Fisher's combined probability test for cross-sectional independence in panel data models with serial correlation

Hongfei Wang[1], Binghui Liu[2*],Long Feng[3*], Yanyuan Ma[4]

[1] *Nanjing Audit University*

[2] *Northeast Normal University*

[3] *Nankai University*

[4] *Pennsylvania State University*

Supplementary Material

The Supplementary Material is organized as follows: Section S1 contains additional numerical experiments, followed by an additional empirical application in Section S2. The proofs of all theoretical results are provided in Section S3, and the proofs of some lemmas used in Section S3 are given in Section S4.

Notations. For any square matrix $\mathbf{A}$, $(\mathbf{A})_{ij}$ denotes the (i, j)-th entry of $\mathbf{A}$, $\mathrm{tr}(\mathbf{A})$ denotes the trace of $\mathbf{A}$, $\|\mathbf{A}\|_{\mathrm{F}}$ denotes the Frobenius norm of

*co-corresponding authors.

matrix $\mathbf{A}$, $\|\mathbf{A}\|$ denotes the operator norm of $\mathbf{A}$, $\lambda_{\max}(\mathbf{A})$ denotes the largest eigenvalue of $\mathbf{A}$. Let $\mathbf{A}$ and $\mathbf{B}$ be the two matrices, we define $\mathbf{A} \otimes \mathbf{B}$ as the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$; when $\mathbf{A}$ and $\mathbf{B}$ are two square matrices of the same order, define $\mathbf{A} \circ \mathbf{B}$ as the Hadamard product of $\mathbf{A}$ and $\mathbf{B}$, that is, $(\mathbf{A} \circ \mathbf{B})_{ij} = (\mathbf{A})_{ij}(\mathbf{B})_{ij}$. For $1 \leq i \leq T$, let e_i be a T-dimensional vector, where the i-th element is 1 and the rest are 0. Let $\tau_T = (1, 1, \ldots, 1)' \in \mathbb{R}^T$. The notation $\mathbf{I}_T$ denotes the $T \times T$ identity matrix. For any two real numbers x and y, let $x \vee y = \max(x, y)$ and $x \wedge y = \min(x, y)$. For any vector $v \in \mathbb{R}^T$, $\|v\|$ denotes the Euclidean norm of v; Throughout the paper, C, $C', c, c', c_1, c_2, \cdots$, denote positive absolute constants.

S1 Additional simulation results

S1.1 Simulation results in asymmetric situation

Assumption 1 assumes that the density function of $(\mathbf{Z})_{it}$ is symmetric, which in turn implies that the density functions of ϵ_{it} and $\hat{\epsilon}_{it}$ are also symmetric under the null hypothesis. Consequently, we have

$$E(\hat{\rho}_{ij}) = E\left\{ \frac{\sum_{t=1}^{T} \hat{\epsilon}_{it}\hat{\epsilon}_{jt}}{(\sum_{t=1}^{T} \hat{\epsilon}_{it}^2 \sum_{t=1}^{T} \hat{\epsilon}_{jt}^2)^{1/2}} \right\} = 0, \tag{S1.1}$$

where the last equality may not be true for an asymmetric distribution. To demonstrate this result, we conduct numerical simulations using the data generating process outlined in (3.16). Let $\epsilon_{i\cdot} = \mathbf{\Gamma} z_{i\cdot}$, where $z_{i\cdot} = (z_{i1}, \ldots, z_{iT})$, for $1 \leq i \leq N$, $\mathbf{\Gamma} = \mathbf{O}^\top \mathbf{B} \mathbf{O} \in \mathbb{R}^{T\times T}$, $\mathbf{B} = \mathrm{diag}\{b\}$ and $b = \{b_1, \ldots, b_T\}$. In particular, the first $T-1$ elements of b are independent and identically distributed from $U[1, 1.1]$, while the remaining one is set to be 3. The construction of matrix $\mathbf{O} \in \mathbb{R}^{T\times T}$ is related to $\mathbf{S} \in \mathbb{R}^{T\times T}$, where each component of $\mathbf{S}$ is independent and identically distributed from $\mathcal{N}(0,1)$. We define the QR decomposition of $\mathbf{S}$ as $\mathbf{S} = \mathbf{U}\mathbf{V}$, where $\mathbf{U}$ is an orthogonal square matrix and $\mathbf{V}$ is a upper-triangular matrix. We then let $\mathbf{O} = \mathbf{U}\mathbf{\Gamma}$, where $\mathbf{\Gamma} = \mathrm{diag}\{\mathrm{sign}(r_1), \ldots, \mathrm{sign}(r_T)\}$, and r_t denotes the (t,t)-th entry of $\mathbf{V}$, for all $1 \leq t \leq T$.

For comparison purposes, we independently generated z_{it} from the following three distributions (1) $\mathcal{N}(0,1)$; (2) $\{Exp(2) - 0.5\}/\sqrt{0.5}$; (3) $(\chi^2_2 - 2)/\sqrt{4}$, where $Exp(2)$ represents an exponential distribution with rate parameter 2. In addition, for the choices of p, N and T, we set $p = 3$, $T = 10$ and $N \in \{200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000\}$.

Figure S.1 displays the average values of the two statistics, S_N and $S_N/\hat{\sigma}_{S_N}$, under three different distributions and the null hypothesis. As N increases while p and T are fixed, the average values of S_N and $S_N/\hat{\sigma}_{S_N}$ remain close to 0 under normal distribution, while the average values of the two statistics move away from 0 under the exponential distribution and

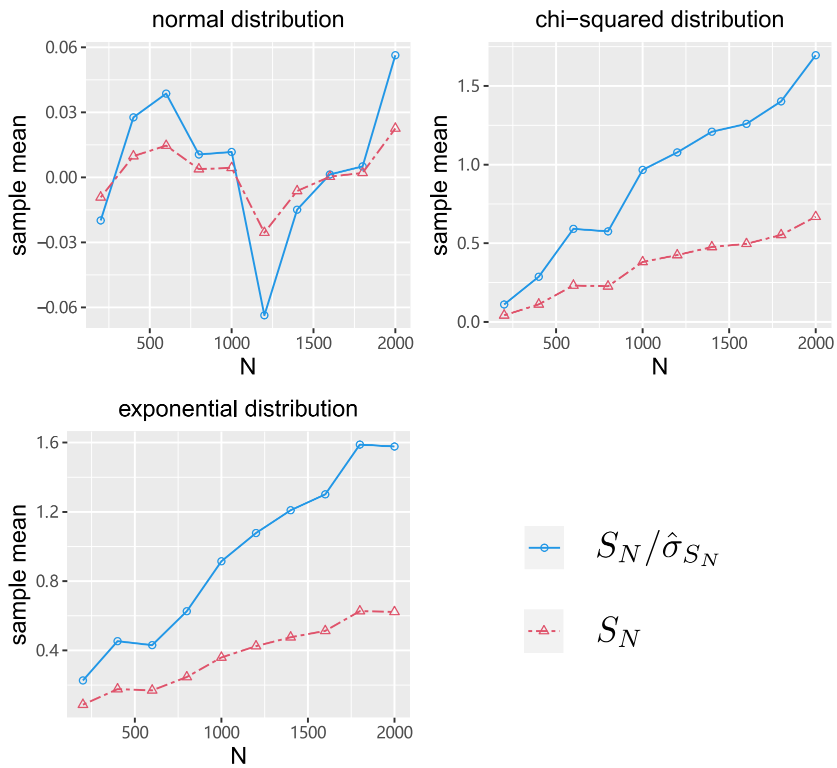

Figure S.1: Average of two statistics obtained under three different distributions based on 500 replications.

the chi-squared distribution. This suggests that when the distribution is asymmetric, the expectation of $\hat{\rho}_{ij}$ may not be exactly 0, which means that the last equation in (S1.1) may not hold when errors follow an asymmetric distribution. As a result, we assume in Assumption 1 that the distribution is symmetric to avoid bias in the test of S_N.

S1.2 Simulation results for very dense alternatives

We use the data generation process in (3.17), where ϵ_{it}^* are generated from the following setting.

(III) Very dense case. Let $(\epsilon_{1\cdot}^*, \cdots, \epsilon_{N\cdot}^*)' = \mathbf{W}^{1/2}(\epsilon_{1\cdot}, \cdots, \epsilon_{N\cdot})'$, where $\epsilon_{i\cdot}^* = (\epsilon_{i1}^*, \cdots, \epsilon_{iT}^*)'$ and $\epsilon_{i\cdot} = (\epsilon_{i1}, \cdots, \epsilon_{iT})'$, for all $1 \leq i \leq N$. Here, $\epsilon_{i\cdot}$ are generated from settings (i)-(ii) with distributions (1)-(3). $\mathbf{W}$ is constructed as follows. Let $(\mathbf{W})_{ii} = 1$, for $1 \leq i \leq N$. The non-zero number of non-diagonal elements of $\mathbf{W}$ is $2\lfloor N^{5/4} \rfloor$, and the number of non-zero positions is randomly selected. When $1 \leq i \neq j \leq N$, if $(\mathbf{W})_{ij} \neq 0$, then $(\mathbf{W})_{ij} \overset{iid}{\sim} U\left[\sqrt{\frac{1}{22}\frac{\log T}{N}}, \sqrt{\frac{3}{22}\frac{\log T}{N}}\right]$. Here, $\lfloor x \rfloor$ denotes the greatest integer less than or equal to $x \in \mathbb{R}$.

Table S.1 presents the empirical power for this very dense situation. When compared to the non-sparse case in Setting (I), Setting (III) is much denser. In this scenario, L_N does not perform as well as S_N and T_C, with S_N outperforming T_C. However, a review of Tables 2, 3, and S.1 reveals that T_C demonstrates robust performance across all these cases.

S1.3 Simulation results in situation of factor structure

To further compare the empirical power of the involved tests, we consider another common error structure with common factors. We consider the

Table S.1: The empirical power of the three tests at 5% level under case (III).

		Setting (i)								Setting (ii)							
p		3				5				3				5			
N	T	200	300	400	500	200	300	400	500	200	300	400	500	200	300	400	500
Normal distribution																	
100	S_N	75.3	89.3	95.8	99.0	73.3	87.8	95.3	99.2	69.0	84.2	93.3	97.4	64.3	81.6	92.8	97.1
	L_N	4.5	8.6	12.1	17.1	4.9	8.2	10.0	19.2	2.9	5.2	11.6	14.0	3.4	5.7	8.2	13.6
	T_C	68.4	86.1	94.3	98.2	65.6	82.9	94.2	97.8	58.8	79.1	91.4	95.4	56.2	76.0	89.1	96.0
200	S_N	58.3	79.0	88.3	94.8	60.0	79.0	89.9	95.0	52.7	71.8	83.1	91.0	54.2	71.7	84.1	84.1
	L_N	2.2	3.6	4.7	6.5	1.9	3.7	4.6	6.5	1.8	2.8	4.4	5.4	1.2	2.5	3.2	3.2
	T_C	47.1	70.6	84.7	92.1	49.9	72.4	83.2	92.4	40.4	62.4	75.2	85.9	41.1	62.7	76.8	76.8
t_6-distribution																	
100	S_N	74.6	89.6	96.1	99.0	73.7	87.4	95.3	98.8	67.5	84.0	93.1	97.2	66.9	82.3	91.0	96.6
	L_N	4.9	7.2	11.0	18.8	3.8	5.5	12.0	17.1	3.3	6.7	8.3	11.2	2.8	4.1	8.4	12.1
	T_C	66.9	85.1	94.6	98.2	64.9	84.2	93.8	98.0	57.0	79.8	90.1	95.5	56.8	76.8	89.7	94.9
200	S_N	61.2	78.2	89.0	95.2	60.9	78.8	89.6	95.2	53.3	72.0	83.3	90.9	54.8	70.5	82.1	92.4
	L_N	3.3	4.4	5.8	6.8	2.9	4.0	5.1	5.9	1.5	2.8	3.9	6.0	2.0	2.8	4.0	4.6
	T_C	50.4	71.5	84.4	91.9	50.6	70.4	82.9	92.3	42.9	62.8	76.8	86.5	42.5	61.9	76.9	86.8
χ_5^2-distribution																	
100	S_N	75.0	90.1	96.3	98.9	76.1	89.7	95.8	99.1	67.3	84.9	94.4	97.3	68.6	83.7	92.4	98.3
	L_N	5.3	7.5	13.8	19.4	5.2	7.9	12.3	19.2	3.7	6.4	9.5	16.4	4.2	5.5	9.3	14.5
	T_C	66.9	86.9	95.1	98.8	69.1	84.1	94.9	98.6	57.4	80.0	91.5	97.0	60.4	76.6	89.4	96.8
200	S_N	61.6	78.9	90.9	97.2	63.3	77.1	91.4	95.2	54.3	71.5	85.5	92.6	56.2	72.4	86.5	91.1
	L_N	2.5	5.1	6.2	8.1	3.5	4.7	8.1	7.9	1.6	3.7	4.9	7.3	1.8	3.4	4.7	5.6
	T_C	53.9	71.7	87.1	94.8	51.9	72.8	86.4	91.5	45.8	61.5	79.4	88.6	44.9	64.3	78.4	86.9

data generating process outlined in (3.17), where

$$\epsilon_{it}^* = \lambda_i f_t + \epsilon_{it}, \text{ for } i = 1, \ldots, N, \, t = 1, \ldots, T$$

are generated from the following two settings.

(a) Non-sparse case. We generate $f_t \overset{iid}{\sim} \mathcal{N}(0,1)$, $\lambda_i = (\log N/T)^{1/4}\xi_i$, for $i = 1, 2, 3$ and $\lambda_i = 0.7/T^{1/4}/N^{1/2}\xi_i u_i$, for $i = 4, \ldots, N$, where $\xi_i \overset{iid}{\sim} \chi_2^2$, $u_i \overset{iid}{\sim} U[\psi^{1/2}, \psi^{2/3}]$, and ϵ_{it} are generated from settings (i)-(ii) with distributions (1)-(3). Here, $\psi = T/N$.

Table S.2: The empirical power of the three tests at 5% level under case (a).

		Setting (i)								Setting (ii)							
p		3				5				3				5			
N	T	200	300	400	500	200	300	400	500	200	300	400	500	200	300	400	500
Normal distribution																	
100	S_N	53.6	81.7	94.5	98.4	53.7	80.5	93.4	98.3	35.5	62.4	79.6	89.9	36	60.9	79.6	91.1
	L_N	31.9	35.6	40.4	45.3	32.7	36.3	42.1	43.9	20.9	23.1	28.5	30.4	21.3	25.9	27.4	28.5
	T_C	58	80.7	92.6	96.7	59.4	79	91.5	97.5	37.6	60.5	76.1	86.9	39.4	58.6	76.2	87.9
200	S_N	17.8	33.1	49.2	63.3	18.7	35.3	49.9	67.4	12.3	20.1	29.1	40.8	13	21.6	30.7	44.2
	L_N	28.2	34.2	36.2	36.2	30	35.2	34.6	35.8	18.1	23.4	25.9	25.8	17.5	23.3	22.4	24.8
	T_C	33.9	45.6	56.9	67.2	36	48.7	57.6	69.5	21.8	29.7	37.4	46	22.3	32.8	37.7	47.3
t_6-distribution																	
100	S_N	55.8	83.4	93.3	97.6	56.3	82.3	94.1	98.7	34.9	61.2	79.6	90.9	35.7	59.8	80.2	91.4
	L_N	33.2	35.9	37.1	45.1	30.1	34.5	41.3	41.7	22.1	24.3	25.3	29.9	21.1	24.1	28.6	28.4
	T_C	58.3	80	91.5	97	58.3	79.9	92.5	97.3	40	59.6	76	88.9	39	59.7	74.7	87.4
200	S_N	17.7	30.6	47.6	67.5	19.7	31.8	49.5	67.2	11.3	19	28.4	44.2	13.1	20	31.8	46.1
	L_N	28.9	29.9	35.9	39.4	29.3	32.6	33.3	37	17.1	19.9	23.8	27.8	18.8	21.6	23.2	26.2
	T_C	33.8	41.4	54.7	70.7	34.8	44	57.9	70.1	22.1	27.6	37.1	50.3	22.4	29.9	37.7	48.1
χ^2_5-distribution																	
100	S_N	56.2	81.5	95.4	98.9	58.7	83.4	95.0	98.7	37.6	63.1	80.7	90.6	41	60.1	80.3	93.6
	L_N	31.9	40.4	39.1	45.4	31.7	38.4	39.6	43.6	21.5	28.6	28	29.6	22	25.3	26.3	30.3
	T_C	57.4	81.7	92.4	97.7	61.2	82.6	93.2	98.2	39	62.6	77.3	87.8	44	59	77.9	89.2
200	S_N	20.9	31.2	51.2	67.7	20.5	35.9	50.4	64.5	14	20.7	31.3	43.8	14.2	21.8	32.8	42.2
	L_N	30.5	35	35.5	40.5	29.1	34.1	36.7	38.1	20.9	25.5	25.6	28	17.6	24.7	27.9	27
	T_C	37.7	48.7	57.3	72	36.1	47.9	58.7	68.3	25.7	32.8	39.4	50.9	22.3	32.1	40.1	47.6

(b) Sparse case. We generate $f_t \overset{iid}{\sim} \mathcal{N}(0,1)$, $\lambda_i = 2.5(\log N/T)^{1/4}\xi_i u_i$, for $i = 1, \dots \lfloor N^{1/2} \rfloor$, and $\lambda_i = 0$, for $i = \lfloor N^{1/2} \rfloor + 1, \dots, N$, where $\xi_i \overset{iid}{\sim} \chi^2_2$, $u_i \overset{iid}{\sim} U[0.2, 0.4]$, and ϵ_{it} are generated from settings (i)-(ii) with distributions (1)-(3).

The results in Tables S.2 and S.3 are similar to those in Tables 2 and 3, respectively. In general, S_N performs better than L_N in the non-sparse case, and L_N performs better than S_N in the sparse case. It is worth noting

Table S.3: The empirical power of the three tests at 5% level under case (b)

		Setting (i)								Setting (ii)							
p		3				5				3				5			
N	T	200	300	400	500	200	300	400	500	200	300	400	500	200	300	400	500
Normal distribution																	
100	S_N	13.3	14.6	14	17.1	15.4	15.3	14.9	15.5	9.5	10.8	10.3	11.9	12.7	12.5	11.1	11
	L_N	74.4	78	77.8	80.5	75.5	74.9	78.7	82.7	55.8	61	62.3	66.6	56.4	59.8	63.6	68.9
	T_C	71.5	76.6	76.7	79.9	72.5	74.4	77.9	80.8	52.7	58.8	60.9	62.9	55	59.2	62.1	65.8
200	S_N	13.6	14.5	15.8	18.3	14.6	15.3	17.4	18.3	10.4	11.2	11	13.5	10.9	12.1	13.3	13.5
	L_N	84.8	87.7	92.4	91.1	89.1	89.8	91.3	91.2	66.5	74.4	80.3	80.6	72.6	74.9	79.4	80.6
	T_C	83	86.4	90	91.1	86.1	87.3	90.6	89.9	63.5	71.7	77.6	79.3	68.5	72.7	79	78.8
t_6-distribution																	
100	S_N	14.1	15.8	15	14.9	13.9	17	14.5	15.1	10.4	11.8	10.7	11.1	11.1	12.2	11	12.5
	L_N	77	77.3	79.7	80.8	72.1	76.5	79.4	83.8	57.7	58.2	64.8	68.7	54.7	60.8	63.9	68.9
	T_C	74.6	74.7	78	79.5	70.8	73.9	78.1	83	55.6	56.8	62.4	66.8	52.1	59.2	61.4	66.9
200	S_N	13.9	15.2	14.8	18.3	13.6	16	16.6	17.1	10	11.5	10	13.4	11.1	12.1	12.7	12.5
	L_N	83.6	88.8	90.3	90.7	86.9	89.3	91.4	91.8	66.7	74.8	79	80.5	68.4	76.1	79.1	80
	T_C	83.1	87.9	89	89.3	84.3	88.2	89.7	91.6	63.9	70.9	77.3	78.9	65.7	75.1	77.2	79.4
χ^2_5-distribution																	
100	S_N	13.3	15.7	15.8	15.5	14.8	14.1	14	15.9	10.4	11.6	12.1	11.6	11.6	10.7	10.2	12.2
	L_N	71.6	79	79.7	79.6	73	77.4	78.5	79.5	53.5	62.9	64.5	66	55.3	63	63.4	65.1
	T_C	69.6	76.7	76.9	78.6	71.6	74.9	75.6	78.6	51.5	59.1	62	63.7	52.8	61.6	60.4	62.6
200	S_N	16.6	16.2	18.2	19.4	15.8	17.2	16.3	17.2	12.7	11.6	13.5	14.7	12.2	12.9	12.6	12.7
	L_N	88.3	87.4	91.2	91.7	85.1	92.2	90.1	92.9	70.2	75	82.2	80.1	70.2	78.8	76.8	80.5
	T_C	86	87.1	89.9	90.4	83.4	90.9	88.5	90.9	66.6	72.9	79	77.6	68.3	76.3	74.6	79.6

that T_C performs well in both cases.

S2 Additional empirical application

We now apply the proposed tests to analyze the securities in the S&P 500 index. To account for the changes to the composition of the index over time, we compiled weekly returns on all the securities that constitute the S&P 500 index that have been listed over the period from January 2005 to September 2018. Because the securities that make up the index change over time, we only consider $N = 424$ securities that were included in the S&P 500 index during the entire period. A total of $T = 716$ consecutive observations were obtained. The time series data on the safe rate of return, and the market factors are obtained from Ken French's data library web page. The one-week US treasury bill rate is chosen as the risk-free rate (r_{ft}). The value-weighted return on all NYSE, AMEX, and NASDAQ stocks from CRSP is used as a proxy for the market return (r_{mt}). The average return on the three small portfolios minus the average return on the three big portfolios (SMB_t), and the average return on two value portfolios minus the average return on two growth portfolios (HML_t) are calculated based on the stocks listed on the NYSE, AMEX and NASDAQ.

We use the Fama-French three-factor model proposed by Fama and French (1993) to describe the above panel data, which is given by

$$Y_{it} = r_{it} - r_{ft} = \alpha_i + \beta_{i1}(r_{mt} - r_{ft}) + \beta_{i2}SMB_t + \beta_{i3}HML_t + \epsilon_{it}, \quad \text{(S2.2)}$$

for each $1 \leq i \leq N$ and $1 \leq t \leq T$, where $r_{mt} - r_{ft}$ is referred to as the market factor. Since cross-sectional dependence among the errors may lead to misspecified inference (Bernard, 1987), we are interested in testing

$$H_0 : \epsilon_{1\cdot}, \epsilon_{2\cdot}, \ldots, \epsilon_{N\cdot} \text{ are independent random vectors.} \quad \text{(S2.3)}$$

Before applying the proposed tests to the above panel data under the Fama-French three-factor model, we need to investigate whether there exists serial correlation in the residuals under such model. To this end, we applied the Box-Pierce test, a traditional test for autocorrelation, to the residual sequence of each security under the Fama-French three-factor model. Figure S.2 is the histogram of the p-values of the residual sequences, which suggests that for many securities the Box-Pierce tests are rejected. Furthermore, we applied a high-dimensional white noise test proposed by Li et al. (2019) to the sequences of the residual vector and the resulting p-value is 0, which also suggests rejecting. These results indicate that there exists serial correlation in the residuals under the Fama-French three-factor model. This may be because the model fails to take into account all the useful factors with serial correlation. See Schwartz and Whitcomb (1977) and Rosenberg and Rudd (1982) for further discussions on causes of serial correlation in the residuals. Hence, it is reasonable and necessary for us to use a test that allows serial correlation for the above panel data under the Fama-French three-factor model.

To accomodate the residual serial correlation, we used S_N, L_N and T_C to test the null hypothesis in (S2.3), and obtained the p-values 0, 0.92 and 0 respectively. This seems to suggest a dense residual correlation matrix

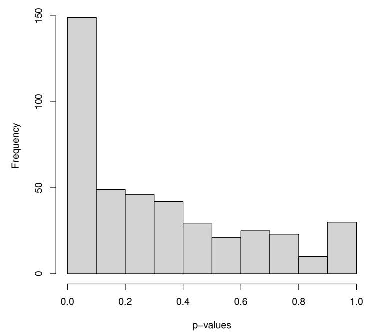

Figure S.2: Histogram of the p-values of the Box-Pierce tests for all residual sequences under the Fama-French three-factor model.

with weak signals overall. To verify this, we plotted the histogram and heat map of the absolute values of the cross-sectional correlations between all pairwise residual sequences of the securities in Figure S.3, which indicate that there are extensive non-zero correlations with small absolute values. This explains why the sum based test S_N and Fisher's combined probability test T_C are able to detect the deviation from H_0. In conclusion, because both S_N and T_C reject the null hypothesis, we reject the null hypothesis, and have the knowledge that the error correlation is nondiagonal with off diagonal values dense with mostly small magnitude.

Furthermore, we analyze the stock data using the Fama-French five-factor model. The model is defined as follows:

$$Y_{it} = r_{it} - r_{ft} = \alpha_i + \beta_{i1}(r_{mt} - r_{ft}) + \beta_{i2} SMB_t + \beta_{i3} HML_t$$

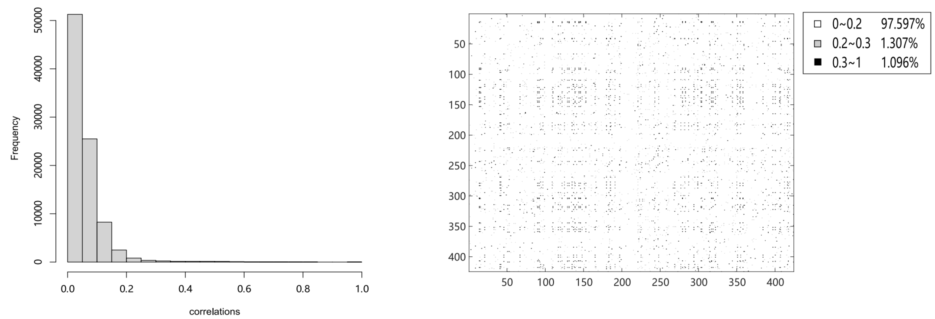

Figure S.3: Histogram and heat map of the absolute values of the correlations between all pairwise residual sequences of the securities under the Fama-French three-factor model.

$$+ \beta_{i4} RMW_t + \beta_{i5} CMA_t + \epsilon_{it}, \quad \text{(S2.4)}$$

Here, the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios (RMW_t), and the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios (CMA_t) are also calculated based on the stocks listed on the NYSE, AMEX and NASDAQ.

Similarly, we used the Box-Pierce test to examine autocorrelation within the residual series. The histogram of the p-values for the residual sequences, depicted in Figure S.4, also suggests that the Box-Pierce tests are rejected for many securities. We utilized S_N, L_N and T_C to test the null hypothe-

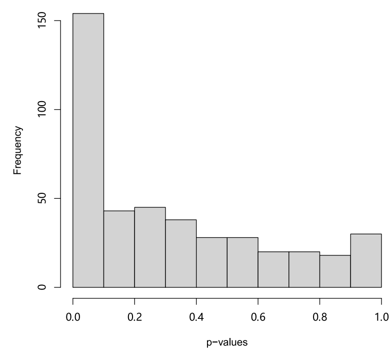

Figure S.4: Histogram of the p-values of the Box-Pierce tests for all residual sequences under the Fama-French five-factor model.

sis in (S2.3), which yields the p-values of 0.00, 0.93 and 0.00, respectively. Moreover, in Figure S.5, we plotted the histogram and heatmap of the absolute values of the cross-sectional correlations between all pairwise residuals. The results indicate that there are extensive non-zero correlations with small absolute values, which are very similar to the results obtained under the Fama-French three-factor model.

The above results imply that in addition to the three or five observed factors, there may be some unobserved factors that lead to cross-section correlations. Hence, we may assume that $\epsilon_{it} = \lambda_i' f_t + u_{it}$ in (S2.2), where f_t is a $r \times 1$ unobserved factor time series, λ_i is a $r \times 1$ unknown constant factor loading matrix and r is the number of factors. Similar to Bai (2009), we estimated the unobserved factors and their loadings based on the residuals. Specifically, we applied the method proposed in Lam et al. (2011) to the

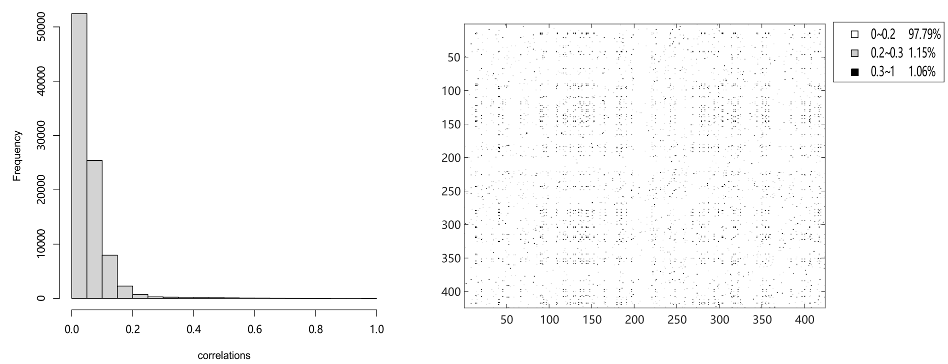

Figure S.5: Histogram and heat map of the absolute values of the correlations between all pairwise residual sequences of the securities under the Fama-French Five-factor model.

residuals at different values of r to obtain the estimates of f_t and λ_i, where the time lag is set to 25, and removed the estimates of $\lambda_i' f_t$ from the residuals to get the further residuals. Then, we applied the S_N, L_N and T_C tests to the further residuals. For example, when $r = 80$, the p-values obtained from the S_N, L_N and T_C tests are 0.69, 0.00 and 0.00, respectively. In fact, the L_N and T_C tests rejected H_0 because there are a few number of very large cross-sectional correlations, such as 0.93, -0.60. This situation occurs even when r is larger. We plotted the histogram and heatmap of the absolute values of the cross-sectional correlations of the further residuals in Figure S.6, which suggest that most cross-sectional correlation coefficients are very small. To further investigate the differences between these tests, we applied them to the data within the sliding windows constructed by 50 consecutive

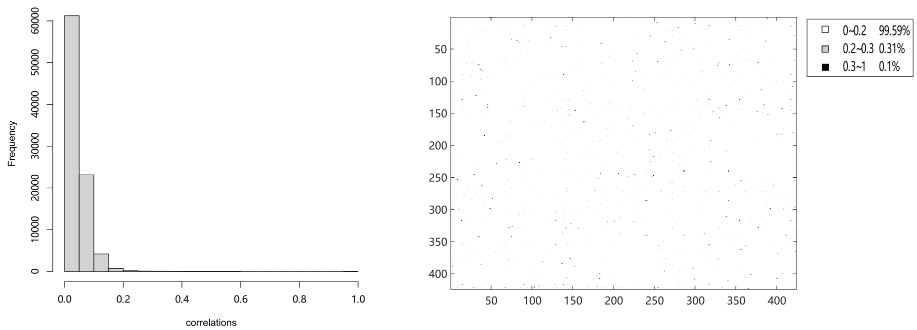

Figure S.6: Histogram and Heat map of the absolute values of the correlations between all pairwise residual sequences of the securities under the Fama-French three-factor model.

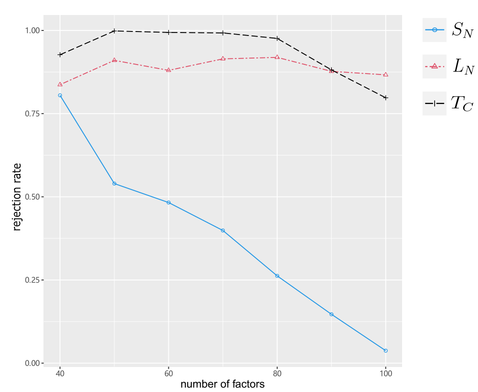

Figure S.7: The curves of the rejection rates of the S_N, L_N and T_C tests.

weeks with $r \in \{40, 50, 60, 70, 80, 90, 100\}$. The curves of the rejection rates of the S_N, L_N and T_C tests were presented in Figure S.7, which indicate that the L_N and T_C tests outperform the S_N test, and the T_C test exhibits very robust performance in terms of the number of unobserved factors.

S3 Proofs of the theorems

This section introduces a total of 14 lemmas. The first four lemmas are mainly used as proof tools. Next, We will concisely elucidate the role of the remaining ten lemmas in sequence.

Lemma S.5, which was proven by Feng et al. (2022), provides an upper bound on the distance between the sample correlation coefficient of the residual and that of the error.

Under Assumptions 1-3 and the null hypothesis H_0, Lemma S.6 establishes that $\hat{\sigma}_{ij}$ converges to $(\mathbf{\Sigma})_{ij}$ in distribution uniformly for all $1 \leq i, j \leq T$. Meanwhile, Lemma S.7 shows that $\hat{P}_N$ converges to 1 in distribution.

Similarly, under Assumptions 2-5 and local alternative hypotheses, Lemma S.8 proves that $\hat{\sigma}_{ij}$ converges to $(\mathbf{\Sigma})_{ij}$ in distribution uniformly for all $1 \leq i, j \leq T$. Furthermore, Lemma S.9 demonstrates that $\hat{P}_N$ converges to $\|\mathbf{\Phi}\|_{\mathrm{F}}^2/N$ in distribution.

Under Assumptions 1-3 and H_0, Lemma S.10 presents the asymptotic null distribution of S_N, and Lemma S.11 shows that $\hat{\sigma}_{S_N}$ is a ratio-consistent estimator of the variance of S_N. Moreover, Lemma S.12 establishes the premise that if $\tilde{T}_{\max}$ and S_N are asymptotically independent, it can be inferred that L_N and S_N are also asymptotically independent, where $\tilde{T}_{\max} = \max_{1\leq i<j\leq N}(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2/\left\|\mathbf{\Sigma}\right\|_{\mathrm{F}}^2$. Lastly, Lemma S.13 and Lemma S.14 respectively prove the two key results of asymptotic independence, which are formulas (2.13) and (2.14).

Let $m_k = E\left(Z_{it}^k\right)$, where Z_{it} is the element at row i and column t of matrix $\mathbf{Z}$. In order to analyze the moments of the quadratic form, we define some notations:

$$m_1 = 0, \quad m_2 = 1, \quad m_3 = \gamma_1, \quad m_4 = \gamma_2 + 3, \; m_5 = \gamma_3 + 10\gamma_1,$$

$$m_6 = \gamma_4 + 15\gamma_2 + 10\gamma_1^2 + 15, m_7 = \gamma_5 + 21\gamma_3 + 35\gamma_2\gamma_1 + 105\gamma_1,$$

$$m_8 = \gamma_6 + 28\gamma_4 + 56\gamma_3\gamma_1 + 35\gamma_2^2 + 210\gamma_2 + 280\gamma_1^2 + 105.$$

Note that by Assumption 1, we can conclude that $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6$ are finite.

To facilitate the theoretical derivations, we recall some useful lemmas as follows.

Lemma S.1. *(Lieberman, 1994; Pesaran and Yamagata, 2012) Let $X = (X_1 \cdots, X_T)'$ be a T-dimensional random vector. Let $\mathbf{A}$ and $\mathbf{B}$ be $T \times T$ nonstochastic matrices, where $\mathbf{A}$ is symmetric, $\mathbf{B}$ is semi-positive definite, and suppose that $X_1 \cdots, X_T$ are independently and identically distributed (iid) observations with zero mean and one variance. Denote the l-th cumu-*

lant of $X'\mathbf{B}X$ by κ_l and the $m+1$ order, $m+k$ degree generalised cumulant of the product of $(X'\mathbf{A}X)^k$ and $X'\mathbf{B}X$ by κ_{km}. If the following three conditions are true: (i) For $l = 1, 2, \ldots,$ we have $\kappa_l = O(T)$; (ii) for $k = 1, 2, \ldots,$ we have $\kappa_{k0} = E\{ (X'\mathbf{A}X)^k \} = O\left(T^k\right)$; (iii) for $k, m = 1, 2, \ldots,$ we have $\kappa_{km} = O\left(T^l\right)$, with $l \leqslant k$; then, we have

$$E\left\{\left(\frac{X'\mathbf{A}X}{X'\mathbf{B}X}\right)^k\right\} = \frac{E\{ (X'\mathbf{A}X)^k \}}{\{E(X'\mathbf{B}X)\}^k} + b_{k1} + O\left(T^{-2}\right), \tag{S3.5}$$

where

$$b_{k1} = \frac{k(k+1)}{2}\left[\frac{E\left\{(X'\mathbf{A}X)^k\right\}\kappa_2}{\{E\left(X'\mathbf{B}X\right)\}^{k+2}}\right] - k\left[\frac{\kappa_{k1}}{\{E\left(X'\mathbf{B}X\right)\}^{k+1}}\right] = O(T^{-1}),$$

$$\kappa_{k1} = E\left\{(X'\mathbf{A}X)^k\left(X'\mathbf{B}X\right)\right\} - E\left\{(X'\mathbf{A}X)^k\right\}E\left(X'\mathbf{B}X\right).$$

Remark S.1. Lieberman (1994) proposed the Laplace approximation of moments of the ratio of quadratic forms, where requires $\mathbf{B}$ to be a positive definite matrix. Pesaran and Yamagata (2012) relaxed this condition and allowed $\mathbf{B}$ to be a positive semi-definite matrix.

