Stockholms universitet

On phase transitions for the trace of squared sample correlation matrices in high dimension

Felix Seo

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2024:10
Matematisk statistik
Juni 2024

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

ON PHASE TRANSITIONS FOR THE TRACE OF SQUARED SAMPLE CORRELATION MATRICES IN HIGH DIMENSION

FELIX SEO

Abstract. We provide limit theory for the trace of the squared sample correlation matrix $\mathbf{R}$, constructed from n observations of a p-dimensional random vector with iid components. If the entries have finite fourth moment and p and n grow proportionally, it is known that $\mathrm{tr}(\mathbf{R}^2)$ satisfies a central limit theorem (CLT) and the centering and scaling sequences are universal in the sense that they do not depend on the entry distribution. Under a symmetry and a regular variation assumption with index α and any growth rate of the dimension, we prove that the universal CLT remains valid for $\alpha > 3$. Moreover, for $\alpha \leq 3$ we establish a non-universal CLT with norming sequences depending on the value of α. Our findings are illustrated in a small simulation study.

1. Introduction

Measuring the dependence between random variables has always been a fundamental task in statistics. Starting with the early works of Pearson [18], Kendall [15], Hoeffding [13] and Blum [6], several measures of dependence or association have been introduced and analyzed by numerous authors. An outstanding role is played by Pearson's correlation coefficient, a measure of the linear dependency of two random variables, about which most students learn early on in their studies. Motivated by its importance for statistical inference and estimation, many works are devoted to its stochastic properties in different frameworks. For example, in time series analysis, the notion of correlation plays a vital role in multivariate statistical analysis for parameter estimation, goodness-of-fit tests, change-point detection, etc.; see for example the classical monographs [7, 19].

With the rapid advancements of data collection devices, many modern fields such as biological engineering, telecommunications and finance require the analysis of high-dimensional data sets where the dimension p and the sample size n are of comparable magnitude. As a result traditional results from multivariate analysis, which rely on the assumption that the dimension remains fixed and thus is negligible compared to the sample size, are typically not applicable in high-dimensional regimes. Driven by such challenges, random matrix theory - as outlined in the monographs [2, 20] - aims to provide a deeper understanding of differences that arise when p is assumed to grow with n. A standard assumption is that the ratio p/n approaches some positive constant. It is worth mentioning that a regime where $p = \sqrt{n}$ might lead to completely different asymptotic theory than (say) $p = n$. In practical applications, however, p/n is always some positive number and it is therefore non-trivial to distinguish between various regimes.

1.1. Our Model.

Consider a p-dimensional population $\mathbf{x} = (X_1, \dots, X_p)^\top \in \mathbb{R}^p$, where the components X_i are independent and identically distributed (iid), non-degenerated random variables with mean zero. For a sample $\mathbf{x}_1, \dots, \mathbf{x}_n$ from the population we construct the data matrix

2020 *Mathematics Subject Classification.* Primary 60B20; Secondary 60F05.

$\mathbf{X} = \mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n) = (X_{ij})_{1\le i\le p; 1\le j\le n}$, the sample covariance matrix $\mathbf{S} = \mathbf{S}_n = n^{-1}\mathbf{X}\mathbf{X}^\top$ and the sample correlation matrix $\mathbf{R}$,

$$\mathbf{R} = \mathbf{R}_n = \{\mathrm{diag}(\mathbf{S}_n)\}^{-1/2}\,\mathbf{S}_n\{\mathrm{diag}(\mathbf{S}_n)\}^{-1/2} = \mathbf{Y}\mathbf{Y}^\top\,. \tag{1.1}$$

Here the standardized matrix $\mathbf{Y} = \mathbf{Y}_n = (Y_{ij})_{1\le i\le p; 1\le j\le n}$ for the sample correlation matrix has entries

$$Y_{ij} = Y_{ij}^{(n)} = \frac{X_{ij}}{\sqrt{X_{i1}^2 + \cdots + X_{in}^2}}\,, \tag{1.2}$$

which depend on n. Throughout the paper, we often suppress the dependence on n in our notation. Since Y_{ij} is invariant with respect to a scaling of the X_{ij}'s, we will assume without loss of generality that $\mathbb{E}[X_{11}^2] = 1$ whenever $\mathbb{E}[X_{11}^2]$ is finite. In this paper, we will often assume that $|X_{11}|$ has a regularly varying tail with index $\alpha > 0$, that is

$$\mathbb{P}(|X_{11}| > x) = x^{-\alpha}\,L(x)\,, \qquad x > 0\,, \tag{1.3}$$

for a function L that is slowly varying at infinity. Thus, regularly varying distributions possess power-law tails and moments of $|X_{11}|$ of higher order than α are infinite. Typical examples include the Pareto distribution with parameter α and the t-distribution with α degrees of freedom. In addition, we assume that the distribution of X_{11} is symmetric, that is, $X_{11} \overset{\mathrm{d}}{=} -X_{11}$. We consider the high-dimensional regime

$$p = p_n \to \infty, \qquad \text{as} \quad n \to \infty\,.$$

1.2. **Background.** Many popular test statistics, such as the likelihood ratio statistic for testing independence of a normal population, can be expressed as a function of the eigenvalues of the sample covariance matrix $\mathbf{S}$ or the sample correlation matrix $\mathbf{R}$. For a function $f : \mathbb{R} \to \mathbb{R}$ and a random matrix $\mathbf{A}$ with p real eigenvalues $\lambda_1(\mathbf{A}) \ge \cdots \ge \lambda_p(\mathbf{A})$, we call $\sum_{i=1}^p f(\lambda_i(\mathbf{A}))$ a linear spectral statistic of $\mathbf{A}$. In the proportional regime $p/n \to \gamma \in (0,\infty)$, the spectral properties of the sample covariance matrix $\mathbf{S}$ have been well studied in random matrix theory since the pioneering work [17], where it is shown that the empirical distribution of the eigenvalues $\lambda_i(\mathbf{S})$ converges weakly to the celebrated Marčenko–Pastur law. Moreover, the paper [3] established asymptotic normality of suitably centered and normalized linear spectral statistics of $\mathbf{S}$ in the proportional regime with finite fourth moment $\mathbb{E}[X_{11}^4]$.

In contrast, the study of the sample correlation matrix $\mathbf{R}$ is more recent and more limited. A fundamental reason is that compared to the original data matrix $\mathbf{X}$, the entries Y_{ij} of the standardized matrix $\mathbf{Y}$ are no longer independent within the same row (the different rows remain iid). This makes the sample correlation matrix more challenging to study. Assuming the proportional regime, Lemma 2 in [4] asserts that $\mathbb{E}[X_{11}^4] < \infty$ is equivalent to $\|\mathrm{diag}(\mathbf{S}) - \mathbf{I}\| \overset{\mathrm{a.s.}}{\to} 0\,, n \to \infty$, where $\|\cdot\|$ is the spectral norm and $\mathbf{I}$ the identity matrix; see also [11, Theorem 1.2] for a similar result in the dependent case. Therefore, under finite fourth moment the normalization $\{\mathrm{diag}(\mathbf{S})\}^{-1/2}$ in (1.1) can be replaced with $\mathbf{I}$ and consequently $\max_i |\lambda_i(\mathbf{R}) - \lambda_i(\mathbf{S})| \le \|\mathbf{R} - \mathbf{S}\|$ converges to zero almost surely as $n \to \infty$. It turns out that a modification of this trick can be used to obtain central limit theorems (CLTs) for linear spectral statistics of $\mathbf{R}$ from CLTs for linear spectral statistics of $\mathbf{S}$ [21, 22].

To the best of our knowledge, the only available result under infinite fourth moment concerns the function $f = \log$ for which $\sum_{i=1}^p \log(\lambda_i(\mathbf{R})) = \log\det\mathbf{R}$, the log-determinant of the sample correlation matrix.

Theorem 1.1. *[12, Theorem 2.1] Assume $p/n \to \gamma \in (0,1)$ and that the distribution of X_{11} is symmetric and regularly varying with index $\alpha \in (3,4)$. Then, as $n \to \infty$, we have*

$$\frac{\log\det\mathbf{R} - (p-n+\frac{1}{2})\log(1-\frac{p}{n}) + p - \frac{p}{n}}{\sqrt{-2\log(1-p/n) - 2p/n}} \stackrel{\mathrm{d}}{\to} N(0,1)\,.$$

This work was motivated by the functions $f_k(x) := x^k$, $k \geq 2$, for which

$$\sum_{i=1}^p f_k(\lambda_i(\mathbf{R})) = \operatorname{tr}(\mathbf{R}^k)\,.$$

(Note that the case $k=1$ is degenerate since $\operatorname{tr}(\mathbf{R}) = p$ is non-random.) CLTs for linear spectral statistics of $\mathbf{R}$ for more general functions f can be obtained by approximating f through polynomials, that is, linear combinations of f_k's. For technical reasons and for the sake of clarity, we restrict ourselves to the case $k=2$.[1]

1.3. **Our contributions.**

The novel contributions of this paper are outlined below.

- We provide CLTs for $\operatorname{tr}(\mathbf{R}^2)$ under general growth rates of p relative to n and investigate the influence of the tail index α. If $p \asymp n^\delta$ for some $\delta > 0$, we determine a region for (α,δ) where $\operatorname{tr}(\mathbf{R}^2)$ satisfies a CLT. For any pair (α,δ) outside the closure of this region, we prove that moment convergence fails.
- At $\alpha = 3$ (corresponding to the boundary of finite and infinite third moment), we discover a transition in the variance of $\operatorname{tr}(\mathbf{R}^2)$. As a consequence, if $\alpha > 3$, no restriction on the growth of p is required for the validity of the CLT.
- To the best of our knowledge, this work is the first that provides a CLT for a (non-trivial) linear spectral statistic of $\mathbf{R}$ in the case of infinite third moment $\mathbb{E}|X_{11}|^3 = \infty$.

1.4. **Structure of this paper.**

This paper is structured as follows. In Section 2, we derive a decomposition of $\operatorname{tr}(\mathbf{R}^2)$, the trace of the squared sample correlation matrix. Based on this decomposition, Theorem 2.6 provides precise conditions for the convergence of the fourth moment of the standardized $\operatorname{tr}(\mathbf{R}^2)$ to the fourth moment of a standard normal variable. In Section 3, we present central limit theorems for $\operatorname{tr}(\mathbf{R}^2)$ (Theorem 3.1) under very general growth rates on the dimension p and the tail index α. The results are then illustrated by means of a small simulation study. Section 4 contains the proof of Theorem 3.1 using martingale theory, while Section 5 is devoted to the proofs of the results of Section 2. Finally, the appendix consists of facts for sums of regularly varying random variables.

[1] An extension to general k is a topic for future research.

1.5. **Notation.** Convergence in distribution (resp. probability) is denoted by $\xrightarrow{\mathrm{d}}$ (resp. $\xrightarrow{\mathbb{P}}$), equality in distribution by $\stackrel{\mathrm{d}}{=}$, and unless explicitly stated otherwise all limits are for $n \to \infty$. For sequences $(a_n)_n$ and $(b_n)_n$ we write $a_n = O(b_n)$ if $a_n/b_n \le C$ for some constant $C > 0$ and every $n \in \mathbb{N}$, and $a_n = o(b_n)$ if $\lim_{n\to\infty} a_n/b_n = 0$. Additionally, we use the notation $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$, and $a_n = \omega(b_n)$ if $\lim_{n\to\infty} a_n/b_n = \infty$. We write $a_n \lesssim b_n$ if there exists a positive constant C not depending on n such that $a_n \le C\, b_n$ for sufficiently large n. The notation $a_n \stackrel{\text{sl.v.}}{=} b_n$ means that $a_n = b_n \ell(n)$ for some function ℓ that is slowly varying (at infinity). A function $\ell : (0,\infty) \to (0,\infty)$ is said to be slowly varying (at infinity) if $\lim_{x\to\infty} \ell(tx)/\ell(x) = 1$ for any $t > 0$.

2. Preliminaries

The main motivation of this work is to investigate the fluctuations of the trace of powers of the sample correlation matrix. That is, we aim to prove limit theorems for

$$\operatorname{tr}(\mathbf{R}^k) = \sum_{i_1,\ldots,i_k=1}^{p} \sum_{t_1,\ldots,t_k=1}^{n} Y_{i_1t_1}Y_{i_1t_2}Y_{i_2t_1}Y_{i_2t_2}\cdots Y_{i_kt_k}Y_{i_{k+1}t_k}, \qquad k \ge 2. \tag{2.1}$$

(Here the convention $i_{k+1} = i_1$ is used.) It is worth mentioning that $\operatorname{tr}(\mathbf{R}) = p$ since all diagonal elements of $\mathbf{R}$ are one. In this paper, we will focus on $\operatorname{tr}(\mathbf{R}^2)$, but our approach naturally extends to traces of higher powers of $\mathbf{R}$.

Unless explicitly stated otherwise, the (X_{it}) are iid and symmetric throughout this paper, which implies that the Y_{it} are symmetric as well. The following properties of the matrix $\mathbf{Y} = (Y_{it})$ will be repeatedly used in this paper.

(1) By symmetry of the entry distribution we have for $s \le n$ that $\mathbb{E}[Y_{i1}^{m_1}\cdots Y_{is}^{m_s}] = 0$ if at least one exponent $m_j \in \mathbb{N}$ is odd.
(2) $\mathbf{Y}$ has independent rows.
(3) By definition, $\sum_{t=1}^n Y_{it}^2 = 1$ for each row i.

The first step is to study $\operatorname{tr}(\mathbf{R}^2)$ for a wide range of distributions of X_{11} and growth rates of the dimension p relative to the sample size n. But before we delve deeper into this we need to introduce some important notation and results about the moments of Y_{ij}'s that will be crucial in our analysis.

First, define for all positive integers $k_1,\ldots,k_r$

$$\beta_{2k_1,\ldots,2k_r} := \mathbb{E}[Y_{11}^{2k_1}Y_{12}^{2k_2}\cdots Y_{1r}^{2k_r}],$$

where we recall the definition of Y_{ij} from (1.2). Since $\beta_{2k_1,\ldots,2k_r} = \beta_{2k_{\pi(1)},\ldots,2k_{\pi(r)}}$ for any permutation π on $\{1,\ldots,r\}$ we will write the indices in decreasing order, e.g., instead of $\beta_{2,4}$ we prefer writing $\beta_{4,2}$. The following key lemma reveals the asymptotic behavior of $\beta_{2k_1,\ldots,2k_r}$.

Lemma 2.1. *Define the Y_{ij}'s as in* (1.2) *and let L be a slowly varying function (at infinity). For integers $k_1,\ldots,k_r \ge 1$, set $k = k_1 + \cdots + k_r$ and $N_1 = \#\{1 \le i \le r : k_i = 1\}$, and let $\Gamma(\cdot)$ denote the gamma function.*

(a) *If $\alpha \in (0,2)$ and $\mathbb{P}(|X_{11}| > x) = x^{-\alpha}L(x)$ for $x > 0$, then it holds*

$$\lim_{n\to\infty} n^r \beta_{2k_1,\ldots,2k_r} = \frac{\left(\frac{\alpha}{2}\right)^{r-1}\Gamma(r)\prod_{j=1}^r \Gamma(k_j - \alpha/2)}{\Big(\Gamma(1-\alpha/2)\Big)^r\,\Gamma(k)}\,.$$

In particular, we have

$$\lim_{n\to\infty} n\beta_{2k} = \frac{\Gamma(k-\alpha/2)}{\Gamma(1-\alpha/2)\Gamma(k)}\,, \qquad k \ge 1\,.$$

(b) *If* $\alpha \in [2,4)$, $\mathbb{E}[X_{11}^2]=1$ *and* $\mathbb{P}(|X_{11}|>x) = x^{-\alpha}L(x)$ *for* $x>0$*, then it holds*

$$\lim_{n\to\infty} \frac{n^{N_1(1-\alpha/2)+r\alpha/2}}{L^{r-N_1}(n^{1/2})}\beta_{2k_1,\ldots,2k_r} = \frac{(\alpha/2)^{r-N_1}\Gamma(N_1(1-\alpha/2)+r\alpha/2)\,\prod_{i:k_i\ge 2}\Gamma(k_i-\alpha/2)}{\Gamma(k)}\,. \quad (2.2)$$

In particular, we have

$$\lim_{n\to\infty} \frac{n^{\alpha/2}}{L(n^{1/2})}\beta_{2k} = \frac{\alpha\Gamma(\alpha/2)\Gamma(k-\alpha/2)}{2\Gamma(k)}\,, \qquad k\ge 2\,.$$

(c) *Assume* $\mathbb{P}(|X_{11}|>x) = x^{-\alpha}L(x)$*. If* $\{\alpha=2$ *and* $\mathbb{E}[X_{11}^2]=\infty\}$ *or* $\{\alpha=4, \mathbb{E}[X_{11}^2]=1$ *and* $\mathbb{E}[X_{11}^4]=\infty\}$*, then* (2.2) *remains valid if we multiply its left-hand side with some slowly varying function (that depends on* L *and* $k_1,\ldots,k_r$*).*

(d) *If* $\mathbb{E}[X_{11}^{2\max_i k_i}]<\infty$ *and* $\mathbb{E}[X_{11}^2]=1$*, then it holds*

$$\lim_{n\to\infty} n^k\beta_{2k_1,\ldots,2k_r} = \prod_{i=1}^r \mathbb{E}[X_{11}^{2k_i}]\,.$$

Proof. For a proof of part (a), see [1, p. 4]. Regarding part (b), we remark that (2.2) was proved in [12, Lemma 4.1] for $\alpha\in(2,4)$. For our case let $\beta=\alpha/2$ and $X \stackrel{d}{=} X_{11}$. From [1, p. 7], we have

$$\mathbb{E}[Y_{11}^{2k_1}\cdots Y_{1r}^{2k_r}] = \frac{(-1)^k}{n\Gamma(k)}\int_0^\infty \left(\tfrac{t}{n}\right)^{k-1}\varphi^{n-r}\left(\tfrac{t}{n}\right)\prod_{i=1}^r \varphi^{(k_i)}\left(\tfrac{t}{n}\right)\mathrm{d}t\,, \quad (2.3)$$

where $\varphi(s)=\mathbb{E}[\mathrm{e}^{-sX^2}]$, $s>0$, and $\varphi^{(m)}(s) = \frac{\mathrm{d}^m}{\mathrm{d}s^m}\varphi(s)$. By [1], we have

$$\lim_{n\to\infty}\varphi^{n-r}\left(\tfrac{t}{n}\right) = \mathrm{e}^{-t}\,, \qquad t>0\,, \quad (2.4)$$

provided that $\mathbb{E}[X^2]=1$. For regularly varying $|X|$ with index β, [16, Lemma 2] asserts that the asymptotic behavior of $\varphi^{(m)}(s)$, $m\in\mathbb{N}$, at the origin is given by

$$(-1)^m\varphi^{(m)}(s) \sim \begin{cases} \beta\Gamma(m-\beta)s^{\beta-m}L(s^{-1/2})\,, & \text{if } m>\beta,\\ \beta\ell(s^{-1})\,, & \text{if } m=\beta \text{ and } \mathbb{E}[X^{2m}]=\infty,\\ \mathbb{E}[X^{2m}]\,, & \text{if } m\le\beta \text{ and } \mathbb{E}[X^{2m}]<\infty,\end{cases} \qquad s\downarrow 0\,, \quad (2.5)$$

where $\ell(x)=\int_0^x L(u^{1/2})/u\,\mathrm{d}u$ is a slowly varying function (at infinity).

