Statistics and Probability Letters 182 (2022) 109304

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Robust sphericity test in the panel data model

Xiaoxu Zhang [a], Ping Zhao [b], Long Feng [b,*]

[a] School of Mathematics and Statistics & KLAS, Northeast Normal University, Changchun 130024, Jilin Province, China
[b] School of Statistics and Data Science, LPMC and KLMDASR, Nankai University, Tianjin, China

ARTICLE INFO

Article history:
Received 8 December 2019
Received in revised form 3 November 2021
Accepted 9 November 2021
Available online 18 November 2021

MSC:
62G05
62G08
62G20

Keywords:
Cross-sectional dependence
High dimension
Multivariate-sign
Panel data
Sphericity

ABSTRACT

This paper proposes a test for the null of sphericity in the panel data model. We proposed a novel multivariate sign test for sphericity based on sample splitting and leave out method in the panel data model. The limiting distribution of the proposed test statistic is derived under the null and alternative hypothesis. Simulation studies also demonstrate the advantage of our method.

1. Introduction

Testing for sphericity is an important work in the fixed effects panel data model, see an overview in Baltagi et al. (2011). Using the Random Matrix Theory-based approach of Ledoit and Wolf (2002), Baltagi et al. (2011) propose a test for the null of sphericity of the remainder disturbances with large dimension N and sample sizes T. To consider the trace of higher-order of the covariance matrix, Mao (2014) extends a new sphericity test recently developed by Fisher et al. (2010) to the fixed effects panel data model. Both the above two methods need the normality assumption of the disturbances. So Baltagi et al. (2015) extend Chen et al. (2010)'s test and proposed a new test without assuming normality of the disturbances. However, the above three tests can only be used in fixed effects panel data model and perform poorly if the disturbances are from the heavy-tailed distributions, such as the multivariate t-distribution, the multivariate mixture normal distributions.

In this article, we consider testing for sphericity in a more general model—panel data model. We allow the slope parameters β_i to vary across i. Under the panel data model, there are two main drawbacks of the above three test procedures. First, if we directly use the above three test procedures in the panel data model, there would be a non-negligible bias term in their test statistics, even under the normality assumption. It is not surprised because the coefficient β_i is only $\sqrt{T}$-consistent in the panel data model, while β is $\sqrt{NT}$-consistent in the fixed effect panel data model. Second, the above three tests perform poorly if the disturbances are from the heavy-tailed distributions, such as the multivariate t-distribution, the multivariate mixture normal distributions. Both the above two distributions do not satisfy the distribution Assumption 1 in Baltagi et al. (2015), see Appendix 3 in Zou et al. (2014).

* Corresponding author.
E-mail address: flnankai@gmail.com (L. Feng).

https://doi.org/10.1016/j.spl.2021.109304

To overcome the above two issues, we adopt the multivariate-sign method to construct a robust test for sphericity in the panel data model. The multivariate sign- and rank-based methods are very popular in the classic statistics (Oja, 2010). In an important work, Zou et al. (2014) proposed a multivariate-sign-based high-dimensional test for sphericity for the raw data. They show that the multivariate sign based methods also are very robust and effective in handling non-normal data in high-dimensional settings. A natural method to extend Zou et al. (2014)'s method is replacing the disturbances with its estimators. However, there would be a non-negligible bias term in the corresponding testing statistics when the dimension N gets larger, which will limit the scope of application. So we adopt the sample-splitting and leave-out method and propose a novel test procedure for the sphericity in the panel data model. In the theoretical analysis, we show that the proposed test statistic is not biased under the null hypothesis. The asymptotic distributions of this test statistic under the null and alternative hypothesis are derived. Simulation results suggest that the proposed tests outperform the other tests under the heavy-tailed distributions.

The rest of the article is organized as follows. We introduce the proposed test for the panel data model in Section 2. The numerical performance of the proposed test is demonstrated in Section 3. Finally, we relegate the technical proofs to the Appendix.

2. Our method

Here we state the problem of testing for sphericity in the panel data model. We consider the following heterogeneous panel data model:

$$y_{it} = x_{it}^\top \beta_i + u_{it}, \ i = 1, \ldots, N, \ t = 1, \ldots, T, \tag{2.1}$$

where i indexes the cross-sectional units and t the time series. y_{it} is the dependent variable, and x_{it} is the exogenous regressors of dimension $p \times 1$ with slope parameters β_i that is allowed to vary across i. The error u_{it} is allowed to be cross-sectionally dependent but is uncorrelated with x_{it}. Let $\mathbf{x}_i = (x_{i1}, \ldots, x_{iT})^\top$ and $\mathbf{y}_i = (y_{i1}, \ldots, y_{iT})^\top$. Let $\boldsymbol{u}_t = (u_{1t}, \ldots, u_{Nt})^\top$ and assume that $\boldsymbol{u}_t$'s are i.i.d. over time t. And the covariance matrix of $\boldsymbol{u}_t$ is $\boldsymbol{\Sigma}$. The null hypothesis of interest is sphericity:

$$H_0 : \boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_N, \quad \text{vs} \quad H_1 : \boldsymbol{\Sigma} \neq \sigma^2 \mathbf{I}_N. \tag{2.2}$$

The alternative hypothesis allows cross-sectional dependence or heteroskedasticity or both. To the best of our knowledge, there are no methods of testing sphericity for the panel data model.