Lemma S.2. *(Baltagi et al., 2016) Set $\mathbf{M}_i = \mathbf{P}_i\mathbf{\Sigma}\mathbf{P}_i$ for all $1 \leq i \leq N$. For any fixed positive number k, we have (1) $\frac{1}{T}\operatorname{tr}\left(\mathbf{\Sigma}^k\right) = O(1)$; (2) $\frac{1}{T}\operatorname{tr}\left(\mathbf{M}_i^k\right) = O(1)$; (3) for $1 \leq i_1, i_2, \cdots, i_k \leq N$, $\operatorname{tr}\left(\mathbf{M}_{i_1}\mathbf{M}_{i_2}\cdots\mathbf{M}_{i_k}\right) = O(T)$.*

Lemma S.3. *(Abadir and Magnus, 2005) Let $\mathbf{A}$ and $\mathbf{B}$ be the same order square matrices.*

(1) $\text{tr}(\mathbf{A} \circ \mathbf{B}) \le \sqrt{\text{tr}(\mathbf{A}^2)\text{tr}(\mathbf{B}^2)}$, $\text{tr}(\mathbf{A} \circ \mathbf{A}) \le \text{tr}(\mathbf{A}^2)$, *when* $\mathbf{A}$, $\mathbf{B}$ *are symmetric;*

(2) $\text{tr}(\mathbf{A} \circ \mathbf{B}) \le C\text{tr}(\mathbf{B})$ *when* $\mathbf{A}$, $\mathbf{B}$ *are non-negative definite matrices and* $\lambda_{\max}(\mathbf{A}) \le C$ *for some constant* $C \ge 0$;

(3) $\mathbf{A} \circ \mathbf{B}$ *is positive semidefinite when* $\mathbf{A}$, $\mathbf{B}$ *are positive semidefinite (See Theorem 7.5.3 in Bernstein (2009));*

(4) $\tau_T' \mathbf{A} \circ \mathbf{B} \tau_T = \text{tr}(\mathbf{A}\mathbf{B}')$ *and* $\text{tr}^2(\mathbf{A}\mathbf{B}') \le \text{tr}(\mathbf{A}'\mathbf{A})\text{tr}(\mathbf{B}'\mathbf{B})$ *for any the same order square matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *(See Exercise 12.5 and 12.32 in Abadir and Magnus (2005)).*

Lemma S.4. *(Bao and Ullah, 2010) Suppose that* $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_T)'$, *and* $\xi_1, \xi_2, \ldots, \xi_T$ *are iid observations with zero mean, one variance,* $\gamma_1 = E\left(\xi_t^3\right)$, $\gamma_2 = E\left(\xi_t^4\right) - 3, \gamma_3 = E\left(\xi_t^5\right) - 10\gamma_1, \gamma_4 = E\left(\xi_t^6\right) - 15\gamma_2 - 10\gamma_1^2 - 15, \gamma_5 = E\left(\xi_t^7\right) - 21\gamma_3 - 35\gamma_2\gamma_1 - 105\gamma_1$ *and* $\gamma_6 = E\left(\xi_t^8\right) - 28\gamma_4 - 56\gamma_3\gamma_1 - 35\gamma_2^2 - 210\gamma_2 - 280\gamma_1^2 - 105$, *for all* $t = 1, 2, \ldots, T$. *Suppose that* $\mathbf{A}_j$, $j = 1, 2, 3, 4$ *are* $T \times T$ *real symmetric matrices, and* $\tau_T = (1, 1, \cdots, 1)'$ *is a* T*-dimensional vector. Then*

$$E\left(\boldsymbol{\xi}'\mathbf{A}_1\boldsymbol{\xi}\right) = \text{tr}\left(\mathbf{A}_1\right),$$

$$E\left[\left(\boldsymbol{\xi}'\mathbf{A}_1\boldsymbol{\xi}\right)\left(\boldsymbol{\xi}'\mathbf{A}_2\boldsymbol{\xi}\right)\right] = \gamma_2\text{tr}\left[\left(\mathbf{A}_1 \circ \mathbf{A}_2\right)\right] + \text{tr}\left(\mathbf{A}_1\right)\text{tr}\left(\mathbf{A}_2\right) + 2\text{tr}\left(\mathbf{A}_1\mathbf{A}_2\right),$$

$$E\left[\left(\boldsymbol{\xi}'\mathbf{A}_1\boldsymbol{\xi}\right)\left(\boldsymbol{\xi}'\mathbf{A}_2\boldsymbol{\xi}\right)\left(\boldsymbol{\xi}'\mathbf{A}_3\boldsymbol{\xi}\right)\right] = \gamma_4\text{tr}\left(\mathbf{A}_1 \circ \mathbf{A}_2 \circ \mathbf{A}_3\right) + \gamma_2\text{tr}\left(\mathbf{A}_1\right)\text{tr}\left(\mathbf{A}_2 \circ \mathbf{A}_3\right)$$

$$+\gamma_2 \mathrm{tr}(\mathbf{A}_2)\,\mathrm{tr}(\mathbf{A}_1\circ\mathbf{A}_3)+\gamma_2 \mathrm{tr}(\mathbf{A}_3)\,\mathrm{tr}(\mathbf{A}_1\circ\mathbf{A}_2)+4\gamma_2 \mathrm{tr}\left[\mathbf{A}_1\circ(\mathbf{A}_2\mathbf{A}_3)\right]$$

$$+4\gamma_2 \mathrm{tr}\left[\mathbf{A}_2\circ(\mathbf{A}_1\mathbf{A}_3)\right]+4\gamma_2 \mathrm{tr}\left[\mathbf{A}_3\circ(\mathbf{A}_1\mathbf{A}_2)\right]$$

$$+2\gamma_1^2\left[\tau_T'(\mathbf{I}_T\circ\mathbf{A}_1)\,\mathbf{A}_2\,(\mathbf{I}_T\circ\mathbf{A}_3)\,\tau_T\right]$$

$$+2\gamma_1^2\left[\tau_T'(\mathbf{I}_T\circ\mathbf{A}_1)\,\mathbf{A}_3\,(\mathbf{I}_T\circ\mathbf{A}_2)\,\tau_T\right]+2\gamma_1^2\left[\tau_T'(\mathbf{I}_T\circ\mathbf{A}_2)\,\mathbf{A}_1\,(\mathbf{I}_T\circ\mathbf{A}_3)\,\tau_T\right]$$

$$+4\gamma_1^2\left[\tau_T'(\mathbf{A}_1\circ\mathbf{A}_2\circ\mathbf{A}_3)\,\tau_T\right]+\mathrm{tr}(\mathbf{A}_1)\,\mathrm{tr}(\mathbf{A}_2)\,\mathrm{tr}(\mathbf{A}_3)+2\mathrm{tr}(\mathbf{A}_1)\,\mathrm{tr}(\mathbf{A}_2\mathbf{A}_3)$$

$$+2\mathrm{tr}(\mathbf{A}_2)\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_3)+2\mathrm{tr}(\mathbf{A}_3)\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_2)+8\mathrm{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3)\,,$$

$$E\left[(\boldsymbol{\xi}'\mathbf{A}_1\boldsymbol{\xi})\,(\boldsymbol{\xi}'\mathbf{A}_2\boldsymbol{\xi})\,(\boldsymbol{\xi}'\mathbf{A}_3\boldsymbol{\xi})\,(\boldsymbol{\xi}'\mathbf{A}_4\boldsymbol{\xi})\right]=\mathrm{tr}(\mathbf{A}_1)\,\mathrm{tr}(\mathbf{A}_2)\,\mathrm{tr}(\mathbf{A}_3)\,\mathrm{tr}(\mathbf{A}_4)$$

$$+2\left[\mathrm{tr}(\mathbf{A}_1)\,\mathrm{tr}(\mathbf{A}_2)\,\mathrm{tr}(\mathbf{A}_3\mathbf{A}_4)+\mathrm{tr}(\mathbf{A}_1)\,\mathrm{tr}(\mathbf{A}_3)\,\mathrm{tr}(\mathbf{A}_2\mathbf{A}_4)\right.$$

$$+\mathrm{tr}(\mathbf{A}_1)\,\mathrm{tr}(\mathbf{A}_4)\,\mathrm{tr}(\mathbf{A}_2\mathbf{A}_3)+\mathrm{tr}(\mathbf{A}_2)\,\mathrm{tr}(\mathbf{A}_3)\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_4)$$

$$\left.+\mathrm{tr}(\mathbf{A}_2)\,\mathrm{tr}(\mathbf{A}_4)\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_3)+\mathrm{tr}(\mathbf{A}_3)\,\mathrm{tr}(\mathbf{A}_4)\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_2)\right]$$

$$+4\left[\mathrm{tr}(\mathbf{A}_1\mathbf{A}_2)\,\mathrm{tr}(\mathbf{A}_3\mathbf{A}_4)+\mathrm{tr}(\mathbf{A}_1\mathbf{A}_3)\,\mathrm{tr}(\mathbf{A}_2\mathbf{A}_4)+\mathrm{tr}(\mathbf{A}_1\mathbf{A}_4)\,\mathrm{tr}(\mathbf{A}_2\mathbf{A}_3)\right]$$

$$+8\left[\mathrm{tr}(\mathbf{A}_1)\,\mathrm{tr}(\mathbf{A}_2\mathbf{A}_3\mathbf{A}_4)+\mathrm{tr}(\mathbf{A}_2)\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_3\mathbf{A}_4)+\mathrm{tr}(\mathbf{A}_3)\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_4)\right.$$

$$\left.+\mathrm{tr}(\mathbf{A}_4)\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3)\right]+16\left[\mathrm{tr}(\mathbf{A}_1\mathbf{A}_3\mathbf{A}_4\mathbf{A}_2)+\mathrm{tr}(\mathbf{A}_1\mathbf{A}_4\mathbf{A}_2\mathbf{A}_3)\right.$$

$$\left.+\mathrm{tr}(\mathbf{A}_1\mathbf{A}_4\mathbf{A}_3\mathbf{A}_2)\right]+\gamma_2 f_{\gamma_2}+\gamma_4 f_{\gamma_4}+\gamma_6 f_{\gamma_6}+\gamma_1^2 f_{\gamma_1^2}+\gamma_2^2 f_{\gamma_2^2}+\gamma_1\gamma_3 f_{\gamma_1\gamma_3},$$

where according to Bao and Ullah (2010), we can obtain the expressions for $f_{\gamma_2}, f_{\gamma_4}, f_{\gamma_6}, f_{\gamma_1^2}, f_{\gamma_2^2}$ *and* $f_{\gamma_1\gamma_3}$.

Now, we are ready to present the proofs of the theorems in Section 2.

S3.1 Proof of Theorem 1

Recall that for any $1 \le i \le N$, $\epsilon_{i\cdot} = \mathbf{R} Z_{i\cdot}$,

$$\tilde{T}_{ij} \doteq \frac{\epsilon_{i\cdot}' \epsilon_{j\cdot}}{\|\boldsymbol{\Lambda}\|_{\mathrm{F}}} = \frac{Z_{i\cdot}' \boldsymbol{\Lambda} Z_{j\cdot}}{\|\boldsymbol{\Lambda}\|_{\mathrm{F}}}, \qquad \rho_{ij} = \frac{\epsilon_{i\cdot}' \epsilon_{j\cdot}}{\|\epsilon_{i\cdot}\| \times \|\epsilon_{j\cdot}\|},$$

where $\boldsymbol{\Lambda} = \mathbf{R}'\mathbf{R}$. Because $\boldsymbol{\Sigma} = \mathbf{R}\mathbf{R}'$, we have $\operatorname{tr}(\boldsymbol{\Sigma}) = \operatorname{tr}(\boldsymbol{\Lambda})$, $\operatorname{tr}(\boldsymbol{\Sigma}^2) = \operatorname{tr}(\boldsymbol{\Lambda}^2)$, and the matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$ have the same eigenvalues. Hence, we have $\hat{\epsilon}_{i\cdot} = \mathbf{P}_i \mathbf{R} Z_{i\cdot}$.

Lemma S.5. *(Feng et al., 2022) Set* $\mathbf{B}_i = \mathbf{x}_i \left(\mathbf{x}_i'\mathbf{x}_i\right)^{-1} \mathbf{x}_i'$, *for* $1 \le i \le N$. *Then,*

$$\max_{1 \le i < j \le N} |\hat{\rho}_{ij} - \rho_{ij}| \le 14 \cdot \left(\max_{1 \le i \le j \le N} \frac{\epsilon_{j\cdot}' \mathbf{B}_i \epsilon_{j\cdot}}{\epsilon_{j\cdot}' \epsilon_{j\cdot}} \right).$$

We now are ready to derive the asymptotic null distribution of L_N. First, we will show that for $y \in \mathbb{R}$, under H_0,

$$P\left(\max_{1 \le i < j \le N} \tilde{T}_{ij}^2 - 4\log N + \log\log N \le y \right) \to \exp\left\{ -\frac{1}{\sqrt{8\pi}} \exp\left(-\frac{y}{2} \right) \right\}. \tag{S3.6}$$

By Theorem 1 in Arratia et al. (1989), we have

$$\left| P\left(\max_{1 \le i < j \le N} \tilde{T}_{ij}^2 \le t_N \right) - e^{-\tau_N} \right| \le b_{1N} + b_{2N} + b_{3N}$$

where $t_N = 4\log N - \log\log N + y, \tau_N = \sum\sum_{1 \le i < j \le N} P\left(\tilde{T}_{12}^2 > t_N \right)$,

$$b_{1N} = \sum\sum_{1 \le i < j \le N} \sum_{(k,l) \in B_{ij}} P\left(\tilde{T}_{ij}^2 > t_N \right) P\left(\tilde{T}_{kl}^2 > t_N \right) \le N^3 \left[P\left(\tilde{T}_{12}^2 > t_N \right) \right]^2,$$

$$b_{2N} = \sum\sum_{1 \le i < j \le N} \sum_{(k,l) \in B_{ij} \backslash \{(i,j)\}} P\left(\tilde{T}_{ij}^2 > t_N, \tilde{T}_{kl}^2 > t_N\right)$$

$$\le N^3 P\left(\tilde{T}_{12}^2 > t_N, \tilde{T}_{13}^2 > t_N\right),$$

$$b_{3N} = \sum\sum_{1 \le i < j \le N} E\left| P\left\{\tilde{T}_{ij}^2 > t_N | \sigma(\tilde{T}_{kl}^2 : (k,l) \notin B_{ij})\right\} - P\left(\tilde{T}_{12}^2 > t_N\right)\right| = 0,$$

for all $1 \le i < j \le N$, $B_{ij} \doteq \{(k,l) : 1 \le k < l \le N, \{k,l\} \cap \{i,j\} \ne \varnothing\}$, and the third term b_{3N} on the right side of the inequality is equal to zero because for four different indices i, j, k, l, $\tilde{T}_{ij}$ and $\tilde{T}_{kl}$ are independent. Note that according to Theorem 1.1 in Rudelson and Vershynin (2013), we have for any large $M > 0$, there exists some $C_1 > 0$ such that

$$P\left(\frac{Z_{2\cdot}'\mathbf{\Lambda}^2 Z_{2\cdot}}{\mathrm{tr}(\mathbf{\Lambda}^2)} > 1 + \varepsilon_1\right) \le 2N^{-M},$$

where $\varepsilon_1 = C_1\sqrt{\log N/\mathrm{tr}(\mathbf{\Lambda}^2)}$. By Corollary 3.1 in Saulis and Statulevicius (1991) and $\mathrm{tr}(\mathbf{\Lambda}^2) = \|\mathbf{\Lambda}\|_{\mathrm{F}}^2$, we have

$$\begin{aligned} P\left(\tilde{T}_{12}^2 > t_N\right) =& P\left(\left|\tilde{T}_{12}\right| > \sqrt{t_N}\right) \\ =& P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda} Z_{2\cdot}|}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}^2)}} > \sqrt{t_N}\right\} \\ \le& P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda} Z_{2\cdot}|}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}^2)}} > \sqrt{t_N}, \frac{Z_{2\cdot}'\mathbf{\Lambda}^2 Z_{2\cdot}}{\mathrm{tr}(\mathbf{\Lambda}^2)} \le 1 + \varepsilon_1\right\} \\ &+ P\left(\frac{Z_{2\cdot}'\mathbf{\Lambda}^2 Z_{2\cdot}}{\mathrm{tr}(\mathbf{\Lambda}^2)} > 1 + \varepsilon_1\right) \\ \le& P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda} Z_{2\cdot}|}{\sqrt{Z_{2\cdot}'\mathbf{\Lambda}^2 Z_{2\cdot}}} > \frac{\sqrt{t_N}}{1 + \varepsilon_1}\right\} + 2N^{-M} \end{aligned} \tag{S3.7}$$

$$
\begin{aligned}
&\le E\left[E\left\{I\left(\frac{|Z_{1\cdot}'\mathbf{\Lambda}Z_{2\cdot}|}{\sqrt{Z_{2\cdot}'\mathbf{\Lambda}^2Z_{2\cdot}}}>\frac{\sqrt{t_N}}{1+\varepsilon_1}\right)\Big|Z_{2\cdot}\right\}\right]+2N^{-M}\\
&\le E\left\{P\left(\frac{|Z_{1\cdot}'\mathbf{\Lambda}Z_{2\cdot}|}{\sqrt{Z_{2\cdot}'\mathbf{\Lambda}^2Z_{2\cdot}}}>\frac{\sqrt{t_N}}{1+\varepsilon_1}\Big|Z_{2\cdot}\right)\right\}+2N^{-M}\\
&=\{1+o(1)\}\frac{2}{\sqrt{2\pi t_N}}\exp^{-t_N/2}+2N^{-M}\\
&=O(N^{-2}), \qquad\qquad (S3.8)
\end{aligned}
$$

where last equality holds due to Proposition 2.1.2 in Vershynin (2018), $I(\cdot)$ denotes indicative function, and M is sufficiently large. Similarly, we have

$$
\begin{aligned}
P\left(\tilde{T}_{12}^2\le t_N\right)&=P\left(\left|\tilde{T}_{12}\right|\le\sqrt{t_N}\right)\\
&=P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda}Z_{2\cdot}|}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}^2)}}\le\sqrt{t_N}\right\}\\
&\le P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda}Z_{2\cdot}|}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}^2)}}\le\sqrt{t_N},\frac{Z_{2\cdot}'\mathbf{\Lambda}^2Z_{2\cdot}}{\mathrm{tr}(\mathbf{\Lambda}^2)}\ge1-\varepsilon_1\right\}\\
&\quad+P\left(\frac{Z_{2\cdot}'\mathbf{\Lambda}^2Z_{2\cdot}}{\mathrm{tr}(\mathbf{\Lambda}^2)}<1-\varepsilon_1\right)\\
&\le P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda}Z_{2\cdot}|}{\sqrt{Z_{2\cdot}'\mathbf{\Lambda}^2Z_{2\cdot}}}\le\frac{\sqrt{t_N}}{1-\varepsilon_1}\right\}+2N^{-M},
\end{aligned}
$$

then

$$
1-P\left(\tilde{T}_{12}^2<t_N\right)\ge1-P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda}Z_{2\cdot}|}{\sqrt{Z_{2\cdot}'\mathbf{\Lambda}^2Z_{2\cdot}}}<\frac{\sqrt{t_N}}{1-\varepsilon_1}\right\}-2N^{-M}.
$$

Hence, we have

$$
\begin{aligned}
&P\left(\tilde{T}_{12}^2>t_N\right)\\
&\ge P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda}Z_{2\cdot}|}{\sqrt{Z_{2\cdot}'\mathbf{\Lambda}^2Z_{2\cdot}}}>\frac{\sqrt{t_N}}{1-\varepsilon_1}\right\}-2N^{-M}
\end{aligned}
$$

$$
\begin{aligned}
=&\{1-o(1)\}\frac{2}{\sqrt{2\pi t_N}}\exp^{-t_N/2}-2N^{-M}\\
=&O(N^{-2}).
\end{aligned}
$$

This shows that $\tau_N \sim \frac{1}{\sqrt{8\pi}}e^{-y/2}$ and $b_{1N} \le CN^{-1}$. For b_{2N}, we have

$$
P\Big(\tilde{T}_{12}^2 > t_N, \tilde{T}_{13}^2 > t_N\Big) \le P\Big(\big|\tilde{T}_{12}-\tilde{T}_{13}\big| \ge 2\sqrt{t_N}\Big)+P\Big(\big|\tilde{T}_{12}+\tilde{T}_{13}\big| \ge 2\sqrt{t_N}\Big).
$$

Again, by Theorem 1.1 in Rudelson and Vershynin (2013), for any large $M > 0$, there exists some $C_2 > 0$ such that

$$
P\left(\frac{(Z_{2\cdot}-Z_{3\cdot})'\mathbf{\Lambda}^2(Z_{2\cdot}-Z_{3\cdot})}{\mathrm{tr}(\mathbf{\Lambda}^2)} > 2(1+\varepsilon_2)\right) \le 2N^{-M},
$$

where $\varepsilon_2 = C_2\sqrt{\log N/\mathrm{tr}(\mathbf{\Lambda}^2)}$. By Corollary 3.1 (Cramér type moderate deviation results) in Saulis and Statulevicius (1991), we have

$$
\begin{aligned}
&P\left(\left|\tilde{T}_{12}-\tilde{T}_{13}\right| \ge 2\sqrt{t_N}\right)\\
=&P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda}\,(Z_{2\cdot}-Z_{3\cdot})|}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}^2)}} \ge 2\sqrt{t_N}\right\}\\
\le&P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda}\,(Z_{2\cdot}-Z_{3\cdot})|}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}^2)}} \ge 2\sqrt{t_N}, \frac{(Z_{2\cdot}-Z_{3\cdot})'\mathbf{\Lambda}^2(Z_{2\cdot}-Z_{3\cdot})}{\mathrm{tr}(\mathbf{\Lambda}^2)} < 2(1+\varepsilon_2)\right\}\\
&+P\left\{\frac{(Z_{2\cdot}-Z_{3\cdot})'\mathbf{\Lambda}^2(Z_{2\cdot}-Z_{3\cdot})}{\mathrm{tr}(\mathbf{\Lambda}^2)} > 2(1+\varepsilon_2)\right\}\\
\le&P\left\{\frac{|Z_{1\cdot}'\mathbf{\Lambda}\,(Z_{2\cdot}-Z_{3\cdot})|}{\sqrt{(Z_{2\cdot}-Z_{3\cdot})'\mathbf{\Lambda}^2(Z_{2\cdot}-Z_{3\cdot})}} \ge \frac{\sqrt{2t_N}}{1+\varepsilon_2}\right\}+2N^{-M}\\
\le&\{1+o(1)\}\frac{\sqrt{\log N}}{2\sqrt{\pi}}N^{-4}e^{-y},
\end{aligned}
$$

where M is sufficiently large. Similarly,

$$P\Big(\big|\tilde{T}_{12}+\tilde{T}_{13}\big| \geq 2\sqrt{t_n}\Big) = \{1+o(1)\}\frac{\sqrt{\log N}}{2\sqrt{\pi}}N^{-4}e^{-y}.$$

Combining these inequalities, we have $b_{2N} \leq CN^{-1}\sqrt{\log N}$ and

$$P\left(\max_{1\leq i<j\leq N}\frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2} - 4\log N + \log\log N \leq y\right) \tag{S3.9}$$

$$\to G(y) = \exp\left\{-\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{y}{2}\right)\right\}. \tag{S3.10}$$

Similarly, due to Theorem 1.1 in Rudelson and Vershynin (2013), we have

$$P\left(\max_{1\leq i\leq N}|\epsilon_{i\cdot}'\epsilon_{i\cdot} - \mathrm{tr}\left(\mathbf{\Lambda}\right)| \geq C'\sqrt{T\log N}\right) = O\left(N^{-M}\right), \tag{S3.11}$$

where $M>0$ is sufficiently large, C' is a constant that depends on M. Due to (S3.9), we have

$$P\left(\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{T^2\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\max_{1\leq i<j\leq N}\frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{T^2}} - 4\log N + \log\log N \leq y\right)$$

$$\to \exp\left\{-\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{y}{2}\right)\right\}. \tag{S3.12}$$

Set

$$\Omega = \left\{\max_{1\leq i<j\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\mathrm{tr}\left(\mathbf{\Lambda}\right)}{T^2} + \frac{\epsilon_{j\cdot}'\epsilon_{j\cdot}\mathrm{tr}\left(\mathbf{\Lambda}\right)}{T^2} - \frac{2\mathrm{tr}^2(\mathbf{\Lambda})}{T^2}\right| \leq 2C_1\sqrt{\frac{\log N}{T}}\right\}.$$

Obviously, due to (S3.11), we have

$$\begin{aligned}P\left(\Omega^c\right) \leq & P\left\{\max_{1\leq i<j\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}}{T} + \frac{\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T} - \frac{2\mathrm{tr}\left(\mathbf{\Lambda}\right)}{T}\right| > \frac{2C_1T}{\mathrm{tr}\left(\mathbf{\Lambda}\right)}\sqrt{\frac{\log N}{T}}\right\}\\ \leq & P\left\{\max_{1\leq i<j\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}}{T} + \frac{\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T} - \frac{2\mathrm{tr}\left(\mathbf{\Lambda}\right)}{T}\right| > \frac{2C_1}{\lambda_{\max}(\mathbf{\Lambda})}\sqrt{\frac{\log N}{T}}\right\}\end{aligned}$$

$$
\begin{aligned}
\leq & P\left\{\max_{1\leq i\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}}{T}-\frac{\operatorname{tr}(\mathbf{\Lambda})}{T}\right|>\frac{C_1}{\lambda_{\max}(\mathbf{\Lambda})}\sqrt{\frac{\log N}{T}}\right\} \\
& + P\left\{\max_{1\leq j\leq N}\left|\frac{\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T}-\frac{\operatorname{tr}(\mathbf{\Lambda})}{T}\right|>\frac{C_1}{\lambda_{\max}(\mathbf{\Lambda})}\sqrt{\frac{\log N}{T}}\right\}=O\left(N^{-M}\right).
\end{aligned}
$$

We claim that for some constant $C_3>0$ such that

$$
P\left(\max_{1\leq i<j\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T^2}-\frac{\operatorname{tr}^2(\mathbf{\Lambda})}{T^2}\right|>C_3\sqrt{\frac{\log N}{T}}\right)\to 0
$$

as N,T are sufficiently large. Notice that

$$
\begin{aligned}
& P\left(\max_{1\leq i<j\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T^2}-\frac{\operatorname{tr}^2(\mathbf{\Lambda})}{T^2}\right|>C_3\sqrt{\frac{\log N}{T}}\right) \\
\leq & P\left(\max_{1\leq i<j\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T^2}-\frac{\operatorname{tr}^2(\mathbf{\Lambda})}{T^2}\right|>C_3\sqrt{\frac{\log N}{T}},\Omega\right)+P\left(\Omega^c\right),
\end{aligned}
$$

we just need to prove that

$$
P\left(\max_{1\leq i<j\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T^2}-\frac{\operatorname{tr}^2(\mathbf{\Lambda})}{T^2}\right|>C_3\sqrt{\frac{\log N}{T}},\Omega\right)\to 0.
$$

In fact, due to (S3.11), for sufficiently large C_3 satisfying $C_3-2C_1>C_3\sqrt{\frac{1}{2}}$ and $\sqrt{\frac{C_3}{2}}>C_1$, we have

$$
\begin{aligned}
& P\Big(\max_{1\leq i<j\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T^2}-\frac{\operatorname{tr}^2(\mathbf{\Lambda})}{T^2}\right|>C_3\sqrt{\frac{\log N}{T}},\Omega\Big) \\
\leq & P\Big\{\max_{1\leq i<j\leq N}\left|\Big(\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}}{T}-\frac{\operatorname{tr}(\mathbf{\Lambda})}{T}\Big)\Big(\frac{\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T}-\frac{\operatorname{tr}(\mathbf{\Lambda})}{T}\Big)\right|>C_3\sqrt{\frac{\log N}{2T}}\Big\} \\
\leq & P\Big\{\max_{1\leq i\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}}{T}-\frac{\operatorname{tr}(\mathbf{\Lambda})}{T}\right|>\sqrt{C_3}\Big(\frac{\log N}{2T}\Big)^{1/4}\Big\} \\
& + P\Big\{\max_{1\leq j\leq N}\left|\frac{\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T}-\frac{\operatorname{tr}(\mathbf{\Lambda})}{T}\right|>\sqrt{C_3}\Big(\frac{\log N}{2T}\Big)^{1/4}\Big\} \\
\leq & P\Big\{\max_{1\leq i\leq N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}}{T}-\frac{\operatorname{tr}(\mathbf{\Lambda})}{T}\right|>\sqrt{\frac{C_3\log N}{2T}}\Big\}
\end{aligned}
$$

$$
\begin{aligned}
&+P\Big\{\max_{1\le j\le N}\left|\frac{\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T}-\frac{\mathrm{tr}\,(\boldsymbol{\Lambda})}{T}\right|>\sqrt{\frac{C_3\log N}{2T}}\Big\}\\
=&O\Big(N^{-M}\Big),
\end{aligned}
$$

where the last inequality holds due to (S3.11). Then, we can conclude that

$$
P\left(\max_{1\le i<j\le N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T^2}-\frac{\mathrm{tr}^2(\boldsymbol{\Lambda})}{T^2}\right|>C_3\sqrt{\frac{\log N}{T}}\right)=O\left(N^{-M}\right)
$$

and

$$
P\left(\max_{1\le i<j\le N}\left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}/T^2}{\mathrm{tr}^2(\boldsymbol{\Lambda})/T^2}-1\right|<\frac{C_3}{\lambda_{\min}^2(\boldsymbol{\Lambda})}\sqrt{\frac{\log N}{T}}\right)\to 1.
$$

Thus, with probability tending to one, we have

$$
\begin{aligned}
&\max_{1\le i<j\le N}\left|\frac{\mathrm{tr}^2(\boldsymbol{\Lambda})/T^2}{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}/T^2}-1\right|\\
\le&\max_{1\le i<j\le N}\frac{T^2}{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}\max_{1\le i<j\le N}\left|\frac{\mathrm{tr}^2(\boldsymbol{\Lambda})}{T^2}-\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T^2}\right|\\
\le&\,2\max_{1\le i<j\le N}\frac{T^2}{\mathrm{tr}^2(\boldsymbol{\Lambda})}\max_{1\le i<j\le N}\left|\frac{\mathrm{tr}^2(\boldsymbol{\Lambda})}{T^2}-\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T^2}\right|\\
=&\,O_p\left\{\sqrt{\frac{\log N}{T}}\right\}.
\end{aligned}
$$

This, together with (S3.12), we have

$$
\begin{aligned}
&P\left(\frac{\mathrm{tr}^2(\boldsymbol{\Lambda})}{\|\boldsymbol{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}\frac{\left(\epsilon_{i\cdot}'\epsilon_{j\cdot}\right)^2}{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}-4\log N+\log\log N\le y\right)\\
=&P\left(\frac{\mathrm{tr}^2(\boldsymbol{\Lambda})}{\|\boldsymbol{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}\rho_{ij}^2-4\log N+\log\log N\le y\right)\\
&\to\exp\left\{-\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{y}{2}\right)\right\},
\end{aligned}
\tag{S3.13}
$$

for any $y\in\mathbb{R}$.

Then, we want to prove that

$$P\Big(\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}\hat{\rho}_{ij}^2-4\log N+\log\log N\le y\Big)\to\exp\Big\{-\frac{\exp\big(-y/2\big)}{\sqrt{8\pi}}\Big\}.$$

So, we just need to show that for any $\epsilon>0$

$$P\left(\max_{1\le i<j\le N}|\hat{\rho}_{ij}-\rho_{ij}|>\frac{\epsilon\sqrt{\mathrm{tr}\,(\mathbf{\Lambda}^2)}}{\sqrt{\log N}\mathrm{tr}\,(\mathbf{\Lambda})}\right)\to 0.$$

This is because if it holds, we have

$$\sqrt{\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\log N}\cdot\left(\max_{1\le i<j\le N}|\hat{\rho}_{ij}|-\max_{1\le i<j\le N}|\rho_{ij}|\right)\to 0$$

in probability. Set $\Delta=\max_{1\le i<j\le N}|\hat{\rho}_{ij}|-\max_{1\le i<j\le N}|\rho_{ij}|$. Then

$$\begin{aligned}&\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}|\hat{\rho}_{ij}|^2\\ =&\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\left(\max_{1\le i<j\le N}|\rho_{ij}|+\Delta\right)^2\\ =&\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}|\rho_{ij}|^2+2\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}|\rho_{ij}|\,\Delta+\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\Delta^2.\end{aligned}$$

The Slutsky lemma and (S3.13) say that

$$\frac{(\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2/\log N)^{1/2}}\max_{1\le i<j\le N}|\rho_{ij}|\to 2$$

in probability. Consequently,

$$\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}|\rho_{ij}|\,\Delta=\sqrt{\frac{\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}}{\log N}}\max_{1\le i<j\le N}|\rho_{ij}|\cdot\Big(\sqrt{\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\log N}\Delta\Big)\to 0$$

$$\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\Delta^2=\Big[\sqrt{\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\log N}\Delta\Big]^2\cdot\frac{1}{\log N}\to 0$$

in probability. Then, we can obtain that

$$\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}|\hat{\rho}_{ij}|^2=\frac{\mathrm{tr}^2(\mathbf{\Lambda})}{\|\mathbf{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}|\rho_{ij}|^2+o_p(1).$$

So, we will prove

$$P\left(\max_{1\le i<j\le N}|\hat{\rho}_{ij}-\rho_{ij}|>\frac{\epsilon\sqrt{\mathrm{tr}\left(\mathbf{\Lambda}^2\right)}}{\sqrt{\log N}\mathrm{tr}\left(\mathbf{\Lambda}\right)}\right)\to 0. \tag{S3.14}$$

By Lemma S.5,

$$\begin{aligned}
&P\left(\max_{1\le i<j\le N}|\hat{\rho}_{ij}-\rho_{ij}|>\frac{\epsilon\sqrt{\mathrm{tr}\left(\mathbf{\Lambda}^2\right)}}{\sqrt{\log N}\mathrm{tr}\left(\mathbf{\Lambda}\right)}\right)\\
\le& N^2\max_{1\le i<j\le N}P\left(\frac{\epsilon_{j\cdot}'\mathbf{B}_i\epsilon_{j\cdot}}{\epsilon_{j\cdot}'\epsilon_{j\cdot}}>\frac{\epsilon\sqrt{\mathrm{tr}\left(\mathbf{\Lambda}^2\right)}}{14\sqrt{\log N}\mathrm{tr}\left(\mathbf{\Lambda}\right)}\right)\\
\le& N^2\max_{1\le i\le N}P\left(\frac{\epsilon_{1\cdot}'\mathbf{B}_i\epsilon_{1\cdot}}{\epsilon_{1\cdot}'\epsilon_{1\cdot}}>\frac{\epsilon\sqrt{\mathrm{tr}\left(\mathbf{\Lambda}^2\right)}}{14\sqrt{\log N}\mathrm{tr}\left(\mathbf{\Lambda}\right)}\right)\\
\le& N^2\max_{1\le i\le N}\left[P\left(\frac{\epsilon_{1\cdot}'\mathbf{B}_i\epsilon_{1\cdot}}{\epsilon_{1\cdot}'\epsilon_{1\cdot}}>\frac{\epsilon\sqrt{\mathrm{tr}\left(\mathbf{\Lambda}^2\right)}}{14\sqrt{\log N}\mathrm{tr}\left(\mathbf{\Lambda}\right)},\epsilon_{1\cdot}'\epsilon_{1\cdot}>\frac{\mathrm{tr}(\mathbf{\Lambda})}{2}\right)\right.\\
&\left.+P\left(\epsilon_{1\cdot}'\epsilon_{1\cdot}<\frac{\mathrm{tr}(\mathbf{\Lambda})}{2}\right)\right]\\
\le& N^2\max_{1\le i\le N}\left[P\left(\epsilon_{1\cdot}'\mathbf{B}_i\epsilon_{1\cdot}>\frac{\epsilon\sqrt{\mathrm{tr}\left(\mathbf{\Lambda}^2\right)}}{28\sqrt{\log N}}\right)+P\left(Z_{1\cdot}'\mathbf{\Lambda}Z_{1\cdot}<\frac{\mathrm{tr}(\mathbf{\Lambda})}{2}\right)\right]\\
\le& N^2\max_{1\le i\le N}\left[P\left(\epsilon_{1\cdot}'\mathbf{B}_i\epsilon_{1\cdot}>\frac{\epsilon\sqrt{\mathrm{tr}\left(\mathbf{\Lambda}^2\right)}}{28\sqrt{\log N}}\right)+2\exp\left(-\eta_0T\right)\right].
\end{aligned}$$

Here, by Theorem 1.1 in Rudelson and Vershynin (2013), there exists a constant η_0, $C>0$ such that

$$P\left(Z_{1\cdot}'\mathbf{\Lambda}Z_{1\cdot}<\frac{\mathrm{tr}(\mathbf{\Lambda})}{2}\right)\le C\exp\left(-\eta_0T\right).$$

Note that $\mathrm{tr}(\mathbf{R}'\mathbf{B}_i\mathbf{R}) = \mathrm{tr}(\mathbf{B}_i\boldsymbol{\Sigma}) \le p\lambda_{\max}(\boldsymbol{\Sigma})$ and

$$\mathrm{tr}(\mathbf{R}'\mathbf{B}_i\boldsymbol{\Sigma}\mathbf{B}_i\mathbf{R}) \le \mathrm{tr}(\mathbf{B}_i\boldsymbol{\Sigma}^2) \le p\lambda_{\max}(\boldsymbol{\Sigma})^2.$$

Again, by Theorem 1.1 in Rudelson and Vershynin (2013) and Assumption 3, for some constant $C > 0$, we have

$$\begin{aligned}
&P\left(\epsilon_{1\cdot}'\mathbf{B}_i\epsilon_{1\cdot} > \frac{\epsilon\sqrt{\mathrm{tr}\,(\boldsymbol{\Lambda}^2)}}{28\sqrt{\log N}}\right) \\
\le& P\left(\epsilon_{1\cdot}'\mathbf{B}_i\epsilon_{1\cdot} > \mathrm{tr}(\mathbf{R}'\mathbf{B}_i\mathbf{R}) + \frac{\epsilon\sqrt{\mathrm{tr}\,(\boldsymbol{\Lambda}^2)}}{56\sqrt{\log N}}\right) \\
\le& 2\exp\left\{-C\sqrt{T/\log N}\right\},
\end{aligned}$$

where the first inequality holds because that for sufficiently large N, T, for all $i = 1, \cdots, N$, $\mathrm{tr}(\mathbf{R}'\mathbf{B}_i\mathbf{R}) \le p\lambda_{\max}(\boldsymbol{\Sigma}) \le \epsilon\sqrt{\mathrm{tr}\,(\boldsymbol{\Sigma}^2)}/\{56\sqrt{\log N}\}$. As $p > 0$ is fixed, we have

$$\begin{aligned}
&P\left(\max_{1\le i<j\le N}|\hat{\rho}_{ij} - \rho_{ij}| > \frac{\epsilon\sqrt{\mathrm{tr}\,(\boldsymbol{\Lambda}^2)}}{\sqrt{\log N}\mathrm{tr}\,(\boldsymbol{\Lambda})}\right) \\
\le& 2N^2\exp\left(-C\sqrt{T/\log N}\right) + 2N^2\exp\left(-\eta_0 T\right) \to 0.
\end{aligned}$$

By the slutsky lemma again, we can obtain that

$$\begin{aligned}
&P\left(\frac{\mathrm{tr}^2(\boldsymbol{\Lambda})}{\|\boldsymbol{\Lambda}\|_{\mathrm{F}}^2}\max_{1\le i<j\le N}\hat{\rho}_{ij}^2 - 4\log N + \log\log N \le y\right) \\
&\to \exp\left\{-\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{y}{2}\right)\right\}.
\end{aligned} \tag{S3.15}$$

Because $\mathrm{tr}^2(\boldsymbol{\Lambda}) = \mathrm{tr}^2(\boldsymbol{\Sigma})$ and $\|\boldsymbol{\Lambda}\|_{\mathrm{F}}^2 = \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2$, then, we complete the proof.

