^2] \leq C\alpha^2\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3\theta_s\theta_t, & \text{case (i)}\\
C\alpha^2(\theta_i\theta_j\theta_k\theta_\ell)(\theta_s\theta_\ell\theta_k\theta_j)\mathbb{E}[W_{js}^2 W_{kt}^2 W_{\ell i}^2] \leq C\alpha^2\theta_i^2\theta_j^3\theta_k^3\theta_\ell^3\theta_s^2\theta_t, & \text{case (ii)}\\
C\alpha^2(\theta_i\theta_j\theta_k\theta_\ell)(\theta_s\theta_k\theta_\ell\theta_j)\mathbb{E}[W_{js}^2 W_{kt}^2 W_{\ell i}^2] \leq C\alpha^2\theta_i^2\theta_j^3\theta_k^3\theta_\ell^3\theta_s^2\theta_t, & \text{case (iii)}\\
0, & \text{otherwise.}
\end{cases}
\end{aligned}
$$

As a result,

$$
\begin{aligned}
\mathrm{Var}(Z_{3a}^*) &\leq \frac{C}{\|\theta\|_1^4}\Big(\sum_{i,j,k,\ell,s,t}\alpha^2\theta_i^3\theta_j^3\theta_k^3\theta_\ell^3\theta_s\theta_t + \sum_{i,j,k,\ell,s,t}\alpha^2\theta_i^2\theta_j^3\theta_k^3\theta_\ell^3\theta_s^2\theta_t\Big)\\
&\leq \frac{C\alpha^2}{\|\theta\|_1^4}\big(\|\theta\|_3^{12}\|\theta\|_1^2 + \|\theta\|^4\|\theta\|_3^9\|\theta\|_1\big)\\
&\leq \frac{C\alpha^2\|\theta\|_3^{12}}{\|\theta\|_1^2}.
\end{aligned}
$$

Combining the variance of $\widetilde{Z}_{3a}$ and Z^*_{3a} gives

$$\mathrm{Var}(Z_{3a}) \le \frac{C\alpha^2\|\theta\|_3^{12}}{\|\theta\|_1^2}. \tag{108}$$

Second, we study Z_{3b}. It is seen that

$$\begin{aligned}
Z_{3b} &= \sum_{i,j,k,\ell(dist)} \eta_i\Big(-\frac{1}{\sqrt{v}}\sum_{s\neq j} W_{js}\Big)\Big(-\frac{1}{\sqrt{v}}\sum_{t\neq j} W_{jt}\Big)\eta_k\widetilde{\Omega}_{k\ell}W_{\ell i} \\
&= \frac{1}{v}\sum_{\substack{i,j,\ell(dist)\\ s\neq j, t\neq j}}\Big(\sum_{k\notin\{i,j,\ell\}}\eta_i\eta_k\widetilde{\Omega}_{k\ell}\Big)W_{js}W_{jt}W_{\ell i} \\
&\equiv \frac{1}{v}\sum_{\substack{i,j,\ell(dist)\\ s\neq j, t\neq j}}\beta_{ij\ell}W_{js}W_{jt}W_{\ell i},
\end{aligned}$$

where by (74) and (81),

$$|\beta_{ij\ell}| \le \sum_{k\notin\{i,j,\ell\}}|\eta_i\eta_k\widetilde{\Omega}_{k\ell}| \le \sum_k C\alpha\theta_i\theta_k^2\theta_\ell \le C\alpha\|\theta\|^2\cdot\theta_i\theta_\ell. \tag{109}$$

We further decompose Z_{3b} into

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

It is easy to see that both terms have mean zero. It follows that

$$\mathbb{E}[Z_{3b}] = 0. \tag{110}$$

To calculate the variance of $\widetilde{Z}_{3b}$, we note that

$$\begin{aligned}
&\beta_{ij\ell}\beta_{i'j'\ell'}\cdot\mathbb{E}[W_{js}^2W_{\ell i}\cdot W_{j's'}^2W_{\ell'i'}] \\
&\le C\alpha^2\|\theta\|^4\theta_i\theta_{i'}\theta_\ell\theta_{\ell'}\cdot\mathbb{E}[W_{js}^2W_{\ell i}\cdot W_{j's'}^2W_{\ell'i'}] \\
&\le \begin{cases}
C\alpha^2\|\theta\|^4\theta_i^2\theta_\ell^2\cdot\mathbb{E}[W_{js}^4W_{\ell i}^2] \le C\alpha^2\|\theta\|^4\theta_i^3\theta_j\theta_\ell^3\theta_s & \text{if } \{\ell',i'\}=\{\ell,i\},\ \{j',s'\}=\{j,s\} \\
C\alpha^2\|\theta\|^4\theta_i^2\theta_\ell^2\cdot\mathbb{E}[W_{js}^2W_{\ell i}^2W_{j's'}^2] \le C\alpha^2\|\theta\|^4\theta_i^3\theta_j\theta_\ell^3\theta_s\theta_{j'}\theta_{s'}, & \text{if } \{\ell',i'\}=\{\ell,i\},\ \{j',s'\}\neq\{j,s\}; \\
C\alpha^2\|\theta\|^4\theta_i\theta_\ell\theta_j\theta_s\cdot\mathbb{E}[W_{js}^3W_{\ell i}^3] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_j^2\theta_\ell^2\theta_s^2, & \text{if } \{\ell',i'\}=\{j,s\},\ \{j',s'\}=\{\ell,i\}; \\
0, & \text{otherwise.}
\end{cases}
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathrm{Var}(\widetilde{Z}_{3b}) &\le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^4}\Big(\sum_{i,j,\ell,s}\theta_i^3\theta_j\theta_\ell^3\theta_s + \sum_{i,j,\ell,s,j',s'}\theta_i^3\theta_j\theta_\ell^3\theta_s\theta_{j'}\theta_{s'} + \sum_{i,j,\ell,s,j',s'}\theta_i^2\theta_j^2\theta_\ell^2\theta_s^2\Big) \\
&\le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^4}\big(\|\theta\|_3^6\|\theta\|_1^2 + \|\theta\|_3^6\|\theta\|_1^4 + \|\theta\|^8\big) \\
&\le C\alpha^2\|\theta\|^4\|\theta\|_3^6.
\end{aligned}$$

To calculate the variance of Z^*_{3b}, we note that $\mathrm{E}[W_{js}W_{jt}W_{\ell i}\cdot W_{j's'}W_{j't'}W_{\ell'i'}]$ is nonzero only if $j'=j$, $\{s',t'\}=\{s,t\}$ and $\{\ell',i'\}=\{\ell,i\}$. Combining it with (112) gives

$$\mathrm{Var}(Z^*_{3b}) \le \frac{C}{v^2}\sum_{\substack{i,j,\ell(dist)\\ s,t(dist)\notin\{j\}}}\beta_{ij\ell}^2\cdot\mathbb{E}[W_{js}^2W_{jt}^2W_{\ell i}^2]$$

$$\le \frac{C}{\|\theta\|_1^4} \sum_{i,j,\ell,s,t} (\alpha\|\theta\|^2\theta_i\theta_\ell)^2 \cdot \theta_j^2\theta_s\theta_t\theta_\ell\theta_i$$

$$\le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^4} \sum_{i,j,\ell,s,t} \theta_i^3\theta_j^2\theta_\ell^3\theta_s\theta_t$$

$$\le \frac{C\alpha^2\|\theta\|^6\|\theta\|_3^6}{\|\theta\|_1^2}.$$

Since $\|\theta\|^6 \le \|\theta\|^4\|\theta\|^2 \ll \|\theta\|^4\|\theta\|_1$, the variance of $\widetilde{Z}_{3b}$ dominates the variance of Z_{3b}^*. Combining the above gives

(111) $$\mathrm{Var}(Z_{3b}) \le 2\mathrm{Var}(\widetilde{Z}_{3b}) + 2\mathrm{Var}(Z_{3b}^*) \le C\alpha^2\|\theta\|^4\|\theta\|_3^6.$$

Third, we study Z_{3c}. It is seen that

$$Z_{3c} = \sum_{i,j,k,\ell(dist)} \Big(-\frac{1}{\sqrt{v}}\sum_{s\neq i} W_{is}\Big)\eta_j^2\Big(-\frac{1}{\sqrt{v}}\sum_{t\neq k} W_{kt}\Big)\widetilde{\Omega}_{k\ell}W_{\ell i}$$

$$= \frac{1}{v}\sum_{\substack{i,k,\ell(dist)\\ s\neq i, t\neq k}} \Big(\sum_{j\notin\{i,k,\ell\}} \eta_j^2\widetilde{\Omega}_{k\ell}\Big) W_{is}W_{kt}W_{\ell i}$$

$$\equiv \frac{1}{v}\sum_{\substack{i,k,\ell(dist)\\ s\neq i, t\neq k}} \beta_{ik\ell}W_{is}W_{kt}W_{\ell i},$$

where by (74) and (81),

(112) $$|\beta_{ik\ell}| \le \sum_{j\notin\{i,k,\ell\}} |\eta_j^2\widetilde{\Omega}_{k\ell}| \le \sum_j C\alpha\theta_j^2\theta_k\theta_\ell \le C\alpha\|\theta\|^2\theta_k\theta_\ell.$$

There are two cases for the indices: $i=\ell$ and $i\neq\ell$. We further decompose Z_{3c} into

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

It is easy to see that both terms have zero mean. Hence,

(113) $$\mathbb{E}[Z_{3c}] = 0.$$

To calculate the variance of $\widetilde{Z}_{3c}$, we note that $W_{i\ell}^2W_{kt}$ and $W_{i'\ell'}^2W_{k't'}$ are correlated only when (i) $\{k',t'\}=\{k,t\}$ or (ii) $\{k',t'\}=\{i,\ell\}$ and $\{i',\ell'\}=\{k,t\}$. By direct calculations,

$$\beta_{ik\ell}\beta_{i'k'\ell'} \cdot \mathbb{E}[W_{i\ell}^2W_{kt}\cdot W_{i'\ell'}^2W_{k't'}]$$
$$\le C\alpha^2\|\theta\|^4\theta_k\theta_{k'}\theta_\ell\theta_{\ell'} \cdot \mathbb{E}[W_{i\ell}^2W_{kt}\cdot W_{i'\ell'}^2W_{k't'}]$$

$$\le \begin{cases} C\alpha^2\|\theta\|^4\theta_k^2\theta_\ell^2\,\mathbb{E}[W_{i\ell}^4W_{kt}^2] \le C\alpha^2\|\theta\|^4\theta_i\theta_k^3\theta_\ell^3\theta_t, & \text{if } (k',t')=(k,t),\ (i',\ell')=(i,\ell);\\ C\alpha^2\|\theta\|^4\theta_k^2\theta_\ell\theta_i\,\mathbb{E}[W_{i\ell}^4W_{kt}^2] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^3\theta_\ell^2\theta_t, & \text{if } (k',t')=(k,t),\ (i',\ell')=(\ell,i);\\ C\alpha^2\|\theta\|^4\theta_k\theta_\ell^2\theta_t\,\mathbb{E}[W_{i\ell}^4W_{kt}^2] \le C\alpha^2\|\theta\|^4\theta_i\theta_k^2\theta_\ell^3\theta_t^2, & \text{if } (k',t')=(t,k),\ (i',\ell')=(i,\ell);\\ C\alpha^2\|\theta\|^4\theta_k\theta_t\theta_\ell\theta_i\,\mathbb{E}[W_{i\ell}^4W_{kt}^2] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^2\theta_\ell^2\theta_t^2, & \text{if } (k',t')=(t,k),\ (i',\ell')=(\ell,i);\\ C\alpha^2\|\theta\|^4\theta_k^2\theta_\ell\theta_{\ell'}\,\mathbb{E}[W_{i\ell}^2W_{kt}^2W_{i'\ell'}^2] \le C\alpha^2\|\theta\|^4\theta_i\theta_k^3\theta_\ell^2\theta_t\theta_{i'}\theta_{\ell'}^2, & \text{if } (k',t')=(k,t),\ \{i',\ell'\}\neq\{i,\ell\};\\ C\alpha^2\|\theta\|^4\theta_k\theta_t\theta_\ell\theta_{\ell'}\,\mathbb{E}[W_{i\ell}^2W_{kt}^2W_{i'\ell'}^2] \le C\alpha^2\|\theta\|^4\theta_i\theta_k^2\theta_\ell^2\theta_t^2\theta_{i'}\theta_{\ell'}^2, & \text{if } (k',t')=(t,k),\ \{i',\ell'\}\neq\{i,\ell\};\\ C\alpha^2\|\theta\|^4\theta_k\theta_i\theta_\ell\theta_t\,\mathbb{E}[W_{i\ell}^3W_{kt}^3] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^2\theta_\ell^2\theta_t^2, & \text{if } (k',t')=(i,\ell),\ (i',\ell')=(k,t);\\ C\alpha^2\|\theta\|^4\theta_k^2\theta_i\theta_\ell\,\mathbb{E}[W_{i\ell}^3W_{kt}^3] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^3\theta_\ell^2\theta_t, & \text{if } (k',t')=(i,\ell),\ (i',\ell')=(t,k);\\ C\alpha^2\|\theta\|^4\theta_k\theta_\ell^2\theta_t\,\mathbb{E}[W_{i\ell}^3W_{kt}^3] \le C\alpha^2\|\theta\|^4\theta_i\theta_k^2\theta_\ell^3\theta_t^2, & \text{if } (k',t')=(\ell,i),\ (i',\ell')=(k,t);\\ C\alpha^2\|\theta\|^4\theta_k^2\theta_\ell^2\,\mathbb{E}[W_{i\ell}^3W_{kt}^3] \le C\alpha^2\|\theta\|^4\theta_i\theta_k^3\theta_\ell^3\theta_t, & \text{if } (k',t')=(\ell,i),\ (i',\ell')=(t,k);\\ 0, & \text{otherwise.} \end{cases}$$

There are only five types on the right hand side. It follows that

$$\begin{aligned} \mathrm{Var}(\widetilde{Z}_{3c}) &\le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^4}\Big(\sum_{i,k,\ell,t}\theta_i\theta_k^3\theta_\ell^3\theta_t + \sum_{i,k,\ell,t}\theta_i^2\theta_k^3\theta_\ell^2\theta_t + \sum_{i,k,\ell,t}\theta_i^2\theta_k^2\theta_\ell^2\theta_t^2 \\ &\qquad + \sum_{i,k,\ell,t,i',\ell'}\theta_i\theta_k^3\theta_\ell^2\theta_t\theta_{i'}\theta_{\ell'}^2 + \sum_{i,k,\ell,t,i',\ell'}\theta_i\theta_k^2\theta_\ell^2\theta_t^2\theta_{i'}\theta_{\ell'}^2\Big) \\ &\le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^4}\big(\|\theta\|_3^6\|\theta\|_1^2 + \|\theta\|^4\|\theta\|_3^3\|\theta\|_1 + \|\theta\|^8 + \|\theta\|^4\|\theta\|_3^3\|\theta\|_1^3 + \|\theta\|^8\|\theta\|_1^2\big) \\ &\le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1}, \end{aligned}$$

where the last inequality is obtained as follows: Among the five terms in the brackets, the first and third terms are dominated by the last term, and the second term is dominated by the fourth term; it remains to compare the fourth term and the last term, where the fourth term dominated because $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$. To calculate the variance of Z_{3c}^*, we write

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

Regarding the first term, we note that

$$\begin{aligned} &\beta_{ik\ell}\beta_{i'k'\ell'}\cdot\mathbb{E}[W_{ik}^2W_{\ell i}\cdot W_{i'k'}^2W_{\ell' i'}] \\ \le& C\alpha^2\|\theta\|^4\theta_k\theta_\ell\theta_{k'}\theta_{\ell'}\cdot\mathbb{E}[W_{ik}^2W_{\ell i}\cdot W_{i'k'}^2W_{\ell' i'}] \\ \le& \begin{cases} C\alpha^2\|\theta\|^4\theta_k^2\theta_\ell^2\,\mathbb{E}[W_{ik}^4W_{\ell i}^2] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^3\theta_\ell^3, & \text{if } (\ell',i')=(\ell,i),\ k'=k;\\ C\alpha^2\|\theta\|^4\theta_k\theta_\ell^2\theta_{k'}\,\mathbb{E}[W_{ik}^2W_{\ell i}^2W_{ik'}^2] \le C\alpha^2\|\theta\|^4\theta_i^3\theta_k^2\theta_\ell^3\theta_{k'}^2, & \text{if } (\ell',i')=(\ell,i),\ k'\neq k;\\ C\alpha^2\|\theta\|^4\theta_i\theta_k\theta_\ell\theta_{k'}\,\mathbb{E}[W_{ik}^2W_{\ell i}^2W_{\ell k'}^2] \le C\alpha^2\|\theta\|^4\theta_i^3\theta_k^2\theta_\ell^3\theta_{k'}^2, & \text{if } (\ell',i')=(i,\ell);\\ C\alpha^2\|\theta\|^4\theta_k^2\theta_\ell^2\,\mathbb{E}[W_{ik}^3W_{\ell i}^3] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^3\theta_\ell^3, & \text{if } (\ell',i')=(k,i),\ k'=\ell;\\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

It follows that

$$\mathrm{Var}\Big(\frac{1}{v}\sum_{i,k,\ell(dist)}\beta_{ik\ell}W_{ik}^2W_{\ell i}\Big) \le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^4}\Big(\sum_{i,k,\ell}\theta_i^2\theta_k^3\theta_\ell^3 + \sum_{i,k,\ell,k'}\theta_i^3\theta_k^2\theta_\ell^3\theta_{k'}^2\Big)$$

$$\le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^4}\big(\|\theta\|^2\|\theta\|_3^6 + \|\theta\|^4\|\theta\|_3^6\big) \le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^6}{\|\theta\|_1^4}.$$

Regarding the second term, we note that

$$\beta_{ik\ell}\beta_{i'k'\ell'}\cdot\mathbb{E}[W_{is}W_{kt}W_{\ell i}\cdot W_{i's'}W_{k't'}W_{\ell' i'}]$$
$$\le C\alpha^2\|\theta\|^4\theta_k\theta_{k'}\theta_\ell\theta_{\ell'}\cdot\mathbb{E}[W_{is}W_{kt}W_{\ell i}\cdot W_{i's'}W_{k't'}W_{\ell' i'}]$$
$$\le \begin{cases} C\alpha^2\|\theta\|^4\theta_k^2\theta_\ell^2\,\mathbb{E}[W_{is}^2W_{kt}^2W_{\ell i}^2] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^3\theta_\ell^3\theta_s\theta_t, & \text{if } (i',s',\ell')=(i,s,\ell),\ (k',t')=(k,t);\\ C\alpha^2\|\theta\|^4\theta_k\theta_t\theta_\ell^2\,\mathbb{E}[W_{is}^2W_{kt}^2W_{\ell i}^2] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^2\theta_\ell^3\theta_s\theta_t^2, & \text{if } (i',s',\ell')=(i,s,\ell),\ (k',t')=(t,k);\\ C\alpha^2\|\theta\|^4\theta_k^2\theta_\ell\theta_s\,\mathbb{E}[W_{is}^2W_{kt}^2W_{\ell i}^2] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^3\theta_\ell^2\theta_s^2\theta_t, & \text{if } (i',s',\ell')=(i,\ell,s),\ (k',t')=(k,t);\\ C\alpha^2\|\theta\|^4\theta_k\theta_t\theta_\ell\theta_s\,\mathbb{E}[W_{is}^2W_{kt}^2W_{\ell i}^2] \le C\alpha^2\|\theta\|^4\theta_i^2\theta_k^2\theta_\ell^2\theta_s^2\theta_t^2, & \text{if } (i',s',\ell')=(i,\ell,s),\ (k',t')=(t,k);\\ 0, & \text{otherwise.}\end{cases}$$

It follows that

$$\mathrm{Var}\Big(\frac{1}{v}\sum_{\substack{i,k,\ell(dist)\\ s\notin\{i,\ell\},t\neq k,\\ (s,t)\neq(k,i)}}\beta_{ik\ell}W_{is}W_{kt}W_{\ell i}\Big) \le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^4}\sum_{\substack{i,k,\ell,\\ s,t}}(\theta_i^2\theta_k^3\theta_\ell^3\theta_s\theta_t + \theta_i^2\theta_k^2\theta_\ell^3\theta_s\theta_t^2 + \theta_i^2\theta_k^2\theta_\ell^2\theta_s^2\theta_t^2)$$
$$\le \frac{C\alpha^2\|\theta\|^4}{\|\theta\|_1^4}\big(\|\theta\|^2\|\theta\|_3^6\|\theta\|_1^2 + \|\theta\|^6\|\theta\|_3^3\|\theta\|_1 + \|\theta\|^{10}\big)$$
$$\le \frac{C\alpha^2\|\theta\|^6\|\theta\|_3^6}{\|\theta\|_1^2}.$$

We plug the above results into Z^*_{3c}. Since $\|\theta\|^2 \le \|\theta\|_1\theta_{\max} \ll \|\theta\|_1^2$, we have $\frac{C\alpha^2\|\theta\|^8\|\theta\|_3^6}{\|\theta\|_1^4} \ll \frac{C\alpha^2\|\theta\|^6\|\theta\|_3^6}{\|\theta\|_1^2}$. It follows that

$$\mathrm{Var}(Z^*_{3c}) \le \frac{C\alpha^2\|\theta\|^6\|\theta\|_3^6}{\|\theta\|_1^2}.$$

Since $\|\theta\|_3^6 \ll \|\theta\|_3^3\|\theta\|_1$, the variance of Z^*_{3c} is dominated by the variance of $\widetilde{Z}_{3c}$. It follows that

$$\mathrm{Var}(Z_{3c}) \le 2\mathrm{Var}(\widetilde{Z}_{3c}) + 2\mathrm{Var}(Z^*_{3c}) \le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1}. \tag{114}$$

Last, we study Z_{3d}. In the definition of Z_{3d}, if we switch the two indices (j,k), then it becomes

$$Z_{3d} = \sum_{\substack{i,j,k,\ell\\(dist)}}(\eta_i-\tilde\eta_i)\eta_k(\eta_k-\tilde\eta_k)\eta_j\widetilde{\Omega}_{j\ell}W_{\ell i} = \sum_{\substack{i,j,k,\ell\\(dist)}}(\eta_k\eta_j\widetilde{\Omega}_{j\ell})(\eta_i-\tilde\eta_i)(\eta_k-\tilde\eta_k).$$

At the same time, we recall that

$$Z_{3c} = \sum_{\substack{i,j,k,\ell\\(dist)}}(\eta_i-\tilde\eta_i)\eta_j^2(\eta_k-\tilde\eta_k)\widetilde{\Omega}_{k\ell}W_{\ell i} = \sum_{\substack{i,j,k,\ell\\(dist)}}(\eta_j^2\widetilde{\Omega}_{k\ell})(\eta_i-\tilde\eta_i)(\eta_k-\tilde\eta_k).$$

Here, Z_{3d} has a similar structure as Z_{3c}. Moreover, in deriving the bound for $\mathrm{Var}(Z_{3c})$, we have used $|\eta_j^2\widetilde{\Omega}_{k\ell}| \le C\alpha\theta_j^2\theta_k\theta_\ell$. In the expression of Z_{3d} above, we also have $|\eta_k\eta_j\widetilde{\Omega}_{j\ell}| \le C\alpha\theta_j^2\theta_k\theta_\ell$. Therefore, we can use (113) and (114) to directly get

$$\mathbb{E}[Z_{3d}] = 0, \qquad \mathrm{Var}(Z_{3d}) \le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1} \tag{115}$$

Now, we combine (107), (110), (113) and (114) to get

$$\mathbb{E}[Z_3] = 0.$$

We also combine (108), (111), (114)-(115). Since $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$, the right hand side of (114)-(115) is dominated by the right hand side of (111); since $\|\theta\|_3^6 \ll \|\theta\|_1^2$, the right hand side of (108) is negligible to the right hand side of (111). It follows that

$$\mathrm{Var}(Z_3) \leq C\alpha^2\|\theta\|^4\|\theta\|_3^6.$$

This proves the claims of Z_3.

Next, we analyze Z_4. Since $\delta_{ij} = \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_i - \tilde{\eta}_i)$,

$$Z_4 = \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)W_{\ell i} + \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell W_{\ell i}$$
$$+ \sum_{i,j,k,\ell(dist)} (\eta_i - \tilde{\eta}_i)\eta_j\widetilde{\Omega}_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)W_{\ell i} + \sum_{i,j,k,\ell(dist)} (\eta_i - \tilde{\eta}_i)\eta_j\widetilde{\Omega}_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell W_{\ell i}.$$

If we relabel (i,j,k,ℓ) as (ℓ',k',j',i') in the last sum, it is equal to the first sum. Therefore,

$$\begin{aligned} Z_4 = 2 &\sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)W_{\ell i} \\ &+ \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell W_{\ell i} \\ &+ \sum_{i,j,k,\ell(dist)} (\eta_i - \tilde{\eta}_i)\eta_j\widetilde{\Omega}_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)W_{\ell i} \\ &\equiv Z_{4a} + Z_{4b} + Z_{4c}. \end{aligned} \tag{116}$$

First, we study Z_{4a} and Z_{4b}. We show that they have the same structure as Z_{3c} and Z_{3a}, respectively. In Z_{4a}, by relabeling (i,j,k,ℓ) as (ℓ,k,j,i), we have

$$Z_{4a} = 2\sum_{\substack{i,j,k,\ell \\ (dist)}} \eta_\ell(\eta_k - \tilde{\eta}_k)\widetilde{\Omega}_{kj}\eta_j(\eta_i - \tilde{\eta}_i)W_{i\ell} = 2\sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_j\eta_\ell\widetilde{\Omega}_{kj})(\eta_i - \tilde{\eta}_i)(\eta_k - \tilde{\eta}_k)W_{\ell i}.$$

At the same time, we recall the definition of Z_{3c} in (106):

$$Z_{3c} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i - \tilde{\eta}_i)\eta_j^2(\eta_k - \tilde{\eta}_k)\widetilde{\Omega}_{k\ell}W_{\ell i} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_j^2\widetilde{\Omega}_{k\ell})(\eta_i - \tilde{\eta}_i)(\eta_k - \tilde{\eta}_k)W_{\ell i}.$$

It is seen that Z_{4a} has a similar structure as Z_{3c} does. Also, by (74) and (81), in the expression of Z_{4a}, we have $|\eta_j\eta_\ell\widetilde{\Omega}_{kj}| \leq C\alpha\theta_j^2\theta_k\theta_\ell$, while in the expression of Z_{3d}, we have $|\eta_j^2\widetilde{\Omega}_{k\ell}| \leq C\alpha\theta_j^2\theta_k\theta_\ell$. As a result, if we use similar calculation as before, we will get the same conclusion for Z_{4a} and Z_{3d}. Hence, we use (113)-(114) to conclude that

$$\mathbb{E}[Z_{4a}] = 0, \qquad \mathrm{Var}(Z_{4a}) \leq \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1}. \tag{117}$$

For Z_{4b}, we note that

$$Z_{4b} = \sum_{\substack{i,j,k,\ell \\ (dist)}} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell W_{\ell i} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i\eta_\ell\widetilde{\Omega}_{jk})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k)W_{\ell i},$$

where $|\eta_i\eta_\ell\widetilde{\Omega}_{jk}| \leq C\alpha\theta_i\theta_j\theta_k\theta_\ell$. At the same time, we recall the definition of Z_{3a} in (106):

$$Z_{3a} = \sum_{\substack{i,j,k,\ell \\ (dist)}} \eta_i(\eta_j - \tilde{\eta}_j)\eta_j(\eta_k - \tilde{\eta}_k)\widetilde{\Omega}_{k\ell}W_{\ell i} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i\eta_j\widetilde{\Omega}_{k\ell})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k)W_{\ell i},$$

where $|\eta_i\eta_j\widetilde{\Omega}_{k\ell}| \leq C\alpha\theta_i\theta_j\theta_k\theta_\ell$. Therefore, we can quote the results for Z_{3a} in (107)-(108) to get

$$\mathbb{E}[Z_{4b}] = 0, \qquad \mathrm{Var}(Z_{4b}) \leq \frac{C\alpha^2\|\theta\|_3^{12}}{\|\theta\|_1^2}. \tag{118}$$

Second, we study Z_{4c}. It is seen that

$$\begin{aligned} Z_{4c} &= \sum_{i,j,k,\ell(dist)} \Big(-\frac{1}{\sqrt{v}}\sum_{s\neq i} W_{is}\Big)\eta_j\widetilde{\Omega}_{jk}\eta_k\Big(-\frac{1}{\sqrt{v}}\sum_{t\neq \ell} W_{\ell t}\Big)W_{\ell i} \\ &= \frac{1}{v}\sum_{\substack{i,\ell(dist)\\ s\neq i, t\neq \ell}} \Big(\sum_{j,k(dist)\notin\{i,\ell\}} \eta_j\eta_k\widetilde{\Omega}_{jk}\Big)W_{is}W_{\ell t}W_{\ell i} \\ &\equiv \frac{1}{v}\sum_{\substack{i,\ell(dist)\\ s\neq i, t\neq \ell}} \beta_{i\ell}W_{is}W_{\ell t}W_{\ell i}, \end{aligned}$$

where

$$|\beta_{i\ell}| \leq \sum_{j,k(dist)\notin\{i,\ell\}} |\eta_j\eta_k\widetilde{\Omega}_{jk}| \leq \sum_{j,k} C\alpha\theta_j^2\theta_k^2 \leq C\alpha\|\theta\|^4. \tag{119}$$

We divide the summands into four groups: (i) $s = \ell$, $t = i$; (ii) $s = \ell$, $t \neq i$; (iii) $s \neq \ell$, $t = i$; (iv) $s \neq \ell$, $t \neq i$. By symmetry, the sum of group (ii) and the sum of group (iii) are equal. It yields that

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

Only $\widetilde{Z}_{4c}$ has a nonzero mean. By (80) and (119),

$$\big|\mathbb{E}[Z_{4c}]\big| = \big|\mathbb{E}[\widetilde{Z}_{4c}]\big| \leq \frac{C}{\|\theta\|_1^2}\sum_{i,\ell} \alpha\|\theta\|^4\theta_i\theta_\ell \leq C\alpha\|\theta\|^4. \tag{120}$$

We now compute the variances of these terms. It is seen that

$$\mathrm{Var}(\widetilde{Z}_{4c}) \leq \frac{C}{v^2}\sum_{i,\ell(dist)} \beta_{i\ell}^2\mathrm{Var}(W_{i\ell}^3) \leq \frac{C\alpha^2\|\theta\|^8}{\|\theta\|_1^4}\sum_{i,\ell}\theta_i\theta_\ell \leq \frac{C\alpha^2\|\theta\|^8}{\|\theta\|_1^2}.$$

For Z_{4c}^*, by direct calculations,

$$\begin{aligned} &\beta_{i\ell}\beta_{i'\ell'} \cdot \mathbb{E}[W_{is}W_{\ell i}^2 \cdot W_{i's'}W_{\ell' i'}^2] \\ &\leq C\alpha^2\|\theta\|^8 \cdot \mathbb{E}[W_{is}W_{\ell i}^2 \cdot W_{i's'}W_{\ell' i'}^2] \\ &\leq \begin{cases} C\alpha^2\|\theta\|^8 \cdot \mathbb{E}[W_{is}^2W_{\ell i}^4] \leq C\alpha^2\|\theta\|^8\theta_i^2\theta_\ell\theta_s, & \text{if } i' = i,\ s' = s,\ \ell' = \ell; \\ C\alpha^2\|\theta\|^8 \cdot \mathbb{E}[W_{is}^2W_{\ell i}^2W_{\ell' i}^2] \leq C\alpha^2\|\theta\|^8\theta_i^3\theta_\ell\theta_s\theta_{\ell'}, & \text{if } i' = i,\ s' = s,\ \ell' \neq \ell; \\ C\alpha^2\|\theta\|^8 \cdot \mathbb{E}[W_{is}^3W_{\ell i}^3] \leq C\alpha^2\|\theta\|^8\theta_i^2\theta_\ell\theta_s, & \text{if } i' = i,\ s' = \ell,\ \ell' = s; \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

It follows that

$$\begin{aligned}\mathrm{Var}(Z_{4c}^*) &\le \frac{C\alpha^2\|\theta\|^8}{\|\theta\|_1^4}\Big(\sum_{i,\ell,s}\theta_i^2\theta_\ell\theta_s + \sum_{i,\ell,s,\ell'}\theta_i^3\theta_\ell\theta_s\theta_{\ell'}\Big)\\ &\le \frac{C\alpha^2\|\theta\|^8}{\|\theta\|_1^4}\big(\|\theta\|^2\|\theta\|_1^2 + \|\theta\|_3^3\|\theta\|_1^3\big)\\ &\le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1},\end{aligned}$$

where, to get the last line, we have used $\|\theta\|^2 \ll \|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$. Regarding the variance of $Z_{4c}^\dagger$, we note that $W_{is}W_{\ell t}W_{\ell i}$ and $W_{i's'}W_{\ell't'}W_{\ell'i'}$ are correlated only when the two undirected paths s-i-ℓ-t and s'-i'-ℓ'-t' are the same. Mimicking the argument in (85) or (90), we can derive that

$$\begin{aligned}\mathrm{Var}(Z_{4c}^\dagger) &\le \frac{C}{v^2}\sum_{\substack{i,\ell(dist)\\ s\notin\{i,\ell\},t\notin\{\ell,i\}}}\beta_{i\ell}^2\cdot\mathrm{Var}(W_{is}W_{\ell t}W_{\ell i})\\ &\le \frac{C\alpha^2\|\theta\|^8}{\|\theta\|_1^4}\sum_{i,\ell,s,t}\theta_i^2\theta_\ell^2\theta_s\theta_t\\ &\le \frac{C\alpha^2\|\theta\|^{12}}{\|\theta\|_1^2}.\end{aligned}$$

Since $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$, the variance of $Z_{4c}^\dagger$ is dominated by the variance of Z_{4c}^*. Since $\|\theta\| \to \infty$, we have $\|\theta\|_3^3 \gg 1/\|\theta\|_1$; it follows that the variance of $\widetilde{Z}_{4c}$ is dominated by the variance of Z_{4c}^*. Combining the above gives

$$\mathrm{Var}(Z_{4c}) \le 3\mathrm{Var}(\widetilde{Z}_{4c}) + 3\mathrm{Var}(Z_{4c}^*) + 3\mathrm{Var}(Z_{4c}^\dagger) \le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1}. \tag{121}$$

We combine (117), (118) and (120) to get

$$\big|\mathbb{E}[Z_4]\big| \le C\alpha\|\theta\|^4 = o(\alpha^4\|\theta\|^8).$$

We then combine (117), (118) and (121). Since $\|\theta\|_3^6 \le (\theta_{\max}^2\|\theta\|_1)(\theta_{\max}\|\theta\|^2) = o(\|\theta\|_1\|\theta\|^2)$, the variance of Z_{4b} is negligible compared to the variances of Z_{4a} and Z_{4c}. It follows that

$$\mathrm{Var}(Z_4) \le \frac{C\alpha^2\|\theta\|^8\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8).$$

This proves the claims of Z_4.

