Supplementary material: Testing High-dimensional Multinomials with Applications to Text Analysis

T. Tony Cai, Zheng Tracy Ke, and Paxton Turner

December 29, 2023

Contents

Notational conventions: We write $A \lesssim B$ (respectively, $A \gtrsim B$) if there exists an absolute constant $C > 0$ such that $A \leq C \cdot B$ (respectively $A \geq C \cdot B$). If both $A \lesssim B$ and $B \lesssim A$, we write $A \asymp B$. The implicit constant C may vary from line to line. For sequences a_t, b_t indexed by an integer $t \in \mathbb{N}$, we write $a_t \ll b_t$ if $b_t/a_t \to \infty$ as $t \to \infty$, and we write $a_t \gg b_t$ if $a_t/b_t \to \infty$ as $t \to \infty$. We also may write $a_t = o(b_t)$ to denote $a_t \ll b_t$. In particular, we write $a_t = (1+o(1))b_t$ if $a_t/b_t \to 1$ as $t \to \infty$. Given a positive integer T, define $[T] = \{1, 2, \ldots, T\}$.

A Additional simulation results

We present some simulation results that are not included in the main paper for space constraint.

A.1 Power diagrams of DELVE+

In Experiment 2 of Section 5, we investigate the power of the DELVE test. We now present the power diagrams for DELVE+. Please see Figure S1, where the simulation settings are the same as those in Figure 2. Comparing these two figures, we observe that DELVE+ and DELVE have similar power on simulated data. This is consistent with our theory in Section 2.2.

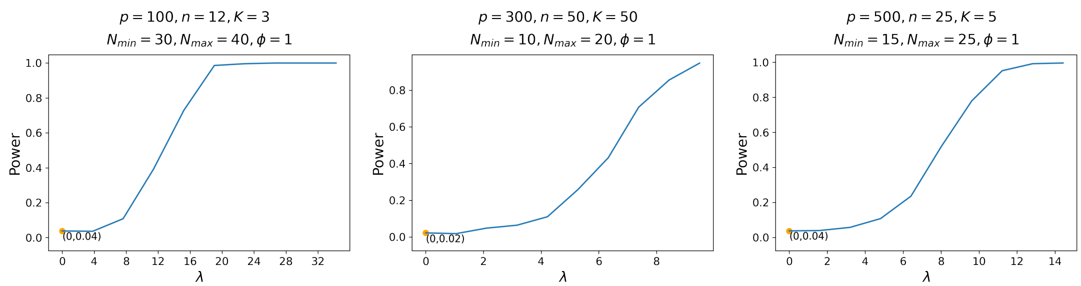

Figure S1: Power of the level-5% DELVE+ test (x-axis represents the SNR $\lambda(\tau_n) = \frac{n\bar{N}\|\mu\|\tau_n^2}{\sqrt{K}}$).

A.2 More comparison between LR and DELVE+

In Experiment 3 of Section 5, we compare the power of DELVE+ with that of the likelihood ratio (LR) test. We recall that in our general setting (3), both the null and alternative hypotheses are highly composite, because Ω_i's are allowed to be unequal within each group. It is impossible to compute the LR test statistic, except in the special setting where all of the Ω_i's in group k are equal to μ_k. In this special setting, the LR test statistic takes the form

$$LR := \sum_k n_k \bar{N}_k \sum_j \hat{\mu}_{kj} \log\big(\frac{\hat{\mu}_{kj}}{\hat{\mu}_j}\big), \tag{A.1}$$

where

$$\hat{\mu}_k = \frac{1}{n_k \bar{N}_k} \sum_{i \in S_k} X_i, \qquad \text{and} \qquad \hat{\mu} = \frac{1}{n\bar{N}} \sum_{k=1}^{K} n_k \bar{N}_k \hat{\mu}_k = \frac{1}{n\bar{N}} \sum_{i=1}^{n} X_i. \tag{A.2}$$

To ensure that LR is well-defined in the case of zero-counts (ie, $\hat{\mu}_{kj} = 0$), we define $\log(0/0) = 0$.

In Figure 3 of the main paper, we have seen the power diagrams of LR and DELVE+ for two values of $(p, n, K, N_{\min}, N_{\max}, \phi)$. Results for some other values of $(p, n, K, N_{\min}, N_{\max}, \phi)$ are in Figure S2. These results suggest that when p is relatively large, DELVE+ outperforms LR in terms of power. In theory, DELVE+ attains the optimal detection boundary, but the asymptotic behavior of LR for large-p is unclear. There are cases where LR performs somewhat better than DELVE+, but they seem to be limited to the smaller-p regime.

B Supplementary results from real data

B.1 The pairwise Z-score of another author

In Section 6.1, we give a pair-wise Z-score plot for a representative author (denoted by Author A). We can produce such a plot for any author in our data set. Here we show another example (this

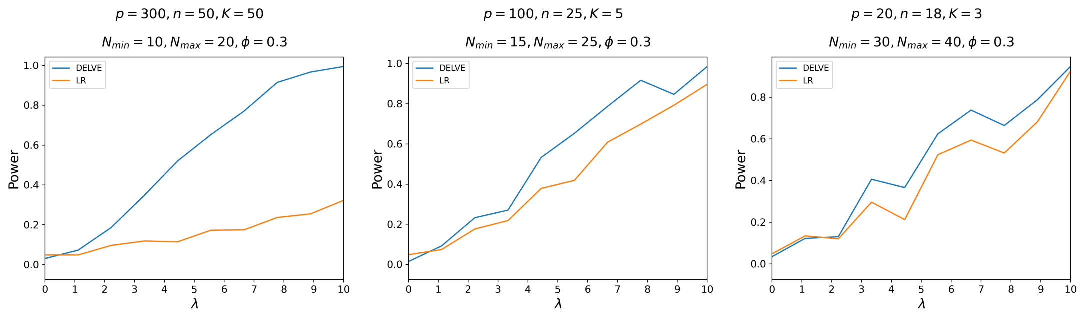

Figure S2: Power curves for DELVE+ (blue) and LR (orange) versus SNR λ for two different settings of $(p, n, K, N_{\min}, N_{\max}, \phi)$.

author is denoted by Author B). Compared to Author A, the publication years of Author B's papers are less evenly distributed. We divide Author B's abstracts into 6 groups, and the time window sizes for 6 groups are unequal, to guarantee that all groups have roughly equal numbers of abstracts. The pairwise Z-score plot for Author B is in the right panel of Figure S3. We also include the pairwise Z-score plot for Author A in the left panel of this figure (which is the same as the right panel of Figure 4).

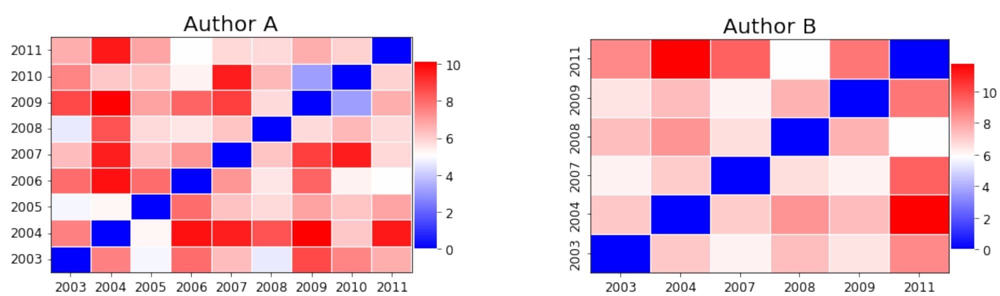

Figure S3: Pairwise Z-score plots for Author A (left) and Author B (right). In the cell (x, y), we compare the corpus of an author's abstracts from time x with the corpus of that author's abstracts from time y. The heatmap shows the value of DELVE+ with $K = 2$ for each cell.

There are some interesting temporal patterns. For Author A, the group consisting of 2004-2005 abstracts has comparably large Z-scores in the pairwise comparison with other groups, and similarly for Author B, the group of 2011-2012 abstracts have relatively large Z-scores. To gain further insight, we collected the titles and abstracts of each author's papers and manually inspected them. We found that Author A extensively studied topics related to bandwidth selection in the context of nonparametric estimation. For Author B, the time period 2011-2012 reveals a more intense focus on variable selection, compared to this author's papers in other years within this data set.

B.2 Checking the applicability of our asymptotic result on real data

The properties of the DELVE test are established in the asymptotic regime of $n^2\bar{N}^2/(Kp) \to \infty$ (see Section 3). We check if this "asymptotics" is reasonable for real applications. To this end, define the *dimension ratio* as

$$DR := n^2\bar{N}^2K^{-1}p^{-1}. \tag{B.1}$$

Author	Total papers (n)	Average abstract length ($\bar{N}$)	Vocab size (p)	DR ($n\bar{N}^2/p$)
1	81	75.90	1103	423.07
2	40	81.78	801	333.94
3	39	75.38	758	292.39
4	32	68.66	562	268.39
5	30	98.77	672	435.48
6	27	85.74	698	284.37
7	27	72.59	592	240.34
8	24	65.58	471	219.17
9	22	61.23	415	198.73
10	20	73.55	463	233.68
11	20	84.15	502	282.12
12	19	114.53	617	403.90
13	19	52.47	361	144.92
14	18	77.06	459	232.85
15	18	59.17	369	170.77

Figure S4: Summary statistics and DR values of the corpora of the top 15 most prolific authors.

	2003	2004	2005	2006	2007	2008	2009	2010
2003	—							
2004	1411	—						
2005	1313	1518	—					
2006	1986	2208	2107	—				
2007	1408	1541	1470	2216				
2008	1448	1615	1547	2263	1615			
2009	1887	2088	1981	2753	2065	2223		
2010	1506	1714	1650	2395	1714	1758	2293	
2011	1393	1576	1499	2213	1617	1631	2160	1762

Time \ Time	2003	2004	2007	2008	2009
2003					
2004	1145				
2007	859	1636			
2008	784	1548	1263		
2009	1226	2064	1675	1597	
2011	963	1694	1347	1358	1843

Figure S5: The DR values for cells the pairwise Z-score plots in Figure S3, where the left table is for Author A and the right table is for Author B.

The larger DR, the more appropriate to apply our asymptotic theory. We report the DR values of all the corpora used in the analysis of statistics abstracts. In the first experiment of Section 6.1, for each author, we take all his/her abstracts as the corpus and apply DELVE with $K = n$. Each author is associated with a corpus. Figure S4 displays the DR values for the corpora of the 15 most prolific authors. In the second experiment of Section 6.1, we take the abstracts written by an author (Author A), divide them by year into 9 groups, and apply DELVE with $K = 2$ to each pair of groups. There are a total of $(9 \times 8)/2 = 36$ corpora for this experiment, whose DR values are shown in the left panel of Figure S5. In Section B.1, we conduct similar analysis for another author (Author B). The DR values in this experiment are in the right panel of Figure S5. These DR values are large, suggesting that our asymptotic setting is relevant for real applications and that the Z-scores obtained in these experiments are trustworthy.

C Some analysis of the naive ANOVA test

In Section 2, we introduced a native estimator of ρ^2 as

$$\widetilde{T} = \sum_{k=1}^{K} n_k \bar{N}_k \|\hat{\mu}_k - \hat{\mu}\|^2.$$

Consider a $K \times p$ "contingency table" whose (k,j)th cell is $\sum_{i\in S_k} X_i(j)$. Then, $\widetilde{T}$ is an ANOVA-type statistic associated with this contingency table. It is interesting to investigate the test based on $\widetilde{T}$ and compare it with our proposed DELVE test.

In the proof of Lemma D.1, we will show that

$$\mathbb{E}[\widetilde{T}] = \rho^2 + J_5, \qquad \text{where} \quad J_5 = \sum_{k=1}^{K}\sum_{i\in S_k}\sum_{j}\Big(1-\frac{n_k\bar{N}_k}{n\bar{N}}\Big)\frac{N_i\Omega_{ij}(1-\Omega_{ij})}{n_k\bar{N}_k}. \tag{C.1}$$

Here, ρ^2 is the signal of interest, and J_5 characterizes the bias in $\widetilde{T}$. To gain some insight about the order of these two terms, we consider a simple case where (i) groups have equal size, (ii) N_i's are equal, (iii) $\Omega_{ij} = O(p^{-1})$, (iv) under H_1, $\min_k \|\mu_k - \mu\| \geq c_0\|\mu\|$, for a constant $c_0 > 0$. It holds that

$$J_5 \asymp K/p \ \text{ under } H_0 \text{ and } H_1, \quad \text{and} \quad \rho^2 \asymp n\bar{N}/p^2 \ \text{ under } H_1. \tag{C.2}$$

The bias term is negligible if $n\bar{N} \ll Kp$. This is a stronger condition than the optimal detection boundary, which only requires $n^2N^2 \gg Kp$. In particular, when

$$Kp \ \ll \ n^2\bar{N}^2 \ \ll \ K^2p^2,$$

the bias term dominates the "signal" term, so the test based on $\widetilde{T}$ may lose power. In comparison, the DELVE statistic T in (9) is a de-biased version of $\widetilde{T}$, hence, it has no such issue.

An example where $\widetilde{T}$ is powerless. Suppose $K = n$, both n and p are even, and $N_i \equiv N$. Take two vectors $\sigma \in \{-1,1\}^p$ and $\varepsilon \in \{-1,1\}^n$ such that $\sum_{j=1}^p \sigma_j = 0$ and $\sum_{i=1}^n \varepsilon_i = 0$. Under H_0, let $\Omega = p^{-1}\mathbf{1}_p\mathbf{1}_n'$. Under H_1, let $\Omega_{ij} = p^{-1} + \alpha p^{-1}\varepsilon_i\sigma_j$, for some $\alpha \in (0,1)$. We can easily check that each Ω_i is indeed a PMF. For this example,

$$\begin{aligned} J_5^{alt} - J_5^{null} &= (1-\frac{1}{n})\sum_{i,j}\frac{1}{p}(1+\alpha\varepsilon_i\sigma_j)(1-\frac{1}{p}-\frac{1}{p}\alpha\varepsilon_i\sigma_j) - (1-\frac{1}{n})\sum_{i,j}\frac{1}{p}(1-\frac{1}{p}) \\ &= -(1-\frac{1}{n})\sum_{i,j}\frac{1}{p^2}\alpha^2\varepsilon_i^2\sigma_j^2 = -(1-\frac{1}{n})\frac{\alpha^2 n}{p}. \end{aligned}$$

Moreover, $\rho^2_{null} = 0$, and $\rho^2_{alt} = O(nN/p^2)$. When $p \gg N$ and α is lower bounded by a constant,

$$\mathbb{E}_1[\widetilde{T}] - \mathbb{E}_0[\widetilde{T}] = \rho^2_{alt} + J_5^{alt} - J_5^{null} = O\Big(\frac{nN}{p^2}\Big) - (1-\frac{1}{n})\frac{\alpha^2 n}{p} \leq -\frac{\alpha^2 n}{2p}.$$

Since $\mathbb{E}_1[\widetilde{T}]$ is smaller than $\mathbb{E}_0[\widetilde{T}]$, the test based on $\widetilde{T}$ is powerless.

D Properties of T and V

This section is a preparation for the proofs of our main theorems. We recall that

$$X_i \sim \text{Multinomial}(N_i, \Omega_i), \qquad 1 \leq i \leq n. \tag{D.1}$$

For each $1 \leq k \leq K$, define

$$\mu_k = \frac{1}{n_k\bar{N}_k}\sum_{i\in S_k} N_i\Omega_i \ \in \ \mathbb{R}^p, \qquad \Sigma_k = \frac{1}{n_k\bar{N}_k}\sum_{i\in S_k} N_i\Omega_i\Omega_i' \ \in \ \mathbb{R}^{p\times p}. \tag{D.2}$$

Moreover, let

$$\mu = \frac{1}{n\bar{N}}\sum_{k=1}^K n_k\bar{N}_k\mu_k = \frac{1}{n\bar{N}}\sum_{i=1}^n N_i\Omega_i\,, \quad \Sigma = \frac{1}{n\bar{N}}\sum_{k=1}^K n_k\bar{N}_k\Sigma_k = \frac{1}{n\bar{N}}\sum_{i=1}^n N_i\Omega_i\Omega_i' \tag{D.3}$$

The DELVE test statistic is $\psi = T/\sqrt{V}$, where T is as in (9) and V is as in (11). As a preparation for the main proofs, in this section, we study T and V separately.

D.1 The decomposition of T

It is well-known that a multinomial with the number of trials equal to N can be equivalently written as the sum of N independent multinomials each with the number of trials equal to 1. This inspires us to introduce a set of independent, mean-zero random vectors:

$$\{Z_{ir}\}_{1\le i\le n, 1\le r\le N_i}, \qquad \text{with } Z_{ir} = B_{ir} - \mathbb{E}B_{ir}, \text{ and } B_{ir} \sim \text{Multinomial}(1, \Omega_i). \tag{D.4}$$

We use them to get a decomposition of T into mutually uncorrelated terms:

Lemma D.1. *Let $\{Z_{ir}\}_{1\le i\le n,1\le r\le N_i}$ be as in (D.4). For each $Z_{ir} \in \mathbb{R}^p$, let $\{Z_{ijr}\}_{1\le j\le p}$ denote its p coordinates. Recall that $\rho^2 = \sum_{k=1}^K n_k \bar{N}_k \|\mu_k - \mu\|^2$. For $1 \le j \le p$, define*

$$\begin{aligned}
U_{1j} &= 2\sum_{k=1}^K \sum_{i\in S_k} \sum_{r=1}^{N_i} (\mu_{kj} - \mu_j) Z_{ijr}, \\
U_{2j} &= \sum_{k=1}^K \sum_{i\in S_k} \sum_{1\le r\ne s\le N_i} \Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big) \frac{N_i}{N_i - 1} Z_{ijr} Z_{ijs}, \\
U_{3j} &= -\frac{1}{n\bar{N}} \sum_{1\le k\ne \ell\le K} \sum_{i\in S_k} \sum_{m\in S_\ell} \sum_{r=1}^{N_i} \sum_{s=1}^{N_m} Z_{ijr} Z_{mjs}, \\
U_{4j} &= \sum_{k=1}^K \sum_{\substack{i\in S_k, m\in S_k \\ i\ne m}} \sum_{r=1}^{N_i} \sum_{s=1}^{N_m} \Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big) Z_{ijr} Z_{mjs}.
\end{aligned}$$

Then, $T = \rho^2 + \sum_{\kappa=1}^4 \mathbf{1}_p' U_\kappa$. Moreover, $\mathbb{E}[U_\kappa] = \mathbf{0}_p$ and $\mathbb{E}[U_\kappa U_\zeta'] = \mathbf{0}_{p\times p}$ for $1 \le \kappa \ne \zeta \le 4$.

D.2 The variance of T

By Lemma D.1, the four terms $\{\mathbf{1}_p' U_\kappa\}_{1\le\kappa\le 4}$ are uncorrelated with each other. Therefore,

$$\text{Var}(T) = \text{Var}(\mathbf{1}_p' U_1) + \text{Var}(\mathbf{1}_p' U_2) + \text{Var}(\mathbf{1}_p' U_3) + \text{Var}(\mathbf{1}_p' U_4).$$

It suffices to study the variance of each of these four terms.

Lemma D.2. *Let U_1 be the same as in Lemma D.1. Define*

$$\Theta_{n1} = 4\sum_{k=1}^K n_k \bar{N}_k \big\|\text{diag}(\mu_k)^{1/2}(\mu_k - \mu)\big\|^2 \tag{D.5}$$

$$L_n = 4\sum_{k=1}^K n_k \bar{N}_k \big\|\Sigma_k^{1/2}(\mu_k - \mu)\big\|^2 \tag{D.6}$$

Then $\text{Var}(\mathbf{1}_p' U_1) = \Theta_{n1} - L_n$. Furthermore, if $\max_{1\le k\le K} \|\mu_k\|_\infty = o(1)$, then $\text{Var}(\mathbf{1}_p' U_1) = o(\rho^2)$.

Lemma D.3. *Let U_2 be the same as in Lemma D.1. Define*

$$\Theta_{n2} = 2\sum_{k=1}^K \Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 \sum_{i\in S_k} \frac{N_i^3}{N_i - 1} \|\Omega_i\|^2 \tag{D.7}$$

$$A_n = 2\sum_{k=1}^K \Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 \sum_{i\in S_k} \frac{N_i^3}{N_i - 1} \|\Omega_i\|_3^3 \tag{D.8}$$

Then

$$\Theta_{n2} - A_n \le \text{Var}(\mathbf{1}_p' U_2) \le \Theta_{n2}.$$

Furthermore, if

$$\max_{1\le k\le K}\Big\{\frac{\sum_{i\in S_k} N_i^2\|\Omega_i\|_3^3}{\sum_{i\in S_k} N_i^2\|\Omega_i\|^2}\Big\} = o(1), \tag{D.9}$$

then $\mathrm{Var}(\mathbf{1}_p' U_2) = [1+o(1)]\cdot\Theta_{n2}$.

Lemma D.4. *Let* U_3 *be the same as in Lemma D.1. Define*

$$\Theta_{n3} = \frac{2}{n^2\bar{N}^2}\sum_{k\neq \ell}\sum_{i\in S_k}\sum_{m\in S_\ell}\sum_j N_i N_m \Omega_{ij}\Omega_{mj} \tag{D.10}$$

$$B_n = 2\sum_{k\neq\ell}\frac{n_k n_\ell \bar{N}_k\bar{N}_\ell}{n^2\bar{N}^2}\mathbf{1}_p'(\Sigma_k\circ\Sigma_\ell)\mathbf{1}_p \tag{D.11}$$

Then

$$\Theta_{n3} - B_n \le \mathrm{Var}(\mathbf{1}_p' U_3) \le \Theta_{n3} + B_n.$$

Lemma D.5. *Let* U_4 *be the same as in Lemma D.1. Define*

$$\Theta_{n4} = 2\sum_{k=1}^K \sum_{\substack{i\in S_k, m\in S_k\\ i\neq m}}\sum_j\Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 N_i N_m\Omega_{ij}\Omega_{mj}. \tag{D.12}$$

$$E_n = 2\sum_k \sum_{\substack{i\in S_k, m\in S_k,\\ i\neq m}}\sum_{1\le j,j'\le p}\Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 N_i N_m \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'} \tag{D.13}$$

Then

$$\Theta_{n4} - E_n \le \mathrm{Var}(\mathbf{1}_p' U_4) \le \Theta_{n4} + E_n$$

.

Using Lemmas D.2-D.5, we derive regularity conditions such that the first term in $\mathrm{Var}(\mathbf{1}_p' U_\kappa)$ is the dominating term. Observe that $\Theta_n = \Theta_{n1} + \Theta_{n2} + \Theta_{n3} + \Theta_{n4}$, where the quantity Θ_n is defined in (10). The following intermediate result is useful.

Lemma D.6. *Suppose that* (21) *holds. Then*

$$\Theta_{n2} + \Theta_{n3} + \Theta_{n4} \asymp \sum_k \|\mu_k\|^2. \tag{D.14}$$

Moreover, under the null hypothesis, $\Theta_n \asymp K\|\mu\|^2$.

The next result is useful in proving that our variance estimator V is asymptotically unbiased.

Lemma D.7. *Suppose that* (21) *holds, and recall the definition of* Θ_n *in* (10)*. Define*

$$\beta_n = \frac{\max\Big\{\sum_k\sum_{i\in S_k}\frac{N_i^2}{n_k^2\bar{N}_k^2}\|\Omega_i\|_3^3,\ \sum_k\|\Sigma_k\|_F^2\Big\}}{K\|\mu\|^2}. \tag{D.15}$$

If $\beta_n = o(1)$*, then under the null hypothesis,* $\mathrm{Var}(T) = [1+o(1)]\cdot\Theta_n$.

We also study the case of $K=2$ more explicitly. In the lemmas below we use the notation from Section 3.4. First we have an intermediate result analogous to Lemma D.6 that holds under weaker conditions.

Lemma D.8. *Consider $K = 2$ and suppose that $\min N_i \geq 2$, $\min M_i \geq 2$ Then*

$$\Theta_{n2} + \Theta_{n3} + \Theta_{n4} \asymp \left\| \frac{m\bar{M}}{n\bar{N} + m\bar{M}}\eta + \frac{n\bar{N}}{n\bar{N} + m\bar{M}}\theta \right\|^2.$$

Moreover, under the null hypothesis, $\Theta_n \asymp \|\mu\|^2$.

The next result is a version of Lemma D.7 for the case $K = 2$ that holds under weaker conditions.

Lemma D.9. *Suppose that $\min_i N_i \geq 2$ and $\min_i M_i \geq 2$. Define*

$$\beta_n^{(2)} = \frac{\max\left\{ \sum_i N_i^2 \|\Omega_i\|^3,\ \sum_i M_i^2 \|\Gamma_i\|^3,\ \|\Sigma_1\|_F^2 + \|\Sigma_2\|_F^2 \right\}}{\|\mu\|^2}. \tag{D.16}$$

If $\beta_n^{(2)} = o(1)$, then under the null hypothesis, $\mathrm{Var}(T) = [1 + o(1)] \cdot \Theta_n$.

D.3 The decomposition of V

Lemma D.10. *Let $\{Z_{ir}\}_{1\leq i\leq n, 1\leq r\leq N_i}$ be as in* (D.4)*. Recall that*

$$V = 2\sum_{k=1}^{K}\sum_{i\in S_k}\sum_{j=1}^{p}\Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2\Big[\frac{N_i X_{ij}^2}{N_i - 1} - \frac{N_i X_{ij}(N_i - X_{ij})}{(N_i-1)^2}\Big] \tag{D.17}$$

$$+ \frac{2}{n^2\bar{N}^2}\sum_{1\leq k\neq \ell\leq K}\sum_{i\in S_k}\sum_{m\in S_\ell}\sum_{j=1}^{p} X_{ij}X_{mj} + 2\sum_{k=1}^{K}\sum_{\substack{i\in S_k, m\in S_k,\\ i\neq m}}\sum_{j=1}^{p}\Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 X_{ij}X_{mj}.$$

Define

$$\theta_i = \Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2\frac{N_i^3}{N_i - 1} \quad \textit{for } i \in S_k\ , \quad \textit{and let}$$

$$\alpha_{im} = \begin{cases} \frac{2}{n^2\bar{N}^2} & \textit{if } i\in S_k, m\in S_\ell, k\neq \ell \\ 2\big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\big)^2 & \textit{if } i, m\in S_k \end{cases}$$

If we let

$$A_1 = \sum_i\sum_{r=1}^{N_i}\sum_j\Big[\frac{4\theta_i\Omega_{ij}}{N_i} + \sum_{m\in[n]\setminus\{i\}} 2\alpha_{im}N_m\Omega_{mj}\Big]Z_{ijr}, \tag{D.18}$$

$$A_2 = \sum_i\sum_{r\neq s\in[N_i]}\frac{2\theta_i}{N_i(N_i-1)}\Big(\sum_j Z_{ijr}Z_{ijs}\Big) \tag{D.19}$$

$$A_3 = \sum_{i\neq m}\sum_{r=1}^{N_i}\sum_{s=1}^{N_m}\alpha_{im}\Big(\sum_j Z_{ijr}Z_{mjs}\Big), \tag{D.20}$$

then these terms are mean zero, are mutually uncorrelated, and satisfy

$$V = A_1 + A_2 + A_3 + \Theta_{n2} + \Theta_{n3} + \Theta_{n4}. \tag{D.21}$$

D.4 Properties of V

First we control the variance of V.

Lemma D.11. *Let A_1, A_2, and A_3 be defined as in Lemma D.10. Then*

$$\begin{aligned}
\mathrm{Var}(A_1) &\lesssim \frac{1}{n\bar{N}}\|\mu\|_3^3 + \sum_k \frac{\|\mu_k\|_3^3}{n_k \bar{N}_k} \lesssim \sum_k \frac{\|\mu_k\|_3^3}{n_k \bar{N}_k} \\
\mathrm{Var}(A_2) &\lesssim \sum_k \sum_{i \in S_k} \frac{N_i^2 \|\Omega_i\|_2^2}{n_k^4 \bar{N}_k^4} \lesssim \sum_k \frac{\|\mu_k\|^2}{n_k^2 \bar{N}_k^2} \\
\mathrm{Var}(A_3) &\lesssim \sum_k \frac{\|\mu_k\|^2}{n_k^2 \bar{N}_k^2} + \frac{1}{n^2 \bar{N}^2}\|\mu\|^2 \lesssim \sum_k \frac{\|\mu_k\|^2}{n_k^2 \bar{N}_k^2}.
\end{aligned}$$

Next we show consistency of V under the null, which is crucial in properly standardizing our test statistic and establishing asymptotic normality.

Proposition D.12. *Recall the definition of β_n in* (D.15)*. Suppose that $\beta_n = o(1)$ and that the condition* (21) *holds. If under the null hypothesis we have*

$$K^2 \|\mu\|^4 \gg \sum_k \frac{\|\mu\|^2}{n_k^2 \bar{N}_k^2} \vee \sum_k \frac{\|\mu\|_3^3}{n_k \bar{N}_k}, \tag{D.22}$$

then $V/\mathrm{Var}T \to 1$ in probability.

To later control the type II error, we must also show that V does not dominate the true variance under the alternative. We first state an intermediate result that is useful throughout.

Lemma D.13. *Suppose that, under either the null or alternative, $\max_i \|\Omega_i\|_\infty \leq 1 - c_0$ holds for an absolute constant $c_0 > 0$. Then*

$$\mathrm{Var}(T) \gtrsim \Theta_{n2} + \Theta_{n3} + \Theta_{n4}. \tag{D.23}$$

Proposition D.14. *Suppose that under the alternative* (21) *holds and*

$$\Big(\sum_k \|\mu_k\|^2\Big)^2 \gg \sum_k \frac{\|\mu_k\|^2}{n_k^2 \bar{N}_k^2} \vee \sum_k \frac{\|\mu_k\|_3^3}{n_k \bar{N}_k}. \tag{D.24}$$

Then $V = O_{\mathbb{P}}(\mathrm{Var}(T))$ under the alternative.

We also require versions of Proposition D.12 and Proposition D.14 that hold under weaker conditions in the special case $K = 2$. We omit the proofs as they are similar. Below we use the notation of Section 3.4.

Proposition D.15. *Suppose that $K = 2$ and recall the definition of $\beta_n^{(2)}$ in D.16. Suppose that $\beta_n^{(2)} = o(1)$, $\min_i N_i \geq 2, \min_i M_i \geq 2$, and $\max_i \|\Omega_i\|_\infty \leq 1 - c_0, \max_i \|\Gamma_i\|_\infty \leq 1 - c_0$. If under the null hypothesis*

$$\|\mu\|^4 \gg \max\Big\{ \big(\frac{\|\mu\|_2^2}{n^2 \bar{N}^2} + \frac{\|\mu\|_2^2}{m^2 \bar{M}_2^2}\big), \big(\frac{\|\mu\|_3^3}{n\bar{N}} + \frac{\|\mu\|_3^3}{m\bar{M}}\big) \Big\}, \tag{D.25}$$

then $V/\mathrm{Var}(T) \to 1$ in probability.

Under the alternative we have the following.

Proposition D.16. *Suppose that $K = 2$, $\min_i N_i \geq 2, \min_i M_i \geq 2$, and $\max_i \|\Omega_i\|_\infty \leq 1 - c_0, \max_i \|\Gamma_i\|_\infty \leq 1 - c_0$. If under the alternative*

$$\Big\| \frac{m\bar{M}}{n\bar{N} + m\bar{M}}\eta + \frac{n\bar{N}}{n\bar{N} + m\bar{M}}\theta \Big\|^4 \gg \max\Big\{ \big(\frac{\|\eta\|_2^2}{n^2 \bar{N}^2} + \frac{\|\theta\|_2^2}{m^2 \bar{M}_2^2}\big), \big(\frac{\|\eta\|_3^3}{n\bar{N}} + \frac{\|\theta\|_3^3}{m\bar{M}}\big) \Big\}, \tag{D.26}$$

then $V = O_{\mathbb{P}}(\mathrm{Var}(T))$.

In the setting of $K = n$ and utilize the variance estimator V^*. The next results capture the behavior of V^* under the null and alternative. The proofs are given later in this section.

Proposition D.17. *Define*

$$\beta_n^{(n)} = \frac{\sum_i \|\Omega_i\|^3}{n\|\mu\|^2}. \tag{D.27}$$

Suppose that (21) *holds,* $\beta_n^{(n)} = o(1)$*, and*

$$n^2\|\mu\|^4 \gg \sum_i \frac{\|\mu\|^2}{N_i^2} \vee \sum_i \frac{\|\mu\|_3^3}{N_i}. \tag{D.28}$$

Then $V^*/\mathrm{Var}(T) \to 1$ *in probability as* $n \to \infty$*.*

Proposition D.18. *Suppose that under the alternative* (21) *holds and*

$$\big(\sum_i \|\Omega_i\|^2\big)^2 \gg \sum_i \frac{\|\Omega_i\|^2}{N_i^2} \vee \sum_i \frac{\|\Omega_i\|_3^3}{N_i}. \tag{D.29}$$

Then $V^* = O_{\mathbb{P}}(\mathrm{Var}(T))$ *under the alternative.*

D.5 Proof of Lemma D.1

We first show that $\mathbb{E}[U_\kappa] = \mathbf{0}_p$ and $\mathbb{E}[U_\kappa U_\zeta'] = \mathbf{0}_{p\times p}$ for $\kappa \neq \zeta$. Note that $\{Z_{ir}\}_{1\leq i\leq n, 1\leq r\leq N_i}$ are independent mean-zero random vectors. It follows that each U_κ is a mean-zero random vector. We then compute $\mathbb{E}[U_{\kappa j_1} U_{\zeta j_2}]$ for $\kappa \neq \zeta$ and all $1 \leq j_1, j_2 \leq p$. By direct calculations,

$$\mathbb{E}[U_{1j}U_{2j_2}] = 2\sum_{(k,i,r,s)}\sum_{(k',i',r')}\Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)(\mu_{k'j} - \mu_j)\frac{N_i}{N_i - 1}\mathbb{E}[Z_{ij_2r}Z_{ij_2s}Z_{i'j_1r'}].$$

If $i' \neq i$, or if $i' = i$ and $r' \notin \{r, s\}$, then $Z_{i'j_1r'}$ is independent of $Z_{ij_2r}Z_{ij_2s}$, and it follows that $\mathbb{E}[Z_{ij_2r}Z_{ij_2s}Z_{i'j_1r'}] = 0$. If $i' = i$ and $r = r'$, then $\mathbb{E}[Z_{ij_2r}Z_{ij_2s}Z_{i'j_1r'}] = \mathbb{E}[Z_{ij_2r}Z_{ij_1r}]\cdot\mathbb{E}[Z_{ij_2s}]$; since $r \neq s$, we also have $\mathbb{E}[Z_{ij_2r}Z_{ij_2s}Z_{i'j_1r'}] = 0$. This proves $\mathbb{E}[U_{1j}U_{2j_*}] = 0$. Since this holds for all $1 \leq j_1, j_2 \leq p$, we immediately have

$$\mathbb{E}[U_1U_2'] = \mathbf{0}_{p\times p}.$$

We can similarly show that $\mathbb{E}[U_\kappa U_\zeta'] = \mathbf{0}_{p\times p}$, for other $\kappa \neq \zeta$. The proof is omitted.

It remains to prove the desirable decomposition of T. Recall that $T = \sum_{j=1}^p T_j$. Write $\rho^2 = \sum_{j=1}^p \rho_j^2$, where $\rho_j^2 = 2\sum_{k=1}^K n_k\bar{N}_k(\mu_{kj} - \mu_j)^2$. It suffices to show that

$$T_j = \rho_j^2 + U_{1j} + U_{2j} + U_{3j} + U_{4j}, \qquad \text{for all } 1 \leq j \leq p. \tag{D.30}$$

To prove (D.30), we need some preparation. Define

$$Y_{ij} := \frac{X_{ij}}{N_i} - \Omega_{ij} = \frac{1}{N_i}\sum_{r=1}^{N_i} Z_{ijr}, \qquad Q_{ij} := Y_{ij}^2 - \mathbb{E}Y_{ij}^2 = Y_{ij}^2 - \frac{\Omega_{ij}(1-\Omega_{ij})}{N_i}. \tag{D.31}$$

With these notations, $X_{ij} = N_i(\Omega_{ij} + Y_{ij})$ and $N_iY_{ij}^2 = N_iQ_{ij} + \Omega_{ij}(1 - \Omega_{ij})$. Moreover, we can use (D.31) to re-write Q_{ij} as a function of $\{Z_{ijr}\}_{1\leq r\leq N_i}$ as follows:

$$Q_{ij} = \frac{1}{N_i^2}\sum_{r=1}^{N_i}[Z_{ijr}^2 - \Omega_{ij}(1-\Omega_{ij})] + \frac{1}{N_i^2}\sum_{1\leq r\neq s\leq N_i} Z_{ijr}Z_{ijs}.$$

Note that $Z_{ijr} = B_{ijr} - \Omega_{ij}$, where B_{ijr} can only take values in $\{0, 1\}$. Hence, $(Z_{ijr} + \Omega_{ij})^2 = (Z_{ijr} + \Omega_{ij})$ always holds. Re-arranging the terms gives $Z_{ijr}^2 - \Omega_{ij}(1 - \Omega_{ij}) = (1 - 2\Omega_{ij})Z_{ijr}$. It follows that

$$Q_{ij} = (1 - 2\Omega_{ij})\frac{Y_{ij}}{N_i} + \frac{1}{N_i^2}\sum_{1\leq r\neq s\leq N_i} Z_{ijr}Z_{ijs}. \tag{D.32}$$

This is a useful equality which we will use in the proof below.

