Submitted to Bernoulli

Asymptotic power of Rao's score test for independence in high dimensions

DENNIS LEUNG[1] and QIMAN SHAO[2]

[1] *Department of Data Sciences and Operations, University of Southern California*
E-mail: leungden@usc.edu

[2] *Department of Statistics, Chinese University of Hong Kong*
E-mail: qmshao@cuhk.edu.hk

Let $\mathbf{R}$ be the Pearson correlation matrix of m normal random variables. The Rao's score test for the independence hypothesis $H_0 : \mathbf{R} = \mathbf{I}_{\rm m}$, where $\mathbf{I}_{\rm m}$ is the identity matrix of dimension m, was first considered by Schott [19] in the high dimensional setting. In this paper, we study the exact power function of this test, under an asymptotic regime in which both m and the sample size n tend to infinity with the ratio m/n upper bounded by a constant. In particular, our result implies that the Rao's score test is minimax rate-optimal for detecting the dependency signal $\| \mathbf{R} - \mathbf{I}_{\rm m} \|_F$ of order $\sqrt{m/n}$, where $\| \cdot \|_F$ is the matrix Frobenius norm.

1. Introduction

Let $(X_1, \dots, X_m)'$ be an m-variate normal vector with population Pearson correlation matrix denoted by $\mathbf{R} = (\rho_{pq})_{1 \le p,q \le m}$. Suppose we observe n independent samples $X_{p1}, \dots, X_{pn}$ for each component X_p, $1 \le p \le m$. When the dimension m can be larger than the sample size n, Schott [19] was the first to consider the Rao's score statistic

$$T = \sum_{1 \le p < q \le m} \hat{\rho}_{pq}^2, \tag{1.1}$$

for testing the independence null hypothesis

$$H_0 : \mathbf{R} = \mathbf{I}_m, \tag{1.2}$$

where $\hat{\rho}_{pq}$, $1 \le p \ne q \le m$ is the sample correlation of the pair (X_p, X_q) computed from the data, and $\mathbf{I}_m$ is the m-by-m identity matrix. It was shown to be asymptotically normal under H_0 as both m and n go to infinity with the ratio m/n converging to a positive constant. The purpose of this paper is to complement the theoretical study of T by investigating its power under alternatives of the form

$$H_1 : \mathbf{R} \in \Theta(b),$$

where for any constant $b > 0$ and matrix Frobenius norm $\| \cdot \|_F$, we define the set of Pearson correlation matrices

$$\Theta(b) := \{\mathbf{R} : \|\mathbf{R} - \mathbf{I}_m\|_F \ge b\sqrt{m/n}, \quad \mathrm{diag}(\mathbf{R}) = \mathbf{I}_{\rm m}\}, \tag{1.3}$$

which comprises a composite alternative hypothesis delineated by a signal size $\|\mathbf{R}-\mathbf{I}_m\|_F$ of order no less than $\sqrt{m/n}$.

There are three major approaches to testing independence with growing dimension m in the literature, to the best of our knowledge. The first is the statistic T considered in this paper. Being a "sum" of squared pairwise sample correlation as in (1.1), it is good at detecting diffuse dependency among many pairs of variables. Such dependency is most naturally described by the signal $\|\mathbf{R}-\mathbf{I}_m\|_F$. In fact, the main result in this paper will show that T is minimax rate optimal for detecting such signal. The second approach considers the "max" statistic,

$$\max_{1\le p<q\le m} \hat{\rho}^2_{pq}.$$

Following many previous works [12, 18, 16, 21, 17], Cai and Jiang [3] showed that it admits an asymptotic Gumbel distribution under H_0 in the ultra high dimensional regime when m can be as large as e^{n^c} for some constant $0 < c < 1$, as $m, n \longrightarrow \infty$. Naturally, it is good at detecting a structured alternative whose population correlation matrix $\mathbf{R}$ has sparse non-zero off-diagonal entries with considerable magnitudes. Both the "sum" and "max" approaches base their test on forming intuitive statistics that measure the overall dependency among the m variables, with their respective non-parametric extensions; see Leung and Drton [14] and Han and Liu [7]. The third is likelihood ratio test (LRT), which is well-known to give implementable test only if the dimension m is smaller than n. Despite this limitation, Jiang and Qi [13] showed the LRT statistic to be asymptotically normal when $m, n \longrightarrow \infty$, as long as $m + 4$ is less than n.

We remark that the derivation of (1.1) as the Rao's score statistic involves taking derivatives of the log-normal likelihood with respect to the mean vector and the precision matrix. The interested reader is referred to Appendix A in Leung and Drton [14] for those calculations.

2. Notations and main results

For any positive integer k, $[k]$ is defined as the set $\{1, \dots, k\}$. $\mathcal{S}_k$ is the symmetric group of order k. Depending on the context, its elements will sometimes be treated as permutation functions on k elements, or simply permutations of the set $[k]$. C always denotes a positive constant that is universal, i.e, its value may change from place to place but does not depend on m and n. "$a \lesssim b$" means that $a \le Cb$ for some constant $C > 0$. $\mathbb{E}[\cdot]$, $\mathrm{Var}[\cdot]$ and $P[\cdot]$ are expectation, variance and probability operators respectively.

In this paper we shall *always* assume that, for all $1 \le p \le m$, $\mathrm{Var}[X_p] = 1$ and $\mathbb{E}[X_p] = 0$. Thus, for a duple $(p, q) \in [m] \times [m]$, $\mathbb{E}[X_p X_q] = \rho_{pq}$, and its corresponding squared sample correlation is defined as

$$\hat{\rho}^2_{pq} := \frac{S^2_{pq}}{S_{pp} S_{qq}} = f\left(S_{pp}, S_{qq}, S_{pq}\right), \tag{2.1}$$

where $f : \mathbb{R}^2_{>0} \times \mathbb{R} \longrightarrow \mathbb{R}$ is the function

$$f(u_1, u_2, u_3) := u_1^{-1} u_2^{-1} u_3^2, \tag{2.2}$$

and

$$S_{pq} := \frac{\sum_{i=1}^{n} X_{pi} X_{qi}}{n}. \tag{2.3}$$

We will also use

$$\bar{S}_{pq} := S_{pq} - \rho_{pq}$$

to denote the centered sample covariance. Imposing the assumption $\mathrm{Var}[X_p] = 1$ is always permitted, even if we use the more general form of Pearson correlations with all sample covariances S_{pq} defined alternatively as

$$\frac{\sum_{i=1}^{n} (X_{pi} - n^{-1} \sum_{j=1}^{n} X_{pj})(X_{qi} - n^{-1} \sum_{j=1}^{n} X_{qj})}{n-1} \tag{2.4}$$

in (2.1), since the distribution of $\hat{\rho}_{pq}$ is invariant to the scaling of variables. Under normality, the restrictions $\mathbb{E}[X_p] = 0$ and (2.3) can be still be assumed without forgoing any generality of our results to follow; see the classical result in Anderson [1, Theorem 3.3.2].

According to Chen and Shao [5, Theorem 2.2] who refined the asymptotic result of [19] under H_0, for a given $\alpha \in (0, 1)$, a test of asymptotic level α based on (1.1) is given as

$$\psi = I\left(T - \frac{m(m-1)}{2n} > \frac{m}{n} z_\alpha\right), \tag{2.5}$$

where $I(\cdot)$ is the indicator function , $z_\alpha := \bar{\Phi}^{-1}(\alpha)$, and Φ and $\bar{\Phi}(x) := 1 - \Phi(x)$ are respectively the cumulative distribution function and tail probability of a standard normal variate. Below, $\mathbb{E}_{\mathbf{R}}[\cdot]$ simply emphasizes that the expectation is taken with respect to a particular correlation matrix $\mathbf{R} \in \Theta(b)$.

Theorem 2.1 (Main result: asymptotic power)**.** *Suppose $m, n \longrightarrow \infty$ such that $\frac{m}{n} \le \kappa$ for some constant $\kappa < \infty$. For any significance level $\alpha \in (0, 1)$, the asymptotic power of ψ is given as*

$$\lim_{n \to \infty} \inf_{\Theta(b)} \mathbb{E}_{\mathbf{R}}[\psi] = \bar{\Phi}(z_\alpha - 2^{-1} b^2).$$

This theorem resembles Cai and Ma [4, Theorem 4], in which the different problem of testing $H_0 : \mathbf{\Sigma} = \mathbf{I}_{\mathrm{m}}$, where $\mathbf{\Sigma}$ is the *covariance* matrix of $(X_1, \ldots, X_m)'$, is studied. Despite this, Theorem 1 and Remark 1 in their paper indicate that a matching lower bound on the detectable signal size as measured by $\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F$ can be established for our problem (1.2), which we restate next for our readers' convenience. We add that Theorem 2.1 is slightly weaker than the parallel result of Cai and Ma [4] in that an upper bound on the ratio m/n is imposed, which we believe to be merely a proof artifact not necessary for the theorem to hold. Discussion on this will be deferred later.

Theorem 2.2 (Matching lower bound, Cai and Ma [4])**.** *Let $0 < \alpha < \beta < 1$. Suppose $m, n \longrightarrow \infty$ such that $\frac{m}{n} \le \kappa$ for some constant $\kappa < \infty$. Then there exists a constant $b = b(\kappa, \beta - \alpha) < 1$, such that*

$$\limsup_{n \longrightarrow \infty} \inf_{\Theta(b)} \mathbb{E}_{\mathbf{R}}[\phi] < \beta$$

for any test ϕ with significance level α for testing H_0.