Next, we analyze Z_5. By plugging in the definition of δ_{ij}, we have

$$\begin{aligned}Z_5 = &\sum_{i,j,k,\ell(dist)}\eta_i(\eta_j-\tilde{\eta}_j)\eta_j(\eta_k-\tilde{\eta}_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} + \sum_{i,j,k,\ell(dist)}\eta_i(\eta_j-\tilde{\eta}_j)^2\eta_k\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\\ &+\sum_{i,j,k,\ell(dist)}(\eta_i-\tilde{\eta}_i)\eta_j^2(\eta_k-\tilde{\eta}_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} + \sum_{i,j,k,\ell(dist)}(\eta_i-\tilde{\eta}_i)\eta_j(\eta_j-\tilde{\eta}_j)\eta_k\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\\ = &\,2\sum_{i,j,k,\ell(dist)}\eta_i(\eta_j-\tilde{\eta}_j)\eta_j(\eta_k-\tilde{\eta}_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} + \sum_{i,j,k,\ell(dist)}\eta_i(\eta_j-\tilde{\eta}_j)^2\eta_k\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\\ &+\sum_{i,j,k,\ell(dist)}(\eta_i-\tilde{\eta}_i)\eta_j^2(\eta_k-\tilde{\eta}_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\end{aligned}$$

(122)

$$\equiv Z_{5a} + Z_{5b} + Z_{5c}.$$

First, we study Z_{5a}. By definition, $(\tilde{\eta}_i - \eta_i)$ has the expression in (77). It follows that

$$\begin{aligned} Z_{5a} &= 2 \sum_{i,j,k,\ell(dist)} \eta_i \Big(-\frac{1}{\sqrt{v}} \sum_{s\neq j} W_{js}\Big) \eta_j \Big(-\frac{1}{\sqrt{v}} \sum_{t\neq k} W_{kt}\Big) \widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} \\ &= \frac{2}{v} \sum_{\substack{j,k(dist)\\ s\neq j, t\neq k}} \Big(\sum_{i,\ell(dist)\notin\{j,k\}} \eta_i\eta_j \widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}\Big) W_{js}W_{kt} \\ &\equiv \frac{2}{v} \sum_{\substack{j,k(dist)\\ s\neq j, t\neq k}} \beta_{jk} W_{js}W_{kt}, \end{aligned}$$

where

$$|\beta_{jk}| \leq \sum_{i,\ell(dist)\notin\{j,k\}} |\eta_i\eta_j\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}| \leq \sum_{i,\ell} (C\theta_i\theta_j)(C\alpha^2\theta_k\theta_\ell^2\theta_i) \leq C\alpha^2\|\theta\|^4\theta_j\theta_k. \tag{123}$$

In Z_{5a}, the summand has a nonzero mean only if $(s,t) = (k,j)$. We further decompose Z_{5a} into

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

Only the first term has a nonzero mean. By (80) and (123), we have

$$\big|\mathbb{E}[Z_{5a}]\big| = \big|\mathbb{E}[\widetilde{Z}_{5a}]\big| \leq \frac{C}{\|\theta\|_1^2}\sum_{j,k}(\alpha^2\|\theta\|^4\theta_j\theta_k)(\theta_j\theta_k) \leq \frac{C\alpha^2\|\theta\|^8}{\|\theta\|_1^2}. \tag{124}$$

We then compute the variances. In each of $\widetilde{Z}_{5a}$ and Z_{5a}^*, two summands are uncorrelated unless they are exactly the same variables (e.g., when $(j',k') = (k,j)$ in $\widetilde{Z}_{5a}$). Mimicking the argument in (85) or (90), we can derive that

$$\mathrm{Var}(\widetilde{Z}_{5a}) \leq \frac{C}{v^2}\sum_{j,k(dist)} \beta_{jk}^2\,\mathrm{Var}(W_{jk}^2) \leq \frac{C\alpha^4\|\theta\|^8}{\|\theta\|_1^4}\sum_{j,k}(\theta_j^2\theta_k^2)\theta_j\theta_k \leq \frac{C\alpha^4\|\theta\|^8\|\theta\|_3^6}{\|\theta\|_1^4},$$

$$\mathrm{Var}(Z_{5a}^*) \leq \frac{C}{v^2}\sum_{\substack{j,k(dist)\\ s\neq j, t\neq k,\\ (s,t)\neq(k,j)}} \beta_{jk}^2\,\mathrm{Var}(W_{js}W_{kt}) \leq \frac{C\alpha^4\|\theta\|^8}{\|\theta\|_1^4}\sum_{j,k}(\theta_j^2\theta_k^2)\theta_j\theta_s\theta_k\theta_t \leq \frac{C\alpha^4\|\theta\|^8\|\theta\|_3^6}{\|\theta\|_1^2}.$$

It immediately leads to

$$\mathrm{Var}(Z_{5a}) \leq 2\mathrm{Var}(\widetilde{Z}_{5a}) + 2\mathrm{Var}(Z_{5a}^*) \leq \frac{C\alpha^4\|\theta\|^8\|\theta\|_3^6}{\|\theta\|_1^2}. \tag{125}$$

Second, we study Z_{5b}. It is seen that

$$Z_{5b} = \sum_{i,j,k,\ell(dist)} \eta_i \Big(-\frac{1}{\sqrt{v}}\sum_{s\neq j} W_{js}\Big)\Big(-\frac{1}{\sqrt{v}}\sum_{t\neq j} W_{jt}\Big)\eta_k\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}$$

$$= \frac{1}{v} \sum_{j, s \neq j, t \neq j} \Big(\sum_{i,k,\ell(dist) \notin \{j\}} \eta_i \eta_k \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i} \Big) W_{js} W_{jt}$$

$$\equiv \frac{1}{v} \sum_{j, s \neq j, t \neq j} \beta_j W_{js} W_{jt},$$

where

$$|\beta_j| \leq \sum_{i,k,\ell(dist) \notin \{j\}} |\eta_i \eta_k \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i}| \leq \sum_{i,k,\ell} (C\theta_i\theta_k)(C\alpha^2 \theta_i \theta_k \theta_\ell^2) \leq C\alpha^2 \|\theta\|^6. \tag{126}$$

In Z_{5b}, the summand has a nonzero mean only if $s = t$. We further decompose Z_{5b} into

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

Only $\widetilde{Z}_{5b}$ has a nonzero mean. By (80) and (126),

$$\big|\mathbb{E}[Z_{5b}]\big| = \big|\mathbb{E}[\widetilde{Z}_{5b}]\big| \leq \frac{C}{\|\theta\|_1^2} \sum_{j,s} (\alpha^2 \|\theta\|^6) \theta_j \theta_s \leq C\alpha^2 \|\theta\|^6. \tag{127}$$

To compute the variance, we note that in each of $\widetilde{Z}_{5b}$ and Z_{5b}^*, two summands are uncorrelated unless they are exactly the same random variables (e.g., when $\{j', s'\} = \{s, j\}$ in $\widetilde{Z}_{5b}$, and when $j' = j$ and $\{s', t'\} = \{s, t\}$ in Z_{5b}^*). Mimicking the argument in (85) or (90), we can derive that

$$\mathrm{Var}(\widetilde{Z}_{5b}) \leq \frac{C}{v^2} \sum_{j,s(dist)} \beta_j^2 \,\mathrm{Var}(W_{js}^2) \leq \frac{C\alpha^4 \|\theta\|^{12}}{\|\theta\|_1^4} \sum_{j,s} \theta_j \theta_s \leq \frac{C\alpha^4 \|\theta\|^{12}}{\|\theta\|_1^2},$$

$$\mathrm{Var}(Z_{5b}^*) \leq \frac{C}{v^2} \sum_{\substack{j \\ s,t(dist) \notin \{j\}}} \beta_j^2 \,\mathrm{Var}(W_{js} W_{jt}) \leq \frac{C\alpha^4 \|\theta\|^{12}}{\|\theta\|_1^4} \sum_{j,s,t} \theta_j^2 \theta_s \theta_t \leq \frac{C\alpha^4 \|\theta\|^{14}}{\|\theta\|_1^2}.$$

Combining the above gives

$$\mathrm{Var}(Z_{5b}) \leq 2\mathrm{Var}(\widetilde{Z}_{5b}) + 2\mathrm{Var}(Z_{5b}^*) \leq \frac{C\alpha^4 \|\theta\|^{14}}{\|\theta\|_1^2}. \tag{128}$$

Third, we study Z_{5c}. If we relabel $(i, j, k, \ell) = (j, i, k, \ell)$, then Z_{5c} becomes

$$Z_{5c} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_j - \tilde{\eta}_j) \eta_i^2 (\eta_k - \tilde{\eta}_k) \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell j} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i^2 \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell j})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k),$$

where $|\eta_i^2 \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell j}| \leq C\alpha^2 \theta_i^2 \theta_j \theta_k \theta_\ell^2$. At the same time, we recall that

$$Z_{5a} = 2 \sum_{\substack{i,j,k,\ell \\ (dist)}} \eta_i (\eta_j - \tilde{\eta}_j) \eta_j (\eta_k - \tilde{\eta}_k) \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i \eta_j \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k),$$

where $|\eta_i \eta_j \widetilde{\Omega}_{k\ell} \widetilde{\Omega}_{\ell i}| \leq C\alpha^2 \theta_i^2 \theta_j \theta_k \theta_\ell^2$. It is easy to see that Z_{5c} has a similar structure as Z_{5c}. As a result, from (124)-(125), we immediately have

$$\big|\mathbb{E}[Z_{5c}]\big| \leq \frac{C\alpha^2 \|\theta\|^8}{\|\theta\|_1^2}, \qquad \mathrm{Var}(Z_{5c}) \leq \frac{C\alpha^4 \|\theta\|^8 \|\theta\|_3^6}{\|\theta\|_1^2}. \tag{129}$$

We now combine the results for Z_{5a}-Z_{5c}. Since $\|\theta\|^2 \leq \theta_{\max}\|\theta\|_1 \ll \|\theta\|_1^2$, $\mathbb{E}[Z_{5a}]$ and $\mathbb{E}[Z_{5c}]$ are of a smaller order than the the right hand side of (127). Since $\|\theta\|_3^6 \leq \theta_{\max}^2\|\theta\|^4 \ll \|\theta\|^6$, $\mathrm{Var}(Z_{5a})$ and $\mathrm{Var}(Z_{5c})$ are of a smaller order than the right hand side of (128). It follows that

$$\big|\mathbb{E}[Z_5]\big| \leq C\alpha^2\|\theta\|^6 = o(\alpha^4\|\theta\|^8), \qquad \mathrm{Var}(Z_5) \leq \frac{C\alpha^4\|\theta\|^{14}}{\|\theta\|_1^2} = o(\alpha^6\|\theta\|^8\|\theta\|_3^6).$$

We briefly explain why $\mathrm{Var}(Z_5) = o(\alpha^6\|\theta\|^8\|\theta\|_3^6)$: since $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$, we immediately have $\|\theta\|^{14} \leq \|\theta\|^6(\|\theta\|_1\|\theta\|_3^3)^2$; it follows that the bound for $\mathrm{Var}(Z_5)$ is $\leq C\alpha^4\|\theta\|^6\|\theta\|_3^6$; note that $\alpha\|\theta\| \to \infty$, we immediately have $\alpha^4\|\theta\|^6\|\theta\|_3^6 = o(\alpha^6\|\theta\|^8\|\theta\|_3^6)$. This proves the claims of Z_5.

Last, we analyze Z_6. Plugging in the definition of δ_{ij}, we have

$$\begin{aligned} Z_6 = & \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)\widetilde{\Omega}_{\ell i} + \sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell\widetilde{\Omega}_{\ell i} \\ & + \sum_{i,j,k,\ell(dist)} (\eta_i - \tilde{\eta}_i)\eta_j\widetilde{\Omega}_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)\widetilde{\Omega}_{\ell i} + \sum_{i,j,k,\ell(dist)} (\eta_i - \tilde{\eta}_i)\eta_j\widetilde{\Omega}_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell\widetilde{\Omega}_{\ell i} \\ = & \, 2\sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}\eta_k(\eta_\ell - \tilde{\eta}_\ell)\widetilde{\Omega}_{\ell i} + 2\sum_{i,j,k,\ell(dist)} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell\widetilde{\Omega}_{\ell i} \\ \equiv & \, Z_{6a} + Z_{6b}. \end{aligned}$$

By relabeling (i,j,k,ℓ) as (i,j,ℓ,k), we can write

$$Z_{6a} = 2\sum_{\substack{i,j,k,\ell \\ (dist)}} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{j\ell}\eta_\ell(\eta_k - \tilde{\eta}_k)\widetilde{\Omega}_{ki} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i\eta_\ell\widetilde{\Omega}_{j\ell}\widetilde{\Omega}_{ki})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k),$$

where $|\eta_i\eta_\ell\widetilde{\Omega}_{j\ell}\widetilde{\Omega}_{ki}| \leq C\alpha^2\theta_i^2\theta_j\theta_k\theta_\ell^2$. Also, we write

$$Z_{6b} = 2\sum_{\substack{i,j,k,\ell \\ (dist)}} \eta_i(\eta_j - \tilde{\eta}_j)\widetilde{\Omega}_{jk}(\eta_k - \tilde{\eta}_k)\eta_\ell\widetilde{\Omega}_{\ell i} = 2\sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i\eta_\ell\widetilde{\Omega}_{jk}\widetilde{\Omega}_{\ell i})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k).$$

where $|\eta_i\eta_\ell\widetilde{\Omega}_{jk}\widetilde{\Omega}_{\ell i}| \leq C\alpha^2\theta_i^2\theta_j\theta_k\theta_\ell^2$. At the same time, we recall that

$$Z_{5a} = 2\sum_{\substack{i,j,k,\ell \\ (dist)}} \eta_i(\eta_j - \tilde{\eta}_j)\eta_j(\eta_k - \tilde{\eta}_k)\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i} = \sum_{\substack{i,j,k,\ell \\ (dist)}} (\eta_i\eta_j\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i})(\eta_j - \tilde{\eta}_j)(\eta_k - \tilde{\eta}_k),$$

where $|\eta_i\eta_j\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}| \leq C\alpha^2\theta_i^2\theta_j\theta_k\theta_\ell^2$. It is clear that both Z_{6a} and Z_{6b} have a similar structure as Z_{5a}. From (124)-(125), we immediately have

$$\big|\mathbb{E}[Z_6]\big| \leq \frac{C\alpha^2\|\theta\|^8}{\|\theta\|_1^2} = o(\alpha^4\|\theta\|^8), \qquad \mathrm{Var}(Z_6) \leq \frac{C\alpha^4\|\theta\|^8\|\theta\|_3^6}{\|\theta\|_1^2} = o(\|\theta\|^8).$$

This proves the claims of Z_6.

G.4.8. *Proofs of Lemmas G.8 and G.9.* Recall that $\lambda_1, \lambda_2, \ldots, \lambda_K$ are all the nonzero eigenvalues of Ω, arranged in the descending order in magnitude. Write for short $\alpha = |\lambda_2|/|\lambda_1|$. We shall repeatedly use the following results, which are proved in (74), (80), and (81):

$$v \asymp \|\theta\|_1^2, \qquad 0 < \eta_i < C\theta_i, \qquad |\widetilde{\Omega}_{ij}| \leq C\alpha\theta_i\theta_j.$$

Recall that $U_c = 4T_1 + F$, under the null hypothesis; $U_c = 4T_1 + 4T_2 + F$ under the alternative hypothesis. By definition,

$$T_1 = \sum_{i_1,i_2,i_3,i_4(dist)} \delta_{i_1i_2}\delta_{i_2i_3}\delta_{i_3i_4}W_{i_4i_1},$$

$$T_2 = \sum_{i_1,i_2,i_3,i_4(dist)} \delta_{i_1i_2}\delta_{i_2i_3}\delta_{i_3i_4}\widetilde{\Omega}_{i_4i_1},$$

$$F = \sum_{i_1,i_2,i_3,i_4(dist)} \delta_{i_1i_2}\delta_{i_2i_3}\delta_{i_3i_4}\delta_{i_4i_1},$$

where $\delta_{ij} = \eta_i(\eta_j - \widetilde{\eta}_j) + \eta_j(\eta_i - \widetilde{\eta}_i)$, for $1 \le i, j \le n$, $i \ne j$. By symmetry and elementary algebra, we further write

$$T_1 = 2T_{1a} + 2T_{1b} + 2T_{1c} + 2T_{1d}, \tag{130}$$

where

$$T_{1a} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_1} - \tilde{\eta}_{i_1})(\eta_{i_2} - \tilde{\eta}_{i_2})(\eta_{i_3} - \tilde{\eta}_{i_3})\big] \cdot W_{i_4i_1},$$

$$T_{1b} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}^2\big[(\eta_{i_1} - \tilde{\eta}_{i_1})(\eta_{i_2} - \tilde{\eta}_{i_2})(\eta_{i_4} - \tilde{\eta}_{i_4})\big] \cdot W_{i_4i_1},$$

$$T_{1c} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_2} - \tilde{\eta}_{i_2})^2(\eta_{i_3} - \tilde{\eta}_{i_3})\big] \cdot W_{i_4i_1},$$

$$T_{1d} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}^2\big[(\eta_{i_2} - \tilde{\eta}_{i_2})^2(\eta_{i_4} - \tilde{\eta}_{i_4})\big] \cdot W_{i_4i_1}.$$

Similarly, we write

$$T_2 = 2T_{2a} + 2T_{2b} + 2T_{2c} + 2T_{2d}, \tag{131}$$

where

$$T_{2a} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_1} - \tilde{\eta}_{i_1})(\eta_{i_2} - \tilde{\eta}_{i_2})(\eta_{i_3} - \tilde{\eta}_{i_3})\big] \cdot \widetilde{\Omega}_{i_4i_1},$$

$$T_{2b} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}^2\big[(\eta_{i_1} - \tilde{\eta}_{i_1})(\eta_{i_2} - \tilde{\eta}_{i_2})(\eta_{i_4} - \tilde{\eta}_{i_4})\big] \cdot \widetilde{\Omega}_{i_4i_1},$$

$$T_{2c} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_2} - \tilde{\eta}_{i_2})^2(\eta_{i_3} - \tilde{\eta}_{i_3})\big] \cdot \widetilde{\Omega}_{i_4i_1},$$

$$T_{2d} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}^2\big[(\eta_{i_2} - \tilde{\eta}_{i_2})^2(\eta_{i_4} - \tilde{\eta}_{i_4})\big] \cdot \widetilde{\Omega}_{i_4i_1}.$$

Also, similarly, we have

$$F = 2F_a + 12F_b + 2F_c, \tag{132}$$

where

$$F_a = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_2}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_1} - \tilde{\eta}_{i_1})(\eta_{i_2} - \tilde{\eta}_{i_2})(\eta_{i_3} - \tilde{\eta}_{i_3})(\eta_{i_4} - \tilde{\eta}_{i_4})\big],$$

$$F_b = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}^2\eta_{i_4}\big[(\eta_{i_1}-\tilde{\eta}_{i_1})^2(\eta_{i_2}-\tilde{\eta}_{i_2})(\eta_{i_4}-\tilde{\eta}_{i_4})\big],$$

$$F_c = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}^2\eta_{i_4}^2\big[(\eta_{i_1}-\tilde{\eta}_{i_1})^2(\eta_{i_3}-\tilde{\eta}_{i_3})^2\big].$$

To show the lemmas, it is sufficient to show the following 11 items (a)-(k), corresponding to $T_{1a}, T_{1b}, T_{1c}, T_{1d}, T_{2a}, T_{2b}, T_{2c}, T_{2d}, F_a, F_b, F_c$, respectively. Item (a) claims that both under the null and the alternative,

$$|\mathbb{E}[T_{1a}]| \leq C\|\theta\|^6/\|\theta\|_1^2, \qquad \mathrm{Var}(T_{1a}) \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2. \tag{133}$$

Item (b) claims that both under the null and the alternative,

$$|\mathbb{E}[T_{1b}]| \leq C\|\theta\|^6/\|\theta\|_1^2, , \qquad \mathrm{Var}(T_{1b}) \leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1. \tag{134}$$

Item (c) claims that both under the null and the alternative,

$$\mathbb{E}[T_{1c}] = 0, \qquad \mathrm{Var}(T_{1c}) \leq C\|\theta\|_3^9/\|\theta\|_1, \tag{135}$$

Item (d) claims that

$$\mathbb{E}[T_{1d}] \asymp -\|\theta\|^4 \text{ under the null,}$$

$$|\mathbb{E}[T_{1d}]| \leq C\|\theta\|^4 \text{ under the alternative,} \tag{136}$$

and that both under the null and the alternative,

$$\mathrm{Var}(T_{1d}) \leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1. \tag{137}$$

Next, for item (e)-(h), we recall that under the null, $T_2 = 0$, and correspondingly $T_{2a} = T_{2b} = T_{2c} = T_{2d} = 0$, so we only need to consider the alternative. Recall that $\alpha = |\lambda_2/\lambda_1|$. Item (e) claims that under the alternative,

$$\mathbb{E}[T_{2a}] = 0, \qquad \mathrm{Var}(T_{2a}) \leq C\alpha^2 \cdot \|\theta\|^4\|\theta\|_3^9/\|\theta\|_1^3. \tag{138}$$

Item (f) claims that under the alternative,

$$\mathbb{E}[T_{2b}] = 0, \qquad \mathrm{Var}(T_{2b}) \leq C\alpha^2 \cdot \|\theta\|^{12}\|\theta\|_3^3/\|\theta\|_1^5, \tag{139}$$

Item (g) claims that under the alternative,

$$|\mathbb{E}[T_{2c}]| \leq C\alpha\|\theta\|^6/\|\theta\|_1^3, \qquad \mathrm{Var}(T_{2c}) \leq C\alpha^2 \cdot \|\theta\|^8\|\theta\|_3^3/\|\theta\|_1. \tag{140}$$

Item (h) claims that both under the null and the alternative,

$$|\mathbb{E}[T_{2d}]| \leq C\alpha\|\theta\|^6/\|\theta\|_1^3, \qquad \mathrm{Var}(T_{2d}) \leq C\alpha^2 \cdot \|\theta\|^8\|\theta\|_3^3/\|\theta\|_1. \tag{141}$$

Finally, for items (i)-(k). Item (i) claims that both under the null and the alternative,

$$|\mathbb{E}[F_a]| \leq C\|\theta\|^8/\|\theta\|_1^4, \qquad \mathrm{Var}(F_a) \leq C\|\theta\|_3^{12}/\|\theta\|_1^4. \tag{142}$$

Item (j) claims that both under the null and the alternative,

$$|\mathbb{E}[F_b]| \leq C\|\theta\|^6/\|\theta\|_1^2, \qquad \mathrm{Var}(F_b) \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2. \tag{143}$$

Item (k) claims that

$$\mathbb{E}[F_c] \asymp \|\theta\|^4 \text{ under the null,}$$

$$|\mathbb{E}[F_c]| \leq C\|\theta\|^4 \text{ under the alternative,} \tag{144}$$

and that under both under the null and the alternative,

$$\mathrm{Var}(F_3) \leq C\|\theta\|^{10}/\|\theta\|_1^2. \tag{145}$$

We now show Lemmas G.4 and G.5 follow once (a)-(k) are proved. In detail, first, we note that $\|\theta\|^6/\|\theta\|_1^2 = o(\|\theta\|^4)$. Inserting (136) and the first equation in each of (133)-(135) into (130) gives that

$$\mathbb{E}[T_1] \asymp -2\|\theta\|^4 \text{ under the null}, \qquad |\mathbb{E}[T_1]| \leq C\|\theta\|^4 \text{ under the alternative},$$

and inserting (137) and the second equation in each of (133)-(135) into (130) gives that both under the null and the alternative,

$$\mathrm{Var}(T_1) \leq C[\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2 + \|\theta\|^6\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|_3^9/\|\theta\|_1 + \|\theta\|^6\|\theta\|_3^3/\|\theta\|_1],$$

where since $\|\theta\|_3^3/\|\theta\|^2 = o(1)$ and $\|\theta\|^2/\|\theta\|_1 = o(1)$, the right hand side

$$\leq C[\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1^2 + \|\theta\|^6\|\theta\|_3^3/\|\theta\|_1] \leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1.$$

Second, inserting the first equation in each of (138)-(141) into (131) gives that under the alternative (recall that $T_2 = 0$ under the null),

$$|\mathbb{E}[T_2]| \leq C\alpha\|\theta\|^6/\|\theta\|_1^3,$$

and inserting the second equation in each of (138)-(141) into (131) gives

$$\mathrm{Var}(T_2) \leq C\alpha^2[\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|^{12}\|\theta\|_3^3/\|\theta\|_1^5] \leq C\alpha^2\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1,$$

where we have used $\|\theta\|^2 = o(\|\theta\|_1^2)$. Third, note that $\|\theta\|^8/\|\theta\|_1^4 = o(\|\theta\|^4)$ and $\|\theta\|^6/\|\theta\|_1^2 = o(\|\theta\|^4)$. Inserting (144) and the first equation in each of (142)-(143) into (132) gives

$$\mathbb{E}[F] \sim 2\|\theta\|^4 \text{ under the null}, \qquad |\mathbb{E}[F]| \leq C\|\theta\|^4 \text{ under the alternative},$$

and inserting (145) and the second equation in each of (142)-(143) into (132) gives that both under the null and the alternative,

$$\mathrm{Var}(F) \leq C[\|\theta\|_3^{12}/\|\theta\|_1^4 + \|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2 + \|\theta\|^{10}/\|\theta\|_1^2] \leq C\|\theta\|^{10}/\|\theta\|_1^2,$$

where we have used $\|\theta\|_3^3 \ll \theta\|^2 \ll \|\theta\|_1$ and $\|\theta\|_3^3/\|\theta\|^2 = o(1)$.

We now combine the above results for T_1, T_2 and F. First, since that $U_c = 4T_1 + F$ under the null, it follows that under the null,

$$\mathbb{E}[U_c] \sim -6\|\theta\|^4,$$

and

$$\mathrm{Var}(U_c) \leq C[\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|^{10}/\|\theta\|_1^2] \leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1,$$

where we have used $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$ (a direct use of Cauchy-Schwartz inequality). Second, since $U_c = 4T_1 + 4T_2 + F$ under the alternative, it follows that under the alternative,

$$|\mathbb{E}[U_c]| \leq C\|\theta\|^4,$$

and

$$\mathrm{Var}(U_c) \leq C[\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1 + \alpha^2\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|^{10}/\|\theta\|_1^2] \leq C\|\theta\|^6\|\theta\|_3^3(\alpha^2\|\theta\|^2 + 1)/\|\theta\|_1,$$

where we have used $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$ and basic algebra. Combining the above gives all the claims in Lemmas G.4 and G.5.

It remains to show the 11 items (a)-(k). We consider them separately.

Consider Item (a). The goal is to show (133). Recall that

$$T_{1a} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_1}-\tilde{\eta}_{i_1})(\eta_{i_2}-\tilde{\eta}_{i_2})(\eta_{i_3}-\tilde{\eta}_{i_3})\big]\cdot W_{i_4 i_1},$$

and that

$$\widetilde{\eta}-\eta = v^{-1/2}W1_n. \tag{146}$$

Plugging (146) into T_{11} gives

$$\begin{aligned} T_{1a} &= -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}\eta_{i_4}\Big(\sum_{j_1,j_1\neq i_1} W_{i_1 j_1}\Big)\Big(\sum_{j_2,j_2\neq i_2} W_{i_2 j_2}\Big)\Big(\sum_{j_3,j_3\neq i_3} W_{i_3 j_3}\Big)W_{i_4 i_1} \\ &= -\frac{1}{v^{3/2}} \sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ j_1\neq i_1, j_2\neq i_2, j_3\neq i_3}} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1 j_1}W_{i_2 j_2}W_{i_3 j_3}W_{i_1 i_4}. \end{aligned}$$

By basic combinatorics and careful observations, we have

$$W_{i_1 j_1}W_{i_2 j_2}W_{i_3 j_3}W_{i_1 i_4} = \begin{cases} W_{i_1 i_4}^2 W_{i_2 i_3}^2, & \text{if } j_1 = i_4, (j_2,j_3) = (i_3,i_2),\\ W_{i_1 i_4}^2 W_{i_2 j_2} W_{i_3 j_3}, & \text{if } j_1 = i_4, (j_2,j_3) \neq (i_3,i_2),\\ W_{i_2 i_3}^2 W_{i_1 j_1} W_{i_1 i_4}, & \text{if } j_1 \neq i_4, (j_2,j_3) = (i_3,i_2),\\ W_{i_1 i_2}^2 W_{i_3 j_3} W_{i_1 i_4}, & \text{if } (j_1,j_2) = (i_2,i_1),\\ W_{i_1 i_3}^2 W_{i_2 j_2} W_{i_1 i_4}, & \text{if } (j_1,j_3) = (i_3,i_1),\\ W_{i_1 j_1}W_{i_2 j_2}W_{i_3 j_3}W_{i_1 i_4}, & \text{otherwise.} \end{cases} \tag{147}$$

This allows us to further split T_{11} into 6 different terms:

$$T_{1a} = X_a + X_{b1} + X_{b2} + X_{b3} + X_{b4} + X_c, \tag{148}$$

where

$$X_a = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1 i_4}^2 W_{i_2 i_3}^2,$$

$$X_{b1} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_2,j_3\\ (j_2,j_3)\neq\{i_3,i_2\}}} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1 i_4}^2 W_{i_2 j_2} W_{i_3 j_3},$$

$$X_{b2} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_1(j_1\neq i_4)} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_2 i_3}^2 W_{i_1 j_1} W_{i_1 i_4},$$

$$X_{b3} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_3(j_3\neq i_3)} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1 i_2}^2 W_{i_3 j_3} W_{i_1 i_4},$$

$$X_{b4} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2(j_2\neq i_2)} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1 i_3}^2 W_{i_2 j_2} W_{i_1 i_4},$$

$$X_c = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_1,j_2,j_3\\ j_1\notin\{i_1,i_4\},(j_2,j_3)\neq(i_3,i_2)\\ (j_1,j_2)\neq(i_2,i_1),(j_1,j_3)\neq(i_3,i_1)}} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1 j_1}W_{i_2 j_2}W_{i_3 j_3}W_{i_1 i_4}.$$

We now show (133). Consider the first claim of (133). It is seen that out of the 6 terms on the right hand side of (148), the mean of all terms are 0, except for the first term. Note that for any $1 \le i, j \le n$, $i \ne j$, $\mathbb{E}[W_{ij}^2] = \Omega_{ij}(1 - \Omega_{ij})$, where Ω_{ij} are upper bounded by $o(1)$ uniformly for all such i, j. It follows

$$\begin{aligned}\mathbb{E}[X_a] &= -v^{-3/2} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}\eta_{i_4}\mathbb{E}[W_{i_1 i_4}^2]\mathbb{E}[W_{i_2 i_3}^2] \\ &= -(1 + o(1)) \cdot v^{-3/2} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}\eta_{i_4}\Omega_{i_1 i_4}\Omega_{i_2 i_3}.\end{aligned}$$

Since for any $1 \le i, j \le n$, $i \ne j$, $0 < \eta_i \le C\theta_i$, $\Omega_{ij} \le C\theta_i\theta_j$ and $v \asymp \|\theta\|_1^2$,

$$|\mathbb{E}[X_a]| \le C(\|\theta\|_1)^{-3} \sum_{i_1,i_2,i_3,i_4(dist)} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}^2 \le C\|\theta\|^6/\|\theta\|_1^2.$$

Inserting these into (148) gives

$$|\mathbb{E}[T_{1a}]| \le C\|\theta\|^6/\|\theta\|_1^2, \tag{149}$$

and the first claim of (133) follows.

Consider the second claim of (133) next. By (148) and Cauchy-Schwartz inequality,

$$\begin{aligned}\mathrm{Var}(T_{1a}) &\le C\mathrm{Var}(X_a) + \mathrm{Var}(X_{b1}) + \mathrm{Var}(X_{b2}) + \mathrm{Var}(X_{b3}) + \mathrm{Var}(X_{b4}) + \mathrm{Var}(X_c)) \\ &\le C(\mathrm{Var}(X_a) + \mathbb{E}[X_{b1}^2] + \mathbb{E}[X_{b2}^2] + \mathbb{E}[X_{b3}^2] + \mathbb{E}[X_{b4}^2] + \mathbb{E}[X_c^2]). \qquad (150)\end{aligned}$$

We now consider $\mathrm{Var}(X_a)$, $\mathbb{E}[X_{b1}^2] + \mathbb{E}[X_{b2}^2] + \mathbb{E}[X_{b3}^2] + \mathbb{E}[X_{b4}^2]$, and $\mathbb{E}[X_c^2]$, separately.

Consider $\mathrm{Var}(X_a)$. Write $\mathrm{Var}(X_a)$ as

$$v^{-3} \sum_{\substack{i_1,\cdots,i_4(dist) \\ i_1',\cdots,i_4'(dist)}} \eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'}$$

$$\mathbb{E}\big[(W_{i_1 i_4}^2 W_{i_2 i_3}^2 - \mathbb{E}[W_{i_1 i_4}^2 W_{i_2 i_3}^2])(W_{i_1' i_4'}^2 W_{i_2' i_3'}^2 - \mathbb{E}[W_{i_1' i_4'}^2 W_{i_2' i_3'}^2])\big]. \tag{151}$$

In the sum, a term is nonzero only when one of the following cases happens.

- (A). $\{W_{i_1 i_4}, W_{i_2 i_3}, W_{i_1' i_4'}, W_{i_2' i_3'}\}$ has 2 distinct random variables.
- (B). $\{W_{i_1 i_4}, W_{i_2, i_3}, W_{i_1' i_4'}, W_{i_2' i_3'}\}$ has 3 distinct random variables. This has 4 sub-cases: (B1) $W_{i_1 i_4} = W_{i_1' i_4'}$, (B2) $W_{i_1 i_4} = W_{i_2' i_3'}$, (B3) $W_{i_2 i_3} = W_{i_1' i_4'}$, and (B4) $W_{i_2 i_3} = W_{i_2' i_3'}$.

For Case (A), the two sets $\{i_1, i_2, i_3, i_4\}$ and $\{i_1', i_2', i_3', i_4'\}$ are identical. By basic statistics and independence between $W_{i_1 i_4}$ and $W_{i_2 i_3}$,

$$\begin{aligned}&\mathbb{E}[(W_{i_1 i_4}^2 W_{i_2 i_3}^2 - \mathbb{E}[W_{i_1 i_4}^2 W_{i_2 i_3}^2])(W_{i_1' i_4'}^2 W_{i_2' i_3'}^2 - \mathbb{E}[W_{i_1' i_4'}^2 W_{i_2' i_3'}^2])] \\ =&\mathbb{E}[(W_{i_1 i_4}^2 W_{i_2 i_3}^2 - \mathbb{E}[W_{i_1 i_4}^2 W_{i_2 i_3}^2])^2] \\ =&\mathbb{E}[W_{i_1 i_4}^4]\mathbb{E}[W_{i_2 i_3}^4] - (\mathbb{E}[W_{i_1 i_4}^2])^2(\mathbb{E}[W_{i_2 i_3}^2])^2 \\ \le&\mathbb{E}[W_{i_1 i_4}^4]\mathbb{E}[W_{i_2 i_3}^4], \qquad (152)\end{aligned}$$

where by basic statistics and that $\Omega_{ij} \le C\theta_i\theta_j$ for all $1 \le i, j \le n$, $i \le j$, the right hand side

$$\le C\Omega_{i_1 i_4}\Omega_{i_2 i_3} \le C\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4}.$$

Combining these with (151) and noting that $v \sim \|\theta\|_1^2$ and that $0 < \eta_i \le C\theta_i$ for all $1 \le i \le n$, the contribution of this case to $\mathrm{Var}(X_a)$ is no more than

$$C(\|\theta\|_1)^{-6} \sum_{i_1,\cdots,i_4(dist)} \sum_a \theta_{i_1}^{a_1+1}\theta_{i_2}^{a_2+2}\theta_{i_3}^{a_3+2}\theta_{i_4}^{a_4+2}, \tag{153}$$

where $a = (a_1, a_2, a_3, a_4)$ and each a_i is either 0 and 1, satisfying $a_1 + a_2 + a_3 + a_4 = 3$. Note that the right hand side of (153) is no greater than

$$C(\|\theta\|_1)^{-6} \max\{\|\theta\|_1 \|\theta\|_3^9, \|\theta\|^4 \|\theta\|_3^6\} \leq C\|\theta\|_3^9 / \|\theta\|_1^5,$$

where we have used $\|\theta\|^4 \leq \|\theta\|_1 \|\theta\|_3^3$.

Next, consider (B1). By independence between $W_{i_1 i_4}$, $W_{i_2 i_3}$, and $W_{i_2' i_3'}$ and basic algebra,

$$\begin{aligned}
&\mathbb{E}[(W_{i_1 i_4}^2 W_{i_2 i_3}^2 - \mathbb{E}[W_{i_1 i_4}^2 W_{i_2 i_3}^2])(W_{i_1' i_4'}^2 W_{i_2' i_3'}^2 - \mathbb{E}[W_{i_1' i_4'}^2 W_{i_2' i_3'}^2])] \\
=&\mathbb{E}[(W_{i_1 i_4}^2 W_{i_2 i_3}^2 - \mathbb{E}[W_{i_1 i_4}^2 W_{i_2 i_3}^2])(W_{i_1 i_4}^2 W_{i_2' i_3'}^2 - \mathbb{E}[W_{i_1 i_4}^2 W_{i_2' i_3'}^2])] \\
=&\mathbb{E}[W_{i_1 i_4}^4]\mathbb{E}[W_{i_2 i_3}^2]\mathbb{E}[W_{i_2' i_3'}^2] - (\mathbb{E}[W_{i_1 i_4}^2])^2 \mathbb{E}[W_{i_2 i_3}^2]\mathbb{E}[W_{i_2' i_3'}^2] \\
=&\mathrm{Var}(W_{i_1 i_4}^2)\mathbb{E}[W_{i_2 i_3}^2]\mathbb{E}[W_{i_2' i_3'}^2], \qquad (154)
\end{aligned}$$

where by similar arguments, the last term

$$\leq C\Omega_{i_1 i_4}\Omega_{i_2 i_3}\Omega_{i_2' i_3'} \leq C\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4}\theta_{i_2'}\theta_{i_3'}.$$

Combining this with (151) and using similar arguments, the contribution of this case to $\mathrm{Var}(X_a)$

$$\leq C(\|\theta\|_1)^{-6} \sum_{\substack{i_1, i_2, i_3, i_4 (dist) \\ i_2', i_3' (dist)}} C\theta_{i_1}^{b_1+1}\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4}^{b_2+2}\theta_{i_2'}^2\theta_{i_3'}^2, \qquad (155)$$

where similarly b_1, b_2 are either 0 or 1 and $b_1 + b_2 = 1$. By similar argument, the right hand side

$$\leq C\|\theta\|_1 \|\theta\|^8 \|\theta\|_3^3 / \|\theta\|_1^6 = C\|\theta\|^8 \|\theta\|_3^3 / \|\theta\|_1^5.$$

The discussion for (B2), (B3), and (B4) are similar so is omitted, and their contribution to $\mathrm{Var}(X_a)$ are respectively

$$\leq C\|\theta\|^8 \|\theta\|_3^3 / \|\theta\|_1^5, \qquad (156)$$

$$\leq C\|\theta\|^8 \|\theta\|_3^3 / \|\theta\|_1^5, \qquad (157)$$

and

$$\leq C\|\theta\|^4 \|\theta\|_3^6 / \|\theta\|_1^4. \qquad (158)$$

Finally, inserting (153), (155), (156), (157), and (158) into (151) gives

$$\mathrm{Var}(X_a) \leq C[\|\theta\|_3^9/\|\theta\|_1^5 + \|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^5 + \|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4] \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4, \qquad (159)$$

where we have used $\|\theta\|_3^3 \ll \|\theta\|^2$ and $\|\theta\|^4 \leq \|\theta\|_1 \|\theta\|_3^3$.

Consider $\mathbb{E}[X_{b1}^2] + \mathbb{E}[X_{b21}^2] + \mathbb{E}[X_{b3}^2] + \mathbb{E}[X_{b4}^2]$. We claim that both under the null and the alternative,

$$\mathbb{E}[X_{b1}^2] \leq C\|\theta\|^4 \|\theta\|_3^6 / \|\theta\|_1^2, \qquad (160)$$

$$\mathbb{E}[X_{b2}^2] \leq C\|\theta\|^8 \|\theta\|_3^3 / \|\theta\|_1^3, \qquad (161)$$

$$\mathbb{E}[X_{b3}^2] \leq C\|\theta\|^6 \|\theta\|_3^6 / \|\theta\|_1^4, \qquad (162)$$

$$\mathbb{E}[X_{b4}^2] \leq C\|\theta\|^6 \|\theta\|_3^6 / \|\theta\|_1^4, \qquad (163)$$

where the last two terms are seen to be negligible compared to the other two. Therefore,

$$\mathbb{E}[X_{b1}^2] + \mathbb{E}[X_{b2}^2] + \mathbb{E}[X_{b3}^2] + \mathbb{E}[X_{b4}^2] \leq C[\|\theta\|^4 \|\theta\|_3^6 / \|\theta\|_1^2 + \|\theta\|^8 \|\theta\|_3^3 / \|\theta\|_1^3], \qquad (164)$$

where since $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$ (Cauchy-Schwartz inequality) the right hand side

$$\leq C[\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2.$$

We now prove (160)-(163). Since the study for $\mathbb{E}[X_{b1}^2], \mathbb{E}[X_{b2}^2], \mathbb{E}[X_{b3}^2]$ and $\mathbb{E}[X_{b4}^2]$ are similar, we only present the proof for $\mathbb{E}[X_{b1}^2]$. Write $\mathbb{E}[X_{b1}^2]$ as

$$v^{-3} \sum_{\substack{i_1,i_2,i_3,i_4 (dist)\\ i_1',i_2',i_3',i_4' (dist)}} \sum_{\substack{j_2,j_3\\ (j_2,j_3)\neq(i_3,i_2)}} \sum_{\substack{j_2',j_3'\\ (j_2',j_3')\neq(i_3',i_2')}} \eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'} W_{i_1i_4}^2 W_{i_2j_2} W_{i_3j_3} W_{i_1'i_4'}^2 W_{i_2'j_2'} W_{i_3'j_3'}.$$

Consider the term

$$W_{i_1i_4}^2 W_{i_2j_2} W_{i_3j_3} W_{i_1'i_4'}^2 W_{i_2'j_2'} W_{i_3'j_3'}.$$

In order for the mean to be nonzero, we have two cases

- Case A. The two sets of random variables $\{W_{i_1i_4}, W_{i_2j_2}, W_{i_3j_3}\}$ and $\{W_{i_1'i_4'}, W_{i_2'j_2'}, W_{i_3'j_3'}\}$ are identical.
- Case B. The two sets $\{W_{i_2j_2}, W_{i_3j_3}\}$ and $\{W_{i_2'j_2'}, W_{i_3'j_3'}\}$ are identical.