For testing the sphericity of the variance–covariance matrix of the disturbances, Zou et al. (2014) proposed a test statistic based on the multivariate-sign method, i.e.

$$T_Z = \frac{2N}{T(T-1)} \sum_{1 \le t_1 < t_2 \le T} (U(\boldsymbol{u}_{t_1})^\top U(\boldsymbol{u}_{t_2}))^2 - 1.$$

They showed that the above procedure has good size and power for a wide range of dimensions, sample sizes and distributions. So, building upon the work of Zou et al. (2014), a natural idea is replacing the disturbances $\boldsymbol{u}_t$ with its estimators $\hat{\boldsymbol{u}}_t = (\hat{u}_{i1}, \ldots, \hat{u}_{iN})$ where $\hat{u}_{it} = y_{it} - x_{it}\hat{\beta}_i$ and $\hat{\beta}_i$ is the corresponding least-square estimator of β_i, i.e.

$$\tilde{T}_Z = \frac{2N}{T(T-1)} \sum_{1 \le t_1 < t_2 \le T} (U(\hat{\boldsymbol{u}}_{t_1})^\top U(\hat{\boldsymbol{u}}_{t_2}))^2 - 1.$$

However, due to the only $\sqrt{T}$-consistency of the estimator $\hat{\beta}_i$ and the non-independence between the two estimators $\hat{\boldsymbol{u}}_{t_1}$ and $\hat{\boldsymbol{u}}_{t_2}$, there would be a non-negligible bias term in $\tilde{T}_Z$ when the dimension N gets larger than T. So to avoid this bias term in $\tilde{T}_Z$, we adopt the sample-splitting and leave-out method in our new test procedure.

Define the set $A_{t_1t_2} = \{1, \ldots, T\} \setminus \{t_1, t_2\}$. And the first half subset of $A_{t_1t_2}$ is $A_{1t_1t_2}$, the second half subset of $A_{t_1t_2}$ is $A_{2t_1t_2}$. So $|A_{1t_1t_2}| = |A_{2t_1t_2}| = \frac{T-2}{2}$ and $A_{1t_1t_2} \cap A_{2t_1t_2} = \emptyset$. $\hat{\boldsymbol{\beta}}_{1k(t_1,t_2)}$ is the corresponding least square estimator of $\boldsymbol{\beta}_k$ based on the sample $\{(x_{kt}, y_{kt})\}_{t \in A_{1t_1t_2}}$, $k = 1, \ldots, N$, respectively. $\hat{\boldsymbol{\beta}}_{2k(t_1,t_2)}$ is the corresponding least square estimator of $\boldsymbol{\beta}_k$ based on the sample $\{(x_{kt}, y_{kt})\}_{t \in A_{2t_1t_2}}$, $k = 1, \ldots, N$, respectively. We propose the following test statistic

$$J_S = \frac{2N}{T(T-1)} \sum_{1 \le t_1 < t_2 \le T} (\tilde{\mathbf{U}}_{t_1,t_2}^\top \tilde{\mathbf{U}}_{t_2,t_1})^2 - 1$$

where $\tilde{\mathbf{U}}_{t_1,t_2} = U(\tilde{\boldsymbol{u}}_{t_1,t_2})$, $\tilde{\mathbf{U}}_{t_2,t_1} = U(\tilde{\boldsymbol{u}}_{t_2,t_1})$. $\tilde{\boldsymbol{u}}_{t_1,t_2} = (\tilde{u}_{1t_1}, \ldots, \tilde{u}_{Nt_1})$, $\tilde{u}_{kt_1} = y_{kt_1} - x_{kt_1}^\top \tilde{\boldsymbol{\beta}}_{1k(t_1,t_2)}$, $\tilde{\boldsymbol{u}}_{t_2,t_1} = (\tilde{u}_{1t_2}, \ldots, \tilde{u}_{Nt_2})$, $\tilde{u}_{kt_2} = y_{kt_2} - x_{kt_2}^\top \tilde{\boldsymbol{\beta}}_{2k(t_1,t_2)}$. In the new test statistics J_S, $\tilde{\mathbf{U}}_{t_1,t_2}$ are independent of $\tilde{\mathbf{U}}_{t_2,t_1}$. So the expectation of J_S would be negligible to the standard deviation of J_S, i.e. $E(J_S) = o(\sqrt{\mathrm{var}(J_S)})$, under the null hypothesis. And then there would be no bias term in our test statistics under the null hypothesis.

To facilitate our analysis, we require the following assumptions:

(A1) The error vectors $\{\boldsymbol{u}_1, \dots \boldsymbol{u}_T\}$ are independently and identically distributed (i.i.d.) from the N-variate mean zero elliptical distribution with probability density function:

$$\det(\boldsymbol{\Sigma})^{-1/2} g_N(\|\boldsymbol{\Sigma}^{-1/2}\boldsymbol{u}\|), \tag{2.3}$$

where $\|\boldsymbol{u}\| = (\boldsymbol{u}^T\boldsymbol{u})^{1/2}$ is the Euclidean length of the vector $\boldsymbol{u}$, $\boldsymbol{\Sigma} \in \mathcal{R}^{N\times N}$ is the symmetric positive define scatter matrix. The moments $E(r_t^{-k})$ for $k = 1, \dots, 4$ exist for large enough N where $r_t = \|\boldsymbol{u}_t\|$. And for $k = 2, 3, 4$, $E(r_t^{-k})/(E(r_t^{-1}))^k \to d_k \in [1, +\infty)$ as $N \to \infty$ where d_k are constants.

(A2) The regressors x_{it}, $i = 1, \dots, N$, $t = 1, \dots, T$ are independent of the idiosyncratic disturbances u_{it}, $i = 1, \dots, N$, $t = 1, \dots, T$. The regressors x_{it} have finite fourth moments.

The above assumptions are very common. Assumption (A1) is the same as Assumption 1 in Zou et al. (2014). Assumption (A2) is the same as Assumption 2 in Baltagi et al. (2015).

Theorem 2.1. *Under Assumptions (A1)–(A2) and H_0, as $N \to \infty$ and $T \to \infty$, we have $J_S/\sigma_{T1} \to^d N(0, 1)$, where $\sigma_{T1}^2 = 4/T(T-1)$.*

According to Theorem 2.1, we will reject the null hypothesis at the significant level α if $J_S/\sigma_{T1} > z_\alpha$ where z_α is the $1-\alpha$ quantile of the standard normal distribution $N(0, 1)$.

Next, we consider the asymptotic distribution of J_S under the alternative $H_1 : \boldsymbol{\Lambda}_N = I_N + D_{N,T}$ where $\boldsymbol{\Lambda} = \frac{1}{\mathrm{tr}(\boldsymbol{\Sigma})}\boldsymbol{\Sigma}$. Define $\sigma_{T2}^2 = \sigma_{T1}^2 + T^{-2}N^{-2}\{8N\mathrm{tr}(D_{N,T}^2) + 4\mathrm{tr}^2(D_{N,T}^2)\} + T^{-1}N^{-2}8\mathrm{tr}(\boldsymbol{\Lambda}_N^4)$.

