An Exact Test about the Covariance Matrix

Arjun K. Gupta[a,*], Taras Bodnar[b]

[a]*Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403, USA*
[b]*Department of Mathematics, Humboldt-University of Berlin, D-10099 Berlin, Germany*

Abstract

In the present paper, we propose an exact test on the structure of the covariance matrix. In its development the properties of the Wishart distribution are used. Unlike the classical likelihood-ratio type tests and the tests based on the empirical distance, whose statistics depend on the total variance and the generalized variance only, the proposed approach provides more information about the changes in the covariance matrix. Via an extensive simulation study the new approach is compared with the existent asymptotic tests.

Keywords: covariance matrix, Wishart distribution, inference procedure
2010 MSC: 62H15, 62H10, 62E15, 62H12, 62F03

---

1. Introduction

Tests about the covariance matrix have significantly increased its popularity recently. Historically, the first test on the covariance matrix was suggested by Mauchly (1940) that is based on the likelihood ratio approach. Because the statistic of this test depends on the determinant and the trace of the sample covariance matrix, the so-called generalized and total variances respectively, it requires that the sample covariance matrix is non-singular which is the case with probability one when the sample size is larger than the process dimension. Gupta and Xu (2006) extended the likelihood-ratio test to non-normal distributions by deriving the asymptotic expansion of the test statistic under the null hypothesis, while Bai et al. (2009) considered a modification of the likelihood-ratio test. The second approach considered

*Corresponding author. Phone: +419372 2820. Email: gupta@bgsu.edu

in statistical literature is based on the empirical distance initially suggested by John (1971) and Nagao (1974). These test statistics with some modifications can also be applied for testing the covariance matrix in case of high-dimensional data (cf. Ledoit and Wolf (2002), Chen at al. (2010)) even when the sample size is smaller than the process dimension. Other approaches are based on the largest eigenvalue of the covariance matrix (Johnstone (2001, 2008)) or they are derived by using the methods of random matrix theory (cf. Cai and Jiang (2011)).

In this paper we derive an exact test about the covariance matrix. In the development of this test the properties of the Wishart distribution are used. Since an exact test is developed it always correctly sized. Moreover, the suggested test can also be applied if the sample size is much smaller than the dimension of the process. Via an extensive simulation study we show that the new approach performs very well if changes in a few elements of the covariance matrix take places.

The rest of the paper is structured as follows. In Section 2, we introduce a test about the covariance matrix. The distribution of the test statistic is derived under both the null and alternative hypothesis. In Section 3, an extension of the test is provided. A very useful stochastic representation of the test statistic is obtained under H_0 which shows that under the null hypothesis the distribution is independent of the target matrix specified under H_0. In Section 4, the distributional properties of the test statistic under H_1 are studied via an extensive Monte-Carlo study. Some proofs are given in the appendix (Section 5).

2. Test based on a column of the covariance matrix

Let $\mathbf{X}_1, ..., \mathbf{X}_n \sim ii\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $p > 2$, be an independent sample from the multivariate normal distribution with known mean vector $\boldsymbol{\mu}$. Without loss of generality we assume that $\boldsymbol{\mu} = \mathbf{0}_p$, where $\mathbf{0}_p$ stands for the p-dimensional vector of zeroes. The covariance matrix $\boldsymbol{\Sigma}$ is estimated by

$$\mathbf{S} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i' \,. \tag{1}$$

If $\boldsymbol{\mu}$ is an unknown quantity then instead of (1) we use the sample covariance matrix for estimating $\mathbf{\Sigma}$ expressed as

$$\tilde{\mathbf{S}} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})' \quad \text{with} \quad \bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i \,. \tag{2}$$

The two approaches differ only slightly from each other since it holds that (see, e.g. Muirhead (1982, p. 90))

$$\mathbf{S} \sim W_p(n, \mathbf{\Sigma}) \quad \text{and} \quad \tilde{\mathbf{S}} \sim W_p(n-1, \mathbf{\Sigma})\,,$$

where the symbol $W_p(n, \mathbf{\Sigma})$ stands for the p-dimensional Wishart distribution with n degrees of freedom and covariance matrix $\mathbf{\Sigma}$ (cf. Muirhead (1982), Srivastava (2003)). Moreover, both estimators are unbiased as well as asymptotically normally distributed (Muirhead (1982, p. 90-91)). The last result is usually used for the derivation of asymptotic tests on the covariance matrix.

In this paper we consider an alternative approach that is based on the distributional properties of the Wishart distribution and the singular Wishart distribution (Srivastava (2003), Bodnar and Okhrin (2008)). First, an exact test is proposed which is based on a column of the sample covariance matrix and then it is generalized. In the derivation, no assumption on p, like $n \geq p$, is imposed. The results hold in all possible cases, i.e. for $n \geq p$ and $n < p$. While the properties of the Wishart distribution are applied for $n \geq p$, we make use of the distributional results derived for the singular Wishart distribution in the case of $n < p$.

We assume that $\mathbf{\Sigma} = \mathbf{\Sigma}_0$ under H_0 and $\mathbf{\Sigma} = \mathbf{\Sigma}_1$ under H_1. The matrices $\mathbf{\Sigma}_0$, $\mathbf{\Sigma}_1$, and $\mathbf{S}$ are partitioned as follows

$$\mathbf{\Sigma}_0 = \begin{bmatrix} \xi_0 & \boldsymbol{\nu}_0' \\ \boldsymbol{\nu}_0 & \mathbf{\Xi}_0 \end{bmatrix}, \mathbf{\Sigma}_1 = \begin{bmatrix} \xi_1 & \boldsymbol{\nu}_1' \\ \boldsymbol{\nu}_1 & \mathbf{\Xi}_1 \end{bmatrix}, \text{ and } \mathbf{S} = \begin{bmatrix} v & \mathbf{t}' \\ \mathbf{t} & \mathbf{W} \end{bmatrix}. \tag{3}$$

Let $\mathbf{\Upsilon}_0 = \mathbf{\Xi}_0 - \boldsymbol{\nu}_0\boldsymbol{\nu}_0'/\xi_0$ and $\mathbf{\Upsilon}_1 = \mathbf{\Xi}_1 - \boldsymbol{\nu}_1\boldsymbol{\nu}_1'/\xi_1$. Without loss of generality we now present a test based on the first column of the covariance matrix $\mathbf{\Sigma}_0$. In case of the i-th column the test statistic can be derived similarly. Here, instead of the partitions (3), we construct the partition for the (i,i)-th element of the matrices $\mathbf{\Sigma}_0$, $\mathbf{\Sigma}_1$, and $\mathbf{S}$ as follows. Let $\xi_{0,i}$ denote the (i,i)-th element of the matrix $\mathbf{\Sigma}_0$, $i = 1, ..., p$. By $\boldsymbol{\nu}_{0,i}$ we denote the i-th column of the matrix $\mathbf{\Sigma}_0$ without $\xi_{0,i}$. Let $\mathbf{\Xi}_{0,i}$ denote a square matrix of order $p-1$, which is obtained from the matrix $\mathbf{\Sigma}_0$ by deleting the i-th row and the i-th column. Finally, $\mathbf{\Upsilon}_{0,i} = \mathbf{\Xi}_{0,i} - \boldsymbol{\nu}_{0,i}\boldsymbol{\nu}_{0,i}'/\xi_{0,i}$ is calculated. In the same way we define $\xi_{1,i}$, $\boldsymbol{\nu}_{1,i}$, $\boldsymbol{\Xi}_{1,i}$, $\boldsymbol{\Upsilon}_{1,i}$, v_i, $\mathbf{t}_i$, and $\mathbf{W}_i$ by splitting $\boldsymbol{\Sigma}_1$ and $\mathbf{S}$ correspondingly. For presentation purposes we drop the index i in the notations if $i = 1$.

The hypotheses are given by

$$H_0: \ \boldsymbol{\Sigma} = d\boldsymbol{\Sigma}_0 \qquad \text{against} \qquad H_1: \ \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1 \neq d\boldsymbol{\Sigma}_0 \,, \tag{4}$$

where $d > 0$ denotes an arbitrary (un)known constant. We define

$$\boldsymbol{\eta}_1 = \sqrt{n}\boldsymbol{\Upsilon}_0^{-1/2}\left(\frac{\mathbf{t}}{v} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right)v^{1/2} \tag{5}$$

Let $\phi_k(.;\boldsymbol{\mu},\boldsymbol{\Sigma})$ stand for the density function of the k-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. In Theorem 1 we derive the distributions of the random vector $\boldsymbol{\eta}_1$ under both H_0 and H_1 hypotheses.