Consider Case A. In this case, $\{i_2', i_3', i_4'\}$ are three distinct indices in $\{i_1, i_2, i_3, i_4, j_2, j_3\}$, and for some integers satisfying $0 \leq a_1, a_2, \ldots, a_6 \leq 1$, $a_1 + a_2 + \ldots + a_6 = 3$,

$$\eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'} = \eta_{i_1}^{a_1}\eta_{i_2}^{1+a_2}\eta_{i_3}^{1+a_3}\eta_{i_4}^{1+a_4}\eta_{j_2}^{a_5}\eta_{j_3}^{a_6}$$

and for some integers satisfying $0 \leq b_1, b_2, b_3 \leq 1$, and $b_1 + b_2 + b_3 = 1$,

$$W_{i_1i_4}^2 W_{i_2j_2} W_{i_3j_3} W_{i_1'i_4'}^2 W_{i_2'j_2'} W_{i_3'j_3'} = W_{i_1i_4}^{b_1+3} W_{i_2j_2}^{b_2+2} W_{i_3j_3}^{b_3+2}.$$

Similarly, by $v \sim \|\theta\|_1^2$, $0 < \eta_i \leq C\theta_i$, and uniformly for all b_1, b_2, b_3 above,

$$0 < \mathbb{E}[W_{i_1i_4}^{b_1+3} W_{i_2j_2}^{b_2+2} W_{i_3j_3}^{b_3+2}] \leq C\Omega_{i_1i_4}\Omega_{i_2j_2}\Omega_{i_3j_3} \leq C\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4}\theta_{j_2}\theta_{j_3}.$$

Therefore under both the null and the alternative, the contribution of Case A to the variance is

$$\leq C(\|\theta\|_1)^{-6} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_2,j_3\\ j_2\neq i_2, j_3\neq i_3, (j_2,j_3)\neq(i_3,i_2)}} [\sum_a \theta_{i_1}^{a_1+1}\theta_{i_2}^{a_2+2}\theta_{i_3}^{a_3+2}\theta_{i_4}^{a_4+2}\theta_{j_2}^{a_5+1}\theta_{j_3}^{a_6+1}], \quad (165)$$

where $a = (a_1, a_2, \ldots, a_6)$ and a_i satisfies the above constraints. Note that the right hand size

$$\leq C(\|\theta\|_1)^{-6} \cdot \max\{\|\theta\|_1^3\|\theta\|_3^9, \|\theta\|_1^2\|\theta\|^4\|\theta\|_3^6, \|\theta\|_1\|\theta\|^8\|\theta\|_3^3, \|\theta\|^{12}\} \leq C\|\theta\|_3^9/\|\theta\|_1^3.$$

Here in the last inequality we have used $\|\theta\|^2 \leq \sqrt{\|\theta\|_1\|\theta\|_3^3}$.

Consider Case B. In this case, $\{i_2, i_3, j_2, j_3\} = \{i_2', i_3', j_2', j_3'\}$, and for some integers $0 \leq c_1, c_2, c_3, c_4 \leq 1$, $c_1 + c_2 + c_3 + c_4 = 2$,

$$\eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'} = \eta_{i_2}^{c_1+1}\eta_{i_3}^{c_2+1}\eta_{i_4}\eta_{j_2}^{c_3}\eta_{j_3}^{c_4}\eta_{i_4'},$$

and

$$W_{i_1i_4}^2 W_{i_2j_2} W_{i_3j_3} W_{i_1'i_4'}^2 W_{i_2'j_2'} W_{i_3'j_3'} = W_{i_1i_4}^2 W_{i_2j_2}^2 W_{i_3j_3}^2 W_{i_1'i_4'}^2,$$

where the four W terms on the right are independent of each other. Similarly, by $v \sim \|\theta\|_1^2$, $0 < \eta_i \leq C\theta_i$,

$$0 < \mathbb{E}[W_{i_1i_4}^2 W_{i_2j_2}^2 W_{i_3j_3}^2 W_{i_1'i_4'}^2] \leq C\Omega_{i_1i_4}\Omega_{i_2j_2}\Omega_{i_3j_3}\Omega_{i_1'i_4'} \leq C\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4}\theta_{j_2}\theta_{j_3}\theta_{i_1'}\theta_{i_4'},$$

we have that under both the null and the alternative, the contribution of Case B to the variance

$$\leq C(\|\theta\|_1)^{-6} \sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ i_1',i_4'(dist)}} \sum_{\substack{j_2,j_3\\ (j_2,j_3)\neq(i_3,i_2)}} \theta_{i_1}\theta_{i_2}^{c_1+2}\theta_{i_3}^{c_2+2}\theta_{i_4}^2\theta_{j_2}^{c_3+1}\theta_{j_3}^{c_4+1}\theta_{i_1'}\theta_{i_4'}^2,$$

where the right hand size

(166)
$$\leq C(\|\theta\|_1)^{-6}\cdot\|\theta\|_1^2\|\theta\|^4\cdot\max\{\|\theta\|_1^2\|\theta\|_3^6, \|\theta\|_1\|\theta\|^4\|\theta\|_3^3, \|\theta\|^8\}\leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2.$$

Here we have again used $\|\theta\|^2\leq\sqrt{\|\theta\|_1\|\theta\|_3^3}$.

Finally, combining (165) and (166) gives

$$\mathbb{E}[X_{b1}^2]\leq C(\|\theta\|_3^9/\|\theta\|_1^3+\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2)\leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2,$$

which proves (160).

Consider $\mathbb{E}[X_c^2]$. Consider the terms in the sum,

$$\eta_{i_2}\eta_{i_3}\eta_{i_4}W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}W_{i_1i_4}, \qquad \text{and} \qquad \eta_{i_2'}\eta_{i_3'}\eta_{i_4'}W_{i_1'j_1'}W_{i_2'j_2'}W_{i_3'j_3'}W_{i_1'i_4'}.$$

Each term has a mean 0, and two terms are uncorrelated with each other if only if the two sets of random variables $\{W_{i_1j_1}, W_{i_2j_2}, W_{i_3j_3}, W_{i_1i_4}\}$ and $\{W_{i_1'j_1'}, W_{i_2'j_2'}, W_{i_3'j_3'}, W_{i_1'i_4'}\}$ are identical (however, it is possible that $W_{i_1j_1}$ does not equal to $W_{i_1j_1'}$ but equals to $W_{i_2'j_2'}$, say). Additionally, the indices $i_2', i_3', i_4'\in\{i_1,i_2,i_3,i_4,j_1,j_2,j_3\}$, and it follows

$$\mathbb{E}[X_c^2]\leq Cv^{-3}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{\substack{j_1,j_2,j_3\\ j_1\notin\{i_1,i_4\},(j_1,j_3)\neq(i_3,i_1)\\ (j_2,j_3)\neq(i_3,i_2),(j_2,j_1)\neq(i_2,i_1)}} [\sum_a \eta_{i_1}^{a_1}\eta_{i_2}^{a_2+1}\eta_{i_3}^{a_3+1}\eta_{i_4}^{a_4+1}\eta_{j_1}^{a_5}\eta_{j_2}^{a_6}\eta_{j_3}^{a_7}]\cdot\mathbb{E}[W_{i_1j_1}^2W_{i_2j_2}^2W_{i_3j_3}^2W_{i_1j_1}^2],$$

where $a=(a_1,a_2,\cdots,a_7)$ and the power $0\leq a_1,a_2,\cdots,a_7\leq 1$, and $a_1+a_2+\cdots+a_7=3$. Note that $W_{i_1j_1}, W_{i_2j_2}, W_{i_3j_3}$ and $W_{i_1i_4}$ are independent and $\mathbb{E}(W_{ij}^2)\leq\Omega_{ij}\leq C\theta_i\theta_j$, $1\leq i,j\leq n$, $i\neq j$,

$$\mathbb{E}[W_{i_1j_1}^2W_{i_2j_2}^2W_{i_3j_3}^2W_{i_1i_4}^2]\leq\Omega_{i_1j_1}\Omega_{i_2j_2}\Omega_{i_3j_3}\Omega_{i_1i_4}\leq C\theta_{i_1}^2\theta_{i_2}\theta_{i_3}\theta_{i_4}\theta_{j_1}\theta_{j_2}\theta_{j_3}.$$

Also, recall that both under the null and the alternative, $v\asymp\|\theta\|_1^2$ and $0<\eta_i\leq C\theta_i$, $1\leq i\leq n$. Combining these gives

$$\mathbb{E}[X_c^2]\leq C(\|\theta\|_1)^{-6}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{\substack{j_1,j_2,j_3\\ j_1\notin\{i_1,i_4\},(j_1,j_3)\neq(i_3,i_1)\\ (j_2,j_3)\neq(i_3,i_2),(j_2,j_1)\neq(i_2,i_1)}} [\sum_a \eta_{i_1}^{a_1+2}\eta_{i_2}^{a_2+2}\eta_{i_3}^{a_3+2}\eta_{i_4}^{a_4+2}\eta_{j_1}^{a_5+1}\eta_{j_2}^{a_6+1}\eta_{j_3}^{a_7+1}],$$

where the last term

$$\leq C\sum_a\|\theta\|_{a_1+2}^{a_1+2}\cdot\|\theta\|_{a_2+2}^{a_2+2}\cdot\|\theta\|_{a_3+2}^{a_3+2}\cdot\|\theta\|_{a_4+2}^{a_4+2}\|\theta\|_{a_5+1}^{a_5+1}\|\theta\|_{a_6+1}^{a_6+1}\|\theta\|_{a_7+1}^{a_7+1}/\|\theta\|_1^6.$$

Since $a_1,a_2,\cdots,a_7$ have to take values from $\{0,1\}$ and their sum is 3, the above term

$$\leq C\|\theta\|^2\|\theta\|_3^9/\|\theta\|_1^3=o(\|\theta\|_3^3),$$

where we have used $\|\theta\|_3^3\ll\|\theta\|_2^2\ll\|\theta\|_1$. Combining these gives

(167)
$$\mathbb{E}[X_c^2]\leq C\|\theta\|^2\|\theta\|_3^9/\|\theta\|_1^3.$$

Finally, inserting (159), (164), and (167) into (148) gives that both under the null and the alternative,

$$\mathrm{Var}(T_{11}) \leq C[\|\theta\|^8/\|\theta\|_1^4 + \|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2 + \|\theta\|^2\|\theta\|_3^9/\|\theta\|_1^3] \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2,$$

where we have used $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$ and $\|\theta\|_3^3/\|\theta\|_1 = o(1)$. This gives (133) and completes the proof for Item (a).

Consider Item (b). The goal is to show (134). Recall that

$$T_{1b} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}^2 \big[(\eta_{i_1} - \tilde{\eta}_{i_1})(\eta_{i_2} - \tilde{\eta}_{i_2})(\eta_{i_4} - \tilde{\eta}_{i_4})\big] \cdot W_{i_4 i_1},$$

and that

$$\widetilde{\eta} - \eta = v^{-1/2}W1_n.$$

Plugging this into T_{1b} gives

$$\begin{aligned} T_{1b} &= -v^{-3/2} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}^2 \Big(\sum_{j_1\neq i_1} W_{i_1j_1}\Big)\Big(\sum_{j_2\neq i_2} W_{i_2j_2}\Big)\Big(\sum_{j_4\neq i_4} W_{i_4j_4}\Big)W_{i_1i_4} \\ &= -\frac{1}{v^{3/2}} \sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ j_1\neq i_1, j_2\neq i_2, j_4\neq i_4}} \eta_{i_2}\eta_{i_3}^2 W_{i_1j_1}W_{i_2j_2}W_{i_4j_4}W_{i_1i_4}. \end{aligned}$$

By basic combinatorics and careful observations, we have

$$W_{i_1j_1}W_{i_2j_2}W_{i_4j_4}W_{i_1i_4} = \begin{cases} W_{i_1i_4}^3 W_{i_2j_2}, & \text{if } j_1 = i_4, j_4 = i_1, \\ W_{i_1i_2}^2 W_{i_1i_4}^2, & \text{if } j_1 = i_2, j_2 = i_1, j_4 = i_1, \\ W_{i_1i_4}^2 W_{i_2i_4}^2, & \text{if } j_1 = i_4, j_2 = i_4, j_4 = i_2, \\ W_{i_1i_2}^2 W_{i_4j_4} W_{i_1i_4}, & \text{if } j_1 = i_2, j_2 = i_1, \\ W_{i_1i_4}^2 W_{i_1j_1} W_{i_2j_2}, & \text{if } j_4 = i_1, \\ W_{i_1i_4}^2 W_{i_2j_2} W_{i_4j_4}, & \text{if } j_1 = i_4, \{i_2, j_2\} \neq \{i_4, j_4\}, \\ W_{i_2i_4}^2 W_{i_1j_1} W_{i_1i_4}, & \text{if } j_2 = i_4, j_4 = i_2, \\ W_{i_1j_1}W_{i_2j_2}W_{i_4j_4}W_{i_1i_4}, & \text{otherwise.} \end{cases} \tag{168}$$

This allows us to further split T_{1b} into 8 different terms:

$$T_{1b} = Y_{a1} + Y_{a2} + Y_{a3} + Y_{b1} + Y_{b2} + Y_{b3} + Y_{b4} + Y_c, \tag{169}$$

where

$$Y_{a1} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2(j_2\neq i_2)} \eta_{i_2}\eta_{i_3}^2 W_{i_1i_4}^3 W_{i_2j_2},$$

$$Y_{a2} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}^2 W_{i_1i_2}^2 W_{i_1i_4}^2,$$

$$Y_{a3} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}^2 W_{i_1i_4}^2 W_{i_2i_4}^2,$$

$$Y_{b1} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_4(j_4\neq i_4)} \eta_{i_2}\eta_{i_3}^2 W_{i_1i_2}^2 W_{i_4j_4} W_{i_1i_4},$$

$$Y_{b2} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_1 (j_1 \neq i_1), j_2 (j_2 \neq i_2) \\ \{i_1,j_1\} \neq \{i_2,j_2\}}} \eta_{i_2}\eta_{i_3}^2 W_{i_1 i_4}^2 W_{i_1 j_1} W_{i_2 j_2},$$

$$Y_{b3} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_2 (j_2 \neq i_2), j_4 (j_4 \neq i_4) \\ \{i_2,j_2\} \neq \{i_4,j_4\}}} \eta_{i_2}\eta_{i_3}^2 W_{i_1 i_4}^2 W_{i_2 j_2} W_{i_4 j_4},$$

$$Y_{b4} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{j_1 (j_1 \neq i_1)} \eta_{i_2}\eta_{i_3}^2 W_{i_2 i_4}^2 W_{i_1 j_1} W_{i_1 i_4},$$

$$Y_c = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_1,j_2,j_4 \\ j_1 \notin \{i_2,i_4\}, j_2 \notin \{i_1,i_4\}, j_4 \notin \{i_1,i_2\}}} \eta_{i_2}\eta_{i_3}^2 W_{i_1 j_1} W_{i_2 j_2} W_{i_4 j_4} W_{i_1 i_4}.$$

We now show the two claims in (134) separately.

Consider the first claim of (134). It is seen that out of the 8 terms on the right hand side of (196), the mean of all terms are 0, except that of the Y_{a2} and Y_{a3}. Note that for any $1 \leq i, j \leq n$, $i \neq j$, $\mathbb{E}[W_{ij}^2] = \Omega_{ij}(1 - \Omega_{ij})$, where Ω_{ij} are upper bounded by $o(1)$ uniformly for all such i, j. It follows

$$\mathbb{E}[Y_{a2}] = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_2}\eta_{i_3}^2 \mathbb{E}[W_{i_1 i_2}^2]\mathbb{E}[W_{i_1 i_4}^2]$$
$$= -(1 + o(1)) \cdot v^{-3/2} \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_2}\eta_{i_3}^2 \Omega_{i_1 i_2}\Omega_{i_1 i_4}.$$

Since for any $1 \leq i, j \leq n$, $i \neq j$, $0 < \eta_i \leq C\theta_i$, $\Omega_{ij} \leq C\theta_i\theta_j$ and $v \asymp \|\theta\|_1^2$,

$$|\mathbb{E}[Y_{a2}]| \leq C(\|\theta\|_1)^{-3} \sum_{i_1,i_2,i_3,i_4 (dist)} \theta_{i_1}^2\theta_{i_2}^2\theta_{i_3}^2\theta_{i_4} \leq C\|\theta\|^6/\|\theta\|_1^2.$$

Therefore,

$$|\mathbb{E}[Y_{a2}]| \leq C\|\theta\|^6/\|\theta\|_1^2. \tag{170}$$

By symmetry, we similarly find

$$|\mathbb{E}[Y_{a3}]| \leq C\|\theta\|^6/\|\theta\|_1^2. \tag{171}$$

Combining (170) and (171) gives

$$\mathbb{E}[|T_{1b}|] \leq C\|\theta\|^6/\|\theta\|_1^2.$$

This completes the proof of the first claim of (134).

We now show the second claim of (134) . By Cauchy-Schwartz inequality,

$$\mathrm{Var}(T_{1b}) \leq C(\mathrm{Var}(Y_{a1}) + \mathrm{Var}(Y_{a2}) + \mathrm{Var}(Y_{a3}) + \sum_{s=1}^{4} \mathrm{Var}(Y_{bs}) + \mathrm{Var}(Y_c))$$
$$\leq C(\mathrm{Var}(Y_{a1}) + \mathrm{Var}(Y_{a2}) + \mathrm{Var}(Y_{a3}) + \sum_{s=1}^{4} \mathbb{E}[Y_{bs}^2] + \mathbb{E}[Y_c^2]). \tag{172}$$

We now show $\mathrm{Var}(Y_{a1})$, $\mathrm{Var}(Y_{a2})$, $\mathrm{Var}(Y_{a3})$, $\sum_{s=1}^{4} \mathbb{E}[Y_{bs}^2]$, and $\mathbb{E}[Y_c^2]$, separately.

Consider $\mathrm{Var}(Y_{a1})$. Recalling $\mathbb{E}[Y_{a1}] = 0$, we write $\mathrm{Var}(Y_{a1})$ as

$$v^{-3} \sum_{\substack{i_1,i_2,i_3,i_4(dist) \\ i_1',i_2',i_3',i_4'(dist)}} \sum_{j_2(j_2\neq i_2)} \sum_{j_2'(j_2'\neq i_2')} \eta_{i_2}\eta_{i_3}^2 \eta_{i_2'}\eta_{i_3'}^2 \mathbb{E}\big[W_{i_1i_4}^3 W_{i_2j_2} W_{i_1'i_4'}^3 W_{i_2'j_2'}\big]. \tag{173}$$

In the sum, a term is nonzero only when one of the following cases happens.

- (A). $\{W_{i_1i_4}, W_{i_2j_2}, W_{i_1'i_4'}, W_{i_2'j_2'}\}$ has 2 distinct random variables.
- (B). $\{W_{i_1i_4}, W_{i_2j_2}, W_{i_1'i_4'}, W_{i_2'j_2'}\}$ has 3 distinct random variables. While it may seem we have 4 possibilities in this case, but the only one that has a nonzero mean is when $W_{i_2j_2} = W_{i_2'j_2'}$.

For Case (A), the two sets $\{i_1, i_2, i_4, j_2\}$ and $\{i_1', i_2', i_4', j_2'\}$ are identical, and so for two integers $0 \le b_1, b_2 \le 1$ and $b_1 + b_2 = 1$,

$$W_{i_1i_4}^3 W_{i_2j_2} W_{i_1'i_4'}^3 W_{i_2'j_2'} = W_{i_1i_4}^{4+2b_1} W_{i_2j_2}^{2+2b_2},$$

and so

$$\mathbb{E}[W_{i_1i_4}^3 W_{i_2j_2} W_{i_1'i_4'}^3 W_{i_2'j_2'}] = \mathbb{E}[W_{i_1i_4}^{4+2b_1} W_{i_2j_2}^{2+2b_2}] = \mathbb{E}[W_{i_1i_4}^{4+2b_1}]\mathbb{E}[W_{i_2j_2}^{2+2b_2}],$$

Note that for any integer $2 \le b \le 6$,

$$0 < \mathbb{E}[W_{ij}^b] \le C\Omega_{ij},$$

where note that $\Omega_{ij} \le C\theta_i\theta_j$ for all $1 \le i, j \le n$, $i \le j$. Recall that $v \sim \|\theta\|_1^2$, and that $0 < \eta_i \le C\theta_i$ for all $1 \le i \le n$. Combining these that, the contribution of Case (A) to $\mathrm{Var}(Y_{a1})$ is no more than

$$C(\|\theta\|_1)^{-6} \sum_{i_1,\cdots,i_4(dist)} \sum_{i_3',j_2} \sum_{a} \theta_{i_1}^{a_1+1}\theta_{i_2}^{a_2+2}\theta_{i_3}^2\theta_{i_4}^{a_3+1}\theta_{i_3'}^2\theta_{j_2}^{a_4+1}, \tag{174}$$

where $a = (a_1, a_2, a_3, a_4)$ and each a_i is either 0 and 1, satisfying $a_1 + a_2 + a_3 + a_4 = 1$. Note that the right hand side of (174) is no greater than

$$C(\|\theta\|_1)^{-6}\max\{\|\theta\|_1^3\|\theta\|^4\|\theta\|_3^3, \|\theta\|_1^2\|\theta\|^8\} \le C\|\theta\|^4\|\theta\|_3^3/\|\theta\|_1^3,$$

where we have used $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$.

Next, consider Case (B). In this case, $\{i_2, j_2\} = \{i_2', j_2'\}$ and

$$W_{i_1i_4}^3 W_{i_2j_2} W_{i_1'i_4'}^3 W_{i_2'j_2'} = W_{i_1i_4}^3 W_{i_2j_2}^2 W_{i_1'i_4'}^3,$$

and by similar argument,

$$0 < \mathbb{E}[W_{i_1i_4}^3 W_{i_2j_2}^2 W_{i_1'i_4'}^3] \le C\Omega_{i_1i_4}\Omega_{i_2j_2}\Omega_{i_1'i_4'}. \tag{175}$$

Recall that $\Omega_{ij} \le C\theta_i\theta_j$ for all $1 \le i, j \le n$, $i \le j$, that $v \sim \|\theta\|_1^2$, and that $0 < \eta_i \le C\theta_i$ for all $1 \le i \le n$. Combining this with (173), the contribution of this case to $\mathrm{Var}(Y_{a1})$

$$\le C(\|\theta\|_1)^{-6} \sum_{\substack{i_1,i_2,i_3,i_4(dist) \\ i_1',i_3',i_4'(dist)}} \sum_{j_2} C\theta_{i_1}\theta_{i_2}^{2+b_1}\theta_{i_3}^2\theta_{i_4}\theta_{i_1'}\theta_{i_3'}^2\theta_{i_4'}\theta_{j_2}^{1+b_2}, \tag{176}$$

where similarly b_1, b_2 are either 0 or 1 and $b_1 + b_2 = 1$. By similar argument, the right hand side

$$\le C\|\theta\|_1^{-6} \cdot [\|\theta\|_1^5\|\theta\|^4\|\theta\|_3^3 + \|\theta\|_1^4\|\theta\|^8] \le C\|\theta\|^4\|\theta\|_3^3/\|\theta\|_1,$$

where we've used Cauchy-Schwartz inequality that $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$.

Now, inserting (174) and (176) into (173) gives

(177) $$\mathrm{Var}(Y_{a1}) \leq C[\|\theta\|^4\|\theta\|_3^3/\|\theta\|_1^3 + \|\theta\|^4\|\theta\|_3^3/\|\theta\|_1] \leq C\|\theta\|^4\|\theta\|_3^3/\|\theta\|_1,$$

where we have used $\|\theta\|_1 \to \infty$ and $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$. This shows

(178) $$\mathrm{Var}(Y_{a1}) \leq C\|\theta\|^4\|\theta\|_3^3/\|\theta\|_1.$$

Next, we consider $\mathrm{Var}(Y_{a2})$ and $\mathrm{Var}(Y_{a3})$. The proofs are similar to that of $\mathrm{Var}(X_a)$ of Item (a), so we skip the detail, but claim that

(179) $$\mathrm{Var}(Y_{a2}) \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4,$$

and

(180) $$\mathrm{Var}(Y_{a3}) \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4.$$

Combining (178), (179), and (180) gives

(181)
$$\mathrm{Var}(Y_{a1})+\mathrm{Var}(Y_{a2})+\mathrm{Var}(Y_{a3}) \leq C[\|\theta\|^4\|\theta\|_3^3/\|\theta\|_1+\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4] \leq C\|\theta\|^4\|\theta\|_3^3/\|\theta\|_1,$$

where we have used the universal inequality that $\|\theta\|_3^3 \leq \|\theta\|_1^3$.

Next, consider $\sum_{s=1}^4 \mathbb{E}[Y_{bs}^2]$. For each $1 \leq s \leq 4$, the study of $\mathbb{E}[Y_{bs}^2]$ is similar to that of $\mathbb{E}[X_{b1}^2]$ in Item (a), so we skip the details. We have that both under the null and the alternative,

(182) $$\mathbb{E}[Y_{b1}^2] \leq C\|\theta\|^{12}/\|\theta\|_1^4,$$

(183) $$\mathbb{E}[Y_{b2}^2] \leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1,$$

(184) $$\mathbb{E}[Y_{b3}^2] \leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1,$$

(185) $$\mathbb{E}[Y_{b4}^2] \leq C\|\theta\|^{12}/\|\theta\|_1^4.$$

Therefore,

(186) $$\sum_{s=1}^4 \mathbb{E}[Y_{bs}^2] \leq C[\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|^{12}/\|\theta\|_1^4] \leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1.$$

Third, we consider $\mathbb{E}[Y_c^2]$. The proof is very similar to that of $\mathbb{E}[X_c^2]$ and we have that both under the null and the alternative,

(187) $$\mathbb{E}[Y_c^2] \leq C\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^3.$$

Finally, combining (181), (186), and (187) with (172) gives

(188)
$$\mathrm{Var}(T_{1b}) \leq C[\|\theta\|^4\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|^6\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^3] \leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1,$$

where we have used $\|\theta\| \to \infty$ and $\|\theta\|^2 \ll \|\theta\|_1$. This completes the proof of (134).

Consider Item (c). The goal is to show (135). Recall that

$$T_{1c} = \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_1}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_2} - \tilde{\eta}_{i_2})^2(\eta_{i_3} - \tilde{\eta}_{i_3})\big] \cdot W_{i_4 i_1},$$

and that

$$\widetilde{\eta} - \eta = v^{-1/2}W\mathbf{1}_n.$$

Plugging this into T_{1c} gives

$$T_{1c} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}\eta_{i_4}\Big(\sum_{j_2\neq i_2} W_{i_2j_2}\Big)\Big(\sum_{\ell_2\neq i_2} W_{i_2\ell_2}\Big)\Big(\sum_{j_3\neq i_3} W_{i_3j_3}\Big)W_{i_1i_4}$$
$$= -\frac{1}{v^{3/2}} \sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ j_2\neq i_2,\ell_2\neq i_2,j_3\neq i_3}} \eta_{i_1}\eta_{i_3}\eta_{i_4} W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}W_{i_1i_4}.$$

By basic combinatorics and careful observations, we have

$$W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}W_{i_1i_4} = \begin{cases} W_{i_2i_3}^3 W_{i_1i_4}, & \text{if } j_2=\ell_2=i_3, j_3=i_2,\\ W_{i_2j_2}^2 W_{i_3j_3}W_{i_1i_4}, & \text{if } j_2=\ell_2, (j_3,j_2)\neq(i_2,i_3),\\ W_{i_2i_3}^2 W_{i_2\ell_2}W_{i_1i_4}, & \text{if } j_2=i_3, j_3=i_2, \ell_2\neq i_3,\\ W_{i_2i_3}^2 W_{i_2j_2}W_{i_1i_4}, & \text{if } \ell_2=i_3, j_3=i_2, j_2\neq i_3,\\ W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}W_{i_1i_4}, & \text{otherwise.}\end{cases} \tag{189}$$

This allows us to further split T_{1c} into 5 different terms:

$$T_{1c} = Z_a + Z_{b1} + Z_{b2} + Z_{b3} + Z_c, \tag{190}$$

where

$$Z_a = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}\eta_{i_4} W_{i_2i_3}^3 W_{i_1i_4},$$

$$Z_{b1} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2,(j_3,j_2)\neq(i_2,i_3)} \eta_{i_1}\eta_{i_3}\eta_{i_4} W_{i_2j_2}^2 W_{i_3j_3}W_{i_1i_4},$$

$$Z_{b2} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_2=i_3,j_3=i_2\\ \ell_2\neq i_3}} \eta_{i_1}\eta_{i_3}\eta_{i_4} W_{i_2i_3}^2 W_{i_2\ell_2}W_{i_1i_4},$$

$$Z_{b3} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{\ell_2=i_3,j_3=i_2\\ j_2\neq i_3}} \eta_{i_1}\eta_{i_3}\eta_{i_4} W_{i_2i_3}^2 W_{i_2j_2}W_{i_1i_4},$$

$$Z_c = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_2,\ell_2,j_3\\ j_2\neq \ell_2, j_2,\ell_2\neq i_3, j_3\neq i_2}} \eta_{i_1}\eta_{i_3}\eta_{i_4} W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}W_{i_1i_4}.$$

We now show the two claims in (135) separately. The proof of the first claim is trivial, so we only show the second claim of (135).

Consider the second claim of (135). By Cauchy-Schwartz inequality,

$$\mathrm{Var}(T_{1c}) \leq C(\mathrm{Var}(Z_a) + \mathrm{Var}(Z_{b1}) + \mathrm{Var}(Z_{b2}) + \mathrm{Var}(Z_{b3}) + \mathrm{Var}(Z_c))$$
$$\leq C(\mathbb{E}[Z_a^2] + \sum_{s=1}^{3} \mathbb{E}[Z_{bs}^2] + \mathbb{E}[Z_c^2]). \tag{191}$$

Note that

- The proof of $\mathrm{Var}(Z_a)$ is similar to that of $\mathrm{Var}(Y_a)$ in Item (b).
- The proof of $\sum_{s=1}^{3} \mathbb{E}[Z_{bs}^2]$ is similar to that of $\sum_{s=1}^{4} \mathbb{E}[X_{bs}^2]$ in Item (a).
- The proof of $\mathbb{E}[Z_c^2]$ is similar to that of $\mathbb{E}[X_c^2]$ in Item (a).

For these reasons, we omit the proof details and only state the claims. We have that under both the null and the alternative,

$$\mathrm{Var}(Z_a) \le C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4, \tag{192}$$

$$\sum_{s=1}^{3}\mathbb{E}[Z_{bs}^2] \le C\|\theta\|_3^9/\|\theta\|_1, \tag{193}$$

and

$$\mathbb{E}[Z_c^2] \le C\|\theta\|^2\|\theta\|_3^9/\|\theta\|_1^3. \tag{194}$$

Inserting (192), (193), and (194) into (191) gives

$$\mathrm{Var}(T_{1c}) \le C[\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4 + \|\theta\|_3^9/\|\theta\|_1 + \|\theta\|^2\|\theta\|_3^9/\|\theta\|_1^3] \le C\|\theta\|_3^9/\|\theta\|_1,$$

where we have used $\|\theta\|_3^3 \ll \|\theta\|^2 \ll \|\theta\|_1$, $\|\theta\|^4 \le \|\theta\|_1\|\theta\|_3^3$ and $\|\theta\|_1 \to \infty$. This proves (135).

Consider Item (d). The goal is to show (136) and (137). Recall that

$$T_{1d} = -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\eta_{i_1}\eta_{i_3}^2\left[(\eta_{i_2}-\tilde{\eta}_{i_2})^2(\eta_{i_4}-\tilde{\eta}_{i_4})\right]\cdot W_{i_4i_1}.$$

and that

$$\widetilde{\eta} - \eta = v^{-1/2}W1_n.$$

Plugging this into T_{1d} gives

$$\begin{aligned}T_{1d} &= -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\eta_{i_1}\eta_{i_3}^2\Big(\sum_{j_2\neq i_2}W_{i_2j_2}\Big)\Big(\sum_{\ell_2\neq i_2}W_{i_2\ell_2}\Big)\Big(\sum_{j_4\neq i_4}W_{i_4j_4}\Big)W_{i_1i_4}\\ &= -\frac{1}{v^{3/2}}\sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ j_2\neq i_2,\ell_2\neq i_2,j_4\neq i_4}}\eta_{i_1}\eta_{i_3}^2W_{i_2j_2}W_{i_2\ell_2}W_{i_4j_4}W_{i_1i_4}.\end{aligned}$$

By basic combinatorics and careful observations, we have

$$W_{i_2j_2}W_{i_2\ell_2}W_{i_4j_4}W_{i_1i_4} = \begin{cases} W_{i_2i_4}^3W_{i_1i_4}, & \text{if } j_2=\ell_2=i_4, j_4=i_2,\\ W_{i_2j_2}^2W_{i_1i_4}^2, & \text{if } j_2=\ell_2, j_4=i_1,\\ W_{i_2j_2}^2W_{i_4j_4}W_{i_1i_4}, & \text{if } j_2=\ell_2, j_4\neq i_1, (j_2,j_4)\neq(i_4,i_2),\\ W_{i_2j_2}W_{i_2i_4}^2W_{i_1i_4}, & \text{if } \ell_2=i_4, j_4=i_2, j_2\neq i_4,\\ W_{i_2\ell_2}W_{i_2i_4}^2W_{i_1i_4}, & \text{if } j_2=i_4, j_4=i_2, \ell_2\neq i_4,\\ W_{i_2j_2}W_{i_2\ell_2}W_{i_1i_4}^2, & \text{if } j_4=i_1, j_2\neq\ell_2,\\ W_{i_2j_2}W_{i_2\ell_2}W_{i_4j_4}W_{i_1i_4}, & \text{otherwise.}\end{cases} \tag{195}$$

This allows us to further split T_{14} into 7 different terms:

$$T_{1d} = U_{a1} + U_{a2} + U_{b1} + U_{b2} + U_{b3} + U_{b4} + U_c, \tag{196}$$

where

$$U_{a1} = -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\eta_{i_1}\eta_{i_3}^2W_{i_2i_4}^3W_{i_1i_4},$$

$$U_{a2} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2} \eta_{i_1}\eta_{i_3}^2 W_{i_2j_2}^2 W_{i_1i_4}^2,$$

$$U_{b1} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_2(j_2\neq i_2),j_4(j_4\neq i_4)\\ j_4\neq i_1,(j_2,j_4)\neq(i_4,i_2)}} \eta_{i_1}\eta_{i_3}^2 W_{i_2j_2}^2 W_{i_4j_4} W_{i_1i_4},$$

$$U_{b2} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2(j_2\neq i_4)} \eta_{i_1}\eta_{i_3}^2 W_{i_2j_2} W_{i_2i_4}^2 W_{i_1i_4},$$

$$U_{b3} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\ell_2(\ell_2\neq i_4)} \eta_{i_1}\eta_{i_3}^2 W_{i_2\ell_2} W_{i_2i_4}^2 W_{i_1i_4},$$

$$U_{b4} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2\neq\ell_2} \eta_{i_1}\eta_{i_3}^2 W_{i_2j_2} W_{i_2\ell_2} W_{i_1i_4}^2,$$

$$U_{c} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2,\ell_2,j_4,W\text{dist}} \eta_{i_1}\eta_{i_3}^2 W_{i_2j_2} W_{i_2\ell_2} W_{i_4j_4} W_{i_1i_4}.$$

We now show (136) and (137) separately.

Consider (136). It is seen that out of the 7 terms on the right hand side of (190), all terms are mean 0, except for the second term U_{a2}. Note that for any $1 \leq i,j \leq n$, $i \neq j$, $\mathbb{E}[W_{ij}^2] = \Omega_{ij}(1-\Omega_{ij})$, where Ω_{ij} are upper bounded by $o(1)$ uniformly for all such i,j. It follows

$$\mathbb{E}[U_{a2}] = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2} \eta_{i_1}\eta_{i_3}^2 \mathbb{E}[W_{i_2j_2}^2]\mathbb{E}[W_{i_1i_4}^2]$$
$$= -(1+o(1))\cdot v^{-3/2} \sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ j_2}} \eta_{i_1}\eta_{i_3}^2 \Omega_{i_2j_2}\Omega_{i_1i_4}.$$

Under null, for any $1 \leq i,j \leq n$, $i \neq j$, $\eta_i = (1+o(1))\theta_i$, $\Omega_{ij} = (1+o(1))\theta_i\theta_j$ and $v \asymp \|\theta\|_1^2$,

$$\mathbb{E}[U_{a2}] = (\|\theta\|_1)^{-3} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2} \theta_{i_1}^2\theta_{i_2}\theta_{i_3}^2\theta_{i_4}\theta_{j_2} = -(1+o(1))\|\theta\|^4,$$

and under alternative, a similar arguments yields

$$|\mathbb{E}[U_{a1}]| \leq C\|\theta\|^4. \tag{197}$$

This proves (136).

We now consider (137). By Cauchy-Schwartz inequality,

$$\mathrm{Var}(T_{1d}) \leq C(\mathrm{Var}(U_{a1}) + \mathrm{Var}(U_{a2}) + \sum_{s=1}^4 \mathrm{Var}(U_{bs}) + \mathrm{Var}(U_c))$$
$$\leq C(\mathrm{Var}(U_{a1}) + \mathrm{Var}(U_{a2}) + \sum_{s=1}^4 \mathbb{E}[U_{bs}^2] + \mathbb{E}[U_c^2]). \tag{198}$$

Note that

- The proof of U_{a1} is similar to that of Y_{a1} in Item (b).
- The proof of U_{a2} is similar to that of X_{a1} in Item (a).
- The proof of U_{bs}, $1 \leq s \leq 4$, is similar to that of X_{b1} in Item (a).

- The proof of U_c is similar to that of X_c in Item (a).

For these reasons, we omit the proof details, and claim that

$$\mathrm{Var}(U_{a1}) \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4, \tag{199}$$

$$\mathrm{Var}(U_{a2}) \leq C\|\theta\|^4\|\theta\|_3^3/\|\theta\|_1, \tag{200}$$

$$\sum_{s=1}^4 \mathbb{E}[U_{bs}^2] \leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1, \tag{201}$$

and

$$\mathrm{Var}(U_c) \leq C\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^3, \tag{202}$$

Inserting (199), (200), (201), and (202) into (198) gives

$$\mathrm{Var}(T_{1d}) \leq C[\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4 + \|\theta\|^4\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|^6\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^3] \tag{203}$$

$$\leq C\|\theta\|^6\|\theta\|_3^3/\|\theta\|_1, \tag{204}$$

where we have used $\|\theta\| \to \infty$ and $\|\theta\|_3^3 \leq \|\theta\|_1^3$. This proves (137).