By (2.3), Potter's theorem and the dominated convergence theorem (for more details see [1] or [9]), we obtain in view of (2.4) and (2.5) that, as $n \to \infty$,

$$\begin{aligned}
\mathbb{E}[Y_{11}^{2k_1} \cdots Y_{1r}^{2k_r}] &= \frac{(-1)^k}{n\Gamma(k)} \int_0^\infty \left(\tfrac{t}{n}\right)^{k-1} \varphi^{n-r}\left(\tfrac{t}{n}\right) \left(\varphi^{(1)}\left(\tfrac{t}{n}\right)\right)^{N_1} \prod_{i:k_1 \geq 2} \varphi^{(k_i)}\left(\tfrac{t}{n}\right) \mathrm{d}t \\
&\sim \frac{1}{n\Gamma(k)} \int_0^\infty \left(\tfrac{t}{n}\right)^{k-1} \mathrm{e}^{-t} \left(\mathbb{E}[X^2]\right)^{N_1} \prod_{i:k_i \geq 2} \beta\Gamma(k_i - \beta) \left(\tfrac{t}{n}\right)^{\beta - k_i} \underbrace{L\left(\left(\tfrac{t}{n}\right)^{-1/2}\right)}_{\sim L(n^{1/2})} \mathrm{d}t \\
&\sim \Big(\prod_{i:k_i \geq 2} \Gamma(k_i - \beta)\Big) \frac{\beta^{r-N_1} L^{r-N_1}(n^{1/2})}{n^{N_1(1-\beta)+\beta r}\Gamma(k)} \int_0^\infty \mathrm{e}^{-t}\, t^{N_1(1-\beta)+\beta r - 1} \,\mathrm{d}t \\
&= \frac{L^{r-N_1}(n^{1/2})}{n^{N_1(1-\beta)+\beta r}} \frac{\beta^{r-N_1}\Gamma(N_1(1-\beta)+\beta r) \prod_{i:k_i \geq 2} \Gamma(k_i - \beta)}{\Gamma(k)} \,.
\end{aligned}$$

Rearranging yields (2.2) and completes the proof of part (b).

The proof of part (c) is very similar. To this end, note that the three lines in (2.5) only differ by a slowly varying function. In the case $\{\alpha = 2 \text{ and } \mathbb{E}[X_{11}^2] = \infty\}$, one has to use the middle line in (2.5) for $\varphi^{(1)}$ instead of the first one, combined with equation (14) in [1] instead of (2.4). In the case $\{\alpha = 4, \mathbb{E}[X_{11}^2] = 1 \text{ and } \mathbb{E}[X_{11}^4] = \infty\}$, one needs to use the middle line in (2.5) for $\varphi^{(2)}$ instead of the last one. For brevity we omit details.

Regarding part (d), we analogously get, as $n \to \infty$,

$$\begin{aligned}
\mathbb{E}[Y_{11}^{2k_1} \cdots Y_{1r}^{2k_r}] &= \frac{(-1)^k}{n\Gamma(k)} \int_0^\infty \left(\tfrac{t}{n}\right)^{k-1} \varphi^{n-r}\left(\tfrac{t}{n}\right) \prod_{i=1}^r \varphi^{(k_i)}\left(\tfrac{t}{n}\right) \mathrm{d}t \\
&\sim \frac{1}{n\Gamma(k)} \int_0^\infty \left(\tfrac{t}{n}\right)^{k-1} \mathrm{e}^{-t} \prod_{i=1}^r \mathbb{E}[X^{2k_i}] \,\mathrm{d}t = \frac{1}{n^k} \prod_{i=1}^r \mathbb{E}[X^{2k_i}] \,.
\end{aligned}$$

□

Now we are ready to calculate the mean and variance of $\mathrm{tr}(\mathbf{R}^2)$. From (2.1) it is simple to see that

$$\begin{aligned}
\mathrm{tr}(\mathbf{R}^2) &= \sum_{i_1,i_2=1}^p \sum_{t_1,t_2=1}^n Y_{i_1t_1} Y_{i_1t_2} Y_{i_2t_1} Y_{i_2t_2} \\
&= \sum_{i=1}^p \sum_{t_1,t_2=1}^n Y_{it_1}^2 Y_{it_2}^2 + \sum_{\substack{i_1,i_2=1 \\ i_1 \neq i_2}}^p \sum_{t_1,t_2=1}^n Y_{i_1t_1} Y_{i_1t_2} Y_{i_2t_1} Y_{i_2t_2} \\
&= p + \sum_{\substack{i_1,i_2=1 \\ i_1 \neq i_2}}^p \sum_{t=1}^n Y_{i_1t}^2 Y_{i_2t}^2 + \sum_{\substack{i_1,i_2=1 \\ i_1 \neq i_2}}^p \sum_{\substack{t_1,t_2=1 \\ t_1 \neq t_2}}^n Y_{i_1t_1} Y_{i_1t_2} Y_{i_2t_1} Y_{i_2t_2} \,,
\end{aligned} \tag{2.6}$$

where the property $Y_{i1}^2 + \cdots + Y_{in}^2 = 1$ was used for the last equality. Since $\mathbb{E}[Y_{it}^2] = 1/n$, we deduce that

$$\mu_n := \mathbb{E}[\mathrm{tr}(\mathbf{R}^2)] = p + \frac{p(p-1)}{n}. \tag{2.7}$$

An interesting observation is that the mean does not depend on the distribution of X_{11}. In view of the identity $\sum_{t=1}^n \big(Y_{i_1t}^2 Y_{i_2t}^2 - \frac{1}{n^2}\big) = \sum_{t=1}^n \big(Y_{i_1t}^2 - \frac{1}{n}\big)\big(Y_{i_2t}^2 - \frac{1}{n}\big)$, we have derived the nice decomposition

$$\operatorname{tr}(\mathbf{R}^2) - \mathbb{E}[\operatorname{tr}(\mathbf{R}^2)] = T_1 + T_2\,, \tag{2.8}$$

where

$$T_1 := \sum_{\substack{i_1,i_2=1\\ i_1\neq i_2}}^p \sum_{t=1}^n \Big(Y_{i_1t}^2 - \frac{1}{n}\Big)\Big(Y_{i_2t}^2 - \frac{1}{n}\Big) \quad \text{and} \quad T_2 := \sum_{\substack{i_1,i_2=1\\ i_1\neq i_2}}^p \sum_{\substack{t_1,t_2=1\\ t_1\neq t_2}}^n Y_{i_1t_1}Y_{i_1t_2}Y_{i_2t_1}Y_{i_2t_2} \tag{2.9}$$

are two sums of centered and uncorrelated random variables. Using (2.8), the variance of $\operatorname{tr}(\mathbf{R}^2)$ is given by

$$\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2)) = \mathbb{E}[T_1^2] + \mathbb{E}[T_2^2], \tag{2.10}$$

since $\mathbb{E}[T_1T_2] = 0$ as it only contains moments of odd powers of Y_{it}'s. The next lemma gives $\mathbb{E}[T_1^2]$ and $\mathbb{E}[T_2^2]$ in terms of $\beta_4 = \mathbb{E}[Y_{11}^4]$. Its proof as well as the proofs of the following results in this section are presented in Section 5.

Lemma 2.2. *For any symmetric distribution of X_{11} and $p = p_n \to \infty$, it holds*

$$\mathbb{E}[T_1^2] = \frac{p(p-1)n(2n-1)}{n-1}\Big(\beta_4 - \frac{1}{n^2}\Big)^2 \sim 2p^2 n\Big(\beta_4 - \frac{1}{n^2}\Big)^2,$$
$$\mathbb{E}[T_2^2] = \frac{4p(p-1)n}{n-1}\Big(\frac{1}{n} - \beta_4\Big)^2 \sim 4p^2\Big(\frac{1}{n} - \beta_4\Big)^2, \qquad n\to\infty\,.$$

From (2.10) and Lemma 2.2 we deduce that

$$\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2)) = 2np(p-1)\bigg(\frac{n-0.5}{n-1}\Big(\beta_4 - \frac{1}{n^2}\Big)^2 + \frac{2}{n-1}\Big(\frac{1}{n} - \beta_4\Big)^2\bigg). \tag{2.11}$$

Remark 2.3. From a theoretical point of view, it is more interesting to discuss and interpret the findings of this work for distributions with infinite fourth moments. To this end, we typically impose the regular variation assumption (1.3) with index $\alpha \in (0,4)$. We would like to mention that the case of finite fourth moment is much simpler from a technical point of view since it only requires part (d) of Lemma 2.1, whereas the regular variation setup with $\alpha \in (0,4)$ contains fascinating transitions since the orders of the β's depend on α as showcased in parts (a) and (b) of Lemma 2.1.

Roughly speaking, the variance in (2.11) is essentially determined by $\mathbb{E}[T_1^2]$ if $\mathbb{E}|X_{11}|^3 = \infty$ and by $\mathbb{E}[T_2^2]$ otherwise. This means that the variance undergoes a transition at $\alpha = 3$. To make this point more precise, we will start by analyzing the formulas in Lemma 2.2 for $\alpha \in (2,4)$. In this case, Lemma 2.1 asserts that $\beta_4 \sim n^{-\alpha/2}L(n^{1/2})C_{4,\alpha}$, as $n\to\infty$, where $C_{4,\alpha} := \Gamma(1+\alpha/2)\Gamma(2-\alpha/2)$. In combination with Lemma 2.2, we deduce that, as $n \to \infty$,

$$\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2)) \sim \begin{cases} \mathbb{E}[T_1^2] \sim 2p^2 n \beta_4^2 & \text{if } \alpha < 3,\\ \mathbb{E}[T_2^2] \sim 4p^2 n^{-2} & \text{if } \alpha > 3. \end{cases} \tag{2.12}$$

The next lemma provides more detailed information about the asymptotic behavior of $\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))$. It shows that

$$\sigma_n^2 := 2p^2 n\big(\beta_4^2 + 2n^{-3}\big) \tag{2.13}$$

is asymptotically equivalent to $\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))$, and thus efficiently captures the effect of p, n and the distribution of X_{11}.

Lemma 2.4. *For any symmetric distribution of X_{11} and $p = p_n \to \infty$, it holds*

$$\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2)) \sim \sigma_n^2 , \qquad n \to \infty .$$

Moreover, σ_n^2 satisfies

$$\sigma_n^2 \sim \begin{cases} 4p^2 n^{-2} & \text{if } \mathbb{E}[X_{11}^4] < \infty \text{ or } \alpha \in (3,4), \\ 2p^2 n^{-2}(L^2(n^{1/2}) C_{4,\alpha}^2 + 2) & \text{if } \alpha = 3, \\ 2p^2 n^{1-\alpha} L^2(n^{1/2}) C_{4,\alpha}^2 & \text{if } \alpha \in [2,3) \text{ and } \mathbb{E}[X_{11}^2] = 1, \\ 2p^2 n^{-1}(1-\alpha/2)^2 & \text{if } \alpha \in (0,2), \end{cases}$$

where $C_{4,\alpha} := \Gamma(1+\alpha/2)\,\Gamma(2-\alpha/2)$.

In view of (2.12), it follows from Markov's inequality that

$$\frac{\mathrm{tr}(\mathbf{R}^2) - \mathbb{E}[\mathrm{tr}(\mathbf{R}^2)]}{\sigma_n} = \begin{cases} T_1/\sigma_n + o_{\mathbb{P}}(1) & \text{if } \alpha < 3, \\ T_2/\sigma_n + o_{\mathbb{P}}(1) & \text{if } \alpha > 3. \end{cases} \tag{2.14}$$

2.1. **Moments of T_1 and T_2.** Throughout this subsection, we assume that the distribution of X_{11} is symmetric and regularly varying with index $\alpha \in (0,4)$, unless explicitly stated otherwise. In order to get a first idea for which combinations of α, p and n we might have a CLT for $(\mathrm{tr}(\mathbf{R}^2) - \mathbb{E}[\mathrm{tr}(\mathbf{R}^2)])/\sigma_n \stackrel{\mathrm{d}}{\to} N(0,1)$, we compute the fourth moment of these random variables and check when they converge to 3, the fourth moment of a standard normal variable.

In view of (2.14), it suffices to study T_1 for $\alpha \le 3$ and T_2 for $\alpha \ge 3$. Careful combinatorial considerations yield the following important technical result.

Lemma 2.5. *For $\alpha \in (0,4)$ it holds*

$$\mathbb{E}[T_1^4] \sim 12\beta_4^4 n^2 p^4 + 8\beta_8^2 n p^2 + 64\beta_6^2 \beta_4 n p^3 , \qquad n \to \infty .$$

For $4 > \alpha \ge 3$ it holds

$$\mathbb{E}[T_2^4] \sim 48\, n^{-4} p^4 , \qquad n \to \infty .$$

In order to capture the interplay between p, n and α, we introduce the function δ^*,

$$\delta^*(\alpha) := \begin{cases} (5-\alpha)/2 & \text{if } \alpha \in (3,4), \\ (\alpha-1)/2 & \text{if } \alpha \in [2,3], \\ 1/2 & \text{if } \alpha \in (0,2). \end{cases} \tag{2.15}$$

A condition of the form $p = \omega(n^\delta)$ for some $\delta > \delta^*(\alpha)$ will turn out to play in important role in convergence of the fourth moment and in the CLT for $\mathrm{tr}(\mathbf{R}^2)$. The next theorem sheds light on the convergence of the fourth moment.

Theorem 2.6. *For $\alpha \in (0,3) \cup (3,4)$ and assuming that $p = \omega(n^\delta)$ for some $\delta > \delta^*(\alpha)$, we have*

$$\lim_{n\to\infty} \frac{\mathbb{E}\big[(\mathrm{tr}(\mathbf{R}^2) - \mathbb{E}[\mathrm{tr}(\mathbf{R}^2)])^4\big]}{\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))^2} = 3 . \tag{2.16}$$

Moreover, if $p = o(n^\delta)$ for some $\delta < \delta^*(\alpha)$, then it holds

$$\lim_{n\to\infty} \frac{\mathbb{E}\big[(\operatorname{tr}(\mathbf{R}^2) - \mathbb{E}[\operatorname{tr}(\mathbf{R}^2)])^4\big]}{\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))^2} = \infty\,.$$

Theorem 2.6 reveals a transition in the convergence of moments at $p = n^{\delta^*(\alpha)}$. The limit in the case $p = n^{\delta^*(\alpha)}$ depends in a delicate way on certain slowly varying functions[2] that can be expressed in terms of the function L in $\mathbb{P}(|X_{11}| > x) = x^{-\alpha}L(x)$. It should be pointed out that the requirement $\delta > \delta^*(\alpha)$ for (2.16) is specifically tailored to the convergence of the fourth moment. In general, for higher moments a slightly different condition will be required. The next result shows that we can obtain a larger region for p by only considering the leading term (in the sense of convergence in distribution) of the decomposition $T_1 + T_2$. Recall that $T_1/\sigma_n = o_{\mathbb{P}}(1)$ if $\alpha > 3$, and $T_2/\sigma_n = o_{\mathbb{P}}(1)$ if $\alpha < 3$. Interestingly, in the case $\alpha \in (3,4)$, the crucial condition $\delta > \delta^*(\alpha)$ can be dropped if one focuses on the leading term only, as the following result shows.

