



Integral transform methods in goodness-of-fit testing, II: the Wishart distributions

Elena Hadjicosta¹ · Donald Richards¹

Received: 1 May 2019 / Revised: 12 September 2019 / Published online: 20 November 2019
© The Institute of Statistical Mathematics, Tokyo 2019

Abstract

We initiate the study of goodness-of-fit testing for data consisting of positive definite matrices. Motivated by the appearance of positive definite matrices in numerous applications, including factor analysis, diffusion tensor imaging, volatility models for financial time series, wireless communication systems, and polarimetric radar imaging, we apply the method of Hankel transforms of matrix argument to develop goodness-of-fit tests for Wishart distributions with given shape parameter and unknown scale matrix. We obtain the limiting null distribution of the test statistic and a corresponding covariance operator, show that the eigenvalues of the operator satisfy an interlacing property, and apply our test to some financial data. We establish the consistency of the test against a large class of alternative distributions and derive the asymptotic distribution of the test statistic under a sequence of contiguous alternatives. We obtain the Bahadur and Pitman efficiency properties of the test statistic and establish a modified version of Wieand's condition.

Keywords Bahadur slope · Bessel function of matrix argument · Contiguous alternative · Diffusion tensor imaging · Factor analysis · Gaussian random field · Pitman efficiency · Zonal polynomial

1 Introduction

The problem of testing that a random sample of positive definite matrices follows a Wishart distribution arose in factor analysis over fifty years ago; Browne (1968, p. 278)

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10463-019-00737-z>) contains supplementary material, which is available to authorized users.

✉ Donald Richards
richards@stat.psu.edu

Elena Hadjicosta
exh963@psu.edu

¹ Department of Statistics, Pennsylvania State University, University Park, PA 16801, USA

noted the difficulty of performing such a test, but no such results have appeared since then. More recently, random positive definite matrix data have appeared in numerous applications, e.g., diffusion tensor imaging, financial time series, wireless communication, and polarimetric radar images.

Positive definite random matrix data are especially important in medical research, specifically in diffusion tensor imaging (DTI) (Dryden et al. 2009; Jian et al. 2007; Jian and Vemuri 2007; Kim et al. 2011; Lee and Schwartzman 2017; Schwartzman 2006; Schwartzman et al. 2005, 2008). DTI is a magnetic resonance imaging method that has attracted much interest in the study of brain diseases. DTI is based on the observation that water molecules in vivo are always in motion; by modeling the diffusion of the water molecules at any location by a three-dimensional Brownian motion, the resulting diffusion tensor image is represented by the 3×3 positive definite matrix of the local diffusion process at the given location.

DTI, although noninvasive, enables the study of deep brain white-matter fibers. Thus, DTI has been used to study epileptic seizures, Alzheimer's disease, traumatic brain injuries, white-matter abnormalities, developmental disorders, and psychiatric conditions (Neumann-Haefelin et al. 2000; Rosenbloom et al. 2003; Pomara et al. 2001; Matthews and Arnold 2001), and also to study the pathology of organs or tissues such as the breast, cardiac, kidney, lingual, skeletal muscles, and spinal cord (Damon et al. 2002). The Wishart distribution with known degrees of freedom and unknown scale matrix has appeared in several articles on DTI data (Dryden et al. 2009; Jian et al. 2007; Jian and Vemuri 2007).

The Wishart distributions with known degrees of freedom also arise in stochastic volatility models (Asai et al. 2006; Gouriéroux and Sufana 2010; Ku and Bloomfield 2010). Here, the problem is to estimate the covariance matrix of the joint capital returns on several financial assets, with the goal of predicting returns, devising portfolio allocations, and estimating risk.

The *complex* Wishart distributions with known degrees of freedom arise in the spectral analysis of multivariate Gaussian time series (Goodman 1963), wireless communications (Siriteanu et al. 2016, 2015; Tulino and Verdú 2004), and the analysis of polarimetric synthetic aperture radar (Anfinsen and Eltoft 2011; Anfinsen et al. 2011). The results to follow can be extended, with obvious changes, to the complex Wishart distributions (James 1964, p. 488) and even to Wishart distributions on symmetric cones (Faraut and Korányi 1994).

Motivated by these applications, we develop goodness-of-fit tests for the Wishart distributions, extending results for the exponential distributions (Baringhaus and Taherizadeh 2010; Taherizadeh 2009) and the gamma distributions (Hadjicosta 2019; Hadjicosta and Richards 2019). The technical material needed to develop such tests includes mathematical analysis on the cone of positive definite matrices (Herz 1955; Maass 1971), the Bessel and Laguerre polynomials of matrix argument and their zonal polynomial expansions (Gross and Richards 1987; Herz 1955; James 1964; Muirhead 1982), and the Hankel transforms of matrix argument (Herz 1955). To simplify the exposition, we present numerous proofs as supplementary material in Sects. S.10–S.13.

The non-commutative nature of matrix multiplication leads us to impose on the distribution of the sample data an orthogonal invariance condition. The Frobenius,

spectral, and operator norms appear in the matrix case, and several inequalities between them will be needed. There is also the surprising appearance of Schur's lemma in Sect. 2.3, a result well known in linear algebra and representation theory (Shilov 1977) but which we have not seen before now in statistical inference.

In Sect. 2, we provide properties of the Wishart distribution, Bessel function, Hankel transform, confluent hypergeometric function, and generalized Laguerre polynomial, all of matrix argument. We also provide a uniqueness theorem and an inversion formula for the Hankel transform and some limit theorems. We present a generalized hypergeometric function of two matrix arguments, define the orthogonally invariant Hankel transform, and provide some of their properties.

In Sect. 3, we define the statistic T_n^2 for goodness-of-fit testing for the Wishart distributions. We obtain the asymptotic distribution of T_n^2 under the null hypothesis as an integral of the square of a centered Gaussian random field \mathcal{Z} . In Sect. 4, we derive the covariance operator corresponding to \mathcal{Z} and show that the eigenvalues of \mathcal{S} satisfy an interlacing property. It remains an open problem to determine the multiplicity of the eigenvalues of the operator.

In Sect. 5, we apply the test to financial data, and we establish in Sect. 6 the consistency of the test against numerous alternatives. In Sect. 7, we derive the asymptotic distribution of T_n^2 under certain sequences of contiguous alternatives to the null hypothesis, such as Wishart alternatives with varying shape or scale parameters and some contaminated Wishart models.

Finally, in Sect. 8, we establish the Bahadur and Pitman efficiency properties of the statistic T_n^2 . We investigate the approximate Bahadur slope of T_n^2 under local alternatives, and we show the validity of a modified Wieand's condition. A complete extension of Wieand's condition, under which the Bahadur and Pitman efficiencies coincide, remains an open problem.

2 Wishart distributions and Hankel transforms of matrix argument

2.1 Preliminary results for the Wishart distributions

Throughout the paper, all needed results on the zonal polynomials and on the special functions of matrix argument are provided by Herz (1955), Muirhead (1982), or Richards (2010), and we will generally conform to their notation. We denote any zero matrix by 0, the order being determined by the context; also, I_m denotes the $m \times m$ identity matrix. We denote by $\mathbb{R}^{m \times m}$ the space of $m \times m$ (real) matrices, by $\mathcal{S}^{m \times m}$ the space of $m \times m$ symmetric matrices, by $\mathcal{P}_+^{m \times m}$ the cone of $m \times m$ positive definite matrices, and by $O(m)$ the group of $m \times m$ orthogonal matrices. To specify that $Y \in \mathcal{P}_+^{m \times m}$, we usually write $Y > 0$; more generally, we write $Y_1 > Y_2$ whenever $Y_1 - Y_2 > 0$. We also denote the trace of Y by $\text{tr}(Y)$, the determinant of Y by $\det(Y)$, and $\exp(\text{tr } Y)$ by $\text{etr}(Y)$.

The *multivariate gamma function* is defined by

$$\Gamma_m(a) = \int_{R>0} (\det R)^{a-\frac{1}{2}(m+1)} \operatorname{etr}(-R) \, dR,$$

for $a \in \mathbb{C}, \operatorname{Re}(a) > \frac{1}{2}(m - 1)$; this integral is well known to have the explicit formula,

$$\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{j=1}^m \Gamma\left(a - \frac{1}{2}(j - 1)\right).$$

An $m \times m$ positive definite random matrix X is said to have a *Wishart distribution* if its probability density function (p.d.f.) is of the form

$$f(X) = \frac{1}{\Gamma_m(\alpha)} (\det \Sigma)^\alpha (\det X)^{\alpha-\frac{1}{2}(m+1)} \operatorname{etr}(-\Sigma X), \tag{1}$$

$X > 0$, where $\alpha > \frac{1}{2}(m - 1)$ and $\Sigma > 0$. We write $X \sim W_m(\alpha, \Sigma)$ whenever (1) holds. The parameter α is called the *shape parameter* and Σ is called the *scale matrix* of X . If α is a half-integer, then 2α is called the *degrees of freedom of X* . In general, $E(X) = \alpha \Sigma^{-1}$; also, if M is a $q \times m$ matrix of rank q , where $q \leq m$, then $MXM' \sim W_q(\alpha, (M\Sigma^{-1}M')^{-1})$ (Muirhead 1982, p. 92).

A *partition* $\kappa = (k_1, \dots, k_m)$ is a vector of nonnegative integers, listed in non-increasing order. The *weight* of κ is $|\kappa| = k_1 + \dots + k_m$, and the *length*, $\ell(\kappa)$, of κ is the number of nonzero $k_j, j = 1, \dots, m$. For $a \in \mathbb{C}$ and $k = 0, 1, 2, \dots$, the *shifted factorial* is defined as $(a)_k = a(a + 1)(a + 2) \dots (a + k - 1)$. For any partition $\kappa = (k_1, \dots, k_m)$, the *partitional shifted factorial* is defined as

$$[a]_\kappa = \prod_{j=1}^m \left(a - \frac{1}{2}(j - 1)\right)_{k_j}.$$

For $Y \in \mathcal{S}^{m \times m}$, we denote by $\det_j(Y)$ the j th principal minor of $Y, j = 1, \dots, m$. For any partition κ , the *zonal polynomial* $C_\kappa(Y)$ is defined as

$$C_\kappa(Y) = C_\kappa(I_m) (\det Y)^{k_m} \int_{O(m)} \prod_{j=1}^{m-1} (\det_j(HYH^{-1}))^{k_j-k_{j+1}} \, dH, \tag{2}$$

where dH is the normalized Haar measure on $O(m)$ (Richards 2010, (35.4.2)). By (2), $C_\kappa(Y)$ is homogeneous of degree $|\kappa|$.

It follows from the invariance of the Haar measure that $C_\kappa(HYH') = C_\kappa(Y)$ for all $H \in O(m)$ and $Y \in \mathcal{S}^{m \times m}$; hence, $C_\kappa(Y)$ depends only on the eigenvalues of Y and it is a symmetric function of the eigenvalues. Suppose that $Z \in \mathcal{S}^{m \times m}$ and that $Y^{1/2}$ denotes the unique positive definite square root of $Y \in \mathcal{P}_+^{m \times m}$. Since the matrices $Y^{1/2}ZY^{1/2}, YZ$, and ZY all have the same eigenvalues, we will follow standard

convention, writing $C_\kappa(YZ)$ or $C_\kappa(ZY)$ for $C_\kappa(Y^{1/2}ZY^{1/2})$; throughout the paper, we retain this convention for all orthogonally invariant functions of matrix argument.

With the normalization

$$C_\kappa(I_m) = 2^{2|\kappa|} |\kappa|! [m/2]_\kappa \frac{\prod_{i < j}^{\ell(\kappa)} (2k_i - 2k_j - i + j)}{\prod_{i=1}^{\ell(\kappa)} (2k_i + \ell(\kappa) - i)!}, \tag{3}$$

the zonal polynomials satisfy the identity,

$$(\text{tr } Y)^k = \sum_{|\kappa|=k} C_\kappa(Y), \tag{4}$$

$k = 0, 1, 2, \dots$ (see [Muirhead 1982](#), p. 228, Eq. (iii) or [Richards 2010](#), Eq. (35.4.6)). For $Y > 0$ and $Z \in \mathcal{S}^{m \times m}$, the zonal polynomials satisfy the *mean-value property* ([Muirhead 1982](#), p. 243),

$$\int_{O(m)} C_\kappa(HYH'Z) dH = \frac{C_\kappa(Y) C_\kappa(Z)}{C_\kappa(I_m)}. \tag{5}$$

We will also need the identity,

$$\sum_{|\kappa|=k} C_\kappa(I_m)[a]_\kappa = (m a)_k, \tag{6}$$

$a \in \mathbb{C}, k = 0, 1, 2, \dots$. This result is established by applying a power series identity,

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{|\kappa|=k} C_\kappa(I_m)[a]_\kappa = (\det(I_m - tI_m))^{-a}, \tag{7}$$

$|t| < 1$; see [James \(1964, p. 495, Eq. \(143\)\)](#), [Muirhead \(1982, p. 259, Eq. \(4\)\)](#). Writing

$$(\det(I_m - tI_m))^{-a} \equiv (1 - t)^{-ma} = \sum_{k=0}^{\infty} \frac{t^k}{k!} (m a)_k, \tag{8}$$

then (6) is obtained by comparing the coefficients of t^k in (7) and (8).

The zonal polynomials also satisfy a Laplace transform identity ([Muirhead 1982](#), p. 248): For $\text{Re}(a) > \frac{1}{2}(m - 1)$, $Z > 0$, and $M \in \mathcal{S}^{m \times m}$,

$$\int_{R>0} C_\kappa(MR)(\det R)^{a-\frac{1}{2}(m+1)} \text{etr}(-RZ) dR = [a]_\kappa \Gamma_m(a)(\det Z)^{-a} C_\kappa(MZ^{-1}). \tag{9}$$

For $\kappa = 0$, this result reduces to

$$\int_{R>0} (\det R)^{a-\frac{1}{2}(m+1)} \text{etr}(-RZ) dR = \Gamma_m(a)(\det Z)^{-a}, \tag{10}$$

from which we confirm that (1) is a probability density function (Muirhead 1982, p. 61).

2.2 Bessel functions and Laguerre polynomials of matrix argument

The Bessel function of matrix argument, first treated in detail by Herz (1955), can be defined in several ways. Let $\nu \in \mathbb{C}$ be such that $-\nu + \frac{1}{2}(j - m) \notin \mathbb{N}$ for all $j = 1, \dots, m$; these restrictions ensure that $[\nu + \frac{1}{2}(m + 1)]_\kappa \neq 0$ for all partitions κ . Following Muirhead (1982, Chapter 7), the *Bessel function (of the first kind) of order ν* is defined for $Y \in \mathcal{S}^{m \times m}$ as

$$A_\nu(Y) = \frac{1}{\Gamma_m(\nu + \frac{1}{2}(m + 1))} \sum_{k=0}^\infty \frac{(-1)^k}{k!} \sum_{|\kappa|=k} \frac{1}{[\nu + \frac{1}{2}(m + 1)]_\kappa} C_\kappa(Y). \tag{11}$$

We also refer to Faraut and Korányi (1994), Gross and Richards (1987), James (1964), Richards (2010) for further details of these Bessel functions. In particular, the series (11) converges absolutely for all $Y \in \mathcal{S}^{m \times m}$ (Gross and Richards 1987, Theorem 6.3).

For $\text{Re}(\nu) > \frac{1}{2}(m - 2)$, the Bessel function A_ν also satisfies Herz’s generalization of the classical *Poisson integral* (Herz 1955, Eq. (3.6’)): For any $m \times m$ matrix V ,

$$A_\nu(V'V) = \frac{1}{\pi^{m^2/2} \Gamma_m(\nu + \frac{1}{2})} \int_{Q'Q < I_m} \text{etr}(2iV'Q) (\det(I_m - Q'Q))^{\nu - \frac{1}{2}m} dQ, \tag{12}$$

where $i = \sqrt{-1}$ and the integral is with respect to Lebesgue measure on the set $\{Q \in \mathbb{R}^{m \times m} : Q'Q < I_m\}$. This result leads to an inequality that will arise repeatedly in the sequel.

Lemma 1 For $\text{Re}(\nu) > \frac{1}{2}(m - 2)$ and $V \in \mathbb{R}^{m \times m}$,

$$|A_\nu(V'V)| \leq \frac{1}{\Gamma_m(\nu + \frac{1}{2}(m + 1))}. \tag{13}$$

For $\text{Re}(\nu) > -1$, M symmetric, and $Z > 0$, the Bessel function of matrix argument satisfies the Laplace transform identity,

$$\int_{R>0} \text{etr}(-RZ) A_\nu(MR) (\det R)^\nu dR = \text{etr}(-MZ^{-1}) (\det Z)^{-\nu - \frac{1}{2}(m+1)}. \tag{14}$$

Indeed, this identity is Herz’s original definition of $A_\nu(R)$ (Herz 1955, Eq. (2.5)).

Herz (1955, Eq. (5.8)) proved a remarkable generalization of a classical formula called *Weber’s second exponential integral*: For $\text{Re}(\nu) > -1$, $m \times m$ symmetric matrices Λ and M , and $Z > 0$,

$$\int_{R>0} \text{etr}(-RZ) A_\nu(\Lambda R) A_\nu(MR) (\det R)^\nu \, dR = (\det Z)^{-\nu - \frac{1}{2}(m+1)} \text{etr}(-(\Lambda + M)Z^{-1}) A_\nu(-\Lambda Z^{-1} M Z^{-1}). \tag{15}$$

Let $a, b \in \mathbb{C}$ where $-b + \frac{1}{2}(j + 1) \notin \mathbb{N}$, $j = 1, \dots, m$. The *confluent hypergeometric function of matrix argument* is defined, for $Y \in \mathcal{S}^{m \times m}$, as

$${}_1F_1(a; b; Y) = \sum_{k=0}^\infty \frac{1}{k!} \sum_{|\kappa|=k} \frac{[a]_\kappa}{[b]_\kappa} C_\kappa(Y).$$

We will make repeated use of *Kummer’s formula* (see Herz 1955, Eq. (2.8); Muirhead 1982, p. 265; Richards 2010, Section 35.8):

$${}_1F_1(a; b; Y) = \text{etr}(Y) {}_1F_1(b - a; b; -Y). \tag{16}$$

The Laplace transform relationship between the functions A_ν and ${}_1F_1$ is that for $\text{Re}(a) > \frac{1}{2}(m - 1)$, symmetric M , and $Z > 0$,

$$\Gamma_m(\nu + \frac{1}{2}(m + 1)) \int_{R>0} A_\nu(MR) (\det R)^{a - \frac{1}{2}(m+1)} \text{etr}(-RZ) \, dR = \Gamma_m(a) (\det Z)^{-a} {}_1F_1(a; \nu + \frac{1}{2}(m + 1); -MZ^{-1}); \tag{17}$$

see Herz (1955, p. 489). This result can also be proved by expressing $A_\nu(MR)$ as a series of zonal polynomials and then applying (9) to integrate term by term.

