Biometrika (2014), **101**, 3, *pp.* 748–754
Printed in Great Britain
doi: 10.1093/biomet/asu023
Advance Access publication 14 July 2014

Inference on multiple correlation coefficients with moderately high dimensional data

BY SHURONG ZHENG

School of Mathematics & Statistics, Northeast Normal University, Renmin Street No. 5268, Changchun 130024, Jilin Province, China

zhengsr@nenu.edu.cn

DANDAN JIANG

School of Mathematics, Jilin University, Qianjing Street No. 2699, Changchun 130012, Jilin Province, China

jiangdandan@jlu.edu.cn

ZHIDONG BAI

School of Mathematics & Statistics, Northeast Normal University, Renmin Street 5268, Changchun 130024, Jilin Province, China

baizd@nenu.edu.cn

AND XUMING HE

Department of Statistics, University of Michigan, 1085 South University, Ann Arbor, Michigan 48109-1107, U.S.A.

xmhe@umich.edu

SUMMARY

When the multiple correlation coefficient is used to measure how strongly a given variable can be linearly associated with a set of covariates, it suffers from an upward bias that cannot be ignored in the presence of a moderately high dimensional covariate. Under an independent component model, we derive an asymptotic approximation to the distribution of the squared multiple correlation coefficient that depends on a simple correction factor. We show that this approximation enables us to construct reliable confidence intervals on the population coefficient even when the ratio of the dimension to the sample size is close to unity and the variables are non-Gaussian.

Some key words: Independent component model; Multiple correlation; Testing.

1. INTRODUCTION

The multiple correlation coefficient is important in multivariate analysis and measures how strongly a given variable can be correlated with a linear combination of other variables. In this article, we use ρ for the population version of the multiple correlation between a one-dimensional random variable y and a $(p-1)$-dimensional random vector x, and R for the sample multiple correlation coefficient from a sample of size n. In linear regression of y on x, the coefficient of determination R^2 is routinely used to measure the strength of the linear relationship.

Multiple correlation has a long history and has been widely used. The distribution of R^2 was first investigated by Fisher (1928) under the assumption of multivariate normality. Various methods of deriving its distribution can be found in Wilks (1932), Garding (1941), Moran (1950), Williams (1978), Gurland & Asiribo (1991), Johnson et al. (1995), Nandi & Choudhury (2005), and Anderson (2003). It is well recognized that R^2 is a biased estimate of ρ^2, and an adjusted estimate of ρ^2 can be found in Anderson (2003, p. 153). Recent uses of multiple correlation can be found in Pena & Rodruguez (2002) for constructing a portmanteau test of fit and in Dutilleul et al. (2008) for assessing correlation in spatial processes.

Although the exact density of R^2 under normality is known, it is given by an infinite series when $\rho^2 > 0$ and thus is inconvenient to use. More importantly, the convergence of the series is slow, especially when n is large. For non-Gaussian variables, the density of R^2 is unavailable. Efforts have been made to allow inference on multiple correlation without normality. For example, Muirhead (1982) obtains the exact null distribution of R^2 under spherical symmetry, Ali & Nagar (2002) give the null distribution of multiple correlation for normal mixture models, and Ogasawara (2006) obtains approximations to the R^2 distribution by using the Edgeworth expansion. Existing work, however, does not provide effective inferences on multiple correlation when p is moderately large.

In this article, we provide an approximation for inference on multiple correlation coefficients under an independent component model. Our approximation is shown to work well for a wide range of dimensions $p < n$. The analysis of multiple correlation in high or ultra-high dimensions with $p \geqslant n$ is not attempted.

2. A quick review of multiple correlation

Let $Y_1, \ldots, Y_n$ be a random sample from a p-dimensional population with covariance matrix Σ. Let $\bar{Y}$ and $\hat{\Sigma}$ be the usual sample mean and the sample covariance matrix. We consider the case $2 \leqslant p < n$. The covariance matrices are partitioned into

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{(1)}^{\mathrm{T}} \\ \sigma_{(1)} & \Sigma_{22} \end{pmatrix}, \quad \hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_{11} & a_{(1)}^{\mathrm{T}} \\ a_{(1)} & A \end{pmatrix},$$

where σ_{11} and $\hat{\sigma}_{11}$ are scalars. We further assume that Σ_{22} and A are nonsingular. Then the squared multiple correlation coefficients between the first component of Y and the other components, both for the population and for the sample, are

$$\rho^2 = \frac{\sigma_{(1)}^{\mathrm{T}} \Sigma_{22}^{-1} \sigma_{(1)}}{\sigma_{11}}, \quad R^2 = \frac{a_{(1)}^{\mathrm{T}} A^{-1} a_{(1)}}{\hat{\sigma}_{11}}.$$

If $Y_i \sim N(\mu, \Sigma)$, then for the hypothesis testing problem

$$H_0 : \rho = 0, \quad H_1 : \rho \neq 0,$$

the rejection region at significance level α is

$$\frac{R^2}{1 - R^2} \frac{n - p}{p - 1} > F_{p-1,n-p}(1 - \alpha)$$

where $F_{p-1,n-p}(1 - \alpha)$ is the $1 - \alpha$ quantile of the F distribution with degrees of freedom $p - 1$ and $n - p$. Under the alternative hypothesis H_1, the density $g(t; \rho^2)$ of R^2 (Anderson, 2003) is

$$\begin{aligned} g(t; \rho^2) &= \frac{(1 - t)^{(n-p-2)/2} (1 - \rho^2)^{(n-1)/2}}{\Gamma\{(n - p)/2\} \Gamma\{(n - 1)/2\}} \sum_{\mu=0}^{+\infty} \frac{\rho^{2\mu}\, t^{(p-1)/2+\mu-1} \Gamma^2\{(n - 1)/2 + \mu\}}{\mu! \Gamma\{(p - 1)/2 + \mu\}} \\ &= \frac{(1 - \rho^2)^{(n-1)/2}\, t^{(p-3)/2} (1 - t)^{(n-p)/2}}{B\{(n - p)/2, (p - 1)/2\}} F\{(n - 1)/2, (n - 1)/2, (p - 1)/2, t\rho^2\} \end{aligned} \quad (1)$$

where $F(\cdot,\cdot,\cdot,\cdot)$ is the hypergeometric function, $B(\cdot,\cdot)$ is the beta function and $\Gamma(\cdot)$ is the Gamma function. So a confidence interval for ρ^2 at the level of $(1-\alpha)100\%$ has lower and upper limits

$$L_A = \min_{\rho^2}\{\alpha/2 \leqslant G(R^2;\rho^2) \leqslant 1-\alpha/2\},$$
$$U_A = \max_{\rho^2}\{\alpha/2 \leqslant G(R^2;\rho^2) \leqslant 1-\alpha/2\}, \tag{2}$$

where $G(x;\rho^2) = \int_0^x g(t;\rho^2)\,\mathrm{d}t$ is the distribution function of R^2. Even under the Gaussian model, R^2 has an upward bias for ρ^2 when the true correlation is small.

THEOREM 1. *Under the normality assumption, $E(R^r) > \rho^r$ for any $r \geqslant 2$, when $(p-1)/(n-1) \geqslant \rho^2$ and $n > p$.*

Motivated by Gaussian models, Anderson (2003) (p. 153) considered an adjusted multiple correlation. By truncating the adjusted multiple correlation at zero, we use

$$R^{*2} = \max\left\{R^2 - \frac{p-1}{n-p}(1-R^2),\quad 0\right\}$$

as the adjusted multiple correlation in this paper. We shall show that R^{*2} is consistent under a broader class of distributions and even when p grows proportionally to n. More importantly, we shall obtain a limiting distribution that enables us to construct approximate confidence intervals for ρ^2 for a wide range of p and n with $n > p$.

3. A NEW APPROXIMATION FRAMEWORK

We now assume the following independent component model for Y, that is, there exists a representation

$$Y_i = \begin{pmatrix} d_1 \\ D_2 \end{pmatrix} X_i \quad (i = 1,\ldots,n), \tag{3}$$

where d_1 is a $1 \times m$ row vector, D_2 is a $(p-1)\times m$ matrix of rank $p-1$ with $m \geqslant p$, and X_i is a random sample of m-random vectors with independently and identically distributed components.

Because the multiple correlation coefficient is invariant under location changes, any positive scale change on the first component of Y_i and any nonsingular affine transformation on the rest of the components, we assume without loss of generality that $E(X_i) = 0$, $\mathrm{cov}(X_i) = I_m$, $\|d_1\| = 1$, $D_2 D_2^{\mathrm{T}} = I_{p-1}$, and $d_1 D_2^{\mathrm{T}} = (\rho, 0, \ldots, 0)$. We study the asymptotic properties of R^2 under the setting $n \to \infty$ and $q_n = p/n \to q \in [0,1)$. To begin, let

$$\sigma^2(t) = 2\{q + (1-q)t\}^2 - 2\{-2(1-q)t^2 + 4(1-q)t + 2q\}\{q + (1-q)t - 1/2\}. \tag{4}$$

We now state our main results, followed by a discussion on the assumptions.

THEOREM 2. *Under the independent component model (3), and if X_i has finite fourth moment and the L_∞-norm $\|d_1\|_\infty = o(1)$, then $R^{*2} \to \rho^2$ almost surely, and*

$$n^{1/2}\{R^2 - q_n - (1-q_n)\rho^2\} \to N\{0, \sigma^2(\rho^2)\}$$

in distribution as $n \to \infty$ and $q_n = p/n \to q < 1$. As a consequence, we have, if $\rho^2 > 0$, then

$$n^{1/2}(R^{*2} - \rho^2) \to N\{0, \sigma^2(\rho^2)/(1-q)^2\}, \tag{5}$$
$$n^{1/2}\{g(R^{*2}) - g(\rho^2)\} \to N(0,1) \tag{6}$$

in distribution, where $g(x) = \int_0^x (1-q)/\sigma(t)\,\mathrm{d}t$ for $x \geqslant 0$. If $\rho^2 = 0$, we have $n^{1/2}(R^2 - q_n) \to N\{0, 2q(1-q)\}$, and

$$(2n)^{1/2}\{\arccos(R) - \arccos(q_n^{1/2})\} \to N(0,1)\,. \tag{7}$$

The variance $\sigma^2(\rho^2)$ in (5) can be estimated by (4) with $t = R^{*2}$. Since $g(x)$ is monotone and can be computed numerically, Theorem 2 allows the construction of two asymptotic confidence intervals for ρ^2, one based on (5) and the other based on the variance-stabilizing transformation (6). For example, an $(1-\alpha)100\%$ asymptotic confidence interval for ρ^2 based on (6) is (L, U), where

$$L = \min\left\{0 \leqslant x \leqslant 1 : g(x) \geqslant g(R^{*2}) - n^{-1/2} z_{\alpha/2}\right\}$$
$$U = \max\left\{0 \leqslant x \leqslant 1 : g(x) \leqslant g(R^{*2}) + n^{-1/2} z_{\alpha/2}\right\}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ quantile of the standard normal. For testing the null hypothesis that $\rho^2 = 0$, the asymptotic distribution of (7) may be used.

The independent component model generalizes the multivariate Gaussian model, but the representation (3) is not required to be unique. The latent variable X_i could be normal, but need not be. The assumption that $\|d_1\|_\infty = o(1)$ in Theorem 2 requires that the first component of Y_i is not dominated by a small number of components in X_i as p increases, but this condition can be removed if each component of X_i has fourth moment equal to 3 as in the Gaussian case.

Our main result is obtained under the assumption that $q_n = p/n \to q < 1$. A natural question is why we are interested in this limit, but not others, such as $p \propto n^r$ for some positive r. In a given application where both n and p are fixed, we cannot distinguish one value of r from another. What matters is which limiting distribution provides a good approximation. We find that the framework of $p \propto n$ has the advantage of an asymptotic normal approximation that works well for a wide range of (n, p) values with $n > p$. Similar observations have been made in Bai & Saranadasa (1996), Bai et al. (2009), Chen & Qin (2010), and Chen et al. (2010), and we shall confirm this through simulation in § 4.

4. Empirical evidence

In this section, we evaluate the performance of the asymptotic confidence intervals for ρ^2 based on (5) and (6), and compare them to confidence intervals based on the Edgeworth expansion method of Ogasawara (2006), the classical method of Anderson (2003) in (2), and ABC bootstrap confidence intervals as described in Davison & Hinkley (1997 § 5).

Bootstrap methods are ill-suited for datasets with p/n close to unity, because many bootstrap samples will not have enough distinct observations, and when the number of distinct observations is p or less, the value of R^2 cannot be computed. For example, if $(n, p) = (25, 20)$, only 1 out of 100 bootstrapped samples is expected to have more than p distinct observations. In this paper, we adopt the strategy of adding a small random perturbation to each observation in the bootstrapped samples before the R^{*2} are calculated.

The data are generated under the independent component model (3), with the following specifications, with $k = p(1+\rho)/2$ for the given values of p and ρ considered in our study. In our study, k is an integer between 2 and $p-1$.

(i) The vector d_1 has all its elements equal to $p^{-1/2}$.
(ii) The first row of D_2 has its first k elements equal to $p^{-1/2}$ and the rest equal to $-p^{-1/2}$.
(iii) For $l = 2, \ldots, k$, the lth row of D_2 has the first $l-1$ elements equal to $l^{-1/2}(l-1)^{-1/2}$, the lth element equal to $-l^{-1/2}(l-1)^{1/2}$ and the rest equal to 0.
(iv) For $l = k+1, \ldots, p$, the lth row of D_2 has the first k elements equal to 0, the next $l-k-1$ elements equal to $(l-k+1)^{-1/2}(l-k)^{-1/2}$, the lth element equal to $-(l-k+1)^{-1/2}(l-k)^{1/2}$, and the rest equal to 0.

We consider two scenarios of $q_n = p/n = 0{\cdot}2$ and $0{\cdot}8$, three values of $\rho = 0{\cdot}2, 0{\cdot}6, 0{\cdot}8$, and three values of moderate to high dimensions $p = 20, 30, 60$ at $q = 0{\cdot}2$ and $p = 20, 60, 200$ at $q = 0{\cdot}8$. With the above construction of d_1 and D_2, the desired ρ values are achieved. Table 1 reports the average coverage probability for nominal 95% intervals and interval length of ρ^2 obtained from 10 000 Monte Carlo samples

Table 1. *Coverage (%) and average length (in parentheses) of 95% confidence intervals for ρ^2 when the X_{ij} are uniformly distributed*

p		$\rho = 0{\cdot}2$	$\rho = 0{\cdot}6$	$\rho = 0{\cdot}8$
		$q = 0{\cdot}2$		
20	M1	95·9 (0·19)	94·9 (0·33)	94·9 (0·26)
	M2	98·7 (0·21)	95·3 (0·34)	96·6 (0·25)
	M3	71·6 (0·39)	90·9 (0·39)	93·7 (0·24)
	A1	90·5 (0·06)	0·0 (0·07)	0·0 (0·07)
	B1	43·6 (0·05)	67·8 (0·30)	76·9 (0·24)
30	M1	95·2 (0·17)	94·7 (0·27)	95·2 (0·21)
	M2	98·6 (0·18)	95·1 (0·28)	96·3 (0·20)
	M3	47·7 (0·34)	85·1 (0·34)	89·0 (0·22)
	A1	69·5 (0·04)	0·0 (0·05)	0·0 (0·05)
	B1	28·1 (0·02)	58·5 (0·26)	67·7 (0·20)
60	M1	95·4 (0·13)	95·3 (0·20)	95·0 (0·15)
	M2	98·4 (0·13)	94·8 (0·20)	95·1 (0·14)
	M3	5·6 (0·26)	62·8 (0·27)	78·4 (0·18)
	A1	0·0 (0·02)	0·0 (0·02)	0·0 (0·02)
	B1	7·4 (0·01)	32·8 (0·19)	45·6 (0·15)
		$q = 0{\cdot}8$		
p		$\rho = 0{\cdot}2$	$\rho = 0{\cdot}6$	$\rho = 0{\cdot}8$
20	M1	93·2 (0·52)	92·4 (0·68)	93·7 (0·83)
	M2	95·2 (0·76)	94·3 (0·78)	94·4 (0·75)
	M3	0·7 (0·62)	3·1 (0·44)	4·5 (0·26)
	A1	95·6 (0·20)	0·0 (0·23)	0·0 (0·25)
	B1	25·6 (0·12)	26·2 (0·18)	26·6 (0·20)
60	M1	95·1 (0·47)	94·6 (0·60)	94·6 (0·69)
	M2	95·7 (0·56)	95·7 (0·63)	94·4 (0·55)
	M3	0·0 (0·45)	0·0 (0·32)	0·0 (0·18)
	A1	90·5 (0·07)	0·0 (0·09)	0·0 (0·09)
	B1	9·7 (0·02)	10·9 (0·07)	13·2 (0·10)
200	M1	95·0 (0·29)	95·0 (0·46)	94·6 (0·33)
	M2	96·5 (0·37)	95·0 (0·46)	94·5 (0·30)
	M3	0·0 (0·30)	0·0 (0·22)	0·0 (0·13)
	A1	0·0 (0·03)	0·0 (0·03)	0·0 (0·03)
	B1	1·4 (0·01)	1·0 (0·02)	1·5 (0·04)

M1 and M2, asymptotic confidence intervals based on Theorem 2; M3, Edgeworth expansion method of Ogasawara (2006); A1, the classical method of Anderson (2003) in (2); B1, ABC bootstrap method.

when each component of X_i is uniformly distributed. A similar table is given in the Supplementary Material when X_i is normally distributed. It is clear from the tables that our proposed methods attain decent coverage probabilities in all cases. The competing methods fail badly for different reasons. More specifically, the method based on Edgeworth expansion of Ogasawara (2006) uses R^2 as the estimate of ρ^2 with an upward bias approximately $q(1-\rho^2)$. At $q = 0{\cdot}9$ and $\rho^2 = 0{\cdot}1$, the bias is more than 0·8, which leads to nearly zero coverage. The ABC method is biased because the bootstrap distribution of R^{*2} is unduly skewed towards the value 1 as we explained earlier, leading very low coverage probabilities.