The lower bound result says that no α-level test for H_0 can achieve a preset target power if the signal size $\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F$ falls below a certain threshold modulo the separation rate $\sqrt{m/n}$. Our main result in Theorem 2.1 hence suggests that our test ψ is "rate" optimal when the ratio m/n is bounded, since the asymptotic power $\lim_{n\longrightarrow\infty} \inf_{\Theta(b)} \mathbb{E}_{\mathbf{R}}[\psi]$ tends to one as $b \longrightarrow \infty$.

Although the result in Theorem 2.1 is neat, its proof, which occupies the rest of this paper, is quite involved. As it will become clear later, this is because our statistic T is constructed with Pearson correlations whose higher order moment properties involve a lot of computations to be understood; see Hotelling [10, Section 7] for classical work on this. At some point in this paper we will use `mathematica` to help us with certain symbolic calculations. We shall begin with a Taylor expansion of the expression for $\hat{\rho}^2_{pq}$ in terms of the function f in (2.1). We need the multi-index notations: For a vector $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_k)$ of k non-negative integers, $\boldsymbol{\lambda}! = \lambda_1! \dots \lambda_k!$ and $|\boldsymbol{\lambda}| = \lambda_1 + \cdots + \lambda_k$, and if $g = g(u_1, \dots, u_k)$ is a function in k arguments, $\partial^{\boldsymbol{\lambda}} g(\tilde{u}_1, \dots, \tilde{u}_k) = \frac{\partial^{|\boldsymbol{\lambda}|} g}{\partial u_1^{\lambda_1} \dots \partial u_k^{\lambda_k}}\Big|_{u_i = \tilde{u}_i}$ is its partial derivative with respect to $\boldsymbol{\lambda}$ evaluated at the point $(\tilde{u}_1, \dots \tilde{u}_k)$. Since $\rho^2_{pq} = f(1, 1, \rho_{pq}) = f(\rho_{pp}, \rho_{qq}, \rho_{pq})$, by Taylor's theorem, for each pair $1 \leq p \neq q \leq m$,

$$\hat{\rho}^2_{pq} - \rho^2_{pq} = \sum_{\substack{\boldsymbol{\lambda} \in \mathbb{N}^3_{\geq 0}: \\ 1 \leq |\boldsymbol{\lambda}| \leq 4}} \frac{\partial^{\boldsymbol{\lambda}} f(1, 1, \rho_{pq})}{\boldsymbol{\lambda}!} \bar{S}^{\lambda_1}_{pp} \bar{S}^{\lambda_2}_{qq} \bar{S}^{\lambda_3}_{pq} + III_{pq} \quad \text{a.s.}, \tag{2.6}$$

where

$$III_{pq} := \sum_{\substack{\boldsymbol{\lambda} \in \mathbb{N}^3_{\geq 0}: \\ |\boldsymbol{\lambda}| = 5}} \frac{\left(\rho_{pq} + k_{pq} \bar{S}_{pq}\right)^{2 - \lambda_1} \bar{S}^{\lambda_1}_{pp} \bar{S}^{\lambda_2}_{qq} \bar{S}^{\lambda_3}_{pq}}{\left(1 + k_{pq} \bar{S}_{pp}\right)^{1 + \lambda_2} \left(1 + k_{pq} \bar{S}_{qq}\right)^{1 + \lambda_3}}, \tag{2.7}$$

for some $k_{pq} = k_{pq}(S_{pp}, S_{qq}, S_{pq}) \in (0, 1)$, is the remainder in Lagrange's form. The "almost surely" qualifier is in (2.6) because on an event of measure zero, either S_{pp} or S_{qq} may be zero, in which case the Taylor's theorem doesn't apply since f is defined on $\mathbb{R}^2_{>0} \times \mathbb{R}$. Our proof depends crucially on recognizing that, when $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3) = (0, 0, 2)$,

$$\begin{aligned} &\frac{\partial^{\boldsymbol{\lambda}} f(1, 1, \rho_{pq})}{\boldsymbol{\lambda}!} \bar{S}^{\lambda_1}_{pp} \bar{S}^{\lambda_2}_{qq} \bar{S}^{\lambda_3}_{pq} = \bar{S}^2_{pq} \\ &= \frac{\sum_{i=1}^n (X_{pi} X_{qi} - \rho_{pq})^2}{n^2} + \frac{2 \sum_{1 \leq i < j \leq n} (X_{pi} X_{qi} - \rho_{pq})(X_{pj} X_{qj} - \rho_{pq})}{n^2}, \end{aligned}$$

in light of Lemma A.1 which specifies the partial derivatives of f. One can then equivalently write (2.6) as

$$\hat{\rho}^2_{pq} - \rho^2_{pq} = I_{pq} + II_{pq} + III_{pq}, \tag{2.8}$$

where

$$I_{pq} := \frac{2 \sum_{1 \leq i < j \leq n} (X_{pi} X_{qi} - \rho_{pq})(X_{pj} X_{qj} - \rho_{pq})}{n^2}, \text{ and} \tag{2.9}$$

$$II_{pq} \quad := \quad \frac{\sum_{i=1}^{n}(X_{pi}X_{qi}-\rho_{pq})^2}{n^2} \quad + \sum_{\substack{\boldsymbol{\lambda}\in\mathbb{N}^3_{\geq 0}:\\ 1\leq|\boldsymbol{\lambda}|\leq 4\\ \boldsymbol{\lambda}\neq(0,0,2)}} \frac{\partial^{\boldsymbol{\lambda}} f(1,1,\rho_{pq})}{\boldsymbol{\lambda}!}\bar{S}_{pp}^{\lambda_1}\bar{S}_{qq}^{\lambda_2}\bar{S}_{pq}^{\lambda_3}. \quad (2.10)$$

Defining $I := \sum_{1\leq p<q\leq m} I_{pq}$, $II := \sum_{1\leq p<q\leq m} II_{pq}$ and $III := \sum_{1\leq p<q\leq m} III_{pq}$ by summing over all $1 \leq p < q \leq m$, from (2.8) one can write

$$T - \frac{m(m-1)}{2n} - 2^{-1}\|\mathbf{R}-\mathbf{I}_{\mathrm{m}}\|_F^2 = I + \left(II - \frac{m(m-1)}{2n}\right) + III, \quad (2.11)$$

realizing that $2^{-1}\|\mathbf{R}-\mathbf{I}_m\|_F^2 = \sum_{1\leq p<q\leq m}\rho_{pq}^2$. We are now in the position to introduce three supporting lemmas that are the building blocks of Theorem 2.1. The first lemma gives a Berry-Esseen bound for the cumulative distribution function of the term I with $\Phi(\cdot)$ after standardization. This will ultimately drive the form of our power function in Theorem 2.1. The next two lemmas control the variability of the extra terms, $(II - \frac{m(m-1)}{2n})$ and III. From now on for the rest of this paper all the big O, little o notations are with respect to our considered asymptotic regime $m, n \longrightarrow \infty$, $m/n \leq \kappa$.

Lemma 2.3 (Berry Esseen theorem for I). *The following are true for I:*

i. Variance:

$$\mathit{Var}[I] = \mathbb{E}[I^2] = \frac{m^2}{n^2} + o\left(\frac{m^{2(1-\gamma)}}{n^2}\right)\sum_{k=0}^{2}\|\,\mathbf{R}-\mathbf{I}_{\mathrm{m}}\,\|_F^{2k}$$

for any $0 < \gamma < 1/2$.

ii. Berry-Esseen bound:

$$\sup_{t\in\mathbb{R}}\left|P\left(\frac{I}{\sqrt{\mathit{Var}(I)}} \leq t\right) - \Phi(t)\right| \lesssim \left\{\frac{o(m^4/n^4)\sum_{k=0}^{8}\|\,\mathbf{R}-\mathbf{I}_{\mathrm{m}}\,\|_F^k}{\mathit{Var}(I)^2}\right\}^{1/5}.$$

Lemma 2.4 (Bound on the 2nd moment of $II - \frac{m(m-1)}{2n}$).

$$\mathbb{E}\left[\left(II - \frac{m(m-1)}{2n}\right)^2\right] \lesssim \frac{\|\,\mathbf{R}-\mathbf{I}_{\mathrm{m}}\,\|_F^2 + \|\,\mathbf{R}-\mathbf{I}_{\mathrm{m}}\,\|_F^4}{n} + o\left(\frac{m^{2(1-\gamma)}}{n^2}\right)\sum_{k=0}^{4}\|\,\mathbf{R}-\mathbf{I}_{\mathrm{m}}\,\|_F^k, \quad (2.12)$$

for any fixed $0 < \gamma < 1/2$.

Lemma 2.5 (Probability bound for III). *For any $0 < c < \frac{1}{2}$, there exists $C > 0$ such that*

$$P\left(|III| > C\frac{m^2}{n^{5c}}\right) \lesssim (n^{c-1}\log m + n^{c-1/2}\sqrt{\log m})$$

for large enough m, n.

The proofs of Lemmas 2.3 and 2.4 are separately given in the next two sections. Lemma 2.5 is proved by a standard maximal inequality in Appendix A. With these tools we can now establish Theorem 2.1 based on the general approach laid out in Cai and Ma [4].