For partitions κ and σ , we denote by $\binom{\kappa}{\sigma}$ the *generalized binomial coefficient* (see Muirhead 1982, pp. 267–269; Richards 2010, Eq. (35.6.3)). For $\gamma > -1$ and $Y \in \mathcal{S}^{m \times m}$, the (*generalized*) *Laguerre polynomial* $L_\kappa^{(\gamma)}(Y)$, corresponding to κ , is defined as

$$L_\kappa^{(\gamma)}(Y) = [\gamma + \frac{1}{2}(m + 1)]_\kappa C_\kappa(I_m) \sum_{s=0}^{|\kappa|} \sum_{|\sigma|=s} \binom{\kappa}{\sigma} \frac{C_\sigma(-Y)}{[\gamma + \frac{1}{2}(m + 1)]_\sigma C_\sigma(I_m)}. \tag{18}$$

Setting $Y = 0$ in (18), we obtain

$$L_\kappa^{(\gamma)}(0) = [\gamma + \frac{1}{2}(m + 1)]_\kappa C_\kappa(I_m).$$

The *normalized (generalized) Laguerre polynomial* corresponding to κ is defined by

$$\mathcal{L}_\kappa^{(\gamma)}(Y) := \left(|\kappa|! L_\kappa^{(\gamma)}(0) \right)^{-1/2} L_\kappa^{(\gamma)}(Y), \tag{19}$$

$Y \in \mathcal{S}^{m \times m}$. By Muirhead (1982, p. 281), the polynomials $\mathcal{L}_\kappa^{(\gamma)}$ are orthonormal with respect to the Wishart distribution $W(\gamma + \frac{1}{2}(m + 1), I_m)$:

$$\frac{1}{\Gamma_m(\gamma + \frac{1}{2}(m + 1))} \int_{Y>0} \mathcal{L}_\kappa^{(\gamma)}(Y)\mathcal{L}_\sigma^{(\gamma)}(Y) (\det Y)^\gamma \text{etr}(-Y) dY = \begin{cases} 1, & \kappa = \sigma \\ 0, & \kappa \neq \sigma \end{cases} \tag{20}$$

By Muirhead (1982, p. 282), for $\gamma > -1$ and $Z > 0$, there holds the Laplace transform,

$$\begin{aligned} & \int_{Y>0} \text{etr}(-YZ)(\det Y)^\gamma L_\kappa^{(\gamma)}(Y) dY \\ &= [\gamma + \frac{1}{2}(m + 1)]_\kappa \Gamma_m(\gamma + \frac{1}{2}(m + 1))(\det Z)^{-\gamma - \frac{1}{2}(m+1)} C_\kappa(I_m - Z^{-1}). \end{aligned} \tag{21}$$

Further, by Muirhead (1982, p. 284), for $\gamma > -1$ and $Z \in \mathcal{S}^{m \times m}$,

$$\text{etr}(-Z)L_\kappa^{(\gamma)}(Z) = \int_{Y>0} \text{etr}(-Y)(\det Y)^\gamma C_\kappa(Y) A_\gamma(ZY) dY. \tag{22}$$

Lemma 2 *Let $Z > 0$ and $\gamma > -1$, then*

$$|L_\kappa^{(\gamma)}(Z)| \leq \text{etr}(Z) [\gamma + \frac{1}{2}(m + 1)]_\kappa C_\kappa(I_m). \tag{23}$$

Also, for $v \in \mathbb{R}$, $v > 0$,

$$\begin{aligned} & \int_{Y>0} \text{etr}(-vY)(\text{tr } Y)(\det Y)^\gamma L_\kappa^{(\gamma)}(Y) dY \\ &= [\gamma + \frac{1}{2}(m + 1)]_\kappa \Gamma_m(\gamma + \frac{1}{2}(m + 1)) C_\kappa(I_m) \\ & \quad \times (v - 1)^{|\kappa| - 1} v^{-[m(\gamma + (m+1)/2) + |\kappa| + 1]} (m(\gamma + \frac{1}{2}(m + 1))(v - 1) - |\kappa|). \end{aligned} \tag{24}$$

For $\nu \in \mathbb{C}$ such that $-\nu + \frac{1}{2}(j - m) \notin \mathbb{N}$, for all $j = 1, \dots, m$, and $X, Y \in \mathcal{S}^{m \times m}$, the Bessel function (of the first kind) of order ν with two matrix arguments is defined as the infinite series

$$A_\nu(X, Y) = \frac{1}{\Gamma_m(\nu + \frac{1}{2}(m + 1))} \sum_{k=0}^\infty \frac{(-1)^k}{k!} \sum_{|\kappa|=k} \frac{C_\kappa(X)C_\kappa(Y)}{[\nu + \frac{1}{2}(m + 1)]_\kappa C_\kappa(I_m)}.$$

It is straightforward from (5) and (11) to see that

$$A_\nu(X, Y) = \int_{O(m)} A_\nu(HXH'Y) dH, \tag{25}$$

$X > 0, Y \in \mathcal{S}^{m \times m}$ (Muirhead 1982, p. 260). Also, by applying (13), for $X, Y > 0$, we obtain

$$|A_\nu(X, Y)| \leq \frac{1}{\Gamma_m(\nu + \frac{1}{2}(m + 1))}. \tag{26}$$

2.3 Lipschitz properties of the Bessel functions of matrix argument and a related expectation

For $X, Y \in \mathcal{S}^{m \times m}$, the inner product between X and Y is defined by $\langle X, Y \rangle = \text{tr}(XY)$, and the Frobenius norm of X is defined by $\|X\|_F^2 = \langle X, X \rangle = \text{tr}(X^2)$. The Frobenius norm satisfies the triangle inequality, $\|X + Y\|_F \leq \|X\|_F + \|Y\|_F$, and it is also sub-multiplicative, $\|XY\|_F \leq \|X\|_F \cdot \|Y\|_F$ (Horn and Johnson 1990, p. 291).

The following result, which provides a Lipschitz property of the Bessel function A_ν , will be needed in Sect. 6 to establish the consistency of the test statistic T_n^2 .

Lemma 3 For $T > 0, Y_1 > 0$, and $Y_2 > 0$,

$$\left\| A_\nu(T, Y_1) - A_\nu(T, Y_2) \right\|_F \leq 2m^{3/4} \|T\|_F^{1/2} \|Y_1 - Y_2\|_F^{1/2} / \Gamma_m(\alpha). \tag{27}$$

Proof From the integral representation (12) for A_ν and the triangle inequality, we obtain

$$\begin{aligned} & |A_\nu(Y_1) - A_\nu(Y_2)| \\ & \leq \frac{1}{\pi^{m^2/2} \Gamma_m(\alpha - \frac{1}{2}m)} \int_{Q'Q < I_m} |\text{etr}(2iY_1^{1/2}Q) - \text{etr}(2iY_2^{1/2}Q)| d\mu(Q), \end{aligned}$$

where $d\mu(Q) := (\det(I_m - Q'Q))^{\alpha - \frac{1}{2}(2m+1)} dQ$. Setting $\theta_j := 2 \text{tr}(Y_1^{1/2}Q)$, $j = 1, 2$, and using the identity

$$|e^{i\theta_1} - e^{i\theta_2}|^2 = 4 \sin^2\left(\frac{1}{2}(\theta_1 - \theta_2)\right),$$

we obtain

$$|A_\nu(Y_1) - A_\nu(Y_2)| \leq \frac{2}{\pi^{m^2/2} \Gamma_m(\alpha - \frac{1}{2}m)} \int_{Q'Q < I_m} |\sin(\text{tr}(Y_1^{1/2} - Y_2^{1/2})Q)| d\mu(Q).$$

By the well-known inequality, $|\sin t| \leq |t|, t \in \mathbb{R}$; the sub-multiplicative property of the Frobenius norm; and the Cauchy–Schwarz inequality, we have

$$\begin{aligned} |\sin(\text{tr}(Y_1^{1/2} - Y_2^{1/2})Q)| & \leq |\text{tr}(Y_1^{1/2} - Y_2^{1/2})Q| \\ & \leq \left\| Y_1^{1/2} - Y_2^{1/2} \right\|_F \cdot (\text{tr}(QQ'))^{1/2} \leq m^{1/2} \left\| Y_1^{1/2} - Y_2^{1/2} \right\|_F, \end{aligned}$$

due to the fact that $Q'Q < I_m$. Further, by Wihler (2009, Eq. (3.2)),

$$\left\| Y_1^{1/2} - Y_2^{1/2} \right\|_F \leq m^{1/4} \|Y_1 - Y_2\|_F^{1/2}.$$

Combining the above inequalities, we obtain

$$\begin{aligned} \left| A_v(Y_1) - A_v(Y_2) \right| &\leq \frac{2m^{3/4}}{\pi^{m^2/2}\Gamma_m(\alpha - \frac{1}{2}m)} \|Y_1 - Y_2\|_F^{1/2} \int_{Q'Q < I_m} d\mu(Q) \\ &= \frac{2m^{3/4}}{\Gamma_m(\alpha)} \|Y_1 - Y_2\|_F^{1/2}, \end{aligned} \tag{28}$$

since

$$\int_{Q'Q < I_m} d\mu(Q) = \frac{\pi^{m^2/2}\Gamma_m(\alpha - \frac{1}{2}m)}{\Gamma_m(\alpha)}.$$

By (25), (28), and the sub-multiplicative property of the Frobenius norm, we obtain

$$\begin{aligned} \left| A_v(T, Y_1) - A_v(T, Y_2) \right| &\leq \int_{O(m)} |A_v(HTH'Y_1) - A_v(HTH'Y_2)| dH \\ &\leq \frac{2m^{3/4}}{\Gamma_m(\alpha)} \int_{O(m)} \|HTH'Y_1 - HTH'Y_2\|_F^{1/2} dH \\ &\leq \frac{2m^{3/4}}{\Gamma_m(\alpha)} \int_{O(m)} \|HTH'\|_F^{1/2} \|Y_1 - Y_2\|_F^{1/2} dH. \end{aligned}$$

Since $\|HTH'\|_F = \|T\|_F$ for $H \in O(m)$, and $\int_{O(m)} dH = 1$, we obtain the desired result. □

We will also need some Lipschitz properties of the gradient of the Bessel functions A_v . We use the usual notation for Kronecker’s delta, viz. $\delta_{ij} = 1$ or 0 for $i = j$ or $i \neq j$, respectively. For $Z = (z_{ij}) \in \mathcal{S}^{m \times m}$, the *gradient operator* is the $m \times m$ matrix

$$\nabla_Z = \left(\frac{1}{2}(1 + \delta_{ij}) \frac{\partial}{\partial z_{ij}} \right)_{i,j=1,\dots,m}.$$

Let $F : \mathcal{S}^{m \times m} \rightarrow \mathbb{C}$ be a C^1 function; that is, F is differentiable of order one and its partial derivatives are continuous. The *Taylor expansion of order one* of F , at $Z_0 \in \mathcal{S}^{m \times m}$, is

$$F(Z) = F(Z_0) + \langle Z - Z_0, \nabla_U F(U) \rangle, \tag{29}$$

where $U = tZ + (1 - t)Z_0$, for some $t \in [0, 1]$.

Lemma 4 For $T, Z > 0$,

$$\nabla_Z A_v(T, Z) = \int_{O(m)} M^{1/2} \left[\nabla_Y A_v(Y) \right] M^{1/2} dH, \tag{30}$$

where $M := HTH'$ and $Y := M^{1/2} Z M^{1/2}$.

All of our interchanges of derivatives and integrals are justifiable *via* Burkil and Burkil (2002, p. 289, Theorem 8.72), so we will perform such interchanges without further citation. Also, various positive constants arise in the calculations, and we denote them generically by $c, c_j, C_j, j \geq 1$.

Recall from Bishop et al. (2017, p. 28) the *multilinear operator norm* $\|\cdot\|$ which, in our context, is defined as follows: Let K_{ij} be the (i, j) th element of a $m \times m$ matrix K and $(V_{ij})_{kl}$ be the (k, l) th element of $V_{ij} := (\nabla_Y \otimes Y^{1/2})_{ij}$, the (i, j) th block in the tensor product $\nabla_Y \otimes Y^{1/2}$, then

$$((\nabla_Y \otimes Y^{1/2}) \cdot K)_{kl} = \sum_i \sum_j K_{ij} (V_{ij})_{kl},$$

and we define

$$\left\| \nabla_Y \otimes Y^{1/2} \right\| := \sup_{\|K\|_F=1} \|(\nabla_Y \otimes Y^{1/2}) \cdot K\|_F.$$

Lemma 5 Let Q be an $m \times m$ matrix such that $0 < QQ' < I_m$. Also, let Y be an $m \times m$ positive definite matrix. Then, there exists a constant $c > 0$ such that

$$\|\nabla_Y(\text{tr } QY^{1/2})\|_F \leq c (\lambda_{\min}(Y))^{-1/2}. \tag{31}$$

Lemma 6 For $T, Z > 0$, there exists a constant $C > 0$ such that

$$\|\nabla_Z A_v(T, Z)\|_F \leq C \|T\|_F (\lambda_{\min}(T))^{-1/2} (\lambda_{\min}(Z))^{-1/2}. \tag{32}$$

In the following result, we present a Lipschitz property of the Bessel functions of matrix argument. As the proof uses techniques (from Billingsley 1979; Del Moral and Niclas 2018; Kågström 1977) that are significantly different from those appearing generally in classical multivariate statistical analysis, we present the details in Sect. S.10.

Proposition 1 For $T, Z_1, Z_2 > 0$, there exist constants $C_1, C_2 > 0$ such that

$$\begin{aligned} & \left\| \nabla_{Z_1} A_v(T, Z_1) - \nabla_{Z_2} A_v(T, Z_2) \right\|_F \\ & \leq \frac{\|Z_1 - Z_2\|_F^{1/2} \|T\|_F^{3/2}}{\lambda_{\min}(Z_2^{1/2})} \left[\frac{C_1}{\lambda_{\min}(T)\lambda_{\min}(Z_1^{1/2})} + \frac{C_2}{\lambda_{\min}(T^{1/2})} \right]. \end{aligned} \tag{33}$$

Throughout the paper, if X is a random entity, we denote expectation with respect to the distribution of X by E_X or simply by E whenever the context is clear.

Let X be a Wishart-distributed random matrix, $X \sim W_m(\alpha, I_m)$, and define for $m \times m$ positive definite matrices T the *matrix-valued* function

$$g(T) = E \left[\alpha^{-1} X^{1/2} \nabla_Z A_\nu(T, Z) X^{1/2} \Big|_{Z=\alpha^{-1}X} \right]. \tag{34}$$

Lemma 7 For $T > 0$,

$$\text{tr } g(T) = -\frac{\alpha^{-1}}{\Gamma_m(\alpha)} (\text{tr } T) \text{etr}(-\alpha^{-1}T). \tag{35}$$

Proposition 2 For $T > 0$,

$$g(T) = -\frac{\alpha^{-1}}{m \Gamma_m(\alpha)} (\text{tr } T) \text{etr}(-\alpha^{-1}T) I_m. \tag{36}$$

Proof For $Y > 0$, define the function

$$\phi(Y) := \nabla_Z A_\nu(T, Z) \Big|_{Z=\alpha^{-1}Y}. \tag{37}$$

By (34), $g(T) = E \left[\alpha^{-1} X^{1/2} \phi(X) X^{1/2} \right]$, where $X \sim W_m(\alpha, I_m)$. Since the distribution of X is orthogonally invariant, i.e., $X \stackrel{d}{=} H'XH$ for all $H \in O(m)$, then

$$\begin{aligned} Hg(T)H' &= HE \left[\alpha^{-1} (H'XH)^{1/2} \phi(H'XH) (H'XH)^{1/2} \right] H' \\ &= E \left[\alpha^{-1} X^{1/2} H \phi(H'XH) H' X^{1/2} \right]. \end{aligned} \tag{38}$$

By (37),

$$\begin{aligned} \phi(H'XH) &= \nabla_Z A_\nu(T, Z) \Big|_{Z=\alpha^{-1}H'XH} \\ &= \nabla_{H'ZH} A_\nu(T, H'ZH) \Big|_{H'ZH=\alpha^{-1}H'XH}. \end{aligned}$$

By Maass (1971, p. 64), $\nabla_{H'ZH} = H' \nabla_Z H$, and it follows that

$$\begin{aligned} \phi(H'XH) &= H' \nabla_Z H A_\nu(T, H'ZH) \Big|_{Z=\alpha^{-1}X} \\ &= H' \nabla_Z A_\nu(T, H'ZH) \Big|_{Z=\alpha^{-1}X} H. \end{aligned}$$

However, $A_\nu(T, H'ZH) = A_\nu(T, Z)$ for all $H \in O(m)$; therefore,

$$\phi(H'XH) = H' \nabla_Z A_\nu(T, Z) \Big|_{Z=\alpha^{-1}X} H = H' \phi(X)H.$$

Substituting this result into (38) we obtain, for all $H \in O(m)$,

$$Hg(T)H' = E \left[\alpha^{-1} X^{1/2} \phi(X) X^{1/2} \right] = g(T).$$

Since $Hg(T)H' = g(T)$ for all $H \in O(m)$ then, by *Schur's Lemma* (Shilov 1977, p. 315), $g(T)$ is a scalar matrix: $g(T) = \gamma_1 I_m$ for some scalar γ_1 . By taking traces and applying (35), we obtain

$$m\gamma_1 = \text{tr } \gamma_1 I_m = \text{tr } g(T) = -\frac{\alpha^{-1}}{\Gamma_m(\alpha)} (\text{tr } T) \text{etr}(-\alpha^{-1}T);$$

therefore,

$$\gamma_1 = -\frac{\alpha^{-1}}{m \Gamma_m(\alpha)} (\text{tr } T) \text{etr}(-\alpha^{-1}T).$$

The proof is now complete. □

2.4 Hankel transforms of matrix argument

Let $X > 0$ be a random matrix with probability density function $f(X)$. For $\text{Re}(\nu) > \frac{1}{2}(m - 2)$, we define the *Hankel transform of order ν of X* as the function

$$\mathcal{H}_{X,\nu}(T) = E_X \left[\Gamma_m(\nu + \frac{1}{2}(m + 1)) A_\nu(TX) \right], \tag{39}$$

$T > 0$. The Hankel transform satisfies the following properties:

Lemma 8 For $\text{Re}(\nu) > \frac{1}{2}(m - 2)$, $|\mathcal{H}_{X,\nu}(T)| \leq 1$ for all $T > 0$, and $\mathcal{H}_{X,\nu}(T)$ is a continuous function of T .

Example 1 Let $X \sim W_m(\alpha, \Sigma)$, $\alpha > \frac{1}{2}(m - 1)$, $\Sigma > 0$. For $T > 0$, it follows from definition (39) of the Hankel transform that

$$\mathcal{H}_{X,\nu}(T) = \frac{\Gamma_m(\nu + \frac{1}{2}(m + 1))}{\Gamma_m(\alpha)} (\det \Sigma)^\alpha \int_{X>0} A_\nu(TX) (\det X)^{\alpha - \frac{1}{2}(m + 1)} \text{etr}(-\Sigma X) dX.$$

Applying (17) to calculate this integral, we obtain

$$\mathcal{H}_{X,\nu}(T) = {}_1F_1\left(\alpha; \nu + \frac{1}{2}(m + 1); -T\Sigma^{-1}\right). \tag{40}$$

For the case in which $\nu = \alpha - \frac{1}{2}(m + 1)$, (40) reduces to

$$\mathcal{H}_{X, \alpha - \frac{1}{2}(m+1)}(T) = {}_1F_1(\alpha; \alpha; -T\Sigma^{-1}) = \text{etr}(-T\Sigma^{-1}).$$

Example 2 Let $Z \sim W_m(\alpha, I_m)$ and $X > 0$ be an $m \times m$ random matrix that is independent of Z . For $T > 0$,

$$\begin{aligned} E_Z \mathcal{H}_{X, \nu}(T^{1/2} Z T^{1/2}) &= E_X \mathcal{H}_{Z, \nu}(T^{1/2} X T^{1/2}) \\ &= E_X {}_1F_1(\alpha; \nu + \frac{1}{2}(m + 1); -TX). \end{aligned} \tag{41}$$

To prove this result, we again apply (39) and the independence of X and Z , obtaining

$$\begin{aligned} E_Z \mathcal{H}_{X, \nu}(T^{1/2} Z T^{1/2}) &= E_{Z, X} \Gamma_m(\nu + \frac{1}{2}(m + 1)) A_\nu(T^{1/2} Z T^{1/2} X) \\ &= E_X E_Z \Gamma_m(\nu + \frac{1}{2}(m + 1)) A_\nu(T^{1/2} Z T^{1/2} X). \end{aligned}$$

Since $A_\nu(T^{1/2} Z T^{1/2} X) = A_\nu(T^{1/2} X T^{1/2} Z)$, we have

$$E_Z \mathcal{H}_{X, \nu}(T^{1/2} Z T^{1/2}) = E_X \mathcal{H}_{Z, \nu}(T^{1/2} X T^{1/2}). \tag{42}$$

Applying Example 1, we obtain

$$E_X \mathcal{H}_{Z, \nu}(T^{1/2} X T^{1/2}) = E_X {}_1F_1(\alpha; \nu + \frac{1}{2}(m + 1); -TX). \tag{43}$$

Combining (42) and (43), we obtain (41).

In particular, if $\nu = \alpha - \frac{1}{2}(m + 1)$ then, by Kummer’s formula (16), we obtain

$$\begin{aligned} E_Z \mathcal{H}_{X, \alpha - \frac{1}{2}(m+1)}(T^{1/2} Z T^{1/2}) &= E_X \mathcal{H}_{Z, \alpha - \frac{1}{2}(m+1)}(T^{1/2} X T^{1/2}) \\ &= E_X \text{etr}(-TX), \end{aligned}$$

the Laplace transform of X .

Throughout the remainder of the paper, if X and Y are random entities we write $X \stackrel{d}{=} Y$ whenever X and Y have the same distribution. If $\{X_n, n \geq 1\}$ is a sequence of random entities, we write $X_n \xrightarrow{d} X$ whenever X_n converges in distribution to X .

Theorem 1 (Uniqueness of the Hankel transform). *Let X and Y be $m \times m$ positive definite random matrices with Hankel transforms $\mathcal{H}_{X, \nu}$ and $\mathcal{H}_{Y, \nu}$, respectively. If $\mathcal{H}_{X, \nu} = \mathcal{H}_{Y, \nu}$, then $X \stackrel{d}{=} Y$.*

Proof Suppose that $\mathcal{H}_{X, \nu} = \mathcal{H}_{Y, \nu}$. Let Ψ_X and Ψ_Y be the Laplace transforms of X and Y , respectively, and let $Z \sim W_m(\nu + \frac{1}{2}(m + 1), I_m)$, independently of X and Y . Applying Example 2 twice, we obtain for all $T > 0$,

$$\begin{aligned} \Psi_X(T) &= E_X \operatorname{etr}(-TX) = E_Z \mathcal{H}_{X,v}(T^{1/2} Z T^{1/2}) \\ &= E_Z \mathcal{H}_{Y,v}(T^{1/2} Z T^{1/2}) = E_Y \operatorname{etr}(-TY) = \Psi_Y(T). \end{aligned}$$

By the uniqueness of multivariate Laplace transforms (Farrell 1985, p. 16), we obtain $X \stackrel{d}{=} Y$. □

We denote by L_v^2 the space of functions $\phi : \mathcal{P}_+^{m \times m} \rightarrow \mathbb{C}$ such that

$$\int_{\mathcal{P}_+^{m \times m}} |\phi(X)|^2 (\det X)^{-v} dX < \infty.$$

The following inversion theorem is obtained by applying the Hankel inversion theory of Herz (1955, Section 3). We refer to Hadjicosta (2019) for full details.