In the Supplementary Material, we include the power curve for testing the null hypothesis that $\rho^2 = 0$ versus $\rho^2 > 0$ based on the distribution (7) at $p = 60$ and $q_n = 0{\cdot}2$, 0·6 and 0·8. The significance levels of the tests are close to the nominal level of 0·05, but the power of the test decreases quickly with q_n, reflecting the challenges of statistical inference in high dimensions. In the Supplementary Material, we also provide additional simulation results for two other models with different choices of d_1 and D_2.

Finally, we demonstrate the merit of the proposed method on the diffuse large-B-cell lymphoma microarray data of Rosenwald et al. (2002). The dataset contains the gene expression measurements of over 7399 genes as well as the survival times of 240 patients. We first select 120 genes that have the highest marginal correlation with the survival time, and then estimate the multiple correlation between survival time and the expressions of those genes. In this case, we have $q_n \approx 0{\cdot}5$, and the 95% confidence intervals are (0·15, 0·46) and (0·16, 0·47) from our proposed methods. The ABC bootstrap method would yield a confidence interval (0·12, 0·16), which appears narrow but biased. When we performed the same analysis by taking 120 randomly selected genes, the 95% confidence interval based on our proposed method (5) was [0, 0·29), with $R^2 = 0{\cdot}57$ and $R^{*2} = 0{\cdot}14$. The bootstrap distribution of R^{*2} has nearly all of its probability mass above 0·40, indicating again the problem of the bootstrap in analysing data of high dimensions. Our proposed method is able to produce a trustworthy confidence interval with minimal computational cost.

When the dimension of variables p is moderately large relative to the sample size n, classical methods of inference on the multiple correlation coefficients fail. In this paper, we propose a simple framework for approximating the distribution of the sample multiple correlation. The asymptotic approximation is easy to use but remains effective for a wide range of scenarios in terms of (n, p) as long as p/n is bounded away from 1. The problem of even higher dimensions, $p \geqslant n$, is not considered in this paper.

Acknowledgement

The authors thank the reviewers for their helpful comments and suggestions that led to an improvement of this article. This research was partially supported by the National Science Foundation, U.S.A., the National Natural Science Foundation of China, the Program for New Century Excellent Talents in University and the Program for Changjiang Scholars and Innovative Research Teams in University.

Supplementary material

Supplementary material available at *Biometrika* online includes the proof of Theorem 2 and additional simulation results, including the distributions of the p-values from various tests on ρ^2.

Appendix

Proof of Theorem 1

Our first step is to show

$$E(R^2) > \rho^2. \tag{A1}$$

Following (1), we have

$$E(R^2) = \frac{(1-\rho^2)^{(n-1)/2}}{\Gamma\{(n-1)/2\}} \sum_{\mu=0}^{\infty} \frac{\rho^{2\mu}\Gamma\{(n-1)/2+\mu\}}{\mu!} \frac{p-1+2\mu}{n-1+2\mu}.$$

Moreover, we can write

$$\rho^2 = \frac{(1-\rho^2)^{(n-1)/2}}{\Gamma\{(n-1)/2\}} \sum_{\mu=0}^{\infty} \frac{\rho^{2\mu}\Gamma\{(n-1)/2+\mu\}}{\mu!} \rho^2.$$

If $(p-1)/(n-1) \geqslant \rho^2$, then (A1) is obviously true because $(p-1+2\mu)/(n-1+2\mu) > (p-1)/(n-1) \geqslant \rho^2$ for all $\mu \geqslant 1$. By Jensen's inequality we then have $E(R^r) > \rho^r$ for any $r \geqslant 2$.

References

Ali, H. & Nagar, D. K. (2002). Null distribution of multiple correlation coefficient under mixture normal model. *Int. J. Math. Math. Sci.* **30**, 249–55.

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis.* New York: Wiley.

Bai, Z. D. & Saranadasa, H. (1996) Effect of high dimension comparison of significance tests for a high dimensional two sample problem. *Statist. Sinica* **6**, 311–29.

Bai, Z. D., Jiang, D. D., Yao, J. F. & Zheng, S. R. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann. Statist.* **37**, 3822–40.

Chen, S. X. & Qin, Y. L. (2010). A two sample test for high dimensional data with application to gene-set testing. *Ann. Statist.* **38**, 808–35.

Chen, S. X., Zhang, L. X. & Zhong, P. S. (2010) Testing high-dimensional covariance matrices. *J. Am. Statist. Assoc.* **105**, 810–9.

Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap Methods and Their Application*. New York: Cambridge University Press.

Dutilleul, P., Pelletier, B. & Alpargu, G. (2008). Modified F tests for assessing the multiple correlation between one spatial process and several others. *J. Statist. Plan. Infer.* **138**, 1402–15.

Fisher, R. A. (1928). The general sampling distribution of the multiple correlation coefficient. *Proc. R. Soc. Lond.* **121**, 654–73.

Garding, L. (1941). The distributions of the first and second order moments, the partial correlation coefficients and the multiple correlation coefficient in samples from a normal multivariate population. *Skand. Aktuar.* **24**, 185–202.

Gurland, J. & Asiribo, O. (1991). On the distribution of the multiple correlation coefficient and the kinked chi-square random variable. *Statist. Sinica* **1**, 493–502.

Johnson, N. L., Kotz, S. & Balakrishnan, N. (1995). *Continuous Univariate Distributions*. New York: Wiley.

Moran, P. A. P. (1950). The distribution of the multiple correlation coefficient. *Proc. Camb. Phil. Soc.* **46**, 521–2.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. New York: Wiley.

Nandi, S. B. & Choudhury, S. (2005) Series representation of non-null distribution of the square of sample multiple correlation coefficient by use of the Mellin integral transform. *Commun. Statist.* A **34**, 679–85.

Ogasawara, H. (2006). Asymptotic expansion and conditional robustness for the sample multiple correlation coefficient under nonnormality. *Commun. Statist.* B **35**, 177–99.

Pena, D. & Rodruguez, J. (2002). A powerful portmanteau test of lack of fit for time series. *J. Am. Statist. Assoc.* **97**, 601–10.

Rosenwald, A., Wright, G., Chan, W. C., Connors, J. M., Campo, E., Fisher, R. I., Gascoyne, R. D., Muller-Hermelink, H. K., Smeland, E. B., Giltnane, J. M., Hurt, E. M., Zhao, H., Averett, L., Yang, L., Wilson, W. H., Jaffe, E. S., Simon, R., Klausner, R. D., Powell, J., Duffey, P. L. et al. (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *New Engl. J. Med.* **346**, 1937–47.

Wilks, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika* **24**, 471–94.

Williams, E. J. (1978). A simple derivation of the distribution of the multiple correlation coefficient. *Commun. Statist.* A **7**, 1413–20.

[*Received January* 2013. *Revised April* 2014]

Supplementary material to Inference on multiple correlation coefficients with moderately high dimensional data by Zheng et al.

This supplementary material is organized as follows. We start with assumptions, notations and a truncation on the variables. We then give the proof for the almost sure convergence and the asymptotic normality result of Theorem 2 in Section 1 and Section 2, respectively. Some technical lemmas needed for the proof of Theorem 2 are verified in Section 3. Last, additional simulation results are given in Section 4.

Assumptions: In Theorem 2, we assume that $X_j = (X_{1j}, \ldots, X_{mj})'$ has the finite 4th moment, $p/n = q_n \to q \in [0, 1)$ and the L_∞-norm $\|d_1\|_\infty = o(1)$, which can be removed if each component of X_i has the fourth moment equal to 3. Moreover, X_{ij} has mean 0 and variance 1 without loss of generality.

Notation: Let $\mathcal{B}_n = \{\lambda_{\min}(A_0) \geq a\}$, $\mathcal{B}_{ni} = \{\lambda_{\min}(A_i) \geq a\}$, and $\mathcal{B}_{nij} = \{\lambda_{\min}(A_{ij}) \geq a\}$ where a is a positive constant less than $(1-\sqrt{q})^2$, $\lambda_{\min}(\cdot)$ is the minimum eigenvalue, $A_0 = \sum_{k=1}^n \gamma_k\gamma_k'$, $A_i = A_0 - \gamma_i\gamma_i'$ and $A_{ij} = A_0 - \gamma_i\gamma_i' - \gamma_j\gamma_j'$ with $\gamma_j = n^{-1/2}D_2X_j$ for $i, j = 1, \ldots, n, i \neq j$. Then we have $\mathcal{B}_n \supseteq \mathcal{B}_{ni} \supseteq \mathcal{B}_{nij}$. For brevity, in our proofs, we shall use notations $\mathcal{B}_n, \mathcal{B}_{ni}, \mathcal{B}_{nij}$ and their complements for the random events and also for their indicator functions $I(\mathcal{B}_n), I(\mathcal{B}_{ni}), I(\mathcal{B}_{nij}), I(\mathcal{B}_n^c), I(\mathcal{B}_{ni}^c)$ and $I(\mathcal{B}_{nij}^c)$. Moreover, let $s_i = n^{-1/2}d_1X_i$ for $i = 1, \ldots, n$. Thus we have $\bar{y}_1 = n^{-1}\sum_{i=1}^n d_1X_i = n^{-1/2}\sum_{i=1}^n s_i$, $\bar{y}_2 = n^{-1}\sum_{i=1}^n D_2X_i = n^{-1/2}\sum_{i=1}^n \gamma_i$, $\hat{\sigma}_{11} = \sum_{j=1}^n |s_j|^2 - \bar{y}_1^2$, $C = \sum_{k=1}^n s_k\gamma_k$ and $a_{(1)} = C - \bar{y}_1\bar{y}_2$ and $A = A_0 - \bar{y}_2\bar{y}_2'$. Also note that our model specifies $Y_i = \begin{pmatrix} d_1 \\ D_2 \end{pmatrix} X_i$.

Truncation: Let $\tilde{R}$ be the sample multiple correlation coefficient obtained by

$$X_{ij}I_{(|X_{ij}| \leq \min\{d_i^{-1/2}\eta_n n^{1/4}, d_{1i}^{-1/2}\eta_n n^{1/4}, \eta_n n^{1/2}\})}, i = 1, \ldots, m, j = 1, \ldots, n$$

where $d_1 = (d_1, \ldots, d_m)$, $D_2 = (d_{ij})$ and $\{\eta_n\}$ is a sequence satisfying the truncation condition

$$\eta_n^{-4} n^{-1} \sum_j E\{|X_{1j}|^4 I(|X_{1j}| > \eta_n n^{1/4})\} \to 0. \tag{0.1}$$

As shown in Bai and Silverstein (2010, p. 95), we note that $\{\eta_n\}$ satisfying (0.1) exists under the finite 4th moment assumption on X_{ij}. Because $\sum_i d_i^2 = 1$ and $\sum_i d_{1i}^2 = 1$, we have

$$\begin{aligned} P(R \neq \tilde{R}) &\leq \sum_{ij} \Big(P(|X_{ij}| > \eta_n d_i^{-1/2} n^{1/4}) + P(|X_{ij}| > \eta_n d_{1i}^{-1/2} n^{1/4}) + P(|X_{ij}| > \eta_n n^{1/2}) \Big) \\ &\leq \sum_{ij} \left[d_i^2 \eta_n^{-4} n^{-1} E\{|X_{ij}|^4 I(|X_{ij}| > \eta_n n^{1/4})\} + d_{1i}^2 \eta_n^{-4} n^{-1} E\{|X_{ij}|^4 I(|X_{ij}| > \eta_n n^{1/4})\} \right] \\ &\quad + \eta_n^{-4} n^{-2} \sum_{ij} E\{|X_{ij}|^4 I(|X_{ij}| > \eta_n n^{1/2})\} \to 0. \end{aligned}$$

Therefore we shall give the proof of Theorem 2 assuming that X_{ij} has already been truncated. With this in mind, we have

$$\mathrm{P}(\mathcal{B}_n^c) \leq \mathrm{P}(\mathcal{B}_{ni}^c) \leq \mathrm{P}(\mathcal{B}_{nij}^c) = o(n^{-t}) \tag{0.2}$$

for any fixed $t > 0$ (see Bai and Silverstein (2010)).

1 Almost sure convergence

By the law of large numbers, we have $\hat{\sigma}_{11} \to 1,\ a.s.$ Then one needs to show only that $R^{*2} \to \rho^2$ a.s., that is, $a'_{(1)} A^{-1} a_{(1)} - \{q_n + (1-q_n)\rho^2\} \to 0\ a.s.$ Because $a_{(1)} = C - \bar{y}_1 \bar{y}_2$ and $A = A_0 - \bar{y}_2 \bar{y}_2'$, we first consider the convergence of $C' A_0{}^{-1} C$. We have

$$C' A_0{}^{-1} C = \sum\nolimits_{k=1}^{n} s_k^2 \gamma_k' A_k^{-1} \gamma_k \beta_k + \sum\nolimits_{k_1 \neq k_2} s_{k_1} s_{k_2} \gamma_{k_1}' A_{k_1 k_2}^{-1} \gamma_{k_2} \beta_{k_1} \beta_{k_2(k_1)} \tag{1.1}$$

where $\beta_{k_1} = \frac{1}{1+\gamma_{k_1}' A_{k_1}^{-1} \gamma_{k_1}}$, $\beta_{k_2(k_1)} = \frac{1}{1+\gamma_{k_2}' A_{k_1 k_2}^{-1} \gamma_{k_2}}$. By Lemmas 3.4–3.7 and 3.3, we have

$$\begin{aligned} &\sum_{k_1 \neq k_2} s_{k_1} s_{k_2} \gamma_{k_1}' A_0^{-1} \gamma_{k_2} \mathcal{B}_n \\ = &\sum_{k_1 \neq k_2} s_{k_1} s_{k_2} \gamma_{k_1}' A_0^{-1} \gamma_{k_2} (\mathcal{B}_n \Delta \mathcal{B}_{nk_1k_2}) + \sum_{k_1 \neq k_2} s_{k_1} s_{k_2} \gamma_{k_1}' A_{k_1 k_2}^{-1} \gamma_{k_2} \beta_{k_1} \beta_{k_2(k_1)} \mathcal{B}_{nk_1k_2} \\ = &(1-q)^2 \sum_{k_1 \neq k_2} s_{k_1} s_{k_2} \gamma_{k_1}' A_{k_1 k_2}^{-1} \gamma_{k_2} \mathcal{B}_{nk_1k_2} + o_{a.s.}(1) \end{aligned} \tag{1.2}$$

where $A \Delta B = AB^c + BA^c$ denotes the symmetric difference of sets A and B, and we have used the fact that

$$\mathrm{P}\left\{ \sum_{k_1 \neq k_2} s_{k_1} s_{k_2} \gamma_{k_1}' A_0^{-1} \gamma_{k_2} (\mathcal{B}_n \Delta \mathcal{B}_{nk_1k_2}) \neq 0 \right\} \leq \mathrm{P}(\mathcal{B}_n^c) + \sum_{k_1 \neq k_2} \mathrm{P}(\mathcal{B}_{nk_1k_2}^c) = o\big\{ n^{-t}(1+n^2) \big\}$$

which is summable for sufficiently large t. Note that

$$s_{k_1}s_{k_2}\gamma'_{k_1}A^{-1}_{k_1k_2}\gamma_{k_2} = n^{-1}x'_{k_1}D'_2A^{-1}_{k_1k_2}s_{k_2}\gamma_{k_2}d_1x_{k_1},$$

$$\mathrm{tr}(D'_2A^{-1}_{k_1k_2}s_{k_2}\gamma_{k_2}d_1) = n^{-1}x'_{k_2}d'_1\gamma'A^{-1}_{k_1k_2}D_2x_{k_2},$$

where $\gamma = D_2d'_1 = (\rho, 0, \ldots, 0)'$. Employing Lemmas 3.4–3.7 again as well as Lemma 6 of Bai, Liu and Wong (2011), we obtain

$$\begin{array}{cl} & \sum_{k_1\neq k_2} s_{k_1}s_{k_2}\gamma'_{k_1}A_0^{-1}\gamma_{k_2}\mathcal{B}_n \\ = & \frac{(1-q)^2}{n}\sum_{k_1\neq k_2} s_{k_2}\gamma'A^{-1}_{k_1k_2}\gamma_{k_2}\mathcal{B}_{nk_1k_2} + o_{a.s.}(1) \\ = & \frac{(1-q)^2}{n^2}\sum_{k_1\neq k_2} \gamma'A^{-1}_{k_1k_2}\gamma\mathcal{B}_{nk_1k_2} + o_{a.s.}(1) \\ = & (1-q)^2\gamma'A_0^{-1}\gamma\mathcal{B}_n + o_{a.s.}(1) \to (1-q)\rho^2,\ a.s. \end{array} \tag{1.3}$$