We now show (D.30). Fix j and write $T_j = R_j - D_j$, where

$$R_j = \sum_{k=1}^{K} n_k\bar{N}_k(\hat{\mu}_{kj} - \hat{\mu}_j)^2, \quad \text{and} \quad D_j = \sum_{k=1}^{K}\sum_{i\in S_k} \xi_k \frac{X_{ij}(N_i - X_{ij})}{n_k\bar{N}_k(N_i - 1)}, \quad \text{with } \xi_k = 1 - \frac{n_k\bar{N}_k}{n\bar{N}}$$

First, we study D_j. Note that $X_{ij}(N_{ij} - X_{ij}) = N_i^2(\Omega_{ij} + Y_{ij})(1 - \Omega_{ij} - Y_{ij}) = N_i^2\Omega_{ij}(1 - \Omega_{ij}) - N_i^2 Y_{ij}^2 + N_i^2(1 - 2\Omega_{ij})Y_{ij}$, where $Y_{ij}^2 = Q_{ij} + N_i^{-1}\Omega_{ij}(1 - \Omega_{ij})$. It follows that

$$\frac{X_{ij}(N_{ij} - X_{ij})}{N_i(N_i - 1)} = \Omega_{ij}(1 - \Omega_{ij}) - \frac{N_i Q_{ij}}{N_i - 1} + \frac{N_i}{N_i - 1}(1 - 2\Omega_{ij})Y_{ij}.$$

We apply (D.32) to get

$$\frac{X_{ij}(N_{ij} - X_{ij})}{N_i(N_i - 1)} = \Omega_{ij}(1 - \Omega_{ij}) + (1 - 2\Omega_{ij})Y_{ij} - \frac{1}{N_i(N_i - 1)}\sum_{1\leq r\neq s\leq N_i} Z_{ijr}Z_{ijs}. \tag{D.33}$$

It follows that

$$\begin{aligned} D_j = &\sum_{k=1}^{K}\sum_{i\in S_k}\frac{\xi_k N_i}{n_k\bar{N}_k}\Omega_{ij}(1 - \Omega_{ij}) + \sum_{k=1}^{K}\sum_{i\in S_k}\frac{\xi_k N_i}{n_k\bar{N}_k}(1 - 2\Omega_{ij})Y_{ij} \\ &- \sum_{k=1}^{K}\sum_{i\in S_k}\frac{\xi_k}{n_k\bar{N}_k(N_i - 1)}\sum_{1\leq r\neq s\leq N_i} Z_{ijr}Z_{ijs}. \end{aligned} \tag{D.34}$$

Next, we study R_j. Note that $n_k\bar{N}_k(\hat{\mu}_{kj} - \hat{\mu}_j) = \sum_{i\in S_k}(X_{ij} - \bar{N}_k\hat{\mu}_j)$. It follows that

$$R_j = \sum_{k=1}^{K}\frac{1}{n_k\bar{N}_k}\Big[\sum_{i\in S_k}(X_{ij} - \bar{N}_k\hat{\mu}_j)\Big]^2.$$

Recall that $X_{ij} = N_i(\Omega_{ij} + Y_{ij})$. By direct calculations, $\sum_{i\in S_k} X_{ij} = n_k\bar{N}_k\mu_{kj} + \sum_{i\in S_k} N_i Y_{ij}$, and $\hat{\mu}_j = \mu_j + (n\bar{N})^{-1}\sum_{m=1}^{n} N_m Y_{mj}$. We then have the following decomposition:

$$\sum_{i\in S_k}(X_{ij} - \bar{N}_k\hat{\mu}_j) = n_k\bar{N}_k(\mu_{kj} - \mu_j) + \sum_{i\in S_k} N_i Y_{ij} - \frac{n_k\bar{N}_k}{n\bar{N}}\Big(\sum_{m=1}^{n} N_m Y_{mj}\Big).$$

Using this decomposition, we can expand $[\sum_{i\in S_k}(X_{ij} - \bar{N}_k\hat{\mu}_j)]^2$ to a total of 6 terms, where 3 are quadratic terms and 3 are cross terms. It yields a decomposition of R_j into 6 terms:

$$\begin{aligned} R_j = &\sum_{k=1}^{K} n_k\bar{N}_k(\mu_{kj} - \mu_j)^2 + \sum_{k=1}^{K}\frac{1}{n_k\bar{N}_k}\Big(\sum_{i\in S_k} N_i Y_{ij}\Big)^2 + \sum_{k=1}^{K}\frac{n_k\bar{N}_k}{n^2\bar{N}^2}\Big(\sum_{m=1}^{n} N_m Y_{mj}\Big)^2 \\ &+ 2\sum_{k=1}^{K}(\mu_{kj} - \mu_j)\Big(\sum_{i\in S_k} N_i Y_{ij}\Big) - 2\sum_{k=1}^{K}\frac{n_k\bar{N}_k}{n\bar{N}}(\mu_{kj} - \mu_j)\Big(\sum_{m=1}^{n} N_m Y_{mj}\Big) \\ &- \frac{2}{n\bar{N}}\sum_{k=1}^{K}\Big(\sum_{i\in S_k} N_i Y_{ij}\Big)\Big(\sum_{m=1}^{n} N_m Y_{mj}\Big) \end{aligned}$$

$$\equiv I_1 + I_2 + I_3 + I_4 + I_5 + I_6. \tag{D.35}$$

By definition, $\sum_{k=1}^K n_k \bar{N}_k = n\bar{N}$ and $\sum_{k=1}^K n_k \bar{N}_k \mu_{kj} = n\bar{N}\mu_j$. It follows that

$$I_3 = \frac{1}{n\bar{N}}\Big(\sum_{m=1}^n N_m Y_{mj}\Big)^2, \qquad I_5 = 0, \qquad I_6 = -\frac{2}{n\bar{N}}\Big(\sum_{m=1}^n N_m Y_{mj}\Big)^2 = -2I_3.$$

It follows that

$$R_j = I_1 + I_2 - I_3 + I_4. \tag{D.36}$$

We further simplify I_3. Recall that $\xi_k = 1 - (n\bar{N})^{-1} n_k \bar{N}_k$. By direct calculations,

$$\begin{aligned}
I_3 &= \frac{1}{n\bar{N}}\Big(\sum_{m=1}^n N_m Y_{mj}\Big)^2 = \frac{1}{n\bar{N}}\Big[\sum_{k=1}^K\Big(\sum_{i\in S_k} N_i Y_{ij}\Big)\Big]^2 \\
&= \frac{1}{n\bar{N}}\sum_{k=1}^K\Big(\sum_{i\in S_k} N_i Y_{ij}\Big)^2 + \frac{1}{n\bar{N}}\sum_{1\le k\neq \ell\le K}\Big(\sum_{i\in S_k} N_i Y_{ij}\Big)\Big(\sum_{m\in S_\ell} N_m Y_{mj}\Big) \\
&= \sum_{k=1}^K (1-\xi_k)\frac{1}{n_k\bar{N}_k}\Big(\sum_{i\in S_k} N_i Y_{ij}\Big)^2 + \underbrace{\frac{1}{n\bar{N}}\sum_{k\neq\ell}\sum_{i\in S_k}\sum_{m\in S_\ell} N_i N_m Y_{ij} Y_{mj}}_{J_1} \\
&= I_2 - \sum_{k=1}^K\sum_{i\in S_k}\frac{\xi_k}{n_k\bar{N}_k}\Big(\sum_{i\in S_k} N_i Y_{ij}\Big)^2 + J_1 \\
&= I_2 + J_1 - \sum_{k=1}^K \frac{\xi_k}{n_k\bar{N}_k}\Big(\sum_{i\in S_k} N_i^2 Y_{ij}^2\Big) - \underbrace{\sum_{k=1}^K \frac{\xi_k}{n_k\bar{N}_k}\sum_{\substack{i\in S_k, m\in S_k\\ i\neq m}} N_i N_m Y_{ij} Y_{mj}}_{J_2}.
\end{aligned} \tag{D.37}$$

By (D.31), $N_i Y_{ij}^2 = N_i Q_i + \Omega_{ij}(1-\Omega_{ij})$. We further apply (D.32) to get

$$N_i^2 Y_{ij}^2 = N_i(1-2\Omega_{ij})Y_{ij} + \sum_{1\le r\neq s\le N_i} Z_{ijr}Z_{ijs} + N_i\Omega_{ij}(1-\Omega_{ij}).$$

It follows that

$$\begin{aligned}
\sum_{k=1}^K \frac{\xi_k}{n_k\bar{N}_k}\Big(\sum_{i\in S_k} N_i^2 Y_{ij}^2\Big) &= \underbrace{\sum_{k=1}^K\sum_{i\in S_k}\frac{\xi_k N_i}{n_k\bar{N}_k}(1-2\Omega_{ij})Y_{ij}}_{J_3} \\
&\quad + \underbrace{\sum_{k=1}^K\sum_{i\in S_k}\frac{\xi_k}{n_k\bar{N}_k}\sum_{r\neq s} Z_{ijr}Z_{ijs}}_{J_4} + \underbrace{\sum_{k=1}^K\sum_{i\in S_k}\frac{\xi_k N_i}{n_k\bar{N}_k}\Omega_{ij}(1-\Omega_{ij})}_{J_5}.
\end{aligned} \tag{D.38}$$

We plug (D.38) into (D.37) to get $I_3 = I_2 + J_1 - J_2 - J_3 - J_4 - J_5$. Further plugging I_3 into the expression of R_j in (D.36), we have

$$R_j = I_1 + I_4 - J_1 + J_2 + J_3 + J_4 + J_5, \tag{D.39}$$

where I_1 and I_4 are defined in (D.35), J_1-J_2 are defined in (D.37), and J_3-J_5 are defined in (D.38).

Finally, we combine the expressions of D_j and R_j. By (D.34) and the definitions of J_1-J_5,

$$
\begin{aligned}
D_j &= J_5 + J_3 - \sum_{k=1}^{K} \sum_{i \in S_k} \frac{\xi_k}{n_k \bar{N}_k (N_i - 1)} \sum_{r \neq s} Z_{ijr} Z_{ijs} \\
&= J_5 + J_3 + J_4 - \underbrace{\sum_{k=1}^{K} \sum_{i \in S_k} \frac{\xi_k N_i}{n_k \bar{N}_k (N_i - 1)} \sum_{r \neq s} Z_{ijr} Z_{ijs}}_{J_6} .
\end{aligned}
$$

Combining it with (D.39) gives $T_j = R_j - D_j = I_1 + I_4 - J_1 + J_2 + J_6$. We further plug in the definition of each term. It follows that

$$
\begin{aligned}
T_j = \sum_{k=1}^{K} n_k \bar{N}_k (\mu_{kj} - \mu_j)^2 &+ 2 \sum_{k=1}^{K} \sum_{i \in S_k} (\mu_{kj} - \mu_j) N_i Y_{ij} - \frac{1}{n \bar{N}} \sum_{k \neq \ell} \sum_{i \in S_k, m \in S_\ell} N_i N_m Y_{ij} Y_{mj} \\
&+ \sum_{k=1}^{K} \sum_{\substack{i \in S_k, m \in S_k \\ i \neq m}} \frac{\xi_k}{n_k \bar{N}_k} N_i N_m Y_{ij} Y_{mj} + \sum_{k=1}^{K} \sum_{i \in S_k} \frac{\xi_k N_i}{n_k \bar{N}_k (N_i - 1)} \sum_{r \neq s} Z_{ijr} Z_{ijs}.
\end{aligned} \tag{D.40}
$$

We plug in $Y_{ij} = N_i^{-1} \sum_{r=1}^{N_i} Z_{ijr}$ and take a sum of $1 \leq j \leq p$. It gives (D.30) immediately. The proof is now complete. □

D.6 Proof of Lemma D.2

Recall that $\{Z_{ir}\}_{1 \leq i \leq n, 1 \leq r \leq N_i}$ are independent random vectors. Write

$$
\mathbf{1}_p' U_1 = 2 \sum_{k=1}^{K} \sum_{i \in S_k} \sum_{r=1}^{N_i} (\mu_k - \mu)' Z_{ir}.
$$

The covariance matrix of Z_{ir} is $\text{diag}(\Omega_i) - \Omega_i \Omega_i'$. It follows that

$$
\begin{aligned}
\text{Var}(\mathbf{1}_p' U_1) &= 4 \sum_{k=1}^{K} \sum_{i \in S_k} \sum_{r=1}^{N_i} (\mu_k - \mu)' \big[\text{diag}(\Omega_i) - \Omega_i \Omega_i' \big] (\mu_k - \mu) \\
&= 4 \sum_{k} (\mu_k - \mu)' \Big[\text{diag}\Big(\sum_{i \in S_k} N_i \Omega_i \Big) - \Big(\sum_{i \in S_k} N_i \Omega_i \Omega_i' \Big) \Big] (\mu_k - \mu) \\
&= 4 \sum_{k} (\mu_k - \mu)' \Big[\text{diag}(n_k \bar{N}_k \mu_k) - n_k \bar{N}_k \Sigma_k \Big] (\mu_k - \mu) \\
&= 4 \sum_{k} n_k \bar{N}_k \big\| \text{diag}(\mu_k)^{1/2} (\mu_k - \mu) \big\|^2 - 4 \sum_{k} n_k \bar{N}_k \big\| \Sigma_k^{1/2} (\mu_k - \mu) \big\|^2 .
\end{aligned} \tag{D.41}
$$

This proves the first claim. Furthermore, by (D.41),

$$
\text{Var}(\mathbf{1}_p' U_1) \leq 4 \sum_{k} n_k \bar{N}_k \big\| \text{diag}(\mu_k)^{1/2} (\mu_k - \mu) \big\|^2 \leq 4 \sum_{k} n_k \bar{N}_k \|\text{diag}(\mu_k)\| \|\mu_k - \mu\|^2 .
$$

Note that $\|\text{diag}(\mu_k)\| = \|\mu_k\|_\infty$. Therefore, if $\max_k \|\mu_k\|_\infty = o(1)$, the right hand side above is $o(1) \cdot 4 \sum_k n_k \bar{N}_k \|\mu_k - \mu\|^2 = o(\rho^2)$. This proves the second claim. □

D.7 Proof of Lemma D.3

For each $1 \leq k \leq K$, define a set of index triplets: $\mathcal{M}_k = \{(i, r, s) : i \in S_k, 1 \leq r < s \leq N_i\}$. Let $\mathcal{M} = \cup_{k=1}^K \mathcal{M}_k$. Write for short $\theta_i = (\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}})^2 \frac{N_i^3}{N_i - 1}$, for $i \in S_k$. It is seen that

$$\mathbf{1}_p' U_2 = 2 \sum_{(i,r,s)\in\mathcal{M}} \frac{\sqrt{\theta_i}}{\sqrt{N_i(N_i-1)}} W_{irs}, \qquad \text{with} \quad W_{irs} = \sum_{j=1}^p Z_{ijr} Z_{ijs}.$$

For W_{irs} and $W_{i'r's'}$, if $i \neq i'$, or if $i = i'$ and $\{r, s\} \cap \{r', s'\} = \emptyset$, then these two variables are independent; if $i = i'$, $r = r'$ and $s \neq s'$, then $\mathbb{E}[W_{irs} W_{irs'}] = \sum_{j,j'} \mathbb{E}[Z_{ijr} Z_{ijs} Z_{ij'r} Z_{ij's'}] = \sum_{j,j'} \mathbb{E}[Z_{ijr} Z_{ij'r}] \cdot \mathbb{E}[Z_{ijs}] \cdot \mathbb{E}[Z_{ij's'}] = 0$. Therefore, $\{W_{irs}\}_{(i,r,s)\in\mathcal{M}}$ is a collection of mutually uncorrelated variables. It follows that

$$\text{Var}(\mathbf{1}_p' U_2) = 4 \sum_{(i,r,s)\in\mathcal{M}} \frac{\theta_i}{N_i(N_i-1)} \text{Var}(W_{irs}).$$

It remains to calculate the variance of each W_{irs}. By direction calculations,

$$\begin{aligned}
\text{Var}(W_{irs}) &= \sum_j \mathbb{E}[Z_{ijr}^2 Z_{ijs}^2] + 2 \sum_{j<\ell} \mathbb{E}[Z_{ijr} Z_{ijs} Z_{i\ell r} Z_{i\ell s}] \\
&= \sum_j [\Omega_{ij}(1-\Omega_{ij})]^2 + 2 \sum_{j<\ell} (-\Omega_{ij}\Omega_{i\ell})^2 \\
&= \sum_j \Omega_{ij}^2 - 2 \sum_j \Omega_{ij}^3 + \Big(\sum_j \Omega_{ij}^2\Big)^2 \\
&= \|\Omega_i\|^2 - 2\|\Omega_i\|_3^3 + \|\Omega_i\|^4
\end{aligned} \tag{D.42}$$

Since $\max_{ij} \Omega_{ij} \leq 1$, we have

$$\|\Omega_i\|^2 - \|\Omega_i\|_3^3 \leq \text{Var}(W_{irs}) \leq \|\Omega_i\|^2.$$

Therefore,

$$\begin{aligned}
\text{Var}(\mathbf{1}_p' U_2) &= 4 \sum_{k=1}^K \sum_{i\in S_k} \sum_{1\leq r<s\leq N_i} \frac{\theta_i}{N_i(N_i-1)} \text{Var}(W_{irs}) \\
&= 2 \sum_{k=1}^K \sum_{i\in S_k} \theta_i \text{Var}(W_{irs}) \geq 2 \sum_{k=1}^K \sum_{i\in S_k} \theta_i \big[\|\Omega_i\|^2 - \|\Omega_i\|_3^3\big] = \Theta_{n2} - A_n,
\end{aligned}$$

and similarly $\text{Var}(\mathbf{1}_p' U_2) \leq \Theta_{n2}$, which proves the first claim. To prove the second claim, note that $\text{Var}(\mathbf{1}_p' U_2) = \Theta_{n2} + O(A_n)$. By (D.9) and the assumption $\min N_i \geq 2$, we have

$$\begin{aligned}
A_n &\lesssim \sum_k \Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 \sum_{i\in S_k} N_i^2 \|\Omega_i\|_3^3 \\
&= \sum_k \Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 \cdot o\Big(\sum_{i\in S_k} N_i^2 \|\Omega_i\|^2\Big) = o(\Theta_{n2}),
\end{aligned}$$

which implies that $\text{Var}(\mathbf{1}_p U_2) = [1 + o(1)]\Theta_{n2}$, as desired.

$\square$

D.8 Proof of Lemma D.4

For each $1 \leq k < \ell \leq K$, define a set of index quadruples: $\mathcal{J}_{k\ell} = \{(i, r, m, s) : i \in S_k, j \in S_\ell, 1 \leq r \leq N_i, 1 \leq s \leq N_m\}$. Let $\mathcal{J} = \cup_{(k,\ell):1\leq k<\ell\leq K} \mathcal{J}_{k\ell}$. It is seen that

$$\mathbf{1}_p' U_3 = -\frac{2}{n\bar{N}} \sum_{(i,r,m,s)\in\mathcal{J}} V_{irms}, \qquad \text{where } V_{irms} = \sum_{j=1}^{p} Z_{ijr} Z_{mjs}.$$

For V_{irms} and $V_{i'r'm's'}$, if $\{(i,r),(m,s)\} \cap \{(i',r'),(m',s')\} = \emptyset$, then the two variables are independent of each other. If $(i,r) = (i',r')$ and $(m,s) \neq (m',s')$, then $\mathbb{E}[V_{irms} V_{irm's'}] = \sum_{j,j'} \mathbb{E}[Z_{ijr} Z_{mjs} Z_{ij'r} Z_{m'j's'}] = \sum_{j,j'} \mathbb{E}[Z_{ijr} Z_{ij'r}] \cdot \mathbb{E}[Z_{mjs}] \cdot \mathbb{E}[Z_{m'js'}] = 0$. Therefore, the only correlated case is when $(i,r,m,s) = (i',r',m',s')$. This implies that $\{V_{irms}\}_{(i,r,m,s)\in\mathcal{J}}$ is a collection of mutually uncorrelated variables. Therefore,

$$\text{Var}(\mathbf{1}_p' U_3) = \frac{4}{n^2 \bar{N}^2} \sum_{(i,r,m,s)\in\mathcal{J}} \text{Var}(V_{irms}).$$

Note that $\text{Var}(V_{irms}) = \mathbb{E}[(\sum_j Z_{ijr} Z_{mjs})^2] = \sum_{j,j'} \mathbb{E}[Z_{ijr} Z_{mjs} Z_{ij'r} Z_{mj's}]$; also, the covariance matrix of Z_{ir} is $\text{diag}(\Omega_i) - \Omega_i \Omega_i'$. It follows that

$$\begin{aligned}
\text{Var}(V_{irms}) &= \sum_j \mathbb{E}[Z_{ijr}^2] \cdot \mathbb{E}[Z_{mjs}^2] + \sum_{j\neq j'} \mathbb{E}[Z_{ijr} Z_{ij'r}] \cdot \mathbb{E}[Z_{mjs} Z_{mj's}] \\
&= \sum_j \Omega_{ij}(1-\Omega_{ij})\Omega_{mj}(1-\Omega_{mj}) + \sum_{j\neq j'} \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'} \\
&= \sum_j \Omega_{ij}\Omega_{mj} - 2\sum_j \Omega_{ij}^2 \Omega_{mj}^2 + \sum_{j,j'} \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'}. \qquad \text{(D.43)}
\end{aligned}$$

Write for short $\delta_{im} = -2\sum_j \Omega_{ij}^2\Omega_{mj}^2 + \sum_{j,j'} \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'}$.Combining the above gives

$$\begin{aligned}
\text{Var}(\mathbf{1}_p' U_3) &= \frac{4}{n^2\bar{N}^2} \sum_{k<\ell} \sum_{i\in S_k} \sum_{m\in S_\ell} \sum_{r=1}^{N_i} \sum_{s=1}^{N_m} \Big(\sum_j \Omega_{ij}\Omega_{mj} + \delta_{im} \Big) \\
&= \frac{2}{n^2\bar{N}^2} \sum_{k\neq\ell} \sum_{i\in S_k} \sum_{m\in S_\ell} \sum_j N_i N_m \Omega_{ij}\Omega_{mj} + \frac{2}{n^2\bar{N}^2} \sum_{k\neq\ell} \sum_{i\in S_k} \sum_{m\in S_\ell} N_i N_m \delta_{im}. \qquad \text{(D.44)}
\end{aligned}$$

It is easy to see that $|\delta_{im}| \leq \sum_{j,j'} \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'}$. Also, by the definition of Σ_k in (D.2), we have $\Sigma_k(j,j') = \frac{1}{n_k \bar{N}_k} \sum_{i\in S_k} N_i \Omega_{ij}\Omega_{ij'}$. Using these results, we immediately have

$$\begin{aligned}
\Big| \frac{2}{n^2\bar{N}^2} \sum_{k\neq\ell} \sum_{i\in S_k} \sum_{m\in S_\ell} N_i N_m \delta_{im} \Big| &\leq \frac{2}{n^2\bar{N}^2} \sum_{k\neq\ell} \sum_{i\in S_k} \sum_{m\in S_\ell} \sum_{j,j'} N_i N_m \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'} \\
&= \frac{2}{n^2\bar{N}^2} \sum_{j,j'} \sum_{k\neq\ell} \Big(\sum_{i\in S_k} N_i \Omega_{ij}\Omega_{ij'} \Big) \Big(\sum_{m\in S_\ell} N_i \Omega_{mj}\Omega_{mj'} \Big) \\
&= \frac{2}{n^2\bar{N}^2} \sum_{j,j'} \sum_{k\neq\ell} n_k \bar{N}_k \Sigma_k(j,j') \cdot n_\ell \bar{N}_\ell \Sigma_\ell(j,j') \\
&= 2 \sum_{k\neq\ell} \frac{n_k n_\ell \bar{N}_k \bar{N}_\ell}{n^2\bar{N}^2} \mathbf{1}_p' (\Sigma_k \circ \Sigma_\ell) \mathbf{1}_p =: B_n \qquad \text{(D.45)}
\end{aligned}$$

as desired.

□

D.9 Proof of Lemma D.5

For $1 \leq k \leq K$, define a set of index quadruples: $\mathcal{Q}_k = \{(i,r,m,s) : i \in S_k, m \in S_k, i < m, 1 \leq r \leq N_i, 1 \leq s \leq N_m\}$. Let $\mathcal{Q} = \cup_{k=1}^K \mathcal{Q}_k$. Write $\kappa_{im} = (\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}})^2 N_i N_m$, for $i \in S_k$ and $m \in S_k$. It is seen that

$$\mathbf{1}_p' U_4 = 2 \sum_{(i,r,m,s)\in\mathcal{Q}} \frac{\sqrt{\kappa_{im}}}{\sqrt{N_i N_m}} V_{irms}, \qquad \text{where} \quad V_{irms} = \sum_{j=1}^p Z_{ijr} Z_{mjs}.$$

It is not hard to see that V_{irms} and $V_{i'r'm's'}$ are correlated only if $(i,r,m,s) = (i',r',m',s')$. It follows that

$$\text{Var}(\mathbf{1}_p' U_4) = 4 \sum_{(i,r,m,s)\in\mathcal{Q}} \frac{\kappa_{im}}{N_i N_m} \text{Var}(V_{irms}).$$

In the proof of Lemma D.4, we have studied $\text{Var}(V_{irms})$. In particular, by (D.43), we have

$$\text{Var}(V_{irms}) = \sum_j \Omega_{ij}\Omega_{mj} + \delta_{im}, \qquad \text{with} \quad |\delta_{im}| \leq \sum_{j,j'} \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'}.$$

Thus

$$\begin{aligned}
\text{Var}(\mathbf{1}_p' U_4) &= 4 \sum_{k=1}^K \sum_{\substack{i\in S_k, m\in S_k \\ i<m}} \sum_{i=1}^{N_i} \sum_{r=1}^{N_m} \frac{\kappa_{im}}{N_i N_m} \text{Var}(V_{irms}) \\
&= 4 \sum_{k=1}^K \sum_{\substack{i\in S_k, m\in S_k \\ i<m}} \kappa_{im} \Big(\sum_j \Omega_{ij}\Omega_{mj} + \delta_{im} \Big) \\
&= 2 \sum_{k=1}^K \sum_{\substack{i\in S_k, m\in S_k \\ i\neq m}} \sum_j \kappa_{im}\Omega_{ij}\Omega_{mj} \pm 2 \sum_k \sum_{i\neq m\in S_k} \kappa_{im} \sum_{j,j'} \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'}, \\
&= \Theta_{n3} \pm E_n.
\end{aligned} \tag{D.46}$$

which proves the lemma.

□

D.10 Proof of Lemma D.6

By assumption (21), $N_i^3/(N_i - 1) \asymp N_i$ and $\Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 \asymp \frac{1}{n_k^2 \bar{N}_k^2}$. First, observe that

$$\begin{aligned}
\Theta_{n2} + \Theta_{n4} &= 2 \sum_{k=1}^K \Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}} \Big)^2 \sum_{i\in S_k} \frac{N_i^3}{N_i - 1} \|\Omega_i\|^2 \\
&\quad + 2 \sum_{k=1}^K \sum_{\substack{i\in S_k, m\in S_k \\ i\neq m}} \sum_j \Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}} \Big)^2 N_i N_m \Omega_{ij}\Omega_{mj} \\
&\asymp \sum_{k=1} \Big(\frac{1}{n_k \bar{N}_k} \Big)^2 \sum_j \sum_{i,m\in S_k} N_i\Omega_{ij} \cdot N_m\Omega_{ij} = \sum_k \|\mu_k\|^2.
\end{aligned} \tag{D.47}$$

Recall the definitions of μ_k and μ in (D.2)-(D.3). By direct calculations, we have

$$\Theta_{n3} = 2 \sum_j \sum_{k\neq \ell} \Big(\frac{1}{n\bar{N}} \sum_{i\in S_k} N_i\Omega_{ij} \Big) \Big(\frac{1}{n\bar{N}} \sum_{m\in S_\ell} N_m\Omega_{mj} \Big)$$

$$= 2\sum_j \sum_{k\neq \ell} \frac{n_k \bar{N}_k}{n\bar{N}} \mu_{kj} \cdot \frac{n_\ell \bar{N}_\ell}{n\bar{N}} \mu_{\ell j}$$

$$= 2\sum_{k\neq \ell} \frac{n_k n_\ell \bar{N}_k \bar{N}_\ell}{n^2 \bar{N}^2} \cdot \mu_k{}' \mu_\ell$$

$$\leq 2\sum_j \Big(\sum_k \frac{n_k \bar{N}_k}{n\bar{N}} \mu_{kj} \Big)^2 = 2\sum_j \mu_j^2 = 2\|\mu\|^2. \tag{D.48}$$

By Cauchy–Schwarz,

$$\begin{aligned} \|\mu\|^2 &= \sum_j \left(\sum_k (\frac{n_k \bar{N}_k}{n\bar{N}}) \mu_{kj} \right)^2 \\ &\leq \sum_j \left(\sum_k (\frac{n_k \bar{N}_k}{n\bar{N}})^2 \right) \cdot \left(\sum_k \mu_{kj}^2 \right) \\ &\leq \sum_j \left(\sum_k (\frac{n_k \bar{N}_k}{n\bar{N}}) \right) \cdot \left(\sum_k \mu_{kj}^2 \right) = \sum_j \sum_k \mu_{kj}^2 = \sum_k \|\mu_k\|^2. \end{aligned} \tag{D.49}$$

Combining (D.47), (D.48), and (D.49) yields

$$c\big(\sum_k \|\mu_k\|^2\big) \leq \Theta_{n2} + \Theta_{n3} + \Theta_{n4} \leq C\big(\sum_k \|\mu_k\|^2\big),$$

for absolute constants $c, C > 0$. This completes the proof. □

D.11 Proof of Lemma D.7

By (21), it holds that

$$(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}})^2 \asymp \frac{1}{(n_k \bar{N}_k)^2}, \tag{D.50}$$

and moreover, for all $i \in \{1, 2, \dots, n\}$,

$$\frac{N_i^3}{N_i - 1} \asymp N_i^2. \tag{D.51}$$

Recall the definitions of A_n, B_n, and E_n in (D.8), (D.11), and (D.13), respectively. Note that these are the remainder terms in Lemmas D.3, D.4, and D.5, respectively. Under the null hypothesis (recall $\Theta_{n1} \equiv 0$ under the null),

$$\mathrm{Var}(T) = \Theta_{n2} + \Theta_{n3} + \Theta_{n4} + O(A_n + B_n + E_n). \tag{D.52}$$

It holds that

$$A_n \leq \sum_{k=1}^K \Big(\frac{1}{n_k \bar{N}_k} \Big)^2 \sum_{i \in S_k} N_i^2 \|\Omega_i\|_3^3. \tag{D.53}$$

Next, by linearity and the definition of Σ_k, Σ in (D.2), (D.3), respectively,

$$\begin{aligned} B_n &\leq 2\sum_{k,\ell} \frac{n_k n_\ell \bar{N}_k \bar{N}_\ell}{n^2 \bar{N}^2} \mathbf{1}_p'(\Sigma_k \circ \Sigma_\ell)\mathbf{1}_p \\ &\leq 2\mathbf{1}_p' \bigg(\frac{1}{n\bar{N}} \sum_k n_k \bar{N}_k \Sigma_k \bigg) \circ \bigg(\frac{1}{n\bar{N}} \sum_\ell n_\ell \bar{N}_\ell \Sigma_k \ell \bigg) \mathbf{1}_p \end{aligned}$$

$$= 2\mathbf{1}_p'(\Sigma \circ \Sigma)\mathbf{1}_p = 2\|\Sigma\|_F^2$$

By Cauchy–Schwarz,

$$\begin{aligned}
B_n \le \|\Sigma\|_F^2 &= \sum_{j,j'} \left(\sum_k (\frac{n_k \bar{N}_k}{n\bar{N}} \Sigma_k(j,j') \right)^2 \\
&\le \sum_{j,j'} \left(\sum_k (\frac{n_k \bar{N}_k}{n\bar{N}})^2 \right) \cdot \left(\sum_k \Sigma_k(j,j')^2 \right) \\
&\le \sum_{j,j'} \left(\sum_k \frac{n_k \bar{N}_k}{n\bar{N}} \right) \cdot \left(\sum_k \Sigma_k(j,j')^2 \right) = \sum_{j,j'} \sum_k \Sigma_k(j,j')^2 = \sum_k \|\Sigma_k\|_F^2. \qquad \text{(D.54)}
\end{aligned}$$

Next by the definition of Σ_k in (D.2), we have $\Sigma_k(j,j') = \frac{1}{n_k \bar{N}_k} \sum_{i \in S_k} N_i \Omega_{ij} \Omega_{ij'}$. It follows that

$$\begin{aligned}
E_n &\le \sum_k \sum_{j,j'} \Big(\frac{1}{n_k \bar{N}_k} \sum_{i \in S_k} N_i \Omega_{ij} \Omega_{ij'} \Big) \Big(\frac{1}{n_k \bar{N}_k} \sum_{m \in S_k} N_m \Omega_{mj} \Omega_{mj'} \Big) \\
&= \sum_k \sum_{j,j'} \Sigma_k^2(j,j') = \sum_k \|\Sigma_k\|_F^2. \qquad \text{(D.55)}
\end{aligned}$$

Next, Lemma D.6 implies that

$$\Theta_{n2} + \Theta_{n3} + \Theta_{n4} \asymp \sum_k \|\mu_k\|^2 = K\|\mu\|^2, \qquad \text{(D.56)}$$

where we use that the null hypothesis holds. By assumption of the lemma, we have

$$\beta_n = \frac{\max\Big\{ \sum_k \sum_{i \in S_k} \frac{N_i^2}{n_k^2 \bar{N}_k^2} \|\Omega_i\|_3^3 ,\ \sum_k \|\Sigma_k\|_F^2 \Big\}}{K\|\mu\|^2} = o(1)$$

Combining this with (D.52), (D.53), (D.54), (D.55),and (D.56) completes the proof of the first claim. The second claim follows plugging in $\mu_k = \mu$ for all $k \in \{1, 2, \dots, K\}$.

□

D.12 Proof of Lemma D.8

By assumption, $N_i^3/(N_i - 1) \asymp N_i$, $M_i^3/(M_i - 1) \asymp M_i$. By direct calculation,

$$\begin{aligned}
\Theta_{n2} + \Theta_{n4} &\asymp \Big[\frac{m\bar{M}}{(n\bar{N} + m\bar{M})n\bar{N}}\Big]^2 \sum_{i,m,j} N_i N_m \Omega_{ij} \Omega_{mj} + \Big[\frac{n\bar{N}}{(n\bar{N} + m\bar{M})m\bar{M}}\Big]^2 \sum_{i,m} N_i N_m \Gamma_{ij} \Gamma_{mj} \\
&= \frac{1}{(n\bar{N} + m\bar{M})^2} \Big((m\bar{M})^2 \|\eta\|^2 + n\bar{N}^2 \|\theta\|^2 \Big). \qquad \text{(D.57)}
\end{aligned}$$

Next

$$\begin{aligned}
\Theta_{n3} &= \frac{4}{(n\bar{N} + m\bar{M})^2} \sum_{i \in S_1} \sum_{m \in S_2} \sum_j N_i \Omega_{ij} \cdot N_m \Gamma_{mj} \\
&= \frac{4}{(n\bar{N} + m\bar{M})^2} \cdot n\bar{N} m\bar{M} \langle \theta, \eta \rangle. \qquad \text{(D.58)}
\end{aligned}$$

Combining (D.57) and (D.58) yields

$$\begin{aligned}
\Theta_{n2} + \Theta_{n3} + \Theta_{n4} &\asymp \frac{1}{(n\bar{N} + m\bar{M})^2} \big((m\bar{M})^2 \|\eta\|^2 + 2n\bar{N} m\bar{M} \langle \theta, \eta \rangle + n\bar{N}^2 \|\theta\|^2 \big) \\
&= \left\| \frac{m\bar{M}}{n\bar{N} + m\bar{M}} \eta + \frac{n\bar{N}}{n\bar{N} + m\bar{M}} \theta \right\|^2,
\end{aligned}$$

which proves the first claim. The second follows by plugging in $\theta = \eta = \mu$ under the null. □

D.13 Proof of Lemma D.9

As in (D.52), we have under the null that

$$\mathrm{Var}(T) = \Theta_{n2} + \Theta_{n3} + \Theta_{n4} + O(A_n + B_n + E_n). \tag{D.59}$$

For general K, observe that the proofs of the bounds

$$\begin{aligned} A_n &\le \sum_{k=1}^{K} \Big(\frac{1}{n_k \bar{N}_k}\Big)^2 \sum_{i \in S_k} N_i^2 \|\Omega_i\|_3^3 \\ B_n &\le \sum_{k=1}^{K} \|\Sigma_k\|_F^2 \\ E_n &\le \sum_{k=1}^{K} \|\Sigma_k\|_F^2 \end{aligned}$$

derived in (D.53), (D.54), and (D.55), only use the assumption that $N_i, M_i \ge 2$ for all i.

Translating these bounds to the notation of the $K = 2$ case, we have

$$\begin{aligned} A_n &\le \sum_i N_i^2 \|\Omega_i\|^3 + \sum_i M_i^2 \|\Gamma_i\|^3 \\ B_n &\le \|\Sigma_1\|_F^2 + \|\Sigma_2\|_F^2 \\ E_n &\le \|\Sigma_1\|_F^2 + \|\Sigma_2\|_F^2. \end{aligned} \tag{D.60}$$

Furthermore, we know that $\Theta_n \ge c\|\mu\|^2$ under the null by Lemma D.8, for an absolute constant $c > 0$. Combining this with (D.59) and (D.60) completes the proof. □

D.14 Proof of Lemma D.10

Define

$$\begin{aligned} V_1 &= 2\sum_{k=1}^{K} \sum_{i \in S_k} \sum_{j=1}^{p} \Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 \Big[\frac{N_i X_{ij}^2}{N_i - 1} - \frac{N_i X_{ij}(N_i - X_{ij})}{(N_i - 1)^2}\Big] \\ V_2 &= \frac{2}{n^2 \bar{N}^2} \sum_{1 \le k \ne \ell \le K} \sum_{i \in S_k} \sum_{m \in S_\ell} \sum_{j=1}^{p} X_{ij} X_{mj} \\ V_3 &= 2\sum_{k=1}^{K} \sum_{\substack{i \in S_k, m \in S_k, \\ i \ne m}} \sum_{j=1}^{p} \Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 X_{ij} X_{mj}. \end{aligned}$$

Observe that $V_1 + V_2 + V_3 = V$. Also define

$$A_{11} = \sum_i \sum_{r=1}^{N_i} \sum_j \Big[\frac{4\theta_i \Omega_{ij}}{N_i}\Big] Z_{ijr} \tag{D.61}$$

$$A_{12} = 2\sum_i \sum_{r=1}^{N_i} \sum_j \Big[\sum_{m \in [n] \setminus \{i\}} \alpha_{im} N_m \Omega_{mj}\Big] Z_{ijr} \tag{D.62}$$

and observe that $A_{11} + A_{12} = A_1$.

First, we derive the decomposition of V_1. Recall that

$$Y_{ij} := \frac{X_{ij}}{N_i} - \Omega_{ij} = \frac{1}{N_i}\sum_{r=1}^{N_i} Z_{ijr}, \qquad Q_{ij} := Y_{ij}^2 - \mathbb{E}Y_{ij}^2 = Y_{ij}^2 - \frac{\Omega_{ij}(1 - \Omega_{ij})}{N_i}. \tag{D.63}$$

With these notations, $X_{ij} = N_i(\Omega_{ij} + Y_{ij})$ and $N_i Y_{ij}^2 = N_i Q_{ij} + \Omega_{ij}(1 - \Omega_{ij})$.

Write

$$V_1 = 2\sum_{i=1}^{n}\sum_{i=1}^{n} \frac{\theta_i}{N_i}\Delta_{ij}, \qquad \text{where} \quad \Delta_{ij} := \frac{X_{ij}^2}{N_i} - \frac{X_{ij}(N_i - X_{ij})}{N_i(N_i-1)}. \tag{D.64}$$

Note that $X_{ij} = N_i(\Omega_{ij} + Y_{ij})$ and $Y_{ij}^2 = Q_{ij} + N_i^{-1}\Omega_{ij}(1 - \Omega_{ij})$. It follows that

$$\frac{X_{ij}^2}{N_i} = N_i\Omega_{ij}^2 + 2N_i\Omega_{ij}Y_{ij} + N_iQ_{ij} + \Omega_{ij}(1-\Omega_{ij}).$$

In (D.32), we have shown that $Q_{ij} = (1-2\Omega_{ij})\frac{Y_{ij}}{N_i} + \frac{1}{N_i^2}\sum_{1\le r\neq s\le N_i} Z_{ijr}Z_{ijs}$. It follows that

$$\frac{X_{ij}^2}{N_i} = N_i\Omega_{ij}^2 + 2N_i\Omega_{ij}Y_{ij} + (1-2\Omega_{ij})Y_{ij} + \frac{1}{N_i}\sum_{1\le r\neq s\le N_i} Z_{ijr}Z_{ijs} + \Omega_{ij}(1-\Omega_{ij}).$$

Additionally, by (D.33),

$$\frac{X_{ij}(N_{ij} - X_{ij})}{N_i(N_i-1)} = \Omega_{ij}(1-\Omega_{ij}) + (1-2\Omega_{ij})Y_{ij} - \frac{1}{N_i(N_i-1)}\sum_{1\le r\neq s\le N_i} Z_{ijr}Z_{ijs}.$$

Combining the above gives

$$\begin{aligned}\Delta_{ij} &= N_i\Omega_{ij}^2 + 2N_i\Omega_{ij}Y_{ij} + \frac{1}{N_i-1}\sum_{1\le r\neq s\le N_i} Z_{ijr}Z_{ijs} \\ &= N_i\Omega_{ij}^2 + 2\Omega_{ij}\sum_{r=1}^{N_i} Z_{ijr} + \frac{1}{N_i-1}\sum_{1\le r\neq s\le N_i} Z_{ijr}Z_{ijs}. \end{aligned} \tag{D.65}$$

Recall the definition of Θ_{n2} in (D.7), A_2 in (D.19), and A_{11} in (D.61). We have

$$\begin{aligned} V_1 &= 2\sum_{k,i\in S_k}\sum_j \frac{\theta_i}{N_i}\big[N_i\Omega_{ij}^2 + 2\Omega_{ij}\sum_{r=1}^{N_i} Z_{ijr} + \frac{1}{N_i-1}\sum_{1\le r\neq s\le N_i} Z_{ijr}Z_{ijs}\big]. \\ &= \Theta_{n2} + \sum_{k,i\in S_k}\sum_j \frac{4\theta_i\Omega_{ij}}{N_i}\sum_{r=1}^{N_i} Z_{ijr} + \sum_{k,i\in S_k}\sum_j \frac{2\theta_i}{N_i(N_i-1)}\sum_{1\le r\neq s\le N_i} Z_{ijr}Z_{ijs} \\ &= \Theta_{n2} + A_{11} + A_2 \end{aligned} \tag{D.66}$$

Next, we have

$$\begin{aligned} V_2 + V_3 &= \sum_{i\neq m}\alpha_{im}N_iN_m\sum_j\Big[(Y_{ij}+\Omega_{ij})(Y_{mj}+\Omega_{mj})\Big] \\ &= \sum_{i\neq m}\alpha_{im}N_iN_m\sum_j Y_{ij}Y_{mj} + 2\sum_{i\neq m}\alpha_{im}N_iN_m\sum_j Y_{ij}\Omega_{mj} + \sum_{i\neq m}\alpha_{im}N_iN_m\sum_j\Omega_{ij}\Omega_{mj} \\ &= \sum_{i\neq m}\sum_{r=1}^{N_i}\sum_{s=1}^{N_m}\alpha_{im}\big(\sum_j Z_{ijr}Z_{mjs}\big) + 2\sum_i\sum_{r=1}^{N_i}\sum_j\big[\sum_{m\in[n]\setminus\{i\}}\alpha_{im}N_m\Omega_{mj}\big]Z_{ijr} + \Theta_{n3} + \Theta_{n4} \\ &= A_3 + A_{12} + \Theta_{n3} + \Theta_{n4}. \end{aligned}$$

Hence

$$A_1 + A_2 + A_3 + \Theta_{n2} + \Theta_{n3} + \Theta_{n4} = V,$$

which verifies (D.21). By inspection, we also see that $\mathbb{E}A_b = 0$ for $b \in \{1,2,3\}$. That A_1, A_2, A_3 are mutually uncorrelated follows immediately from the linearity of expectation and the fact that the random variables $\{Z_{ijr}\}_{i,r} \cup \{Z_{ijr}Z_{mjs}\}_{(i,r)\neq(m,s)}$ are mutually uncorrelated.