Theorem 2 (Inversion of the Hankel transform). *Let $X > 0$ be a random matrix with Hankel transform $\mathcal{H}_{X,v}$, and with a probability density function $f \in L_v^2$. Then,*

$$f(X) = \frac{1}{\Gamma_m(v + \frac{1}{2}(m + 1))} \int_{\mathcal{P}_+^{m \times m}} A_v(TX) (\det TX)^v \mathcal{H}_{X,v}(T) dT.$$

Theorem 3 (Hankel Continuity). *Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of $m \times m$ positive definite random matrices with corresponding Hankel transforms $\{\mathcal{H}_{X_n}, n \in \mathbb{N}\}$. If there exists an $m \times m$ positive semi-definite random matrix X with Hankel transform \mathcal{H}_X such that $X_n \xrightarrow{d} X$, then for each $T > 0$,*

$$\lim_{n \rightarrow \infty} \mathcal{H}_{X_n}(T) = \mathcal{H}_X(T). \tag{44}$$

Conversely, suppose there exists a function $\mathcal{H} : \mathcal{P}_+^{m \times m} \rightarrow \mathbb{R}$ such that $\mathcal{H}(T) \rightarrow 1$ as $T \rightarrow 0$, \mathcal{H} is continuous at 0, and (44) holds. Then \mathcal{H} is the Hankel transform of an $m \times m$ positive semi-definite random matrix X , and $X_n \xrightarrow{d} X$.

The next result constitutes a characterization of the Wishart distributions using the Hankel transform $\mathcal{H}_{X,v}$, where $\operatorname{Re}(v) > \frac{1}{2}(m - 2)$. The result enables the extension, to the Wishart case, of some results of Baringhaus and Taherizadeh (2013) on a supremum norm test statistic.

Theorem 4 *Let X be an $m \times m$ positive definite random matrix with an orthogonally invariant distribution and Hankel transform $\mathcal{H}_{X,v}$. If there exist $\epsilon > 0$ and $\alpha > \frac{1}{2}(m - 1)$ such that for all T satisfying $0 < T \leq \epsilon I_m$,*

$$\mathcal{H}_{X,v}(T) = {}_1F_1(\alpha; v + \frac{1}{2}(m + 1); -T),$$

then $X \sim W_m(\alpha, I_m)$.

We refer to Hadjicosta (2019), who gave three proofs of this result. We provide in the supplementary Sect. S.10 the briefest of those proofs, which uses the principle of analytic continuation.

2.5 Orthogonally invariant Hankel transforms of matrix argument

Definition 1 Let X be an $m \times m$ positive definite random matrix with p.d.f. $f(X)$. For $\text{Re}(v) > \frac{1}{2}(m - 2)$ and $T > 0$, we define the *orthogonally invariant Hankel transform of order v of X* as

$$\tilde{\mathcal{H}}_{X,v}(T) = E_X \left[\Gamma_m(v + \frac{1}{2}(m + 1)) A_v(T, X) \right]. \tag{45}$$

Remark 1 By (25) and definition (39) of $\mathcal{H}_{X,v}$, we have

$$\tilde{\mathcal{H}}_{X,v}(T) = \int_{O(m)} \mathcal{H}_{X,v}(HTH') \, dH. \tag{46}$$

Further, since $\int_{O(m)} dH = 1$, then $\tilde{\mathcal{H}}_{X,v}$ also satisfies the properties in Lemma 8.

Let $a, b \in \mathbb{C}$, where $-b + \frac{1}{2}(j - 1) \notin \mathbb{N}$, for all $j = 1, \dots, m$. The *confluent hypergeometric function of two matrix arguments* is defined, for $X, Y \in \mathcal{S}^{m \times m}$, as the infinite series,

$${}_1F_1(a; b; X, Y) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} \frac{[a]_{\kappa}}{[b]_{\kappa}} \frac{C_{\kappa}(X)C_{\kappa}(Y)}{C_{\kappa}(I_m)}.$$

It is clear from the definition that ${}_1F_1(a; b; X, I_m) = {}_1F_1(a; b; X)$. Similar to (25), it follows from (5) that for $X, Y \in \mathcal{S}^{m \times m}$,

$${}_1F_1(a; b; X, Y) = \int_{O(m)} {}_1F_1(a; b; HXH'Y) \, dH. \tag{47}$$

Example 3 Let $X \sim W_m(\alpha, \Sigma)$ where $\alpha > \frac{1}{2}(m - 1)$ and $\Sigma > 0$. For $T > 0$, it follows from Example 1, (46), and (47) that

$$\begin{aligned} \tilde{\mathcal{H}}_{X,v}(T) &= \int_{O(m)} {}_1F_1(\alpha; v + \frac{1}{2}(m + 1); -HTH'\Sigma^{-1}) \, dH \\ &= {}_1F_1(\alpha; v + \frac{1}{2}(m + 1); -T, \Sigma^{-1}). \end{aligned}$$

Theorem 5 (Uniqueness of orthogonally invariant Hankel transforms). *Let X and Y be $m \times m$ positive definite random matrices with orthogonally invariant distributions and orthogonally invariant Hankel transforms $\tilde{\mathcal{H}}_{X,v}$ and $\tilde{\mathcal{H}}_{Y,v}$, respectively. Then $\tilde{\mathcal{H}}_{X,v} = \tilde{\mathcal{H}}_{Y,v}$ if and only if $X \stackrel{d}{=} Y$.*

3 The test statistic and its limiting null distribution

Let X_1, \dots, X_n be independent, identically distributed (i.i.d.), $m \times m$ positive definite random matrices, each with probability density function $f(X)$ and positive definite

mean $\mu = E(X_1)$. We assume also that the density function of X_1 is of the form

$$f(X_1) = f_0(\mu^{-1/2}X_1\mu^{-1/2}), \tag{48}$$

where f_0 is orthogonally invariant.

Lemma 9 *Under the assumption (48), the distribution of $\mu^{-1/2}X_1\mu^{-1/2}$ is orthogonally invariant.*

Proof Let $\tilde{Y} = \mu^{-1/2}X_1\mu^{-1/2}$; then, $X_1 = \mu^{1/2}\tilde{Y}\mu^{1/2}$ and the Jacobian of the transformation from X_1 to \tilde{Y} is $(\det \mu)^{(m+1)/2}$ (Muirhead 1982, p. 58). Therefore, the p.d.f. of \tilde{Y} is

$$g(\tilde{Y}) = (\det \mu)^{(m+1)/2} f(\mu^{1/2}\tilde{Y}\mu^{1/2}) = (\det \mu)^{(m+1)/2} f_0(\tilde{Y}).$$

Since f_0 is orthogonally invariant, then it follows that g is orthogonally invariant. \square

We denote by P the distribution of X_1 . On the basis of the random sample X_1, \dots, X_n , we wish to test the null hypothesis, $H_0 : P \in \{W_m(\alpha, \Sigma), \Sigma > 0\}$, against the alternative, $H_1 : P \notin \{W_m(\alpha, \Sigma), \Sigma > 0\}$, where α is known.

Since Σ is unspecified by H_0 , the data X_1, \dots, X_n cannot be used to construct a test statistic. Let $\bar{X}_n = n^{-1} \sum_{j=1}^n X_j$ be the sample mean, and define $Y_j = \bar{X}_n^{-1/2} X_j \bar{X}_n^{-1/2}$, $j = 1, \dots, n$. Under H_0 , the distribution of Y_1, \dots, Y_n does not depend on Σ , so a test statistic can be based on them. Let P_0 denote the probability measure corresponding to the $W_m(\alpha, I_m)$ distribution. For $\text{Re}(v) > \frac{1}{2}(m - 2)$, define the empirical orthogonally invariant Hankel transform of order v of Y_1, \dots, Y_n as

$$\tilde{\mathcal{H}}_{n,v}(T) = \Gamma_m(v + \frac{1}{2}(m + 1)) \frac{1}{n} \sum_{j=1}^n A_v(T, Y_j), \tag{49}$$

$T > 0$. Further, define the test statistic

$$T_{n,v}^2 = n \int_{T>0} [\tilde{\mathcal{H}}_{n,v}(T) - {}_1F_1(\alpha; v + \frac{1}{2}(m + 1); -T/\alpha)]^2 dP_0(T). \tag{50}$$

Suppose that H_0 is valid; then, $E(X_1) = \alpha\Sigma^{-1}$ and, for large n , we can expect that $Y_j = \bar{X}_n^{-1/2} X_j \bar{X}_n^{-1/2} \simeq \alpha^{-1}\Sigma^{1/2} X_j \Sigma^{1/2}$, almost surely. By the continuous mapping theorem, the sequence $A_v(T, Y_j)$ should approximate the i.i.d. sequence $A_v(T, \alpha^{-1}\Sigma^{1/2} X_j \Sigma^{1/2})$, $j = 1, \dots, n$, for each $T > 0$ and for sufficiently large n . Applying to (49), the strong law of large numbers, we can expect that, for large n , $\tilde{\mathcal{H}}_{n,v}(T) \simeq \tilde{\mathcal{H}}_{\alpha^{-1}\Sigma^{1/2}X_1\Sigma^{1/2},v}(T)$, almost surely.

By Example 3, we deduce that

$$\begin{aligned} \tilde{\mathcal{H}}_{\alpha^{-1}\Sigma^{1/2}X_1\Sigma^{1/2},v}(T) &= {}_1F_1(\alpha; v + \frac{1}{2}(m + 1); -\alpha^{-1}T, I_m) \\ &= {}_1F_1(\alpha; v + \frac{1}{2}(m + 1); -\alpha^{-1}T), \end{aligned}$$

for $T > 0$. Therefore, by Lemma 9 and Theorem 5, small values of $T_{n,\nu}^2$ provide strong evidence in support of H_0 , so we will reject H_0 for large values of $T_{n,\nu}^2$.

For the remainder of the paper, we set

$$\nu = \alpha - \frac{1}{2}(m + 1). \tag{51}$$

Since $\nu > \frac{1}{2}(m - 2)$ then $\alpha > \frac{1}{2}(2m - 1)$. We also denote $T_{n,\nu}^2$ and $\tilde{\mathcal{H}}_{n,\nu}$ by T_n^2 and $\tilde{\mathcal{H}}_n$, respectively. By Kummer’s formula (16), statistic (50) becomes

$$T_n^2 = n \int_{T>0} \left[\tilde{\mathcal{H}}_n(T) - \text{etr}(-T/\alpha) \right]^2 dP_0(T). \tag{52}$$

This integral represents T_n^2 as a weighted integral of the squared difference between the empirical orthogonally invariant Hankel transform $\tilde{\mathcal{H}}_n$ and its almost sure limit under the null hypothesis.

We now evaluate the statistic T_n^2 for a given random sample.

Proposition 3 *The test statistic (52) is a V-statistic of order 2. Specifically,*

$$T_n^2 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n h(Y_i, Y_j)$$

where, for $X, Y > 0$,

$$h(X, Y) = \Gamma_m(\alpha) \text{etr}(-X - Y) A_\nu(-X, Y) - \left(\frac{\alpha}{\alpha + 1}\right)^{m\alpha} \left[\text{etr}\left(-\frac{\alpha}{\alpha + 1}X\right) + \text{etr}\left(-\frac{\alpha}{\alpha + 1}Y\right) \right] + \left(\frac{2}{\alpha} + 1\right)^{-m\alpha}.$$

Proof By expanding the integrand in (52), we find three integrals to be computed. First,

$$\begin{aligned} \int_{T>0} \tilde{\mathcal{H}}_n^2(T) dP_0(T) &= \frac{1}{n^2} \int_{T>0} \left(\sum_{i=1}^n \Gamma_m(\alpha) A_\nu(T, Y_i) \right)^2 dP_0(T) \\ &= \frac{\Gamma_m(\alpha)}{n^2} \sum_{i=1}^n \sum_{j=1}^n \int_{T>0} A_\nu(T, Y_i) A_\nu(T, Y_j) (\det T)^\nu \text{etr}(-T) dT. \end{aligned}$$

By (25) and Fubini’s theorem,

$$\begin{aligned} &\int_{T>0} A_\nu(T, Y_i) A_\nu(T, Y_j) (\det T)^\nu \text{etr}(-T) dT \\ &= \int_{O(m)} \int_{O(m)} \int_{T>0} A_\nu(HTH'Y_i) A_\nu(KTK'Y_j) (\det T)^\nu \text{etr}(-T) dT dH dK. \end{aligned} \tag{53}$$

Writing $A_\nu(HTH'Y_j) = A_\nu(H'Y_jHT)$, $j = 1, \dots, n$, and applying Herz's generalization (15) of Weber's second exponential integral, we find that (53) equals

$$\begin{aligned} & \int_{O(m)} \int_{O(m)} \text{etr}(-H'Y_iH - K'Y_jK) A_\nu(-H'Y_iHK'Y_jK) dH dK \\ &= \text{etr}(-Y_i - Y_j) \int_{O(m)} \int_{O(m)} A_\nu(-H'Y_iHK'Y_jK) dH dK. \end{aligned} \tag{54}$$

On the right-hand side of (54), we replace H by HK and apply the group invariance of the Haar measure and its normalization; then, we find that (54) reduces to

$$\int_{O(m)} A_\nu(-H'Y_iHY_j) dH \equiv A_\nu(-Y_i, Y_j).$$

Therefore,

$$\int_{T>0} \tilde{\mathcal{H}}_n^2(T) dP_0(T) = \frac{\Gamma_m(\alpha)}{n^2} \sum_{i=1}^n \sum_{j=1}^n \text{etr}(-Y_i - Y_j) A_\nu(-Y_i, Y_j).$$

The second integral to be calculated is

$$\begin{aligned} & \int_{T>0} \tilde{\mathcal{H}}_n(T) \text{etr}(-T/\alpha) dP_0(T) \\ &= \frac{\Gamma_m(\alpha)}{n} \sum_{i=1}^n \int_{T>0} A_\nu(T, Y_i) (\det T)^\nu \text{etr}(-(I_m + \alpha^{-1}I_m)T) dT. \end{aligned} \tag{55}$$

Similar to the previous calculation, we use (25) to express $A_\nu(T, Y_i)$ as an average over $O(m)$ and apply Fubini's theorem to reverse the order of integration. The resulting integral is a special case of (14), and we find that (55) equals

$$\begin{aligned} & \frac{\Gamma_m(\alpha)}{n} \left(\frac{\alpha}{\alpha + 1}\right)^{m\alpha} \sum_{i=1}^n \text{etr}\left(-\frac{\alpha}{\alpha + 1}Y_i\right) \\ &\equiv \frac{\Gamma_m(\alpha)}{2n^2} \left(\frac{\alpha}{\alpha + 1}\right)^{m\alpha} \sum_{i=1}^n \sum_{j=1}^n \left[\text{etr}\left(-\frac{\alpha}{\alpha + 1}Y_i\right) + \text{etr}\left(-\frac{\alpha}{\alpha + 1}Y_j\right) \right]. \end{aligned}$$

The third and last integral, which we evaluate using the multivariate gamma integral (10), is

$$\begin{aligned} \int_{T>0} \text{etr}(-2T/\alpha) dP_0(T) &= (\det(2\alpha^{-1}I_m + I_m))^{-\alpha} \\ &= (2\alpha^{-1} + 1)^{-m\alpha} \equiv \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \left(\frac{2}{\alpha} + 1\right)^{-m\alpha}. \end{aligned}$$

Collecting together the three terms, we obtain the stated result. □

We denote by $L^2 = L^2(P_0)$ the space of (equivalence classes of) orthogonally invariant Borel measurable functions $f : \mathcal{P}_+^{m \times m} \rightarrow \mathbb{C}$ that are square-integrable with respect to the probability measure P_0 , i.e., for which $\int_{X>0} |f(X)|^2 dP_0(X) < \infty$. The space L^2 is a separable Hilbert space when equipped with the inner product

$$\langle f, g \rangle_{L^2} = \int_{X>0} f(X) \overline{g(X)} dP_0(X),$$

and the norm $\|f\|_{L^2} = \sqrt{\langle f, f \rangle_{L^2}}$, for $f, g \in L^2$. Moreover, the set of normalized Laguerre polynomials $\{\mathcal{L}_\kappa^{(v)}\}$, with κ ranging over all partitions, defined in (19), forms an orthonormal basis for the space L^2 ; see Herz (1955, p. 502) and Constantine (1966, Section 3).

We now define the stochastic process

$$\mathcal{Z}_n(T) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \left[\Gamma_m(\alpha) A_v(T, Y_j) - \text{etr}(-T/\alpha) \right], \tag{56}$$

$T > 0$. We view the random field $\mathcal{Z}_n := \{\mathcal{Z}_n(T), T > 0\}$ as a random element in L^2 since, as we now show, its sample paths are in L^2 . The following result is a consequence of (49), (52), and (56).

Lemma 10 *The test statistic (52) can be written as*

$$\mathbf{T}_n^2 = \int_{T>0} (\mathcal{Z}_n(T))^2 dP_0(T) = \|\mathcal{Z}_n\|_{L^2}^2.$$

In particular, $\|\mathcal{Z}_n\|_{L^2}^2 < \infty$.

Remark 2 By Gupta and Richards (1987, Example 1.4) (Y_1, \dots, Y_n) has a matrix Liouville distribution, of the second kind, that does not depend on Σ . Therefore, without loss of generality, we will set $\Sigma = I_m$ in deriving the limiting null distribution of \mathbf{T}_n^2 .

For $j = 1, \dots, n$, $Y_j = \bar{X}_n^{-1/2} X_j \bar{X}_n^{-1/2}$ and $Z_j = X_j^{1/2} \bar{X}_n^{-1} X_j^{1/2}$ have the same spectrum; this is proved by showing that Y_j and Z_j have the same characteristic polynomial. Consequently,

$$A_v(T, Y_j) = A_v(T, Z_j), \tag{57}$$

$j = 1, \dots, n$, so we can replace Y_j by Z_j in definition (49) of the test statistic.

We now state the main result of this section.

Theorem 6 *Let $m \geq 2$, $\alpha > \max\{\frac{1}{2}(2m - 1), \frac{1}{2}(m + 3)\}$, and X_1, \dots, X_n be i.i.d. P_0 -distributed random matrices. Also, let $\mathcal{Z}_n := (\mathcal{Z}_n(T), T > 0)$ be the random field*

defined in (56). Then, there exists a centered Gaussian random field $\mathcal{Z} := (\mathcal{Z}(T), T > 0)$, with sample paths in L^2 and with covariance function,

$$K(S, T) = \text{etr}(-\alpha^{-1}(S+T)) \left[\Gamma_m(\alpha) A_\nu(-\alpha^{-2}S, T) - \frac{1}{\alpha^3 m} (\text{tr } S)(\text{tr } T) - 1 \right], \tag{58}$$

$S, T > 0$, such that $\mathcal{Z}_n \xrightarrow{d} \mathcal{Z}$ in L^2 as $n \rightarrow \infty$. Moreover,

$$T_n^2 \xrightarrow{d} \int_{T>0} \mathcal{Z}^2(T) \, dP_0(T).$$

The remainder of this section is devoted to proving Theorem 6, so readers who wish to postpone reading the detailed derivation may continue directly to Sect. 4.