Theorem 1. *a) Let* $\mathbf{X}_i \sim iid\mathcal{N}_p(\mathbf{0}_p, \boldsymbol{\Sigma}_1)$, $i = 1, ..., n$. *Then the density function of* $\boldsymbol{\eta}_1$ *is given by*

$$f_{\boldsymbol{\eta}_1}(\mathbf{x}) = 2\frac{\sqrt{\pi}n^{(n-1)/2}}{2^{(n-1)/2}\xi_1^{(n-1)/2}\Gamma\left(\frac{n}{2}\right)}\phi_{p-1}\left(\mathbf{x};\mathbf{0}_{p-1},\boldsymbol{\Omega}+\xi_1\boldsymbol{\Delta}\boldsymbol{\Delta}'\right) \tag{6}$$
$$\times \quad \int_0^\infty y^{n-1}\phi_1\left(y;\frac{\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\mathbf{x}}{\sqrt{n}(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})}, n^{-1}(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})^{-1}\right)dy\,,$$

where

$$\boldsymbol{\Delta} = \boldsymbol{\Upsilon}_0^{-1/2}\left(\frac{\boldsymbol{\nu}_1}{\xi_1} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right) \qquad \text{and} \qquad \boldsymbol{\Omega} = \boldsymbol{\Upsilon}_0^{-1/2}\boldsymbol{\Upsilon}_1\boldsymbol{\Upsilon}_0^{-1/2}\,.$$

b) Let $\mathbf{X}_i \sim iid\mathcal{N}_p(\mathbf{0}_p, \boldsymbol{\Sigma}_0)$, $i = 1, ..., n$. *Then* $\boldsymbol{\eta}_1 \sim \mathcal{N}_{p-1}(\mathbf{0}_{p-1}, \mathbf{I}_{p-1})$, *where* $\mathbf{I}_k$ *denotes a* $k \times k$ *identity matrix.*

The part a) of Theorem 1 is proved in the appendix. The part b) follows directly from the part a) by noting that $\boldsymbol{\Delta} = \mathbf{0}_{p-1}$ and $\boldsymbol{\Omega} = \mathbf{I}_{p-1}$ in this case as well as using the fact that

$$\frac{n^{n/2}}{2^{n/2}\xi_1^{n/2}\Gamma\left(\frac{n}{2}\right)}t^{n-1}\exp\left(-\frac{n}{2\xi_1}t^2\right)$$

is the density function of a squared gamma-distributed random variable.

For testing (4), we introduce the test statistic given by

$$T_1 = \boldsymbol{\eta}_1' \boldsymbol{\eta}_1 \,, \tag{7}$$

which is motivated by the distributional properties of $\boldsymbol{\eta}_1$. The distribution of T_1 is derived in Theorem 2.

Theorem 2. *a) Let* $\mathbf{X}_i \sim iid\mathcal{N}_p(\mathbf{0}_p, \boldsymbol{\Sigma}_1)$, $i = 1, ..., n$. *Let* $\boldsymbol{\Omega} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'$ *be the eigenvalue decomposition of* $\boldsymbol{\Omega}$ *where* $\boldsymbol{\Lambda} = diag(\lambda_1, ..., \lambda_{p-1})$ *is the diagonal matrix of eigenvalues and* $\mathbf{P}$ *is the corresponding orthogonal matrix of eigenvectors. Then the distribution function of* T_1 *is given by*

$$F_{T_1}(x) = \frac{1}{2} - \frac{n^{n/2}}{2^{n/2}\xi_1^{n/2}\pi\Gamma\left(\frac{n}{2}\right)} \int_0^\infty \int_0^\infty y^{n/2-1} \exp\left(-\frac{n}{2}\left(\frac{y}{\xi_1}\right)\right) \frac{\sin\theta(u,y)}{u\rho(u,y)} du dy \,, \tag{8}$$

where

$$\mathbf{b}(y) = (b_1(y), ..., b_{p-1}(y))' = \sqrt{n}\mathbf{P}'\boldsymbol{\Omega}^{-1/2}\boldsymbol{\Delta}\sqrt{y} \,, \tag{9}$$

$$\theta(u,y) = \frac{1}{2}\sum_{j=1}^{p-1}\left(\frac{1}{\tan(\lambda_j u)} + \frac{b_j(y)^2\lambda_j u}{1+\lambda_j^2 u^2}\right) - \frac{ux}{2} \,, \tag{10}$$

$$\rho(u,y) = \prod_{j=1}^{p-1}(1+\lambda_j^2 u^2)^{1/4} \exp\left(\frac{1}{2}\frac{b_j(y)^2\lambda_j^2 u^2}{1+\lambda_j^2 u^2}\right) . \tag{11}$$

$$\tag{12}$$

b) Let $\mathbf{X}_i \sim iid\mathcal{N}_p(\mathbf{0}_p, \boldsymbol{\Sigma}_0)$, $i = 1, ..., n$. *Then* $T_1 \sim \chi^2_{p-1}$.

The proof of Theorem 2a) is given in the appendix. The part b) follows directly from Theorem 1a). Under an additional assumption imposed on $\boldsymbol{\Omega}$ the density of T_1 can be presented in an analytical form. This result is formulated as Corollary 1 below.

In the following we also use the generalized hypergeometric function (cf. Muirhead (1982, p. Ch. 1.3)), i.e.

$${}_pF_q(a_1, a_2, ..., a_p; c_1, c_2, ..., c_q; x) = \frac{\prod_{l=1}^q \Gamma(c_l)}{\prod_{j=1}^p \Gamma(a_j)} \sum_{i=0}^\infty \frac{\prod_{j=1}^p \Gamma(a_j + i)}{\prod_{l=1}^q \Gamma(c_l + i)} \frac{z^i}{i!} \,.$$

The technical computation of a hypergeometric function is a standard routine within many mathematical software packages like, e.g., in Mathematica. In the statement of Corollary 1 we make use of ${}_1F_1(.;.;.)$, while in its proof (see appendix) ${}_0F_1(.;.)$ is used.

Corollary 1. *Under the assumptions of Theorem 2 if* $\mathbf{\Omega} = \mathbf{I}_{p-1}$ *then the density function of* T_1 *is given by*

$$f_{T_1}(x) = \frac{1}{(1+\lambda)^{n/2}} f_{p-1}(x)\ {}_1F_1\left(\frac{n}{2}; \frac{p-1}{2}; \frac{\lambda x}{2(1+\lambda)}\right) \tag{13}$$

where f_k *denotes the density of the* χ_k^2*-distribution with* k *degrees of freedom and*

$$\lambda = \xi_1 \mathbf{\Delta}'\mathbf{\Delta} = \xi_1 \left(\frac{\boldsymbol{\nu}_1}{\xi_1} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right)' \mathbf{\Upsilon}_0^{-1} \left(\frac{\boldsymbol{\nu}_1}{\xi_1} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right), \tag{14}$$

which is a measure of the population distance between $\boldsymbol{\nu}_1/\xi_1$ *and* $\boldsymbol{\nu}_0/\xi_0$.

The proposed test on the covariance matrix possesses several advantages. First, the test statistic (7) has a classical distribution under H_0. Hence, the test is easy to perform by comparing the values of the test statistic with the quantile of the χ^2-distribution. Second, if $\mathbf{\Omega} = \mathbf{I}_{p-1}$ then the result of Corollary 1 allows us to study the power of the suggested test, which appears to be a function of one parameter λ only.

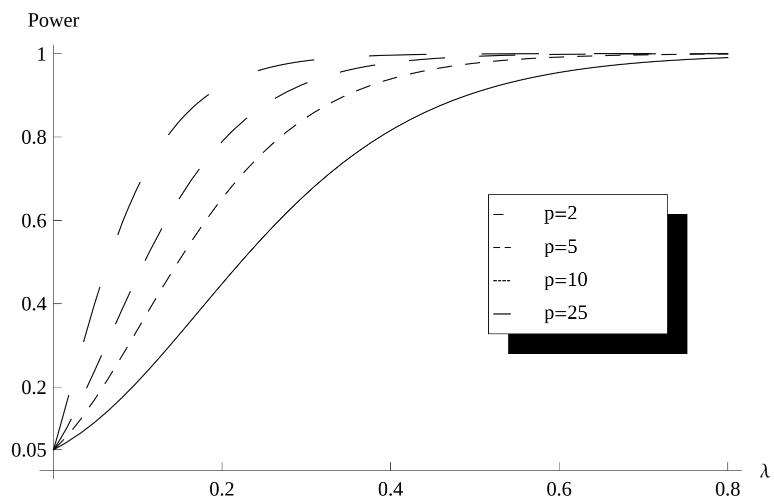

Figure 1: Probability of rejection of the null hypothesis as a function of λ ($\alpha = 0.05$) for $p \in \{2, 5, 10, 25\}$, $n = 60$.