Proposition 2.7. *If $\alpha \in (0,3]$ then*

$$\lim_{n\to\infty} \frac{\mathbb{E}[T_1^4]}{\mathbb{E}[T_1^2]^2} = \begin{cases} 3 & \text{if } p = \omega(n^\delta) \text{ for some } \delta > \delta^*(\alpha)\,,\\ \infty & \text{if } p = o(n^\delta) \text{ for some } \delta < \delta^*(\alpha)\,, \end{cases}$$

with $\delta^*(\alpha)$ defined in (2.15). If $\alpha \in [3,4)$, we have

$$\lim_{n\to\infty} \frac{\mathbb{E}[T_2^4]}{\mathbb{E}[T_2^2]^2} = 3\,. \tag{2.17}$$

Following the lines of the proof of Proposition 2.7, one can show that (2.17) remains valid if the regular variation assumption is replaced by $\mathbb{E}[X_{11}^4] < \infty$. The difference between the conditions on δ in Theorem 2.6 and Proposition 2.7, respectively, is due to the fact that the $o_{\mathbb{P}}(1)$-term T_1/σ_n might have a diverging fourth moment for $\alpha \in (3,4)$ which the following lemma asserts.

Lemma 2.8. *For $\alpha \in (3,4)$ we have*

$$\lim_{n\to\infty} \frac{\mathbb{E}[T_1^4]}{\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))^2} = \begin{cases} 0 & \text{if } p = \omega(n^\delta) \text{ for some } \delta > \delta^*(\alpha)\,,\\ \infty & \text{if } p = o(n^\delta) \text{ for some } \delta < \delta^*(\alpha)\,. \end{cases}$$

For $\alpha \in (0,3)$ it holds

$$\lim_{n\to\infty} \frac{\mathbb{E}[T_2^4]}{\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))^2} = 0.$$

3. Main results

Now we are ready to present a CLT for $\operatorname{tr}(\mathbf{R}^2)$ as the main result of this paper. Again we assume that the distribution of X_{11} is symmetric and regularly varying with index $\alpha \in (0,4)$. From (2.7) and (2.13) recall that

$$\mu_n = p + \frac{p(p-1)}{n} \quad \text{and} \quad \sigma_n^2 = 2p^2 n\big(\beta_4^2 + 2n^{-3}\big)\,.$$

[2]In our proofs, we often use the Potter bounds which guarantee that $\lim_{n\to\infty} \ell(n)/n^\varepsilon \to 0$ and $\lim_{n\to\infty} \ell(n)n^\varepsilon \to \infty$ for any positive slowly varying function ℓ and $\varepsilon > 0$ (see for instance [5]).

Theorem 3.1. *Let $\alpha \in (0,4)$ and $p = p_n \to \infty$. If $\alpha \in (0,3]$, additionally assume $p = \omega(n^\delta)$ for some $\delta > \delta^*(\alpha)$ with $\delta^*(\alpha)$ defined in* (2.15)*. Then, as $n \to \infty$, $\operatorname{tr}(\mathbf{R}^2)$ satisfies*

$$\frac{\operatorname{tr}(\mathbf{R}^2) - \mu_n}{\sigma_n} \stackrel{\mathrm{d}}{\to} N(0,1)\,.$$

Theorem 3.1 reveals the dependence on α in the variance but also that we need a minimal growth rate for p in the case $\alpha \in (0,3)$. This is solely because of T_1 giving rise to high order terms in the martingale differences in Lemmas 4.5 and 4.6 that are essential in the proof of Theorem 3.1 which can be found in Section 4. What is interesting is that we acquire less strict conditions for convergence in distribution to the standard normal distribution than for convergence of the centralized fourth moment divided by the squared variance; see Theorem 2.6 for details. The extra condition on p when $\alpha \in [3,4)$ that is present in Theorem 2.6 but not in Theorem 3.1 is due to the fact that we do not need to consider T_1 when $\alpha > 3$ in the martingale differences, while T_1 needs to be considered in the fourth moment calculations. More specifically, we need convergence to zero for $\mathbb{E}[T_1^4]/\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))^2$ when $\alpha \in [3,4)$ as stated in Lemma 2.8 for (2.16) to hold.

Since $\sigma_n^2 = 4p^2n^{-2}$ for $\alpha \in (3,4)$, we obtain the following corollary of Theorem 3.1.

Corollary 3.2. *Let $\alpha \in (3,4)$ and $p = p_n \to \infty$, then as $n \to \infty$ we have*

$$\frac{\operatorname{tr}(\mathbf{R}^2) - \mu_n}{2pn^{-1}} \stackrel{\mathrm{d}}{\to} N(0,1)\,.$$

This means that for $\alpha > 3$ the normalizing sequences in the CLT are universal and p may tend to infinity at arbitrary speeds. The situation is quite different for $\alpha < 3$, when both the variance $\sigma_n^2 \stackrel{\text{sl.v.}}{=} p^2 n^{1-\max\{\alpha,2\}}$ (Lemma 2.4) and the minimal growth of p depend on α.

Furthermore, a careful inspection of the proof of Theorem 3.1 in the case $\alpha \in (3,4)$ yields the following result.

Theorem 3.3. *Let X_{11} follow a symmetric distribution with finite fourth moment and assume $p = p_n \to \infty$. Then it holds*

$$\frac{\operatorname{tr}(\mathbf{R}^2) - \mu_n}{2pn^{-1}} \stackrel{\mathrm{d}}{\to} N(0,1)\,, \qquad n \to \infty\,.$$

With slightly more effort it is possible to remove the symmetry assumption on X_{11} in Theorem 3.3 at the cost of imposing the condition $\mathbb{E}[|X_{11}|^k] < \infty$ for a suitably large $k \geq 4$ that might depend on the relationship between p and n. However, this result will not be pursued in this work.

3.1. **Simulations.** To illustrate the role of α and the need of symmetry we will perform some simulations. In order to use Lemma 2.4 in the case $\alpha \in (2,3]$, we need the slowly varying function L or, more precisely, the value $L(n^{1/2})$. Let Z_1, Z_2 be iid Pareto random variables with parameter $\alpha > 2$, that is, $\mathbb{P}(Z_1 > x) = x^{-\alpha}$, $x > 1$. To generate a symmetric X_{11} with unit variance, we let $X_{11} \stackrel{\mathrm{d}}{=} (Z_1 - Z_2)\operatorname{Var}(Z_1 - Z_2)^{-0.5}$. Since Z_1, Z_2 are iid and Pareto-distributed we have

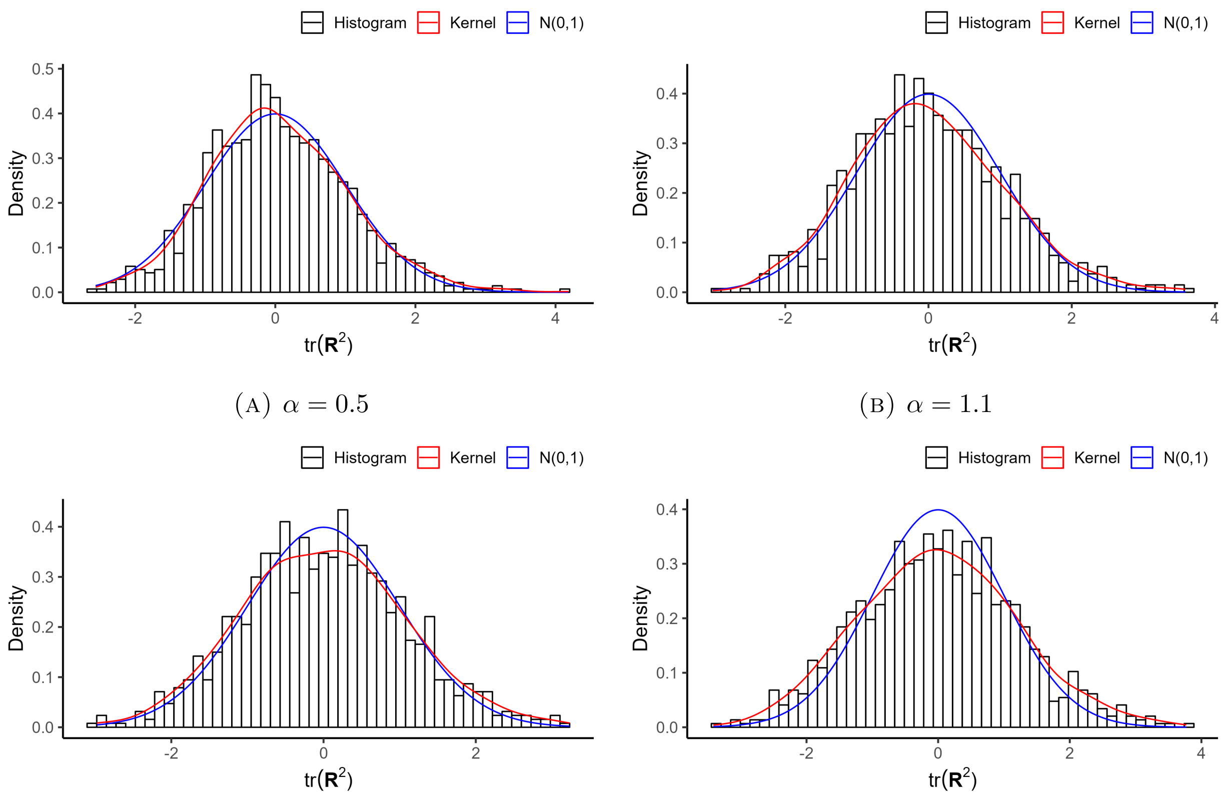
$$\operatorname{Var}(Z_1 - Z_2) = \frac{2\alpha}{(\alpha-1)^2(\alpha-2)} =: K_\alpha.$$

Using Lemma A.1, it is staightforward to show that the distribution of X_{11} has regularly varying tail with index α. In the notation of Lemma A.1, we let $S_2 = Z_1 + (-Z_2)$ and since, for $x > 0$,

$\mathbb{P}(Z_1 > x) = \overline{F}(x)$ and $\mathbb{P}(-Z_2 > x) = 0$ we get $c_1^+ = 1$ and $c_2^+ = 0$. We also find that $c_1^- = 0$ and $c_2^- = 1$ since $\mathbb{P}(Z_1 \le -x) = 0$ and $\mathbb{P}(-Z_2 \le -x) = \overline{F}(x)$. This satisfies (A.7) and it is obvious that also (A.8) holds. Now Lemma A.1 implies that $\mathbb{P}(Z_1 - Z_2 > x) \sim x^{-\alpha}$ and $\mathbb{P}(-Z_1 + Z_2 > x) \sim x^{-\alpha}$ yielding that $\mathbb{P}(|Z_1 - Z_2| > x) \sim 2x^{-\alpha}$, as $x \to \infty$. Finally, we get by using the latter

$$\mathbb{P}(|X_{11}| > x) = \mathbb{P}(|Z_1 - Z_2| > xK_\alpha^{0.5}) \sim x^{-\alpha} 2K_\alpha^{-0.5\alpha}$$

meaning that $L(x) \sim 2K_\alpha^{-0.5\alpha}$. Thus, $L(n^{1/2}) \approx 2K_\alpha^{-0.5\alpha}$ and we can approximate σ_n^2 by using Lemma 2.4. Figure 1 provides the simulation results and we observe a good fit of both the histogram and kernel density to the density of the standard normal distribution for all values of α.



(A) $\alpha = 0.5$ (B) $\alpha = 1.1$

(C) $\alpha = 2.8$ (D) $\alpha = 3.9$

FIGURE 1. Simulations of the central limit theorem for $\mathrm{tr}(\mathbf{R}^2)$ for different values of α and $p = 400$, $n = 1000$ with 1000 repetitions. Here $X_{11} \stackrel{\mathrm{d}}{=} (Z_1 - Z_2)K_\alpha^{-0.5}$ is symmetric with Z_1, Z_2 iid Pareto-distributed random variables.

To get some understanding of the behavior in the tails we present the Q-Q-plots seen in Figure 2. In the cases with small α the tails deviate the most from the straight line, which is not unexpected since the tails get heavier for smaller value of the shape parameter α in a Pareto distribution. Generally the Q-Q-plots support our theoretical findings, but also indicate that the convergence rates might be slower in more heavy-tailed scenarios.

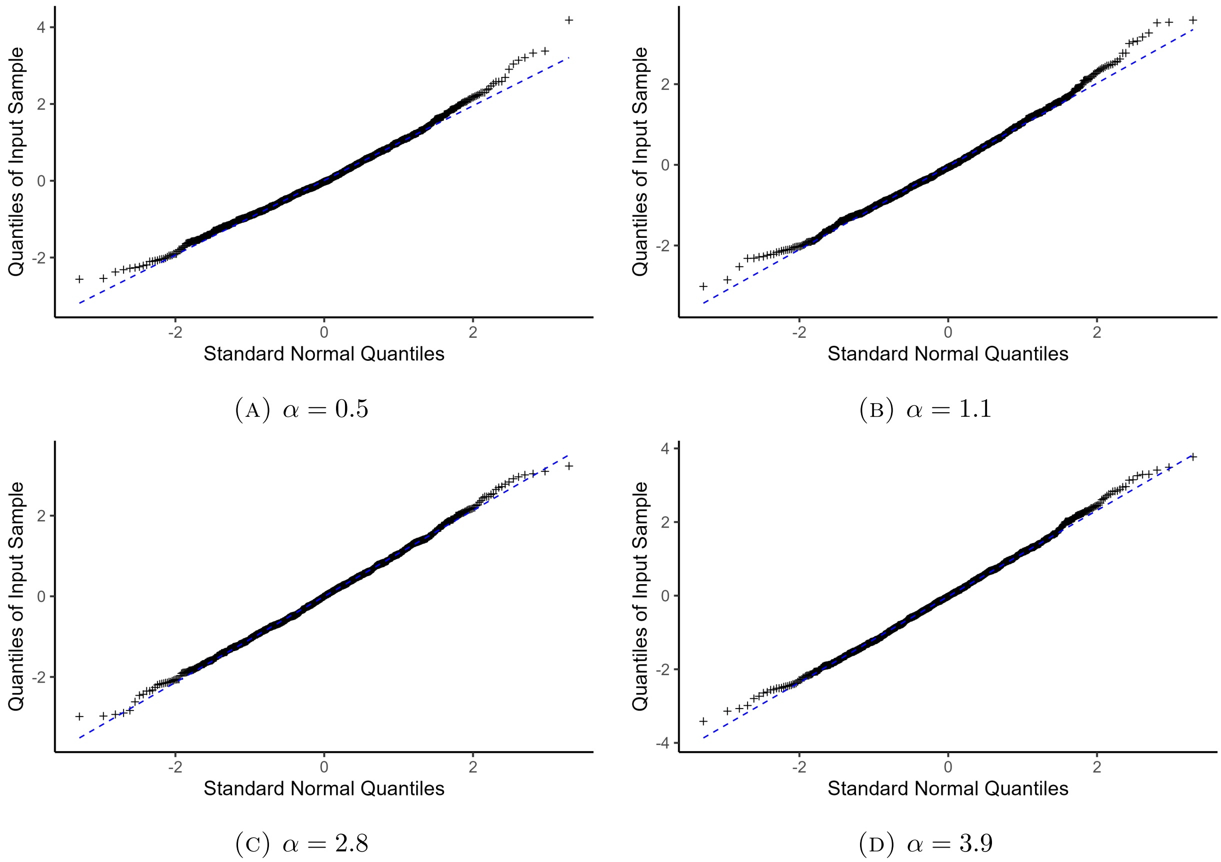
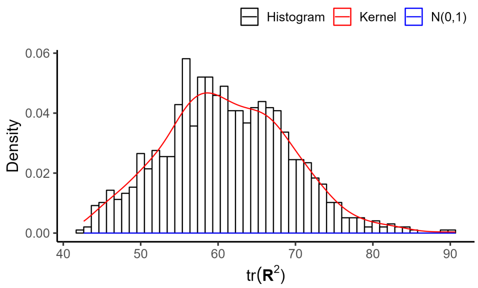
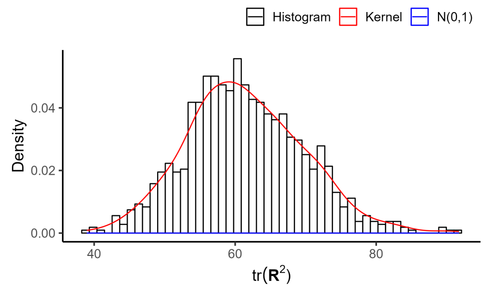


FIGURE 2. Q-Q-plots for the CLT of $\mathrm{tr}(\mathbf{R}^2)$ for different values of α and $p = 400$, $n = 1000$ with 1000 repetitions. Here the input sample $X_{11} \stackrel{\mathrm{d}}{=} (Z_1 - Z_2)K_\alpha^{-0.5}$ is symmetric where Z_1, Z_2 are iid Pareto-distributed random variables.