□

D.15 Proof of Lemma D.11

Define

$$\gamma_{irj} = \frac{4\theta_i \Omega_{ij}}{N_i} + \sum_{m\in[n]\setminus\{i\}} 2\alpha_{im} N_m \Omega_{mj} \tag{D.67}$$

and recall that $A_1 = \sum_i \sum_{r\in[N_i]} \sum_j \gamma_{irj} Z_{ijr}$. First we develop a bound on γ_{irj}. Suppose that $i \in S_k$. Then we have

$$\begin{aligned}
\gamma_{irj} &\lesssim \frac{N_i\Omega_{ij}}{n_k^2 \bar{N}_k^2} + \sum_{m\in S_k, m\neq i} \frac{N_m \Omega_{mj}}{n_k^2 \bar{N}_k^2} + \sum_{k'\in[K]\setminus\{k\}} \sum_{m\in S_{k'}} \frac{N_m\Omega_{mj}}{n^2\bar{N}^2} \\
&\lesssim \frac{\mu_{kj}}{n_k\bar{N}_k} + \frac{\mu_j}{n\bar{N}}.
\end{aligned}$$

Next using properties of the covariance matrix of a multinomial vector, we have

$$\begin{aligned}
\mathrm{Var}(A_1) &= \sum_{i,r\in[N_i]} \mathrm{Var}(\gamma'_{ir:} Z_{i:r}) = \sum_{i,r\in[N_i]} \gamma'_{ir:}\mathrm{Cov}(Z_{i:r})\gamma_{ir:} \\
&\leq \sum_{i,r\in[N_i]} \gamma'_{ir:}\mathrm{diag}(\Omega_{i:})\gamma_{ir:} = \sum_{i,r\in[N_i]} \sum_j \Omega_{ij}\gamma_{irj}^2 \\
&\lesssim \sum_{k,j} (\frac{\mu_{kj}}{n_k\bar{N}_k} + \frac{\mu_j}{n\bar{N}})^2 \sum_{i\in S_k, r\in[N_i]} \Omega_{ij} \\
&\lesssim \sum_{k,j} (\frac{\mu_{kj}}{n_k\bar{N}_k})^2 n_k\bar{N}_k\mu_{kj} + \sum_{k,j} (\frac{\mu_j}{n\bar{N}})^2 n_k\bar{N}_k\mu_{kj} \\
&= (\sum_k \frac{\|\mu_k\|_3^3}{n_k\bar{N}_k}) + \frac{\|\mu\|_3^3}{n\bar{N}} \lesssim \sum_k \frac{\|\mu_k\|_3^3}{n_k\bar{N}_k},
\end{aligned} \tag{D.68}$$

which proves the first claim. The last inequality follows because by Jensen's inequality (noting that the function $x \mapsto x^3$ is convex for $x \geq 0$),

$$\|\mu\|_3^3 = \sum_j \left(\sum_k (\frac{n_k\bar{N}_k}{n\bar{N}})\mu_{kj} \right)^3 \leq \sum_j \sum_k (\frac{n_k\bar{N}_k}{n\bar{N}})\mu_{kj}^3 \leq \sum_k \|\mu_k\|_3^3.$$

Next observe that

$$A_2 = \sum_i \sum_{r\neq s} \frac{2\theta_i}{N_i(N_i-1)} W_{irs} \tag{D.69}$$

where recall $W_{irs} = \sum_j Z_{ijr}Z_{ijs}$. Also recall that W_{irs} and $W_{i'r's'}$ are uncorrelated unless $i = i'$ and $\{r,s\} = \{r',s'\}$. By (D.42),

$$\begin{aligned}
\mathrm{Var}(A_2) &= \sum_i \sum_{r\neq s} \frac{4\theta_i^2}{N_i^2(N_i-1)^2} \mathrm{Var}(W_{irs}) \\
&\lesssim \sum_i \sum_{r\neq s} \frac{4\theta_i^2}{N_i^2(N_i-1)^2} \|\Omega_i\|^2 \\
&\lesssim \sum_k \sum_{i\in S_k} \cdot(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}})^4 \frac{N_i^6}{(N_i-1)^2} \cdot \frac{1}{N_i(N_i-1)} \|\Omega_i\|^2 \\
&\lesssim \sum_k \sum_{i\in S_k} \frac{N_i^2}{n_k^4\bar{N}_k^4} \|\Omega_i\|^2
\end{aligned} \tag{D.70}$$

Also observe that

$$\sum_k \frac{1}{n_k^4 \bar{N}_k^4} \sum_{i \in S_k} N_i^2 \|\Omega_i\|_2^2 \leq \sum_k \frac{1}{n_k^2 \bar{N}_k^2} \sum_{i,m \in S_k} \left\langle (\frac{N_i}{n_k \bar{N}_k})\Omega_i, (\frac{N_m}{n_m \bar{N}_m})\Omega_m \right\rangle$$
$$= \sum_k \frac{1}{n_k^2 \bar{N}_k^2} \|\mu_k\|^2.$$

This establishes the second claim.

Last we study A_3. Observe that

$$A_3 = \sum_{i \neq m} \sum_{r=1}^{N_i} \sum_{s=1}^{N_m} \alpha_{im} V_{irms}$$

where recall $V_{irms} = \sum_j Z_{ijr} Z_{mjs}$. Recall that V_{irms} and $V_{i'r'm's'}$ are uncorrelated unless $(r, s) = (r', s')$ and $\{i, m\} = \{i', m'\}$ .By (D.43),

$$\begin{aligned}
\mathrm{Var}(A_3) &\lesssim \sum_{i \neq m} \alpha_{im}^2 N_i N_m \sum_j \Omega_{ij} \Omega_{mj} \\
&\lesssim \sum_k \sum_{i \neq m \in S_k} \frac{1}{n_k^4 \bar{N}_k^4} \langle N_i \Omega_i, N_m \Omega_m \rangle + \sum_{k \neq \ell} \sum_{i \in S_k, m \in S_\ell} \frac{1}{n^4 \bar{N}^4} \langle N_i \Omega_i, N_m \Omega_m \rangle \\
&\lesssim \sum_k \frac{\|\mu_k\|^2}{n_k^2 \bar{N}_k^2} + \sum_{k,\ell} \frac{1}{n^4 \bar{N}^4} \langle n_k \bar{N}_k \mu_k, n_\ell \bar{N}_\ell \mu_\ell \rangle \\
&\lesssim \sum_k \frac{\|\mu_k\|^2}{n_k^2 \bar{N}_k^2} + \frac{\|\mu\|^2}{n^2 \bar{N}^2} \lesssim \sum_k \frac{\|\mu_k\|^2}{n_k^2 \bar{N}_k^2}.
\end{aligned} \tag{D.71}$$

In the last line we use that $\|\mu\|^2 \leq 2 \sum \|\mu_k\|^2$ as shown in (D.49). This proves all required claims. □

D.16 Proof of Proposition D.12

Under the null hypothesis, we have $\Theta_{n1} \equiv 0$. Thus, $\mathbb{E}V = \Theta_n$ under the null by Lemma D.10. Under (21), we have $\mathrm{Var}(T) = [1 + o(1)]\Theta_n$. Therefore,

$$\mathbb{E}V = [1 + o(1)]\mathrm{Var}(T), \tag{D.72}$$

so V is asymptotically unbiased under the null. Furthermore, by Lemma D.6, we have

$$\Theta_n \asymp K\|\mu\|^2. \tag{D.73}$$

In Lemma D.11, we showed that

$$\mathrm{Var}(A_2) \lesssim \sum_k \sum_{i \in S_k} \frac{N_i^2 \|\Omega_i\|_2^2}{n_k^4 \bar{N}_k^4}$$

We conclude by Lemma D.11 that under the null

$$\mathrm{Var}(V) \lesssim \sum_k \frac{\|\mu\|^2}{n_k^2 \bar{N}_k^2} \vee \sum_k \frac{\|\mu\|_3^3}{n_k \bar{N}_k}. \tag{D.74}$$

By Chebyshev's inequality, (D.73), (D.74), and assumption (D.22) of the theorem statement, we have

$$\frac{|V - \mathbb{E}V|}{\mathrm{Var}(T)} \asymp \frac{|V - \mathbb{E}V|}{K\|\mu\|^2} = o_{\mathbb{P}}(1).$$

Thus by (D.72),

$$\frac{V}{\mathrm{Var}(T)} = \frac{(V - \mathbb{E}V)}{\mathrm{Var}(T)} + \frac{\mathbb{E}V}{\mathrm{Var}(T)} = o_{\mathbb{P}}(1) + [1 + o(1)],$$

as desired. $\square$

D.17 Proof of Lemma D.13

By Lemmas D.1–D.5, we have

$$\mathrm{Var}(T) = \sum_{a=1}^{4} \mathrm{Var}(\mathbf{1}_p' U_a) \geq (\sum_{a=2}^{4} \Theta_{na}) - (A_n + B_n + E_n). \tag{D.75}$$

Using that $\max_i \|\Omega_i\|_\infty \leq 1 - c_0$, we have $\|\Omega_i\|^3 \leq (1 - c_0)\|\Omega_i\|^2$, which implies that

$$A_n \leq (1 - c_0)\Theta_{n2}. \tag{D.76}$$

Again using $\max_i \|\Omega_i\|_\infty \leq 1 - c_0$, as well as $\sum_{j'} \Omega_{ij'} = 1$, we have

$$\begin{aligned} B_n &= \frac{2}{n^2 \bar{N}^2} \sum_{k \neq \ell} \sum_{i \in S_k} \sum_{m \in S_\ell} \sum_{j,j'} N_i N_m \Omega_{ij} \Omega_{ij'} \Omega_{mj} \Omega_{mj'} \\ &\leq (1 - c_0) \cdot \frac{2}{n^2 \bar{N}^2} \sum_{k \neq \ell} \sum_{i \in S_k} \sum_{m \in S_\ell} \sum_{j,j'} N_i N_m \Omega_{ij} \Omega_{ij'} \Omega_{mj} \\ &= (1 - c_0) \cdot \frac{2}{n^2 \bar{N}^2} \sum_{k \neq \ell} \sum_{i \in S_k} \sum_{m \in S_\ell} \sum_{j} N_i N_m \Omega_{ij} \Omega_{mj} \\ &\leq (1 - c_0) \cdot \Theta_{n3}. \end{aligned} \tag{D.77}$$

Similarly to control E_n, we again use $\max_i \|\Omega_i\|_\infty \leq 1 - c_0$ and obtain

$$\begin{aligned} E_n &= 2 \sum_k \sum_{\substack{i \in S_k, m \in S_k, \\ i \neq m}} \sum_{1 \leq j,j' \leq p} \Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n \bar{N}} \Big)^2 N_i N_m \Omega_{ij} \Omega_{ij'} \Omega_{mj} \Omega_{mj'} \\ &\leq (1 - c_0) \cdot 2 \sum_k \sum_{\substack{i \in S_k, m \in S_k, \\ i \neq m}} \sum_{1 \leq j,j' \leq p} \Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n \bar{N}} \Big)^2 N_i N_m \Omega_{ij} \Omega_{ij'} \Omega_{mj} \\ &\leq (1 - c_0) \cdot 2 \sum_k \sum_{\substack{i \in S_k, m \in S_k, \\ i \neq m}} \sum_{1 \leq j \leq p} \Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n \bar{N}} \Big)^2 N_i N_m \Omega_{ij} \Omega_{mj} \\ &\leq (1 - c_0) \cdot \Theta_{n4}. \end{aligned} \tag{D.78}$$

Combining (D.75), (D.76), (D.77), and (D.78) finishes the proof.

$\square$

D.18 Proof of Proposition D.14

By Lemmas D.6 and D.13,

$$\mathrm{Var}(T) \gtrsim \Theta_{n2} + \Theta_{n3} + \Theta_{n4} \gtrsim \sum_k \|\mu_k\|^2. \tag{D.79}$$

By Lemma D.11,

$$\mathrm{Var}(V) \lesssim \sum_k \frac{\|\mu_k\|^2}{n_k^2 \bar{N}_k^2} \vee \sum_k \frac{\|\mu_k\|_3^3}{n_k \bar{N}_k} \tag{D.80}$$

Using a similar argument based on Chebyshev's inequality as in the proof of Proposition D.12 and applying (D.79) and (D.80), we have

$$\frac{|V - \mathbb{E}V|}{\mathrm{Var}(T)} \gtrsim \frac{|V - \mathbb{E}V|}{\sum_k \|\mu_k\|^2} = o_{\mathbb{P}}(1). \tag{D.81}$$

Next, by Lemma D.10 and (D.79),

$$\mathbb{E}V = \Theta_{n2} + \Theta_{n3} + \Theta_{n4} \lesssim \mathrm{Var}(T). \tag{D.82}$$

Combining (D.81) and (D.82) finishes the proof. □

D.19 Proof of Proposition D.17

From the proof of Lemma D.10, we have

$$V^* = V_1 = \Theta_{n2} + A_{11} + A_2,$$

and the terms on the right-hand-side are mutually uncorrelated. From (D.68), we have

$$\mathrm{Var}(A_{11}) \lesssim \sum_i \frac{\|\Omega_i\|_3^3}{N_i}$$
$$\mathrm{Var}(A_2) \lesssim \sum_i \frac{\|\Omega_i\|^2}{N_i^2}.$$

Hence

$$\mathbb{E}V^* = \Theta_{n2}$$
$$\mathrm{Var}(V^*) \lesssim \sum_i \frac{\|\Omega_i\|_3^3}{N_i} \vee \sum_i \frac{\|\Omega_i\|^2}{N_i^2}. \tag{D.83}$$

Since $K = n$ and the null hypothesis holds, we have $\Theta_{n1} \equiv \Theta_{n4} \equiv 0$. Moreover, by (D.48), we have

$$\Theta_{n3} \lesssim \|\mu\|^2 \ll \Theta_{n2} \asymp n\|\mu\|^2.$$

It follows that

$$\mathrm{Var}(T) = [1 + o(1)]\Theta_{n2} \asymp n\|\mu\|^2. \tag{D.84}$$

Thus by (D.83) and Chebyshev's inequality, we have

$$\frac{V^*}{\mathrm{Var}(T)} = \frac{V^* - \mathbb{E}V^*}{\mathrm{Var}(T)} + \frac{\mathbb{E}V^*}{\mathrm{Var}(T)} = o_{\mathbb{P}}(1) + 1 + o(1),$$

as desired.

□

D.20 Proof of Proposition D.18

By Lemmas D.6 and D.13,

$$\mathrm{Var}(T) \gtrsim \Theta_{n2} + \Theta_{n3} \gtrsim \sum_i \|\Omega_i\|^2. \tag{D.85}$$

By (D.83),

$$\mathrm{Var}(V^*) \lesssim \sum_i \frac{\|\Omega_i\|^2}{N_i^2} \vee \sum_i \frac{\|\Omega_i\|_3^3}{N_i} \tag{D.86}$$

Using a similar argument based on Chebyshev's inequality as in the proof of Proposition D.12 and applying (D.85) and (D.86), we have

$$\frac{|V^* - \mathbb{E}V^*|}{\mathrm{Var}(T)} \gtrsim \frac{|V^* - \mathbb{E}V^*|}{\sum_i \|\Omega_i\|^2} = o_{\mathbb{P}}(1). \tag{D.87}$$

Next, by Lemma D.10 and (D.85),

$$\mathbb{E}V^* = \Theta_{n2} \lesssim \mathrm{Var}(T). \tag{D.88}$$

Combining (D.81) and (D.88) finishes the proof. □

E Proofs of asymptotic normality results

The goal of this section is to prove Theorems 1 and 2. The argument relies on the martingale central limit theorem and the lemmas stated below. As a preliminary, we describe a martingale decomposition of T under the null.

Define

$$U = \mathbf{1}_p'(U_3 + U_4), \qquad \text{and} \quad S = \mathbf{1}_p' U_2.$$

By Lemma D.1, we have $T = U + S$ under the null hypothesis. It holds that

$$U = \sum_{i<i'} \sigma_{i,i'} \sum_{r=1}^{N_i} \sum_{s=1}^{N_{i'}} \Big(\sum_j Z_{ijr} Z_{i'js} \Big). \tag{E.1}$$

where we define

$$\sigma_{i,i'} = \begin{cases} 2\big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}}\big) & \text{if } i, i' \in S_k \text{ for some } k \\ -\frac{2}{n\bar{N}} & \text{else.} \end{cases}$$

Define a sequence of random variables

$$D_{\ell,s} = \sum_{i \in [\ell-1]} \sigma_{i,\ell} \sum_{r=1}^{N_i} \sum_j Z_{ijr} Z_{\ell js} \tag{E.2}$$

indexed by $(\ell, s) \in \{(i,r)\}_{1 \le i \le n, 1 \le r \le N_i}$, where these tuples are placed in lexicographical order. Precisely, we define

$$(\ell_1, s_1) \prec (\ell_2, s_2)$$

if either

- $\ell_1 < \ell_2$, or
- $\ell_1 = \ell_2$ and $s_1 < s_2$.

Observe that

$$\sum_{\ell,s} D_{\ell,s} = U.$$

Next define $\mathcal{F}_{\prec(\ell,s)}$ to be the σ-field generated by $\{Z_{i:r}\}_{(i,r)\prec(\ell,s)}$. Observe that

$$\mathbb{E}[D_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}] = 0,$$

and hence $\{D_{\ell,s}\}$ is a martingale difference sequence. Turning to S, we have

$$S = \sum_{i=1}^{n} \sigma_i \sum_{r<s} \sum_j Z_{ijr} Z_{ijs}. \tag{E.3}$$

where we define

$$\sigma_i = 2\Big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}}\Big)\frac{N_i}{N_i - 1}$$

if $i \in S_k$. Define

$$E_{\ell,s} = \sigma_\ell \sum_{r\in[s-1]} \sum_j Z_{\ell jr} Z_{\ell js}. \tag{E.4}$$

Note that $E_{\ell,1} = 0$. Order (ℓ, s) lexicographically as above, and recall that $\mathcal{F}_{\prec(\ell,s)}$ is the σ-field generated by $\{Z_{i:r}\}_{(i,r)\prec(\ell,s)}$. Observe that

$$\mathbb{E}[E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}] = 0,$$

and hence $\{E_{\ell,s}\}$ is a martingale difference sequence. We have

$$\sum_{(\ell,s)} \sigma_\ell \sum_{r\in[s-1]} \sum_j Z_{\ell jr} Z_{\ell js} = \sum_{\ell=1}^{n} \sum_{s=1}^{N_\ell} \sigma_\ell \sum_{r\in[s-1]} \sum_j Z_{\ell jr} Z_{\ell js} = S.$$

Define

$$\mathcal{M}_{\ell,s} = D_{\ell,s} + E_{\ell,s}, \quad \widetilde{\mathcal{M}}_{\ell,s} = \frac{\mathcal{M}_{\ell,s}}{\sqrt{\mathrm{Var}(T)}}. \tag{E.5}$$

Thus we obtain the martingale decomposition:

$$T = U + S = \sum_{(\ell,s)} [D_{\ell,s} + E_{\ell,s}] = \sum_{(\ell,s)} \mathcal{M}_{\ell,s}. \tag{E.6}$$

The technical results below are crucial to the proof of Theorem 1 given in Section E.1. Theorem 2 then follows easily from Theorem 1 and Theorem D.12.

Lemma E.1. *Let $\widetilde{\mathcal{M}}_{\ell,s}$ be defined as in* (E.5)*. It holds that*

$$\mathbb{E}\Big[\sum_{(\ell,s)} \mathrm{Var}\big(\widetilde{\mathcal{M}}_{\ell,s}\big|\mathcal{F}_{\prec(\ell,s)}\big)\Big] = 1.$$

Lemma E.2. *Suppose that* $\min N_i \geq 2$ *and* $\max \|\Omega_i\|_\infty \leq 1 - c_0$*. Under the null hypothesis, it holds that*

$$\mathrm{Var}\Big(\sum_{(\ell,s)} \mathrm{Var}(D_{\ell,s}|\mathcal{F}_{\prec(\ell,s)})\Big) \lesssim \big(\sum_k \frac{1}{n_k \bar{N}_k}\big)\|\mu\|_3^3 + K\|\mu\|_4^4.$$

Lemma E.3. *Suppose that* $\min N_i \geq 2$ *and* $\max \|\Omega_i\|_\infty \leq 1 - c_0$*. Under the null hypothesis, it holds that*

$$\sum_{(\ell,s)} \mathbb{E} D_{\ell,s}^4 \lesssim \big(\sum_k \frac{1}{n_k^2 \bar{N}_k^2}\big)\|\mu\|^2 + \big(\sum_k \frac{1}{n_k \bar{N}_k}\big)\|\mu\|_3^3\,,$$

Lemma E.4. *Suppose that* $\min N_i \geq 2$ *and and* $\max \|\Omega_i\|_\infty \leq 1 - c_0$*. Then we have*

$$\mathrm{Var}\Big(\sum_{(\ell,s)} \mathrm{Var}(\tilde{E}_{\ell,s}|\mathcal{F}_{\prec(\ell,s)})\Big) \lesssim \sum_k \sum_{i\in S_k} \frac{N_i^3 \|\Omega_i\|_3^3}{n_k^4 \bar{N}_k^4} \vee \sum_k \sum_{i\in S_k} \frac{N_i^4 \|\Omega_i\|_4^4}{n_k^4 \bar{N}_k^4} \tag{E.7}$$

Lemma E.5. *Suppose that* $\min N_i \geq 2$ *and and* $\max \|\Omega_i\|_\infty \leq 1 - c_0$*. Then we have*

$$\sum_{(\ell,s)} \mathbb{E}\, E_{\ell,s}^4 \lesssim \sum_k \sum_{i\in S_k} \frac{N_i^2 \|\Omega_i\|^2}{n_k^4 \bar{N}_k^4} \vee \sum_k \sum_{i\in S_k} \frac{N_i^3 \|\Omega_i\|_3^3}{n_k^4 \bar{N}_k^4}$$

Lemma E.6. *Under either the null or alternative, it holds that*

$$\sum_k \sum_{i\in S_k} \frac{N_i^2 \|\Omega_i\|^2}{n_k^4 \bar{N}_k^4} \leq \sum_k \frac{1}{n_k^2 \bar{N}_k^2}\|\mu_k\|^2$$
$$\sum_k \sum_{i\in S_k} \frac{N_i^3 \|\Omega_i\|_3^3}{n_k^4 \bar{N}_k^4} \leq \sum_k \frac{1}{n_k \bar{N}_k}\|\mu_k\|_3^3$$
$$\sum_k \sum_{i\in S_k} \frac{N_i^4 \|\Omega_i\|_4^4}{n_k^4 \bar{N}_k^4} \leq \sum_k \|\mu_k\|_4^4$$

E.1 Proof of Theorem 1

By the martingale central limit theorem (see e.g. Hall and Heyde [2014]), we have that $T/\sqrt{\mathrm{Var}(T)} \Rightarrow N(0,1)$ if the following conditions are satisfied:

$$\sum_{(\ell,s)} \mathrm{Var}\big(\widetilde{\mathcal{M}}_{\ell,s}\big|\mathcal{F}_{\prec(\ell,s)}\big) \xrightarrow{\mathbb{P}} 1 \tag{E.8}$$

$$\sum_{(\ell,s)} \mathbb{E}\big[\widetilde{\mathcal{M}}_{\ell,s}^2 \mathbf{1}_{|\widetilde{\mathcal{M}}_{\ell,s}|>\varepsilon}\big|\mathcal{F}_{\prec(\ell,s)}\big] \xrightarrow{\mathbb{P}} 0, \quad \text{for any } \varepsilon > 0. \tag{E.9}$$

It is known that (E.9), which is a Lindeberg-type condition, is implied by the Lyapunov-type condition

$$\sum_{(\ell,s)} \mathbb{E}\widetilde{\mathcal{M}}_{\ell,s}^4 = o(1). \tag{E.10}$$

See e.g. Jin et al. [2018].

Since (21) holds,

$$\mathrm{Var}(T) \gtrsim \Theta = \Theta_{n2} + \Theta_{n3} + \Theta_{n4} \gtrsim K\|\mu\|^2. \tag{E.11}$$

Recall that

$$\widetilde{\mathcal{M}}_{\ell,s} = \frac{\mathcal{M}_{\ell,s}}{\mathrm{Var}(T)} = \frac{D_{\ell,s} + E_{\ell,s}}{\mathrm{Var}(T)},$$

Note that (E.8) holds if

$$\mathbb{E}\Big[\mathrm{Var}\big(\widetilde{\mathcal{M}}_{\ell,s}\big|\mathcal{F}_{\prec(\ell,s)}\big)\Big] \to 1, \quad \text{and} \tag{E.12}$$

$$\operatorname{Var}\Big(\operatorname{Var}\big(\widetilde{\mathcal{M}}_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}\big)\Big) \to 0. \tag{E.13}$$

Recall that (E.12) holds by Lemma E.1.

Next note that

$$\mathbb{E}(D_{\ell,s}E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) = 0,$$

by inspection of the expressions for $D_{\ell,s}$ and $E_{\ell,s}$ in (E.2) and (E.4). Therefore

$$\operatorname{Var}(\mathcal{M}_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) = \operatorname{Var}(D_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) + \operatorname{Var}(E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}).$$

Hence by (E.11); Lemmas E.2, E.4 , and E.6; and the assumption (24), under the null hypothesis, we have

$$\begin{aligned}\operatorname{Var}\Big(\operatorname{Var}\big(\widetilde{\mathcal{M}}_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}\big)\Big) &\le \frac{1}{\operatorname{Var}(T)^2}\Big[\operatorname{Var}\Big(\operatorname{Var}\big(D_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}\big)\Big) + \operatorname{Var}\Big(\operatorname{Var}\big(E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}\big)\Big)\Big] \\ &\lesssim \frac{1}{K^2\|\mu\|^4}\Big[(\sum_k \frac{1}{n_k\bar{N}_k})\|\mu\|_3^3 + K\|\mu\|_4^4)\|\mu\|^2\Big] = o(1).\end{aligned}$$

This proves (E.13). Thus, (E.12) and (E.13) are established, which proves (E.8).

Similarly, (E.10) (and thus (E.9)) holds by (E.11); Lemmas (E.3), (E.5), and (E.6), and the assumption (24). Combining (E.8) and (E.9) verifies the conditions of the martingale central limit theorem, so we conclude that $T/\sqrt{\operatorname{Var}(T)} \Rightarrow N(0,1)$. Since $\operatorname{Var}(T) = [1+o(1)]\Theta_n$ by (24) and Lemma D.7, the proof is complete. □

We record a useful proposition that records the weaker conditions under which $T/\sqrt{\operatorname{Var}(T)}$ is asymptotically normal.

Proposition E.7. *Recall that α_n is defined as*

$$\alpha_n := \max\left\{\sum_{k=1}^K \frac{\|\mu_k\|_3^3}{n_k\bar{N}_k}, \quad \sum_{k=1}^K \frac{\|\mu_k\|^2}{n_k^2\bar{N}_k^2}\right\} \Big/ \Big(\sum_{k=1}^K \|\mu_k\|^2\Big)^2 \tag{E.14}$$

in (22)*. If under the null hypothesis,*

$$\alpha_n = \max\left\{\sum_{k=1}^K \frac{\|\mu_k\|_3^3}{n_k\bar{N}_k}, \quad \sum_{k=1}^K \frac{\|\mu_k\|^2}{n_k^2\bar{N}_k^2}\right\} \Big/ \Big(K\|\mu\|^2\Big)^2 \to 0, \qquad \text{and} \quad \frac{\|\mu\|_4^4}{K\|\mu\|^4} \to 0, \tag{E.15}$$

then $T/\sqrt{\operatorname{Var}(T)} \Rightarrow N(0,1)$.

E.2 Proof of Theorem 2

By our assumptions, Proposition D.12 holds, and $V/\operatorname{Var}(T) \to 1$. Thus the variance estimate V is consistent under the null. Theorem 2 then follows from Slutsky's theorem and Theorem 1. □

E.3 Proof of Lemma E.1

By Lemma D.1, S and U are uncorrelated, and it holds that

$$\operatorname{Var}(T) = \operatorname{Var}(S) + \operatorname{Var}(U). \tag{E.16}$$

Next note that

$$\mathbb{E}(D_{\ell,s}E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) = 0,$$

by inspection of the expressions for $D_{\ell,s}$ and $E_{\ell,s}$ in (E.2) and (E.4). Therefore

$$\mathrm{Var}(\mathcal{M}_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) = \mathrm{Var}(D_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) + \mathrm{Var}(E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}).$$

Observe that

$$\begin{aligned}
\mathbb{E}\Big[\sum_{(\ell,s)} \mathrm{Var}(E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)})\Big] &= \sum_{(\ell,s)} \mathbb{E}E_{\ell,s}^2 = \sum_{(\ell,s)} \sigma_\ell^2 \sum_{r,r'\in[s-1]} \sum_{j,j'} \mathbb{E}[Z_{\ell jr}Z_{\ell js}Z_{\ell j'r'}Z_{\ell j's}] \\
&= \sum_{(\ell,s)} \sigma_\ell^2 \sum_{r\in[s-1]} \sum_{j,j'} \mathbb{E}[Z_{\ell jr}Z_{\ell j'r}Z_{\ell js}Z_{\ell j's}] \\
&= \sum_{\ell=1}^{n} \sigma_\ell^2 \sum_{s\in[N_\ell]} \sum_{r\in[s-1]} \mathbb{E}\big(\sum_j Z_{\ell jr}Z_{\ell js}\big)^2 \\
&= \mathrm{Var}(S).
\end{aligned} \tag{E.17}$$

The last line is obtained noting that S as defined in (E.3) is a sum of uncorrelated terms over (i, r, s).

Similarly, we have

$$\begin{aligned}
\mathbb{E}\Big[\sum_{(\ell,s)} \mathrm{Var}(D_{\ell,s}|\mathcal{F}_{\prec(\ell,s)})\Big] &= \mathbb{E}\Big[\sum_{(\ell,s)} \mathbb{E}\big[D_{\ell,s}^2|\mathcal{F}_{\prec(\ell,s)}\big]\Big] = \sum_{(\ell,s)} \mathbb{E}\big[D_{\ell,s}^2\big] \\
&= \sum_{(\ell,s)} \sum_{i\in[\ell-1]} \sigma_{i,\ell}^2 \mathrm{Var}\big(\sum_{r=1}^{N_i} \sum_j Z_{ijr}Z_{\ell js}\big) \\
&= \sum_{\ell} \sum_{i\in[\ell-1]} \sigma_{i,\ell}^2 \mathrm{Var}\big(\sum_{r=1}^{N_i} \sum_{s=1}^{N_\ell} Z_{ijr}Z_{\ell js}\big) \\
&= \mathrm{Var}(U).
\end{aligned} \tag{E.18}$$

The lemma follows by combining (E.16)–(E.18). □

E.4 Proof of Lemma E.2

Let $M_k = n_k\bar{N}_k$ and $M = n\bar{N}$. Define

$$\Sigma = \frac{1}{M}\sum_k M_k\Sigma_k = \frac{1}{M}\sum_{\ell\in[n]} N_\ell \Omega_{\ell j_1}\Omega_{\ell j_2}. \tag{E.19}$$

Our main goal is to control the conditional variance process. Define

$$\delta_{jj'\ell} = \mathbb{E}Z_{\ell jr}Z_{\ell j'r} = \begin{cases} \Omega_{\ell j}(1-\Omega_{\ell j}) & \text{if } j = j' \\ -\Omega_{\ell j}\Omega_{\ell j'} & \text{else.} \end{cases} \tag{E.20}$$

Observe that

$$\begin{aligned}
\mathrm{Var}(D_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) &= \mathbb{E}[\sum_{i,i'\in[\ell-1]} \sum_{r,r'} \sum_{j_1,j_2} \sigma_{i\ell}\sigma_{i'\ell} Z_{ij_1r}Z_{\ell j_1 s}Z_{i'j_2r'}Z_{\ell j_2 s}|\mathcal{F}_{\prec(\ell,s)}] \\
&= \sum_{i,i'\in[\ell-1]} \sum_{r,r'} \sum_{j_1,j_2} \sigma_{i\ell}\sigma_{i'\ell} Z_{ij_1r}Z_{i'j_2r'}\mathbb{E}[Z_{\ell j_1 s}Z_{\ell j_2 s}] \\
&= \sum_{i,i'\in[\ell-1]} \sum_{r,r'} \sigma_{i\ell}\sigma_{i'\ell} \sum_{j_1,j_2} \delta_{j_1j_2\ell} Z_{ij_1r}Z_{i'j_2r'}
\end{aligned}$$

Define

$$\alpha_{ii'j_1j_2} = \sum_{\ell > i'} N_\ell \sigma_{i\ell}\sigma_{i'\ell}\delta_{j_1j_2\ell}. \tag{E.21}$$

Thus

$$\begin{aligned}
\sum_{(\ell,s)} \mathrm{Var}(D_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) &= \sum_{\ell,s} \sum_{i,i'\in[\ell-1]} \sum_{r=1}^{N_i} \sum_{r'=1}^{N_{i'}} \sigma_{i\ell}\sigma_{i'\ell} \sum_{j_1,j_2} \delta_{j_1j_2\ell}\, Z_{ij_1r} Z_{i'j_2r'} \\
&= \sum_i \sum_{r=1}^{N_i} \sum_{r'=1}^{N_i} \sum_{j_1,j_2} \Big(\sum_{\ell>i} N_\ell \sigma_{i\ell}^2 \delta_{j_1j_2\ell} \Big) Z_{ij_1r} Z_{i'j_2r'} \\
&\quad + 2 \sum_{i<i'} \sum_{r=1}^{N_i} \sum_{r'=1}^{N_{i'}} \sum_{j_1,j_2} \Big(\sum_{\ell>i'} N_\ell \sigma_{i\ell}\sigma_{i'\ell} \delta_{j_1j_2\ell} \Big) Z_{ij_1r} Z_{i'j_2r'} \\
&= \sum_i \sum_{r=1}^{N_i} \sum_{r'=1}^{N_i} \sum_{j_1,j_2} \alpha_{iij_1j_2} Z_{ij_1r} Z_{i'j_2r'} \\
&\quad + 2 \sum_{i<i'} \sum_{r=1}^{N_i} \sum_{r'=1}^{N_{i'}} \sum_{j_1,j_2} \alpha_{ii'j_1j_2} Z_{ij_1r} Z_{i'j_2r'}.
\end{aligned}$$

Define

$$\zeta_{iri'r'} = \sum_{j_1,j_2} \alpha_{ii'j_1j_2} Z_{ij_1r} Z_{i'j_2r'}. \tag{E.22}$$

Then

$$\begin{aligned}
\sum_{(\ell,s)} \mathrm{Var}(D_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) &= \sum_i \sum_{r\in[N_i]} \zeta_{irir} + \Big(2 \sum_i \sum_{r<r'\in[N_i]} \zeta_{irir'} + 2 \sum_{i<i'} \sum_{r=1}^{N_i} \sum_{r'=1}^{N_{i'}} \zeta_{iri'r'} \Big) \\
&=: V_1 + V_2
\end{aligned}$$

With this decomposition, Lemma E.2 follows directly from Lemmas E.8 and E.9 stated below and proved in the next remainder of this subsection.

Lemma E.8. *It holds that*

$$\mathrm{Var}(V_1) \lesssim \big(\sum_k \frac{1}{M_k}\big)\|\mu\|_3^3.$$

Lemma E.9. *It holds that*

$$\mathrm{Var}(V_2) \lesssim K\|\mu\|_4^4$$

□

E.4.1 Statement and proof of Lemma E.10

The proofs of Lemmas E.8 and E.9 heavily rely on the following intermediate result that bounds the coefficients $\alpha_{ii'j_1j_2}$ in all cases.