In Figures 1 and 2, we plot the power function of the test T_1 in case of $\mathbf{\Omega} = \mathbf{I}_{p-1}$ as a function of λ for different values of $p \in \{2, 5, 10, 25\}$ as well as

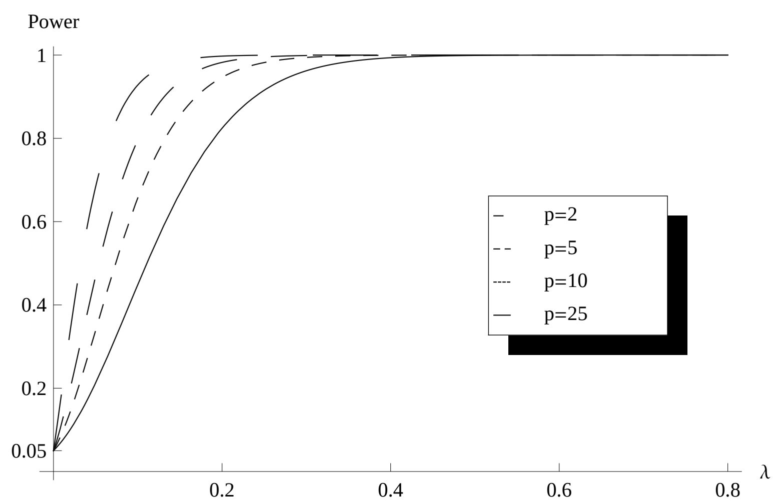

Figure 2: Probability of rejection of the null hypothesis as a function of λ ($\alpha = 0.05$) for $p \in \{2, 5, 10, 25\}$, $n = 120$.

$n = 60$ (Figure 1) and $n = 120$ (Figure 2). Note that the power function is a decreasing function of p for a fixed value of λ and an increasing function with respect to n. It is quite large in the case $p = 2$ for both values of n, while for larger values of p the power becomes smaller. On the other hand we note that the parameter λ is, usually, larger for larger values of p.

3. Test on the covariance matrix

In this section we extend the results of the previous section by taking the maximum over the individual test statistics T_j's calculated for the jth column of the sample covariance matrix. Namely, for testing (4) we consider the following test statistic expressed as

$$T = \max_{j \in \{1,\dots,p\}} \{T_j\} \,. \tag{15}$$

The application of the test statistic (15) is motivated by the observation that T_j's are influenced in different ways by different changes in the covariance matrix. In order to ensure that similar changes in the columns of the covariance matrix are treated in the same way, the maximum of the individual test statistics is calculated.

In the following theorem we derive a very important stochastic representation of the test statistic (15) under H_0 which appears to be independent of $\mathbf{\Sigma}_0$.

Theorem 3. *Let* $\mathbf{X}_i \sim iid\mathcal{N}_p(\mathbf{0}_p, \mathbf{\Sigma}_0)$*,* $i = 1, ..., n$*,* $\mathbf{U}_i \sim ii\mathcal{N}_p(\mathbf{0}_p, \mathbf{I}_p)$*,* $i = 1, ..., n$*, and let*

$$\mathbf{S}_{\mathbf{U}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{U}_i \mathbf{U}_i' . \tag{16}$$

Then under H_0 *the test statistic* T *is equal in distribution to*

$$T \overset{d}{=} \max_{j \in \{1,...,p\}} \left\{ n \left(\frac{\mathbf{S}'_{\mathbf{U};j}\mathbf{S}_{\mathbf{U};j}}{s_{\mathbf{U};jj}} - s_{\mathbf{U};jj} \right) \right\} , \tag{17}$$

where $\mathbf{S}_{\mathbf{U};j}$ *is the* j*-th column of* $\mathbf{S}_{\mathbf{U}}$ *and* $s_{\mathbf{U};jj}$ *is the* j*-th diagonal element of* $\mathbf{S}_{\mathbf{U}}$*.*

The proof of Theorem 3 is given in the appendix. Later on, the distribution of the test statistic T under H_0 we denote by $T(p, n)$, whereas the symbol $T_\beta(p, n)$ stands for its β-quantile. The null hypothesis in (4) is rejected if $T > T_{1-\alpha}(p, n)$. The result of Theorem 3 allows us to derive the critical value for the test $T_{1-\alpha}(p, n)$ which appears to be independent of $\mathbf{\Sigma}_0$ and, consequently, it can be applied for different choices of $\mathbf{\Sigma}_0$. It is remarkable to note that the critical value of the test depends only on the dimension of $\mathbf{X}_i$ and the sample size n.

Next, we study this point in detail. In Table 1, we present the upper quantiles $T_{1-\alpha}(p, n)$ of the distribution $T(p, n)$ for different values of p and n. These values are calculated by applying the result of Theorem 3 and they are based on the 10^5 independently simulated random vectors $\mathbf{U}_i$. We observe that the values of n have only a minor impact on the corresponding quantiles for fixed p.

In Table 2 we provide further analysis of this observation. Here, the estimated probabilities $P_{H_0}(T > T_{1-\alpha}(p, 30))$ (the first and third rows in each panel) are presented as well as the relative differences (the second and fourth row in each panel) calculated as

$$r_{p,\alpha} = \frac{\alpha - P_{H_0}(T > T_{1-\alpha}(p, 30))}{\alpha} .$$

We observe that all estimated probabilities are very close to the corresponding significant levels α. The relative differences are also small with highest

Table 1: The estimated quantiles $T_{1-\alpha}(p, n)$ of the distribution $T(p, n)$ for $n = 30$ (the first row in each panel), $n = 60$ (the second row in each panel), and $n = 120$ (the third row in each panel) in case of $\alpha \in \{0.1, 0.05, 0.01\}$ and $p \in \{2, 4, 6, 8, 10, 12, 14\}$.

$\alpha \backslash p$	2	4	6	8	10	12	14	
	3.081197	8.711363	13.34424	17.30745	21.06636	24.59540	28.00077	$n = 30$
0.1	3.016870	8.725965	13.21182	17.27303	21.20263	24.69749	28.17860	$n = 60$
	2.873459	8.540466	13.13826	17.35949	20.99666	24.81323	28.24585	$n = 120$
	4.286687	10.427932	15.06843	19.36323	23.18504	26.90867	30.26823	$n = 30$
0.05	4.279034	10.275211	14.91522	19.26061	23.22965	27.36723	30.56959	$n = 60$
	4.223091	10.251880	15.01471	19.34528	23.15833	26.84275	30.60427	$n = 120$
	7.266405	14.218312	18.91408	23.70290	27.37060	31.43515	35.24717	$n = 30$
0.01	7.184371	13.840018	19.08365	23.47222	27.68378	31.63213	35.46858	$n = 60$
	7.121917	14.018406	18.92412	23.46961	27.85884	31.64072	35.58516	$n = 120$

values achieved for $\alpha = 0.01$ with maximum smaller than 0.2. This observation shows that the critical values are not significantly influenced by the sample size if it is large enough compared to p. For instance, even for 10-dimensional case it is enough to calculate the critical values for $n = 30$ and use them for larger values of n.

Table 2: The estimated probabilities $P_{H_0}(T > T_{1-\alpha}(p, 30))$ (the first and third rows in each panel) and the relative differences $r_{p,\alpha} = (\alpha - P_{H_0}(T > T_{1-\alpha}(p, 30)))/\alpha$ (the second and fourth row in each panel) for $n \in \{60, 120\}$, $\alpha \in \{0.1, 0.05, 0.01\}$, and $p \in \{2, 4, 6, 8, 10, 12, 14\}$.

$\alpha \backslash p$	2	4	6	8	10	12	14	
0.1	0.0963	0.0986	0.0935	0.0972	0.0989	0.1065	0.1049	n=60
	0.037	0.014	0.065	0.028	0.011	-0.065	-0.049	n=60
	0.0833	0.0945	0.0964	0.0993	0.1051	0.1023	0.1060	n=120
	0.167	0.055	0.036	0.007	-0.051	-0.023	-0.060	n=120
0.05	0.0481	0.0489	0.0496	0.0499	0.0462	0.0475	0.0507	n=60
	0.038	0.022	0.008	0.002	0.076	0.050	-0.014	n=60
	0.0489	0.0467	0.0541	0.0473	0.0502	0.0488	0.0488	n=120
	0.022	0.066	-0.082	0.054	-0.004	0.024	0.024	n=120
0.01	0.0100	0.0081	0.0103	0.0084	0.0106	0.0112	0.0102	n=60
	0.0	0.190	-0.030	0.160	-0.060	-0.120	-0.020	n=60
	0.0085	0.0086	0.0117	0.0091	0.0115	0.0114	0.0102	n=120
	0.150	0.140	-0.170	0.090	-0.150	-0.140	-0.020	n=120

4. Simulation study

In this section we study the size and power of the test (15) and compare it with the existent asymptotic tests for the covariance matrix. It is noted that both n and p can be considerably large, i.e. $p \to \infty$ and $n \to \infty$. Moreover, we impose no restriction on the ratio p/n. Whenever it is necessary the limit $p/n \to c \in [0, \infty)$ is used for $n, p \to \infty$.