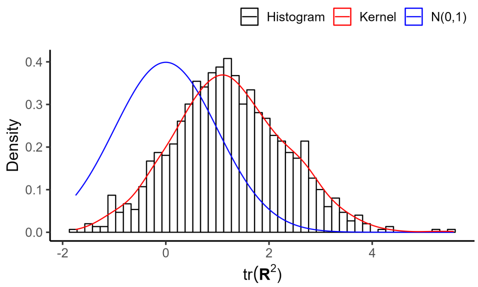
3.2. **Some comments on the non-symmetric case.** The assumption that X_{11} is symmetric is not only of significant technical importance in our proofs, but also indispensable for the validity of the CLT presented in Theorem 3.1 in general. To analyze the requirement of symmetry and the role that specific distributions may have, we simulate from two different non-symmetric distributions. For the first case, define the entries $X_{11} \stackrel{\mathrm{d}}{=} Z_1 - \mathbb{E}[Z_1]$ of the data matrix where Z_1 follows a Pareto distribution with shape parameter α which is equal to the tail index α in (1.3), and scale $x_m = 1$. For the second case, generate instead the entries $X_{11} \stackrel{\mathrm{d}}{=} T^2 - \mathbb{E}[T^2]$, where T follows a t-distribution with 2α degrees of freedom. Note that in both cases X_{11} is regularly varying with index α. The simulation results are shown in Figure 3. An interesting observation is that in both cases of different non-symmetric distributions, the same shift in the mean appears to be present. The main consequence of symmetry violation is that the limiting distribution resembles a normal distribution but with larger variance and some considerable shift in the mean is seen in the top left plots where α is small. The shift in mean for small values of α can be partially explained.



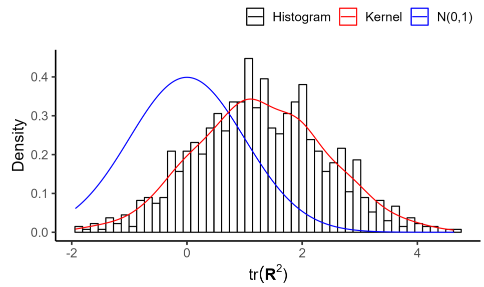
(A) $\alpha = 1.1$



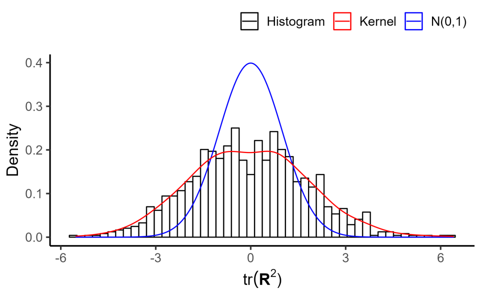
(B) $\alpha = 1.1$



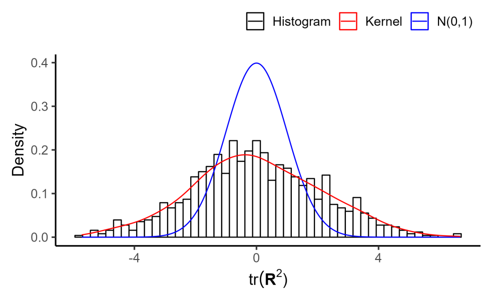
(C) $\alpha = 1.3$



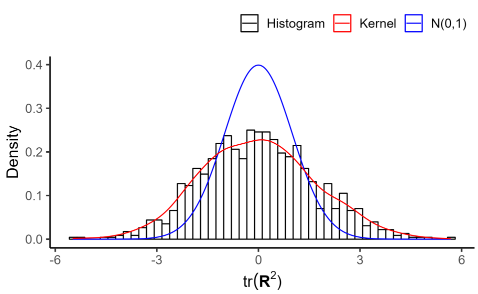
(D) $\alpha = 1.3$



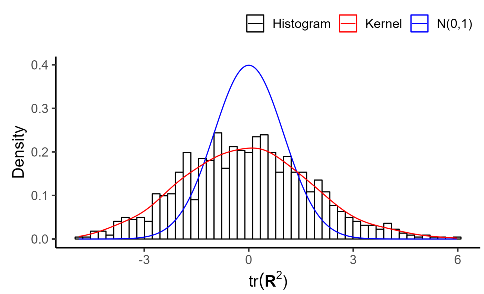
(E) $\alpha = 3.2$



(F) $\alpha = 3.2$



(G) $\alpha = 3.5$



(H) $\alpha = 3.5$

FIGURE 3. Simulations of the central limit theorem for $\mathrm{tr}(\mathbf{R}^2)$ for different values of α and $p = 400$, $n = 1000$ with 1000 repetitions. Here the plots on the left side have $X_{11} \stackrel{\mathrm{d}}{=} Z_1 - \mathbb{E}[Z_1]$, where Z_1 is Pareto-distributed with scale $x_m = 1$ and shape α, while the plots on the right side have $X_{11} \stackrel{\mathrm{d}}{=} T^2 - \mathbb{E}[T^2]$, where T is t-distributed with 2α degrees of freedom.

Some major differences arise when X_{11} does not have a symmetric distribution. To start with, from (2.6) we get

$$\begin{aligned}\mathbb{E}[\operatorname{tr}(\mathbf{R}^2)] &= \mu_n + \sum_{\substack{i_1,i_2=1\\ i_1\neq i_2}}^{p}\sum_{\substack{t_1,t_2=1\\ t_1\neq t_2}}^{n} \mathbb{E}\big[Y_{i_1t_1}Y_{i_1t_2}Y_{i_2t_1}Y_{i_2t_2}\big]\\ &= \mu_n + p(p-1)n(n-1)(\mathbb{E}[Y_{11}Y_{12}])^2\,.\end{aligned}$$

For all $y, \beta > 0$, we have

$$\frac{1}{y^\beta} = \frac{1}{\Gamma(\beta)}\int_0^\infty \exp(-ty)t^{\beta-1}\,\mathrm{d}t\,,$$

where Γ denotes the Gamma function. Combining this representation with Fubini's theorem, we deduce

$$\mathbb{E}[Y_{11}Y_{12}] = \int_0^\infty \big(\mathbb{E}\left[X_{11}\exp(-sX_{11}^2)\right]\big)^2\varphi^{n-2}(s)ds, \tag{3.1}$$

where $\varphi(s) = \mathbb{E}[\exp(-sX_{11}^2)], s > 0$, denotes the Laplace transform of X_{11}^2. We conclude that $\mathbb{E}[Y_{11}Y_{12}] \geq 0$ with equality if and only if the distribution of X_{11} is symmetric. In general, the exact calculation of $\mathbb{E}[Y_{11}Y_{12}]$ from (3.1) is rather involved and highly dependent on the specific non-symmetric distribution at hand. Assuming $\mathbb{E}[|X_{11}|^\eta] < \infty$ for some $\eta > 1$, Lemma 4.2 in [8] guarantees that $\mathbb{E}[Y_{11}Y_{12}] = o\big(n^{-\eta'}\big)$ for any $\eta' < \eta$. This implies that

$$\mathbb{E}[\operatorname{tr}(\mathbf{R}^2)] = \mu_n + p^2\, n^{-2(\eta'-1)}\, o(1)\,.$$

In the special case $\limsup_{n\to\infty} p/n < \infty$, we conclude that $\mathbb{E}[\operatorname{tr}(\mathbf{R}^2)] = \mu_n + o(1)$ if $\mathbb{E}[|X_{11}|^\eta] < \infty$ for some $\eta > 2$. This partially explains the shift in mean in our simulations for the non-symmetric case with small values of α. Finally, we would like to point out that non-symmetry of X_{11} also affects the variance $\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))$ and Lemma 2.4 no longer holds.

4. Proof of the main theorem

4.1. **Preparations.** We will need a few results about the entries of $\mathbf{Y}$ defined in (1.2). For integers $k_1, \ldots, k_r$, we recall the notation $\beta_{2k_1,\ldots,2k_r} := \mathbb{E}[Y_{11}^{2k_1}\cdots Y_{1r}^{2k_r}]$. The following lemma is a special case of [12, Lemma 3.2].

Lemma 4.1. *For any distribution of X_{11}, it holds that $\beta_2 = 1/n$ and*

$$\begin{aligned}\beta_4 &= \frac{1}{n} - (n-1)\beta_{2,2}\,, \qquad \beta_{4,2} = \frac{1}{2}\beta_{2,2} - \frac{n-2}{2}\beta_{2,2,2}\,,\\ \beta_6 &= \frac{1}{n} - \frac{3(n-1)}{2}\beta_{2,2} + \frac{(n-1)(n-2)}{2}\beta_{2,2,2}\,,\\ \beta_{6,2} &= \frac{1}{2}\beta_{2,2} - \frac{5(n-2)}{6}\beta_{2,2,2} + \frac{(n-2)(n-3)}{3}\beta_{2,2,2,2} - \beta_{4,4}\,,\\ \beta_{4,2,2} &= \frac{1}{3}\beta_{2,2,2} + \frac{3-n}{3}\beta_{2,2,2,2}\,,\\ \beta_8 &= \frac{1}{n} + 2(1-n)\beta_{2,2} + \Big(\frac{4n^2}{3} - 4n + \frac{8}{3}\Big)\beta_{2,2,2}\end{aligned}$$

$$+\Big(\frac{-n^3}{3}+2n^2-\frac{11n}{3}+2\Big)\beta_{2,2,2,2}+(n-1)\beta_{4,4}.$$

Throughout this paper, we will use the notation

$$\widetilde{\beta}_{2k_1,\ldots,2k_r}:=\mathbb{E}[(Y_{11}^2-n^{-1})^{k_1}\cdots(Y_{1r}^2-n^{-1})^{k_r}]\,,$$

where $k_1,\ldots,k_r$ are positive integers. It is easy to see that $\widetilde{\beta}_2=0$. The next lemma studies the asymptotic behavior of $\widetilde{\beta}_{2k_1,\ldots,2k_r}$.

Lemma 4.2. *Let $\alpha\in(0,4)$ and assume that $\mathbb{P}(|X_{11}|>x)=x^{-\alpha}L(x)$ for $x>0$, where L is a slowly varying function. If $\mathbb{E}[X_{11}^2]$ is finite, we additionally assume that $\mathbb{E}[X_{11}^2]=1$. Define the Y_{kn}'s as in* (1.2) *and consider integers $k_1,\ldots,k_r\ge 1$. Then it holds, as $n\to\infty$,*

$$\widetilde{\beta}_{2k_1,\ldots,2k_r}\begin{cases}\sim\beta_{2k_1,\ldots,2k_r}\,, & \text{if } \min(k_1,\ldots,k_r)\ge 2\,,\\ =O(\beta_{2k_1,\ldots,2k_r})\,, & \text{if } \min(k_1,\ldots,k_r)=1\,.\end{cases}$$

Proof. For any real numbers $b_1,\ldots,b_k$ we have the identity

$$\prod_{i=1}^k(b_i-\tfrac{1}{n})=\sum_{m=0}^k\sum_{\substack{S\subseteq\{1,\ldots,k\}\\|S|=m}}(-1)^m n^{-m}\prod_{i\in S^c}b_i\,,$$

where $S^c=\{1,\ldots,k\}\backslash S$ denotes the complement of S. Applying this identity for $k=k_1+\cdots+k_r$ with positive integers $k_1,\ldots,k_r$ and

$$b_i=\begin{cases}Y_{11}^2\,, & \text{if } i=1,\ldots,k_1\\ Y_{12}^2\,, & \text{if } i=k_1+1,\ldots,k_1+k_2\\ \vdots & \\ Y_{1r}^2\,, & \text{if } i=k_1+\cdots+k_{r-1}+1,\ldots,k_1+\cdots+k_r\,,\end{cases}$$

we obtain

$$\begin{aligned}\widetilde{\beta}_{2k_1,\ldots,2k_r}&=\mathbb{E}\prod_{j=1}^r\prod_{\ell=1}^{k_j}(Y_{1j}^2-\tfrac{1}{n})=\mathbb{E}\prod_{i=1}^k(b_i-\tfrac{1}{n})\\&=\sum_{m=0}^k\sum_{\substack{S\subseteq\{1,\ldots,k\}\\|S|=m}}(-1)^m n^{-m}\,\mathbb{E}\prod_{i\in S^c}b_i\,.\end{aligned}$$

Using the shorthand notation $b(S^c):=\mathbb{E}\prod_{i\in S^c}b_i$ and observing that $b(S)=\beta_{2k_1,\ldots,2k_r}$, we deduce that

$$\widetilde{\beta}_{2k_1,\ldots,2k_r}=\beta_{2k_1,\ldots,2k_r}+\sum_{m=1}^k(-1)^m\sum_{\substack{S\subseteq\{1,\ldots,k\}\\|S|=m}}n^{-m}\,b(S^c)\,.$$

Without loss of generality we assume that the k_i's are ordered, that is, $k_1\ge\cdots\ge k_r$.

If $k_r\ge 2$, one can see from Lemma 2.1 that $n^{-m}\,b(S^c)=o(\beta_{2k_1,\ldots,2k_r})$ for any $S\subseteq\{1,\ldots,k\}$ with cardinality $1\le m\le k$. Therefore, we conclude that

$$\widetilde{\beta}_{2k_1,\ldots,2k_r}\sim\beta_{2k_1,\ldots,2k_r}\,,\qquad n\to\infty\,,k_r\ge 2.$$

If $k_r = 2$, it one can analogously get from Lemma 2.1 that $n^{-m}\, b(S^c) = O(\beta_{2k_1,\dots,2k_r})$ for any $S \subseteq \{1,\dots,k\}$ with cardinality $1 \le m \le k$. Note that, for example, for $S = \{k\}$ we have $n^{-1}\, b(S^c) = n^{-1}\, b(\{1,\dots,k-1\}) = n^{-1}\beta_{2k_1,\dots,2k_{r-1}}$, which is of the same order as $\beta_{2k_1,\dots,2k_{r-1},2}$. We conclude that

$$\widetilde{\beta}_{2k_1,\dots,2k_r} = O(\beta_{2k_1,\dots,2k_r})\,, \qquad n \to \infty\,, k_r = 1,$$

completing the proof of the lemma. □

4.2. **Proof of Theorem 3.1.** It will be convenient to write our statistics as a sum of martingale differences. In what follows, the notation

$$\tilde{\mathbf{x}}_i = (X_{i1},\dots,X_{in})\,, \qquad i \in \{1,\dots,p\}$$

will be helpful. We consider the filtration $(\mathcal{F}_j)_{j\ge 0}$, where $\mathcal{F}_j$ is the σ-algebra generated by $\{\tilde{\mathbf{x}}_1,\dots,\tilde{\mathbf{x}}_j\}$. For $j = 1,\dots,p$, define

$$M_{j,1} := \mathbb{E}[T_1|\mathcal{F}_j] - \mathbb{E}[T_1|\mathcal{F}_{j-1}] \quad \text{ and } \quad M_{j,2} := \mathbb{E}[T_2|\mathcal{F}_j] - \mathbb{E}[T_2|\mathcal{F}_{j-1}]$$

and observe that $T_i = \sum_{j=1}^p M_{j,i}$ for $i = 1, 2$.

Then $(M_{j,1})_{j\ge1}$, $(M_{j,2})_{j\ge1}$ and $(M_{j,3})_{j\ge1} := (M_{j,1} + M_{j,2})_{j\ge1}$ are martingale difference sequences with respect to the filtration $(\mathcal{F}_j)_{j\ge0}$. The following lemma provides explicit formulas for $M_{j,1}$ and $M_{j,2}$.

Lemma 4.3. *For $j = 1,\dots,p$ it holds*

$$M_{j,1} = 2\sum_{i=1}^{j-1}\sum_{t=1}^{n}(Y_{it}^2 - n^{-1})(Y_{jt}^2 - n^{-1}) \quad \text{ and } \quad M_{j,2} = 4\sum_{i=1}^{j-1}\sum_{1\le t_1<t_2\le n} Y_{it_1}Y_{it_2}Y_{jt_1}Y_{jt_2}.$$

Proof. We start by considering $M_{j,1}$. Using the definition of T_1, we have for $k \in \{0,\dots,p\}$

$$\mathbb{E}[T_1|\mathcal{F}_k] = 2\sum_{1\le i_1<i_2\le p}\sum_{t=1}^{n}\mathbb{E}\big[(Y_{i_1t}^2 - n^{-1})(Y_{i_2t}^2 - n^{-1})\big|\mathcal{F}_k\big]\,.$$

Since $Y_{it}^2 - n^{-1}$ is centered and independent of $\mathcal{F}_k$ for $k < i$, and $\mathcal{F}_k$-measurable for $k \ge i$ we have $\mathbb{E}\big[(Y_{i_1t}^2 - n^{-1})(Y_{i_2t}^2 - n^{-1})\big|\mathcal{F}_k\big] = 0$ if $i_2 > k$. Therefore, we conclude that

$$\mathbb{E}[T_1|\mathcal{F}_k] = 2\sum_{1\le i_1<i_2\le k}\sum_{t=1}^{n}(Y_{i_1t}^2 - n^{-1})(Y_{i_2t}^2 - n^{-1})\,,$$

from which we easily deduce

$$\begin{aligned} M_{j,1} &= \mathbb{E}[T_1|\mathcal{F}_j] - \mathbb{E}[T_1|\mathcal{F}_{j-1}] \\ &= 2\Bigg(\sum_{1\le i_1<i_2\le j} - \sum_{1\le i_1<i_2\le j-1}\Bigg)\sum_{t=1}^{n}(Y_{i_1t}^2 - n^{-1})(Y_{i_2t}^2 - n^{-1}) \\ &= 2\sum_{i=1}^{j-1}\sum_{t=1}^{n}(Y_{i_1t}^2 - n^{-1})(Y_{i_2t}^2 - n^{-1})\,. \end{aligned}$$

The proof for $M_{j,2}$ is completely analogous. □

We will use the following CLT for martingale differences.