Proof of Theorem 2.1. From (2.5) and (2.11) the power of our test can be written as

$$\mathbb{E}[\psi] = P\left(I + II + III - \frac{m(m-1)}{2n} > \frac{m}{n}z_\alpha - 2^{-1}\|\mathbf{R} - \mathbf{I}_{\mathrm{m}}\|_F^2\right). \tag{2.13}$$

By dividing the set $\Theta(b)$ into two subsets

$$\Theta(b, B) = \{\mathbf{R} : B\sqrt{m/n} > \|\mathbf{R} - \mathbf{I}_{\mathrm{m}}\|_F \geq b\sqrt{m/n}\}$$

and

$$\Theta(B) = \{\mathbf{R} : \|\mathbf{R} - \mathbf{I}_{\mathrm{m}}\|_F \geq B\sqrt{m/n}\},$$

where B is a sufficiently large constant depending on (α, b, κ), it suffices to show

$$\liminf_{n\to\infty} \inf_{\Theta(B)} \mathbb{E}_{\mathbf{R}}[\psi] \geq \bar{\Phi}\left(z_\alpha - \frac{b^2}{2}\right) \tag{2.14}$$

and

$$\sup_{\Theta(b,B)} \left|\mathbb{E}_{\mathbf{R}}\psi - \bar{\Phi}\left(z_\alpha - \frac{\|\mathbf{R} - \mathbf{I}_{\mathrm{m}}\|_F^2}{2m/n}\right)\right| \longrightarrow 0 \tag{2.15}$$

as $m, n \longrightarrow \infty$, $m/n \leq \kappa$. Together, they lead to the theorem since (2.15) implies that

$$\lim_{n\to\infty} \inf_{\Theta(b,B)} \mathbb{E}_{\mathbf{R}}\psi = \lim_{n\to\infty} \inf_{\Theta(b,B)} \bar{\Phi}\left(z_\alpha - \frac{\|\mathbf{R} - \mathbf{I}_{\mathrm{m}}\|_F^2}{2m/n}\right) = \bar{\Phi}\left(z_\alpha - \frac{b^2}{2}\right).$$

To prove (2.14) we first suppose that B is larger than $\sqrt{3z_\alpha}$, and let δ be any positive constant satisfying $0 < \delta \leq 4^{-1}z_\alpha$. By definition, for any $\mathbf{R} \in \Theta(B)$, it must be the case that $\|\mathbf{R} - \mathbf{I}_{\mathrm{m}}\|_F = \tau\sqrt{m/n}$ for some $\tau \geq B$. Together with the fact that $mn^{-1}z_\alpha - 2^{-1}\|\mathbf{R} - \mathbf{I}_{\mathrm{m}}\|_F^2 \leq -\frac{m\tau^2}{n6}$ and $\delta \leq 12^{-1}\tau^2$ which are consequences of the choice of B, by a union bound and Chebyshev's inequality we continue from (2.13) and obtain

$$\begin{aligned}
1 - \mathbb{E}[\psi] &\leq P\left(\left|I + II - \frac{m(m-1)}{2n}\right| \geq \frac{\tau^2 m}{6n} - \delta\frac{m}{n}\right) + P\left(|III| > \delta\frac{m}{n}\right) \\
&\leq 288\tau^{-4}n^2m^{-2}\left(\mathbb{E}[I^2] + \mathbb{E}\left[\left(II - \frac{m(m-1)}{2n}\right)^2\right]\right) + P\left(|III| > \delta\frac{m}{n}\right).
\end{aligned} \tag{2.16}$$

Substituting $\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F$ for $\tau\sqrt{m/n}$ into the bounds for $\mathbb{E}[I^2]$ and $\mathbb{E}[(II - \frac{m(m-1)}{2n})^2]$ in Lemmas 2.3 and 2.4, it is seen that the first term in (2.16) is bounded by a term of order

$$\tau^{-4} + o(1)\left(\sum_{k=0}^{4} \tau^{-k}\right)$$

Moreover, the second term in (2.16) converges to 0 as $m, n \longrightarrow \infty$ by Lemma 2.5 since $\delta m/n$ is larger than m^2/n^{5c} asymptotically for any constant $2/5 < c < 1/2$, given that $m/n \leq \kappa$. They together imply that the constant $B = B(\alpha, b, \kappa)$ can be taken large enough so that

$$1 - \inf_{\Theta(B)} \mathbb{E}_{\mathbf{R}}[\psi] \leq \Phi\left(z_\alpha - \frac{b^2}{2}\right) \text{ as } m, n \longrightarrow \infty,$$

which is equivalent to (2.14).

To show (2.15), the uniform convergence of power on the "stripe" of alternatives with the signal $\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F$ bounded from above and below in size, we shall first establish that

$$P\left(|\tilde{I}| \geq \frac{m^{1-\gamma}}{n}\right) = o(1) \quad \text{as} \quad m, n \longrightarrow \infty \quad \text{and} \quad m/n \leq \kappa, \tag{2.17}$$

uniformly over the set $\Theta(b, B)$, where

$$\tilde{I} := II - \frac{m(m-1)}{2n} + III.$$

and γ is any number such that $0 < \gamma < 1/2$. By a union bound we have

$$\begin{aligned} P\left(|\tilde{I}| \geq \frac{m^{1-\gamma}}{n}\right) &\leq P\left(|III| \geq \frac{m^{1-\gamma}}{2n}\right) + P\left(\left|II - \frac{m(m-1)}{2n}\right| \geq \frac{m^{1-\gamma}}{2n}\right) \\ &\lesssim n^{c-1}\log m + n^{c-1/2}\sqrt{\log m} + \frac{n^2}{m^{2(1-\gamma)}}\mathbb{E}\left[\left(II - \frac{m(m-1)}{2n}\right)^2\right] \end{aligned} \tag{2.18}$$

for any $(2+\gamma)/5 < c < 1/2$ and large enough m, n. The last inequality comes from the Chebyshev inequality and the fact that, by taking $(2+\gamma)/5 < c < 1/2$ in Lemma 2.5, for large enough m, n, under $m/n \leq \kappa$, we have

$$P\left(|III| \geq \frac{m^{1-\gamma}}{2n}\right) \leq P\left(|III| \geq \frac{m^2}{2\kappa^{1+\gamma}n^{2+\gamma}}\right) \leq P\left(|III| \geq C\frac{m^2}{n^{5c}}\right),$$

where the constant C is same as the one in Lemma 2.5. Since $\mathbf{R} \in \Theta(b, B)$, it must be that $\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F = \tau\sqrt{m/n}$ for some $b \leq \tau \leq B$, and substituting this into the variance bound in Lemma 2.4 it can be easily seen that

$$\frac{n^2}{m^{2(1-\gamma)}}\mathbb{E}\left[\left(II - \frac{m(m-1)}{2n}\right)^2\right] \longrightarrow 0 \tag{2.19}$$

uniformly over $\Theta(b, B)$ as $m, n \longrightarrow \infty$, $m/n \leq \kappa$. This gives (2.17) since $c < 1/2$ in (2.18).

To finish the proof of (2.15), by union bound arguments one has

$$\mathbb{E}[\psi] \leq P\left(I \geq \frac{mz_\alpha}{n} - \frac{\| \mathbf{R} - \mathbf{I}_\mathrm{m} \|_F^2}{2} - \frac{m^{1-\gamma}}{n}\right) + P\left(|\tilde{I}| \geq \frac{m^{1-\gamma}}{n}\right)$$

and

$$\mathbb{E}[\psi] \geq P\left(I \geq \frac{mz_\alpha}{n} - \frac{\| \mathbf{R} - \mathbf{I}_\mathrm{m} \|_F^2}{2} + \frac{m^{1-\gamma}}{n}\right) - P\left(|\tilde{I}| \geq \frac{m^{1-\gamma}}{n}\right),$$

which collectively imply

$$\begin{aligned} &\left|\mathbb{E}[\psi] - \bar{\Phi}\left(\frac{mz_\alpha n^{-1} - 2^{-1}\| \mathbf{R} - \mathbf{I}_\mathrm{m} \|_F^2}{\sqrt{\mathrm{Var}(I)}}\right)\right| \\ &\qquad \leq \sup_{t\in\mathbb{R}} \left| P\left(\frac{I}{\sqrt{\mathrm{Var}(I)}} \geq t\right) - \bar{\Phi}\left(t\right)\right| + 2P\left(|\tilde{I}| \geq \frac{m^{1-\gamma}}{n}\right) + \frac{2m^{1-\gamma}n^{-1}}{\sqrt{\mathrm{Var}(I)}} \end{aligned} \tag{2.20}$$

since $|\bar{\Phi}(x \pm \epsilon) - \bar{\Phi}(x)| \leq \epsilon$ for any $x \in \mathbb{R}$ and $\epsilon \geq 0$. Moreover, all three terms on the right hand side of (2.20) are of order $o(1)$ uniformly over $\Theta(b, B)$. The first two terms are so by Lemma 2.3(ii) and (2.17), and the last term is so since by Lemma 2.3(i), $\sqrt{\mathrm{Var}(I)} = m/n + o(m^{1-\gamma}/n)$ where the $o(m^{1-\gamma}/n)$ term is also uniform over $\Theta(b, B)$. Finally, by Lemma 2.3(i) as $m, n \longrightarrow \infty$, $m/n \leq \kappa$, we also have

$$\sup_{\Theta(b,B)} \left|\frac{\mathrm{Var}(I)}{m^2/n^2} - 1\right| \longrightarrow 0, \tag{2.21}$$

and it is not hard to see that this implies

$$\sup_{\Theta(b,B)} \left|\bar{\Phi}\left(z_\alpha - \frac{\| \mathbf{R} - \mathbf{I}_\mathrm{m} \|_F^2}{2m/n}\right) - \bar{\Phi}\left(\frac{mz_\alpha n^{-1} - 2^{-1}\| \mathbf{R} - \mathbf{I}_\mathrm{m} \|_F^2}{\sqrt{\mathrm{Var}(I)}}\right)\right| \longrightarrow 0.$$

Applying these facts to (2.20) leads to (2.15). □

In establishing the normal tail form of our power function, perhaps the most important step is singling out I as the main term that drives the asymptotic normality of the left hand side in (2.11) under the "stripe" of alternative $\Theta(b, B)$ via the Berry-Esseen bound in Lemma 2.3(ii). We note that I is already a rather simple term to handle, but proving Lemma 2.3(ii) for it still takes considerable effort in the next section. Moreover, $m/n \leq \kappa$ has been used at different places, the convergences in (2.19) and (2.21) for instances. However, the assumption is mostly a convenient one for such statements regarding terms I and II, since the estimates presented in Lemmas 2.3 and 2.4 are not the sharpest possible, for either aesthetic purpose or saving us some effort on refining them in the next two sections.