Lemma E.10. *It holds that*

$$\alpha_{ii'j_1j_2} \lesssim \begin{cases} \frac{1}{M_k}\mu_{j_1} & \text{if } i,i' \in S_k, j_1 = j_2 \\ \frac{1}{M_k}\Sigma_{kj_1j_2} + \frac{1}{M}\Sigma_{j_1j_2} & \text{if } i,i' \in S_k, j_1 \neq j_2 \\ \frac{1}{M}\mu_{j_1} & \text{if } i \in S_{k_1}, i' \in S_{k_2}, k_1 \neq k_2, j_1 = j_2 \\ \frac{1}{M}\sum_{a=1}^2 \Sigma_{k_aj_1j_2} + \frac{1}{M}\Sigma_{j_1j_2} & \text{if } i \in S_{k_1}, i' \in S_{k_2}, k_1 \neq k_2, j_1 \neq j_2 \end{cases}$$

Proof. If $j_1 = j_2$ and $i, i' \in S_k$, we have

$$|\alpha_{ii'j_1j_1}| = |\sum_{\ell > i'} N_\ell \sigma_{i\ell}\sigma_{i'\ell}\delta_{j_1j_1\ell}| \le \sum_{k'=1}^{K}\sum_{\ell\in S_{k'}} N_\ell \sigma_{i\ell}\sigma_{i'\ell}\delta_{j_1j_1\ell}$$
$$\lesssim \frac{1}{M_k}\cdot\frac{1}{M_k}\sum_{\ell\in S_k} N_\ell\Omega_{\ell j_1} + \frac{1}{M}\cdot\frac{1}{M}\sum_{\ell\in[n]} N_\ell\Omega_{\ell j_1} \lesssim \frac{1}{M_k}\mu_{j_1} + \frac{1}{M}\mu_{j_1} \lesssim \frac{1}{M_k}\mu_{j_1}.$$

If $j_1 \neq j_2$ and $i, i' \in S_k$, we have

$$|\alpha_{ii'j_1j_2}| = |\sum_{\ell > i'} N_\ell \sigma_{i\ell}\sigma_{i'\ell}\delta_{j_1j_2\ell}| \le \sum_{\ell\in[n]} N_\ell|\sigma_{i\ell}\sigma_{i'\ell}|\Omega_{\ell j_1}\Omega_{\ell j_2}$$
$$\lesssim \frac{1}{M_k}\cdot\frac{1}{M_k}\sum_{\ell\in S_k} N_\ell\Omega_{\ell j_1}\Omega_{\ell j_2} + \frac{1}{M}\cdot\frac{1}{M}\sum_{\ell\in[n]} N_\ell\Omega_{\ell j_1}\Omega_{\ell j_2} \lesssim \frac{1}{M_k}\Sigma_{kj_1j_2} + \frac{1}{M}\Sigma_{j_1j_2}.$$

If $i \neq i'$, $j_1 = j_2$, and $i \in S_{k_1}, i' \in S_{k_2}$ where $k_1 \neq k_2$, we have

$$|\alpha_{ii'j_1j_1}| = |\sum_{\ell > i'} N_\ell \sigma_{i\ell}\sigma_{i'\ell}\delta_{j_1j_1\ell}| \le \sum_{\ell} N_\ell|\sigma_{i\ell}\sigma_{i'\ell}|\Omega_{\ell j_1}$$
$$\lesssim \frac{1}{M}\cdot\sum_{a=1}^{2}\frac{1}{M_{k_a}}\sum_{\ell\in S_{k_a}} N_\ell\Omega_{\ell j_1} + \frac{1}{M}\cdot\frac{1}{M}\sum_{\ell\in[n]} N_\ell\Omega_{\ell j_1} = \frac{3}{M}\mu_{j_1}.$$

If $i \neq i'$, $j_1 \neq j_2$, and $i \in S_{k_1}, i' \in S_{k_2}$ where $k_1 \neq k_2$, we have

$$|\alpha_{ii'j_1j_2}| = |\sum_{\ell > i'} N_\ell \sigma_{i\ell}\sigma_{i'\ell}\delta_{j_1j_2\ell}| \lesssim \sum_{\ell} N_\ell\sigma_{i\ell}\sigma_{i'\ell}\Omega_{\ell j_1}\Omega_{\ell j_2}$$
$$\lesssim \frac{1}{M}\cdot\sum_{a=1}^{2}\frac{1}{M_{k_a}}\sum_{\ell\in S_{k_a}} N_\ell\Omega_{\ell j_1}\Omega_{\ell j_2} + \frac{1}{M}\cdot\frac{1}{M}\sum_{\ell\in[n]} N_\ell\Omega_{\ell j_1}\Omega_{\ell j_2}$$
$$\le \frac{1}{M}\sum_{a=1}^{2}\Sigma_{k_aj_1j_2} + \frac{1}{M}\Sigma_{j_1j_2}.$$

□

E.4.2 Proof of Lemma E.8

We have

$$\mathrm{Var}(V_1) = \sum_{i,r}\mathbb{E}\zeta_{irir}^2.$$

Next by symmetry,

$$\mathbb{E}\zeta_{irir}^2 = \sum_{j_1,j_2,j_3,j_4}\alpha_{iij_1j_2}\alpha_{iij_3j_4}\,\mathbb{E}Z_{ij_1r}Z_{ij_3r}Z_{ij_2r}Z_{ij_4r}$$
$$\lesssim \sum_{j_1}\alpha_{iij_1j_1}^2\Omega_{ij_1} + \sum_{j_1\neq j_4}\alpha_{iij_1j_1}\alpha_{iij_1j_4}\,\Omega_{ij_1}\Omega_{ij_4}$$
$$+ \sum_{j_1\neq j_3}\alpha_{iij_1j_1}\alpha_{iij_3j_3}\,\Omega_{ij_1}\Omega_{ij_3} + \sum_{j_1\neq j_2}\alpha_{iij_1j_2}^2\,\Omega_{ij_1}\Omega_{ij_2}$$
$$+ \sum_{j_1,j_3,j_4(dist.)}\alpha_{iij_1j_1}\alpha_{iij_3j_4}\,\Omega_{ij_1}\Omega_{ij_3}\Omega_{ij_4} + \sum_{j_1,j_2,j_4(dist.)}\alpha_{iij_1j_2}\alpha_{iij_1j_4}\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_4}$$
$$+ \sum_{j_1,j_2,j_3,j_4(dist.)}\alpha_{iij_1j_2}\alpha_{iij_3j_4}\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_3}\Omega_{ij_4} =: \sum_{a=1}^{7}B_{a,i,r}$$

Thus

$$\mathrm{Var}(V_1) \lesssim \sum_a \underbrace{\Big(\sum_{i,r} B_{a,i,r} \Big)}_{=:B_a}.$$

We analyze B_1– B_7 separately, bounding the $\alpha_{ii'j_rj_s}$ coefficients using Lemma E.10.

For B_1,

$$\begin{aligned} B_1 &\lesssim \sum_{i,r} \sum_{j_1} \alpha^2_{iij_1j_2} \Omega_{ij_1} \lesssim \sum_{k=1}^{k} \sum_{i \in S_k} \sum_{r \in [N_i]} \sum_{j_1} (\frac{1}{M_k}\mu_{j_1})^2 \Omega_{ij_1} \\ &\lesssim \sum_k \sum_{j_1} (\frac{1}{M_k}\mu_{j_1})^2 M_k \mu_{j_1} \lesssim \big(\sum_k \frac{1}{M_k}\big)\|\mu\|_3^3. \end{aligned} \tag{E.23}$$

For B_2,

$$\begin{aligned} B_2 &\lesssim \sum_{i,r} \sum_{j_1 \neq j_4} \alpha_{iij_1j_1}\alpha_{iij_1j_4} \Omega_{ij_1}\Omega_{ij_4} \\ &\lesssim \sum_k \sum_{i \in S_k} \sum_{r \in [N_i]} \sum_{j_1 \neq j_4} \frac{1}{M_k}\mu_{j_1} \cdot \big(\frac{1}{M_k}\Sigma_{kj_1j_4} + \frac{1}{M}\Sigma_{j_1j_4}\big) \cdot \Omega_{ij_1}\Omega_{ij_4} \\ &\lesssim \sum_k \sum_{j_1 \neq j_4} \frac{1}{M_k}\mu_{j_1} \cdot \big(\frac{1}{M_k}\Sigma_{kj_1j_4} + \frac{1}{M}\Sigma_{j_1j_4}\big) \cdot M_k \Sigma_{kj_1j_4} \\ &\lesssim \sum_k \frac{1}{M_k} \sum_{j_1 \neq j_4} \Sigma^2_{kj_1j_4}\mu_{j_1} + \sum_k \frac{1}{M} \sum_{j_1 \neq j_4} \Sigma_{kj_1j_4}\Sigma_{j_1j_4}\mu_{j_1} \\ &\lesssim \sum_k \frac{\mathbf{1}'\Sigma_k^{\circ 2}\mu}{M_k} + \sum_k \frac{\mathbf{1}'(\Sigma_k \circ \Sigma)\mu}{M} = \sum_k \frac{\mathbf{1}'\Sigma_k^{\circ 2}\mu}{M_k} \end{aligned}$$

Next,

$$\begin{aligned} \sum_{j_1 \neq j_4} \Sigma^2_{kj_1j_4}\mu_{j_1} &= \sum_{j_1 \neq j_4} \frac{1}{M_k^2} \sum_{i,i' \in S_k} N_i N_{i'} \Omega_{ij_1}\Omega_{i'j_1}\Omega_{ij_4}\Omega_{i'j_4} \cdot \mu_{j_1} \\ &\leq \sum_{j_1} \frac{1}{M_k^2} \sum_{i,i' \in S_k} N_i N_{i'} \Omega_{ij_1}\Omega_{i'j_1}\mu_{j_1} \cdot \big(\sum_{j_4} \Omega_{ij_4}\Omega_{i'j_4}\big) \\ &\leq \sum_{j_1} \frac{1}{M_k^2} \sum_{i,i' \in S_k} N_i N_{i'} \Omega_{ij_1}\Omega_{i'j_1} \cdot \mu_{j_1} \\ &\leq \sum_{j_1} \mu^3_{j_1} = \|\mu\|_3^3, \end{aligned} \tag{E.24}$$

and similarly

$$\begin{aligned} \sum_{j_1 \neq j_4} \Sigma_{kj_1j_4}\Sigma_{j_1j_4}\mu_{j_1} &= \sum_{j_1 \neq j_4} \frac{1}{M_k M} \sum_{i \in S_k, i' \in [n]} N_i N_{i'} \Omega_{ij_1}\Omega_{i'j_1}\Omega_{ij_4}\Omega_{i'j_4} \cdot \mu_{j_1} \\ &\leq \sum_{j_1} \frac{1}{M_k M} \sum_{i \in S_k, i' \in [n]} N_i N_{i'} \Omega_{ij_1}\Omega_{i'j_1}\mu_{j_1} \\ &= \sum_{j_1} \mu^3_{j_1} = \|\mu\|_3^3. \end{aligned}$$

Thus

$$B_2 \lesssim \big(\sum_k \frac{1}{M_k}\big)\|\mu\|_3^3. \tag{E.25}$$

For B_3,

$$B_3 \lesssim \sum_{i,r}\sum_{j_1\neq j_3} \alpha_{iij_1j_1}\alpha_{iij_3j_3}\,\Omega_{ij_1}\Omega_{ij_3}$$
$$\lesssim \sum_k \sum_{i\in S_k}\sum_{r\in[N_i]}\sum_{j_1\neq j_3} \frac{1}{M_k}\mu_{j_1}\cdot\frac{1}{M_k}\mu_{j_3}\cdot\Omega_{ij_1}\Omega_{ij_3}$$
$$\lesssim \sum_k\sum_{j_1\neq j_3}\frac{1}{M_k}\mu_{j_1}\cdot\frac{1}{M_k}\mu_{j_3}\cdot M_k\Sigma_{kj_1j_3} \lesssim \sum_k \frac{\mu'\Sigma_k\mu}{M_k}.$$

We have by Cauchy-Schwarz,

$$\begin{aligned}\mu'\Sigma_k\mu &= \frac{1}{M_k}\sum_{i\in S_k} N_i\mu'\Omega_i\Omega_{i'}'\mu \\ &= \frac{1}{M_k}\sum_{i\in S_k} N_i\big(\sum_j \mu_j\Omega_{ij}\big)^2 \\ &\le \frac{1}{M_k}\sum_{i\in S_k} N_i\big(\sum_j \Omega_{ij}\big)\big(\sum_j \mu_j^2\Omega_{ij}\big) \\ &= \sum_j \mu_j^3 = \|\mu\|_3^3. \end{aligned} \tag{E.26}$$

Thus

$$B_3 \lesssim \big(\sum_k \frac{1}{M_k}\big)\|\mu\|_3^3 \tag{E.27}$$

For B_4,

$$B_4 \lesssim \sum_{i,r}\sum_{j_1\neq j_2}\alpha^2_{iij_1j_2}\,\Omega_{ij_1}\Omega_{ij_2} \lesssim \sum_k\sum_{i\in S_k}\sum_{r\in[N_i]}\sum_{j_1\neq j_2}\big(\frac{1}{M_k}\Sigma_{kj_1j_2}+\frac{1}{M}\Sigma_{j_1j_2}\big)^2\,\Omega_{ij_1}\Omega_{ij_2}$$
$$\lesssim \sum_k\sum_{j_1\neq j_2}\big(\frac{1}{M_k}\Sigma_{kj_1j_2}+\frac{1}{M}\Sigma_{j_1j_2}\big)^2\cdot M_k\Sigma_{kj_1j_2} \lesssim \sum_k \frac{\mathbf{1}'(\Sigma_k^{\circ 3})\mathbf{1}}{M_k} + \sum_k \frac{M_k}{M^2}\mathbf{1}'(\Sigma_k\circ\Sigma^{\circ 2})\mathbf{1}$$
$$\lesssim \big(\sum_k \frac{\mathbf{1}'(\Sigma_k^{\circ 3})\mathbf{1}}{M_k}\big) + \frac{1}{M}\mathbf{1}'(\Sigma^{\circ 3})\mathbf{1}.$$

First,

$$\begin{aligned}\mathbf{1}'(\Sigma_k^{\circ 3})\mathbf{1} &= \frac{1}{M_k^3}\sum_{i_1,i_2,i_3\in S_k} N_{i_1}N_{i_2}N_{i_3}\big(\sum_j \Omega_{i_1j}\Omega_{i_2j}\Omega_{i_3j}\big)^2 \\ &\le \frac{1}{M_k^3}\sum_{i_1,i_2,i_3\in S_k} N_{i_1}N_{i_2}N_{i_3}\cdot\sum_j \Omega_{i_1j}\Omega_{i_2j}\Omega_{i_3j} = \sum_j \mu_j^3 = \|\mu\|_3^3,\end{aligned}$$

and similarly,

$$\mathbf{1}'(\Sigma^{\circ 3})\mathbf{1} = \frac{1}{M^3}\sum_{i_1,i_2,i_3\in[n]} N_{i_1}N_{i_2}N_{i_3}\big(\sum_j \Omega_{i_1j}\Omega_{i_2j}\Omega_{i_3j}\big)^2 \le \|\mu\|_3^3.$$

Thus

$$B_4 \lesssim \big(\sum_k \frac{1}{M_k}\big)\|\mu\|_3^3 \tag{E.28}$$

For B_5,

$$
\begin{aligned}
B_5 &\lesssim \sum_{i,r} \sum_{j_1,j_3,j_4(dist.)} \alpha_{iij_1j_1}\alpha_{iij_3j_4}\,\Omega_{ij_1}\Omega_{ij_3}\Omega_{ij_4} \\
&\lesssim \sum_k \sum_{i\in S_k} N_i \sum_{j_1,j_3,j_4} \frac{1}{M_k}\mu_{j_1}\cdot\big(\frac{1}{M_k}\Sigma_{kj_3j_4}+\frac{1}{M}\Sigma_{j_3j_4}\big)\cdot\,\Omega_{ij_1}\Omega_{ij_3}\Omega_{ij_4} \\
&\lesssim \sum_k \sum_{i\in S_k} \sum_{j_1,j_3,j_4} \frac{N_i\mu_{j_1}\Sigma_{kj_3j_4}\Omega_{ij_1}\Omega_{ij_3}\Omega_{ij_4}}{M_k^2} + \sum_k \sum_{i\in S_k} \sum_{j_1,j_3,j_4} \frac{N_i\mu_{j_1}\Sigma_{j_3j_4}\Omega_{ij_1}\Omega_{ij_3}\Omega_{ij_4}}{M_kM} \\
&=: B_{51}+B_{52}.
\end{aligned}
$$

We have

$$
\begin{aligned}
B_{51} &= \sum_k \frac{1}{M_k^3} \sum_{i_1,i_2\in S_k} \sum_{j_1,j_3,j_4} N_{i_1}N_{i_2}\mu_{j_1}\Omega_{i_1j_1}\Omega_{i_1j_3}\Omega_{i_2j_3}\Omega_{i_1j_4}\Omega_{i_2j_4} \\
&= \sum_k \frac{1}{M_k^3} \sum_{i_1,i_2\in S_k} N_{i_1}N_{i_2}(\Omega_{i_1}'\mu)\cdot(\Omega_{i_1}'\Omega_{i_2})^2 \\
&\le \sum_k \frac{1}{M_k^3} \sum_{i_1,i_2\in S_k} N_{i_1}N_{i_2}\cdot\Omega_{i_1}'\mu\cdot\Omega_{i_1}'\Omega_{i_2} \\
&= \sum_k \frac{1}{M_k^2} \sum_{i_1} N_{i_1}\mu'\Omega_{i_1}\Omega_{i_1}'\mu = \frac{1}{M_k}\mu'\Sigma_k\mu \le \sum_k \frac{1}{M_k}\|\mu\|_3^3.
\end{aligned} \tag{E.29}
$$

In the last line we apply (E.26). Similarly,

$$
\begin{aligned}
B_{52} &= \sum_k \frac{1}{M_kM^2} \sum_{i_1\in S_k, i_2\in[n]} \sum_{j_1,j_3,j_4} N_{i_1}N_{i_2}\mu_{j_1}\Omega_{i_1j_1}\Omega_{i_1j_3}\Omega_{i_2j_3}\Omega_{i_1j_4}\Omega_{i_2j_4} \\
&\le \sum_k \frac{1}{M_kM^2} \sum_{i_1\in S_k, i_2\in[n]} N_{i_1}N_{i_2}\cdot\Omega_{i_1}'\mu\cdot\Omega_{i_1}'\Omega_{i_2} \\
&\le \sum_k \frac{1}{M_kM} \sum_{i_1\in S_k} N_{i_1}\mu'\Omega_{i_1}\Omega_{i_1}'\mu \le \sum_k \frac{1}{M}\|\mu\|_3^3.
\end{aligned} \tag{E.30}
$$

Thus

$$
B_5 \lesssim \big(\sum_k \frac{1}{M_k}\big)\|\mu\|_3^3. \tag{E.31}
$$

For B_6,

$$
\begin{aligned}
B_6 &\lesssim \sum_k \sum_{i\in S_k} \sum_{r\in[N_i]} \sum_{j_1,j_2,j_4(dist.)} \big(\frac{1}{M_k}\Sigma_{kj_1j_2}+\frac{1}{M}\Sigma_{j_1j_2}\big)\big(\frac{1}{M_k}\Sigma_{kj_1j_4}+\frac{1}{M}\Sigma_{j_1j_4}\big)\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_4} \\
&\lesssim \sum_k \sum_{i\in S_k} \sum_{r\in[N_i]} \sum_{j_1,j_2,j_4} \frac{\Sigma_{kj_1j_2}^2\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_4}}{M_k^2} + 2\sum_k \sum_{i\in S_k} \sum_{r\in[N_i]} \sum_{j_1,j_2,j_4} \frac{\Sigma_{kj_1j_2}\Sigma_{j_1j_2}\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_4}}{M_kM} \\
&\quad + \sum_k \sum_{i\in S_k} \sum_{r\in[N_i]} \sum_{j_1,j_2,j_4} \frac{\Sigma_{j_1j_2}^2\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_4}}{M^2} =: B_{61}+B_{62}+B_{63}.
\end{aligned}
$$

First,

$$
B_{61} \le \sum_k \sum_{i\in S_k} \sum_{r\in[N_i]} \sum_{j_1,j_2,j_4} \frac{\Sigma_{kj_1j_2}^2\Omega_{ij_1}}{M_k^2} = \sum_k \frac{1}{M_k}\mathbf{1}'\Sigma_k^{\circ 2}\mu \le \sum_k \frac{1}{M_k}\|\mu\|_3^3,
$$

where we applied (E.24). Similarly,

$$B_{62} \lesssim \sum_k \frac{1}{M_k}\|\mu\|_3^3, \text{ and}$$
$$B_{63} \lesssim \sum_k \frac{1}{M_k}\|\mu\|_3^3.$$

Thus

$$B_6 \lesssim \big(\sum_k \frac{1}{M_k}\big)\|\mu\|_3^3. \tag{E.32}$$

For B_7, we have

$$\begin{aligned}
B_7 \lesssim & \sum_{j_1,j_2,j_3,j_4(dist.)} \big(\frac{1}{M_k}\Sigma_{kj_1j_2} + \frac{1}{M}\Sigma_{j_1j_2}\big)\big(\frac{1}{M_k}\Sigma_{kj_3j_4} + \frac{1}{M}\Sigma_{j_3j_4}\big)\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_3}\Omega_{ij_4} \\
\lesssim & \sum_k \sum_{i\in S_k} \sum_{r\in[N_i]} \sum_{j_1,j_2,j_3,j_4} \frac{\Sigma_{kj_1j_2}\Sigma_{kj_3j_4}\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_3}\Omega_{ij_4}}{M_k^2} \\
& + 2\sum_k \sum_{i\in S_k} \sum_{r\in[N_i]} \sum_{j_1,j_2,j_3,j_4} \frac{\Sigma_{kj_1j_2}\Sigma_{j_3j_4}\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_3}\Omega_{ij_4}}{M_k M} \\
& + \sum_k \sum_{i\in S_k} \sum_{r\in[N_i]} \sum_{j_1,j_2,j_3,j_4} \frac{\Sigma_{j_1j_2}\Sigma_{j_3j_4}\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_3}\Omega_{ij_4}}{M^2} =: B_{71} + B_{72} + B_{73}.
\end{aligned}$$

Note that

$$\begin{aligned}
\Sigma_{kj_1j_2} &= \frac{1}{M_k}\sum_{i\in S_k} N_i\Omega_{ij_1}\Omega_{ij_2} \le \frac{1}{M_k}\sum_{i\in S_k} N_i\Omega_{ij_1} = \mu_{j_1}, \text{ and} \\
\Sigma_{j_1j_2} &= \frac{1}{M}\sum_{i\in[n]} N_i\Omega_{ij_1}\Omega_{ij_2} \le \frac{1}{M}\sum_{i\in[n]} N_i\Omega_{ij_1} = \mu_{j_1}.
\end{aligned} \tag{E.33}$$

Thus

$$\begin{aligned}
B_{71} &\le \sum_k \sum_{i\in S_k} \sum_{r\in[N_i]} \sum_{j_1,j_2,j_3,j_4} \frac{\mu_{j_1}\Sigma_{kj_3j_4}\Omega_{ij_1}\Omega_{ij_2}\Omega_{ij_3}\Omega_{ij_4}}{M_k^2} \\
&\le \sum_k \sum_{i\in S_k} \sum_{j_1,j_3,j_4} \frac{N_i\mu_{j_1}\Sigma_{kj_3j_4}\Omega_{ij_1}\Omega_{ij_3}\Omega_{ij_4}}{M_k^2} \le \sum_k \frac{1}{M_k}\|\mu\|_3^3
\end{aligned}$$

where we applied (E.29). Similarly,

$$B_{72} \lesssim \sum_k \frac{1}{M_k}\|\mu\|_3^3, \text{ and}$$
$$B_{73} \lesssim \sum_k \frac{1}{M_k}\|\mu\|_3^3.$$

Thus

$$B_7 \lesssim \big(\sum_k \frac{1}{M_k}\big)\|\mu\|_3^3. \tag{E.34}$$

Combining the results for B_1–B_7 concludes the proof.

□

E.4.3 Proof of Lemma E.9

We have

$$\mathrm{Var}(V_2) \lesssim 4 \sum_{(i,r)\neq(i',r')} \mathbb{E}\zeta_{iri r'}^2,$$

where $r \in [N_i]$ and $r \in [N_{i'}]$ in the summation above.

By symmetry, if $(i,r) \neq (i',r')$,

$$\begin{aligned}
\mathbb{E}\zeta_{iri'r'}^2 &= \sum_{j_1,j_2,j_3,j_4} \alpha_{ii'j_1j_2}\alpha_{ii'j_3j_4}\,\mathbb{E}Z_{ij_1r}Z_{ij_3r}\,\mathbb{E}Z_{i'j_2r'}Z_{i'j_4r'} \\
&\lesssim \sum_{j_1} \alpha_{ii'j_1j_1}^2\,\Omega_{ij_1}\Omega_{i'j_1} + \sum_{j_1\neq j_4} \alpha_{ii'j_1j_1}\alpha_{ii'j_1j_4}\,\Omega_{ij_1}\Omega_{i'j_1}\Omega_{i'j_4} \\
&\quad + \sum_{j_1\neq j_3} \alpha_{ii'j_1j_1}\alpha_{ii'j_3j_3}\,\Omega_{ij_1}\Omega_{ij_3}\Omega_{i'j_1}\Omega_{i'j_3} + \sum_{j_1\neq j_2} \alpha_{ii'j_1j_2}^2\,\Omega_{ij_1}\Omega_{i'j_2} \\
&\quad + \sum_{j_1,j_3,j_4(dist.)} \alpha_{ii'j_1j_1}\alpha_{ii'j_3j_4}\,\Omega_{ij_1}\Omega_{ij_3}\Omega_{i'j_1}\Omega_{i'j_4} + \sum_{j_1,j_2,j_4(dist.)} \alpha_{ii'j_1j_2}\alpha_{ii'j_1j_4}\,\Omega_{ij_1}\Omega_{i'j_2}\Omega_{i'j_4} \\
&\quad + \sum_{j_1,j_2,j_3,j_4(dist.)} \alpha_{ii'j_1j_2}\alpha_{ii'j_3j_4}\Omega_{ij_1}\Omega_{ij_3}\Omega_{i'j_2}\Omega_{i'j_4} =: \sum_{a}^{7} C_{a,i,r}. \qquad \text{(E.35)}
\end{aligned}$$

Thus

$$\mathrm{Var}(V_2) \lesssim \sum_{a=1}^{7} \sum_{(i,r)\neq(i',r')} C_{a,i,r} \lesssim \sum_{a=1}^{7} \underbrace{\sum_{i,i'} N_iN_{i'}C_{a,i,r}}_{=:C_a}.$$

Next we analyze $C_1,\ldots,C_7$, bounding the $\alpha_{ii'j_rj_s}$ coefficients using Lemma E.10.

For C_1,

$$\begin{aligned}
C_1 &\lesssim \sum_k \sum_{i,i'\in S_k} \sum_{j_1} N_iN_{i'}\alpha_{ii'j_1j_1}^2\Omega_{ij_1}\Omega_{i'j_1} + \sum_{k\neq k'} \sum_{i\in S_k,i'\in S_{k'}} \sum_{j_1} N_iN_{i'}\alpha_{ii'j_1j_1}^2\Omega_{ij_1}\Omega_{i'j_1} \\
&\lesssim \sum_k \sum_{i,i'\in S_k} \sum_{j_1} N_iN_{i'}(\frac{1}{M_k}\mu_{j_1})^2\Omega_{ij_1}\Omega_{ij_1} + \sum_{k\neq k'} \sum_{i\in S_k,i'\in S_{k'}} \sum_{j_1}(\frac{1}{M}\mu_{j_1})^2\Omega_{ij_1}\Omega_{i'j_1} \\
&\lesssim \sum_k \sum_{j_1} \mu_{j_1}^4 + \sum_{k\neq k'} \sum_{j_1} \frac{M_kM_{k'}}{M^2}\mu_{j_1}^4 \lesssim K\|\mu\|_4^4. \qquad \text{(E.36)}
\end{aligned}$$

For C_2,

$$\begin{aligned}
C_2 &\lesssim \sum_k \sum_{i,i'\in S_k} N_iN_{i'} \sum_{j_1\neq j_4} \alpha_{ii'j_1j_1}\alpha_{ii'j_1j_4}\,\Omega_{ij_1}\Omega_{i'j_1}\Omega_{i'j_4} \\
&\quad + \sum_{k\neq k'} \sum_{i\in S_k,i'\in S_{k'}} N_iN_{i'} \sum_{j_1\neq j_4} \alpha_{ii'j_1j_1}\alpha_{ii'j_1j_4}\,\Omega_{ij_1}\Omega_{i'j_1}\Omega_{i'j_4} \\
&\lesssim \sum_k \sum_{i,i'\in S_k} N_iN_{i'} \sum_{j_1\neq j_4} \frac{1}{M_k}\mu_{j_1}\cdot\big(\frac{1}{M_k}\Sigma_{kj_1j_4} + \frac{1}{M}\Sigma_{j_1j_4}\big)\,\Omega_{ij_1}\Omega_{i'j_1}\Omega_{i'j_4} \\
&\quad + \sum_{k\neq k'} \sum_{i\in S_k,i'\in S_{k'}} N_iN_{i'} \sum_{j_1\neq j_4} \frac{1}{M}\mu_{j_1}\cdot\big(\frac{1}{M}\sum_{a\in\{k,k'\}} \Sigma_{aj_1j_4} + \frac{1}{M}\Sigma_{j_1j_4}\big)\,\Omega_{ij_1}\Omega_{i'j_1}\Omega_{i'j_4} \\
&\lesssim \sum_k \sum_{i,i'\in S_k} N_iN_{i'} \sum_{j_1\neq j_4} \frac{1}{M_k}\mu_{j_1}\cdot\big(\frac{1}{M_k}\mu_{j_1} + \frac{1}{M}\mu_{j_1}\big)\,\Omega_{ij_1}\Omega_{i'j_1}\Omega_{i'j_4}
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{k\neq k'}\sum_{i\in S_k, i'\in S_{k'}} N_i N_{i'} \sum_{j_1\neq j_4} \frac{1}{M}\mu_{j_1}\cdot\big(\frac{2}{M}\mu_{j_1}+\frac{1}{M}\mu_{j_1}\big)\,\Omega_{ij_1}\Omega_{i'j_1}\Omega_{i'j_4} \\
&\lesssim \sum_k\sum_{j_1}\big(\mu_{j_1}^4+\frac{M_k}{M}\mu_{j_1}^4\big)+\sum_{k\neq k'}\sum_{j_1}\frac{M_kM_{k'}}{M^2}\mu_{j_1}^4 \lesssim K\|\mu\|_4^4. \qquad \text{(E.37)}
\end{aligned}
$$

where we applied (E.33).

For C_3,

$$
\begin{aligned}
C_3 &\lesssim \Big(\sum_k\sum_{i,i'\in S_k} N_iN_{i'} + \sum_{k\neq k'}\sum_{i\in S_k, i'\in S_{k'}} N_iN_{i'}\Big)\sum_{j_1\neq j_3}\alpha_{ii'j_1j_1}\alpha_{ii'j_3j_3}\,\Omega_{ij_1}\Omega_{ij_3}\Omega_{i'j_1}\Omega_{i'j_3} \\
&\lesssim \sum_k\sum_{i,i'\in S_k} N_iN_{i'}\sum_{j_1\neq j_3}\frac{1}{M_k}\mu_{j_1}\cdot\frac{1}{M_k}\mu_{j_3}\cdot\Omega_{ij_1}\Omega_{ij_3}\Omega_{i'j_1}\Omega_{i'j_3} \\
&\quad+\sum_{k\neq k'}\sum_{i\in S_k, i'\in S_{k'}} N_iN_{i'}\sum_{j_1\neq j_3}\frac{1}{M}\mu_{j_1}\cdot\frac{1}{M}\mu_{j_3}\,\Omega_{ij_1}\Omega_{ij_3}\Omega_{i'j_1}\Omega_{i'j_3} \\
&= \sum_k\sum_{j_1\neq j_3}\mu_{j_1}\mu_{j_3}\Sigma_{kj_1j_3}^2 + \sum_{k\neq k'}\sum_{j_1\neq j_3}\frac{M_kM_{k'}}{M^2}\mu_{j_1}\mu_{j_3}\Sigma_{kj_1j_3}\Sigma_{k'j_1j_3} \\
&\le \big(\sum_k \mu'\Sigma_k^{\circ 2}\mu\big)+\mu'\Sigma^{\circ 2}\mu.
\end{aligned}
$$

First, by Cauchy–Schwarz,

$$
\begin{aligned}
\mu'\Sigma_k^{\circ 2}\mu &= \frac{1}{M_k^2}\sum_{i,i'\in S_k} N_iN_{i'}\big(\sum_j \mu_j\Omega_{ij}\Omega_{i'j}\big)^2 \\
&= \frac{1}{M_k^2}\sum_{i,i'\in S_k} N_iN_{i'}\big(\sum_j \Omega_{ij}\Omega_{i'j}\big)\sum_j \mu_j^2\Omega_{ij}\Omega_{i'j} \\
&\le \frac{1}{M_k^2}\sum_{i,i'\in S_k} N_iN_{i'}\sum_j \mu_j^2\Omega_{ij}\Omega_{i'j} = \sum_j \mu_j^4 = \|\mu\|_4^4. \qquad \text{(E.38)}
\end{aligned}
$$

Similarly

$$
\mu'\Sigma^{\circ 2}\mu \lesssim \|\mu\|_4^4. \qquad \text{(E.39)}
$$

Hence

$$
C_3 \lesssim K\|\mu\|_4^4. \qquad \text{(E.40)}
$$

For C_4,

$$
\begin{aligned}
C_4 &\lesssim \Big(\sum_k\sum_{i,i'\in S_k} N_iN_{i'} + \sum_{k\neq k'}\sum_{i\in S_k, i'\in S_{k'}} N_iN_{i'}\Big)\sum_{j_1\neq j_2}\alpha_{ii'j_1j_2}^2\,\Omega_{ij_1}\Omega_{i'j_2} \\
&\lesssim \sum_k\sum_{i,i'\in S_k} N_iN_{i'}\sum_{j_1\neq j_2}\big(\frac{1}{M_k}\Sigma_{kj_1j_2}+\frac{1}{M}\Sigma_{j_1j_2}\big)^2\,\Omega_{ij_1}\Omega_{i'j_2} \\
&\quad+\sum_{k\neq k'}\sum_{i\in S_k, i'\in S_{k'}} N_iN_{i'}\sum_{j_1\neq j_2}\big(\frac{1}{M}\sum_{a\in\{k,k'\}}^{2}\Sigma_{aj_1j_2}+\frac{1}{M}\Sigma_{j_1j_2}\big)^2\,\Omega_{ij_1}\Omega_{i'j_2} \\
&\lesssim \sum_k\sum_{i,i'\in S_k} N_iN_{i'}\sum_{j_1\neq j_2}\big(\frac{1}{M_k^2}\Sigma_{kj_1j_2}^2+\frac{1}{M^2}\Sigma_{j_1j_2}^2\big)\,\Omega_{ij_1}\Omega_{i'j_2} \\
&\quad+\sum_{k\neq k'}\sum_{i\in S_k, i'\in S_{k'}} N_iN_{i'}\sum_{j_1\neq j_2}\big(\frac{1}{M^2}\sum_{a\in\{k,k'\}}^{2}\Sigma_{aj_1j_2}^2+\frac{1}{M^2}\Sigma_{j_1j_2}^2\big)\,\Omega_{ij_1}\Omega_{i'j_2} =: C_{41}+C_{42}
\end{aligned}
$$

First,

$$C_{41} \lesssim \sum_k \sum_{i,i' \in S_k} N_i N_{i'} \sum_{j_1 \neq j_2} \frac{1}{M_k^2} \Sigma_{kj_1j_2}^2 \Omega_{ij_1} \Omega_{i'j_2} + \sum_k \sum_{i,i' \in S_k} N_i N_{i'} \sum_{j_1 \neq j_2} \frac{1}{M^2} \Sigma_{j_1j_2}^2 \Omega_{ij_1} \Omega_{i'j_2}$$
$$\lesssim \sum_k \sum_{j_1 \neq j_2} \Sigma_{kj_1j_2}^2 \mu_{j_1} \mu_{j_2} + \sum_k \sum_{j_1 \neq j_2} \frac{M_k^2}{M^2} \Sigma_{j_1j_2}^2 \mu_{j_1} \mu_{j_2} \leq \sum_k \mu' \Sigma_k^{\circ 2} \mu + \sum_k \frac{M_k^2}{M^2} \mu' \Sigma^{\circ 2} \mu.$$

Similarly,

$$C_{42} \lesssim \sum_{k \neq k'} \sum_{j_1 \neq j_2} \frac{M_k M_{k'}}{M^2} \Sigma_{kj_1j_2}^2 \mu_{j_1} \mu_{j_2} + \sum_{k \neq k'} \sum_{j_1 \neq j_2} \frac{M_k M_{k'}}{M^2} \Sigma_{j_1j_2}^2 \mu_{j_1} \mu_{j_2}$$
$$\lesssim \sum_{k \neq k'} \frac{M_k M_{k'}}{M^2} \big(\mu' \Sigma_k^{\circ 2} \mu + \mu' \Sigma^{\circ 2} \mu \big)$$

Combining the previous two displays and applying (E.38) and (E.39), we have

$$C_4 \lesssim K \|\mu\|_4^4. \tag{E.41}$$

For C_5,

$$C_5 \lesssim \Bigg(\sum_k \sum_{i,i' \in S_k} N_i N_{i'} + \sum_{k \neq k'} \sum_{i \in S_k, i' \in S_{k'}} N_i N_{i'} \Bigg) \sum_{j_1, j_3, j_4 (dist.)} \alpha_{ii'j_1j_1} \alpha_{ii'j_3j_4} \Omega_{ij_1} \Omega_{ij_3} \Omega_{i'j_1} \Omega_{i'j_4}$$
$$\lesssim \sum_k \sum_{i,i' \in S_k} N_i N_{i'} \sum_{j_1, j_3, j_4} \frac{1}{M_k} \mu_{j_1} \cdot \big(\frac{1}{M_k} \Sigma_{kj_3j_4} + \frac{1}{M} \Sigma_{j_3j_4} \big) \Omega_{ij_1} \Omega_{ij_3} \Omega_{i'j_1} \Omega_{i'j_4}$$
$$+ \sum_{k \neq k'} \sum_{i \in S_k, i' \in S_{k'}} N_i N_{i'} \sum_{j_1, j_3, j_4} \frac{1}{M} \mu_{j_1} \big(\frac{1}{M} \sum_{a \in \{k, k'\}}^{2} \Sigma_{aj_3j_4} + \frac{1}{M} \Sigma_{j_3j_4} \big) \Omega_{ij_1} \Omega_{ij_3} \Omega_{i'j_1} \Omega_{i'j_4}$$
$$= \sum_k \sum_{j_1, j_3, j_4} \mu_{j_1} \Sigma_{kj_3j_4} \Sigma_{kj_1j_3} \Sigma_{kj_1j_4} + \sum_k \sum_{j_1, j_3, j_4} \frac{M_k}{M} \mu_{j_1} \Sigma_{j_3j_4} \Sigma_{kj_1j_3} \Sigma_{kj_1j_4}$$
$$+ 2 \sum_{k \neq k'} \sum_{j_1, j_3, j_4} \frac{M_k M_{k'}}{M^2} \mu_{j_1} \Sigma_{kj_3j_4} \Sigma_{kj_1j_3} \Sigma_{k'j_1j_4} + \sum_{k \neq k'} \sum_{j_1, j_3, j_4} \frac{M_k M_{k'}}{M^2} \mu_{j_1} \Sigma_{j_3j_4} \Sigma_{kj_1j_3} \Sigma_{k'j_1j_4}$$
$$= C_{51} + C_{52} + 2C_{53} + C_{54}$$

For C_{51}, we have

$$C_{51} = \sum_k \frac{1}{M_k^3} \sum_{i_1, i_2, i_3 \in S_k} N_{i_1} N_{i_2} N_{i_3} \langle \mu \circ \Omega_{i_1}, \Omega_{i_2} \rangle \langle \Omega_{i_1}, \Omega_{i_3} \rangle \langle \Omega_{i_2}, \Omega_{i_3} \rangle$$
$$= \sum_k \frac{1}{M_k^2} \sum_{i_1, i_2 \in S_k} N_{i_1} N_{i_2} \langle \mu \circ \Omega_{i_1}, \Omega_{i_2} \rangle \cdot \langle \Omega_{i_1}, \Sigma_k \Omega_{i_2} \rangle$$
$$\leq \sum_k \Bigg(\frac{1}{M_k^2} \sum_{i_1, i_2 \in S_k} N_{i_1} N_{i_2} \langle \mu \circ \Omega_{i_1}, \Omega_{i_2} \rangle^2 \Bigg)^{1/2} \Bigg(\frac{1}{M_k^2} \sum_{i_1, i_2 \in S_k} N_{i_1} N_{i_2} \langle \Omega_{i_1}, \Sigma_k \Omega_{i_2} \rangle^2 \Bigg)^{1/2}$$
$$=: \sum_k C_{511k}^{1/2} \cdot C_{512k}^{1/2}. \tag{E.42}$$

We have by Cauchy–Schwarz that

$$C_{511k} = \frac{1}{M_k^2} \sum_{i_1, i_2 \in S_k} N_{i_1} N_{i_2} \big(\sum_j \mu_j \Omega_{i_1j} \Omega_{i_2j} \big)^2$$
$$\leq \frac{1}{M_k^2} \sum_{i_1, i_2 \in S_k} N_{i_1} N_{i_2} \big(\sum_j \mu_j^2 \Omega_{i_1j} \Omega_{i_2j} \big) \big(\sum_j \Omega_{i_1j} \Omega_{i_2j} \big) \leq \|\mu\|_4^4,$$

and similarly

$$
\begin{aligned}
C_{512k} &= \frac{1}{M_k^2} \sum_{i_1,i_2 \in S_k} N_{i_1} N_{i_2} \big(\sum_{j_1,j_2} \Omega_{i_1 j_1} \Sigma_{k j_1 j_2} \Omega_{i_2 j_2} \big)^2 \\
&= \frac{1}{M_k^2} \sum_{i_1,i_2} N_{i_1} N_{i_2} \big(\sum_{j_1,j_2} \Omega_{i_1 j_1} \Sigma_{k j_1 j_2}^2 \Omega_{i_2 j_2} \big) \big(\sum_{j_1,j_2} \Omega_{i_1 j_1} \Omega_{i_2 j_2} \big) \\
&\leq \frac{1}{M_k^2} \sum_{i_1,i_2} N_{i_1} N_{i_2} \big(\sum_{j_1,j_2} \Omega_{i_1 j_1} \Sigma_{k j_1 j_2}^2 \Omega_{i_2 j_2} \big) = \mu' \, \Sigma_k^{\circ 2} \, \mu
\end{aligned}
\tag{E.43}
$$