4.1. Benchmark asymptotic tests

For testing the hypothesis

$$H_0 : \mathbf{\Sigma} = \sigma^2 \mathbf{I} \qquad \text{against} \qquad H_0 : \mathbf{\Sigma} \neq \sigma^2 \mathbf{I} \tag{18}$$

several asymptotic tests are suggested in literature. Concerning the approach used in their derivation, the tests can be divided in three groups. The first group consists of the tests that are obtained from the application of the likelihood-ratio approach. The empirical distance method is implied in the derivation of the tests from the second group, while the testing procedures from the third group are based on the methods of the random matrix theory. Below, we briefly review the most important tests from each group and study their size. In the next section, the power functions are evaluated and compared with each other and with the derived exact test of Section 3.

Historically, the first test for the hypotheses (18) was suggested by Mauchly (1940), whose test statistic is given by

$$\tilde{T}_{LR} = n\left(p \ln|\sigma^2| - \ln|\mathbf{S}| + tr(\mathbf{S}/\sigma^2) - p\right) . \tag{19}$$

For the values of n that are of moderate size a further modification of the test statistic (19) is expressed as (see, e.g., Rencher (2002, Ch. 7.2))

$$T_{LR} = \left(1 - \frac{1}{6n-1}\left(2p + 1 - \frac{2}{p+1}\right)\right) T_{LR} . \tag{20}$$

Both the test statistics $\tilde{T}_{LR}$ and T_{LR} are asymptotically $\chi^2_{p(p+1)/2}$-distributed under H_0. The exact moments of these test statistics are given by Muirhead (1982, Ch. 8.4).

Bai et al. (2009) suggested a further modification of the likelihood ratio test for the case when the dimension is large compared to the sample size which is derived by using the central limit theorem for linear spectral statistics of sample covariance matrices (cf. Bai and Silverstein (2004)). The test statistic of this test is given by

$$T_{BJYZ} = v(c)^{-1/2} \left(\frac{\tilde{T}_{LR}}{n} - pF^{y_n}(c) - m(c) \right) , \tag{21}$$

with $m(c) = -ln(1-c)/2$, $v(c) = -2ln(1-c) - 2c$ where c is defined at the beginning of Section 3. The symbol $F^{y_n}(.)$ denotes the Marčenko-Pastur law of index $y_n = p/n$ (cf. Bai et al. (2009)). Under the null hypothesis, the test statistic (21) has the standard normal distribution.

The main disadvantage of the tests based on the likelihood approach is that they are based on the determinant of the sample covariance matrix and as a result they can only be used if $n > p$ since in the opposite case the sample covariance matrix is singular. In order to overcome this difficulty, alternative approaches are considered in literature which are based on the empirical distance between the sample covariance matrix and the corresponding target value. For testing (18) John (1971) suggested

$$T_{Jn} = \frac{1}{p} tr \left[\left(\frac{\mathbf{S}}{(1/p)tr(\mathbf{S})} - \mathbf{I} \right)^2 \right] . \tag{22}$$

Ledoit and Wolf (2002) showed that the asymptotic distribution of T_{Jn}, which is the $\chi^2_{p(p+1)/2-1}$-distribution, is also valid in the case of the (p, n)-asymptotic, i.e. when $p \to \infty$ and $n \to \infty$. In contrast to this result, Ledoit and Wolf (2002) proved that the asymptotic distribution of the Nagao's (1973) statistic in terms of (p, n)-asymptotic does not coincide with the asymptotic distribution in case of the n-asymptotic.

A further test for the high-dimensional covariance matrix is discussed in Srivastava (2005) with the statistic given by

$$T_S = \frac{n^2}{(n-1)(n+2)} \frac{1}{p} \left[tr(\mathbf{S}^2) - \frac{1}{n} (tr\mathbf{S})^2 \right] / (tr\mathbf{S}/p)^2 - 1 \, , \tag{23}$$

while Fisher et al. (2010) considered

$$T_{FSG} = \frac{n}{\sqrt{8(8+12c+c^2)}} \tag{24}$$

$$\times \left\{ \frac{\tau}{p} \frac{[tr\mathbf{S}^4 + btr\mathbf{S}^3 tr\mathbf{S} + c^*(tr\mathbf{S}^2)^2 + dtr\mathbf{S}^2(tr\mathbf{S})^2 + e(tr\mathbf{S})^4]}{\frac{n^4}{(n-1)^2(n+2)^2} \frac{1}{p^2} \left[tr\mathbf{S}^2 - \frac{1}{n}(tr\mathbf{S})^2 \right]^2} - 1 \right\} ,$$

where $p/n \to c$ and the constants τ, b, c^*, d, and e are defined in Theorem 1 of Fisher et al. (2010).

Let $\sum^*$ denote the summation over mutually different indices (see Chen et al. (2010, p. 811)) and let

$$
\begin{aligned}
Y_1 &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i\,, & Y_2 &= \frac{1}{P_n^2}\sum_{i\neq j} \mathbf{X}_i'\mathbf{X}_j\,,\\
Y_3 &= \tfrac{1}{P_n^2}\textstyle\sum_{i\neq j} \mathbf{X}_i'\mathbf{X}_i\mathbf{X}_j'\mathbf{X}_j\,, & Y_4 &= \frac{1}{P_n^2}\sum_{i\neq j} (\mathbf{X}_i'\mathbf{X}_j)^2\,,\\
Y_5 &= \tfrac{1}{P_n^3}\textstyle\sum_{i,j,k}^{*} \mathbf{X}_i'\mathbf{X}_j\mathbf{X}_j'\mathbf{X}_k\,, & Y_6 &= \frac{1}{P_n^4}\sum_{i,j,k,l}^{*} \mathbf{X}_i'\mathbf{X}_j\mathbf{X}_k'\mathbf{X}_l\,,
\end{aligned}
$$

with $P_n^k = n!/(n-k)!$. The test statistic of the test suggested by Chen et al. (2010) is given by

$$T_{CZZ} = p\left[\frac{Y_4 - 2Y_5 + Y_6}{(Y_1 - Y_2)^2}\right] - 1\,, \tag{25}$$

while Ahmad (2010) considered

$$T_A = p\frac{Y_4}{Y_3} - 1\,. \tag{26}$$

The test statistics (23), (24), (25), and (26) are asymptotically standard normally distributed under H_0. Moreover, from the derivation of these tests, it follows that they all are one-sided.

In the comparison study we also consider two tests derived using the methods of the random matrix theory. The first one was suggested by Johnstone (2001) which is based on the largest eigenvalue of the sample covariance matrix. The test statistic is given by

$$T_J = \frac{l_1 - \mu_{np}}{\sigma_{np}}\,, \tag{27}$$

where l_1 denotes the largest eigenvalue of $n\mathbf{S}$ and

$$\mu_{np} = (\sqrt{n-1} + \sqrt{p})^2 \quad \text{and} \quad \sigma_{np} = (\sqrt{n-1} + \sqrt{p})\left(\frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{p}}\right)^{1/3}.$$

Johnstone (2001) derived the asymptotic distribution of T_J under H_0 which follows the Tracy-Widom law that is tabulated. Its important critical values are given, for example, in Johnstone (2008).

Cai and Jiang (2011) considered a further test. Let r_{ij} be the sample correlation coefficient and let

$$L_n = \max_{1 \le i < j \le p} r_{ij}\,. \tag{28}$$

The test statistic is given by

$$T_{CJ} = nL_n^2 - 4\ln p + \ln\ln p\,, \tag{29}$$

which under the null hypothesis of zero population correlations has an extreme distribution of type I with the distribution function expressed as

$$F(y) = e^{-(1/\sqrt{8\pi})e^{-y/2}} \quad \text{for} \quad y \in I\!R\,.$$

Next, we analyze the size of the test derived in this paper as well as of the benchmark asymptotic tests. The analysis is performed by carrying out a Monte Carlo study. In each case, a sample of 10^5 independent observations is simulated. Then, the relative frequencies of the rejection of H_0 are calculated. In the simulation study we chose $\mathbf{\Sigma}_0 = \mathbf{I}_p$. The size of each test is presented for $p \in \{5, 10, 25\}$ with $n \in \{30, ..., 250\}$ in Figure 3 and for $n \in \{10, 20\}$ with $p \in \{10, ..., 100\}$ in Figure 4. Because in most of the cases considered in Figure 4 the determinant of the sample covariance matrix is equal zero, we drop the tests based on the generalized variances (T_{LR} and T_{BJYZ} statistics) from the discussion.