Theorem 2.2. *Suppose that $T\mathrm{tr}(D_{N,T}^2)/N = O(1)$. Under H_1 and (A1)–(A2), as $N \to \infty$ and $T \to \infty$,we have $\{J_S - N^{-1}\mathrm{tr}(D_{N,T}^2)\}/\sigma_{T2} \to^d N(0, 1)$,*

According to Corollary 1 in Zou et al. (2014), our proposed test is consistent. If $T\mathrm{tr}(D_{N,T}^2)/N \to \infty$, $P(J_S/\sigma_{T1} > z_\alpha) \to 1$. In the fixed effect panel data model, Baltagi et al. (2015) showed that the power function of J_u is the same as Chen et al. (2010). By Corollary 2 in Zou et al. (2014), our test J_S is asymptotically as efficient as J_u under the normal distributions.

3. Simulation

We consider the data generating process used in Pesaran et al. (2008), which is specified as

$$y_{it} = \alpha_i + \sum_{l=1}^{p} x_{lit}\beta_i + u_{it},\ i = 1, \dots, N,\ t = 1, \dots, T, \tag{3.4}$$

where $\alpha_i \sim N(0, 1)$, $\beta_i \sim N(1, 0.04)$. The covariates are generated as

$$x_{lit} = 0.6x_{lit-1} + v_{lit},\ i = 1, \dots, N,\ t = -50, \dots, T, l = 2, \dots, p,$$

with $x_{li,-51} = 0$, where $v_{lit} \sim N(0, \zeta_{li}^2/(1-0.6^2))$, $\zeta_{li}^2 \sim \chi^2(6)/6$. Here we choose $p = 3$. We consider five models for the errors u_{it}:

(I) Multivariate normal distribution: $\boldsymbol{u}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma})$;
(II) $\boldsymbol{u}_t = \boldsymbol{\Sigma}^{1/2}\mathbf{z}_t$, $\mathbf{z}_t = (z_{1t}, \dots, z_{Nt})$, $z_{it} \sim t(5)/\sqrt{5/3}$;
(III) $\boldsymbol{u}_t = \boldsymbol{\Sigma}^{1/2}\mathbf{z}_t$, $\mathbf{z}_t = (z_{1t}, \dots, z_{Nt})$, $z_{it} \sim (\chi_4^2 - 4)/\sqrt{8}$;
(IV) Multivariate t-distribution: $\boldsymbol{u}_t \sim t_N(\mathbf{0}, \boldsymbol{\Sigma}, 4)$;
(V) Multivariate mixture normal distribution: $\boldsymbol{u}_t$'s are generated from $\gamma N_N(\mathbf{0}, \boldsymbol{\Sigma}) + (1-\gamma)N_N(\mathbf{0}, 9\boldsymbol{\Sigma})$, denoted by $\mathrm{MN}_{N,\gamma,9}(\mathbf{0}, \boldsymbol{\Sigma})$; γ is fixed to be 0.8.

First, to show the necessity of sample-splitting and leave-out, we summarize the simulation results using the mean–standard deviation-ratio (MDR) $E(T)/\sqrt{\mathrm{var}(T)}$ under the null hypothesis, $\boldsymbol{\Sigma} = \mathbf{I}_N$. Because the explicit forms of $E(T)$ and $\mathrm{var}(T)$ are difficult to calculate, we estimate them by simulation. Table 1 shows the MDR of $\tilde{T}_Z$ and J_S with different sample sizes and dimensions. We observe that the MDR of J_S is close to zero, while the MDR of $\tilde{T}_Z$ is significantly larger than zero. So there is a non-negligible bias term of $\tilde{T}_Z$.

Next, we choose $\boldsymbol{\Sigma} = (0.4^{|i-j|})_{1\le i,j\le N}$ under the alternative hypothesis. We compare our test with the tests proposed by Baltagi et al. (2011) (abbreviated as J_{BFK}), Mao (2014) (abbreviated as J_M), and Baltagi et al. (2015) (abbreviated as J_u). Table 2 summarizes the results of each test. Our test J_S also controls the empirical sizes very well in all cases. However, the empirical sizes of J_{BFK} and J_u are larger than the nominal level in all cases. It is not surprising because there is a non-negligible bias term in their test statistics in the panel data model. And the J_M test also cannot control the empirical sizes, even under the normal distributions — Scenario (I). We also proposed a size-corrected power comparison for these tests. In the size-corrected power comparison procedure, the critical values of each test are calculated by simulation under the null hypothesis. In Scenarios (I)–(III), the power of our test J_S is a little less than J_{BFK} and J_u. However, our test J_S outperforms the other three tests in Scenarios (IV)–(V). It shows that our test procedure is very robust and efficient in testing sphericity in the panel data model.

Table 1
The mean–standard deviation-ratio (MDR) of tests under scenarios (I)–(V).

	$T=30$						$T=50$					
	J_S			$\tilde{T}_Z$			J_S			$\tilde{T}_Z$		
N	50	100	200	50	100	200	50	100	200	50	100	200
(I)	0.03	0.04	0.03	1.2	2.24	4.48	−0.08	−0.02	0.01	0.55	1.22	2.48
(II)	−0.04	−0.06	0.06	1.4	2.52	4.83	0.02	−0.04	−0.02	0.8	1.36	2.64
(III)	−0.06	0.05	−0.01	1.31	2.61	4.84	−0.04	0.02	−0.01	0.82	1.48	2.66
(IV)	−0.04	−0.03	0.02	1.87	3.70	6.80	0.03	0.01	0.01	1.18	2.19	4.39
(V)	0.06	0.01	−0.01	2.03	3.91	7.75	−0.01	−0.04	−0.07	1.17	2.34	4.70

Table 2
Sizes and size adjusted power of tests for panel data model under scenarios (I)–(V).