Similarly, we have have

$$\begin{array}{rcl} \sum_{k=1}^n s_k^2\gamma'_kA_0^{-1}\gamma_k\mathcal{B}_n & = & \sum_{k=1}^n s_k^2\gamma'_kA_0^{-1}\gamma_k\mathcal{B}_n\mathcal{B}^c_{nk} + \sum_{k=1}^n s_k^2\gamma'_kA_k^{-1}\gamma_k\beta_k\mathcal{B}_{nk} \\ & = & q\sum_{k=1}^n s_k^2 + o_{a.s.}(1) \to q,\ a.s. \end{array} \tag{1.4}$$

That is, $C'A_0^{-1}C \to q + (1-q)\rho^2\ a.s.$ or

$$C'A_0^{-1}C - \{q_n + (1-q_n)\rho^2\} \to 0\ a.s. \tag{1.5}$$

Moreover we have

$$\begin{array}{rcl} a'_{(1)}A^{-1}a_{(1)} & = & (C-\bar{y}_1\bar{y}_2)'(A_0-\bar{y}_2\bar{y}'_2)^{-1}(C-\bar{y}_1\bar{y}_2) \\ & = & (C-\bar{y}_1\bar{y}_2)'\left(A_0^{-1} + \frac{A_0^{-1}\bar{y}_2\bar{y}'_2A_0^{-1}}{1-\bar{y}'_2A_0^{-1}\bar{y}_2}\right)(C-\bar{y}_1\bar{y}_2) \\ & = & C'A_0{}^{-1}C - 2\bar{y}_1C'A_0{}^{-1}\bar{y}_2 + C'\frac{A_0^{-1}\bar{y}_2\bar{y}'_2A_0^{-1}}{1-\bar{y}'_2A_0^{-1}\bar{y}_2}C - 2\bar{y}_1C'\frac{A_0^{-1}\bar{y}_2\bar{y}'_2A_0^{-1}}{1-\bar{y}'_2A_0^{-1}\bar{y}_2}\bar{y}_2 \\ & & + \bar{y}_1^2\bar{y}'_2A_0{}^{-1}\bar{y}_2 + \bar{y}_1^2\bar{y}'_2\frac{A_0^{-1}\bar{y}_2\bar{y}'_2A_0^{-1}}{1-\bar{y}'_2A_0^{-1}\bar{y}_2}\bar{y}_2 \\ & = & C'A_0{}^{-1}C + \frac{(C'A_0^{-1}\bar{y}_2)^2}{1-\bar{y}'_2A_0^{-1}\bar{y}_2} - 2\bar{y}_1\cdot\frac{C'A_0^{-1}\bar{y}_2}{1-\bar{y}'_2A_0^{-1}\bar{y}_2} + \bar{y}_1^2\cdot\frac{\bar{y}'_2A_0^{-1}\bar{y}_2}{1-\bar{y}'_2A_0^{-1}\bar{y}_2}, \end{array}$$

where $(A_0-\bar{y}_2\bar{y}'_2)^{-1} = A_0^{-1} + \frac{A_0^{-1}\bar{y}_2\bar{y}'_2A_0^{-1}}{1-\bar{y}'_2A_0^{-1}\bar{y}_2}$. Because

$$\begin{array}{rcl} C'A_0^{-1}\bar{y}_2 & = & n^{-1/2}\sum_{k_1} s_{k_1}\gamma'_{k_1}A_0^{-1}\sum_{k_2}\gamma_{k_2} \\ & = & n^{-1/2}\left(\sum_{k=1}^n s_k\gamma'_kA_k^{-1}\gamma_k\beta_k + \sum_{k_1\neq k_2} s_{k_1}\gamma'_{k_1}A^{-1}_{k_1k_2}\gamma_{k_2}\beta_{k_1}\beta_{k_2(k_1)}\right), \end{array}$$

then we have $C'A_0^{-1}\bar{y}_2 = O_{a.s.}(n^{-1/2})$, similar to the proofs of (1.3) and (1.4). Similar calculations also lead to $\bar{y}'_2A_0^{-1}\bar{y}_2 = O_{a.s.}(n^{-1/2})$. Thus we obtain

$$a'_{(1)}A^{-1}a_{(1)} = C'A_0{}^{-1}C + O_{a.s.}(n^{-1/2}). \tag{1.6}$$

By (1.5) and (1.6), we have

$$a'_{(1)}A^{-1}a_{(1)} - \{q_n + (1-q_n)\rho^2\} \to 0 \ a.s.$$

That is, $R^2 - \{q_n + (1-q_n)\rho^2\} \to 0$ or $R^{*2} \to \rho^2 \ a.s.$ ■

2 Asymptotic normality

By (1.6), $\hat{\sigma}_{11} = \sum_{i=1}^n |s_i|^2 - \bar{y}_1^2 = O_{a.s.}(1)$ and $\bar{y}_1^2 = o_{a.s.}(1/n)$, we have

$$\begin{aligned} & n^{1/2}\left\{R^2 - (q_n + (1-q_n)\rho^2)\right\} \\ = \ & n^{1/2}\left\{\frac{a'_{(1)}A^{-1}a_{(1)}}{\hat{\sigma}_{11}} - (q_n + (1-q_n)\rho^2)\right\} \\ = \ & n^{1/2}\left\{\frac{C'\mathbf{A}_0^{-1}C}{\sum_{i=1}^n |s_i|^2} - (q_n + (1-q_n)\rho^2)\right\} + O_{a.s.}(1/n^{1/2}\} \end{aligned} \tag{2.1}$$

where $R^2 = a'_{(1)}\mathbf{A}^{-1}a_{(1)}/\hat{\sigma}_{11}$. To obtain the asymptotic normality of R^2, we just need to consider the asymptotic distribution of $n^{1/2}\left\{C'\mathbf{A}_0^{-1}C/\sum_{i=1}^n |s_i|^2 - (q_n + (1-q_n)\rho^2\right\}$. By the δ-method and the decomposition

$$\begin{aligned} & n^{1/2}\left\{\frac{C'\mathbf{A}_0^{-1}C}{\sum_{i=1}^n |s_i|^2} - (q_n + (1-q_n)\rho^2)\right\} \\ = \ & \frac{1}{\sum_{i=1}^n |s_i|^2}\left[n^{1/2}\{C'\mathbf{A}_0^{-1}C - (q_n + (1-q_n)\rho^2)\}\right] - \frac{q_n + (1-q_n)\rho^2}{\sum_{i=1}^n |s_i|^2}\left\{n^{1/2}\left(\sum_{i=1}^n |s_i|^2 - 1\right)\right\}, \end{aligned}$$

it suffices to show that

$$\begin{cases} n^{1/2}[C'\mathbf{A}_0^{-1}C - \{q_n + (1-q_n)\rho^2\}] \\ n^{1/2}(\sum_{i=1}^n |s_i|^2 - 1) \end{cases}$$

tends to a bivariate normal distribution and find out the asymptotic mean, variance, and covariance. To this end, we do it in two steps. First, we show that

$$\begin{cases} n^{1/2}(C'\mathbf{A}_0^{-1}C\mathcal{B}_n - EC'\mathbf{A}_0^{-1}C\mathcal{B}_n) \\ n^{1/2}(\sum_{i=1}^n |s_i|^2 - 1) \end{cases} \tag{2.2}$$

tends to a bivariate normal distribution, and second, we show that

$$n^{1/2}[E(C'\mathbf{A}_0^{-1}C\mathcal{B}_n) - \{q_n + (1-q_n)\rho^2\}] \to 0. \tag{2.3}$$

The proof of (2.2) is given in Section 2.2 and the proof of (2.3) in Section 2.3.

2.1 Simplification of certain expressions

Now, consider the martingale decomposition

$$C'A_0^{-1}C\mathcal{B}_n - E(C'A_0^{-1}C\mathcal{B}_n) = \sum_j Q_j, \quad Q_j = (E_j - E_{j-1})(C'A_0^{-1}C\mathcal{B}_n)$$

where E_j denotes the conditional expectation on $\mathbf{x}_j, \mathbf{x}_{j-1}, \ldots, \mathbf{x}_1$ and E_0 denotes the expectation. Then $\{Q_j, j = 1, \ldots, n\}$ forms a martingale difference sequence. Split $Q_j = Q_{j1} + Q_{j2} + Q_{j3}$ where

$$\begin{cases} Q_{j1} = (E_j - E_{j-1})(C'A_j^{-1}C\mathcal{B}_{nj}), \\ Q_{j2} = (E_j - E_{j-1})(C'(A_0^{-1} - A_j^{-1})C\mathcal{B}_{nj}) \\ Q_{j3} = (E_j - E_{j-1})(C'A_0^{-1}C\mathcal{B}_n\mathcal{B}_{nj}^c), \end{cases}$$

and the reader is reminded that $\mathcal{B}_n \supset \mathcal{B}_{nj}$. Note that

$$E\left|\sum_{j=1}^n Q_{j3}\right| \le K\sum_{j=1}^n E|C'C\mathcal{B}_{nj}^c| = o(n^{-t})$$

for any fixed $t > 0$. Here the sum $\sum_{j=1}^n Q_{j3}$ is negligible in the proof. Let

$$\begin{array}{lll} C_j = C - s_j\gamma_j, & \varepsilon_j = |s_j^2|\Delta_j, & \Delta_{jc} = \gamma_j'A_j^{-1}C_j, \\ \Delta_j = \gamma_j'A_j^{-1}\gamma_j - \frac{1}{n}\mathrm{tr}A_j^{-1}, & \bar\beta_j = \left(1 + \frac{1}{n}\mathrm{tr}A_j^{-1}\right)^{-1}, & \Delta_{jd} = |\Delta_{jc}|^2 - \frac{1}{n}C_j'A_j^{-2}C_j, \\ \delta_j = ns_j^2 - 1, & \Delta_{je} = s_j\Delta_{jc} - \frac{1}{n}\gamma' A_j^{-1}C_j, & \varepsilon_{jh} = 2\bar\beta_j\Delta_{je}\Delta_j, \\ \varepsilon_{jf} = -\bar\beta_j\beta_j\Delta_j|\Delta_{jc}|^2, & & \end{array}$$

$$\varepsilon_{je} = -\bar\beta_j\beta_j\Delta_j|s_j|^2(\gamma_j'A_j^{-1}\gamma_j)^2 + \bar\beta_j|s_j|^2\left(\Delta_j^2 + \frac{2}{n}\Delta_j\mathrm{tr}A_j^{-1}\right),$$

$$\varepsilon_{jg} = -2\bar\beta_j\beta_j\Delta_j s_j\gamma_j'A_j^{-1}\gamma_j\Delta_{jc} + \frac{1}{n}\bar\beta_j\gamma'A_j^{-1}C_j\cdot\gamma_j'A_j^{-1}\gamma_j.$$

Under these notations and by (1.15) of Bai and Silverstein (2004), we decompose Q_{j1} as

$$Q_{j1} \;=\; (E_j - E_{j-1})(C'A_j^{-1}C\mathcal{B}_{nj}) = E_j\left(2\Delta_{je} + \frac{1}{n^2}\delta_j\mathrm{tr}A_j^{-1}\mathrm{B}_{nj} + (E_j - E_{j-1})\varepsilon_j\mathcal{B}_{nj}.\right. \qquad (2.5)$$

Decompose Q_{j2} as

$$\begin{aligned} Q_{j2} &= -E_j \left\{ \frac{\delta_j \left(\mathrm{tr}A_j^{-1}\right)^2}{n^3} + \Delta_{jd} + \frac{2}{n}\Delta_{je}\mathrm{tr}A_j^{-1} \right\} \bar{\beta}_j \mathcal{B}_{nj} \\ &\quad - (E_j - E_{j-1})(\varepsilon_{je} + \varepsilon_{jf} + \varepsilon_{jg} + \varepsilon_{jh})\mathcal{B}_{nj}, \end{aligned} \tag{2.6}$$

where $\beta_j - \bar{\beta}_j = -\bar{\beta}_j \beta_j \Delta_j$.

Lemma 2.1 *We have*

$$\sum_{j=1}^{n} n^{1/2} Q_j - \sum_{j=1}^{n} n^{1/2} \tilde{Q}_j = o_p(1), \tag{2.7}$$

where

$$\begin{aligned} \tilde{Q}_j &= E_j \left[2\Delta_{je} + \tfrac{\delta_j}{n} \left\{ \tfrac{1}{n}\mathrm{tr}A_j^{-1} - \tfrac{\bar{\beta}_j}{n^2} \left(\mathrm{tr}A_j^{-1}\right)^2 \right\} - \bar{\beta}_j \Delta_{jd} - \tfrac{2\bar{\beta}_j}{n} \Delta_{je} \mathrm{tr}A_j^{-1} \right] \mathcal{B}_{nj} \\ &= E_j \left[2\Delta_{je} \left(1 - \tfrac{\bar{\beta}_j}{n}\mathrm{tr}\mathbf{A}_j^{-1}\right) + \tfrac{\delta_j}{n} \left\{ \tfrac{1}{n}\mathrm{tr}A_j^{-1} - \tfrac{\bar{\beta}_j}{n^2} \left(\mathrm{tr}A_j^{-1}\right)^2 \right\} - \bar{\beta}_j \Delta_{jd} \right] \mathcal{B}_{nj}. \end{aligned}$$

PROOF. Comparing (2.4) and (2.6), to prove (2.7), one just needs to show that

$$n \sum_{j=1}^{n} \left\{ E(|\varepsilon_j|^2 \mathcal{B}_{nj}) + E(|\varepsilon_{je}|^2 \mathcal{B}_{nj}) + E(|\varepsilon_{jf}|^2 \mathcal{B}_{nj}) + E(|\varepsilon_{jg}|^2 \mathcal{B}_{nj}) + E(|\varepsilon_{jh}|^2 \mathcal{B}_{nj}) \right\} \to 0. \tag{2.8}$$

First, let $D_2' A_j^{-1} D_2 = (a_{ik})$, then we have

$$\begin{aligned} & n \sum_{j=1}^{n} E(|\varepsilon_j|^2 \mathcal{B}_{nj}) = n \sum_{j=1}^{n} E|s_j^4||\Delta_j|^2 \mathcal{B}_{nj} \\ \leq\ & K n^{-3} \sum_{j=1}^{n} \left\{ \left(\sum_{i=1}^{m} d_i^2 \right)^2 \left(\sum_{i=1}^{m} |a_{ii}|^2 + \sum_{i<k} |a_{ik}^2| \right) \right. \\ & \left. + \left(\sum_{i=1}^{m} d_i^2 \right) \sum_{i=1}^{m} d_i^2 |a_{ii}|^2 E|x_{1j}^6| + \sum_{i=1}^{m} d_i^4 a_{ii}^2 E|x_{1j}^8| + \sum_{i_1<i_2} a_{i_1 i_2}^2 d_i^4 E|x_{1j}^6| \right\} \mathcal{B}_{nj} \\ \leq\ & K n^{-3} \sum_{j=1}^{n} E\left\{ \mathrm{tr}(A_j^{-2}\mathcal{B}_{nj}) + \eta_n^4 n^2 + \eta_n^2 n \right\} \to 0 \ (X_{11} \text{ is truncated at } \eta_n n^{1/2}). \end{aligned} \tag{2.9}$$

By similar approaches, we may verify that for all $\ell = e, f, g$,

$$n^{1/2} \sum_{j=1}^{n} E(|\varepsilon_{j\ell}|^2 \, \mathcal{B}_{nj}) \to 0. \tag{2.10}$$

The proof of assertion (2.7) is completed. ■

2.2 Asymptotic normality for (2.2)

To obtain the asymptotic normality of expression (2.2), we need to verify the Lyapunoff condition and obtain the asymptotic variance and covariance in the following steps:

(a) Verify the Lyapunoff condition, i.e. $\sum_{j=1}^{n} E(n^2|\tilde{Q}_j|^4\mathcal{B}_{nj}) \to 0$.

(b) Find the limit of $\sum_{j=1}^{n} E_{j-1}(n\tilde{Q}_j^2\mathcal{B}_{nj})$.

(c) Find the limit of $\sum_{j=1}^{n} E_{j-1}(\delta_j\tilde{Q}_j\mathcal{B}_{nj})$.