We now consider Item (e) and Item (f). Since the proof is similar, we only prove Item (e). The goal is to show (138). Recall that

$$T_{2a} = \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_2}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_1}-\tilde{\eta}_{i_1})(\eta_{i_2}-\tilde{\eta}_{i_2})(\eta_{i_3}-\tilde{\eta}_{i_3})\big]\cdot \widetilde{\Omega}_{i_4 i_1}, \tag{205}$$

and

$$\widetilde{\eta} - \eta = v^{-1/2} W 1_n. \tag{206}$$

Plugging (206) into (205) gives

$$\begin{aligned} T_{2a} &= -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_2}\eta_{i_3}\eta_{i_4}\Big(\sum_{j_1\neq i_1} W_{i_1 j_1}\Big)\Big(\sum_{j_2\neq i_2} W_{i_2 j_2}\Big)\Big(\sum_{j_3\neq i_3} W_{i_3 j_3}\Big)\widetilde{\Omega}_{i_4 i_1} \\ &= -\frac{1}{v^{3/2}} \sum_{\substack{i_1,i_2,i_3,i_4 (dist)\\ j_1\neq i_1, j_2\neq i_2, j_3\neq i_3}} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1 j_1}W_{i_2 j_2}W_{i_3 j_3}\widetilde{\Omega}_{i_1 i_4}. \end{aligned}$$

By basic combinatorics and careful observations, we have

$$W_{i_1 j_1}W_{i_2 j_2}W_{i_3 j_3} = \begin{cases} W_{i_1 i_2}^2 W_{i_3 j_3}, & \text{if } j_1 = i_2, j_2 = i_1, \\ W_{i_1 i_3}^2 W_{i_2 j_2}, & \text{if } j_1 = i_3, j_3 = i_1, \\ W_{i_2 i_3}^2 W_{i_1 j_1}, & \text{if } j_2 = i_3, j_3 = i_2, \\ W_{i_1 j_1}W_{i_2 j_2}W_{i_3 j_3}, & \text{otherwise.} \end{cases} \tag{207}$$

This allows us to further split T_{2a} into 4 different terms:

$$T_{2a} = X_{a1} + X_{a2} + X_{a3} + X_b, \tag{208}$$

where

$$X_{a1} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_3 \neq i_3} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1i_2}^2 W_{i_3j_3} \widetilde{\Omega}_{i_1i_4},$$

$$X_{a2} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_2 \neq i_2} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1i_3}^2 W_{i_2j_2} \widetilde{\Omega}_{i_1i_4},$$

$$X_{a3} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_1 \neq i_1} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_2i_3}^2 W_{i_1j_1} \widetilde{\Omega}_{i_1i_4},$$

$$X_{b} = -\frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_1,j_2,j_3 \\ j_k \neq i_\ell, k,\ell=1,2,3}} \eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1j_1} W_{i_2j_2} W_{i_3j_3} \widetilde{\Omega}_{i_1i_4}.$$

We now consider the two claims of (138) separately. Since the mean of $X_{a1}, X_{a2}, X_{a3}, X_b$ are all 0, the first claim of (138) follows trivially, so all remains to show is the second claim of (138).

We now consider the second claim of (138). By Cauchy-Schwartz inequality,

$$\begin{aligned} \mathrm{Var}(T_{2a}) &\leq C\mathrm{Var}(X_{a1}) + \mathrm{Var}(X_{a2}) + \mathrm{Var}(X_{a3}) + \mathrm{Var}(X_b)) \\ &\leq C(\mathbb{E}[X_{a1}^2] + \mathbb{E}[X_{a2}^2] + \mathbb{E}[X_{a3}^2] + \mathbb{E}[X_b^2]). \end{aligned} \tag{209}$$

We now consider $\mathbb{E}[X_{a1}^2] + \mathbb{E}[X_{a2}^2] + \mathbb{E}[X_{a3}^2]$, and $\mathbb{E}[X_b^2]$, separately.

Consider $\mathbb{E}[X_{a1}^2] + \mathbb{E}[X_{a2}^2] + \mathbb{E}[X_{a3}^2]$. We claim that both under the null and the alternative,

$$\mathbb{E}[X_{a1}^2] \leq C\alpha^2 \|\theta\|^{12} \|\theta\|_3^3 / \|\theta\|_1^5, \tag{210}$$

$$\mathbb{E}[X_{a2}^2] \leq C\alpha^2 \|\theta\|^{12} \|\theta\|_3^3 / \|\theta\|_1^5, \tag{211}$$

$$\mathbb{E}[X_{a3}^2] \leq C\alpha^2 \|\theta\|^{12} \|\theta\|_3^3 / \|\theta\|_1^5. \tag{212}$$

Combining these gives that both under the null and the alternative,

$$\mathbb{E}[X_{a1}^2] + \mathbb{E}[X_{a2}^2] + \mathbb{E}[X_{a3}^2] \leq C\alpha^2 \|\theta\|^{12} \|\theta\|_3^3 / \|\theta\|_1^5. \tag{213}$$

It remains to show (210)-(212). Since the proofs are similar, we only prove (210). Write

$$\mathbb{E}[X_{a1}^2] = v^{-3} \sum_{\substack{i_1,i_2,i_3,i_4(dist) \\ i_1',i_2',i_3',i_4'(dist)}} \sum_{\substack{j_3,j_3' \\ j_3 \neq i_3, j_3' \neq i_3'}} \eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'} \mathbb{E}[W_{i_1i_2}^2 W_{i_3j_3} W_{i_1'i_2'}^2 W_{i_3'j_3'}] \widetilde{\Omega}_{i_1i_4} \widetilde{\Omega}_{i_1'i_4'}.$$

Consider the term

$$W_{i_1i_2}^2 W_{i_3j_3} W_{i_1'i_2'}^2 W_{i_3'j_3'}.$$

In order for the mean is nonzero, we have three cases

- Case A. $W_{i_1i_2} = W_{i_3'j_3'}$ and $W_{i_3j_3} = W_{i_1'i_2'}$.
- Case B. $W_{i_3j_3} = W_{i_3'j_3'}$ and $W_{i_1i_2} = W_{i_1'i_2'}$.
- Case C. $W_{i_3j_3} = W_{i_3'j_3'}$ and $W_{i_1i_2} \neq W_{i_1'i_2'}$.

Consider Case A. In this case, $\{i_1', i_2', i_3'\}$ are three distinct indices in $\{i_1, i_2, i_3, j_3\}$. In this case,

$$W_{i_1i_2}^2 W_{i_3j_3} W_{i_1'i_2'}^2 W_{i_3'j_3'} = W_{i_1i_2}^3 W_{i_3j_3}^3,$$

where by similar arguments as before

$$0 < \mathbb{E}[W_{i_1i_2}^3 W_{i_3j_3}^3] \leq C\Omega_{i_1i_2}\Omega_{i_3j_3} \leq C\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{j_3}.$$

At the same time, recall that that $0 < \eta_i \leq C\theta_i$ for any $1 \leq i \leq n$, and that $|\widetilde{\Omega}_{ij}| \leq C\alpha\theta_i\theta_j$ for any $1 \leq i, j \leq n$, $i \neq j$, where $\alpha = |\lambda_2/\lambda_1|$ with λ_k being the k-th largest (in magnitude) eigenvalue of Ω, $1 \leq k \leq K$. By basic algebra,

$$|\eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'}\widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i_1'i_4'}| \leq C\alpha^2\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4}^2\theta_{i_1'}\theta_{i_2'}\theta_{i_3'}\theta_{i_4'}^2.$$

Note that in the current case, $\{i_1, i_2\} = \{i_3', j_3'\}$ and $\{i_3, j_3\} = \{i_1', i_2'\}$, so for some integers $0 \leq b_1, b_2 \leq 1$ and $b_1 + b_2 = 1$,

$$\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4}^2\theta_{i_1'}\theta_{i_2'}\theta_{i_3'}\theta_{i_4'}^2 = \theta_{i_1}^{1+b_1}\theta_{i_2}^{1+b_2}\theta_{i_3}^2\theta_{j_3}\theta_{i_4}^2\theta_{i_4'}^2.$$

Recall that $v \asymp \|\theta\|_1^2$. Combining these, the contribution of Case (A) to $\mathbb{E}[X_{a1}^2]$ is no greater than

$$C\alpha^2(\|\theta\|_1)^{-6} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{i_4'} \sum_{j_3(j_3\neq i_3)} \sum_{b_1,b_2(b_1+b_2=1)} \theta_{i_1}^{2+b_1}\theta_{i_2}^{2+b_2}\theta_{i_3}^3\theta_{j_3}^2\theta_{i_4}^2\theta_{i_4'}^2,$$

where the right hand side $\leq C\alpha^2 \cdot \|\theta\|^8\|\theta\|_3^6/\|\theta\|_1^6$. This shows that the contribution of Case (A) to $\mathbb{E}[X_{a1}^2]$ is no greater than

$$C\alpha^2 \cdot \|\theta\|^8\|\theta\|_3^6/\|\theta\|_1^6. \tag{214}$$

Consider Case B. By similar arguments,

$$W_{i_1i_2}^2 W_{i_3j_3} W_{i_1'i_2'}^2 W_{i_3'j_3'} = W_{i_1i_2}^6 W_{i_3j_3}^2,$$

where

$$\mathbb{E}[W_{i_1i_2}^6 W_{i_3j_3}^2] \leq C\Omega_{i_1i_2}\Omega_{i_3j_3} \leq C\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{j_3},$$

Also, by similar arguments,

$$|\eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'}\widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i_1'i_4'}| \leq C\alpha^2\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4}^2\theta_{i_1'}\theta_{i_2'}\theta_{i_3'}\theta_{i_4'}^2,$$

where as $W_{i_1i_2} = W_{i_1'i_2'}$ and $W_{i_3j_3} = W_{i_3'j_3'}$, the right hand side

$$\leq C\alpha^2\theta_{i_1}^2\theta_{i_2}^2\theta_{i_3}^{1+c_1}\theta_{j_3}^{c_2}\theta_{i_4}^2\theta_{i_4'}^2,$$

where $0 < c_1, c_2 \leq$ are integers satisfying $c_1 + c_2 = 1$. Recall $v \sim \|\theta\|_1^2$. Combining these, the contribution of Case (B) to $\mathbb{E}[X_{a1}^2]$

$$\leq C\alpha^2(\|\theta\|_1)^{-6} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{i_4'} \sum_{j_3(j_3\neq i_3)} \sum_{b_1,b_2(b_1+b_2=1)} \theta_{i_1}^3\theta_{i_2}^3\theta_{i_3}^{2+c_1}\theta_{j_3}^{1+c_2}\theta_{i_4}^2\theta_{i_4'}^2,$$

where by $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$, the above term

$$\leq C\alpha^2[\|\theta\|^4\|\theta\|_3^9/\|\theta\|_1^5, \|\theta\|^8\|\theta\|_3^6/\|\theta\|_1^6] \leq C\alpha^2\|\theta\|^4\|\theta\|_3^9/\|\theta\|_1^5.$$

This shows that the contribution of Case (B) to $\mathbb{E}[X_{a1}^2]$ is no greater than

$$C\|\theta\|^4\|\theta\|_3^9/\|\theta\|_1^5. \tag{215}$$

Consider Case (C). In this case,

$$W_{i_1i_2}^2 W_{i_3j_3} W_{i_1'i_2'}^2 W_{i_3'j_3'} = W_{i_1i_2}^2 W_{i_3j_3}^2 W_{i_1'i_2'}^2,$$

where by similar arguments,

$$\mathbb{E}[W_{i_1i_2}^2 W_{i_3j_3}^2 W_{i_1'i_2'}^2] \leq C\Omega_{i_1i_2}\Omega_{i_3j_3}\Omega_{i_1'i_2'} \leq C\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{j_3}\theta_{i_1'}\theta_{i_2'}.$$

Also, by similar arguments,

$$|\eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'}\widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i_1'i_4'}| \leq C\alpha^2\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4}^2\theta_{i_1'}\theta_{i_2'}\theta_{i_3'}\theta_{i_4'}^2,$$

where as $W_{i_3 j_3} = W_{i_3' j_3'}$, the right hand side

$$\leq C\alpha^2 \theta_{i_1}\theta_{i_2}\theta_{i_3}^{1+c_1}\theta_{j_3}^{c_2}\theta_{i_4}^2\theta_{i_4'}^2,$$

with the same c_1, c_2 as in the proof of Case B. Combining these and using $v \asymp \|\theta\|_1^2$, we have that under both the null and the alternative, the contribution of Case (C) to $\mathbb{E}[X_{a1}^2]$

$$\leq C\alpha^2(\|\theta\|_1)^{-6} \sum_{\substack{i_1,i_2,i_3,i_4 (dist)\\ i_1',i_2',i_4' (dist)}} \sum_{j_3 (j_3 \neq i_3)} \theta_{i_1}^2\theta_{i_2}^2\theta_{i_3}^{2+c_1}\theta_{j_3}^{1+c_2}\theta_{i_4}^2\theta_{i_1'}^2\theta_{i_2'}^2\theta_{i_4'}^2,$$

where the right hand size

$$\leq C\alpha^2 \cdot [\|\theta\|^{12}\|\theta\|_3^3/\|\theta\|_1^5 + \|\theta\|^{12}\|\theta\|_3^6/\|\theta\|_1^6] \leq C\alpha^2\|\theta\|^{12}\|\theta\|_3^3/\|\theta\|_1^5. \tag{216}$$

Here we have again used $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$.

Combining (214), (215), and (216) gives

$$\mathbb{E}[X_{a1}^2] \leq C\alpha^2(\|\theta\|^8\|\theta\|_3^6/\|\theta\|_1^6 + \|\theta\|^4\|\theta\|_3^9/\|\theta\|_1^5 + \|\theta\|^8\|\theta\|_3^9/\|\theta\|^5] \leq C\alpha^2\|\theta\|^8\|\theta\|_3^9/\|\theta\|_1^5,$$

where we have used $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$ and $\|\theta\| \to \infty$. This proves (210).

We now consider $\mathbb{E}[X_b^2]$. Write

$$\mathbb{E}[X_b^2] = v^{-3} \sum_{\substack{i_1,i_2,i_3,i_4 (dist)\\ i_1',i_2',i_3',i_4' (dist)}} \sum_{\substack{j_3,j_3'\\ j_3\neq i_3, j_3'\neq i_3'}} \eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'}$$
$$\mathbb{E}[W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}W_{i_1'j_1'}W_{i_2'j_2'}W_{i_3'j_3'}]\widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i_1'i_4'}.$$

Consider

$$W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}, \qquad \text{and} \qquad W_{i_1'j_1'}W_{i_2'j_2'}W_{i_3'j_3'}.$$

Each term has a mean 0, and two terms are uncorrelated with each other if and only if the two sets of random variables $\{W_{i_1j_1}, W_{i_2j_2}, W_{i_3j_3}\}$ and $\{W_{i_1'j_1'}, W_{i_2'j_2'}, W_{i_3'j_3'}\}$ are identical (however, it is possible that $W_{i_1j_1}$ does not equal to $W_{i_1'j_1'}$ but equals to $W_{i_2'j_2'}$, say). When this happens, first, $\{i_1, i_2, i_3, j_1, j_2, j_3\} = \{i_1', i_2', i_3', j_1', j_2', j_3'\}$. Recall that $|\widetilde{\Omega}_{ij}| \leq C\alpha\theta_i\theta_j$ for all $1 \leq i, j \leq n$, $i \neq j$, and that $0 < \eta_i \leq C\theta_i$ for all $1 \leq i \leq n$. For integers $a_i \in \{0, 1\}$, $1 \leq i \leq 4$, that satisfy $\sum_{i=1}^6 a_i = 3$, we have

$$|\eta_{i_2}\eta_{i_3}\eta_{i_4}\eta_{i_2'}\eta_{i_3'}\eta_{i_4'}\widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i_1'i_4'}| \leq C\eta_{i_1}^{a_1}\eta_{j_1}^{a_2}\eta_{i_2}^{1+a_3}\eta_{j_2}^{a_4}\eta_{i_3}^{1+a_5}\eta_{j_3}^{a_6}\eta_{i_4}\eta_{i_4'}|\widetilde{\Omega}_{i_1i_4}||\widetilde{\Omega}_{i_1'i_4'}|$$
$$\leq C\alpha^2\theta_{i_1}^{1+a_1}\eta_{j_1}^{a_2}\eta_{i_2}^{1+a_3}\eta_{j_2}^{a_4}\eta_{i_3}^{1+a_5}\eta_{j_3}^{a_6}\eta_{i_4}^2\eta_{i_4'}^2.$$

Second,

$$\mathbb{E}[W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}W_{i_1'j_1'}W_{i_2'j_2'}W_{i_3'j_3'}] = \mathbb{E}[W_{i_1j_1}^2W_{i_2j_2}^2W_{i_3j_3}^2],$$

where by similar arguments, the right hand side

$$\leq C\Omega_{i_1j_1}\Omega_{i_2j_2}\Omega_{i_3j_3} \leq C\theta_{i_1}\theta_{j_1}\theta_{i_2}\theta_{j_2}\theta_{i_3}\theta_{j_3}.$$

Recall that $v \sim \|\theta\|_1^2$. Combining these gives

$$\mathbb{E}[X_b^2] \leq C\alpha^2\|\theta\|_1^{-6} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{i_4'} \sum_{\substack{j_1,j_2,j_3\\ j_1\neq i_1, j_2\neq i_2, j_3\neq i_3}} \sum_a \theta_{i_1}^{2+a_1}\eta_{j_1}^{1+a_2}\eta_{i_2}^{2+a_3}\eta_{j_2}^{1+a_4}\eta_{i_3}^{2+a_5}\eta_{j_3}^{1+a_6}\eta_{i_4}^2\eta_{i_4'}^2,$$

where $a = (a_1, a_2, \ldots, a_6)$ as above. By the way a_i are defined, the right hand side

$$\leq C\alpha^2\|\theta\|^4(\sum_a \|\theta\|_{a_1+2}^{a_1+2} \cdot \|\theta\|_{a_2+1}^{a_2+1} \cdot \|\theta\|_{a_3+2}^{a_3+2} \cdot \|\theta\|_{a_4+1}^{a_4+1}\|\theta\|_{a_5+2}^{a_5+2}\|\theta\|_{a_6+1}^{a_6+1})/\|\theta\|_1^6,$$

which by $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$, the term in the bracket does not exceed

$$C\max\{\|\theta\|^{12}, \|\theta\|_1\|\theta\|^8\|\theta\|_3^3, \|\theta\|_1^2\|\theta\|^4\|\theta\|_3^6, \|\theta\|_1^3\|\theta\|_3^9\} \leq C\|\theta\|_1^3\|\theta\|_3^9.$$

Combining these gives

$$\mathbb{E}[X_b^2] \leq C\alpha^2\|\theta\|^4\|\theta\|_3^9/\|\theta\|_1^3. \tag{217}$$

Finally, inserting (213)-(217) into (209) gives

$$\mathrm{Var}(T_{2a}) \leq C\alpha^2[\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^5 + \|\theta\|^4\|\theta\|_3^9/\|\theta\|_1^3] \leq C\alpha^2\|\theta\|^4\|\theta\|_3^9/\|\theta\|_1^3,$$

and (138) follows.

Consider Item (g) and Item (h). The proof are similar, so we only show Item (g). The goal is to show (140). Recall that

$$T_{2c} = \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}\eta_{i_4}\left[(\eta_{i_2}-\tilde{\eta}_{i_2})^2(\eta_{i_3}-\tilde{\eta}_{i_3})\right]\cdot\widetilde{\Omega}_{i_4i_1}, \tag{218}$$

and

$$\widetilde{\eta} - \eta = v^{-1/2}W1_n.$$

Plugging this into T_{2c} gives (note symmetry in $\widetilde{\Omega}$)

$$\begin{aligned} T_{2c} &= -\frac{1}{v^{2/3}}\sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}\eta_{i_4}\Big(\sum_{j_2\neq i_2} W_{i_2j_2}\Big)\Big(\sum_{\ell_2\neq i_2} W_{i_2\ell_2}\Big)\Big(\sum_{j_3\neq i_3} W_{i_3j_3}\Big)\widetilde{\Omega}_{i_4i_1} \\ &= -\frac{1}{v^{3/2}}\sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ j_1\neq i_1, j_2\neq i_2, j_3\neq i_3}} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}\widetilde{\Omega}_{i_1i_4}. \end{aligned}$$

By basic combinatorics and careful observations, we have

$$W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3} = \begin{cases} W_{i_2i_3}^3, & \text{if } j_1=\ell_2=i_3, j_3=i_2, \\ W_{i_2j_2}^2W_{i_3j_3}, & \text{if } j_1=\ell_2, (j_2,j_3)\neq(i_3,i_2), \\ W_{i_2i_3}^2W_{i_2\ell_2}, & \text{if } j_2=i_3, j_3=i_2, \ell_2\neq i_3, \\ W_{i_2i_3}^2W_{i_2j_2}, & \text{if } \ell_2=i_3, j_3=i_2, j_2\neq i_3, \\ W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}, & \text{otherwise.} \end{cases} \tag{219}$$

This allows us to further split T_{2c} into 4 different terms:

$$T_{2c} = Y_a + Y_{b1} + Y_{b2} + Y_{b3} + Y_c, \tag{220}$$

$$Y_a = -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{j_3\neq i_3} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2i_3}^3\widetilde{\Omega}_{i_1i_4},$$

$$Y_{b1} = -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{j_3\neq i_3} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2j_2}^2W_{i_3j_3}\widetilde{\Omega}_{i_1i_4},$$

$$Y_{b2} = -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{j_2\neq i_2} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2i_3}^2W_{i_2\ell_2}\widetilde{\Omega}_{i_1i_4},$$

$$Y_{b3} = -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{j_1\neq i_1} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2i_3}^2W_{i_2j_2}\widetilde{\Omega}_{i_1i_4},$$

$$Y_c = -\frac{1}{v^{3/2}}\sum_{i_1,i_2,i_3,i_4(dist)}\sum_{\substack{j_2,\ell_2,j_3\\ j_2\neq i_2,\ell_2\neq i_2,j_3\neq i_3\\ j_2\neq i_3,\ell_2\neq i_3,j_3\neq i_2}} \eta_{i_1}\eta_{i_3}\eta_{i_4}W_{i_2j_2}W_{i_2\ell_2}W_{i_3j_3}\widetilde{\Omega}_{i_1i_4}.$$

We now show the two claims in (140) separately. Consider the first claim. It is seen that out of the 5 terms on the right hand side of (220), the mean of all terms are 0, except for the first one. Note that for any $1 \le i, j \le n, i \ne j$, $\mathbb{E}[W_{ij}^3] \le C\Omega_{ij}$. Together with $\Omega_{ij} \le C\theta_i\theta_j$, $\widetilde{\Omega}_{ij} \le C\alpha\theta_i\theta_j$, $0 < \eta_i < C\theta_i$ and $v \sim \|\theta\|_1^2$, it follows

$$\begin{aligned}\mathbb{E}[|Y_a|] &\le \frac{1}{v^{3/2}} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_1}\eta_{i_3}\eta_{i_4}\Omega_{i_2i_3}\widetilde{\Omega}_{i_1i_4} \\ &\le C\alpha \cdot \frac{1}{\|\theta\|_1^3} \sum_{i_1,i_2,i_3,i_4(dist)} \theta_{i_1}^2\theta_{i_2}\theta_{i_3}^2\eta_{i_4}^2,\end{aligned}$$

where the last term is no greater than $C\alpha \cdot \|\theta\|^6/\|\theta\|_1^3$, and the first claim of (140) follows.

Consider the second claim of (140). By Cauchy-Schwartz inequality,

$$\begin{aligned}\mathrm{Var}(T_{2c}) &\le C(\mathrm{Var}(Y_a) + \mathrm{Var}(Y_{b1}) + \mathrm{Var}(Y_{b2}) + \mathrm{Var}(Y_{b3}) + \mathrm{Var}(Y_c)) \\ &\le C(\mathrm{Var}(Y_a) + \mathbb{E}[Y_{b1}^2] + \mathbb{E}[Y_{b2}^2] + \mathbb{E}[Y_{b3}^2] + \mathbb{E}[Y_c^2]). \quad (221)\end{aligned}$$

We now study $\mathrm{Var}(Y_a)$. Write

$$\mathrm{Var}(Y_a) = v^{-3} \sum_{\substack{i_1,i_2,i_3,i_4(dist) \\ i_1',i_2',i_3',i_4'(dist)}} \eta_{i_1}\eta_{i_3}\eta_{i_4}\eta_{i_1'}\eta_{i_3'}\eta_{i_4'}\mathbb{E}[(W_{i_2i_3}^3 - \mathbb{E}[W_{i_2i_3}^3])(W_{i_2'i_3'}^3 - \mathbb{E}[W_{i_2'i_3'}^3])] \cdot \widetilde{\Omega}_{i_1i_4}\widetilde{\Omega}_{i_1'i_4'}.$$

Fix a term $(W_{i_2i_3}^3 - \mathbb{E}[W_{i_2i_3}^3])(W_{i_2'i_3'}^3 - \mathbb{E}[W_{i_2'i_3'}^3])$. When the mean is nonzero, we must have $\{i_2, i_3\} = \{i_2', i_3'\}$, and when this happens,

$$\mathbb{E}[(W_{i_2i_3}^3 - \mathbb{E}[W_{i_2i_3}^3])(W_{i_2'i_3'}^3 - \mathbb{E}[W_{i_2'i_3'}^3])] = \mathrm{Var}(W_{i_2i_3}^3).$$

For a random variable X, we have $\mathrm{Var}(X) \le \mathbb{E}[X^2]$, and it follows that

$$\mathrm{Var}(W_{i_2i_3}^3) \le \mathbb{E}[W_{i_2i_3}^6] \le \mathbb{E}[W_{i_2i_3}^2],$$

where we have used the property that $0 \le W_{i_2i_3}^2 \le 1$. Notice that $\mathbb{E}[W_{i_2i_3}^2] \le C\theta_{i_2}\theta_{i_3}$, and recall that $v \asymp \|\theta\|_1^2$, $\widetilde{\Omega}_{ij} \le C\alpha\theta_i\theta_j$ and $0 < \eta_i \le C\theta_i$ for all $1 \le i \le n$. Combining these gives

$$\mathrm{Var}(Y_a) \le C\alpha^2(\|\theta\|_1^{-6}) \cdot \sum_{\substack{i_1,i_2,i_3,i_4(dist) \\ i_1',i_4'(dist)}} \theta_{i_1}^2\theta_{i_2}\theta_{i_3}^3\theta_{i_4}^2\theta_{i_1'}^2\theta_{i_4'}^2 \le C\alpha^2\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^5. \quad (222)$$

Additionally, note that

- The proof of Y_{b1}, Y_{b2}, and Y_{b3} is similar to that of X_{a1} in Item (e).
- The proof of Y_c is similar to that of X_b in Item (e).

For these reasons, we skip the proof details, but only to state that, both under the null and the alternative,

$$\mathbb{E}[Y_{b1}^2] \le C\alpha^2\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1, \quad (223)$$

$$\mathbb{E}[Y_{b2}^2] \le C\alpha^2\|\theta\|^{12}\|\theta\|_3^3/\|\theta\|_1^5, \quad (224)$$

$$\mathbb{E}[Y_{b3}^2] \le C\alpha^2\|\theta\|^{12}\|\theta\|_3^3/\|\theta\|_1^5, \quad (225)$$

and therefore,

$$\mathbb{E}[Y_{b1}^2] + \mathbb{E}[Y_{b2}^2] + \mathbb{E}[Y_{b3}^2] \le C\alpha^2\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1. \quad (226)$$

At the same time, both under the null and the alternative,

$$\mathbb{E}[Y_c^2] \le C\alpha^2 \cdot \|\theta\|^{10}\|\theta\|_3^3/\|\theta\|_1^3. \tag{227}$$

Inserting (226) and (227) into (221) gives

$$\mathbb{E}[T_{2c}^2] \le C\alpha^2[\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^5 + \|\theta\|^8\|\theta\|_3^3/\|\theta\|_1 + \|\theta\|^{10}\|\theta\|_3^3/\|\theta\|_1^3] \le C\alpha^2\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1.$$

This proves (140).

Consider Item (i). The goal is to show (142). Recall that

$$F_a = \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_1}\eta_{i_2}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_1}-\tilde{\eta}_{i_1})(\eta_{i_2}-\tilde{\eta}_{i_2})(\eta_{i_3}-\tilde{\eta}_{i_3})(\eta_{i_4}-\tilde{\eta}_{i_4})\big], \tag{228}$$

and that for any $1 \le i \le n$,

$$\tilde{\eta}_i - \eta_i = v^{-1/2}\sum_{j\neq i}^{n} W_{ij}.$$

Inserting it into (228) gives

$$F_a = \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_1}\eta_{i_2}\eta_{i_3}\eta_{i_4}\big[(\eta_{i_1}-\tilde{\eta}_{i_1})(\eta_{i_2}-\tilde{\eta}_{i_2})(\eta_{i_3}-\tilde{\eta}_{i_3})(\eta_{i_4}-\tilde{\eta}_{i_4})\big],$$

By basic combinatorics and basic algebra, we have

$$W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}W_{i_4j_4} = \begin{cases} W_{i_1i_2}^2 W_{i_3i_4}^2, & \text{if } (i_1,j_1)=(j_2,i_2), (i_3,j_3)=(j_4,i_4),\\ W_{i_1i_3}^2 W_{i_2i_4}^2, & \text{if } (i_1,j_1)=(j_3,i_3), (i_2,j_2)=(j_4,i_4),\\ W_{i_1i_4}^2 W_{i_2i_3}^2, & \text{if } (i_1,i_4)=(j_4,i_1), (i_2,j_2)=(j_3,i_3),\\ W_{i_1i_2}^2 W_{i_3j_3}W_{i_4j_4}, & \text{if } (i_1,j_1)=(j_2,i_2), (j_4,j_3)\neq(i_3,i_4),\\ W_{i_1i_3}^2 W_{i_2j_2}W_{i_4j_4}, & \text{if } (i_1,j_1)=(j_3,i_3), (j_4,j_2)\neq(i_2,i_4),\\ W_{i_1i_4}^2 W_{i_2j_2}W_{i_3j_4}, & \text{if } (i_1,j_1)=(j_4,i_4), (j_3,j_2)\neq(i_2,i_3),\\ W_{i_2i_3}^2 W_{i_1j_1}W_{i_4j_4}, & \text{if } (i_2,j_2)=(j_3,i_3), (j_4,j_1)\neq(i_1,i_4),\\ W_{i_2i_4}^2 W_{i_1j_1}W_{i_3j_3}, & \text{if } (i_2,j_2)=(j_4,i_4), (j_3,j_1)\neq(i_1,i_3),\\ W_{i_3i_4}^2 W_{i_1j_1}W_{i_2j_2}, & \text{if } (i_3,j_3)=(j_4,i_4), (j_2,j_1)\neq(i_1,i_2).\\ W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}W_{i_4j_4}, & \text{otherwise.} \end{cases}$$

By symmetry, it allows us to further split F_1 into 3 different terms:

$$F_1 = 3X_a + 6X_b + X_c, \tag{229}$$

where

$$X_a = v^{-2}\sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_1}\eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1i_2}^2 W_{i_3i_4}^2,$$

$$X_b = v^{-2}\sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_3,j_4\\ (j_3,j_4)\neq(i_4,i_3)}} \eta_{i_1}\eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1i_2}^2 W_{i_3j_3}W_{i_4j_4},$$

and

$$X_c = v^{-2}\sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_1,j_2,j_3,j_4\\ j_k\neq i_\ell, k,\ell=1,2,3,4}} \eta_{i_1}\eta_{i_2}\eta_{i_3}\eta_{i_4} W_{i_1j_1}W_{i_2j_2}W_{i_3j_3}W_{i_4j_4}.$$

We now show the two claims in (142) separately. Consider the first claim of (142). Note that $\mathbb{E}[X_b] = \mathbb{E}[X_c] = 0$. Recall that both under the null and the alternative, for any $i \neq j$, $\mathbb{E}[W_{ij}^2] = \Omega_{ij}(1 - \Omega_{ij}) \leq C\theta_i\theta_j$, and that $0 < \eta_i \leq C\theta_i$, and that $v \asymp \|\theta\|_1^2$, Therefore,

$$0 < \mathbb{E}[X_a] \leq v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4}\theta_{i_1}\theta_{i_2}\theta_{i_3}\theta_{i_4} \leq C\|\theta\|^8/\|\theta\|_1^4.$$

Inserting into (229) gives

$$\mathbb{E}[|F_1|] \leq C\|\theta\|^8/\|\theta\|_1^4,$$

and the first claim (142) follows.

Consider the second claim (142) next. By (229) and Cauchy-Schwarz inequality,

$$\text{Var}(F_1) \leq C(\text{Var}(X_a) + \text{Var}(X_b) + \text{Var}(X_c)) \leq C(\text{Var}(X_a) + \mathbb{E}[X_b^2] + \mathbb{E}[X_c^2]). \tag{230}$$

We now consider $\text{Var}(X_a)$, $\mathbb{E}[X_b^2]$, and $\mathbb{E}[X_c^2]$, separately. Note that

- The proof of $\text{Var}(X_a)$ is similar to that of $\text{Var}(X_a)$ in Item (a).
- The proof of $\mathbb{E}[X_b^2]$ is similar to that of $\sum_{s=1}^4 \mathbb{E}[X_{bs}^2]$ in Item (a).
- The proof of $\mathbb{E}[X_c^2]$ is similar to that of $\mathbb{E}[X_c^2]$ in Item (a).

For these reasons, we omit the proof details and only state the claims. We have that under both the null and the alternative,

$$\text{Var}(X_a) \leq C\|\theta\|^8\|\theta\|_3^6/\|\theta\|_1^8. \tag{231}$$

$$\text{Var}(X_b^2) + \text{Var}(Y_{a3}) \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4, \tag{232}$$

$$\mathbb{E}[X_c^2] \leq C\|\theta\|_3^{12}/\|\theta\|_1^4, \tag{233}$$

Finally, inserting (231), (232), and (233) into (229) gives that, both under the null and the alternative,

$$\text{Var}(F_1) \leq C[\|\theta\|^8\|\theta\|_3^6/\|\theta\|_1^8 + \|\theta\|^8\|\theta\|_3^6/\|\theta\|_1^6 + \|\theta\|_3^{12}/\|\theta\|_1^4] \leq C\|\theta\|^8\|\theta\|_3^6/\|\theta\|_1^6,$$

where we have used $\|\theta\| \to \infty$ and $\|\theta\|_3^3 \ll \|\theta\|^2 \ll \|\theta\|_1$. This gives (142) and completes the proof for Item (i).

Consider Item (j). The goal is to show (143). Recall that

$$F_b = \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_2}\eta_{i_3}^2\eta_{i_4}\big[(\eta_{i_1} - \tilde{\eta}_{i_1})^2(\eta_{i_2} - \tilde{\eta}_{i_2})(\eta_{i_4} - \tilde{\eta}_{i_4})\big],$$

and that

$$\widetilde{\eta} - \eta = v^{-1/2}W1_n.$$

Plugging this into F_b, we have

$$F_b = v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_1,\ell_1,j_2,j_4 \\ j_1 \neq i_1, \ell_1 \neq i_1, j_2 \neq i_2, j_4 \neq i_4}} \eta_{i_2}\eta_{i_3}^2\eta_{i_4}W_{i_1j_1}W_{i_1\ell_1}W_{i_2j_2}W_{i_4j_4}.$$

By basic combinatorics and basic algebra, we have

$$W_{i_1j_1}W_{i_1\ell_1}W_{i_2j_2}W_{i_4j_4} = \begin{cases} W_{i_1i_2}^3 W_{i_4j_4}, & \text{if } j_1,\ell_1 = i_2, j_2 = i_1, \\ W_{i_1i_4}^3 W_{i_2j_2}, & \text{if } j_1,\ell_1 = i_4, j_4 = i_1, \\ W_{i_1i_2}^2 W_{i_1i_4}^2, & \text{if } (j_1,j_2) = (i_2,i_1), (\ell_1,j_4) = (i_4,i_1), \\ W_{i_1i_2}^2 W_{i_1i_4}^2, & \text{if } (\ell_1,j_2) = (i_2,i_1), (j_1,j_4) = (i_4,i_1), \\ W_{i_1i_4}^2 W_{i_1i_2}^2, & \text{if } (j_1,j_4) = (i_4,i_1), (\ell_1,j_2) = (i_2,i_1), \\ W_{i_1i_4}^2 W_{i_1i_2}^2, & \text{if } (\ell_1,j_4) = (i_4,i_1), (j_1,j_2) = (i_2,i_1), \\ W_{i_1j_1}^2 W_{i_2i_4}^2, & \text{if } j_1 = \ell_1, (j_2,j_4) = (i_4,i_2), \\ W_{i_1i_2}^2 W_{i_1j_1} W_{i_4j_4}, & \text{if } \ell_1 = i_2, j_2 = i_1, j_1 \neq i_2, i_4, \\ W_{i_1i_2}^2 W_{i_1\ell_1} W_{i_4j_4}, & \text{if } j_1 = i_2, j_2 = i_1, \ell_1 \neq i_2, i_4, \\ W_{i_1i_4}^2 W_{i_1j_1} W_{i_2j_2}, & \text{if } \ell_1 = i_4, j_4 = i_1, \ell_1 \neq i_2, i_4, \\ W_{i_1i_4}^2 W_{i_1\ell_1} W_{i_2j_2}, & \text{if } j_1 = i_4, j_4 = i_1, j_1 \neq i_2, i_4, \\ W_{i_2i_4}^2 W_{i_1j_1} W_{i_1\ell_1}, & \text{if } j_1 \neq \ell_1, (j_2,j_4) = (i_4,i_2). \\ W_{i_1j_1}^2 W_{i_2j_2} W_{i_4j_4}, & \text{if } j_1 = \ell_1, (j_1,j_2) \neq (i_2,i_1), (j_1,j_4) \neq (i_4,i_1), \\ W_{i_1j_1}W_{i_1\ell_1}W_{i_2j_2}W_{i_4j_4}, & \text{otherwise.} \end{cases}$$

By these and symmetry, we can further split F_b into 7 different terms,

We decompose

$$F_b = 2Y_{a1} + 4Y_{a2} + Y_{a3} + 4Y_{b1} + Y_{b2} + Y_{b3} + Y_c, \tag{234}$$

where

$$Y_{a1} = v^{-2} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_4, j_4 \neq i_4} \eta_{i_2}\eta_{i_3}^2\eta_{i_4} W_{i_1i_2}^3 W_{i_4j_4},$$

$$Y_{a2} = v^{-2} \sum_{i_1,i_2,i_3,i_4(dist)} \eta_{i_2}\eta_{i_3}^2\eta_{i_4} W_{i_1i_2}^2 W_{i_1i_4}^2,$$

$$Y_{a3} = v^{-2} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_1, j_1 \neq i_1} \eta_{i_2}\eta_{i_3}^2\eta_{i_4} W_{i_1j_1}^2 W_{i_2i_4}^2,$$

$$Y_{b1} = v^{-2} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_1,j_4 \\ j_1 \neq i_1, j_4 \neq i_4}} \eta_{i_2}\eta_{i_3}^2\eta_{i_4} W_{i_1i_2}^2 W_{i_1j_1} W_{i_4j_4},$$

$$Y_{b2} = v^{-2} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_1,\ell_1 \\ j_1,\ell_1 \neq i_1}} \eta_{i_2}\eta_{i_3}^2\eta_{i_4} W_{i_2i_4}^2 W_{i_1j_1} W_{i_1\ell_1},$$

$$Y_{b3} = v^{-2} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_1,j_2,j_4 \\ j_1 \neq i_1, j_2 \neq i_2, j_4 \neq i_4}} \eta_{i_2}\eta_{i_3}^2\eta_{i_4} W_{i_1j_1}^2 W_{i_2j_2} W_{i_4j_4},$$

$$Y_c = v^{-2} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{\substack{j_1,\ell_1,j_2,j_4 \\ j_1,\ell_1 \notin \{i_1,i_2,i_4\} \\ j_2 \notin \{i_1,i_4\}, j_4 \notin \{i_1,i_2\}}} \eta_{i_2}\eta_{i_3}^2\eta_{i_4} W_{i_1j_1}W_{i_1\ell_1}W_{i_2j_2}W_{i_4j_4},$$

We now consider the two claims in (143) separately. Consider the first claim. It is seen that only the second and the third terms above have non-zero mean. Recall that both under the null and the alternative, for any $i \neq j$, $\mathbb{E}[W_{ij}^2] = \Omega_{ij}(1-\Omega_{ij}) \leq C\theta_i\theta_j$, $0 < \eta_i \leq C\theta_i$, and that $v \asymp \|\theta\|_1^2$. It follows

$$0 < \mathbb{E}[Y_{a2}] \leq v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \theta_{i_2}\theta_{i_3}^2\theta_{i_4} \cdot \theta_{i_1}^2\theta_{i_2}\theta_{i_4} \leq C\|\theta\|^8/\|\theta\|_1^4. \tag{235}$$

and

$$0 < \mathbb{E}[Y_{a3}] \leq v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{j_1} \theta_{i_2}\theta_{i_3}^2\theta_{i_4} \cdot \theta_{i_1}\theta_{i_2}\theta_{j_1}\theta_{i_4} \leq C\|\theta\|^6/\|\theta\|_1^2. \tag{236}$$

Combining (235), (236) with (234) gives

$$\mathbb{E}[|F_2|] \leq C[\|\theta\|^8/\|\theta\|_1^4 + \|\theta\|^6/\|\theta\|_1^2] \leq C\|\theta\|^6/\|\theta\|_1^2,$$

where we've used the universal inequality that $\|\theta\|^2 \leq \|\theta\|_1$. It follows the first claim of (143).