In Figure 3 we observe that the tests based on the T_J, T_{CJ}, and T_{FSG} statistics are considerably undersized, whereas the tests based on T_A, T_{BJYZ}, and T_{CZZ} are usually oversized. In contrast to these tests the rest of the competitors are properly sized in almost all of the cases considered. The situation changes if n becomes smaller than p. In these cases non of the asymptotic tests is properly sized, especially for $n = 10$. The tests T_A and T_{CZZ} are significantly oversized, while the results for the tests based on T_J, T_{CJ}, and T_{FSG} lie always significantly below the desired significance level.

4.2. Comparison study

Because it is rather difficult to derive some analytical results concerning the distributions of the considered test statistics under H_1, we will carry

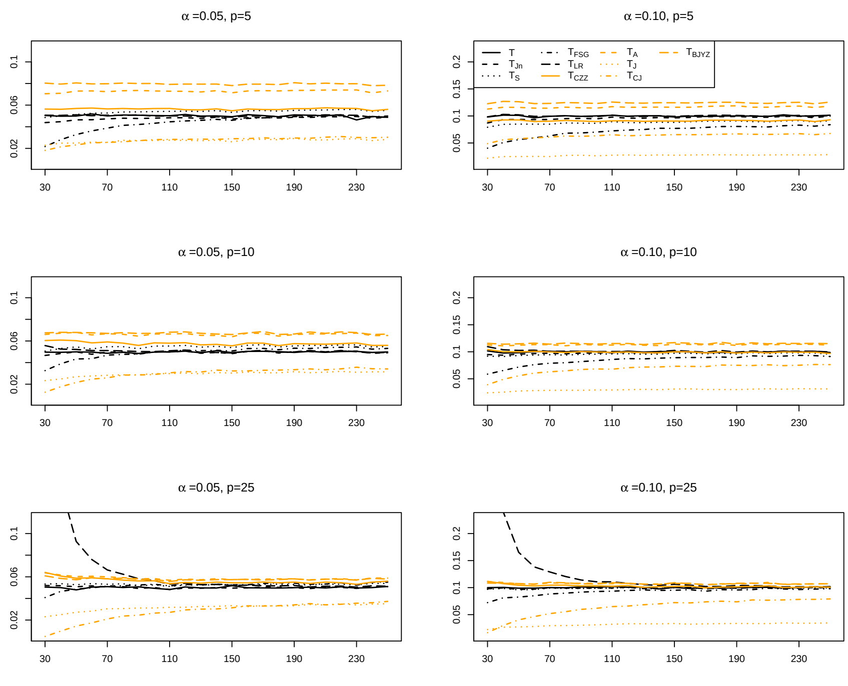

Figure 3: The estimated probabilities $P_{H_0}(T > T_{1-\alpha})$ for $p \in \{5, 10, 25\}$ and $n \in \{30, ..., 250\}$.

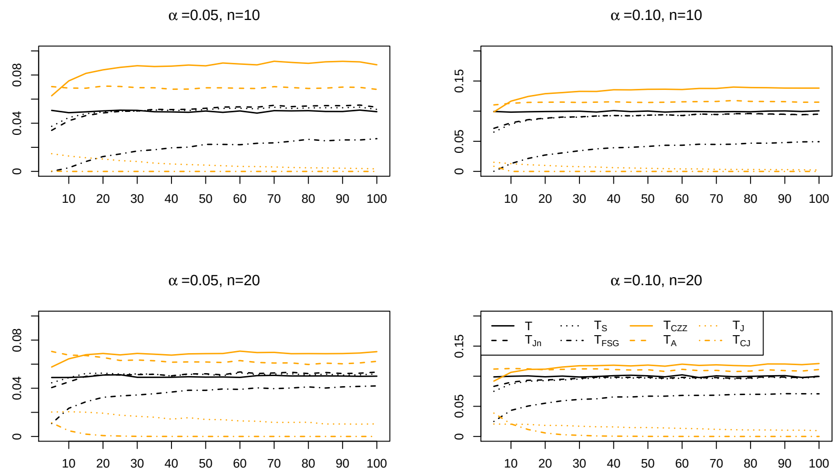

Figure 4: The estimated probabilities $P_{H_0}(T > T_{1-\alpha})$ for $p \in \{10, ..., 100\}$, $n \in \{10, 20\}$.

out a Monte Carlo study. In each case, a sample of 10000 independent observations is simulated. Then, the relative frequencies of the rejection of H_0 are calculated. In the simulation study we chose $\mathbf{\Sigma}_0 = \mathbf{I}_p$. The matrix $\mathbf{\Sigma}_1$ is obtained from $\mathbf{\Sigma}_0$ by changing its $(1, 2)$th and $(2, 1)$th elements given by $\nu_{1,1} \in \{-0.9, -0.8, ..., -0.1, 0, 0.1, ..., 0.9\}$. The significance level is fixed at $\alpha = 0.05$. The values of the power function are approximated for several dimensions of the random vectors $\mathbf{X}_i$, i.e. $p = 5, 10, 25$, and for several sample sizes $n = 30, 60, 120, 250$.

In Figures 5-7, the results of the simulation study are given. The case of $p = 5$ is treated in Figure 5, while Figures 6-7 show the values of the power functions for $p = 10$ and $p = 25$, respectively. We observe a very good performance of the new test which is ranked second. The best performance is reached by the test suggested in Cai and Jiang (2011). Although it is significantly undersized this test is ranked first in almost all of the considered cases. Such a results is not surprising since changes in covariances are considered and the test of Cai and Jiang (2011) was especially derived for these types of changes in the covariance matrix. On the third and the fourth places the tests based on the T_{LR} and T_{BJYZ} are present. It is also noted that the test derived from the likelihood ratio approach is significantly oversized for $p = 25$, which is not the case for its improved version suggested by Bai

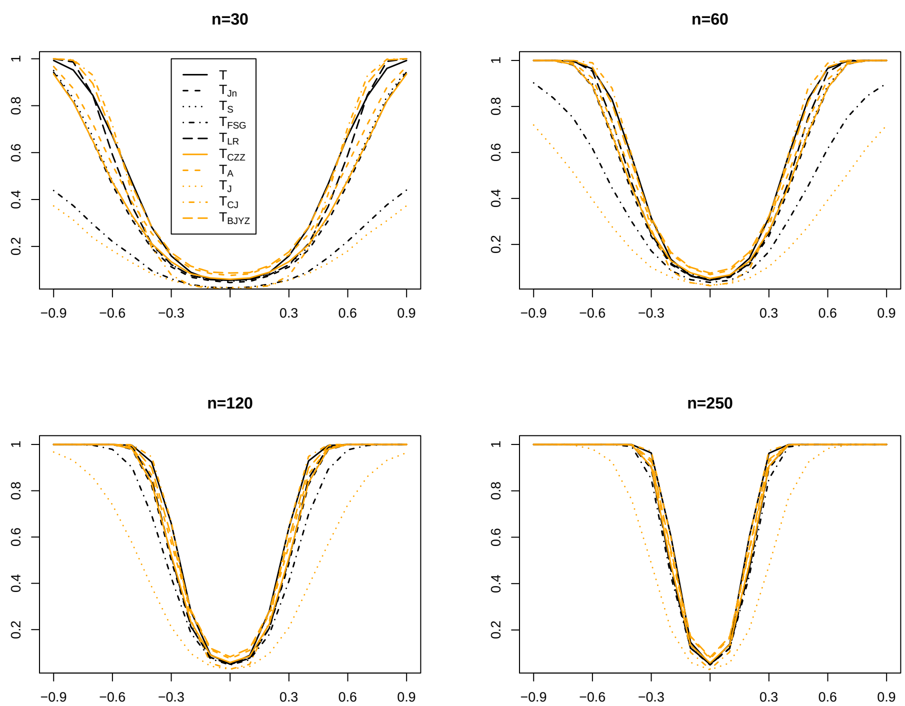

Figure 5: The estimated probabilities $P_{H_1}(T > T_{1-\alpha})$ for $p = 5$, $n \in \{30, 60, 120, 250\}$, and $\nu_{1,1} \in \{-0.9, -0.8, ..., -0.1, 0, 0.1, ..., 0.9\}$.

et al. (2009). The worst results are observed for the tests based on the T_{FSG} and T_J statistics, although the findings for the T_J statistic are expected since this test is based on the maximum eigenvalue of the covariance matrix and, as a result, it is not powerful enough to detect changes that have a minor impact on this quantity. A change in one of the correlation coefficients can be one of such possible situations. Finally, we note that the results obtained are robust with respect to the sample size and the dimension of $\mathbf{X}_i$.