Lemma 4.4 (e.g. Hall and Heyde [10]). *Let* $\{S_{ni} = \sum_{j=1}^{i} Z_{nj}, \mathcal{F}_{ni}, 1 \le i \le k_n, n \ge 1\}$ *be a zero-mean, square integrable martingale array with differences* Z_{nj}*. Suppose that* $\mathbb{E}[\max_{j=1,\dots,k_n} Z_{nj}^2]$ *is bounded in* n *and that*

$$\max_{j=1,\dots,k_n} |Z_{nj}| \xrightarrow{\mathbb{P}} 0 \quad \text{and} \quad \sum_{j=1}^{k_n} Z_{nj}^2 \xrightarrow{\mathbb{P}} 1\,.$$

Then we have $S_{nk_n} \xrightarrow{\mathrm{d}} N(0,1)$ *as* $n \to \infty$*.*

For convenience of notation, we set $T_3 = T_1 + T_2$. In view of (2.14), we will apply Lemma 4.4 to $\sigma_n^{-1} T_i = \sum_{j=1}^p \sigma_n^{-1} M_{j,i}$ with martingale differences $\sigma_n^{-1} M_{j,i}$, by considering $i=1$ when $\alpha < 3$, $i = 2$ when $\alpha \in (3,4)$ and $i=3$ when $\alpha = 3$. By definition, we have $\sigma_n^{-1}\mathbb{E}[T_i] = 0$ and (2.12) implies that $\sigma_n^{-2}\mathbb{E}[T_i^2] \to 1$, fulfilling the square integrability condition. Also note that $\mathbb{E}[\max_{j=1,\dots,p} \sigma_n^{-2} M_{j,i}^2]$ is bounded since

$$\mathbb{E}[\max_{j=1,\dots,p} \sigma_n^{-2} M_{j,i}^2] \le \sigma_n^{-2} \sum_{j=1}^p \mathbb{E}[M_{j,i}^2] = \frac{\mathbb{E}[T_i^2]}{\sigma_n^2} \to 1, \qquad n \to \infty.$$

Note that in the case $\alpha \in (0,3)$, we have assumed in Theorem 3.1 that $p = \omega(n^\delta)$ for some $\delta > \delta^*(\alpha)$ with $\delta^*(\alpha)$ defined in (2.15). The two conditions of Lemma 4.4 left to show are

$$\max_{j=1,\dots,p} |\sigma_n^{-1} M_{j,i}| \xrightarrow{\mathbb{P}} 0 \quad \text{and} \quad \sigma_n^{-2} \sum_{j=1}^p M_{j,i}^2 \xrightarrow{\mathbb{P}} 1\,, \qquad i = 1,2,3\,, \tag{4.1}$$

which is the content of Section 4.3 below. This concludes the proof of Theorem 3.1.

4.3. **Proof of (4.1).** We start with the first part of (4.1). To prove $\max_j |\sigma_n^{-1} M_{j,i}| \xrightarrow{\mathbb{P}} 0$, we note that by the union bound and Markov's inequality

$$\mathbb{P}\Big(\max_{j=1,\dots,p} |\sigma_n^{-1} M_{j,i}| > \varepsilon\Big) \le \sum_{j=1}^p \mathbb{P}(|\sigma_n^{-1} M_{j,i}| > \varepsilon) \le \varepsilon^{-4} \sum_{j=1}^p \frac{\mathbb{E}[M_{j,i}^4]}{\sigma_n^4}\,, \qquad \varepsilon > 0.$$

The following lemma and (4.4) assert that the right-hand side tends to zero.

Lemma 4.5. *If* $\alpha \in (0,3]$ *and* $p = \omega(n^\delta)$ *for some* $\delta > \delta^*(\alpha)$*, where* $\delta^*(\alpha)$ *is defined in* (2.15)*, we have*

$$\frac{1}{\sigma_n^4} \sum_{j=1}^p \mathbb{E}[M_{j,1}^4] \to 0\,, \qquad n \to \infty\,, \tag{4.2}$$

If $\alpha \in [3,4)$*, we have*

$$\frac{1}{\sigma_n^4} \sum_{j=1}^p \mathbb{E}[M_{j,2}^4] \to 0\,, \qquad n \to \infty\,. \tag{4.3}$$

Note that in the case $\alpha = 3$ (corresponding to $i = 3$), the convergence $\sigma_n^{-4} \sum_{j=1}^p \mathbb{E}[(M_{j,1} + M_{j,2})^4] \to 0$ follows directly since

$$\frac{1}{\sigma_n^4} \sum_{j=1}^p \mathbb{E}[(M_{j,1} + M_{j,2})^4] \le \frac{16}{\sigma_n^4} \sum_{j=1}^p \mathbb{E}[M_{j,1}^4] + \frac{16}{\sigma_n^4} \sum_{j=1}^p \mathbb{E}[M_{j,2}^4] \tag{4.4}$$

and we deduce from Lemma 4.5 that the above tends to zero if $\delta > \delta^*(\alpha)$ for $\alpha = 3$.

Proof of Lemma 4.5. Direct calculation yields

$$\mathbb{E}[M_{j,1}^4] = \mathbb{E}\Big[\Big(2\sum_{i=1}^{j-1}\sum_{t=1}^{n}\overline{Y}_{it}\overline{Y}_{jt}\Big)^4\Big]$$
$$= 16\sum_{i_1,i_2,i_3,i_4=1}^{j-1}\sum_{t_1,t_2,t_3,t_4=1}^{n}\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{i_2t_2}\overline{Y}_{i_3t_3}\overline{Y}_{i_4t_4}]\mathbb{E}[\overline{Y}_{jt_1}\overline{Y}_{jt_2}\overline{Y}_{jt_3}\overline{Y}_{jt_4}]\,.$$

To get nonzero summands we have two possibilities to pair the i's. Either all are equal, $i_1 = i_2 = i_3 = i_4$, or we get two pairs, e.g., $i_1 = i_2$ and $i_3 = i_4$ but $i_1 \neq i_3$.

$\mathbb{E}[\overline{Y}_{jt_1}\overline{Y}_{jt_2}\overline{Y}_{jt_3}\overline{Y}_{jt_4}]$	$\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{i_1t_2}\overline{Y}_{i_1t_3}\overline{Y}_{i_1t_4}]$	$\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{i_1t_2}]\mathbb{E}[\overline{Y}_{i_2t_3}\overline{Y}_{i_2t_4}]$	
Case 1	**Case 1**	**Case 1**	**Case 2**
$\widetilde{\beta}_8$	$\widetilde{\beta}_8$	$\widetilde{\beta}_4^2$	
$\widetilde{\beta}_{6,2}$	$\widetilde{\beta}_{6,2}$	$\widetilde{\beta}_4\widetilde{\beta}_{2,2}$	
$\widetilde{\beta}_{4,4}$	$\widetilde{\beta}_{4,4}$	$\widetilde{\beta}_{2,2}^2$	$\widetilde{\beta}_4^2$
$\widetilde{\beta}_{4,2,2}$	$\widetilde{\beta}_{4,2,2}$	$\widetilde{\beta}_4\widetilde{\beta}_{2,2}$	$\widetilde{\beta}_{2,2}^2$
$\widetilde{\beta}_{2,2,2,2}$	$\widetilde{\beta}_{2,2,2,2}$	$\widetilde{\beta}_{2,2}^2$	

TABLE 1. Possible terms of $\widetilde{\beta}_{2k_1,\dots,2k_r}$ that occur in $\mathbb{E}[M_{j,1}^4]$. Case 1 and Case 2 corresponds to the fact that depending on which t's that pair up we can get different terms. e.g. $t_1 = t_2$, $t_3 = t_4$ and $t_1 = t_2$, $t_3 = t_4$ will yield different results in $\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{i_1t_2}]\mathbb{E}[\overline{Y}_{i_2t_3}\overline{Y}_{i_2t_4}]$.

Table 1 shows the possible (nonzero) terms in $\mathbb{E}[M_{j,1}^4]$. Each term is a product of certain $\widetilde{\beta}_{2k_1,\dots,2k_r}$ and - to get an upper bound on the order of its contribution to $\sum_{j=1}^p \mathbb{E}[M_{j,1}^4]$ - should be multiplied by a factor $p^{h+1}n^d$, where d denotes the number of distinct t's. The p^{h+1} factor comes from the number h of distinct i's and then summing from $j = 1$ to p in (4.2). Thus, using Lemma 2.1 combined with Lemma 4.2 one can check that the highest order term for $\alpha \in [2, 3]$ is $\widetilde{\beta}_8^2 np^2$, which (up to a slowly varying function) behaves like $n^{-\alpha+1}p^2$. Now using Lemma 2.4 for $\alpha \in [2, 3]$ we get

$$\frac{\widetilde{\beta}_8^2 np^2}{\sigma_n^4} \overset{\text{sl.v.}}{=} n^{\alpha-1}p^{-2}$$

which goes to 0 if $p = \omega(n^\delta)$ for some $\delta > \delta^*(\alpha)$ with $\delta^*(\alpha)$ defined in (2.15). Similarly one can find for the other case with $i_1 = i_2$ and $i_3 = i_4$, but $i_1 \neq i_3$, that the dominating term is $\widetilde{\beta}_8\widetilde{\beta}_4^2 n^2p^3$. Combining Lemmas 2.4 and 2.1 one can find that $\widetilde{\beta}_8\widetilde{\beta}_4^2 n^2p^3/\sigma_n^4 \to 0$ if $p = \omega(n^\delta)$ for some $\delta > \alpha/2 - 1$ with $\alpha \in [2, 3]$.

Now for $\alpha \in (0, 2)$ the same terms will dominate in the case where all i's are equal and here $\beta_8^2 np^2 \sim n^{-1}p^2C_{8,\alpha}^2$ and $\sigma_n^4 \sim 4p^4n^{-2}(1-\alpha/2)^4$ by Lemmas 2.1 and 2.4, respectively. Thus it is easy to see that $\beta_8^2np^2/\sigma_n^4 \to 0$ if $p = \omega(n^\delta)$ for some $\delta > \delta^*(\alpha)$. In the case $i_1 = i_2$ and $i_3 = i_4$, but $i_1 \neq i_3$, one of $\widetilde{\beta}_8\widetilde{\beta}_4^2 np^3$ and $\widetilde{\beta}_{4,4}\widetilde{\beta}_4^2 n^2p^3$ is the dominating term but still σ_n^4 is of higher order. This completes the proof of (4.2)

Regarding (4.3) we proceed similarly. By direct calculation we have

$$\mathbb{E}[M_{j,2}^4] = \mathbb{E}\Big[\Big(4\sum_{i=1}^{j-1}\sum_{\substack{t_1,t_2=1\\t_1<t_2}}^{n} Y_{it_1}Y_{it_2}Y_{jt_1}Y_{jt_2}\Big)^4\Big]$$

$$= 256\sum_{i_1,\ldots,i_4=1}^{j-1}\sum_{\substack{t_1,\ldots,t_8=1\\t_1<t_2,\ldots,t_7<t_8}}^{n} \mathbb{E}[Y_{i_1t_1}Y_{i_1t_2}Y_{i_2t_3}Y_{i_2t_4}Y_{i_3t_5}Y_{i_3t_6}Y_{i_4t_7}Y_{i_4t_8}]\mathbb{E}[Y_{jt_1}Y_{jt_2}Y_{jt_3}Y_{jt_4}Y_{jt_5}Y_{jt_6}Y_{jt_7}Y_{jt_8}]$$

and as before in order to get nonzero expectation we want no odd powers; so either all the i's are equal or we establish four pairs between the i's. Table 2 shows the possible terms that occur in $\mathbb{E}[M_{j,2}^4]$.

$\mathbb{E}[Y_{jt_1}\cdots Y_{jt_8}]$	$\mathbb{E}[Y_{i_1t_1}Y_{i_1t_2}\cdots Y_{i_1t_7}Y_{i_1t_8}]$	$\mathbb{E}[Y_{i_1t_1}Y_{i_1t_2}Y_{i_1t_3}Y_{i_1t_4}]\mathbb{E}[Y_{i_2t_5}Y_{i_2t_6}Y_{i_2t_7}Y_{i_2t_8}]$
	Case: All i's equal	**Case: Four pairs between i's**
$\beta_{4,4}$	$\beta_{4,4}$	$\beta_{2,2}^2$
$\beta_{4,2,2}$	$\beta_{4,4,2}$	$\beta_{2,2}^2$
$\beta_{2,2,2,2}$	$\beta_{2,2,2,2}$	$\beta_{2,2}^2$

TABLE 2. Possible terms of $\beta_{2k_1,\ldots,2k_r}$ that occur in $\mathbb{E}[M_{j,2}^4]$. Here we both include the case where all i's equal ($i_1 = i_2 = i_3 = i_4$) and the case where four pairs are formed between the i's (e.g. $i_1 = i_2$ and $i_3 = i_4$).

The highest order term in the case where all i's are equal is $\beta_{2,2,2,2}^2 n^4p^2$ if $\alpha \in (3,4)$ and combining Lemmas 2.1 and 2.4 we get

$$\frac{\beta_{2,2,2,2}^2 n^4p^2}{\sigma_n^4} \sim \frac{n^{-4}p^2\widetilde{C}_{2,2,2,2,\alpha}^2}{16p^4n^{-4}} \to 0.$$

In the case $\alpha = 3$ then all the terms will have the same order and the squared variance will instead become $\sigma_n^4 \sim 4p^4n^{-4}(C_{\alpha,4}^2L^2(n^{1/2})+2)^2$, so that the same conclusion holds. Now in the case of four pairs among the i's and $\alpha \in [3,4)$, the term $\beta_{2,2,2,2}\beta_{2,2}^2n^4p^3 \sim n^{-4}p^3$ dominates. Using Lemma 2.4 we have that $\sigma_n^4 \sim 16p^4n^{-4}$ for $\alpha \in (3,4)$ and it is obvious that $\beta_{2,2,2,2}\beta_{2,2}^2n^4p^3/\sigma_n^4 \to 0$ when $\alpha \in [3,4)$. This finishes the proof of (4.3). □

Now we turn to the second part of (4.1).

Lemma 4.6. *If $\alpha \in (0,3)$, it holds*

$$\frac{1}{\sigma_n^2}\sum_{j=1}^{p} M_{j,1}^2 \xrightarrow{\mathbb{P}} 1\,, \qquad n\to\infty\,, \tag{4.5}$$

where for $\alpha \in [2,3)$ we additionally assume $p = \omega(n^\delta)$ for some $\delta > \alpha/2 - 1$. If $\alpha \in (3,4)$, then it holds

$$\frac{1}{\sigma_n^2}\sum_{j=1}^{p} M_{j,2}^2 \xrightarrow{\mathbb{P}} 1\,, \qquad n\to\infty. \tag{4.6}$$

Finally, for $\alpha = 3$ it holds

$$\frac{1}{\sigma_n^2}\sum_{j=1}^{p} M_{j,3}^2 = \frac{1}{\mathbb{E}[T_1^2]+\mathbb{E}[T_2^2]}\sum_{j=1}^{p}\left(M_{j,1}^2+M_{j,2}^2+M_{j,1}M_{j,2}\right) \xrightarrow{\mathbb{P}} 1\,. \tag{4.7}$$

Proof. We start by first proving (4.5) and (4.6) for $\alpha \in (0,3)$ and $\alpha \in (3,4)$ respectively. Since $\sigma_n^{-2}\sum_{j=1}^{p}\mathbb{E}[M_{j,k}^2] \to 1$ for $k \in \{1,2\}$, it is sufficient to show

$$\frac{1}{\sigma_n^4}\mathbb{E}\Big[\Big(\sum_{j=1}^{p} M_{j,k}^2\Big)^2\Big] \to 1, \qquad n \to \infty\,.$$

Using Lemma 4.5 and Lemma 4.3, we have

$$\begin{aligned}\mathbb{E}\Big[\Big(\sum_{j=1}^{p} M_{j,1}^2\Big)^2\Big] &= o(\sigma_n^4) + \sum_{\substack{j_1,j_2=1\\ j_1\neq j_2}}^{p} \mathbb{E}\Big[M_{j_1,1}^2 M_{j_2,1}^2\Big] \\ &= o(\sigma_n^4) + 16\sum_{\substack{j_1,j_2=1\\ j_1\neq j_2}}^{p}\sum_{i_1,i_2=1}^{j_1-1}\sum_{i_3,i_4=1}^{j_2-1}\sum_{t_1,t_2=1}^{n}\sum_{t_3,t_4=1}^{n}\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{i_2t_2}\overline{Y}_{i_3t_3}\overline{Y}_{i_4t_4}\overline{Y}_{j_1t_1}\overline{Y}_{j_1t_2}\overline{Y}_{j_2t_3}\overline{Y}_{j_2t_4}].\end{aligned} \tag{4.8}$$

For the above expectation to be nonzero three cases are possible (up to a permutation of the i indices). Either all the i's are the same (which implies that the j indices take distinct values than the i's), or exactly 2 pairs are formed, e.g., $i_1 = i_2 \neq i_3 = i_4$ which amounts to two cases depending on if some i's may coincide with a j or not. The latter two cases we will call the second and third case respectively.