We now show the second claim of (143). By Cauchy-Schwarz inequality,

$$\begin{aligned} \mathrm{Var}(F_b) &\leq C\big(\mathrm{Var}(Y_{a1}) + \mathrm{Var}(Y_{a2}) + \mathrm{Var}(Y_{a3}) + \mathrm{Var}(Y_{b1}) + \mathrm{Var}(Y_{b2}) + \mathrm{Var}(Y_{b3}) + \mathrm{Var}(Y_c)\big) \\ &\leq C\big(\mathrm{Var}(Y_{a1}) + \mathrm{Var}(Y_{a2}) + \mathrm{Var}(Y_{a3}) + \mathbb{E}[Y_{b1}^2] + \mathbb{E}[Y_{b2}^2] + \mathbb{E}[Y_{b3}^2] + \mathbb{E}[Y_c^2]\big). \end{aligned} \tag{237}$$

We now consider $\mathrm{Var}(Y_{a1})$, $\mathrm{Var}(Y_{a2}) + \mathrm{Var}(Y_{a3})$, $\mathbb{E}[Y_{b1}^2] + \mathbb{E}[Y_{b2}^2] + \mathbb{E}[Y_{b3}^2]$, and $\mathbb{E}[Y_c^2]$, separately. Note that

- The proof of $\mathrm{Var}(Y_{a1})$ is similar to that of $\mathrm{Var}(Y_a)$ in Item (b).
- The proof of $\mathrm{Var}(Y_{a2})$ and $\mathrm{Var}(Y_{a3})$ are similar to that of $\mathrm{Var}(X_a)$ in Item (a).
- The proof of $\sum_{s=1}^3 \mathbb{E}[Y_{b_s}^2]$ is similar to that of $\sum_{s=1}^4 \mathbb{E}[X_{bs}^2]$ in Item (a).
- The proof of $\mathbb{E}[Y_c^2]$ is similar to that of $\mathbb{E}[X_c^2]$ in Item (a).

For these reasons, we omit the proof details and only state the claims. We have that under both the null and the alternative,

$$\mathrm{Var}(Y_{a1}) \leq C\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^5. \tag{238}$$

$$\mathrm{Var}(Y_{a2}) + \mathrm{Var}(Y_{a3}) \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4, \tag{239}$$

$$\sum_{s=1}^{3} \mathbb{E}[Y_{b_s}^2] \leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2, \tag{240}$$

$$\mathbb{E}[Y_c^2] \leq C\|\theta\|^6\|\theta\|_3^6/\|\theta\|_1^4. \tag{241}$$

Finally, inserting (238), (239), (240), and (241) into (237) gives

$$\begin{aligned} \mathrm{Var}(F_2) &\leq C[\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^5 + \|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4 + \|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^2 + \|\theta\|^6\|\theta\|_3^6/\|\theta\|_1^4] \\ &\leq C\|\theta\|^4\|\theta\|_3^6/\|\theta\|_1^4, \end{aligned} \tag{242}$$

where we have used $\|\theta\|_3^3 \ll \|\theta\|^2 \ll \|\theta\|_1$, $\|\theta\| \to \infty$ and $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$. This completes the proof of (143).

Consider Item (k). The goal is to show (144) and (145). Recall that

$$F_c = \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_2}^2\eta_{i_4}^2 \left[(\eta_{i_1} - \tilde{\eta}_{i_1})^2(\eta_{i_3} - \tilde{\eta}_{i_3})^2\right],$$

and that $\widetilde{\eta} - \eta = v^{-1/2} W 1_n$. Plugging this into F_3 gives

$$F_c = v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_1,\ell_1,j_2,j_4 \\ j_1 \neq i_1, \ell_1 \neq i_1, j_3 \neq i_3, \ell_3 \neq i_3}} \eta_{i_2}^2 \eta_{i_4}^2 W_{i_1 j_1} W_{i_1 \ell_1} W_{i_3 j_3} W_{i_3 \ell_3}.$$

By basic combinatorics and basic algebra, we have

$$W_{i_1 j_1} W_{i_1 \ell_1} W_{i_3 j_3} W_{i_3 \ell_3} = \begin{cases} W_{i_1 i_3}^4, & \text{if } j_1 = \ell_1 = i_1, j_3 = \ell_3 = i_1, \\ W_{i_1 i_3}^3 W_{i_1 j_1}, & \text{if } j_3 = \ell_3 = i_1, \ell_1 = i_3, \\ W_{i_1 i_3}^3 W_{i_1 \ell_1}, & \text{if } j_3 = \ell_3 = i_1, j_1 = i_3, \\ W_{i_1 i_3}^3 W_{i_3 j_3}, & \text{if } j_1 = \ell_1 = i_3, \ell_3 = i_1, \\ W_{i_1 i_3}^3 W_{i_3 \ell_3}, & \text{if } j_1 = \ell_1 = i_3, j_3 = i_1, \\ W_{i_1 j_1}^2 W_{i_3 j_3}^2, & \text{if } j_1 = \ell_1, j_3 = \ell_3, \\ W_{i_1 j_1}^2 W_{i_3 j_3} W_{i_3 \ell_3}, & \text{if } j_1 = \ell_1 \neq i_3, j_3 \neq \ell_3, \\ W_{i_3 j_3}^2 W_{i_1 j_1} W_{i_1 \ell_1}, & \text{if } j_3 = \ell_3 \neq i_1, j_1 \neq \ell_1, \\ W_{i_1 i_3}^2 W_{i_1 \ell_1} W_{i_3 \ell_3}, & \text{if } j_1 = i_3, j_3 = i_1, \\ W_{i_1 i_3}^2 W_{i_1 j_1} W_{i_3 j_3}, & \text{if } \ell_1 = i_3, \ell_3 = i_1, \\ W_{i_1 i_3}^2 W_{i_1 j_1} W_{i_3 \ell_3}, & \text{if } \ell_1 = i_3, j_3 = i_1, \\ W_{i_1 i_3}^2 W_{i_1 \ell_1} W_{i_3 j_3}, & \text{if } j_1 = i_3, \ell_3 = i_1, \\ W_{i_1 j_1} W_{i_1 \ell_1} W_{i_3 j_3} W_{i_3 \ell_3}, & \text{otherwise.} \end{cases}$$

By these and symmetry, we can further split F_3 into 6 different terms:

$$F_c = Z_a + 4Z_{b1} + Z_{b2} + 2Z_{c1} + 4Z_{c2} + Z_d, \tag{243}$$

where

$$Z_a = v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \eta_{i_2}^2 \eta_{i_4}^2 W_{i_1 i_3}^4,$$

$$Z_{b1} = v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{j_4, j_4 \neq i_4} \eta_{i_2}^2 \eta_{i_4}^2 W_{i_1 i_3}^3 W_{i_3 j_3},$$

$$Z_{b2} = v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{j_1, j_1 \neq i_1, j_3, j_3 \neq i_3} \eta_{i_2}^2 \eta_{i_4}^2 W_{i_1 j_1}^2 W_{i_3 j_3}^2,$$

$$Z_{c1} = v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_1, j_3, \ell_3 \\ j_1 \notin \{i_1, i_3\}, j_3, \ell_3}} \eta_{i_2}^2 \eta_{i_4}^2 W_{i_1 j_1}^2 W_{i_3 j_3} W_{i_3 \ell_3},$$

$$Z_{c2} = v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{\ell_1, \ell_3 \\ \ell_1 \neq i_1, \ell_3 \neq i_3}} \eta_{i_2}^2 \eta_{i_4}^2 W_{i_1 i_3}^2 W_{i_1 \ell_1} W_{i_3 \ell_3},$$

$$Z_d = v^{-2} \sum_{i_1,i_2,i_3,i_4 (dist)} \sum_{\substack{j_1, \ell_1, j_3, \ell_3 \\ j_1 \neq \ell_1, j_3 \neq \ell_3 \\ j_1, \ell_1 \neq i_3, j_3, \ell_3 \neq i_1}} \eta_{i_2}^2 \eta_{i_4}^2 W_{i_1 j_1} W_{i_1 \ell_1} W_{i_3 j_3} W_{i_3 \ell_3}.$$

We now show (144) and (145) separately. Consider (144) first. It is among all the 6 Z-terms, only Z_a and Z_{b2} have non-zero means. We now consider $\mathbb{E}[Z_a]$ and $\mathbb{E}[Z_{b2}]$ separately.

First, consider $\mathbb{E}[Z_a]$. By similar arguments, both under the null and the alternative,

$$\mathbb{E}[W_{i_1 i_3}^4] \leq C\Omega_{i_1 i_3} \leq C\theta_{i_1}\theta_{i_3}.$$

Recalling that $0 < \eta_i \leq C\theta_i$ and $v \asymp \|\theta\|^2$, it is seen that

$$\mathbb{E}[Z_a] \leq C(\|\theta\|_1)^{-4} \sum_{i_1,i_2,i_3,i_4(dist)} \theta_{i_2}^2\theta_{i_4}^2\theta_{i_1}\theta_{i_3} \leq C\|\theta\|^4/\|\theta\|_1^2. \tag{244}$$

Next, consider $\mathbb{E}[Z_{b2}]$. First, recall that under the null, $\Omega = \theta\theta'$, $v = 1_n'(\Omega - \text{diag}(\Omega))1_n$, and $\eta = v^{-1/2}(\Omega - \text{diag}(\Omega)1_n$. It is seen $v \sim \|\theta\|_1^2$, $\eta_i = (1 + o(1)\theta_i$, $1 \leq i \leq n$, where $o(1) \to 0$ uniformly for all $1 \leq i \leq n$, and for any $i \neq j$, $\mathbb{E}[W_{ij}^2] = (1 + o(1))\theta_i\theta_j$, where $o(1) \to 0$ uniformly for all $1 \leq i, j \leq n$. It follows

$$\mathbb{E}[Z_{b2}] = v^{-2} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_1,j_1\neq i_1,j_3,j_3\neq i_3} \eta_{i_2}^2\eta_{i_4}^2\mathbb{E}[W_{i_1j_1}^2 W_{i_3j_3}^2], \tag{245}$$

which

$$\sim (\|\theta\|_1)^{-4} \sum_{i_1,i_2,i_3,i_4(dist)} \sum_{j_1,j_1\neq i_1,j_3,j_3\neq i_3} \theta_{i_1}\theta_{i_2}^2\theta_{i_3}\theta_{i_4}^2\theta_{j_1}\theta_{j_3} \sim \|\theta\|^4.$$

Second, under the alternative, by similar argument, we have that $v \asymp \|\theta\|_1^2$, $0 < \eta_i < C\theta_i$ for all $1 \leq i \leq n$, and $\mathbb{E}[W_{ij}^2] \leq C\theta_i\theta_j$ for all $1 \leq i, j \leq n$, $i \neq j$. Similar to that under the null, we have

$$0 < |\mathbb{E}[Z_{b2}]| \leq C\|\theta\|^4. \tag{246}$$

Inserting (244), (245), and (246) into (243) and recalling that the mean of all other Z terms are 0,

$$\mathbb{E}[F_3] \sim \|\theta\|^4, \qquad \text{under the null,}$$

and

$$\mathbb{E}[F_3] \leq C\|\theta\|^4, \qquad \text{under the alternative,}$$

where we have used $\|\theta\|_1 \to \infty$. This proves (144).

We now consider (145). By Cauchy-Schwarz inequality,

$$\begin{aligned} \text{Var}(F_c) &\leq C\big(\text{Var}(Z_a) + \text{Var}(Z_{b1}) + \text{Var}(Z_{b2}) + \text{Var}(Z_{c1}) + \text{Var}(Z_{c2}) + \text{Var}(Z_d)\big) \\ &\leq C\big(\text{Var}(Z_a) + \mathbb{E}[Z_{b1}^2] + \text{Var}(Z_{b2}) + \mathbb{E}[Z_{c1}^2] + \mathbb{E}[Z_{c2}^2] + \mathbb{E}[Z_d^2]\big). \end{aligned} \tag{247}$$

Consider $\text{Var}(Z_a)$. Write

$$\text{Var}(Z_a) = v^{-4} \sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ i_1',i_2',i_3',i_4'(dist)}} \eta_{i_2}^2\eta_{i_4}^2\eta_{i_2'}^2\eta_{i_4'}^2\mathbb{E}[(W_{i_1i_3}^4 - \mathbb{E}[W_{i_1i_3}^4])(W_{i_1'i_3'}^4 - \mathbb{E}[W_{i_1'i_3'}^4])].$$

Fix a term $(W_{i_1i_3}^4 - \mathbb{E}[W_{i_1i_3}^4])(W_{i_1'i_3'}^4 - \mathbb{E}[W_{i_1'i_3'}^4])$. When the mean is nonzero, we must have $\{i_1, i_3\} = \{i_1', i_3'\}$, and when this happens,

$$\mathbb{E}[(W_{i_1i_3}^4 - \mathbb{E}[W_{i_1i_3}^4])(W_{i_1'i_3'}^4 - \mathbb{E}[W_{i_1'i_3'}^4])] = \text{Var}(W_{i_1i_3}^4).$$

For a random variable X, we have $\text{Var}(X) \leq \mathbb{E}[X^2]$, and it follows that

$$\text{Var}(W_{i_1i_3}^4) \leq \mathbb{E}[W_{i_1i_3}^8] \leq \mathbb{E}[W_{i_1i_3}^2],$$

where we have used the property that $0 \leq W_{i_1i_3}^2 \leq 1$; note that $\mathbb{E}[W_{i_1i_3}^2] \leq C\theta_{i_1}\theta_{i_3}$. Recall that $v \asymp \|\theta\|_1^2$ and $0 < \eta_i \leq C\theta_i$ for all $1 \leq i \leq n$. Combining these gives

$$\text{Var}(Z_a) \leq C(\|\theta\|_1^{-8}) \cdot \sum_{\substack{i_1,i_2,i_3,i_4(dist)\\ i_2',i_4'(dist)}} \theta_{i_2}^2\theta_{i_4}^2\theta_{i_2'}^2\theta_{i_4'}^2\theta_{i_1}\theta_{i_3} \leq C\|\theta\|^8/\|\theta\|_1^6. \tag{248}$$

We now consider all other terms on the right hand side of (247). Note that

- The proof of $\mathbb{E}[Z_{b1}^2]$ is similar to that of Y_{a1} in Item (b).
- The proof of $\mathrm{Var}(Z_{b2})$ is similar to that of X_a in Item (a).
- The proof of $\mathbb{E}[Z_{c1}^2]$ and $\mathbb{E}[Z_{c2}^2]$ are similar to that of X_b in Item (a).
- The proof of $\mathbb{E}[Z_d^2]$ is similar to that of X_c in Item (a).

For these reasons, we skip the proof details. We have that, under both the null and the alternative,

$$\mathbb{E}[Z_{b1}^2] \le C\|\theta\|^8\|\theta\|_3^3/\|\theta\|_1^5, \tag{249}$$

$$\mathrm{Var}(Z_{b2}) \le C\|\theta\|^8/\|\theta\|_1^2, \tag{250}$$

$$\mathbb{E}[Z_{c1}^2] + \mathbb{E}[Z_{c2}^2] \le C\|\theta\|^{10}/\|\theta\|_1^2, \tag{251}$$

and

$$\mathbb{E}[Z_d^2] \le C\|\theta\|^{12}/\|\theta\|_1^4. \tag{252}$$

Inserting (248), (249), (250), (251) and (252) into (247) gives

$$\begin{aligned}\mathrm{Var}(F_c) &\le C[\|\theta\|^8/\|\theta\|_1^6 + \|\theta\|^8/\|\theta\|_1^2 + \|\theta\|^{10}/\|\theta\|_1^2 + \|\theta\|^{12}/\|\theta\|_1^4] \\ &\le C\|\theta\|^{10}/\|\theta\|_1^2,\end{aligned}$$

which completes the proof of (145).

G.4.9. *Proof of Lemma G.10.* Define an event D as

$$D = \{|V - v| \le \|\theta\|_1 \cdot x_n\}, \qquad \text{for} \quad \sqrt{\log(\|\theta\|_1)} \ll x_n \ll \|\theta\|_1.$$

We aim to show that

$$\mathbb{E}[(Q_n - Q_n^*)^2 \cdot I_{D^c}] = o(\|\theta\|^8). \tag{253}$$

First, we bound the tail probability of $|V - v|$. Write

$$V - v = 2\sum_{i<j}(A_{ij} - \Omega_{ij}).$$

The variables $\{A_{ij} - \Omega_{ij}\}_{1\le i<j\le n}$ are mutually independent with mean zero. They satisfy $|A_{ij} - \Omega_{ij}| \le 1$ and $\sum_{i<j}\mathrm{Var}(A_{ij} - \Omega_{ij}) \le \sum_{i<j}\Omega_{ij} \le 1_n'\Omega 1_n/2 \le \|\theta\|_1^2/2$. Applying the Bernstein's inequality, for any $t > 0$,

$$\mathbb{P}\Big(\Big|2\sum_{i<j}(A_{ij} - \Omega_{ij})\Big| > t\Big) \le 2\exp\Big(-\frac{t^2/2}{2\|\theta\|_1^2 + t/3}\Big).$$

We immediately have that, for some positive constants $C_1, C_2 > 0$,

$$\mathbb{P}(|V - v| > t) \le \begin{cases} 2\exp\big(-\frac{C_1}{\|\theta\|_1^2}t^2\big), & \text{when } x_n\|\theta\|_1 \le t \le \|\theta\|_1^2, \\ 2\exp\big(-C_2 t\big), & \text{when } t > \|\theta\|_1^2. \end{cases} \tag{254}$$

Especially, letting $t = x_n\|\theta\|_1$, we have

$$\mathbb{P}(D^c) \le 2\exp(-C_1 x_n^2). \tag{255}$$

Next, we derive an upper bound of $(Q_n - Q_n^*)^2$ in terms of V. Recall that V is the total number of edges and that $Q_n = \sum_{i,j,k,\ell(dist)} M_{ij}M_{jk}M_{k\ell}M_{\ell i}$, where $M_{ij} = A_{ij} - \hat\eta_i\hat\eta_j$. If one node of i,j,k,ℓ has a zero degree (say, node i), then $A_{ij} = 0$ and $\hat\eta_i = 0$, and it follows that $M_{ij}=0$ and $M_{ij}M_{jk}M_{k\ell}M_{\ell i}=0$. Hence, only when (i,j,k,ℓ) all have nonzero degrees, this quadruple has a contribution to Q_n. Since V is the total number of edges, there are at most V nodes that have a nonzero degree. It follows that

$$|Q_n| \le CV^4.$$

Moreover, $Q_n^* = \sum_{i,j,k,\ell(dist)} M_{ij}^* M_{jk}^* M_{k\ell}^* M_{\ell i}^*$, where $M_{ij}^* = \widetilde{\Omega}_{ij} + W_{ij} + \delta_{ij}$. Re-write $M_{ij}^* = A_{ij} - \eta_i^*\eta_j^* + \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_j - \tilde{\eta}_j)$. First, since $\eta_i^* \le C\theta_i$ and $\eta_i \le C\theta_i$ (see (81)), $|M_{ij}^*| \le A_{ij} + C\theta_i\theta_j + C\theta_i|\eta_j - \tilde{\eta}_j| + C\theta_j|\eta_i - \tilde{\eta}_i|$. Second, note that $\tilde{\eta}_i$ equals to $v^{-1/2}$ times degree of node i, where $v \asymp \|\theta\|_1^2$ according to (80). It follows that $|\eta_i - \tilde{\eta}_i| \le C(\theta_i + \|\theta\|_1^{-1}V)$. Therefore,

$$|M_{ij}^*| \le A_{ij} + C\theta_i\theta_j + C\|\theta\|_1^{-1}V(\theta_i + \theta_j).$$

We plug it into the definition of Q_n^* and note that there are at most V pairs of (i,j) such that $A_{ij} \neq 0$. By elementary calculation,

$$|Q_n^*| \le C(V^4 + \|\theta\|_1^4).$$

Combining the above gives

$$(Q_n - Q_n^*)^2 \le 2Q_n^2 + 2(Q_n^*)^2 \le C(V^8 + \|\theta\|_1^8). \tag{256}$$

Last, we show (253). By (256) and that $V^8 \le Cv^8 + C|V-v|^8$, we have

$$\begin{aligned} \mathbb{E}[(Q_n - Q_n^*)^2 \cdot I_{D^c}] &\le C\mathbb{E}[|V-v|^8 \cdot I_{D^c}] + C(v^8 + \|\theta\|_1^8) \cdot \mathbb{P}(D^c) \\ &\le C\mathbb{E}[|V-v|^8 \cdot I_{D^c}] + C\|\theta\|_1^{16} \cdot \mathbb{P}(D^c), \end{aligned} \tag{257}$$

where the second line is from $v \asymp \|\theta\|_1^2$. Note that $x_n \gg \sqrt{\log(\|\theta\|_1)}$. For n sufficiently large, $x_n^2 \ge 17C_1^{-1}\log(\|\theta\|_1)$. Combining it with (255), we have

$$\|\theta\|_1^{16} \cdot \mathbb{P}(D^c) \le \|\theta\|_1^{16} \cdot 2e^{-C_1 x_n^2} \le \|\theta\|_1^{16} \cdot 2e^{-17\|\theta\|_1} = o(1). \tag{258}$$

We then bound $\mathbb{E}[|V-v|^8 \cdot I_{D^c}]$. Let $f(t)$ and $F(t)$ be the probability density and CDF of $|V-v|$, and write $\bar{F}(t) = 1 - F(t)$. Using integration by part, for any continuously differentiable function $g(t)$ and $x>0$, $\int_x^\infty g(t)f(t)dt = g(x)\bar{F}(x) + \int_x^\infty g'(t)\bar{F}(t)dt$. We apply the formula to $g(t)=t^8$ and $x = x_n\|\theta\|_1$. It yields

$$\begin{aligned} \mathbb{E}[|V-v|^8 \cdot I_{D^c}] &= (x_n\|\theta\|_1)^8 \cdot \mathbb{P}(D^c) + \int_{x_n\|\theta\|_1}^\infty 8t^7 \cdot \mathbb{P}(|V-v|>t)dt \\ &\equiv I + II. \end{aligned}$$

Consider I. By (258) and $x_n \ll \|\theta\|_1$,

$$I \ll \|\theta\|_1^{16} \cdot \mathbb{P}(D^c) = o(1).$$

Consider II. By (254), (258), and elementary probability,

$$\begin{aligned} II &\le 8(\|\theta\|_1^2)^7 \cdot \mathbb{P}\big(x_n\|\theta\|_1 < |V-v| \le \|\theta\|_1^2\big) + \int_{\|\theta\|_1^2} 8t^7 \cdot \mathbb{P}(|V-v|>t)dt \\ &\le C\|\theta\|_1^{14} \cdot \mathbb{P}(D^c) + \int_{\|\theta\|_1^2} 8t^7 \cdot 2e^{-C_2 t}dt \\ &= o(1), \end{aligned}$$

where in the last line we have used (258) and the fact that $\int_x^\infty t^7 e^{-C_2 t}dt \to 0$ as $x \to \infty$. Combining the bounds for I and II gives

$$\mathbb{E}[|V-v|^8 \cdot I_{D^c}] = o(1). \tag{259}$$

Then, (253) follows by plugging (258)-(259) into (257).

TABLE G.4
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

G.4.10. *Proof of Lemma G.11.* There are 175 post-expansion sums in $(\widetilde{Q}_n^* - Q_n^*)$. They divide into 34 different types, denoted by R_1-R_{34} as shown in Table G.4. It suffices to prove that, for each $1 \le k \le 34$, under the null hypothesis,

$$|\mathbb{E}[R_k]| = o(\|\theta\|^4), \qquad \mathrm{Var}(R_k) = o(\|\theta\|^8), \tag{260}$$

and under the alternative hypothesis,

$$|\mathbb{E}[R_k]| = o(\alpha^4 \|\theta\|^8), \qquad \mathrm{Var}(R_k) = O(\|\theta\|^8 + \alpha^6 \|\theta\|^8 \|\theta\|_3^6). \tag{261}$$

We need some preparation. First, recall that $\tilde{r}_{ij} = -\frac{v}{V}(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)$. It follows that each post-expansion sum has the form

$$\left(\frac{v}{V}\right)^{N_{\tilde{r}}} \sum_{i,j,k,\ell(dist)} a_{ij} b_{jk} c_{k\ell} d_{\ell i}, \tag{262}$$

where a_{ij} takes values in $\{\widetilde{\Omega}_{ij}, W_{ij}, \delta_{ij}, -(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)\}$ and $b_{jk}, c_{k\ell}, d_{\ell i}$ are similar. The variable $\frac{v}{V}$ has a complicated correlation with each summand, so we want to get rid of it. Denote the variable in (262) by Y. Write $m = N_{\tilde{r}}$ and

$$Y = \left(\frac{v}{V}\right)^m X, \qquad \text{where} \quad X = \sum_{i,j,k,\ell (dist)} a_{ij} b_{jk} c_{k\ell} d_{\ell i}. \tag{263}$$

We compare the mean and variance of X and Y. By assumption, $\sqrt{\log(\|\theta\|_1)} \ll \|\theta\|_1/\|\theta\|^2$. Then, there exists a sequence x_n such that

$$\sqrt{\log(\|\theta\|_1)} \ll x_n \ll \|\theta\|_1/\|\theta\|^2, \qquad \text{as } n \to \infty.$$

We introduce an event

$$D = \big\{|V - v| \le \|\theta\|_1 x_n\big\}.$$

In Lemma G.10, we have proved $\mathbb{E}[(Q_n - Q_n^*)^2 \cdot I_{D^c}] = o(1)$. By similar proof, we can show: as long as $|Y - X|$ is bounded by a polynomial of V and $\|\theta\|_1$,

$$\mathbb{E}[(Y - X)^2 \cdot I_{D^c}] = o(1). \tag{264}$$

Additionally, on the event D, since $v \asymp \|\theta\|_1^2 \gg \|\theta\|_1 x_n$, we have $|V - v| = o(v)$. It follows that $\frac{|V-v|}{V} \lesssim \frac{|V-v|}{v} \le C\|\theta\|^{-1} x_n = o(1)$. For any fixed $m \ge 1$, $(1+x)^m \le 1 + Cx$ for x being close to 0. Hence, $|1 - \frac{v^m}{V^m}| \le C|1 - \frac{v}{V}| \le C\|\theta\|_1^{-1} x_n = o(\|\theta\|^{-2})$. It implies

$$|Y - X| = o(\|\theta\|^{-2}) \cdot |X|, \qquad \text{on the event } D. \tag{265}$$

By (264)-(265) and elementary probability,

$$\begin{aligned}
|\mathbb{E}[Y - X]| &\le |\mathbb{E}[(Y - X) \cdot I_D]| + |\mathbb{E}[(Y - X) \cdot I_{D^c}]| \\
&\le o(\|\theta\|^{-2}) \cdot \mathbb{E}[|X| \cdot I_D] + \sqrt{\mathbb{E}[(Y - X)^2 \cdot I_{D^c}]} \\
&\le o(\|\theta\|^{-2}) \sqrt{\mathbb{E}[X^2]} + o(1),
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}(Y) &\le 2\mathrm{Var}(X) + 2\mathrm{Var}(Y - X) \\
&\le 2\mathrm{Var}(X) + 2\mathbb{E}[(Y - X)^2] \\
&= 2\mathrm{Var}(X) + 2\mathbb{E}[(Y - X)^2 \cdot I_D] + 2\mathbb{E}[(Y - X)^2 \cdot I_{D^c}] \\
&\le 2\mathrm{Var}(X) + o(\|\theta\|^{-4}) \cdot \mathbb{E}[X^2] + o(1).
\end{aligned}$$

Under the null hypothesis, suppose we can prove that

$$\mathbb{E}[X^2] = o(\|\theta\|^8). \tag{266}$$

Since $\mathbb{E}[X^2] = (\mathbb{E}[X])^2 + \mathrm{Var}(X)$, it implies $|\mathbb{E}[X]| = o(\|\theta\|^4)$ and $\mathrm{Var}(X) = o(\|\theta\|^8)$. Therefore,

$$\begin{aligned}
&|\mathbb{E}[Y]| \le |\mathbb{E}[X]| + |\mathbb{E}[Y - X]| = o(\|\theta\|^4), \\
&\mathrm{Var}(Y) \le C\mathrm{Var}(X) + o(\|\theta\|^{-4}) \cdot \mathbb{E}[X^2] + o(1) = o(\|\theta\|^8).
\end{aligned}$$

Under the alternative hypothesis, suppose we can prove that

$$|\mathbb{E}[X]| = O(\alpha^2 \|\theta\|^6), \qquad \mathrm{Var}(X) = o(\|\theta\|^8 + \alpha^6 \|\theta\|^8 \|\theta\|_3^6). \tag{267}$$

Since $\mathbb{E}[X^2] = (\mathbb{E}[X])^2 + \mathrm{Var}(X)$, we have $\mathbb{E}[X^2] = O(\alpha^4\|\theta\|^{12})$. Then,

$$|\mathbb{E}[Y]| \le O(\alpha^2\|\theta\|^6) + o(\|\theta\|^{-2}) \cdot O(\alpha^2\|\theta\|^6) = o(\alpha^4\|\theta\|^8),$$
$$\mathrm{Var}(Y) \le o(\|\theta\|^8 + \alpha^6\|\theta\|^8\|\theta\|_3^6) + o(\|\theta\|^{-4}) \cdot O(\alpha^4\|\theta\|^{12}) = o(\|\theta\|^8 + \alpha^6\|\theta\|^8\|\theta\|_3^6).$$

In conclusion, to prove that Y satisfies the requirement in (260)-(261), it is sufficient to prove that X satisfies (266)-(267). We remark that (267) puts a more stringent requirement on the mean of the variable, compared to (261).

From now on, in the analysis of each R_k of the form (262), we shall always neglect the factor $(\frac{v}{V})^{N_{\tilde{r}}}$, and show that, after this factor is removed, the random variable satisfies (266)-(267). This is equivalent to pretending

$$\tilde{r}_{ij} = -(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)$$

and proving each R_k satisfies (266)-(267). Unless mentioned, we stick to this mis-use of notation $\tilde{r}_{ij}$ in the proof below.

Second, we divide 34 terms into several groups using the *intrinsic order of* W defined below. Note that $\tilde{r}_{ij} = -(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)$, $\delta_{ij} = \eta_i(\eta_j - \tilde{\eta}_j) + \eta_j(\eta_i - \tilde{\eta}_i)$, and $\tilde{\eta}_i - \eta_i = \frac{1}{\sqrt{v}}\sum_{s\neq i} W_{is}$. We thus have

$$\tilde{r}_{ij} = -\frac{1}{v}\Big(\sum_{s\neq i} W_{is}\Big)\Big(\sum_{t\neq j} W_{jt}\Big), \qquad \delta_{ij} = -\frac{1}{\sqrt{v}}\eta_i\Big(\sum_{t\neq j} W_{jt}\Big) - \frac{1}{\sqrt{v}}\eta_j\Big(\sum_{s\neq i} W_{is}\Big).$$

Each $\tilde{r}_{ij}$ is a weighted sum of terms like $W_{is}W_{jt}$, and each δ_{ij} is a weighted sum of terms like W_{jt}. Intuitively, we view $\tilde{r}$-term as an “order-2 W-term" and view δ-term as “order-1 W-term." It motivates the definition of *intrinsic order of* W as

$$N_W^* = N_W + N_\delta + 2N_{\tilde{r}}. \tag{268}$$

We group 34 terms by the value of N_W^*; see the last column of Table G.4.

G.4.10.1. Analysis of post-expansion sums with $N_W^* \le 4$. There are 14 such terms, including R_2-R_6, R_9-R_{13}, R_{16}-R_{17}, and R_{21}-R_{22}. They all equal to zero under the null hypothesis, so it is sufficient to show that they satisfy (267) under the alternative hypothesis. We prove by comparing each R_k to some previously analyzed terms. Take R_9 for example. Plugging in the definition of $\tilde{r}_{ij}$ and δ_{ij} gives

$$\begin{aligned} R_9 &= \sum_{i,j,k,\ell (dist)} [(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)][(\tilde{\eta}_j - \eta_j)\eta_k + \eta_j(\tilde{\eta}_k - \eta_k)]\widetilde{\Omega}_{k\ell}W_{\ell i} \\ &= R_{9a} + R_{9b}, \end{aligned}$$

where

$$R_{9a} = \sum_{i,j,k,\ell (dist)} \eta_k\widetilde{\Omega}_{k\ell} \cdot [(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)^2 W_{\ell i}],$$
$$R_{9b} = \sum_{i,j,k,\ell (dist)} \eta_j\widetilde{\Omega}_{k\ell} \cdot [(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)(\tilde{\eta}_k - \eta_k)W_{\ell i}]. \tag{269}$$

At the same time, we recall that T_1 in Lemmas G.8-G.9 is defined as

$$T_1 = \sum_{i,j,k,\ell (dist)} \delta_{ij}\delta_{jk}\delta_{k\ell}W_{\ell i} = \sum_{i,j,k,\ell (dist)} \delta_{\ell j}\delta_{jk}\delta_{ki}W_{i\ell}.$$

In the proof of the above two lemmas, we express T_1 as the weighted sum of T_{1a}-T_{1d}; see (130). Note that T_{1a} and T_{1d} in (130) can be re-written as

$$T_{1d} = \sum_{i,j,k,\ell(dist)} [\eta_\ell(\tilde{\eta}_j - \eta_j)][(\tilde{\eta}_j - \eta_j)\eta_k][\eta_k(\tilde{\eta}_i - \eta_i)]W_{i\ell}$$
$$= \sum_{i,j,k,\ell(dist)} \eta_k^2\eta_\ell \cdot [(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)^2 W_{\ell i}],$$
$$T_{1a} = \sum_{i,j,k,\ell(dist)} [\eta_\ell(\tilde{\eta}_j - \eta_j)][\eta_j(\tilde{\eta}_k - \eta_k)][\eta_k(\tilde{\eta}_i - \eta_i)]W_{i\ell}$$
$$= \sum_{i,j,k,\ell(dist)} \eta_j\eta_k\eta_\ell \cdot [(\tilde{\eta}_i - \eta_i)(\tilde{\eta}_j - \eta_j)(\tilde{\eta}_k - \eta_k)W_{i\ell}]. \tag{270}$$

Compare (269) and (270). It is seen that R_{9a} and T_{1d} have the same structure, where the non-stochastic coefficients in the summand satisfy $|\eta_k\widetilde{\Omega}_{k\ell}| \le C\alpha\theta_k^2\theta_\ell$ and $|\eta_k^2\eta_\ell| \le C\theta_k^2\theta_\ell$, respectively. This means we can bound $|\mathbb{E}(R_{9a})|$ and $\mathrm{Var}(R_{9a})$ in the same way as we bound $|\mathbb{E}[T_{1d}]|$ and $\mathrm{Var}(T_{1d})$, and the bounds have an extra factor of α and α^2, respectively. In detail, in the proof of Lemmas G.8-G.9, we have shown

$$|\mathbb{E}[T_{1d}]| \le C\|\theta\|^4, \qquad \mathrm{Var}(T_{1d}) \le \frac{C\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1}.$$

It follows immediately that

$$|\mathbb{E}[R_{9a}]| \le C\alpha\|\theta\|^4 = o(\alpha^2\|\theta\|^6), \qquad \mathrm{Var}(T_{1d}) \le \frac{C\alpha^2\|\theta\|^6\|\theta\|_3^3}{\|\theta\|_1} = o(\|\theta\|^8).$$

Similarly, since we have proved

$$|\mathbb{E}[T_{1a}]| \le \frac{C\|\theta\|^6}{\|\theta\|_1^2}, \qquad \mathrm{Var}(T_{1a}) \le \frac{C\|\theta\|^4\|\theta\|_3^6}{\|\theta\|_1^2},$$

it follows immediately that

$$|\mathbb{E}[R_{9b}]| \le \frac{C\alpha\|\theta\|^6}{\|\theta\|_1^2} = o(\alpha^2\|\theta\|^6), \qquad \mathrm{Var}(R_{9b}) \le \frac{C\alpha^2\|\theta\|^4\|\theta\|_3^6}{\|\theta\|_1^2} = o(\|\theta\|^8).$$

This proves (267) for $X = R_{9a}$.