Lyapunoff condition. Given the structure of $\tilde{Q}_j$ in (2.7), one needs to verify only

$$n^2\sum_{j=1}^{n} E(|a_j|^4\mathcal{B}_{nj}) \to 0, \tag{2.11}$$

for $a_j = \delta_j/n, \Delta_{jd}$, and Δ_{je}. The limit (2.11) is true for $a_j = \delta_j/n$ because

$$E\delta_j^4 \le K\left\{\sum_i d_i^8 E|X_{ij}^8| + \sum_{i<k} d_i^2 d_k^6 E|X_{ij}|^6 + \left(\sum d_i^2 d_j^2\right)^2\right\} \le Kn\eta_n^4 = o(n).$$

where X_{ij} is truncated at $d_i^{-1/2}\eta_n n^{1/4}$.

Next let $D_2'A_j^{-1}C_jC_j'A_j^{-1}D_2 = (w_{k_1k_2})$, then we have

$$\begin{aligned}
E(|\Delta_{jd}|^4\mathcal{B}_{nj}) &= \frac{1}{n^4}E\left(\sum\nolimits_{k_1\ne k_2} X_{jk_1}X_{jk_2}w_{k_1k_2} + \sum\nolimits_k (X_{jk}^2-1)w_{kk}\right)^4\mathcal{B}_{nj}\\
&\le \frac{K}{n^4}\left[\left\{E(C_j'C_j)^2\right\}^2 + E\{(C_j'C_j)^4\} + O\left(\sum\nolimits_k d_{1k}^8 E(X_{jk}^2-1)^4\right)\right] = o(n^{-3}),
\end{aligned}$$

where X_{kj} is truncated at $d_{1k}^{-1/2}\eta_n n^{1/4}$, $E\{(C_j'C_j)^2\} \le K$, $E\{(C_j'C_j)^4\} \le K$, $\alpha_k = D_2 e_k$, e_k is the zero vector expect the kth element being one and

$$\begin{aligned}
w_{kk} &= \mathbf{e}_k'D_2'A_j^{-1}C_jC_j'A_j^{-1}D_2\mathbf{e}_k\mathcal{B}_{nj} \le K\mathbf{e}_k'D_2'C_jC_j'D_2\mathbf{e}_k = K\left|C_j'\boldsymbol{\alpha}_k\right|^2\\
E(w_{kk}^4) &\le E(|C_j'\boldsymbol{\alpha}_k|^8) = E\left(\left|\sum\nolimits_{k_1\ne j} s_{k_1}\gamma_{k_1}'\boldsymbol{\alpha}_i\right|^8\right)\\
&= \frac{1}{n^8}E\left\{\left|\sum\nolimits_{k_1\ne j}(X_{k_1}'D_2'\alpha_i d_1 X_{k_1} - \gamma'\alpha_k) + (n-1)\gamma'\alpha_k\right|^8\right\} = O(d_{1k}^8)
\end{aligned}$$

because $E(|X_{11}|^8) \leq n^2\eta_n^4 E(|X_{11}|^4)$ with X_{ij} is truncated at $n^{1/2}\eta_n$. Then (2.11) holds for $a_j = \Delta_{jd}$. At last, we have $E(|\Delta_{je}|^4 \mathcal{B}_{nj}) \leq (K\eta_n^4/n^3)(\gamma'\gamma) = o(n^{-3})$ by similar proofs. Then (2.11) holds for $a_j = \Delta_{je}$. So the Lyapunoff condition is verified. ■

Completing step (b) To find the limit of $\sum_{j=1}^n E_{j-1}(n\tilde{Q}_j^2\mathcal{B}_{nj})$, first we have

$$\sum\nolimits_{i=1}^{m}[(d_1'd_1)_{ii}]^2 = \sum\nolimits_{i=1}^{m} d_i^4 \leq \max d_i^2. \tag{2.12}$$

Second, we have

$$\begin{aligned}
&\frac{1}{n}\sum\nolimits_{j=1}^{n}\sum\nolimits_{i=1}^{m}[\{E_j(d_1'C_j'A_j^{-1}D_2\mathcal{B}_{nj})\}_{ii}]^2 = \frac{1}{n}\sum\nolimits_{j=1}^{n}\sum\nolimits_{i=1}^{m} d_i^2\{E_j(C_j'A_j^{-1}D_2\mathbf{e}_i\mathcal{B}_{nj})\}^2 \\
\leq\ & \max d_i^2 \cdot \frac{1}{n}\sum\nolimits_{j=1}^{n} E_j(C_j'A_j^{-1})E_j(A_j^{-1}C_j\mathcal{B}_{nj}) = \max d_i^2 \cdot O_p(1) \text{ (by Lemma 3.15)}. \quad (2.13)
\end{aligned}$$

Third, we have

$$\begin{aligned}
&\frac{1}{n}\sum\nolimits_{j=1}^{n}\sum\nolimits_{i=1}^{m} E[\{(E_jD_2'A_j^{-1}C_jC_j'A_j^{-1}D_2\mathcal{B}_{nj})_{ii}\}^2] \\
\leq\ & \max d_i^4 \frac{1}{n}\sum\nolimits_{j=1}^{n} E\left[\left(\frac{1}{n}X_j'D_2'A_j^{-1}D_2X_j\right)^2\right] = O(\max d_i^4) \qquad (2.14)
\end{aligned}$$

That is, $\frac{1}{n}\sum_{j=1}^n\sum_{i=1}^m[\{E_j(D_2'A_j^{-1}C_jC_j'A_j^{-1}D_2\mathcal{B}_{nj})\}_{ii}]^2 = O(\max d_i^4)$. Moreover, by the Holder inequality, we have

$$\sum\nolimits_{i=1}^{m}(d_1'd_1)_{ii}(E_jd_1'C_j'A_j^{-1}D_2)_{ii}\mathcal{B}_{nj} = O_p(\max d_i^2) \tag{2.15}$$

$$\sum\nolimits_{i=1}^{m}(d_1'd_1)_{ii}(E_jD_2'A_j^{-1}C_jC_j'A_j^{-1}D_2\mathcal{B}_{nj})_{ii} = O_p(\max d_i^4). \tag{2.16}$$

Recall (2.7) and by (1.15) of Bai and Silverstein (2004), we have

$$\begin{aligned}
&\sum_{j=1}^{n} nE_{j-1}(\tilde{Q}_j\mathcal{B}_{nj})^2 = n\sum_{j=1}^{n}\Bigg[4E_{j-1}\left\{E_j\Big(1-\frac{\bar{\beta}_j}{n}\mathrm{tr}A_j^{-1}\Big)\Delta_{je}\mathcal{B}_{nj}\right\}^2 \\
&+ E_{j-1}\frac{\delta_j^2}{n^2}\left\{E_j\left(\frac{1}{n}\mathrm{tr}A_j^{-1} - \frac{\bar{\beta}_j}{n^2}\left(\mathrm{tr}A_j^{-1}\right)^2\right)\mathcal{B}_{nj}\right\}^2 \\
&+ 4E_{j-1}\frac{\delta_j}{n}\left\{E_j\Big(1-\frac{\bar{\beta}_j}{n}\mathrm{tr}A_j^{-1}\Big)\Delta_{je}\mathcal{B}_{nj}\right\}E_j\left\{\frac{1}{n}\mathrm{tr}A_j^{-1} - \frac{\bar{\beta}_j}{n^2}\left(\mathrm{tr}A_j^{-1}\right)^2\right\}\mathcal{B}_{nj} \\
&+ \bar{\beta}_j^2E_{j-1}(E_j\Delta_{jd}\mathcal{B}_{nj})^2 - 4E_{j-1}\left\{E_j\Big(1-\frac{\bar{\beta}_j}{n}\mathrm{tr}A_j^{-1}\Big)\Delta_{je}\mathcal{B}_{nj}\right\}\bar{\beta}_jE_j\Delta_{jd}\mathcal{B}_{nj} \\
&-2E_{j-1}\frac{\delta_j}{n}\left\{\frac{1}{n}E_j\mathrm{tr}A_j^{-1} - \frac{\bar{\beta}_j}{n^2}E_j\left(\mathrm{tr}A_j^{-1}\right)^2\mathcal{B}_{nj}\right\}\bar{\beta}_jE_j\Delta_{jd}\mathcal{B}_{nj}\Bigg]
\end{aligned}$$

$$
\begin{aligned}
= \quad & \frac{1}{n}\sum_{j=1}^{n}\Bigg\{4(1-q)^2\left(E_jC_j'A_j^{-1}\mathcal{B}_{nj}\right)\left(E_jA_j^{-1}C_j\mathcal{B}_{nj}\right)+4(1-q)^2\left(E_jC_j'A_j^{-1}\gamma\mathcal{B}_{nj}\right)^2 \\
& + E(Y_{11}^2-1)^2q^2+8q(1-q)\left(E_j\gamma'A_j^{-1}C_j\mathcal{B}_{nj}\right)+2(1-q)^2\mathrm{tr}\left(E_jA_j^{-1}C_jC_j'A_j^{-1}\mathcal{B}_{nj}\right)^2 \\
& - 8(1-q)^2\left(\gamma'E_jA_j^{-1}C_jC_j'A_j^{-1}\mathcal{B}_{nj}\right)\left(E_jA_j^{-1}C_j\mathcal{B}_{nj}\right) \\
& - 4q(1-q)E_j(\gamma'A_j^{-1}C_j)^2\mathcal{B}_{nj}\Bigg\}+o_p(1) \\
& \text{(see (2.12)–(2.16))}.
\end{aligned}
$$

Because by Lemma 3.4, we have $1-\bar{\beta}_jn^{-1}\mathrm{tr}\mathbf{A}_j^{-1}=1-q+o(n^{-t})$ and $n^{-1}\mathrm{tr}\mathbf{A}_j^{-1}\left(1-\bar{\beta}_jn^{-1}\mathrm{tr}\mathbf{A}_j^{-1}\right)=q+o(n^{-t})$ uniformly where t is any positive number. For brevity, let $\breve{A}_j$ be the analogue of the matrix A_j with vectors $\mathbf{x}_{j+1},\ldots,\mathbf{x}_n$ replaced by their i.i.d. copies $\breve{\mathbf{x}}_{j+1},\ldots,\breve{\mathbf{x}}_n$. The same for the other notations. Using this notation system, we may simplify $\sum_{j=1}^n E_{j-1}n\tilde{Q}_j^2\mathcal{B}_{nj}$ as

$$
\begin{aligned}
= \quad & \frac{1}{n}\sum_{j=1}^{n}\Bigg\{4(1-q)^2\left(E_jC_j'A_j^{-1}\mathcal{B}_{nj}\breve{A}_j^{-1}\breve{C}_j\breve{\mathcal{B}}_{nj}\right)+4(1-q)^2\left(E_jC_j'A_j^{-1}\gamma\mathcal{B}_{nj}\right)^2 \\
& + E(Y_{11}^2-1)^2q^2+8q(1-q)\left(E_j\gamma'A_j^{-1}C_j\mathcal{B}_{nj}\right)+2(1-q)^2\mathrm{tr}\left(E_jA_j^{-1}C_jC_j'A_j^{-1}\mathcal{B}_{nj}\right)^2 \\
& - 8(1-q)^2\left(\gamma'E_jA_j^{-1}C_jC_j'A_j^{-1}\mathcal{B}_{nj}\right)\left(E_jA_j^{-1}C_j\mathcal{B}_{nj}\right)-4q(1-q)E_j(\gamma'A_j^{-1}C_j)^2\mathcal{B}_{nj}\Bigg\} \\
& \text{(by Lemmas 3.15–3.20)} \\
= \quad & -2(1-q)(\rho^2)^2+4(1-q)\rho^2+2q(1-q)+\kappa q^2+o_p(1).
\end{aligned}
$$

where $\kappa=E(d_1X_i-1)^2=\mathrm{var}((d_1X_i)^2)/\sigma_{11}^4\to 2$. That is,

$$
\begin{aligned}
\sum_{j=1}^{n} nE_{j-1}(\tilde{Q}_j\mathcal{B}_{nj})^2 &= \underbrace{-2(1-q)(\rho^2)^2+4(1-q)\rho^2+2q(1-q)+\kappa q^2}_{\sigma_1^2}+o_p(1) \\
&= \underbrace{-2(1-q)(\rho^2)^2+4(1-q)\rho^2+2q(1-q)+2q^2}_{\sigma_1^2}+o_p(1). \quad (2.17)
\end{aligned}
$$

So the limit of $\sum_{j=1}^n E_{j-1}(n\tilde{Q}_j^2\mathcal{B}_{nj})$ is obtained. ■

Competing step (c). We have $\sum_{j=1}^n E_{j-1}(\delta_j \tilde{Q}_j \mathcal{B}_{nj})$

$$\begin{aligned}
&= n\sum\nolimits_{j=1}^n E_{j-1}\left\{2\Delta_{je}\frac{\delta_j}{n}\left(1-\frac{\bar{\beta}_j}{n}\mathrm{tr}\mathbf{A}_j^{-1}\right)+\frac{\delta_j^2}{n^2}\left(\frac{1}{n}\mathrm{tr}A_j^{-1}-\frac{\bar{\beta}_j}{n^2}\left(\mathrm{tr}A_j^{-1}\right)^2\right)-\bar{\beta}_j\Delta_{jd}\frac{\delta_j}{n}\right\}\mathcal{B}_{nj} \\
&= \underbrace{-2(1-q)(\rho^2)^2+4(1-q)\rho^2+\kappa q}_{\sigma_{12}}+o_p(1) \\
&= \underbrace{-2(1-q)(\rho^2)^2+4(1-q)\rho^2+2q}_{\sigma_{12}}+o_p(1). \qquad (2.18)
\end{aligned}$$

So the limit of $\sum_{j=1}^n E_{j-1}\delta_j\tilde{Q}_j$ is obtained. ■

By the properties of a martingale difference sequence and by (2.17)–(2.18), we know that

$$\left\{\begin{array}{l} n^{1/2}\{C'\mathbf{A}_0^{-1}C\mathcal{B}_n - E(C'\mathbf{A}_0^{-1}C\mathcal{B}_n)\} \\ n^{1/2}(\sum_{i=1}^n |s_i|^2-1) \end{array}\right. \qquad (2.19)$$

tends to a bivariate normal distribution $N\left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{pmatrix}\right\}$ where $\sigma_1^2=-2(1-q)(\rho^2)^2+4(1-q)\rho^2+2q(1-q)+2q^2$, $\sigma_2^2=2$ and $\sigma_{12}=-2(1-q)(\rho^2)^2+4(1-q)\rho^2+2q$.

2.3 Proofs of (2.3): $n^{1/2}\{E(C'A_0^{-1}C\mathcal{B}_n)-q_n-(1-q_n)\rho^2\}=o(1)$

First, we have

$$E(C'A_0^{-1}C\mathcal{B}_n) = n(n-1)E(s_1s_2\boldsymbol{\gamma}_1'A_{12}^{-1}\boldsymbol{\gamma}_2\beta_1\beta_{2(1)}\mathcal{B}_n)+nE(s_1^2\boldsymbol{\gamma}_1'A_1^{-1}\boldsymbol{\gamma}_1\beta_1\mathcal{B}_n).$$

By (0.2), for any positive constant t, we have

$$\begin{aligned}
nE(s_1^2\boldsymbol{\gamma}_1'A_1^{-1}\boldsymbol{\gamma}_1\beta_1\mathcal{B}_n) &\le nE(s_1^2\boldsymbol{\gamma}_1'A_1^{-1}\boldsymbol{\gamma}_1\beta_1\mathcal{B}_{n1})+nE(s_1^2\mathcal{B}_{n1}^c) \\
&= nE(s_1^2\boldsymbol{\gamma}_1'A_1^{-1}\boldsymbol{\gamma}_1\beta_1\mathcal{B}_{n1})+o(n^{-t}).
\end{aligned}$$

Similarly, $n(n-1)\{E(s_1s_2\boldsymbol{\gamma}_1'A_{12}^{-1}\boldsymbol{\gamma}_2\beta_1\beta_{2(1)}\mathcal{B}_n)-E(s_1s_2\boldsymbol{\gamma}_1'A_{12}^{-1}\boldsymbol{\gamma}_2\beta_1\beta_{2(1)}\mathcal{B}_{n12})\}=o(n^{-t})$. We have

$$\begin{aligned}
&|nE(s_1^2\gamma_1'A_1^{-1}\gamma_1\beta_1\mathcal{B}_{n1})-q| \\
\le\ & n\left|E\left(\frac{\delta_j}{n}\Delta_j\beta_1\mathcal{B}_{n1}\right)\right|+\left|nE\left(\frac{\delta_j}{n}\frac{1}{n}\mathrm{tr}A_1^{-1}\beta_1\mathcal{B}_{n1}\right)\right|+\left|E\left\{\frac{1}{1+\frac{1}{n}\mathrm{tr}A_1^{-1}}-(1-q)\right\}\right|.
\end{aligned}$$