In the first case where all i's are equal, the possible terms are summarised in Table 3 using Lemma 2.1, Lemma 4.2 and equations (5.3), (5.4), (5.5) when $\alpha \in [2,3)$. Each term should also be multiplied by a factor n^d where d stands for the number of distinct t indices in $\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{i_1t_2}\overline{Y}_{i_1t_3}\overline{Y}_{i_1t_4}]$ in Table 3. Note also that in the first case the order of (4.8) with respect to p is p^3.

$\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{i_1t_2}\overline{Y}_{i_1t_3}\overline{Y}_{i_1t_4}]\mathbb{E}[\overline{Y}_{j_1t_1}\overline{Y}_{j_1t_2}]\mathbb{E}[\overline{Y}_{j_2t_3}\overline{Y}_{j_2t_4}]$	**Order**	
Term	$\alpha \in (0,2)$	$\alpha \in [2,3)$
$\widetilde{\beta}_8\widetilde{\beta}_4^2 n$	n^{-2}	$n^{-3\alpha/2+1}$
$\widetilde{\beta}_{6,2}\widetilde{\beta}_4\widetilde{\beta}_{2,2}n^2$	n^{-3}	$-n^{-3\alpha/2}$
$\widetilde{\beta}_{4,4}\widetilde{\beta}_4^2 n^2$	n^{-2}	$n^{-2\alpha+2}$
$\widetilde{\beta}_{4,4}\widetilde{\beta}_{2,2}^2 n^2$	n^{-4}	$n^{-2\alpha}$
$\widetilde{\beta}_{4,2,2}\widetilde{\beta}_4\widetilde{\beta}_{2,2}n^3$	n^{-3}	$-n^{-3\alpha/2}$
$\widetilde{\beta}_{4,2,2}\widetilde{\beta}_{2,2}^2 n^3$	n^{-4}	$-n^{-3\alpha/2-1}$
$\widetilde{\beta}_{2,2,2,2}\widetilde{\beta}_{2,2}^2 n^4$	n^{-4}	$O(n^{-3\alpha/2})$

TABLE 3. Terms of $\widetilde{\beta}_{2k_1,\ldots,2k_r}$ that occur in (4.8) when $i_1 = i_2 = i_3 = i_4$ and their order with respect to n. The $O(n^{-3\alpha/2})$ is due to the fact that we only have an upper bound for the order by (5.5).

The highest order term when $\alpha \in [2,3)$ is $\widetilde{\beta}_8\widetilde{\beta}_4^2np^3 \sim n^{-3\alpha/2+1}p^3C_{8,\alpha}C_{4,\alpha}^2L^3(n^{1/2})$ and also $\widetilde{\beta}_{4,4}\widetilde{\beta}_4^2n^2$ which obtains the same order when $\alpha = 2$. By Lemma 2.4, $\sigma_n^4 \sim 4p^4n^{2-2\alpha}L^4(n^{1/2})C_{4,\alpha}^4$ which yields $\widetilde{\beta}_8\widetilde{\beta}_4^2np^3/\sigma_n^4(\alpha) \to 0$ assuming $p = \omega(n^\delta)$ with $\delta > \alpha/2 - 1$. If $\alpha \in (0,2)$, then Lemma 2.4 yields $\sigma_n^4 \sim 4p^4n^{-2}(1-\alpha/2)^4$ which has greater order than all the terms in Table 3. Thus, we have shown that σ_n^{-4} times the contribution of the first case to (4.8) tends to zero.

Next, we turn to the second case. In this case, since we assume that the sets of j's and i's are disjoint, we can have either $i_1 = i_2 \neq i_3 = i_4$, or $i_1 = i_3 \neq i_2 = i_4$ that yield different terms (the case $i_1 = i_4$ and $i_2 = i_3$ yields the same terms as in $i_1 = i_3$ and $i_2 = i_4$ because of symmetry of the t's). This means that the terms in the sum of (4.8) can take two forms:

$$\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{i_1t_2}]\mathbb{E}[\overline{Y}_{i_3t_3}\overline{Y}_{i_3t_4}]\mathbb{E}[\overline{Y}_{j_1t_1}\overline{Y}_{j_1t_2}]\mathbb{E}[\overline{Y}_{j_2t_3}\overline{Y}_{j_2t_4}],$$
$$\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{i_1t_3}]\mathbb{E}[\overline{Y}_{i_2t_2}\overline{Y}_{i_2t_4}]\mathbb{E}[\overline{Y}_{j_1t_1}\overline{Y}_{j_1t_2}]\mathbb{E}[\overline{Y}_{j_2t_3}\overline{Y}_{j_2t_4}].$$

The general form of either of these products of expectations is $\widetilde{\beta}_4^{k_1}\widetilde{\beta}_{2,2}^{k_2}$ with $k_1 + k_2 = 4$. Since now we allow two pairs of the i's we get that in this case the contribution of such terms to (4.8) will be $\widetilde{\beta}_4^{k_1}\widetilde{\beta}_{2,2}^{k_2}n^{k_3}p^4$ where k_3 stands for how many of the t's that may vary freely, e.g. if $t_1 = t_2 = t_3 = t_4$ then $k_3 = 1$. Using (5.2) we find that $\widetilde{\beta}_{2,2} \sim -n^{-\alpha/2-1}L(n^{1/2})C_{4,\alpha}$ when $\alpha \in [2,3)$ which in conjunction with Lemma 2.4 yields

$$\frac{\widetilde{\beta}_4^{k_1}\widetilde{\beta}_{2,2}^{k_2}n^{k_3}p^4}{\sigma_n^4} \sim \frac{(-1)^{k_2}n^{-k_1\alpha/2}n^{-k_2(\alpha/2+1)}n^{k_3}p^4C_{4,\alpha}^{k_1+k_2}L^{k_1+k_2}(n^{1/2})}{4p^4n^{2-2\alpha}L^4(n^{1/2})C_{4,\alpha}^4} = (-1)^{k_2}n^{-k_2-2+k_3}\frac{1}{4}$$

where the last equality follows by $k_1 + k_2 = 4$. If $k_3 = 4$ this means that all the t's are pairwise different and hence $k_2 = 4$. If $k_3 = 3$ then two t's are equal but then $k_2 \geq 2$ and if $k_3 = 1$ then it is obvious that the exponent of n above is negative. Only when $k_3 = 2$ and $k_1 = 4$ with $i_1 = i_2$ and $i_3 = i_4$ is when we may get nonzero results. This results in the term $\widetilde{\beta}_4^4$. In view of (4.8) the asymptotic behavior of p in the second case is given by

$$\sum_{j_1,j_2=1}^{p}(j_1-1)(j_2-1) - \sum_{j=1}^{p}(j-1)^2 = \frac{1}{4}(p-1)^2p^2 - \frac{1}{6}p(2p^2-3p+1) \sim \frac{1}{4}p^4 \tag{4.9}$$

and now we get using (4.8) and Lemma 4.2 that

$$16\frac{\frac{1}{4}\widetilde{\beta}_4^4n^2p^4}{\sigma_n^4(\alpha)} \sim \frac{n^{-2\alpha+2}p^4L^4(n^{1/2})C_{4,\alpha}^4}{n^{-2\alpha+2}p^4L^4(n^{1/2})C_{4,\alpha}^4} = 1, \qquad n \to \infty.$$

When $\alpha \in (0,2)$ then $\widetilde{\beta}_4^{k_1}\widetilde{\beta}_{2,2}^{k_2}n^{k_3}p^4/\sigma_n^4(\alpha) \lesssim n^{-2-k_2+k_3}$ which is the same expression analyzed before. The dominating term is again $\widetilde{\beta}_4^4$ and similarly we get $\widetilde{\beta}_4^4n^2p^4/\sigma_n^4(\alpha) \to 1$ if $\alpha \in (0,2)$ by Lemmas 2.4 and 2.1. This concludes our analysis of the second case.

Now we turn to the third case, where we form two pairs among the i's and have some i's coincide with j_1 or j_2. If $i_1 = i_2$ and $i_3 = i_4$, then we must have either $j_1 = i_3$ or $j_2 = i_1$. If $j_2 = i_1$, then we get in the sum of (4.8)

$$E[\overline{Y}_{i_1t_1}\overline{Y}_{i_2t_2}\overline{Y}_{j_2t_3}\overline{Y}_{j_2t_4}]E[\overline{Y}_{i_3t_3}\overline{Y}_{i_3t_4}]E[\overline{Y}_{j_1t_1}\overline{Y}_{j_1t_2}]$$

which essentially yields the same cases as summarised in Table 3 but with the same or less order in p. Now if $i_1 \neq i_2$ then we get nonzero only if $i_2 = i_3 = i_4$ and the cases can be summarised in Table 4 by use of Lemma 2.1 and equations (5.3), (5.4) and (5.5).

$\mathbb{E}[\overline{Y}_{i_1t_2}\overline{Y}_{i_1t_3}\overline{Y}_{i_1t_4}]\mathbb{E}[\overline{Y}_{i_3t_1}\overline{Y}_{i_3t_3}\overline{Y}_{i_3t_4}]\mathbb{E}[\overline{Y}_{j_1t_1}\overline{Y}_{j_1t_2}]$	Order	
Term	$\alpha \in (0,2)$	$\alpha \in [2,3)$
$\widetilde{\beta}_6^2\widetilde{\beta}_4 n$	n^{-2}	$n^{-3\alpha/2+1}$
$\widetilde{\beta}_6\widetilde{\beta}_{4,2}\widetilde{\beta}_{2,2}n^2$	n^{-3}	$-n^{-3\alpha/2}$
$\widetilde{\beta}_{4,2}^2\widetilde{\beta}_{2,2}n^3$	n^{-3}	$-n^{-3\alpha/2}$
$\widetilde{\beta}_{4,2}^2\widetilde{\beta}_4 n^2$	n^{-3}	$n^{-3\alpha/2}$
$\widetilde{\beta}_{2,2,2}^2\widetilde{\beta}_4 n^3$	n^{-4}	$n^{-3\alpha/2-1}$
$\widetilde{\beta}_{4,2}\widetilde{\beta}_{2,2,2}\widetilde{\beta}_{2,2}n^3$	n^{-4}	$-n^{-3\alpha/2-1}$
$\widetilde{\beta}_{2,2,2}^2\widetilde{\beta}_{2,2}n^4$	n^{-4}	$n^{-3\alpha/2-1}$

TABLE 4. Terms of $\widetilde{\beta}_{2k_1,\dots,2k_r}$ that occur in (4.8) when $i_1 = j_2 \neq i_2$ but $i_2 = i_3 = i_4$ and their order with respect to n.

Since now $j_2 = i_1 \neq i_2$ and $i_2 = i_3 = i_4$ the order of (4.8) in p is p^3. Hence it is clear that when $\alpha \in (0,2)$ every term in Table 4 has lower order than $\sigma_n^4 \sim 4p^4n^{-2}(1-\alpha/2)^4$. When $\alpha \in [2,3)$ then $\widetilde{\beta}_6^2\widetilde{\beta}_4 np^3$ is of highest order but $\widetilde{\beta}_6^2\widetilde{\beta}_4 np^3/\sigma_n^4(\alpha) \to 0$ if $p = \omega(n^\delta)$ for some $\delta > \alpha/2 - 1$. This concludes the Third and last case and hence we have completed the proof of (4.5).

Finally, we start proving (4.6) and we obtain using Lemma 4.3 and Lemma 4.5 that

$$\mathbb{E}\Bigg[\bigg(\sum_{j=1}^{p} M_{j,2}^2\bigg)^2\Bigg] = o(\sigma_n^4)$$
$$+ 16 \sum_{\substack{j_1,j_2=1\\ j_1\neq j_2}}^{p} \sum_{i_1,i_2=1}^{j_1-1} \sum_{i_3,i_4=1}^{j_2-1} \sum_{\substack{t_1,\dots,t_8=1\\ t_1<t_2,\dots,t_7<t_8}}^{n} \mathbb{E}[Y_{i_1t_1}Y_{i_1t_2}Y_{i_2t_3}Y_{i_2t_4}\cdots Y_{j_2t_5}Y_{j_2t_6}Y_{j_2t_7}Y_{j_2t_8}]\,. \tag{4.10}$$

We have nonzero expectation in (4.10) if either all i's are equal (which implies that the j's are distinct from the i's) or if we form two pairs among the i's (but here some i's may coincide with j's). Note that when the sets of i's and j's are disjoint, then we must have that $t_1 = t_3$, $t_2 = t_4$, $t_5 = t_7$ and $t_6 = t_8$ to get nonzero. Starting with $i_1 = i_2 = i_3 = i_4$ the terms that occur in (4.10) can be summarised in Table 5.

$\mathbb{E}[Y_{i_1t_1}^2Y_{i_1t_2}^2Y_{i_2t_3}^2Y_{i_2t_4}^2]\mathbb{E}[Y_{j_1t_1}^2Y_{j_1t_2}^2]\mathbb{E}[Y_{j_2t_5}^2Y_{j_2t_6}^2]$	Order
Term	$\alpha \in (3,4)$
$\beta_{4,4}\beta_{2,2}^2n^2$	$n^{-\alpha-2}$
$\beta_{4,2,2}\beta_{2,2}^2n^3$	$n^{-\alpha/2-3}$
$\beta_{2,2,2,2}\beta_{2,2}^2n^4$	n^{-4}

TABLE 5. Terms of $\beta_{2k_1,\dots,2k_r}$ that occur in (4.10) when $i_1 = i_2 = i_3 = i_4$ and their order with respect to n.

Since the order in p is p^3 for (4.10) we easily see that the terms in Table 5 multiplied by $p^3\sigma_n^{-4}$ go to zero in view of Lemmas 2.1 and 2.4 for $\alpha \in (3,4)$. Investigating the cases where two pairs are formed, we have three cases with respect to the i's. If either $i_1 = i_3$ or $i_1 = i_4$ then we would need to impose the extra conditions ($t_1 = t_5$ and $t_2 = t_6$) or ($t_1 = t_5$ and $t_2 = t_6$) respectively. This leads to a lower order in n than in the pairing $i_1 = i_2$ and $i_3 = i_4$ and in this case we only get the term $\beta_{2,2}^4 n^4$. The order in p is the same as (4.9) and we get for $\alpha \in (3,4)$ that

$$16\frac{\frac{1}{4}\beta_{2,2}^4 n^4 p^4}{\sigma_n^4} \sim \frac{4n^{-4}p^4}{4n^{-4}p^4} = 1, \qquad n \to \infty.$$

It is left to consider the overlap between some i's and j's in the case $i_1 = i_2$ and $i_3 = i_4$. If $i_1 = i_2 = j_2$ then we have that the expectation in (4.10) is

$$\begin{aligned}&\mathbb{E}[Y_{i_1t_1}Y_{i_1t_2}Y_{i_1t_3}Y_{i_1t_4}Y_{i_1t_5}Y_{i_1t_6}Y_{i_1t_7}Y_{i_1t_8}]\mathbb{E}[Y_{i_3t_5}Y_{i_3t_6}Y_{i_3t_7}Y_{i_3t_8}]\mathbb{E}[Y_{j_1t_1}Y_{j_1t_2}Y_{j_1t_3}Y_{j_1t_4}] =\\&\mathbb{E}[Y_{i_1t_1}^2Y_{i_1t_2}^2Y_{i_1t_5}^2Y_{i_1t_6}^2]\mathbb{E}[Y_{i_3t_5}^2Y_{i_3t_6}^2]\mathbb{E}[Y_{j_1t_1}^2Y_{j_1t_2}^2] = \mathbb{E}[Y_{i_1t_1}^2Y_{i_1t_2}^2Y_{i_1t_5}^2Y_{i_1t_6}^2]\beta_{2,2}^2.\end{aligned}$$

The second equality is easiest to see by looking at the last two factors and consider the possible pairing of the t's that yield nonzero results. Then we get the pairing in the first factor for free as a consequence. These are the same terms as in Table 5 so we get that this case goes to zero as well. Lastly, we look at $i_1 = j_2$ but $i_1 \neq i_2$. This case yields nonzero if $i_2 = i_3$ which then yields in (4.10)

$$\mathbb{E}[Y_{i_1t_1}Y_{i_1t_2}Y_{i_1t_5}Y_{i_1t_6}Y_{i_1t_7}Y_{i_1t_8}]\mathbb{E}[Y_{i_2t_3}Y_{i_2t_4}Y_{i_2t_5}Y_{i_2t_6}Y_{i_2t_7}Y_{i_2t_8}]\mathbb{E}[Y_{j_1t_1}Y_{j_1t_2}Y_{j_1t_3}Y_{j_1t_4}].$$

Note that we must have in the last factor above $t_1 = t_3$ and $t_2 = t_4$. For the first two factors, they can only yield the possible term $\beta_{2,2,2}$ to be nonzero each. But for this to happen we need to set t_1 and t_2 to be equal to some other t's than only t_3 and t_4 respectively. This implies that an upper bound for the order would be $\beta_{2,2,2}^2 n^3 p^4 \sim n^{-5}p^4$ by Lemma 2.1. Now Lemma 2.4 yields that $n^{-5}p^4\sigma_n^{-4} \to 0$. By symmetry of indices we get the same results for $i_3 = j_1$ and we have thus shown (4.6) for $\alpha \in (3,4)$.