	Size						Size Adjusted Power					
	$T=30$			$T=50$			$T=30$			$T=50$		
N	50	100	200	50	100	200	50	100	200	50	100	200
	(I)											
J_S	5.3	4.5	5.9	5.5	5.5	6.3	78.6	83.4	83.7	99.8	99.9	100
J_{BFK}	11.1	15.6	29.3	7.7	8.7	12.5	97.4	98.6	98.4	100	100	100
J_u	10.5	15.5	30	7.2	8.6	12.4	96.1	98.3	98.2	100	100	100
J_M	3.9	8.5	7.5	5.1	5.5	3.6	53.1	34.2	26.6	97.1	87.3	77.2
	(II)											
J_S	5.7	3.9	4.9	5.0	5.7	4.6	77.2	80.2	83.4	99.7	100	100
J_{BFK}	58.1	68.8	83.3	56.9	64.9	73.5	54.6	68.1	70	99.2	99.9	100
J_u	22.3	27.4	41	13.8	15.5	19	88.9	93.7	93.8	100	100	100
J_M	30.1	32.5	26.3	42.1	40.5	42.6	20.1	14.5	5.6	4.4	8.8	15.9
	(III)											
J_S	5.1	5.8	5.4	5.7	5.4	4.5	77.8	82.5	84.7	100	100	100
J_{BFK}	48.4	57.1	73.7	48	49.9	58.5	88.6	93.6	95.4	100	100	100
J_u	18.1	22	39.6	11.5	13.3	18.3	92.1	95.5	96.9	100	100	100
J_M	15.2	12.9	10.4	29.2	23.6	17.5	47.8	38.7	31.1	86.5	70.2	54.1
	(IV)											
J_S	5.8	4.8	4.7	5.6	4.8	4.8	66	69.6	73.1	98.9	99.7	99.7
J_{BFK}	100	100	100	100	100	100	4.8	3.5	5.1	6	5.1	5.4
J_u	25	39	56.4	18.9	25.6	36.3	53	32.6	10.2	86.1	83	55.2
J_M	99.2	97.6	100	100	100	100	5.3	4.7	4.6	12.1	7.6	6.3
	(V)											
J_S	5.5	5.5	4.5	4.5	4.2	5.0	60.9	59.7	64.2	98.7	99.6	99.9
J_{BFK}	99.9	99.7	99.9	100	100	100	17.5	12.4	8.1	44.5	24.4	16.2
J_u	28.3	41.2	61.7	19	23.8	34.1	51.3	40.4	32.5	90.3	87.6	84.9
J_M	100	98	95	100	100	100	3.3	12.1	13.3	17.1	6.6	7.6

Acknowledgments

This work is supported by the National Natural Science Foundation of China grants 11501092, 11571068 and 11671073.

Appendix. Proof of theorems

Define $\mathbf{U}_{t_i} = U(\boldsymbol{u}_{t_i})$, $r_{t_i} = \|\boldsymbol{u}_{t_i}\|$, $\tilde{\mathbf{e}}_{t_1,t_2} = \tilde{\boldsymbol{u}}_{t_1,t_2} - \boldsymbol{u}_{t_1}$. First we restate the Lemma 4 in Zou et al. (2014).

Lemma A.1. *Suppose that* $\mathbf{U}$ *is uniformly distributed on the unit N sphere. For any* $N \times N$ *symmetric matrix* $\mathbf{A}$*, we have*

$$E(\mathbf{U}^T\mathbf{A}\mathbf{U})^2 = \{\mathrm{tr}^2(\mathbf{A}) + 2\mathrm{tr}(\mathbf{A}^2)\}/(N^2+2N),$$
$$E(\mathbf{U}^T\mathbf{A}\mathbf{U})^4 = \{3\mathrm{tr}^2(\mathbf{A}^2) + 6\mathrm{tr}(\mathbf{A}^4)\}/N(N+2)(N+4)(N+6).$$

Second, we restate Lemma 3 in Zou et al. (2014).

Lemma A.2. *Under* H_0*, as* $N \to \infty$ *and* $T \to \infty$, $\{\frac{N}{T(T-1)}\sum_{t_1\neq t_2}(\mathbf{U}_{t_1}^T\mathbf{U}_{t_2})^2 - 1\}/\sigma_{s0} \to^d N(0,1)$*, where* $\sigma_{s0}^2 = 4(N-1)/T(T-1)(N+2)$*.*

Below we will propose the following lemma, based on which we can directly obtain the proof of Theorem 3.1.

Lemma A.3. *Under the conditions given in Theorem 2.1, we have*

$$J_S = \frac{N}{T(T-1)} \sum_{t_1 \neq t_2} (\mathbf{U}_{t_1}^T \mathbf{U}_{t_2})^2 - 1 + o_p(\sigma_{s0}).$$

Proof of Lemma A.3.

$$\begin{aligned}\tilde{\mathbf{U}}_{t_1,t_2} &= U(\boldsymbol{u}_{t_1} + r_{t_1}^{-1}\tilde{\mathbf{e}}_{t_1,t_2})(1 + 2r_{t_1}^{-1}U_{t_1}^T\tilde{\mathbf{e}}_{t_1,t_2} + r_{t_i}^{-2}\tilde{\mathbf{e}}_{t_1,t_2}^T\tilde{\mathbf{e}}_{t_1,t_2})^{-1/2} \\ &\doteq U(\boldsymbol{u}_{t_1} + r_{t_1}^{-1}\tilde{\mathbf{e}}_{t_1,t_2})(1+\gamma_{t_1,t_2})^{-1/2},\end{aligned}$$

where $\gamma_{t_1,t_2} = 2r_{t_1}^{-1}\mathbf{U}_{t_1}^T\tilde{\mathbf{e}}_{t_1,t_2} + r_{t_1}^{-2}\tilde{\mathbf{e}}_{t_1,t_2}^T\tilde{\mathbf{e}}_{t_1,t_2}$. And

$$\begin{aligned}E(r_{t_1}^{-1}\mathbf{U}_{t_1}^T\tilde{\mathbf{e}}_{t_1,t_2})^2 &= O(N^{-1})\mathrm{tr}\{E(\mathbf{U}_{t_1}\mathbf{U}_{t_1}^T)E(\tilde{\mathbf{e}}_{t_1,t_2}\tilde{\mathbf{e}}_{t_1,t_2}^T)\} = O(N^{-1}T^{-1}),\\ E(r_{t_1}^{-2}\tilde{\mathbf{e}}_{t_1,t_2}^T\tilde{\mathbf{e}}_{t_1,t_2})^2 &= E(r_{t_1}^{-4})E(\tilde{\mathbf{e}}_{t_1,t_2}^T\tilde{\mathbf{e}}_{t_1,t_2})^2 = O(N^{-1}T^{-2}),\end{aligned}$$