With $\beta_1 = \bar\beta_1 - \bar\beta_1\Delta_1 + \bar\beta_1^2\beta_1\Delta_1^2$, we have

$$E\left(\frac{\delta_1}{n}\Delta_1\beta_1\mathcal{B}_{n1}\right) = E\left(\frac{\delta_1}{n}\Delta_1\bar\beta_1\mathcal{B}_{n1}\right) - E\left(\frac{\delta_1}{n}\Delta_1^2\bar\beta_1\mathcal{B}_{n1}\right) + E\left(\frac{\delta_1}{n}\Delta_1^3\bar\beta_1^2\beta_1\mathcal{B}_{n1}\right) = o(n^{-3/2})$$

because $\sum_{i=1}^m (d_1'd_1)_{ii}(D_2'A_1^{-1}D_2)_{ii}\mathcal{B}_{n1} \le \max_{i=1,\dots,m}(D_2'A_1^{-1}D_2)_{ii}\mathcal{B}_{n1} = |A_1^{-1}\mathcal{B}_{n1}D_2\mathbf{e}_i\mathbf{e}_i'D_2'| \le |A_1^{-1}\mathcal{B}_{n1}| \le K$, and

$$\begin{aligned}
& E\left(\frac{\delta_j}{n}\frac{1}{n}\mathrm{tr}A_1^{-1}\beta_1\mathcal{B}_{n1}\right) \\
= \; & E\left(\frac{\delta_j}{n}\bar\beta_1\frac{1}{n}\mathrm{tr}A_1^{-1}\mathcal{B}_{n1}\right) - E\left(\frac{\delta_j}{n}\Delta_1\bar\beta_1\frac{1}{n}\mathrm{tr}A_1^{-1}\mathcal{B}_{n1}\right) + E\left(\frac{\delta_j}{n}\Delta_1^2\bar\beta_1^2\beta_1\frac{1}{n}\mathrm{tr}A_1^{-1}\mathcal{B}_{n1}\right) \\
= \; & -E\left(\frac{\delta_j}{n}\Delta_1\bar\beta_1\frac{1}{n}\mathrm{tr}A_1^{-1}\mathcal{B}_{n1}\right) + E\left(\frac{\delta_j}{n}\Delta_1^2\bar\beta_1^2\beta_1\frac{1}{n}\mathrm{tr}A_1^{-1}\mathcal{B}_{n1}\right) \\
= \; & o(n^{-3/2}).
\end{aligned}$$

We also have $\left|E(1+n^{-1}\mathrm{tr}A_1^{-1})^{-1} - (1-q)\right| = O(n^{-1})$ because $E(n^{-1}\mathrm{tr}A_1^{-1} - (1-q_n)/q_n) = O(n^{-1})$. Then $n^{1/2}|nE(s_1^2\gamma_1'A_1^{-1}\gamma_1\beta_1\mathcal{B}_{n1}) - q| = O(1)$,

$$E\{s_1s_2\boldsymbol{\gamma}_1'A_{12}^{-1}\boldsymbol{\gamma}_2(\bar\beta_1 - \bar\beta_1\Delta_1 + \bar\beta_1^2\beta_1\Delta_1^2)\beta_{2(1)}\} = o(n^{-3/2}) + (1-q)\rho^2,$$

and therefore

$$n^{1/2}\{E(C'A_0^{-1}C\mathcal{B}_n) - q_n - (1-q_n)\rho^2\} = o(1). \tag{2.20}$$

■

Proof of the asymptotic normality in Theorem 2. By (2.19) and (2.20), we have

$$\left\{\begin{array}{l} n^{1/2}\{C'\mathbf{A}_0^{-1}C\mathcal{B}_n - q_n - (1-q_n)\rho^2\} \\ n^{1/2}(\sum_{i=1}^n |s_i|^2 - 1) \end{array}\right. \tag{2.21}$$

tends to a bivariate normal distribution $N\left\{\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{array}\right)\right\}$ where $\sigma_1^2 = -2(1-q)(\rho^2)^2 + 4(1-q)\rho^2 + 2q(1-q) + 2q^2$, $\sigma_2^2 = 2$ and $\sigma_{12} = -2(1-q)(\rho^2)^2 + 4(1-q)\rho^2 + 2q$. Then by the δ method and (2.1), we have that

$$n^{1/2}\{R^2 - q_n - (1-q_n)\rho^2\}$$

tends to a normal distribution $N\{0, \sigma^2(\rho^2)\}$ where

$$\begin{aligned}
& \sigma^2(\rho^2) \\
= \; & \sigma_1^2 + \{q + (1-q)\rho^2\}^2\sigma_2^2 - 2\{q + (1-q)\rho^2\}\sigma_{12} \\
= \; & -2(1-q)(\rho^2)^2 + 4(1-q)\rho^2 + 2q(1-q) + 2q^2 + 2\{q + (1-q)\rho^2\}^2 \\
& -2\{q + (1-q)\rho^2\}\{-2(1-q)(\rho^2)^2 + 4(1-q)\rho^2 + 2q\} \\
= \; & 2\{q + (1-q)\rho^2\}^2 - 2\{-2(1-q)(\rho^2)^2 + 4(1-q)\rho^2 + 2q\}\{q + (1-q)\rho^2 - 1/2\} \\
& +q(1-q)(2-2) \\
= \; & 2\{q + (1-q)\rho^2\}^2 - 2\{-2(1-q)(\rho^2)^2 + 4(1-q)\rho^2 + 2q\}\left\{q + (1-q)\rho^2 - \frac{1}{2}\right\}. \quad (2.22)
\end{aligned}$$

Let $R^{**2} = R^2 - \{(p-1)/(n-p)\}(1-R^2)$. By the delta method, we have $n^{1/2}(R^{**2} - \rho^2) \to N\{0, \sigma^2(\rho^2)/(1-q)^2\}$. Moreover, $P\{n^{1/2}(R^{**2} - \rho^2) \le x\} = P\{R^{**2} < 0, n^{1/2}(R^{**2} - \rho^2) \le x\} + P\{R^{**2} \ge 0, n^{1/2}(R^{**2} - \rho^2) \le x\}$ and $P\{n^{1/2}(R^{*2} - \rho^2) \le x\}$ have the same limiting distribution because $P(R^{**2} < 0) \to 0$ for $\rho^2 > 0$. That is

$$n^{1/2}(R^{*2} - \rho^2) \to N\{0, \sigma^2(\rho^2)/(1-q)^2\}$$

for $\rho^2 > 0$. Moreover, we have for $\rho^2 > 0$

$$n^{1/2}\{g(R^{*2}) - g(\rho^2)\} \to N[0, \sigma^2(\rho^2)\{g'\left(\rho^2\right)\}^2/(1-q)^2]$$

where $g(\cdot)$ is a nondecreasing differential function. By choosing

$$g(x) = \int_0^x \frac{1-q}{\sigma_t}\,\mathrm{d}t,$$

where $\sigma_t^2 = 2\{q + (1-q)t\}^2 - 2\{-2(1-q)t^2 + 4(1-q)t + 2q\}\{q + (1-q)t - 1/2\}$, we obtain

$$n^{1/2}\{g(R^{*2}) - g(\rho^2)\} \to N\Big(0, 1\Big).$$

When $\rho^2 = 0$, we have $\sigma^2(\rho^2) = 2q(1-q)$. Then we have

$$n^{1/2}(R^2 - q_n) \to N\{0, 2q(1-q)\}.$$

Then by delta method, we have $n^{1/2}\{g(R^2) - g(q_n)\} \to N(0,1)$ where $g(x) = \sqrt{2}\arccos(x)$. That is,

$$\sqrt{2n}\{\arccos(R) - \arccos(\sqrt{q_n})\} \to N\left(0,1\right).$$

The proof of Theorem 2 is now completed. ■

3 Some lemmas

For simplicity, the following lemmas and corollaries are written to hold for truncation at $\eta_n n^{1/2}$ of X_{ij}. They also hold for the truncation $\min\left\{d_i^{-\frac{1}{2}}\eta_n n^{\frac{1}{4}}, d_{1i}^{-\frac{1}{2}}\eta_n n^{\frac{1}{4}}, \eta_n n^{1/2}\right\}$ of X_{ij}.

Lemma 3.1 (Bai and Silverstein (2010, p. 225)) *Suppose that $X_i,\ i = 1,\ldots,n$ are independent, with $EX_i = 0$, $E|X_i|^2 = 1$, $\sup E|X_i|^4 = \nu < +\infty$ and $|X_i| \le \eta n^{1/2}$ with $\eta > 0$. Assume that $\mathbf{M}$ is a nonrandom complex matrix. Then, for any given $2 \le \ell \le b\log(n\nu^{-1}\eta^4)$ and $b > 1$, we have*

$$E\left(|\boldsymbol{\alpha}'\mathbf{M}\boldsymbol{\alpha} - \mathrm{tr}(\mathbf{M})|^\ell\right) \le \nu n^\ell (n\eta^4)^{-1}(40b^2\|\mathbf{M}\|\eta^2)^\ell$$

where $\boldsymbol{\alpha} = (X_1,\ldots,X_n)'$.

Lemma 3.2 *Suppose that $X_i,\ i = 1,\ldots,n$ are independent random variables with $E(X_i) = 0$ and finite b-th moments $(b \ge 2)$, then*

$$\left|E\left(\sum_{i=1}^n X_i\right)^b\right| \le \sum_{1\le t\le b/2}\ \sum_{\substack{i_1+\cdots+i_t=b \\ \min(i_1,\cdots,i_t)\ge 2}} \frac{b!}{i_1!\cdots i_t!t!}\prod_{h=1}^{t}{}' \left(\sum_{i=1}^n |E(X_i^{i_h})|\right).$$

PROOF. By multinomial expansion

$$\left(\sum_{i=1}^n X_i\right)^{2\ell} = \sum_{i_1+\cdots i_n=b} b!i_1!\cdots i_n!X_1^{i_1}\cdots X_n^{i_n}.$$

Taking expectation on both sides, skip all terms with some $i_h = 1$, then we have

$$E\left\{\left(\sum_{i=1}^{n} X_i\right)^b\right\} = \sum_{1\le t\le b/2}\sum_{i_1+\cdots+i_t=b}\frac{b!}{i_1!\cdots i_t!}\sum_{1\le j_1<\cdots<j_t\le n} E(X_{j_1}^{i_1})\cdots E(X_{j_t}^{i_t})$$

$$\le \sum_{1\le t\le b/2}\sum_{i_1+\cdots i_t=b}\frac{b!}{i_1!\cdots i_t!t!}\sum_{1\le j_1,\ldots,j_t\le n}|E(X_{j_1}^{i_1})|\ldots|E(x_{j_t}^{i_t})|$$

$$= \sum_{1\le t\le b/2}\sum_{i_1+\cdots i_t=b}\frac{b!}{i_1!\cdots i_t!t!}\sum_{j=1}^{n}|E(X_j^{i_1})|\cdots\sum_{j=1}^{n}|E(X_j^{i_t})|.$$

The proof is completed. ■

Corollary 3.1 *If x_{ij} are independent and identically distributed random variables with mean zero, variance 1, $|x_{ij}| \le \eta_n n^{1/2}$, $\eta_n \to 0$, and $\mathbf{M}$ be an $n\times n$ nonrandom matrix with bounded operator norm, then for any $\varepsilon > 0$,*

$$\mathrm{P}(|\mathbf{x}_1'\mathbf{M}\mathbf{x}_2| \ge \varepsilon n) = o(n^{-t}),$$

where $\mathbf{x}_i = (x_{i1},\ldots,x_{in})'$.

PROOF. By the Markov inequality, we have

$$\mathrm{P}(|\mathbf{x}_1'\mathbf{M}\mathbf{x}_2| \ge n\varepsilon) \le \frac{1}{(n\varepsilon)^{2\ell}}E\left(|\mathbf{x}_1'\mathbf{M}\mathbf{x}_2|^{2\ell}\right). \tag{3.23}$$

Write $a_i = x_1'a_i$, $M = (a_1,\ldots,a_n)$. Employing Lemma 3.2 with $X_i = a_i x_{2i}$ and $\ell = [\log n]$, we obtain

$$E\left(|x_1'Mx_2|^{2\ell}\right) \le \sum_{t=1}^{\ell}\sum_{\substack{i_1+\ldots+i_t=2\ell\\ i_1,\ldots,i_t\ge 2}}\frac{(2\ell)!}{i_1!\ldots i_t!t!}\sum_{j=1}^{n}E(a_j^{i_1})\cdots\sum_{j=1}^{n}E(a_j^{i_t})(\eta_n n^{1/2})^{2\ell-2t}.$$

Employing Lemma 3.2 again, we have that for $i_h \ge 2$ and all large n,

$$|E(a_j^{i_h})| \le \sum_{1\le s\le i_h/2}\sum_{\substack{k_1+\ldots+k_s=i_h\\ k_1,\ldots,k_s\ge 2}}\frac{(i_h)!}{k_1!\ldots k_s!s!}\sum_{i=1}^{n}a_{ij}^{k_1}\ldots\sum_{i=1}^{n}a_{ij}^{k_s}(\eta_n n^{1/2})^{i_h-2s}$$

$$\le \|\mathbf{M}\|^{i_h}\sum_{1\le s\le i_h/2}\frac{s^{i_h}}{s!}\eta_n^{i_h-2s}n^{\frac{1}{2}i_h-s} \le \|\mathbf{M}\|^{i_h}\frac{i_h}{2}\eta_n^{i_h-2}n^{\frac{1}{2}i_h-1}.$$

Combining the two inequalities above, we obtain for all large n,

$$E\left(|\mathbf{x}_1'\mathbf{M}\mathbf{x}_2|^{2\ell}\right) \le \|\mathbf{M}\|^{2\ell} n^{2\ell} \eta_n^{4\ell} \sum_{t=1}^{\ell} \frac{t^{2\ell}}{t!\eta_n^{2t} n'} \ell' \le \|\mathbf{M}\|^{2\ell} n^{2\ell-1} \eta_n^{4\ell-2} \ell.$$

Then the lemma follows from this estimation and (3.23). ■

In the proof of the main theorems, to simplify the expressions of the target quantities, we will frequently replace some components with simpler ones. To this end, we establish the following lemma.

Lemma 3.3 (Substitution Lemma) *Suppose that* $\sum_{i=1}^n E(|a_i|) \le B$, $\sum_{i=1}^n |a_i(b_i - c_i)| \le Cn^\alpha$ *or* $\sum_{i=1}^n \sqrt{E\{a_i^2(b_i - c_i)^2\}} \le Cn^\alpha$*, and for any fixed* $\varepsilon > 0$ *and* $t > 0$*, we have*

$$\mathrm{P}(|b_i - c_i| \ge \varepsilon) = o(n^{-t}),$$

then $\sum_{i=1}^n a_i b_i$ *and* $\sum_{i=1}^n a_i c_i$ *have the same weak limit simultaneously.*

PROOF. The lemma follows from the fact that for any $\varepsilon > 0$,

$$E\left(\left|\sum_{i=1}^n a_i(b_i - c_i)\right|\right) \le \varepsilon \sum_{i=1}^n E(|a_i|) + Cn^\alpha \cdot \mathrm{P}(|b_i - c_i| > \varepsilon) \to \varepsilon B.$$

The proof is completed. ■

Lemma 3.4 *For any* $\varepsilon > 0$*, we have*

$$\mathrm{P}\left(\left|\gamma_k' A_k^{-h} \gamma_k - \frac{1}{n}\mathrm{tr} A_k^{-h}\right| \mathcal{B}_{nk} \ge \varepsilon\right) = o(n^{-t}).$$

where $h > 0$ *and* $\mathcal{B}_{nk}$ *also be used as an indicator function for brevity.*

PROOF. Recall that $\gamma_k' A_k^{-h} \gamma_k = \mathbf{x}_k' D_2' A_k^{-h} D_2 \mathbf{x}_k$ and notice that $\|D_2' A_k^{-h} D_2 \mathcal{B}_{nk}\| \le a^{-h}$. Then the lemma is consequence of an application of Lemma 3.1 with $\ell = [\log n]$. The truth of the second conclusion can be seen by the similar reasons.

Lemma 3.5 *For any* $\varepsilon > 0$*, we have*

$$\mathrm{P}\left(\left|n^{-1}(\mathrm{tr} A_k^{-1} - \mathrm{tr} A_0^{-1})\right| \mathcal{B}_{nk} \ge \varepsilon\right) = o(n^{-t}).$$

PROOF. Recall that $\mathrm{tr}A_k^{-1} - \mathrm{tr}A_0^{-1} = \gamma_k' A_k^{-2} \gamma_k \beta_k$. Then the lemma is proved by applying Lemma 3.4 with $h = 2$ and the fact that $n^{-2}\mathrm{tr}A_k^{-2}\mathcal{B}_{nk} \leq a^{-2}n^{-1}$.

By a similar approach, we can verify

Lemma 3.6 *We have*

$$\mathrm{P}\left(\frac{1}{n}|E_j\mathrm{tr}(A_j^{-1}\breve{A}_j^{-1}) - E_j\mathrm{tr}(A_{jl}^{-1}\breve{A}_{jl}^{-1})|\mathcal{B}_{njl} \geq \varepsilon\right) \leq o(n^{-t}))$$

for any positive number t.