At this point only the case $\alpha = 3$ is left. A careful inspection of the above arguments in the case $\alpha \in (0,3) \cup (3,4)$ shows that (4.5) respectively (4.6) still hold for $\alpha = 3$ if σ_n^2 is replaced by $\mathbb{E}[T_1^2]$ and $\mathbb{E}[T_2^2]$, respectively. That is, if $\alpha = 3$ and $p = \omega(n^\delta)$ for some $\delta > \alpha/2 - 1$, it holds

$$\frac{1}{\mathbb{E}[T_1^2]}\sum_{j=1}^p M_{j,1}^2 \xrightarrow{\mathbb{P}} 1 \quad \text{and} \quad \frac{1}{\mathbb{E}[T_2^2]}\sum_{j=1}^p M_{j,2}^2 \xrightarrow{\mathbb{P}} 1. \tag{4.11}$$

Now we turn to the proof of (4.7). By virtue of (4.11), we get

$$\begin{aligned}\frac{1}{\mathbb{E}[T_1^2]+\mathbb{E}[T_2^2]}\sum_{j=1}^p \left(M_{j,1}^2 + M_{j,2}^2\right) &= \frac{\mathbb{E}[T_1^2]}{\mathbb{E}[T_1^2]+\mathbb{E}[T_2^2]}\frac{1}{\mathbb{E}[T_1^2]}\sum_{j=1}^p M_{j,1}^2 + \frac{\mathbb{E}[T_2^2]}{\mathbb{E}[T_1^2]+\mathbb{E}[T_2^2]}\frac{1}{\mathbb{E}[T_2^2]}\sum_{j=1}^p M_{j,2}^2\\&= \frac{\mathbb{E}[T_1^2]}{\mathbb{E}[T_1^2]+\mathbb{E}[T_2^2]}(1+o_{\mathbb{P}}(1)) + \frac{\mathbb{E}[T_2^2]}{\mathbb{E}[T_1^2]+\mathbb{E}[T_2^2]}(1+o_{\mathbb{P}}(1))\\&= 1 + o_{\mathbb{P}}(1)\,,\end{aligned}$$

where $o_{\mathbb{P}}(1)$ is a generic notation for a term that tends to zero in probability. In order to establish (4.7), it remains to show that

$$\frac{1}{\sigma_n^2}\sum_{j=1}^p M_{j,1}M_{j,2} \xrightarrow{\mathbb{P}} 0\,. \tag{4.12}$$

To this end, an application of Markov's inequality yields for $\varepsilon > 0$,

$$\mathbb{P}\Big(\sigma_n^{-2}\sum_{j=1}^{p} M_{j,1}M_{j,2} > \varepsilon\Big) \lesssim \sigma_n^{-4}\sum_{j,k=1}^{p}\mathbb{E}[M_{k,1}M_{k,2}M_{j,1}M_{j,2}].$$

By Lemma 4.3 and letting $k < j$ we have that $\mathbb{E}[M_{k,1}M_{k,2}M_{j,1}M_{j,2}]$ is equal to

$$64\sum_{i_1,i_2=1}^{k-1}\sum_{i_3,i_4=1}^{j-1}\sum_{t_1,t_4=1}^{n}\sum_{1\le t_2<t_3\le n}\sum_{1\le t_5<t_6\le n}\mathbb{E}[\overline{Y}_{i_1t_1}\overline{Y}_{kt_1}Y_{i_2t_2}Y_{i_2t_3}Y_{kt_2}Y_{kt_3}\overline{Y}_{i_3t_4}\overline{Y}_{jt_4}Y_{i_4t_5}Y_{i_4t_6}Y_{jt_5}Y_{jt_6}].$$

Since $k < j$ the inner expectation can be broken into two factors whereof one is $\mathbb{E}[\overline{Y}_{jt_4}Y_{jt_5}Y_{jt_6}]$. The aforementioned factor is nonzero only if $t_5 = t_6$ which cannot happen. We conclude that

$$\sigma_n^{-4}\sum_{j,k=1}^{p}\mathbb{E}[M_{k,1}M_{k,2}M_{j,1}M_{j,2}] = \sigma_n^{-4}\sum_{j=1}^{p}\mathbb{E}[M_{j,1}^2M_{j,2}^2] \le \sigma_n^{-4}\sum_{j=1}^{p}\Big(\mathbb{E}[M_{j,1}^4]+\mathbb{E}[M_{j,2}^4]\Big) \to 0,$$

as $n\to\infty$, by Lemma 4.5 if $p=\omega(n^\delta)$ for some $\delta > \delta^*(3) = 1$. This establishes (4.12) and completes the proof of the lemma.

□

5. Proofs for Section 2

In our proofs we will repeatedly use the notation $\overline{Y}_{it} := \big(Y_{it}^2 - \frac{1}{n}\big)$ *and*

$$C_{2k,\alpha} := \frac{\alpha\Gamma(\alpha/2)\Gamma(k-\alpha/2)}{2\Gamma(k)},$$

$$\widetilde{C}_{2k_1,\dots,2k_r,\alpha} := \frac{(\alpha/2)^{r-N_1}\Gamma(N_1(1-\alpha/2)+r\alpha/2)\prod_{i:k_i\ge 2}\Gamma(k_i-\alpha/2)}{\Gamma(k_1+\cdots+k_r)}.$$

5.1. **Proof of Lemma 2.2.**

Proof of Lemma 2.2. Let T_1 and T_2 be defined by (2.9). We get

$$\begin{aligned}\mathbb{E}[T_1^2] &= \mathbb{E}\left[\left(2\sum_{i_1<i_2}^{p}\sum_{t=1}^{n}\overline{Y}_{i_1,t}\overline{Y}_{i_2,t}\right)^2\right]\\ &= 4\sum_{i_1<i_2}^{p}\sum_{t=1}^{n}\mathbb{E}[\overline{Y}_{i_1,t}^2]\mathbb{E}[\overline{Y}_{i_2,t}^2] + 4\sum_{i_1<i_2}^{p}\sum_{t_1<t_2}^{n}\mathbb{E}[\overline{Y}_{i_1,t_1}\overline{Y}_{i_1,t_2}]\mathbb{E}[\overline{Y}_{i_2,t_1}\overline{Y}_{i_2,t_2}]\\ &= 2p(p-1)n\left(\beta_4-\frac{1}{n^2}\right)^2 + p(p-1)n(n-1)\left(\beta_{2,2}-\frac{1}{n^2}\right)^2\\ &= p(p-1)n\frac{2n-1}{n-1}\left(\beta_4-\frac{1}{n^2}\right)^2.\end{aligned}$$

For the last equality we used Lemma 4.1 to express $\beta_{2,2}$ in terms of β_4. Similarly we get for T_2,

$$\mathbb{E}[T_2^2] = \mathbb{E}\left[\left(\sum_{\substack{i_1,i_2=1\\ i_1\ne i_2}}^{p}\sum_{\substack{t_1,t_2=1\\ t_1\ne t_2}}^{n} Y_{i_1t_1}Y_{i_1t_2}Y_{i_2t_1}Y_{i_2t_2}\right)^2\right] = 16\,\mathbb{E}\left[\sum_{i_1<i_2}^{p}\sum_{t_1<t_2}^{n} Y_{i_1t_1}^2Y_{i_1t_2}^2Y_{i_2t_1}^2Y_{i_2t_2}^2\right] \quad (5.1)$$

$$= 4p(p-1)n(n-1)\beta_{2,2}^2 = 4p(p-1)\frac{n}{n-1}\left(\frac{1}{n}-\beta_4\right)^2.$$

□

5.2. **Proof of Lemma 2.4.**

Proof of Lemma 2.4. From (2.11) we obtain, as $n \to \infty$,

$$\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2)) \sim 2np^2\left(\left(\beta_4 - \frac{1}{n^2}\right)^2 + \frac{2}{n}\left(\beta_4 - \frac{1}{n}\right)^2\right).$$

We start with the case $\mathbb{E}[X^4] < \infty$, where Lemma 2.1 yields $\beta_4 \sim n^{-2}\mathbb{E}[X^4]$. It is easy to see that

$$\left(\beta_4 - \frac{1}{n^2}\right)^2 = O(n^{-4}), \quad \frac{2}{n}\left(\beta_4 - \frac{1}{n}\right)^2 \sim 2n^{-3},$$

so the asymptotic behavior of $\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))$ is the same as of σ_n^2 and

$$\sigma_n^2 \sim 2np^2 \cdot 2n^{-3} \sim 4\frac{p^2}{n^2}.$$

For the case $\alpha \in [2,4)$ with $\mathbb{E}[X_{11}^2] = 1$, Lemma 2.1 yields

$$\beta_4 \sim n^{-\alpha/2}L(n^{1/2})\frac{\alpha\Gamma(\alpha/2)\Gamma(2-\alpha/2)}{2\Gamma(2)} = n^{-\alpha/2}L(n^{1/2})C_{4,\alpha}.$$

In view of the Potter bounds for the slowly varying function L, we have $L(n) = O(n^{\varepsilon})$ for any $\varepsilon > 0$ so that $\beta_4 = O(n^{-\alpha/2+\varepsilon})$. If $\mathbb{E}[X_{11}^2] = \infty$ then we get another slowly varying function by Lemma 2.1, which we bound analogously. Hence, we do not need to distinguish between $\mathbb{E}[X_{11}^2] = \infty$ and $\mathbb{E}[X_{11}^2] = 1$ in the case $\alpha \in [2,4)$. Use of Lemma 2.1 now yields

$$\left(\beta_4 - \frac{1}{n^2}\right)^2 \sim \beta_4^2 = O(n^{-\alpha+2\varepsilon}),$$
$$\frac{2}{n}\left(\beta_4 - \frac{1}{n}\right)^2 = \frac{2}{n}\left(O(n^{-\alpha/2+\varepsilon}) - \frac{1}{n}\right)^2 \sim 2n^{-3}.$$

For $\alpha \in (3,4)$, the $2n^{-3}$ term dominates and therefore, $\sigma_n^2 \sim 2np^2 \cdot 2n^{-3} = 4p^2n^{-2}$. While for $\alpha \in (2,3)$ the β_4^2 term dominates and therefore,

$$\sigma_n^2 \sim 2np^2 \cdot \beta_4^2 \sim 2np^2 \cdot (n^{-\alpha/2}L(n^{1/2})C_{4,\alpha})^2 = 2p^2n^{1-\alpha}L^2(n^{1/2})C_{4,\alpha}^2.$$

In the case $\alpha = 3$ both terms in (2.11) might dominate (depending on L) and applying Lemma 2.1 one gets $\sigma_n^2 \sim 2p^2n^{-2}(C_{\alpha,4}^2L^2(n^{1/2}) + 2)$. Finally, when $\alpha \in (0,2)$, Lemma 2.1 yields

$$\beta_4 \sim n^{-1}\frac{\Gamma(2-\alpha/2)}{\Gamma(1-\alpha/2)\Gamma(2)} = n^{-1}\left(1 - \frac{\alpha}{2}\right)$$

from which we conclude

$$\left(\beta_4 - \frac{1}{n^2}\right)^2 \sim \left(1 - \frac{\alpha}{2}\right)^2 n^{-2},$$
$$\frac{2}{n}\left(\beta_4 - \frac{1}{n}\right)^2 \sim \frac{2}{n}\left(\frac{1-\alpha/2}{n} - \frac{1}{n}\right)^2 = \frac{\alpha^2}{2}n^{-3},$$

which implies $\sigma_n^2 \sim 2p^2n^{-1}(1-\alpha/2)^2$. For completeness, we note that the fact that $\sigma_n^2 \sim \mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))$ is easily deduced from the above considerations in all cases. □

5.3. **Proof of Lemma 2.5.**

Proof of Lemma 2.5. We need to find the types of $\beta's$ that occur in $\mathbb{E}[T_i^4]$ for $i = 1, 2$. Let us start with $\mathbb{E}[T_1^4]$ which is given by

$$\mathbb{E}[T_1^4] = 16 \sum_{\substack{i_1,i_2=1\\ i_1<i_2}}^{p} \cdots \sum_{\substack{i_7,i_8=1\\ i_3<i_4}}^{p} \sum_{t_1,\dots,t_4=1}^{n} \mathbb{E}\left[\overline{Y}_{i_1,t_1}\overline{Y}_{i_2,t_1}\overline{Y}_{i_3,t_2}\overline{Y}_{i_4,t_2}\overline{Y}_{i_5,t_3}\overline{Y}_{i_6,t_3}\overline{Y}_{i_7,t_4}\overline{Y}_{i_8,t_4}\right].$$

We proceed to the tedious task of finding all the combinations of indices that lead to non-zero contributions in the form of products of certain $\widetilde{\beta}_{2k_1,\dots,2k_r}$, whose order can be determined using Lemma 4.2 yielding that any $\widetilde{\beta}_{2k_1,\dots,2k_r}$ with no k_i equal to 1, will behave as $\beta_{2k_1,\dots,2k_r}$ asymptotically. It is important to note that any term including $\widetilde{\beta}_2$ is zero. In a methodical manner one can check from big indices (of $\widetilde{\beta}$) to small, terms with 2 up to 4 factors of β, which combinations of β's are possible. For example, the term with largest possible index is β_8^2 and this happens only if $i_1 = i_3 = i_5 = i_7$ and $i_2 = i_4 = i_6 = i_8$ with all the t indices equal. Accounting for multiplicities, the contribution of such terms to $\mathbb{E}[T_1^4]$ will be equal to

$$\frac{p(p-1)}{2} n\beta_8^2 \sim \frac{1}{2} p^2 n \beta_8^2.$$

Since we are only interested in the asymptotic behavior of $\mathbb{E}[T_1^4]$, the exact multiplicities are not required. Hence, it is not necessary to know the exact number of occurring β_8^2 and similarly for all other products of $\widetilde{\beta}_{2k_1,\dots,2k_r}$. Then we look at how many of the $t's$ and i's must not be equal respectively and hence get the order of the term by setting the corresponding amount in the exponent of n and p. Finally, we focus on the terms with highest order.

For our analysis we need the exact asymptotic behavior of terms $\widetilde{\beta}_{2k_1,\dots,2k_r}$ where $k_1 = \cdots = k_r = 1$. Using Lemma 4.1 we get $\beta_{2,2} = \frac{1-n\beta_4}{n(n-1)}$, which yields

$$\widetilde{\beta}_{2,2} = \beta_{2,2} - n^{-2} = \frac{1}{n^2(n-1)} - \frac{\beta_4}{n-1}. \tag{5.2}$$

Since $\beta_4 \sim n^{-\alpha/2} L(n^{1/2}) C_{4,\alpha}$ for $\alpha \in (2, 4)$ by Lemma 2.1, it follows that

$$\widetilde{\beta}_{2,2} \sim -n^{-\alpha/2-1} L(n^{1/2}) C_{4,\alpha}\,. \tag{5.3}$$

Similar but more tedious calculations yield

$$\widetilde{\beta}_{2,2,2} \sim 2n^{-\alpha/2-2} L(n^{1/2})(C_{4,\alpha} + C_{6,\alpha}), \tag{5.4}$$

$$\widetilde{\beta}_{2,2,2,2} = O(n^{-\alpha/2-2}). \tag{5.5}$$

The fourth moment of T_2

$$\mathbb{E}[T_2^4] = 4^4\, \mathbb{E}\left[\left(\sum_{i_1<i_2}^{p} \sum_{t_1<t_2}^{n} Y_{i_1t_1}Y_{i_1t_2}Y_{i_2t_1}Y_{i_2t_2}\right)^4\right]$$

is analyzed in a similar way. Table 6 below includes all the products of $\beta_{2k_1,\dots,2k_r}$ that occur in either $\mathbb{E}[T_1^4]$ or $\mathbb{E}[T_2^4]$ and gives their order respective orders (utilizing Lemma 2.1) up to some slowly varying function for $\alpha \in (0, 4]$. Strictly speaking, in the case $\alpha = 2$ we need to distinguish between finite or infinite second moment of X_{11}; luckily both cases yield the same asymptotic results apart from a slowly different function.