In the sequel, we will use for various symmetric matrices V the shorthand notation

$$\nabla A_\nu(T, V) \equiv \nabla_Z A_\nu(T, Z) \Big|_{Z=V}.$$

Proof of Theorem 6 By (29), the Taylor expansion of the Bessel function $A_\nu(T, Z)$ at (T, Z_0) is

$$A_\nu(T, Z) = A_\nu(T, Z_0) + \langle Z - Z_0, \nabla A_\nu(T, U) \rangle, \tag{59}$$

where $U = tZ + (1-t)Z_0, t \in [0, 1]$. Setting $Z = Z_j$ and $Z_0 = \alpha^{-1}X_j, j = 1, \dots, n$, in (59), we obtain the Taylor expansion of order one of $A_\nu(T, Z_j)$ at $(T, \alpha^{-1}X_j)$:

$$A_\nu(T, Z_j) = A_\nu(T, \alpha^{-1}X_j) + \langle Z_j - \alpha^{-1}X_j, \nabla A_\nu(T, U_j) \rangle, \tag{60}$$

where $U_j = tZ_j + (1-t)\alpha^{-1}X_j, t \in [0, 1]$. Define $M_n = \bar{X}_n^{-1/2}(\alpha I_m - \bar{X}_n)\bar{X}_n^{-1/2}$; then (60) becomes

$$A_\nu(T, Z_j) = A_\nu(T, \alpha^{-1}X_j) + \langle \alpha^{-1}X_j^{1/2} M_n X_j^{1/2}, \nabla A_\nu(T, U_j) \rangle.$$

Adding and subtracting $\langle \alpha^{-1}X_j^{1/2} M_n X_j^{1/2}, \nabla A_\nu(T, \alpha^{-1}X_j) \rangle$ on the right-hand side, we obtain

$$\begin{aligned} A_\nu(T, Z_j) &= A_\nu(T, \alpha^{-1}X_j) + \langle \alpha^{-1}X_j^{1/2} M_n X_j^{1/2}, \nabla A_\nu(T, U_j) \rangle \\ &\quad + \langle \alpha^{-1}X_j^{1/2} M_n X_j^{1/2}, \nabla A_\nu(T, U_j) - \nabla A_\nu(T, \alpha^{-1}X_j) \rangle \\ &= A_\nu(T, \alpha^{-1}X_j) + \langle M_n, \alpha^{-1}X_j^{1/2} \nabla A_\nu(T, \alpha^{-1}X_j) X_j^{1/2} \rangle \\ &\quad + \langle M_n, \alpha^{-1}X_j^{1/2} [\nabla A_\nu(T, U_j) - \nabla A_\nu(T, \alpha^{-1}X_j)] X_j^{1/2} \rangle, \tag{61} \end{aligned}$$

where the second equality is obtained by permuting terms cyclically in the inner product. For $T > 0$ and $X_j > 0, j = 1, \dots, n$, define the function

$$g(T, X_j) := \alpha^{-1} X_j^{1/2} \nabla A_\nu(T, \alpha^{-1}X_j) X_j^{1/2}.$$

We remark that as X_1, \dots, X_n are i.i.d. then $E_{X_j}g(T, X_j)$ does not depend on j ; hence,

$$g(T) := E_{X_j}g(T, X_j) = E(\alpha^{-1} X_j^{1/2} \nabla A_v(T, \alpha^{-1} X_j) X_j^{1/2})$$

is a function evaluated earlier; by (36),

$$g(T) = -\frac{\alpha^{-1}}{m \Gamma_m(\alpha)} (\text{tr } T) \text{etr}(-\alpha^{-1} T) I_m.$$

Define the random fields $\mathcal{Z}_{n,1}(T)$, $\mathcal{Z}_{n,2}(T)$ and $\mathcal{Z}_{n,3}(T)$, $T > 0$, by

$$\begin{aligned} \mathcal{Z}_{n,1}(T) &= \frac{\Gamma_m(\alpha)}{\sqrt{n}} \sum_{j=1}^n \left[A_v(T, \alpha^{-1} X_j) + \langle M_n, g(T, X_j) \rangle - \frac{\text{etr}(-\alpha^{-1} T)}{\Gamma_m(\alpha)} \right], \\ \mathcal{Z}_{n,2}(T) &= \frac{\Gamma_m(\alpha)}{\sqrt{n}} \sum_{j=1}^n \left[A_v(T, \alpha^{-1} X_j) + \langle M_n, g(T) \rangle - \frac{\text{etr}(-\alpha^{-1} T)}{\Gamma_m(\alpha)} \right], \\ \mathcal{Z}_{n,3}(T) &= \frac{\Gamma_m(\alpha)}{\sqrt{n}} \sum_{j=1}^n \left[A_v(T, \alpha^{-1} X_j) + \langle \alpha^{-1}(\alpha I_m - X_j), g(T) \rangle - \frac{\text{etr}(-\alpha^{-1} T)}{\Gamma_m(\alpha)} \right], \end{aligned}$$

The random fields $\mathcal{Z}_{n,k}$, $k = 1, 2, 3$ arise as follows. To define $\mathcal{Z}_{n,1}(T)$, we use the first two terms in (61). To define $\mathcal{Z}_{n,2}(T)$, we use the same expression from $\mathcal{Z}_{n,1}(T)$ except that the term $g(T, X_j)$ is replaced by its expected value $g(T)$, which is given by (36). To define $\mathcal{Z}_{n,3}(T)$, we replace the term M_n in $\mathcal{Z}_{n,2}(T)$ by a constant multiple of $\alpha I_m - X_j$, the constant being obtained by applying the law of large numbers to $\bar{X}_n^{-1/2}$. We will show that

$$\mathcal{Z}_{n,3} \xrightarrow{d} \mathcal{Z} \text{ in } L^2, \tag{62}$$

$$\|\mathcal{Z}_n - \mathcal{Z}_{n,1}\|_{L^2} \xrightarrow{p} 0, \tag{63}$$

$$\|\mathcal{Z}_{n,1} - \mathcal{Z}_{n,2}\|_{L^2} \xrightarrow{p} 0, \tag{64}$$

$$\|\mathcal{Z}_{n,2} - \mathcal{Z}_{n,3}\|_{L^2} \xrightarrow{p} 0. \tag{65}$$

By writing \mathcal{Z}_n as

$$\mathcal{Z}_n = \mathcal{Z}_n - \mathcal{Z}_{n,1} + \mathcal{Z}_{n,1} - \mathcal{Z}_{n,2} + \mathcal{Z}_{n,2} - \mathcal{Z}_{n,3} + \mathcal{Z}_{n,3},$$

it will follow that $\mathcal{Z}_n \xrightarrow{d} \mathcal{Z}$ in L^2 (Billingsley 1968, p. 25).

To establish (62), define for $T > 0$,

$$\mathcal{Z}_{n,3,j}(T) := \Gamma_m(\alpha) A_v(T, \alpha^{-1} X_j) + \Gamma_m(\alpha) \langle \alpha^{-1}(\alpha I_m - X_j), g(T) \rangle - \text{etr}(-\alpha^{-1} T), \tag{66}$$

$j = 1, \dots, n$. Since $X_j \sim W_m(\alpha, I_m)$ then $E(X_j - \alpha I_m) = 0$, and therefore, since the trace and the expectation are linear operators, we deduce that

$$E\left[\langle \alpha^{-1}(\alpha I_m - X_j), g(T) \rangle\right] = \text{tr}\left[\alpha^{-1} E(\alpha I_m - X_j) \cdot g(T)\right] = 0.$$

Also, by Example 3 and (16), we have $E\left[\Gamma_m(\alpha)A_v(T, \alpha^{-1}X_j)\right] = \text{etr}(-\alpha^{-1}T)$; so $E(\mathcal{Z}_{n,3,j}(T)) = 0$ for all $T > 0$ and $j = 1, \dots, n$, and it is also clear that $\mathcal{Z}_{n,3,1}, \dots, \mathcal{Z}_{n,3,n}$ are independent and identically distributed random elements in L^2 .

We now show that $E(\|\mathcal{Z}_{n,3,j}\|_{L^2}^2) < \infty$ for $j = 1, \dots, n$. We have

$$\begin{aligned} E(\|\mathcal{Z}_{n,3,j}\|_{L^2}^2) &= E \int_{T>0} \mathcal{Z}_{n,3,j}^2(T) dP_0(T) \\ &= E \int_{T>0} \left[\Gamma_m(\alpha)A_v(T, \alpha^{-1}X_j) \right. \\ &\quad \left. + \Gamma_m(\alpha)\langle \alpha^{-1}(\alpha I_m - X_j), g(T) \rangle - \text{etr}(-\alpha^{-1}T) \right]^2 dP_0(T). \end{aligned}$$

By the Cauchy–Schwarz inequality, $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$ for $a, b, c \in \mathbb{R}$; so to prove that $E(\|\mathcal{Z}_{n,3,j}\|_{L^2}^2) < \infty$, it suffices to prove that

$$E \int_{T>0} [\Gamma_m(\alpha)A_v(T, \alpha^{-1}X_j)]^2 dP_0(T) < \infty, \tag{67}$$

$$E \int_{T>0} [\Gamma_m(\alpha)\langle \alpha^{-1}(\alpha I_m - X_j), g(T) \rangle]^2 dP_0(T) < \infty, \tag{68}$$

and

$$E \int_{T>0} \text{etr}(-2\alpha^{-1}T) dP_0(T) < \infty. \tag{69}$$

To establish (67), we apply (26) to obtain

$$E \int_{T>0} \left[\Gamma_m(\alpha)A_v(T, \alpha^{-1}X_j) \right]^2 dP_0(T) \leq E \int_{T>0} 1 \cdot dP_0(T) = 1.$$

To prove (68), write

$$\begin{aligned} \langle \alpha I_m - X_j, g(T) \rangle^2 &= (\text{tr}[(\alpha I_m - X_j) \cdot g(T)])^2 \\ &= \left(\frac{\alpha^{-1}}{m\Gamma_m(\alpha)} \right)^2 (\text{tr}(\alpha I_m - X_j))^2 (\text{tr} T)^2 \text{etr}(-2\alpha^{-1}T); \end{aligned}$$

therefore, the integral in (68) is a constant multiple of

$$E(\text{tr}(\alpha I_m - X_j))^2 \cdot \int_{T>0} (\text{tr} T)^2 \text{etr}(-2\alpha^{-1}T) dP_0(T).$$

Since $(\text{tr}(\alpha I_m - X_j))^2$ is a polynomial in X_j , its expectation is finite because the moment-generating function of X exists. As for

$$\int_{T>0} (\text{tr } T)^2 \text{etr}(-2\alpha^{-1}T) \, dP_0(T),$$

again this integral is finite because $(\text{tr } T)^2$ is a polynomial and $\text{etr}(-2\alpha^{-1}T) \, dP_0(T)$, after normalization, is a Wishart measure. For the same reason, (69) is valid.

In summary, for $T > 0$ and $j = 1, \dots, n$, $\mathcal{Z}_{n,3,1}, \dots, \mathcal{Z}_{n,3,n}$ are i.i.d. random elements in L^2 with $E(\mathcal{Z}_{n,3,j}(T)) = 0$ and $E(\|\mathcal{Z}_{n,3,j}\|_{L^2}^2) < \infty$. Therefore, by the central limit theorem in L^2 ,

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n \mathcal{Z}_{n,3,j} \xrightarrow{d} \mathcal{Z},$$

where $\mathcal{Z} := (\mathcal{Z}(T), T > 0)$ is a centered Gaussian random element in L^2 . Moreover, \mathcal{Z} has the same covariance operator as $\mathcal{Z}_{n,3,1}$. It is well known that the covariance operator of the random element $\mathcal{Z}_{n,3,1}$ is uniquely determined by the covariance function of the random field $\mathcal{Z}_{n,3,1}$ (Gikhman and Skorokhod 1980, pp. 218–219).

We now show that the function $K(S, T)$ in (58) is the covariance function of $\mathcal{Z}_{n,3,1}$. Noting that $E[\mathcal{Z}_{n,3,1}(T)] = 0$ for all $T > 0$, we obtain

$$\begin{aligned} K(S, T) &= \text{Cov}[\mathcal{Z}_{n,3,1}(S), \mathcal{Z}_{n,3,1}(T)] \\ &= \text{Cov}[\mathcal{Z}_{n,3,1}(S) + \text{etr}(-\alpha^{-1}S), \mathcal{Z}_{n,3,1}(T) + \text{etr}(-\alpha^{-1}T)] \\ &= E[(\mathcal{Z}_{n,3,1}(S) + \text{etr}(-\alpha^{-1}S)) \cdot (\mathcal{Z}_{n,3,1}(T) + \text{etr}(-\alpha^{-1}T))] \\ &\quad - \text{etr}(-\alpha^{-1}(S + T)). \end{aligned}$$

By (66),

$$\begin{aligned} &E[(\mathcal{Z}_{n,3,1}(S) + \text{etr}(-\alpha^{-1}S)) \cdot (\mathcal{Z}_{n,3,1}(T) + \text{etr}(-\alpha^{-1}T))] \\ &= E \left[\Gamma_m(\alpha) A_\nu(S, \alpha^{-1}X_1) + \Gamma_m(\alpha) \langle \alpha^{-1}(\alpha I_m - X_1), g(S) \rangle \right] \\ &\quad \times E \left[\Gamma_m(\alpha) A_\nu(T, \alpha^{-1}X_1) + \Gamma_m(\alpha) \langle \alpha^{-1}(\alpha I_m - X_1), g(T) \rangle \right], \end{aligned} \tag{70}$$

so the calculation of $K(S, T)$ reduces to evaluating the four terms obtained by expanding the product on the right-hand side of (70).

The first term in the product in (70) is

$$\begin{aligned} &E[\Gamma_m(\alpha)]^2 A_\nu(S, \alpha^{-1}X_1) A_\nu(T, \alpha^{-1}X_1) \\ &= \Gamma_m(\alpha) \int_{X>0} A_\nu(S, \alpha^{-1}X) A_\nu(T, \alpha^{-1}X) (\det X)^\nu \text{etr}(-X) \, dX. \end{aligned} \tag{71}$$

By (15), (25), and Fubini’s theorem, we find that this term equals

$$\begin{aligned} & \Gamma_m(\alpha) \int_{O(m)} \int_{O(m)} \int_{X>0} A_\nu(\alpha^{-1} HSH'X) A_\nu(\alpha^{-1} KTK'X) \\ & \quad \times (\det X)^\nu \operatorname{etr}(-X) \, dX \, dH \, dK \\ &= \Gamma_m(\alpha) \int_{O(m)} \int_{O(m)} A_\nu(-\alpha^{-2} HSH' KTK') \operatorname{etr}(-\alpha^{-1} (HSH' + KTK')) \, dH \, dK. \end{aligned}$$

Since $\operatorname{etr}(-\alpha^{-1} (HSH' + KTK')) = \operatorname{etr}(-\alpha^{-1} (S + T))$, and

$$\int_{O(m)} A_\nu(-\alpha^{-2} HSH' KTK') \, dH = A_\nu(-\alpha^{-2} S, KTK') = A_\nu(-\alpha^{-2} S, T),$$

we conclude that the first term equals

$$\Gamma_m(\alpha) \operatorname{etr}(-\alpha^{-1} (S + T)) A_\nu(-\alpha^{-2} S, T). \tag{72}$$

The second term in the product in (70) is

$$\begin{aligned} & \alpha^{-1} [\Gamma_m(\alpha)]^2 E \left[A_\nu(S, \alpha^{-1} X_1) \cdot \langle (\alpha I_m - X_1), g(T) \rangle \right] \\ &= \frac{\Gamma_m(\alpha)}{\alpha^2 m} E \left[A_\nu(S, \alpha^{-1} X_1) \cdot \langle (X_1 - \alpha I_m), (\operatorname{tr} T) \operatorname{etr}(-\alpha^{-1} T) I_m \rangle \right] \\ &= \frac{1}{\alpha^2 m} \Gamma_m(\alpha) (\operatorname{tr} T) \operatorname{etr}(-\alpha^{-1} T) E \left[\left((\operatorname{tr} X_1) - 1 \right) \cdot A_\nu(S, \alpha^{-1} X_1) \right]. \end{aligned} \tag{73}$$

We have seen earlier that

$$\Gamma_m(\alpha) E A_\nu(S, \alpha^{-1} X_1) = \operatorname{etr}(-\alpha^{-1} S). \tag{74}$$

Also, by (25),

$$E(\operatorname{tr} X_1) A_\nu(S, \alpha^{-1} X_1) = \int_{O(m)} \operatorname{tr} E \left(X_1 \cdot A_\nu(\alpha^{-1} HSH' X_1) \right) \, dH. \tag{75}$$

By Muirhead (1982, p. 442), $E \left(X_1 \cdot A_\nu(\alpha^{-1} HSH' X_1) \right)$ is a multiple of the expected value of a non-central Wishart-distributed random matrix, with distribution denoted by $W_m(\alpha, I_m, \Omega)$, where, in our setting, the matrix of non-centrality parameters is $\Omega = -\alpha^{-1} HSH'$. Hence,

$$\begin{aligned} E \left((\operatorname{tr} X_1) A_\nu(\alpha^{-1} HSH' X_1) \right) &= \operatorname{tr} E \left(X_1 A_\nu(\alpha^{-1} HSH' X_1) \right) \\ &= \frac{1}{\Gamma_m(\alpha)} \operatorname{tr} (\alpha I_m - \alpha^{-1} \Omega) \operatorname{etr}(-\alpha^{-1} \Omega) \\ &= \frac{1}{\Gamma_m(\alpha)} (\alpha m - \alpha^{-1} \operatorname{tr} S) \operatorname{etr}(-\alpha^{-1} S). \end{aligned}$$

Substituting this result into (75), we obtain

$$E \left((\text{tr } X_1) A_v(S, \alpha^{-1} X_1) \right) = \frac{1}{\Gamma_m(\alpha)} \text{etr}(-\alpha^{-1} S) [\alpha m - \alpha^{-1} \text{tr } S]. \tag{76}$$

Substituting (74) and (76) into (73), we find that the second term equals

$$-(\alpha^3 m)^{-1} (\text{tr } S)(\text{tr } T) \text{etr}(-\alpha^{-1}(S + T)).$$

The third term in the product in (70) is $\alpha^{-1}[\Gamma_m(\alpha)]^2 E[A_v(T, \alpha^{-1} X_1) \cdot \langle (\alpha I_m - X_1), g(S) \rangle]$, which is the same as the second term but with S and T interchanged.

The fourth term in the product in (70) is

$$\begin{aligned} & [\alpha^{-1} \Gamma_m(\alpha)]^2 E \left[\langle (\alpha I_m - X_1), g(S) \rangle \cdot \langle (\alpha I_m - X_1), g(T) \rangle \right] \\ &= [\alpha^{-1} \Gamma_m(\alpha)]^2 E \left[\text{tr}((\alpha I_m - X_1) \cdot g(S)) \cdot \text{tr}((\alpha I_m - X_1) \cdot g(T)) \right]. \end{aligned} \tag{77}$$

Using the explicit formula for $g(T)$ from (36), we find that the expectation on the right-hand side of (77) equals

$$\begin{aligned} & \frac{1}{[\Gamma_m(\alpha)]^2} \left[(\text{tr } S) \text{etr}(-\alpha^{-1} S)(\text{tr } T) \text{etr}(-\alpha^{-1} T) \right. \\ & \quad - 2(m\alpha)^{-1} (\text{tr } S) \text{etr}(-\alpha^{-1} S)(\text{tr } T) \text{etr}(-\alpha^{-1} T) E(\text{tr } X_1) \\ & \quad \left. + (m\alpha)^{-2} (\text{tr } S) \text{etr}(-\alpha^{-1} S)(\text{tr } T) \text{etr}(-\alpha^{-1} T) E(\text{tr } X_1)^2 \right]. \end{aligned} \tag{78}$$

Applying (3), (4), and (9), we obtain $E(\text{tr } X_1) = \alpha m$ and $E[(\text{tr } X_1)^2] = \alpha m(\alpha m + 1)$. Substituting these results into (78), we find that the fourth term equals $(\alpha^3 m)^{-1} (\text{tr } S)(\text{tr } T) \text{etr}(-\alpha^{-1}(S + T))$. Combining all four terms, we obtain (58).

To establish (63), we show first that $\text{tr}[(\sqrt{n} M_n)^2] \equiv \|\bar{X}_n^{-1/2} \sqrt{n}(\alpha I_m - \bar{X}_n) \bar{X}_n^{-1/2}\|_F^2$ converges in distribution to a random variable with finite variance. By the multivariate central limit theorem, $\sqrt{n} \text{vech}(\alpha I_m - \bar{X}_n)$ converges in distribution to a normal random vector. Also, by the law of large numbers, $\bar{X}_n^{-1} \xrightarrow{P} \alpha^{-1} I_m$. Therefore, by Slutsky's theorem, $\sqrt{n} \text{vech}(M_n)$ converges in distribution to a normal random vector, so it follows from the continuous mapping theorem that $\text{tr}[(\sqrt{n} M_n)^2]$ converges in distribution to a random variable with finite variance.

By the Taylor expansion (61),

$$\begin{aligned} Z_n - Z_{n,1} &= \frac{\Gamma_m(\alpha)}{\sqrt{n}} \sum_{j=1}^n \left\langle M_n, \alpha^{-1} X_j^{1/2} \left(\nabla A_v(T, U_j) - \nabla A_v(T, \alpha^{-1} X_j) \right) X_j^{1/2} \right\rangle \\ &= \frac{\alpha^{-1} \Gamma_m(\alpha)}{n} \sum_{j=1}^n \left\langle \sqrt{n} M_n, X_j^{1/2} \left(\nabla A_v(T, U_j) - \nabla A_v(T, \alpha^{-1} X_j) \right) X_j^{1/2} \right\rangle. \end{aligned}$$

Define

$$V_n := \frac{1}{n^2} \int_{T>0} \text{tr} \left[\sum_{j=1}^n X_j^{1/2} \left(\nabla A_v(T, U_j) - \nabla A_v(T, \alpha^{-1} X_j) \right) X_j^{1/2} \right]^2 dP_0(T).$$

By the Cauchy–Schwarz inequality,

$$\|Z_n - Z_{n,1}\|_{L^2}^2 \leq [\alpha^{-1} \Gamma_m(\alpha)]^2 \text{tr} \left[(\sqrt{n} M_n)^2 \right] \cdot V_n; \tag{79}$$

so we will establish (63) by proving that $V_n \xrightarrow{P} 0$.

By the triangle inequality and the sub-multiplicative property of the Frobenius norm, we have

$$\begin{aligned} & \text{tr} \left[\sum_{j=1}^n X_j^{1/2} \left(\nabla A_v(T, U_j) - \nabla A_v(T, \alpha^{-1} X_j) \right) X_j^{1/2} \right]^2 \\ &= \left\| \sum_{j=1}^n X_j^{1/2} \left(\nabla A_v(T, U_j) - \nabla A_v(T, \alpha^{-1} X_j) \right) X_j^{1/2} \right\|_F^2 \\ &\leq \left(\sum_{j=1}^n \|X_j\|_F \|\nabla A_v(T, U_j) - \nabla A_v(T, \alpha^{-1} X_j)\|_F \right)^2. \end{aligned}$$

Applying (33), we obtain

$$\begin{aligned} & \|\nabla A_v(T, U_j) - \nabla A_v(T, \alpha^{-1} X_j)\|_F \\ &\leq \frac{\|T\|_F^{3/2} \|U_j - \alpha^{-1} X_j\|_F^{1/2}}{\lambda_{\min}(X_j^{1/2})} \left[\frac{C_1}{\lambda_{\min}(T) \lambda_{\min}(U_j^{1/2})} + \frac{C_2}{\lambda_{\min}(T^{1/2})} \right]. \end{aligned}$$

Also, since $U_j = X_j^{1/2} [\alpha^{-1} I_m + t(\bar{X}_n^{-1} - \alpha^{-1} I_m)] X_j^{1/2}$, $t \in [0, 1]$, then

$$\begin{aligned} \|U_j - \alpha^{-1} X_j\|_F^{1/2} &= \|X_j^{1/2} [t(\bar{X}_n^{-1} - \alpha^{-1} I_m)] X_j^{1/2}\|_F^{1/2} \\ &= \|t X_j (\bar{X}_n^{-1} - \alpha^{-1} I_m)\|_F^{1/2} \leq \|X_j\|_F^{1/2} \|\bar{X}_n^{-1} - \alpha^{-1} I_m\|_F^{1/2}. \end{aligned}$$

Define

$$\begin{aligned} V_{n,1} &:= C_1^2 \|\bar{X}_n^{-1} - \alpha^{-1} I_m\|_F \left(\frac{1}{n} \sum_{j=1}^n \frac{\|X_j\|_F^{3/2}}{\lambda_{\min}(X_j^{1/2}) \lambda_{\min}(U_j^{1/2})} \right)^2 \\ &\cdot \int_{T>0} \frac{\|T\|_F^3}{[\lambda_{\min}(T)]^2} dP_0(T), \end{aligned}$$

and

$$V_{n,2} := C_2^2 \|\bar{X}_n^{-1} - \alpha^{-1} I_m\|_F \left(\frac{1}{n} \sum_{j=1}^n \frac{\|X_j\|_F^{3/2}}{\lambda_{\min}(X_j^{1/2})} \right)^2 \cdot \int_{T>0} \frac{\|T\|_F^3}{\lambda_{\min}(T)} dP_0(T).$$

By the Cauchy–Schwarz inequality, $V_n \leq V_{n,1} + V_{n,2}$, so it suffices to show that $V_{n,1}, V_{n,2} \xrightarrow{p} 0$.