It is the remainder term III that truly prevents us from removing the upper bound on m/n. In order to show it tends to zero in probability, as in (2.18), we applied the crude tail bound in Lemma 2.5 based on a maximal inequality (see Appendix A). Such an estimate doesn't take the correlations among the constituent summands III_{pq} into account, as was done for the II_{pq}'s with respective to $II-(m-1)m(2n)^{-1}$ via explicitly estimating its second moment in Lemma 2.4. The major obstacle to computing $\mathbb{E}[III^2]$ is the *random* coefficients

$$\frac{(\rho_{pq}+k_{pq}\bar{S}_{pq})^{2-\lambda_1}}{(1+k_{pq}\bar{S}_{pp})^{1+\lambda_2}(1+k_{pq}\bar{S}_{qq})^{1+\lambda_3}} \tag{2.22}$$

attached to the products $\bar{S}_{pp}^{\lambda_1}\bar{S}_{qq}^{\lambda_2}\bar{S}_{pq}^{\lambda_3}$ in definition (2.7). Unlike II, where the constituents II_{pq} have *constant* coefficients, not only is the coefficient in (2.22) a *rational* functions in $\bar{S}_{pp}$, $\bar{S}_{pq}$, $\bar{S}_{qq}$, but it also involves the intractable random quantity $k_{pq}=k_{pq}(\bar{S}_{pp},\bar{S}_{pq},\bar{S}_{qq})\in(0,1)$. As such, there is no straightforward way of applying Isserlis's theorem (Theorem A.2) to compute the moment $\mathbb{E}[III^2]$ like we did for $\mathbb{E}[(II-(m-1)m(2n)^{-1})^2]$ in Section 4. In fact, even with the help of `mathematica`, it still took us substantial effort to get our bound in Lemma 2.4 as seen later. At this moment, we cannot think of other ways to control term III.

3. The Berry Esseen bound for I

We will prove Lemma 2.3 in this section. For our presentation, given a finite set D and $|D|$ duples $(p_d,q_d)\in[m]\times[m]$ indexed by a subscript d that ranges over D, we define the central moment quantities

$$\mathcal{M}_{\substack{(p_d,q_d)\\ d\in D}} := \mathbb{E}\left[\prod_{d\in D}(X_{p_d}X_{q_d}-\rho_{p_dq_d})\right].$$

Recall that I is defined as $\sum_{p<q} I_{pq}$, where each I_{pq} is given in (2.9). We first observe that I has a natural martingale structure: For each $i=1,\dots,n$, let $\mathcal{F}_i$ be the sigma-algebra generated by $\{X_{pj}:1\le p\le m; 1\le j\le i\}$ and $\mathcal{F}_0$ be the trivial sigma algebra, and define

$$Y_i := \frac{2}{n^2}\sum_{p<q}\sum_{j<i}(X_{pi}X_{qi}-\rho_{pq})(X_{pj}X_{qj}-\rho_{pq}) \text{ for } i=2,\dots,n \tag{3.1}$$

as well as

$$Y_0 = Y_1 := 0. \tag{3.2}$$

Then $I=\sum_{i=0}^n Y_i$, and $(Y_i)_{i=0}^n$ is a the sequence of martingale differences since

$$\mathbb{E}[Y_i|\mathcal{F}_{i-1}] = \sum_{p<q}\frac{2}{n^2}\sum_{j<i}(X_{pj}X_{qj}-\rho_{pq})\mathbb{E}[X_{pi}X_{qi}-\rho_{pq}] = 0$$

for $i \geq 2$, where $\mathbb{E}[Y_i|\mathcal{F}_{i-1}] = 0$ is trivial for $i = 0, 1$.

With the observations just made it is easy to see that $\mathbb{E}[I] = 0$ and

$$\mathrm{Var}[I] = \mathbb{E}[I^2] = \sum_{i=2}^{n} \mathbb{E}[Y_i^2]. \tag{3.3}$$

By the i.i.d.'ness of the samples, for each $i = 2, \dots, n$,

$$\begin{aligned}\mathbb{E}[Y_i^2] &= \frac{4}{n^4} \sum_{\substack{1 \leq p_d < q_d \leq m \\ d=1,2}} \mathcal{M}_{\substack{(p_d,q_d) \\ d\in[2]}} \left(\sum_{1 \leq j,j' < i} \mathbb{E}\left[(X_{p_1 j'} X_{q_1 j'} - \rho_{p_1 q_1})(X_{p_2 j} X_{q_2 j} - \rho_{p_2 q_2})\right] \right) \\ &= \frac{4(i-1)}{n^4} \sum_{\substack{1 \leq p_d < q_d \leq m \\ d=1,2}} \mathcal{M}^2_{\substack{(p_d,q_d) \\ d\in[2]}}, \end{aligned} \tag{3.4}$$

where, to clarify, $\sum_{\substack{1 \leq p_d < q_d \leq m \\ d=1,2}}$ means a summation over all pairs of duples $\{(p_1, q_1), (p_2, q_2)\}$ such that $1 \leq p_d < q_d \leq m$ for each $d = 1, 2$. We have the equality in (3.4) because $\mathbb{E}[(X_{p_1 j'} X_{q_1 j'} - \rho_{p_1 q_1})(X_{p_2 j} X_{q_2 j} - \rho_{p_2 q_2})]$ equals $\mathcal{M}_{\substack{(p_d,q_d) \\ d\in\{1,2\}}}$ when $j = j'$ and zero otherwise. For $k = 2, 3, 4$, let

$$\mathbb{S}(k) := \sum_{\substack{1 \leq p_d < q_d \leq m \\ d=1,2 \\ |\cup_{d=1}^2 \{p_d,q_d\}| = k}} \mathcal{M}^2_{\substack{(p_d,q_d) \\ d\in[2]}} \tag{3.5}$$

correspond to a sum over all duples $1 \leq p_d < q_d \leq m$, $d = 1, 2$ such that as a set $\cup_{d=1}^2 \{p_d, q_d\}$ has cardinality k. From (3.3) and (3.4) we can write

$$\mathrm{Var}[I] = \frac{2n(n-1)}{n^4} \sum_{k=2}^{4} \mathbb{S}(k). \tag{3.6}$$

since $\sum_{i=2}^{n} (i-1) = 2^{-1}(n^2 - n)$. In Appendix C of our supplement [15], we will show the following estimates hold:

$$\mathbb{S}(2) = 2^{-1} m(m-1) + O(\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^2) \tag{3.7}$$

$$\mathbb{S}(3) = O(m \| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^2 + \| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^4) \tag{3.8}$$

$$\mathbb{S}(4) = O(\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^4) \tag{3.9}$$

Substituting these into (3.6) results in Lemma 2.3(i). In fact, this general strategy of decomposing a sum according to the cardinality of an index set as in (3.5) and forming separate estimates will be employed repeatedly in the sequel.

We shall now prove the normal approximation in Lemma 2.3(ii). With a Berry-Esseen theorem for martingale central limit theorem in Heyde and Brown [8], it suffices to verify the fourth moment conditions

$$\sum_{i=2}^{n} \mathbb{E}[Y_i^4] = o(m^4/n^4) \sum_{k=0}^{4} \| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^k \tag{3.10}$$

and

$$\mathbb{E}\left[\left(\sum_{i=2}^{n}\mathbb{E}[Y_i^2|\mathcal{F}_{i-1}]-\mathrm{Var}(I)\right)^2\right]=\mathbb{E}\left[\left(\sum_{i=2}^{n}\mathbb{E}[Y_i^2|\mathcal{F}_{i-1}]\right)^2\right]-\mathrm{Var}(I)^2$$
$$=o(m^4/n^4)\sum_{k=0}^{8}\|\,\mathbf{R}-\mathbf{I}_{\mathrm{m}}\,\|_F^k. \tag{3.11}$$

Note that the equality before (3.11) holds because $\mathbb{E}[\sum_{i=2}^n \mathbb{E}[Y_i^2|\mathcal{F}_{i-1}]] = \mathbb{E}[\sum_{i=2}^n Y_i^2] = \mathrm{Var}(I)$.