Since by Cauchy–Schwarz,

$$
\begin{aligned}
\mu' \, \Sigma_k^{\circ 2} \, \mu &= \sum_{j_1,j_2} \mu_{j_1} \mu_{j_2} \big(\frac{1}{M_k} \sum_{i \in S_k} N_i \Omega_{i j_1} \Omega_{i j_2} \big)^2 = \frac{1}{M_k^2} \sum_{j_1,j_2} \mu_{j_1} \mu_{j_2} \sum_{i,i' \in S_k} N_i N_{i'} \Omega_{i j_1} \Omega_{i j_2} \Omega_{i' j_1} \Omega_{i' j_2} \\
&= \frac{1}{M_k^2} \sum_{i,i' \in S_k} \big(\sum_j \mu_j \Omega_{ij} \Omega_{i'j} \big)^2 \leq \frac{1}{M_k^2} \sum_{i,i' \in S_k} \sum_j \mu_j^2 \Omega_{ij} \Omega_{i'j} \leq \|\mu\|_4^4
\end{aligned}
\tag{E.44}
$$

we have in total $C_{512k} \lesssim K\|\mu\|_4^4$. Combining the result with the bound for C_{511k} implies that

$$
C_{51} \lesssim K\|\mu\|_4^4.
$$

Next we study C_{52} using a similar argument.

$$
\begin{aligned}
C_{52} &= \sum_k \sum_{j_1,j_3,j_4} \frac{M_k}{M} \mu_{j_1} \Sigma_{j_3 j_4} \Sigma_{k j_1 j_3} \Sigma_{k j_1 j_4} \\
&= \sum_k \sum_{j_1,j_3,j_4} \frac{M_k}{M} \mu_{j_1} \big(\frac{1}{M} \sum_{i_1 \in [n]} N_{i_1} \Omega_{i_1 j_3} \Omega_{i_1 j_4} \big) \big(\frac{1}{M_k} \sum_{i_2 \in S_k} N_{i_2} \Omega_{i_2 j_1} \Omega_{i_2 j_3} \big) \big(\frac{1}{M_k} \sum_{i_3 \in S_k} N_{i_3} \Omega_{i_3 j_1} \Omega_{i_3 j_4} \big) \\
&= \sum_k \frac{1}{M^2 M_k} \sum_{j_1,j_2,j_3} \sum_{\substack{i_1 \in [n] \\ i_2, i_3 \in S_k}} N_{i_1} N_{i_2} N_{i_3} \langle \mu \circ \Omega_{i_2}, \Omega_{i_3} \rangle \langle \Omega_{i_1}, \Omega_{i_3} \rangle \langle \Omega_{i_1}, \Omega_{i_2} \rangle \\
&= \sum_k \frac{1}{M^2} \sum_{i_2,i_3 \in [S_k]} N_{i_2} N_{i_3} \langle \mu \circ \Omega_{i_2}, \Omega_{i_3} \rangle \langle \Omega_{i_3}, \Sigma \Omega_{i_2} \rangle \\
&\leq \sum_k \left(\frac{1}{M^2} \sum_{i_2,i_3 \in [S_k]} N_{i_2} N_{i_3} \langle \mu \circ \Omega_{i_2}, \Omega_{i_3} \rangle^2 \right)^{1/2} \left(\frac{1}{M^2} \sum_{i_2,i_3 \in [S_k]} N_{i_2} N_{i_3} \langle \Omega_{i_3}, \Sigma \Omega_{i_2} \rangle \right)^{1/2} \\
&=: \sum_k C_{521k}^{1/2} C_{522k}^{1/2}.
\end{aligned}
\tag{E.45}
$$

Observe that $C_{521k} = C_{511k}$, and thus $C_{521} \lesssim \|\mu\|^4$ by (E.43). With a similar argument as in (E.44) we obtain $C_{522k} \lesssim \|\mu\|_4^4$. Hence we obtain

$$
C_{52} \leq \sum_k C_{521k}^{1/2} C_{522k}^{1/2} \lesssim K\|\mu\|_4^4.
$$

For C_{53}, we have

$$
\begin{aligned}
C_{53} &= \sum_{k \neq k'} \sum_{j_1,j_3,j_4} \frac{M_k M_{k'}}{M^2} \mu_{j_1} \Sigma_{k j_3 j_4} \Sigma_{k j_1 j_3} \Sigma_{k' j_1 j_4} \\
&\leq \sum_k \sum_{j_1,j_3,j_4} \frac{M_k}{M} \mu_{j_1} \Sigma_{k j_3 j_4} \Sigma_{k j_1 j_3} \Sigma_{j_1 j_4} \\
&= \sum_k \sum_{j_1,j_3,j_4} \frac{M_k}{M} \mu_{j_1} \big(\frac{1}{M_k} \sum_{i_1 \in S_k} N_{i_1} \Omega_{i_1 j_3} \Omega_{i_1 j_4} \big) \big(\frac{1}{M_k} \sum_{i_2 \in S_k} N_{i_2} \Omega_{i_2 j_1} \Omega_{i_2 j_3} \big) \big(\frac{1}{M} \sum_{i_3 \in [n]} N_{i_3} \Omega_{i_3 j_1} \Omega_{i_3 j_4} \big)
\end{aligned}
$$

$$= \sum_k \frac{1}{M^2 M_k} \sum_{\substack{i_1,i_2\in S_k\\ i_3\in[n]}} N_{i_1}N_{i_2}N_{i_3}\langle \mu\circ\Omega_{i_2},\Omega_{i_3}\rangle\langle\Omega_{i_1},\Omega_{i_2}\rangle\langle\Omega_{i_1},\Omega_{i_3}\rangle$$

$$= \sum_k \frac{1}{M^2} \sum_{i_2\in S_k, i_3\in[n]} N_{i_2}N_{i_3}\langle \mu\circ\Omega_{i_2},\Omega_{i_3}\rangle\langle\Omega_{i_2},\Sigma_k\Omega_{i_3}\rangle. \tag{E.46}$$

We then upper bound the last line using a similar strategy as in that we used for C_{51} and C_{52}, respectively. We omit the details and state the final bound:

$$C_{53} \lesssim K\|\mu\|_4^4 \tag{E.47}$$

Finally for C_{54}, summing over k, k' we obtain

$$C_{54} \le \sum_{j_1,j_3,j_4} \mu_{j_1}\Sigma_{j_3j_4}\Sigma_{j_1j_3}\Sigma_{j_1j_4} = \frac{1}{M^3}\sum_{i_1,i_2,i_3\in[n]} N_{i_1}N_{i_2}N_{i_3}\langle \mu\circ\Omega_{i_2},\Omega_{i_3}\rangle\langle\Omega_{i_1},\Omega_{i_2}\rangle\langle\Omega_{i_1},\Omega_{i_3}\rangle. \tag{E.48}$$

We then proceed as in (E.46) to control the right-hand side. We omit the details and state the final bound:

$$C_{54} \lesssim K\|\mu\|_4^4. \tag{E.49}$$

Combining the results for $C_{51},\ldots,C_{54}$, we see that

$$C_5 \lesssim K\|\mu\|^4.$$

For C_6, we have

$$\begin{aligned}
C_6 &\le \Big(\sum_k \sum_{i,i'\in S_k} N_iN_{i'} + \sum_{k\neq k'}\sum_{i\in S_k, i'\in S_{k'}} N_iN_{i'}\Big)\sum_{j_1,j_2,j_4} \alpha_{ii'j_1j_2}\alpha_{ii'j_1j_4}\,\Omega_{ij_1}\Omega_{i'j_2}\Omega_{i'j_4}\\
&\lesssim \sum_k\sum_{i,i'\in S_k} N_iN_{i'} \sum_{j_1,j_2,j_4}\big(\frac{1}{M_k}\Sigma_{kj_1j_2}+\frac{1}{M}\Sigma_{j_1j_2}\big)\big(\frac{1}{M_k}\Sigma_{kj_1j_4}+\frac{1}{M}\Sigma_{j_1j_4}\big)\,\Omega_{ij_1}\Omega_{i'j_2}\Omega_{i'j_4}\\
&+ \sum_{\substack{k\neq k'\\ i\in S_k,i'\in S_{k'}\\ j_1,j_2,j_4}} N_iN_{i'}\big(\frac{1}{M}\sum_{a\in\{k,k'\}}^{2}\Sigma_{aj_1j_2}+\frac{1}{M}\Sigma_{j_1j_2}\big)\big(\frac{1}{M}\sum_{a\in\{k,k'\}}^{2}\Sigma_{aj_1j_4}+\frac{1}{M}\Sigma_{j_1j_4}\big)\,\Omega_{ij_1}\Omega_{i'j_2}\Omega_{i'j_4}\\
&=: C_{61}+C_{62}.
\end{aligned}$$

For C_{61}, we have

$$\begin{aligned}
C_{61} &= \sum_k\sum_{i'\in S_k} N_{i'}\sum_{j_1,j_2,j_4}\frac{1}{M_k}\Sigma_{kj_1j_2}\Sigma_{kj_1j_4}\mu_{j_1}\Omega_{i'j_2}\Omega_{i'j_4}\\
&+2\sum_k\sum_{i'\in S_k} N_{i'}\sum_{j_1,j_2,j_4}\frac{1}{M}\Sigma_{kj_1j_2}\Sigma_{j_1j_4}\mu_{j_1}\Omega_{i'j_2}\Omega_{i'j_4}\\
&+\ \sum_k\sum_{i'\in S_k} N_{i'}\sum_{j_1,j_2,j_4}\frac{M_k}{M^2}\Sigma_{j_1j_2}\Sigma_{j_1j_4}\mu_{j_1}\Omega_{i'j_2}\Omega_{i'j_4} =: C_{611}+2C_{612}+C_{613}.
\end{aligned}$$

Relabeling indices, we see that

$$C_{611} = \sum_k\sum_{j_1,j_2,j_4}\mu_{j_1}\Sigma_{kj_1j_2}\Sigma_{kj_1j_4}\Sigma_{kj_2j_4} = C_{51}$$

Hence, $C_{611} \lesssim K\|\mu\|_4^4$. Next,

$$C_{612} \le \sum_k \frac{M_k}{M} \sum_{j_1,j_2,j_4} \mu_{j_1} \Sigma_{kj_1j_2} \Sigma_{j_1j_4} \Sigma_{kj_2j_4} \lesssim K\|\mu\|^4,$$

where we applied (E.46). Similarly,

$$C_{613} = \sum_k \frac{M_k^2}{M^2} \sum_{j_1,j_2,j_4} \mu_{j_1} \Sigma_{j_1j_2} \Sigma_{j_1j_4} \Sigma_{kj_2j_4} \le \sum_{j_1,j_2,j_4} \mu_{j_1} \Sigma_{j_1j_2} \Sigma_{j_1j_4} \Sigma_{j_2j_4} \lesssim K\|\mu\|^4,$$

where in the final bound we apply (E.48) and (E.49). Combining the results above for $C_{611}, C_{612}, C_{613}$, we obtain

$$C_{61} \lesssim K\|\mu\|_4^4 \tag{E.50}$$

The argument for C_{62} is very similar, so we omit proof and state the final bound. We have

$$C_{62} \lesssim K\|\mu\|_4.$$

Thus

$$C_6 \lesssim K\|\mu\|_4^4$$

For C_7, we have

$$\begin{aligned}
C_7 &\lesssim \Big(\sum_k \sum_{i,i'\in S_k} N_i N_{i'} + \sum_{k\neq k'} \sum_{i\in S_k, i'\in S_{k'}} N_i N_{i'} \Big) \sum_{j_1,j_2,j_3,j_4} \alpha_{ii'j_1j_2} \alpha_{ii'j_3j_4} \Omega_{ij_1} \Omega_{ij_3} \Omega_{i'j_2} \Omega_{i'j_4} \\
&\lesssim \sum_k \sum_{i,i'\in S_k} N_i N_{i'} \sum_{j_1,j_2,j_3,j_4} \big(\frac{1}{M_k} \Sigma_{kj_1j_2} + \frac{1}{M} \Sigma_{j_1j_2}\big)\big(\frac{1}{M_k} \Sigma_{kj_3j_4} + \frac{1}{M} \Sigma_{j_3j_4}\big) \Omega_{ij_1} \Omega_{ij_3} \Omega_{i'j_2} \Omega_{i'j_4} \\
&+ \sum_{k\neq k'} \sum_{\substack{j_1,j_2,j_3,j_4 \\ i\in S_k, i'\in S_{k'}}} N_i N_{i'} \big(\frac{1}{M} \sum_{a\in\{k,k'\}}^{2} \Sigma_{aj_1j_2} + \frac{1}{M} \Sigma_{j_1j_2}\big)\big(\frac{1}{M} \sum_{a\in\{k,k'\}}^{2} \Sigma_{aj_3j_4} + \frac{1}{M} \Sigma_{j_3j_4}\big) \Omega_{ij_1} \Omega_{ij_3} \Omega_{i'j_2} \Omega_{i'j_4} \\
&=: C_{71} + C_{72}
\end{aligned}$$

Write

$$\begin{aligned}
C_{71} &= \sum_k \sum_{i,i'\in S_k} N_i N_{i'} \sum_{j_1,j_2,j_3,j_4} \frac{1}{M_k^2} \Sigma_{kj_1j_2} \Sigma_{kj_3j_4} \Omega_{ij_1} \Omega_{ij_3} \Omega_{i'j_2} \Omega_{i'j_4} \\
&\quad + 2 \sum_k \sum_{i,i'\in S_k} N_i N_{i'} \sum_{j_1,j_2,j_3,j_4} \frac{1}{M_k M} \Sigma_{j_1j_2} \Sigma_{kj_3j_4} \Omega_{ij_1} \Omega_{ij_3} \Omega_{i'j_2} \Omega_{i'j_4} \\
&\quad + \sum_k \sum_{i,i'\in S_k} N_i N_{i'} \sum_{j_1,j_2,j_3,j_4} \frac{1}{M^2} \Sigma_{j_1j_2} \Sigma_{j_3j_4} \Omega_{ij_1} \Omega_{ij_3} \Omega_{i'j_2} \Omega_{i'j_4} =: C_{711} + 2C_{712} + C_{713}.
\end{aligned}$$

For C_{711}, we have

$$\begin{aligned}
C_{711} &= \sum_k \sum_{j_1,j_2,j_3,j_4} \Sigma_{kj_1j_2} \Sigma_{kj_3j_4} \Sigma_{kj_1j_3} \Sigma_{kj_2j_4} \\
&= \sum_k \frac{1}{M_k^4} \sum_{i_1,i_2,i_3,i_4\in S_k} N_{i_1} N_{i_2} N_{i_3} N_{i_4} \langle \Omega_{i_1}, \Omega_{i_3} \rangle \langle \Omega_{i_1}, \Omega_{i_4} \rangle \langle \Omega_{i_2}, \Omega_{i_3} \rangle \langle \Omega_{i_2}, \Omega_{i_4} \rangle \\
&= \frac{1}{M_k^2} \sum_k \sum_{i_3,i_4} N_{i_3} N_{i_4} \big(\Omega_{i_3}' \Sigma_k \Omega_{i_4}\big)^2 = \sum_k \frac{1}{M_k^2} \sum_{i_3,i_4} N_{i_3} N_{i_4} \big(\sum_{j,j'} \Omega_{i_3j}' \Sigma_{kjj'} \Omega_{i_4j'} \big)^2
\end{aligned}$$

$$\leq \sum_k \frac{1}{M_k^2} \sum_{i_3,i_4} N_{i_3}N_{i_4} \sum_{j,j'} \Omega'_{i_3 j}\Sigma^2_{kjj'}\Omega_{i_4 j'} \leq \sum_k \sum_{j,j'} \mu_j \Sigma^2_{kjj'}\mu_{j'} \lesssim K\|\mu\|_4^4. \tag{E.51}$$

In the last line we applied Cauchy–Schwarz and (E.44). For C_{712}, we have similarly

$$\begin{aligned}
C_{712} &= \sum_k \frac{M_k}{M} \sum_{j_1,j_2,j_3,j_4} \Sigma_{j_1j_2}\Sigma_{kj_3j_4}\Sigma_{kj_1j_3}\Sigma_{kj_2j_4} \\
&= \sum_k \frac{1}{M^2 M_k} \sum_{\substack{i_1\in[n] \\ i_2,i_3,i_4\in S_k}} N_{i_1}N_{i_2}N_{i_3}N_{i_4}\langle\Omega_{i_1},\Omega_{i_3}\rangle\langle\Omega_{i_1},\Omega_{i_4}\rangle\langle\Omega_{i_2},\Omega_{i_3}\rangle\langle\Omega_{i_2},\Omega_{i_4}\rangle \\
&= \sum_k \frac{M_k}{M^2} \sum_{i_1\in[n],i_2\in S_k} N_{i_1}N_{i_2}\langle\Omega_{i_1},\Sigma_k\Omega_{i_2}\rangle^2 \leq \sum_k \frac{M_k}{M^2}\sum_{i_1\in[n],i_2\in S_k} N_{i_1}N_{i_2}\sum_{j,j'}\Omega_{i_1j}\Sigma^2_{kjj'}\Omega_{i_2j'} \\
&\leq \sum_k \frac{M_k^2}{M^2}\sum_{j,j'}\mu_j\Sigma^2_{kjj'}\mu_{j'} \lesssim K\|\mu\|_4^4.
\end{aligned} \tag{E.52}$$

Next,

$$\begin{aligned}
C_{713} &= \sum_k \frac{M_k^2}{M^2} \sum_{j_1,j_2,j_3,j_4} \Sigma_{j_1j_2}\Sigma_{j_3j_4}\Sigma_{kj_1j_3}\Sigma_{kj_2j_4} \\
&= \sum_k \frac{1}{M^4} \sum_{\substack{i_1,i_2\in[n] \\ i_3,i_4\in S_k}} N_{i_1}N_{i_2}N_{i_3}N_{i_4}\langle\Omega_{i_1},\Omega_{i_3}\rangle\langle\Omega_{i_1},\Omega_{i_4}\rangle\langle\Omega_{i_2},\Omega_{i_3}\rangle\langle\Omega_{i_2},\Omega_{i_4}\rangle,
\end{aligned}$$

and applying a similar strategy as in (E.51), (E.52) leads to the bound $C_{713} \lesssim K\|\mu\|_4^4$. Thus

$$C_{71} \lesssim K\|\mu\|_4^4.$$

Next , by symmetry and summing over $i \in S_k, i' \in S_{k'}$, we have

$$\begin{aligned}
C_{72} &= \sum_{k\neq k'} \frac{M_kM_{k'}}{M^2} \sum_{j_1,j_2,j_3,j_4} \Big[2\Sigma_{kj_1j_2}\Sigma_{kj_3j_4} + 2\Sigma_{k'j_1j_2}\Sigma_{kj_3j_4} + 4\Sigma_{kj_1j_2}\Sigma_{j_3j_4} + \Sigma_{j_1j_2}\Sigma_{j_3j_4}\Big]\Sigma_{kj_1j_3}\Sigma_{k'j_2j_4} \\
&=: 2C_{721} + 2C_{722} + 4C_{723} + C_{724}
\end{aligned}$$

First,

$$C_{721} \leq \sum_k \frac{M_k}{M}\sum_{j_1,j_2,j_3,j_4}\Sigma_{kj_1j_2}\Sigma_{kj_3j_4}\Sigma_{kj_1j_3}\Sigma_{j_2j_4} = C_{712} \lesssim K\|\mu\|_4^4$$

by (E.52). Next,

$$\begin{aligned}
C_{722} &= \sum_{k\neq k'} \frac{M_kM_{k'}}{M^2} \sum_{j_1,j_2,j_3,j_4} \Sigma_{k'j_1j_2}\Sigma_{kj_3j_4}\Sigma_{kj_1j_3}\Sigma_{k'j_2j_4} \\
&\leq \sum_{k,k'} \frac{1}{M^2M_kM_{k'}} \sum_{\substack{i_1,i_2\in S_k \\ i_3,i_4\in S_{k'}}} N_{i_1}N_{i_2}N_{i_3}N_{i_4}\langle\Omega_{i_1},\Omega_{i_3}\rangle\langle\Omega_{i_1},\Omega_{i_4}\rangle\langle\Omega_{i_2},\Omega_{i_3}\rangle\langle\Omega_{i_2},\Omega_{i_4}\rangle \\
&= \sum_{k,k'} \frac{M_k}{M^2M_{k'}} \sum_{i_3,i_4\in S_{k'}} N_{i_3}N_{i_4}\langle\Omega_{i_3},\Sigma_k\Omega_{i_4}\rangle^2 \leq \sum_{k,k'}\frac{M_k}{M^2M_{k'}}\sum_{i_3,i_4\in S_{k'}}N_{i_3}N_{i_4}\sum_{j,j'}\Omega_{i_3j}\Sigma^2_{kjj'}\Omega_{i_4j'} \\
&\leq \sum_{k,k'}\frac{M_kM_{k'}}{M^2}\mu'\Sigma_k^{\circ 2}\mu \leq \|\mu\|_4^4,
\end{aligned} \tag{E.53}$$

where we applied Cauchy-Schwarz in the penultimate line and (E.44) in the last line.

For C_{723}, we have

$$\begin{aligned}
C_{723} &= \sum_{k\neq k'} \frac{M_k M_{k'}}{M^2} \sum_{j_1,j_2,j_3,j_4} \Sigma_{kj_1j_2}\Sigma_{j_3j_4}\Sigma_{kj_1j_3}\Sigma_{k'j_2j_4} \le \sum_k \frac{M_k}{M} \sum_{j_1,j_2,j_3,j_4} \Sigma_{kj_1j_2}\Sigma_{j_3j_4}\Sigma_{kj_1j_3}\Sigma_{j_2j_4} \\
&= \sum_k \frac{1}{M^3 M_k} \sum_{\substack{i_1,i_3\in S_k \\ i_2,i_4\in[n]}} N_{i_1}N_{i_2}N_{i_3}N_{i_4}\langle\Omega_{i_1},\Omega_{i_3}\rangle\langle\Omega_{i_1},\Omega_{i_4}\rangle\langle\Omega_{i_2},\Omega_{i_3}\rangle\langle\Omega_{i_2},\Omega_{i_4}\rangle \\
&= \sum_k \frac{1}{M^2} \sum_{i_3\in S_k, i_4\in[n]} N_{i_3}N_{i_4}\langle\Omega_{i_3},\Sigma_k\Omega_{i_4}\rangle\langle\Omega_{i_3},\Sigma\Omega_{i_4}\rangle \\
&\le \frac{1}{2}\sum_k \frac{1}{M^2} \sum_{i_3\in S_k, i_4\in[n]} N_{i_3}N_{i_4}\big(\langle\Omega_{i_3},\Sigma_k\Omega_{i_4}\rangle^2 + \langle\Omega_{i_3},\Sigma\Omega_{i_4}\rangle^2\big)
\end{aligned}$$

Using a similar technique as in (E.51)–(E.53) and applying (E.38), (E.39) we obtain

$$C_{723} \lesssim \|\mu\|_4^4.$$

Finally, for C_{724} we have

$$\begin{aligned}
C_{724} &= \sum_{k\neq k'} \frac{M_k M_{k'}}{M^2} \sum_{j_1,j_2,j_3,j_4} \Sigma_{j_1j_2}\Sigma_{j_3j_4}\Sigma_{kj_1j_3}\Sigma_{k'j_2j_4} \le \sum_{j_1,j_2,j_3,j_4} \Sigma_{j_1j_2}\Sigma_{j_3j_4}\Sigma_{j_1j_3}\Sigma_{j_2j_4} \\
&= \frac{1}{M^4} \sum_{i_1,i_2,i_3,i_4\in[n]} N_{i_1}N_{i_2}N_{i_3}N_{i_4}\langle\Omega_{i_1},\Omega_{i_3}\rangle\langle\Omega_{i_1},\Omega_{i_4}\rangle\langle\Omega_{i_2},\Omega_{i_3}\rangle\langle\Omega_{i_2},\Omega_{i_4}\rangle
\end{aligned}$$

The details are very similar to (E.51)–(E.53), so we omit them and simply state the final bound:

$$C_{724} \lesssim \|\mu\|_4^4$$

Combining the bounds for $C_{721}, C_{722}, C_{723}$, and C_{724} yields

$$C_7 \lesssim K\|\mu\|_4^4.$$

Combining the bounds for C_1–C_7 proves the result.

□

E.5 Proof of Lemma E.3

We have

$$\begin{aligned}
\mathbb{E}D_{\ell,s}^4 &= \mathbb{E}\Big[\big(\sum_{i\in[\ell-1]} \sigma_{i,\ell} \sum_{r=1}^{N_i}\sum_j Z_{ijr}Z_{\ell js}\big)^4\Big] \\
&= \sum_{i_1,i_2,i_3,i_4\in[\ell-1]} \sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell} \sum_{\substack{r_1,r_2,r_3,r_4 \\ j_1,j_2,j_3,j_4}} \mathbb{E}\big[Z_{i_1j_1r_1}Z_{\ell j_1 s}Z_{i_2j_2r_2}Z_{\ell j_2 s}Z_{i_3j_3r_3}Z_{\ell j_3 s}Z_{i_4j_4r_4}Z_{\ell j_4 s}\big] \\
&= \sum_{i_1,i_2,i_3,i_4\in[\ell-1]} \sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell} \sum_{\substack{r_1,r_2,r_3,r_4 \\ j_1,j_2,j_3,j_4}} \mathbb{E}\big[Z_{i_1j_1r_1}Z_{i_2j_2r_2}Z_{i_3j_3r_3}Z_{i_4j_4r_4}\big]\mathbb{E}\big[Z_{\ell j_1 s}Z_{\ell j_2 s}Z_{\ell j_3 s}Z_{\ell j_4 s}\big] \\
&= \sum_{j_1,j_2,j_3,j_4} \mathbb{E}[Z_{\ell j_1 s}Z_{\ell j_2 s}Z_{\ell j_3 s}Z_{\ell j_4 s}] \sum_{\substack{i_1,i_2,i_3,i_4\in[\ell-1] \\ r_1,r_2,r_3,r_4}} \sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{i_1j_1r_1}Z_{i_2j_2r_2}Z_{i_3j_3r_3}Z_{i_4j_4r_4}] \\
&=: \sum_{j_1,j_2,j_3,j_4} \mathbb{E}[Z_{\ell j_1 s}Z_{\ell j_2 s}Z_{\ell j_3 s}Z_{\ell j_4 s}]A_{j_1,j_2,j_3,j_4} \qquad \text{(E.54)}
\end{aligned}$$

In the summations above, r_t ranges over $[N_{i_t}]$.

Observe that

$$|\mathbb{E}[Z_{\ell j_1 s}Z_{\ell j_2 s}Z_{\ell j_3 s}Z_{\ell j_4 s}]| \lesssim \begin{cases} \Omega_{\ell j_1} & \text{if } j_1 = j_2 = j_3 = j_4 \\ \Omega_{\ell j_1}\Omega_{\ell j_4} & \text{if } j_1 = j_2 = j_3, j_4 \neq j_1 \\ \Omega_{\ell j_1}\Omega_{\ell j_3} & \text{if } j_1 = j_2, j_3 = j_4, j_1 \neq j_3 \\ \Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4} & \text{if } j_1 = j_2, j_1, j_3, j_4 \ dist. \\ \Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4} & \text{if } j_1, j_2, j_3, j_4 \ dist. \end{cases} \tag{E.55}$$

Up to permutation of the indices $j_1, \dots, j_4$, this accounts for all possible cases.

To proceed we also bound A_{j_1,j_2,j_3,j_4} by casework on the number of distinct j indices. For brevity we define $\omega_t = (i_t, r_t)$ and slightly abuse notation, letting $Z_{\omega_t, j} = Z_{i_t j r_t}$. Further let $\mathcal{I}_\ell = \{\omega = (i, r) : i \in [\ell], 1 \leq r \leq N_i\}$. Our goal is to control

$$A_{j_1,j_2,j_3,j_4} = \sum_{\omega_1,\omega_2,\omega_3,\omega_4 \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{\omega_1 j_1}Z_{\omega_2 j_2}Z_{\omega_3 j_3}Z_{\omega_4 j_4}]. \tag{E.56}$$

To do this, we study (E.56) in five cases that cover all possibilities (up to permutation of the indices $j_1, \dots, j_4$).

Case 1: $j_1 = j_2 = j_3 = j_4$. Define $j = j_1$. It holds that

$$\begin{aligned} &\sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{\omega_1 j}Z_{\omega_2 j}Z_{\omega_3 j}Z_{\omega_4 j}] \\ &\quad = \begin{cases} \sigma_{i_1\ell}^4 \, \mathbb{E}Z_{\omega_1 j}^4 \lesssim \sigma_{i_1\ell}^4 \, \Omega_{i_1 j} & \text{if } \omega_1 = \omega_2 = \omega_3 = \omega_4 \\ \sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\mathbb{E}Z_{\omega_1 j}^2\mathbb{E}Z_{\omega_3 j}^2 \lesssim \sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\Omega_{i_1 j}\Omega_{i_3 j} & \text{if } \omega_1 = \omega_2, \omega_3 = \omega_4, \omega_1 \neq \omega_3 \end{cases} \end{aligned} \tag{E.57}$$

Up to permutation of the indices $\omega_1, \dots, \omega_4$, this accounts for all cases such that (E.57) is nonvanishing. To be precise, by symmetry, it also holds that for all permutations $\pi : [4] \to [4]$ that if $\omega_{\pi(1)} = \omega_{\pi(2)}, \omega_{\pi(3)} = \omega_{\pi(4)}, \omega_{\pi(1)} \neq \omega_{\pi(3)}$, then

$$\sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{\omega_1 j}Z_{\omega_2 j}Z_{\omega_3 j}Z_{\omega_4 j}] \lesssim \sigma_{i_{\pi(1)}\ell}^2\sigma_{i_{\pi(3)}\ell}^2\Omega_{i_{\pi(1)} j}\Omega_{i_{\pi(3)} j}.$$

In all other cases besides those considered above, we have

$$\sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{\omega_1 j}Z_{\omega_2 j}Z_{\omega_3 j}Z_{\omega_4 j}] = 0$$

by independence.

Therefore,

$$A_{jjjj} \lesssim \sum_{\omega \in \mathcal{I}_{\ell-1}} \sigma_{i\ell}^4\Omega_{ij} + \sum_{\omega_1 \neq \omega_3 \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\Omega_{i_1 j}\Omega_{i_3 j} \tag{E.58}$$

In the remaining Cases 2–6, we follow the same strategy of writing out bounds for

$$\sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{\omega_1 j_1}Z_{\omega_2 j_2}Z_{\omega_3 j_3}Z_{\omega_4 j_4}]$$

that cover all nonzero cases, up to permutation of the indices $\omega_1, \dots, \omega_4$.

Case 2: $j_1 = j_2 = j_3, j_1 \neq j_4$. It holds that

$$\begin{aligned} &\sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{\omega_1 j_1}Z_{\omega_2 j_1}Z_{\omega_3 j_1}Z_{\omega_4 j_4}] \\ &\quad = \begin{cases} \sigma_{i_1\ell}^4\mathbb{E}[Z_{\omega_1 j_1}^3 Z_{\omega_1 j_4}] \lesssim \sigma_{i_1\ell}^4 \, \Omega_{i_1 j_1}\Omega_{i_1 j_4} & \text{if } \omega_1 = \omega_2 = \omega_3 = \omega_4 \\ \sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\mathbb{E}Z_{\omega_1 j_1}^2\mathbb{E}Z_{\omega_3 j_1}Z_{\omega_3 j_4} \lesssim \sigma_{i_1\ell}^2\sigma_{i_3\ell}^2 \, \Omega_{i_1 j_1}\Omega_{i_3 j_1}\Omega_{i_3 j_4} & \text{if } \omega_1 = \omega_2, \omega_3 = \omega_4, \omega_1 \neq \omega_3 \end{cases} \end{aligned} \tag{E.59}$$

Up to permutation of the indices $\omega_1, \ldots, \omega_4$, this accounts for all cases such that (E.59) is nonvanishing. Thus

$$A_{j_1,j_1,j_1,j_4} \lesssim \sum_{\omega \in \mathcal{I}_{\ell-1}} \sigma_{i\ell}^4 \Omega_{i_1 j_1} \Omega_{i_1 j_4} + \sum_{\omega_1 \neq \omega_3 \in \mathcal{I}_{\ell-1}} \sigma_{i_1 \ell}^2 \sigma_{i_3 \ell}^2 \, \Omega_{i_1 j_1} \Omega_{i_3 j_1} \Omega_{i_3 j_4} \tag{E.60}$$

Case 3: $j_1 = j_2, j_3 = j_4, j_1 \neq j_3$. It holds that

$$\begin{aligned} &\sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{\omega_1 j_1} Z_{\omega_2 j_1} Z_{\omega_3 j_3} Z_{\omega_4 j_3}] \\ &= \begin{cases} \sigma_{i_1\ell}^4 \mathbb{E} Z_{\omega_1 j_1}^2 Z_{\omega_1 j_3}^2 \lesssim \sigma_{i_1\ell}^4 \, \Omega_{i_1 j_1} \Omega_{i_1 j_3} & \text{if } \omega_1 = \omega_2 = \omega_3 = \omega_4 \\ \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \mathbb{E} Z_{\omega_1 j_1}^2 \mathbb{E} Z_{\omega_3 j_3}^2 \lesssim \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \, \Omega_{i_1 j_1} \Omega_{i_3 j_3} & \text{if } \omega_1 = \omega_2, \omega_3 = \omega_4, \omega_1 \neq \omega_3 \\ \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \mathbb{E} Z_{\omega_1 j_1} Z_{\omega_1 j_3} \mathbb{E} Z_{\omega_2 j_1} Z_{\omega_2 j_3} \lesssim \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \Omega_{i_1 j_1} \Omega_{i_1 j_3} \Omega_{i_2 j_1} \Omega_{i_2 j_3} & \text{if } \omega_1 = \omega_3, \omega_2 = \omega_4, \omega_1 \neq \omega_2. \end{cases} \end{aligned} \tag{E.61}$$

Up to permutation of the indices $\omega_1, \ldots, \omega_4$, this accounts for all cases such that (E.61) is nonvanishing. Thus by symmetry,

$$\begin{aligned} A_{j_1,j_1,j_3,j_3} \lesssim &\sum_{\omega \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^4 \, \Omega_{i_1 j_1} \Omega_{i_1 j_3} + \sum_{\omega_1 \neq \omega_3 \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \, \Omega_{i_1 j_1} \Omega_{i_3 j_3} \\ &+ \sum_{\omega_1 \neq \omega_3 \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \Omega_{i_1 j_1} \Omega_{i_1 j_3} \Omega_{i_3 j_1} \Omega_{i_3 j_3} \end{aligned} \tag{E.62}$$

Case 4: $j_1 = j_2$ and j_1, j_3, j_4 distinct. We have

$$\begin{aligned} &\sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{\omega_1 j_1} Z_{\omega_2 j_1} Z_{\omega_3 j_3} Z_{\omega_4 j_4}] \\ &= \begin{cases} \sigma_{i_1\ell}^4 \mathbb{E} Z_{\omega_1 j_1}^2 Z_{\omega_1 j_3} Z_{\omega_1 j_4} \lesssim \sigma_{i_1\ell}^4 \, \Omega_{i_1 j_1} \Omega_{i_1 j_3} \Omega_{i_1 j_4} & \text{if } \omega_1 = \omega_2 = \omega_3 = \omega_4 \\ \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \mathbb{E} Z_{\omega_1 j_1}^2 \mathbb{E} Z_{\omega_3 j_3} Z_{\omega_3 j_4} \lesssim \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \, \Omega_{i_1 j_1} \Omega_{i_3 j_3} \Omega_{i_3 j_4} & \text{if } \omega_1 = \omega_2, \omega_3 = \omega_4, \omega_1 \neq \omega_3 \\ \sigma_{i_1\ell}^2 \sigma_{i_2\ell}^2 \mathbb{E} Z_{\omega_1 j_1} Z_{\omega_1 j_3} \mathbb{E} Z_{\omega_2 j_1} Z_{\omega_2 j_4} \lesssim \sigma_{i_1\ell}^2 \sigma_{i_2\ell}^2 \, \Omega_{i_1 j_1} \Omega_{i_1 j_3} \Omega_{i_2 j_1} \Omega_{i_2 j_4} & \text{if } \omega_1 = \omega_3, \omega_2 = \omega_4, \omega_1 \neq \omega_2 \end{cases} \end{aligned} \tag{E.63}$$

Up to permutation of the indices $\omega_1, \ldots, \omega_4$, this accounts for all cases such that (E.63) is nonvanishing. Thus

$$\begin{aligned} A_{j_1,j_1,j_3,j_4} \lesssim &\sum_{\omega \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^4 \, \Omega_{i_1 j_1} \Omega_{i_1 j_3} \Omega_{i_1 j_4} + \sum_{\omega_1 \neq \omega_3 \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \, \Omega_{i_1 j_1} \Omega_{i_3 j_3} \Omega_{i_3 j_4} \\ &\sum_{\omega_1 \neq \omega_3 \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^2 \sigma_{i_2\ell}^2 \, \Omega_{i_1 j_1} \Omega_{i_1 j_3} \Omega_{i_3 j_1} \Omega_{i_3 j_4}. \end{aligned} \tag{E.64}$$

Case 5: j_1, j_2, j_3, j_4 distinct. For this final case, it holds that

$$\begin{aligned} &\sigma_{i_1\ell}\sigma_{i_2\ell}\sigma_{i_3\ell}\sigma_{i_4\ell}\mathbb{E}[Z_{\omega_1 j_1} Z_{\omega_2 j_2} Z_{\omega_3 j_3} Z_{\omega_4 j_4}] \\ &= \begin{cases} \sigma_{i_1\ell}^4 \mathbb{E} Z_{\omega_1 j_1} Z_{\omega_1 j_2} Z_{\omega_1 j_3} Z_{\omega_1 j_4} \lesssim \sigma_{i_1\ell}^4 \, \Omega_{i_1 j_1} \Omega_{i_1 j_2} \Omega_{i_1 j_3} \Omega_{i_1 j_4} & \text{if } \omega_1 = \omega_2 = \omega_3 = \omega_4 \\ \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \mathbb{E} Z_{\omega_1 j_1} Z_{\omega_1 j_2} \mathbb{E} Z_{\omega_3 j_3} Z_{\omega_3 j_4} \lesssim \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \, \Omega_{i_1 j_1} \Omega_{i_1 j_2} \Omega_{i_3 j_3} \Omega_{i_3 j_4} & \text{if } \omega_1 = \omega_2, \omega_3 = \omega_4, \omega_1 \neq \omega_3 \end{cases} \end{aligned}$$

The above accounts for all nonzero cases, up to permutation of $\omega_1, \omega_2, \omega_3, \omega_4$. Hence

$$A_{j_1,j_2,j_3,j_4} \lesssim \sum_{\omega \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^4 \, \Omega_{i_1 j_1} \Omega_{i_1 j_2} \Omega_{i_1 j_3} \Omega_{i_1 j_4} + \sum_{\omega_1 \neq \omega_3 \in \mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^2 \sigma_{i_3\ell}^2 \, \Omega_{i_1 j_1} \Omega_{i_1 j_2} \Omega_{i_3 j_3} \Omega_{i_3 j_4}. \tag{E.65}$$