We first establish that $V_{n,1} \xrightarrow{p} 0$. By the Cauchy–Schwarz inequality,

$$\begin{aligned} \left(\frac{1}{n} \sum_{j=1}^n \frac{\|X_j\|_F^{3/2}}{\lambda_{\min}(X_j^{1/2})\lambda_{\min}(U_j^{1/2})} \right)^2 &\leq \frac{1}{n} \sum_{j=1}^n \frac{\|X_j\|_F^3}{\lambda_{\min}(X_j)\lambda_{\min}(U_j)} \\ &= \frac{1}{n} \sum_{j=1}^n \frac{(\text{tr}(X_j^2))^{3/2}}{\lambda_{\min}(X_j)\lambda_{\min}(U_j)}. \end{aligned}$$

By Weyl’s inequality for the smallest eigenvalue of the sum of two symmetric matrices,

$$\begin{aligned} \lambda_{\min}(U_j) &\geq t\lambda_{\min}(X_j\bar{X}_n^{-1}) + (1-t)\alpha^{-1}\lambda_{\min}(X_j) \\ &\geq t\lambda_{\min}(X_j)\lambda_{\min}(\bar{X}_n^{-1}) + (1-t)\alpha^{-1}\lambda_{\min}(X_j) \\ &\geq \lambda_{\min}(X_j)\min\{\lambda_{\min}(\bar{X}_n^{-1}), \alpha^{-1}\}; \end{aligned}$$

therefore,

$$\frac{1}{n} \sum_{j=1}^n \frac{(\text{tr}(X_j^2))^{3/2}}{\lambda_{\min}(X_j)\lambda_{\min}(U_j)} \leq \frac{1}{\min\{\lambda_{\min}(\bar{X}_n^{-1}), \alpha^{-1}\}} \frac{1}{n} \sum_{j=1}^n \frac{(\text{tr}(X_j^2))^{3/2}}{[\lambda_{\min}(X_j)]^2}.$$

By the law of large numbers and the continuous mapping theorem, we have

$$\frac{\|\bar{X}_n^{-1} - \alpha^{-1} I_m\|_F}{\min\{\lambda_{\min}(\bar{X}_n^{-1}), \alpha^{-1}\}} \xrightarrow{p} 0.$$

Again by the law of large numbers,

$$\frac{1}{n} \sum_{j=1}^n \frac{(\text{tr}(X_j^2))^{3/2}}{[\lambda_{\min}(X_j)]^2} \xrightarrow{p} E_{P_0} \left(\frac{(\text{tr}(X^2))^{3/2}}{[\lambda_{\min}(X)]^2} \right).$$

Therefore, to complete the proof that $V_{n,1} \xrightarrow{p} 0$, we need to establish that

$$\int_{T>0} \frac{\|T\|_F^3}{[\lambda_{\min}(T)]^2} dP_0(T) < \infty \quad \text{and} \quad E_{P_0} \left(\frac{(\text{tr}(X^2))^{3/2}}{[\lambda_{\min}(X)]^2} \right) < \infty.$$

Since $\|T\|_F^3 = (\text{tr } T^2)^{3/2}$, then these criteria are the same, so we show that the first one holds.

For $T > 0$, we have $\text{tr } T^2 \leq (\text{tr } T)^2$ and hence $(\text{tr } T^2)^{3/2} \leq (\text{tr } T)^3$. By a result of Khatri (1966, Lemma 7, Eq. (20)),

$$\int_{T>0} \frac{(\text{tr } T)^3}{[\lambda_{\min}(T)]^2} dP_0(T) < \infty,$$

for $\alpha > \frac{1}{2}(m + 3)$, so it follows that $V_{n,1} \xrightarrow{P} 0$.

As for $V_{n,2} \xrightarrow{P} 0$, the proof is similar. By the Cauchy–Schwarz inequality,

$$\left(\frac{1}{n} \sum_{j=1}^n \frac{\|X_j\|_F^{3/2}}{\lambda_{\min}(X_j^{1/2})}\right)^2 \leq \frac{1}{n} \sum_{j=1}^n \frac{\|X_j\|_F^3}{\lambda_{\min}(X_j)} = \frac{1}{n} \sum_{j=1}^n \frac{(\text{tr } X_j^2)^{3/2}}{\lambda_{\min}(X_j)}.$$

Applying the law of large numbers, we obtain $\|\bar{X}_n^{-1} - \alpha^{-1}I_m\|_F \xrightarrow{P} 0$ and

$$\frac{1}{n} \sum_{j=1}^n \frac{(\text{tr } X_j^2)^{3/2}}{\lambda_{\min}(X_j)} \xrightarrow{P} E_{P_0} \left(\frac{(\text{tr } X^2)^{3/2}}{\lambda_{\min}(X)} \right).$$

Thus, to complete the proof of $V_{n,2} \xrightarrow{P} 0$, we need to establish that

$$\int_{T>0} \frac{\|T\|_F^3}{\lambda_{\min}(T)} dP_0(T) < \infty \quad \text{and} \quad E_{P_0} \left(\frac{(\text{tr } X^2)^{3/2}}{\lambda_{\min}(X)} \right) < \infty,$$

which are identical criteria. Since $\|T\|_F^3 = (\text{tr } T^2)^{3/2}$, it suffices to show that

$$\int_{T>0} \frac{(\text{tr } T^2)^{3/2}}{\lambda_{\min}(T)} dP_0(T) < \infty.$$

However, $\text{tr } T^2 \leq (\text{tr } T)^2$ so $(\text{tr } T^2)^{3/2} \leq (\text{tr } T)^3$ so, by Khatri (1966, Lemma 7, Eq. (20)),

$$\int_{T>0} \frac{(\text{tr } T)^3}{\lambda_{\min}(T)} dP_0(T) < \infty$$

for all $\alpha > \frac{1}{2}(m + 1)$. Therefore, $V_{n,2} \xrightarrow{P} 0$ for all $\alpha > \frac{1}{2}(2m - 1)$.

Since $0 \leq V_n \leq V_{n,1} + V_{n,2}$ then we obtain $V_n \xrightarrow{P} 0$ for all $\alpha > \max\{\frac{1}{2}(2m - 1), \frac{1}{2}(m + 3)\}$. By Slutsky’s theorem, $[\alpha^{-1}\Gamma_m(\alpha)]^2 \text{tr} [(\sqrt{n} M_n)^2] \cdot V_n \xrightarrow{d} 0$, and therefore $[\alpha^{-1}\Gamma_m(\alpha)]^2 \text{tr} [(\sqrt{n} M_n)^2] \cdot V_n \xrightarrow{P} 0$. Hence, by (79), $\|\mathcal{Z}_n - \mathcal{Z}_{n,1}\|_{L^2} \xrightarrow{P} 0$, for $\alpha > \max\{\frac{1}{2}(2m - 1), \frac{1}{2}(m + 3)\}$.

To establish (64), define $V_j := g(T, X_j) - g(T)$ for $T > 0$ and $j = 1, \dots, n$. Then,

$$\mathcal{Z}_{n,1} - \mathcal{Z}_{n,2} = \Gamma_m(\alpha) \left\langle M_n, \frac{1}{\sqrt{n}} \sum_{j=1}^n V_j \right\rangle$$

and therefore

$$\|\mathcal{Z}_{n,1} - \mathcal{Z}_{n,2}\|_{L^2}^2 \leq [\Gamma_m(\alpha)]^2 \text{tr}(M_n^2) \cdot \int_{T>0} \left[\text{tr} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n V_j \right)^2 \right] dP_0(T). \tag{80}$$

By the law of large numbers and the continuous mapping theorem, $\text{tr}(M_n^2) \xrightarrow{P} 0$. Since $g(T) = E[g(T, X_j)]$ then $E(V_j) = 0, j = 1, \dots, n$; also, V_1, \dots, V_n are i.i.d.

We now show that $E_{X_j} E_T \|V_j\|_F^2 < \infty$. First,

$$E_{X_j} E_T (\|V_j\|_F^2) = E_{X_j} \left(\int_{T>0} \|g(T, X_j) - g(T)\|_F^2 dP_0(T) \right).$$

By the triangle inequality,

$$\begin{aligned} \|g(T, X_j) - g(T)\|_F^2 &\leq \left(\|g(T, X_j)\|_F + \|g(T)\|_F \right)^2 \\ &\leq 2 \left(\|g(T, X_j)\|_F^2 + \|g(T)\|_F^2 \right). \end{aligned}$$

Therefore, it suffices to show that $E_{X_j} E_T \|g(T, X_j)\|_F^2$ and $E_T \|g(T)\|_F^2$ are finite.

Applying the sub-multiplicative property of the Frobenius norm and (32), we have

$$\begin{aligned} \|g(T, X_j)\|_F^2 &= \|X_j^{1/2} \nabla A_v(T, \alpha^{-1} X_j) X_j^{1/2}\|_F^2 \\ &\leq \|X_j\|_F^2 \|\nabla A_v(T, \alpha^{-1} X_j)\|_F^2 \\ &= c (\text{tr } X_j^2) (\lambda_{\min}(X_j))^{-1} (\text{tr } T^2) (\lambda_{\min}(T))^{-1}, \end{aligned}$$

$c > 0$; therefore,

$$E_{X_j} E_T \|g(T, X_j)\|_F^2 \leq c E_{X_j} \left[(\text{tr } X_j^2) (\lambda_{\min}(X_j))^{-1} \right] E_T \left[(\text{tr } T^2) (\lambda_{\min}(T))^{-1} \right].$$

By Khatri (1966, Lemma 7, Eq. (20)), $E_T \left[(\text{tr } T^2) (\lambda_{\min}(T))^{-1} \right] < \infty$ for $\alpha > \frac{1}{2}(m + 1)$. Since $X_j \sim W_m(\alpha, I_m)$ then also $E_{X_j} \left[(\text{tr } X_j^2) (\lambda_{\min}(X_j))^{-1} \right] < \infty, \alpha > \frac{1}{2}(m + 1)$. Thus, $E_{X_j} E_T \|g(T, X_j)\|_F^2 < \infty$ for all $\alpha > \frac{1}{2}(2m - 1)$. Also, since $\|g(T)\|_F^2 = \text{tr}[(g(T))^2]$ is a polynomial in T then $E_T \|g(T)\|_F^2 < \infty$ for $T \sim W_m(\alpha, I_m)$ since the moment-generating function of T exists.

We now vectorize the matrices V_1, \dots, V_n , denoting by $\text{vech}(V_1), \dots, \text{vech}(V_n)$ the corresponding vectors. Then, $\text{vech}(V_1), \dots, \text{vech}(V_n)$ are i.i.d. zero-mean random vectors with finite covariance matrices. By the central limit theorem, $n^{-1/2} \sum_{j=1}^n \text{vech}(V_j)$ converges in distribution to a normal random vector. Define

$$\mathcal{V}(T) := \left\| \frac{1}{\sqrt{n}} \sum_{j=1}^n V_j \right\|_F,$$

for $T > 0$; we regard \mathcal{V} as a random element in L^2 . Since $\|\cdot\|_F$ is a continuous function, it follows from the continuous mapping theorem that \mathcal{V} converges to a random element in L^2 and also that

$$\|\mathcal{V}\|_{L^2}^2 := \int_{T>0} \left\| \frac{1}{\sqrt{n}} \sum_{j=1}^n V_j \right\|_F^2 dP_0(T) = \int_{T>0} \text{tr} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n V_j \right)^2 dP_0(T)$$

converges in distribution to a random variable that has finite variance. Since $\text{tr}(M_n^2) \xrightarrow{P} 0$, by (80) and Slutsky’s theorem, we obtain $\|\mathcal{Z}_{n,1} - \mathcal{Z}_{n,2}\|_{L^2}^2 \xrightarrow{d} 0$; therefore $\|\mathcal{Z}_{n,1} - \mathcal{Z}_{n,2}\|_{L^2} \xrightarrow{P} 0$.

To establish (65), we observe that

$$\begin{aligned} \mathcal{Z}_{n,2} - \mathcal{Z}_{n,3} &= \frac{\Gamma_m(\alpha)}{\sqrt{n}} \sum_{j=1}^n (\langle M_n, g(T) \rangle - \langle \alpha^{-1}(\alpha I_m - X_j), g(T) \rangle) \\ &= \frac{\Gamma_m(\alpha)}{\sqrt{n}} (\langle nM_n, g(T) \rangle - \langle \alpha^{-1} \sum_{j=1}^n (\alpha I_m - X_j), g(T) \rangle) \\ &= \Gamma_m(\alpha) \text{tr} [(\bar{X}_n^{-1/2} \sqrt{n}(\alpha I_m - \bar{X}_n) \bar{X}_n^{-1/2} - \alpha^{-1} \sqrt{n}(\alpha I_m - \bar{X}_n))g(T)]. \end{aligned}$$

Substituting the now-familiar explicit formula for $g(T)$ from (36), we obtain

$$\begin{aligned} \|\mathcal{Z}_{n,2} - \mathcal{Z}_{n,3}\|_{L^2}^2 &= \frac{1}{\alpha^2 m^2} [\text{tr}(\bar{X}_n^{-1/2} \sqrt{n}(\alpha I_m - \bar{X}_n) \bar{X}_n^{-1/2} - \alpha^{-1} \sqrt{n}(\alpha I_m - \bar{X}_n))]^2 \\ &\quad \times \int_{T>0} (\text{tr } T)^2 \text{etr}(-2\alpha^{-1}T) dP_0(T), \end{aligned}$$

and as we have seen before, the latter integral is finite. Now, we observe that

$$\begin{aligned} \bar{X}_n^{-1/2} \sqrt{n}(\alpha I_m - \bar{X}_n) \bar{X}_n^{-1/2} - \alpha^{-1} \sqrt{n}(\alpha I_m - \bar{X}_n) \\ \equiv \sqrt{n}(\alpha I_m - \bar{X}_n)(\bar{X}_n^{-1} - \alpha^{-1} I_m). \end{aligned}$$

By the central limit theorem, $\sqrt{n} \text{vech}(\alpha I_m - \bar{X}_n)$ converges in distribution to a multivariate normal random vector; by the law of large numbers for random vectors,

$\bar{X}_n^{-1} \xrightarrow{P} \alpha^{-1} I_m$. By Slutsky’s theorem, $\sqrt{n}(\alpha I_m - \bar{X}_n)(\bar{X}_n^{-1} - \alpha^{-1} I_m) \xrightarrow{d} 0$, and so $\sqrt{n}(\alpha I_m - \bar{X}_n)(\bar{X}_n^{-1} - \alpha^{-1} I_m) \xrightarrow{P} 0$. Hence, by the continuous mapping theorem,

$$\left[\text{tr} \left(\bar{X}_n^{-1/2} \sqrt{n}(\alpha I_m - \bar{X}_n) \bar{X}_n^{-1/2} - \alpha^{-1} \sqrt{n}(\alpha I_m - \bar{X}_n) \right) \right]^2 \xrightarrow{P} 0;$$

and so $\|\mathcal{Z}_{n,2} - \mathcal{Z}_{n,3}\|_{L^2} \xrightarrow{P} 0$.

Finally, by the continuous mapping theorem in L^2 (Billingsley 1968, p. 31), we deduce that $\|\mathcal{Z}_n\|_{L^2}^2 \xrightarrow{d} \|\mathcal{Z}\|_{L^2}^2$, i.e.,

$$T_n^2 = \int_{T>0} \mathcal{Z}_n^2(T) \, dP_0(T) \xrightarrow{d} \int_{T>0} \mathcal{Z}^2(T) \, dP_0(T).$$

The proof now is complete. □

4 Eigenvalues and eigenfunctions of the covariance operator

The covariance operator $\mathcal{S} : L^2 \rightarrow L^2$ of the random element \mathcal{Z} is defined for $S > 0$ and $f \in L^2$ by

$$\mathcal{S}f(S) = \int_{S>0} K(S, T) f(T) \, dP_0(T),$$

where $K(S, T)$ is the covariance function defined in equation (58). Let $\{\delta_k : k \geq 1\}$ be the positive eigenvalues, listed in non-increasing order according to their multiplicities, of \mathcal{S} ; also, let $\{\chi_{1k}^2 : k \geq 1\}$ be i.i.d. χ_1^2 -distributed random variables. It is well known that the integrated squared process, $\int_{T>0} \mathcal{Z}^2(T) \, dP_0(T)$, has the same distribution as $\sum_{k=1}^\infty \delta_k \chi_{1k}^2$. This result follows from the Karhunen–Loève expansion of the Gaussian random field $\mathcal{Z}(T)$; see Le Maître and Knio (2010, Chapter 2). Therefore, the limiting null distribution of T_n^2 is the same as $\sum_{k=1}^\infty \delta_k \chi_{1k}^2$. Let us also denote by $\tilde{\delta}_k, k \geq 1$, an enumeration, listed in non-increasing order, of the *distinct* values of the eigenvalues δ_k . Further, we denote by $N(\tilde{\delta}_k)$ the corresponding multiplicities of the distinct eigenvalues $\tilde{\delta}_k$. Then, $T_n^2 \xrightarrow{d} \sum_{k \geq 1} \tilde{\delta}_k \chi_{N(\tilde{\delta}_k)}^2$, where $\{\chi_{N(\tilde{\delta}_k)}^2\}$ are independent χ^2 -distributed random variables and $\chi_{N(\tilde{\delta}_k)}^2$ has $N(\tilde{\delta}_k)$ degrees of freedom.

For $S, T > 0$, define

$$K_0(S, T) = \Gamma_m(\alpha) \text{etr}(-\alpha^{-1}(S + T)) A_\nu(-\alpha^{-2}S, T), \tag{81}$$

the first term in the covariance function defined in equation (58); by (71) and (72),

$$K_0(S, T) = [\Gamma_m(\alpha)]^2 \int_{X>0} A_\nu(S, \alpha^{-1}X) A_\nu(T, \alpha^{-1}X) \, dP_0(X). \tag{82}$$

We will first find the eigenvalues and eigenfunctions of the integral operator $\mathcal{S}_0 : L^2 \rightarrow L^2$, defined for $S > 0$ and f in L^2 by

$$\mathcal{S}_0 f(S) = \int_{T>0} K_0(S, T) f(T) \, dP_0(T). \tag{83}$$

Recall that $m \geq 2$ and $\alpha > \max\{\frac{1}{2}(2m - 1), \frac{1}{2}(m + 3)\}$. Throughout the remainder of this work, we use the notation

$$\beta = \left(\frac{\alpha + 4}{\alpha}\right)^{1/2} \quad \text{and} \quad b_\alpha = \left(1 + \frac{1}{2}\alpha(1 - \beta)\right)^{1/2}. \tag{84}$$

We also set

$$\rho_\kappa = \alpha^{m\alpha} b_\alpha^{4|\kappa|+2m\alpha} \tag{85}$$

for κ ranging over all partitions, and

$$\mathfrak{L}_\kappa^{(v)}(S) := \beta^{m\alpha/2} \operatorname{etr}((1 - \beta)S/2) \mathcal{L}_\kappa^{(v)}(\beta S). \tag{86}$$

Theorem 7 *The collection $\{(\rho_\kappa, \mathfrak{L}_\kappa^{(v)})\}$, where κ ranges over the set of all partitions, is a complete enumeration of the eigenvalues and eigenfunctions, respectively, of the operator \mathcal{S}_0 . Further, the eigenfunctions $\{\mathfrak{L}_\kappa^{(v)}\}$, for κ ranging over all partitions, form an orthonormal basis in L^2 , and \mathcal{S}_0 is positive and of trace class.*

The proof of the following theorem is similar to the proof of Theorem 7, and the complete details are provided by Hadjicosta (2019).

Theorem 8 *Let $\mathcal{S} : L^2 \rightarrow L^2$, the covariance operator of the random element \mathcal{Z} , be defined as*

$$\mathcal{S} f(S) = \int_{T>0} K(S, T) f(T) \, dP_0(T)$$

for all $S > 0$ and for all functions f in L^2 , where $K(S, T)$ is the covariance function defined in equation (58). Then, \mathcal{S} is positive and of trace class.

Recall that a non-trivial function $\phi \in L^2$ is an eigenfunction of \mathcal{S} if there exists an eigenvalue $\delta \in \mathbb{C}$ such that $\mathcal{S}\phi = \delta\phi$. As \mathcal{S} is self-adjoint and positive, its eigenvalues are real and nonnegative. In the next result, we find the positive eigenvalues (that are not eigenvalues of \mathcal{S}_0) and corresponding eigenfunctions of \mathcal{S} , and we will show later that 0 is not an eigenvalue of \mathcal{S} .