Lemma 3.7 *We have*

$$\frac{1}{n}(\mathrm{tr}A_0^{-1}) - \frac{q}{1-q} \to 0, \quad a.s. \quad E\frac{1}{1+\frac{1}{n}(\mathrm{tr}A_0^{-1})} - (1-q) = O(n^{-1/2}).$$

This lemma follows from the work of Bai and Silverstein (2010, p. 278).

Lemma 3.8 *We have $\gamma' A_0^{-1}\gamma \to \rho^2/(1-q)$ almost surely.*

The conclusion remains true when A is replaced by A_j for any j, and the convergence is uniform in $j \leq n$.

Lemma 3.9 *For any $\varepsilon > 0$, we have $\mathrm{P}\left(\left|\gamma_{k_1}' A_{k_1,k_2}^{-1}\gamma_{k_2}\right| \mathcal{B}_{nk_1k_2} \geq \varepsilon\right) = o(n^{-t})$.*

PROOF. This is an easy consequence of Corollary 3.1. ■

Lemma 3.10 *For any nonrandom vector b and integer k, we have*

$$E\left(\left|s_k\gamma_k' b - \frac{1}{n}\gamma' b\right|^2\right) \leq \frac{K}{n^2}b'b, \quad E\left(\left|s_k\gamma_k' b - \frac{1}{n}\gamma' b\right|^l\right) \leq Kn^{-1}\eta_n^{(2l-4)}\gamma' b$$

PROOF. Note that $s_k\gamma_k' b = w_k' D_2' b d_1 w_k$ and $\gamma' b = \mathrm{tr}(D_2' b d_1)$, by the covariance formula of two quadratic forms (see (1.15) of Bai and Silverstein (2004))

$$E\left(\left|s_k\gamma_k' b - \frac{1}{n}\gamma' b\right|^2\right) \leq \frac{K}{n^2}\{b'b + (\gamma' b)^2\} \leq \frac{K}{n^2}b'b.$$

Moreover, $E\left(\left|s_k\gamma_k' b - \frac{1}{n}\gamma'\mathbf{b}\right|^l\right) \leq Kn^{-1}\eta_n^{(2l-4)}\gamma'\mathbf{b}$ is easily obtained by Lemma 3.1. ■

Now, we analyze the quantity $C_j' A_j^{-1}\breve{A}_j^{-1}\breve{C}_j$.

Lemma 3.11 *For any j, we have*

$$E_j(C_j'A_j^{-1}\breve{A}_j^{-1}\breve{C}_j) = \left(\frac{(1-q)^2(j-1)}{n^2} - \frac{2(1-q)^2\rho^2(n-j)(j-1)}{n^3}\right)E_j\{\mathrm{tr}(A_j^{-1}\breve{A}_j^{-1}\cdot\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj})\}$$
$$+(1-q)^2\gamma'E_j(A_j^{-1}\breve{A}_j^{-1}\gamma\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj})$$
$$-\frac{2(1-q)^3(j-1)^2}{n^3}E_j\{\gamma'A_j^{-1}\gamma\mathrm{tr}(\breve{A}_j^{-1}A_j^{-1}\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj})\} + o_p(1) \tag{3.24}$$

where $o_p(1)$ is uniform in j.

PROOF. We first consider

$$\sum_{k>j}E_j(s_k\gamma_k'A_j^{-1}\breve{A}_j^{-1}\breve{\gamma}_k\breve{s}_k\mathcal{B}_{jk}\breve{\mathcal{B}}_{jk}) = \frac{(1-q)^2(n-j)}{n^2}\gamma'E_j(A_j^{-1}\breve{A}_j^{-1}\gamma\mathcal{B}_j\breve{\mathcal{B}}_j) + o_p(1) = o_p(1)$$

because $s_k\gamma_k$ is independent of A_{kj}, $\breve{A}_{kj}$ and $\breve{\gamma}_k\breve{s}_k$ and $\beta_{k(j)} = (1-q)+o_p(n^{-t})$, $\breve{\beta}_{k(j)} = (1-q) = o_p(n^{-t})$ hold uniformly.

Next, using the Substitution Lemma 3.3, we obtain

$$\sum_{k<j}s_k\gamma_k'E_j(A_j^{-1}\breve{A}_j^{-1}\breve{\gamma}_k\breve{s}_k\mathcal{B}_{njk}\breve{\mathcal{B}}_{njk}) = \frac{(1-q)^2(j-1)}{n^2}\{\mathrm{tr}E_j(A_j^{-1}\breve{A}_j^{-1}\mathcal{B}_{njk}\breve{\mathcal{B}}_{njk})\} + o_p(1) \tag{3.25}$$

because $P(|\gamma_k'(E_jA_j^{-1}\breve{A}_j^{-1})\breve{\gamma}_k - \frac{1}{n}\mathrm{tr}E_jA_j^{-1}\breve{A}_j^{-1}| \geq \varepsilon) = o(n^{-t})$ and $P(||s_k|^2 - \frac{1}{n}| \geq \varepsilon) = o(n^{-t})$ for any positive number t by Lemma 3.4 and Lemma 3.6.

Third, we can similarly prove that

$$\sum_{k\neq l>j}E_j(s_k\gamma_k'A_j^{-1}\breve{A}_j^{-1}\breve{\gamma}_l\breve{s}_l\mathcal{B}_{njk}\breve{\mathcal{B}}_{njl}) = \frac{(1-q)^2(n-j)^2}{n^2}\gamma'\{E_j(A_j^{-1}\breve{A}_j^{-1})\}\gamma\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj} + o_p(1) \tag{3.26}$$

by Lemma 3.6. Also, we have

$$\sum_{k>j>l}E_j(s_k\gamma_k'A_j^{-1}\breve{A}_j^{-1}\breve{\gamma}_l\breve{s}_l\mathcal{B}_{njk}\breve{\mathcal{B}}_{njl}) = -\frac{(1-q)^2\rho^2(n-j)(j-1)}{n^3}E_j\left\{\mathrm{tr}(A_j^{-1}\breve{A}_j^{-1}\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj})\right\} +$$
$$\frac{(1-q)^2(n-j)(j-1)}{n^2}\gamma'A_j^{-1}\breve{A}_j^{-1}\gamma\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj} + o_p(1) \tag{3.27}$$

because

$$P\left\{\left|\gamma'E_j(A_{jl}^{-1}\breve{A}_{jl}^{-1}\gamma_l s_l) - \frac{1}{n}\gamma'E_j(A_{jl}^{-1}\breve{A}_{jl}^{-1}\gamma)\right| \geq \varepsilon\right\} = o(n^{-t})$$

and

$$\gamma' E_j(A_{jl}^{-1}\breve{A}_{jl}^{-1}\gamma_l s_l) = x_k' \left\{ \frac{1}{n} D_k' E_j(A_{jl}^{-1}\breve{A}_{jl}^{-1}\gamma d_1) \right\} x_k$$

and by Lemma 3.4. Similarly, we have

$$\sum_{l>j>k} E_j(s_k\gamma_k' A_j^{-1}\breve{A}_j^{-1}\breve{\gamma}_l\breve{s}_l\mathcal{B}_{njk}\breve{\mathcal{B}}_{njl}) = -\frac{(1-q)^2\rho^2(n-j)(j-1)}{n^3} E_j\left\{\mathrm{tr}(A_j^{-1}\breve{A}_j^{-1}\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj})\right\} + \frac{(1-q)^2(n-j)(j-1)}{n^2}\gamma' A_j^{-1}\breve{A}_j^{-1}\gamma\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj} + o_p(1) \quad (3.28)$$

Finally, we have

$$\sum_{k\neq l<j} s_k\gamma_k^{'}(E_j A_j^{-1}\breve{A}_j^{-1})\breve{\gamma}_l\breve{s}_l\mathcal{B}_{njk}\breve{\mathcal{B}}_{njl} = \frac{(1-q)^2(j-1)^2}{n^2} E_j(\gamma' A_j^{-1}\breve{A}_j^{-1}\gamma\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}) - \frac{2(1-q)^3(j-1)^2}{n^3} E_j(\gamma' A_j^{-1}\gamma\mathrm{tr}\breve{A}_j^{-1}A_j^{-1}\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}). \quad (3.29)$$

Summing up the estimates (3.25)–(3.29), we obtain

$$\begin{aligned}
& E_j\Big(C_j A_j^{-1}\breve{A}_j^{-1}\breve{C}_j\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}\Big) \\
= & \left\{\frac{(1-q)^2(j-1)}{n^2} - \frac{2(1-q)^2\rho^2(n-j)(j-1)}{n^3}\right\} E_j\{\mathrm{tr}(A_j^{-1}\breve{A}_j^{-1}\cdot\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj})\} \\
& +(1-q)^2\cdot\gamma' E_j(A_j^{-1}\breve{A}_j^{-1}\gamma\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}) \\
& -\frac{2(1-q)^3(j-1)^2}{n^3} E_j(\gamma' A_j^{-1}\gamma)\mathrm{tr}(\breve{A}_j^{-1}A_j^{-1}\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}) + o_p(1). \qquad (3.30)
\end{aligned}$$

The proof of Lemma 3.11 is completed. ■

Lemma 3.12 *We have*

$$\mathrm{tr}(\mathbf{A}_j^{-1}\breve{\mathbf{A}}_j^{-1}\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}) = \frac{nq/(1-q)^2 + o_p(n)}{1-q(j-1)/n} \qquad (3.31)$$

where $o_p(n)$ is uniform in j.

PROOF. Let

$$b_j = \frac{1}{1+\frac{1}{n}E\{\mathrm{tr}(A_j^{-1}\mathcal{B}_{nj})\}}, \qquad \mathbf{B}_{1j} = -\sum_{k\neq j} A_{kj}^{-1}\left(\gamma_k\gamma_k' - \frac{1}{n}\mathbf{I}\right)\mathcal{B}_{njk},$$

$$\mathbf{B}_{2j} = -b_j^{-1}\sum_{k\neq j}(\beta_{k(j)} - b_j)A_{kj}^{-1}\gamma_k\gamma_k'\mathcal{B}_{njk}, \qquad \mathbf{B}_{3j} = -n^{-1}\sum_{k\neq j}(A_{kj}^{-1} - A_j^{-1})\mathcal{B}_{njk}.$$

Note that

$$\frac{n-1}{n}A_j^{-1}\mathcal{B}_j \;\;=\;\; b_j^{-1}\mathbf{I}+\mathbf{B}_{1j}+\mathbf{B}_{2j}+\mathbf{B}_{3j}, \tag{3.32}$$

where then we have

$$\begin{aligned} E\{|\mathrm{tr}(\mathbf{B}_{2j}\breve{A}_j^{-1}\breve{\mathcal{B}}_{nj})|\} &\leq \sum\nolimits_{k\neq j}\{E(|\beta_{kj}b_j^{-1}-1|^2\mathcal{B}_{njk})\}^{1/2}\{E(|\gamma_k'\mathbf{A}_{kj}^{-1}\breve{A}_j^{-1}\gamma_k|^2\breve{\mathcal{B}}_{nj}\mathcal{B}_{njk})\}^{1/2} \\ &= K\sum\nolimits_{k\neq j}\left\{E\left|\gamma_k'A_{jk}^{-1}\gamma_k-\frac{1}{n}E\mathrm{tr}A_j^{-1}\right|^2\mathcal{B}_{njk}\right\}^{1/2}\{E(|\gamma_k'\gamma_k|^2)\}^{1/2}\leq Kn^{1/2}. \end{aligned} \tag{3.33}$$

Using $A_j^{-1}=A_{jk}^{-1}-A_{jk}^{-1}\gamma_k\gamma_k'A_{jk}^{-1}/(1+\gamma_k'A_{jk}^{-1}\gamma_k)$, we similarly have

$$E\{|\mathrm{tr}(\mathbf{B}_{3j}\breve{A}_j^{-1}\breve{\mathcal{B}}_{nj})|\}\leq K. \tag{3.34}$$

Now we consider $\mathrm{tr}(\mathbf{B}_{1j}\breve{A}_j^{-1}\breve{\mathcal{B}}_{nj})$. We have

$$E\left(\left|\mathrm{tr}\sum\nolimits_{k>j}\left(\gamma_k\gamma_k'-\frac{1}{n}\mathbf{I}\right)\mathbf{A}_{kj}^{-1}\mathcal{B}_{njk}\breve{A}_j^{-1}\breve{\mathcal{B}}_{njk}\right|^2\right)=E\left(\left|\sum_{k>j}\left(\gamma_k'A_{jk}^{-1}\breve{A}_j\gamma_k-\frac{1}{n}\mathrm{tr}(A_{jk}^{-1}\breve{A}_j)\right)\right|^2\right)=o(1)$$

where the last step can be proved by expanding the square of the sum. The squares terms can be treated easily, the estimation of the cross terms needs one more expansion $A_{jk}^{-1}=A_{njkl}^{-1}-A_{njkl}^{-1}\gamma_l\gamma_l'A_{njkl}^{-1}\beta_{l(jk)}$. The details are omitted. The readers are referred to Bai and Silverstein (2010) if they are interested the details of the proof. Consequently, we have

$$\mathrm{tr}(\mathbf{B}_{1j}\breve{A}_j^{-1}\breve{\mathcal{B}}_{nj}) \;\;=\;\; C_{1j}+C_{2j}+C_{3j}++o_p(1), \tag{3.35}$$

where

$$C_{1j}=-\sum\nolimits_{k<j}(\gamma_k'\mathbf{A}_{kj}^{-1}\mathcal{B}_{njk})(\breve{A}_j^{-1}-\breve{A}_{kj}^{-1})\breve{\mathcal{B}}_{njk}\gamma_k,$$

$$C_{2j}=\frac{1}{n}\sum\nolimits_{k<j}\{\mathrm{tr}(\mathbf{A}_{kj}^{-1}\mathcal{B}_{njk})\}(\breve{\mathbf{A}}_j^{-1}-\breve{\mathbf{A}}_{kj}^{-1})\breve{\mathcal{B}}_{njk}\gamma_k,$$

$$C_{3j}=-\sum\nolimits_{k<j}\{\mathrm{tr}(\mathbf{A}_{kj}^{-1}\mathcal{B}_{njk})\}\left(\gamma_k\gamma_k'-\frac{1}{n}\mathbf{I}\right)\breve{\mathbf{A}}_{kj}^{-1}\breve{\mathcal{B}}_{njk}.$$

It is not difficult to show that

$$E(|C_{3j}|)\leq Kn^{1/2},\quad E(|C_{2j}|)\leq K. \tag{3.36}$$

Further expanding C_{1j}, we have

$$\begin{aligned}C_{1j} &= \sum\nolimits_{k<j}(\gamma_k'\mathbf{A}_{kj}^{-1}\mathcal{B}_{njk})(\breve{\mathbf{A}}_{kj}^{-1}\gamma_k\gamma_k'\breve{\mathbf{A}}_{kj}^{-1}\beta_{kj}\breve{\mathcal{B}}_{njk})\gamma_k \\ &= \frac{1}{n^2}\sum\nolimits_{k<j}\mathrm{tr}(\mathbf{A}_{kj}^{-1}\mathcal{B}_{njk})(\breve{\mathbf{A}}_{kj}^{-1}\mathrm{tr}\breve{\mathbf{A}}_{kj}^{-1}\breve{\beta}_{kj}\breve{\mathcal{B}}_{njk})+O(1).\end{aligned}$$

Using $n^{-1}\mathrm{tr}\mathbf{A}_{kj}^{-1}=q/(1-q)+o_p(1), b_j=(1-q)+o_p(1)$, we obtain

$$\begin{aligned}C_{1j} &= \frac{q}{n}\sum\nolimits_{k<j}\mathrm{tr}(\mathbf{A}_{kj}^{-1}\mathcal{B}_{njk})(\breve{\mathbf{A}}_{kj}^{-1}\breve{\mathcal{B}}_{njk})+O_p(1) \\ &= \frac{q(j-1)}{n}\mathrm{tr}(\mathbf{A}_j^{-1}\mathcal{B}_{nj})(\breve{\mathbf{A}}_j^{-1}\breve{\mathcal{B}}_{nj})+O_p(1).\end{aligned} \tag{3.37}$$

Summing up the estimates (3.33)–(3.37), we obtain

$$\mathrm{tr}(\mathbf{B}_{1j}+\mathbf{B}_{2j}+\mathbf{B}_{3j})\breve{A}_j^{-1}\breve{\mathcal{B}}_{nj}=\frac{q(j-1)}{n}\mathrm{tr}A_j^{-1}\breve{A}_j^{-1}\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}+O_p(n^{1/2}) \tag{3.38}$$