due to the fact that $E(r_{t_1}^{-2}) = O(N^{-1})$, $E(\tilde{\mathbf{e}}_{t_1,t_2}\tilde{\mathbf{e}}_{t_1,t_2}^T) = O(NT^{-1})$ and $E(r_{t_1}^{-4}) = O(N^{-2})$. Thus, we have $\gamma_{t_1,t_2} = O_p(N^{-1/2}T^{-1/2})$. So

$$\begin{aligned}J_S &= \frac{N}{T(T-1)} \sum_{t_1 \neq t_2} (\tilde{\mathbf{U}}_{t_1,t_2}^T \tilde{\mathbf{U}}_{t_2,t_1})^2 - 1 \\
&= \frac{N}{T(T-1)} \sum_{t_1 \neq t_2} (\mathbf{U}_{t_1}^T \mathbf{U}_{t_2})^2 - 1 + \frac{N}{T(T-1)} \sum_{t_1 \neq t_2} (\mathbf{U}_{t_1}^T \mathbf{U}_{t_2})^2 \{(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} - 1\} \\
&\quad + 2\frac{N}{T(T-1)} \sum_{t_1 \neq t_2} r_{t_2}^{-1}\mathbf{U}_{t_2}^T\mathbf{U}_{t_1}\mathbf{U}_{t_1}^T\tilde{\mathbf{e}}_{t_2,t_1}(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} \\
&\quad + 2\frac{N}{T(T-1)} \sum_{t_1 \neq t_2} r_{t_1}^{-1}\mathbf{U}_{t_1}^T\mathbf{U}_{t_2}\mathbf{U}_{t_2}^T\tilde{\mathbf{e}}_{t_1,t_2}(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} \\
&\quad + \frac{N}{T(T-1)} \sum_{t_1 \neq t_2} r_{t_2}^{-2}\mathbf{U}_{t_1}^T\tilde{\mathbf{e}}_{t_2,t_1}\tilde{\mathbf{e}}_{t_2,t_1}^T\mathbf{U}_{t_2}(1+\gamma_{t_1.t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} \\
&\quad + \frac{N}{T(T-1)} \sum_{t_1 \neq t_2} r_{t_1}^{-2}\mathbf{U}_{t_2}^T\tilde{\mathbf{e}}_{t_1,t_2}\tilde{\mathbf{e}}_{t_1,t_2}^T\mathbf{U}_{t_1}(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} \\
&\quad + 2\frac{N}{T(T-1)} \sum_{t_1 \neq t_2} r_{t_1}^{-1}r_{t_2}^{-1}\mathbf{U}_{t_2}^T\tilde{\mathbf{e}}_{t_1,t_2}\tilde{\mathbf{e}}_{t_1,t_2}^T\mathbf{U}_{t_2}(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} \\
&\quad + 2\frac{N}{T(T-1)} \sum_{t_1 \neq t_2} r_{t_1}^{-1}r_{t_2}^{-1}\mathbf{U}_{t_1}^T\tilde{\mathbf{e}}_{t_2,t_1}\tilde{\mathbf{e}}_{t_2,t_1}^T\mathbf{U}_{t_1}(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} \\
&\quad + 2\frac{N}{T(T-1)} \sum_{t_1 \neq t_2} r_{t_1}^{-1}r_{t_2}^{-2}\mathbf{U}_{t_1}^T\tilde{\mathbf{e}}_{t_2,t_1}\tilde{\mathbf{e}}_{t_2,t_1}^T\tilde{\mathbf{e}}_{t_1,t_2}(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} \\
&\quad + 2\frac{N}{T(T-1)} \sum_{t_1 \neq t_2} r_{t_1}^{-2}r_{t_2}^{-1}\mathbf{U}_{t_2}^T\tilde{\mathbf{e}}_{t_1,t_2}\tilde{\mathbf{e}}_{t_1,t_2}^T\tilde{\mathbf{e}}_{t_2,t_1}(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} \\
&\quad \frac{N}{T(T-1)} \sum_{t_1 \neq t_2} r_{t_1}^{-2}r_{t_2}^{-2}\tilde{\mathbf{e}}_{t_2,t_1}^T\tilde{\mathbf{e}}_{t_1,t_2}\tilde{\mathbf{e}}_{t_1,t_2}^T\tilde{\mathbf{e}}_{t_2,t_1}(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1} \\
&\doteq \frac{N}{T(T-1)} \sum_{t_1 \neq t_2} (\mathbf{U}_{t_1}^T \mathbf{U}_{t_2})^2 - 1 + A_1 + A_2 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8 + A_9 + A_{10}.\end{aligned}$$

Here, we only prove $A_1 = o_p(\sigma_{s0})$, $A_2 = o_p(\sigma_{s0})$ these two items. The rest of $A_i = o_p(\sigma_{s0})$, for $i = 3, 4, 5 \ldots 10$ are similar. In fact

$$\begin{aligned}
E(A_1^2) = O(N^2T^{-4})2\sum\sum_{t_1\neq t_2} & E[(\mathbf{U}_{t_1}^T\mathbf{U}_{t_2})^4\{(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1}-1\}^2] \\
& + 4\sum\sum\sum_{t_1\neq t_2\neq t_3} E[(\mathbf{U}_{t_1}^T\mathbf{U}_{t_2})^2(\mathbf{U}_{t_1}^T\mathbf{U}_{t_3})^2\{(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1}-1\} \\
& \quad \times \{(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_3,t_1})^{-1}-1\}] \\
& \quad + \sum\sum\sum\sum_{t_1\neq t_2\neq t_3\neq t_4} E[(\mathbf{U}_{t_1}^T\mathbf{U}_{t_2})^2(\mathbf{U}_{t_4}^T\mathbf{U}_{t_3})^2\{(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1}-1\} \\
& \quad \times \{(1+\gamma_{t_3,t_4})^{-1}(1+\gamma_{t_4,t_3})^{-1}-1\}] \\
\doteq & A_{11}+A_{12}+A_{13}.
\end{aligned}$$