Theorem 9 *Let $\delta \in \mathbb{R}$ with $\delta \neq \rho_\kappa$ for any partition κ . Also, denote by $\tilde{\rho}_k, k \geq 1$, an enumeration, listed in non-increasing order, of the distinct values of the eigenvalues ρ_κ and define the functions*

$$A(\delta) = 1 - \beta^{m\alpha} m \sum_{k=0}^{\infty} \frac{(m\alpha)_k}{k!(\tilde{\rho}_k - \delta)} \tilde{\rho}_k^2,$$

$$B(\delta) = 1 - \alpha\beta^{m\alpha} m \sum_{k=0}^{\infty} \frac{(m\alpha)_k}{k!(\tilde{\rho}_k - \delta)} \tilde{\rho}_k^2 (b_\alpha^2 - m^{-1}k\beta)^2,$$

and

$$D(\delta) = \alpha^2\beta^{m\alpha} m \sum_{k=0}^{\infty} \frac{(m\alpha)_k}{k!(\tilde{\rho}_k - \delta)} \tilde{\rho}_k^2 (b_\alpha^2 - m^{-1}k\beta).$$

Then, the positive eigenvalues of \mathcal{S} are the positive roots of $G(\delta) = \alpha^3 A(\delta)B(\delta) - D^2(\delta)$. The eigenfunction corresponding to an eigenvalue δ has Fourier–Laguerre expansion

$$\beta^{m\alpha/2} \sum_{k=0}^{\infty} \frac{\tilde{\rho}_k}{\sqrt{k!(\tilde{\rho}_k - \delta)}} (C_1 + C_2\alpha^{-1}(b_\alpha^2 - m^{-1}k\beta)) \sum_{|\kappa|=k} (C_\kappa(I_m) [\alpha]_\kappa)^{1/2} \mathfrak{L}_\kappa^{(v)},$$

where $C_1 C_2 \neq 0$, $\alpha^3 C_1 A(\delta) = C_2 D(\delta)$ and $C_2 B(\delta) = C_1 D(\delta)$.

Remark 3 In an earlier paper (Hadjicosta and Richards 2019), we studied goodness-of-fit testing for the gamma distributions and conjectured that, for all α , the eigenvalues of \mathcal{S} are not eigenvalues of \mathcal{S}_0 . As shown in the next subsection, this is not valid in the case of the Wishart distributions.

A problem with the eigenvalues δ_k is that they have no closed-form expression; hence, there is no simple formula for N , the number of terms in the truncated series $\sum_{k=1}^N \delta_k \chi_{1k}^2$ that should be used in practice to approximate the asymptotic distribution, $\sum_{k=1}^{\infty} \delta_k \chi_{1k}^2$, of the test statistic T_n^2 .

Since \mathcal{S}_0 is of trace class, then by Brislawn (1991, p. 237, Corollary 3.2), $Tr(\mathcal{S}_0)$ can be calculated by integrating the kernel K_0 or by evaluating the sum of all eigenvalues ρ_κ :

$$\int_{S>0} K_0(S, S) dP_0(S) = Tr(\mathcal{S}_0) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} \rho_\kappa = \alpha^{m\alpha} b_\alpha^{2m\alpha} \prod_{k=1}^m (1 - b_\alpha^{4k})^{-1}. \tag{87}$$

Since \mathcal{S} also is of trace class, then

$$\begin{aligned} \sum_{k=1}^{\infty} \delta_k &= Tr(\mathcal{S}) \\ &= \int_{S>0} K(S, S) dP_0(S) \\ &= \int_{S>0} \left[K_0(S, S) - (\alpha^{-3} m^{-1} (\text{tr } S)^2 + 1) \text{etr}(-2\alpha^{-1} S) \right] dP_0(S) \end{aligned}$$

$$\begin{aligned}
 &= \alpha^{m\alpha} b_\alpha^{2m\alpha} \prod_{k=1}^m (1 - b_\alpha^{4k})^{-1} - \alpha^{-3} m^{-1} \sum_{|\kappa|=2} \int_{S>0} \text{etr}(-2\alpha^{-1}S) C_\kappa(S) \, dP_0(S) \\
 &\quad - \int_{S>0} \text{etr}(-2\alpha^{-1}S) \, dP_0(S).
 \end{aligned}$$

All of these integrals can be evaluated using (9) and (10), and the resulting sum can be simplified using Lemma 6. Consequently, we obtain

$$\sum_{k=1}^\infty \delta_k = \alpha^{m\alpha} b_\alpha^{2m\alpha} \prod_{k=1}^m (1 - b_\alpha^{4k})^{-1} - \left(\frac{\alpha}{\alpha + 2}\right)^{m\alpha} \left(1 + \frac{m\alpha + 1}{(\alpha + 2)^2}\right). \tag{88}$$

To determine the number of terms in the truncated series $\sum_{k=1}^N \delta_k \chi_{1k}^2$ that should be used in practice to approximate the asymptotic distribution of T_n^2 , we derive bounds for the eigenvalues δ_k in terms of the ρ_κ and then obtain a general formula for N as a function of α . We refer to the ratio $(\sum_{k=1}^N \delta_k)/Tr(S)$ as the N th *scree ratio* for T_n^2 .

Since S is compact and positive, then its spectrum is countable and contains only nonnegative values (Young 1998, Theorem 8.12, p. 98). The next result shows that the eigenvalues are positive.

Proposition 4 *The operators S and S_0 are injective; that is, $Sf = Sg$ if and only if $f = g$, and the same holds for S_0 . In particular, 0 is not an eigenvalue of S or S_0 .*

Proof By linearity, it suffices to assume that $g = 0$. So, suppose that $Sf = 0$, i.e.,

$$\int_{S>0} K(S, T) f(T) \, dP_0(T) = 0$$

for all $S > 0$. Then for $U > I_m$, by Fubini’s theorem,

$$\begin{aligned}
 0 &= \int_{S>0} \text{etr}(-(U - I_m)S/\alpha) (\det S)^{\alpha - \frac{1}{2}(m+1)} \int_{T>0} K(S, T) f(T) \, dP_0(T) \, dS \\
 &= \int_{T>0} f(T) \left[\int_{S>0} \text{etr}(-(U - I_m)S/\alpha) (\det S)^{\alpha - \frac{1}{2}(m+1)} K(S, T) \, dS \right] \, dP_0(T).
 \end{aligned} \tag{89}$$

By the definition of the covariance function K in (58),

$$\begin{aligned}
 &\int_{S>0} \text{etr}(-(U - I_m)S/\alpha) (\det S)^{\alpha - \frac{1}{2}(m+1)} K(S, T) \, dS \\
 &= \text{etr}(-T/\alpha) \int_{S>0} \text{etr}(-US/\alpha) (\det S)^{\alpha - \frac{1}{2}(m+1)} \\
 &\quad \times \left[\Gamma_m(\alpha) A_\nu(-\alpha^{-2}S, T) - \alpha^{-3} m^{-1} (\text{tr } S)(\text{tr } T) - 1 \right] \, dS.
 \end{aligned}$$

By (25), (14), and Fubini’s theorem, we have

$$\begin{aligned} & \int_{S>0} \text{etr}(-US/\alpha)(\det S)^{\alpha-\frac{1}{2}(m+1)} A_v(-\alpha^{-2}S, T) \, dS \\ &= \int_{O(m)} \int_{S>0} \text{etr}(-US/\alpha)(\det S)^{\alpha-\frac{1}{2}(m+1)} A_v(-\alpha^{-2}HSH'T) \, dS \, dH \\ &= \alpha^{m\alpha}(\det U)^{-\alpha} \int_{O(m)} \text{etr}(\alpha^{-1}H'THU^{-1}) \, dH. \end{aligned}$$

Also, by (4) and (9), we have

$$\int_{S>0} \text{etr}(-US/\alpha)(\det S)^{\alpha-\frac{1}{2}(m+1)} (\text{tr } S) \, dS = \alpha^{m\alpha+2} \Gamma_m(\alpha)(\det U)^{-\alpha} \text{tr}(U^{-1}),$$

and, by (10),

$$\int_{S>0} \text{etr}(-US/\alpha)(\det S)^{\alpha-\frac{1}{2}(m+1)} \, dS = \alpha^{m\alpha} \Gamma_m(\alpha)(\det U)^{-\alpha}.$$

Substituting these results into (89) and discarding extraneous factors, we obtain

$$\begin{aligned} & \int_{T>0} \left[\int_{O(m)} \text{etr}(\alpha^{-1}H'THU^{-1}) \, dH - \alpha^{-1}m^{-1} \text{tr}(U^{-1})(\text{tr } T) - 1 \right] \\ & \qquad \qquad \qquad \times \text{etr}(-T/\alpha) f(T) \, dP_0(T) = 0. \end{aligned} \tag{90}$$

Replacing U by U^{-1} , we find that (90) is equivalent to

$$\begin{aligned} & \int_{T>0} \left[\int_{O(m)} \text{etr}(\alpha^{-1}H'THU) \, dH - 1 \right] \text{etr}(-T/\alpha) f(T) \, dP_0(T) \\ & \qquad \qquad \qquad = \alpha^{-1}m^{-1}(\text{tr } U) \int_{T>0} (\text{tr } T) \text{etr}(-T/\alpha) f(T) \, dP_0(T). \end{aligned} \tag{91}$$

Differentiating both sides of (91) with respect to U , we obtain

$$\begin{aligned} & \int_{O(m)} \int_{T>0} \text{etr}(\alpha^{-1}H'THU)(\alpha^{-1}H'TH) f(T) \, dP_0(T) \, dH \\ & \qquad \qquad \qquad = \alpha^{-1}m^{-1} I_m \int_{T>0} (\text{tr } T) \text{etr}(-T/\alpha) f(T) \, dP_0(T). \end{aligned}$$

Since $T \stackrel{d}{=} HTH'$ for all $H \in O(m)$, and $f(HTH') = f(T)$, then

$$\int_{T>0} \text{etr}(\alpha^{-1}H'THU)(\alpha^{-1}H'TH) f(T) \, dP_0(T)$$

$$= \int_{T>0} \text{etr}(\alpha^{-1}UT)(\alpha^{-1}T)f(T) \, dP_0(T).$$

Therefore,

$$\begin{aligned} &\int_{T>0} \text{etr}(\alpha^{-1}UT)(\alpha^{-1}T)f(T) \, dP_0(T) \\ &= \alpha^{-1}m^{-1}I_m \int_{T>0} (\text{tr } T) \text{etr}(-T/\alpha)f(T) \, dP_0(T). \end{aligned} \tag{92}$$

Differentiating both sides of (92) with respect to U , we find that

$$\alpha^{-2} \int_{T>0} \text{etr}(\alpha^{-1}UT) (T \otimes T) f(T) \, dP_0(T) = 0.$$

As this latter integral is a Laplace transform, we obtain $f = 0$, P_0 -almost everywhere. Also, the same argument may be used in the case of \mathcal{S}_0 . Consequently, 0 is not an eigenvalue of \mathcal{S} . □

We now provide an interlacing property of the eigenvalues δ_k and ρ_k . To state this property, denote by $\xi_k, k = 1, 2, 3 \dots$ the partitions of all nonnegative integers, listed in increasing lexicographic order, e.g., $\xi_1 = (0), \xi_2 = (1), \xi_3 = (2), \xi_4 = (1^2), \xi_5 = (3), \xi_6 = (21), \xi_7 = (1^3), \dots$

Proposition 5 *For all $k \geq 1, \rho_{\xi_k} \geq \delta_k \geq \rho_{\xi_{k+2}}$. Further, for $k \geq 3$, every eigenvalue of \mathcal{S}_0 is an eigenvalue of \mathcal{S} with multiplicity $p_m(k) - 2, p_m(k) - 1$, or $p_m(k)$.*

Proof Define the kernels $k_0(S, T) = -\text{etr}(-(S + T)/\alpha)$ and

$$k_1(S, T) = -\alpha^{-3}m^{-1} \text{etr}(-(S + T)/\alpha)(\text{tr } S)(\text{tr } T),$$

where $S, T > 0$. Also, define on L^2 the corresponding integral operators,

$$\mathcal{U}_j f(S) = \int_{T>0} k_j(S, T) f(T) \, dP_0(T),$$

$j = 0, 1, S > 0$. Then it follows from (58) that $\mathcal{S} = \mathcal{S}_0 + \mathcal{U}_0 + \mathcal{U}_1$.

It is clear that each \mathcal{U}_j is self-adjoint, and also of rank one, i.e., the range of \mathcal{U}_j is a one-dimensional subspace of L^2 . Also, $\mathcal{S}_0 + \mathcal{U}_0$ is self-adjoint, and by following the same steps as in Theorem 8, we see that it is positive and compact.

By the same argument as in the proof of Proposition 4, we find that the operator $\mathcal{S}_0 + \mathcal{U}_0$ is injective; hence, the eigenvalues of $\mathcal{S}_0 + \mathcal{U}_0$ are positive.

Denote by $\omega_k, k \geq 1$, the eigenvalues of $\mathcal{S}_0 + \mathcal{U}_0$, where $\omega_1 \geq \omega_2 \geq \dots$, listed repeatedly according to multiplicity. Since \mathcal{S}_0 is compact, self-adjoint, and injective, and \mathcal{U}_0 is self-adjoint and of rank one, it follows from Hochstadt (1973) or Dancis and Davis (1987) that the eigenvalues of \mathcal{S}_0 interlace the eigenvalues of $\mathcal{S}_0 + \mathcal{U}_0$, i.e., $\rho_{\xi_1} \geq \omega_1 \geq \rho_{\xi_2} \geq \omega_2 \geq \rho_{\xi_3} \geq \omega_3 \geq \rho_{\xi_4} \geq \dots$. Further, by Hochstadt (1973), every eigenvalue of multiplicity $p_m(k), k \geq 2$, of \mathcal{S}_0 , where $p_m(k)$ denotes the number of

Table 1 Values of the lower bounds on r and N for $m = 2$

α	2.5	3	5	10	20	50	100
r	8	7	6	4	3	3	2
N	23	18	14	7	4	4	2

Table 2 Values of the lower bounds on r and N for $m = 3$

α	3	4	5	10	20	50	100
r	8	7	6	4	3	3	2
N	39	29	21	9	5	5	2

partitions of k in at most m parts, is an eigenvalue of $\mathcal{S}_0 + \mathcal{U}_0$ with multiplicity $p_m(k)$ or $p_m(k) - 1$.

Since \mathcal{U}_1 is self-adjoint and of rank one then, repeating the above argument, we find that the eigenvalues of $\mathcal{S}_0 + \mathcal{U}_0$ interlace the eigenvalues of $\mathcal{S}_0 + \mathcal{U}_0 + \mathcal{U}_1 \equiv \mathcal{S}$, i.e., $\omega_k \geq \delta_k \geq \omega_{k+1}, k \geq 1$.

Combining the conclusions of the preceding paragraphs, we deduce that $\rho_{\xi_k} \geq \delta_k \geq \rho_{\xi_{k+2}}, k \geq 1$. Further, by Hochstadt (1973), for $k \geq 3$, every eigenvalue of \mathcal{S}_0 is an eigenvalue of \mathcal{S} with multiplicity $p_m(k) - 2, p_m(k) - 1$, or $p_m(k)$. □

For $\epsilon \in (0, 1)$, we can now determine a value for N such that the N th sree ratio of T_n^2 exceeds $1 - \epsilon$. Applying the interlacing inequalities for δ_k , we obtain $\sum_{k=1}^N \delta_k \geq \sum_{2 \leq |\kappa| \leq r} \rho_\kappa$, where $N = \sum_{k=2}^r p_m(k)$. Since $Tr(\mathcal{S}_0) > Tr(\mathcal{S})$, we advise that N be chosen so that

$$\sum_{0 \leq |\kappa| \leq r} \rho_\kappa \geq (1 - \epsilon)Tr(\mathcal{S}_0).$$

This criterion leads to a value for N that is readily applicable in the analysis of data. Substituting $\rho_\kappa = \alpha^{m\alpha} b_\alpha^{4|\kappa|+2m\alpha}$ and the value of $Tr(\mathcal{S}_0)$ from (87), we obtain

$$\alpha^{m\alpha} b_\alpha^{2m\alpha} \sum_{k=0}^r b_\alpha^{4k} p_m(k) \geq (1 - \epsilon)Tr(\mathcal{S}_0) = (1 - \epsilon)\alpha^{m\alpha} b_\alpha^{2m\alpha} \prod_{k=1}^m (1 - b_\alpha^{4k})^{-1}. \tag{93}$$

For $m = 2, 3$ and $\epsilon = 10^{-10}$, which represents accuracy to ten decimal places, we present in Tables 1 and 2 the values of the lower bounds on r and N for various values of α . As indicated by Tables 1 and 2, fewer eigenvalues appear to be needed to approximate the asymptotic distribution of T_n^2 as α increases. As we show in the following result, which is partly a consequence of the interlacing property of the eigenvalues, all but one of the δ_k and ρ_κ converge to 0 as $\alpha \rightarrow \infty$, a result that is consistent with the decreasing values of r and N in the tables.

Corollary 1 As $\alpha \rightarrow \infty, \rho_\kappa \rightarrow 0$ for all $\kappa \neq (0), \delta_k \rightarrow 0$ for all $k \geq 2$, and $\delta_1 \rightarrow e^{-m}(1 - e^{-m})$.

Proof By (84), $\beta = (1 + 4\alpha^{-1})^{1/2}$. Expanding this expression as a power series in α^{-1} , we obtain $\alpha b_\alpha^2 = \alpha(1 + \frac{1}{2}\alpha(1 - \beta)) = 1 - \alpha^{-1} + O(\alpha^{-2})$. Therefore, $(\alpha b_\alpha^2)^\alpha \rightarrow e^{-1}$ and $b_\alpha \rightarrow 0$ as $\alpha \rightarrow \infty$. By (85), $\rho_\kappa = (\alpha b_\alpha^2)^{m\alpha} b_\alpha^{4|\kappa|}$, so it follows that if $\kappa \neq (0)$ then $\rho_\kappa \rightarrow 0$.

By Proposition 5, $\delta_2 \leq \rho_{(1)}$, so it follows that $\delta_2 \rightarrow 0$ as $\alpha \rightarrow \infty$. Since the δ_k are positive and listed in non-increasing order, then it follows that, as $\alpha \rightarrow \infty$, $\delta_k \rightarrow 0$ for all $k \geq 2$.

Finally, the limiting value of δ_1 is obtained by taking limits in (88). □

5 An application to financial data

In applying our test to financial data, we partially follow Haff et al. (2011, Example 5.3). Denote by $S_{j,k}$, $k = 1, 2, 3$ the daily closing stock prices of Johnson & Johnson (JNJ), Berkshire Hathaway Inc., Class B (BRK-B), and JPMorgan Chase & Co. (JPM), respectively, from November 26, 2017, to November 23, 2018. Were a day a trading holiday, we repeated the observation of the previous day; thus, we had 260 observations in total. Next, we computed the daily logarithmic returns, $\log(S_{j+1,k}/S_{j,k})$, $j = 1, \dots, 260$, $k = 1, 2, 3$; graphs of these logarithmic returns are given in Figure 1. Finally, we partitioned the daily logarithmic returns into biweekly periods and calculated the 3×3 unnormalized covariance matrix for each biweekly period, resulting in matrices X_1, \dots, X_{26} .

A common assumption in research on stochastic volatility models is that the vectors of logarithmic returns, $(\log(S_{j+1,1}/S_{j,1}), \log(S_{j+1,2}/S_{j,2}), \log(S_{j+1,3}/S_{j,3}))$, $j = 1, \dots, 260$, are i.i.d. trivariate normally distributed. If this assumption were valid, then the corresponding biweekly covariance matrices would be i.i.d. Wishart-distributed. We remark that the spikiness of the graphs of the daily logarithmic returns indicates that those logarithmic returns may contain substantial numbers of potential outliers; this leads us to surmise that the data are not normally distributed. Nevertheless, we will test the hypothesis that the biweekly covariance matrices are Wishart-distributed with 9 degrees of freedom, i.e., $\alpha = 4.5$, and unspecified scale matrix Σ . To apply the statistic T_n^2 to test this hypothesis, we use an algorithm of Koev and Edelman (2006) to evaluate the Bessel functions of two matrix arguments, and then, we find that the observed value of the test statistic T_n^2 is 0.127.

We conducted a simulation study to approximate $T_{n;0.05}^2$, the 95th-percentile of the null distribution of T_n^2 . We generated 10,000 random samples of size $n = 26$ from the Wishart distribution with $\alpha = 4.5$ and scale matrix $\Sigma = I_3$, calculated the value of T_n^2 for each sample, and recorded the 95th-percentile of all 10,000 simulated values of T_n^2 . We repeated this process ten times, finally approximating $T_{n;0.05}^2$ as the mean of all 10 simulated 95th-percentiles, viz. $T_{n;0.05}^2 = 0.002$. Since the observed value of T_n^2 exceeds the critical value, then we reject the null hypothesis at the 5% level of significance. We also derived from the simulation study an approximate P value of 0.000 for the test, so we have strong evidence that the three-dimensional vectors of logarithmic returns do not have a trivariate normal distribution or are not mutually independent.