$\mathbb{E}[T_1^4]$			$\mathbb{E}[T_2^4]$		
Term	**Order** $\alpha \in (0,2)$	**Order** $\alpha \in [2,4)$	**Term**	**Order** $\alpha \in (0,2)$	**Order** $\alpha \in [2,4)$
$\beta_8^2 np^2$	$n^{-1}p^2$	$n^{-\alpha+1}p^2$	$\beta_4\beta_{2,2}^3 n^3p^3$	$n^{-4}p^3$	$n^{-\alpha/2-3}p^3$
$\beta_{4,4}^2 n^2p^2$	$n^{-2}p^2$	$n^{-2\alpha+2}p^2$	$\beta_{4,4}^2 n^2p^2$	$n^{-2}p^2$	$n^{-2\alpha+2}p^2$
$\widetilde{\beta}_{6,2}^2 n^2p^2$	$n^{-2}p^2$	$n^{-\alpha}p^2$	$\beta_{2,2,2,2}^2 n^4p^2$	$n^{-4}p^2$	$n^{-4}p^2$
$\widetilde{\beta}_{2,2,2,2}^2 n^4p^2$	$n^{-4}p^2$	$O(n^{-\alpha})p^2$	$\beta_{2,2}^4 n^4p^4$	$n^{-4}p^4$	$n^{-4}p^4$
$\widetilde{\beta}_{6,2}\beta_4\widetilde{\beta}_{2,2}n^2p^3$	$n^{-3}p^3$	$-n^{-3\alpha/2}p^3$	$\beta_{2,2,2,2}\beta_{2,2}^2 n^4p^3$	$n^{-4}p^3$	$n^{-4}p^3$
$\widetilde{\beta}_{2,2,2}^2\widetilde{\beta}_{2,2}n^4p^3$	$n^{-4}p^3$	$-n^{-3\alpha/2-1}p^3$			
$\widetilde{\beta}_{2,2,2}^2\beta_4 n^3p^3$	$n^{-4}p^3$	$n^{-3\alpha/2-1}p^3$			
$\beta_6^2\beta_4 np^3$	$n^{-2}p^3$	$n^{-3\alpha/2+1}p^3$			
$\beta_{4,4}\widetilde{\beta}_{2,2}^2 n^2p^3$	$n^{-4}p^3$	$n^{-2\alpha}p^3$			
$\beta_4^4 n^2p^4$	$n^{-2}p^4$	$n^{-2\alpha+2}p^4$			
$\beta_4^2\widetilde{\beta}_{2,2}^2 n^3p^4$	$n^{-3}p^4$	$n^{-2\alpha+1}p^4$			
$\widetilde{\beta}_{2,2}^4 n^4p^4$	$n^{-4}p^4$	$n^{-2\alpha}p^4$			

TABLE 6. The order of the terms in $\mathbb{E}[T_1^4]$ and $\mathbb{E}[T_2^4]$ for $\alpha \in (0,4)$. The $O(n^{-\alpha})$ is due to the fact that we only have an upper bound for the order by (5.5).

For $\mathbb{E}[T_1^4]$ the dominating terms for $\alpha \in [2,4)$ are $\beta_4^4 n^2p^4$, $\beta_8^2 np^2$ and $\beta_6^2\beta_4 np^3$. Note that the growth of p relative to n is an important detail that needs to be taken into consideration. We recall that our results are valid for general growth rates of p. One can check that the same terms will also dominate when $\alpha \in (0,2)$.

Finding the exact multiplicities of β_4^4, β_8^2 and $\beta_6^2\beta_4$ in $\mathbb{E}[T_1^4]$ is not too difficult and will yield

$$\frac{1}{16}\,\mathbb{E}[T_1^4] \sim 3\Big(\frac{p(p-1)}{2}n\Big)^2\beta_4^4 + \frac{p(p-1)}{2}pn\beta_8^2 + 8\,\frac{p(p-1)}{2}pn\beta_6^2\beta_4$$
$$\sim \frac{3}{4}n^2p^4\beta_4^4 + \frac{1}{2}\beta_8^2 np^2 + 4\beta_6^2\beta_4 np^3.$$

Now we turn to $\mathbb{E}[T_2^4]$. We find that the highest order term for $\alpha \in [3,4)$ is $\beta_{2,2}^4 n^4p^4$ by looking in Table 6. The multiplicity of $\beta_{2,2}^4 n^4p^4$ is 3 which then for $\alpha \in [3,4)$ yields

$$\frac{1}{256}\,\mathbb{E}[T_2^4] \sim \frac{3}{16}\,\beta_{2,2}^4 n^4p^4 \sim \frac{3}{16}\,n^{-4}p^4.$$

□

5.4. **Proof of Theorem 2.6.**

Proof of Theorem 2.6. First, note that by the binomial theorem we have

$$\mathbb{E}\big[(\mathrm{tr}(\mathbf{R}^2) - \mathbb{E}[\mathrm{tr}(\mathbf{R}^2)])^4\big] = \mathbb{E}\big[(T_1+T_2)^4\big] = \mathbb{E}[T_1^4] + 4\,\mathbb{E}[T_1^3T_2] + 6\,\mathbb{E}[T_1^2T_2^2] + 4\,\mathbb{E}[T_1T_2^3] + \mathbb{E}[T_2^4].$$

Recalling the definitions of T_1 and T_2 in (2.9), we observe that T_2 contains only odd powers of Y_{it}'s and T_1 only even powers and thus, we immediately see that $\mathbb{E}[T_1^3T_2] = 0$. By Lemma 2.8, we only need to study $\mathbb{E}[T_1^4]$ for $\alpha < 3$ and $\mathbb{E}[T_2^4]$ for $\alpha > 3$. Using Cauchy-Schwarz inequality we have

$$\frac{\mathbb{E}[T_1^2 T_2^2]}{\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))^2} \le \left(\frac{\mathbb{E}[T_1^4]}{\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))^2}\right)^{1/2}\left(\frac{\mathbb{E}[T_2^4]}{\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))^2}\right)^{1/2}, \tag{5.6}$$

which goes to zero as $n \to \infty$ by Proposition 2.7 and Lemma 2.8 assuming $p = \omega(n^\delta)$ for some $\delta > \delta^*(\alpha)$ with $\alpha \in (0,3) \cup (3,4)$. Similarly, using Hölder's inequality, we get

$$\frac{\left|\mathbb{E}[T_1 T_2^3]\right|}{\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))^2} \le \frac{\mathbb{E}[T_1^4]^{1/4}\mathbb{E}[T_2^4]^{3/4}}{\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))^2} = \left(\frac{\mathbb{E}[T_1^4]}{\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))^2}\right)^{1/4}\left(\frac{\mathbb{E}[T_2^4]}{\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))^2}\right)^{3/4},$$

which goes to 0 as $n \to \infty$ by Lemma 2.8 and Proposition 2.7 if $p = \omega(n^\delta)$ for some $\delta > \delta^*(\alpha)$ when $\alpha \in (0,3) \cup (3,4)$. Another application of Proposition 2.7 for $\mathbb{E}[T_1^4]$ and $\mathbb{E}[T_2^4]$ directly yields the desired result for $\alpha \in (0,3) \cup (3,4)$.

Finally, we note that, for example, (5.6) tends to infinity if $p = o(n^\delta)$ for some $\delta < \delta^*(\alpha)$ by Proposition 2.7 and Lemma 2.8, which highlights the importance of our condition $\delta > \delta^*(\alpha)$. □

5.5. **Proof of Proposition 2.7.**

Proof of Proposition 2.7. From Lemmas 2.5 and 2.2 we have for $\alpha \in (0,4)$

$$\mathbb{E}[T_1^4] \sim 12\beta_4^4 n^2 p^4 + 8\beta_8^2 np^2 + 64\beta_6^2\beta_4 np^3, \quad \mathbb{E}[T_1^2] \sim 4\beta_4^4 p^4 n^2$$

which yields

$$\frac{\mathbb{E}[T_1^4]}{\mathbb{E}[T_1^2]^2} \sim 3 + \frac{2\beta_8^2}{\beta_4^4 p^2 n} + \frac{16\beta_6^2}{\beta_4^3 pn}.$$

By Lemma 2.1, we see that $\beta_4, \beta_6, \beta_8$ are of the same order up to some slowly varying function. Thus, if $\beta_4^2 p^2 n\ell(n) \to \infty$ and $\beta_4 pn\ell(n) \to \infty$ as $n \to \infty$ for any slowly varying function ℓ, then $\mathbb{E}[T_1^4]/\mathbb{E}[T_1^2]^2 \sim 3$. This clearly holds if $\alpha \in (0,3]$ and $p = \omega(n^\delta)$ for some $\delta > \delta^*(\alpha)$ since by Lemma 2.1 β_4 behaves like $n^{-\max(1,\alpha/2)}$ (up to a slowly varying function). If instead $p = o(n^\delta)$ for some $\delta < \delta^*(\alpha)$, then we have

$$\lim_{n\to\infty} \frac{\mathbb{E}[T_1^4]}{\mathbb{E}[T_1^2]^2} = \infty.$$

Finally, if $\alpha \in [3,4)$ then by (5.1) and Lemma 2.1 we have

$$\mathbb{E}[T_2^2]^2 = \left(4p(p-1)\frac{n}{n-1}\left(\beta_4 - \frac{1}{n}\right)^2\right)^2 \sim 16p^4 n^{-4}$$

and we find by combining with Lemma 2.5 that $\mathbb{E}[T_2^4]/\mathbb{E}[T_2^2]^2 \to 3$ for $\alpha \in [3,4)$. □

5.6. **Proof of Lemma 2.8.**

Proof of Lemma 2.8. Combining Lemmas 2.4 and 2.5 we have for $\alpha \in (3,4)$

$$\frac{\mathbb{E}[T_1^4]}{\mathrm{Var}(\mathrm{tr}(\mathbf{R}^2))^2} \sim \frac{12\,\beta_4^4 n^2 p^4 + 8\beta_8^2 np^2 + 64\beta_6^2\beta_4 np^3}{16n^{-4}p^4}, \qquad n \to \infty.$$

By Lemma 2.1, we get

$$\frac{12\,\beta_4^4 n^2 p^4}{16n^{-4}p^4} \sim \frac{3}{4} n^{-2(\alpha-3)}\, C_{4,\alpha}^4 L(n^{1/2})^4 \to 0, \qquad n \to \infty.$$

Moreover, assuming $p = \omega(n^\delta)$ for some $\delta > (5-\alpha)/2$, we have that

$$\frac{8\beta_8^2 np^2}{16n^{-4}p^4} \sim \frac{1}{2} n^{-\alpha+5} p^{-2} L(n^{1/2})^2 C_{8,\alpha}^2 \to 0.$$

Similarly we have

$$\frac{64\beta_6^2\beta_4 np^3}{16n^{-4}p^4} \sim \frac{1}{4} n^{-3\alpha/2+5} p^{-1} L(n^{1/2})^3 C_{6,\alpha}^2 C_{4,\alpha} \to 0, \qquad n \to \infty,$$

as long as $\delta > 5 - 3\alpha/2$. We conclude that

$$\frac{\mathbb{E}[T_1^4]}{\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))^2} \to 0, \qquad n \to \infty$$

for $\alpha \in (3,4)$ assuming $p = \omega(n^\delta)$ for some $\delta > (5-\alpha)/2$. For $\alpha \in [2,3)$ and $\mathbb{E}[X_{11}^2] = 1$, the possible terms in $\mathbb{E}[T_2^4]$ that dominate are $\beta_4\beta_{2,2}^3 n^3p^3$, $\beta_{4,4}^2 n^2p^2$ and $\beta_{2,2}^4 n^4p^4$ (see Table 6). Using the variance in Lemma 2.4 combined with Lemma 2.1 we get

$$\frac{\mathbb{E}[T_2^4]}{\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))^2} \lesssim \frac{n^{-\alpha/2-3}p^3 C_{4,\alpha}\widetilde{C}_{2,2,\alpha}^3 L(n^{1/2}) + n^{-2\alpha+2}p^2\widetilde{C}_{4,4,\alpha}^2 L^4(n^{1/2}) + n^{-4}p^4\widetilde{C}_{2,2,\alpha}^4}{4p^4 n^{2-2\alpha} L^4(n^{1/2}) C_{4,\alpha}^4} \to 0.$$

If for $\alpha = 2$ we instead have $\mathbb{E}[X_{11}^2] = \infty$, then the above still holds but with a different slowly varying function in view of Lemma 2.1. Now for $\alpha \in (0,2)$ using Lemmas 2.4, 2.5 and 2.1 we get

$$\frac{\mathbb{E}[T_2^4]}{\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))^2} \sim \frac{48n^{-4}p^4}{4p^4n^{-2}(1-\alpha/2)^4}\left(\frac{\alpha}{2}\right)^4 \to 0, \qquad n \to \infty,$$

which means that $\frac{\mathbb{E}[T_2^4]}{\operatorname{Var}(\operatorname{tr}(\mathbf{R}^2))^2} \to 0$, as $n \to \infty$, for all $\alpha \in (0,3)$. □

Appendix A. Sums of regularly varying random variables

Lemma A.1 (Jessen and Mikosch [14]). *Assume $|X_1|$ is regularly varying with index $\alpha \geq 0$ and distribution function $F = 1 - \overline{F}$. Assume $X_1, ..., X_n$ are random variables satisfying*

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_i > x)}{\overline{F}(x)} = c_i^+ \quad \text{and} \quad \lim_{x\to\infty} \frac{\mathbb{P}(X_i \leq -x)}{\overline{F}(x)} = c_i^-, \quad i = 1, \dots, n, \tag{A.7}$$

for some non-negative numbers $c_i^\pm$ and

$$\begin{aligned}\lim_{x\to\infty} \frac{\mathbb{P}(X_i > x, X_j > x)}{\overline{F}(x)} &= \lim_{x\to\infty} \frac{\mathbb{P}(X_i \leq -x, X_j > x)}{\overline{F}(x)} \\ &= \lim_{x\to\infty} \frac{\mathbb{P}(X_i \leq -x, X_j \leq -x)}{\overline{F}(x)} = 0, \quad i \neq j.\end{aligned} \tag{A.8}$$

Then

$$\lim_{x\to\infty} \frac{\mathbb{P}(S_n > x)}{\overline{F}(x)} = c_1^+ + \cdots + c_n^+ \quad \text{and} \quad \lim_{x\to\infty} \frac{\mathbb{P}(S_n \leq x)}{\overline{F}(x)} = c_1^- + \cdots + c_n^-,$$

with $S_n = X_1 + \cdots + X_n$, $n \geq 1$.

Acknowledgments. This work is the outcome of a semester-long collaboration between Felix Seo and Johannes Heiny, and (in an extended version) will be submitted to a mathematical journal. I (Felix Seo) would especially want to thank Johannes Heiny for his incredible support and guidance throughout this period. Without his helpful feedback and patience the results presented in this work would not have been possible.

References

[1] Albrecher, H., and Teugels, J. L. Asymptotic analysis of a measure of variation. *Theory of Probability and Mathematical Statistics 74* (2007), 1–10.

[2] Bai, Z., and Silverstein, J. W. *Spectral Analysis of Large Dimensional Random Matrices*, second ed. Springer Series in Statistics. Springer, New York, 2010.

[3] Bai, Z. D., and Silverstein, J. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab. 32* (2004), 553–605.

[4] Bai, Z. D., and Yin, Y. Q. Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann. Probab. 21*, 3 (1993), 1275–1294.

[5] Bingham, N. H., Goldie, C. M., and Teugels, J. L. *Regular Variation*, vol. 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1987.

[6] Blum, J. R., Kiefer, J., and Rosenblatt, M. Distribution Free Tests of Independence Based on the Sample Distribution Function. *Ann. Math. Statist. 32*, 2 (1961), 485 – 498.

[7] Brockwell, P. J., and Davis, R. A. *Time series: theory and methods*, second ed. Springer Series in Statistics. Springer-Verlag, New York, 1991.

[8] Dörnemann, N., and Heiny, J. Limiting spectral distribution for large sample correlation matrices. *arXiv preprint arXiv:2208.14948* (2022).

[9] Fuchs, A., Joffe, A., and Teugels, J. Expectation of the ratio of the sum of squares to the square of the sum: exact and asymptotic results. *Teor. Veroyatnost. i Primenen. 46*, 2 (2001), 297–310.

[10] Hall, P., and Heyde, C. C. *Martingale limit theory and its application.* Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980.

[11] Heiny, J. Large sample correlation matrices: a comparison theorem and its applications. *Electron. J. Probab. 27* (2022), 1–20.

[12] Heiny, J., and Parolya, N. Log determinant of large correlation matrices under infinite fourth moment. *Ann. Inst. Henri Poincaré Probab. Stat. (forthcoming), arXiv:2112.15388* (2023).

[13] Hoeffding, W. A Non-Parametric Test of Independence. *The Annals of Mathematical Statistics 19*, 4 (1948), 546 – 557.

[14] Jessen, A. H., and Mikosch, T. Regularly varying functions. *Publications De L'institut Mathematique 80* (2006), 171–192.

[15] Kendall, M. G. A new measure of rank correlation. *Biometrika 30*, 1/2 (1938), 81–93.

[16] Ladoucette, S. A. Asymptotic behavior of the moments of the ratio of the random sum of squares to the square of the random sum. *Statist. Probab. Lett. 77*, 10 (2007), 1021–1033.

[17] Marčenko, V. A., and Pastur, L. A. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.) 72 (114)* (1967), 507–536.

[18] Pearson, K. Notes on the history of correlation. *Biometrika 13*, 1 (1920), 25–45.

[19] Priestley, M. B. *Spectral analysis and time series. Vols. 1 and 2.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1981. Univariate series, Probability

and Mathematical Statistics.

[20] Yao, J., Zheng, S., and Bai, Z. *Large sample covariance matrices and high-dimensional data analysis*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, New York, 2015.

[21] Yin, Y., Li, C., Tian, G.-L., and Zheng, S. Spectral properties of rescaled sample correlation matrix. *Statist. Sinica 32*, 4 (2022), 2007–2022.

[22] Yin, Y., Zheng, S., and Zou, T. Central limit theorem of linear spectral statistics of high-dimensional sample correlation matrices. *Bernoulli 29*, 2 (2023), 984–1006.

Department of Mathematics, Stockholm University, Albano hus 1, 10691 Stockholm, Sweden
Email address: `fese8867@student.su.se`