We use the same strategy to bound all other terms with $N_W^* \le 4$. The details are in Table G.5. In each row of the table, the left column displays a targeting variable X, and the right column displays a previously analyzed variable, which we call X^*, that has a similar structure as X. It is not hard to see that we can obtain upper bounds for $|\mathbb{E}[X]|$ and $\mathrm{Var}(X)$ from multiplying the upper bounds of $|\mathbb{E}[X^*]|$ and $\mathrm{Var}(X^*)$ by α^m and α^{2m}, respectively, where m is a nonnegative integer (e.g., $m = 1$ in the analysis of R_9). Using our previous results, each X^* in the right column satisfies

$$|\mathbb{E}[X^*]| = O(\alpha^2\|\theta\|^6), \qquad \mathrm{Var}(X^*) = o(\|\theta\|^8 + \alpha^6\|\theta\|^8\|\theta\|_3^6).$$

So, each X in the left column satisfies (267).

G.4.10.2. Analysis of post-expansion sums with $N_W^* = 5$. There are 10 such terms, including R_1, R_7-R_8, R_{14}-R_{15}, R_{18}, and R_{25}-R_{28}. Using the the notation

$$G_i = \tilde{\eta}_i - \eta_i,$$

TABLE G.5

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. Definitions of the terms in the right column can be found in (94), (100), (106), (116), (122), (130), (131), and (132). 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	$\sum\eta_i(\tilde\eta_j-\eta_j)\eta_j(\tilde\eta_k-\eta_k)\eta_k(\tilde\eta_\ell-\eta_\ell)\eta_\ell(\tilde\eta_i-\eta_i)$

we get the following expressions (note: factors of $(\frac{v}{V})^m$ have been removed; see explanations in (266)-(267)):

$$R_1 = \sum_{i,j,k,\ell(dist)} G_iG_jW_{jk}W_{k\ell}W_{\ell i},$$

$$\begin{aligned}R_7 &= \sum_{i,j,k,\ell(dist)} G_iG_j\eta_jG_kW_{k\ell}W_{\ell i} + \sum_{i,j,k,\ell(dist)} G_iG_j^2\eta_kW_{k\ell}W_{\ell i}\\ &= \sum_{i,j,k,\ell(dist)} \eta_j(G_iG_jG_kW_{k\ell}W_{\ell i}) + \sum_{i,j,k,\ell(dist)} \eta_k(G_iG_j^2W_{k\ell}W_{\ell i}),\end{aligned}$$

$$R_8 = 2\sum_{i,j,k,\ell(dist)} G_iG_jW_{jk}\eta_kG_\ell W_{\ell i} = 2\sum_{i,j,k,\ell(dist)} \eta_k(G_iG_jG_\ell W_{jk}W_{\ell i}),$$

$$\begin{aligned}R_{14} &= \sum_{\substack{i,j,k,\ell\\(dist)}} G_iG_j^2\eta_k^2G_\ell W_{\ell i} + 2\sum_{\substack{i,j,k,\ell\\(dist)}} G_iG_j^2\eta_kG_k\eta_\ell W_{\ell i} + \sum_{\substack{i,j,k,\ell\\(dist)}} G_iG_j\eta_jG_k\eta_kG_\ell W_{\ell i}\\ &= \sum_{\substack{i,j,k,\ell\\(dist)}} \eta_k^2(G_iG_j^2G_\ell W_{\ell i}) + 2\sum_{\substack{i,j,k,\ell\\(dist)}} \eta_k\eta_\ell(G_iG_j^2G_kW_{\ell i}) + \sum_{\substack{i,j,k,\ell\\(dist)}} \eta_j\eta_k(G_iG_jG_kG_\ell W_{\ell i}),\end{aligned}$$

$$
\begin{aligned}
R_{15} &= \sum_{\substack{i,j,k,\ell\\(dist)}} G_iG_j\eta_jG_kW_{k\ell}G_\ell\eta_i + 2\sum_{\substack{i,j,k,\ell\\(dist)}} G_iG_j^2\eta_kW_{k\ell}G_\ell\eta_i + \sum_{\substack{i,j,k,\ell\\(dist)}} G_iG_j^2\eta_kW_{k\ell}\eta_\ell G_i\\
&= \sum_{\substack{i,j,k,\ell\\(dist)}} \eta_i\eta_j(G_iG_jG_kG_\ell W_{k\ell}) + 2\sum_{\substack{i,j,k,\ell\\(dist)}} \eta_i\eta_k(G_iG_j^2G_\ell W_{k\ell}) + \sum_{\substack{i,j,k,\ell\\(dist)}} \eta_k\eta_\ell(G_i^2G_j^2W_{k\ell}),\\
R_{18} &= 4\sum_{i,j,k,\ell(dist)} \eta_j\eta_k\eta_\ell(G_i^2G_jG_kG_\ell) + 4\sum_{i,j,k,\ell(dist)} \eta_k\eta_\ell^2(G_i^2G_j^2G_k),\\
R_{25} &= \sum_{i,j,k,\ell(dist)} G_iG_j^2G_k\widetilde{\Omega}_{k\ell}W_{\ell i} = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{k\ell}(G_iG_j^2G_kW_{\ell i}),\\
R_{26} &= \sum_{i,j,k,\ell(dist)} G_iG_j\widetilde{\Omega}_{jk}G_kG_\ell W_{\ell i} = \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{jk}(G_iG_jG_kG_\ell W_{\ell i}),\\
R_{27} &= \sum_{i,j,k,\ell(dist)} G_iG_j^2G_k\eta_kG_\ell\widetilde{\Omega}_{\ell i} + \sum_{i,j,k,\ell(dist)} G_iG_j^2G_k^2\eta_\ell\widetilde{\Omega}_{\ell i}\\
&= \sum_{i,j,k,\ell(dist)} \eta_k\widetilde{\Omega}_{\ell i}(G_iG_j^2G_kG_\ell) + \sum_{i,j,k,\ell(dist)} \eta_\ell\widetilde{\Omega}_{\ell i}(G_iG_j^2G_k^2),\\
R_{28} &= 2\sum_{i,j,k,\ell(dist)} G_iG_j\eta_jG_k^2G_\ell\widetilde{\Omega}_{\ell i} = 2\sum_{i,j,k,\ell(dist)} \eta_j\widetilde{\Omega}_{\ell i}(G_iG_jG_k^2G_\ell).
\end{aligned}
$$

Each expression above belongs to one of the following types:

$$
\begin{aligned}
J_1 &= \sum_{i,j,k,\ell(dist)} G_iG_jW_{jk}W_{k\ell}W_{\ell i}, & J_2 &= \sum_{i,j,k,\ell(dist)} \eta_j(G_iG_jG_kW_{k\ell}W_{\ell i}),\\
J_3 &= \sum_{i,j,k,\ell(dist)} \eta_k(G_iG_jG_\ell W_{jk}W_{\ell i}), & J_4 &= \sum_{i,j,k,\ell(dist)} \eta_k(G_iG_j^2W_{k\ell}W_{\ell i}),\\
J_5 &= \sum_{i,j,k,\ell(dist)} \eta_j\eta_k(G_iG_jG_kG_\ell W_{\ell i}), & J_5' &= \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{jk}(G_iG_jG_kG_\ell W_{\ell i}),\\
J_6 &= \sum_{i,j,k,\ell(dist)} \eta_k\eta_\ell(G_iG_j^2G_kW_{\ell i}), & J_6' &= \sum_{i,j,k,\ell(dist)} \widetilde{\Omega}_{k\ell}(G_iG_j^2G_kW_{\ell i}),\\
J_7 &= \sum_{i,j,k,\ell(dist)} \eta_k^2(G_iG_j^2G_\ell W_{\ell i}), & J_8 &= \sum_{i,j,k,\ell(dist)} \eta_k\eta_\ell(G_i^2G_j^2W_{k\ell}),\\
J_9 &= \sum_{i,j,k,\ell(dist)} \eta_k\widetilde{\Omega}_{\ell i}(G_iG_j^2G_kG_\ell), & J_{10} &= \sum_{i,j,k,\ell(dist)} \eta_\ell\widetilde{\Omega}_{\ell i}(G_iG_j^2G_k^2).
\end{aligned}
$$

Since $|\eta_j\eta_k| \le C\theta_j\theta_k$ and $|\widetilde{\Omega}_{jk}| \le C\alpha\theta_j\theta_k$, the study of J_5 and J_5' are similar. Also, the study of J_6 and J_6' are similar. We now study J_1-J_{10}. Consider J_1. It is seen that

$$
J_1 = \frac{1}{v}\sum_{i,j,k,\ell(dist)} \Big(\sum_{s\neq i} W_{is}\Big)\Big(\sum_{t\neq j} W_{jt}\Big)W_{jk}W_{k\ell}W_{\ell i} = \frac{1}{v}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq i, t\neq j}} W_{is}W_{i\ell}W_{jt}W_{jk}W_{k\ell}.
$$

Since s can be equal to ℓ and t can be equal to k, there are three different types:

$$J_{1a} = \frac{1}{v} \sum_{i,j,k,\ell(dist)} W_{i\ell}^2 W_{jk}^2 W_{k\ell}, \qquad J_{1b} = \frac{1}{v} \sum_{\substack{i,j,k,\ell(dist)\\ t\notin\{j,k\}}} W_{i\ell}^2 W_{jt} W_{jk} W_{k\ell},$$

$$J_{1c} = \frac{1}{v} \sum_{\substack{i,j,k,\ell(dist)\\ s\notin\{i,\ell\}, t\notin\{j,k\}}} W_{is} W_{i\ell} W_{jt} W_{jk} W_{k\ell}.$$

We now calculate $\mathbb{E}[J_{1a}^2]$-$\mathbb{E}[J_{1c}^2]$. Take J_{1a} for example. In order to get nonzero $\mathbb{E}[W_{i\ell}^2 W_{jk}^2 W_{k\ell} W_{i'\ell'}^2 W_{j'k'}^2 W_{k'\ell'}]$, we need either $W_{k\ell} = W_{k'\ell'}$ or each of the two variables $(W_{k\ell}, W_{k',\ell'})$ equals to another squared-W term. The leading term of $\mathbb{E}[J_{1a}^2]$ comes from the first case. In this case, we have $W_{k\ell} = W_{k'\ell'}$ but allow for $W_{i\ell} \neq W_{i'\ell'}$ and $W_{jk} \neq W_{j'k'}$. It has to be the case of either $(k', \ell') = (k, \ell)$ or $(k', \ell') = (\ell, k)$. Therefore, we have $\mathbb{E}[W_{i\ell}^2 W_{jk}^2 W_{k\ell} W_{i'\ell'}^2 W_{j'k'}^2 W_{k'\ell'}] = \mathbb{E}[W_{i\ell}^2 W_{jk}^2 W_{i'\ell'}^2 W_{j'k'}^2 W_{k\ell}^2]$. Using similar arguments, we have the following results, where details are omitted, as they are similar to the calculations in the proof of Lemmas G.4-G.9.

$$\mathbb{E}[J_{1a}^2] \leq \frac{C}{v^2} \sum_{\substack{i,j,k,\ell\\ i',j'}} \mathbb{E}[W_{i\ell}^2 W_{jk}^2 W_{i'\ell}^2 W_{j'k}^2 W_{k\ell}^2] \leq \frac{C}{\|\theta\|_1^4} \sum_{\substack{i,j,k,\ell\\ i',j'}} \theta_i \theta_j \theta_k^3 \theta_\ell^3 \theta_{i'} \theta_{j'} \leq C\|\theta\|_3^6,$$

$$\mathbb{E}[J_{1b}^2] \leq \frac{C}{v^2} \sum_{\substack{i,j,k,\ell,t\\ i'}} \mathbb{E}[W_{i\ell}^2 W_{i'\ell}^2 W_{jt}^2 W_{jk}^2 W_{k\ell}^2] \leq \frac{C}{\|\theta\|_1^4} \sum_{\substack{i,j,k,\ell,t\\ i'}} \theta_i \theta_j^2 \theta_k^2 \theta_\ell^3 \theta_t \theta_{i'} \leq \frac{C\|\theta\|^4 \|\theta\|_3^3}{\|\theta\|_1},$$

$$\mathbb{E}[J_{1c}^2] \leq \frac{C}{v^2} \sum_{i,j,k,\ell,s,t} \mathbb{E}[W_{is}^2 W_{i\ell}^2 W_{jt}^2 W_{jk}^2 W_{k\ell}^2] \leq \frac{C}{\|\theta\|_1^4} \sum_{i,j,k,\ell,s,t} \theta_i^2 \theta_j^2 \theta_k^2 \theta_\ell^2 \theta_s \theta_t \leq \frac{C\|\theta\|^8}{\|\theta\|_1^2}.$$

The right hand sides are all $o(\|\theta\|^8)$. It follows that

$$\mathbb{E}[J_1^2] = o(\|\theta\|^8), \qquad \text{under both hypotheses.}$$

Consider J_2-J_4. By definition,

$$J_2 = \frac{1}{v\sqrt{v}} \sum_{\substack{i,j,k,\ell(dist)\\ s\neq i, t\neq j, q\neq k}} \eta_j W_{is} W_{jt} W_{kq} W_{k\ell} W_{\ell i}, \qquad J_3 = \frac{1}{v\sqrt{v}} \sum_{\substack{i,j,k,\ell(dist)\\ s\neq i, t\neq j, q\neq \ell}} \eta_k W_{is} W_{jt} W_{\ell q} W_{jk} W_{\ell i},$$

$$J_4 = \frac{1}{v\sqrt{v}} \sum_{\substack{i,j,k,\ell(dist)\\ s\neq i, t\neq j, q\neq j}} \eta_k W_{is} W_{jt} W_{jq} W_{k\ell} W_{\ell i}.$$

The analysis is summarized in Table G.6. In the first column of this table, we study different types of summands. For example, in the expression of J_2, $W_{is} W_{kq} W_{k\ell} W_{\ell i}$ have four different cases: (a) $W_{k\ell}^2 W_{\ell i}^2$, (b) $W_{k\ell}^2 W_{\ell i} W_{is}$ or $W_{k\ell} W_{\ell i}^2 W_{kq}$, (c) $W_{k\ell} W_{\ell i} W_{ik}^2$, and (d) $W_{k\ell} W_{\ell i} W_{is} W_{kq}$. In cases (b) and (d), W_{is} or W_{kq} may further equal to W_{jt}. Having explored all variants and considered index symmetry, we end up with 6 different cases, as listed in the first column of Table G.6. In the second column, we study the mean of the squares of the sum of each type of summands. Take the first row for example. We aim to study

$$\mathrm{E}\Big[\Big(\sum_{\substack{i,j,k,\ell(dist)\\ t\neq j}} \eta_j (W_{k\ell}^2 W_{\ell i}^2) W_{jt} \Big)\Big].$$

The naive expansion gives the sum of $\eta_j\eta_{j'}\,\mathbb{E}[W_{k\ell}^2W_{\ell i}^2W_{jt}W_{k'\ell'}^2W_{\ell' i'}^2W_{j't'}]$ over $(i,j,k,\ell,t,i',j',k',\ell',t')$. However, for this term to be nonzero, all single-W terms have to be paired (either with another single-W term or with a squared-W term). The main contribution is from the case of $W_{jt}=W_{j't'}$. This is satisfied only when $(j',s')=(j,s)$ or $(j',s')=(s,j)$. By calculations which are omitted here, we can show that $(j',s')=(j,s)$ yields a larger bound. Therefore, it reduces to the sum of $\eta_j^2\,\mathbb{E}[(W_{jt}^2)W_{k\ell}^2W_{\ell i}^2W_{k'\ell'}^2W_{\ell' i'}^2]$ over $(i,j,k,\ell,t,i',k',\ell')$, which is displayed in the second column of the table. In the last column, we sum the quantity in the second column over indices; it gives rise to a bound for the mean of the square of sum. See the table for details. Recall that the definition of J_2-J_4 contains a factor of $\frac{1}{v\sqrt{v}}$ in front of the sum, where $v\asymp\|\theta\|_1^2$ by (80). Hence, to get a desired bound, we only need that each row in the third column of Table G.6 is

$$o(\|\theta\|^8\|\theta\|_1^6).$$

This is true. We thus conclude that

$$\max\left\{\mathbb{E}[J_2^2],\,\mathbb{E}[J_3^2],\,\mathbb{E}[J_4^2]\right\}=o(\|\theta\|^8),\qquad\text{under both hypotheses.}$$

TABLE G.6

Analysis of J_2-J_4. In the second column, the variables in brackets are paired W terms.

	Types of summand	Terms in mean-squared	Bound
	$\eta_j(W_{k\ell}^2W_{\ell i}^2)W_{jt}$	$\eta_j^2\,\mathbb{E}[(W_{jt}^2)W_{k\ell}^2W_{\ell i}^2W_{k'\ell'}^2W_{\ell' i'}^2]\le\theta_i\theta_j^3\theta_k\theta_\ell^2\theta_t\theta_{i'}\theta_{k'}\theta_{\ell'}^2$	$\|\theta\|^4\|\theta\|_3^3\|\theta\|_1^5$
	$\eta_j(W_{k\ell}W_{\ell i}W_{ik}^2)W_{jt}$	$\eta_j^2\,\mathbb{E}[(W_{k\ell}^2W_{\ell i}^2W_{jt}^2)W_{ik}^4]\le C\theta_i^2\theta_j^3\theta_k^2\theta_\ell^2\theta_t$	$\|\theta\|^6\|\theta\|_3^3\|\theta\|_1$
J_2	$\eta_j(W_{k\ell}^2W_{\ell i}W_{is})W_{jt}$	$\eta_j^2\,\mathbb{E}[(W_{\ell i}^2W_{is}^2W_{jt}^2)W_{k\ell}^2W_{k'\ell}^2]\le C\theta_i^2\theta_j^3\theta_k\theta_\ell^3\theta_s\theta_t\theta_{k'}$	$\|\theta\|^2\|\theta\|_3^6\|\theta\|_1^4$
	$\eta_j(W_{k\ell}^2W_{\ell i})W_{ij}^2$	$\eta_j\eta_{j'}\,\mathbb{E}[(W_{\ell i}^2)W_{k\ell}^2W_{ij}^2W_{k'\ell}^2W_{ij'}^2]\le C\theta_i^3\theta_j^2\theta_k\theta_\ell^3\theta_{j'}^2\theta_{k'}$	$\|\theta\|^4\|\theta\|_3^6\|\theta\|_1^2$
	$\eta_j(W_{k\ell}W_{\ell i}W_{kq}W_{is})W_{jt}$	$\eta_j^2\,\mathbb{E}[(W_{k\ell}^2W_{\ell i}^2W_{kq}^2W_{is}^2W_{jt}^2)]\le C\theta_i^2\theta_j^3\theta_k^2\theta_\ell^2\theta_s\theta_t\theta_q$	$\|\theta\|^6\|\theta\|_3^3\|\theta\|_1^3$
	$\eta_j(W_{k\ell}W_{\ell i})W_{kq}W_{ij}^2$	$\eta_j\eta_{j'}\,\mathbb{E}[(W_{k\ell}^2W_{\ell i}^2W_{kq}^2)W_{ij}^2W_{ij'}^2]\le C\theta_i^3\theta_j^2\theta_k^2\theta_\ell^2\theta_q\theta_{j'}^2$	$\|\theta\|^8\|\theta\|_3^3\|\theta\|_1$
	$\eta_kW_{\ell i}^3W_{jk}^2$	$\eta_k\eta_{k'}\,\mathbb{E}[W_{\ell i}^3W_{jk}^2W_{\ell' i'}^3W_{j'k'}^2]\le C\theta_i\theta_j\theta_k^2\theta_\ell\theta_{i'}\theta_{j'}\theta_{k'}^2\theta_{\ell'}$	$\|\theta\|^4\|\theta\|_1^6$
	$\eta_kW_{\ell i}^3(W_{jk}W_{jt})$	$\eta_k^2\,\mathbb{E}[(W_{jk}^2W_{jt}^2)W_{\ell i}^3W_{\ell' i'}^3]\le C\theta_i\theta_j^2\theta_k^3\theta_\ell\theta_t\theta_{i'}\theta_{\ell'}$	$\|\theta\|^2\|\theta\|_3^3\|\theta\|_1^5$
	$\eta_k(W_{\ell i}^2W_{is})W_{jk}^2$	$\eta_k\eta_{k'}\,\mathbb{E}[(W_{is}^2)W_{\ell i}^2W_{jk}^2W_{\ell' i}^2W_{j'k'}^2]\le C\theta_i^3\theta_j\theta_k^2\theta_\ell\theta_s\theta_{j'}\theta_{k'}^2\theta_{\ell'}$	$\|\theta\|^4\|\theta\|_3^3\|\theta\|_1^5$
J_3	$\eta_k(W_{\ell i}^2W_{is})W_{jk}W_{jt}$	$\eta_k^2\,\mathbb{E}[(W_{is}^2W_{jk}^2W_{jt}^2)W_{\ell i}^2W_{\ell' i}^2]\le C\theta_i^3\theta_j^2\theta_k^3\theta_\ell\theta_s\theta_t\theta_{\ell'}$	$\|\theta\|^2\|\theta\|_3^6\|\theta\|_1^4$
	$\eta_kW_{\ell i}^2W_{ij}^2W_{jk}$	$\eta_k^2\,\mathbb{E}[(W_{jk}^2)W_{\ell i}^2W_{ij}^2W_{\ell' i'}^2W_{i'j}^2]\le C\theta_i^2\theta_j^3\theta_k^3\theta_\ell\theta_{i'}^2\theta_{\ell'}$	$\|\theta\|^4\|\theta\|_3^6\|\theta\|_1^2$
	$\eta_k(W_{\ell i}W_{is}W_{\ell q})W_{jk}^2$	$\eta_k\eta_{k'}\,\mathbb{E}[(W_{\ell i}^2W_{is}^2W_{\ell q}^2)W_{jk}^2W_{j'k'}^2]\le C\theta_i^2\theta_j\theta_k^2\theta_\ell^2\theta_s\theta_q\theta_{j'}\theta_{k'}^2$	$\|\theta\|^8\|\theta\|_1^4$
	$\eta_k(W_{\ell i}W_{is}W_{\ell q})W_{jk}W_{jt}$	$\eta_k^2\,\mathbb{E}[(W_{\ell i}^2W_{is}^2W_{\ell q}^2W_{jk}^2W_{jt}^2)]\le C\theta_i^2\theta_j^2\theta_k^3\theta_\ell^2\theta_s\theta_t\theta_q$	$\|\theta\|^6\|\theta\|_3^3\|\theta\|_1^3$
	$\eta_kW_{\ell i}W_{ij}^2W_{\ell q}W_{jk}$	$\eta_k^2\,\mathbb{E}[(W_{\ell i}^2W_{\ell q}^2W_{jk}^2)W_{ij}^4]\le C\theta_i^2\theta_j^2\theta_k^3\theta_\ell^2\theta_q$	$\|\theta\|^6\|\theta\|_3^3\|\theta\|_1$
	$\eta_k(W_{k\ell}W_{\ell i}^2)W_{jt}^2$	$\eta_k^2\,\mathbb{E}[(W_{k\ell}^2)W_{\ell i}^2W_{jt}^2W_{\ell i'}^2W_{j't'}^2]\le C\theta_i\theta_j\theta_k^3\theta_\ell^3\theta_t\theta_{i'}\theta_{j'}\theta_{t'}$	$\|\theta\|_3^6\|\theta\|_1^6$
	$\eta_k(W_{k\ell}W_{\ell i}^2)W_{jt}W_{jq}$	$\eta_k^2\,\mathbb{E}[(W_{k\ell}^2W_{jt}^2W_{jq}^2)W_{\ell i}^2W_{\ell i'}^2]\le C\theta_i\theta_j^2\theta_k^3\theta_\ell^3\theta_t\theta_q\theta_{i'}$	$\|\theta\|^2\|\theta\|_3^6\|\theta\|_1^4$
J_4	$\eta_k(W_{k\ell}W_{\ell i}W_{is})W_{jt}^2$	$\eta_k^2\,\mathbb{E}[(W_{k\ell}^2W_{\ell i}^2W_{is}^2)W_{jt}^2W_{j't'}^2]\le C\theta_i^2\theta_j\theta_k^3\theta_\ell^2\theta_s\theta_t\theta_{j'}\theta_{t'}$	$\|\theta\|^4\|\theta\|_3^3\|\theta\|_1^5$
	$\eta_kW_{k\ell}W_{\ell i}W_{ij}^3$	$\eta_k^2\,\mathbb{E}[(W_{k\ell}^2W_{\ell i}^2)W_{ij}^3W_{ij'}^3]\le C\theta_i^3\theta_j\theta_k^3\theta_\ell^2\theta_{j'}$	$\|\theta\|^2\|\theta\|_3^6\|\theta\|_1^2$
	$\eta_k(W_{k\ell}W_{\ell i}W_{is})W_{jt}W_{jq}$	$\eta_k^2\,\mathbb{E}[(W_{k\ell}^2W_{\ell i}^2W_{is}^2W_{jt}^2W_{jq}^2]\le C\theta_i^2\theta_j^2\theta_k^3\theta_\ell^2\theta_s\theta_t\theta_q$	$\|\theta\|^6\|\theta\|_3^3\|\theta\|_1^3$
	$\eta_kW_{k\ell}W_{\ell i}W_{ij}^2W_{jq}$	$\eta_k^2\,\mathbb{E}[(W_{k\ell}^2W_{\ell i}^2W_{jq}^2)W_{ij}^4]\le C\theta_i^2\theta_j^2\theta_k^3\theta_\ell^2\theta_q$	$\|\theta\|^6\|\theta\|_3^3\|\theta\|_1$

Consider J_5-J_8. It is seen that

$$J_5=\frac{1}{v^2}\sum_{i,j,k,\ell(dist)}\eta_j\eta_kW_{is}W_{jt}W_{kq}W_{\ell m}W_{\ell i},\quad J_6=\frac{1}{v^2}\sum_{i,j,k,\ell(dist)}\eta_k\eta_\ell W_{is}W_{jt}W_{jq}W_{km}W_{\ell i},$$

$$J_7=\frac{1}{v^2}\sum_{i,j,k,\ell(dist)}\eta_k^2W_{is}W_{jt}W_{jq}W_{\ell m}W_{\ell i},\qquad J_8=\frac{1}{v^2}\sum_{i,j,k,\ell(dist)}\eta_k\eta_\ell W_{is}W_{it}W_{jq}W_{jm}W_{k\ell},$$

The analysis is summarized in Table G.7. We note that J_7 can be written as

$$J_7 = \frac{1}{v^2} \sum_{i,j,\ell (dist)} \beta_{ij\ell} W_{is} W_{jt} W_{jq} W_{\ell m} W_{\ell i}, \qquad \text{where} \quad \beta_{ij\ell} \equiv \sum_{k \notin \{i,j,\ell\}} \eta_k^2.$$

Although the values of $\beta_{ij\ell}$ change with indices, they have a common upper bound of $C\|\theta\|^2$. We treat $\beta_{ij\ell}$ as $\|\theta\|^2$ in Table G.7, as this doesn't change the bounds but simplifies notations. Recall that the definition of J_5-J_8 contains a factor of $\frac{1}{v^2}$ in front of the sum, where $v \asymp \|\theta\|_1^2$ by (80). Hence, to get a desired bound, we only need that each row in the third column of Table G.6 is

$$o(\|\theta\|^8 \|\theta\|_1^8).$$

This is true. We thus conclude that

$$\max\left\{\mathbb{E}[J_5^2],\, \mathbb{E}[J_6^2],\, \mathbb{E}[J_7^2],\, \mathbb{E}[J_8^2]\right\} = o(\|\theta\|^8), \qquad \text{under both hypotheses.}$$

Consider J_9-J_{10}. They can be analyzed in the same way as we did for J_1-J_8. To save space, we only give a simplified proof for the case of $\|\theta\| \gg \alpha[\log(n)]^{5/2}$. For $1 \ll \|\theta\| \leq C\alpha[\log(n)]^{5/2}$, the proof is similar to those in Tables G.6-G.7, which is omitted. For a constant $C_0 > 0$ to be decided, we introduce an event

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

Recall that $\sqrt{v}G_i = \sqrt{v}(\tilde{\eta}_i - \eta_i) = \sum_{j \neq i}(A_{ij} - \mathbb{E}A_{ij})$. The variables $\{A_{ij}\}_{j \neq i}$ are mutually independent, satisfying that $|A_{ij} - \mathbb{E}A_{ij}| \leq 1$ and $\sum_j \mathrm{Var}(A_{ij}) \leq \sum_j \theta_i\theta_j \leq \theta_i\|\theta\|_1$. By Bernstein's inequality, for large n, the probability of E_i^c is $O(n^{-C_0/4.1})$. Applying the probability union bound, we find that the probability of E^c is $O(n^{-C_0/2.01})$. Recall that $V = \sum_{i,j:i\neq j} A_{ij}$. On the event E^c, if $V = 0$ (i.e., the network has no edges), then $\widetilde{Q}_n^* = Q_n^* = 0$; otherwise, $V \geq 1$ and $|\widetilde{Q}_n^* - Q_n^*| \leq n^4$. Combining these results gives

$$\mathbb{E}\big[|\widetilde{Q}_n^* - Q_n^*|^2 \cdot I_{E^c}\big] \leq n^4 \cdot O(n^{-C_0/2.01}).$$

With an properly large C_0, the right hand side is $o(\|\theta\|^8)$. Hence, it suffices to focus on the event E. On the event E,

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

TABLE G.7

Analysis of J_5-J_8. In the second column, the variables in brackets are paired W terms.