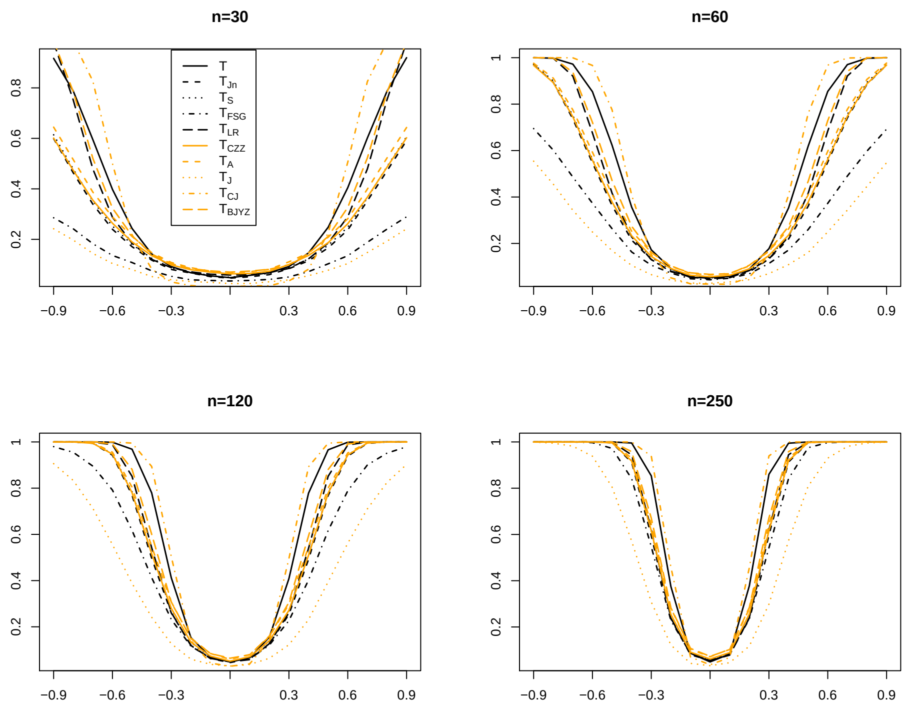

Figure 6: The estimated probabilities $P_{H_1}(T > T_{1-\alpha})$ for $p = 10$, $n \in \{30, 60, 120, 250\}$, and $\nu_{1,1} \in \{-0.9, -0.8, ..., -0.1, 0, 0.1, ..., 0.9\}$.

5. Appendix

In this section the proofs of Theorems 1a, 2a, 3 and of Corollary 1 are presented. In the case of $n \geq p$ the proofs are given by applying Theorem 3.2.10 of Muirhead (1982). For $n < p$ similar results are derived by using Lemma

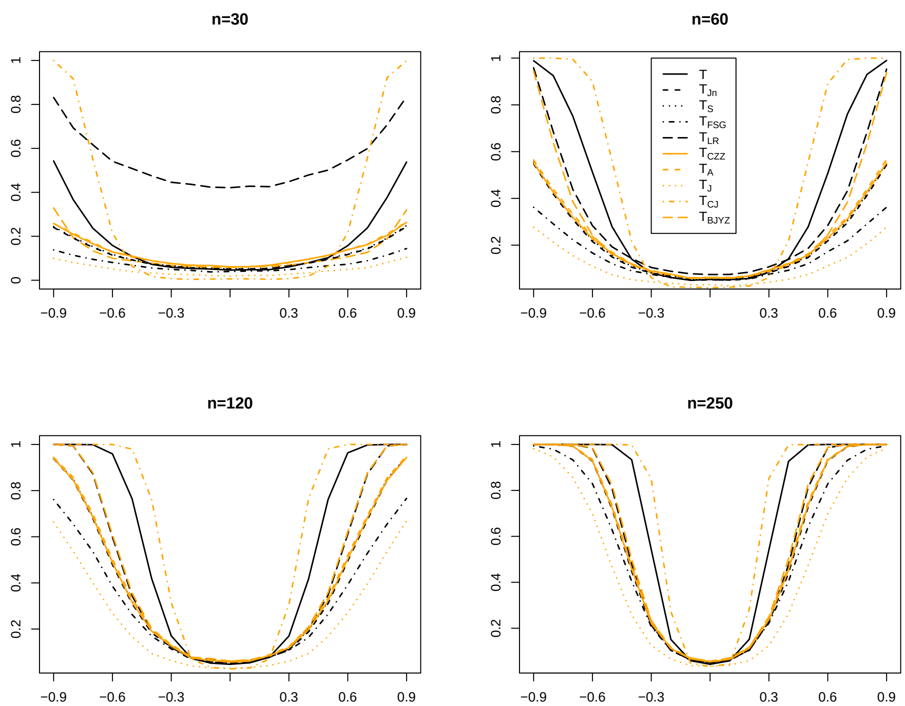

Figure 7: The estimated probabilities $P_{H_1}(T > T_{1-\alpha})$ for $p = 25$, $n \in \{30, 60, 120, 250\}$, and $\nu_{1,1} \in \{-0.9, -0.8, ..., -0.1, 0, 0.1, ..., 0.9\}$.

1 of Bodnar and Okhrin (2008) that extends Theorem 3.2.10 of Muirhead to the case of $n < p$. For this reason we next present the detailed proofs in the case of $n \geq p$ and note, wherever it is necessary, what has to be done for $n < p$.

Proof of Theorem 1a:
Application of Theorem 3.2.10 of Muirhead (1982) leads to

$$\mathbf{t}|v \sim \mathcal{N}_{p-1}\left(\boldsymbol{\nu}_1\xi_1^{-1}v, \frac{1}{n}\boldsymbol{\Upsilon}_1 v\right).$$

If $n < p$ then the last identity also holds following Lemma 1b) of Bodnar and Okhrin (2008).

It holds that

$$\frac{\mathbf{t}}{v}|v \sim \mathcal{N}_{p-1}\left(\frac{\boldsymbol{\nu}_1}{\xi_1}, \frac{1}{n}\frac{\boldsymbol{\Upsilon}_1}{v}\right)$$

and, hence,

$$\boldsymbol{\eta}_1|v \sim \mathcal{N}_{p-1}\left(\sqrt{n}\boldsymbol{\Upsilon}_0^{-1/2}\left(\frac{\boldsymbol{\nu}_1}{\xi_1} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right)\sqrt{v}, \boldsymbol{\Upsilon}_0^{-1/2}\boldsymbol{\Upsilon}_1\boldsymbol{\Upsilon}_0^{-1/2}\right). \tag{30}$$

Let

$$\boldsymbol{\Delta} = \boldsymbol{\Upsilon}_0^{-1/2}\left(\frac{\boldsymbol{\nu}_1}{\xi_1} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right) \qquad \text{and} \qquad \boldsymbol{\Omega} = \boldsymbol{\Upsilon}_0^{-1/2}\boldsymbol{\Upsilon}_1\boldsymbol{\Upsilon}_0^{-1/2}.$$

Because $nv/\xi_1 \sim \chi_n^2$ (see, e.g. Muirhead (1982, Theorem 3.2.10) for $n \geq p$ and Lemma 1a) of Bodnar and Okhrin (2008) for $n < p$) the unconditional density of $\boldsymbol{\eta}_1$ is given by

$$f_{\boldsymbol{\eta}_1}(\mathbf{x}) = \frac{\pi^{-(p-1)/2}2^{-(p-1)/2}}{|\boldsymbol{\Omega}|^{1/2}}\frac{n^{n/2}}{2^{n/2}\xi_1^{n/2}\Gamma\left(\frac{n}{2}\right)}$$
$$\times \int_0^\infty z^{n/2-1}\exp\left(-\frac{1}{2}\left(\frac{nz}{\xi_1} + (\mathbf{x} - \sqrt{n}\boldsymbol{\Delta}\sqrt{z})'\boldsymbol{\Omega}^{-1}(\mathbf{x} - \sqrt{n}\boldsymbol{\Delta}\sqrt{z})\right)\right)dz.$$