□

S3.2 Proof of Theorem 2

Lemma S.6. *For any $\varepsilon \in (0,1)$ and sufficiently large T,*

$$P\left(|\hat{\sigma}_{ij} - \sigma_{ij}| \geq x\sqrt{\frac{\sigma_{ij}^2 + (1+|\gamma_2|)\sigma_{ii}\sigma_{jj}}{N}} \right) \leq C \exp\left(-\frac{x^2}{2}(1-\varepsilon) \right),$$

uniformly for $x \in (0, N^{\frac{1}{2}\wedge(\frac{1}{\tau}-\frac{1}{2})})$, where C does not depend on i, j.

Lemma S.7. *Under the same assumptions as in Theorem 2. Let*

$$\hat{\boldsymbol{\Gamma}} = (\frac{\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot}}{T})_{1\leq i,j\leq N}, \hat{\gamma}_N = \|\hat{\boldsymbol{\Gamma}}\|_{\mathrm{F}}^2 - \frac{1}{T}\{\mathrm{tr}(\hat{\boldsymbol{\Gamma}})\}^2,$$

and $\gamma_N = N\left(\frac{\mathrm{tr}(\boldsymbol{\Sigma})}{T}\right)^2$. We have

$$\frac{\hat{\gamma}_N}{N} = \frac{\gamma_N}{N} a_N + b_N, \tag{S3.16}$$

where $\{a_N\}$ are real numbers satisfying $1 - c_1/T \leq a_N \leq 1 + c_2 N/T$ for some constant $c_1, c_2 > 0$, $\{b_N\}$ are random variables satisfying

$$E\left(b_N^2\right) = O\left(\frac{1}{NT}\right).$$

With preparations earlier, we are now ready to prove Theorem 2. According to (S3.15), we have

$$\mathrm{tr}^2(\boldsymbol{\Sigma})/\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2 \max_{1\leq i<j\leq N} \hat{\rho}_{ij}^2 = O_p(\log N).$$

Notice that

$$\left| \frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} \max_{1\leq i<j\leq N} \hat{\rho}_{ij}^2 - \frac{\mathrm{tr}^2(\boldsymbol{\Sigma})}{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2} \max_{1\leq i<j\leq N} \hat{\rho}_{ij}^2 \right|$$

$$\leq \frac{\operatorname{tr}^2(\boldsymbol{\Sigma})}{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2} \max_{1\leq i<j\leq N} \hat{\rho}_{ij}^2 \left| \frac{\operatorname{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} \frac{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2}{\operatorname{tr}^2(\boldsymbol{\Sigma})} - 1 \right|.$$

To obtain the limit null distribution of L_N, we only need to show that

$$\frac{\operatorname{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} \frac{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2}{\operatorname{tr}^2(\boldsymbol{\Sigma})} - 1 = o_p\{\log^{-1} N\}.$$

So, similar to the proof of Theorem 2.1 in Chen and Liu (2018), we can obtain that

$$\frac{\operatorname{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} \frac{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2}{\operatorname{tr}^2(\boldsymbol{\Sigma})} = 1 + O_p\left\{ \left(\sqrt{\frac{\log T}{N}} \right)^{\min(1,2-\tau)} \right\} = 1 + o_p\{\log^{-1} N\}, \tag{S3.17}$$

due to Lemmas S.6 and S.7. Then, we complete the proof and obtain that

$$P\left(\frac{\operatorname{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N - 4\log N + \log\log N \leq y \right) \to \exp\left\{ -\frac{1}{\sqrt{8\pi}} \exp\left(-\frac{y}{2} \right) \right\}.$$

□

S3.3 Proof of Theorem 3

We first establish two similar lemmas.

Lemma S.8. *Under the same assumptions as in Theorem 3. For any $\varepsilon \in (0,1)$ and sufficiently large T,*

$$P\left(|\hat{\sigma}_{ij} - \sigma_{ij}| \geq x \sqrt{\frac{\operatorname{tr}(\boldsymbol{\Phi}^2)(\sigma_{ii}\sigma_{jj} + \sigma_{ij}^2)}{N^2}} \right) \leq C \exp\left(-C' x^2 (1-\varepsilon) \right),$$

uniformly for $x \in (0, N^{\frac{1}{2} \wedge (\frac{1}{\tau} - \frac{1}{2})})$, where C and C' do not depend on i, j. Here, for any $1 \leq i, j \leq T$, $\sigma_{ij} = (\boldsymbol{\Sigma})_{ij}$.

Lemma S.9. *Under the same assumptions as in Theorem 3. Let*

$$\hat{\boldsymbol{\Gamma}} = (\frac{\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot}}{T})_{1\le i,j\le N}, \hat{\gamma}_N = \|\hat{\boldsymbol{\Gamma}}\|_{\mathrm{F}}^2 - \frac{1}{T}\{\mathrm{tr}(\hat{\boldsymbol{\Gamma}})\}^2,$$

and $\gamma_N = \left(\frac{\mathrm{tr}(\boldsymbol{\Sigma})}{T}\right)^2 \|\boldsymbol{\Phi}\|_{\mathrm{F}}^2$. *We have*

$$\frac{\hat{\gamma}_N}{N} = \frac{\gamma_N}{N} a_N + b_N, \tag{S3.18}$$

where $\{a_N\}$ *are real numbers satisfying* $1 - c_1/T \le a_N \le 1 + c_2 N/T$ *for some constant* $c_1, c_2 > 0, \{b_N\}$ *are random variables satisfying*

$$E\left(b_N^2\right) = O\left(\frac{1}{T^{\frac{1}{2}\vee\left(\frac{1}{\tau}-\frac{1}{2}\right)}}\right).$$

We are now ready to prove Theorem 3. Because for any $1 \le i \le j \le N$,

$$\epsilon_{i\cdot}'\epsilon_{j\cdot} = Z'\left[\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\} \otimes \boldsymbol{\Lambda}\right] Z$$

and $\epsilon_{j\cdot}'\mathbf{B}_i\epsilon_{j\cdot} = Z'\left[\{\phi_{j\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\} \otimes (\mathbf{R}'\mathbf{B}_i\mathbf{R})\right] Z$, where $\phi_{l\cdot}^{1/2} \in \mathbb{R}^N$ denotes the l-th row vector of matrix $\mathbf{L}$ and

$$Z = (Z_{11}, \ldots, Z_{1T}, Z_{21}, \ldots, Z_{2T}, \ldots, Z_{N1}, \ldots, Z_{NT}).$$

Note that $\mathrm{tr}(\boldsymbol{\Lambda}) = \mathrm{tr}(\boldsymbol{\Sigma})$ and $\mathrm{tr}(\boldsymbol{\Lambda}^2) = \mathrm{tr}(\boldsymbol{\Sigma}^2)$. Similar to (S3.14), we can use the similar approach to prove that

$$P\left(\max_{1\le i<j\le N} |\hat{\rho}_{ij} - \rho_{ij}| > \frac{\epsilon\sqrt{\mathrm{tr}\left(\boldsymbol{\Sigma}^2\right)}}{\sqrt{\log N}\mathrm{tr}\left(\boldsymbol{\Sigma}\right)}\right) \to 0,$$

and

$$P\left(\max_{1\le i<j\le N} \left|\frac{\epsilon_{i\cdot}'\epsilon_{i\cdot}\epsilon_{j\cdot}'\epsilon_{j\cdot}}{T^2} - \frac{\mathrm{tr}^2(\boldsymbol{\Sigma})}{T^2}\right| > C_3\sqrt{\frac{\log N}{T}}\right) = O\left(N^{-M}\right).$$

In addition, similar to the proof of Theorem 2.1 in Chen and Liu (2018), we can also prove that

$$\frac{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2}{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}\frac{\mathrm{tr}^2(\boldsymbol{\Sigma})}{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2} = 1 + O_p\left\{\left(\sqrt{\frac{\log T}{N}}\right)^{\min(1,2-\tau)}\right\}, \tag{S3.19}$$

due to Lemmas S.8 and S.9. Note that

$$\mathrm{Var}\left(\epsilon_{i\cdot}'\epsilon_{j\cdot}\right) = \|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2\left(\phi_{ii}\phi_{jj}+\phi_{ij}^2\right) + \gamma_2\mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\circ\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\right]\mathrm{tr}(\boldsymbol{\Lambda}\circ\boldsymbol{\Lambda}),$$

where ϕ_{ij} denotes the (i,j)-th element of matrix $\boldsymbol{\Phi}$. Note that we assume that $\phi_{ii} = (\boldsymbol{\Phi})_{ii} = 1$, for all $1 \le i \le N$, so, we have

$$\mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\circ\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\right]\mathrm{tr}(\boldsymbol{\Lambda}\circ\boldsymbol{\Lambda}) \le \mathrm{tr}(\boldsymbol{\Sigma}^2).$$

Without loss of generality, we can assume that $\phi_{12} \ge \delta\sqrt{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2/\mathrm{tr}^2(\boldsymbol{\Sigma})\log N}$ for some constant $\delta > 2$. Note that for some constant $C > 0$, by Theorem 1.1 in Rudelson and Vershynin (2013) and Assumption 3, we have

$$P\left(\sqrt{\frac{\log N}{T}}\frac{\epsilon_{1\cdot}'\epsilon_{2\cdot} - \phi_{12}\mathrm{tr}(\boldsymbol{\Sigma})}{\sqrt{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2}} > \varepsilon\right) \le 2\exp\left\{-C\varepsilon\log N\right\} \to 0,$$

for all $\varepsilon > 0$. Then, there exist a constant $M > 0$ satisfying

$$\begin{aligned}
&P\left(\sqrt{\frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2}}\max_{1\le i<j\le N}\hat{\rho}_{ij} > \sqrt{4\log N - \log\log N + w_\alpha}\right)\\
\ge &P\left(\sqrt{\frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2}}\hat{\rho}_{12} > \sqrt{4\log N - \log\log N + w_\alpha}\right)\\
\ge &P\left(\sqrt{\frac{\mathrm{tr}^2(\boldsymbol{\Sigma})}{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2}}\hat{\rho}_{12} > \sqrt{4\log N - \log\log N + w_\alpha}\left(1+\frac{M\sqrt{\log N}}{\sqrt{N}}\right)\right) + o(1)
\end{aligned}$$

$$
\begin{aligned}
\geq & P\left(\sqrt{\frac{\mathrm{tr}^2(\mathbf{\Sigma})}{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2}}\rho_{12} > \sqrt{4\log N - \log\log N + w_\alpha} + o(1)\right) + o(1) \\
\geq & P\left(\sqrt{\frac{\mathrm{tr}^2(\mathbf{\Sigma})}{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2}}\frac{\epsilon_{1\cdot}'\epsilon_{2\cdot} - \mathrm{tr}(\mathbf{\Sigma})\phi_{12}}{\sqrt{\epsilon_{1\cdot}'\epsilon_{1\cdot}\epsilon_{2\cdot}'\epsilon_{2\cdot}}} > -(\delta-2)\sqrt{\log N}/2\right) + o(1), \\
\geq & P\left(\frac{\epsilon_{1\cdot}'\epsilon_{2\cdot} - \mathrm{tr}(\mathbf{\Sigma})\phi_{12}}{\sqrt{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2}} > -(\delta-2)\sqrt{\log N}/2\right) + o(1) \\
= & 1 - P\left(\frac{\epsilon_{1\cdot}'\epsilon_{2\cdot} - \mathrm{tr}(\mathbf{\Sigma})\phi_{12}}{\sqrt{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2}} \leq -(\delta-2)\sqrt{\log N}/2\right) + o(1) \\
\geq & 1 - P\left(\left|\frac{\epsilon_{1\cdot}'\epsilon_{2\cdot} - \mathrm{tr}(\mathbf{\Sigma})\phi_{12}}{\sqrt{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2}}\right| \geq (\delta-2)\sqrt{\log N}/2\right) + o(1) \\
= & 1 - O(N^{-\varepsilon_1}) + o(1) \to 1,
\end{aligned}
$$

where $\varepsilon_1 = C\{(\delta-2)/2\}^2 > 0$, for some constant $C > 0$. Consequently, we complete the proof of this theorem. □

S3.4 Proof of Theorem 4

Define

$$
\begin{aligned}
&\tilde{T}_{\max} = \max_{1\leq i<j\leq N} \frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2} \\
&\Lambda_N = \{(i,j); 1 \leq i < j \leq N\} \\
&A_N = \{S_N/\sigma_{S_N} \leq x\} \quad \text{and} \quad B_I = \{|\epsilon_{i\cdot}'\epsilon_{j\cdot}| \geq l_N\} \qquad \text{(S3.20)}
\end{aligned}
$$

for any $I = (i,j) \in \Lambda_N$, where $a_N = 4\log N - \log\log N + y$ and

$$
l_N = \sqrt{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2\,[4\log N - \log\log N + y]} = \sqrt{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2\, a_N}.
$$

To make a clear presentation, we impose a trivial ordering for elements in Λ_N. For any $I_1 = (i_1, j_1) \in \Lambda_N$ and $I_2 = (i_2, j_2) \in \Lambda_N$, we say $I_1 < I_2$ if $i_1 < i_2$ or $i_1 = i_2$ but $j_1 < j_2$.

Recall that under certain assumptions, Baltagi et al. (2016) established the asymptotic property of S_N that under the null hypothesis, $S_N/\hat{\sigma}_{S_N} \to \mathcal{N}(0,1)$ in distribution when $\min(N,T) \to \infty$. Since the assumptions in this paper are slightly different from those in Baltagi et al. (2016), we reconsider the asymptotic properties of S_N under our Assumptions 1-3. Similar to Theorems 2-3 in Baltagi et al. (2016), Lemma S.10 presents the asymptotic null distribution of S_N, and Lemma S.11 presents that $\hat{\sigma}_{S_N}$ is a ratio-consistent estimator of the variance of S_N.

Lemma S.10. *Under Assumptions 1-3 and H_0, we have $S_N/\sigma_{S_N} \to \mathcal{N}(0,1)$ in distribution, as $\min(N,T) \to \infty$ with $\lim_{\min(N,T)\to\infty} N/T = \gamma \in (0, +\infty)$. Here, $\mathbf{M}_i = \mathbf{P}_i \boldsymbol{\Sigma} \mathbf{P}_i$, for any $1 \le i \le N$, and*

$$\sigma_{S_N}^2 = \frac{2}{N(N-1)} \sum_{1\le i<j\le N} \operatorname{tr}\left(\mathbf{M}_i \mathbf{M}_j\right) / \{\operatorname{tr}\left(\mathbf{M}_i\right) \operatorname{tr}\left(\mathbf{M}_j\right)\}.$$

Recall that

$$\hat{\sigma}_{S_N}^2 = \frac{2}{N(N-1)} \sum_{1\le i<j\le N} v_j' \left(v_i - \bar{v}_{ij}\right) v_i' \left(v_j - \bar{v}_{ij}\right), \tag{S3.21}$$

$\bar{v}_{ij} = \sum_{1<k\neq i,j<N} v_k/(N-2)$ and $v_k = \hat{\epsilon}_{k\cdot}/\|\hat{\epsilon}_{k\cdot}\|$ for all $1 \le k \le N$.

Lemma S.11. *Under Assumptions 1-3 and* H_0*, with* $\lim_{\min(N,T)\to\infty} N/T = \gamma \in (0, +\infty)$*, we have* $\hat{\sigma}^2_{S_N}/\sigma^2_{S_N} \to 1$ *in probability, as* $\min(N,T) \to \infty$*.*

Next, the following lemmas are provided for establishing the asymptotic independence between the two test statistics S_N and L_N.

Lemma S.12. *Assume that Assumptions 1-3 hold, under* H_0*, if* S_N/σ_{S_N} *and* $\tilde{T}_{\max} - 4\log N + \log\log N$ *are asymptotically independent, then* $S_N/\hat{\sigma}_{S_N}$ *and* $L_N \mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})/\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2 - 4\log N + \log\log N$ *are also asymptotically independent when* $\min(N,T) \to \infty$ *and* $N/T \to \gamma$*.*

Lemma S.13. *Under* H_0 *and same assumptions in Lemma S.12, let*

$$H(N,k) = \sum_{I_1 < I_2 < \cdots < I_k \in \Lambda_N} P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right).$$

Then $\lim_{k\to\infty} \limsup_{\min(N,T)\to\infty} H(N,k) = 0$ *when* $\min(N,T) \to \infty$ *and* $N/T \to \gamma$*.*

Lemma S.14. *Under the assumptions of Lemma S.12,*

$$\sum_{I_1 < I_2 < \cdots < I_k \in \Lambda_N} \left[P\left(A_N B_{I_1} B_{I_2} \cdots B_{I_k}\right) - P\left(A_N\right) \cdot P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right)\right] \to 0$$

as $\min(N,T) \to \infty$ *and* $N/T \to \gamma$ *for each* $k \geq 1$*.*

Next, we are ready to prove asymptotic independence stated in Theorem 4. By (S3.9) and Lemma S.10, the following hold,

$$\max_{1\le i<j\le N} \frac{\left(\epsilon_{i\cdot}'\epsilon_{j\cdot}\right)^2}{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2} - 4\log N + \log\log N \to G(y) \text{ in distribution;} \tag{S3.22}$$

$$\frac{S_N}{\sigma_{S_N}} \to N(0,1) \text{ in distribution,} \tag{S3.23}$$

where $G(y) = \exp\left\{-\exp\left(-y/2\right)/\sqrt{8\pi}\right\}$. To show asymptotic independence, according to Lemma S.12, it is enough to show the limit of

$$\lim_{\min(N,T)\to\infty} P\left(\frac{S_N}{\sigma_{S_N}} \le x, \max_{1\le i<j\le N} \frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2} \le a_N\right) = \Phi(x)\cdot G(y), \tag{S3.24}$$

for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$, where $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-t^2/2}dt$ and

$$a_N = 4\log N - \log\log N + y,$$

which makes sense for large N. Because of (S3.22) and (S3.23), the above is equivalent to that

$$\lim_{\min(N,T)\to\infty} P\left(\frac{S_N}{\sigma_{S_N}} \le x, \max_{1\le i<j\le N} \frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2} > a_N\right) = \Phi(x)\cdot\{1-G(y)\}, \tag{S3.25}$$

for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$. Review notation Λ_N, A_N and B_I for any $I = (i,j) \in \Lambda_N$ in (S3.20). Write

$$P\left(\frac{S_N}{\sigma_{S_N}} \le x, \max_{1\le i<j\le N} \frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\boldsymbol{\Sigma}\|_{\mathrm{F}}^2} > a_N\right) = P\left(\bigcup_{I\in\Lambda_N} A_N B_I\right). \tag{S3.26}$$

Here the notation $A_N B_I$ stands for $A_N \cap B_I$. From the inclusion-exclusion principle,

$$P\left(\bigcup_{I\in\Lambda_N} A_N B_I\right) \le \sum_{I_1\in\Lambda_N} P\left(A_N B_{I_1}\right) - \sum_{I_1<I_2\in\Lambda_N} P\left(A_N B_{I_1} B_{I_2}\right) + \cdots +$$

$$\sum_{I_1<I_2<\cdots<I_{2l+1}\in\Lambda_N} P\left(A_N B_{I_1} B_{I_2}\cdots B_{I_{2l+1}}\right) \qquad \text{(S3.27)}$$

and

$$P\left(\bigcup_{I\in\Lambda_N} A_N B_I\right) \geq \sum_{I_1\in\Lambda_N} P\left(A_N B_{I_1}\right) - \sum_{I_1<I_2\in\Lambda_N} P\left(A_N B_{I_1} B_{I_2}\right) + \cdots - \qquad \text{(S3.28)}$$

$$\sum_{I_1<I_2<\cdots<I_{2l}\in\Lambda_N} P\left(A_N B_{I_1} B_{I_2}\cdots B_{I_{2l}}\right)$$

for any integer $l \geq 1$. Reviewing the definition

$$H(N,k) = \sum_{I_1<I_2<\cdots<I_k\in\Lambda_N} P\left(B_{I_1} B_{I_2}\cdots B_{I_k}\right),$$

for $k \geq 1$ in Lemma $S.13$, we have from the lemma that

$$\lim_{k\to\infty} \limsup_{\min(N,T)\to\infty} H(N,k) = 0. \qquad \text{(S3.29)}$$

Set

$$\zeta(N,k) = \sum_{I_1<I_2<\cdots<I_k\in\Lambda_N} \left\{P\left(A_N B_{I_1} B_{I_2}\cdots B_{I_k}\right) - P\left(A_N\right)\cdot P\left(B_{I_1} B_{I_2}\cdots B_{I_k}\right)\right\},$$

for $T \geq 1$. By Lemma S.14,

$$\lim_{\min(N,T)\to\infty} \zeta(N,k) = 0, \qquad \text{(S3.30)}$$

for each $k \geq 1$. The assertion (S3.27) implies that

$$P\left(\bigcup_{I\in\Lambda_N} A_N B_I\right)$$

$$\leq P(A_N)\left\{\sum_{I_1\in\Lambda_N} P(B_{I_1}) - \sum_{I_1<I_2\in\Lambda_N} P(B_{I_1}B_{I_2}) + \cdots - \right.$$

$$\left.\sum_{I_1<I_2<\cdots<I_{2l}\in\Lambda_N} P(B_{I_1}B_{I_2}\cdots B_{I_{2l}})\right\} + \left\{\sum_{k=1}^{2l}\zeta(N,k)\right\} + H(N,2l+1)$$

$$\leq P(A_N)\cdot P\left(\bigcup_{I\in\Lambda_N} B_I\right) + \left\{\sum_{k=1}^{2l}\zeta(N,k)\right\} + H(N,2l+1), \tag{S3.31}$$

where the inclusion-exclusion formula is used again in the last inequality, that is

$$P\left(\bigcup\nolimits_{I\in\Lambda_N} B_I\right) \geq \left\{\sum\nolimits_{I_1\in\Lambda_N} P(B_{I_1}) - \sum\nolimits_{I_1<I_2\in\Lambda_N} P(B_{I_1}B_{I_2}) + \cdots - \sum\nolimits_{I_1<I_2<\cdots<I_{2l}\in\Lambda_N} P(B_{I_1}B_{I_2}\cdots B_{I_{2l}})\right\},$$

for all $l \geq 1$. By the definition of a_N and (S3.22),

$$\begin{aligned} P\left(\bigcup_{I\in\Lambda_N} B_I\right) &= P\left(\max_{1\leq i<j\leq N}\frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2} > a_N\right) \\ &= P\left\{\max_{1\leq i<j\leq N}\frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2} - 4\log N + \log\log N > y\right\} \\ &\to 1 - G(y), \end{aligned}$$

as $\min(N,T)\to\infty$. By (S3.23), $P(A_N)\to\Phi(x)$ as $\min(N,T)\to\infty$. From (S3.26), by fixing l first and sending $\min(N,T)\to\infty$, we get from (S3.30) that

$$\begin{aligned} &\limsup_{\min(N,T)\to\infty} P\left(\frac{S_N}{\sigma_{S_N}} \leq x, \max_{1\leq i<j\leq N}\frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2} > a_N\right) \\ \leq &\Phi(x)\cdot\{1 - G(y)\} + \limsup_{\min(N,T)\to\infty} H(N,2l+1). \end{aligned}$$

Now, let $l \to \infty$ and use (S3.29) to see

$$\limsup_{\min(N,T)\to\infty} P\left(\frac{S_N}{\sigma_{S_N}} \le x, \max_{1\le i<j\le N} \frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2} > a_N\right) \le \Phi(x)\cdot\{1-G(y)\}\,. \tag{S3.32}$$

By applying the same argument to (S3.28), we see that the counterpart of (S3.31) becomes

$$\begin{aligned} P\left(\bigcup_{I\in\Lambda_N} A_N B_I\right) \ge & P(A_N)\left\{\sum_{I_1\in\Lambda_N} P(B_{I_1}) - \sum_{I_1<I_2\in\Lambda_N} P(B_{I_1}B_{I_2}) + \cdots + \right. \\ & \left. \sum_{I_1<I_2<\cdots<I_{2l-1}\in\Lambda_N} P\left(B_{I_1}B_{I_2}\cdots B_{I_{2l-1}}\right)\right\} + \\ & \left\{\sum_{k=1}^{2l-1}\zeta(N,k)\right\} - H(N,2l) \\ \ge & P(A_N)\cdot P\left(\bigcup_{I\in\Lambda_N} B_I\right) + \left\{\sum_{k=1}^{2l-1}\zeta(N,k)\right\} - H(N,2l), \end{aligned}$$

where in the last step we use the inclusion-exclusion principle such that

$$\begin{aligned} P\left(\bigcup_{I\in\Lambda_N} B_I\right) \le & \left\{\sum_{I_1\in\Lambda_N} P(B_{I_1}) - \sum_{I_1<I_2\in\Lambda_N} P(B_{I_1}B_{I_2}) + \cdots + \right. \\ & \left. \sum_{I_1<I_2<\cdots<I_{2l-1}\in\Lambda_N} P\left(B_{I_1}B_{I_2}\cdots B_{I_{2l-1}}\right)\right\}, \end{aligned}$$

for all $l \ge 1$. Review (S3.26) and repeat the earlier procedure to see

$$\limsup_{\min(N,T)\to\infty} P\left(\frac{S_N}{\sigma_{S_N}} \le x, \max_{1\le i<j\le N} \frac{(\epsilon_{i\cdot}'\epsilon_{j\cdot})^2}{\|\mathbf{\Sigma}\|_{\mathrm{F}}^2} > a_N\right) \ge \Phi(x)\cdot\{1-G(y)\}\,,$$

by sending $\min(N,T) \to \infty$ and then sending $l \to \infty$. This and (S3.32) yield (S3.25). The proof is completed due to Lemma S.12. □

S4 Proofs of the lemmas

In this section, we will prove the lemmas used in the previous section.

S4.1 Proof of Lemma S.6

Recall that $\mathbf{M}_i = \mathbf{P}_i \boldsymbol{\Sigma} \mathbf{P}_i$, $\mathbf{B}_i = \mathbf{x}_i \left(\mathbf{x}_i' \mathbf{x}_i\right)^{-1} \mathbf{x}_i'$. $(\mathbf{A})_{ij}$ is the element at row i and column j of matrix $\mathbf{A}$, for any square matrix $\mathbf{A}$. Hence, we define the element at row i and column j of matrix $\mathbf{M}_l = \mathbf{P}_l \boldsymbol{\Sigma} \mathbf{P}_l$ as $\sigma_{ij,l}$, $1 \le i \le j \le N$. Let $\omega_{ij,l} = \mathrm{tr}\{(\mathbf{R}'\mathbf{P}_l e_i e_j' \mathbf{P}_l \mathbf{R}) \circ (\mathbf{R}'\mathbf{P}_l e_i e_j' \mathbf{P}_l \mathbf{R})\} \le \sigma_{ii,l}\sigma_{jj,l}$, for $1 \le i < j \le T; 1 \le l \le N$.

Then, we claim that

$$\left|\left(\mathbf{B}_l \boldsymbol{\Sigma}\right)_{ij}\right|,\ \left|\left(\mathbf{B}_l \boldsymbol{\Sigma} \mathbf{B}_s\right)_{ij}\right|,\ \left|\left(\boldsymbol{\Sigma} \mathbf{B}_l\right)_{ij}\right| = O(T^{-1 \vee -\frac{1}{\tau}}) \tag{S4.33}$$

for any $1 \le l, s \le N$, $1 \le i, j \le T$. To simplify the expression, we let $b_{ij,l} \doteq (\mathbf{B}_l)_{ij}$, for any $1 \le l \le N$, $1 \le i, j \le T$. Under Assumption 2, we have $\max_{1 \le j \le N} |b_{ij,l}| = O(T^{-1})$.

$$\begin{aligned} \left|\left(\mathbf{B}_l \boldsymbol{\Sigma}\right)_{ij}\right| = \left|\sum_{k=1}^{T} b_{ik,l}\sigma_{kj}\right| &\le \sum_{k=1}^{T} |b_{ik,l}||\sigma_{kj}| \\ &\le O(T^{-1}) \sum_{k=1}^{T} |\sigma_{kj}|. \end{aligned}$$

Due to Assumption 3, we have $\sum_{k=1}^{T} |\sigma_{jk}|^{\tau} \le C$ for some $0 < \tau < 2$ and $1 \le j \le T$. So, if $0 < \tau < 1$, we have $\sum_{k=1}^{T} |\sigma_{kj}| \le \sum_{k=1}^{T} |\sigma_{kj}|^{\tau} |\sigma_{kj}|^{1-\tau} =$

$O(1)$. If $1 < \tau < 2$, from a convex inequality we have

$$\left\{ \Big(\sum_{k=1}^{T} |\sigma_{kj}| \Big)^{\tau} \right\}^{1/\tau} \leq \left\{ T^{\tau-1} \sum_{k=1}^{T} |\sigma_{kj}|^{\tau} \right\}^{1/\tau} \leq O(T^{1-\frac{1}{\tau}}).$$

So, $\left| \Big(\mathbf{B}_l \boldsymbol{\Sigma} \Big)_{ij} \right| = O(T^{-1 \vee -\frac{1}{\tau}})$. Similarly, we have $\left| \Big(\boldsymbol{\Sigma} \mathbf{B}_l \Big)_{ij} \right| = O(T^{-1 \vee -\frac{1}{\tau}})$ and $\left| \Big(\mathbf{B}_l \boldsymbol{\Sigma} \mathbf{B}_s \Big)_{ij} \right| = O(T^{-1 \vee -\frac{1}{\tau}})$. Then, due to $\mathbf{P}_l \boldsymbol{\Sigma} \mathbf{P}_l = \boldsymbol{\Sigma} - \mathbf{B}_l \boldsymbol{\Sigma} - \boldsymbol{\Sigma} \mathbf{B}_l + \mathbf{B}_l \boldsymbol{\Sigma} \mathbf{B}_l$, we can conclude that

$$\left| \sum_{l=1}^{N} \frac{\sigma_{ij,l} - \sigma_{ij}}{N} \right| = O(T^{-1 \vee -\frac{1}{\tau}})$$

and

$$\left| \sum_{l=1}^{N} \frac{\sigma_{ii,l}\sigma_{jj,l} + \sigma_{ij,l}^2 - \sigma_{ii}\sigma_{jj} - \sigma_{ij}^2}{N} \right| = O(T^{-1 \vee -\frac{1}{\tau}}).$$

Recall that $\bar{\hat{\epsilon}}_{\cdot j} = \sum_{l=1}^{N} \hat{\epsilon}_{lj}/N$, $1 \leq j \leq N$, we have $\hat{\sigma}_{ij} = \frac{1}{N-1} \sum_{l=1}^{N} \hat{\epsilon}_{li}\hat{\epsilon}_{lj} - \frac{N}{N-1} \bar{\hat{\epsilon}}_{\cdot i} \bar{\hat{\epsilon}}_{\cdot j}$ Since $\mathrm{Cov}\left(\hat{\epsilon}_{li}, \hat{\epsilon}_{lj}\right) = \sigma_{ij,l}$, we obtain that $\mathrm{Var}\left(\hat{\epsilon}_{li}\hat{\epsilon}_{lj}\right) = \sigma_{ij,l}^2 + \sigma_{ii,l}\sigma_{jj,l} + \gamma_2 \omega_{ij,l}$. By classical Cramér type large deviation results for independent random variables (see Corollary 3.1 in Saulis and Statulevicius (1991)), we have for any $\varepsilon > 0$,

$$P\Big(\Big| \frac{\sum_{l=1}^{N} \left(\hat{\epsilon}_{li}\hat{\epsilon}_{lj} - \sigma_{ij,l} \right)}{\sqrt{\sum_{l=1}^{N} \left(\sigma_{ij,l}^2 + (1 + |\gamma_2|)\sigma_{ii,l}\sigma_{jj,l} \right)}} \Big| \geq x \Big) \tag{S4.34}$$

$$\leq P\Big(\Big| \frac{\sum_{l=1}^{N} \left(\hat{\epsilon}_{li}\hat{\epsilon}_{lj} - \sigma_{ij,l} \right)}{\sqrt{\sum_{l=1}^{N} \left(\sigma_{ii,l}\sigma_{jj,l} + \sigma_{ij,l}^2 + \gamma_2 \omega_{ij,l} \right)}} \Big| \geq x \Big) \tag{S4.35}$$

$$\leq C \exp\Big\{ - \frac{x^2}{2}(1 - \varepsilon) \Big\} \tag{S4.36}$$

uniformly in $x \in \big[0, o(\sqrt{N})\big)$. For $\bar{\hat{\epsilon}}_{\cdot j}$, we have $\mathrm{Var}\left(\bar{\hat{\epsilon}}_{\cdot j}\right) = \sum_{l=1}^{N} \sigma_{jj,l}/N^2$. By (S4.33), $\mathrm{Var}\left(\bar{\hat{\epsilon}}_{\cdot j}\right) \leq CN^{-1}$, uniformly in $1 \leq j \leq T$. Again, by classical

Cramér type large deviation results for independent random variables, we have for any $\varepsilon > 0$,

$$P\left(\left|\bar{\hat{\epsilon}}_{\cdot j}\right| \geq x\sqrt{\operatorname{Var}\left(\bar{\hat{\epsilon}}_{\cdot j}\right)}\right) \leq C\exp\Big\{-\frac{x^2}{2}(1-\varepsilon)\Big\}$$

uniformly in $x \in \big[0, o(\sqrt{N})\big)$. So

$$P\left(\left|\bar{\hat{\epsilon}}_{\cdot i}\bar{\hat{\epsilon}}_{\cdot j}\right| \geq x^2\sqrt{\operatorname{Var}\left(\bar{\hat{\epsilon}}_{\cdot i}\right)\operatorname{Var}\left(\bar{\hat{\epsilon}}_{\cdot j}\right)}\right) \leq 2C\exp\Big\{-\frac{x^2}{2}(1-\varepsilon)\Big\}$$

uniformly in $x \in \big[0, o(\sqrt{N})\big)$. We have, uniformly for $x \in \big[0, o(\sqrt{N})\big)$, $x^2\sqrt{\operatorname{Var}\left(\bar{\hat{\epsilon}}_{\cdot i}\right)\operatorname{Var}\left(\bar{\hat{\epsilon}}_{\cdot i}\right)} = o(x/\sqrt{N})$. So for any $\delta > 0$ and large N

$$P\left(\left|\bar{\hat{\epsilon}}_{\cdot i}\bar{\hat{\epsilon}}_{\cdot j}\right| \geq \delta\frac{x}{\sqrt{N}}\right) \leq 2C\exp\Big\{-\frac{x^2}{2}(1-\varepsilon)\Big\} \tag{S4.37}$$

uniformly for $x \in \big[0, o(\sqrt{N})\big)$. Hence, the lemma follows from (S4.34) and (S4.37). □

S4.2 Proof of Lemma S.7

So, according to the definitions in Lemma S.7, we have

$$\|\hat{\mathbf{\Gamma}}\|_{\mathrm{F}}^2 = \frac{1}{T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2,$$

$$\{\operatorname{tr}(\hat{\mathbf{\Gamma}})\}^2 = \frac{1}{T^2}\left(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2,$$

$$T^2 a_N = \frac{1}{\gamma_N}\left[\sum_{i=1}^{N}\sum_{j=1}^{N}E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2 - \frac{1}{T}E\left(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\right],$$

$$T^2 b_N = \frac{1}{N}\Big[\sum_{i=1}\sum_{j=1}(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2 - \frac{1}{T}\left(\sum_{i=1}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2 - \sum_{i=1}\sum_{j=1}E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2$$

$$+\frac{1}{T}E\left(\sum_{i=1}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\Big].$$

It is easy to verify that a_N and b_N will make the equation (S3.16) true. In the following, we will prove that a_N, b_N satisfy the properties in the lemma.