We will first show (3.10). For any $2 \leq i \leq n$, on raising Y_i to the 4th power and taking expectation, by the i.i.d.'ness of samples, we have

$$\begin{aligned}
&\mathbb{E}[Y_i^4]\\
&=\frac{16}{n^8}\sum_{\substack{1\le p_d<q_d\le m\\ d=1,2,3,4}}\left\{\mathbb{E}\left[\prod_{d=1}^{4}(X_{p_d i}X_{q_d i}-\rho_{p_d q_d})\right]\sum_{\substack{1\le j_d<i\\ d=1,2,3,4}}\mathbb{E}\left[\prod_{d=1}^{4}(X_{p_d j_d}X_{q_d j_d}-\rho_{p_d q_d})\right]\right\}\\
&=\frac{16}{n^8}\sum_{\substack{1\le p_d<q_d\le m\\ d=1,2,3,4}}\left\{\mathcal{M}_{\substack{(p_d,q_d)\\ d\in[4]}}\sum_{\substack{1\le j_d<i\\ d=1,2,3,4}}\mathbb{E}\left[\prod_{d=1}^{4}(X_{p_d j_d}X_{q_d j_d}-\rho_{p_d q_d})\right]\right\}\\
&=O\left(\frac{i^2}{n^8}\right)\sum_{\substack{1\le p_d<q_d\le m\\ d=1,2,3,4}}\mathcal{M}_{\substack{(p_d,q_d)\\ d\in[4]}}, \qquad (3.12)
\end{aligned}$$

where the summations $\sum_{\substack{1\le p_d<q_d\le m\\ d=1,2,3,4}}$ and $\sum_{\substack{1\le j_d<i\\ d=1,2,3,4}}$ are defined similarly as the one in (3.4). The last equality in (3.12) is explained as follows: For a fixed i and a given set of variables index pairs $\{(p_d, q_d) : d = 1, \dots, 4\}$, with any choice of the sample indices $j_1, \dots, j_4$ in order for the expectation

$$\mathbb{E}\left[\prod_{d=1}^{4}(X_{p_d j_d}X_{q_d j_d}-\rho_{p_d q_d})\right] \tag{3.13}$$

to be non-zero, by independence it must be true that there exists a permutation function $\pi \in \mathcal{S}_4$ so that

$$j_{\pi(1)} = j_{\pi(2)}, \quad j_{\pi(3)} = j_{\pi(4)}. \tag{3.14}$$

Since the condition in (3.14) implies that $|\cup_{d=1}^4 \{j_d\}| \leq 2$, at most $O(\binom{i-1}{2}) = O(i^2)$ many expectations in (3.13) can be non-zero. This leads to (3.12) since the expectations in (3.13), when they are non-zero, can be uniformly bounded regardless of the choice for $\{(p_d, q_d, j_d); d = 1, \dots, 4\}$, owing to our assumptions at the beginning of Section 2 and Isserlis' theorem (Theorem A.2) on higher order normal moments. Provided that $\sum_{i=2}^{n} i^2 = 6^{-1}(2n^3+3n^2+n-6)$, with (3.12) we further write

$$\sum_{i=2}^{n} \mathbb{E}[Y_i^4] = O(n^{-5}) \sum_{\substack{1\le p_d<q_d\le m\\ d=1,2,3,4}} \mathcal{M}_{\substack{(p_d,q_d)\\ d\in[4]}}. \tag{3.15}$$

Now the last term in (3.15) can be decomposed, according to the cardinality of the set of duples $\cup_{d=1}^{4}\{p_d,q_d\}$, as

$$\sum_{\substack{1\le p_d<q_d\le m\\ d=1,2,3,4}} \mathcal{M}_{\substack{(p_d,q_d)\\ d\in[4]}} = \sum_{k=5}^{8} \mathbb{T}(k) + O(m^4), \tag{3.16}$$

where for $k=2,\dots,8$,

$$\mathbb{T}(k) := \sum_{\substack{1\le p_d<q_d\le m\\ d=1,2,3,4\\ |\cup_{d=1}^{4}\{p_d,q_d\}|=k}} \mathcal{M}_{\substack{(p_d,q_d)\\ d\in[4]}}$$

and the $O(m^4)$ term comes from the fact that there are only $O(m^4)$ many uniformly bounded extra summands under the restriction $|\cup_{k=1}^{4}\{p_d,q_d\}| \le 4$. In Appendix C [15] we will show that

$$\mathbb{T}(k) = O(m^4)\|\,\mathbf{R}-\mathbf{I}_{\mathrm{m}}\,\|_F^{k-4} \tag{3.17}$$

for each $k=5,\dots,8$. Collecting (3.15), (3.16) and (3.17) we get (3.10).

To show (3.11) it suffices to understand the term $\mathbb{E}[(\sum_{i=1}^{n}\mathbb{E}[Y_i^2|\mathcal{F}_{i-1}])^2]$ since the form of $\mathrm{Var}(I)$ has been proven in Lemma 2.3(i). On expansion,

$$\sum_{i=2}^{n} \mathbb{E}[Y_i^2|\mathcal{F}_{i-1}] = \frac{4}{n^4}\sum_{i=2}^{n} \sum_{\substack{1\le p_d<q_d\le m\\ d=1,2}} \mathcal{M}_{\substack{(p_d,q_d)\\ d\in[2]}} \left[\sum_{1\le j,k<i} (X_{p_1 j}X_{q_1 j}-\rho_{p_1q_1})(X_{p_2 k}X_{q_2 k}-\rho_{p_2q_2})\right]. \tag{3.18}$$

Proceeding with our calculations,

$$\mathbb{E}\left[\left(\sum_{i=2}^{n}\mathbb{E}[Y_i^2|\mathcal{F}_{i-1}]\right)^2\right] = \frac{16}{n^8}\sum_{\substack{1\le p_d<q_d\le m\\ d=1,2,3,4}} \left\{\mathbb{P}_1 \times \sum_{2\le i,j\le n} \sum_{\substack{1\le i_1,i_2<i\\ 1\le i_3,i_4<j}} \mathbb{E}\left[\prod_{d=1}^{4}(X_{p_d i_d}X_{q_d i_d}-\rho_{p_dq_d})\right]\right\}, \tag{3.19}$$

where

$$\mathbb{P}_1 = \mathbb{P}_1(p_1, q_1, \dots, p_4, q_4) := \mathcal{M}_{\substack{(p_d,q_d)\\ d\in\{1,2\}}} \mathcal{M}_{\substack{(p_d,q_d)\\ d\in\{3,4\}}} \tag{3.20}$$

By independence, we note that the expression

$$\mathbb{E}\left[\prod_{d=1}^{4}(X_{p_d i_d}X_{q_d i_d} - \rho_{p_d q_d})\right]$$

on the right hand side of (3.19) can be non-zero only if the four sample indices $i_1, \dots, i_4$ are such that either

$$i_1 = \cdots = i_4, \tag{3.21}$$

$$i_1 = i_2, \quad i_3 = i_4, \quad |\{i_1, \dots, i_4\}| = 2, \tag{3.22}$$

$$i_1 = i_3, \quad i_2 = i_4, \quad |\{i_1, \dots, i_4\}| = 2 \tag{3.23}$$

or

$$i_1 = i_4, \quad i_2 = i_3, \quad |\{i_1, \dots, i_4\}| = 2. \tag{3.24}$$

For any fixed given pair $2 \le i, j \le n$, by simple counting, there are, respectively, $i \wedge j - 1$, $(i \wedge j - 1)(i \vee j - 2)$, $(i \wedge j - 1)(i \wedge j - 2)$, $(i \wedge j - 1)(i \wedge j - 2)$ combinations of (i_1, i_2, i_3, i_4) that satisfy (3.21), (3.22), (3.23), (3.24) for which $1 \le i_1, i_2 < i$ and $1 \le i_3, i_4 < j$, where $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$. Hence,

$$\begin{aligned}
&\sum_{2\le i,j\le n} \sum_{\substack{1\le i_1,i_2<i\\ 1\le i_3,i_4<j}} \mathbb{E}\left[\prod_{d=1}^{4}(X_{p_d i_d}X_{q_d i_d} - \rho_{p_d q_d})\right] \\
&= \mathcal{M}_{\substack{(p_d,q_d)\\ d\in[4]}} \underbrace{\left\{\sum_{2\le i,j\le n}(i\wedge j-1)\right\}}_{=6^{-1}(2n^3-3n^2+n)} + \mathbb{P}_1 \underbrace{\sum_{2\le i,j\le n}(i\wedge j-1)(i\vee j-2)}_{=12^{-1}(-2n+9n^2-10n^3+3n^4)} \\
&\quad + (\mathbb{P}_2+\mathbb{P}_3)\underbrace{\left\{\sum_{2\le i,j\le n}(i\wedge j-1)(i\wedge j-2)\right\}}_{=6^{-1}(n^4-4n^3+5n^2-2n)} \\
&= \mathcal{M}_{\substack{(p_d,q_d)\\ d\in[4]}} O(n^3) + \mathbb{P}_1\left(\frac{n^4}{4} + O(n^3)\right) + (\mathbb{P}_2+\mathbb{P}_3)\, O(n^4),
\end{aligned} \tag{3.25}$$

where

$$\mathbb{P}_2 = \mathbb{P}_2(p_1, q_1, \dots, p_4, q_4) := \mathcal{M}_{\substack{(p_d,q_d)\\ d\in\{1,3\}}} \mathcal{M}_{\substack{(p_d,q_d)\\ d\in\{2,4\}}}$$

$$\mathbb{P}_3 = \mathbb{P}_3(p_1, q_1, \dots, p_4, q_4) := \mathcal{M}_{\substack{(p_d,q_d)\\ d\in\{1,4\}}} \mathcal{M}_{\substack{(p_d,q_d)\\ d\in\{2,3\}}}$$

are the value of $\mathbb{E}[\prod_{d=1}^4 (X_{p_d i_d} X_{q_d i_d} - \rho_{p_d q_d})]$ when $i_1, \ldots, i_4$ satisfy the criteria (3.23) and (3.24) respectively. Substituting (3.25) into (3.19) gives

$$\mathbb{E}\left[\left(\sum_{i=2}^n \mathbb{E}[Y_i^2|\mathcal{F}_{i-1}]\right)^2\right] =$$

$$O(n^{-5}) \sum_{\substack{1\le p_d<q_d\le m \\ d=1,2,3,4}} \mathbb{P}_1 + \left(\frac{4}{n^4} + O(n^{-5})\right) \sum_{\substack{1\le p_d<q_d\le m \\ d=1,2,3,4}} \mathbb{P}_1^2 + O(n^{-4}) \sum_{\substack{1\le p_d<q_d\le m \\ d=1,2,3,4}} \sum_{u=2}^3 \mathbb{P}_1 \mathbb{P}_u, \tag{3.26}$$

where the terms $\mathcal{M}_{\substack{(p_d,q_d) \\ d\in[4]}}$ in(3.25) are absorbed into the first $O(n^{-5})$ term because they are uniformly bounded regardless of the choice of $p_1, q_1, \ldots, p_4, q_4$, again by our assumptions and Theorem A.2. From this it remains to show the estimates

$$\sum_{\substack{1\le p_d<q_d\le m \\ d=1,2,3,4}} \mathbb{P}_1 = O(m^4) \sum_{k=0}^4 \| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^k, \tag{3.27}$$

$$\sum_{\substack{1\le p_d<q_d\le m \\ d=1,2,3,4}} \mathbb{P}_1^2 = \frac{m^4}{4} + O(m^3) \sum_{k=0}^8 \| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^k, \tag{3.28}$$

and

$$\sum_{\substack{1\le p_d<q_d\le m \\ d=1,2,3,4}} \sum_{u=2}^3 \mathbb{P}_1 \mathbb{P}_u = O(m^3) \sum_{k=0}^8 \| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^k, \tag{3.29}$$

which, together with Lemma 2.3(i) and (3.26), imply (3.11). The proofs of these estimates will, again, be deferred to Appendix C [15].