	Types of summand	Terms in mean-squared	Bound
	$\eta_j\eta_k W_{\ell i}^3 W_{jk}^2$	$\eta_j\eta_k\eta_{j'}\eta_{k'}\,\mathbb{E}[W_{\ell i}^3 W_{jk}^2 W_{\ell' i'}^3 W_{j'k'}^2] \leq C\theta_i\theta_j^2\theta_k^2\theta_\ell\theta_{i'}\theta_{j'}^2\theta_{k'}^2\theta_{\ell'}$	$\Vert\theta\Vert^8\Vert\theta\Vert_1^4$
	$\eta_j\eta_k W_{\ell i}^3(W_{jt}W_{kq})$	$\eta_j^2\eta_k^2\,\mathbb{E}[(W_{jt}^2 W_{kq}^2)W_{\ell i}^3 W_{\ell' i'}^3] \leq C\theta_i\theta_j^3\theta_k^3\theta_\ell\theta_t\theta_q\theta_{i'}\theta_{\ell'}$	$\Vert\theta\Vert_3^6\Vert\theta\Vert_1^6$
	$\eta_j\eta_k(W_{\ell i}^2 W_{is})W_{jk}^2$	$\eta_j\eta_k\eta_{j'}\eta_{k'}\,\mathbb{E}[(W_{is}^2)W_{\ell i}^2 W_{jk}^2 W_{\ell' i}^2 W_{j'k'}^2] \leq C\theta_i^3\theta_j^2\theta_k^2\theta_\ell\theta_s\theta_{j'}^2\theta_{k'}^2\theta_{\ell'}$	$\Vert\theta\Vert^8\Vert\theta\Vert_3^3\Vert\theta\Vert_1^3$
J_5	$\eta_j\eta_k(W_{\ell i}^2 W_{is})(W_{jt}W_{kq})$	$\eta_j^2\eta_k^2\,\mathbb{E}[(W_{is}^2)W_{jt}^2 W_{kq}^2)W_{\ell i}^2 W_{\ell' i}^2] \leq C\theta_i^3\theta_j^3\theta_k^3\theta_\ell\theta_s\theta_t\theta_q\theta_{\ell'}$	$\Vert\theta\Vert_3^9\Vert\theta\Vert_1^5$
	$\eta_j\eta_k W_{\ell i}^2 W_{ij}^2 W_{kq}$	$\eta_j\eta_k^2\eta_{j'}\,\mathbb{E}[(W_{kq}^2)W_{\ell i}^2 W_{ij}^2 W_{\ell' i'}^2 W_{i'j'}^2] \leq C\theta_i^2\theta_j^2\theta_k^3\theta_\ell\theta_q\theta_{i'}^2\theta_{j'}^2\theta_{\ell'}$	$\Vert\theta\Vert^8\Vert\theta\Vert_3^3\Vert\theta\Vert_1^3$
	$\eta_j\eta_k(W_{\ell i}W_{is}W_{\ell m})W_{jk}^2$	$\eta_j\eta_k\eta_{j'}\eta_{k'}\,\mathbb{E}[(W_{\ell i}^2 W_{is}^2 W_{\ell m}^2)W_{jk}^2 W_{j'k'}^2] \leq C\theta_i^2\theta_j^2\theta_k^2\theta_\ell^2\theta_s\theta_m\theta_{j'}^2\theta_{k'}^2$	$\Vert\theta\Vert^{12}\Vert\theta\Vert_1^2$
	$\eta_j\eta_k(W_{\ell i}W_{is}W_{\ell m})(W_{jt}W_{kq})$	$\eta_j^2\eta_k^2\,\mathbb{E}[(W_{\ell i}^2 W_{is}^2 W_{\ell m}^2 W_{jt}^2 W_{kq}^2)] \leq C\theta_i^2\theta_j^3\theta_k^3\theta_\ell^2\theta_s\theta_t\theta_q\theta_m$	$\Vert\theta\Vert^4\Vert\theta\Vert_3^6\Vert\theta\Vert_1^4$
	$\eta_j\eta_k W_{\ell i}W_{ij}^2 W_{\ell m}W_{kq}$	$\eta_j\eta_k^2\eta_{j'}\,\mathbb{E}[(W_{\ell i}^2 W_{\ell m}^2 W_{kq}^2)W_{ij}^2 W_{ij'}^2] \leq C\theta_i^3\theta_j^2\theta_k^3\theta_\ell^2\theta_q\theta_m\theta_{j'}^2$	$\Vert\theta\Vert^6\Vert\theta\Vert_3^6\Vert\theta\Vert_1^2$
	$\eta_k\eta_\ell W_{\ell i}^2 W_{jt}^2 W_{km}$	$\eta_k^2\eta_\ell\eta_{\ell'}\,\mathbb{E}[(W_{km}^2)W_{\ell i}^2 W_{jt}^2 W_{\ell' i'}^2 W_{j't'}^2] \leq C\theta_i\theta_j\theta_k^3\theta_\ell^2\theta_t\theta_m\theta_{i'}\theta_{j'}\theta_{\ell'}^2\theta_{t'}$	$\Vert\theta\Vert^4\Vert\theta\Vert_3^3\Vert\theta\Vert_1^7$
	$\eta_k\eta_\ell W_{\ell i}^2 W_{jk}^3$	$\eta_k\eta_\ell\eta_{k'}\eta_{\ell'}\,\mathbb{E}[W_{\ell i}^2 W_{jk}^3 W_{\ell' i'}^2 W_{j'k'}^3] \leq C\theta_i\theta_j\theta_k^2\theta_\ell^2\theta_{i'}\theta_{j'}\theta_{k'}^2\theta_{\ell'}^2$	$\Vert\theta\Vert^8\Vert\theta\Vert_1^4$
	$\eta_k\eta_\ell W_{\ell i}^2(W_{jt}W_{jq})W_{km}$	$\eta_k^2\eta_\ell\eta_{\ell'}\,\mathbb{E}[(W_{jt}^2 W_{jq}^2 W_{km}^2)W_{\ell i}^2 W_{\ell' i'}^2] \leq C\theta_i\theta_j^2\theta_k^3\theta_\ell^2\theta_t\theta_q\theta_m\theta_{i'}\theta_{\ell'}^2$	$\Vert\theta\Vert^6\Vert\theta\Vert_3^3\Vert\theta\Vert_1^5$
	$\eta_k\eta_\ell W_{\ell i}^2 W_{jk}^2 W_{jq}$	$\eta_k\eta_\ell\eta_{k'}\eta_{\ell'}\,\mathbb{E}[(W_{jq}^2)W_{\ell i}^2 W_{jk}^2 W_{\ell' i'}^2 W_{jk'}^2] \leq C\theta_i\theta_j^3\theta_k^2\theta_\ell^2\theta_q\theta_{i'}\theta_{k'}^2\theta_{\ell'}^2$	$\Vert\theta\Vert^8\Vert\theta\Vert_3^3\Vert\theta\Vert_1^3$
	$\eta_k\eta_\ell(W_{\ell i}W_{is})W_{jt}^2 W_{km}$	$\eta_k^2\eta_\ell^2\,\mathbb{E}[(W_{\ell i}^2 W_{is}^2 W_{km}^2)W_{jt}^2 W_{j't'}^2] \leq C\theta_i^2\theta_j\theta_k^3\theta_\ell^3\theta_s\theta_t\theta_m\theta_{j'}\theta_{t'}$	$\Vert\theta\Vert^2\Vert\theta\Vert_3^6\Vert\theta\Vert_1^6$
J_6	$\eta_k\eta_\ell W_{\ell i}W_{ij}^3 W_{km}$	$\eta_k^2\eta_\ell^2\,\mathbb{E}[(W_{\ell i}^2 W_{km}^2)W_{ij}^3 W_{ij'}^3] \leq C\theta_i^3\theta_j\theta_k^3\theta_\ell^3\theta_m\theta_{j'}$	$\Vert\theta\Vert_3^9\Vert\theta\Vert_1^3$
	$\eta_k\eta_\ell W_{\ell i}W_{is}W_{jk}^3$	$\eta_k\eta_\ell^2\eta_{k'}\,\mathbb{E}[(W_{\ell i}^2 W_{is}^2)W_{jk}^3 W_{j'k'}^3] \leq C\theta_i^2\theta_j\theta_k^2\theta_\ell^3\theta_s\theta_{j'}\theta_{k'}^2$	$\Vert\theta\Vert^6\Vert\theta\Vert_3^3\Vert\theta\Vert_1^3$
	$\eta_k\eta_\ell W_{\ell i}W_{ik}^2 W_{jt}^2$	$\eta_k\eta_\ell^2\eta_{k'}\,\mathbb{E}[(W_{\ell i}^2)W_{ik}^2 W_{jt}^2 W_{ik'}^2 W_{j't'}^2] \leq C\theta_i^3\theta_j\theta_k^2\theta_\ell^3\theta_t\theta_{j'}\theta_{k'}^2\theta_{t'}$	$\Vert\theta\Vert^4\Vert\theta\Vert_3^6\Vert\theta\Vert_1^4$
	$\eta_k\eta_\ell(W_{\ell i}W_{is})(W_{jt}W_{jq})W_{km}$	$\eta_k^2\eta_\ell^2\,\mathbb{E}[(W_{\ell i}^2 W_{is}^2 W_{jt}^2 W_{jq}^2 W_{km}^2)] \leq C\theta_i^2\theta_j^2\theta_k^3\theta_\ell^3\theta_s\theta_t\theta_q\theta_m$	$\Vert\theta\Vert^4\Vert\theta\Vert_3^6\Vert\theta\Vert_1^4$
	$\eta_k\eta_\ell W_{\ell i}W_{ij}^2 W_{jq}W_{km}$	$\eta_k^2\eta_\ell^2\,\mathbb{E}[(W_{\ell i}^2 W_{jq}^2 W_{km}^2)W_{ij}^4] \leq C\theta_i^2\theta_j^2\theta_k^3\theta_\ell^3\theta_q\theta_m$	$\Vert\theta\Vert^4\Vert\theta\Vert_3^6\Vert\theta\Vert_1^2$
	$\eta_k\eta_\ell W_{\ell i}W_{is}W_{jk}^2 W_{jq}$	$\eta_k\eta_\ell^2\eta_{k'}\,\mathbb{E}[(W_{\ell i}^2 W_{is}^2 W_{jq}^2)W_{jk}^2 W_{jk'}^2] \leq C\theta_i^2\theta_j^3\theta_k^2\theta_\ell^3\theta_s\theta_q\theta_{k'}^2$	$\Vert\theta\Vert^6\Vert\theta\Vert_3^6\Vert\theta\Vert_1^2$
	$\eta_k\eta_\ell W_{\ell i}W_{ik}^2 W_{jt}W_{jq}$	$\eta_k\eta_\ell^2\eta_{k'}\,\mathbb{E}[(W_{\ell i}^2 W_{jt}^2 W_{jq}^2)W_{ik}^2 W_{ik'}^2] \leq C\theta_i^3\theta_j^2\theta_k^2\theta_\ell^3\theta_t\theta_q\theta_{k'}^2$	$\Vert\theta\Vert^6\Vert\theta\Vert_3^6\Vert\theta\Vert_1^2$
	$\Vert\theta\Vert^2 W_{\ell i}^3 W_{jt}^2$	$\Vert\theta\Vert^4\,\mathbb{E}[W_{\ell i}^3 W_{jt}^2 W_{\ell' i'}^3 W_{j't'}^2] \leq C\Vert\theta\Vert^4\theta_i\theta_j\theta_\ell\theta_t\theta_{i'}\theta_{j'}\theta_{\ell'}\theta_{t'}$	$\Vert\theta\Vert^4\Vert\theta\Vert_1^8$
	$\Vert\theta\Vert^2 W_{\ell i}^3(W_{jt}W_{jq})$	$\Vert\theta\Vert^4\,\mathbb{E}[(W_{jt}^2 W_{jq}^2)W_{\ell i}^3 W_{\ell' i'}^3] \leq C\Vert\theta\Vert^4\theta_i\theta_j^2\theta_\ell\theta_t\theta_q\theta_{i'}\theta_{\ell'}$	$\Vert\theta\Vert^6\Vert\theta\Vert_1^6$
	$\Vert\theta\Vert^2(W_{\ell i}^2 W_{is})W_{jt}^2$	$\Vert\theta\Vert^4\,\mathbb{E}[(W_{is}^2)W_{\ell i}^2 W_{jt}^2 W_{\ell' i}^2 W_{j't'}^2] \leq C\Vert\theta\Vert^4\theta_i^3\theta_j\theta_\ell\theta_s\theta_t\theta_{j'}\theta_{\ell'}\theta_{t'}$	$\Vert\theta\Vert^4\Vert\theta\Vert_3^3\Vert\theta\Vert_1^7$
	$\Vert\theta\Vert^2 W_{\ell i}^2 W_{ij}^3$	$\Vert\theta\Vert^4\,\mathbb{E}[W_{\ell i}^2 W_{ij}^3 W_{\ell' i'}^2 W_{i'j'}^3] \leq C\Vert\theta\Vert^4\theta_i^2\theta_j\theta_\ell\theta_{i'}^2\theta_{j'}\theta_{\ell'}$	$\Vert\theta\Vert^8\Vert\theta\Vert_1^4$
J_7	$\Vert\theta\Vert^2(W_{\ell i}^2 W_{is})(W_{jt}W_{jq})$	$\Vert\theta\Vert^4\,\mathbb{E}[(W_{is}^2 W_{jt}^2 W_{jq}^2)W_{\ell i}^2 W_{\ell' i}^2] \leq C\Vert\theta\Vert^4\theta_i^3\theta_j^2\theta_\ell\theta_s\theta_t\theta_q\theta_{\ell'}$	$\Vert\theta\Vert^6\Vert\theta\Vert_3^3\Vert\theta\Vert_1^5$
	$\Vert\theta\Vert^2 W_{\ell i}^2 W_{ij}^2 W_{jq}$	$\Vert\theta\Vert^4\,\mathbb{E}[(W_{jq}^2)W_{\ell i}^2 W_{ij}^2 W_{\ell' i'}^2 W_{i'j}^2] \leq C\Vert\theta\Vert^4\theta_i^2\theta_j^3\theta_\ell\theta_q\theta_{i'}^2\theta_{\ell'}$	$\Vert\theta\Vert^8\Vert\theta\Vert_3^3\Vert\theta\Vert_1^3$
	$\Vert\theta\Vert^2(W_{\ell i}W_{is}W_{\ell m})W_{jt}^2$	$\Vert\theta\Vert^4\,\mathbb{E}[(W_{\ell i}^2 W_{is}^2 W_{\ell m}^2)W_{jt}^2 W_{j't'}^2] \leq C\Vert\theta\Vert^4\theta_i^2\theta_j\theta_\ell^2\theta_s\theta_t\theta_m\theta_{j'}\theta_{t'}$	$\Vert\theta\Vert^8\Vert\theta\Vert_1^6$
	$\Vert\theta\Vert^2 W_{\ell i}W_{ij}^3 W_{\ell m}$	$\Vert\theta\Vert^4\,\mathbb{E}[(W_{\ell i}^2 W_{\ell m}^2)W_{ij}^3 W_{ij'}^3] \leq C\Vert\theta\Vert^4\theta_i^3\theta_j\theta_\ell^2\theta_m\theta_{j'}$	$\Vert\theta\Vert^6\Vert\theta\Vert_3^3\Vert\theta\Vert_1^3$
	$\Vert\theta\Vert^2(W_{\ell i}W_{is}W_{\ell m})(W_{jt}W_{jq})$	$\Vert\theta\Vert^4\,\mathbb{E}[(W_{\ell i}^2 W_{is}^2 W_{\ell m}^2 W_{jt}^2 W_{jq}^2)] \leq C\Vert\theta\Vert^4\theta_i^2\theta_j^2\theta_\ell^2\theta_s\theta_t\theta_q\theta_m$	$\Vert\theta\Vert^{10}\Vert\theta\Vert_1^4$
	$\Vert\theta\Vert^2 W_{\ell i}W_{ij}^2 W_{\ell m}W_{jq}$	$\Vert\theta\Vert^4\,\mathbb{E}[(W_{\ell i}^2 W_{\ell m}^2 W_{jq}^2)W_{ij}^4] \leq C\Vert\theta\Vert^4\theta_i^2\theta_j^2\theta_\ell^2\theta_q\theta_m$	$\Vert\theta\Vert^{10}\Vert\theta\Vert_1^2$
	$\Vert\theta\Vert^2 W_{\ell i}W_{ij}^2 W_{\ell j}^2$	$\Vert\theta\Vert^4\,\mathbb{E}[(W_{\ell i}^2)W_{ij}^2 W_{\ell j}^2 W_{ij'}^2 W_{\ell j'}^2] \leq C\Vert\theta\Vert^4\theta_i^3\theta_j^2\theta_\ell^3\theta_{j'}^2$	$\Vert\theta\Vert^8\Vert\theta\Vert_3^6$
	$\eta_k\eta_\ell W_{ij}^4 W_{k\ell}$	$\eta_k^2\eta_\ell^2\,\mathbb{E}[(W_{k\ell}^2)W_{ij}^4 W_{i'j'}^4] \leq C\theta_i\theta_j\theta_k^3\theta_\ell^3\theta_{i'}\theta_{j'}$	$\Vert\theta\Vert_3^6\Vert\theta\Vert_1^4$
J_8	$\eta_k\eta_\ell(W_{ij}^3 W_{is})W_{k\ell}$	$\eta_k^2\eta_\ell^2\,\mathbb{E}[(W_{is}^2 W_{k\ell}^2)W_{ij}^3 W_{ij'}^3] \leq C\theta_i^3\theta_j\theta_k^3\theta_\ell^3\theta_s\theta_{j'}$	$\Vert\theta\Vert_3^9\Vert\theta\Vert_1^3$
	$\eta_k\eta_\ell(W_{ij}^2 W_{is}W_{jq})W_{k\ell}$	$\eta_k^2\eta_\ell^2\,\mathbb{E}[(W_{is}^2 W_{jq}^2 W_{k\ell}^2)W_{ij}^4] \leq C\theta_i^2\theta_j^2\theta_k^3\theta_\ell^3\theta_s\theta_q$	$\Vert\theta\Vert^4\Vert\theta\Vert_3^6\Vert\theta\Vert_1^2$
	$\eta_k\eta_\ell(W_{is}W_{it}W_{jq}W_{jm})W_{k\ell}$	$\eta_k^2\eta_\ell^2\,\mathbb{E}[(W_{is}^2 W_{it}^2 W_{jq}^2 W_{jm}^2 W_{k\ell}^2)] \leq C\theta_i^2\theta_j^2\theta_k^3\theta_\ell^3\theta_s\theta_t\theta_q\theta_m$	$\Vert\theta\Vert^4\Vert\theta\Vert_3^6\Vert\theta\Vert_1^4$
	$\eta_k\eta_\ell W_{is}^2 W_{jq}W_{jm}W_{k\ell}$	$\eta_k^2\eta_\ell^2\mathbb{E}[(W_{jq}^2 W_{jm}^2 W_{k\ell}^2)W_{is}^2 W_{i's'}^2] \leq C\theta_i\theta_j^2\theta_k^3\theta_\ell^3\theta_s\theta_q\theta_m\theta_{i'}\theta_{s'}$	$\Vert\theta\Vert^2\Vert\theta\Vert_3^6\Vert\theta\Vert_1^6$
	$\eta_k\eta_\ell W_{is}^2 W_{jq}^2 W_{k\ell}$	$\eta_k^2\eta_\ell^2\mathbb{E}[(W_{k\ell}^2)W_{is}^2 W_{jq}^2 W_{i's'}^2 W_{j'q'}^2] \leq C\theta_i\theta_j\theta_k^3\theta_\ell^3\theta_s\theta_q\theta_{i'}\theta_{j'}\theta_{s'}\theta_{q'}$	$\Vert\theta\Vert_3^6\Vert\theta\Vert_1^8$

where the second last line is from the Cauchy-Schwarz inequality. Since $\Vert\theta\Vert \gg \alpha[\log(n)]^{5/2}$, the right hand side is $o(\Vert\theta\Vert^4)$, which implies that $|J_9|^2 = o(\Vert\theta\Vert^8)$. Similarly, on the event E,

$$|J_{10}| \leq \sum_{i,j,k,\ell} |\eta_\ell \widetilde{\Omega}_{\ell i}||G_i G_j^2 G_k^2|$$

$$\leq C \sum_{i,j,k,\ell} (\alpha\theta_i\theta_\ell^2)\frac{\sqrt{\theta_i\theta_j^2\theta_k^2\Vert\theta\Vert_1^5[\log(n)]^5}}{\sqrt{v^5}}$$

$$
\begin{aligned}
&\le \frac{C\alpha[\log(n)]^{5/2}}{\sqrt{\|\theta\|_1^5}}\Big(\sum_i \theta_i^{3/2}\Big)\Big(\sum_j \theta_j\Big)\Big(\sum_k \theta_k\Big)\Big(\sum_\ell \theta_\ell^2\Big)\\
&\le \frac{C\alpha[\log(n)]^{5/2}}{\sqrt{\|\theta\|_1^5}}\big(\|\theta\|\sqrt{\|\theta\|_1}\big)\|\theta\|_1^2\|\theta\|^2\\
&\le C\alpha[\log(n)]^{5/2}\|\theta\|^3;
\end{aligned}
$$

again, the right hand side is $o(\|\theta\|^4)$. Combining the above gives

$$
\max\big\{\mathbb{E}[J_9^2],\, \mathbb{E}[J_{10}^2]\big\} = o(\|\theta\|^8), \qquad \text{under both hypotheses.}
$$

So far, we have proved: for each R_k with $N_W^* = 5$, it satisfies $\mathbb{E}[R_k^2] = o(\|\theta\|^8)$. This is sufficient to guarantee (266)-(267) for $X = R_k$.

G.4.10.3. Analysis of post-expansion sums with $N_W^* = 6$. There are 7 such terms, including R_{19}-R_{20}, R_{23}-R_{24}, R_{29}-R_{30}, and R_{32}. We plug in the definition of $\tilde r_{ij}$ and δ_{ij} and neglect all factors of $\frac{v}{V}$ (see the explanation in (266)-(267)). It gives ($G_i = \tilde\eta_i - \eta_i$):

$$
\begin{aligned}
R_{19} &= \sum_{i,j,k,\ell(dist)} G_i G_j^2 G_k W_{k\ell} W_{\ell i},\\
R_{20} &= \sum_{i,j,k,\ell(dist)} G_i G_j W_{jk} G_k G_\ell W_{\ell i},\\
R_{23} &= \sum_{i,j,k,\ell(dist)} G_i G_j^2 G_k(\eta_k G_\ell^2 \eta_i + 2G_k\eta_\ell G_\ell \eta_i + G_k \eta_\ell^2 G_i)\\
&= \sum_{i,j,k,\ell(dist)} \eta_i\eta_k G_i G_j^2 G_k G_\ell^2 + 2\sum_{i,j,k,\ell(dist)} \eta_i\eta_\ell G_i G_j^2 G_k^2 G_\ell + \sum_{i,j,k,\ell(dist)} \eta_\ell^2 G_i^2 G_j^2 G_k^2\\
&= 3\sum_{i,j,k,\ell(dist)} \eta_i\eta_k G_i G_j^2 G_k G_\ell^2 + \sum_{i,j,k,\ell(dist)} \eta_\ell^2 G_i^2 G_j^2 G_k^2,\\
R_{24} &= \sum_{i,j,k,\ell(dist)} G_i G_j(\eta_j G_k + G_j\eta_k) G_k G_\ell(\eta_\ell G_i + G_\ell \eta_i)\\
&= 4\sum_{i,j,k,\ell(dist)} \eta_j\eta_\ell G_i^2 G_j G_k^2 G_\ell,\\
R_{29} &= \sum_{i,j,k,\ell(dist)} G_i G_j^2 G_k(\eta_k G_\ell + G_k\eta_\ell) W_{\ell i}\\
&= \sum_{i,j,k,\ell(dist)} \eta_k G_i G_j^2 G_k G_\ell W_{\ell i} + \sum_{i,j,k,\ell(dist)} \eta_\ell G_i G_j^2 G_k^2 W_{\ell i},\\
R_{30} &= 2\sum_{i,j,k,\ell(dist)} G_i G_j(\eta_j G_k) G_k G_\ell W_{\ell i} = 2\sum_{i,j,k,\ell(dist)} \eta_j G_i G_j G_k^2 G_\ell W_{\ell i},\\
R_{32} &= \sum_{i,j,k,\ell(dist)} \widetilde\Omega_{\ell i} G_i G_j^2 G_k^2 G_\ell.
\end{aligned}
$$

Each expression above belongs to one of the following types:

$$
K_1 = \sum_{i,j,k,\ell(dist)} G_i G_j^2 G_k W_{k\ell} W_{\ell i}, \qquad K_2 = \sum_{i,j,k,\ell(dist)} G_i G_j G_k G_\ell W_{jk} W_{\ell i},
$$

$$K_3 = \sum_{i,j,k,\ell(dist)} \eta_k G_i G_j^2 G_k G_\ell W_{\ell i}, \qquad K_4 = \sum_{i,j,k,\ell(dist)} \eta_\ell G_i G_j^2 G_k^2 W_{\ell i},$$

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

$$K_6 = \sum_{i,j,k,\ell(dist)} \eta_\ell^2 G_i^2 G_j^2 G_k^2.$$

Since $|\eta_i \eta_k| \leq C\theta_i\theta_k$ and $|\widetilde{\Omega}_{ik}| \leq C\alpha\theta_i\theta_k$, the study of K_5 and K_5' are similar; we thus omit the analysis of K_5'. We now study K_1-K_6.

Consider K_1. Re-write

$$K_1 = \frac{1}{v^2} \sum_{\substack{i,j,k,\ell(dist)\\ s\neq i, t\neq j, q\neq j, m\neq k}} W_{is}W_{jt}W_{jq}W_{km}W_{k\ell}W_{\ell i}.$$

Note that $W_{km}W_{k\ell}W_{\ell i}W_{is}$ has four different cases: (a) $W_{k\ell}^2 W_{\ell i}^2$, (b) $W_{k\ell}^2 W_{\ell i}W_{is}$, (c) $W_{k\ell}W_{\ell i}W_{ik}^2$, and (d) $W_{k\ell}W_{\ell i}W_{km}W_{is}$. At the same time, $W_{jt}W_{jq}$ has two cases: (i) W_{jk}^2 and (ii) $W_{jt}W_{jq}$. This gives at least $4 \times 2 = 8$ cases. Each case may have sub-cases, e.g., for $(W_{k\ell}^2 W_{\ell i} W_{is})W_{jt}^2$, if $(s,t) = (j,i)$, it becomes $W_{k\ell}^2 W_{\ell i} W_{ij}^3$. By direct calculations, all possible cases of the summand are as follows:

$$(W_{k\ell}^2 W_{\ell i}^2)W_{jt}^2, \quad (W_{k\ell}^2 W_{\ell i}^2)(W_{jt}W_{jq}), \quad (W_{k\ell}^2 W_{\ell i} W_{is})W_{jt}^2,$$
$$W_{k\ell}^2 W_{\ell i} W_{ij}^3, \quad (W_{k\ell}^2 W_{\ell i} W_{is})(W_{jt}W_{jq}), \quad W_{k\ell}^2 W_{\ell i} W_{ij}^2 W_{jq},$$
$$(W_{k\ell}W_{\ell i}W_{ik}^2)W_{jt}^2, \quad (W_{k\ell}W_{\ell i}W_{ik}^2)(W_{jt}W_{jq}),$$
$$(W_{k\ell}W_{\ell i}W_{km}W_{is})W_{jt}^2, \quad W_{k\ell}W_{\ell i}W_{km}W_{ij}^3,$$
$$(W_{k\ell}W_{\ell i}W_{km}W_{is})(W_{jt}W_{jq}), \quad W_{k\ell}W_{\ell i}W_{km}W_{ij}^2 W_{jq},$$
$$W_{k\ell}W_{\ell i}W_{kj}^2 W_{ij}^2. \tag{272}$$

Take the second type for example. We aim to bound $\mathbb{E}[(\sum_{i,j,k,\ell,t,q} W_{k\ell}^2 W_{\ell i}^2 W_{jt}W_{jq})^2]$, which is equal to

$$\sum_{\substack{i,j,k,\ell,t,q\\ i',j',k',\ell',t',q'}} \mathbb{E}[W_{k\ell}^2 W_{\ell i}^2 W_{jt}W_{jq}W_{k'\ell'}^2 W_{\ell' i'}^2 W_{j't'}W_{j'q'}].$$

For the expectation to be nonzero, each single W term has to be paired with another term. The main contribution comes from the case that $W_{j't'}W_{j'q'} = W_{jt}W_{jq}$. It implies $(j',t',q') = (j,t,q)$ or $(j',t',q') = (j,q,t)$. Then, the expression above becomes

$$\sum_{\substack{i,j,k,\ell,t,q\\ i',k',\ell'}} \mathbb{E}[(W_{jt}^2 W_{jq}^2)W_{k\ell}^2 W_{\ell i}^2 W_{k'\ell'}^2 W_{\ell' i'}^2] \leq C \sum_{\substack{i,j,k,\ell,t,q\\ i',k',\ell'}} \theta_i\theta_j^2\theta_k\theta_\ell^2\theta_t\theta_q\theta_{i'}\theta_{k'}\theta_{\ell'}^2$$
$$\leq C\|\theta\|^6\|\theta\|_1^6.$$

There are a total of 9 indices in this sum, which are $(i,j,k,\ell,t,q,i',k',\ell')$. Similarly, for each type of summand, when we bound the expectation of the square of its sum, we count how many indices appear in the ultimate sum. This number equals to twice of the total number of indices appearing in the summand, minus the total number of indices appearing in single W terms. For the above example, all indices appearing in the summand are (i,j,k,ℓ,t,q), while indices appearing in single W terms are (j,t,q); so, the aforementioned number is $2\times 6-3=9$. If this number if m_0, then the expectation of the square of sum of this type is bounded by $C\|\theta\|_1^{m_0}$. We note that K_1 has a factor $\frac{1}{v^2}$ in front of the sum, which brings in a factor of $\frac{C}{\|\theta\|_1^8}$ in the bound. Therefore, for any type of summand with $m_0\leq 8$, the expectation of the square of its sum is $O(1)$, which is $o(\|\theta\|^8)$. As a result, among the types in (272), we only need to consider those with $m_0\geq 9$. We are left with

$$(W_{k\ell}^2W_{\ell i}^2)W_{jt}^2, \qquad (W_{k\ell}^2W_{\ell i}^2)(W_{jt}W_{jq}), \qquad (W_{k\ell}^2W_{\ell i}W_{is})W_{jt}^2.$$

We have proved that the expectation of the square of sum of the second type of summands is bounded by $C\|\theta\|^2\|\theta\|_1^6=o(\|\theta\|^8\|\theta\|_1^8)$. For the other two types, by direct calculations,

$$\begin{aligned}
\mathbb{E}\Big[\Big(\sum_{\substack{i,j,k,\ell(dist)\\ t\neq j}} W_{k\ell}^2W_{\ell i}^2W_{jt}^2\Big)^2\Big] &\leq \sum_{\substack{i,j,k,\ell,t\\ i',j',k',\ell',t'}} \mathbb{E}[W_{k\ell}^2W_{\ell i}^2W_{jt}^2W_{k'\ell'}^2W_{\ell' i'}^2W_{j't'}^2]\\
&\leq \sum_{\substack{i,j,k,\ell,t\\ i',j',k',\ell',t'}} \theta_i\theta_j\theta_k\theta_\ell^2\theta_t\theta_{i'}\theta_{j'}\theta_{k'}\theta_{\ell'}^2\theta_{t'}\\
&\leq C\|\theta\|^4\|\theta\|_1^8=o(\|\theta\|^8\|\theta\|_1^8),\\
\mathbb{E}\Big[\Big(\sum_{\substack{i,j,k,\ell(dist)\\ s\notin\{i,\ell\},t\neq j,\\ (s,t)\neq(j,i)}} W_{k\ell}^2W_{\ell i}W_{is}W_{jt}^2\Big)^2\Big] &\leq \sum_{\substack{i,j,k,\ell,s,t\\ j',k',t'}} \mathbb{E}[(W_{\ell i}^2W_{is}^2)W_{k\ell}^2W_{jt}^2W_{k'\ell}^2W_{j't'}^2]\\
&\leq C\sum_{\substack{i,j,k,\ell,s,t\\ j',k',t'}} \theta_i^2\theta_j\theta_k\theta_\ell^3\theta_s\theta_t\theta_{j'}\theta_{k'}\theta_{t'}\\
&\leq C\|\theta\|^2\|\theta\|_3^3\|\theta\|_1^7=o(\|\theta\|^8\|\theta\|_1^8).
\end{aligned}$$

Combining the above gives

$$\mathbb{E}[K_1^2]=o(\|\theta\|^8), \qquad \text{under both hypotheses.}$$

Consider K_2. Re-write

$$K_2=\frac{1}{v^2}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq i,t\neq j,q\neq k,m\neq\ell}} W_{is}W_{jt}W_{kq}W_{\ell m}W_{jk}W_{\ell i}.$$

Note that $W_{qk}W_{kj}W_{jt}$ has three cases: (a) W_{kj}^3, (b) $W_{kj}^2W_{jt}$ (or $W_{qk}W_{kj}^2$), and (c) $W_{qk}W_{kj}W_{jt}$. Simiarly, $W_{m\ell}W_{\ell i}W_{is}$ has three cases: (a) $W_{\ell i}^3$, (b) $W_{\ell i}^2W_{is}$ (or $W_{m\ell}W_{\ell i}^2$), and (c) $W_{m\ell}W_{\ell i}W_{is}$. By index symmetry, this gives $3+2+1=6$ different cases. Some case may have sub-cases, due to that (s,t) may equal to (j,i), say. By direct calculations, all possible cases of the summand are as follows:

$$\begin{aligned}
&W_{kj}^3W_{\ell i}^3, \quad W_{kj}^3(W_{\ell i}^2W_{is}), \quad W_{kj}^3(W_{m\ell}W_{\ell i}W_{is}), \quad (W_{kj}^2W_{jt})(W_{\ell i}^2W_{is}),\\
&W_{kj}^2W_{ji}^2W_{\ell i}^2, \quad (W_{kj}^2W_{jt})(W_{m\ell}W_{\ell i}W_{is}), \quad W_{kj}^2W_{ji}^2W_{m\ell}W_{\ell i},\\
&(W_{qk}W_{kj}W_{jt})(W_{m\ell}W_{\ell i}W_{is}), \quad W_{qk}W_{kj}W_{ji}^2W_{m\ell}W_{\ell i}, \quad W_{kj}W_{ji}^2W_{k\ell}^2W_{\ell i}.
\end{aligned}$$

As in the analysis of (272), we count the effective number of indices, m_0, which equals to twice of the total number of indices appearing in the summand minus the total number of indices appearing in all single-W terms. For the above types of summand, m_0 equals to

$8, 8, 8, 8, 8, 8, 7, 8, 6, 4$, respectively. None is larger than 8. We conclude that the expectation of the square of sum of each type of summand is bounded by $C\|\theta\|_1^8$. We immediately have

$$\mathbb{E}[K_2^2] = \frac{1}{v^4} \cdot C\|\theta\|_1^8 = O(1) = o(\|\theta\|^8), \qquad \text{under both hypotheses.}$$

Consider K_3. Re-write

$$K_3 = \frac{1}{v^2\sqrt{v}} \sum_{\substack{i,j,k,\ell (dist) \\ s\neq i, t\neq j, q\neq j, m\neq k, p\neq \ell}} \eta_k W_{is} W_{jt} W_{jq} W_{km} W_{\ell p} W_{\ell i}$$

Note that $W_{jt}W_{jq}W_{km}$ has four cases: (a) W_{jk}^3, (b) $W_{jk}^2 W_{jt}$ (or $W_{jk}^2 W_{jq}$), (c) $W_{jt}^2 W_{km}$, and (d) $W_{jt}W_{jq}W_{km}$. At the same time, $W_{is}W_{\ell p}W_{\ell i}$ has three cases: (a) $W_{\ell i}^3$, (b) $W_{\ell i}^2 W_{is}$ (or $W_{\ell i}^2 W_{\ell p}$), and (c) $W_{\ell i}W_{is}W_{\ell p}$. This gives $4 \times 3 = 12$ different cases. Each case may have sub-cases. For example, in the case of $\eta_k(W_{jk}^2 W_{jt})(W_{\ell i}^2 W_{is})$, if $(s,t) = (j,i)$, it becomes $\eta_k W_{jk}^2 W_{ji}^2 W_{\ell i}^2$. By direct calculations, we obtain all possible cases of summands as follows:

$$\begin{aligned}
&\eta_k W_{jk}^3 W_{\ell i}^3, \quad \eta_k W_{jk}^3 (W_{\ell i}^2 W_{is}), \quad \eta_k W_{jk}^3 (W_{\ell i} W_{is} W_{\ell p}), \quad \eta_k (W_{jk}^2 W_{jt}) W_{\ell i}^3, \\
&\eta_k (W_{jk}^2 W_{jt})(W_{\ell i}^2 W_{is}), \quad \eta_k W_{jk}^2 W_{ji}^2 W_{\ell i}^2, \quad \eta_k (W_{jk}^2 W_{jt})(W_{\ell i} W_{is} W_{\ell p}), \\
&\eta_k W_{jk}^2 W_{ji}^2 W_{\ell i} W_{\ell p}, \quad \eta_k (W_{jt}^2 W_{km}) W_{\ell i}^3, \quad \eta_k (W_{jt}^2 W_{km})(W_{\ell i}^2 W_{is}), \quad \eta_k W_{jt}^2 W_{ki}^2 W_{\ell i}^2, \\
&\eta_k (W_{jt}^2 W_{km})(W_{\ell i} W_{is} W_{\ell p}), \quad \eta_k W_{jt}^2 W_{ki}^2 W_{\ell i} W_{\ell p}, \quad \eta_k (W_{jt} W_{jq} W_{km}) W_{\ell i}^3, \\
&\eta_k (W_{jt} W_{jq} W_{km})(W_{\ell i}^2 W_{is}), \quad \eta_k W_{jt} W_{ji}^2 W_{km} W_{\ell i}^2, \quad \eta_k W_{jt} W_{jq} W_{ki}^2 W_{\ell i}^2, \\
&\eta_k (W_{jt} W_{jq} W_{km})(W_{\ell i} W_{is} W_{\ell p}), \quad \eta_k W_{jt} W_{ji}^2 W_{km} W_{\ell i} W_{\ell p}, \quad \eta_k W_{jt} W_{jq} W_{ki}^2 W_{\ell i} W_{\ell p}.
\end{aligned}$$

Same as before, let m_0 be the effective number of indices for each type of summand, which equals to twice of number of distinct indices appearing in the summand minus the number of distinct indices appearing in single-W terms (see (272) and text therein). By direct calculations, $m_0 \leq 10$ for all types above. It follows that, for each type of summand, the expectation of the square of their sums is bounded by

$$\frac{1}{(v\sqrt{v})^2} \cdot C\|\theta\|_1^{m_0} \leq C\|\theta\|_1^{m_0 - 10} = O(1) = o(\|\theta\|^8).$$

We immediately have

$$\mathbb{E}[K_3^2] = o(\|\theta\|^8), \qquad \text{under both hypotheses.}$$

Consider K_4. Re-write

$$K_4 = \frac{1}{v^2\sqrt{v}} \sum_{\substack{i,j,k,\ell (dist) \\ s,t,q,m,p}} \eta_\ell W_{is} W_{jt} W_{jq} W_{km} W_{kp} W_{\ell i}.$$

Note that $W_{is}W_{\ell i}$ has two cases: (a) $W_{\ell i}^2$ and (b) $W_{\ell i}W_{is}$. Moreover, there are a total of six cases for $W_{jt}W_{jq}W_{km}W_{kp}$: (a) W_{jk}^4, (b) $W_{jk}^3 W_{jt}$, (c) $W_{jk}^2 W_{jt} W_{km}$, (d) $W_{jt}^2 W_{km}^2$, (e) $W_{jt}W_{jq}W_{km}^2$, and (f) $W_{jt}W_{jq}W_{km}W_{kp}$. It gives $2 \times 6 = 12$ different cases. Each case may have some sub-cases. It turns out all different types of summand are as follows:

$$\begin{aligned}
&\eta_\ell W_{\ell i}^2 W_{jk}^4, \quad \eta_\ell W_{\ell i}^2 (W_{jk}^3 W_{jt}), \quad \eta_\ell W_{\ell i}^2 (W_{jk}^2 W_{jt} W_{km}), \quad \eta_\ell W_{\ell i}^2 (W_{jt}^2 W_{km}^2), \\
&\eta_\ell W_{\ell i}^2 (W_{jt} W_{jq} W_{km}^2), \quad \eta_\ell W_{\ell i}^2 (W_{jt} W_{jq} W_{km} W_{kp}), \quad \eta_\ell (W_{\ell i} W_{is}) W_{jk}^4, \\
&\eta_\ell (W_{\ell i} W_{is})(W_{jk}^3 W_{jt}), \quad \eta_\ell W_{\ell i} W_{jk}^3 W_{ji}^2, \quad \eta_\ell (W_{\ell i} W_{is})(W_{jk}^2 W_{jt} W_{km}),
\end{aligned}$$

$$\eta_\ell W_{\ell i} W_{jk}^2 W_{ji}^2 W_{km}, \quad \eta_\ell (W_{\ell i} W_{is})(W_{jt}^2 W_{km}^2), \quad \eta_\ell W_{\ell i} W_{ij}^3 W_{km}^2,$$
$$\eta_\ell (W_{\ell i} W_{is})(W_{jt} W_{jq} W_{km}^2), \quad \eta_\ell W_{\ell i} W_{ij}^2 W_{jq} W_{km}^2, \quad \eta_\ell W_{\ell i} W_{jt} W_{jq} W_{ki}^3,$$
$$\eta_\ell (W_{\ell i} W_{is})(W_{jt} W_{jq} W_{km} W_{kp}), \quad \eta_\ell W_{\ell i} W_{ij}^2 W_{jq} W_{km} W_{kp}.$$

Same as before, for each type, let m_0 be the effective number of indices. It suffices to focus on cases where $m_0 \geq 11$. We are left with

$$\eta_\ell W_{\ell i}^2 (W_{jt}^2 W_{km}^2), \qquad \eta_\ell W_{\ell i}^2 (W_{jt} W_{jq} W_{km}^2), \qquad \eta_\ell (W_{\ell i} W_{is})(W_{jt}^2 W_{km}^2).$$

By direct calculations,

$$\begin{aligned}
\mathbb{E}\Big[\Big(\sum_{\substack{i,j,k,\ell (dist)\\ t\neq j, m\neq k}} \eta_\ell W_{\ell i}^2 W_{jt}^2 W_{km}^2\Big)\Big] &\leq \sum_{\substack{i,j,k,\ell,t,m\\ i',j',k',\ell',t',m'}} \eta_\ell \eta_{\ell'}\, \mathbb{E}[W_{\ell i}^2 W_{jt}^2 W_{km}^2 W_{\ell' i'}^2 W_{j't'}^2 W_{k'm'}^2] \\
&\leq C \sum_{\substack{i,j,k,\ell,t,m\\ i',j',k',\ell',t',m'}} \theta_i \theta_j \theta_k \theta_\ell^2 \theta_t \theta_m \theta_{i'} \theta_{j'} \theta_{k'} \theta_{\ell'}^2 \theta_{t'} \theta_{m'} \\
&\leq C\|\theta\|^4 \|\theta\|_1^{10} = o(\|\theta\|^8 \|\theta\|_1^{10}),
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\Big[\Big(\sum_{\substack{i,j,k,\ell (dist)\\ t\neq j, q\neq j, m\neq k\\ t\neq q}} \eta_\ell W_{\ell i}^2 W_{jt} W_{jq} W_{km}^2\Big)\Big] &\leq \sum_{\substack{i,j,k,\ell,t,q,m\\ i',k',\ell',m'}} \eta_\ell \eta_{\ell'}\, \mathbb{E}[(W_{jt}^2 W_{jq}^2) W_{\ell i}^2 W_{km}^2 W_{\ell' i'}^2 W_{k'm'}^2] \\
&\leq C \sum_{\substack{i,j,k,\ell,t,q,m\\ i',k',\ell',m'}} \theta_i \theta_j^2 \theta_k \theta_\ell^2 \theta_t \theta_q \theta_m \theta_{i'} \theta_{k'} \theta_{\ell'}^2 \theta_{m'} \\
&\leq C\|\theta\|^6 \|\theta\|_1^{8} = o(\|\theta\|^8 \|\theta\|_1^{10}),
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\Big[\Big(\sum_{\substack{i,j,k,\ell (dist)\\ s\neq i, t\neq j, m\neq k\\ (s,t)\neq (j,i), (s,m)\neq (k,i)}} \eta_\ell W_{\ell i} W_{is} W_{jt}^2 W_{km}^2\Big)\Big] &\leq C \sum_{\substack{i,j,k,\ell,s,t,m\\ j',k',t',m'}} \eta_\ell^2\, \mathbb{E}[(W_{\ell i}^2 W_{is}^2) W_{jt}^2 W_{km}^2 W_{j't'}^2 W_{k'm'}^2] \\
&\leq C \sum_{\substack{i,j,k,\ell,s,t,m\\ j',k',t',m'}} \theta_i^2 \theta_j \theta_k \theta_\ell^3 \theta_s \theta_t \theta_m \theta_{j'} \theta_{k'} \theta_{t'} \theta_{m'} \\
&\leq C\|\theta\|^2 \|\theta\|_3^3 \|\theta\|_1^{9} = o(\|\theta\|^8 \|\theta\|_1^{10}).
\end{aligned}$$

It follows that

$$\mathbb{E}[K_4^2] \leq \frac{1}{(v^2\sqrt{v})^2} \cdot o(\|\theta\|^8 \|\theta\|_1^{10}) = o(\|\theta\|^8), \qquad \text{under both hypotheses.}$$

Consider K_5-K_6. To save space, we only present the proof for the case of $\|\theta\| \gg [\log(n)]^{3/2}$. When $1 \ll \|\theta\| \leq C[\log(n)]^{3/2}$, we can bound $\mathbb{E}[K_5^2]$ and $\mathbb{E}[K_6^2]$ in the same way as in the study of J_1-J_8, so the proof is omitted. Let E be the event defined in (271). We have argued that it suffices to focus on the event E. On this event, $|G_i| \leq C\sqrt{\theta_i \|\theta\|_1 \log(n)/v}$. It follows that

$$|K_5| \leq C \sum_{i,j,k,\ell} (\theta_i \theta_k) \frac{\sqrt{\theta_i \theta_j^2 \theta_k \theta_\ell^2}\, \|\theta\|_1^3 [\log(n)]^3}{v^3}$$

$$
\begin{aligned}
&\le \frac{C[\log(n)]^3}{\|\theta\|_1^3}\Big(\sum_i \theta_i^{3/2}\Big)\Big(\sum_j \theta_j\Big)\Big(\sum_k \theta_k^{3/2}\Big)\Big(\sum_\ell \theta_\ell\Big)\\
&\le \frac{C[\log(n)]^3}{\|\theta\|_1^3}\big(\|\theta\|\sqrt{\|\theta\|_1}\big)^2\|\theta\|_1^2\\
&\le C[\log(n)]^3\|\theta\|^2,
\end{aligned}
$$

where we have used the Cauchy-Schwarz inequality $(\sum_i \theta_i^{3/2}) \le \|\theta\|\sqrt{\|\theta\|_1}$. Similarly,

$$
\begin{aligned}
|K_6| &\le C\sum_{i,j,k,\ell} \theta_\ell^2 \cdot \frac{\theta_i\theta_j\theta_k\|\theta\|_1^3[\log(n)]^3}{v^3}\\
&\le \frac{C[\log(n)]^3}{\|\theta\|_1^3}\sum_{i,j,k,\ell}\theta_i\theta_j\theta_k\theta_\ell^2\\
&\le C[\log(n)]^3\|\theta\|^2.
\end{aligned}
$$

When $\|\theta\| \gg [\log(n)]^{3/2}$, both right hand sides are $o(\|\theta\|^4)$. We immediately have

$$
\max\big\{\mathbb{E}[K_5^2],\, \mathbb{E}[K_6^2]\big\} = o(\|\theta\|^8).
$$

We have proved: Each R_k with $N_W^* = 6$ satisfies $\mathbb{E}[R_k^2] = o(\|\theta\|^8)$. This is sufficient to guarantee (266)-(267) for $X = R_k$.