We first deal with the term a_N. Recall that $\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot}$ and $\mathbf{P}_i\boldsymbol{\Sigma}\mathbf{P}_i = \mathbf{M}_i$, we have

$$E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2 = \begin{cases} \mathrm{tr}(\mathbf{M}_i\mathbf{M}_j), & i \neq j, \\ 2\mathrm{tr}(\mathbf{M}_i^2) + \mathrm{tr}^2(\mathbf{M}_i) & \\ \quad + \gamma_2\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{R}'\mathbf{P}_i\mathbf{R})\}, & i = j. \end{cases}$$

Therefore,

$$\sum_{i=1}^N\sum_{j=1}^N E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2 = \sum_{i\neq j}^N\sum^N \mathrm{tr}(\mathbf{M}_i\mathbf{M}_j) + \sum_{i=1}^N 2\mathrm{tr}(\mathbf{M}_i^2) + \mathrm{tr}^2(\mathbf{M}_i)$$
$$+ \gamma_2\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{R}'\mathbf{P}_i\mathbf{R})\}.$$

Moreover, by Lemma S.4,

$$\begin{aligned} E\left(\sum_{i=1}^N \hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2 &= \sum_{i=1}^N\sum_{j=1}^N E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}) \\ &= \sum_{i=1}^N E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}) + \sum_{i\neq j}^N\sum^N E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}) \\ &= \sum_{i=1}^N 2\mathrm{tr}(\mathbf{M}_i^2) + \mathrm{tr}^2(\mathbf{M}_i) + \sum_{i\neq j}^N\sum^N \mathrm{tr}(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j) \\ &\quad + \gamma_2\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{R}'\mathbf{P}_i\mathbf{R})\}. \end{aligned}$$

So we have

$$
\begin{aligned}
& T^2\gamma_N a_N \\
=& \left[\sum_{i=1}^N\sum_{j=1}^N E(\hat{\epsilon}_{i.}'\hat{\epsilon}_{j.})^2 - \frac{1}{T}E\left(\sum_{i=1}^N \hat{\epsilon}_{i.}'\hat{\epsilon}_{i.}\right)^2\right] \\
=& \sum_{i\neq j}^N\sum^N \mathrm{tr}(\mathbf{M}_i\mathbf{M}_j) + \sum_{i=1}^N \Big[2\mathrm{tr}(\mathbf{M}_i^2) + \mathrm{tr}^2(\mathbf{M}_i) \\
& + \gamma_2 \mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\mathbf{R})\}\Big] \\
& - \frac{1}{T}\Big[\sum_{i=1}^N 2\mathrm{tr}(\mathbf{M}_i^2) + \mathrm{tr}^2(\mathbf{M}_i) + \gamma_2\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\mathbf{R})\} \\
& \qquad + \sum_{i\neq j}^N\sum^N \mathrm{tr}(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)\Big] \\
\geq& \sum_{i=1}^N \mathrm{tr}^2(\mathbf{M}_i)\left(1-\frac{1}{T}\right) + \sum_{i\neq j}^N\sum^N \left\{\mathrm{tr}(\mathbf{M}_i\mathbf{M}_j) - \frac{1}{T}\mathrm{tr}(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)\right\} \\
& - (1-\frac{1}{T})|\gamma_2|\sum_{i=1}^N \mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\mathbf{R})\} \\
\geq& \big(\frac{T-1}{T}\big)\sum_{i=1}^N\big\{\mathrm{tr}^2(\mathbf{\Sigma}) - 2\mathrm{tr}(\mathbf{\Sigma})\mathrm{tr}(\mathbf{B}_i\mathbf{\Sigma})\big\} + \sum_{i\neq j}^N\sum^N\big\{\mathrm{tr}(\mathbf{\Sigma}^2) - 2\mathrm{tr}(\mathbf{\Sigma}^2\mathbf{B}_j) \\
& - 2\mathrm{tr}(\mathbf{\Sigma}^2\mathbf{B}_i) - 2\mathrm{tr}(\mathbf{\Sigma}\mathbf{B}_i\mathbf{B}_j\mathbf{\Sigma}\mathbf{B}_j) - 2\mathrm{tr}(\mathbf{\Sigma}\mathbf{B}_j\mathbf{B}_i\mathbf{\Sigma}\mathbf{B}_i) + 2\mathrm{tr}(\mathbf{B}_i\mathbf{\Sigma}^2\mathbf{B}_j)\big\} \\
& - \frac{1}{T}\sum_{i\neq j}^N\sum^N\big\{\mathrm{tr}^2(\mathbf{\Sigma}) + \mathrm{tr}(\mathbf{B}_i\mathbf{\Sigma})\mathrm{tr}(\mathbf{B}_j\mathbf{\Sigma})\big\} - |\gamma_2|\sum_{i=1}^N\{\mathrm{tr}(\mathbf{\Sigma}^2) + \mathrm{tr}(\mathbf{B}_i\mathbf{\Sigma}\mathbf{B}_i\mathbf{\Sigma})\} \\
\geq& N\mathrm{tr}^2(\mathbf{\Sigma}) - 2pN\mathrm{tr}(\mathbf{\Sigma})C - \frac{N\mathrm{tr}^2(\mathbf{\Sigma})}{T} + \sum_{i\neq j}^N\sum^N\left\{\mathrm{tr}(\mathbf{\Sigma}^2) - \frac{\mathrm{tr}^2(\mathbf{\Sigma})}{T}\right\} \\
& - 10pC^2N(N-1) - \frac{p^2C^2N(N-1)}{T} - |\gamma_2|C^2NT - |\gamma_2|NpC^2 \\
\geq& T^2\gamma_N\left(1-\frac{c_1}{T}\right),
\end{aligned}
$$

for some constant $c_1 > 0$. Moreover, we have

$$
\begin{aligned}
&T^2\gamma_N a_N \\
=&\left[\sum_{i=1}^N \sum_{j=1}^N E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2 - \frac{1}{T}E\left(\sum_{i=1}^N \hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\right] \\
=&\sum_{i\neq j}^N \sum^N \mathrm{tr}(\mathbf{M}_i\mathbf{M}_j) + \sum_{i=1}^N \left[2\mathrm{tr}(\mathbf{M}_i^2) + \mathrm{tr}^2(\mathbf{M}_i) + \gamma_2\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\mathbf{R})\}\right] \\
&- \frac{1}{T}\Big[\sum_{i=1}^N 2\mathrm{tr}(\mathbf{M}_i^2) + \mathrm{tr}^2(\mathbf{M}_i) + \gamma_2\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\mathbf{R})\} \\
&\qquad + \sum_{i\neq j}^N \sum^N \mathrm{tr}(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)\Big] \\
\leq&\sum_{i\neq j}^N \sum^N \big\{\mathrm{tr}(\mathbf{\Sigma}^2) + 2\mathrm{tr}(\mathbf{\Sigma}\mathbf{B}_i\mathbf{\Sigma}\mathbf{B}_j) + 2\mathrm{tr}(\mathbf{B}_i\mathbf{\Sigma}^2\mathbf{B}_j) + \mathrm{tr}(\mathbf{\Sigma}\mathbf{B}_j\mathbf{\Sigma}\mathbf{B}_j) \\
&+\mathrm{tr}(\mathbf{\Sigma}\mathbf{B}_i\mathbf{\Sigma}\mathbf{B}_i) + \mathrm{tr}(\mathbf{B}_i\mathbf{\Sigma}\mathbf{B}_i\mathbf{B}_j\mathbf{\Sigma}\mathbf{B}_j) - 2\mathrm{tr}(\mathbf{\Sigma}\mathbf{B}_i\mathbf{B}_j\mathbf{\Sigma}\mathbf{B}_j) \\
&-2\mathrm{tr}(\mathbf{\Sigma}\mathbf{B}_j\mathbf{B}_i\mathbf{\Sigma}\mathbf{B}_i)\big\} + \sum_{i=1}^N \big\{\mathrm{tr}^2(\mathbf{\Sigma}) + \mathrm{tr}^2(\mathbf{\Sigma}\mathbf{B}_i)\big\} \\
&+ (1-\frac{1}{T})|\gamma_2|\sum_{i=1}^N \{\mathrm{tr}(\mathbf{\Sigma}^2) + \mathrm{tr}(\mathbf{B}_i\mathbf{\Sigma}\mathbf{B}_i\mathbf{\Sigma})\} + 2C^2NT \\
\leq&C^2N(N-1)T + 11pC^2N(N-1) + 2C^2NT + N\mathrm{tr}^2(\mathbf{\Sigma}) + p^2C^2N \\
&+ |\gamma_2|C^2NT + |\gamma_2|NpC^2 \\
\leq&T^2\gamma_N\left(1+\frac{c_2N}{T}\right),
\end{aligned}
$$

for some constant $c_2 > 0$. This proves that a_N satisfies the inequality in the lemma.

It remains to calculate b_N. We have

$$\text{Var}\bigg\{\sum_{i=1}^N\sum_{j=1}^N(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{j\cdot})^2\bigg\} = E\Bigg[\bigg\{\sum_{i=1}^N\sum_{j=1}^N(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{j\cdot})^2\bigg\}^2\Bigg] - E^2\bigg\{\sum_{i=1}^N\sum_{j=1}^N(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{j\cdot})^2\bigg\}.$$

Note that $i\neq j\neq k$ means that i,j,k are not equal, and $i\neq j\neq k\neq l$ means that i,j,k,l are not equal. Let $\mathbf{\Pi}_i = \mathbf{R}'\mathbf{P}_i\mathbf{R}$ for any $1\leq i\leq N$. First, by Lemma S.4, we have

$$\begin{aligned}
&E\Bigg[\bigg\{\sum_{i=1}^N\sum_{j=1}^N(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{j\cdot})^2\bigg\}^2\Bigg]\\
&=\sum_{i=1}^N E\left\{(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{i\cdot})^4\right\} + 2\sum_{i\neq j}^N\sum^N E\left\{(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{j\cdot})^4\right\} + \sum_{i\neq j}^N\sum^N E\left\{(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{i\cdot})^2(\hat{\epsilon}'_{j\cdot}\hat{\epsilon}_{j\cdot})^2\right\}\\
&\quad + 4\sum_{i\neq j}^N\sum^N E\left\{(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{j\cdot})^2(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{i\cdot})^2\right\} + 4\sum^N\sum_{i\neq j\neq k}^N\sum^N E\left\{(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{j\cdot})^2(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{k\cdot})^2\right\}\\
&\quad + 2\sum^N\sum_{i\neq j\neq k}^N\sum^N E\left\{(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{i\cdot})^2(\hat{\epsilon}'_{j\cdot}\hat{\epsilon}_{k\cdot})^2\right\} + \sum^N\sum^N\sum_{i\neq j\neq k\neq l}^N\sum E\left\{(\hat{\epsilon}'_{i\cdot}\hat{\epsilon}_{j\cdot})^2(\hat{\epsilon}'_{k\cdot}\hat{\epsilon}_{l\cdot})^2\right\}\\
&=\sum_{i=1}^N\Big\{48\text{tr}(\mathbf{M}_i^4) + 32\text{tr}(\mathbf{M}_i^3)\text{tr}(\mathbf{M}_i) + 12\text{tr}^2(\mathbf{M}_i^2) + 12\text{tr}(\mathbf{M}_i^2)\text{tr}^2(\mathbf{M}_i)\\
&\quad + \text{tr}^4(\mathbf{M}_i) + \gamma_2 f_{\gamma_2} + \gamma_4 f_{\gamma_4} + \gamma_6 f_{\gamma_6} + \gamma_1^2 f_{\gamma_1^2} + \gamma_2^2 f_{\gamma_2^2} + \gamma_1\gamma_3 f_{\gamma_1\gamma_3}\Big\}\\
&\quad + 2\sum_{i\neq j}^N\sum^N\Big[6\text{tr}(\mathbf{M}_i\mathbf{M}_j)^2 + 3\text{tr}^2(\mathbf{M}_i\mathbf{M}_j)\\
&\quad + 3\gamma_2\text{tr}\{(\mathbf{R}'\mathbf{P}_j\mathbf{M}_i\mathbf{P}_j\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_j\mathbf{M}_i\mathbf{P}_j\mathbf{R})\}\\
&\quad + \gamma_2 E\text{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\epsilon_{j\cdot}\epsilon'_{j\cdot}\mathbf{P}_j\mathbf{P}_i\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\epsilon_{j\cdot}\epsilon'_{j\cdot}\mathbf{P}_j\mathbf{P}_i\mathbf{R})\}\Big]\\
&\quad + \sum_{i\neq j}^N\sum^N\Big(\left[2\text{tr}(\mathbf{M}_i^2) + \text{tr}^2(\mathbf{M}_i) + \gamma_2\text{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\right]\\
&\qquad\qquad \times\left[2\text{tr}(\mathbf{M}_j^2) + \text{tr}^2(\mathbf{M}_j) + \gamma_2\text{tr}\{(\mathbf{\Pi}_j)\circ(\mathbf{\Pi}_j)\}\right]\Big)
\end{aligned}$$

$$+ 4\sum_{i \neq j}^{N}\sum^{N}\Big[\gamma_4 \mathrm{tr}\left\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i) \circ (\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R})\right\}$$

$$+ 2\gamma_2 \mathrm{tr}\left(\mathbf{\Pi}_i\right) \mathrm{tr}\left\{(\mathbf{\Pi}_i) \circ (\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R})\right\}$$

$$+ \gamma_2 \mathrm{tr}\left(\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R}\right) \mathrm{tr}\left\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\right\}$$

$$+ 8\gamma_2 \mathrm{tr}\left[(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R})\right] + 4\gamma_2 \mathrm{tr}\left[(\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R}) \circ (\mathbf{\Pi}_i\mathbf{\Pi}_i)\right]$$

$$+ 4\gamma_1^2\left[\tau_T'\left\{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\right\}\mathbf{\Pi}_i\left\{\mathbf{I}_T \circ (\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R})\right\}\tau_T\right]$$

$$+ 2\gamma_1^2\left[\tau_T'\left\{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\right\}\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R}\left\{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\right\}\tau_T\right]$$

$$+ 4\gamma_1^2\left[\tau_T'\left\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i) \circ (\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R})\right\}\tau_T\right]$$

$$+ \mathrm{tr}\left(\mathbf{\Pi}_i\right) \mathrm{tr}\left(\mathbf{\Pi}_i\right) \mathrm{tr}\left(\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R}\right) + 4\mathrm{tr}\left(\mathbf{\Pi}_i\right) \mathrm{tr}\left(\mathbf{\Pi}_i\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R}\right)$$

$$+ 2\mathrm{tr}\left(\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R}\right) \mathrm{tr}\left(\mathbf{\Pi}_i\mathbf{\Pi}_i\right) + 8\mathrm{tr}\left(\mathbf{\Pi}_i\mathbf{\Pi}_i\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R}\right)\Big]$$

$$+ 4\sum_{i \neq j \neq k}^{N}\sum^{N}\sum^{N}\Big[2\mathrm{tr}(\mathbf{M}_j\mathbf{M}_i\mathbf{M}_k\mathbf{M}_i) + \mathrm{tr}(\mathbf{M}_j\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_k\mathbf{M}_i)$$

$$+ \gamma_2 \mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\mathbf{R}) \circ (\mathbf{R}'\mathbf{P}_i\mathbf{M}_k\mathbf{P}_i\mathbf{R})\} + \mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}(\mathbf{M}_j\mathbf{M}_k)\Big]$$

$$+ 2\sum_{i \neq j \neq k}^{N}\sum^{N}\sum^{N}\left[\mathrm{tr}^2(\mathbf{M}_i) + \gamma_2 \mathrm{tr}\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\}\right]\mathrm{tr}(\mathbf{M}_j\mathbf{M}_k)$$

$$+ \sum_{i \neq j \neq k \neq l}^{N}\sum^{N}\sum^{N}\sum^{N}\left\{\mathrm{tr}(\mathbf{M}_i\mathbf{M}_j)\mathrm{tr}(\mathbf{M}_k\mathbf{M}_l)\right\}.$$

Here, because the diagonal elements of $(\mathbf{\Pi}_i)_{ii}$ are greater than 0 and less than C, we can obtain that

$$f_{\gamma_2} = 6\mathrm{tr}\left(\mathbf{\Pi}_i\right) \mathrm{tr}\left(\mathbf{\Pi}_i\right) \mathrm{tr}\left\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\right\}$$

$$
\begin{aligned}
&+ 12\tau_T' \{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\} \tau_T \mathrm{tr} \{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\} \\
&+ 48\mathrm{tr}(\mathbf{\Pi}_i) \mathrm{tr} \{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i\mathbf{\Pi}_i)\} + 96\mathrm{tr}[\{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\} \mathbf{\Pi}_i\mathbf{\Pi}_i\mathbf{\Pi}_i] \\
&+ 48\tau_T' \{\mathbf{I}_T \circ (\mathbf{\Pi}_i\mathbf{\Pi}_i)\} \{\mathbf{I}_T \circ (\mathbf{\Pi}_i\mathbf{\Pi}_i)\} \tau_T \\
\leq & 6\mathrm{tr}(\mathbf{\Pi}_i) \mathrm{tr}(\mathbf{\Pi}_i) \mathrm{tr}\left\{(\mathbf{\Pi}_i)^2\right\} + 12\mathrm{tr}^2\left\{(\mathbf{\Pi}_i)^2\right\} \\
&+ 48\mathrm{tr}(\mathbf{\Pi}_i) \sqrt{\mathrm{tr}\left\{(\mathbf{\Pi}_i)^2\right\} \mathrm{tr}\left\{(\mathbf{\Pi}_i)^4\right\}} \\
&+ 96C\mathrm{tr}\left\{(\mathbf{\Pi}_i)^3\right\} + 48\mathrm{tr}^2(\mathbf{\Pi}_i\mathbf{\Pi}_i) \\
= & O(T^3), \\
f_{\gamma_4} = & 4\mathrm{tr}(\mathbf{\Pi}_i) \mathrm{tr}\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\} \\
&+ 24\mathrm{tr}\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i\mathbf{\Pi}_i)\} = O(T^2), \\
f_{\gamma_6} = & \mathrm{tr}\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\} = O(T), \\
f_{\gamma_1^2} = & 24\tau_T' \{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\} \mathbf{\Pi}_i \{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\} \tau_T \mathrm{tr}(\mathbf{\Pi}_i) \\
&+ 48\tau_T' \{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\} \mathbf{\Pi}_i\mathbf{\Pi}_i \{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\} \tau_T \\
&+ 16\tau_T' \{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\} \tau_T \mathrm{tr}(\mathbf{\Pi}_i) \\
&+ 96\tau_T' \{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\} \mathbf{\Pi}_i \{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\} \tau_T \\
&+ 96\mathrm{tr}[\mathbf{\Pi}_i \{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\} \mathbf{\Pi}_i] \\
\leq & 24C^2 \sum_{s=1}^{T} \sum_{t=1}^{T} |(\mathbf{\Pi}_i)_{st}| \mathrm{tr}(\mathbf{\Pi}_i) + 48C^2 \sum_{s=1}^{T} \sum_{t=1}^{T} |(\mathbf{\Pi}_i\mathbf{\Pi}_i)|_{st} \\
&+ 16 \sum_{s=1}^{T} \sum_{t=1}^{T} |(\mathbf{\Pi}_i\mathbf{\Pi}_i)|_{st}^3 \mathrm{tr}(\mathbf{\Pi}_i)
\end{aligned}
$$

$$+ 96\sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{k=1}^{T}(\mathbf{\Pi}_i)_{sk}^2(\mathbf{\Pi}_i)_{kt}(\mathbf{\Pi}_i)_{tt}$$

$$+ 96\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{\Pi}_i)_{st}^2(\mathbf{\Pi}_i\mathbf{\Pi}_i)_{ts} = O(T^3),$$

$$f_{\gamma_2^2} = 24\tau_T'\{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\{\mathbf{I}_T\circ(\mathbf{\Pi}_i)\}\tau_T$$

$$3\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}$$

$$+ 8\tau_T'\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\tau_T$$

$$\le 3\mathrm{tr}^2(\mathbf{\Pi}_i\mathbf{\Pi}_i) + 24\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{\Pi}_i)_{st}^2(\mathbf{\Pi}_i)_{ss}(\mathbf{\Pi}_i)_{tt} + 8\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{\Pi}_i)_{st}^4$$

$$= O(T^2),$$

$$f_{\gamma_1\gamma_3} = 24\tau_T'\{\mathbf{I}_T\circ(\mathbf{\Pi}_i)\}\mathbf{\Pi}_i\{\mathbf{I}_T\circ(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\tau_T$$

$$+ 32\tau_T'\{\mathbf{I}_T\circ(\mathbf{\Pi}_i)\}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\tau_T$$

$$\le 24\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{\Pi}_i)_{st}(\mathbf{\Pi}_i)_{ss}(\mathbf{\Pi}_i)_{tt}^2 + 32\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{\Pi}_i)_{st}^3(\mathbf{\Pi}_i)_{ss}$$

$$= O(T^2).$$

Then, by Lemma S.4,

$$E^2\Big\{\sum_{i=1}^{N}\sum_{j=1}^{N}(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\Big\}$$

$$= \left[\sum_{1\le i\ne j\le N}\mathrm{tr}(\mathbf{M}_i\mathbf{M}_j) + \sum_{i=1}^{N}2\mathrm{tr}(\mathbf{M}_i^2) + \mathrm{tr}^2(\mathbf{M}_i) + \gamma_2\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\right]^2$$

$$= \sum_{i=1}^{N}\Big[4\mathrm{tr}^2(\mathbf{M}_i^2) + \mathrm{tr}^4(\mathbf{M}_i) + 4\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}^2(\mathbf{M}_i) + \gamma_2^2\mathrm{tr}^2\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}$$

$$+ 4\gamma_2 \mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\} + 2\gamma_2 \mathrm{tr}^2(\mathbf{M}_i)\mathrm{tr}\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\}\Big]$$

$$+\sum_{i\neq j}^{N}\sum^{N}\Big[4\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}(\mathbf{M}_j^2) + 2\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}^2(\mathbf{M}_j) + 2\mathrm{tr}(\mathbf{M}_j^2)\mathrm{tr}^2(\mathbf{M}_i)$$

$$+ \mathrm{tr}^2(\mathbf{M}_i)\mathrm{tr}^2(\mathbf{M}_j) + \gamma_2^2 \mathrm{tr}\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\}\mathrm{tr}\{(\mathbf{\Pi}_j) \circ (\mathbf{\Pi}_j)\}$$

$$+ 2\gamma_2 \mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}\{(\mathbf{\Pi}_j) \circ (\mathbf{\Pi}_j)\}$$

$$+ 2\gamma_2 \mathrm{tr}\{(\mathbf{M}_j^2)\mathrm{tr}(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\}$$

$$+ \gamma_2 \mathrm{tr}^2\{(\mathbf{M}_i)\mathrm{tr}(\mathbf{\Pi}_j) \circ (\mathbf{\Pi}_j)\}$$

$$+ \gamma_2 \mathrm{tr}^2(\mathbf{M}_j)\mathrm{tr}\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\}\Big]$$

$$+ \sum_{i\neq j}^{N}\sum^{N}\Big[2\mathrm{tr}^2(\mathbf{M}_i\mathbf{M}_j) + 8\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}(\mathbf{M}_i\mathbf{M}_j) + 4\mathrm{tr}(\mathbf{M}_i\mathbf{M}_j)\mathrm{tr}^2(\mathbf{M}_i)$$

$$+ 4\gamma_2 \mathrm{tr}(\mathbf{M}_i\mathbf{M}_j)\mathrm{tr}\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\}\Big]$$

$$+\sum_{i\neq j\neq k}^{N}\sum^{N}\sum^{N}\Big[4\mathrm{tr}(\mathbf{M}_i\mathbf{M}_j)\mathrm{tr}(\mathbf{M}_i\mathbf{M}_k) + 4\mathrm{tr}(\mathbf{M}_i\mathbf{M}_j)\mathrm{tr}(\mathbf{M}_k^2)$$

$$+ 2\mathrm{tr}(\mathbf{M}_i\mathbf{M}_j)\mathrm{tr}^2(\mathbf{M}_k) + 2\gamma_2 \mathrm{tr}(\mathbf{M}_i\mathbf{M}_j)\mathrm{tr}\{\mathbf{\Pi}_k \circ \mathbf{\Pi}_k\}\Big]$$

$$+ \sum_{i\neq j\neq k\neq l}^{N}\sum^{N}\sum^{N}\sum^{N}\{\mathrm{tr}(\mathbf{M}_i\mathbf{M}_j)\mathrm{tr}(\mathbf{M}_k\mathbf{M}_l)\}.$$

So, we have

$$\mathrm{Var}\left\{\sum_{i=1}^{N}\sum_{j=1}^{N}(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\right\} = O\left(NT^3 + N^2T^2 + N^2T + N^3T\right).$$

Similarly, we can obtain that

$$E\Big(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\Big)^4$$

$$
\begin{aligned}
=&E\Big(\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{k\cdot}'\hat{\epsilon}_{k\cdot}\hat{\epsilon}_{l\cdot}'\hat{\epsilon}_{l\cdot}\Big)\\
=&\sum_{i=1}^{N}E\left(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^4+3\sum_{i\neq j}^{N}\sum^{N}E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^2(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot})^2+4\sum_{i\neq j}^{N}\sum^{N}E\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^3\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}\}\\
&+6\sum^{N}\sum_{i\neq j\neq k}^{N}\sum^{N}E\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^2\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{k\cdot}'\hat{\epsilon}_{k\cdot}\}+\sum^{N}\sum_{i\neq j\neq k\neq l}^{N}\sum^{N}\sum^{N}E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{k\cdot}'\hat{\epsilon}_{k\cdot}\hat{\epsilon}_{l\cdot}'\hat{\epsilon}_{l\cdot})\\
=&\sum_{i=1}^{N}\Big\{48\text{tr}(\mathbf{M}_i^4)+32\text{tr}(\mathbf{M}_i^3)\text{tr}(\mathbf{M}_i)+12\text{tr}^2(\mathbf{M}_i^2)+12\text{tr}(\mathbf{M}_i^2)\text{tr}^2(\mathbf{M}_i)\\
&+\text{tr}^4(\mathbf{M}_i)+\gamma_2 f_{\gamma_2}+\gamma_4 f_{\gamma_4}+\gamma_6 f_{\gamma_6}+\gamma_1^2 f_{\gamma_1^2}+\gamma_2^2 f_{\gamma_2^2}+\gamma_1\gamma_3 f_{\gamma_1\gamma_3}\Big\}\\
&+3\sum_{i\neq j}^{N}\sum^{N}\Big(\left[2\text{tr}(\mathbf{M}_i^2)+\text{tr}^2(\mathbf{M}_i)+\gamma_2\text{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\right]\\
&\qquad\times\left[2\text{tr}(\mathbf{M}_j^2)+\text{tr}^2(\mathbf{M}_j)+\gamma_2\text{tr}\{(\mathbf{\Pi}_j)\circ(\mathbf{\Pi}_j)\}\right]\Big)\\
&+4\sum_{i\neq j}^{N}\sum^{N}\Big(\gamma_4\text{tr}\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\right\}\\
&+3\gamma_2\text{tr}\left(\mathbf{\Pi}_i\right)\text{tr}\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\right\}\\
&+12\gamma_2\text{tr}\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i\mathbf{\Pi}_i)\right\}\\
&+6\gamma_1^2\left[\tau_T'\left\{\mathbf{I}_T\circ(\mathbf{\Pi}_i)\right\}\mathbf{\Pi}_i\left\{(\mathbf{I}_T)\circ(\mathbf{\Pi}_i)\right\}\tau_T\right]\\
&+4\gamma_1^2\left[\tau_T'\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\right\}\tau_T\right]\\
&+\text{tr}^3\left(\mathbf{\Pi}_i\right)+6\text{tr}\left(\mathbf{\Pi}_i\right)\text{tr}\left(\mathbf{\Pi}_i\mathbf{\Pi}_i\right)\\
&+8\text{tr}\left(\mathbf{\Pi}_i\mathbf{\Pi}_i\mathbf{\Pi}_i\right)\Big)\text{tr}(\mathbf{M}_j)\\
&+6\sum^{N}\sum_{i\neq j\neq k}^{N}\sum^{N}[2\text{tr}(\mathbf{M}_i^2)+\text{tr}^2(\mathbf{M}_i)+\gamma_2\text{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}]\text{tr}(\mathbf{M}_j)\text{tr}(\mathbf{M}_k)
\end{aligned}
$$

$$+\sum^{N}\sum^{N}_{i\neq j\neq k\neq l}\sum^{N}\sum^{N}\mathrm{tr}(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)\mathrm{tr}(\mathbf{M}_k)\mathrm{tr}(\mathbf{M}_l),$$

where f_{γ_2}, f_{γ_4}, f_{γ_6}, $f_{\gamma_1^2}$, $f_{\gamma_2^2}$, $f_{\gamma_1\gamma_3}$ are the same as above, and

$$\begin{aligned}
&\Big\{E\Big(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\Big)^2\Big\}^2\\
=&\Big[\sum_{i=1}^{N}2\mathrm{tr}(\mathbf{M}_i^2)+\mathrm{tr}^2(\mathbf{M}_i)+\gamma_2\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}+\sum^{N}_{i\neq j}\sum^{N}\mathrm{tr}(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)\Big]^2\\
=&\sum_{i=1}^{N}\Big[4\mathrm{tr}^2(\mathbf{M}_i^2)+\mathrm{tr}^4(\mathbf{M}_i)+4\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}^2(\mathbf{M}_i)+\gamma_2^2\mathrm{tr}^2\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\\
&+4\gamma_2\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}+2\gamma_2\mathrm{tr}^2(\mathbf{M}_i)\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\Big]\\
+&\sum^{N}_{i\neq j}\sum^{N}\Big[4\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}(\mathbf{M}_j^2)+2\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}^2(\mathbf{M}_j)+2\mathrm{tr}(\mathbf{M}_j^2)\mathrm{tr}^2(\mathbf{M}_i)\\
&+\mathrm{tr}^2(\mathbf{M}_i)\mathrm{tr}^2(\mathbf{M}_j)+\gamma_2^2\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\mathrm{tr}\{(\mathbf{\Pi}_j)\circ(\mathbf{\Pi}_j)\}\\
&+2\gamma_2\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}\{(\mathbf{\Pi}_j)\circ(\mathbf{\Pi}_j)\}\\
&+2\gamma_2\mathrm{tr}(\mathbf{M}_j^2)\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\\
&+\gamma_2\mathrm{tr}^2\{(\mathbf{M}_i)\mathrm{tr}(\mathbf{\Pi}_j)\circ(\mathbf{\Pi}_j)\}\\
&+\gamma_2\mathrm{tr}^2(\mathbf{M}_j)\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\Big]\\
&+\sum^{N}_{i\neq j}\sum^{N}\Big[2\mathrm{tr}^2(\mathbf{M}_i)\mathrm{tr}^2(\mathbf{M}_j)+8\mathrm{tr}(\mathbf{M}_i^2)\mathrm{tr}(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)+4\mathrm{tr}^3(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)\\
&+4\gamma_2\mathrm{tr}(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)\mathrm{tr}\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\}\Big]\\
+&\sum^{N}\sum^{N}_{i\neq j\neq k}\sum^{N}\Big[4\mathrm{tr}^2(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)\mathrm{tr}(\mathbf{M}_k)+4\mathrm{tr}(\mathbf{M}_i)\mathrm{tr}(\mathbf{M}_j)\mathrm{tr}(\mathbf{M}_k^2)
\end{aligned}$$

$$+ 2\text{tr}(\mathbf{M}_i)\text{tr}(\mathbf{M}_j)\text{tr}^2(\mathbf{M}_k) + 2\gamma_2\text{tr}(\mathbf{M}_i)\text{tr}(\mathbf{M}_j)\text{tr}\{\mathbf{\Pi}_k \circ \mathbf{\Pi}_k\}\Big]$$
$$+ \sum_{i \neq j \neq k \neq l}^{N}\sum^{N}\sum^{N}\sum^{N} \{\text{tr}(\mathbf{M}_i)\text{tr}(\mathbf{M}_j)\text{tr}(\mathbf{M}_k)\text{tr}(\mathbf{M}_l)\}.$$

So, we have

$$\text{Var}\left\{\Big(\sum_{i=1}^{N} \hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\Big)^2\right\} = O(NT^3 + N^2T^3 + N^3T^3).$$

Finally, we have

$$\begin{aligned} E(b_N^2) \leq & \frac{1}{N^2T^4}\left[2\text{Var}\left\{\sum_{i=1}^{N}\sum_{j=1}^{N}(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\right\} + \frac{2}{T^2}\text{Var}\left\{\Big(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\Big)^2\right\}\right] \\ = & O\left(\frac{1}{NT}\right). \end{aligned}$$

□

S4.3 Proof of Lemma S.8

Recall that $\mathbf{E} = \mathbf{LZR}'$, then we have

$$\hat{\epsilon}_{li}\hat{\epsilon}_{lj} = Z'\left[\{\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\} \otimes \{m_i^l(m_j^l)'\}\right] Z,$$

where $Z = (Z_{11}, \cdots, Z_{1T}, Z_{21}, \cdots, Z_{2T}, \cdots, Z_{N1}, \cdots, Z_{NT})' \in \mathbb{R}^{NT}$, $\phi_{l\cdot}^{1/2} \in \mathbb{R}^N$ denotes the l-th row vector of matrix $\mathbf{L}$, and $m_i^l \in \mathbb{R}^T$ represents the i-th row vector of matrix $\mathbf{P}_l\mathbf{R}$. Hence,

$$\sum_{l=1}^{N}\hat{\epsilon}_{li}\hat{\epsilon}_{lj} = Z'\left[\sum_{l=1}^{N}\{\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\} \otimes \{m_i^l(m_j^l)'\}\right] Z.$$