4. The second moment bound for $\boldsymbol{II - \frac{m(m-1)}{2n}}$

We will now prove Lemma 2.4. Recall that $II := \sum_{p<q} II_{pq}$, and from the definition of II_{pq} in (2.10) we can equivalently write it as

$$II_{pq} = II_{pq,1} + II_{pq,2},$$

where

$$II_{pq,1} \quad := \quad \frac{\sum_{i=1}^n (X_{pi}X_{qi} - \rho_{pq})^2}{n^2} \quad + \sum_{\substack{\boldsymbol{\lambda}\in\mathbb{N}_{\ge 0}^3: \\ 3\le|\boldsymbol{\lambda}|\le 4 \\ \lambda_3=2 \\ \boldsymbol{\lambda}\ne(1,1,2)}} \frac{\partial^{\boldsymbol{\lambda}} f(1,1,\rho_{pq})}{\boldsymbol{\lambda}!} \bar{S}_{pp}^{\lambda_1} \bar{S}_{qq}^{\lambda_2} \bar{S}_{pq}^{\lambda_3} \tag{4.1}$$

and

$$II_{pq,2} \quad := \quad \frac{\partial^{(1,1,2)} f(1,1,\rho_{pq})}{1!1!2!} \bar{S}_{pp} \bar{S}_{qq} \bar{S}_{pq}^2 \; + \sum_{\substack{\boldsymbol{\lambda} \in \mathbb{N}^3_{\geq 0}: \\ 1 \leq |\boldsymbol{\lambda}| \leq 4 \\ \lambda_3 \neq 2}} \frac{\partial^{\boldsymbol{\lambda}} f(1,1,\rho_{pq})}{\boldsymbol{\lambda}!} \bar{S}_{pp}^{\lambda_1} \bar{S}_{qq}^{\lambda_2} \bar{S}_{pq}^{\lambda_3}. \quad (4.2)$$

We form this grouping of terms for reasons that will be explained later. As such, by defining $II_1 := \sum_{p<q} II_{pq,1}$ and $II_2 := \sum_{p<q} II_{pq,2}$, one can write

$$II = II_1 + II_2.$$

To finish the proof of Lemma 2.4, it suffices to bound the second moments of $II_1 - \frac{m(m-1)}{2n}$ and II_2 respectively in terms of $\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F$.

Lemma 4.1 (Bound on the second moment of $II_1 - \frac{m(m-1)}{2n}$)**.**

$$\mathbb{E}\left[\left(II_1 - \frac{m(m-1)}{2n}\right)^2\right] \lesssim o\left(\frac{m^{2(1-\gamma)}}{n^2}\right) \sum_{k=0}^{4} \| \mathbf{R} - I \|_F^k$$

for any $0 < \gamma < 1/2$.

Lemma 4.2 (Bound on the second moment of II_2)**.**

$$\mathbb{E}\left[(II_2)^2\right] \lesssim \frac{\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^2 + \| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^4}{n} + o\left(\frac{m^{2(1-\gamma)}}{n^2}\right) \sum_{k=0}^{2} \| \mathbf{R} - I \|_F^k \quad (4.3)$$

for any $0 < \gamma < 1/2$.

Using Lemmas 4.1 and 4.2, Lemma 2.4 immediately follows from i. $II^2 = (II_1 - \frac{m(m-1)}{2})^2 + II_2^2 + 2(II_1 - \frac{m(m-1)}{2}) II_2$ and ii. $2|(II_1 - \frac{m(m-1)}{2}) II_2| \leq (II_1 - \frac{m(m-1)}{2})^2 + II_2^2$.

For each pair $p < q$, the main difference between $II_{pq,1}$ and $II_{pq,2}$ is that when $\lambda_3 \neq 2$, all the coefficients $\frac{\partial^{\boldsymbol{\lambda}} f(1,1,\rho_{pq})}{\boldsymbol{\lambda}!}$ appearing in the second term of (4.2) can be bounded by either $|\rho_{pq}|$ or ρ_{pq}^2 up to some multiplicative constants. This makes proving the useful bound for $\mathbb{E}[II_2^2]$ in terms of the norm $\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F$ amenable to the straightforward approach of squaring and taking expectation. Thus we shall defer the proof of Lemma 4.2 to Appendix D of our supplement [15] and address the bound in Lemma 4.1 for the rest of this section.

We will start with the fact that

$$\mathbb{E}\left[\left(II_1 - \frac{m(m-1)}{2n}\right)^2\right] \leq 2\left\{\mathrm{Var}[II_1] + \left(\mathbb{E}[II_1] - \frac{m(m-1)}{2n}\right)^2\right\} \quad (4.4)$$

and form estimates for the terms on the right hand side. To understand the mean and variance of II_1, it is more instructive to first recognize that each term in (4.1) can be written as a U-statistic of degree 4. For instance, for any four distinct indices $1 \leq i, j, k, l \leq n$, if we only treat $\mathbf{X}_{pq,i} = (X_{pi}, X_{qi})', \ldots, \mathbf{X}_{pq,l} = (X_{pl}, X_{ql})'$ as a four tuple in $\mathbb{R}^2$, the function

$$h_{1,pq}(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l}) := \frac{\binom{n}{4}}{n^2\binom{n-1}{3}} \sum_{i' \in \{i,j,k,l\}} \left\{(X_{pi'}X_{qi'} - \rho_{pq})^2\right\}, \tag{4.5}$$

is symmetric in its four arguments, and the first term in (4.1) can be written as the U-statistic

$$n^{-2}\sum_{i=1}^n (X_{pi}X_{qi} - \rho_{pq})^2 = \binom{n}{4}^{-1} \sum h_{1,pq}(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l}) \tag{4.6}$$

where the summation on the right hand side is over all distinct unordered qradruples i, j, k, l that can be formed from $[n]$. We note that the factor $\binom{n-1}{3}$ appears as a denominator in (4.5) because for each $i \in \{1, \ldots, n\}$, the summand $(X_{pi}X_{qi} - \rho_{pq})^2$ will appear only once on the left hand side of (4.6), while by the definition of $h_{1,pq}$ it will appear in $\binom{n-1}{3}$ kernels that are summed over on the right hand side of (4.6) (Since for each i, there will be $\binom{n-1}{3}$ choices of j, k, l to form a quadruple (i, j, k, l) from $\{1, \ldots, n\}$). Thus, the factor $\binom{n-1}{3}$ appears as a denominator in definition (4.5) to account for the multiple counting.

Note that the other terms of the form $\frac{\partial^{\boldsymbol{\lambda}} f(1,1,\rho_{pq})}{\boldsymbol{\lambda}!}\bar{S}_{pp}^{\lambda_1}\bar{S}_{qq}^{\lambda_2}\bar{S}_{pq}^{\lambda_3}$ in (4.1) are indexed by $\boldsymbol{\lambda}$ equal to $(1,0,2)$, $(0,1,2)$, $(2,0,2)$, $(0,2,2)$. These terms can be represented as U-statistics of degree 4 using a similar strategy: With four *distinct* indices i, j, k, l from $[n]$, by defining the symmetric kernel function

$$\begin{aligned}
&h_{2,pq}(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l}) \qquad (4.7)\\
&:= \underbrace{\binom{n}{4}\frac{n^{-3}}{\binom{n-3}{1}}}_{O(1)} \sum_{\substack{\{i',j',k'\} \\ \subset\{i,j,k,l\} \\ i',j',k' \text{ distinct} \\ \text{and unordered}}} \sum_{\pi \in \mathcal{S}_3} \left\{(X^2_{p\pi(i')} - 1)(X_{p\pi(j')}X_{q\pi(j')} - \rho_{pq})(X_{p\pi(k')}X_{q\pi(k')} - \rho_{pq})\right\}\\
&+ \underbrace{\binom{n}{4}\frac{n^{-3}}{\binom{n-2}{2}}}_{O(n^{-1})} \sum_{\substack{\{i',j'\} \\ \subset\{i,j,k,l\} \\ i',j' \text{ distinct} \\ \text{and unordered}}} \sum_{\pi \in \mathcal{S}_2} \Bigg\{(X^2_{p\pi(i')} - 1)(X_{p\pi(j')}X_{q\pi(j')} - \rho_{pq})^2\\
&\qquad + 2(X^2_{p\pi(i')} - 1)(X_{p\pi(i')}X_{q\pi(i')} - \rho_{pq})(X_{p\pi(j')}X_{q\pi(j')} - \rho_{pq})\Bigg\}\\
&+ \underbrace{\binom{n}{4}\frac{n^{-3}}{\binom{n-1}{3}}}_{O(n^{-2})} \sum_{i' \in \{i,j,k,l\}} \left\{(X^2_{pi'} - 1)(X_{pi'}X_{qi'} - \rho_{pq})^2\right\},
\end{aligned}$$