Finally we control the fourth moment using the casework above. By (E.54) and symmetry,

$$\mathbb{E} D_{\ell,s}^4 \lesssim \sum_j \mathbb{E}[Z_{\ell js} Z_{\ell js} Z_{\ell js} Z_{\ell js}] A_{j,j,j,j} + \sum_{j_1 \neq j_4} \mathbb{E}[Z_{\ell j_1 s} Z_{\ell j_1 s} Z_{\ell j_1 s} Z_{\ell j_4 s}] A_{j_1,j_1,j_1,j_4}$$

$$
\begin{aligned}
&+\sum_{j_1\neq j_3}\mathbb{E}[Z_{\ell j_1 s}Z_{\ell j_1 s}Z_{\ell j_3 s}Z_{\ell j_3 s}]A_{j_1,j_1,j_3,j_3}+\sum_{j_1,j_3,j_4\,dist.}\mathbb{E}[Z_{\ell j_1 s}Z_{\ell j_1 s}Z_{\ell j_3 s}Z_{\ell j_4 s}]A_{j_1,j_1,j_3,j_4}\\
&+\sum_{j_1,j_2,j_3,j_4\,dist.}\mathbb{E}[Z_{\ell j_1 s}Z_{\ell j_2 s}Z_{\ell j_3 s}Z_{\ell j_4 s}]A_{j_1,j_2,j_3,j_4}\\
=:&\ F_{1\ell s}+F_{2\ell s}+F_{3\ell s}+F_{4\ell s}+F_{5\ell s} \qquad \text{(E.66)}
\end{aligned}
$$

By (E.55), (E.58), (E.60) ,(E.62), (E.64), and (E.65),

$$
\begin{aligned}
F_{1\ell s}&\lesssim\sum_j\Omega_{\ell j}\bigg(\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i\ell}^4\Omega_{ij}+\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\Omega_{i_1 j}\Omega_{i_3 j}\bigg)\\
F_{2\ell s}&\lesssim\sum_{j_1\neq j_4}\Omega_{\ell j_1}\Omega_{\ell j_4}\bigg(\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i\ell}^4\Omega_{i_1 j_1}\Omega_{i_1 j_4}+\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\,\Omega_{i_1 j_1}\Omega_{i_3 j_1}\Omega_{i_3 j_4}\bigg)\\
F_{3\ell s}&\lesssim\sum_{j_1\neq j_3}\Omega_{\ell j_1}\Omega_{\ell j_3}\bigg(\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^4\,\Omega_{i_1 j_1}\Omega_{i_1 j_3}+\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\,\Omega_{i_1 j_1}\Omega_{i_3 j_3}\\
&\qquad+\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\Omega_{i_1 j_1}\Omega_{i_1 j_3}\Omega_{i_3 j_1}\Omega_{i_3 j_3}\bigg)\\
F_{4\ell s}&\lesssim\sum_{j_1,j_3,j_4\,dist.}\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4}\bigg(\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^4\,\Omega_{i_1 j_1}\Omega_{i_1 j_3}\Omega_{i_1 j_4}+\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\,\Omega_{i_1 j_1}\Omega_{i_3 j_3}\Omega_{i_3 j_4}\\
&\qquad+\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\,\Omega_{i_1 j_1}\Omega_{i_1 j_3}\Omega_{i_3 j_1}\Omega_{i_3 j_4}.\bigg)\\
F_{5\ell s}&\lesssim\sum_{j_1,j_2,j_3,j_4\,dist.}\Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4}\bigg(\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^4\,\Omega_{i_1 j_1}\Omega_{i_1 j_2}\Omega_{i_1 j_3}\Omega_{i_1 j_4}\\
&\qquad+\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\,\Omega_{i_1 j_1}\Omega_{i_1 j_2}\Omega_{i_3 j_3}\Omega_{i_3 j_4}\bigg).
\end{aligned}
$$

Define

$$
\begin{aligned}
F_{11\ell s}&=\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i\ell}^4\sum_j\Omega_{\ell j}\Omega_{ij}\\
F_{21\ell s}&=\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i\ell}^4\sum_{j_1\neq j_4}\Omega_{\ell j_1}\Omega_{\ell j_4}\Omega_{i_1 j_1}\Omega_{i_1 j_4}\\
F_{31\ell s}&=\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^4\sum_{j_1\neq j_3}\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{i_1 j_1}\Omega_{i_1 j_3}\\
F_{41\ell s}&=\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^4\sum_{j_1,j_3,j_4\,dist.}\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4}\Omega_{i_1 j_1}\Omega_{i_1 j_3}\Omega_{i_1 j_4}\\
F_{51\ell s}&=\sum_{\omega\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^4\sum_{j_1,j_2,j_3,j_4\,dist.}\Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4}\Omega_{i_1 j_1}\Omega_{i_1 j_2}\Omega_{i_1 j_3}\Omega_{i_1 j_4}
\end{aligned}
$$

and

$$
\begin{aligned}
F_{12\ell s}&=\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\sum_j\Omega_{\ell j}\Omega_{i_1 j}\Omega_{i_3 j}\\
F_{22\ell s}&=\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\sum_{j_1\neq j_4}\Omega_{\ell j_1}\Omega_{\ell j_4}\Omega_{i_1 j_1}\Omega_{i_3 j_1}\Omega_{i_3 j_4}\\
F_{32\ell s}&=\sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}}\sigma_{i_1\ell}^2\sigma_{i_3\ell}^2\sum_{j_1\neq j_3}\big[\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{i_1 j_1}\Omega_{i_3 j_3}+\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{i_1 j_1}\Omega_{i_1 j_3}\Omega_{i_3 j_1}\Omega_{i_3 j_3}\big]
\end{aligned}
$$

$$
\begin{aligned}
F_{42\ell s} = & \sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^2\sigma_{i_3\ell}^2 \sum_{j_1,j_3,j_4\, dist.} \big[\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4}\Omega_{i_1j_1}\Omega_{i_3j_3}\Omega_{i_3j_4} \\
& \qquad\qquad\qquad\qquad\qquad + \Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4}\Omega_{i_1j_1}\Omega_{i_1j_3}\Omega_{i_3j_1}\Omega_{i_3j_4}\big] \\
F_{52\ell s} = & \sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^2\sigma_{i_3\ell}^2 \sum_{j_1,j_2,j_3,j_4\, dist.} \Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4}\Omega_{i_1j_1}\Omega_{i_1j_2}\Omega_{i_3j_3}\Omega_{i_3j_4}
\end{aligned}
$$

Note that $\sum_{x=1}^2 F_{tx\ell s} = F_{t\ell s}$ for all $t \in [5]$. Using the fact that $\sum_j \Omega_{ij} = 1$, we have

$$
\sum_t F_{t1\ell s} \lesssim F_{11\ell s} = \sum_{\omega\in\mathcal{I}_{\ell-1}} \sigma_{i\ell}^4 \sum_j \Omega_{\ell j}\Omega_{ij} = \sum_{\omega\in\mathcal{I}_{\ell-1}} \sigma_{i\ell}^4 \langle\Omega_\ell, \Omega_i\rangle. \tag{E.67}
$$

To control $\sum_t F_{t2\ell s}$, observe that, since $\Omega_{ij} \le 1$ for all i, j,

$$
\begin{aligned}
&\sum_j \Omega_{\ell j}\Omega_{i_1 j} = \langle\Omega_\ell, \Omega_{i_1}\circ\Omega_{i_3}\rangle \\
&\sum_{j_1\neq j_4} \Omega_{\ell j_1}\Omega_{\ell j_4}\Omega_{i_1j_1}\Omega_{i_3j_1}\Omega_{i_3j_4} \le \langle\Omega_\ell, \Omega_{i_1}\circ\Omega_{i_3}\rangle\cdot\langle\Omega_\ell, \Omega_{i_3}\rangle \\
&\sum_{j_1\neq j_3} \big[\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{i_1j_1}\Omega_{i_3j_3} + \Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{i_1j_1}\Omega_{i_1j_3}\Omega_{i_3j_1}\Omega_{i_3j_3}\big] \le 2\langle\Omega_\ell, \Omega_{i_1}\rangle\cdot\langle\Omega_\ell, \Omega_{i_3}\rangle \\
&\sum_{j_1,j_3,j_4\, dist.} \big[\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4}\Omega_{i_1j_1}\Omega_{i_3j_3}\Omega_{i_3j_4} + \Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4}\Omega_{i_1j_1}\Omega_{i_1j_3}\Omega_{i_3j_1}\Omega_{i_3j_4}\big] \le 2\langle\Omega_\ell, \Omega_{i_1}\rangle\langle\Omega_\ell, \Omega_{i_3}\rangle^2 \\
&\sum_{j_1,j_2,j_3,j_4\, dist.} \Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4}\Omega_{i_1j_1}\Omega_{i_1j_2}\Omega_{i_3j_3}\Omega_{i_3j_4} \le \langle\Omega_\ell, \Omega_{i_1}\rangle^2\langle\Omega_\ell, \Omega_{i_3}\rangle^2.
\end{aligned}
$$

These bounds are relatively sharp, and it is clear that the first and third lines dominate. Furthermore as. Hence,

$$
\sum_t F_{t2\ell s} \lesssim F_{12\ell s} + F_{32\ell s} \lesssim \sum_{\omega_1\neq\omega_3\in\mathcal{I}_{\ell-1}} \sigma_{i_1\ell}^2\sigma_{i_3\ell}^2 \big[\langle\Omega_\ell, \Omega_{i_1}\circ\Omega_{i_3}\rangle + \langle\Omega_\ell, \Omega_{i_1}\rangle\cdot\langle\Omega_\ell, \Omega_{i_3}\rangle\big]. \tag{E.68}
$$

Observe that if $\ell \in S_k$, then

$$
\sum_\omega \sigma_{i\ell}^4 \Omega_{ij} \le \sum_{i\in S_k} \frac{1}{n_k^4 \bar N_k^4} N_i\Omega_{ij} + \sum_{k'=1}^K \sum_{i\in S_{k'}} \frac{1}{n^4\bar N^4} N_i\Omega_{ij} \tag{E.69}
$$

$$
\le \frac{1}{n_k^3\bar N_k^3}\mu_{kj} + \frac{1}{n^3\bar N^3}\mu_j, \tag{E.70}
$$

and

$$
\begin{aligned}
\sum_\omega \sigma_{i\ell}^2 \Omega_{ij} &\le \sum_{i\in S_k} \frac{1}{n_k^2 \bar N_k^2} N_i\Omega_{ij} + \sum_{k'=1}^K \sum_{i\in S_{k'}} \frac{1}{n\bar N} N_i\Omega_{ij} \\
&\le \frac{1}{n_k\bar N_k}\mu_{kj} + \frac{1}{n\bar N}\mu_j.
\end{aligned}
$$

Next,

$$
\begin{aligned}
\sum_{(\ell,s)}\sum_t F_{t1\ell s} &\lesssim \sum_{(\ell,s)}\sum_{\omega\in\mathcal{I}_{\ell-1}} \sigma_{i\ell}^4\langle\Omega_\ell, \Omega_i\rangle. \\
&\lesssim \sum_{(\ell,s)}\sum_j \Omega_{\ell j}\big(\frac{1}{n_k^3\bar N_k^3}\mu_{kj} + \frac{1}{n^3\bar N^3}\mu_j\big) \\
&\lesssim \sum_j\sum_k \frac{1}{n_k^2\bar N_k^2}\mu_{kj}^2 + \sum_j\sum_k \frac{1}{n^2\bar N^2}\mu_j^2 \lesssim \sum_k \frac{1}{n_k^2\bar N_k^2}\|\mu_k\|^2,
\end{aligned} \tag{E.71}
$$

where we applied that $\|\mu\|^2 \lesssim \sum_k \|\mu_k\|^2$ (see (D.49)). Furthermore,

$$
\begin{aligned}
\sum_{(\ell,s)} \sum_t F_{t2\ell s} &\le \sum_{k=1}^K \sum_{\ell\in S_k} N_\ell \sum_{\omega_1,\omega_3} \sigma^2_{i_1\ell}\sigma^2_{i_3\ell}\big[\langle \Omega_\ell, \Omega_{i_1}\circ\Omega_{i_3}\rangle + \langle \Omega_\ell, \Omega_{i_1}\rangle \cdot \langle \Omega_\ell, \Omega_{i_3}\rangle\big] \\
&\lesssim \sum_k \sum_{\ell\in S_k} N_\ell \Big[\sum_j \Omega_{\ell_j}\big(\frac{1}{n_k\bar N_k}\mu_{kj} + \frac{1}{n\bar N}\mu_j\big)^2 + \Big(\sum_j \Omega_{\ell j}\cdot \big(\frac{1}{n_k\bar N_k}\mu_{kj} + \frac{1}{n\bar N}\mu_j\big)\Big)^2\Big] \\
&\lesssim \sum_k \sum_{\ell\in S_k} N_\ell \sum_j \Omega_{\ell_j}\big(\frac{1}{n_k\bar N_k}\mu_{kj} + \frac{1}{n\bar N}\mu_j\big)^2
\end{aligned}
$$

In the last line we apply Cauchy–Schwarz. Continuing, we have

$$
\begin{aligned}
\sum_{(\ell,s)} \sum_t F_{t2\ell s} &\lesssim \sum_k \sum_{\ell\in S_k} N_\ell \sum_j \Omega_{\ell_j}\big(\frac{1}{n_k\bar N_k}\mu_{kj} + \frac{1}{n\bar N}\mu_j\big)^2 \\
&\lesssim \sum_k \sum_{\ell\in S_k} N_\ell \sum_j \Omega_{\ell_j}\big(\frac{1}{n_k\bar N_k}\mu_{kj}\big)^2 + \sum_k \sum_{\ell\in S_k} N_\ell \sum_j \Omega_{\ell_j}\big(\frac{1}{n\bar N}\mu_j\big)^2 \\
&\lesssim \sum_k \frac{\|\mu_k\|_3^3}{n_k\bar N_k} + \sum_k \frac{\|\mu\|_3^3}{n\bar N} \lesssim \sum_k \frac{\|\mu_k\|_3^3}{n_k\bar N_k}, \qquad \text{(E.72)}
\end{aligned}
$$

where we applied (D.68). Combining (E.66), (E.71) and (E.72), we have

$$
\sum_{(\ell,s)} \mathbb{E}D^4_{\ell,s} \lesssim \sum_{(\ell,s)} \sum_{x=1}^2 \sum_{t=1}^5 F_{tx\ell s} \lesssim \sum_k \frac{\|\mu_k\|^2}{n_k^2\bar N_k^2} + \sum_k \frac{\|\mu_k\|_3^3}{n_k\bar N_k},
$$

as desired.

E.6 Proof of Lemma E.4

$$
\mathrm{Var}\Big[\sum_{(\ell,s)} \mathrm{Var}(\tilde E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)})\Big] \to 0 \qquad \text{(E.73)}
$$

Next we study (E.73). We have

$$
\begin{aligned}
\mathrm{Var}(E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) = \mathbb{E}[E^2_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}] &= \sigma^2_\ell \sum_{r,r'\in[s-1]} \sum_{j,j'} \mathbb{E}\big[Z_{\ell jr}Z_{\ell js}Z_{\ell j'r'}Z_{\ell j's}\big|\mathcal{F}_{\prec(\ell,s)}\big] \\
&= \sigma^2_\ell \sum_{r,r'\in[s-1]} \sum_{j,j'} Z_{\ell jr}Z_{\ell j'r'}\mathbb{E}[Z_{\ell js}Z_{\ell j's}] \\
&= \sigma^2_\ell \sum_{r,r'\in[s-1]} \sum_{j,j'} \delta_{jj'\ell} Z_{\ell jr}Z_{\ell j'r'}, \qquad \text{(E.74)}
\end{aligned}
$$

where we let

$$
\delta_{jj'\ell} = \mathbb{E}Z_{\ell js}Z_{\ell j's} = \begin{cases} \Omega_{\ell j}(1-\Omega_{\ell j}) & \text{if } j=j' \\ -\Omega_{\ell j}\Omega_{\ell j'} & \text{else.}\end{cases} \qquad \text{(E.75)}
$$

Define

$$
\varphi_{\ell r\ell r'} = \sum_{j,j'} \delta_{jj'\ell} Z_{\ell jr}Z_{\ell j'r'}. \qquad \text{(E.76)}
$$

By (E.74) we have

$$
\begin{aligned}
\sum_{(\ell,s)} \mathrm{Var}(E_{\ell,s}|\mathcal{F}_{\prec(\ell,s)}) &= \sum_{\ell=1}^{n}\sum_{s=1}^{N_\ell}\sum_{r,r'\in[s-1]} \sigma_\ell^2\,\varphi_{\ell r\ell r'} \\
&= \sum_{\ell=1}^{n}\sum_{s=1}^{N_\ell}\Big[\sum_{r\in[s-1]} \sigma_\ell^2\,\varphi_{\ell r\ell r} + 2\sum_{r<r'\in[s-1]} \sigma_\ell^2\,\varphi_{\ell r\ell r'}\Big] \\
&= \sum_{\ell=1}^{n}\sum_{r=1}^{N_\ell}\sum_{s\in[N_\ell]:s>r} \sigma_\ell^2\,\varphi_{\ell r\ell r} + 2\sum_{\ell=1}^{s}\sum_{r<r'\in[N_\ell]}\sum_{s\in[N_\ell]:s>r'} \sigma_\ell^2\,\varphi_{\ell r\ell r'} \\
&= \sum_{\ell=1}^{n}\sum_{r=1}^{N_\ell}(N_\ell-r)\sigma_\ell^2\,\varphi_{\ell r\ell r} + 2\sum_{\ell=1}^{s}\sum_{r<r'\in[N_\ell]}(N_\ell-r')\sigma_\ell^2\,\varphi_{\ell r\ell r'} \\
&\equiv S_1 + S_2.
\end{aligned}
$$

Observe that S_1 and S_2 are uncorrelated. In addition, the terms in the summation defining S_1 are uncorrelated; the same holds for S_2 also.

First we study S_2. Next,

$$
\begin{aligned}
\mathbb{E}\varphi_{\ell r\ell r'}^2 &= \sum_{j_1,j_2,j_3,j_4} \delta_{j_1j_2,\ell}\delta_{j_3j_4,\ell}\,\mathbb{E}Z_{\ell j_1r}Z_{\ell j_2r'}Z_{\ell j_3r}Z_{\ell j_4r'} \\
&= \sum_{j_1,j_2,j_3,j_4} \delta_{j_1j_2\ell}\delta_{j_3j_4\ell}\mathbb{E}Z_{\ell j_1r}Z_{\ell j_3r}\mathbb{E}Z_{\ell j_2r'}Z_{\ell j_4r'}.
\end{aligned} \tag{E.77}
$$

First we study V_2. By casework,

$$
|\delta_{j_1j_2\ell}\delta_{j_3j_4\ell}\mathbb{E}Z_{\ell j_1r}Z_{\ell j_3r}\mathbb{E}Z_{\ell j_2r'}Z_{\ell j_4r'}| \tag{E.78}
$$

$$
= \begin{cases}
\delta_{jj\ell}^2\mathbb{E}Z_{\ell jr}^2\mathbb{E}Z_{\ell jr'}^2 \lesssim \Omega_{\ell j}^4 & \text{if } j_1=\cdots=j_4 \\
\delta_{j_1j_1\ell}\delta_{j_1j_4\ell}|\mathbb{E}Z_{\ell j_1r}^2\mathbb{E}Z_{\ell j_1r'}Z_{\ell j_4r'}| \lesssim \Omega_{\ell j_1}^4\Omega_{\ell j_4}^2 & \text{if } j_1=j_2=j_3, j_1\neq j_4 \\
\delta_{j_1j_1\ell}\delta_{j_3j_3\ell}\mathbb{E}Z_{\ell j_1r}Z_{\ell j_3r}\mathbb{E}Z_{\ell j_1r'}Z_{\ell j_3r'} \lesssim \Omega_{\ell j_1}^3\Omega_{\ell j_3}^3 & \text{if } j_1=j_2, j_3=j_4, j_1\neq j_3 \\
\delta_{j_1j_2\ell}^2\mathbb{E}Z_{\ell j_1r}^2\mathbb{E}Z_{\ell j_2r'}^2 \lesssim \Omega_{\ell j_1}^3\Omega_{\ell j_2}^3 & \text{if } j_1=j_3, j_2=j_4, j_1\neq j_2 \\
\delta_{j_1j_1\ell}\delta_{j_3j_4\ell}\mathbb{E}Z_{\ell j_1r}Z_{\ell j_3r}\mathbb{E}Z_{\ell j_1r'}Z_{\ell j_4r'} \lesssim \Omega_{\ell j_1}^3\Omega_{\ell j_3}^2\Omega_{\ell j_4}^2 & \text{if } j_1=j_2, j_1,j_3,j_4 \; dist. \\
\delta_{j_1j_2\ell}\delta_{j_1j_4\ell}\mathbb{E}Z_{\ell j_1r}^2\mathbb{E}Z_{\ell j_2r'}Z_{\ell j_4r'} \lesssim \Omega_{\ell j_1}^3\Omega_{\ell j_2}^2\Omega_{\ell j_4}^2 & \text{if } j_1=j_3, j_1,j_2,j_4 \; dist. \\
\delta_{j_1j_2\ell}\delta_{j_3j_4\ell}\mathbb{E}Z_{\ell j_1r}Z_{\ell j_3r}\mathbb{E}Z_{\ell j_2r'}Z_{\ell j_4r'} \lesssim \Omega_{\ell j_1}^2\Omega_{\ell j_2}^2\Omega_{\ell j_3}^2\Omega_{\ell j_4}^2 & \text{if } j_1,j_2,j_3,j_4 \; dist.
\end{cases}
$$

Up to permutation of the indices $j_1,\dots,j_4$, all nonzero terms of (E.77) take one of the forms above. By (E.78) and Cauchy–Schwarz, we have

$$
\mathbb{E}\varphi_{\ell r\ell r'}^2 \lesssim \|\Omega_\ell\|_4^4 + \|\Omega_\ell\|_4^4\|\Omega_\ell\|^2 + 2\|\Omega_\ell\|_3^6 + 2\|\Omega_\ell\|_3^3\|\Omega_\ell\|^4 + \|\Omega_\ell\|^8 \lesssim \|\Omega_\ell\|_4^4. \tag{E.79}
$$

Recalling that $\{\varphi_{\ell r\ell r'}\}_{\ell,r<r'\in[N_\ell]}$ are mutually uncorrelated, it follows that

$$
\begin{aligned}
\mathrm{Var}(S_2) &\lesssim \sum_{\ell}\sum_{r<r'\in[N_\ell]}(N_\ell-r')^2\sigma_\ell^2\mathbb{E}\varphi_{\ell r\ell r'}^2 \\
&\lesssim \sum_{\ell}\sum_{r<r'\in[N_\ell]}(N_\ell-r')^2\sigma_\ell^4\|\Omega_\ell\|_4^4 \\
&\lesssim \sum_{k}\sum_{\ell\in S_k} N_\ell^4\cdot\frac{1}{n_k^4\bar{N}_k^4}\|\Omega_\ell\|_4^4.
\end{aligned} \tag{E.80}
$$

Next we study S_1. We have

$$\mathbb{E}\varphi_{\ell r\ell r}^2 = \sum_{j_1,j_2,j_3,j_4} \delta_{j_1j_2\ell}\delta_{j_3j_4\ell}\mathbb{E}Z_{\ell j_1r}Z_{\ell j_2r}Z_{\ell j_3r}Z_{\ell j_4r}.$$

We have the following bounds by casework.

$$\begin{aligned}
&|\delta_{j_1j_2\ell}\delta_{j_3j_4\ell}\mathbb{E}Z_{\ell j_1r}Z_{\ell j_2r}Z_{\ell j_3r}Z_{\ell j_4r}| \\
&= \begin{cases}
\delta_{jj\ell}^2\mathbb{E}Z_{\ell jr}^4 \lesssim \Omega_{\ell j}^3 & \text{if } j_1 = \cdots = j_4 \\
\delta_{j_1j_1\ell}\delta_{j_1j_4\ell}|\mathbb{E}Z_{\ell j_1r}^3Z_{\ell j_4r}| \lesssim \Omega_{\ell j_1}^3\Omega_{\ell j_4}^2 & \text{if } j_1 = j_2 = j_3, j_1 \neq j_4 \\
\delta_{j_1j_1\ell}\delta_{j_3j_3\ell}\mathbb{E}Z_{\ell j_1r}^2Z_{\ell j_3r}^2 \lesssim \Omega_{\ell j_1}^2\Omega_{\ell j_3}^2 & \text{if } j_1 = j_2, j_3 = j_4, j_1 \neq j_3 \\
\delta_{j_1j_2\ell}^2\mathbb{E}Z_{\ell j_1r}^2Z_{\ell j_2r}^2 \lesssim \Omega_{\ell j_1}^3\Omega_{\ell j_2}^3 & \text{if } j_1 = j_3, j_2 = j_4, j_1 \neq j_3 \\
\delta_{j_1j_1\ell}\delta_{j_3j_4\ell}|\mathbb{E}Z_{\ell j_1r}^2Z_{\ell j_3r}Z_{\ell j_4r}| \lesssim \Omega_{\ell j_1}^2\Omega_{\ell j_3}^2\Omega_{\ell j_4}^2 & \text{if } j_1 = j_2, j_1, j_3, j_4 \textit{ dist.} \\
\delta_{j_1j_2\ell}\delta_{j_1j_4\ell}|\mathbb{E}Z_{\ell j_1r}^2Z_{\ell j_2r}Z_{\ell j_4}| \lesssim \Omega_{\ell j_1}^3\Omega_{\ell j_2}^2\Omega_{\ell j_4}^2 & \text{if } j_1 = j_3, j_1, j_2, j_4 \textit{ dist.} \\
\delta_{j_1j_2\ell}\delta_{j_3j_4\ell}|\mathbb{E}Z_{\ell j_1r}Z_{\ell j_2r}Z_{\ell j_3r}Z_{\ell j_4r}| \lesssim \Omega_{\ell j_1}^2\Omega_{\ell j_2}^2\Omega_{\ell j_3}^2\Omega_{\ell j_4}^2 & \text{if } j_1, j_2, j_3, j_4 \textit{ dist.}
\end{cases}
\end{aligned} \tag{E.81}$$

Up to symmetry, this accounts for all possible (nonzero) cases. Hence by Cauchy–Schwarz,

$$\mathbb{E}\varphi_{\ell r\ell r}^2 \lesssim \|\Omega_\ell\|_3^3 + \|\Omega_\ell\|_3^3\|\Omega_\ell\|^2 + \|\Omega_\ell\|^4 + \|\Omega_\ell\|_3^6 + \|\Omega_\ell\|^6 + \|\Omega_\ell\|_3^3\|\Omega_\ell\|^4 + \|\Omega_\ell\|^8 \lesssim \|\Omega_\ell\|_3^3. \tag{E.82}$$

Recalling that $\{\varphi_{\ell r\ell r}\}_{\ell, r\in[N_\ell]}$ is an uncorrelated collection of random variables, we have

$$\begin{aligned}
\mathrm{Var}(S_1) &\lesssim \sum_\ell \sum_{r\in[N_\ell]} (N_\ell - r)^2\sigma_\ell^4\mathbb{E}\varphi_{\ell r\ell r}^2 \\
&\lesssim \sum_\ell \sum_{r\in[N_\ell]} (N_\ell - r)^2\sigma_\ell^4\|\Omega_\ell\|_3^3 \\
&\lesssim \sum_k \sum_{\ell\in S_k} N_\ell^3 \cdot \frac{1}{n_k^4\bar{N}_k^4}\|\Omega_\ell\|_3^3.
\end{aligned} \tag{E.83}$$

Combining (E.83) and (E.80) proves the result. □

E.7 Proof of Lemma E.5

We have

$$\begin{aligned}
\mathbb{E}E_{\ell,s}^4 &= \sum_{r_1,r_2,r_3,r_4\in[s-1]} \sigma_\ell^4 \sum_{j_1,j_2,j_3,j_4} \mathbb{E}Z_{\ell j_1r_1}Z_{\ell j_1s}Z_{\ell j_2r_2}Z_{\ell j_2s}Z_{\ell j_3r_3}Z_{\ell j_3s}Z_{\ell j_4r_4}Z_{\ell j_4s} \\
&= \sigma_\ell^4 \sum_{j_1,j_2,j_3,j_4} \Bigg[\mathbb{E}[Z_{\ell j_1s}Z_{\ell j_2s}Z_{\ell j_3s}Z_{\ell j_4s}] \cdot \underbrace{\sum_{r_1,r_2,r_3,r_4\in[s-1]} \mathbb{E}[Z_{\ell j_1r_1}Z_{\ell j_2r_2}Z_{\ell j_3r_3}Z_{\ell j_4r_4}]}_{=:B_{\ell,s;j_1,j_2,j_3,j_4}} \Bigg]
\end{aligned} \tag{E.84}$$

We have by exhaustive casework that

$$|\mathbb{E}[Z_{\ell j_1r_1}Z_{\ell j_2r_2}Z_{\ell j_3r_3}Z_{\ell j_4r_4}]| \tag{E.85}$$

$$
= \begin{cases}
\mathbb{E}Z^4_{\ell j_1 r_1} \lesssim \Omega_{\ell j_1} & \text{if } {\scriptstyle j_1=j_2=j_3=j_4; \atop r_1=r_2=r_3=r_4} \\
\mathbb{E}Z^2_{\ell j_1 r_1}\mathbb{E}Z^2_{\ell j_1 r_3} \lesssim \Omega^2_{\ell j_1} & \text{if } {\scriptstyle j_1=j_2=j_3=j_4; \atop r_1=r_2, r_3=r_4, r_1\neq r_3} \\
|\mathbb{E}[Z^3_{\ell j_1 r_1} Z_{\ell j_4 r_1}]| \lesssim \Omega_{\ell j_1}\Omega_{\ell j_4} & \text{if } {\scriptstyle j_1=j_2=j_3, j_1\neq j_4; \atop r_1=r_2=r_3=r_4} \\
|\mathbb{E}[Z^2_{\ell j_1 r_1}\mathbb{E}Z_{\ell j_1 r_3} Z_{\ell j_4 r_3}]| \lesssim \Omega^2_{\ell j_1}\Omega_{\ell j_4} & \text{if } {\scriptstyle j_1=j_2=j_3, j_1\neq j_4; \atop r_1=r_2, r_3=r_4, r_1\neq r_3} \\
|\mathbb{E}Z^2_{\ell j_1 r_1} Z^2_{\ell j_3 r_1}| \lesssim \Omega_{\ell j_1}\Omega_{\ell j_3} & \text{if } {\scriptstyle j_1=j_2, j_3=j_4, j_1\neq j_3; \atop r_1=r_2=r_3=r_4} \\
|\mathbb{E}[Z^2_{\ell j_1 r_1} Z^2_{\ell j_3 r_3}]| \lesssim \Omega_{\ell j_1}\Omega_{\ell j_3} & \text{if } {\scriptstyle j_1=j_2, j_3=j_4, j_1\neq j_3; \atop r_1=r_2, r_3=r_4, r_1\neq r_3} \\
|\mathbb{E}[Z_{\ell j_1 r_1} Z_{\ell j_3 r_1}\mathbb{E}Z_{\ell j_1 r_2} Z_{\ell j_3 r_2}]| \lesssim \Omega^2_{\ell j_1}\Omega^2_{\ell j_3} & \text{if } {\scriptstyle j_1=j_2, j_3=j_4, j_1\neq j_3; \atop r_1=r_3, r_2=r_4, r_1\neq r_2} \\
|\mathbb{E}[Z^2_{\ell j_1 r_1} Z_{\ell j_3 r_1} Z_{\ell j_4 r_1}]| \lesssim \Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4} & \text{if } {\scriptstyle j_1=j_2, j_1, j_3, j_4 \ dist.; \atop r_1=r_2=r_3=r_4} \\
|\mathbb{E}[Z^2_{\ell j_1 r_1}\mathbb{E}Z_{\ell j_3 r_3} Z_{\ell j_4 r_3}]| \lesssim \Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4} & \text{if } {\scriptstyle j_1=j_2, j_1, j_3, j_4 \ dist.; \atop r_1=r_2, r_3=r_4, r_1\neq r_3} \\
|\mathbb{E}[Z_{\ell j_1 r_1} Z_{\ell j_3 r_1}\mathbb{E}Z_{\ell j_1 r_2} Z_{\ell j_4 r_2}]| \lesssim \Omega^2_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4} & \text{if } {\scriptstyle j_1=j_2, j_1, j_3, j_4 \ dist.; \atop r_1=r_3, r_2=r_4, r_1\neq r_2} \\
|\mathbb{E}[Z_{\ell j_1 r_1} Z_{\ell j_2 r_1} Z_{\ell j_3 r_1} Z_{\ell j_4 r_1}]| \lesssim \Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4} & \text{if } {\scriptstyle j_1, j_2, j_3, j_4 \ dist; \atop r_1=r_2=r_3=r_4} \\
|\mathbb{E}[Z_{\ell j_1 r_1} Z_{\ell j_2 r_1}\mathbb{E}Z_{\ell j_3 r_3} Z_{\ell j_4 r_3}]| \lesssim \Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4} & \text{if } {\scriptstyle j_1, j_2, j_3, j_4 \ dist; \atop r_1=r_2, r_3=r_4, r_1\neq r_3}
\end{cases}
$$

Up to permutation of the indices j_1, j_2, j_3, j_4 and r_1, r_2, r_3, r_4, this accounts for all possible cases such that (E.85) is nonzero. Therefore,

$$
B_{\ell,s;j_1,j_2,j_3,j_4} \lesssim \begin{cases}
s\Omega_{\ell j_1} + s^2\Omega^2_{\ell j_1} & \text{if } j_1 = j_2 = j_3 = j_4 \\
s\Omega_{\ell j_1}\Omega_{\ell j_4} + s^2\Omega^2_{\ell j_1}\Omega_{\ell j_4} & \text{if } j_1 = j_2 = j_3, j_1 \neq j_4 \\
s\Omega_{\ell j_1}\Omega_{\ell j_3} + s^2\Omega_{\ell j_1}\Omega_{\ell j_3} & \text{if } j_1 = j_2, j_3 = j_4, j_1 \neq j_3 \\
s\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4} + s^2\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4} & \text{if } j_1 = j_2, j_1, j_3, j_4 \ dist. \\
s\Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4} + s^2\Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4} & \text{if } j_1, j_2, j_3, j_4 \ dist.
\end{cases}
$$

Up to permutation of j_1, j_2, j_3, j_4, this accounts for all possible cases. Returning to (E.84), we have by applying (E.55) and the previous display that

$$
\begin{aligned}
\mathbb{E}E^4_{\ell,s} \lesssim \sigma^4_\ell \Bigg(& \sum_j \Omega_{\ell j}(s\Omega_{\ell j} + s^2\Omega^2_{\ell j}) + \sum_{j_1 \neq j_4} \Omega_{\ell j_1}\Omega_{\ell j_4}(s\Omega_{\ell j_1}\Omega_{\ell j_4} + s^2\Omega^2_{\ell j_1}\Omega_{\ell j_4}) \\
& + \sum_{j_1 \neq j_3} \Omega_{\ell j_1}\Omega_{\ell j_3}(s\Omega_{\ell j_1}\Omega_{\ell j_3} + s^2\Omega_{\ell j_1}\Omega_{\ell j_3}) \\
& + \sum_{j_1, j_3, j_4 (dist.)} \Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4}(s\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4} + s^2\Omega_{\ell j_1}\Omega_{\ell j_3}\Omega_{\ell j_4}) \\
& + \sum_{j_1, j_2, j_3, j_4 \ dist.} \Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4}(s\Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4} + s^2\Omega_{\ell j_1}\Omega_{\ell j_2}\Omega_{\ell j_3}\Omega_{\ell j_4}) \Bigg) \\
\lesssim \ & s\sigma^4_\ell \|\Omega_\ell\|^2 + s^2\sigma^4_\ell \|\Omega_\ell\|^3_3.
\end{aligned}
$$

In the third line we group the coefficients of s and s^2 and use the fact that $\|\Omega_\ell\|^4 \leq \|\Omega_\ell\|^3_3$ by Cauchy–Schwarz. Therefore

$$
\begin{aligned}
\sum_{(\ell,s)} \mathbb{E}E^4_{\ell,s} &\lesssim \sum_{(\ell,s)} s\sigma^4_\ell \|\Omega_\ell\|^2 + \sum_{(\ell,s)} s^2\sigma^4_\ell \|\Omega_\ell\|^3_3 \\
&= \sum_k \sum_{\ell \in S_k} \sum_{s \in [N_\ell]} s\sigma^4_\ell \|\Omega_\ell\|^2 + \sum_k \sum_{\ell \in S_k} \sum_{s \in [N_\ell]} s^2\sigma^4_\ell \|\Omega_\ell\|^3_3 \\
&\lesssim \sum_k \sum_{\ell \in S_k} N^2_\ell \cdot \frac{1}{n^4_k \bar{N}^4_k} \|\Omega_\ell\|^2 + \sum_k \sum_{\ell \in S_k} N^3_\ell \cdot \frac{1}{n^4_k \bar{N}^4_k} \|\Omega_\ell\|^3_3,
\end{aligned}
$$

as desired. □

E.8 Proof of Lemma E.6

We have

$$\sum_k \sum_{i\in S_k} \frac{N_i^2\|\Omega_i\|^2}{n_k^4\bar{N}_k^4} \le \sum_k \frac{1}{n_k^4\bar{N}_k^4} \sum_{i,m\in S_k} N_i N_m \langle \Omega_i, \Omega_m\rangle$$
$$= \sum_k \frac{1}{n_k^2\bar{N}_k^2}\|\mu_k\|^2,$$

which establishes the first claim.

Similarly,

$$\sum_k \sum_{i\in S_k} \frac{N_i^3\|\Omega_i\|_3^3}{n_k^4\bar{N}_k^4} \le \sum_k \frac{1}{n_k^4\bar{N}_k^4} \sum_{i,m,m'\in S_k} N_i N_m N_{m'} \sum_j \Omega_{ij}\Omega_{mj}\Omega_{m'j}$$
$$\le \sum_k \frac{1}{n_k\bar{N}_k}\|\mu_k\|_3^3,$$

which proves the second claim.

The third claim follows similarly and we omit the proof.