Recall the definition of the matrices $\mathbf{B}_{1j}, \mathbf{B}_{2j}$, and $\mathbf{B}_{3j}$, we find that

$$\begin{aligned}\mathbf{B}_{1j}+\mathbf{B}_{2j}+\mathbf{B}_{3j} &= \frac{1}{n}b_j^{-1}\sum_{k\neq j}\mathcal{B}_{njk}-b_j^{-1}A_j^{-1}\sum_{k\neq j}\gamma_k\gamma_k'\mathcal{B}_{njk} \\ = \frac{n-1}{n}(A_j^{-1}-b_j^{-1}\mathbf{I})\mathcal{B}_{nj} &- \frac{1}{n}b_j^{-1}\sum_{k\neq j}\mathcal{B}_{njk}^c\mathcal{B}_{nj}+b_j^{-1}A_j^{-1}\sum_{k\neq j}\gamma_k\gamma_k'\mathcal{B}_{njk}^c\mathcal{B}_{nj}.\end{aligned}$$

By this relation, it is easy to show that

$$\begin{aligned}\text{LHS of (3.38)} &= \frac{n-1}{n}\mathrm{tr}(A_j^{-1}\mathcal{B}_{nj}\breve{A}_j^{-1}\breve{\mathcal{B}}_{nj})-\frac{n-1}{nb_j}\mathrm{tr}(\breve{A}_j^{-1}\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}) \\ &= A_j^{-1}\mathcal{B}_{nj}\breve{A}_j^{-1}\breve{\mathcal{B}}_{nj}-\frac{nq}{(1-q)^2}+o_p(n).\end{aligned}$$

Therefore,

$$\mathrm{tr}(\mathbf{A}_j^{-1}\breve{\mathbf{A}}_j^{-1}\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj})=\frac{nq/(1-q)^2+o_p(n)}{1-q(j-1)/n}.$$

The proof of Lemma 3.12 is completed. ■

Lemma 3.13 *We have*

$$\gamma' A_j^{-1}\breve{A}_j^{-1}\gamma\mathcal{B}_{njk}\breve{\mathcal{B}}_{njk} = \frac{\rho^2}{(1-q)^2\{1-q(j-1)/n\}}+o_p(1). \tag{3.39}$$

PROOF. We use again the decomposition (3.32) and obtain that

$$E\left(\left|\gamma'\mathbf{B}_{3j}\breve{A}_j^{-1}\gamma\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}\right|\right) \;\leq\; \frac{K}{n^{1/2}} \to 0. \tag{3.40}$$

Next, we shall prove that

$$E\left(\left|\gamma'\mathbf{B}_{2j}\breve{A}_j^{-1}\gamma\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}\right|\right) \;\leq\; K/n^{1/2}. \tag{3.41}$$

To this end, one needs to notice that

$$E\left(|\beta_{k(j)} - b_j|^2\mathcal{B}_{njk}\right) = O(1/n), \quad E\left(|\gamma' A_{kj}^{-1}\gamma_k\mathcal{B}_{njk}|^4\right) = O(1/n^2).$$

At the third step, we estimate

$$\gamma'\mathbf{B}_{1j}\breve{A}_j^{-1}\gamma\mathcal{B}_{nj}\breve{\mathcal{B}}_{nj} \;=\; J_{1j} + J_{2j} + J_{3j} + o_p(1),$$

where

$$\begin{aligned}
J_{1j} &= \sum_{k<j}\gamma' A_{kj}^{-1}\left(\gamma_k\gamma_k' - \frac{1}{n}\mathbf{I}\right)\breve{A}_{kj}^{-1}\gamma\mathcal{B}_{njk}\breve{\mathcal{B}}_{njk},\\
J_{2j} &= -\frac{1}{n}\sum_{k<j}\gamma' A_{kj}^{-1}(\breve{A}_j^{-1} - \breve{A}_{kj}^{-1})\gamma\mathcal{B}_{njk}\breve{\mathcal{B}}_{njk},\\
J_{3j} &= \sum_{k<j}\gamma' A_{kj}^{-1}\gamma_k\gamma_k'(\breve{A}_j^{-1} - \breve{A}_{kj}^{-1})\gamma\mathcal{B}_{njk}\breve{\mathcal{B}}_{njk}.
\end{aligned}$$

By elementary calculation, one concludes that $E|J_{1j}|^2 \leq K/n$. (Where the squares can be estimated by simple calculation, the cross products need one more expansion.)

Furthermore, $E|J_{2j}| \leq K/n$ and $E|J_{3j}| = \rho^2/[(1-q)^2\{1 - q(j-1)/n\}] + o_p(1)$.

The proof of the lemma is completed. ■

Lemma 3.14 *We have $E_j\{(C_j' A_j^{-1}\gamma)^2\} = (\rho^2)^2 + o_p(1)$ where $o_p(1)$ is uniform for j.*

Proof. We have

$$\begin{aligned}
&E_j\{(C_j' A_j^{-1}\gamma)^2\} \qquad (3.42)\\
=\;& \sum\nolimits_{k=l>j} E_j(s_k\gamma_k' A_j^{-1}\gamma\cdot\gamma' A_j^{-1}\gamma_l s_l) + \sum\nolimits_{k=l<j} E_j(s_k\gamma_k' A_j^{-1}\gamma\cdot\gamma' A_j^{-1}\gamma_l s_l)\\
&+ \sum\nolimits_{k>j>l} E_j(s_k\gamma_k' A_j^{-1}\gamma\cdot\gamma' A_j^{-1}\gamma_l s_l) + \sum\nolimits_{l>j>k} E_j(s_k\gamma_k' A_j^{-1}\gamma\cdot\gamma' A_j^{-1}\gamma_l s_l)\\
&+ \sum\nolimits_{k\neq l>j} E_j(s_k\gamma_k' A_j^{-1}\gamma\cdot\gamma' A_j^{-1}\gamma_l s_l) + \sum\nolimits_{k\neq l<j} E_j(s_k\gamma_k' A_j^{-1}\gamma\cdot\gamma' A_j^{-1}\gamma_l s_l) \qquad (3.43)
\end{aligned}$$

Because $\gamma_k' A_j^{-1} = \gamma_k' A_{kj}^{-1} \beta_{k(j)}, A_j^{-1}\gamma_l = A_{lj}^{-1}\gamma_l\beta_{l(j)}$, we obtain $\sum_{k=l>j} E_j(s_k\gamma_k' A_j^{-1}\gamma\gamma' A_j^{-1}\gamma_l s_l)$ $= O_p(n^{-1})$ and $\sum_{k=l<j} E_j(s_k\gamma_k' A_j^{-1}\gamma\gamma' A_j^{-1}\gamma_l s_l \mathcal{B}_{nj}\breve{\mathcal{B}}_{nj}) = O_p(n^{-1})$. Moreover, $\sum_{k>j>l} E_j(s_k\gamma_k' A_j^{-1}\gamma \cdot \gamma' A_j^{-1}\gamma_l s_l) = (n-j)(j-1)(\rho^2)^2/n^2 + o_p(1)$ and

$$\sum\nolimits_{k\neq l>j} E_j(s_k\gamma_k' A_j^{-1}\gamma\gamma' A_j^{-1}\gamma_l s_l) = \frac{(n-j)^2(\rho^2)^2}{n^2} + o_p(1) \tag{3.44}$$

$$\sum\nolimits_{k\neq l<j} E_j(s_k\gamma_k' A_j^{-1}\gamma\gamma' A_j^{-1}\gamma_l s_l) = \frac{(j-1)^2(\rho^2)^2}{n^2} + o_p(1). \tag{3.45}$$

Then we have $E_j\{(C_j' A_j^{-1}\gamma)^2\} = (\rho^2)^2 + o_p(1)$.

Then the proof of Lemma 3.14 is completed. ■

Lemma 3.15 *We have*

$$\frac{1}{n}\sum\nolimits_{j=1}^{n} E_j(C_j' A_j^{-1}\breve{A}_j^{-1}\breve{C}_j) = 2\rho^2 - 1 - (1-\rho^2)\frac{\log(1-q)}{q} + o_p(1).$$

Proof. By Lemmas 3.11, 3.31 and 3.13, we have

$$\begin{aligned}
&\frac{1}{n}\sum\nolimits_{j=1}^{n} C_j' A_j^{-1}\breve{A}_j^{-1}\breve{C}_j \\
=\ & \frac{1}{n}\sum\nolimits_{j=1}^{n}\left((1-q)^2\frac{j-1}{n} - 2(1-q)^2\rho^2\left(1-\frac{j-1}{n}\right)\frac{j-1}{n}\right)\cdot\frac{q}{(1-q)^2}\frac{1}{1-q\frac{j-1}{n}} \\
& + \frac{1}{n}\sum\nolimits_{j=1}^{n}\cdot\frac{\rho^2}{(1-\frac{q(j-1)}{n})} - \frac{1}{n}\sum\nolimits_{j=1}^{n}\frac{2(1-q)^3(j-1)^2}{n^2}\frac{\rho^2}{1-q}\cdot\frac{q}{(1-q)^2}\frac{1}{1-\frac{q(j-1)}{n}} + o_p(1) \\
=\ & \int_0^1 \left(x - 2\rho^2(1-x)x\right)\cdot\frac{q}{1-qx}\,\mathrm{d}x + \int_0^1 \frac{\rho^2}{1-qx}\,\mathrm{d}x - \int_0^1 \frac{2\rho^2 q\cdot x^2}{1-qx} + o_p(1) \\
=\ & 2\rho^2 - 1 - (1-\rho^2)\frac{\log(1-q)}{q} + o_p(1)
\end{aligned} \tag{3.46}$$

Then the proof of Lemma 3.15 is completed. ■

Lemma 3.16 *We have* $n^{-1}\sum_{j=1}^{n} E_j(C_j' A_j^{-1}\gamma) = \rho^2 + o_p(1)$.

Proof. We have

$$
\begin{aligned}
& E_j(C_j'A_j^{-1}\gamma) \\
= \; & \sum\nolimits_{k\neq j} E_j(\gamma_k' s_k A_j^{-1}\gamma) = \sum\nolimits_{k\neq j} E_j(\gamma_k' s_k A_{kj}^{-1}\gamma) - \sum\nolimits_{k\neq j} E_j(\gamma_k' s_k A_{kj}^{-1}\gamma_k\gamma_k' A_{kj}^{-1}\gamma\beta_{k(j)}) \\
= \; & \frac{1}{n}\sum\nolimits_{k\neq j} \gamma' E_j(A_{kj}^{-1}\gamma) - \sum\nolimits_{k\neq j} E_j(\gamma_k' s_k A_{kj}^{-1}\gamma_k\gamma_k' A_{kj}^{-1}\gamma\beta_{k(j)}) + o_p(1)\text{(by Lemma 3.3)} \\
= \; & \frac{1}{n}\sum\nolimits_{k\neq j} \gamma' E_j(A_{kj}^{-1}\gamma) - (1-q)\sum\nolimits_{k>j} E_j\left(\gamma_k' A_{kj}^{-1}\gamma_k - \frac{1}{n}\mathrm{tr}A_{kj}^{-1}\right)\cdot\left(\gamma_k' A_{kj}^{-1}\gamma s_k - \frac{1}{n}\gamma' A_{kj}^{-1}\gamma\right) \\
& -\sum\nolimits_{k>j}\frac{1-q}{n^2}E_j(\mathrm{tr}A_{kj}^{-1}\gamma' A_{kj}^{-1}\gamma) - (1-q)\sum\nolimits_{k<j} E_j(\gamma_k' A_{kj}^{-1}\gamma_k\gamma_k' A_{kj}^{-1}\gamma s_k) + o_p(1) \\
= \; & \gamma' E_j(A_j^{-1}\gamma) - \frac{2(1-q)(n-j)}{n^2}E_j(\gamma')E_j(A_j^{-2}\gamma) - \frac{(1-q)(n-j)}{n^2}E_j\{\mathrm{tr}(A_j^{-1}\gamma' A_j^{-1})\gamma\} \\
& -(1-q)\sum\nolimits_{k<j} E_j(\gamma_k' A_{kj}^{-1}\gamma_k\gamma_k' A_{kj}^{-1}\gamma s_k) + o_p(1) \\
= \; & \rho^2 + o_p(1).
\end{aligned} \tag{3.47}
$$

Then we obtain $n^{-1}\sum_{j=1}^n E_j(C_j'A_j^{-1}\gamma) = \rho^2 + o_p(1)$.

Then the proof of Lemma 3.16 is completed.

Lemma 3.17 *We have* $n^{-1}\sum_{j=1}^n\{E_j(\gamma' A_j^{-1}C_j)\}^2 = (\rho^2)^2 + o_p(1)$.

Lemma 3.18 *We have*

$$
n^{-1}\sum\nolimits_{j=1}^n \mathrm{tr}\left\{E_j(A_j^{-1}C_jC_j'A_j^{-1}\mathcal{B}_{nj})\right\}^2 = (2\rho^2-1)^2 + \frac{(1-\rho^2)^2}{1-q} - 2(2\rho^2-1)(1-\rho^2)\frac{\log(1-q)}{q} + o_p(1)
$$

Proof. We have

$$
\begin{aligned}
& \frac{1}{n}\sum\nolimits_{j=1}^n \mathrm{tr}\left\{E_j(A_j^{-1}C_jC_j'A_j^{-1}\mathcal{B}_{nj})\right\}^2 = \frac{1}{n}\sum\nolimits_{j=1}^n\left\{E_j(C_j'A_j^{-1}\mathcal{B}_{nj}\breve{A}_j^{-1}\breve{C}_j\breve{\mathcal{B}}_{nj})\right\}^2 \\
= \; & \frac{1}{n}\sum\nolimits_{j=1}^n\left[\left(\frac{(1-q)^2(j-1)}{n} - \frac{2(1-q)^2\rho^2(n-j)(j-1)}{n^2}\right)\frac{q/(1-q)^2}{1-q(j-1)/n}\right. \\
& \left. + \frac{\rho^2}{\{1-q(j-1)/n\}} - \frac{2(1-q)^3(j-1)^2}{n^2}\frac{\rho^2}{1-q}\frac{q}{(1-q)^2}\frac{1}{1-q(j-1)/n}\right]^2 + o_p(1) \\
= \; & \int_0^1\left\{(1-2\rho^2)q\cdot\frac{x}{1-qx} + \rho^2\cdot\frac{1}{1-qx}\right\}^2\mathrm{d}x + o_p(1) \\
= \; & (2\rho^2-1)^2 + \frac{(1-\rho^2)^2}{1-q} - 2(2\rho^2-1)(1-\rho^2)\frac{\log(1-q)}{q} + o_p(1).
\end{aligned} \tag{3.48}
$$

Then the proof of Lemma 3.18 is completed. ■

Lemma 3.19 *We have*

$$\left\{\gamma' E_j(A_j^{-1}C_jC_j'A_j^{-1}\mathcal{B}_{nj})\right\}\left\{E_j(A_j^{-1}C_j\mathcal{B}_{nj})\right\} = \rho^2\int_0^1\left\{(1-2\rho^2)\frac{qx}{1-qx}+\frac{\rho^2}{1-qx}\right\}\mathrm{d}x + o_p(1).$$

Proof. We have

$$\begin{aligned}
&\left\{\gamma' E_j(A_j^{-1}C_jC_j'A_j^{-1}\mathcal{B}_{nj})\right\}\left\{E_j(A_j^{-1}C_j\mathcal{B}_{nj})\right\} = \frac{1}{n}\sum\nolimits_{j=1}^n E_j(C_j'A_j^{-1}\breve{A}_j^{-1}\breve{C}_j\breve{C}_j'\breve{A}_j^{-1}\gamma)\\
= &\ \frac{1}{n}\sum\nolimits_{j=1}^n\Bigg[\left\{\frac{(1-q)^2(j-1)}{n}-\frac{2(1-q)^2\rho^2(n-j)(j-1)}{n^2}\right\}\frac{q/(1-q)^2}{1-q(j-1)/n}+\frac{\rho^2}{(1-q(j-1)/n)}\\
&\qquad -\frac{2(1-q)^3(j-1)^2}{n^2}\frac{\rho^2}{1-q}\cdot\frac{q}{(1-q)^2}\frac{1}{1-q(j-1)/n}\Bigg]\cdot\breve{C}_j'\breve{A}_j^{-1}\gamma + o_p(1)\\
= &\ \int_0^1\left\{(1-2\rho^2)q\cdot\frac{x}{1-qx}+\rho^2\cdot\frac{1}{1-qx}\right\}\mathrm{d}x\cdot\rho^2+o_p(1)\\
= &\ \left\{2\rho^2-1-(1-\rho^2)\frac{\log(1-q)}{q}\right\}\rho^2.
\end{aligned}\tag{3.49}$$

Then the proof of Lemma 3.19 is completed.

Lemma 3.20 *We have* $n^{-1}\sum_{j=1}^n\{E_j(\gamma' A_j^{-1}C_j)\}^2 = (\rho^2)^2 + o_p(1)$.

This follows from by Lemma 3.14. ■

4 Simulation results

For the same simulation model used in Section 3 of the main paper except that X is normally distributed, we have the results in Table 1^*. For the hypothesis testing problem of $H_0:\rho^2=0$, we plot the power functions in Figure 1 when $q = 0{\cdot}2, 0{\cdot}6, 0{\cdot}8$ and $p=60$ when X_i are f

f

ly distributed.