To simplify notation, we write $\mathbf{W}_{ij}$ as matrix $\sum_{l=1}^{N}\{\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\}\otimes\{m_i^l(m_j^l)'\}$. Recall that $||\mathbf{A}||$ denotes the operator norm of matrix $\mathbf{A}$. So, by the Hanson-Wright inequality in Rudelson and Vershynin (2013), we can obtain that for every $t>0$,

$$P\left\{\left|Z'\mathbf{W}_{ij}Z-E\big(Z'\mathbf{W}_{ij}Z\big)\right|>t\right\}\le 2\exp\left[-c\min\left(\frac{t^2}{K^4\|\mathbf{W}_{ij}\|_{\mathrm{F}}^2},\frac{t}{K^2\|\mathbf{W}_{ij}\|}\right)\right].$$

Obviously, $E\big(Z'\mathbf{W}_{ij}Z\big)=\mathrm{tr}(\mathbf{W}_{ij})=\sum_{l=1}^{N}\phi_{ll}\sigma_{ij,l}$, where $\phi_{ll}=(\mathbf{\Phi})_{ll}$, for $1\le l\le N$. Recall that we assume that $\phi_{ll}=1$, for $1\le l\le N$. According to the properties of the Kronecker product, we have

$$\mathrm{tr}\{\mathbf{W}_{ij}(\mathbf{W}_{ij})'\}=\sum_{l,s=1}^{N}\sum^{N}\phi_{ls}^2(\mathbf{P}_l\mathbf{\Sigma}\mathbf{P}_s)_{ii}(\mathbf{P}_l\mathbf{\Sigma}\mathbf{P}_s)_{jj}.$$

Due to $(\mathbf{P}_l\mathbf{\Sigma}\mathbf{P}_s)_{jj}=\sigma_{jj}-(\mathbf{B}_l\mathbf{\Sigma})_{jj}-(\mathbf{B}_s\mathbf{\Sigma})_{jj}-(\mathbf{B}_l\mathbf{\Sigma}\mathbf{B}_s)_{jj}$ and (S4.33), we have $\mathrm{tr}\{\mathbf{W}_{ij}(\mathbf{W}_{ij})'\}=\mathrm{tr}(\mathbf{\Phi}^2)\sigma_{ii}\sigma_{jj}\{1+o(1)\}$ and $\mathrm{tr}\{\mathbf{W}_{ij}\mathbf{W}_{ij}\}=\mathrm{tr}(\mathbf{\Phi}^2)\sigma_{ij}^2\{1+o(1)\}$. Then, we focus on $\|\mathbf{W}_{ij}\|$. According to the triangle inequality of norms, we have

$$\begin{aligned}&\|\mathbf{W}_{ij}\|\\ \le&\|\sum_{l=1}^{N}\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\otimes\nu_{i\cdot}(\nu_{j\cdot})'\|+\|\sum_{l=1}^{N}\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\otimes\left\{m_i^l(m_j^l)'-\nu_{i\cdot}(\nu_{j\cdot})'\right\}\|\\ \le&\sqrt{\lambda_{\max}(\mathbf{\Phi}^2)\sigma_{ii}\sigma_{jj}}+\|\sum_{l=1}^{N}\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\otimes\left\{m_i^l(m_j^l)'-\nu_{i\cdot}(\nu_{j\cdot})'\right\}\|,\end{aligned}$$

where $\nu_{i\cdot}\in\mathbb{R}^N$ denotes the i-th row vector of matrix $\mathbf{R}$. Then, using the definition of spectral norm, we can obtain that for any vector $x\in\mathbb{R}^{NT}$ with norm $\|x\|=1$,

$$
\begin{aligned}
&\|\sum_{l=1}^{N}\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\otimes\big\{m_i^l(m_j^l)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\|^2\\
=&\max_{\|x\|=1}x'\Big[\sum_{s=1}^{N}\sum_{l=1}^{N}\phi_{sl}\phi_{s\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\otimes\big\{m_j^s(m_i^s)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^l(m_j^l)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]x\\
\le&\max_{\|x\|=1}x'\Big[\sum_{l=1}^{N}\phi_{ll}\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\otimes\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^l(m_j^l)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]x\\
&+\max_{\|x\|=1}x'\Big[\sum_{l\neq s}^{N}\sum^{N}\phi_{ls}\phi_{l\cdot}^{1/2}(\phi_{s\cdot}^{1/2})'\otimes\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^s(m_j^s)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]x\\
=&\Delta_1+\Delta_2.
\end{aligned}
$$

Next, we first calculate Δ_1,

$$
\begin{aligned}
\Delta_1=&\max_{\|x\|=1}x'\Big[\sum_{l=1}^{N}\phi_{ll}\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\otimes\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^l(m_j^l)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]x\\
\le&\mathrm{tr}\Big[\sum_{l=1}^{N}\phi_{ll}\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\otimes\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^l(m_j^l)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]\\
\le&\sum_{l=1}^{N}\mathrm{tr}\Big[\phi_{ll}\phi_{l\cdot}^{1/2}(\phi_{l\cdot}^{1/2})'\otimes\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^l(m_j^l)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]\\
\le&\sum_{l=1}^{N}\phi_{ll}^2\mathrm{tr}\Big[\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^l(m_j^l)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]\\
\le&\sum_{l=1}^{N}\phi_{ll}^2\Big\{\sigma_{ii,l}\sigma_{jj,l}-2(\mathbf{P}_l\boldsymbol{\Sigma})_{ii}(\mathbf{P}_l\boldsymbol{\Sigma})_{jj}+\sigma_{ii}\sigma_{jj}\Big\}\\
=&O(T^{0\vee(1-\frac{1}{\tau})})
\end{aligned}
$$

where the last inequality holds due to (S4.33). Similarly, we have

$$
\Delta_2=\max_{\|x\|=1}x'\Big[\sum_{l\neq s}^{N}\sum^{N}\phi_{ls}\phi_{l\cdot}^{1/2}(\phi_{s\cdot}^{1/2})'\otimes\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^s(m_j^s)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]x
$$

$$
\begin{aligned}
&\leq \mathrm{tr}\Big[\sum_{l\neq s}^{N}\sum^{N}\phi_{ls}\phi_{l\cdot}^{1/2}(\phi_{s\cdot}^{1/2})'\otimes\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^s(m_j^s)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]\\
&\leq \sum_{l\neq s}^{N}\sum^{N}\mathrm{tr}\Big[\phi_{ls}\phi_{l\cdot}^{1/2}(\phi_{s\cdot}^{1/2})'\otimes\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^s(m_j^s)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]\\
&\leq \sum_{l\neq s}^{N}\sum^{N}\phi_{ls}^2\mathrm{tr}\Big[\big\{m_j^l(m_i^l)'-\nu_{j\cdot}(\nu_{i\cdot})'\big\}\big\{m_i^s(m_j^s)'-\nu_{i\cdot}(\nu_{j\cdot})'\big\}\Big]\\
&\leq \sum_{l\neq s}^{N}\sum^{N}\phi_{ls}^2\Big\{(\mathbf{P}_l\boldsymbol{\Sigma}\mathbf{P}_s)_{ii}(\mathbf{P}_l\boldsymbol{\Sigma}\mathbf{P}_s)_{jj}-(\mathbf{P}_l\boldsymbol{\Sigma})_{ii}(\mathbf{P}_l\boldsymbol{\Sigma})_{jj}-(\mathbf{P}_s\boldsymbol{\Sigma})_{jj}(\mathbf{P}_s\boldsymbol{\Sigma})_{jj}\\
&\qquad+\sigma_{ii}\sigma_{jj}\Big\}\\
&=O(T^{0\vee(1-\frac{1}{\tau})}).
\end{aligned}
$$

So, we can conclude that $\|\mathbf{W}_{ij}\|=O(T^{0\vee(\frac{1}{2}-\frac{1}{2\tau})})$. Then, according to Theorem 1.1 in Rudelson and Vershynin (2013), for every $t\in(0,o(N^{\frac{1}{2}\wedge\frac{1}{2\tau}}))$, there exist a constant $c>0$,

$$
\begin{aligned}
&P\left(\left|\sum_{l=1}^{N}\hat{\epsilon}_{li}\hat{\epsilon}_{lj}-\sum_{l=1}^{N}\phi_{ll}\sigma_{ij,l}\right|>t\sqrt{\mathrm{tr}(\boldsymbol{\Phi}^2)(\sigma_{ii}\sigma_{jj}+\sigma_{ij}^2)}\right) \qquad \text{(S4.38)}\\
&\leq 2\exp\Big(-c\frac{t^2\mathrm{tr}(\boldsymbol{\Phi}^2)(\sigma_{ii}\sigma_{jj}+\sigma_{ij}^2)}{K^4\mathrm{tr}\{\mathbf{W}_{ij}(\mathbf{W}_{ij})'\}}\Big)\\
&\leq 2\exp\Big(-c'\frac{t^2}{K^4}\Big)
\end{aligned}
$$

where $\mathrm{tr}\{\mathbf{W}_{ij}(\mathbf{W}_{ij})'\}=\mathrm{tr}(\boldsymbol{\Phi}^2)\sigma_{ii}\sigma_{jj}\{1+o(1)\}$, and $c'>0$ is a positive constant independent of N and T.

In fact, $\bar{\hat{\epsilon}}_{\cdot j}=\sum_{l=1}^{N}\hat{\epsilon}_{li}/N$ can be seen as a linear combination of Z_{ks}, $1\leq k,s\leq N$, where Z_{ks} is the element at row k and column s of random matrix

$\mathbf{Z}$. Then, by classical Cramér type large deviation results for independent random variables(see Corollary 3.1 in Saulis and Statulevicius (1991)), we have for any $\varepsilon > 0$,

$$P\left(|\bar{\hat{\epsilon}}_{\cdot j}| > x\sqrt{\mathrm{Var}(\bar{\hat{\epsilon}}_{\cdot j})}\right) \leq C\exp\left\{-\frac{x^2}{2}(1-\varepsilon)\right\},$$

uniformly in $x \in [0, o(\sqrt{N}))$. For $\bar{\hat{\epsilon}}_{\cdot j}$, we have

$$\mathrm{Var}(\bar{\hat{\epsilon}}_{\cdot j}) = \frac{1}{N^2}\sum_{l=1}^{N}\phi_{ll}\sigma_{ii,l} + \frac{1}{N^2}\sum_{l\neq s}^{N}\sum^{N}\phi_{ls}(\mathbf{P}_l\mathbf{\Sigma}\mathbf{P}_s)_{ii} = O(N^{-1\vee-\frac{1}{\tau}}),$$

uniformly in $1 \leq j \leq N$. So, for any $\varepsilon > 0$,

$$P\left(\left|\bar{\hat{\epsilon}}_{\cdot j}\bar{\hat{\epsilon}}_{\cdot i}\right| \geq x^2\sqrt{\mathrm{Var}\left(\bar{\hat{\epsilon}}_{\cdot j}\right)\mathrm{Var}\left(\bar{\hat{\epsilon}}_{\cdot i}\right)}\right) \leq 2C\exp\left\{-\frac{x^2}{2}(1-\varepsilon)\right\}$$

uniformly in $x \in [0, o(\sqrt{N}))$. We have, uniformly for $x \in [0, o(N^{\frac{1}{2}\wedge\left(\frac{1}{\tau}-\frac{1}{2}\right)}))$,

$$x^2\sqrt{\mathrm{Var}\left(\bar{\hat{\epsilon}}_{\cdot i}\right)\mathrm{Var}\left(\bar{\hat{\epsilon}}_{\cdot j}\right)} = o(x/\sqrt{N}).$$

So for any $\varepsilon > 0$, large N, and any $\delta > 0$,

$$P\left(\left|\bar{\hat{\epsilon}}_{\cdot i}\bar{\hat{\epsilon}}_{\cdot j}\right| \geq \delta\frac{x}{\sqrt{N}}\right) \leq 2C\exp\left\{-\frac{x^2}{2}(1-\varepsilon)\right\}, \tag{S4.39}$$

uniformly and $x \in [0, o(N^{\frac{1}{2}\wedge\left(\frac{1}{\tau}-\frac{1}{2}\right)}))$. Then, the lemma follows from (S4.38) and (S4.39). □

S4.4 Proof of Lemma S.9

Recall that without loss of generality, we assume that $\phi_{ll} = 1$, for $1 \le l \le N$.

Similarly, we have

$$
\begin{aligned}
\|\hat{\mathbf{\Gamma}}\|_{\mathrm{F}}^2 &= \frac{1}{T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot}^2 \\
\{\mathrm{tr}(\hat{\mathbf{\Gamma}})\}^2 &= \frac{1}{T^2}\left(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2 \\
T^2 a_N &= \frac{1}{\gamma_N}\left[\sum_{i=1}^{N}\sum_{j=1}^{N}E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2 - \frac{1}{T}E\left(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\right] \\
T^2 b_N &= \frac{1}{N}\Bigg[\sum_{i=1}\sum_{j=1}(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2 - \frac{1}{T}\left(\sum_{i=1}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2 - \sum_{i=1}\sum_{j=1}E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2 \\
&\quad + \frac{1}{T}E\left(\sum_{i=1}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\Bigg]
\end{aligned}
$$

it is easy to verify that a_N and b_N will make the equation (S3.18) true.

In the following, we prove that a_N, b_N satisfy the properties in the lemma. We first deal with the term a_N. Recall that

$$Z = (Z_{11}, \cdots, Z_{1T}, Z_{21}, \cdots, Z_{2T}, \cdots, Z_{N1}, \cdots, Z_{NT})'$$

and $\phi_{l\cdot}^{1/2} \in \mathbb{R}^N$ denotes the l-th row vector of matrix $\mathbf{L}$, we have

$$(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot}) = Z'\left[\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\} \otimes (\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})\right]Z.$$

Therefore, due to Lemma S.4,

$$\sum_{1\le i,j\le N} E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2$$

$$= \sum_{1\le i\ne j\le N} \phi_{ij}^2 \mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})^2\} + \phi_{ij}^2 \mathrm{tr}^2(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})$$

$$+ \phi_{ii}\phi_{jj}\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})'\} + \gamma_2 \mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\} \circ \{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\right]$$

$$\times \mathrm{tr}\left\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R}) \circ (\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})\right\}$$

$$+ \sum_{i=1}^{N} 2\phi_{ii}^2 \mathrm{tr}\{(\mathbf{\Pi}_i)^2\} + \phi_{ii}^2 \mathrm{tr}^2(\mathbf{\Pi}_i)$$

$$+ \gamma_2 \mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\} \circ \{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\right] \times \mathrm{tr}\left\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\right\}.$$

Moreover,

$$\begin{aligned}
&E\left(\sum_{i=1}^{N} \hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2 \\
=&\sum_{i=1}^{N}\sum_{j=1}^{N} E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}) \\
=&\sum_{i=1}^{N} E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}) + \sum_{i\ne j}^{N}\sum^{N} E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}) \\
=&\sum_{i=1}^{N} 2\phi_{ii}^2 \mathrm{tr}\{(\mathbf{\Pi}_i)^2\} + \phi_{ii}^2\mathrm{tr}^2(\mathbf{\Pi}_i) \\
&+ \gamma_2 \mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\} \circ \{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\right] \mathrm{tr}\left\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_i)\right\} \\
&+ \sum_{i\ne j}^{N}\sum^{N} 2\phi_{ij}^2 \mathrm{tr}(\mathbf{\Pi}_i\mathbf{\Pi}_j) + \phi_{ii}\phi_{jj}\mathrm{tr}(\mathbf{\Pi}_i)\mathrm{tr}(\mathbf{\Pi}_j) \\
&+ \gamma_2 \mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\} \circ \{\phi_{j\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\right] \mathrm{tr}\left\{(\mathbf{\Pi}_i) \circ (\mathbf{\Pi}_j)\right\}.
\end{aligned}$$

Similarly, due to $\sum_{1\le i\ne j\le N} \mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\} \circ \{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\right] = O(N)$, we have

$$
\begin{aligned}
&T^2\gamma_N a_N\\
&=\left[\sum_{i=1}^{N}\sum_{j=1}^{N}E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2-\frac{1}{T}E\left(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\right]\\
&=\sum_{1\le i\neq j\le N}\phi_{ij}^2\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})^2\}+\phi_{ij}^2\mathrm{tr}^2(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})\\
&\quad+\phi_{ii}\phi_{jj}\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})'\}\\
&\quad+\gamma_2\mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\circ\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\right]\mathrm{tr}\left\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})\right\}\\
&\quad+\sum_{i=1}^{N}2\phi_{ii}^2\mathrm{tr}\{(\mathbf{\Pi}_i)^2\}+\phi_{ii}^2\mathrm{tr}^2(\mathbf{\Pi}_i)\\
&\quad+\gamma_2\mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\circ\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\right]\mathrm{tr}\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\right\}\\
&\quad-\frac{1}{T}\Big[\sum_{i=1}^{N}2\phi_{ii}^2\mathrm{tr}\{(\mathbf{\Pi}_i)^2\}+\phi_{ii}^2\mathrm{tr}^2(\mathbf{\Pi}_i)\\
&\quad+\gamma_2\mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\circ\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\right]\mathrm{tr}\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\right\}\\
&\quad+\sum_{i\neq j}^{N}\sum^{N}2\phi_{ij}^2\mathrm{tr}(\mathbf{\Pi}_i\mathbf{\Pi}_j)+\phi_{ii}\phi_{jj}\mathrm{tr}(\mathbf{\Pi}_i)\mathrm{tr}(\mathbf{\Pi}_j)\\
&\quad+\gamma_2\mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\circ\{\phi_{j\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\right]\mathrm{tr}\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_j)\right\}\Big]\\
&\ge\mathrm{tr}(\mathbf{\Phi}^2)\mathrm{tr}^2(\mathbf{\Sigma})\{1-O(T^{-1})\}.
\end{aligned}
$$

Moreover, we have

$$
T^2\gamma_N a_N=\left[\sum_{i=1}^{N}\sum_{j=1}^{N}E(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2-\frac{1}{T}E\left(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\right]
$$

$$
\begin{aligned}
&= \sum_{1\le i\neq j\le N} \phi_{ij}^2 \mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})^2\} + \phi_{ij}^2 \mathrm{tr}^2(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R}) \\
&\quad + \phi_{ii}\phi_{jj}\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})'\} \\
&\quad + \gamma_2 \mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\circ\{\phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\right]\mathrm{tr}\left\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})\right\} \\
&\quad + \sum_{i=1}^N 2\phi_{ii}^2\mathrm{tr}\{(\mathbf{\Pi}_i)^2\} + \phi_{ii}^2\mathrm{tr}^2(\mathbf{\Pi}_i) \\
&\quad + \gamma_2 \mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\circ\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\right]\mathrm{tr}\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\right\} \\
&\quad - \frac{1}{T}\Big[\sum_{i=1}^N 2\phi_{ii}^2\mathrm{tr}\{(\mathbf{\Pi}_i)^2\} + \phi_{ii}^2\mathrm{tr}^2(\mathbf{\Pi}_i) \\
&\quad + \gamma_2 \mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\circ\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\right]\mathrm{tr}\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_i)\right\} \\
&\quad + \sum_{i\neq j}^N\sum^N 2\phi_{ij}^2\mathrm{tr}(\mathbf{\Pi}_i\mathbf{\Pi}_j) + \phi_{ii}\phi_{jj}\mathrm{tr}(\mathbf{\Pi}_i)\mathrm{tr}(\mathbf{\Pi}_j) \\
&\quad + \gamma_2 \mathrm{tr}\left[\{\phi_{i\cdot}^{1/2}(\phi_{i\cdot}^{1/2})'\}\circ\{\phi_{j\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'\}\right]\mathrm{tr}\left\{(\mathbf{\Pi}_i)\circ(\mathbf{\Pi}_j)\right\}\Big] \\
&\le \mathrm{tr}(\mathbf{\Phi}^2)\mathrm{tr}^2(\mathbf{\Sigma})\{1+O(NT^{-1})\}.
\end{aligned}
$$

This proves that a_N satisfies the inequality in the lemma. It remains to calculate b_N. We have

$$
\mathrm{Var}\left\{\sum_{i=1}^N\sum_{j=1}^N(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\right\} = E\left[\left\{\sum_{i=1}^N\sum_{j=1}^N(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\right\}^2\right] - E^2\left\{\sum_{i=1}^N\sum_{j=1}^N(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\right\}.
$$

First,

$$
\mathrm{Var}\left[\left\{\sum_{i=1}^N\sum_{j=1}^N(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\right\}^2\right]
$$

$$
\begin{aligned}
=&\sum_{i=1}^{N} E\left\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^4\right\} - E^2\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^2\} + 2\sum_{i\neq j}^{N}\sum^{N} E\left\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^4\right\} - E^2\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\} \\
&+ \sum_{i\neq j}^{N}\sum^{N} E\left\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^2(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot})^2\right\} - E\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^2\}E\{(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot})^2\} \\
&+ 4\sum_{i\neq j}^{N}\sum^{N} E\left\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^2\right\} - E\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^2\}E\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\} \\
&+ 2\sum_{i\neq j\neq k}^{N}\sum^{N}\sum^{N} E\left\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^2(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{k\cdot})^2\right\} - E\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot})^2\}E\{(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{k\cdot})^2\} \\
&+ 4\sum_{i\neq j\neq k}^{N}\sum^{N}\sum^{N} E\left\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{k\cdot})^2\right\} - E\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\}E\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{k\cdot})^2\} \\
&+ \sum_{i\neq j\neq k\neq l}^{N}\sum^{N}\sum^{N}\sum^{N} E\left\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2(\hat{\epsilon}_{k\cdot}'\hat{\epsilon}_{l\cdot})^2\right\} - E\{(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\}E\{(\hat{\epsilon}_{k\cdot}'\hat{\epsilon}_{l\cdot})^2\} \\
=&G_{N1} + G_{N2} + G_{N3} + G_{N4} + G_{N5} + G_{N6} + G_{N7}.
\end{aligned}
$$

Recall that $\mathbf{Q}_{ij} = \mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R}$. Let $\boldsymbol{\Omega}_{ij} \doteq \phi_{i\cdot}^{1/2}(\phi_{j\cdot}^{1/2})'$ for $1 \leq i, j \leq N$. First, we focus on G_{N1}. According to Lemma S.4, we have

$$
G_{N1} = G_{1_1} + \gamma_2 G_{1_2} + \gamma_4 G_{1_3} + \gamma_6 G_{1_4} + \gamma_1^2 G_{1_5} + \gamma_2^2 G_{1_6} + \gamma_1\gamma_3 G_{1_7},
$$

where

$$
\begin{aligned}
G_{1_1} =& \sum_{i=1}^{N} 8\phi_{ii}^4 \mathrm{tr}\left(\mathbf{Q}_{ii}^2\right)\mathrm{tr}^2\left(\mathbf{Q}_{ii}\right) + 8\phi_{ii}^4\mathrm{tr}^2\left(\mathbf{Q}_{ii}^2\right) + 32\phi_{ii}^4\mathrm{tr}\left(\mathbf{Q}_{ii}\right)\mathrm{tr}\left(\mathbf{Q}_{ii}^3\right) \\
&+ 48\phi_{ii}^4\mathrm{tr}\left(\mathbf{Q}_{ii}^4\right), \\
G_{1_2} =& \sum_{i=1}^{N} 4\phi_{ii}^2\mathrm{tr}^2\left(\mathbf{Q}_{ii}\right)\mathrm{tr}\left\{\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\right\}\mathrm{tr}\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\right) \\
&+ 8\phi_{ii}^2\mathrm{tr}\left(\mathbf{Q}_{ii}^2\right)\mathrm{tr}\left\{\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\right\}\mathrm{tr}\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\right)
\end{aligned}
$$

$$+ 48\phi_{ii}\mathrm{tr}\left(\mathbf{Q}_{ii}\right)\mathrm{tr}\left\{\boldsymbol{\Omega}_{ii}\circ\phi_{ii}\boldsymbol{\Omega}_{ii}\right\}\mathrm{tr}\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}^{2}\right)$$

$$+ 96\mathrm{tr}\left\{\left(\mathbf{I}_{N}\circ\boldsymbol{\Omega}_{ii}\right)\phi_{ii}^{2}\boldsymbol{\Omega}_{ii}\right\}\mathrm{tr}\left\{\left(\mathbf{I}_{T}\circ\mathbf{Q}_{ii}\right)\mathbf{Q}_{ii}^{3}\right\}$$

$$+ 48\tau_{N}'\left\{\left(\mathbf{I}_{N}\circ\phi_{ii}\boldsymbol{\Omega}_{ii}\right)\left(\mathbf{I}_{N}\circ\phi_{ii}\boldsymbol{\Omega}_{ii}\right)\right\}\tau_{N}\tau_{T}'\left\{\left(\mathbf{I}_{T}\circ\mathbf{Q}_{ii}^{2}\right)\left(\mathbf{I}_{T}\circ\mathbf{Q}_{ii}^{2}\right)\right\}\tau_{T},$$

$$G_{1_3} = \sum_{i=1}^{N} 4\phi_{ii}\mathrm{tr}\left(\mathbf{Q}_{ii}\right)\mathrm{tr}\left\{\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\right\}\mathrm{tr}\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\right)$$

$$+ 24\mathrm{tr}\left\{\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\circ\phi_{ii}\boldsymbol{\Omega}_{ii}\right\}\mathrm{tr}(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}^{2}),$$

$$G_{1_4} = \sum_{i=1}^{N} \mathrm{tr}\left\{\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\right\}\mathrm{tr}\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\right),$$

$$G_{1_5} = \sum_{i=1}^{N} 24\tau_{N}'\left[\left\{\mathbf{I}_{N}\circ\boldsymbol{\Omega}_{ii}\right\}\boldsymbol{\Omega}_{ii}\left\{\mathbf{I}_{N}\circ\boldsymbol{\Omega}_{ii}\right\}\right]\tau_{N}$$

$$\times\tau_{T}'\left[\left\{\mathbf{I}_{T}\circ\mathbf{Q}_{ii}\right\}\mathbf{Q}_{ii}\left\{\mathbf{I}_{T}\circ\mathbf{Q}_{ii}\right\}\right]\tau_{T}\phi_{ii}\mathrm{tr}(\mathbf{Q}_{ii})$$

$$+ 48\tau_{N}'\left\{\left(\mathbf{I}_{N}\circ\boldsymbol{\Omega}_{ii}\right)\phi_{ii}\boldsymbol{\Omega}_{ii}\left\{\mathbf{I}_{N}\circ\boldsymbol{\Omega}_{ii}\right\}\right\}\tau_{N}$$

$$\times\tau_{T}'\left\{\left(\mathbf{I}_{N}\circ\mathbf{Q}_{ii}\right)\mathbf{Q}_{ii}^{2}\left(\mathbf{I}_{N}\circ\mathbf{Q}_{ii}\right)\right\}\tau_{T}$$

$$+ 16\tau_{N}'\left(\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\right)\tau_{N}\tau_{T}'\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\right)\tau_{T}\phi_{ii}\mathrm{tr}\left(\mathbf{Q}_{ii}\right)$$

$$+ 96\tau_{N}'\left(\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\right)\boldsymbol{\Omega}_{ii}\left(\mathbf{I}_{N}\circ\boldsymbol{\Omega}_{ii}\right)\tau_{N}\tau_{T}\prime\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\right)\mathbf{Q}_{ii}\left(\mathbf{I}_{T}\circ\mathbf{Q}_{ii}\right)\tau_{T}$$

$$+ 96\mathrm{tr}\left\{\boldsymbol{\Omega}_{ii}\left(\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\right)\boldsymbol{\Omega}_{ii}\right\}\mathrm{tr}\left\{\mathbf{Q}_{ii}\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\right)\mathbf{Q}_{ii}\right\},$$

$$G_{1_6} = \sum_{i=1}^{N} 2\mathrm{tr}^{2}\left(\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\right)\mathrm{tr}^{2}\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\right)$$

$$+ 24\tau_{N}'\left(\mathbf{I}_{N}\circ\boldsymbol{\Omega}_{ii}\right)\left(\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\right)\left(\mathbf{I}_{N}\circ\boldsymbol{\Omega}_{ii}\right)\tau_{N}$$

$$\times\tau_{T}'\left(\mathbf{I}_{T}\circ\mathbf{Q}_{ii}\right)\left(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\right)\left(\mathbf{I}_{T}\circ\mathbf{Q}_{ii}\right)\tau_{T}$$

$$+ 8\tau_N' \left(\boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right) \tau_N \tau_T' \left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \tau_T,$$

$$G_{1_7} = \sum_{i=1}^{N} 32\tau_N' \left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right) \left(\boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right) \tau_N \tau_T' \left(\mathbf{I}_T \circ \mathbf{Q}_{ii}\right) \left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \tau_T$$

$$+ 24\tau_N' \left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right) \boldsymbol{\Omega}_{ii} \left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right) \tau_N$$

$$\times \tau_T' \left(\mathbf{I}_T \circ \mathbf{Q}_{ii}\right) \mathbf{Q}_{ii} \left(\mathbf{I}_T \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \tau_T.$$

Obviously, $G_{1_1} = O(NT^3)$. Due to Lemma S.3 (1), we have

$$\begin{aligned}
G_{1_2} =& \sum_{i=1}^{N} 4\phi_{ii}^2 \mathrm{tr}^2 \left(\mathbf{Q}_{ii}\right) \mathrm{tr} \left\{\boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right\} \mathrm{tr} \left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \\
& + 8\phi_{ii}^2 \mathrm{tr} \left(\mathbf{Q}_{ii}^2\right) \mathrm{tr} \left\{\boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right\} \mathrm{tr} \left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \\
& + 48\phi_{ii} \mathrm{tr} \left(\mathbf{Q}_{ii}\right) \mathrm{tr} \left\{\boldsymbol{\Omega}_{ii} \circ \phi_{ii}\boldsymbol{\Omega}_{ii}\right\} \mathrm{tr} \left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}^2\right) \\
& + 96 \mathrm{tr} \left\{\left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right) \phi_{ii}^2 \boldsymbol{\Omega}_{ii}\right\} \mathrm{tr} \left\{\left(\mathbf{I}_T \circ \mathbf{Q}_{ii}\right) \mathbf{Q}_{ii}^3\right\} \\
& + 48\tau_N' \left\{\left(\mathbf{I}_N \circ \phi_{ii}\boldsymbol{\Omega}_{ii}\right) \left(\mathbf{I}_N \circ \phi_{ii}\boldsymbol{\Omega}_{ii}\right)\right\} \tau_N \tau_T' \left\{\left(\mathbf{I}_T \circ \mathbf{Q}_{ii}^2\right) \left(\mathbf{I}_T \circ \mathbf{Q}_{ii}^2\right)\right\} \tau_T \\
=& O(NT^3) + O(NT^2) + \sum_{i=1}^{N} 48\phi_{ii} \mathrm{tr} \left(\mathbf{Q}_{ii}\right) \mathrm{tr} \left\{\boldsymbol{\Omega}_{ii} \circ \phi_{ii}\boldsymbol{\Omega}_{ii}\right\} \mathrm{tr} \left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}^2\right) \\
& + 96 \mathrm{tr} \left\{\left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right) \phi_{ii}^2 \boldsymbol{\Omega}_{ii}\right\} \mathrm{tr} \left\{\left(\mathbf{I}_T \circ \mathbf{Q}_{ii}\right) \mathbf{Q}_{ii}^3\right\} \\
& + 48\tau_N' \left\{\left(\mathbf{I}_N \circ \phi_{ii}\boldsymbol{\Omega}_{ii}\right) \left(\mathbf{I}_N \circ \phi_{ii}\boldsymbol{\Omega}_{ii}\right)\right\} \tau_N \tau_T' \left\{\left(\mathbf{I}_T \circ \mathbf{Q}_{ii}^2\right) \left(\mathbf{I}_T \circ \mathbf{Q}_{ii}^2\right)\right\} \tau_T \\
=& O(NT^3) + O(NT^2) + \sum_{i=1}^{N} 48\phi_{ii}\phi_{ii} \mathrm{tr} \left(\mathbf{Q}_{ii}\right) \sum_{s=1}^{N} (\boldsymbol{\Omega}_{ii})_{ss}^2 \sum_{s=1}^{T} (\mathbf{Q}_{ii}^2)_{ss} (\mathbf{Q}_{ii})_{ss} \\
& + 96\phi_{ii}^2 \sum_{s=1}^{N} (\boldsymbol{\Omega}_{ii})_{ss}^2 \sum_{s=1}^{T} (\mathbf{Q}_{ii}^3)_{ss} (\mathbf{Q}_{ii})_{ss} + 48\phi_{ii}^2 \sum_{s=1}^{N} (\boldsymbol{\Omega}_{ii})_{ss}^2 \sum_{s=1}^{T} (\mathbf{Q}_{ii}^2)_{ss}^2 \\
\leq& O(NT^3) + O(NT^2) + \sum_{i=1}^{N} 48\phi_{ii}^4 |\mathrm{tr} \left(\mathbf{Q}_{ii}\right)| |\sum_{s=1}^{T} (\mathbf{Q}_{ii}^2)_{ss} (\mathbf{Q}_{ii})_{ss}|
\end{aligned}$$

$$+ 96\phi_{ii}^4|\sum_{s=1}^{T}(\mathbf{Q}_{ii}^3)_{ss}(\mathbf{Q}_{ii})_{ss}| + 48\phi_{ii}^4|\sum_{s=1}^{T}(\mathbf{Q}_{ii}^2)_{ss}^2|$$

$$=O(NT^3) + O(NT^2) + O(NT).$$

Then,

$$\begin{aligned}
G_{1_3} =& \sum_{i=1}^{N} 4\phi_{ii}\text{tr}\left(\mathbf{Q}_{ii}\right)\text{tr}\left\{\boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right\}\text{tr}\left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \\
&+ 24\text{tr}\left\{\boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii} \circ \phi_{ii}\boldsymbol{\Omega}_{ii}\right\}\text{tr}(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}^2) \\
=& \sum_{i=1}^{N} 4\phi_{ii}\text{tr}\left(\mathbf{Q}_{ii}\right)\sum_{s=1}^{N}(\boldsymbol{\Omega}_{ii})_{ss}^3\sum_{s=1}^{T}(\mathbf{Q}_{ii})_{ss}^3 \\
&+ 24\phi_{ii}\sum_{s=1}^{N}(\boldsymbol{\Omega}_{ii})_{ss}^3\sum_{s=1}^{T}(\mathbf{Q}_{ii})_{ss}^2(\mathbf{Q}_{ii}^2)_{ss} \\
\leq& \sum_{i=1}^{N} 4\phi_{ii}^4|\text{tr}\left(\mathbf{Q}_{ii}\right)||\sum_{s=1}^{T}(\mathbf{Q}_{ii})_{ss}^3| + 24\phi_{ii}^4|\sum_{s=1}^{T}(\mathbf{Q}_{ii})_{ss}^2(\mathbf{Q}_{ii}^2)_{ss}| \\
=& O(NT^2) + O(NT).
\end{aligned}$$

Similarly, we can easily obtain that $G_{1_4} = O(NT)$. Note that for any square matrix $\mathbf{A} \in \mathbb{R}^{T\times T}$,

$$|\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{A})_{st}| \leq \sqrt{T^2\text{tr}\left(\mathbf{A}\mathbf{A}'\right)}.$$