□

F Proofs of other main lemmas and theorems

F.1 Proof of Lemma 1

We start from computing $\mathbb{E}[(\hat{\mu}_{kj} - \hat{\mu}_j)^2]$. Write $X_{ij} = N_i(\Omega_{ij} + Y_{ij})$. It follows by elementary calculation that

$$\hat{\mu}_{kj} - \hat{\mu}_j = \mu_{kj} - \mu_j + \Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)\sum_{i\in S_k} N_i Y_{ij} - \frac{1}{n\bar{N}}\sum_{\ell:\ell\neq k}\sum_{i\in S_\ell} N_i Y_{ij}.$$

For different k, the variables $\sum_{i\in S_k} N_i Y_{ij}$ are independent of each other. It follows that

$$\begin{aligned}
\mathbb{E}[(\hat{\mu}_{kj} - \hat{\mu}_j)^2] &= (\mu_{kj} - \mu_j)^2 + \Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 \mathbb{E}\Big[\Big(\sum_{i\in S_k} N_i Y_{ij}\Big)^2\Big] + \sum_{\ell:\ell\neq k}\frac{1}{n^2\bar{N}^2}\mathbb{E}\Big[\Big(\sum_{i\in S_\ell} N_i Y_{ij}\Big)^2\Big] \\
&= (\mu_{kj} - \mu_j)^2 + \Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 \sum_{i\in S_k} N_i\Omega_{ij}(1-\Omega_{ij}) + \sum_{\ell:\ell\neq k}\frac{1}{n^2\bar{N}^2}\sum_{i\in S_\ell} N_i\Omega_{ij}(1-\Omega_{ij}) \\
&= (\mu_{kj} - \mu_j)^2 + \frac{1}{n_k^2\bar{N}_k^2}\Big(1 - \frac{n_k\bar{N}_k}{n\bar{N}}\Big)\sum_{i\in S_k} N_i\Omega_{ij}(1-\Omega_{ij}) \\
&\quad + \frac{1}{n^2\bar{N}^2}\Big[\Big(1 - \frac{n\bar{N}}{n_k\bar{N}_k}\Big)\sum_{i\in S_k} N_i\Omega_{ij}(1-\Omega_{ij}) + \sum_{\ell:\ell\neq k}\sum_{i\in S_\ell} N_i\Omega_{ij}(1-\Omega_{ij})\Big] \\
&= (\mu_{kj} - \mu_j)^2 + \frac{1}{n_k^2\bar{N}_k^2}\Big(1 - \frac{n_k\bar{N}_k}{n\bar{N}}\Big)\sum_{i\in S_k} N_i\Omega_{ij}(1-\Omega_{ij}) \\
&\quad - \frac{1}{n\bar{N}n_k\bar{N}_k}\underbrace{\Big[\sum_{i\in S_k} N_i\Omega_{ij}(1-\Omega_{ij}) - \frac{n_k\bar{N}_k}{n\bar{N}}\sum_{\ell=1}^K\sum_{i\in S_\ell} N_i\Omega_{ij}(1-\Omega_{ij})\Big]}_{\delta_{kj}}.
\end{aligned} \tag{F.1}$$

Since X_{ij} follows a binomial distribution, it is easy to see that $\mathbb{E}[X_{ij}] = N_i\Omega_{ij}$ and $\mathbb{E}[X_{ij}^2] = (\mathbb{E}[X_{ij}])^2 + \mathrm{Var}(X_{ij}) = N_i^2\Omega_{ij}^2 + N_i\Omega_{ij}(1-\Omega_{ij})$. Combining them gives

$$\mathbb{E}[X_{ij}(N_i - X_{ij})] = N_i(N_i-1)\Omega_{ij}(1-\Omega_{ij}). \tag{F.2}$$

Define

$$\hat{\zeta}_{kj} = (\hat{\mu}_{kj} - \hat{\mu}_j)^2 - \frac{1}{n_k^2\bar{N}_k^2}\Big(1 - \frac{n_k\bar{N}_k}{n\bar{N}}\Big)\sum_{i\in S_k}\frac{X_{ij}(N_i - X_{ij})}{N_i - 1},$$

It follows from (F.1)-(F.2) that

$$\mathbb{E}[\hat{\zeta}_{kj}] = (\mu_{kj} - \mu_j)^2 - \frac{1}{n\bar{N}n_k\bar{N}_k}\delta_{kj}. \tag{F.3}$$

We are ready to compute $\mathbb{E}[T]$. By definition, $T = \sum_{j=1}^p\sum_{k=1}^K n_k\bar{N}_k\hat{\zeta}_{kj}$ and $\rho^2 = \sum_{j,k}(\mu_{kj} - \mu_j)^2$. Consequently,

$$\mathbb{E}[T] = \sum_{j=1}^p\sum_{k=1}^K n_k\bar{N}_k\Big[(\mu_{kj} - \mu_j)^2 - \frac{1}{n\bar{N}n_k\bar{N}_k}\delta_{kj}\Big] = \rho^2 - \frac{1}{n\bar{N}}\sum_{j=1}^p\sum_{k=1}^K\delta_{kj}. \tag{F.4}$$

We use the definition of δ_{kj} in (F.1). It is seen that for each $1\leq j\leq p$,

$$\sum_{k=1}^K\delta_{kj} = \sum_{k=1}^K\sum_{i\in S_k}N_i\Omega_{ij}(1-\Omega_{ij}) - \Big(\sum_{k=1}^K\frac{n_k\bar{N}_k}{n\bar{N}}\Big)\sum_{\ell=1}^K\sum_{i\in S_\ell}N_i\Omega_{ij}(1-\Omega_{ij}) = 0. \tag{F.5}$$

Combining (F.4)-(F.5) gives $\mathbb{E}[T] = \rho^2$. This proves the claim. □

F.2 Proof of Theorem 3

First we show that

$$\mathrm{Var}(T) \lesssim \Theta_n \tag{F.6}$$

Recall

$$\begin{aligned}
\Theta_{n1} &= 4\sum_{k=1}^K\sum_{j=1}^p n_k\bar{N}_k(\mu_{kj} - \mu_j)^2\mu_{kj}\\
\Theta_{n2} &= 2\sum_{k=1}^K\sum_{i\in S_k}\sum_{j=1}^p\Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2\frac{N_i^3}{N_i-1}\Omega_{ij}^2\\
\Theta_{n3} &= \frac{2}{n^2\bar{N}^2}\sum_{1\leq k\neq\ell\leq K}\sum_{i\in S_k}\sum_{m\in S_\ell}\sum_{j=1}^p N_iN_m\Omega_{ij}\Omega_{mj}\\
\Theta_{n4} &= 2\sum_{k=1}^K\sum_{\substack{i\in S_k, m\in S_k,\\ i\neq m}}\sum_{j=1}^p\Big(\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}}\Big)^2 N_iN_m\Omega_{ij}\Omega_{mj}.
\end{aligned}$$

and that $\sum_{a=1}^4\Theta_{na} = \Theta_n$.

By Lemma D.2, we immediately have

$$\mathrm{Var}(\mathbf{1}_p'U_1) \leq \Theta_{n1}. \tag{F.7}$$

For U_2, it is shown in the Proof of Lemma D.3 that

$$\mathrm{Var}(\mathbf{1}_p'U_2) = 4\sum_{k=1}^K\sum_{i\in S_k}\sum_{1\leq r<s\leq N_i}\frac{\theta_i}{N_i(N_i-1)}\big[\|\Omega_i\|^2 + O(\|\Omega_i\|_3^3)\big].$$

Thus

$$\begin{aligned}\mathrm{Var}(\mathbf{1}_p' U_2) &\lesssim 4\sum_{k=1}^K \sum_{i\in S_k} \sum_{1\le r<s\le N_i} \frac{\theta_i}{N_i(N_i-1)}\|\Omega_i\|^2 \\ &= 2\sum_{k=1}^K \sum_{i\in S_k} \theta_i \|\Omega_i\|^2 = \Theta_{n2} \end{aligned} \tag{F.8}$$

Next we study U_3. Using that $\Omega_{mj'} \le 1$ and $\|\Omega_i\|_1 = 1$, we have

$$\begin{aligned}\sum_{k\neq\ell} \frac{n_k n_\ell \bar{N}_k \bar{N}_\ell}{n^2\bar{N}^2} \mathbf{1}_p'(\Sigma_k \circ \Sigma_\ell)\mathbf{1}_p &= \frac{2}{n^2\bar{N}^2}\sum_{k\neq\ell}\sum_{i\in S_k}\sum_{m\in S_\ell}\sum_{j,j'} N_i N_m \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'} \\ &\le \frac{2}{n^2\bar{N}^2}\sum_{k\neq\ell}\sum_{i\in S_k}\sum_{m\in S_\ell}\sum_{j} N_i N_m \Omega_{ij}\Omega_{mj}\sum_{j'}\Omega_{ij'} \\ &= \frac{2}{n^2\bar{N}^2}\sum_{k\neq\ell}\sum_{i\in S_k}\sum_{m\in S_\ell}\sum_{j} N_i N_m \Omega_{ij}\Omega_{mj}. \end{aligned}$$

Therefore by Lemma D.4,

$$\mathrm{Var}(\mathbf{1}_p' U_3) \lesssim \frac{2}{n^2\bar{N}^2}\sum_{1\le k\neq\ell\le K}\sum_{i\in S_k}\sum_{m\in S_\ell}\sum_{j=1}^p N_i N_m \Omega_{ij}\Omega_{mj} = \Theta_{n3}. \tag{F.9}$$

Similarly for U_4, we have by the Proof of Lemma D.5 that

$$\begin{aligned}\mathrm{Var}(\mathbf{1}_p' U_4) &= 4\sum_{k=1}^K \sum_{\substack{i\in S_k, m\in S_k \\ i<m}} \kappa_{im}\Big(\sum_j \Omega_{ij}\Omega_{mj} + \delta_{im}\Big) \\ &\lesssim \sum_{k=1}^K \sum_{\substack{i\in S_k, m\in S_k \\ i<m}} \kappa_{im}\sum_j \Omega_{ij}\Omega_{mj} = \Theta_{n4}. \end{aligned} \tag{F.10}$$

Above we use that $|\delta_{im}| \le \sum_j \Omega_{ij}\Omega_{mj}$ and recall that $\kappa_{im} = (\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}})^2 N_i N_m$.

Observe that by Lemma 1,

$$\Theta_{n1} = 4\sum_{k=1}^K \sum_{j=1}^p n_k \bar{N}_k (\mu_{kj}-\mu_j)^2 \mu_{kj} \lesssim \max_k \|\mu_k\|_\infty \cdot \rho^2 = \max_k \|\mu_k\|_\infty \cdot \mathbb{E}\,T. \tag{F.11}$$

Since (21) holds, Lemma D.6 applies and

$$\Theta_{n_2} + \Theta_{n3} + \Theta_{n4} \asymp \sum_k \|\mu_k\|^2. \tag{F.12}$$

Combining (F.6), (F.11), and (F.12) proves the theorem. □

F.3 Proof of Theorem 4

To prove Theorem 4, we must prove the following claims:

(a) Under the alternative hypothesis, $\psi \to \infty$ in probability.

(b) For any fixed $\kappa \in (0,1)$, the level-κ DELVE test has an asymptotic level of κ and an asymptotic power of 1.

(c) If we choose $\kappa = \kappa_n$ such that $\kappa_n \to 0$ and $1 - \Phi(\mathrm{SNR}_n) = o(\kappa_n)$, where Φ is the CDF of $N(0,1)$, then the sum of type I and type II errors of the DELVE test converges to 0.

We show the first claim, that $\psi \to \infty$, under the alternative hypothesis and the conditions of Theorem 4. In particular, recall we assume that

$$\frac{\rho^2}{\sqrt{\sum_{k=1}^K \|\mu_k\|^2}} = \frac{n\bar{N}\|\mu\|^2\omega_n^2}{\sqrt{\sum_{k=1}^K \|\mu_k\|^2}} \to \infty. \tag{F.13}$$

Our first goal is to show that

$$T/\sqrt{\mathrm{Var}(T)} \xrightarrow{\mathbb{P}} \infty \tag{F.14}$$

under the alternative. By Chebyshev's inequality, it suffices to show that

$$\mathbb{E}T \gg \sqrt{\mathrm{Var}(T)}. \tag{F.15}$$

By Theorem 3,

$$\mathrm{Var}(T) \lesssim \sum_k \|\mu_k\|^2 + \max_k \|\mu_k\|_\infty \cdot \mathbb{E}T = \sum_k \|\mu_k\|^2 + \max_k \|\mu_k\|_\infty \cdot \rho^2 \tag{F.16}$$

By (F.13),

$$\mathbb{E}T = \rho^2 \gg \sqrt{\sum_{k=1}^K \|\mu_k\|^2} \geq \max_{1\le k\le K} \|\mu_k\|_\infty.$$

Therefore,

$$\sqrt{\max_{1\le k\le K} \|\mu_k\|_\infty} \cdot \rho \ll \rho^2 = \mathbb{E}T. \tag{F.17}$$

Moreover, by (F.13),

$$\sum_k \|\mu_k\|^2 \ll \rho^4 = (\mathbb{E}T)^2. \tag{F.18}$$

Combining (F.16), (F.17), and (F.18) implies (F.14).

Next we show that $V > 0$ with high probability (i.e., with probability tending to 1 as $n\bar{N} \to \infty$). Recall that by Lemmas D.6, D.10, and D.11,

$$\mathbb{E}V = \Theta_{n2} + \Theta_{n3} + \Theta_{n4} \gtrsim \sum_k \|\mu_k\|^2 > 0, \quad \text{and} \tag{F.19}$$

$$\mathrm{Var}(V) \lesssim \sum_k \frac{\|\mu_k\|^2}{n_k^2\bar{N}_k^2} \vee \sum_k \frac{\|\mu_k\|_3^3}{n_k\bar{N}_k}. \tag{F.20}$$

Using this, the Markov inequality, and (24), we have

$$\mathbb{P}\big(V < \mathbb{E}[V]/2\big) \le \mathbb{P}\big(|V - \mathbb{E}[V]| \ge \mathbb{E}[V]/2\big) \le \frac{4\mathrm{Var}(V)}{(\mathbb{E}[V])^2} = o(1), \tag{F.21}$$

which implies that $V > 0$ with high probability.

To finish the proof of the first claim, note that the assumptions of Proposition D.14 are satisfied and we have $V/\mathrm{Var}(T) = O_{\mathbb{P}}(1)$. By this, (F.14), and (F.21), we have

$$\psi = \frac{T\mathbf{1}_{V>0}}{\sqrt{V}} = \frac{\sqrt{\mathrm{Var}(T)}}{\sqrt{V}} \cdot \frac{T}{\sqrt{\mathrm{Var}(T)}} \cdot \mathbf{1}_{V>0} \gtrsim \frac{T}{\sqrt{\mathrm{Var}(T)}} \to \infty$$

in probability.

The second claim follows directly from the first claim and Theorem 2.

To prove the third claim, by Chebyshev's inequality and $T/\sqrt{\mathrm{Var}(T)} \to \infty$, it follows that $T > (1/2)\mathbb{E}T = (1/2)\rho^2$ with high probability as $n\bar{N} \to \infty$. By a similar Chebyshev argument as above, it also holds that $V < (3/2)\mathbb{E}V$ with high probability as $n\bar{N} \to \infty$. Recall that $\mathbb{E}V = \Theta_{n2} + \Theta_{n3} + \Theta_{n4} \lesssim \sum_k \|\mu_k\|^2$ by Lemmas D.6 and D.10. Thus, with high probability as $n\bar{N} \to \infty$, we have

$$\psi = T\mathbf{1}_{V>0}/\sqrt{V} \gtrsim \rho^2/\sqrt{\mathbb{E}V} \gtrsim \frac{n\bar{N}\|\mu\|^2\omega_n^2}{\sqrt{\sum_k \|\mu_k\|^2}} = \mathrm{SNR}_n.$$

Choosing α_n as specified yield the third claim. The proof is complete since all three claims are established.

□

F.4 Proof of Theorem 5

Without loss of generality, we assume p is even and write $m = p/2$. Let $\mu \in \mathbb{R}^m$ be a nonnegative vector with $\|\mu\|_1 = 1/2$. Let $\tilde{\mu} = (\mu', \mu')' \in \mathbb{R}^p$. We consider the null hypothesis:

$$H_0: \qquad \Omega_i = \tilde{\mu}, \qquad 1 \le i \le n. \tag{F.22}$$

We pair it with a random alternative hypothesis. Let $b_1, b_2, \ldots, b_m$ be a collection of i.i.d. Rademacher variables. Let $z_1, z_2, \ldots, z_K$ denote an independent collection of i.i.d. Rademacher random variables conditioned on the event $|\sum_k z_k| \le 100\sqrt{K}$. For a properly small sequence $\omega_n > 0$ of positive numbers, let

$$H_1: \qquad \Omega_{ij} = \begin{cases} \mu_j\big(1 + \omega_n (n_k\bar{N}_k)^{-1}\big(\frac{1}{K}\sum_{k\in K} n_k\bar{N}_k\big) z_k b_j\big), & \text{if } 1 \le j \le m, i \in S_k \\ \tilde{\mu}_j\big(1 - \omega_n (n_k\bar{N}_k)^{-1}\big(\frac{1}{K}\sum_{k\in K} n_k\bar{N}_k\big) z_k b_{j-m}\big), & \text{if } m+1 \le j \le 2m, i \in S_k \end{cases} \tag{F.23}$$

In this section we slightly abuse notation, using ω_n to refer to the (deterministic) sequence above and reserving $\omega(\Omega)$ for the random quantity

$$\omega(\Omega) = \sqrt{\frac{1}{n\bar{N}\|\mu\|^2}\sum_{k=1}^K n_k\bar{N}_k\|\mu_k - \mu\|^2}. \tag{F.24}$$

As long as

$$\omega_n \le \frac{\min_k n_k\bar{N}_k}{\frac{1}{K}\sum_{k\in[K]} n_k\bar{N}_k} = \frac{\min_k n_k\bar{N}_k}{n\bar{N}/K},$$

then $\Omega_{ij} \ge 0$ for all $i \in [n], j \in [p]$. Furthermore, for each $1 \le i \le n$, we have $\|\Omega_i\|_1 = 2\|\mu\|_1 = 1$. We suppose there exists a constant $c \in (0,1)$ such that

$$cK^{-1}n\bar{N} \le n_k\bar{N}_k \le c^{-1}K^{-1}n\bar{N} \quad \text{for all } k \in [K] \tag{F.25}$$

With (F.25) in hand, we may assume without loss of generality that

$$\omega_n \le c/2 \tag{F.26}$$

This assumption implies that (F.23) is well-defined and moreover $\Omega_{ij} \asymp \mu_j$.

Next we characterize the random quantity $\omega(\Omega)$ in terms of ω_n.

Lemma F.1. *Let $\omega^2(\Omega)$ be as in* (F.24)*. When Ω follows Model* (F.23)*, there exists a constant $c_1 \in (0,1)$ such that $c_1\omega_n^2 \le \omega^2(\Omega) \le c_1^{-1}\omega_n^2$ with probability* 1.

The proof of Lemma F.1 is given in Section F.4.1. By Lemma F.1, under the model (F.23) it holds with probability 1 that

$$\frac{n\bar{N}\|\mu\|^2\omega^2(\Omega)}{\sqrt{\sum_{k=1}^K \|\mu_k\|^2}} \asymp K^{-1/2} n\bar{N}\|\mu\|\omega_n^2. \tag{F.27}$$

Above we use that $\Omega_{ij} \asymp \mu_j$, since we assume (F.26)

We also require Proposition F.2 below, whose proof is given in Section F.4.2.

Proposition F.2. *Suppose that* (F.25) *and* (F.26) *hold. Consider the pair of hypotheses in* (F.22)-(F.23) *and let* $\mathbb{P}_0$*, and* $\mathbb{P}_1$ *be the respective probability measures. If*

$$\frac{n\bar{N}\|\mu\|^2\omega^2(\Omega)}{\sqrt{\sum_{k=1}^K \|\mu_k\|^2}} \asymp K^{-1/2} n\bar{N}\|\mu\|\omega_n^2 \to 0,$$

then the chi-square distance between $\mathbb{P}_0$ *and* $\mathbb{P}_1$ *converges to* 0*.*

Now we prove Theorem 5. Let δ_n denote an arbitrary sequence tending to 0. Without loss of generality, we may assume that $\delta_n \le c^*$ for a small absolute constant $c^* \in (0,1)$. Note that $K^{-1/2}n\bar{N} \ge 1$ since $K \le n$. Thus for appropriate choice of sequences of $\mu = \mu_n$ and $\omega_n \le c/2$ in models (F.22), (F.23) and applying (F.27), we obtain

$$2\delta_n \ge \frac{n\bar{N}\|\mu\|^2\omega^2(\Omega)}{\sqrt{\sum_{k=1}^K \|\mu_k\|^2}} \ge \delta_n. \tag{F.28}$$

Recall the definitions of $\mathcal{Q}_{0n}^*$ and $\mathcal{Q}_{1n}^*$ in (28). Let Π denote the distribution on $\xi = \{(N_i, \Omega_i, \ell_i)\} \in \mathcal{Q}_{1n}^*$ induced by (F.23). Let ξ_0 denote the parameter associated to the simple null hypothesis in (F.22) associated to our choice of μ and ω_n satisfying (F.28). We have by standard manipulations,

$$\begin{aligned}
\mathcal{R}(\mathcal{Q}_{0n}^*, \mathcal{Q}_{1n}^*) &:= \inf_{\Psi\in\{0,1\}} \Big\{ \sup_{\xi\in\mathcal{Q}_{0n}^*(c_0,\epsilon_n)} \mathbb{P}_\xi(\Psi=1) + \sup_{\xi\in\mathcal{Q}_{1n}^*(\delta_n;c_0,\epsilon_n)} \mathbb{P}_\xi(\Psi=0) \Big\} \\
&= \inf_{\Psi\in\{0,1\}} \Big\{ \sup_{\xi\in\mathcal{Q}_{0n}^*(c_0,\epsilon_n),\xi'\in\mathcal{Q}_{1n}^*(\delta_n;c_0,\epsilon_n)} \big[\mathbb{P}_\xi(\Psi=1) + \mathbb{P}_\xi(\Psi=0)\big] \\
&\ge \inf_{\Psi\in\{0,1\}} \Big\{ \sup_{\xi\in\mathcal{Q}_{0n}^*(c_0,\epsilon_n)} \mathbb{E}_{\xi'\sim\Pi}\Big[\mathbb{P}_\xi(\Psi=1) + \mathbb{P}_{\xi'}(\Psi=0)\Big]\Big\} \\
&\ge \inf_{\Psi\in\{0,1\}} \Big\{ \mathbb{E}_{\xi'\sim\Pi}\Big[\mathbb{P}_{\xi_0}(\Psi=1) + \mathbb{P}_{\xi'}(\Psi=0)\Big]\Big\} \\
&= \inf_{\Psi\in\{0,1\}} \Big\{ \mathbb{P}_0(\Psi=1) + \mathbb{P}_1(\Psi=0) \Big\}.
\end{aligned}$$

In the last line we recall the definition of $\mathbb{P}_0$ and $\mathbb{P}_1$ in (F.22) and (F.23), noting that for all events E,

$$\mathbb{P}_1(E) = \mathbb{E}_{\xi'\sim\pi}\,\mathbb{P}_{\xi'}(E).$$

Next, by the Neyman–Pearson lemma and the standard inequality $\mathrm{TV}(P,Q) \le \sqrt{\chi^2(P,Q)}$ (see e.g. Chapter 2 of Tsybakov [2008]),

$$\begin{aligned}
\mathcal{R}(\mathcal{Q}_{0n}^*, \mathcal{Q}_{1n}^*) &\ge \inf_{\Psi\in\{0,1\}} \Big\{ \mathbb{P}_0(\Psi=1) + \mathbb{P}_1(\Psi=0) \Big\} \\
&= 1 - \mathrm{TV}\big(\mathbb{P}_0, \mathbb{P}_1\big) \ge 1 - \sqrt{\chi^2(\mathbb{P}_0, \mathbb{P}_1)}.
\end{aligned}$$

By Proposition F.2, as $\delta_n \to 0$ we have $\chi^2(\mathbb{P}_0, \mathbb{P}_1) \to 0$ and thus $\mathcal{R}(\mathcal{Q}_{0n}^*, \mathcal{Q}_{1n}^*) \to 1$, as desired. □

F.4.1 Proof of Proposition F.1

Next, we perform a change of parameters that preserves the signal strength and chi-squared distance. The testing problem (F.22) and (F.23) has parameters $\Omega_{ij}, N_i, \bar{N}_k, n_k, n,$ and K. Let $\mathbb{P}_0$ and $\mathbb{P}_1$ denote the distributions corresponding to the null and alternative hypotheses, respectively. For each $k \in [K]$, we combine all documents in sample k to obtain new null and alternative distributions $\tilde{\mathbb{P}}_0$ and $\tilde{\mathbb{P}}_1$ with parameters $\tilde{\Omega}_{ij}, \tilde{N}_i, \bar{\tilde{N}}_i, \tilde{n}_i, \tilde{n}$, and $\tilde{K}$ such that

$$\begin{aligned}
&\tilde{K} = K = \tilde{n} & \\
&\tilde{N}_i = n_i \bar{N}_i & \text{for } i \in [\tilde{K}] \\
&\bar{\tilde{N}}_i \equiv \tilde{N}_i & \text{for } i \in [\tilde{K}] \\
&\tilde{n}_i = 1 & \text{for } i \in [\tilde{K}]\ .
\end{aligned} \tag{F.29}$$

For notational ease, we define $\tilde{N} := \bar{\tilde{N}} = \frac{1}{K}\sum_{k\in[K]} n_k \bar{N}_k$. Furthermore, we have $\tilde{\Omega}_i = \mu$ for all $i \in [\tilde{n}]$ under the null $\tilde{\Omega}_i = \mu_i$ for all $i \in [\tilde{n}]$ under the alternative. Explicitly, in the reparameterized model, we have the null hypothesis

$$H_0: \qquad \Omega_i = \tilde{\mu}, \qquad 1 \le i \le n. \tag{F.30}$$

and alternative hypothesis

$$H_1: \qquad \Omega_{ij} = \begin{cases} \mu_j\big(1 + \omega_n \tilde{N}_i^{-1}\tilde{N} z_i b_j\big), & \text{if } 1 \le j \le m, \\ \tilde{\mu}_j\big(1 - \omega_n \tilde{N}_i^{-1}\tilde{N} z_i b_{j-m}\big), & \text{if } m+1 \le j \le 2m. \end{cases} \tag{F.31}$$

for all $i \in [\tilde{K}] = [K] = [\tilde{n}]$. Observe that the likelihood ratio is preserved: $\frac{d\mathbb{P}_0}{d\mathbb{P}_1} = \frac{d\tilde{\mathbb{P}}_0}{d\tilde{\mathbb{P}}_1}$ and also $\omega(\Omega) = \omega(\widetilde{\Omega})$. For simplicity we work with this reparameterized model in this proof.

If $z_1, \ldots, z_{\tilde{n}}$ are independent Rademacher random variables then with probability at least 1/2 it holds that

$$|\sum_i z_i| \le 100\sqrt{\tilde{n}} \tag{F.32}$$

by Hoeffding's inequality. Recall that our random model is defined in (F.23) where (i) $z_1, \ldots, z_{\tilde{n}}$ are independent Rademacher random variables conditioned on the event $|\sum_i z_i| \le 100\sqrt{\tilde{n}}$, and (ii) $b_1, \ldots, b_m$ are independent Rademacher random variables.

Now we study $\omega^2(\widetilde{\Omega})$. For each $1 \le j \le m$, we have $\widetilde{\Omega}_{ij} = \mu_j(1 + \omega_n \tilde{N}_i^{-1}\tilde{N} z_i b_j)$. Define $\eta_j = (\tilde{n}\tilde{N})^{-1}\sum_{i=1}^{\tilde{n}} \tilde{N}_i \widetilde{\Omega}_{ij} = \mu_j(1 + \omega_n \bar{z} b_j)$ for $1 \le j \le m$ and $\eta_j = (\tilde{n}\tilde{N})^{-1}\sum_{i=1}^{\tilde{n}} \tilde{N}_i \widetilde{\Omega}_{ij} = \tilde{\mu}_j(1 - \omega_n \bar{z} b_j)$ for $m < j \le 2m$. We have

$$\begin{aligned}
\sum_{i=1}^{\tilde{n}}\sum_{j=1}^{p} \tilde{N}_i(\widetilde{\Omega}_{ij} - \eta_j)^2 &= 2\sum_{i=1}^{\tilde{n}}\sum_{j=1}^{m} \tilde{N}_i \cdot \mu_j^2 \omega_n^2 \frac{\tilde{N}^2}{\tilde{N}_i^2}(z_i - \bar{z})^2 b_j^2 \\
&= 2\omega_n^2 \tilde{N}^2 \|\mu\|^2 \sum_{i=1}^{\tilde{n}} \tilde{N}_i^{-1}(z_i - \bar{z})^2.
\end{aligned}$$

By (F.32), $|\bar{z}| \le 100\sqrt{\tilde{n}}$. Thus $|z_i - \bar{z}| \asymp 1$. Write $\tilde{N}_* = (\tilde{n}^{-1}\sum_{i=1}^{\tilde{n}} \tilde{N}_i^{-1})$. It follows that

$$\sum_{i=1}^{\tilde{n}}\sum_{j=1}^{p} \tilde{N}_i(\widetilde{\Omega}_{ij} - \eta_j)^2 \asymp \omega_n^2 \tilde{N}^2 \|\mu\|^2 \cdot \tilde{n}\tilde{N}_*^{-1}.$$

Note that $\tilde{N} \ge \tilde{N}_*$. Additionally, by assumption (F.25), $\tilde{N}_i \asymp \tilde{N} \le c^{-1}\tilde{N}_*$. It follows that

$$\sum_{i=1}^{\tilde{n}}\sum_{j=1}^{p} \tilde{N}_i(\widetilde{\Omega}_{ij} - \eta_j)^2 \asymp \tilde{n}\tilde{N}\|\mu\|^2\omega_n^2. \tag{F.33}$$

Moreover, $\|\eta\|^2 = \sum_{j=1}^{p} \mu_j^2(1+\omega_n \bar{z} b_j)^2$. By our conditioning on the event in (F.32),

$$|\omega_n \bar{z} b_j| \lesssim \omega_n \tilde{n}^{-1/2}.$$

Since $\omega_n \leq 1$ and $\sum_j b_j = 0$, we have

$$\|\eta\|^2 = \|\mu\|^2 + \sum_{j=1}^{p} \mu_j^2 \omega_n^2 \bar{z}^2 = \|\mu\|^2[1 + O(\tilde{n}^{-1})] \asymp \|\mu\|^2. \tag{F.34}$$

Hence

$$\omega^2(\widetilde{\Omega}) = \omega^2(\Omega) \asymp \omega_n^2, \qquad \text{where recall} \quad \omega(\widetilde{\Omega}) = \frac{\sum_{i=1}^{\tilde{n}} \sum_{j=1}^{p} \tilde{N}_i(\widetilde{\Omega}_{ij} - \eta_j)^2}{\tilde{n}\tilde{N}\|\eta\|^2}. \tag{F.35}$$

This finishes the proof. □

F.4.2 Proof of Proposition F.2

In this proof, we continue to employ the reparametrization in (F.29). As discussed there, this reparametrization preserves the likelihood ratio and thus the chi-square distance.

By definition, $\chi^2(\mathbb{P}_0, \mathbb{P}_1) = \int (\frac{d\mathbb{P}_1}{d\mathbb{P}_0})^2 d\mathbb{P}_0 - 1$. It suffices to show that

$$\int \Big(\frac{d\mathbb{P}_1}{d\mathbb{P}_0}\Big)^2 d\mathbb{P}_0 = 1 + o(1). \tag{F.36}$$

From the density of of multinomial distribution, $d\mathbb{P}_0 = \prod_{i,j} \tilde{\mu}_j^{X_{ij}}$, and $d\mathbb{P}_1 = \mathbb{E}_{b,z}[\prod_{i,j} \widetilde{\Omega}_{ij}^{X_{ij}}]$. It follows that

$$\frac{d\mathbb{P}_1}{d\mathbb{P}_0} = \mathbb{E}_{b,z}\Big[\prod_{i=1}^{\tilde{n}} \prod_{j=1}^{p} \Big(\frac{\widetilde{\Omega}_{ij}}{\tilde{\mu}_j}\Big)^{X_{ij}}\Big].$$

Let $b^{(0)} = (b_1^{(0)}, \ldots, b_m^{(0)})'$ and $z^{(0)} = (z_1^{(0)}, \cdots, z_{\tilde{n}}^{(0)})'$ be independent copies of b and z. We construct $\widetilde{\Omega}_{ij}^{(0)}$ similarly as in (F.31). It is seen that

$$\begin{aligned}
\int \Big(\frac{d\mathbb{P}_1}{d\mathbb{P}_0}\Big)^2 d\mathbb{P}_0 &= \mathbb{E}_X \mathbb{E}_{b,z,b^{(0)},z^{(0)}}\Big[\prod_{i=1}^{\tilde{n}} \prod_{j=1}^{p} \Big(\frac{\widetilde{\Omega}_{ij}\widetilde{\Omega}_{ij}^{(0)}}{\tilde{\mu}_j^2}\Big)^{X_{ij}}\Big] \\
&= \mathbb{E}_{b,z,b^{(0)},z^{(0)}}\Big\{\prod_{i=1}^{\tilde{n}} \mathbb{E}_{X_i}\Big[\prod_{j=1}^{p} \Big(\frac{\widetilde{\Omega}_{ij}\widetilde{\Omega}_{ij}^{(0)}}{\tilde{\mu}_j^2}\Big)^{X_{ij}}\Big]\Big\} \\
&= \mathbb{E}_{b,z,b^{(0)},z^{(0)}}\Big\{\prod_{i=1}^{\tilde{n}} \Big(\sum_{j=1}^{p} \tilde{\mu}_j \cdot \frac{\widetilde{\Omega}_{ij}\widetilde{\Omega}_{ij}^{(0)}}{\tilde{\mu}_j^2}\Big)^{\tilde{N}_i}\Big]\Big\} \\
&= \mathbb{E}[\exp(M)], \quad \text{with} \quad M := \sum_{i=1}^{\tilde{n}} \tilde{N}_i \log\Big(\sum_{j=1}^{p} \tilde{\mu}_j^{-1} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{ij}^{(0)}\Big).
\end{aligned} \tag{F.37}$$

Here, the third line follows from the moment generating function of a multinomial distribution. We plug in the expression of $\widetilde{\Omega}_{ij}$ in (F.23). By direct calculations,

$$\begin{aligned}
\sum_{j=1}^{p} \tilde{\mu}_j^{-1} \widetilde{\Omega}_{ij}\widetilde{\Omega}_{ij}^{(0)} &= \sum_{j=1}^{m} \mu_j \big(1 + \omega_n \tilde{N}_i^{-1} \tilde{N} z_i b_j\big)\big(1 + \omega_n \tilde{N}_i^{-1} \tilde{N} z_i^{(0)} b_j^{(0)}\big) \\
&\quad + \sum_{j=1}^{m} \mu_j \big(1 - \omega_n \tilde{N}_i^{-1} \tilde{N} z_i b_j\big)\big(1 - \omega_n \tilde{N}_i^{-1} \tilde{N} z_i^{(0)} b_j^{(0)}\big)
\end{aligned}$$

$$= 2\|\mu\|_1 + 2\sum_{j=1}^m \mu_j \omega_n^2 \tilde{N}_i^{-2}\tilde{N}^2 z_i z_i^{(0)} b_j b_j^{(0)}$$

$$= 1 + 2\sum_{j=1}^m \mu_j \omega_n^2 \tilde{N}_i^{-2}\tilde{N}^2 z_i z_i^{(0)} b_j b_j^{(0)}.$$

We plug it into M and notice that $\log(1+t) \le t$ is always true. It follows that

$$M \le \sum_{i=1}^{\tilde{n}} \tilde{N}_i \cdot 2\sum_{j=1}^m \mu_j \omega_n^2 \frac{\tilde{N}^2}{\tilde{N}_i^2} z_i z_i^{(0)} b_j b_j^{(0)} = 2\tilde{N}\omega_n^2 \Big(\sum_{i=1}^{\tilde{n}} \frac{\tilde{N}}{\tilde{N}_i} z_i z_i^{(0)}\Big)\Big(\sum_{j=1}^m \mu_j b_j b_j^{(0)}\Big) =: M^*. \quad \text{(F.38)}$$

We combine (F.38) with (F.37). It is seen that to show (F.36), it suffices to show that

$$\mathbb{E}[\exp(M^*)] = 1 + o(1). \quad \text{(F.39)}$$

We now show (F.39). Write $M_1 = \sum_{i=1}^{\tilde{n}} (\tilde{N}_i^{-1}\tilde{N}) z_i z_i^{(0)}$ and $M_2 = \sum_{j=1}^p \mu_j b_j b_j^{(0)}$.

Recall that we condition on the event (F.32). By Hoeffding's inequality, Bayes's rule, and (F.32),

$$\begin{aligned}
\mathbb{P}(|M_1| > t) &= \mathbb{P}\Big(|\sum_i \frac{\tilde{N}}{\tilde{N}_i} z_i z_i^{(0)} \ge t \;\Big|\; |\sum_i z_i| \le 100\sqrt{\tilde{n}}, |\sum_i z_i^{(0)}| \le 100\sqrt{\tilde{n}}\Big) \\
&= \frac{\mathbb{P}\big(|\sum_i \frac{\tilde{N}}{\tilde{N}_i} z_i z_i^{(0)}| \ge t\big)}{\mathbb{P}(|\sum_i z_i| \le 100\sqrt{\tilde{n}})\ \mathbb{P}(|\sum_i z_i^{(0)}| \le 100\sqrt{\tilde{n}})} \\
&\le 4 \cdot 2\exp\Big(-\frac{t^2}{8\sum_{i=1}^{\tilde{n}} (\tilde{N}_i^{-1}\tilde{N})^2}\Big) \\
&= 8\exp\Big(-\frac{t^2}{8\tilde{n}}\Big).
\end{aligned}$$

for all $t > 0$. In the last line, we have used the assumption of $\tilde{N}_i \asymp \tilde{N}$. By Hoeffding's inequality again, we also have

$$\mathbb{P}(|M_2| > t) \le 2\exp\Big(-\frac{t^2}{8\sum_{j=1}^p \mu_j^2}\Big) = 2\exp\Big(-\frac{t^2}{8\|\mu\|^2}\Big)$$

for all $t > 0$. Write $s_{\tilde{n}}^2 = \sqrt{\tilde{n}}\tilde{N}\omega_n^2\|\mu\|$. It follows that

$$\begin{aligned}
\mathbb{P}(M^* > t) &= \mathbb{P}\big(2\tilde{N}\omega_n^2 M_1 M_2 > t\big) = \mathbb{P}\big(M_1 M_2 > t \cdot \sqrt{\tilde{n}}\|\mu\| s_{\tilde{n}}^{-2}\big) \\
&\le \mathbb{P}\big(M_1 > \sqrt{t} \cdot \sqrt{\tilde{n}} s_{\tilde{n}}^{-1}\big) + \mathbb{P}\big(M_2 > \sqrt{t} \cdot \|\mu\| s_{\tilde{n}}^{-1}\big) \\
&\le 8\exp\Big(-\frac{t}{8s_{\tilde{n}}^2}\Big) + 2\exp\Big(-\frac{t}{8s_{\tilde{n}}^2}\Big) \\
&\le 4\exp(-c_1 t/s_{\tilde{n}}^2), \qquad \text{(F.40)}
\end{aligned}$$

for some constant $c_1 > 0$. Here, in the last line, we have used the assumption of $\tilde{N}_i \asymp \tilde{N}$.

Let $f(x)$ and $F(x)$ be the density and distribution function of M^*. Write $\bar{F}(x) = 1 - F(x)$. Using integration by part, we have $\mathbb{E}[\exp(M^*)] = \int_0^\infty \exp(x) f(x)dx = -\exp(x)\bar{F}(x)|_0^\infty + \int_0^\infty \exp(x)\bar{F}(x)dx = 1 + \int_0^\infty \exp(x)\bar{F}(x)dx$, provided that the integral exists. As a result, when $s_{\tilde{n}} = o(1)$,

$$\begin{aligned}
\mathbb{E}[\exp(M^*)] - 1 &= \int_0^\infty \exp(t) \cdot \mathbb{P}(M^* > t) \\
&\le 4\int_0^\infty \exp\big(-[c_1 s_{\tilde{n}}^{-2} - 1]t\big)dt
\end{aligned}$$

$$\leq 4(c_1 s_{\tilde{n}}^{-1} - 1)^{-1} = 4s_{\tilde{n}}/(c_1 - s_{\tilde{n}}).$$

It implies $\mathbb{E}[\exp(M^*)] = 1 + o(1)$, which is exactly (F.39). This completes the proof. because

$$s_{\tilde{n}}^2 = \sqrt{\tilde{n}}\tilde{N}\omega_n^2\|\mu\| = \frac{n\bar{N}\|\mu\|\omega_n^2}{\sqrt{K}} \asymp \frac{n\bar{N}\|\mu\|\omega_n^2}{\sqrt{\sum_{k\in K}\|\mu_k\|^2}}.$$

□

F.5 Proof of Theorem 6

First we show that

$$T/\sqrt{\mathrm{Var}(T)} \Rightarrow N(0,1), \quad \text{and} \tag{F.41}$$

$$V/\mathrm{Var}(T) \to 1. \tag{F.42}$$

If (F.41) and (F.42) hold, then by mimicking the proof of Theorem 2, we see that ψ is asymptotically normal and the level-κ DELVE test has asymptotic level κ. We omit the details as they are quite similar.

Recall the martingale decomposition of T described in Section E. Observe that, under our assumptions, Lemmas E.1–E.6 are valid. Moreover, by Lemmas D.8 and D.13

$$\mathrm{Var}(T) \gtrsim \Theta_{n2} + \Theta_{n3} + \Theta_{n4} \gtrsim \left\| \frac{m\bar{M}}{n\bar{N} + m\bar{M}}\eta + \frac{n\bar{N}}{n\bar{N} + m\bar{M}}\theta \right\|^2. \tag{F.43}$$

Combining (F.43) with Lemmas E.1–E.6 and mimicking the argument in Section E.1 implies that $T/\sqrt{V} \Rightarrow N(0,1)$. Thus (F.41) is established.