Below we consider two alternative simulation models by using different D in the model. The first one takes the form

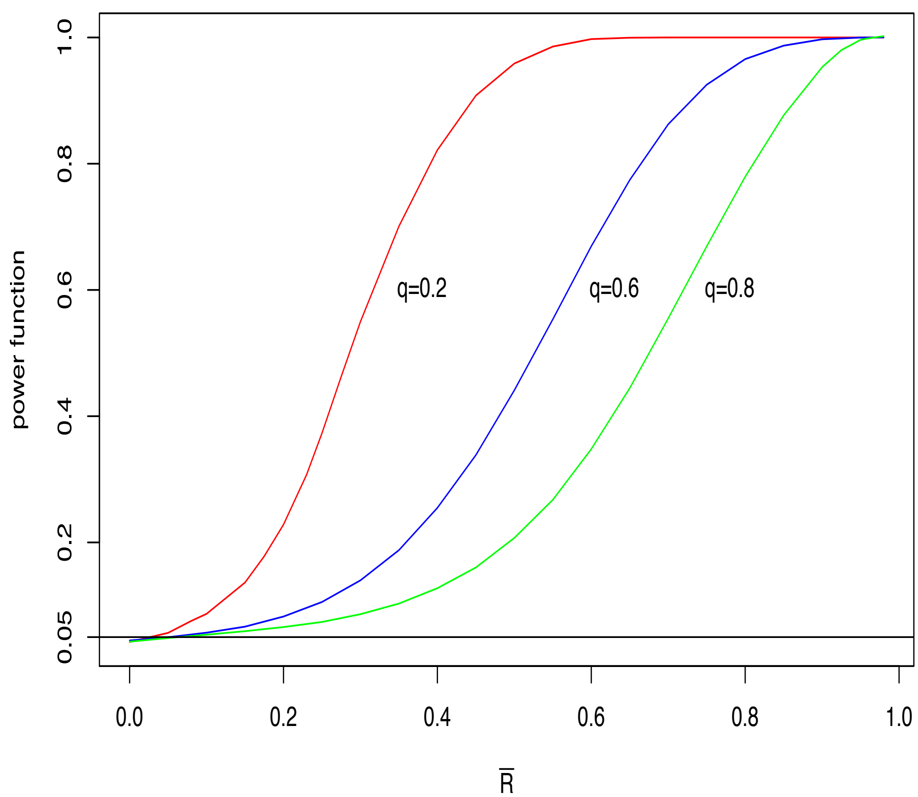

Figure 1: Power functions for testing problem $H_0 : \rho^2 = 0$ with $p = 60$.

$$\begin{pmatrix} d_1 \\ D_2 \end{pmatrix} = \Sigma^{\frac{1}{2}} = \left((r^{|i-j|})_{i,j}\right)^{\frac{1}{2}}$$

where r is chosen to satisfy $\rho = 0{\cdot}2, 0{\cdot}6, 0{\cdot}8$ at $q = 0{\cdot}2, 0{\cdot}8$ and $p = 20, 30, 60, 200$. The simulation results under this model are given in Table 2. The second one takes the form

$$\begin{pmatrix} d_1 \\ D_2 \end{pmatrix} = \Sigma^{\frac{1}{2}} = \left((r^{|i-j|} I_{|i-j|\le 1})_{i,j}\right)^{\frac{1}{2}}$$

where r is chosen to satisfy $\rho = 0{\cdot}2, 0{\cdot}35, 0{\cdot}5$ for different $q = 0{\cdot}2, 0{\cdot}5$ and $p = 20, 30, 60$. The simulation results under this model are listed in Table 3.

Table 2 here.

Table 3 here.

Finally, we examine the Q–Q plots for the simulated p values for testing the hypothesis $H_0 : \rho^2 = 0{\cdot}6^2$ relative to the uniform distribution on (0,1). In the case of $q = 0{\cdot}2, p = 60$ and $\mathbf{\Sigma}$ as defined in Section 3 of the main paper, we have Figure 2 for the tests based on several competing methods, including several variations of the bootstrap methods. The p values from our proposed method (M2) are distributed as expected, but the other methods give p values that are far from the uniform distribution.

Table 1^*. Average coverage (percentages) and length (in parentheses) of 95% confidence intervals for ρ^2 when the X_{ij} are normally distributed.

p \ ρ		$q = 0{\cdot}2$		
		0·2	0·6	0·8
20	M1	95·6 (0·20)	94·6 (0·33)	94·4 (0·26)
	M2	97·8 (0·22)	94·8 (0·34)	95·1 (0·25)
	M3	69·3 (0·40)	91·0 (0·40)	93·8 (0·26)
	A1	91·2 (0·06)	0·0 (0·07)	0·0 (0·07)
	B1	49·9 (0·06)	70·8 (0·31)	77·9 (0·25)
30	M1	95·4 (0·17)	94·8 (0·27)	94·8 (0·21)
	M2	97·9 (0·18)	95·0 (0·28)	95·3 (0·20)
	M3	46·9 (0·35)	85·0 (0·35)	90·3 (0·23)
	A1	72·3 (0·04)	0·0 (0·05)	0·0 (0·05)
	B1	34·2 (0·04)	59·7 (0·26)	70·2 (0·21)
60	M1	95·3 (0·13)	95·2 (0·20)	94·5 (0·15)
	M2	97·9 (0·14)	94·9 (0·20)	95·2 (0·14)
	M3	5·5 (0·26)	62·0 (0·27)	77·4 (0·18)
	A1	0·0 (0·02)	0·0 (0·03)	0·0 (0·03)
	B1	8·8 (0·01)	33·5 (0·19)	48·9 (0·15)

p \ ρ		$q = 0{\cdot}8$		
		0·2	0·6	0·8
20	M1	95·0 (0·55)	93·4 (0·69)	93·3 (0·83)
	M2	94·6 (0·77)	94·1 (0·78)	94·0 (0·75)
	M3	0·5 (0·60)	2·3 (0·43)	4·3 (0·25)
	A1	96·1 (0·21)	0·0 (0·24)	0·0 (0·26)
	B1	28·5 (0·14)	30·1 (0·20)	31·6 (0·21)
60	M1	95·2 (0·43)	94·8 (0·61)	94·5 (0·68)
	M2	95·8 (0·57)	95·9 (0·63)	94·8 (0·54)
	M3	0·0 (0·45)	0·0 (0·31)	0·0 (0·18)
	A1	91·2 (0·07)	0·0 (0·09)	0·0 (0·09)
	B1	10·8 (0·03)	12·2 (0·08)	12·6 (0·10)
200	M1	95·1 (0·30)	95·2 (0·46)	95·0 (0·33)
	M2	96·2 (0·37)	94·6 (0·45)	94·6 (0·30)
	M3	0·0 (0·30)	0·0 (0·22)	0·0 (0·13)
	A1	0·0 (0·02)	0·0 (0·03)	0·0 (0·03)
	B1	1·5 (0·01)	1·6 (0·02)	1·0 (0·04)

M1 and M2, asymptotic confidence intervals based on Theorem 2; M3, Edgeworth expansion method of Ogasawara (2006); A1, the classical method of Anderson (2003); B1, ABC bootstrap method.

Table 2. Average coverage (percentages) and length (in parentheses) of 95% confidence intervals for ρ^2 when the X_{ij} are uniformly distributed.

p \ ρ		$q = 0{\cdot}2$		
		0·2	0·6	0·8
20	M1	94·8 (0·20)	93·7 (0·33)	94·2 (0·26)
	M2	97·8 (0·22)	95·1 (0·34)	94·6 (0·25)
	M3	69·9 (0·40)	91·2 (0·40)	94·6 (0·26)
	A1	91·3 (0·06)	0·0 (0·07)	0·0 (0·07)
	B1	46·9 (0·06)	67·8 (0·33)	77·0 (0·27)
30	M1	95·2 (0·17)	95·9 (0·27)	94·9 (0·21)
	M2	97·6 (0·18)	95·1(0·28)	95·0 (0·20)
	M3	45·8 (0·35)	85·1(0·35)	90·6 (0·23)
	A1	73·0 (0·04)	0·0(0·05)	0·0 (0·04)
	B1	32·1 (0·03)	58·2 (0·28)	69·3 (0·22)
60	M1	94·2 (0·13)	95·4 (0·20)	94·9 (0·14)
	M2	97·6 (0·14)	94·8 (0·20)	95·1 (0·14)
	M3	5·6 (0·27)	63·3 (0·27)	78·4 (0·18)
	A1	0·0 (0·02)	0·0 (0·02)	0·0 (0·02)
	B1	7·5 (0·01)	31·1 (0·21)	48·0 (0·16)

p \ ρ		$q = 0{\cdot}8$		
		0·2	0·6	0·8
20	M1	95·0 (0·59)	93·4 (0·70)	91·4 (0·28)
	M2	94·0 (0·76)	93·7 (0·78)	94·3 (0·75)
	M3	0·4 (0·62)	2·1 (0·43)	5·5 (0·26)
	A1	96·5 (0·20)	0·0 (0·23)	5·5 (0·26)
	B1	22·0 (0·11)	24·5 (0·17)	28·0 (0·20)
60	M1	95·3 (0·43)	94·4 (0·60)	93·0 (0·68)
	M2	95·6 (0·57)	95·9 (0·63)	94·1 (0·54)
	M3	0·0 (0·45)	0·0 (0·31)	0·0 (0·19)
	A1	90·8 (0·07)	0·0 (0·09)	0·0 (0·09)
	B1	9·9 (0·02)	10·1 (0·03)	12·7 (0·12)
200	M1	95·6 (0·31)	96·1 (0·61)	94·8 (0·33)
	M2	97·1 (0·37)	94·8 (0·45)	94·6 (0·30)
	M3	0·0 (0·30)	0·0 (0·21)	0·0 (0·13)
	A1	0·0 (0·02)	0·0 (0·03)	0·0 (0·03)
	B1	1·0 (0·01)	1·3 (0·03)	1·1 (0·06)

M1 and M2, asymptotic confidence intervals based on Theorem 2; M3, Edgeworth expansion method of Ogasawara (2006); A1, the classical method of Anderson (2003); B1, ABC bootstrap method.

Table 3. Average coverage (percentages) and length (in parentheses) of 95% confidence intervals for ρ^2 when the X_{ij} are uniformly distributed.

p \ ρ		$q = 0{\cdot}2$: 0·2	0·35	0·5
20	M1	94·5 (0·20)	96·1(0·27)	95·3 (0·32)
	M2	98·1 (0·22)	97·6 (0·28)	95·8 (0·34)
	M3	69·0 (0·40)	81·2(0·44)	88·5(0·44)
	A1	92·0 (0·06)	0·0 (0·07)	0·0 (0·07)
	B1	46·7 (0·06)	56·1 (0·15)	64·5 (0·27)
30	M1	96·2 (0·17)	95·2 (0·24)	94·8 (0·27)
	M2	97·9 (0·18)	98·1 (0·24)	94·8 (0·28)
	M3	45·8 (0·35)	66·5 (0·38)	79·7 (0·38)
	A1	72·0 (0·04)	0·0 (0·04)	0·0 (0·05)
	B1	28·1 (0·03)	40·5 (0·11)	53·0 (0·23)
60	M1	94·3 (0·13)	96·1 (0·18)	94·4 (0·20)
	M2	98·0 (0·14)	94·9 (0·19)	94·7 (0·20)
	M3	5·3(0·26)	24·7(0·28)	50·1(0·29)
	A1	0·0 (0·02)	0·0 (0·02)	0·0 (0·02)
	B1	7·5 (0·01)	14·4 (0·06)	24·4 (0·18)
p \ ρ		$q = 0{\cdot}5$: 0·2	0·35	0·5
20	M1	95·5 (0·38)	72·7 (0·32)	96·2 (0·50)
	M2	96·9 (0·48)	97·3 (0·51)	96·8 (0·56)
	M3	26·7 (0·76)	35·1 (0·73)	45·0 (0·67)
	A1	96·2 (0·13)	84·2 (0·15)	0·0 (0·16)
	B1	22·0 (0·11)	23·5 (0·12)	22·6 (0·14)
30	M1	94·3 (0·33)	95·6 (0·38)	95·5 (0·46)
	M2	97·2 (0·41)	96·9 (0·45)	97·5 (0·50)
	M3	6·5 (0·66)	12·8 (0·64)	20·7 (0·58)
	A1	94·0 (0·09)	0·0 (0·10)	0·0 (0·11)
	B1	16·9 (0·07)	17·7 (0·08)	17·4 (0·10)
60	M1	94·3 (0·26)	94·1 (0·33)	95·6 (0·39)
	M2	97·4 (0·31)	97·2 (0·35)	94·2 (0·41)
	M3	0·0 (0·49)	0·2 (0·48)	1·1 (0·44)
	A1	81·0 (0·05)	0·0 (0·05)	0·0 (0·06)
	B1	9·5 (0·03)	10·2 (0·04)	11·5 (0·05)

M1 and M2, asymptotic confidence intervals based on Theorem 2; M3, Edgeworth expansion method of Ogasawara (2006); A1, the classical method of Anderson (2003); B1, ABC bootstrap method.

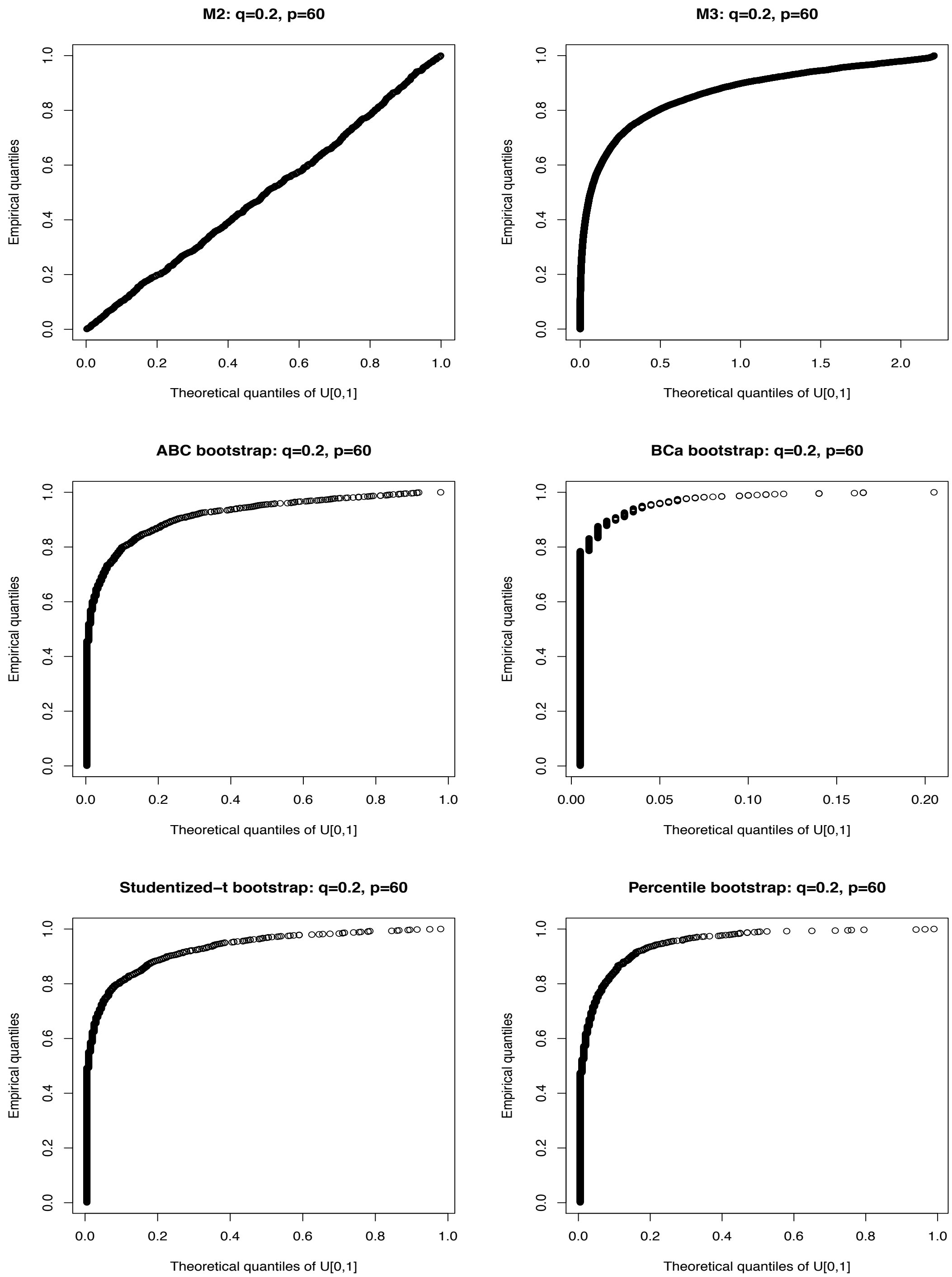

Figure 2: Q–Q plots for simulated p values against the uniform distribution on (0,1) for several methods. M2 is based on our proposed method, M3 refers to the Edgeworth expansion method, and the other bootstrap methods are self-explanatory.