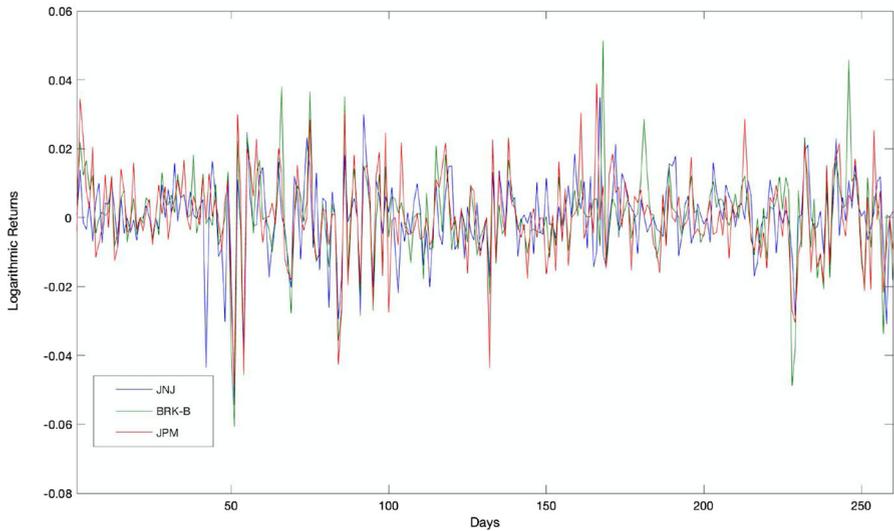


Fig. 1 Graphs of the daily logarithmic returns from the stock prices of JNJ, BRK-B, and JPM over the period November 26, 2017, to November 23, 2018 (color figure online)

For another approach to approximating $T_{n; 0.05}^2$, we can use the limiting distribution of T_n^2 . For $\alpha = 4.5$, from (93), we obtain the approximation $T_n^2 \approx \sum_{k=1}^{21} \delta_k \chi_{1k}^2$. This requires that we first calculate the δ_k (using Theorem 9), and their multiplicities, and then apply the results of Imhof (1961) or Kotz et al. (1967) to derive the distribution of $\sum_{k=1}^{21} \delta_k \chi_{1k}^2$ and carry out the test.

As an alternative to calculating $\delta_1, \dots, \delta_M$, we can apply the interlacing inequalities in Proposition 5 to obtain a stochastic upper bound, $\sum_{k=1}^M \delta_k \chi_{1k}^2 \leq \sum_{0 \leq |\kappa| \leq r} \rho_\kappa \chi_{1\kappa}^2$. Using the upper bound, $\sum_{0 \leq |\kappa| \leq r} \rho_\kappa \chi_{1\kappa}^2$ and applying results of Kotz et al. (1967, loc. cit.) or Imhof (1961) to approximate the critical values, we will obtain a conservative test, i.e., with a level of significance at most 5%.

6 Consistency of the test

Theorem 10 *Let X_1, X_2, \dots be a sequence of $m \times m$ positive definite, i.i.d. random matrices with mean μ . Assume also that the p.d.f. of X_1 is of the form:*

$$f(X_1) = f_0(\mu^{-1/2} X_1 \mu^{-1/2}), \tag{94}$$

where f_0 is orthogonally invariant. Denote by γ the level of significance of the test and by $c_{n,\gamma}$ the $(1 - \gamma)$ -quantile of the statistic T_n^2 under H_0 . If X_1, X_2, \dots are not Wishart-distributed, then $\lim_{n \rightarrow \infty} P(T_n^2 > c_{n,\gamma}) = 1$.

Proof By definition (52) of the test statistic and (57), we have

$$n^{-1}T_n^2 = \int_{T>0} \left[\frac{1}{n} \sum_{j=1}^n \Gamma_m(\alpha) A_v(T, Z_j) - \text{etr}(-\alpha^{-1}T) \right]^2 dP_0(T),$$

where $Z_j = X_j^{1/2} \bar{X}_n^{-1} X_j^{1/2}$. By subtracting and adding the quantity

$$\frac{1}{n} \sum_{j=1}^n \Gamma_m(\alpha) A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2})$$

inside the squared term and then expanding the integrand, we obtain

$$n^{-1}T_n^2 = \int_{T>0} \left[\frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2}) - \text{etr}(-\alpha^{-1}T) \right]^2 dP_0(T) \tag{95}$$

$$+ \int_{T>0} \left[\frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n \left(A_v(T, Z_j) - A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2}) \right) \right]^2 dP_0(T) \tag{96}$$

$$+ 2 \int_{T>0} \left[\frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n \left(A_v(T, Z_j) - A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2}) \right) \right] \times \left[\frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2}) - \text{etr}(-\alpha^{-1}T) \right] dP_0(T). \tag{97}$$

We begin by proving that the integral (96) converges almost surely to 0. By (27), there exists a constant $C > 0$ such that

$$\begin{aligned} & \frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n \left| A_v(T, Z_j) - A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2}) \right| \\ & \leq C \|T\|_F^{1/2} \frac{1}{n} \sum_{j=1}^n \|Z_j - X_j^{1/2} \mu^{-1} X_j^{1/2}\|_F^{1/2} \\ & \leq C \|T\|_F^{1/2} \|\bar{X}_n^{-1} - \mu^{-1}\|_F^{1/2} \frac{1}{n} \sum_{j=1}^n \|X_j\|_F^{1/2}, \end{aligned}$$

since the Frobenius norm is sub-multiplicative. By the triangle inequality, we conclude that the integral (96) is bounded above by

$$C^2 \|\bar{X}_n^{-1} - \mu^{-1}\|_F \left(\frac{1}{n} \sum_{j=1}^n \|X_j\|_F^{1/2} \right)^2 \int_{T>0} \|T\|_F dP_0(T).$$

By the Cauchy–Schwarz inequality, $\left(n^{-1} \sum_{j=1}^n \|X_j\|_F^{1/2}\right)^2 \leq n^{-1} \sum_{j=1}^n \|X_j\|_F$. Since $T > 0$, then $(\text{tr } T^2) \leq (\text{tr } T)^2$, so by (4) and (9), we have

$$\int_{T>0} \|T\|_F \, dP_0(T) = \int_{T>0} (\text{tr } T^2)^{1/2} \, dP_0(T) \leq \int_{T>0} (\text{tr } T) \, dP_0(T) < \infty.$$

By the strong law of large numbers and the continuous mapping theorem, $\|\bar{X}_n^{-1} - \mu^{-1}\|_F \rightarrow 0$, almost surely. Again by the strong law of large numbers, $n^{-1} \sum_{j=1}^n \|X_j\|_F \rightarrow E\|X_1\|_F$, almost surely. It is elementary to verify that $E\|X_1\|_F < \infty$; so (96) converges to 0, almost surely.

Second, we show that (97) tends to 0, almost surely. By (26), the fact that $\text{etr}(-\alpha^{-1}T) \leq 1$ for $T > 0$, and the triangle inequality, we have

$$\left| \frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2}) - \text{etr}(-\alpha^{-1}T) \right| \leq 2.$$

Further, by the triangle inequality, the absolute value of (97) is less than or equal to

$$2 \int_{T>0} \left| \frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n \left(A_v(T, Z_j) - A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2}) \right) \right| \, dP_0(T). \tag{98}$$

By the Cauchy–Schwarz inequality and the fact that $\int_{T>0} dP_0(T) = 1$, (98) is no larger than

$$2 \left(\int_{T>0} \left[\frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n \left(A_v(T, Z_j) - A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2}) \right) \right]^2 \, dP_0(T) \right)^{1/2}.$$

Following the same argument as for (96), we conclude that (97) converges to 0, almost surely.

Since $A_v(T, X_j^{1/2} \mu^{-1} X_j^{1/2}) = A_v(T, \mu^{-1/2} X_j \mu^{-1/2})$, we see that the integral (95) equals

$$\int_{T>0} \left[\frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n A_v(T, \mu^{-1/2} X_j \mu^{-1/2}) - \text{etr}(-\alpha^{-1}T) \right]^2 \, dP_0(T).$$

We subtract and add inside the squared term the quantity $E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2} X_1 \mu^{-1/2})]$ and expand the integrand. Then we find that (95) equals

$$\int_{T>0} \left[\frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n A_v(T, \mu^{-1/2} X_j \mu^{-1/2}) - E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2} X_1 \mu^{-1/2})] \right]^2 \, dP_0(T) \tag{99}$$

$$\begin{aligned}
 &+ \int_{T>0} \left[E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2}X_1\mu^{-1/2})] - \text{etr}(-\alpha^{-1}T) \right]^2 dP_0(T) \\
 &+ 2 \int_{T>0} \left[\frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n A_v(T, \mu^{-1/2}X_j\mu^{-1/2}) - E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2}X_1\mu^{-1/2})] \right] \\
 &\quad \times \left[E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2}X_1\mu^{-1/2})] - \text{etr}(-\alpha^{-1}T) \right] dP_0(T). \tag{100}
 \end{aligned}$$

By the strong law of large numbers in L^2 (Ledoux and Talagrand 1991, p. 189), we conclude that the term (99) converges to 0, almost surely.

Next, we show that (100) converges to 0, almost surely. By (26) and the bound, $\text{etr}(-\alpha^{-1}T) \leq 1$ for $T > 0$, we have

$$\left| E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2}X_1\mu^{-1/2})] - \text{etr}(-\alpha^{-1}T) \right| \leq 2.$$

Therefore, the absolute value of the integral (100) is less than or equal to

$$\begin{aligned}
 &2 \int_{T>0} \left| \frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n A_v(T, \mu^{-1/2}X_j\mu^{-1/2}) - E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2}X_1\mu^{-1/2})] \right| dP_0(T) \\
 &\leq 2 \left(\int_{T>0} \left[\frac{\Gamma_m(\alpha)}{n} \sum_{j=1}^n A_v(T, \mu^{-1/2}X_j\mu^{-1/2}) - E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2}X_1\mu^{-1/2})] \right]^2 dP_0(T) \right)^{1/2},
 \end{aligned}$$

where the latter bound follows from the Cauchy–Schwarz inequality. Again, by the strong law of large numbers in L^2 , we conclude that the integral (100) converges to 0, almost surely.

We have now shown that

$$\frac{1}{n} T_n^2 \xrightarrow{a.s.} \int_{T>0} \left[E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2}X_1\mu^{-1/2})] - \text{etr}(-\alpha^{-1}T) \right]^2 dP_0(T). \tag{101}$$

Denote by Δ the right-hand side of (101); then, $\Delta \geq 0$. Suppose that $\Delta = 0$, then

$$E[\Gamma_m(\alpha)A_v(T, \mu^{-1/2}X_1\mu^{-1/2})] - \text{etr}(-\alpha^{-1}T) = 0,$$

equivalently, $\tilde{\mathcal{H}}_{\mu^{-1/2}X_1\mu^{-1/2}}(T) - \text{etr}(-\alpha^{-1}T) = 0$, P_0 -almost everywhere. By continuity, we obtain $\tilde{\mathcal{H}}_{\mu^{-1/2}X_1\mu^{-1/2}}(T) - \text{etr}(-\alpha^{-1}T) = 0$ for all $T > 0$. By the uniqueness theorem for orthogonally invariant Hankel transforms, it follows that $\mu^{-1/2}X_1\mu^{-1/2}$ has a Wishart distribution. By Muirhead (1982, p. 92), X_1 has also a Wishart distribution, which contradicts the assumption that X_1 does not have a Wishart distribution. Therefore, $\Delta > 0$.

Under H_0 , $n^{-1} T_n^2 \xrightarrow{a.s.} 0$, and therefore $n^{-1} T_n^2 \xrightarrow{P} 0$, i.e., for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} P_{H_0}(n^{-1} T_n^2 \geq \epsilon) = 0.$$

Thus, for any $\epsilon > 0$ and $\gamma > 0$, there exists $n_0(\epsilon, \gamma) \in \mathbb{N}$ such that $P_{H_0}(n^{-1}T_n^2 \geq \epsilon) \leq \gamma$ for all $n \geq n_0(\epsilon, \gamma)$. Let $c_{n,\gamma}$ be the $(1 - \gamma)$ -quantile of T_n^2 under H_0 . Then $0 \leq c_{n,\gamma} \leq n\epsilon$ for all $n \geq n_0(\epsilon)$ since, by definition, $c_{n,\gamma} := \inf\{x \geq 0 : P_{H_0}(T_n^2 > x) \leq \gamma\}$. Therefore, $0 \leq n^{-1}c_{n,\gamma} \leq \epsilon$ for all $n \geq n_0(\epsilon)$. In summary, for any $\epsilon > 0$, there exists $n_0(\epsilon) \in \mathbb{N}$ such that $n^{-1}c_{n,\gamma} \leq \epsilon$ for all $n \geq n_0(\epsilon)$, i.e.,

$$\lim_{n \rightarrow \infty} n^{-1}c_{n,\gamma} = 0. \tag{102}$$

By (101) and (102), we have $n^{-1}T_n^2 - n^{-1}c_{n,\gamma} \xrightarrow{a.s.} \Delta$, and therefore $n^{-1}T_n^2 - n^{-1}c_{n,\gamma} \xrightarrow{P} \Delta$. Thus, by Severini (2005, p. 340, Corollary 11.3 (i)), we have that $n^{-1}T_n^2 - n^{-1}c_{n,\gamma} \xrightarrow{d} \Delta$. Further,

$$\begin{aligned} \lim_{n \rightarrow \infty} P(T_n^2 > c_{n,\gamma}) &= \lim_{n \rightarrow \infty} P(n^{-1}T_n^2 - n^{-1}c_{n,\gamma} > 0) \\ &= 1 - \lim_{n \rightarrow \infty} P(n^{-1}T_n^2 - n^{-1}c_{n,\gamma} \leq 0) = 1 \end{aligned}$$

since $\Delta > 0$. Therefore, $\lim_{n \rightarrow \infty} P(T_n^2 > c_{n,\gamma}) = 1$. □

Remark 4 By applying Theorem 1 of Baringhaus et al. (2017), we also find that under fixed alternatives satisfying (94), $n^{1/2}(n^{-1}T_n^2 - \Delta) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$ as $n \rightarrow \infty$, where σ^2 is a constant that is determined from the alternative distribution.

Remark 5 We show that assumption (94) holds for two alternative distributions. The matrix *F-distribution* (cf. Khatri 1966, Section 4, part (c) or James 1964, Eqs. (65), (72)): Let X be a random matrix with p.d.f.

$$f(X) = \frac{\Gamma_m(a + b)}{\Gamma_m(a)\Gamma_m(b)} (\det X)^{a-(m+1)/2} (\det(I_m + X))^{-(a+b)},$$

$X > 0$, where $a > \frac{1}{2}(m - 1)$ and $b > \frac{1}{2}(m + 1)$. Since $f(X)$ is orthogonally invariant then, by Schur’s Lemma, there exists a constant c such that $\mu = E(X) = cI_m$.

A linear combination of Wishart matrices: Let X be a random matrix with p.d.f.

$$f(X) = \frac{\delta^{mb}(\delta - 1)^{ma}}{\Gamma_m(a + b)} (\det X)^{a+b-(m+1)/2} \text{etr}(-\delta X) {}_1F_1(a; a + b; X),$$

$X > 0$, where $a > \frac{1}{2}(m - 1)$, $b > \frac{1}{2}(m - 1)$, and $\delta > 1$. By Gupta and Richards (1995, Section 4.4), we have $X \stackrel{d}{=} X_1 + \delta^{-1}X_2$, where X_1 and X_2 are independent, $X_1 \sim W_m(a, I_m)$ and $X_2 \sim W_m(b, I_m)$. Again, the distribution of X is orthogonally invariant; thus, it satisfies (94).

7 Contiguous alternatives to the null hypothesis

For $n \in \mathbb{N}$, let X_{n1}, \dots, X_{nn} be a triangular array of row-wise independent $m \times m$ positive definite random matrices. As usual, let $P_0 = W_m(\alpha, I_m)$, $\alpha > \max\{\frac{1}{2}(2m - 1), \frac{1}{2}(m + 3)\}$, and let Q_{n1} be a probability measure dominated by P_0 . We wish to test the null hypothesis

H_0 : The marginal distribution of each X_{ni} , $i = 1, \dots, n$, is P_0 against the alternative

H_1 : The marginal distribution of each X_{ni} , $i = 1, \dots, n$, is Q_{n1} .

We write the Radon–Nikodym derivative of Q_{n1} with respect to P_0 in the form $dQ_{n1}/dP_0 = 1 + n^{-1/2}h_n$. Then we will need two assumptions on the sequence h_n :

Assumptions 1 We assume that:

- (A1) The functions $\{h_n : n \in \mathbb{N}\}$ form a sequence of P_0 -integrable functions converging pointwise, P_0 -almost everywhere, to a function h , and
- (A2) $\sup_{n \in \mathbb{N}} E_{P_0} |h_n|^4 < \infty$.

Since $\int (dQ_{n1}/dP_0) dP_0 = 1$, then we also have $\int h_n dP_0 = 0$, for all $n \in \mathbb{N}$. Denote the indicator function of an event A by $I(A)$; applying (A2), we deduce the uniform integrability of $|h_n|^2$:

$$\begin{aligned} \lim_{k \rightarrow \infty} \sup_{n \in \mathbb{N}} E_{P_0} (|h_n|^2 I(|h_n|^2 > k)) &= \lim_{k \rightarrow \infty} \sup_{n \in \mathbb{N}} \int |h_n|^2 I(|h_n|^2 > k) dP_0 \\ &\leq \lim_{k \rightarrow \infty} k^{-1} \sup_{n \in \mathbb{N}} E_{P_0} |h_n|^4 = 0. \end{aligned}$$

By Bauer (1981, p. 95), the P_0 -almost everywhere convergence of h_n to h implies the P_0 -stochastic convergence of h_n to h . Again by Bauer (1981, p. 104), the uniform integrability of $|h_n|^2$ along with the P_0 -stochastic convergence of h_n to h implies the convergence of h_n in mean square, i.e.,

$$\lim_{n \rightarrow \infty} \int |h_n - h|^2 dP_0 = 0,$$

and therefore

$$\lim_{n \rightarrow \infty} \int |h_n|^2 dP_0 = \int |h|^2 dP_0.$$

By the triangle and the Cauchy–Schwarz inequalities,

$$\begin{aligned} 0 &\leq \lim_{n \rightarrow \infty} \left| \int (h_n - h) dP_0 \right| \\ &\leq \lim_{n \rightarrow \infty} \int |h_n - h| dP_0 \leq \lim_{n \rightarrow \infty} \left(\int |h_n - h|^2 dP_0 \right)^{1/2} = 0, \end{aligned}$$

therefore

$$\lim_{n \rightarrow \infty} \int h_n \, dP_0 = \int h \, dP_0 = 0.$$

We now verify that Assumptions 1 are valid for numerous sequences of contiguous alternatives.

7.1 Wishart alternatives with contiguous scale matrices

Let $Q_{n1} := W_m(\alpha, \Sigma_n)$ with $\alpha > \max\{\frac{1}{2}(2m - 1), \frac{1}{2}(m + 3)\}$ and $\Sigma_n = (1 + \frac{1}{\sqrt{n}})I_m$. Then,

$$\frac{dQ_{n1}}{dP_0} = (1 + n^{-1/2})^{m\alpha} \text{etr}(-n^{-1/2}X) \equiv 1 + n^{-1/2}h_n(X),$$

where

$$h_n(X) = n^{1/2}[(1 + n^{-1/2})^{m\alpha} \text{etr}(-n^{-1/2}X) - 1],$$

for $X > 0$. By applying L'Hôpital's rule, we obtain

$$h(X) := \lim_{n \rightarrow \infty} h_n(X) = m\alpha - \text{tr } X,$$

for $X > 0$. Next, we find $E_{P_0}|h_n^4|$. Define

$$\begin{aligned} R_n(X) &= \text{etr}(-n^{-1/2}X) - (1 - n^{-1/2}(\text{tr } X)) & (103) \\ &= \sum_{k=2}^{\infty} \frac{1}{k!} (-n^{-1/2}(\text{tr } X))^k, \end{aligned}$$

the remainder term of the Taylor series expansion of $\text{etr}(-n^{-1/2}X)$, $X > 0$. Then, by elementary algebraic manipulations, we obtain

$$\begin{aligned} h_n(X) &= n^{1/2}(1 + n^{-1/2})^{m\alpha} \text{etr}(-n^{-1/2}X) - n^{1/2} \\ &= (1 + n^{-1/2})^{m\alpha-1} [1 + (1 + n^{-1/2})R_n(X) - (1 + n^{-1/2})(\text{tr } X)] \\ &\quad + n^{1/2}[(1 + n^{-1/2})^{m\alpha-1} - 1]. \end{aligned}$$

By (103), the triangle inequality, and the Lipschitz continuity of the exponential function, we have

$$\begin{aligned} |R_n(X)| &\leq n^{1/2}|R_n(X)| \leq n^{1/2} [|\text{etr}(-n^{-1/2}X) - 1| + n^{-1/2}(\text{tr } X)] \\ &\leq n^{1/2} [n^{-1/2}(\text{tr } X) + n^{-1/2}(\text{tr } X)] = 2 \text{tr } X, \end{aligned}$$

$X > 0$. Therefore,

$$\begin{aligned} |h_n(X)| &\leq (1 + n^{-1/2})^{m\alpha-1} \left[1 + (1 + n^{1/2})|R_n(X)| + (1 + n^{-1/2})(\text{tr } X) \right] \\ &\quad + n^{1/2} |(1 + n^{-1/2})^{m\alpha-1} - 1| \\ &\leq (1 + n^{-1/2})^{m\alpha-1} (1 + 4 \text{tr } X + 2 \text{tr } X) + n^{1/2} |(1 + n^{-1/2})^{m\alpha-1} - 1| \\ &= (1 + n^{-1/2})^{m\alpha-1} (1 + 6 \text{tr } X) + n^{1/2} |(1 + n^{-1/2})^{m\alpha-1} - 1|. \end{aligned}$$

There exists a constant $C > 0$ such that $(1 + n^{-1/2})^{m\alpha-1} \leq C$ and $|n^{1/2}((1 + n^{-1/2})^{m\alpha-1} - 1)| \leq C$ for all n . Therefore, $|h_n(X)| \leq C(1 + 6 \text{tr } X) + C = C(2 + 6 \text{tr } X)$, $X > 0$, so we obtain

$$E_{P_0} |h_n|^4 \leq C^4 \int_{X>0} (2 + 6 \text{tr } X)^4 dP_0(X),$$

a bound independent of n . By (4) and (9), the latter integral is finite; thus, $\sup_{n \in \mathbb{N}} E_{P_0} |h_n|^4 < \infty$.