The transformation $z = y^2$ yields

$$
\begin{aligned}
f_{\boldsymbol{\eta}_1}(\mathbf{x}) &= \frac{\pi^{-(p-1)/2}2^{-(p-1)/2}}{|\boldsymbol{\Omega}|^{1/2}} \frac{n^{n/2}}{2^{n/2}\xi_1^{n/2}\Gamma\left(\frac{n}{2}\right)} \\
&\times\ 2\int_0^\infty y^{n-1}\exp\left(-\frac{1}{2}\left(\frac{ny^2}{\xi_1} + (\mathbf{x}-\sqrt{n}\boldsymbol{\Delta}y)'\boldsymbol{\Omega}^{-1}(\mathbf{x}-\sqrt{n}\boldsymbol{\Delta}y)\right)\right)dy \\
&= 2\frac{\pi^{-(p-1)/2}2^{-(p-1)/2}n^{n/2}}{|\boldsymbol{\Omega}|^{1/2}2^{n/2}\xi_1^{n/2}\Gamma\left(\frac{n}{2}\right)}\exp\left(-\frac{1}{2}\left(\mathbf{x}'\left(\boldsymbol{\Omega}^{-1} - \frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}}{\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}}\right)\mathbf{x}\right)\right) \\
&\times \int_0^\infty y^{n-1}\exp\left(-\frac{n(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})}{2}\left(y - \frac{\sqrt{n}\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\mathbf{x}}{n(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})}\right)^2\right)dy\,.
\end{aligned}
$$

The last integral is evaluated as

$$
\begin{aligned}
f_{\boldsymbol{\eta}_1}(\mathbf{x}) &= 2\frac{\sqrt{2\pi}n^{n/2}}{|\boldsymbol{\Omega}|^{1/2}2^{n/2}\xi_1^{n/2}\Gamma\left(\frac{n}{2}\right)} \frac{n^{-1/2}(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})^{-1/2}}{\left|\boldsymbol{\Omega}^{-1} - \frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}}{\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}}\right|^{1/2}} \\
&\times\ \phi_{p-1}\left(\mathbf{x};\mathbf{0}_{p-1},\left(\boldsymbol{\Omega}^{-1} - \frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}}{\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}}\right)^{-1}\right) \\
&\times \int_0^\infty y^{n-1}\phi\left(y;\frac{\sqrt{n}\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\mathbf{x}}{n(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})}, n^{-1}(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})^{-1}\right)dy\,.
\end{aligned}
$$

The applications of Theorem 18.1.1 and Theorem 18.2.8 of Harville (1997) leads to

$$
\begin{aligned}
\left|\boldsymbol{\Omega}^{-1} - \frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}}{\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}}\right| &= |\boldsymbol{\Omega}^{-1}|(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})^{-1} \\
&\times\ (\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}-\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}) \\
&= |\boldsymbol{\Omega}|^{-1}(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})^{-1}\xi_1^{-1}\,, \\
\left(\boldsymbol{\Omega}^{-1} - \frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}}{\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}}\right)^{-1} &= \boldsymbol{\Omega} + \frac{\frac{\boldsymbol{\Delta}\boldsymbol{\Delta}'}{\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}}}{1-\frac{\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}}{\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta}}} = \boldsymbol{\Omega} + \xi_1\boldsymbol{\Delta}\boldsymbol{\Delta}'\,.
\end{aligned}
$$

Hence,

$$f_{\boldsymbol{\eta}_1}(\mathbf{x}) = 2\frac{\sqrt{\pi}n^{(n-1)/2}}{2^{(n-1)/2}\xi_1^{(n-1)/2}\Gamma\left(\frac{n}{2}\right)}\phi_{p-1}\left(\mathbf{x};\mathbf{0}_{p-1},\boldsymbol{\Omega}+\xi_1\boldsymbol{\Delta}\boldsymbol{\Delta}'\right)$$
$$\times \int_0^\infty y^{n-1}\phi\left(y;\frac{\sqrt{n}\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\mathbf{x}}{n(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})}, n^{-1}(\xi_1^{-1}+\boldsymbol{\Delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})^{-1}\right)dy\,.$$

The theorem is proved.

Proof of Theorem 2a: From the proof of Theorem 2a we get

$$\boldsymbol{\eta}_1|v=y\sim\mathcal{N}_{p-1}\left(\sqrt{n}\boldsymbol{\Delta}\sqrt{y},\boldsymbol{\Omega}\right)\,.$$

Let $\boldsymbol{\Omega}=\mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'$ be the eigenvalue decomposition of $\boldsymbol{\Omega}$ where $\boldsymbol{\Lambda}=diag(\lambda_1,...,\lambda_{p-1})$ is the diagonal matrix of eigenvalues and $\mathbf{P}$ is the corresponding orthogonal matrix of eigenvectors. Then (cf., Mathai and Provost (1992, Chapter 4))

$$T_1|v=y\overset{d}{=}\sum_{i=1}^{p-1}\lambda_i(Z_i+b_i(y))^2$$

where $\mathbf{b}(y)=(b_1(y),...,b_{p-1}(y))'$ is given in (9) and

$$\mathbf{Z}=(Z_1,...,Z_p)'=\mathbf{P}'\boldsymbol{\Omega}^{-1/2}(\boldsymbol{\eta}_1-\sqrt{n}\boldsymbol{\Delta}\sqrt{y})\,.$$

Using the results of Imhof (1961), we obtain the conditional distribution function of T_1 given $v=y$ expressed as

$$F_{T_1|v=y}(x)=\frac{1}{2}-\frac{1}{\pi}\int_0^\infty\frac{\sin\theta(u,y)}{u\rho(u,y)}du\,, \tag{31}$$

where $\theta(u,y)$ and $u\rho(u,y)$ are defined in (10) and (11). Because $nv/\xi_1\sim\chi_n^2$ (see, e.g. Muirhead (1982, Theorem 3.2.10), multiplying (31) by the density of v and integrating out y leads to the expression given in the statement of Theorem 2a.

Proof of Corollary 1: From the proof of Theorem 2a we get

$$\boldsymbol{\eta}_1|v\sim\mathcal{N}_{p-1}\left(\sqrt{n}\boldsymbol{\Delta}\sqrt{v},\mathbf{I}_{p-1}\right)\,.$$

Thus,

$$T_1|v = \boldsymbol{\eta}_1'\boldsymbol{\eta}_1|v \sim \chi^2_{p-1}(\tilde{\lambda}v)$$

with

$$\tilde{\lambda} = n\boldsymbol{\Delta}'\boldsymbol{\Delta} = n\left(\frac{\boldsymbol{\nu}_1}{\xi_1} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right)' \boldsymbol{\Upsilon}_0^{-1}\left(\frac{\boldsymbol{\nu}_1}{\xi_1} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right).$$

Using the fact that $nv/\xi_1 \sim \chi^2_n$ (see, e.g. Muirhead (1982, Theorem 3.2.10) for $n \geq p$ and Lemma 1a) of Bodnar and Okhrin (2008) for $n < p$) the unconditional density of T_1 is given by

$$f_{T_1}(x) = \frac{n^{n/2}}{2^{n/2}\xi_1^{n/2}\Gamma\left(\frac{n}{2}\right)} \int_0^\infty y^{n/2-1} \exp\left(-\frac{n}{2}\left(\frac{y}{\xi_1}\right)\right) f_{\chi^2_{p-1}(\tilde{\lambda}y)}(x)dy\,.$$

Let f_{p-1} denote the density of the χ^2_{p-1}-distribution. The application of the identity (Muirhead (1982, Theorem 1.3.4))

$$f_{\chi^2_{p-1}(\tilde{\lambda}y)}(x) = \exp\left(-\frac{1}{2}\tilde{\lambda}y\right) {}_0F_1\left(\frac{p-1}{2}; \frac{1}{4}\tilde{\lambda}yx\right) f_{p-1}(x)\,,$$

leads to

$$\begin{aligned} f_{T_1}(x) &= \frac{n^{n/2}}{2^{n/2}\xi_1^{n/2}\Gamma\left(\frac{n}{2}\right)} f_{p-1}(x) \\ &\times \int_0^\infty y^{n/2-1} \exp\left(-\frac{1}{2}(n\xi_1^{-1} + \tilde{\lambda})y\right) {}_0F_1\left(\frac{p-1}{2}; \frac{1}{4}\tilde{\lambda}yx\right) dy\,. \end{aligned}$$

The last integral is evaluated by using Lemma 1.3.3 of Muirhead (1982) and is equal to

$$\begin{aligned} f_{T_1}(x) &= \frac{n^{n/2}}{2^{n/2}\xi_1^{n/2}\Gamma\left(\frac{n}{2}\right)} \\ &\times f_{p-1}(x)\Gamma\left(\frac{n}{2}\right) \frac{2^{n/2}}{(n\xi_1^{-1} + \tilde{\lambda})^{n/2}} {}_1F_1\left(\frac{n}{2}; \frac{p-1}{2}; \frac{\tilde{\lambda}x}{2(n\xi_1^{-1} + \tilde{\lambda})}\right) \\ &= \frac{1}{(1 + n^{-1}\xi_1\tilde{\lambda})^{n/2}} f_{p-1}(x)\, {}_1F_1\left(\frac{n}{2}; \frac{p-1}{2}; \frac{\tilde{\lambda}x}{2(n\xi_1^{-1} + \tilde{\lambda})}\right) \end{aligned}$$

Noting that $\lambda = \xi_1\tilde{\lambda}/n$ completes the proof.