After some tedious calculation, we have

$$\begin{aligned}
& A_{11} \le O(NT^{-3})E(\mathbf{U}_{t_1}^T\mathbf{U}_{t_2})^4 = O(N^{-1}T^{-3}) = o(\sigma_{s0}^2), \\
& A_{12} \le O(NT^{-2})E((\mathbf{U}_{t_1}^T\mathbf{U}_{t_2})^2(\mathbf{U}_{t_1}^T\mathbf{U}_{t_3})^2) = O(N^{-1}T^{-2}) = o(\sigma_{s0}^2), \\
& A_{13} \le O(NT^{-2})E((\mathbf{U}_{t_1}^T\mathbf{U}_{t_2})^2(\mathbf{U}_{t_3}^T\mathbf{U}_{t_4})^2) = O(N^{-1}T^{-2}) = o(\sigma_{s0}^2).
\end{aligned}$$

So $A_1 = o_p(\sigma_{s0})$. And

$$\begin{aligned}
E(A_2^2) &= O(N^2T^{-4})\sum_{t_1\neq t_2} E\{r_{t_2}^{-1}\mathbf{U}_{t_2}^T\mathbf{U}_{t_1}\mathbf{U}_{t_1}^T\tilde{\mathbf{e}}_{t_2,t_1}(1+\gamma_{t_1,t_2})^{-1}(1+\gamma_{t_2,t_1})^{-1}\}^2 \\
&= O(N^2T^{-2})E(r_{t_2}^{-2})E\{\mathbf{U}_{t_2}^T\mathbf{U}_{t_1}\mathbf{U}_{t_1}^T\tilde{\mathbf{e}}_{t_2,t_1}\tilde{\mathbf{e}}_{t_2,t_1}^T\mathbf{U}_{t_1}\mathbf{U}_{t_1}^T\mathbf{U}_{t_2}\} = O(T^{-3}) = o(\sigma_{s0}^2).
\end{aligned}$$

So $A_2 = o_p(\sigma_{s0})$. Here we complete the proof. □

Proof of Theorem 2.1. Based on Lemmas A.2 and A.3, we can easily obtain the result by Slutsky's Theorem. □

Next, we proof Theorem 2.2. Define $\boldsymbol{u}_{t_1}^* = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{u}_{t_1}$, $\mathbf{U}_t^* = \boldsymbol{u}_{t_1}^*/\|\boldsymbol{u}_{t_1}^*\|$, and $r_t^* = \|\boldsymbol{u}_{t_1}^*\|$. $\tilde{\mathbf{e}}_{t_1,t_2}^* = \boldsymbol{\Sigma}^{-1/2}\tilde{\mathbf{e}}_{t_1,t_2}$.

Proof of Theorem 2.2.

$$\begin{aligned}
\tilde{\mathbf{U}}_{t_1,t_2} &= (\boldsymbol{\Lambda}_N^{1/2}\mathbf{U}_{t_1}^* + r_{t_1}^{*-1}\boldsymbol{\Lambda}_N^{1/2}\tilde{\mathbf{e}}_{t_1,t_2}^*)^T\{1+\mathbf{U}_{t_1}^{*T}D_{T,N}\mathbf{U}_{t_1}^* + 2r_{t_1}^{*-1}\mathbf{U}_{t_1}^{*T}\tilde{\mathbf{e}}_{t_1,t_2}^* + r_{t_1}^{*-2}\tilde{\mathbf{e}}_{t_1,t_2}^{*T}\tilde{\mathbf{e}}_{t_1,t_2}^* \\
&\qquad + 2r_{t_1}^{*-1}\mathbf{U}_{t_1}^{*T}D_{T,N}\tilde{\mathbf{e}}_{t_1,t_2}^* + r_{t_1}^{*-2}\tilde{\mathbf{e}}_{t_1,t_2}^{*T}D_{T,N}\tilde{\mathbf{e}}_{t_1,t_2}^*\}^{-1/2} \\
&\doteq (\boldsymbol{\Lambda}_N^{1/2}\mathbf{U}_{t_1}^* + r_{t_1}^{*-1}\boldsymbol{\Lambda}_N^{1/2}\tilde{\mathbf{e}}_{t_1,t_2}^*)^T\{1+\omega_{t_1,t_2}^*\}^{-1/2}
\end{aligned}$$

where $\omega_{t_1}^* = \mathbf{U}_{t_1}^{*T}D_{T,N}\mathbf{U}_{t_1}^* + 2r_{t_1}^{*-1}\mathbf{U}_{t_1}^{*T}\tilde{\mathbf{e}}_{t_1}^* + r_{t_1}^{*-2}\hat{\mathbf{e}}_{t_1}^{*T}\tilde{\mathbf{e}}_{t_1}^* + 2r_{t_1}^{*-1}\mathbf{U}_{t_1}^{*T}D_{T,N}\tilde{\mathbf{e}}_{t_1}^* + r_{t_1}^{*-2}\tilde{\mathbf{e}}_{t_1}^{*T}D_{T,N}\tilde{\mathbf{e}}_{t_1}^*$. And

$$\begin{aligned}
& E\{\mathbf{U}_{t_1}^{*T}D_{T,N}\mathbf{U}_{t_1}^*\}^2 = O(N^{-2}\mathrm{tr}(D_{T,N}^2)) = O(N^{-1}T^{-1}) = o(1), \\
& E\{r_{t_1}^{*-1}\mathbf{U}_{t_1}^{*T}D_{T,N}\tilde{\mathbf{e}}_{t_1,t_2}^*\}^2 = O(N^{-1})E(\mathbf{U}_{t_1}^{*T}D_{T,N}\tilde{\mathbf{e}}_{t_1,t_2}^*)^2 = O(N^{-1}T^{-2}) = o(1), \\
& E\{r_{t_1}^{*-2}\tilde{\mathbf{e}}_{t_1,t_2}^{*T}D_{T,N}\tilde{\mathbf{e}}_{t_1,t_2}^*\}^2 = O(N^{-1}T^{-2}) = o(1).
\end{aligned}$$