So, we next have

$$\begin{aligned}
G_{1_5} =& \sum_{i=1}^{N} 24\tau_N'\left[\left\{\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right\}\boldsymbol{\Omega}_{ii}\left\{\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right\}\right]\tau_N \\
&\times \tau_T'\left[\left\{\mathbf{I}_T \circ \mathbf{Q}_{ii}\right\}\mathbf{Q}_{ii}\left\{\mathbf{I}_T \circ \mathbf{Q}_{ii}\right\}\right]\tau_T\phi_{ii}\text{tr}(\mathbf{Q}_{ii})
\end{aligned}$$

$$+48\tau_N'\left\{(\mathbf{I}_N\circ\boldsymbol{\Omega}_{ii})\,\phi_{ii}\boldsymbol{\Omega}_{ii}\,\{\mathbf{I}_N\circ\boldsymbol{\Omega}_{ii}\}\right\}\tau_N\tau_T'\left\{(\mathbf{I}_N\circ\mathbf{Q}_{ii})\,\mathbf{Q}_{ii}^2\,(\mathbf{I}_N\circ\mathbf{Q}_{ii})\right\}\tau_T$$

$$+16\tau_N'\,(\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii})\,\tau_N\tau_T'\,(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii})\,\tau_T\phi_{ii}\mathrm{tr}\,(\mathbf{Q}_{ii})$$

$$+96\tau_N'\,(\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii})\,\boldsymbol{\Omega}_{ii}\,(\mathbf{I}_N\circ\boldsymbol{\Omega}_{ii})\,\tau_N\tau_T\prime\,(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii})\,\mathbf{Q}_{ii}\,(\mathbf{I}_T\circ\mathbf{Q}_{ii})\,\tau_T$$

$$+96\mathrm{tr}\,\{\boldsymbol{\Omega}_{ii}\,(\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii})\,\boldsymbol{\Omega}_{ii}\}\,\mathrm{tr}\,\{\mathbf{Q}_{ii}\,(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii})\,\mathbf{Q}_{ii}\}$$

$$=\sum_{i=1}^{N}24\sum_{s=1}^{N}\sum_{t=1}^{N}(\boldsymbol{\Omega}_{ii})_{ss}(\boldsymbol{\Omega}_{ii})_{st}(\boldsymbol{\Omega}_{ii})_{tt}\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{Q}_{ii})_{ss}(\mathbf{Q}_{ii})_{st}(\mathbf{Q}_{ii})_{tt}$$

$$+48\phi_{ii}\sum_{s=1}^{N}\sum_{t=1}^{N}(\boldsymbol{\Omega}_{ii})_{ss}(\boldsymbol{\Omega}_{ii})_{st}(\boldsymbol{\Omega}_{ii})_{tt}\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{Q}_{ii})_{ss}(\mathbf{Q}_{ii}^2)_{st}(\mathbf{Q}_{ii})_{tt}$$

$$+16\phi_{ii}\mathrm{tr}\,(\mathbf{Q}_{ii})\sum_{s=1}^{N}\sum_{t=1}^{N}(\boldsymbol{\Omega}_{ii})_{st}^3\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{Q}_{ii})_{st}^3$$

$$+96\sum_{s=1}^{N}\sum_{t=1}^{N}\sum_{q=1}^{N}(\boldsymbol{\Omega}_{ii})_{qt}(\boldsymbol{\Omega}_{ii})_{tt}(\boldsymbol{\Omega}_{ii})_{sq}^2\sum_{s=1}^{N}\sum_{t=1}^{N}\sum_{q=1}^{N}(\mathbf{Q}_{ii})_{qt}(\mathbf{Q}_{ii})_{tt}(\mathbf{Q}_{ii})_{sq}^2$$

$$+96\phi_{ii}\sum_{s=1}^{N}\sum_{t=1}^{N}(\boldsymbol{\Omega}_{ii})_{st}^2(\boldsymbol{\Omega}_{ii})_{ts}\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{Q}_{ii})_{st}^2(\mathbf{Q}_{ii}^2)_{ts}$$

$$\leq\sum_{i=1}^{N}24\phi_{ii}^2\sum_{s=1}^{T}\sum_{t=1}^{T}C|(\mathbf{Q}_{ii})_{st}|+48\phi_{ii}^3\sum_{s=1}^{T}\sum_{t=1}^{T}C|(\mathbf{Q}_{ii}^2)_{st}|$$

$$+16\phi_{ii}^3\mathrm{tr}\,(\mathbf{Q}_{ii})\sum_{s=1}^{T}\sum_{t=1}^{T}C(\mathbf{Q}_{ii})_{st}^2+96\phi_{ii}^3\sum_{t=1}^{N}\sum_{q=1}^{N}|(\mathbf{Q}_{ii})_{qt}|(\mathbf{Q}_{ii})_{tt}(\mathbf{Q}_{ii}^2)_{qq}$$

$$+96\phi_{ii}^3\sum_{s=1}^{T}\sum_{t=1}^{T}C(\mathbf{Q}_{ii})_{st}^2$$

$$=O(NT^{3/2})+O(NT^2)+O(NT).$$

Similarly, for G_{1_6}, we have

$$G_{1_6}=\sum_{i=1}^{N}2\mathrm{tr}^2\,(\boldsymbol{\Omega}_{ii}\circ\boldsymbol{\Omega}_{ii})\,\mathrm{tr}^2\,(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii})$$

$$
\begin{aligned}
&+ 24\tau_N' \left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right) \left(\boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right) \left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right) \tau_N \\
&\qquad\qquad\qquad \times \tau_T' \left(\mathbf{I}_T \circ \mathbf{Q}_{ii}\right) \left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \left(\mathbf{I}_T \circ \mathbf{Q}_{ii}\right) \tau_T \\
&+ 8\tau_N' \left(\boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right) \tau_N \tau_T' \left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \tau_T \\
=& \sum_{i=1}^{N} 2\Big\{ \sum_{s=1}^{N} (\boldsymbol{\Omega}_{ii})_{ss}^2 \Big\}^2 \Big\{ \sum_{s=1}^{N} (\mathbf{Q}_{ii})_{ss}^2 \Big\}^2 + 8 \sum_{s=1}^{N} \sum_{t=1}^{N} (\boldsymbol{\Omega}_{ii})_{st}^4 \sum_{s=1}^{T} \sum_{t=1}^{T} (\mathbf{Q}_{ii})_{st}^4 \\
&+ 24 \sum_{s=1}^{N} \sum_{t=1}^{N} (\boldsymbol{\Omega}_{ii})_{st}^2 (\boldsymbol{\Omega}_{ii})_{ss} (\boldsymbol{\Omega}_{ii})_{tt} \sum_{s=1}^{T} \sum_{t=1}^{T} (\mathbf{Q}_{ii})_{st}^2 (\mathbf{Q}_{ii})_{ss} (\mathbf{Q}_{ii})_{tt} \\
\leq& \sum_{i=1}^{N} 2\phi_{ii}^4 \mathrm{tr}^2(\mathbf{Q}_{ii}^2) + 24\phi_{ii}^4 \sum_{s=1}^{T} \sum_{t=1}^{T} C(\mathbf{Q}_{ii})_{st}^2 + 8\phi_{ii}^4 \sum_{s=1}^{T} \sum_{t=1}^{T} C(\mathbf{Q}_{ii})_{st}^2 \\
=& O(NT^2) + O(NT).
\end{aligned}
$$

Furthermore, we have

$$
\begin{aligned}
G_{1_7} =& \sum_{i=1}^{N} 24\tau_N' \left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right) \boldsymbol{\Omega}_{ii} \left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right) \tau_N \\
&\qquad\qquad\qquad \times \tau_T' \left(\mathbf{I}_T \circ \mathbf{Q}_{ii}\right) \mathbf{Q}_{ii} \left(\mathbf{I}_T \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \tau_T \\
&+ 32\tau_N' \left(\mathbf{I}_N \circ \boldsymbol{\Omega}_{ii}\right) \left(\boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii} \circ \boldsymbol{\Omega}_{ii}\right) \tau_N \tau_T' \left(\mathbf{I}_T \circ \mathbf{Q}_{ii}\right) \left(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}\right) \tau_T \\
=& \sum_{i=1}^{N} 24 \sum_{s=1}^{N} \sum_{t=1}^{N} (\boldsymbol{\Omega}_{ii})_{st} (\boldsymbol{\Omega}_{ii})_{ss} (\boldsymbol{\Omega}_{ii})_{tt}^2 \sum_{s=1}^{T} \sum_{t=1}^{T} (\mathbf{Q}_{ii})_{st} (\mathbf{Q}_{ii})_{ss} (\mathbf{Q}_{ii})_{tt}^2 \\
&+ 32 \sum_{s=1}^{N} \sum_{t=1}^{N} (\boldsymbol{\Omega}_{ii})_{st}^3 (\boldsymbol{\Omega}_{ii})_{ss} \sum_{s=1}^{T} \sum_{t=1}^{T} (\mathbf{Q}_{ii})_{st}^3 (\mathbf{Q}_{ii})_{ss} \\
\leq& \sum_{i=1}^{N} 24\phi_{ii}^3 \sum_{s=1}^{T} \sum_{t=1}^{T} C|(\mathbf{Q}_{ii})_{st}| + 32\phi_{ii}^3 \sum_{s=1}^{T} \sum_{t=1}^{T} C(\mathbf{Q}_{ii})_{st}^2 \\
=& O(NT^{3/2}) + O(NT).
\end{aligned}
$$

Combining the above calculations, we can get $G_{N1} = O(NT^3)$. Similarly, we adopt similar steps for $G_{N2}, G_{N3}, G_{N4}, G_{N5}, G_{N6}$ and G_{N7}, and eventually we can get $G_{N2} = O(NT^3 + N^2T^2)$, $G_{N3} = O(N^2T^3)$, $G_{N4} = O(N^2T^{5/2} + N^{1\vee(2-1/\tau)}T^3)$, $G_{N5} = O(N^{1\vee(3-2/\tau)}T^3) + O(N^3T^{5/2})$, $G_{N6} = O(N^{1\vee(3-2/\tau)}T^3) + O(N^3T^2) + O(N^{2\vee(3-1/\tau)}T^{5/2})$ and $G_{N7} = O(N^{2\vee(4-2/\tau)}T^3) + O(N^{3\vee(4-1/\tau)}T^2) + O(N^{3\vee(4-1/\tau)}T^{5/2}) + O(N^4T^{3/2})$. Combining $G_{N1}, \cdots, G_{N7}$, we can conclude that

$$\begin{aligned}
&\operatorname{Var}\Big[\Big\{\sum_{i=1}^N\sum_{j=1}^N(\hat{\boldsymbol{\epsilon}}_i'\hat{\boldsymbol{\epsilon}}_j)^2\Big\}^2\Big] \\
&= O(N^4T^{3/2} + N^2T^3 + N^3T^{5/2} + N^{2\vee(4-2/\tau)}T^3 + N^{3\vee(4-1/\tau)}T^{5/2}).
\end{aligned}$$

Similarly, for

$$\sum_{i=1}^N \hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot} = \sum_{i=1}^N Z'\left(\mathbf{\Omega}_{ii}\otimes\mathbf{Q}_{ii}\right)Z = Z'\left\{\sum_{i=1}^N\left(\mathbf{\Omega}_{ii}\otimes\mathbf{Q}_{ii}\right)\right\}Z \doteq Z'\mathbf{J}Z.$$

So, we have $\operatorname{Var}\left\{\left(\sum_{i=1}^N \hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\right\} = \operatorname{Var}\left\{\left(Z'\mathbf{J}Z\right)^2\right\}$. Again using Lemma S.4, we have

$$\begin{aligned}
&\operatorname{Var}\left\{\left(Z'\mathbf{J}Z\right)^2\right\} \\
=&E\left(Z'\mathbf{J}Z\right)^4 - E^2\left(Z'\mathbf{J}Z\right)^2 \\
=&K_1 + \gamma_2K_2 + \gamma_4K_3 + \gamma_6K_4 + \gamma_1^2K_5 + \gamma_2^2K_6 + \gamma_1\gamma_3K_7,
\end{aligned}$$

where

$$K_1 = 8\operatorname{tr}^2\left(\mathbf{J}\right)\operatorname{tr}\left(\mathbf{J}^2\right) + 8\operatorname{tr}^2\left(\mathbf{J}^2\right) + 32\operatorname{tr}\left(\mathbf{J}\right)\operatorname{tr}\left(\mathbf{J}^3\right) + 48\operatorname{tr}\left(\mathbf{J}^4\right),$$

$$K_2 = 4\mathrm{tr}\left(\mathbf{J}\right)\mathrm{tr}\left(\mathbf{J}\right)\mathrm{tr}\left(\mathbf{J}\circ\mathbf{J}\right) + 8\tau_{NT}'\left(\mathbf{J}\circ\mathbf{J}\right)\tau_{NT}\mathrm{tr}\left(\mathbf{J}\circ\mathbf{J}\right) + 48\mathrm{tr}\left(\mathbf{J}\right)\mathrm{tr}\left(\mathbf{J}\circ\mathbf{J}^2\right)$$
$$+ 96\mathrm{tr}\left\{\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\mathbf{J}^3\right\} + 48\tau_{NT}'\left(\mathbf{I}_{NT}\circ\mathbf{J}^2\right)\left(\mathbf{I}_{NT}\circ\mathbf{J}^2\right)\tau_{NT},$$
$$K_3 = 4\mathrm{tr}\left(\mathbf{J}\right)\mathrm{tr}\left(\mathbf{J}\circ\mathbf{J}\circ\mathbf{J}\right) + 24\mathrm{tr}\left(\mathbf{J}\circ\mathbf{J}\circ\mathbf{J}^2\right),$$
$$K_4 = \mathrm{tr}\left(\mathbf{J}\circ\mathbf{J}\circ\mathbf{J}\circ\mathbf{J}\right),$$
$$K_5 = 24\tau_{NT}'\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\mathbf{J}\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\tau_{NT}\mathrm{tr}\left(\mathbf{J}\right) + 48\tau_{NT}'\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\mathbf{J}^2\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\tau_{NT}$$
$$+ 16\tau_{NT}'\left(\mathbf{J}\circ\mathbf{J}\circ\mathbf{J}\right)\tau_{NT}\mathrm{tr}\left(\mathbf{J}\right) + 96\tau_{NT}'\left(\mathbf{J}\circ\mathbf{J}\right)\mathbf{J}\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\tau_{NT}$$
$$+ 96\mathrm{tr}\left(\mathbf{J}\left(\mathbf{J}\circ\mathbf{J}\right)\mathbf{J}\right),$$
$$K_6 = 3\mathrm{tr}\left(\mathbf{J}\circ\mathbf{J}\right)\mathrm{tr}\left(\mathbf{J}\circ\mathbf{J}\right) + 24\tau_{NT}'\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\left(\mathbf{J}\circ\mathbf{J}\right)\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\tau_{NT}$$
$$+ 8\tau_{NT}'\left(\mathbf{J}\circ\mathbf{J}\circ\mathbf{J}\circ\mathbf{J}\right)\tau_{NT},$$
$$K_7 = 24\tau_{NT}'\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\mathbf{J}\left(\mathbf{I}_{NT}\circ\mathbf{J}\circ\mathbf{J}\right)\tau_{NT} + 32\tau_{NT}'\left(\mathbf{I}_{NT}\circ\mathbf{J}\right)\left(\mathbf{J}\circ\mathbf{J}\circ\mathbf{J}\right)\tau_{NT}.$$

First, we have

$$\mathrm{tr}\left(\mathbf{J}\right) = \sum_{i=1}^{N}\phi_{ii}\mathrm{tr}(\mathbf{Q}_{ii}) = O(NT),$$
$$\mathrm{tr}\left(\mathbf{J}^2\right) = \sum_{i=1}^{N}\sum_{j=1}^{N}\phi_{ij}^2\mathrm{tr}(\mathbf{Q}_{ii}\mathbf{Q}_{jj}) = O(NT).$$

Then, due to Assumption 3 and

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}|\phi_{ij}\phi_{jk}\phi_{ik}| = O(N^{1\vee(3-2/\tau)}),$$

we have

$$\mathrm{tr}\left(\mathbf{J}^3\right) = \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\phi_{ij}\phi_{jk}\phi_{ik}\mathrm{tr}(\mathbf{Q}_{ii}\mathbf{Q}_{jj}\mathbf{Q}_{kk})$$

$$
\begin{aligned}
=&\mathrm{tr}\left(\boldsymbol{\Phi}^3\right)\mathrm{tr}\left(\boldsymbol{\Sigma}^3\right)+\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N\phi_{ij}\phi_{jk}\phi_{ik}\left\{\mathrm{tr}(\mathbf{Q}_{ii}\mathbf{Q}_{jj}\mathbf{Q}_{kk})-\mathrm{tr}\left(\boldsymbol{\Sigma}^3\right)\right\}\\
\leq&\mathrm{tr}\left(\boldsymbol{\Phi}^3\right)\mathrm{tr}\left(\boldsymbol{\Sigma}^3\right)+O(N^{1\vee(3-2/\tau)})=O(NT).
\end{aligned}
$$

Next, due to Assumption 3 and

$$
\begin{aligned}
&\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N\sum_{l=1}^N|\phi_{ij}\phi_{jk}\phi_{kl}\phi_{li}|\\
\leq&\sum_{i=1}^N\sum_{k=1}^N\left(\sum_{j=1}^N|\phi_{ij}\phi_{jk}|\sum_{l=1}^N|\phi_{kl}\phi_{li}|\right)\\
\leq&\sum_{i=1}^N\sum_{k=1}^N\left(\sum_{j=1}^N\phi_{ij}^2\sum_{j=1}^N\phi_{jk}^2\right)\\
\leq&\sum_{i=1}^N\sum_{k=1}^N C=O(N^2),
\end{aligned}
$$

we can get

$$
\begin{aligned}
\mathrm{tr}\left(\mathbf{J}^4\right)=&\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N\sum_{l=1}^N\phi_{ij}\phi_{jk}\phi_{kl}\phi_{li}\mathrm{tr}(\mathbf{Q}_{ii}\mathbf{Q}_{jj}\mathbf{Q}_{kk}\mathbf{Q}_{ll})\\
=&\mathrm{tr}\left(\boldsymbol{\Phi}^4\right)\mathrm{tr}\left(\boldsymbol{\Sigma}^4\right)\\
&+\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N\sum_{l=1}^N\phi_{ij}\phi_{jk}\phi_{kl}\phi_{li}\left\{\mathrm{tr}(\mathbf{Q}_{ii}\mathbf{Q}_{jj}\mathbf{Q}_{kk}\mathbf{Q}_{ll})-\mathrm{tr}\left(\boldsymbol{\Sigma}^4\right)\right\}\\
\leq&\mathrm{tr}\left(\boldsymbol{\Phi}^4\right)\mathrm{tr}\left(\boldsymbol{\Sigma}^4\right)+O(N^2)=O(NT).
\end{aligned}
$$

Similarly, according to the above formulas, we have $K_1=O(N^3T^3)=K_2$, $K_3=O(N^2T^2)$, $K_4=O(NT)$, $K_5=O(N^3T^3)$, $K_6=O(N^2T^2)=K_7$.

Then, we can conclude that

$$\mathrm{Var}\left\{\left(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\right\}=O(N^3T^3).$$

Finally, by Assumption 3, we can obtain that

$$\begin{aligned}E(b_N^2)\leq&\frac{1}{N^2T^4}\left[2\mathrm{Var}\left\{\sum_{i=1}^{N}\sum_{j=1}^{N}(\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{j\cdot})^2\right\}+\frac{2}{T^2}\mathrm{Var}\left\{\left(\sum_{i=1}^{N}\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}\right)^2\right\}\right]\\=&O\left(\frac{1}{T^{\frac{1}{2}\vee\left(\frac{1}{\tau}-\frac{1}{2}\right)}}\right).\end{aligned}$$

□

S4.5 Proof of Lemma S.10

Recall that $\mathbf{M}_i=\mathbf{P}_i\mathbf{\Sigma}\mathbf{P}_i$. Let $\hat{\epsilon}_{i\cdot}=\mathbf{P}_i\epsilon_{i\cdot}=\mathbf{P}_i\mathbf{R}Z_{i\cdot}$, where $Z_{i\cdot}=(Z_{i1},\cdots,Z_{iT})'$ is the i-th row vector of $\mathbf{Z}$. Define $\mathbf{Q}_{ij}\doteq\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\mathbf{R}$, then, we have $\mathbf{Q}_{ii}=\mathbf{\Pi}_i$. Obviously, by Assumption 1, we have $E\left(S_N\right)=0$. Then, we have

$$\begin{aligned}E\left(S_N^2\right)=&\frac{2}{N(N-1)}\sum\sum_{i<j}\sum\sum_{s<t}E\left(\frac{\epsilon_{i\cdot}'\mathbf{P}_i\mathbf{P}_j\epsilon_{j\cdot}}{\|\mathbf{P}_i\epsilon_{i\cdot}\|\cdot\|\mathbf{P}_j\epsilon_{j\cdot}\|}\frac{\epsilon_s'\mathbf{P}_s\mathbf{P}_t\epsilon_t}{\|\mathbf{P}_s\epsilon_s\|\cdot\|\mathbf{P}_t\epsilon_t\|}\right)\\=&\frac{2}{N(N-1)}\sum\sum_{i<j}E\left\{\frac{(\epsilon_{i\cdot}'\mathbf{P}_i\mathbf{P}_j\epsilon_{j\cdot})^2}{\|\mathbf{P}_i\epsilon_{i\cdot}\|^2\cdot\|\mathbf{P}_j\epsilon_{j\cdot}\|^2}\right\}\\=&\frac{2}{N(N-1)}\sum\sum_{i<j}E\left[\frac{\epsilon_{j\cdot}'\mathbf{P}_j\mathbf{M}_i\mathbf{P}_j\epsilon_{j\cdot}}{\mathrm{tr}(\mathbf{M}_i)\,\|\mathbf{P}_j\epsilon_{j\cdot}\|^2}\right.\\&+\frac{(\epsilon_{j\cdot}'\mathbf{P}_j\mathbf{M}_i\mathbf{P}_j\epsilon_{j\cdot})\cdot\{2\mathrm{tr}(\mathbf{M}_i^2)+\gamma_2\mathrm{tr}(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii})\}}{\mathrm{tr}^3(\mathbf{M}_i)\,\|\mathbf{P}_j\epsilon_{j\cdot}\|^2}+O(T^{-2})\\&\left.-\frac{[2(\epsilon_{j\cdot}'\mathbf{P}_j\mathbf{M}_i^2\mathbf{P}_j\epsilon_{j\cdot})+\gamma_2\mathrm{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\epsilon_{j\cdot}\epsilon_{j\cdot}'\mathbf{P}_j\mathbf{P}_i\mathbf{R})\circ\mathbf{Q}_{ii}\}]}{\mathrm{tr}^2(\mathbf{M}_i)\,\|\mathbf{P}_j\epsilon_{j\cdot}\|^2}\right]\end{aligned}$$

$$= \frac{2}{N(N-1)} \sum\sum_{i<j} \frac{\operatorname{tr}(\mathbf{M}_i \mathbf{M}_j)}{\operatorname{tr}(\mathbf{M}_i) \operatorname{tr}(\mathbf{M}_j)} + O(T^{-2})$$

$$= \sigma_{S_N}^2 \{1 + o(1)\},$$

where the third equality holds due to Lemmas S.1 and S.4, and the last equality above holds because of Lemma S.2, and the last second equality above holds because of Lemmas S.1, S.2 and S.4, $\operatorname{tr}(\mathbf{Q}_{ii} \circ \mathbf{Q}_{ii}) \leq \operatorname{tr}(\mathbf{Q}_{ii}^2) = \operatorname{tr}(\mathbf{M}_i^2) = O(T)$ and

$$0 \leq \frac{\operatorname{tr}\{(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\epsilon_{j\cdot}\epsilon_{j\cdot}'\mathbf{P}_j\mathbf{P}_i\mathbf{R}) \circ \mathbf{Q}_{ii}\}}{\|\mathbf{P}_j\epsilon_{j\cdot}\|} \leq \frac{C\operatorname{tr}(\mathbf{R}'\mathbf{P}_i\mathbf{P}_j\epsilon_{j\cdot}\epsilon_{j\cdot}'\mathbf{P}_j\mathbf{P}_i\mathbf{R})}{\|\mathbf{P}_j\epsilon_{j\cdot}\|} = O(1).$$

The normality of S_N has yet to be proven. Define $\sum_{j=2}^k Z_j$ where $Z_j = \sum_{i=1}^{j-1} \sqrt{\frac{2}{N(N-1)}}\hat{\rho}_{ij}$, Let $\mathcal{F}_j \doteq \sigma\{\epsilon_{1\cdot}, \cdots, \epsilon_{j\cdot}\}$ be the σ-field generated by $\{\epsilon_{i\cdot}, i \leq j\}$, where $\epsilon_{i\cdot} = (\epsilon_{i1}, \cdots, \epsilon_{iT})'$. Because $\epsilon_{1\cdot}, \cdots, \epsilon_{N\cdot}$ are mutually independent under H_0. Obviously, $E(Z_j \mid \mathcal{F}_{j-1}) = 0$ and it follows that $\{\sum_{j=2}^k Z_j, \mathcal{F}_k; 2 \leq k \leq N\}$ is a zero mean martingale. According to the Martingale central limit theorem in Hall and Heyde (2014), we only need to show

$$\frac{\sum_{j=2}^N E\left(Z_j^2 \mid \mathcal{F}_{j-1}\right)}{\sigma_{S_N}^2} \xrightarrow{p} 1, \tag{S4.40}$$

and

$$E\left\{\sum_{j=2}^N E\left(Z_j^4 | \mathcal{F}_{j-1}\right)\right\} = o\{\sigma_{S_N}^4\}. \tag{S4.41}$$

We first focus on (S4.40). It can be shown that

$$\begin{aligned}\sum_{j=2}^{N} E(Z_j^2|\mathcal{F}_{j-1}) =& \sum_{j=2}^{N} E\left[\left\{\sqrt{\frac{2}{N(N-1)}}\sum_{i=1}^{j-1}\hat{\rho}_{ij}\right\}^2\middle|\mathcal{F}_{j-1}\right]\\ =&\sum_{j=2}^{N}\frac{2}{N(N-1)}E\left\{\sum_{i_1=1}^{j-1}\sum_{i_2=1}^{j-1}\hat{\rho}_{i_1j}\hat{\rho}_{i_2j}\middle|\mathcal{F}_{n,j-1}\right\}\\ =&\sum_{j=2}^{N}\frac{2}{N(N-1)}\sum_{i_1=1}^{j-1}\sum_{i_2=1}^{j-1}E\left(\hat{\rho}_{i_1j}\hat{\rho}_{i_2j}\middle|\mathcal{F}_{n,j-1}\right)\\ =&A+B,\end{aligned}$$

where $A=\frac{2}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}E\left\{(\hat{\rho}_{ij})^2\middle|\mathcal{F}_{n,j-1}\right\}$, and

$$B=\frac{4}{N(N-1)}\sum_{j=2}^{N}\sum_{i_1<i_2}^{j-1}\sum^{j-1}E\left(\hat{\rho}_{i_1j}\hat{\rho}_{i_2j}\middle|\mathcal{F}_{n,j-1}\right).$$

First, we consider A. By Lemmas S.1, S.3 and S.4, we have

$$\begin{aligned}A=&\frac{2}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\left\{\frac{\mathrm{tr}\left(\hat{\epsilon}_{i.}'\mathbf{M}_j\hat{\epsilon}_{i.}\right)}{\mathrm{tr}\left(\mathbf{M}_j\right)\hat{\epsilon}_{i.}'\hat{\epsilon}_{i.}}+\frac{E\left(\epsilon_{j.}'\mathbf{P}_j\hat{\epsilon}_{i.}\hat{\epsilon}_{i.}'\mathbf{P}_j\epsilon_{j.}'\right)\mathrm{Var}\left(\epsilon_{j.}'\mathbf{P}_j\epsilon_{j.}\right)}{\mathrm{tr}^3\left(\mathbf{M}_j\right)\hat{\epsilon}_{i.}'\hat{\epsilon}_{i.}}\right.\\ &\left.-\frac{E\left(\epsilon_{j.}'\mathbf{P}_j\hat{\epsilon}_{i.}\hat{\epsilon}_{i.}'\mathbf{P}_j\epsilon_{j.}\epsilon_{j.}'\mathbf{P}_j\epsilon_{j.}\right)-E\left(\epsilon_{j.}'\mathbf{P}_j\hat{\epsilon}_{i.}\hat{\epsilon}_{i.}'\mathbf{P}_j\epsilon_{j.}\right)E\left(\epsilon_{j.}'\mathbf{P}_j\epsilon_{j.}\right)}{\mathrm{tr}^2\left(\mathbf{M}_j\right)\hat{\epsilon}_{i.}'\hat{\epsilon}_{i.}}\right\}+O\left(T^{-2}\right)\\ =&\frac{2}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\left[\frac{\mathrm{tr}\left(\hat{\epsilon}_{i.}'\mathbf{M}_j\hat{\epsilon}_{i.}\right)}{\mathrm{tr}\left(\mathbf{M}_j\right)\hat{\epsilon}_{i.}'\hat{\epsilon}_{i.}}-\frac{\gamma_2\mathrm{tr}[(\mathbf{R}'\mathbf{P}_j\mathbf{P}_i\epsilon_{i.}\epsilon_{i.}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})\circ\mathbf{Q}_{jj}]}{\mathrm{tr}^2\left(\mathbf{M}_j\right)\hat{\epsilon}_{i.}'\hat{\epsilon}_{i.}}\right.\\ &\left.-\frac{2\mathrm{tr}\left(\hat{\epsilon}_{i.}'\mathbf{M}_j^2\hat{\epsilon}_{i.}\right)}{\mathrm{tr}^2\left(\mathbf{M}_j\right)\hat{\epsilon}_{i.}'\hat{\epsilon}_{i.}}+\frac{\mathrm{tr}\left(\hat{\epsilon}_{i.}'\mathbf{M}_j\hat{\epsilon}_{i.}\right)\left\{2\mathrm{tr}\left(\mathbf{M}_j^2\right)+\gamma_2\mathrm{tr}(\mathbf{Q}_{jj}\circ\mathbf{Q}_{jj})\right\}}{\mathrm{tr}^3\left(\mathbf{M}_j\right)\hat{\epsilon}_{i.}'\hat{\epsilon}_{i.}}\right]+O\left(T^{-2}\right)\\ =&A_1+A_2-A_3-A_4+O\left(T^{-2}\right).\end{aligned}$$

By Lemmas S.1 and S.2, we have

$$E\left(A_1\right)=\frac{2}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}E\left(\frac{\mathrm{tr}\left(\hat{\epsilon}_{i.}'\mathbf{M}_j\hat{\epsilon}_{i.}\right)}{\mathrm{tr}\left(\mathbf{M}_j\right)\hat{\epsilon}_{i.}'\hat{\epsilon}_{i.}}\right)$$

$$
\begin{aligned}
&= \frac{2}{N(N-1)} \sum_{j=2}^{N} \sum_{i=1}^{j-1} \frac{\operatorname{tr}\left(\mathbf{M}_i \mathbf{M}_j\right)}{\operatorname{tr}\left(\mathbf{M}_i\right) \operatorname{tr}\left(\mathbf{M}_j\right)} + O\left\{\frac{1}{T \operatorname{tr}\left(\mathbf{M}_j\right)}\right\} \\
&= \sigma_{S_N}^2 \left\{1 + o(1)\right\}
\end{aligned}
$$

and

$$
\begin{aligned}
\operatorname{Var}\left(A_1\right) =& E\left(A_1^2\right) - E^2\left(A_1\right) \\
\leq & \frac{4}{N^2(N-1)^2} \sum_{j=2}^{N} \sum_{i=1}^{j-1} E\left\{\frac{\operatorname{tr}^2\left(\hat{\epsilon}_{i \cdot}' \mathbf{M}_j \hat{\epsilon}_{i \cdot}\right)}{\operatorname{tr}^2\left(\mathbf{M}_j\right)\left(\hat{\epsilon}_{i \cdot}' \hat{\epsilon}_{i \cdot}\right)^2}\right\} \\
& + \frac{8}{N^2(N-1)^2} \sum_{2 \leq j_1 < j_2 \leq N} \sum_{1 \leq i \leq j_1 - 1} E\left\{\frac{\operatorname{tr}\left(\hat{\epsilon}_{i \cdot}' \mathbf{M}_{j_1} \hat{\epsilon}_{i \cdot}\right) \operatorname{tr}\left(\hat{\epsilon}_{i \cdot}' \mathbf{M}_{j_2} \hat{\epsilon}_{i \cdot}\right)}{\operatorname{tr}\left(\mathbf{M}_{j_1}\right) \hat{\epsilon}_{i \cdot}' \hat{\epsilon}_{i \cdot} \operatorname{tr}\left(\mathbf{M}_{j_2}\right) \hat{\epsilon}_{i \cdot}' \hat{\epsilon}_{i \cdot}}\right\} \\
\leq & \frac{4}{N^2(N-1)^2} \sum_{j=2}^{N} \sum_{i=1}^{j-1} E\left\{\frac{\operatorname{tr}^2\left(\hat{\epsilon}_{i \cdot}' \mathbf{M}_j \hat{\epsilon}_{i \cdot}\right)}{\operatorname{tr}^2\left(\mathbf{M}_j\right)\left(\hat{\epsilon}_{i \cdot}' \hat{\epsilon}_{i \cdot}\right)^2}\right\} + \frac{8}{N^2(N-1)^2} \\
& \times \sum_{2 \leq j_1 < j_2 \leq N} \sum_{1 \leq i \leq j_1 - 1} \frac{\sqrt{E\left\{\frac{\operatorname{tr}^2\left(\hat{\epsilon}_{i \cdot}' \mathbf{M}_{j_1} \hat{\epsilon}_{i \cdot}\right)}{\left(\hat{\epsilon}_{i \cdot}' \hat{\epsilon}_{i \cdot}\right)^2}\right\} E\left\{\frac{\operatorname{tr}^2\left(\hat{\epsilon}_{i \cdot}' \mathbf{M}_{j_2} \hat{\epsilon}_{i \cdot}\right)}{\left(\hat{\epsilon}_{i \cdot}' \hat{\epsilon}_{i \cdot}\right)^2}\right\}}}{\operatorname{tr}\left(\mathbf{M}_{j_1}\right) \operatorname{tr}\left(\mathbf{M}_{j_2}\right)} \\
= & O\left(N^{-2} T^{-2}\right) + O\left(N^{-1} T^{-2}\right) \\
= & o\left(\sigma_{S_N}^4\right),
\end{aligned}
$$

where the last second equality holds because

$$
\begin{aligned}
& E\left\{\frac{\operatorname{tr}^2\left(\hat{\epsilon}_{i \cdot}' \mathbf{M}_j \hat{\epsilon}_{i \cdot}\right)}{\operatorname{tr}^2\left(\mathbf{M}_j\right)\left(\hat{\epsilon}_{i \cdot}' \hat{\epsilon}_{i \cdot}\right)^2}\right\} \\
= & \frac{\operatorname{tr}^2\left(\mathbf{M}_i \mathbf{M}_j\right) + 2 \operatorname{tr}\left(\mathbf{M}_i \mathbf{M}_j \mathbf{M}_i \mathbf{M}_j\right)}{\operatorname{tr}^2\left(\mathbf{M}_j\right) \operatorname{tr}^2\left(\mathbf{M}_i\right)} + O\left\{T^{-1} \operatorname{tr}^{-2}\left(\mathbf{M}_j\right)\right\} \\
& + \frac{\gamma_2 \operatorname{tr}\left[\left(\mathbf{R}' \mathbf{P}_i \mathbf{M}_j \mathbf{P}_i \mathbf{R}\right) \circ \left(\mathbf{R}' \mathbf{P}_i \mathbf{M}_j \mathbf{P}_i \mathbf{R}\right)\right]}{\operatorname{tr}^2\left(\mathbf{M}_j\right) \operatorname{tr}^2\left(\mathbf{M}_i\right)}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{\mathrm{tr}^2\left(\mathbf{M}_i\mathbf{M}_j\right)+(2+|\gamma_2|)\mathrm{tr}\left(\mathbf{M}_i\mathbf{M}_j\mathbf{M}_i\mathbf{M}_j\right)}{\mathrm{tr}^2\left(\mathbf{M}_j\right)\mathrm{tr}^2\left(\mathbf{M}_i\right)}+O\left\{T^{-1}\mathrm{tr}^{-2}\left(\mathbf{M}_j\right)\right\}\\
&=O(T^{-2}).
\end{aligned}
$$

Hence, we can conclude that $A_1 = \sigma^2_{S_N} + o_p\left(\sigma^2_{S_N}\right)$. Then, we have

$$
\begin{aligned}
E\left(|A_2|\right) \leq& \frac{2}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}E\left[\frac{\mathrm{tr}\left(\hat{\epsilon}_{i\cdot}'\mathbf{M}_j\hat{\epsilon}_{i\cdot}\right)\left\{2\mathrm{tr}\left(\mathbf{M}_j^2\right)+|\gamma_2|\mathrm{tr}(\mathbf{Q}_{jj}\circ\mathbf{Q}_{jj})\right\}}{\mathrm{tr}^3\left(\mathbf{M}_j\right)\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}}\right]\\
=&\frac{2}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\frac{2\mathrm{tr}\left(\mathbf{M}_j^2\right)+|\gamma_2|\mathrm{tr}(\mathbf{Q}_{jj}\circ\mathbf{Q}_{jj})}{\mathrm{tr}^3\left(\mathbf{M}_j\right)}E\left\{\frac{\mathrm{tr}\left(\hat{\epsilon}_{i\cdot}'\mathbf{M}_j\hat{\epsilon}_{i\cdot}\right)}{\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}}\right\}\\
=&\frac{2}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\frac{2\mathrm{tr}\left(\mathbf{M}_j^2\right)+|\gamma_2|\mathrm{tr}(\mathbf{Q}_{jj}\circ\mathbf{Q}_{jj})}{\mathrm{tr}^3\left(\mathbf{M}_j\right)}\left\{\frac{\mathrm{tr}\left(\mathbf{M}_i\mathbf{M}_j\right)}{\mathrm{tr}\left(\mathbf{M}_i\right)}\right.\\
&\left.+O\left(T^{-1}\right)\right\}\\
=&O\left(T^{-2}\right)=o\left(\sigma^2_{S_N}\right),
\end{aligned}
$$

where the last two equality holds due to $\mathrm{tr}(\mathbf{Q}_{jj}\circ\mathbf{Q}_{jj}) \leq \mathrm{tr}(\mathbf{M}_j\mathbf{M}_j) = O(T)$.