Proof of Theorem 3:
First, we rewrite the expression of the statistic T_i. Without loss of generality we perform the calculations only in the case of T_1 and note that a similar approach can be used for T_i, $i = 1, ..., p$.

Let $\mathbf{X}_i = (X_{1i}\ \mathbf{X}'_{2i})'$. Then it holds that

$$\begin{aligned}
& T_1 = \boldsymbol{\eta}_1'\boldsymbol{\eta}_1 = n\left(\frac{\mathbf{t}}{v} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right)' \boldsymbol{\Upsilon}_0^{-1}\left(\frac{\mathbf{t}}{v} - \frac{\boldsymbol{\nu}_0}{\xi_0}\right) v \\
= & \frac{\left(\sum_{i=1}^n \mathbf{X}_{2i}X_{1i} - \frac{\boldsymbol{\nu}_0}{\xi_0}\sum_{i=1}^n X_{1i}^2\right)' \boldsymbol{\Upsilon}_0^{-1}\left(\sum_{i=1}^n \mathbf{X}_{2i}X_{1i} - \frac{\boldsymbol{\nu}_0}{\xi_0}\sum_{i=1}^n X_{1i}^2\right)}{\sum_{i=1}^n X_{1i}^2} \\
= & \frac{1}{\sum_{i=1}^n X_{1i}^2}\left(\begin{array}{c} \sum_{i=1}^n X_{1i}^2 \\ \sum_{i=1}^n \mathbf{X}_{2i}X_{1i} \end{array}\right)' \left[-\frac{\boldsymbol{\nu}_0}{\xi_0}\ \mathbf{I}_{p-1}\right]' \\
\times & \boldsymbol{\Upsilon}_0^{-1}\left[-\frac{\boldsymbol{\nu}_0}{\xi_0}\ \mathbf{I}_{p-1}\right]\left(\begin{array}{c} \sum_{i=1}^n X_{1i}^2 \\ \sum_{i=1}^n \mathbf{X}_{2i}X_{1i} \end{array}\right) \\
= & \frac{1}{\sum_{i=1}^n X_{1i}^2}\left(\sum_{i=1}^n X_{1i}\mathbf{X}_i\right)' \left[-\frac{\boldsymbol{\nu}_0}{\xi_0}\ \mathbf{I}_{p-1}\right]' \boldsymbol{\Upsilon}_0^{-1}\left[-\frac{\boldsymbol{\nu}_0}{\xi_0}\ \mathbf{I}_{p-1}\right]\left(\sum_{i=1}^n X_{1i}\mathbf{X}_i\right).
\end{aligned}$$

Because

$$\begin{aligned}
& \left[-\frac{\boldsymbol{\nu}_0}{\xi_0}\ \mathbf{I}_{p-1}\right]' \boldsymbol{\Upsilon}_0^{-1}\left[-\frac{\boldsymbol{\nu}_0}{\xi_0}\ \mathbf{I}_{p-1}\right] \\
= & \left(\begin{array}{cc} \left(\frac{\boldsymbol{\nu}_0}{\xi_0}\right)' \boldsymbol{\Upsilon}_0^{-1}\frac{\boldsymbol{\nu}_0}{\xi_0} & -\left(\frac{\boldsymbol{\nu}_0}{\xi_0}\right)' \boldsymbol{\Upsilon}_0^{-1} \\ -\boldsymbol{\Upsilon}_0^{-1}\left(\frac{\boldsymbol{\nu}_0}{\xi_0}\right)' & \boldsymbol{\Upsilon}_0^{-1} \end{array}\right) = \boldsymbol{\Sigma}^{-1} - \left(\begin{array}{cc} \xi_0^{-1} & \mathbf{0}'_{p-1} \\ \mathbf{0}_{p-1} & \mathbf{0}_{p-1,p-1} \end{array}\right),
\end{aligned}$$

where the expression of the inverse of the partitioned matrix was used (see, e.g. Harville (1997, Corollary 8.5.12)), we get

$$T_1 = \frac{1}{\sum_{i=1}^n X_{1i}^2}\left(\left(\sum_{i=1}^n X_{1i}\mathbf{X}_i\right)' \boldsymbol{\Sigma}^{-1}\left(\sum_{i=1}^n X_{1i}\mathbf{X}_i\right) - \left(\sum_{i=1}^n X_{1i}^2\right)^2 / \sigma_{0;11}\right).$$

Let $\mathbf{U}_i = \boldsymbol{\Sigma}_0^{-1/2}\mathbf{X}_i$ where $\boldsymbol{\Sigma}_0^{-1/2}$ is the Cholesky squared root of the matrix $\boldsymbol{\Sigma}_0^{-1}$. Then

$$T_1 = \frac{\left(\sum_{i=1}^n U_{1i}\mathbf{U}_i\right)'\left(\sum_{i=1}^n U_{1i}\mathbf{U}_i\right)}{\sum_{i=1}^n U_{1i}^2} - \sum_{i=1}^n U_{1i}^2\,, \tag{32}$$

where $\mathbf{U}_i = (U_{1i}\ \mathbf{U}'_{2i})'$. Taking the maximum over T_j, $j = 1, ..., p$, we get the statement of Theorem 3.

Acknowledgements

The authors are grateful to the referees and the editor for their suggestions, which have improved the presentation in the paper. The second author is partly supported by the German Science Foundation (DFG) via the Research Unit 1735 "Structural Inference in Statistics: Adaptation and Efficiency". He also appreciates the financial support of the German Science Foundation (DFG) via the projects BO3521/2 and OK103/1 "Wishart Processes in Statistics and Econometrics: Theory and Applications".

References

[1] Ahmad, M.R., (2010), Tests for Covariance Matrices, Particularly for High-Dimensional Data, working paper.

[2] Bai, Z., D. Jiang, J.-F. Yao, and S. Zheng, (2009), Corrections to LRT on large-dimensional covariance matrix by RMT, *The Annals of Statistics* 37, 3822-3840.

[3] Bai, Z. and J. W. Silverstein, (2004), CLT for linear spectral statistics of large-dimensional sample covariance matrices, *The Annals of Probability* 32, 553-605.

[4] Bodnar, T. and Y. Okhrin, (2008), Properties of the singular, inverse and generalized inverse partitioned Wishart distributions, *Journal of Multivariate Analysis* 99, 2389-2405.

[5] Cai, T. and T. Jiang, (2011), Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices, *The Annals of Statistics* 39, 1496-1525.

[6] Chen, S.X., Zhang, L.-X. and P.-S. Zhong, (2010), Testing high dimensional covariance matrices. *Journal of the American Statistical Association* 105, 810-819.

[7] Fisher, T.J., X. Sun, and C. M. Gallagher, (2010), A new test for sphericity of the covariance matrix for high dimensional data, *Journal of Multivariate Analysis* 101, 2554-2570.

[8] Gupta, A.K. and J. Xu, (2006), On some tests of the covariance matrix under general conditions, *Annals of the Institute of Statistical Mathematics* 58, 101-114.

[9] Harville, D.A., (1997), *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag: New York.

[10] John, S., (1971), Some optimal multivariate tests, *Biometrika* 58, 123-127.

[11] Johnstone, I. M., (2001), On the distribution of the largest eigenvalue in principal components analysis, *The Annals of Statistics* 29, 295-327.

[12] Johnstone, I. M., (2008), Multivariate analysis and Jacobi ensembles: Largest eigenvalue, Tracy-Widom limits and rates of convergence, *The Annals of Statistics* 36, 2638-2716.

[13] Ledoit, O. and M. Wolf, (2002), Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size, *The Annals of Statistics* 30, 1081-1102.

[14] Mauchly, J.W., (1940), Significance test for sphericity of a normal N-variate distribution, *The Annals of Mathematical Statistics* 11, 204-209.

[15] Mathai A. M. and S. B. Provost, (1992), *Quadratic Forms in Random Variables: Theory and Applications*, Marcel Dekker: New York.

[16] Nagao, H., (1973), On some test criteria for covariance matrix, *The Annals of Statistics* 1, 700-709.

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

[18] Rencher, A.C., (2002), *Methods of Multivariate Analysis*, Wiley: New York.

[19] Srivastava, M.S., (2003), Singular Wishart and multivariate beta distributions, *The Annals of Statistics* 31, 1537-1560.

[20] Srivastava, M.S., (2005), Some tests concerning the covariance matrix in high dimensional data, *Journal of Japan Statistical Society* 35, 251-272.