Similarly to Lemma A.3, $r_{t_1}^{*-1}\mathbf{U}_{t_1}^{*T}\tilde{\mathbf{e}}_{t_1,t_2}^* + r_{t_1}^{*-2}\tilde{\mathbf{e}}_{t_1,t_2}^{*T}\tilde{\mathbf{e}}_{t_1,t_2}^* = O_p(N^{-1/2}T^{-1/2}) = o_p(1)$ then we have $\omega_{t_1,t_2}^* = O_p(N^{-1/2}T^{-1/2})$. Then we have

$$\begin{aligned}
J_S &= \frac{N}{T(T-1)}\sum_{t_1\neq t_2}(\tilde{\mathbf{U}}_{t_1,t_2}^T\tilde{\mathbf{U}}_{t_2,t_1})^2 - 1 \\
&= \frac{N}{T(T-1)}\sum_{t_1\neq t_2}(\mathbf{U}_{t_1}^{*T}\boldsymbol{\Lambda}_N\mathbf{U}_{t_2}^*)^2 - 1 \\
&\quad + \frac{N}{T(T-1)}\sum_{t_1\neq t_2}(\mathbf{U}_{t_1}^{*T}\boldsymbol{\Lambda}_N\mathbf{U}_{t_2}^*)^2(\{1+\omega_{t_1,t_2}^*\}^{-1}\{1+\omega_{t_2,t_1}^*\}^{-1}-1) \\
&\quad + \frac{N}{T(T-1)}\sum_{t_1\neq t_2}2r_{t_1}^{*-1}\mathbf{U}_{t_1}^{*T}\boldsymbol{\Lambda}_N\mathbf{U}_{t_2}^*\tilde{\mathbf{e}}_{t_1,t_2}^{*T}\mathbf{U}_{t_2}^*\{1+\omega_{t_1,t_2}^*\}^{-1}\{1+\omega_{t_2,t_1}^*\}^{-1} \\
&\quad + \frac{N}{T(T-1)}\sum_{t_1\neq t_2}2r_{t_2}^{*-1}\mathbf{U}_{t_2}^{*T}\boldsymbol{\Lambda}_N\mathbf{U}_{t_1}^*\tilde{\mathbf{e}}_{t_2,t_2}^{*T}\mathbf{U}_{t_1}^*\{1+\omega_{t_1,t_2}^*\}^{-1}\{1+\omega_{t_2,t_1}^*\}^{-1}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{N}{T(T-1)}\sum_{t_1\neq t_2} r_{t_2}^{*-2}(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\tilde{\mathbf{e}}_{t_2,t_1}^{*})^2\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2,t_1}^{*}\}^{-1}\\
&+\frac{N}{T(T-1)}\sum_{t_1\neq t_2} r_{t_1}^{*-2}(\mathbf{U}_{t_2}^{*T}\mathbf{\Lambda}_N\tilde{\mathbf{e}}_{t_1,t_2}^{*})^2\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2,t_1}^{*}\}^{-1}\\
&+\frac{N}{T(T-1)}\sum_{t_1\neq t_2} 2r_{t_1}^{*-1}r_{t_2}^{*-1}\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*}\tilde{\mathbf{e}}_{t_1,t_2}^{*T}\mathbf{\Lambda}_N\tilde{\mathbf{e}}_{t_2,t_1}^{*}\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2,t_1}^{*}\}^{-1}\\
&+\frac{N}{T(T-1)}\sum_{t_1\neq t_2} 2r_{t_1}^{*-1}r_{t_2}^{*-1}\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\tilde{\mathbf{e}}_{t_2,t_1}^{*}\tilde{\mathbf{e}}_{t_1,t_2}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*}\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2,t_1}^{*}\}^{-1}\\
&+\frac{N}{T(T-1)}\sum_{t_1\neq t_2} 2r_{t_1}^{*-1}r_{t_2}^{*-2}\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\tilde{\mathbf{e}}_{t_2,t_1}^{*}\tilde{\mathbf{e}}_{t_1,t_2}^{*T}\mathbf{\Lambda}_N\tilde{\mathbf{e}}_{t_2,t_1}^{*}\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2,t_1}^{*}\}^{-1}\\
&+\frac{N}{T(T-1)}\sum_{t_1\neq t_2} 2r_{t_1}^{*-2}r_{t_2}^{*-1}\mathbf{U}_{t_2}^{*T}\mathbf{\Lambda}_N\tilde{\mathbf{e}}_{t_1,t_2}^{*}\tilde{\mathbf{e}}_{t_1,t_2}^{*T}\mathbf{\Lambda}_N\tilde{\mathbf{e}}_{t_2,t_1}^{*}\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2,t_1}^{*}\}^{-1}\\
&+\frac{N}{T(T-1)}\sum_{t_1\neq t_2} r_{t_1}^{*-2}r_{t_2}^{*-2}(\tilde{\mathbf{e}}_{t_1,t_2}^{*T}\mathbf{\Lambda}_N\tilde{\mathbf{e}}_{t_2,t_1}^{*})^2\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2.t_1}^{*}\}^{-1}\\
&\doteq \frac{N}{T(T-1)}\sum_{t_1\neq t_2}(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^2-1+B_1+B_2+B_3+B_4+B_5+B_6+B_7+B_8+B_9+B_{10}.
\end{aligned}
$$