Hence, we have $A_2 = o_p\left(\sigma^2_{S_N}\right)$ by Markov inequality. Next, we have

$$
\begin{aligned}
E\left(A_3\right) =& \frac{2}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}E\left\{\frac{2\mathrm{tr}\left(\hat{\epsilon}_{i\cdot}'\mathbf{M}_j^2\hat{\epsilon}_{i\cdot}\right)}{\mathrm{tr}^2\left(\mathbf{M}_j\right)\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}}\right\}\\
=&\frac{4}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\frac{1}{\mathrm{tr}^2\left(\mathbf{M}_j\right)}E\left\{\frac{\mathrm{tr}\left(\hat{\epsilon}_{i\cdot}'\mathbf{M}_j^2\hat{\epsilon}_{i\cdot}\right)}{\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}}\right\}\\
=&\frac{4}{N(N-1)}\sum_{j=2}^{N}\sum_{i=1}^{j-1}\frac{1}{\mathrm{tr}^2\left(\mathbf{M}_j\right)}\left\{\frac{\mathrm{tr}\left(\mathbf{M}_i\mathbf{M}_j^2\right)}{\mathrm{tr}\left(\mathbf{M}_i\right)}+O\left(T^{-1}\right)\right\}\\
=&O\left(T^{-2}\right)=o\left(\sigma^2_{S_N}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
E(A_4/\gamma_2) =& \frac{2}{N(N-1)} \sum_{j=2}^{N} \sum_{i=1}^{j-1} E\left\{ \frac{\mathrm{tr}[(\mathbf{R}'\mathbf{P}_j\mathbf{P}_i\epsilon_{i\cdot}\epsilon_{i\cdot}'\mathbf{P}_i\mathbf{P}_j\mathbf{R}) \circ \mathbf{Q}_{jj}]}{\mathrm{tr}^2\left(\mathbf{M}_j\right)\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}} \right\} \\
\leq& \frac{2C}{N(N-1)} \sum_{j=2}^{N} \sum_{i=1}^{j-1} E\left\{ \frac{\mathrm{tr}(\mathbf{R}'\mathbf{P}_j\mathbf{P}_i\epsilon_{i\cdot}\epsilon_{i\cdot}'\mathbf{P}_i\mathbf{P}_j\mathbf{R})}{\mathrm{tr}^2\left(\mathbf{M}_j\right)\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}} \right\} \\
=& \frac{2C}{N(N-1)} \sum_{j=2}^{N} \sum_{i=1}^{j-1} E\left\{ \frac{\mathrm{tr}(\epsilon_{i\cdot}'\mathbf{P}_i\mathbf{M}_j\mathbf{P}_i\epsilon_{i\cdot})}{\mathrm{tr}^2\left(\mathbf{M}_j\right)\hat{\epsilon}_{i\cdot}'\hat{\epsilon}_{i\cdot}} \right\} \\
=& \frac{2C}{N(N-1)} \sum_{j=2}^{N} \sum_{i=1}^{j-1} \left\{ \frac{\mathrm{tr}(\mathbf{M}_j\mathbf{M}_i)}{\mathrm{tr}^2\left(\mathbf{M}_j\right)\mathrm{tr}(\mathbf{M}_i)} + O\left(\frac{1}{T\mathrm{tr}^2\left(\mathbf{M}_j\right)}\right) \right\} \\
=& O\left(T^{-2}\right) = o\left(\sigma_{S_N}^2\right).
\end{aligned}
$$

Because A_3 and A_4/γ_2 are non-negative, we have $A_3 = o_p\left(\sigma_{S_N}^2\right)$ and $A_4/\gamma_2 = o_p\left(\sigma_{S_N}^2\right)$. Second, we focus on B.

$$
\begin{aligned}
B =& \frac{4}{N(N-1)} \sum_{j=2}^{N} \sum_{i_1<i_2}^{j-1}\sum^{j-1} E\left(\hat{\rho}_{i_1 j}\hat{\rho}_{i_2 j}\middle| \mathcal{F}_{n,j-1}\right) \\
=& \frac{4}{N(N-1)} \sum_{j=2}^{N} \left\{ \sum_{i_1<i_2}^{j-1}\sum^{j-1} E\left(\hat{\rho}_{i_1 j}\hat{\rho}_{i_2 j}\middle| \mathcal{F}_{n,j-1}\right) \right\} \\
=& \sum_{j=2}^{N} \left\{ \frac{4}{N(N-1)} \sum_{i_1<i_2}^{j-1}\sum^{j-1} \delta_{i_1 i_2} \right\},
\end{aligned}
$$

where we denote that $\delta_{i_1 i_2} \doteq E\left(\hat{\rho}_{i_1 j}\hat{\rho}_{i_2 j}\middle| \mathcal{F}_{n,j-1}\right)$. Obviously, $E\left(B\right) = 0$ due to the law of total expectation. By Lemmas S.1 and S.3, we have

$$
\begin{aligned}
\delta_{i_1 i_2} =& \frac{\hat{\epsilon}_{i_1\cdot}'\mathbf{M}_j\hat{\epsilon}_{i_2\cdot}}{\sqrt{\hat{\epsilon}_{i_1\cdot}'\hat{\epsilon}_{i_1\cdot}\hat{\epsilon}_{i_2\cdot}'\hat{\epsilon}_{i_2\cdot}}\,\mathrm{tr}\left(\mathbf{M}_j\right)} + \frac{\hat{\epsilon}_{i_1\cdot}'\mathbf{M}_j\hat{\epsilon}_{i_2\cdot}\{2\mathrm{tr}(\mathbf{M}_j^2) + \gamma_2\mathrm{tr}(\mathbf{Q}_{jj}\circ\mathbf{Q}_{jj})\}}{\sqrt{\hat{\epsilon}_{i_1\cdot}'\hat{\epsilon}_{i_1\cdot}\hat{\epsilon}_{i_2\cdot}'\hat{\epsilon}_{i_2\cdot}}\,\mathrm{tr}^3\left(\mathbf{M}_j\right)} \\
&- \frac{2\hat{\epsilon}_{i_1\cdot}'\mathbf{M}_j^2\hat{\epsilon}_{i_2\cdot} + \gamma_2\mathrm{tr}[(\mathbf{R}'\mathbf{P}_j\hat{\epsilon}_{i_1\cdot}\hat{\epsilon}_{i_2\cdot}'\mathbf{P}_j\mathbf{R})\circ\mathbf{Q}_{jj}]}{\sqrt{\hat{\epsilon}_{i_1\cdot}'\hat{\epsilon}_{i_1\cdot}\hat{\epsilon}_{i_2\cdot}'\hat{\epsilon}_{i_2\cdot}}\,\mathrm{tr}^2\left(\mathbf{M}_j\right)} + O\left(T^{-2}\right)
\end{aligned}
$$

for any $1 \leq i_1 < i_2 < j \leq N$. Moreover, we have

$$
\begin{aligned}
& E\left(\delta_{i_1 i_2}^2\right) \\
\leq & 5E\left\{\frac{\left(\hat{\epsilon}_{i_1 .}' \mathbf{M}_j \hat{\epsilon}_{i_2 .}\right)^2}{\hat{\epsilon}_{i_1 .}' \hat{\epsilon}_{i_1 .} \hat{\epsilon}_{i_2 .}' \hat{\epsilon}_{i_2 .} \mathrm{tr}^2\left(\mathbf{M}_j\right)}\right\} + 5E\left\{\frac{(2+|\gamma_2|)^2\left(\hat{\epsilon}_{i_2 .}' \mathbf{M}_j \hat{\epsilon}_{i_1 .}\right)^2 \mathrm{tr}^2\left(\mathbf{M}_j^2\right)}{\hat{\epsilon}_{i_1 .}' \hat{\epsilon}_{i_1 .} \hat{\epsilon}_{i_2 .}' \hat{\epsilon}_{i_2 .} \mathrm{tr}^6\left(\mathbf{M}_j\right)}\right\} \\
& + 5E\left\{\frac{4\left(\hat{\epsilon}_{i_2 .}' \mathbf{M}_j^2 \hat{\epsilon}_{i_1 .}\right)^2}{\hat{\epsilon}_{i_1 .}' \hat{\epsilon}_{i_1 .} \hat{\epsilon}_{i_2 .}' \hat{\epsilon}_{i_2 .} \mathrm{tr}^4\left(\mathbf{M}_j\right)}\right\} + 5E\left\{\frac{\gamma_2^2 \mathrm{tr}^2[(\mathbf{R}' \mathbf{P}_j \hat{\epsilon}_{i_1 .} \hat{\epsilon}_{i_2 .}' \mathbf{P}_j \mathbf{R}) \circ \mathbf{Q}_{jj}]}{\hat{\epsilon}_{i_1 .}' \hat{\epsilon}_{i_1 .} \hat{\epsilon}_{i_2 .}' \hat{\epsilon}_{i_2 .} \mathrm{tr}^4\left(\mathbf{M}_j\right)}\right\} \\
& + O\left(T^{-4}\right) \\
\leq & \frac{5 \mathrm{tr}\left(\mathbf{M}_{i_1} \mathbf{M}_j \mathbf{M}_{i_2} \mathbf{M}_j\right)}{\mathrm{tr}^2\left(\mathbf{M}_j\right) \mathrm{tr}\left(\mathbf{M}_{i_1}\right) \mathrm{tr}\left(\mathbf{M}_{i_2}\right)} + \frac{5(2+|\gamma_2|)^2 \mathrm{tr}\left(\mathbf{M}_{i_1} \mathbf{M}_j \mathbf{M}_{i_2} \mathbf{M}_j\right) \mathrm{tr}^2\left(\mathbf{M}_j^2\right)}{\mathrm{tr}\left(\mathbf{M}_{i_1}\right) \mathrm{tr}\left(\mathbf{M}_{i_2}\right) \mathrm{tr}^6\left(\mathbf{M}_j\right)} \\
& + \frac{20 \mathrm{tr}\left(\mathbf{M}_{i_1} \mathbf{M}_j^2 \mathbf{M}_{i_2} \mathbf{M}_j^2\right)}{\mathrm{tr}\left(\mathbf{M}_{i_1}\right) \mathrm{tr}\left(\mathbf{M}_{i_2}\right) \mathrm{tr}^4\left(\mathbf{M}_j\right)} + \frac{5C^2 \gamma_2^2 T \mathrm{tr}\left(\mathbf{M}_{i_1} \mathbf{M}_j\right) \mathrm{tr}\left(\mathbf{M}_{i_2} \mathbf{M}_j\right)}{\mathrm{tr}\left(\mathbf{M}_{i_1}\right) \mathrm{tr}\left(\mathbf{M}_{i_2}\right) \mathrm{tr}^4\left(\mathbf{M}_j\right)} + O\left(T^{-3}\right) \\
= & O\left(T^{-3}\right),
\end{aligned}
$$

where the last two inequality holds due to

$$
0 \leq \mathrm{tr}^2[(\mathbf{R}' \mathbf{P}_j \epsilon_{i_1} \epsilon_{i_2}' \mathbf{P}_j \mathbf{R}) \circ \mathbf{Q}_{ii}] \leq T C^2 (\epsilon_{i_2}' \mathbf{M}_j \epsilon_{i_2})(\epsilon_{i_1}' \mathbf{M}_j \epsilon_{i_1}),
$$

by Lemma S.3, for some constant $C \geq 0$. So, we can calculate $E(B^2)$,

$$
\begin{aligned}
E\left(B^2\right) \leq & N \sum_{j=2}^{N} \frac{16}{N^2(N-2)^2} E\left(\sum_{i_1<i_2}^{j-1} \sum^{j-1} \sum_{i_3<i_4}^{j-1} \sum^{j-1} \delta_{i_1 i_2} \delta_{i_3 i_4}\right) \\
\leq & \frac{16}{N(N-2)^2} \sum_{j=2}^{N} \sum_{i_1<i_2}^{j-1} \sum^{j-1} E\left(\delta_{i_1 i_2}^2\right) \\
= & O\left(T^{-3}\right) = o\left(\sigma_{S_N}^4\right),
\end{aligned}
$$

then, we obtain that $B = o_p\left(\sigma_{S_N}^2\right)$. Next, we will focus on (S4.41) and prove that

$$\sum_{j=2}^{N} E\left(Z_j^4\right) = o\{\sigma_{S_N}^4\}$$

due to $E\left\{\sum_{j=2}^{N} E\left(Z_j^4|\mathcal{F}_{j-1}\right)\right\} = \sum_{j=2}^{N} E\left(Z_j^4\right)$. We divide $\sum_{j=2}^{N} E\left(Z_j^4\right)$ into two parts:

$$\begin{aligned}
&\sum_{j=2}^{N} E\left(Z_j^4\right) \\
=&\frac{4}{N^2(N-1)^2}\sum_{j=2}^{N}\sum_{i_1=1}^{j-1}\sum_{i_2=1}^{j-1}\sum_{i_3=1}^{j-1}\sum_{i_4=1}^{j-1} E\left(\hat{\rho}_{i_1 j}\hat{\rho}_{i_2 j}\hat{\rho}_{i_3 j}\hat{\rho}_{i_4 j}\right) \\
=&\frac{4}{N^2(N-1)^2}\sum_{j=2}^{N}\sum_{i=1}^{j-1} E\left(\hat{\rho}_{ij}^4\right) + \frac{12}{N^2(N-1)^2}\sum_{j=2}^{N}\sum_{i_1\neq i_2}^{j-1}\sum^{j-1} E\left(\hat{\rho}_{i_1 j}^2\hat{\rho}_{i_2 j}^2\right) \\
=&C+D.
\end{aligned}$$

By Lemma S.1, we have

$$\begin{aligned}
&E\left(\hat{\rho}_{ij}^4\right) \\
=&E\left\{E\left(\hat{\rho}_{ij}^4\middle|\,\epsilon_{j\cdot}\right)\right\} \\
=&E\left[\frac{2\mathrm{tr}\left(\hat{\epsilon}_{j\cdot}'\mathbf{M}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{M}_i\hat{\epsilon}_{j\cdot}\right)+\mathrm{tr}^2\left(\hat{\epsilon}_{j\cdot}'\mathbf{M}_i\hat{\epsilon}_{j\cdot}\right)}{\left(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}\right)^2\mathrm{tr}^2\left(\mathbf{M}_i\right)}\right. \\
&+\frac{\gamma_2\mathrm{tr}[(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R})]}{\left(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}\right)^2\mathrm{tr}^2\left(\mathbf{M}_i\right)} \\
&+\frac{3E\left\{\left(\epsilon_{i\cdot}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\epsilon_{i\cdot}\right)^2|\epsilon_{j\cdot}\right\}\left\{2\mathrm{tr}\left(\mathbf{M}_i^2\right)+\gamma_2\mathrm{tr}(\mathbf{Q}_{ii}\circ\mathbf{Q}_{ii})\right\}}{\left(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}\right)^2\mathrm{tr}^4\left(\mathbf{M}_i\right)}+O\left(T^{-2}\right) \\
&-\frac{2E\left\{\left(\epsilon_{i\cdot}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\epsilon_{i\cdot}\right)^2\left(\epsilon_{i\cdot}'\mathbf{P}_i\epsilon_{i\cdot}\right)|\epsilon_{j\cdot}\right\}}{\left(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}\right)^2\mathrm{tr}^3\left(\mathbf{M}_i\right)}
\end{aligned}$$

$$+\frac{2E\left\{\left(\epsilon_{i\cdot}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\epsilon_{i\cdot}\right)^2|\epsilon_{j\cdot}\right\}E\left(\epsilon_{i\cdot}'\mathbf{P}_i\epsilon_{i\cdot}\right)}{\left(\hat{\epsilon}_{j\cdot}'\hat{\epsilon}_{j\cdot}\right)^2\mathrm{tr}^3\left(\mathbf{M}_i\right)}\Bigg].$$

Then, by Lemma S.4, we have

$$
\begin{aligned}
&E\left\{\left(\epsilon_{i\cdot}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\epsilon_{i\cdot}\right)^2\left(\epsilon_{i\cdot}'\mathbf{P}_i\epsilon_{i\cdot}\right)|\epsilon_{j\cdot}\right\}-E\left\{\left(\epsilon_{i\cdot}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\epsilon_{i\cdot}\right)^2|\epsilon_{j\cdot}\right\}E\left(\epsilon_{i\cdot}'\mathbf{P}_i\epsilon_{i\cdot}\right)\\
=&\gamma_4\mathrm{tr}\left[\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{\Pi}_i\right)\right]\\
&+2\gamma_2\mathrm{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\mathrm{tr}\left[\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{\Pi}_i\right)\right]\\
&+8\gamma_2\mathrm{tr}\left[\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\mathbf{\Pi}_i\right)\right]\\
&+4\gamma_2\mathrm{tr}\left[\left(\mathbf{\Pi}_i\right)\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\right]\\
&+4\gamma_1^2\left[\tau_T'\left\{\mathbf{I}_T\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\right\}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\left\{\mathbf{I}_T\circ\left(\mathbf{\Pi}_i\right)\right\}\tau_T\right]\\
&+2\gamma_1^2\left[\tau_T'\left\{\mathbf{I}_T\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\right\}\mathbf{\Pi}_i\left\{\mathbf{I}_T\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\right\}\tau_T\right]\\
&+4\gamma_1^2\left[\tau_T'\left\{\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{\Pi}_i\right)\right\}\tau_T\right]\\
&+4\mathrm{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\mathrm{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\mathbf{\Pi}_i\right)\\
&+8\mathrm{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\mathbf{\Pi}_i\right)\\
\geq&\gamma_4\mathrm{tr}\left\{\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{\Pi}_i\right)\right\}\\
&+2\gamma_2\mathrm{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\mathrm{tr}\left\{\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{\Pi}_i\right)\right\}\\
&+8\gamma_2\mathrm{tr}\left[\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\mathbf{\Pi}_i\right)\right]\\
&+4\gamma_2\mathrm{tr}\left[\left(\mathbf{\Pi}_i\right)\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\right]\\
&+4\gamma_1^2\left[\tau_T'\left\{\mathbf{I}_T\circ\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\right)\right\}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}_{j\cdot}'\mathbf{P}_i\mathbf{R}\left\{\mathbf{I}_T\circ\left(\mathbf{\Pi}_i\right)\right\}\tau_T\right]
\end{aligned}
$$

$$+ 2\gamma_1^2 \left[\tau_T' \left\{\mathbf{I}_T \circ (\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\} \mathbf{\Pi}_i \left\{\mathbf{I}_T \circ (\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\} \tau_T\right]$$

$$+ 4\gamma_1^2 \left[\tau_T' \left\{(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{\Pi}_i)\right\} \tau_T\right],$$

where the last inequality holds because the last four terms on the right side of the first equation above are all greater than or equal to zero, according to Lemma S.3.

Then, we have

$$\begin{aligned}
&E\left(\hat{\rho}_{ij}^4\right) \\
\leq &E\Bigg(\frac{(3+|\gamma_2|)\left(\hat{\epsilon}_{j.}'\mathbf{M}_i\hat{\epsilon}_{j.}\right)^2}{\left(\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.}\right)^2 \mathrm{tr}^2\left(\mathbf{M}_i\right)} + \frac{3(3+|\gamma_2|)(2+|\gamma_2|)\left(\hat{\epsilon}_{j.}'\mathbf{M}_i\hat{\epsilon}_{j.}\right)^2 \mathrm{tr}\left(\mathbf{M}_i^2\right)}{\left(\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.}\right)^2 \mathrm{tr}^4\left(\mathbf{M}_i\right)} \\
&+ 2\mathrm{tr}^{-3}(\mathbf{M}_i)(\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.})^{-2} \\
&\qquad \times \Big\{\gamma_4 \mathrm{tr}\left\{(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{\Pi}_i)\right\} \\
&+ 2\gamma_2 \mathrm{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}\right) \mathrm{tr}\left\{(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{\Pi}_i)\right\} \\
&+ 8\gamma_2 \mathrm{tr}\left[(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}) \circ \left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}\mathbf{\Pi}_i\right)\right] \\
&+ 4\gamma_2 \mathrm{tr}\left[(\mathbf{\Pi}_i) \circ \left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}\right)\right] + O\left(T^{-2}\right) \\
&+ 4\gamma_1^2 \left[\tau_T' \left\{\mathbf{I}_T \circ (\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\} \mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R} \left\{\mathbf{I}_T \circ (\mathbf{\Pi}_i)\right\} \tau_T\right] \\
&+ 2\gamma_1^2 \left[\tau_T' \left\{\mathbf{I}_T \circ (\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\} \mathbf{\Pi}_i \left\{\mathbf{I}_T \circ (\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\} \tau_T\right] \\
&+ 4\gamma_1^2 \left[\tau_T' \left\{(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}) \circ (\mathbf{\Pi}_i)\right\} \tau_T\right] \Big\}\Bigg) \\
\leq &E\Bigg(\frac{(3+|\gamma_2|)\left(\hat{\epsilon}_{j.}'\mathbf{M}_i\hat{\epsilon}_{j.}\right)^2}{\left(\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.}\right)^2 \mathrm{tr}^2\left(\mathbf{M}_i\right)} + \frac{3(3+|\gamma_2|)(2+|\gamma_2|)\left(\hat{\epsilon}_{j.}'\mathbf{M}_i\hat{\epsilon}_{j.}\right)^2 \mathrm{tr}\left(\mathbf{M}_i^2\right)}{\left(\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.}\right)^2 \mathrm{tr}^4\left(\mathbf{M}_i\right)}
\end{aligned}$$

$$+ 2\text{tr}^{-3}(\mathbf{M}_i)(\hat{\epsilon}'_{j.}\hat{\epsilon}_{j.})^{-2}\Big\{C|\gamma_4|\text{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}\right)$$

$$+ 2C|\gamma_2|\text{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}\right)\text{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}\right)$$

$$+ 8|\gamma_2|(\hat{\epsilon}'_{j.}\mathbf{M}_i\hat{\epsilon}_{j.})^{3/2}(\hat{\epsilon}'_{j.}\mathbf{M}_i^3\hat{\epsilon}_{j.})^{1/2}$$

$$+ 4C|\gamma_2|\text{tr}\left(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}\right) + O\left(T^{-2}\right)$$

$$+ 4\gamma_1^2\left[\tau'_T\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R})\right\}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}\left\{\mathbf{I}_T\circ(\mathbf{\Pi}_i)\right\}\tau_T\right]$$

$$+ 2\gamma_1^2\left[\tau'_T\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R})\right\}\mathbf{\Pi}_i\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R})\right\}\tau_T\right]$$

$$+ 4\gamma_1^2\left[\tau'_T\left\{(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R})\circ(\mathbf{\Pi}_i)\right\}\tau_T\right]\Big\}\Bigg).$$

Because $8|\gamma_2|(\hat{\epsilon}'_{j.}\mathbf{M}_i\hat{\epsilon}_{j.})^{3/2}(\hat{\epsilon}'_{j.}\mathbf{M}_i^3\hat{\epsilon}_{j.})^{1/2}/(\hat{\epsilon}'_{j.}\hat{\epsilon}_{j.})^2 = O(1)$,

$$\begin{aligned}
&\frac{4\gamma_1^2\left[\tau'_T\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R})\right\}\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}\left\{\mathbf{I}_T\circ(\mathbf{\Pi}_i)\right\}\tau_T\right]}{(\hat{\epsilon}'_{j.}\hat{\epsilon}_{j.})^2}\\
\leq&4\gamma_1^2\sum_{s=1}^{T}\sum_{k=1}^{T}\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}'_{j.}\hat{\epsilon}_{j.}}\Big)_{sk}\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}'_{j.}\hat{\epsilon}_{j.}}\Big)_{ss}\big(\mathbf{\Pi}_i\big)_{kk}\\
\leq&4\gamma_1^2O(1)\sqrt{T^2\text{tr}\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}'_{j.}\hat{\epsilon}_{j.}}\Big)^2} = O(T)
\end{aligned}$$

and

$$\begin{aligned}
&4\gamma_1^2\left[\tau'_T\left\{(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R})\circ(\mathbf{\Pi}_i)\right\}\tau_T\right]\Big\}\\
\leq&4\gamma_1^2\sum_{s=1}^{T}\sum_{k=1}^{T}\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}'_{j.}\hat{\epsilon}_{j.}}\Big)^2_{sk}\big(\mathbf{\Pi}_i\big)_{sk}\\
\leq&4\gamma_1^2O(1)\text{tr}\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}'_{j.}\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}'_{j.}\hat{\epsilon}_{j.}}\Big)^2\\
=&O(1),
\end{aligned}$$

we have

$$
\begin{aligned}
&E\left(\hat{\rho}_{ij}^4\right) \\
\leq&\left\{\frac{(3+|\gamma_2|)}{\operatorname{tr}^2(\mathbf{M}_i)}+\frac{3(3+|\gamma_2|)(2+|\gamma_2|)\operatorname{tr}(\mathbf{M}_i^2)}{\operatorname{tr}^4(\mathbf{M}_i)}\right\}E\left\{\frac{\left(\hat{\epsilon}_{j.}'\mathbf{M}_i\hat{\epsilon}_{j.}\right)^2}{\left(\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.}\right)^2}\right\} \\
&+E\Big[\operatorname{tr}^{-3}(\mathbf{M}_i)(\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.})^{-2}\Big\{(C|\gamma_4|+2C|\gamma_2|+4C|\gamma_2|)(\hat{\epsilon}_{j.}'\mathbf{M}_i\hat{\epsilon}_{j.})^2 \\
&+2\gamma_1^2\left[\tau_T'\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\}\mathbf{\Pi}_i\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\}\tau_T\right]\Big\}\Big] \\
&+O\left(T^{-2}\right) \\
\leq&\left\{\frac{(3+|\gamma_2|)}{\operatorname{tr}^2(\mathbf{M}_i)}+\frac{3(3+|\gamma_2|)(2+|\gamma_2|)\operatorname{tr}(\mathbf{M}_i^2)}{\operatorname{tr}^4(\mathbf{M}_i)}+\frac{C|\gamma_4|+2C|\gamma_2|+4C|\gamma_2|}{\operatorname{tr}^3(\mathbf{M}_i)}\right\} \\
&\times\left[\frac{2\operatorname{tr}\left\{(\mathbf{M}_i\mathbf{M}_j)^2\right\}+\operatorname{tr}^2(\mathbf{M}_i\mathbf{M}_j)}{\operatorname{tr}^2(\mathbf{M}_j)}\right. \\
&\left.+\frac{\gamma_2\operatorname{tr}\{(\mathbf{R}'\mathbf{P}_j\mathbf{M}_i\mathbf{P}_j\mathbf{R})\circ(\mathbf{R}'\mathbf{P}_j\mathbf{M}_i\mathbf{P}_j\mathbf{R})\}}{\operatorname{tr}^2(\mathbf{M}_j)}+O\left(T^{-1}\right)\right] \\
&+2\gamma_1^2E\Big\{\frac{\tau_T'\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\}\mathbf{\Pi}_i\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\}\tau_T}{\operatorname{tr}^3(\mathbf{M}_i)(\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.})^2}\Big\} \\
&+O\left(T^{-2}\right) \\
=&O\left(T^{-2}\right),
\end{aligned}
$$

where the last equality holds because of the fact

$$
\begin{aligned}
&\frac{2\gamma_1^2\tau_T'\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\}\mathbf{\Pi}_i\left\{\mathbf{I}_T\circ(\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R})\right\}\tau_T}{(\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.})^2} \\
\leq&2\gamma_1^2\sum_{s=1}^{T}\sum_{k=1}^{T}\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.}}\Big)_{ss}(\mathbf{\Pi}_i)_{sk}\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.}}\Big)_{kk} \\
\leq&2\gamma_1^2O(1)\sum_{s=1}^{T}\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.}}\Big)_{ss}\sum_{k=1}^{T}\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j.}\hat{\epsilon}_{j.}'\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}_{j.}'\hat{\epsilon}_{j.}}\Big)_{kk}
\end{aligned}
$$

$$\leq 2\gamma_1^2 O(1)\mathrm{tr}^2\Big(\frac{\mathbf{R}'\mathbf{P}_i\hat{\epsilon}_{j\cdot}\hat{\epsilon}'_{j\cdot}\mathbf{P}_i\mathbf{R}}{\hat{\epsilon}'_{j\cdot}\hat{\epsilon}_{j\cdot}}\Big) = O(1). \tag{S4.42}$$

So, $C = O\left(N^{-2}T^{-2}\right) = o\left(\sigma_{S_N}^4\right)$. Next, we consider D.

$$E\left(\hat{\rho}_{i_1 j}^2 \hat{\rho}_{i_2 j}^2\right) \leq \sqrt{E\left(\hat{\rho}_{i_1 j}^4\right) E\left(\hat{\rho}_{i_2 j}^4\right)} = O\left(T^{-2}\right).$$

Then, $D = O\left(N^{-1}T^{-2}\right) = o\left(\sigma_{S_N}^4\right)$ and $\sum_{j=2}^N E\left(Z_j^4\right) = o\{\sigma_{S_N}^4\}$. Then, we complete the proof. □

S4.6 Proof of Lemma S.11

Recall that

$$\hat{\sigma}_{S_N}^2 = \frac{2}{N(N-1)} \sum\sum_{1\leq i<j\leq N} v_j'\left(v_i - \bar{v}_{ij}\right) v_i'\left(v_j - \bar{v}_{ij}\right),$$

where $\bar{v}_{ij} = \frac{1}{N-2}\sum_{1<\tau\neq i,j<N} v_\tau$ and $v_\tau = \hat{\epsilon}_{\tau\cdot}/\|\hat{\epsilon}_{\tau\cdot}\|$. We have

$$\begin{aligned}
\hat{\sigma}_{S_N}^2 =& \frac{2}{N(N-1)} \sum\sum_{1\leq i<j\leq N} v_i'\left(v_j - \bar{v}_{ij}\right) v_j'\left(v_i - \bar{v}_{ij}\right) \\
=& \frac{2}{N(N-1)} \sum\sum_{1\leq i<j\leq N} \left(v_i' v_j\right)^2 - \frac{2}{N(N-1)} \sum\sum_{1\leq i<j\leq N} v_i' v_j v_j' \bar{v}_{ij} \\
& - \frac{2}{N(N-1)} \sum\sum_{1\leq i<j\leq N} v_i' \bar{v}_{ij} v_j' v_i + \frac{2}{N(N-1)} \sum\sum_{1\leq i<j\leq N} v_i' \bar{v}_{ij} v_j' \bar{v}_{ij} \\
=& F_1 - F_2 - F_3 + F_4.
\end{aligned}$$

Obviously, $E\left(F_1\right) = \sigma_{S_N}^2\left\{1 + o(1)\right\}$ and $E\left(F_i\right) = 0$ for $i = 2, 3, 4$. Then, we just need to prove that $\mathrm{Var}\left(F_1\right) = o\left(\sigma_{S_N}^4\right)$ and

$$E\left(F_i^2\right) = o\left(\sigma_{S_N}^4\right)$$

for $i = 2, \cdots, 4$.

$$
\begin{aligned}
&\operatorname{Var}\left(F_1^2\right) \\
\leq &\frac{4}{N^2(N-1)^2} \sum\sum_{1 \leq i<j \leq N} E\left\{\left(v_i' v_j\right)^4\right\} \\
&+ \frac{8}{N^2(N-1)^2} \sum\sum\sum_{i_1<i_2<i_3} E\left\{\left(v_{i_1}' v_{i_2}\right)^2\left(v_{i_1}' v_{i_3}\right)^2\right\} \\
&+ \frac{8}{N^2(N-1)^2} \sum\sum\sum_{i_1<i_2<i_3} E\left\{\left(v_{i_1}' v_{i_3}\right)^2\left(v_{i_2}' v_{i_3}\right)^2\right\} \\
&+ \frac{8}{N^2(N-1)^2} \sum\sum\sum_{i_1<i_2<i_3} E\left\{\left(v_{i_1}' v_{i_2}\right)^2\left(v_{i_2}' v_{i_3}\right)^2\right\}.
\end{aligned}
$$

Because $E\left\{\left(v_{i_1}' v_{i_2}\right)^2\left(v_{i_1}' v_{i_3}\right)^2\right\} \leq \sqrt{E\left\{\left(v_{i_1}' v_{i_2}\right)^4\right\} E\left\{\left(v_{i_1}' v_{i_3}\right)^4\right\}}$ and

$$
E\left\{\left(v_{i_1}' v_{i_2}\right)^4\right\} = E\left(\hat{\rho}_{i_1 i_2}^4\right) = O\left(T^{-2}\right),
$$

we have $\operatorname{Var}\left(F_1^2\right) = O\left(N^{-1}T^{-2}\right) = o\left(\sigma_{S_N}^4\right)$. We next deal with F_2.

$$
F_2 = \frac{2}{N(N-1)} \sum\sum_{1 \leq i<j \leq N} v_i' v_j v_j' \bar{v}_{ij} = \frac{2}{N(N-1)(N-2)} \sum\sum_{1 \leq i<j \leq N} \sum_{\tau \neq i,j} v_i' v_j v_j' v_\tau.
$$

So,

$$
\begin{aligned}
&E\left(F_2^2\right) \\
=&\frac{4N^{-2}}{(N-1)^2(N-2)^2} \sum\sum_{1 \leq i_1<j_1 \leq N} \sum_{\tau_1 \neq i_1, j_1} \sum\sum_{1 \leq i_2<j_2 \leq N} \sum_{\tau_2 \neq i_2, j_2} E\left(v_{i_1}' v_{j_1} v_{j_1}' v_{\tau_1} v_{i_2}' v_{j_2} v_{j_2}' v_{\tau_2}\right) \\
=&\frac{4}{N^2(N-1)^2(N-2)^2} \sum\sum_{1 \leq i<j \leq N} \sum_{\tau \neq i,j} E\left(v_i' v_j v_j' v_\tau v_i' v_j v_j' v_\tau\right) \\
+&\frac{8}{N^2(N-1)^2(N-2)^2} \sum\sum\sum_{i<j_1<j_2} \sum_{\tau \neq i, j_1, j_2} E\left(v_i' v_{j_1} v_{j_1}' v_\tau v_i' v_{j_2} v_{j_2}' v_\tau\right).
\end{aligned}
$$

Because

$$E\left(v_i'v_jv_j'v_\tau v_i'v_jv_j'v_\tau\right) = E\left\{\left(v_i'v_jv_j'v_\tau\right)^2\right\} = E\left(\hat{\rho}_{ij}^2\hat{\rho}_{j\tau}^2\right) \leq \sqrt{E\left(\hat{\rho}_{ij}^4\right)E\left(\hat{\rho}_{j\tau}^4\right)}$$
$$= O\left(T^{-2}\right).$$

and

$$E\left(v_i'v_{j_1}v_{j_1}'v_\tau v_i'v_{j_2}v_{j_2}'v_\tau\right) \leq \sqrt{E\left\{\left(v_i'v_{j_1}v_{j_1}'v_\tau\right)^2\right\}E\left\{\left(v_i'v_{j_2}v_{j_2}'v_\tau\right)^2\right\}} = O\left(T^{-2}\right),$$

we have $E\left(F_2^2\right) = O\left(N^{-2}T^{-2}\right) = o\left(\sigma_{S_N}^4\right)$. Similarly, we have $F_3 = o_p\left(\sigma_{S_N}^2\right)$.