for $\boldsymbol{\lambda} = (1, 0, 2)$, where above we interpret π as permutation functions on distinct elements, we have the U-statistic representation of degree 4

$$\begin{aligned}
&\frac{\partial^{(1,0,2)} f(1, 1, \rho_{pq})}{(1, 0, 2)!} \bar{S}_{pp} \bar{S}_{pq}^2 \\
&= -n^{-3} \sum_{\tilde{i}, \tilde{j}, \tilde{k}=1}^{n} (X_{p\tilde{i}}^2 - 1)(X_{p\tilde{j}} X_{q\tilde{j}} - \rho_{pq})(X_{p\tilde{k}} X_{q\tilde{k}} - \rho_{pq}) \qquad (4.8) \\
&= -\binom{n}{4}^{-1} \sum_{\substack{\text{unordered} \\ \text{\& distinct} \\ i,j,k,l \\ \text{from } [n]}} h_{2,pq}(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l}).
\end{aligned}$$

Note that (4.8) simply comes from Lemma A.1. What we have done here is that, for each term $(X_{p\tilde{i}}^2 - 1)(X_{p\tilde{j}} X_{q\tilde{j}} - \rho_{pq})(X_{p\tilde{k}} X_{q\tilde{k}} - \rho_{pq})$ in (4.8) with $\tilde{i}, \tilde{j}, \tilde{k}$ not necessarily distinct, we find any 4 *distinct* indices i, j, k, l that contain $\tilde{i}, \tilde{j}, \tilde{k}$ as sets, and arrange the term into one of the three summands of order $O(1)$, $O(n^{-1})$ and $O(n^{-2})$ in (4.7) according to the actual set cardinality $|\{\tilde{i}, \tilde{j}, \tilde{k}\}|$, which can be equal to 1, 2 or 3. Since there are $\binom{n-|\{\tilde{i},\tilde{j},\tilde{k}\}|}{4-|\{\tilde{i},\tilde{j},\tilde{k}\}|}$ choices of distinct i, j, k, l that contain $\{\tilde{i}, \tilde{j}, \tilde{k}\}$ as sets, to account for the duplications we put the factors $\binom{n-3}{1}$, $\binom{n-2}{2}$, $\binom{n-1}{3}$ as denominators for the three summands in the definition (4.7) of the kernel. By a simple symmetry argument if we define the kernel

$$\bar{h}_{2,pq}(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l}) := h_{2,pq}(\bar{\mathbf{X}}_{pq,i}, \bar{\mathbf{X}}_{pq,j}, \bar{\mathbf{X}}_{pq,k}, \bar{\mathbf{X}}_{pq,l}) \qquad (4.9)$$

where $\bar{\mathbf{X}}_{pq,i} := (X_{qi}, X_{pi})'$, we have

$$\begin{aligned}
&\frac{\partial^{(0,1,2)} f(1, 1, \rho_{pq})}{(0, 1, 2)!} \bar{S}_{qq} \bar{S}_{pq}^2 \\
&= -\binom{n}{4}^{-1} \sum_{\substack{\text{unordered} \\ \text{\& distinct} \\ i,j,k,l \\ \text{from } [n]}} \bar{h}_{2,pq}(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l}).
\end{aligned}$$

In the same vein, for $\boldsymbol{\lambda}$ equals $(2, 0, 2)$ or $(0, 2, 2)$ and four *distinct* indices i, j, k, l from $[n]$, we leave it to the reader to check that one can define a symmetric kernel $h_{3,pq}$ of degree 4 as shown in Appendix D [15] such that

$$\frac{\partial^{(2,0,2)} f(1, 1, \rho_{pq})}{(2, 0, 2)!} \bar{S}_{pp}^2 \bar{S}_{pq}^2 = \binom{n}{4}^{-1} h_{3,pq}(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l})$$

and

$$\frac{\partial^{(0,2,2)} f(1, 1, \rho_{pq})}{(0, 2, 2)!} \bar{S}_{qq}^2 \bar{S}_{pq}^2 = \binom{n}{4}^{-1} \bar{h}_{3,pq}(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l}),$$

where

$$\bar{h}_{3,pq}(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l}) := h_{3,pq}(\bar{\mathbf{X}}_{pq,i}, \bar{\mathbf{X}}_{pq,j}, \bar{\mathbf{X}}_{pq,k}, \bar{\mathbf{X}}_{pq,l}). \tag{4.10}$$

Letting $\mathbf{X}_i = (X_{1i}, \dots, X_{mi})'$ denote the entire i-th sample, we have the degree-4 U-statistic representation for II_1:

$$II_1 = \binom{n}{4}^{-1} \sum_{1 \le i < j < k < l \le n} h(\mathbf{X}_i, \mathbf{X}_j, \mathbf{X}_k, \mathbf{X}_l), \tag{4.11}$$

where

$$h(\mathbf{X}_i, \mathbf{X}_j, \mathbf{X}_k, \mathbf{X}_l) := \sum_{1 \le p < q \le m} (h_{1,pq} - h_{2,pq} - \bar{h}_{2,pq} + h_{3,pq} + \bar{h}_{3,pq})(\mathbf{X}_{pq,i}, \mathbf{X}_{pq,j}, \mathbf{X}_{pq,k}, \mathbf{X}_{pq,l}). \tag{4.12}$$

Hence,

$$\mathbb{E}[II_1] = \sum_{1 \le p < q \le m} \mathbb{E}\left[(h_{1,pq} - h_{2,pq} - \bar{h}_{2,pq} + h_{3,pq} + \bar{h}_{3,pq})(\mathbf{X}_{pq,1}, \mathbf{X}_{pq,2}, \mathbf{X}_{pq,3}, \mathbf{X}_{pq,4})\right].$$

The expectation for each of $h_{1,pq}(\cdot), h_{2,pq}(\cdot), h_{3,pq}(\cdot)$ in the preceding display can be computed by taking expectation for each of the product terms appearing in $\{\cdot\}$ in definitions (4.5), (4.7) as well as the counterparts in the definition of $h_{3,pq}$ in Appendix D [15] (Note that quite a few of these expectations are simply zero due to independence of samples). Exploiting symmetry the same can be done for (4.9) and (4.10). In principle, these higher-order normal moments can all be obtained by repeatedly applying Isserlis's theorem (Theorem A.2) laboriously. With symbolic computational softwares such as `mathematica` they can however be much more effortlessly computed. These computations lead to

$$\begin{aligned}\mathbb{E}[II_1] &= \sum_{1 \le p < q \le m} \frac{16 + n^2 + (80 + 8n + n^2)\rho_{pq}^2}{n^3} \\ &= \frac{m(m-1)}{2n} + O(n^{-1})\| \mathbf{R} - \mathbf{I}_{\mathrm{m}} \|_F^2 + O(m^2/n^3) \end{aligned} \tag{4.13}$$

and further details are given in Appendix D [15]. As a direct consequence of Hoeffding [9]'s classical result on the variance of U-statistics, we also have the bound

$$\mathrm{Var}[II_1] \lesssim \sum_{c=1}^{4} n^{-c} \zeta_c, \tag{4.14}$$

where

$$\zeta_c := \mathbb{E}[g_c(\mathbf{X}_1, \dots, \mathbf{X}_c)^2]$$

and the functions $g_c : (\mathbb{R}^m)^c \longrightarrow \mathbb{R}$, $c = 1, \dots, 4$, are defined as

$$g_c(x_1, \dots, x_c) := \mathbb{E}[h(\mathbf{X}_1, \dots, \mathbf{X}_4)|\mathbf{X}_1 = x_1, \dots, \mathbf{X}_c = x_c] - \mathbb{E}[h(\mathbf{X}_1, \dots, \mathbf{X}_4)]. \quad (4.15)$$

Hence, forming estimates of the quantities $\zeta_1, \dots, \zeta_4$ can lead to an estimate of $\text{Var}[II_1]$.

Lemma 4.3 (Bound for the ζ_c's)**.**

$$\zeta_1 \lesssim \frac{\| \mathbf{R} - \mathbf{I}_\text{m} \|_F^4 + m^2(1 + \| \mathbf{R} - \mathbf{I}_\text{m} \|_F^2)}{n^2} + \frac{m^4}{n^4} \quad (4.16)$$

$$\zeta_2 \lesssim \frac{m^3(1 + \| \mathbf{R} - \mathbf{I}_\text{m} \|_F)}{n^2} + \frac{m^4}{n^4} \quad (4.17)$$

$$\zeta_3 \lesssim \| \mathbf{R} - \mathbf{I}_\text{m} \|_F^4 + m^2(1 + \| \mathbf{R} - \mathbf{I}_\text{m} \|_F^2) + \frac{m^4}{n^2} \quad (4.18)$$

$$\zeta_4 \lesssim m^3(\| \mathbf{R} - \mathbf{I}_\text{m} \|_F + 1) + \frac{m^4}{n^2} \quad (4.19)$$

Again, proving these estimates involves repeatedly applying Theorem A.2 with the help of `mathematica` and the details will be deferred to Appendix D [15]. We note that these estimates are by no means sharp, but suffice for our purpose. Putting Lemma 4.3 and (4.14) together, it is a routine task to check that

$$\text{Var}[II_1] \lesssim o\left(\frac{m^{2(1-\gamma)}}{n^2}\right) \sum_{k=0}^{4} \| \mathbf{R} - I\|_F^k$$

for any $0 < \gamma < 1/2$. This, together with (4.4) and (4.13), proved Lemma 4.1.