G.4.10.4. Analysis of terms with $N_W^* \ge 7$. There are 3 such terms, R_{31}, R_{33} and R_{34}. Consider R_{31}. By definition,

$$
R_{31} = \sum_{i,j,k,\ell(dist)} G_iG_j^2G_k^2G_\ell W_{\ell i} = \frac{1}{v^3}\sum_{\substack{i,j,k,\ell(dist)\\ s\neq i, t\neq j, q\neq j,\\ m\neq k, p\neq k, y\neq \ell}} W_{is}W_{jt}W_{jq}W_{km}W_{kp}W_{\ell y}W_{\ell i}.
$$

We note that $W_{\ell i}W_{is}W_{\ell y}$ has three cases: (a) $W_{\ell i}^3$, (b) $W_{\ell i}^2W_{is}$, and (c) $W_{\ell i}W_{is}W_{\ell y}$. Moreover, $W_{jt}W_{jq}W_{km}W_{kp}$ has six cases: (a) W_{jk}^4, (b) $W_{jk}^3W_{jt}$, (c) $W_{jk}^2W_{jt}W_{km}$, (d) $W_{jt}^2W_{km}^2$, (e) $W_{jt}W_{jq}W_{km}^2$, and (f) $W_{jt}W_{jq}W_{km}W_{kp}$. This gives $3\times 6 = 18$ different cases. Since each case may have sub-cases, we end up with the following different types:

$$
\begin{aligned}
&W_{\ell i}^3W_{jk}^4, \quad W_{\ell i}^3(W_{jk}^3W_{jt}), \quad W_{\ell i}^3(W_{jk}^2W_{jt}W_{km}), \quad W_{\ell i}^3(W_{jt}^2W_{km}^2),\\
&W_{\ell i}^3(W_{jt}W_{jq}W_{km}^2), \quad W_{\ell i}^3(W_{jt}W_{jq}W_{km}W_{kp}), \quad (W_{\ell i}^2W_{is})W_{jk}^4,\\
&(W_{\ell i}^2W_{is})(W_{jk}^3W_{jt}), \quad W_{\ell i}^2W_{jk}^3W_{ji}^2, \qquad (W_{\ell i}^2W_{is})(W_{jk}^2W_{jt}W_{km}),\\
&W_{\ell i}^2W_{jk}^2W_{ji}^2W_{km}, \quad (W_{\ell i}^2W_{is})(W_{jt}^2W_{km}^2), \quad W_{\ell i}^2W_{ij}^3W_{km}^2,\\
&(W_{\ell i}^2W_{is})(W_{jt}W_{jq}W_{km}^2), \quad W_{\ell i}^2W_{ij}^2W_{jq}W_{km}^2, \quad W_{\ell i}^2W_{jt}W_{jq}W_{ki}^3,\\
&(W_{\ell i}^2W_{is})(W_{jt}W_{jq}W_{km}W_{kp}), \quad W_{\ell i}^2W_{ij}^2W_{jq}W_{km}W_{kp},\\
&(W_{\ell i}W_{is}W_{\ell y})W_{jk}^4, \quad (W_{\ell i}W_{is}W_{\ell y})(W_{jk}^3W_{jt}), \quad W_{\ell i}W_{\ell y}W_{jk}^3W_{ji}^2,\\
&(W_{\ell i}W_{is}W_{\ell y})(W_{jk}^2W_{jt}W_{km}), \quad W_{\ell i}W_{\ell y}W_{jk}^2W_{ji}^2W_{km}, \quad W_{\ell i}W_{jk}^2W_{ji}^2W_{k\ell}^2,\\
&(W_{\ell i}W_{is}W_{\ell y})(W_{jt}^2W_{km}^2), \quad W_{\ell i}W_{\ell y}W_{ji}^3W_{km}^2, \quad W_{\ell i}W_{ji}^3W_{k\ell}^3,\\
&(W_{\ell i}W_{is}W_{\ell y})(W_{jt}W_{jq}W_{km}^2), \quad W_{\ell i}W_{\ell y}W_{ji}^2W_{jq}W_{km}^2, \quad W_{\ell i}W_{\ell y}W_{jt}W_{jq}W_{ki}^3,
\end{aligned}
$$

$$W_{\ell i}W_{ji}^2W_{jq}W_{ki}^3, \quad (W_{\ell i}W_{is}W_{\ell y})(W_{jt}W_{jq}W_{km}W_{kp}),$$
$$W_{\ell i}W_{\ell y}W_{ji}^2W_{jq}W_{km}W_{kp}, \quad W_{\ell i}W_{ji}^2W_{jq}W_{k\ell}^2W_{kp}.$$

For each type, we count m_0, the effective number of indices. It equals to twice of the number of distinct indices in the summand, minus the number of distinct indices appearing in all single-W terms. It turns out that $m_0 \leq 12$ for all types above. By similar arguments as in (272), we conclude that

$$\mathbb{E}[R_{31}^2] \leq \frac{1}{v^6} \cdot C\|\theta\|_1^{m_0} \leq C\|\theta\|_1^{m_0-12} = O(1) = o(\|\theta\|^8), \qquad \text{under both hypotheses.}$$

Consider R_{33}-R_{34}. We only give the proof when $\|\theta\|^6 \gg [\log(n)]^7$, as it is much simpler. In the case of $1 \ll \|\theta\|^6 \leq C[\log(n)]^7$, we can follow similar steps above to obtain desired bounds, where details are omitted. On the event E (see (271) for definition),

$$\begin{aligned}
|R_{33}| &\leq \sum_{i,j,k,\ell} |\eta_\ell||G_i^2G_j^2G_k^2G_\ell| \\
&\leq C\sum_{i,j,k,\ell} \theta_\ell \frac{\sqrt{\theta_i^2\theta_j^2\theta_k^2\theta_\ell\|\theta\|_1^7[\log(n)]^7}}{(\sqrt{v})^7} \\
&\leq \frac{C[\log(n)]^{7/2}}{\sqrt{\|\theta\|_1^7}}\Big(\sum_i \theta_i\Big)\Big(\sum_j \theta_j\Big)\Big(\sum_k \theta_k\Big)\Big(\sum_\ell \theta_\ell^{3/2}\Big) \\
&\leq \frac{C[\log(n)]^{7/2}}{\sqrt{\|\theta\|_1^7}} \cdot \|\theta\|_1^3\big(\|\theta\|\sqrt{\|\theta\|_1}\big) \\
&\leq C[\log(n)]^{7/2}\|\theta\|,
\end{aligned}$$

where we have used the Cauchy-Schwarz inequality $\sum_\ell \theta_\ell^{3/2} \leq \|\theta\|\sqrt{\|\theta\|_1}$ in the second last line. When $\|\theta\|^6 \gg [\log(n)]^7$, the right hand side is $o(\|\theta\|^4)$. Similarly,

$$\begin{aligned}
|R_{34}| &\leq \sum_{i,j,k,\ell} |G_i^2G_j^2G_k^2G_\ell^2| \\
&\leq C\sum_{i,j,k,\ell} \frac{\theta_i\theta_j\theta_k\theta_\ell\|\theta\|_1^4[\log(n)]^4}{v^4} \\
&\leq C[\log(n)]^4.
\end{aligned}$$

When $\|\theta\|^6 \gg [\log(n)]^7$, the right hand side is $o(\|\theta\|^4)$. As we have argued in (271), the event E^c has a negligible effect. It follows that

$$\max\big\{\mathbb{E}[R_{31}^2],\, \mathbb{E}[R_{33}^2],\, \mathbb{E}[R_{34}^2]\big\} = o(\|\theta\|^8), \qquad \text{under both hypotheses.}$$

This is sufficient to guarantee (266)-(267) for R_k.

We have analyzed all 34 terms in Table G.4. The proof is now complete.

G.4.11. *Proof of Lemma G.12*. Consider an arbitrary post-expansion sum of the form

$$\sum_{i,j,k,\ell(dist)} a_{ij}b_{jk}c_{k\ell}d_{\ell i}, \qquad \text{where} \quad a,b,c,d \in \{\widetilde{\Omega}, W, \delta, \tilde{r}, \epsilon\}. \tag{273}$$

Let $(N_{\widetilde{\Omega}}, N_W, N_\delta, N_{\tilde{r}}, N_\epsilon)$ be the number of each type in the product, where these numbers have to satisfy $N_{\widetilde{\Omega}} + N_W + N_\delta + N_{\tilde{r}} + N_\epsilon = 4$. As discussed in Section G.3, $(Q_n - Q_n^*)$ equals to the sum of all post-expansion sums such that $N_\epsilon > 0$. Recall that

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

Define

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

Then, $\epsilon_{ij} = \epsilon_{ij}^{(1)} + \epsilon_{ij}^{(2)} + \epsilon_{ij}^{(3)}$. It follows that each post-expansion sum of the form (273) can be further expanded as the sum of terms like

$$\sum_{i,j,k,\ell(dist)} a_{ij}b_{jk}c_{k\ell}d_{\ell i}, \quad \text{where } a, b, c, d \in \{\widetilde{\Omega}, W, \delta, \tilde{r}, \epsilon^{(1)}, \epsilon^{(2)}, \epsilon^{(3)}\}. \tag{274}$$

Let $(N_{\widetilde{\Omega}}, N_W, N_\delta, N_{\tilde{r}})$ have the same meaning as before, and let $N_\epsilon^{(m)}$ be the number of $\epsilon^{(m)}$ term in the product, for $m \in \{1, 2, 3\}$. These numbers have to satisfy $N_{\widetilde{\Omega}} + N_W + N_\delta + N_{\tilde{r}} + N_\epsilon^{(1)} + N_\epsilon^{(2)} + N_\epsilon^{(3)} = 4$. Now, $(Q_n - Q_n^*)$ equals to the sum of all post-expansion sums of the form (274) with

$$N_\epsilon^{(1)} + N_\epsilon^{(2)} + N_\epsilon^{(3)} \geq 1. \tag{275}$$

Fix such a post-expansion sum and denote it by Y. We shall bound $|\mathbb{E}[Y]|$ and $\mathrm{Var}(Y)$.

We need some preparation. First, we derive a bound for $|\epsilon_{ij}^{(1)}|$. By definition, $\eta_i = (1/\sqrt{v})\sum_{j\neq i}\Omega_{ij}$ and $\eta_i^* = (1/\sqrt{v_0})\sum_j \Omega_{ij}$. It follows that

$$\eta_i^* = \frac{\sqrt{v}}{\sqrt{v_0}}\eta_i + \frac{1}{\sqrt{v_0}}\Omega_{ii}.$$

We then have

$$\eta_i^*\eta_j^* = \frac{v}{v_0}\eta_i\eta_j + \frac{\sqrt{v}}{v_0}(\eta_i\Omega_{jj} + \eta_j\Omega_{ii}) + \frac{1}{v_0}\Omega_{ii}\Omega_{jj}.$$

Note that $v = \sum_{i\neq j}\Omega_{ij}$ and $v_0 = \sum_{ij}\Omega_{ij} \asymp \|\theta\|_1^2$. It follows that $v_0 - v = \sum_i \Omega_{ii} \leq \sum_i \theta_i^2 \leq \|\theta\|^2$. Therefore,

$$\begin{aligned}
|\eta_i^*\eta_j^* - \eta_i\eta_j| &\leq \Big|1 - \frac{v}{v_0}\Big|\eta_i\eta_j + \frac{\sqrt{v}}{v_0}(\eta_i\Omega_{jj} + \eta_j\Omega_{ii}) + \frac{1}{v_0}\Omega_{ii}\Omega_{jj} \\
&\leq \frac{C\|\theta\|^2}{\|\theta\|_1^2}\cdot\theta_i\theta_j + \frac{C}{\|\theta\|_1}(\theta_i\theta_j^2 + \theta_j\theta_i^2) + \frac{C}{\|\theta\|_1^2}\cdot\theta_i^2\theta_j^2 \\
&\leq C\theta_i\theta_j\cdot\Big(\frac{\|\theta\|^2}{\|\theta\|_1^2} + \frac{\theta_i + \theta_j}{\|\theta\|_1} + \frac{\theta_i\theta_j}{\|\theta\|_1^2}\Big).
\end{aligned}$$

Since $\|\theta\|^2 \leq \theta_{\max}\|\theta\|_1$, the term in the brackets is bounded by $C\theta_{\max}/\|\theta\|_1$. We thus have

$$|\epsilon_{ij}^{(1)}| \leq \frac{C\theta_{\max}}{\|\theta\|_1}\cdot\theta_i\theta_j, \qquad \text{for all } 1 \leq i \neq j \leq n. \tag{276}$$

Second, in Lemmas G.1-G.11, we have studied all post-expansion sums of the form

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

where $(N_{\widetilde{\Omega}}, N_W, N_\delta, N_{\tilde{r}})$ are the numbers of each type in the product. We hope to take advantage of these results. Using the proved bounds for $|\mathbb{E}[Z]|$ and $\mathrm{Var}(Z)$, we can get

$$\mathbb{E}[Z^2] \le C(\alpha^2)^{N_{\widetilde{\Omega}}} \cdot f(\theta; N_{\widetilde{\Omega}}, N_W, N_\delta, N_{\tilde{r}}), \tag{277}$$

where $\alpha = |\lambda_2|/\lambda_1$ and $f(\theta; m_1, m_2, m_3, m_4)$ is a function of θ whose form is determined by (m_1, m_2, m_3, m_4). For example,

$$\begin{cases} f(\theta;0,4,0,0) = \|\theta\|^8, & \text{by claims of } X_1 \text{ in Lemmas G.1\&G.3;} \\ f(\theta;4,0,0,0) = \|\theta\|^{16}, & \text{by claims of } X_6 \text{ in Lemma G.3;} \\ f(\theta;3,1,0,0) = \|\theta\|^8\|\theta\|_3^6, & \text{by claims of } X_5 \text{ in Lemma G.3;} \\ f(\theta;1,2,1,0) = \|\theta\|^4\|\theta\|_3^6, & \text{by claims of } Y_2, Y_3 \text{ in Lemma G.5;} \\ f(\theta;1,1,1,1) = \|\theta\|^8, & \text{by claims of } R_9\text{-}R_{11} \text{ in the proof of Lemma G.11.} \end{cases}$$

If there are more than one post-expansion sum that corresponds to the same $(N_{\widetilde{\Omega}}, N_W, N_\delta, N_{\tilde{r}})$, we use the largest bound to define $f(\theta; N_{\widetilde{\Omega}}, N_W, N_\delta, N_{\tilde{r}})$. Thanks to previous lemmas, we have known the function $f(\theta; m_1, m_2, m_3, m_4)$ for all possible (m_1, m_2, m_3, m_4).

We now show the claim. Recall that Y is the post-expansion sum in (274). The key is to prove the following argument: For any sequence x_n such that $\sqrt{\log(\|\theta\|_1)} \ll x_n \ll \|\theta\|_1$,

$$\begin{aligned} \mathbb{E}[Y^2] \le C(\alpha^2)^{N_{\widetilde{\Omega}}} \times \left(\frac{\theta_{\max}^2}{\|\theta\|_1^2}\right)^{N_\epsilon^{(1)}} \times \left(\frac{x_n^2}{\|\theta\|_1^2}\right)^{N_\epsilon^{(2)}+N_\epsilon^{(3)}} \\ \times f(\theta; m_1, m_2, m_3, m_4)\Big|_{\substack{m_1 = N_{\widetilde{\Omega}} + N_\epsilon^{(1)} + N_\epsilon^{(2)},\; m_2 = N_W, \\ m_3 = N_\delta + N_\epsilon^{(3)},\; m_4 = N_{\tilde{r}},}} \end{aligned} \tag{278}$$

where $(N_{\widetilde{\Omega}}, N_W, N_\delta, N_{\tilde{r}}, N_\epsilon^{(1)}, N_\epsilon^{(2)}, N_\epsilon^{(3)})$ are the same as in (274)-(275), and $f(\theta; m_1, m_2, m_3, m_4)$ is the known function in (277).

We prove (278). Let D be the event

$$D = \{|V - v| \le \|\theta\|_1 x_n\}.$$

In Lemma G.10, we have proved $\mathbb{E}[(Q_n - Q_n^*)^2 \cdot I_{D^c}] = o(1)$. By similar proof, we can show: when $|Y|$ is bounded by a polynomial of V and $\|\theta\|_1$ (which is always the case here),

$$\mathbb{E}[Y^2 \cdot I_{D^c}] = o(1).$$

It follows that

$$\mathbb{E}[Y^2] \le \mathbb{E}[Y^2 \cdot I_D] + o(1). \tag{279}$$

We then bound $\mathbb{E}[Y^2 \cdot I_D]$. In the definition of Y, each $\epsilon^{(2)}$ term introduces a factor of $(1 - \frac{v}{V})$, and each $\epsilon^{(3)}$ term introduces a factor of $-(1 - \frac{v}{V})$. We bring all these factors to the front and re-write the post-expansion sum as

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

After the factor $(1 - \frac{v}{V})$ is removed, $\epsilon^{(2)}$ becomes $\eta_i\eta_j$; similarly, $\epsilon^{(3)}$ becomes δ_{ij}. Therefore, in the expression of X,

$$\begin{cases} a_{ij}, b_{ij}, c_{ij}, d_{ij} \in \{\widetilde{\Omega}_{ij}, W_{ij}, \delta_{ij}, \tilde{r}_{ij}, \epsilon_{ij}^{(1)}, \eta_i\eta_j\}, \\ \text{number of } \eta_i\eta_j \text{ in the product is } N_\epsilon^{(2)}, \\ \text{number of } \delta_{ij} \text{ in the product is } N_\delta + N_\epsilon^{(3)}, \\ \text{number of any other term in the product is same as before.} \end{cases} \tag{280}$$

On the event D, $|1-\frac{v}{V}|\leq \frac{x_n\|\theta\|_1}{C\|\theta\|_1^2}=O(\frac{x_n}{\|\theta\|_1})$. Hence,

$$|Y|\leq C\Big(\frac{x_n}{\|\theta\|_1}\Big)^{N_\epsilon^{(2)}+N_\epsilon^{(3)}}|X|, \qquad \text{on the event } D.$$

It follows that

$$\mathbb{E}[Y^2\cdot I_D]\leq C\Big(\frac{x_n^2}{\|\theta\|_1^2}\Big)^{N_\epsilon^{(2)}+N_\epsilon^{(3)}}\cdot \mathbb{E}[X^2]. \tag{281}$$

To bound $\mathbb{E}[X^2]$, we compare X and Z. In obtaining (277), the only property of $\widetilde{\Omega}$ we have used is

$$|\widetilde{\Omega}_{ij}|\leq \alpha\cdot C\theta_i\theta_j.$$

In comparison, in the expression of X, we have (by (276) and (81))

$$|\widetilde{\Omega}_{ij}|\leq \alpha\cdot C\theta_i\theta_j, \qquad |\epsilon_{ij}^{(1)}|\leq \frac{\theta_{\max}}{\|\theta\|_1}\cdot C\theta_i\theta_j, \qquad |\eta_i\eta_j|\leq C\theta_i\theta_j. \tag{282}$$

If we consider $(\alpha^{N_{\widetilde{\Omega}}}\cdot(\frac{\theta_{\max}}{\|\theta\|_1})^{N_\epsilon^{(1)}}\cdot 1^{N_\epsilon^{(2)}})^{-1}X$ and $(\alpha^{N_{\widetilde{\Omega}}})^{-1}Z$, we can derive the same upper bound for the second moment of both variables, except that the effective N_δ in X should be $N_\delta+N_\epsilon^{(3)}$ and the effective $N_{\widetilde{\Omega}}$ in X should be $N_{\widetilde{\Omega}}+N_\epsilon^{(1)}+N_\epsilon^{(2)}$. It follows that

$$\mathbb{E}[X^2]\leq C(\alpha^2)^{N_{\widetilde{\Omega}}}\times\Big(\frac{\theta_{\max}^2}{\|\theta\|_1^2}\Big)^{N_\epsilon^{(1)}} \times f(\theta;m_1,m_2,m_3,m_4)\Big|_{\substack{m_1=N_{\widetilde{\Omega}}+N_\epsilon^{(1)}+N_\epsilon^{(2)},\ m_2=N_W,\\ m_3=N_\delta+N_\epsilon^{(3)},\ m_4=N_{\tilde{r}}.}} \tag{283}$$

We plug (283) into (281), and then plug it into (279). It gives (278).

Next, we use (278) to prove the claims of this lemma. Under our assumption, we can choose a sequence x_n such that $\sqrt{\log(\|\theta\|_1)}\ll x_n\ll \|\theta\|_1/\|\theta\|^2$. Also, note that $\|\theta\|_1\geq \theta_{\max}^{-1}\|\theta\|^2\gg\|\theta\|^2$. Then,

$$\frac{\theta_{\max}}{\|\theta\|_1}=o(\|\theta\|^{-2}), \qquad \frac{x_n}{\|\theta\|_1}=o(\|\theta\|^{-2}). \tag{284}$$

As a result, since $N_\epsilon^{(1)}+N_\epsilon^{(2)}+N_\epsilon^{(3)}\geq 1$, (278) implies

$$\mathbb{E}[Y^2]=o(\|\theta\|^{-4})\cdot f(\theta;m_1,m_2,m_3,m_4), \tag{285}$$

for $m_1=N_{\widetilde{\Omega}}+N_\epsilon^{(1)}+N_\epsilon^{(2)}$, $m_2=N_W$, $m_3=N_\delta+N_\epsilon^{(3)}$ and $m_4=N_{\tilde{r}}$. We then extract $f(\theta;m_1,m_2,m_3,m_4)$ from previous lemmas. Recall the following facts:

- Under the null hypothesis, for any previously analyzed post-expansion sum Z, $|\mathbb{E}[Z]|\leq C\|\theta\|^4$ and $\mathrm{Var}(Z)\leq C\|\theta\|^8$.
- Under the alternative hypothesis, except $\sum_{i,j,k,\ell(dist)}\widetilde{\Omega}_{ij}\widetilde{\Omega}_{jk}\widetilde{\Omega}_{k\ell}\widetilde{\Omega}_{\ell i}$, for all previously analyzed post-expansion sum Z , $|\mathbb{E}[Z]|\leq C\alpha^2\|\theta\|^6$ and $\mathrm{Var}(Z)\leq C\|\theta\|^8+C\alpha^6\|\theta\|^8\|\theta\|_3^6$.

Therefore, under both hypotheses, except for $(m_1,m_2,m_3,m_4)=(4,0,0,0)$,

$$f(\theta;m_1,m_2,m_3,m_4)\leq C(\|\theta\|^8+\|\theta\|^{12}+\|\theta\|^8\|\theta\|_3^6)\leq C\|\theta\|^{12}. \tag{286}$$

Consider two cases for Y. The first case is $N_{\widetilde{\Omega}}+N_\epsilon^{(1)}+N_\epsilon^{(2)}\neq 4$. Combining (285)-(286) gives

$$\mathbb{E}[Y^2]=o(\|\theta\|^{-4})\cdot C\|\theta\|^{12}=o(\|\theta\|^8).$$

The claims follow immediately. The second case is $N_{\widetilde{\Omega}} + N_\epsilon^{(1)} + N_\epsilon^{(2)} = 4$. In this case,

$$f(\theta; m_1, m_2, m_3, m_4) = f(\theta; 4, 0, 0, 0) = \|\theta\|^{16}.$$

If $N_\epsilon^{(1)} + N_\epsilon^{(2)} \geq 2$, then by (278) and (284),

$$\mathbb{E}[Y^2] = o(\|\theta\|^{-8}) \cdot C\|\theta\|^{16} = o(\|\theta\|^8).$$

The claims follow. It remains to consider $N_\epsilon^{(1)} + N_\epsilon^{(2)} = 1$ (and so $N_{\widetilde{\Omega}} = 3$). Write for short $S = 1 - \frac{v}{V}$. By (280),

$$Y = S^{N_\epsilon^{(2)}} \cdot X, \qquad \text{where} \quad X = \sum_{i,j,k,\ell(dist)} a_{ij} b_{jk} c_{k\ell} d_{\ell i},$$

and $a_{ij}, b_{ij}, c_{ij}, d_{ij}$ can only take values from $\{\widetilde{\Omega}_{ij}, \epsilon_{ij}^{(1)}, \eta_i\eta_j\}$. So, X is a non-stochastic number. Using (282), we can easily show

$$|X| \leq C\alpha^{N_{\widetilde{\Omega}}} \Big(\frac{\theta_{\max}}{\|\theta\|_1}\Big)^{N_\epsilon^{(1)}} \|\theta\|^8.$$

When $(N_\epsilon^{(1)}, N_\epsilon^{(2)}) = (1, 0)$, we have $Y = X$. By (284), $\frac{\theta_{\max}}{\|\theta\|_1} = o(\|\theta\|^{-2})$. It follows that

$$\mathrm{Var}(Y) = 0, \qquad |\mathbb{E}[Y]| = |X| \leq C\alpha^3 \cdot o(\|\theta\|^{-2}) \cdot \|\theta\|^8 = o(\alpha^4\|\theta\|^8).$$

This gives the desired claims. When $(N_\epsilon^{(1)}, N_\epsilon^{(2)}) = (0, 1)$, we have $Y = S \cdot X$. So,

$$|Y| = |X| \cdot |S| \leq C\alpha^3\|\theta\|^8 \cdot |S|.$$

Note that $S = 1 - \frac{v}{V}$, where $v = \mathbb{E}[V]$. Using the tail bound (254), we can prove $\mathbb{E}[S^2] \leq C\|\theta\|_1^{-2}$. Therefore,

$$\mathbb{E}[Y^2] \leq \frac{C\alpha^6\|\theta\|^{16}}{\|\theta\|_1^2} \leq C\alpha^6\|\theta\|^8\|\theta\|_3^6,$$

where the last inequality is due to $\|\theta\|^4 \leq \|\theta\|_1\|\theta\|_3^3$ (Cauchy-Schwarz). The claims follow immediately. □

APPENDIX H: ADDITIONAL SIMULATION RESULTS

In Section 5 of the main article, we investigated the numerical performance of SgnT and SgnQ tests and compare them with the EZ and GC tests. Due to space limit, we only reported the sum of the percent of type I errors and the percent of type II errors. It does not show the contribution of each type of errors. We now report separately the percent of each type of errors.

Figures H.1-H.3 here are supplement to Figures 3-5 of the main article, corresponding to Experiments 1-3, respectively. Below is a brief summary of the settings in three experiments:

- *Experiment 1*. In this experiment, $K = 2$, and the degree parameters are iid generated from a uniform distribution (Experiment 1a), a two-point mass (Experiment 1b), and a Pareto distribution (Experiment 1c), respectively.
- *Experiment 2*. In this experiment, K is larger ($K \in \{5, 10\}$) and P is more complicated, and the community sizes are either balanced (Experiment 2a) or unbalanced (Experiment 2b).
- *Experiment 3*. In this experiment, we allow for mixed memberships, where the percent of mixed nodes is 0% (Experiment 3a), 10% (Experiment 3b), and 25% (Experiment 3c), respectively.

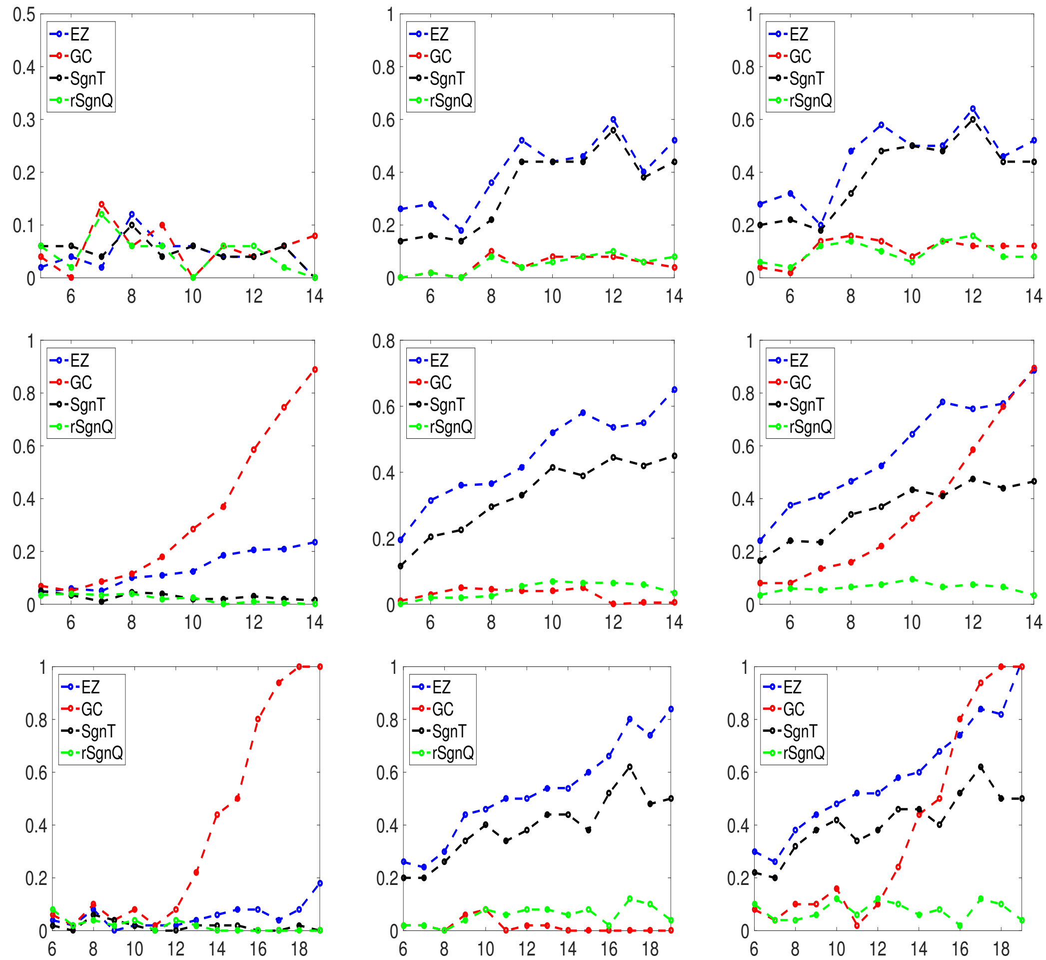

FIG H.1. *Experiment 1 (from top to bottom: Experiment 1a, 1b, and 1c). The x-axis is $\|\theta\|$, and the y-axis is type I error (left), type II error (middle) and the sum (right).*

For each parameter setting, we generate 200 networks under the null hypothesis and 200 networks under the alternative hypothesis, run all the four tests with a target level $\alpha = 5\%$, and record the percent of type I errors, the percent of type II errors, and their sum. In each figure, the plots in the third column are those already shown in the main article.

The results confirm our claims in Section 5. In terms of the type I error, the EZ and GC tests fail to control it at the target level when $\|\theta\|$ is large. It is because the biases of these tests are non-negligible for less sparse networks (the bias of GC is comparably larger). The SgnT and SgnQ tests successfully control the type I error for both sparse and less sparse networks. In terms of the type II error, the order-4 graphlet counting tests have uniformly better power than the order-3 graphlet counting tests. E.g., the type II error of GC is smaller than that of EZ, and the type II error of SgnQ is smaller than that of SgnT.

REFERENCES

[1] BAI, Z. and SILVERSTEIN, J. (2010). *Spectral Analysis of Large Dimensional Random Matrices*. Springer.
[2] HORN, R. and JOHNSON, C. (1985). *Matrix Analysis*. Cambridge University Press.
[3] JIN, J., KE, Z. T. and LUO, S. (2018). Network global testing by counting graphlets. In *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018* 2338–2346.

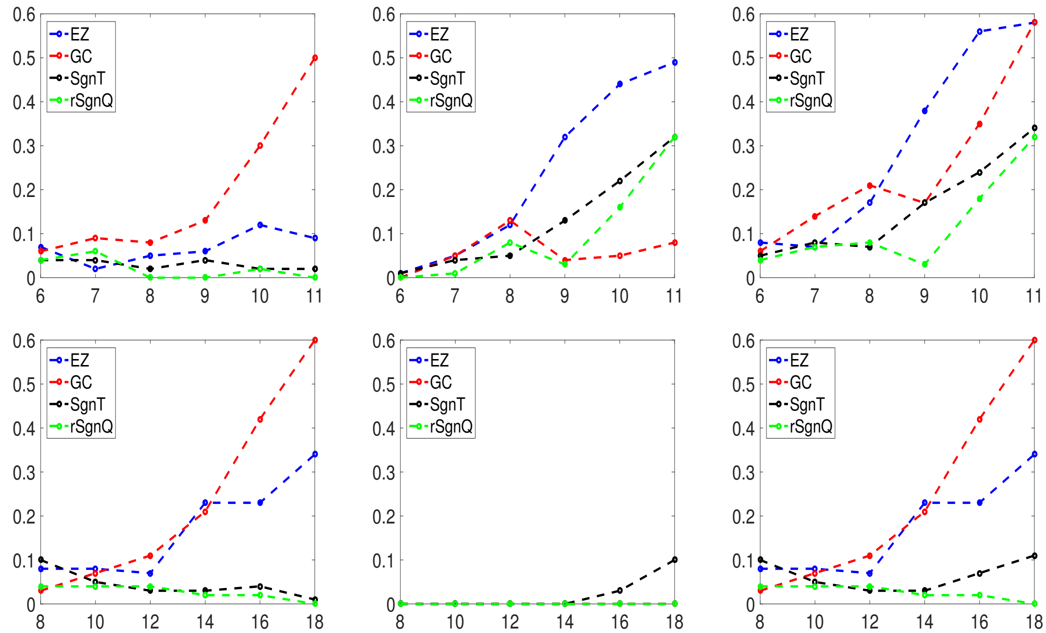

FIG H.2. *Experiment 2 (from top to bottom: Experiment 2a and 2b). The x-axis is $\|\theta\|$, and the y-axis is type I error (left), type II error (middle) and the sum (right).*

[4] JIN, J., KE, Z. T. and LUO, S. (2019). Optimal adaptivity of signed-polygon statistics for network testing. *Manuscript*.

[5] JOHNSON, C. and REAMS, R. (2009). Scaling of symmetric matrices by positive diagonal congruence. *Linear Multilinear A* **57** 123-140.

[6] MARSHALL, A. and OLKIN, I. (1968). Scaling of matrices to achieve specified row and column sums. *Numer. Math.* **12** 83–90.

[7] TROPP, J. (2012). User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.* **12** 389–434.

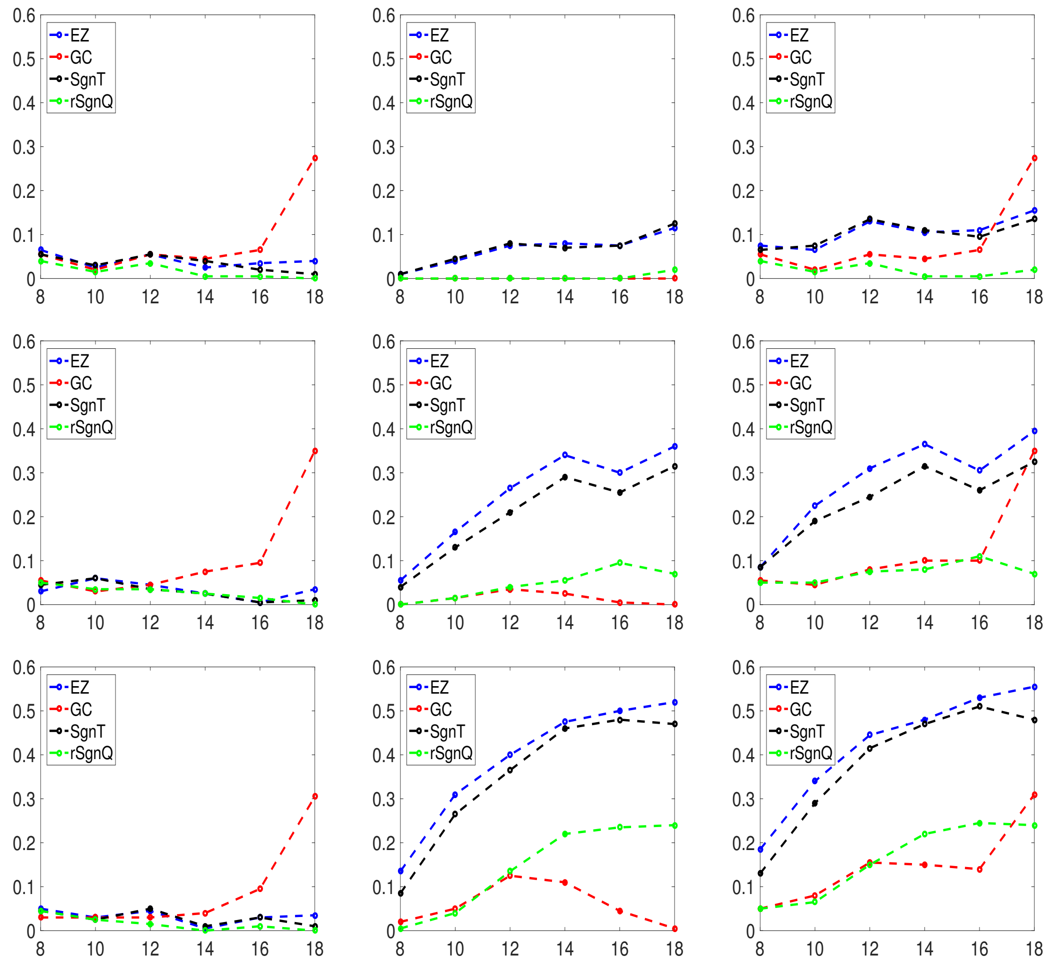

FIG H.3. *Experiment 3 (from top to bottom: Experiment 3a, 3b, and 3c). The x-axis is $\|\theta\|$, and the y-axis is type I error (left), type II error (middle) and the sum (right).*