Now, we consider F_4,

$$\begin{aligned} F_4 =& \frac{2}{N(N-1)}\sum\sum_{1\leq i<j\leq N} v_i'\bar{v}_{ij}v_j'\bar{v}_{ij} \\ =& \frac{2}{N(N-1)(N-2)^2}\sum\sum_{1\leq i<j\leq N}\sum_{\tau_1\neq i,j}\sum_{\tau_2\neq i,j} v_i'v_{\tau_1}v_j'v_{\tau_2} \\ =& \frac{2}{N(N-1)(N-2)^2}\left\{\sum\sum_{1\leq i<j\leq N}\sum_{\tau\neq i,j} v_i'v_\tau v_j'v_\tau + \sum\sum_{1\leq i<j\leq N}\sum_{\tau_1\neq\tau_2\neq i,j} v_i'v_{\tau_1}v_j'v_{\tau_2}\right\} \\ =& F_{41} + F_{42}. \end{aligned}$$

So,

$$\begin{aligned} E\left(F_{41}^2\right) =& \frac{4}{N^2(N-1)^2(N-2)^4}\sum\sum_{1\leq i<j\leq N}\sum_{\tau_1,\tau_2\neq i,j} E\left(v_i'v_{\tau_1}v_j'v_{\tau_1}v_i'v_{\tau_2}v_j'v_{\tau_2}\right) \\ =& O\left(N^{-4}\right)\sqrt{E\left(v_i'v_{\tau_1}v_j'v_{\tau_1}\right)^2 E\left(v_i'v_{\tau_2}v_j'v_{\tau_2}\right)^2} \\ =& O\left(N^{-4}T^{-2}\right). \end{aligned}$$

Similarly, we have $E\left(F_{42}^2\right) = O\left(N^{-4}T^{-2}\right)$. Next, we can conclude that $E\left(F_4^2\right) = o\left(\sigma_{S_N}^4\right)$ due to $E\left(F_4^2\right) \leq 2E\left(F_{41}^2 + F_{42}^2\right)$, which leads to

$$\hat{\sigma}_{S_N}^2/\sigma_{S_N}^2 \to 1,$$

in probability. □

S4.7 Proof of Lemma S.12

By Theorem 1 and Lemma S.10, the following hold,

$$\frac{\mathrm{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N - 4\log N + \log\log N \to G(y) \text{ in distribution;} \tag{S4.43}$$

$$S_N/\hat{\sigma}_{S_N} \to \mathcal{N}(0,1) \text{ in distribution.} \tag{S4.44}$$

First, we will prove that if $\tilde{T}_{\max} - \log N + \log\log N$ and S_N/σ_{S_N} are asymptotically independent, then $\frac{\mathrm{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N - \log N + \log\log N$ and S_N/σ_{S_N} are also asymptotically independent. Due to (S3.9), (S3.13) and (S3.19), we have

$$\frac{\mathrm{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N - \tilde{T}_{\max} \to 0, \tag{S4.45}$$

in probability. Then, to show asymptotic independence, it is enough to show

$$\lim_{\min(N,T)\to\infty} P\Big(S_N/\sigma_{S_N} \leq x, \frac{\mathrm{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N - 4\log N + \log\log N \leq y\Big)$$
$$= \Phi(x)\cdot G(y), \tag{S4.46}$$

for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$, where $\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-t^2/2} dt$ and

$$G(y) = \exp\left\{-\frac{1}{\sqrt{8\pi}} \exp\left(-\frac{y}{2}\right)\right\}.(N-1)(N-2)^2$$

Let

$$a_N \doteq 4 \log N - \log \log N + y.$$

Due to (S4.43) and (S4.44), we know (S4.46) is equivalent to that

$$\lim_{\min(N,T)\to\infty} P\left(S_N/\sigma_{S_N} \le x, \frac{\operatorname{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N\right) = \Phi(x) \cdot \{1 - G(y)\}, \tag{S4.47}$$

for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$. By the assumption, we know that

$$\lim_{\min(N,T)\to\infty} P\left(S_N/\sigma_{S_N} \le x, \tilde{T}_{\max} > a_N\right) = \Phi(x) \cdot \{1 - G(y)\}, \tag{S4.48}$$

for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$. We show next that (S4.48) implies (S4.47). By (S4.45),

$$\left|\frac{\operatorname{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N - \tilde{T}_{\max}\right| \to 0$$

in probability. Given $\epsilon \in (0, 1)$. Set

$$\Omega = \left\{\left|\frac{\operatorname{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N - \tilde{T}_{\max}\right| < \epsilon\right\}.$$

Then

$$\lim_{\min(N,T)\to\infty} P(\Omega) = 1. \tag{S4.49}$$

Now,

$$
\begin{aligned}
&P\left(S_N/\sigma_{S_N} \le x, \frac{\operatorname{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N\right) \\
\le &P\left(S_N/\sigma_{S_N} \le x, \frac{\operatorname{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N, \Omega\right) + P\left(\Omega^c\right).
\end{aligned} \tag{S4.50}
$$

On Ω, if $\frac{\operatorname{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N$ then

$$
\tilde{T}_{\max} \ge \frac{\operatorname{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N - \left|\frac{\operatorname{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N - \tilde{T}_{\max}\right| > a_N - \epsilon. \tag{S4.51}
$$

Define

$$
\tilde{a}_N \doteq 4\log N - \log\log N + y - 2\epsilon,
$$

which makes sense for large T. Thus,

$$
a_N - \tilde{a}_N > \epsilon,
$$

and (S4.51) conclude that

$$
\tilde{T}_{\max} \ge \tilde{a}_N,
$$

as p is sufficiently large. Review (S4.50). We have

$$
\begin{aligned}
&P\left(S_N/\sigma_{S_N} \le x, \frac{\operatorname{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N\right) \\
\le &P\left(S_N/\sigma_{S_N} \le x, \tilde{T}_{\max} \ge \tilde{a}_N\right) + P\left(\Omega^c\right).
\end{aligned}
$$

Immediately from (S4.48) and (S4.49) we get

$$
\limsup_{\min(N,T)\to\infty} P\left(S_N/\sigma_{S_N} \le x, \frac{\operatorname{tr}^2(\tilde{\mathbf{\Sigma}})}{\|\tilde{\mathbf{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N\right) \le \Phi(x) \cdot \left\{1 - G(y - 2\epsilon)\right\},
$$

for any $\epsilon \in (0,1)$. Inspect that the left-hand side of the above does not depend on ϵ. Letting $\epsilon \downarrow 0$, we obtain

$$\limsup_{\min(N,T)\to\infty} P\left(S_N/\sigma_{S_N} \leq x, \frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N\right) \leq \Phi(x)\cdot\{1-G(y)\}\,, \tag{S4.52}$$

for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$. In the following we will show the lower limit. Evidently,

$$\begin{aligned} & P\left(S_N/\sigma_{S_N} \leq x, \frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N\right) \\ \geq & P\left(S_N/\sigma_{S_N} \leq x, \frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N, \Omega\right). \end{aligned} \tag{S4.53}$$

Set

$$\tilde{a}'_N \doteq 4\log N - \log\log N + y + 2\epsilon,$$

Therefore, $\tilde{a}'_N > a_N + \epsilon$. It is straightforward to verify that

$$\left\{\tilde{T}_{\max} > \tilde{a}'_N, \Omega\right\} \subset \left\{\tilde{T}_{\max} > a_N + \epsilon, \Omega\right\} \subset \left\{\frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N, \Omega\right\}, \tag{S4.54}$$

as N is sufficiently large, where the last inclusion follows from the definition of Ω. By (S4.53),

$$\begin{aligned} & P\left(S_N/\sigma_{S_N} \leq x, \frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N\right) \\ \geq & P\left(S_N/\sigma_{S_N} \leq x, \tilde{T}_{\max} > \tilde{a}'_N, \Omega\right). \end{aligned}$$

Thus, from (S4.48) and (S4.49) we get

$$\liminf_{\min(N,T)\to\infty} P\left(S_N/\sigma_{S_N} \le x, \frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N\right) \ge \Phi(x)\cdot\{1 - G(y+2\epsilon)\},$$

for any $\epsilon \in (0,1)$. Sending $\epsilon \downarrow 0$ we see

$$\liminf_{\min(N,T)\to\infty} P\left(S_N/\sigma_{S_N} \le x, \frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N > a_N\right) \ge \Phi(x)\cdot\{1 - G(y)\},$$

for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$.

We have proved that $\frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N - 4\log N + \log\log N$ and S_N/σ_{S_N} are asymptotically independent. Similarly, it is easy to prove that $\frac{\mathrm{tr}^2(\tilde{\boldsymbol{\Sigma}})}{\|\tilde{\boldsymbol{\Sigma}}\|_{\mathrm{F}}^2} L_N - 4\log N + \log N$ and $S_N/\hat{\sigma}_{S_N}$ are also asymptotically independent. Consequently, we complete the proof. □

S4.8 Proof of Lemma S.13

For I_l appeared in $H(N,k)$, write $I_l = (i_l, j_l)$ for $l = 1, \cdots, k$. Now we classify the indices $I_1 < I_2 < \cdots < I_k \in \Lambda_N$ in the definition of $H(N,k)$ into three cases. Let $\Gamma_{N,1}$ be the set of indices $(I_1, \cdots, I_k)$ such that no two of the $2k$ indices $\{i_l, j_l; l = 1, \cdots, k\}$ are identical. Let $\Gamma_{N,2}$ be the set of indices $(I_1, \cdots, I_k)$ such that either $i_1 = \cdots = i_k$ or $j_1 = \cdots = j_k$. Let $\Gamma_{N,3}$ be the set of indices $I_1 < I_2 < \cdots < I_k \in \Lambda_N$ excluding $\Gamma_{N,1} \cup \Gamma_{N,2}$. In the following we will estimate

$$D_j := \sum_{I_1<I_2<\cdots<I_k\in\Gamma_{N,j}} P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right)$$

for $j = 1, 2, 3$ one by one. We will see D_1 contributes essentially the sum in the expression of $H(N, k)$ by an easy argument; the term D_2 is negligible and its computation is trivial; the term D_3 is also negligible but its estimate is most involved.

Step 1: the estimate of D_1. Recall $B_I = \{|\epsilon_{i\cdot}' \epsilon_{j\cdot}| \geq l_N\}$ if $I = (i, j) \in \Lambda_N$. By the definition of $\Gamma_{N,1}$, we know that $B_{I_1}, B_{I_2} \cdots, B_{I_k}$ are independent. For large N,

$$\max_{I \in \Lambda_N} P(B_I) = P\{|\epsilon_{i\cdot}' \epsilon_{j\cdot}| \geq l_N\} \leq \frac{C}{N^2}$$

where the inequality holds due to (S3.7). Then,

$$D_1 \leq \frac{C}{N^{2k}} \cdot \binom{\frac{N(N-1)}{2}}{k} \leq \frac{C}{k!} \tag{S4.55}$$

Step 2: the estimate of D_2. Evidently, the size of $\Gamma_{N,2}$ is no more than $\binom{N}{1} \cdot \binom{N}{k} \cdot 2 \leq 2N^{k+1}$. Recall that $Z_{i\cdot}$ is the i-th row vector of matrix $\mathbf{Z}$. Similarly, there exists some $C_3 > 0$ such that for fixed $k > 0$ and $a_l \in \{-1, 1\}$, $1 \leq l \leq k$, that

$$P\left(\frac{\left(\sum_{l=1}^{k} a_l Z_{l\cdot}\right)' \mathbf{\Lambda}^2 \left(\sum_{l=1}^{k} a_l Z_{l\cdot}\right)}{\mathrm{tr}(\mathbf{\Lambda}^2)} > k(1 + \varepsilon_3)\right) \leq 2N^{-3k},$$

where $\varepsilon_3 = C_3\sqrt{\log N / \mathrm{tr}(\mathbf{\Lambda}^2)}$. Again, by Cramér type moderate deviation results in Statulevičius (1966) and Theorem 1.1 in Rudelson and Vershynin (2013), for $j_1 \neq j_2 \neq \cdots \neq j_k$,

$$P(B_{I_1} B_{I_2} \cdots B_{I_k})$$

$$
\begin{aligned}
=& P\left(|\epsilon_{i_1\cdot}'\epsilon_{j_1\cdot}| > l_N, \cdots, |\epsilon_{i_1\cdot}'\epsilon_{j_k\cdot}| > l_N\right) \\
\leq & \sum_{a_l\in\{-1,1\}} P\Big\{\Big|\epsilon_{i_1\cdot}'\Big(\sum_{l=1}^{k} a_l\epsilon_{j_l\cdot}\Big)\Big| > k\cdot l_N\Big\} \\
\leq & \sum_{a_l\in\{-1,1\}} P\Big\{\Big|Z_{i1\cdot}'\mathbf{\Lambda}\Big(\sum_{l=1}^{k} a_l Z_{j_l\cdot}\Big)\Big| > k\cdot\sqrt{a_N \mathrm{tr}(\mathbf{\Lambda}^2)}\Big\} \\
\leq & \sum_{a_l\in\{-1,1\}} P\Big\{\frac{\Big|Z_{i1\cdot}'\mathbf{\Lambda}\Big(\sum_{l=1}^{k} a_l Z_{j_l\cdot}\Big)\Big|}{\sqrt{\Big(\sum_{l=1}^{k} a_l Z_{j_l\cdot}\Big)'\mathbf{\Lambda}^2\Big(\sum_{l=1}^{k} a_l Z_{j_l\cdot}\Big)}} > \frac{\sqrt{ka_N}}{\sqrt{1+\varepsilon_3}}\Big\} \\
& + \sum_{a_l\in\{-1,1\}} P\Big\{\Big(\sum_{l=1}^{k} a_l Z_{j_l\cdot}\Big)'\mathbf{\Lambda}^2\Big(\sum_{l=1}^{k} a_l Z_{j_l\cdot}\Big) > k(1+\varepsilon_3)\mathrm{tr}(\mathbf{\Lambda}^2)\Big\} \\
\leq & \sum_{a_l\in\{-1,1\}} E\Big[E\Big\{\frac{\big|Z_{i1\cdot}'\mathbf{\Lambda}(\sum_{l=1}^{k} a_l Z_{j_l\cdot})\big|}{\sqrt{(\sum_{l=1}^{k} a_l Z_{j_l\cdot})'\mathbf{\Lambda}^2(\sum_{l=1}^{k} a_l Z_{j_l\cdot})}} > \frac{\sqrt{ka_N}}{\sqrt{1+\varepsilon_3}}|Z_{j_1\cdot}, \cdots, Z_{j_k\cdot}\Big\}\Big] \\
& + 2^{k+1}N^{-3k} \\
\leq & \sum_{a_l\in\{-1,1\}} \{1+o(1)\}\frac{2}{\sqrt{ka_N 2\pi}}\exp\left(-ka_N/2\right) + 2^{k+1}N^{-3k} \\
\leq & \frac{2^{k+1}(2\pi)^{-1/2}}{\sqrt{4k\log N}}\exp\left[-2k\log N + k\log\log N/2 - ky/2\right] + 2^{k+1}N^{-3k} \\
\leq & 2^{k+1}\left(\log N\right)^{k/2-1/2} N^{-2k}\exp(-ky/2) + 2^{k+1}N^{-3k},
\end{aligned} \tag{S4.56}
$$

where $ka_N = 4k\log N - k\log\log N + ky$. So, for $k \geq 2$, we have

$$
\begin{aligned}
D_2 &\leq 2N^{k+1}\cdot P\left(B_{I_1}B_{I_2}\cdots B_{I_k}\right) \\
&= 2N^{k+1}\cdot 2^{k+1}\left(\log N\right)^{k/2-1/2} N^{-2k}\exp(-ky/2) \to 0.
\end{aligned} \tag{S4.57}
$$

Step 3 : the estimate of D_3. Fix a tuple $(I_1, I_2, \cdots, I_l) \in \Gamma_{N,3}$. By the ordering imposed on Λ_N, we see that $i_1 \le i_2 \le \cdots \le i_k$. Let's assume that

$$i_1 = i_2 = \cdots = i_{n_1},$$

$$i_{n_1+1} = i_{n_1+2} = \cdots = i_{n_1+n_2},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$i_{n_1+\cdots+n_{r-1}+1} = i_{n_1+\cdots+n_{r-1}+2} = \cdots = i_{n_1+\cdots+n_r},$$

where $n_1+\cdots+n_r = k$. Let $\mathcal{F}_1$ be the set of random vectors $\{\epsilon_{j_1\cdot}, \epsilon_{i_l\cdot}, \epsilon_{j_l\cdot}; 2 \le l \le k\}$. Similar to (S4.56), for sufficiently large constant $C > 0$, we have

$$\begin{aligned}
& P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right) \\
=& E\left\{P\left(B_{I_1} B_{I_2} \cdots B_{I_k} \mid \mathcal{F}_1\right)\right\} \\
=& E\Bigg[P\bigg\{\min_{1\le l\le n_1} |\epsilon_{i_1\cdot}' \epsilon_{j_l\cdot}| \ge l_N \mid \mathcal{F}_1\bigg\} \cdot \prod_{l=n_1+1}^{k} I\left(B_{I_l}\right)\Bigg] \\
\le& E\Bigg[\sum_{a_l\in\{-1,1\}} P\bigg\{\Big|\epsilon_{i_1\cdot}'(\sum_{l=1}^{n_1} a_l \epsilon_{j_l\cdot})\Big| \ge l_N \mid \mathcal{F}_1\bigg\} \cdot \prod_{l=n_1+1}^{k} I\left(B_{I_l}\right)\Bigg] \\
\le& E\Bigg[\sum_{a_l\in\{-1,1\}} P\bigg\{\Big|\epsilon_{i_1\cdot}'(\sum_{l=1}^{n_1} a_l \epsilon_{j_l\cdot})\Big| \ge l_N \mid \mathcal{F}_1\bigg\} \\
& \quad \times I\bigg\{\frac{\big(\sum_{l=1}^{n_1} a_l Z_{j_l\cdot}\big)' \boldsymbol{\Lambda}^2 \big(\sum_{l=1}^{n_1} a_l Z_{j_l\cdot}\big)}{n_1 \mathrm{tr}(\boldsymbol{\Lambda}^2)} > 1 + C\sqrt{\frac{\log N}{\mathrm{tr}(\boldsymbol{\Lambda}^2)}}\bigg\} \prod_{l=n_1+1}^{k} I\left(B_{I_l}\right)\Bigg] \\
& + E\Bigg[I\bigg\{\frac{\big(\sum_{l=1}^{n_1} a_l Z_{j_l\cdot}\big)' \boldsymbol{\Lambda}^2 \big(\sum_{l=1}^{n_1} a_l Z_{j_l\cdot}\big)}{n_1 \mathrm{tr}(\boldsymbol{\Lambda}^2)} > 1 + C\sqrt{\frac{\log N}{\mathrm{tr}(\boldsymbol{\Lambda}^2)}}\bigg\} \prod_{l=n_1+1}^{k} I\left(B_{I_l}\right)\Bigg] \\
\le& O(N^{-2n_1}) E\Bigg[\prod_{l=n_1+1}^{k} I\left(B_{I_l}\right)\Bigg] + 2N^{-3k} \\
\le& O(N^{-2n_1}) P\left(B_{I_{n_1+1}} \cdots B_{I_k}\right) + 2N^{-3k}.
\end{aligned}$$

In the same way, we finally obtain

$$\begin{aligned}P\left(B_{I_1}B_{I_2}\cdots B_{I_k}\right)=&P\Big(\prod_{l=1}^{n_1}B_{I_l}\Big)\cdots P\Big(\prod_{l=n_{r-1}+1}^{n_r}B_{I_l}\Big)\\ \le&\frac{C}{N^{2n_1+\cdots+2n_r}}\\ \le&\frac{C}{N^{2k}},\end{aligned}$$

for some constant C and sufficiently large N. Recall $I_l=(i_l,j_l)$ for each $1\le l\le k$. In view of the definition of $\Gamma_{N,3}$, there are at least two of the $2k$ indices from $\{(i_l,j_l)\,;1\le l\le k\}$ are identical for any $(I_1,\cdots,I_k)\in\Gamma_{N,3}$. Let $\kappa=|\{i_l,j_l;1\le l\le k\}|$ for such $(I_1,I_2,\cdots,I_k)$. Easily, $k+1\le\kappa\le 2k-1$. To see how many such $(I_1,\cdots,I_k)$ with $|\ \{i_l,j_l;1\le\ l\le k\}\ |=\kappa$, first pick κ many indices from $\{1,2,\cdots,N\}$, which has the total number of ways $\binom{N}{\kappa}\le N^{\kappa}$, then use the κ many indices to make a $(I_1,\cdots,I_k)\in\Gamma_{N,3}$. The total number of ways to do so is no more than κ^{2k}. Therefore,

$$|\Gamma_{N,3}|\le\sum_{\kappa=k+1}^{2k-1}N^{\kappa}\cdot\kappa^{2k}\le(2k)^{2k}\cdot N^{2k-1}.$$

Similar to the (S4.56), for fixed k, we can have $P\left(B_{I_1}B_{I_2}\cdots B_{I_k}\right)=O(N^{-2k})$ and

$$\begin{aligned}D_3\le&(2k)^{2k}\cdot 2^k\left(\log N\right)^{k/2-1/2}N^{-2k}\exp(-ky/2)\\ \le&(2k)^{2k}\cdot 2^k\exp(-ky/2)\frac{\left(\log N\right)^{k/2-1/2}}{N}\to 0\end{aligned}$$

as N is sufficiently large. □

S4.9 Proof of Lemma S.14

For $I_1 < I_2 < \cdots < I_k \in \Lambda_N$, write $I_l = (i_l, j_l)$ for $l = 1, 2, \cdots, k$. Set

$$\Lambda_{N,k} = \{(i_l, j)\,; i_l < j \le N, 1 \le l \le k\} \bigcup \{(i, j_l)\,; 1 \le i < j_l, 1 \le l \le k\}$$

for $k \ge 1$. It is easy to check that $|\Lambda_{N,k}| = \sum_{l=1}^{k} (N - i_l + j_l - 2)$. since $i_l < j_l$ for each l, we see that

$$k(N-1) \le |\Lambda_{N,k}| \le \sum_{l=1}^{k} (N + j_l) \le 2kN.$$

Recall

$$A_N = \left\{ S_N / \sqrt{\sigma_{S_N}^2} \le x \right\} \quad x \in \mathbb{R},$$

for $N \ge 3$ and for $N \ge k \ge 1$,

$$S_N^k = \sqrt{\frac{2}{N(N-1)}} \sum_{(i,j) \in \Lambda_{N,k}} \hat{\rho}_{ij}.$$

Observe that $B_{I_1} B_{I_2} \cdots B_{I_k}$ is an event generated by random vectors $\left\{\hat{\rho}_{ij};\right.$ $\left.(i,j) \in \Lambda_{N,k}\right\}$. A crucial observation is that $S_N - S_N^k$ is independent of $B_{I_1} B_{I_2} \cdots B_{I_k}$. It is easy to see that

$$\begin{aligned} S_N^k = & \sqrt{\frac{2}{N(N-1)}} \sum\sum\nolimits_{\{(i_l,j); i_l < j \le N, 1 \le l \le k\}} \hat{\rho}_{i_l j} \\ & + \sqrt{\frac{2}{N(N-1)}} \sum\sum\nolimits_{\{(i,j_l); 1 \le i < j_l, 1 \le l \le k\}} \hat{\rho}_{i j_l} \\ & - \sqrt{\frac{2}{N(N-1)}} \sum_{s=1}^{k} \sum_{t=1}^{k} \hat{\rho}_{i_s j_t} \end{aligned}$$

$$\doteq D_{N,1} + D_{N,2} - D_{N,3}.$$

Fix $\epsilon \in (0,1)$, for even τ and $\tau > 2$, we have

$$E\left|D_{N,1}\right|^{\tau} \leq k^{\tau-1} \frac{2^{\tau/2}}{N^{\tau/2}(N-1)^{\tau/2}} \sum_{l=1}^{k} E\left(\left|\sum_{j=i_l+1}^{N} \hat{\rho}_{i_l j}\right|^{\tau}\right),$$

where

$$E\left\{\left(\sum_{j=i_l+1}^{N} \hat{\rho}_{i_l j}\right)^{\tau}\right\} = E\left[E\left\{\left(\sum_{j=i_l+1}^{N} \hat{\rho}_{i_l j}\right)^{\tau} \middle| \epsilon_{i_l\cdot}\right\}\right].$$

By Lemma 2 in Feng et al. (2022), there exists a constant $K_\tau > 0$ depending on τ only such that

$$E\left\{\left(\sum_{j=i_l+1}^{N} \hat{\rho}_{i_l j}\right)^{\tau} \middle| \epsilon_{i_l\cdot}\right\} \leq K_\tau (N-i_l)^{\tau/2-1} \sum_{j=i_l+1}^{N} E\left\{\left(\hat{\rho}_{i_l j}\right)^{\tau} \middle| \epsilon_{i_l\cdot}\right\}.$$

So, we have

$$E\left|D_{N,1}\right|^{\tau} \leq k^{\tau-1} \frac{2^{\tau/2}}{N^{\tau/2}(N-1)^{\tau/2}} \sum_{l=1}^{k} K_\tau (N-i_l)^{\tau/2-1} \sum_{j=i_l+1}^{N} E\left[E\left\{\left(\hat{\rho}_{i_l j}\right)^{\tau} \middle| \epsilon_{i_l\cdot}\right\}\right].$$

By Khintchine's inequality (Exercise 2.6.5 in Vershynin (2018)) and Lemma S.1, it's easy to obtain that $E\left[E\left\{\left(\hat{\rho}_{i_l j}\right)^{\tau} \middle| \epsilon_{i_l\cdot}\right\}\right] \leq (\tau/2)^{\tau/4} O\left(T^{-\tau/2}\right)$. Hence, together with $\sigma_{S_N} = O\left(T^{-1/2}\right)$, we obtain that

$$\begin{aligned} P\left(\left|D_{N,1}\right| > \sigma_{S_N}\epsilon\right) &\leq \frac{E\left|D_{N,1}\right|^{\tau}}{\sigma_{S_N}^{\tau}\epsilon^{\tau}} \\ &= \frac{K_\tau k^{\tau} \tau^{\tau/4} 2^{\tau/4}}{\epsilon^{\tau}} \cdot O\left(N^{-\tau/2}\right), \end{aligned}$$

where the last equality holds due to Assumption 2. Similarly,

$$P\left(\left|D_{N,2}\right| > \sigma_{S_N}\epsilon\right) \leq \frac{k^{\tau} \tau^{\tau/4} 2^{\tau/4}}{\epsilon^{\tau}} \cdot O\left(N^{-\tau/2}\right).$$

Lastly, for even τ and $\tau > 2$,

$$\begin{aligned} E\left(|D_{N,3}|^{\tau}\right) \leq & \frac{2^{\tau/2}}{N^{\tau/2}(N-1)^{\tau/2}} \cdot k^{2(\tau-1)} \cdot \sum_{s=1}^{k} \sum_{t=1}^{k} E\left(|\hat{\rho}_{i_s j_t}|^{\tau}\right) \\ \leq & \frac{2^{\tau/2}}{N^{\tau/2}(N-1)^{\tau/2}} \cdot k^{2\tau} E\left(|\hat{\rho}_{i_s j_t}|^{\tau}\right). \end{aligned}$$

So, we have

$$P\left\{\frac{|S_N^k|}{\sigma_{S_N}} \geq \epsilon\right\} \leq C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}},$$

for large N, where C' is a constant depending on ϵ but free of N.

Fix $I_1 < I_2 < \cdots < I_k \in \Lambda_N$. By the definition of A_N,

$$\begin{aligned} & P\left\{A_N(x) B_{I_1} B_{I_2} \cdots B_{I_k}\right\} \\ \leq & P\left\{A_N(x) B_{I_1} B_{I_2} \cdots B_{I_k}, \frac{\left|S_N^k\right|}{\sigma_{S_N}} < \epsilon\right\} + C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}} \\ \leq & P\left\{\frac{S_N - S_N^k}{\sigma_{S_N}} \leq x + \epsilon, B_{I_1} B_{I_2} \cdots B_{I_k}\right\} + C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}} \\ = & P\left\{\frac{S_N - S_N^k}{\sigma_{S_N}} \leq x + \epsilon\right\} \cdot P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right) + C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}}, \end{aligned}$$

by the independence between $S_N - S_N^k$ and $B_{I_1} B_{I_2} \cdots B_{I_k}$. Now

$$\begin{aligned} & P\left\{\frac{S_N - S_N^k}{\sigma_{S_N}} \leq x + \epsilon\right\} \\ \leq & P\left\{\frac{S_N - S_N^k}{\sigma_{S_N}} \leq x + \epsilon, \frac{\left|S_N^k\right|}{\sigma_{S_N}} < \epsilon\right\} + C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}} \\ \leq & P\left\{\frac{S_N}{\sigma_{S_N}} \leq x + 2\epsilon\right\} + C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}} \\ \leq & P\left\{A_N(x + 2\epsilon)\right\} + C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}}. \end{aligned}$$

Combing the two inequalities to get

$$\begin{aligned} &P\left\{A_N(x) B_{I_1} B_{I_2} \cdots B_{I_k}\right\} \\ \leq &P\left\{A_N(x+2\epsilon)\right\} \cdot P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right)+2 C^{\prime} \cdot \frac{k^{2\tau} \tau^{\tau/4}}{N^{\tau/2}}. \end{aligned} \tag{S4.58}$$

Similarly,

$$\begin{aligned} &P\left\{\frac{S_N-S_N^k}{\sigma_{S_N}} \leq x-\epsilon, B_{I_1} B_{I_2} \cdots B_{I_k}\right\} \\ \leq &P\left\{\frac{S_N-S_N^k}{\sigma_{S_N}} \leq x-\epsilon, B_{I_1} B_{I_2} \cdots B_{I_k},\right. \\ &\left.\frac{\left|S_N^k\right|}{\sigma_{S_N}}<\epsilon\right\}+C^{\prime} \cdot \frac{k^{2\tau} \tau^{\tau/4}}{N^{\tau/2}} \\ \leq &P\left\{\frac{S_N}{\sigma_{S_N}} \leq x, B_{I_1} B_{I_2} \cdots B_{I_k}\right\}+C^{\prime} \cdot \frac{k^{2\tau} \tau^{\tau/4}}{N^{\tau/2}}. \end{aligned}$$

In other words, by independence,

$$\begin{aligned} &P\left\{A_N(x) B_{I_1} B_{I_2} \cdots B_{I_k}\right\} \\ \geq &P\left\{\frac{S_N-S_N^k}{\sigma_{S_N}} \leq x-\epsilon\right\} \cdot P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right)-C^{\prime} \cdot \frac{k^{2\tau} \tau^{\tau/4}}{N^{\tau/2}}. \end{aligned}$$

Furthermore,

$$\begin{aligned} &P\left\{\frac{S_N}{\sigma_{S_N}} \leq x-2\epsilon\right\} \\ \leq &P\left\{\frac{S_N}{\sigma_{S_N}} \leq x-2\epsilon, \frac{\left|S_N^k\right|}{\sigma_{S_N}}<\epsilon\right\}+C^{\prime} \cdot \frac{k^{2\tau} \tau^{\tau/4}}{N^{\tau/2}} \\ \leq &P\left\{\frac{S_N-S_N^k}{\sigma_{S_N}} \leq x-\epsilon\right\}+C^{\prime} \cdot \frac{k^{2\tau} \tau^{\tau/4}}{N^{\tau/2}}. \end{aligned} \tag{S4.59}$$

The above two strings of inequalities imply

$$P\left\{A_N(x) B_{I_1} B_{I_2} \cdots B_{I_k}\right\}$$

$$\geq P\left\{\frac{S_N}{\sigma_{S_N}} \leq x - 2\epsilon\right\} \cdot P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right) - 2C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}},$$

which joining with (S4.58) yields

$$\begin{aligned}
&\left|P\left\{A_N(x) B_{I_1} B_{I_2} \cdots B_{I_k}\right\} - P\left\{A_N(x)\right\} \cdot P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right)\right| \\
\leq &\Delta_{N,\epsilon} \cdot P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right) + 4C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}},
\end{aligned}$$

where

$$\Delta_{N,\epsilon} \doteq \left|P\left\{A_N(x)\right\} - P\left\{A_N(x+2\epsilon)\right\}\right| + \left|P\left\{A_N(x)\right\} - P\left\{A_N(x-2\epsilon)\right\}\right|.$$

In particular,

$$\Delta_{N,\epsilon} \to |\Phi(x+2\epsilon) - \Phi(x)| + |\Phi(x-2\epsilon) - \Phi(x)|, \tag{S4.60}$$

as $\min(N,T) \to \infty$ by Theorem 1. As a consequence,

$$\begin{aligned}
\zeta(N,k) \doteq &\sum_{I_1<I_2<\cdots<I_k \in \Lambda_N} \left[P\left(A_N(x) B_{I_1} B_{I_2} \cdots B_{I_k}\right) - \right. \\
&\left. P\left\{A_N(x)\right\} \cdot P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right)\right] \\
\leq &\sum_{I_1<I_2<\cdots<I_k \in \Lambda_N} \left\{\Delta_{N,\epsilon} \cdot P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right) + 4C' \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}}\right\} \\
\leq &\Delta_{N,\epsilon} \cdot H(N,k) + (4C') \cdot \binom{\frac{1}{2}N(N-1)}{k} \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}},
\end{aligned}$$

where

$$H(N,k) = \sum_{I_1<I_2<\cdots<I_k \in \Lambda_k} P\left(B_{I_1} B_{I_2} \cdots B_{I_k}\right)$$

as defined in Lemmas S.13 and S.14, we know $\limsup_{\min(N,T)\to\infty} H(N,k) \leq C/k!$, where C is a universal constant. Picking $\tau = 6k$, and using the trivial fact $\binom{r}{i} \leq r^i$ for any integers $1 \leq i \leq r$, we have that

$$\binom{\frac{1}{2}N(N-1)}{k} \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}} \leq N^{2k} \cdot \frac{k^{2\tau}\tau^{\tau/4}}{N^{\tau/2}} \leq \frac{k^{2\tau}\tau^{\tau/4}}{N^k} \to 0.$$

Hence, from (S4.60)

$$\begin{aligned}\limsup_{\min(N,T)\to\infty} \zeta(N,k) &\leq \frac{C}{k!} \cdot \limsup_{\min(N,T)\to\infty} \Delta_{N,\epsilon} \\ &= \frac{C}{k!} \cdot \left\{|\Phi(x+2\epsilon) - \Phi(x)| + |\Phi(x-2\epsilon) - \Phi(x)|\right\},\end{aligned}$$

for any $\epsilon > 0$. The desired result follows by sending $\epsilon \downarrow 0$. □

References

Abadir, K. M. and J. R. Magnus (2005). *Matrix algebra*, Volume 1. Cambridge University Press.

Arratia, R., L. Goldstein, and L. Gordon (1989). Two moments suffice for poisson approximations: the chen-stein method. *The Annals of Probability 17*(1), 9–25.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica 77*(4), 1229–1279.

Baltagi, B. H., C. Kao, and B. Peng (2016). Testing cross-sectional correlation in large panel data models with serial correlation. *Econometrics 4*(4), 44.

Bao, Y. and A. Ullah (2010). Expectation of quadratic forms in normal and nonnormal variables with applications. *Journal of Statistical Planning and Inference 140*(5), 1193–1205.

Bernard, V. L. (1987). Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research 25*(1), 1–48.

Bernstein, D. S. (2009). *Matrix mathematics: theory, facts, and formulas.* Princeton university press.

Chen, X. and W. Liu (2018). Testing independence with high-dimensional correlated samples. *The Annals of Statistics 46*(2), 866–894.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Feng, L., T. Jiang, B. Liu, and W. Xiong (2022). Max-sum tests for cross-sectional dependence of high-demensional panel data. *The Annals of Statistics 50*(2), 1124–1143.

Hall, P. and C. C. Heyde (2014). *Martingale limit theory and its application.* Academic press.

Lam, C., Q. Yao, and N. Bathia (2011). Estimation of latent factors for high-dimensional time series. *Biometrika 98*(4), 901–918.

Li, Z., C. Lam, J. Yao, and Q. Yao (2019). On testing for high-dimensional white noise. *The Annals of Statistics 47*(6), 3382–3412.

Lieberman, O. (1994). A laplace approximation to the moments of a ratio of quadratic forms. *Biometrika 81*(4), 681–690.

Pesaran, M. H. and T. Yamagata (2012). Testing capm with a large number of assets. In *AFA 2013 San Diego Meetings Paper.*

Rosenberg, B. and A. Rudd (1982). Factor-related and specific returns of common stocks: Serial correlation and market inefficiency. *The Journal of Finance 37*(2), 543–554.

Rudelson, M. and R. Vershynin (2013). Hanson-wright inequality and sub-gaussian concentration. *Electronic Communications in Probability 18*, 1–9.

Saulis, L. and V. Statulevicius (1991). *Limit theorems for large deviations*, Volume 73. Springer Science & Business Media.

Schwartz, R. A. and D. K. Whitcomb (1977). Evidence on the presence and causes of serial correlation in market model residuals. *Journal of financial and quantitative analysis 12*(2), 291–313.

Statulevičius, V. (1966). On large deviations. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 6*(2), 133–144.

Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge University Press.