Moreover, (F.42) is a direct consequence of our assumptions and Proposition D.15. The claims of Theorem 6 regarding the null hypothesis follow.

To prove the claims about the alternative hypothesis, it suffices to show

$$T/\sqrt{\mathrm{Var}(T)} \to \infty, \tag{F.44}$$

$$V > 0 \;\; \text{with high probability, and} \tag{F.45}$$

$$V = O_{\mathbb{P}}(\mathrm{Var}(T)). \tag{F.46}$$

Once these claims are established, we prove that $\psi = T\mathbf{1}_{V>0}/\sqrt{V} \to \infty$ under the alternative by mimicking the last step of the proof of Theorem 4 in Section F.3. We omit the details as they are very similar.

Note that (F.46) follows directly from our assumptions and Proposition D.16.

As in the proof of Theorem 4 in Section F.3, to establish (F.44), it suffices to prove that

$$\mathbb{E}T = \rho^2 \gg \mathrm{Var}(T). \tag{F.47}$$

Our main assumption under the alternative when $K = 2$ is

$$\frac{\|\eta - \theta\|^2}{\left(\frac{1}{n\bar{N}} + \frac{1}{m\bar{M}}\right)\max\{\|\eta\|, \|\theta\|\}} \to \infty. \tag{F.48}$$

As shown in Section F.2, we have that

$$\mathrm{Var}(T) \lesssim \Theta_n = \Theta_{n1} + \sum_{t=2}^{4}\Theta_{nt}. \tag{F.49}$$

Applying (F.17) to the first term and Lemma D.8 to the remaining terms, we have

$$\begin{aligned}\mathrm{Var}(T) &\lesssim \max\{\|\eta\|_\infty,\,\|\theta\|_\infty\}\cdot\rho^2 + \left\|\frac{m\bar{M}}{n\bar{N}+m\bar{M}}\eta + \frac{n\bar{N}}{n\bar{N}+m\bar{M}}\theta\right\|^2 \\ &\lesssim \max\{\|\eta\|,\,\|\theta\|\}\cdot\rho^2 + \max\{\|\eta\|^2,\,\|\theta\|^2\}\end{aligned} \tag{F.50}$$

Next, note that

$$\begin{aligned}\rho^2 &= n\bar{N}\|\eta-\mu\|^2 + m\bar{M}\|\theta-\mu\|^2 \\ &= n\bar{N}\left\|\eta - \big(\frac{n\bar{N}}{n\bar{N}+m\bar{M}}\eta + \frac{m\bar{M}}{n\bar{N}+m\bar{M}}\theta\big)\right\|^2 \\ &\quad + m\bar{M}\left\|\theta - \big(\frac{n\bar{N}}{n\bar{N}+m\bar{M}}\eta + \frac{m\bar{M}}{n\bar{N}+m\bar{M}}\theta\big)\right\|^2 \\ &= n\bar{N}\cdot\big(\frac{m\bar{M}}{n\bar{N}+m\bar{M}}\big)^2\|\eta-\theta\|^2 + m\bar{M}\cdot\big(\frac{n\bar{N}}{n\bar{N}+m\bar{M}}\big)^2\|\eta-\theta\|^2 \\ &= \frac{n\bar{N}m\bar{M}}{(n\bar{N}+m\bar{M})}\|\eta-\theta\|^2 = \big(\frac{1}{n\bar{N}}+\frac{1}{m\bar{M}}\big)^{-1}\|\eta-\theta\|^2.\end{aligned} \tag{F.51}$$

By (F.48), (F.50), and (F.51), we have

$$\begin{aligned}\frac{(\mathbb{E}T)^2}{\mathrm{Var}(T)} &\gtrsim \frac{\rho^4}{\max\{\|\eta\|,\,\|\theta\|\}\cdot\rho^2 + \max\{\|\eta\|^2,\,\|\theta\|^2\}} \\ &\gtrsim \frac{\|\eta-\theta\|^2}{(\frac{1}{n\bar{N}}+\frac{1}{m\bar{M}})\max\{\|\eta\|,\,\|\theta\|\}} + \big(\frac{\|\eta-\theta\|^2}{(\frac{1}{n\bar{N}}+\frac{1}{m\bar{M}})\max\{\|\eta\|,\,\|\theta\|\}}\big)^2 \to \infty,\end{aligned}$$

which proves (F.47) and thus (F.44).

To prove (F.45), we mimick the Markov argument in (F.21) and use that under our assumptions, $\mathrm{Var}(V)/(\mathbb{E}V)^2 = o(1)$. We omit the details as they are similar. Since we have established (F.44), (F.45), and (F.46), the proof is complete. □

F.6 Proof of Theorem 7

Note that $T/\sqrt{\mathrm{Var}(T)} \Rightarrow N(0,1)$ by our assumptions and Proposition E.7. In particular, using that $n\to\infty$ and the monotonicity of the ℓ_p norms we have

$$\frac{\|\mu\|_4^4}{K\|\mu\|^4} = \frac{\|\mu\|_4^4}{n\|\mu\|^4} \le \frac{1}{n}\cdot\frac{\|\mu\|^4}{\|\mu\|^4} = \frac{1}{n}\to 0.$$

Moreover, $V^*/\mathrm{Var}(T)\to 1$ in probability by Proposition D.17. It follows by Slutsky's theorem that $\psi^* = T/\sqrt{V^*}\Rightarrow N(0,1)$ and that the level-κ DELVE test has an asymptotic level κ.

To conclude the proof, it suffices to show that $\psi^*\to\infty$ under the alternative. As in the proof of Theorem 4, this follows immediately if we can show

$$T/\sqrt{\mathrm{Var}(T)}\to\infty, \tag{F.52}$$

$$V^* > 0 \text{ with high probability, and} \tag{F.53}$$

$$V^* = O_{\mathbb{P}}(\mathrm{Var}(T)). \tag{F.54}$$

Note that (F.52) follows from (F.14), and (F.54) is the content of Proposition D.18. Since our assumptions imply that $\mathbb{E}V^*\gg\sqrt{\mathrm{Var}(V^*)}$, (F.53) follows by a Markov argument as in (F.21). □

F.7 Proof of Theorem 8

We apply Theorem 2 to get the asymptotic null distribution. Since $N_i = N$ and $\mu = p^{-1}\mathbf{1}_p$, it is easy to see that Condition 2 is satisfied under our assumption of $p = o(N^2 n)$. Therefore, by Theorem 2, $\psi^* \to N(0,1)$ under H_0.

We now show the asymptotic alternative distribution. By direct calculations and using $\sum_{i=1}^n \delta_{ij} = 0$ and $\sum_{j=1}^p \delta_{ij} = 0$, we have

$$\sum_{i,j} N_i(\Omega_{ij} - \mu_j)^2 = \frac{nN\nu_n^2}{p}, \qquad \sum_{i,j} N_i(\Omega_{ij} - \mu_j)^2\Omega_{ij} = \frac{nN\nu_n^2}{p^2}, \qquad \sum_i \|\Omega_i\|^2 = \frac{n(1+\nu_n^2)}{p}.$$

We apply Lemmas D.1-D.5 and plug in the above expressions. Let $S = \mathbf{1}_p' U_2$. It follows that

$$T = \frac{nN\nu_n^2}{p} + S + O_{\mathbb{P}}\Big(\frac{\sqrt{nN}\nu_n}{p} + \frac{1}{\sqrt{p}}\Big), \qquad \text{where } \mathrm{Var}(S) = 2p^{-1}n[1+o(1)]. \tag{F.55}$$

First, we plug in $\nu_n^2 = a\sqrt{2p}/(N\sqrt{n})$. It gives $p^{-1}nN\nu_n^2 = \sqrt{2n/p}$. Second, $p^{-1}\sqrt{nN}\nu_n \asymp (np)^{-1/4}\sqrt{n/p} = o(\sqrt{n/p})$. It follows that

$$T = a\sqrt{2n/p} + S + o_{\mathbb{P}}\big(\sqrt{n/p}\big), \qquad \text{where } \mathrm{Var}(S) = (2n/p)[1+o(1)]. \tag{F.56}$$

Recall the martingale decomposition $S = \sum_{(\ell,s)} E_{\ell,s}$ where $E_{\ell,s}$ is defined in (E.4). Observe that Lemmas E.4 and E.5 hold (even under the alternative). Define $\widetilde{E}_{\ell,s} = E_{\ell,s}/\sqrt{\mathrm{Var}(S)}$. Using $\mathrm{Var}(S) \gtrsim n\sum_i \|\Omega_i\|^2$ and these lemmas, it is straightforward to verify that the following conditions hold:

$$\sum_{(\ell,s)} \mathrm{Var}\big(\widetilde{E}_{\ell,s} \big| \mathcal{F}_{\prec(\ell,s)}\big) \xrightarrow{\mathbb{P}} 1 \tag{F.57}$$

$$\sum_{(\ell,s)} \mathbb{E}\widetilde{E}_{\ell,s}^4 \xrightarrow{\mathbb{P}} 0. \tag{F.58}$$

As in Section E.1, the martingale CLT applies and we have

$$S/\sqrt{\mathrm{Var}(S)} \Rightarrow N(0,1).$$

By F.55,

$$T/\sqrt{\mathrm{Var}(S)} \quad \to \quad N(a,1). \tag{F.59}$$

By Lemma D.3 and (D.84),

$$\mathrm{Var}(S) = [1+o(1)]\Theta_{n2} = [1+o(1)]\mathrm{Var}(T)$$

By Proposition D.18, we have that $V^*/\mathrm{Var}(T) \to 1$ in probability. As a result,

$$V^*/\mathrm{Var}(S) \quad \to \quad 1, \qquad \text{in probability.} \tag{F.60}$$

We combine (F.59) and (F.60) to conclude that $\psi = T/\sqrt{V^*} \to N(a,1)$.

□

G Proofs of the corollaries for text analysis

G.1 Proof of Corollary 1

Note that Corollary 1 follows immediately from the slightly more general result stated below.

Corollary G.1. *Consider Model* (1) *and suppose that* $\Omega = \mu \mathbf{1}_n'$ *under the null hypothesis and that* Ω *satisfies* (37) *under the alternative hypothesis. Define* $\xi \in \mathbb{R}^n$ *by* $\xi_i = \bar{N}^{-1} N_i$ *and let* $\widetilde{\Omega} = \Omega[\mathrm{diag}(\xi)]^{1/2}$. *Let* $\lambda_1, \ldots, \lambda_M > 0$ *and* $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_M > 0$ *denote the singular values of* Ω *and* $\widetilde{\Omega}$, *respectively, arranged in decreasing order. We further assume that under the alternative hypothesis,*

$$\frac{\bar{N} \cdot \sum_{k=2}^{M} \widetilde{\lambda}_k^2}{\sqrt{\sum_{k=1}^{M} \lambda_k^2}} \to \infty. \tag{G.1}$$

For any fixed $\kappa \in (0,1)$, *the level-*κ *DELVE test has an asymptotic level* κ *and an asymptotic power* 1. *Moreover if* $N_i \asymp \bar{N}$ *for all* i, *we may replace* $\sum_{k=2}^{M} \widetilde{\lambda}_k^2$ *with* $\sum_{k=2}^{M} \lambda_k^2$ *in the numerator of* (G.1).

Proof of Corollary G.1. This is a special case of our testing problem with $K = n$. Moreover, $\mu = n^{-1}\Omega\xi$ matches with the definition of μ in (2). Therefore, we can apply Theorem 7 directly. It remains to verify that the condition

$$\frac{\bar{N} \cdot \sum_{k=2}^{M} \widetilde{\lambda}_k^2}{\sqrt{\sum_{k=1}^{M} \lambda_k^2}} \to \infty \tag{G.2}$$

is sufficient to lead to the condition

$$\frac{n\bar{N}\|\mu\|^2\omega_n^2}{\sqrt{\sum_i \|\Omega_i\|^2}} \to \infty. \tag{G.3}$$

If we show this then Theorem 7 applies directly. We first calculate ω_n^2. Recall $\xi_i = N_i/\bar{N}$ for $1 \le i \le n$. Write

$$\widetilde{\Omega} = \Omega[\mathrm{diag}(\xi)]^{1/2}, \qquad \widetilde{\xi} = [\mathrm{diag}(\xi)]^{1/2}\mathbf{1}_n.$$

For $K = n$, by (33), $\omega_n^2 = \frac{1}{n\bar{N}\|\mu\|^2}\sum_{i=1}^{n} N_i \|\Omega_i - \mu\|^2$. It follows that

$$\omega_n^2 = \frac{1}{n\|\mu\|^2}\left\|(\Omega - \mu\mathbf{1}_n')[\mathrm{diag}(\xi)]^{1/2}\right\|_F^2 = \frac{1}{n\|\mu\|^2}\|\widetilde{\Omega} - \mu\widetilde{\xi}'\|_F^2. \tag{G.4}$$

Recall that $\widetilde{\lambda_1}, \ldots, \widetilde{\lambda}_M$ are the singular values of $\widetilde{\Omega}$. We apply a well-known result in linear algebra [Horn and Johnson, 1985], namely Weyl's inequality: For any rank-1 matrix Δ, $\|\widetilde{\Omega} - \Delta\|_F^2 \ge \sum_{k\neq 1} \widetilde{\lambda}_k^2$. In (G.4), $\mu\widetilde{\xi}'$ is a rank-1 matrix. It follows that

$$\|\widetilde{\Omega} - \mu\widetilde{\xi}'\|_F^2 \ge \sum_{k=2}^{M} \widetilde{\lambda}_k^2. \tag{G.5}$$

Hence

$$\frac{n\bar{N}\|\mu\|^2\omega_n^2}{\sqrt{\sum_i \|\Omega_i\|^2}} \ge \frac{\bar{N} \cdot \sum_{k=2}^{M} \widetilde{\lambda}_k^2}{\|\Omega\|_F} = \frac{\bar{N} \cdot \sum_{k=2}^{M} \widetilde{\lambda}_k^2}{\sqrt{\sum_{k=1}^{M} \lambda_k^2}},$$

which implies (G.3) by our assumption. The first claim is proved.

Next we prove the second claim. Observe that if $N_i \asymp \bar{N}$, then by Weyl's inequality:

$$\begin{aligned}\omega_n^2 &= \frac{1}{\|\mu\|^2 n\bar{N}}\sum_i N_i\|\Omega_i - \mu\|^2 \gtrsim \frac{1}{\|\mu\|^2}\sum_i \|\Omega_i - \mu\|^2 \\ &= \frac{1}{\|\mu\|^2}\|\Omega - \mu\mathbf{1}_n'\|_F^2 \ge \frac{1}{\|\mu\|^2}\sum_{k=2}^{M}\lambda_k^2.\end{aligned}$$

Thus

$$\frac{n\bar{N}\|\mu\|^2\omega_n^2}{\sqrt{\sum_i \|\Omega_i\|^2}} \geq \frac{\bar{N}\cdot\sum_{k=2}^M \lambda_k^2}{\|\Omega\|_F} = \frac{\bar{N}\cdot\sum_{k=2}^M \lambda_k^2}{\sqrt{\sum_{k=1}^M \lambda_k^2}}.$$

We see that the assumption

$$\frac{\bar{N}\cdot\sum_{k=2}^M \lambda_k^2}{\sqrt{\sum_{k=1}^M \lambda_k^2}} \to \infty \tag{G.6}$$

implies (G.3). The second claim is established and the proof is complete. □

G.2 Proof of Corollary 2

Recall the construction of a simple null and simple (random) alternative model from Section F.4.2, specialized below to the case of $K = n$ and $N_i \equiv N$:

$$H_0: \qquad \Omega_i = \tilde{\mu}, \qquad 1 \leq i \leq n. \tag{G.7}$$

$$H_1: \qquad \Omega_{ij} = \begin{cases} \mu_j\big(1+\omega_n z_i b_j\big), & \text{if } 1 \leq j \leq m \\ \tilde{\mu}_j\big(1-\omega_n z_i b_{j-m}\big), & \text{if } m+1 \leq j \leq 2m \end{cases} \tag{G.8}$$

where $b_1, \dots, b_m$ are i.i.d. Rademacher random variables and $z_1, \dots, z_n$ are i.i.d Rademacher random variables conditioned to satisfy $|\sum_i z_i| \leq 100\sqrt{n}$. Define

$$\tilde{b} = (b_1, \dots, b_m, b_1, \dots, b_m)'.$$

To derive the lower bound of Corollary 2, we assume without loss of generality that ω_n is a sufficiently small absolute constant.

We claim that H_1 prescribes a topic model with $M = 2$ topics. To see this, under the alternative,

$$\Omega_i = \begin{cases} \mu \circ (\mathbf{1}_p + \omega_n\,\tilde{b}) & \text{if } z_i = 1 \\ \mu \circ (\mathbf{1}_p - \omega_n\,\tilde{b}) & \text{if } z_i = -1. \end{cases} \tag{G.9}$$

Moreover, we showed in Section F.4.2 that $\Omega_{ij} \geq 0$ for all i, j and that $\|\Omega_{ij}\|_1 = 1$. From (G.9), we see that $\Omega = AW$ where $A \in \mathbb{R}^{p\times 2}$ and $W \in \mathbb{R}^{2\times n}$ are defined as follows:

$$A_{:1} = \mu \circ (\mathbf{1}_p + \omega_n\,\tilde{b}), \quad A_{:2} = \mu \circ (\mathbf{1}_p - \omega_n\,\tilde{b})$$
$$W_{:i} = \begin{cases} (1,0)' & \text{if } z_i = 1 \\ (0,1)' & \text{if } z_i = -1. \end{cases}$$

Moreover, under the null hypothesis, Ω clearly prescribes a topic model with $K = 1$. Therefore Ω follows the topic model (37). Moreover, since $N_i \equiv N$, we have $\Omega[\mathrm{diag}(\xi)]^{1/2} = \Omega$.

By Proposition F.2 specialized to our setting, we know that the χ^2 distance between the null and alternative goes to zero if

$$\sqrt{n}N\|\mu\|\omega_n^2 \to 0.$$

Thus to prove Corollary 2 it suffices to show that

$$\frac{N\sum_{k\geq 2}^M \lambda_k^2}{\sqrt{\sum_{k=1}^M \lambda_k^2}} = \frac{N\lambda_2^2}{\sqrt{\sum_{k=1}^M \lambda_k^2}} \gtrsim \sqrt{n}N\|\mu\|\omega_n^2 \tag{G.10}$$

Accordingly we study the second largest singular value of Ω. First we have some preliminary calculations. Let $U = \{i : z_i = 1\}$, and let $V = \{i : z_i = -1\}$. Define

$$u = \mu \circ (\mathbf{1}_p + \omega_n \tilde{b}), \quad \text{and}$$
$$v = \mu \circ (\mathbf{1}_p - \omega_n \tilde{b}).$$

Observe that

$$\langle u, v\rangle = \|\mu\|^2 - \omega_n^2 \|\mu \circ \tilde{b}\|^2 = \|\mu\|^2 (1 - \omega_n^2).$$

Also, since ω_n is a sufficiently small absolute constant,

$$\|u\|^2 = \|\mu\|^2 + 2\omega_n \langle \mu, \mu \circ \tilde{b}\rangle + \omega_n^2 \|\mu \circ \tilde{b}\|^2 = (1 + \omega_n^2)\|\mu\|^2 + 2\omega_n \sum_j \mu_j^2 \tilde{b}_j \gtrsim \|\mu\|^2, \quad \text{and}$$
$$\|v\|^2 = \|\mu\|^2 - 2\omega_n \langle \mu, \mu \circ \tilde{b}\rangle + \omega_n^2 \|\mu \circ \tilde{b}\|^2 = (1 + \omega_n^2)\|\mu\|^2 - 2\omega_n \sum_j \mu_j^2 \tilde{b}_j \gtrsim \|\mu\|^2. \tag{G.11}$$

Again, since we assume that ω_n is a sufficiently small absolute constant,

$$\begin{aligned}\delta^2 := \frac{\langle u, v\rangle^2}{\|u\|^2 \|v\|^2} &= \frac{\|\mu\|^4 (1 - \omega_n^2)^2}{(1 + \omega_n^2)^2 \|\mu\|^4 - 4\omega_n^2 \langle \mu, \mu \circ b\rangle^2} \le \frac{\|\mu\|^4 (1 - \omega_n^2)^2}{(1 + \omega_n^2)^2 \|\mu\|^4 - 4\omega_n^2 \|\mu\|^4} \\ &= \frac{\|\mu\|^4 (1 - \omega_n^2)^2}{\|\mu\|^4 (1 + 2\omega_n^2 - 3\omega_n^4)} = \frac{(1 - \omega_n^2)^2}{1 + 2\omega_n^2 - 3\omega_n^4}\end{aligned} \tag{G.12}$$

Note that

$$\begin{aligned}\|au + bv\|^2 &= a^2 \|u\|^2 + 2ab\langle u, v\rangle + b^2 \|v\|^2 \ge a^2 \|u\|^2 + b^2 \|v\|^2 - 2ab\delta \|u\|\|v\| \\ &\ge (1 - \delta)\big(a^2 \|u\|^2 + b^2 \|v\|^2\big) + \|au - bv\|^2 \ge (1 - \delta)\big(a^2 \|u\|^2 + b^2 \|v\|^2\big).\end{aligned}$$

By (G.12), we have for ω_n sufficiently small that

$$\begin{aligned}1 - \delta &\ge 1 - \frac{1 - \omega_n^2}{\sqrt{1 + 2\omega_n^2 - 3\omega_n^4}} = \frac{\sqrt{1 + 2\omega_n^2 - 3\omega_n^4} - 1 + \omega_n^2}{\sqrt{1 + 2\omega_n^2 - 3\omega_n^4}} \\ &\ge \frac{\omega_n^2}{\sqrt{1 + 2\omega_n^2 - 3\omega_n^4}} \gtrsim \omega_n^2.\end{aligned}$$

Thus

$$\|au + bv\|^2 \ge \omega_n^2 (a^2 \|u\|^2 + b^2 \|v\|^2) \gtrsim \omega_n^2 \|\mu\|^2 (a^2 + b^2) \tag{G.13}$$

Recall that if M is a rank k matrix, then

$$\lambda_k(M) = \sup_{y : \|y\| = 1,\, y \in \mathrm{Ker}(M)^\perp} \|My\| = \sup_{y : \|y\| = 1,\, y \in \mathrm{Im}(M')} \|My\|. \tag{G.14}$$

We have

$$\Omega\Omega' = \sum_{i \in U} uu' + \sum_{i \in V} vv' = |U| uu' + |V| vv'.$$

Let $y \in \mathbb{R}^n$ satisfy $\|y\| = 1$ and $y = \Omega' x$ for some x. We have

$$\Omega y = \Omega\Omega' x = |U| \langle u, x\rangle u + |V| \langle v, x\rangle v.$$

By the previous equation and (G.13),

$$\|\Omega y\|^2 = \|\Omega\Omega' x\|^2 = \Big\| |U| \langle u, x\rangle u + |V| \langle v, x\rangle v \Big\|^2 \gtrsim \omega_n^2 \|\mu\|^2 \big(|U|^2 \langle u, x\rangle^2 + |V|^2 \langle v, x\rangle^2\big).$$

By our conditioning on z, we have $\min(|U|, |V|) \gtrsim n$. Moreover

$$1 = \|y\|^2 = \|\Omega' x\|^2 = |U|\langle u, x\rangle^2 + |V|\langle v, x\rangle^2.$$

Applying these facts and (G.14), we obtain

$$\lambda_2^2 \geq \|\Omega y\|^2 = \|\Omega\Omega' x\|^2 \gtrsim \omega_n^2 \|\mu\|^2 n\big(|U|\langle u, x\rangle^2 + |V|\langle v, x\rangle^2\big) = \omega_n^2 \|\mu\|^2 n.$$

Next,

$$\sum_{k=1}^{M} \lambda_k^2 = \|\Omega\|_F^2 = \sum_{i\in U} \|u\|^2 + \sum_{i\in V} \|v\|^2 = |U| \cdot \|u\|^2 + |V| \cdot \|v\|^2 \asymp n\|\mu\|^2 \tag{G.15}$$

We conclude that

$$\frac{N \sum_{k\geq 2}^{M} \lambda_k^2}{\sqrt{\sum_{k=1}^{M} \lambda_k^2}} = \frac{N\lambda_2^2}{\sqrt{\sum_{k=1}^{M} \lambda_k^2}} \gtrsim \frac{N \cdot \omega_n^2 \|\mu\|^2 n}{\sqrt{n}\|\mu\|} = \sqrt{n} N \|\mu\| \omega_n^2$$

which establishes (G.10). The proof is complete. □

G.3 Proof of Corollary 3

This is a special case of our testing problem with $K = 2$, we can apply Theorem 6 directly. It remains to verify that the condition

$$\frac{\zeta_n^2 \cdot (\|\eta_S\|_1 + \|\theta_S\|_1)}{\big(\frac{1}{n\bar{N}} + \frac{1}{m\bar{M}}\big) \max\{\|\eta\|, \|\theta\|\}} \to \infty \tag{G.16}$$

is sufficient to yield the condition (31) in Theorem 6. This is done by calculating $\|\eta - \theta\|^2$ directly. By our sparse model (40), for $j \in S$, $|\sqrt{\eta_j} - \sqrt{\theta_j}| \geq \zeta_n$. It follows that for $j \in S$,

$$|\eta_j - \theta_j|^2 = (\sqrt{\eta_j} + \sqrt{\theta_j})^2 (\sqrt{\eta_j} - \sqrt{\theta_j})^2 \geq \zeta_n^2 (\sqrt{\eta_j} + \sqrt{\theta_j})^2 \geq \zeta_n^2 (\eta_j + \theta_j).$$

It follows that

$$\|\eta - \theta\|^2 \geq \zeta_n^2 \sum_{j\in S} (\eta_j + \theta_j) \geq \zeta_n^2 \big(\|\eta_S\|_1 + \|\theta_S\|_1\big). \tag{G.17}$$

We plug it into (31) and see immediately that (G.16) implies this condition. The claim follows directly from Theorem 6. □

H A modification of DELVE for finite p

Below we write out the variance of the terms of the raw DELVE statistic under the null, using the proofs of Lemmas D.3–D.5.

$$\mathrm{Var}(\mathbf{1}_p' U_2) = 2 \sum_{k=1}^{K} \sum_{i\in S_k} \sum_{1\leq r<s\leq N_i} \big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}}\big)^2 \frac{N_i^2}{(N_i - 1)^2} \big[\|\Omega_i\|^2 - 2\|\Omega_i\|_3^3 + \|\Omega_i\|^4\big] \tag{H.1}$$

$$\mathrm{Var}(\mathbf{1}_p' U_3) = \frac{2}{n^2 \bar{N}^2} \sum_{k\neq \ell} \sum_{i\in S_k} \sum_{m\in S_\ell} N_i N_m \Big(\sum_j \Omega_{ij}\Omega_{mj} - 2 \sum_j \Omega_{ij}^2 \Omega_{mj}^2 + \sum_{j,j'} \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'}\Big)$$

$$\mathrm{Var}(\mathbf{1}_p' U_4) = 2 \sum_{k=1}^{K} \sum_{\substack{i\in S_k, m\in S_k \\ i\neq m}} \big(\frac{1}{n_k \bar{N}_k} - \frac{1}{n\bar{N}}\big)^2 N_i N_m \Big(\sum_j \Omega_{ij}\Omega_{mj} - 2\sum_j \Omega_{ij}^2 \Omega_{mj}^2 + \sum_{j,j'} \Omega_{ij}\Omega_{ij'}\Omega_{mj}\Omega_{mj'}\Big).$$

In this section we develop an unbiased estimator for each term above, which leads to an unbiased estimator of $\mathrm{Var}(T)$ by taking their sum. We require some preliminary results proved later in this section. Recall that Lemma H.2 was established in the proof of Lemma D.1.

Lemma H.1. *If $j \neq j'$, an unbiased estimator of $\Omega_{ij}\Omega_{ij'}$ is*

$$\widehat{\Omega_{ij}\Omega_{ij'}} := \frac{X_{ij}X_{ij'}}{N_i(N_i-1)}$$

Lemma H.2. *An unbiased estimator of Ω_{ij}^2 is*

$$\widehat{\Omega_{ij}^2} := \frac{X_{ij}^2 - X_{ij}}{N_i(N_i-1)}. \tag{H.2}$$

Lemma H.3. *If $j \neq j'$, an unbiased estimator for $\Omega_{ij}^2\Omega_{ij'}^2$ is*

$$\widehat{\Omega_{ij}^2\Omega_{ij'}^2} = \frac{(X_{ij}^2 - X_{ij})(X_{ij'}^2 - X_{ij'})}{N_i(N_i-1)(N_i-2)(N_i-3)}$$

Lemma H.4. *An unbiased estimator of Ω_{ij}^3 is*

$$\widehat{\Omega_{ij}^3} := \frac{X_{ij}^3 - 3X_{ij}^2 + 2X_{ij}}{N_i(N_i-1)(N_i-2)}. \tag{H.3}$$

Lemma H.5. *An unbiased estimator of Ω_{ij}^4 is*

$$\widehat{\Omega_{ij}^4} := \frac{X_{ij}^4 - 3X_{ij}^3 - X_{ij}^2 + 3X_{ij}}{N_i(N_i-1)(N_i-2)(N_i-3)}. \tag{H.4}$$

Define

$$\begin{aligned}
\widehat{\|\Omega_i\|^2} &:= \sum_j \widehat{\Omega_{ij}^2} \\
\widehat{\|\Omega_i\|_3^3} &:= \sum_j \widehat{\Omega_{ij}^3} \\
\widehat{\|\Omega_i\|^4} &:= \sum_j \widehat{\Omega_{ij}^4} + \sum_{j\neq j'} \widehat{\Omega_{ij}^2\Omega_{ij'}^2}.
\end{aligned} \tag{H.5}$$

Using Lemmas H.1–H.5 and (H.5), we define an unbiased estimator for each term of (H.1). Let $\widehat{\Omega_{ij}} = X_{ij}/N_i$ and define

$$\widehat{\mathrm{Var}(\mathbf{1}_p' U_2)} = 2\sum_{k=1}^K \sum_{i\in S_k} \sum_{1\le r<s\le N_i} (\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}})^2 \frac{N_i^2}{(N_i-1)^2}\left[\widehat{\|\Omega_i\|^2} - 2\widehat{\|\Omega_i\|_3^3} + \widehat{\|\Omega_i\|^4}\right] \tag{H.6}$$

$$\widehat{\mathrm{Var}(\mathbf{1}_p' U_3)} = \frac{2}{n^2\bar{N}^2} \sum_{k\neq \ell}\sum_{i\in S_k}\sum_{m\in S_\ell} N_iN_m\Big(\sum_j \widehat{\Omega_{ij}}\widehat{\Omega_{mj}} - 2\sum_j \widehat{\Omega_{ij}^2}\widehat{\Omega_{mj}^2} + \sum_{j,j'} \widehat{\Omega_{ij}\Omega_{ij'}}\widehat{\Omega_{mj}\Omega_{mj'}}\Big)$$

$$\widehat{\mathrm{Var}(\mathbf{1}_p' U_4)} = 2\sum_{k=1}^K \sum_{\substack{i\in S_k, m\in S_k \\ i\neq m}} (\frac{1}{n_k\bar{N}_k} - \frac{1}{n\bar{N}})^2 N_iN_m\Big(\sum_j \widehat{\Omega_{ij}}\widehat{\Omega_{mj}} - 2\sum_j \widehat{\Omega_{ij}^2}\widehat{\Omega_{mj}^2} + \sum_{j,j'} \widehat{\Omega_{ij}\Omega_{ij'}}\widehat{\Omega_{mj}\Omega_{mj'}}\Big).$$

Define

$$\widetilde{V} = \widehat{\mathrm{Var}(\mathbf{1}_p' U_2)} + \widehat{\mathrm{Var}(\mathbf{1}_p' U_3)} + \widehat{\mathrm{Var}(\mathbf{1}_p' U_4)}. \tag{H.7}$$

We define *exact DELVE* as $\tilde{\psi} = T/\widetilde{V}^{1/2}$. Combining our results above, we obtain the following.

Proposition H.6. *Consider the statistic $\widetilde{V}$ defined in* (H.7). *Under the null hypothesis, $\widetilde{V}$ is an unbiased estimator for* $\mathrm{Var}(T)$.

With this result in hand, it is possible to derive consistency of $\widetilde{V}$ as an estimator of $\mathrm{Var}(T)$ under certain regularity conditions. We omit the details.

H.1 Proof of Lemma H.1

Recall that B_{ijr} is the Bernoulli random variable $B_{ijr} = Z_{ijr} + \Omega_{ij}$ and satisfies $X_{ijr} = \sum_{r=1}^{N_i} B_{ijr}$. Observe that

$$X_{ij}X_{ij'} = \sum_{r,s} B_{ijr}B_{ij's} = \sum_{r} B_{ijr}B_{ij'r} + \sum_{r\neq s} B_{ijr}B_{ij's} = 0 + \sum_{r\neq s} B_{ijr}B_{ij's}$$

Thus

$$\mathbb{E}X_{ij}X_{ij'} = N_i(N_i-1)\Omega_{ij}\Omega_{ij'},$$

and we obtain

$$\widehat{\Omega_{ij}\Omega_{ij'}} = \frac{X_{ij}X_{ij'}}{N_i(N_i-1)}$$

is an unbiased estimator for $\Omega_{ij}\Omega_{ij'}$, as desired. □

H.2 Proof of Lemma H.3

Note that

$$\begin{aligned}
X_{ij}^2X_{ij'}^2 &= \big(\sum_{r} B_{ijr} + \sum_{r\neq s} B_{ijr}B_{ijs}\big)\big(\sum_{r} B_{ij'r} + \sum_{r\neq s} B_{ij'r}B_{ij's}\big) \\
&= \sum_{r} B_{ijr}B_{ij'r} + \sum_{r_1\neq r_2} B_{ijr}B_{ij's} + \sum_{r_1\neq s} B_{ijr_1}B_{ijs}\sum_{r_2} B_{ij'r_2} + \sum_{r_1\neq s} B_{ij'r_1}B_{ij's}\sum_{r_2} B_{ijr_2} \\
&\quad + \big(\sum_{r\neq s} B_{ijr}B_{ijs}\big)\big(\sum_{r\neq s} B_{ij'r}B_{ij's}\big) \\
&= \sum_{r_1\neq r_2} B_{ijr}B_{ij's} + \sum_{r_1\neq s} B_{ijr_1}B_{ijs}\sum_{r_2} B_{ij'r_2} + \sum_{r_1\neq s} B_{ij'r_1}B_{ij's}\sum_{r_2} B_{ijr_2} \\
&\quad + \big(\sum_{r\neq s} B_{ijr}B_{ijs}\big)\big(\sum_{r\neq s} B_{ij'r}B_{ij's}\big)
\end{aligned}$$

Since $B_{ijr}B_{ij'r} = 0$, note that

$$\begin{aligned}
(X_{ij}^2 - X_{ij})(X_{ij'}^2 - X_{ij'}) &= \sum_{r_1\neq s_1}\sum_{r_2\neq s_2} B_{ijr_1}B_{ijs_1}B_{ij'r_2}B_{ij's_2} \\
&= \sum_{r_1,s_1,r_2,s_2\,dist.} B_{ijr_1}B_{ijs_1}B_{ij'r_2}B_{ij's_2}.
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}(X_{ij}^2 - X_{ij})(X_{ij'}^2 - X_{ij'}) &= \sum_{r_1,s_1,r_2,s_2\,dist.} \mathbb{E}\big[B_{ijr_1}B_{ijs_1}B_{ij'r_2}B_{ij's_2}\big] \\
&= N_i(N_i-1)(N_i-2)(N_i-3)\cdot\Omega_{ij}^2\Omega_{ij'}^2.
\end{aligned}$$

It follows that

$$\widehat{\Omega_{ij}^2\Omega_{ij'}^2} = \frac{(X_{ij}^2 - X_{ij})(X_{ij'}^2 - X_{ij'})}{N_i(N_i-1)(N_i-2)(N_i-3)}$$

is an unbiased estimator for $\Omega_{ij}^2\Omega_{ij'}^2$.

□

H.3 Proof of Lemma H.4

Recall that B_{ijr} is the Bernoulli random variable $B_{ijr} = Z_{ijr} + \Omega_{ij}$ and satisfies $X_{ijr} = \sum_{r=1}^{N_i} B_{ijr}$. Observe that

$$X_{ij}^3 = \sum_r B_{ijr} + 3 \sum_{r_1 \neq r_2} B_{ijr_1} B_{ijr_2} + \sum_{r_1 \neq r_2 \neq r_3} B_{ijr_1} B_{ijr_2} B_{ijr_3}.$$

Thus

$$\mathbb{E} X_{ij}^3 = N_i \Omega_{ij} + 3N_i(N_i - 1)\Omega_{ij}^2 + N_i(N_i - 1)(N_i - 2)\Omega_{ij}^3.$$

Unbiased estimators for Ω_{ij} and Ω_{ij}^2 are

$$\begin{aligned} &\frac{X_{ij}}{N_i} \\ &\frac{X_{ij}^2}{N_i^2} - \frac{X_{ij}(N_i - X_{ij})}{N_i^2(N_i - 1)} = \frac{1}{N_i(N_i - 1)}\left(X_{ij}^2 - X_{ij}\right), \end{aligned}$$

respectively. Hence

$$X_{ij}^3 - X_{ij} - 3(X_{ij}^2 - X_{ij}) = X_{ij}^3 - 3X_{ij}^2 + 2X_{ij}$$

is an unbiased estimator for $N_i(N_i - 1)(N_i - 2)\Omega_{ij}^3$, as desired.

□

H.4 Proof of Lemma H.5

Observe that

$$\begin{aligned} X_{ij}^4 &= \sum_r B_{ijr}^4 + 4 \sum_{r_1 \neq r_2} B_{ijr_1}^3 B_{ijr_2} + 6 \sum_{r_1 \neq r_2} B_{ijr_1}^2 B_{ijr_2}^2 \\ &\quad + 3 \sum_{r_1 \neq r_2 \neq r_3} B_{ijr_1}^2 B_{ijr_2} B_{ijr_3} + \sum_{r_1 \neq r_2 \neq r_3 \neq r_4} B_{ijr_1} B_{ijr_2} B_{ijr_3} B_{ijr_4} \\ &= \sum_r B_{ijr} + 10 \sum_{r_1 \neq r_2} B_{ijr_1} B_{ijr_2} + 3 \sum_{r_1 \neq r_2 \neq r_3} B_{ijr_1} B_{ijr_2} B_{ijr_3} \\ &\quad + \sum_{r_1 \neq r_2 \neq r_3 \neq r_4} B_{ijr_1} B_{ijr_2} B_{ijr_3} B_{ijr_4}. \end{aligned}$$

Thus

$$\begin{aligned} \mathbb{E} X_{ij}^4 &= N_i \Omega_{ij} + 10N_i(N_i - 1)\Omega_{ij}^2 + 3N_i(N_i - 1)(N_i - 2)\Omega_{ij}^3 \\ &\quad + N_i(N_i - 1)(N_i - 2)(N_i - 3)\Omega_{ij}^4. \end{aligned}$$

Plugging in unbiased estimators for the first three terms, we have

$$X_{ij}^4 - X_{ij} - 10(X_{ij}^2 - X_{ij}) - 3(X_{ij}^3 - 3X_{ij}^2 + 2X_{ij}) = X_{ij}^4 - 3X_{ij}^3 - X_{ij}^2 + 3X_{ij}$$

is an unbiased estimator for $N_i(N_i - 1)(N_i - 2)(N_i - 3)$, as desired. □

References

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

Roger Horn and Charles Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

Jiashun Jin, Zheng Ke, and Shengming Luo. Network global testing by counting graphlets. In *International conference on machine learning*, pages 2333–2341. PMLR, 2018.

A.B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer New York, 2008. ISBN 9780387790527. URL `https://books.google.com/books?id=mwB8rUBsbqoC`.