5. Conclusion

In this paper, we studied the exact power of the Rao's score statistic for testing independence, under the asymptotic regime where both the dimension m and sample size n grow to infinity when the ratio m/n is bounded. A consequence of our main result is that the Rao's score test is minimax rate optimal under this regime, with respect to a signal size $\| \mathbf{R} - \mathbf{I}_\text{m} \|_F$ of order $\sqrt{m/n}$.

While previous related work [5] on the null theory only requires the random variables to have finite moments, our power analysis relied on the normality assumption in different ways. Via applications of the Isserlis' theorem on normal moments (Theorem A.2), all the higher moment quantities involved in the calculations for the terms I and II in Sections 3 and 4 can be controlled in terms of $\|\mathbf{R} - \mathbf{I}_m\|_F$, a second moment quantity in the original variables $X_1, \dots, X_m$ per se. It is thus conceivable that one can replace normality with appropriate higher moment conditions by carefully keeping track of these calculations. The tail bound for III in Lemma 2.5 relies on a maximal inequality applicable to sub-exponential random variables, which is true for the centered sample covariances $\bar{S}_{pq}$ when they are formed with normal data (see Appendix A). When normality cannot be assumed, we expect that one can use more general maximal inequalities such as Chernozhukov et al. [6, Lemma 8] along with their consequential moment conditions. A final caveat for pursuing the non-normal generality is that one should consider the more common definition of the sample covariance in (2.4) when constructing their Pearson correlations. Comparing (2.3) with (2.4), the insertion of sample means will likely complicate the calculations to follow under our current proof strategy.

Acknowledgments

We thank the referees for their valuable comments and suggestions. Qi-Man Shao's research is partially supported by the grant Hong Kong RGC GRF14302515.

Appendix A: Probability tail bound of III and two other lemmas

We will first prove the tail bound for III in Lemma 2.5.

Proof. For $1 \leq p, q \leq m$, by a standard trick [2, p.221], for any $t > 0$, one can show the sub-exponential inequality

$$P\left(|\bar{S}_{pq}| > t\right) \leq 4\exp\left(\frac{-t^2}{n^{-1}2(1+\rho_{pq})(2(1+\rho_{pq})+t)}\right)$$

under our assumptions at the beginning of Section 2. Then by the maximal inequality in van der Vaart and Wellner [20, Lemma 2.2.10] and a union bound, we have for any $0 < c < 1/2$,

$$P(\max_{1\leq p,q\leq m} |\bar{S}_{pq}| > n^{-c}) \lesssim n^{c-1}\log m + n^{c-1/2}\sqrt{\log m}. \tag{A.1}$$

Note that by the definition of III,

$$|III| \leq \max_{1\leq p,q\leq m} |\bar{S}_{pq}|^5 \sum_{1\leq p<q\leq m} \sum_{\boldsymbol{\lambda}:|\boldsymbol{\lambda}|=5} \frac{|\rho_{pq}+k_{pq}\bar{S}_{pq}|^{2-\lambda_1}}{|1+k_{pq}\bar{S}_{pp}|^{1+\lambda_2}|1+k_{pq}\bar{S}_{qq}|^{1+\lambda_3}} \tag{A.2}$$

for $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$. If $\max_{1\leq p,q\leq m} |\bar{S}_{pq}| \leq n^{-c}$, for all $1 \leq p, q \leq m$ it must be true that

$$|\rho_{pq}+k_{pq}\bar{S}_{pq}| \leq 1+n^{-c}, |1+k_{pq}\bar{S}_{pp}| \geq 1-n^{-c} \tag{A.3}$$

since $k_{pq} \in (0,1)$ Combining (A.1), (A.2), (A.3), with probability larger than $1 - C(n^{c-1}\log m + n^{c-1/2}\sqrt{\log m})$

$$|III| \leq Cn^{-5c}\frac{m(m-1)}{2}\frac{(1+n^{-c})^2}{(1-n^{-c})^7} \leq C\frac{m^2}{n^{5c}}$$

for large m, n. □

These are two technical lemmas we mentioned in the main text.

Lemma A.1. *Let f be as defined in (2.2). For any $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{N}^3_{\geq 0}$*

$$\partial^{\boldsymbol{\lambda}} f(u_1, u_2, u_3) = \begin{cases} (-1)^{\lambda_1+\lambda_2}\lambda_1!\lambda_2!\frac{u_3^2}{u_1^{1+\lambda_1}u_2^{1+\lambda_2}} & \text{if } \lambda_3 = 0 \\ 2(-1)^{\lambda_1+\lambda_2}\lambda_1!\lambda_2!\frac{u_3^{2-\lambda_3}}{u_1^{1+\lambda_1}u_2^{1+\lambda_2}} & \text{if } \lambda_3 = 1, 2 \\ 0 & \text{if } \lambda_3 > 2 \end{cases}$$

Theorem A.2 (Isserlis [11])**.** *For any natural number $k \geq 1$, let $(Z_1, \dots, Z_{2k})$ be a mean zero normal vector with covariance matrix $\mathbf{R} = (\rho_{pq})_{1\leq p,q\leq 2k}$. Then*

$$\mathbb{E}[Z_1 \dots Z_{2k}] = \sum \rho_{p_1 p_2} \dots \rho_{p_{2k-1} p_{2k}},$$

where the summation is over all possible $\frac{(2k)!}{2^k k!}$ partitions of the indices $1, \dots, 2k$ into k pairs $(p_1, p_2), \dots, (p_{2k-1}, p_{2k})$.

Supplementary Material

Supplement to Asymptotic Power of Rao's Score Test for Independence in High Dimenions
(). Due to space limitations we defer some of the proofs in this paper to our supplement.

References

[1] T. W. Anderson. *An introduction to multivariate statistical analysis*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition, 2003. ISBN 0-471-36091-0.

[2] Peter J. Bickel and Elizaveta Levina. Regularized estimation of large covariance matrices. *Ann. Statist.*, 36(1):199–227, 2008. ISSN 0090-5364. .

[3] T. Tony Cai and Tiefeng Jiang. Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. *Ann. Statist.*, 39(3):1496–1525, 2011. ISSN 0090-5364. .

[4] T. Tony Cai and Zongming Ma. Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli*, 19(5B):2359–2388, 2013. ISSN 1350-7265. .

[5] Yanchu Chen and Qi-Man Shao. Berry-Esseen inequality for unbounded exchangeable pairs. In *Probability approximations and beyond*, volume 205 of *Lecture Notes in Statist.*, pages 13–30. Springer, New York, 2012. .

[6] Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probab. Theory Related Fields*, 162(1-2):47–70, 2015. ISSN 0178-8051. .

[7] Fang Han and Han Liu. Distribution-free tests of independence with applications to testing more structures. *arXiv preprint arXiv:1410.4179*, 2014.

[8] C. C. Heyde and B. M. Brown. On the departure from normality of a certain class of martingales. *Ann. Math. Statist.*, 41:2161–2165, 1970. ISSN 0003-4851.

[9] Wassily Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, 19:293–325, 1948. ISSN 0003-4851.

[10] Harold Hotelling. New light on the correlation coefficient and its transforms. *J. Roy. Statist. Soc. Ser. B.*, 15:193–225; discussion, 225–232, 1953. ISSN 0035-9246.

[11] Leon Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1/2): 134–139, 1918.

[12] Tiefeng Jiang. The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.*, 14(2):865–880, 2004. ISSN 1050-5164. .

[13] Tiefeng Jiang and Yongcheng Qi. Likelihood ratio tests for high-dimensional normal distributions. *Scand. J. Stat.*, 42(4):988–1009, 2015. ISSN 0303-6898.

[14] Dennis Leung and Mathias Drton. Testing independence in high dimensions with sums of squares of rank correlations. *arXiv preprint arXiv:1501.01732*, 2015.

[15] Dennis Leung and Qiman Shao. Supplement to "Asymptotic power of Rao's score test for independence in high dimensions". 2016. .

[16] Deli Li, Wei-Dong Liu, and Andrew Rosalsky. Necessary and sufficient conditions for the asymptotic distribution of the largest entry of a sample correlation matrix. *Probab. Theory Related Fields*, 148(1-2):5–35, 2010. ISSN 0178-8051. .

[17] Deli Li, Yongcheng Qi, and Andrew Rosalsky. On Jiang's asymptotic distribution of the largest entry of a sample correlation matrix. *J. Multivariate Anal.*, 111:256–270, 2012. ISSN 0047-259X. .

[18] Wei-Dong Liu, Zhengyan Lin, and Qi-Man Shao. The asymptotic distribution and Berry-Esseen bound of a new test for independence in high dimension with an application to stochastic optimization. *Ann. Appl. Probab.*, 18(6):2337–2366, 2008. ISSN 1050-5164. .

[19] James R. Schott. Testing for complete independence in high dimensions. *Biometrika*, 92(4):951–956, 2005. ISSN 0006-3444. .

[20] Aad W. van der Vaart and Jon A. Wellner. *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York, 1996. ISBN 0-387-94640-3. . With applications to statistics.

[21] Wang Zhou. Asymptotic distribution of the largest off-diagonal entry of correlation matrices. *Trans. Amer. Math. Soc.*, 359(11):5345–5363, 2007. ISSN 0002-9947. .