Here, we only prove $B_1=o_p(\sigma_{T2})$, $B_2=o_p(\sigma_{T2})$ these two items. The rest of $B_i=o_p(\sigma_{T2})$, for $i=3,4,\ldots 10$ are similar. In fact

$$
\begin{aligned}
E(B_1^2)=&\,O(N^2T^{-4})E\{2\sum_{t_1\neq t_2}(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^4(\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2,t_1}^{*}\}^{-1}-1)^2\\
&+4\sum\sum_{t_1\neq t_2\neq t_3}\sum(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^2(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_3}^{*})^2(\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2,t_1}^{*}\}^{-1}-1)\\
&\qquad\times(\{1+\omega_{t_1,t_3}^{*}\}^{-1}\{1+\omega_{t_3,t_1}^{*}\}^{-1}-1)\\
&+\sum\sum_{t_1\neq t_2\neq t_3\neq t_4}\sum\sum(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^2(\mathbf{U}_{t_3}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_4}^{*})^2(\{1+\omega_{t_1,t_2}^{*}\}^{-1}\{1+\omega_{t_2,t_1}^{*}\}^{-1}-1)\\
&\qquad\times(\{1+\omega_{t_3,t_4}^{*}\}^{-1}\{1+\omega_{t_4,t_3}^{*}\}^{-1}-1)\}\\
\doteq&\, B_{11}+B_{12}+B_{13}.
\end{aligned}
$$

And

$$
\begin{aligned}
&B_{11}\le O(NT^{-3})E(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^4=O(N^{-1}T^{-3}\{3\mathrm{tr}^2(\mathbf{\Lambda}_N^2)+4\mathrm{tr}(\mathbf{\Lambda}_N^4)\})=o(\sigma_{T2}^2),\\
&B_{12}\le O(NT^{-2})E\{(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^2(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_3}^{*})^2\}\le O(N^{-3}T^{-2}\{\mathrm{tr}^2(\mathbf{\Lambda}_N^2)+2\mathrm{tr}(\mathbf{\Lambda}_N^4)\})=o(\sigma_{T2}^2),\\
&B_{13}\le O(NT^{-2})E\{(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^2E(\mathbf{U}_{t_3}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_4}^{*})^2\}=O(N^{-3}T^{-2}\mathrm{tr}^2(\mathbf{\Lambda}_N^2))=o(\sigma_{T2}^2).
\end{aligned}
$$

So $B_1=o_p(\sigma_{T2})$. Next,

$$
\begin{aligned}
E(B_2^2)&\le O(N^2T^{-4})\sum_{t_1\neq t_2}2E(r_{t_1}^{*-2})E(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*}\mathbf{U}_{t_2}^{*T}\tilde{\mathbf{e}}_{t_1,t_2})^2\\
&=O(NT^{-2})E(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*}\mathbf{U}_{t_2}^{*T}\tilde{\mathbf{e}}_{t_1,t_2}^{*}\tilde{\mathbf{e}}_{t_1,t_2}^{*T}\mathbf{U}_{t_2}^{*}\mathbf{U}_{t_2}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_1}^{*})\\
&=O(N^{-1}T^{-3}\mathrm{tr}(\mathbf{\Lambda}_N^2))=O(N^{-1}T^{-3}(N+\mathrm{tr}(D_{N,T}^2)))=o(\sigma_{T2}^2).
\end{aligned}
$$

So $B_2=o_p(\sigma_{T2})$. Thus, $J_S=\frac{N}{T(T-1)}\sum_{t_1\neq t_2}(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^2-1+o_p(\sigma_{T2})$. By Lemma A.1,

$$
\begin{aligned}
&E(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^2=\mathrm{tr}[E(\mathbf{\Lambda}_N^{1/2}\mathbf{U}_{t_1}\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N^{1/2})]^2=N^{-2}\mathrm{tr}(\mathbf{\Lambda}_N^2)=N^{-2}(N+\mathrm{tr}(D_{N,T}^2)),\\
&E(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^4=\{3\mathrm{tr}^2(\mathbf{\Lambda}_N^2)+4\mathrm{tr}(\mathbf{\Lambda}_N^4)\}/\{N^2(N+2)^2\},\\
&E(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^2(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_3}^{*})^2=\{\mathrm{tr}^2(\mathbf{\Lambda}_N^2)+2\mathrm{tr}(\mathbf{\Lambda}_N^4)\}/\{N^3(N+2)\}.
\end{aligned}
$$

Combining all above, we can get

$$
\mathrm{var}\left\{\frac{1}{T(T-1)}\sum_{t_1\neq t_2}(\mathbf{U}_{t_1}^{*T}\mathbf{\Lambda}_N\mathbf{U}_{t_2}^{*})^2\right\}=\left[\frac{4\mathrm{tr}^2(\mathbf{\Lambda}_N^2)}{T(T-1)N^2(N+2)^2}+\frac{8\mathrm{tr}(\mathbf{\Lambda}_N^4)}{TN^2(N+2)^2}\right](1+o(1)).
$$

so we have

$$E(J_S) = N^{-1}\mathrm{tr}(D_{N,T}^2) + o(\sigma_{T2}),$$
$$\mathrm{var}(J_S) = \left[\frac{4\mathrm{tr}^2(\mathbf{\Lambda}_N^2)}{T(T-1)(N+2)^2} + \frac{8\mathrm{tr}(\mathbf{\Lambda}_N^4)}{T(N+2)^2}\right](1 + o(1)).$$

By Theorem 2 of Zou et al. (2014), we can complete this theorem. □

References

Baltagi, B.H., Feng, Q., Kao, C., 2011. Testing for sphericity in a fixed effects panel data model. Econom. J. 14, 25–47.
Baltagi, B.H., Kao, C., Peng, B., 2015. On testing for sphericity with non-normality in a fixed effects panel data model. Statist. Probab. Lett. 98, 123–130.
Chen, S., Zhang, L., Zhong, P., 2010. Tests for high-dimensional covariance matrices. J. Amer. Statist. Assoc. 105, 810–819.
Fisher, T., Sun, X., Gallagher, C.M., 2010. A new test for sphericity of the covariance matrix for high dimensional data. J. Multivariate Anal. 101, 2554–2570.
Ledoit, O., Wolf, M., 2002. Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. Ann. Statist. 30, 1081–1102.
Mao, G., 2014. A note on tests of sphericity and cross-sectional dependence for fixed effects panel model. Econom. Lett. 122, 215–219.
Oja, H., 2010. Multivariate Nonparametric Methods with R. Springer, New York.
Pesaran, M.H., Ullah, A., Yamagata, T., 2008. A bias-adjusted lm test of error cross-section independence. Econom. J. 11, 105–127.
Zou, C., Peng, L., Feng, L., Wang, Z., 2014. Multivariate-sign-based high-dimensional tests for sphericity. Biometrika 101, 229–236.