7.2 Wishart alternatives with contiguous shape parameters

Let $Q_{n1} := W_m(\alpha_n, I_m)$ with $\alpha_n = \alpha + \frac{1}{\sqrt{n}}$, $\alpha > \max\{\frac{1}{2}(2m - 1), \frac{1}{2}(m + 3)\}$. Then,

$$\frac{dQ_{n1}}{dP_0} = \frac{\Gamma_m(\alpha)}{\Gamma_m(\alpha_n)} (\det X)^{1/\sqrt{n}} \equiv 1 + n^{-1/2} h_n(X),$$

where

$$h_n(X) = n^{1/2} \left(\frac{\Gamma_m(\alpha)}{\Gamma_m(\alpha_n)} (\det X)^{1/\sqrt{n}} - 1 \right),$$

$X > 0$. Recall the *multivariate digamma function*

$$\psi_m(z) := \frac{d}{dz} \log \Gamma_m(z) = \frac{\Gamma'_m(z)}{\Gamma_m(z)},$$

$z > 0$. Applying L'Hôpital's rule, we obtain

$$\begin{aligned} h(X) &:= \lim_{n \rightarrow \infty} h_n(X) \\ &= \lim_{n \rightarrow \infty} n^{1/2} \left(\frac{\Gamma_m(\alpha)}{\Gamma_m(\alpha + n^{-1/2})} (\det X)^{1/\sqrt{n}} - 1 \right) \\ &= \log(\det X) - \psi_m(\alpha), \end{aligned}$$

$X > 0$. To calculate $E_{P_0}|h_n|^4$, we apply the binomial expansion, obtaining

$$\begin{aligned} & \left| n^{1/2} \left(\frac{\Gamma_m(\alpha)}{\Gamma_m(\alpha + n^{-1/2})} (\det X)^{1/\sqrt{n}} - 1 \right) \right|^4 \\ &= n^2 \sum_{j=0}^4 (-1)^j \binom{4}{j} \left(\frac{\Gamma_m(\alpha)}{\Gamma_m(\alpha + n^{-1/2})} \right)^j (\det X)^{j/\sqrt{n}}, \end{aligned}$$

thus,

$$E_{P_0}|h_n|^4 = n^2 \sum_{j=0}^4 (-1)^j \binom{4}{j} \left(\frac{\Gamma_m(\alpha)}{\Gamma_m(\alpha + n^{-1/2})} \right)^j \frac{\Gamma_m(\alpha + jn^{-1/2})}{\Gamma_m(\alpha)}. \tag{104}$$

Next, the Taylor expansion of $\Gamma_m(\alpha) / \Gamma_m(\alpha + n^{-1/2})$ for sufficiently large values of n is

$$\frac{\Gamma_m(\alpha)}{\Gamma_m(\alpha + n^{-1/2})} = \sum_{j=0}^4 a_j n^{-j/2} + o(n^{-2}). \tag{105}$$

After lengthy but straightforward calculations, we obtain

$$\begin{aligned} a_0 &= 1, & a_1 &= -\psi_m(\alpha), & a_2 &= \frac{1}{2} \psi_m^2(\alpha) - \frac{1}{2} \psi_m'(\alpha), \\ a_3 &= -\frac{1}{6} \psi_m^3(\alpha) - \frac{1}{6} \psi_m''(\alpha) + \frac{1}{2} \psi_m(\alpha) \psi_m'(\alpha), \\ a_4 &= -\frac{\psi_m'''(\alpha)}{24} + \frac{1}{8} (\psi_m'(\alpha))^2 + \frac{1}{6} \psi_m(\alpha) \psi_m''(\alpha) - \frac{1}{4} \psi_m^2(\alpha) \psi_m'(\alpha) + \frac{1}{24} \psi_m^4(\alpha). \end{aligned}$$

Next, we substitute the Taylor expansion (105) in (104) and then let $n \rightarrow \infty$. Applying L'Hôpital's rule four times then, after some lengthy but straightforward calculations, we obtain

$$\lim_{n \rightarrow \infty} E_{P_0}|h_n|^4 = 9a_1^4 + 24a_2^2 + 24a_1a_3 - 36a_1^2a_2 - 24a_4.$$

Thus, $E_{P_0}|h_n^4|$ is a bounded sequence, and therefore, $\sup_{n \in \mathbb{N}} E_{P_0}|h_n|^4 < \infty$.

7.3 Contaminated Wishart models

Consider the contamination model, $Q_{n1} := (1 - n^{-1/2})P_0 + n^{-1/2}W_m(2\alpha, I_m)$, where, as usual, $\alpha > \max\{\frac{1}{2}(2m - 1), \frac{1}{2}(m + 3)\}$. These contaminated Wishart models appear in the analysis of diffusion tensor images (Jian and Vemuri 2007). We have

$$\frac{dQ_{n1}}{dP_0} = n^{-1/2} \left(\frac{\Gamma_m(\alpha)}{\Gamma_m(2\alpha)} (\det X)^\alpha - 1 \right) + 1 \equiv 1 + n^{-1/2}h_n(X),$$

where

$$h_n(X) = \frac{\Gamma_m(\alpha)}{\Gamma_m(2\alpha)} (\det X)^\alpha - 1,$$

for $X > 0$. As h_n does not depend on n , then $h(X) \equiv h_n(X)$. Since

$$E_{P_0}|h_n^4| = \int_{X>0} \left(\frac{\Gamma_m(\alpha)}{\Gamma_m(2\alpha)} (\det X)^\alpha - 1 \right)^4 dP_0(X)$$

clearly is finite and does not depend on n then $\sup_{n \in \mathbb{N}} E_{P_0}|h_n|^4 < \infty$.

Note that Q_{n1} is a special case of a contamination model $Q_{n2} = (1 - n^{-1/2})P_0 + n^{-1/2}P_1$, with $P_1 \ll P_0$ and $\int (dP_1/dP_0)^4 dP_0 < \infty$. The earlier calculations can be done for many such P_1 . For example, let P_1 be the probability measure for the matrix generalized inverse Gaussian distribution (Butler 1998) with density function $f_1(X) = c_1 (\det X)^{b-\frac{1}{2}(m+1)} \text{etr}(-\Phi X^{-1} - \Psi X)$, $X > 0$, where c_1 is the normalizing constant, Φ and Ψ are positive definite matrices, and $b \in \mathbb{R}$. Then

$$\begin{aligned} & \int (dP_1/dP_0)^4 dP_0 \\ &= c \int_{X>0} (\det X)^{4b-3\alpha-\frac{1}{2}(m+1)} \text{etr}(-4\Phi X^{-1} - (4\Psi - 3I_m)X) dX, \end{aligned}$$

where $c_0 = 1/\Gamma_m(\alpha)$ is the normalizing constant of $W_m(\alpha, I_m)$ and $c = c_1^4/c_0^3$. By Herz (1955, p. 506) and Butler (1998, Eq. (2)), we deduce that $\int (dP_1/dP_0)^4 dP_0 < \infty$ in the following cases: (i) $\Phi \geq 0$, $\Psi - \frac{3}{4}I_m > 0$, $b \geq \frac{1}{4}(3\alpha + \frac{1}{2}m)$; (ii) $\Phi > 0$, $\Psi - \frac{3}{4}I_m > 0$, $b \in \mathbb{R}$; and (iii) $\Phi > 0$, $\Psi - \frac{3}{4}I_m \geq 0$, $b < \frac{1}{4}(3\alpha - \frac{1}{2}(m - 1))$. Therefore, Assumptions 1 also hold for broad classes of the model Q_{n2} .

7.4 The distribution of the test statistic under contiguous alternatives

Let $P_0 = W_m(\alpha, I_m)$, $\alpha > \max\{\frac{1}{2}(2m - 1), \frac{1}{2}(m + 3)\}$. Also, denote by $P_n = P_0 \otimes \dots \otimes P_0$ and $Q_n = Q_{n1} \otimes \dots \otimes Q_{n1}$ the n -fold product probability measures of P_0 and Q_{n1} , respectively.

Theorem 11 *Let $m \geq 2$ and X_{n1}, \dots, X_{nn} , $n \in \mathbb{N}$, be a triangular array of $m \times m$ positive definite row-wise i.i.d. random matrices, where $X_{nj} = X_j$, $j = 1, \dots, n$. We assume that the distribution of X_{nj} is Q_{n1} , for every $j = 1, \dots, n$. Further, let $Z_n = (Z_n(T), T > 0)$ be a random field with*

$$Z_n(T) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \left[\Gamma_m(\alpha) A_v(T, X_{nj}^{1/2} \bar{X}_n^{-1} X_{nj}^{1/2}) - \text{etr}(-T/\alpha) \right],$$

$T > 0$. Under Assumptions 1, there exists a centered Gaussian field $Z := (Z(T), T > 0)$ with sample paths in L^2 and the covariance function $K(S, T)$ in (58), and a function

$$c(T) = \int_{X>0} \left[\Gamma_m(\alpha) A_v(T, \alpha^{-1}X) + \alpha^{-2}(\text{tr } T) \text{etr}(-\alpha^{-1}T)(\text{tr } X) \right] h(X) dP_0(X),$$

$T > 0$, such that $Z_n \xrightarrow{d} Z + c$ in L^2 . Moreover, as $n \rightarrow \infty$,

$$T_n^2 \xrightarrow{d} \int_{T>0} (Z(T) + c(T))^2 dP_0(T).$$

The proof of this theorem can be obtained by following the approach of Taherizadeh (2009, pp. 79–91) and Hadjicosta and Richards (2019, Theorem 4.2). We omit the details, which can also be found in Hadjicosta (2019).

8 The efficiency of the test

We now investigate the approximate Bahadur slope of the statistic T_n^2 under local alternatives. We will prove the validity of a modified version of Wieand's condition. The proof of Wieand's condition, under which the Bahadur and Pitman efficiencies agree, remains an open problem. By applying these results, we will be able to calculate the approximate asymptotic relative efficiency (ARE) of the proposed test relative to potential alternative tests.

For $m \geq 2$, let X_1, X_2, \dots be i.i.d., $m \times m$ positive definite random matrices with unknown distribution P . We assume that P is indexed by a parameter $\theta \in \Theta := (-\eta, \eta)$ or $\Theta := [0, \eta)$, for some $\eta > 0$. We let $\theta \in \Theta_0 = \{\theta_0\} = \{0\}$ to represent the null hypothesis and $\theta \in \Theta_1 = \Theta \setminus \{0\}$ to represent the alternative hypothesis. In Sect. 3, we showed that T_n^2 is scale invariant, i.e., it does not depend on the unknown scale matrix Σ . Thus, under H_0 , we assume that X_1, X_2, \dots are i.i.d., $m \times m$ positive definite P_0 -distributed random matrices and under the local alternatives, represented by $\theta \in \Theta_1$, X_1, X_2, \dots are i.i.d., $m \times m$ positive definite P_θ -distributed random matrices.

The Radon–Nikodym derivative of P_θ with respect to P_0 is $dP_\theta/dP_0 = 1 + \theta h_\theta$. We assume that as $\theta \rightarrow 0$, the function h_θ converges to some function h in mean square, i.e.,

$$\lim_{\theta \rightarrow 0} \int_{X>0} |h_\theta(X) - h(X)|^2 dP_0(X) = 0. \quad (106)$$

Since $\int (dP_\theta/dP_0) dP_0 = 1$ then $\int_{X>0} h_\theta(X) dP_0(X) = 0$, $\theta \in \Theta_1$. We also assume that for $\theta \in \Theta_1$,

$$\int_{X>0} X h_\theta(X) dP_0(X) = 0. \quad (107)$$

Denote by Θ_0 and Θ_1 the null and alternative parameter spaces, respectively. For $\theta \in \Theta_0$, let $F_n(t) = P_\theta(T_n < t)$, $t \in \mathbb{R}$, be the null distribution of T_n ; then, the level attained by T_n is $L_n := 1 - F_n(T_n)$. For $\theta \in \Theta_1$, the exact Bahadur slope of $\{T_n : n \in \mathbb{N}\}$ is $c(\theta) = -2 \lim_{n \rightarrow \infty} n^{-1} \log L_n$, whenever the limit exists (almost surely); for $\theta \in \Theta_0$, this limit exists with $c(\theta) = 0$.

For a sequence $\{U_{j,n} : n \in \mathbb{N}\}$ of test statistics with exact Bahadur slope $c_j(\theta)$, $j = 1, 2$, the exact Bahadur asymptotic relative efficiency of $\{U_{1,n} : n \in \mathbb{N}\}$ with respect to $\{U_{2,n} : n \in \mathbb{N}\}$ is $c_1(\theta)/c_2(\theta)$, $\theta \in \Theta_1$. If $c_1(\theta)/c_2(\theta) > 1$, then the sequence $\{U_{1,n} :$

$n \in \mathbb{N}$ is preferred for the test of hypothesis. We study the approximate (Bahadur) slope of T_n as it is difficult to calculate exact slopes (Bahadur 1971, Theorem 7.2) and since, for $\Theta_0 = \{\theta_0\}$, the approximate slope is close to the exact slope for θ in a neighborhood of θ_0 , i.e., under local alternatives (Bahadur 1960, 1967).

Theorem 12 *The sequence of test statistics $\{T_n : n \in \mathbb{N}\}$ is a standard sequence. The approximate Bahadur slope of the test is $\tilde{\delta}_1^{-1} b^2(\theta)$, where $\tilde{\delta}_1$ is the largest eigenvalue of the operator \mathcal{S} , and*

$$b^2(\theta) = \theta^2 \int_{T>0} \left[\int_{X>0} \Gamma_m(\alpha) A_v(T, \alpha^{-1} X) h_\theta(X) dP_0(X) \right]^2 dP_0(T).$$

Moreover,

$$\lim_{\theta \rightarrow 0} \frac{\tilde{\delta}_1^{-1} b^2(\theta)}{\theta^2} = \tilde{\delta}_1^{-1} \int_{T>0} \left[\int_{X>0} \Gamma_m(\alpha) A_v(T, \alpha^{-1} X) h(X) dP_0(X) \right]^2 dP_0(T).$$

The proof of this theorem is similar to the proof of Hadjicosta and Richards (2019, Theorem 5.1). We will omit the details as they can be found in Hadjicosta (2019).

Wieand (1976) showed that if two standard sequences of test statistics satisfy an additional condition, now called the *Wieand's condition*, then the limiting approximate Bahadur efficiency is in accord with the limiting Pitman efficiency, as the level of significance decreases to 0. For a description of Pitman's asymptotic relative efficiency, see Taherizadeh (2009, Chapter 5) or Hadjicosta and Richards (2019, Section 5). Although a proof of Wieand's condition for the statistics $\{T_n : n \in \mathbb{N}\}$ remains an open problem, we will show that a modified form of the condition holds.

Theorem 13 *There exists a constant $\theta^* > 0$ such that for any $\epsilon > 0$ and $\gamma \in (0, 1)$, there exists a constant $C > 0$ such that*

$$P(|n^{-1/2} T_n - b(\theta)| \leq \epsilon b(\theta)) > 1 - \gamma$$

for any $\theta \in \Theta_1 \cap (-\theta^*, \theta^*)$ and $n^{1/2} > C/b^2(\theta)$.

9 Concluding remarks

The results in this paper give rise to a plethora of open problems. The statistic T_n^2 is the only statistic available for carrying out goodness-of-fit tests for the Wishart distributions, so it will be useful to develop alternative tests so that relative efficiency calculations can be done. Henze et al. (2012) have provided a goodness-of-fit test, for the gamma distributions, based on the Laplace transform; we believe that their results are worthy of study for extension to the Wishart distributions.

The results of Henze et al. (2012) also pertain to the case in which α is unknown, which raises the problem of extending our results to Wishart distributions with unknown shape parameter. One approach to this problem is to insert for each α in

(3) an estimator $\hat{\alpha}$; this is related to the method of inserting a parameter estimator into a V -statistic; cf. Matsui and Takemura (2008).

In constructing the statistic T_n^2 , we fixed $\nu = \alpha - \frac{1}{2}(m + 1)$; see (51). For general ν , the corresponding V -statistic representation will be more complex, involving new integrals for the generalized hypergeometric functions of matrix argument. Also, the limiting distribution of the statistic will be of the usual form, an infinite linear combination of i.i.d. χ_1^2 random variables, but it will be a challenging problem to develop the spectral analysis of the corresponding covariance operator.

As for choosing a weight measure P_0 , it will be useful to study weight measures w of the form $dw(T) = \text{etr}(-BT) dP_0(T)$, where the “tuning parameter” B is a positive definite matrix that enables detection of departures from H_0 due to specific entries in the random data.

Acknowledgements We thank the reviewers and the editors for their constructive comments.

References

- Anfinsen, S. N., Doulgeris, A. P., Eltoft, T. (2011). Goodness-of-fit tests for multilook polarimetric radar data based on the Mellin transform. *IEEE Transactions on Geoscience and Remote Sensing*, *49*, 2764–2781.
- Anfinsen, S. N., Eltoft, T. (2011). Application of the matrix-variate Mellin transform to analysis of polarimetric radar images. *IEEE Transactions on Geoscience and Remote Sensing*, *49*, 2281–2295.
- Asai, M., McAleer, M., Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Reviews*, *25*, 145–175.
- Bahadur, R. R. (1960). Stochastic comparison of tests. *Annals of Mathematical Statistics*, *31*, 276–295.
- Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics. *Annals of Mathematical Statistics*, *38*, 303–324.
- Bahadur, R. R. (1971). *Some limit theorems in statistics*. Philadelphia, PA: SIAM.
- Baringhaus, L., Ebner, B., Henze, N. (2017). The limit distribution of weighted L^2 -goodness-of-fit statistics under fixed alternatives, with applications. *Annals of the Institute of Statistical Mathematics*, *69*, 969–995.
- Baringhaus, L., Taherizadeh, F. (2010). Empirical Hankel transforms and their applications to goodness-of-fit tests. *Journal of Multivariate Analysis*, *101*, 1445–1467.
- Baringhaus, L., Taherizadeh, F. (2013). A K-S type test for exponentiality based on empirical Hankel transforms. *Communications in Statistics – Theory and Methods*, *42*, 3781–3792.
- Bauer, H. (1981). *Probability theory and elements of measure theory, second English edition*. New York, NY: Academic Press.
- Billingsley, P. (1968). *Convergence of probability measures*. New York: Wiley.
- Billingsley, P. (1979). *Probability and measure*. New York: Wiley.
- Bishop, A. N., Del Moral, P., Niclas, A. (2017). An introduction to Wishart matrix moments. *Foundations and Trends in Machine Learning*, *11*, 97–218.
- Brislawn, C. (1991). Traceable integral kernels on countably generated measure spaces. *Pacific Journal of Mathematics*, *150*, 229–240.
- Browne, M. W. (1968). A comparison of factor analytic techniques. *Psychometrika*, *33*, 267–334.
- Burkill, J. C., Burkill, H. (2002). *A second course in mathematical analysis*. New York: Cambridge University Press.
- Butler, R. W. (1998). Generalized inverse Gaussian distributions and their Wishart connections. *Scandinavian Journal of Statistics*, *25*, 69–75.
- Constantine, A. G. (1966). The distribution of Hotelling’s generalised T_0^2 . *Annals of Mathematical Statistics*, *37*, 215–225.
- Damon, B. M., Ding, Z., Anderson, A. W., Freyer, A. S., Gore, J. C. (2002). Validation of diffusion tensor MRI-based muscle fiber tracking. *Magnetic Resonance in Medicine*, *48*, 97–104.

- Dancis, J., Davis, C. (1987). An interlacing theorem for eigenvalues of self-adjoint operators. *Linear Algebra and its Applications*, 88(89), 117–122.
- Del Moral, P., Niclas, A. (2018). A Taylor expansion of the square root matrix functional. *Journal of Mathematical Analysis and Applications*, 465, 259–266.
- Dryden, I. L., Koloydenko, A., Zhou, D. (2009). Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Annals of Applied Statistics*, 3, 1102–1123.
- Faraut, J., Korányi, A. (1994). *Analysis on symmetric cones*. Oxford: Oxford University Press.
- Farrell, R. H. (1985). *Multivariate calculation*. New York: Springer.
- Gikhman, I. I., Skorokhod, A. V. (1980). *The theory of stochastic processes*, Vol. 1. New York: Springer.
- Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *The Annals of Mathematical Statistics*, 34, 152–177.
- Gourieroux, C., Sufana, R. (2010). Derivative pricing with Wishart multivariate stochastic volatility. *Journal of Business & Economic Statistics*, 28, 438–451.
- Gross, K. I., Richards, D. St. P. (1987). Special functions of matrix argument. I. Algebraic induction, zonal polynomials, and hypergeometric functions. *Transactions of the American Mathematical Society*, 301, 781–811.
- Gupta, R. D., Richards, D. St. P. (1987). Multivariate Liouville distributions. *Journal of Multivariate Analysis*, 23, 233–256.
- Gupta, R. D., Richards, D. St. P. (1995). Multivariate Liouville distributions, IV. *Journal of Multivariate Analysis*, 54, 1–17.
- Hadjicosta, E. (2019). *Integral transform methods in goodness-of-fit testing*. Doctoral dissertation, Pennsylvania State University, University Park.
- Hadjicosta, E., Richards, D. (2019). Integral transform methods in goodness-of-fit testing, I: The gamma distributions. *Metrika* (to appear) (Preprint, [arXiv:1810.07138](https://arxiv.org/abs/1810.07138)).
- Haff, L. R., Kim, P. T., Koo, J. Y., Richards, D. St. P. (2011). Minimax estimation for mixtures of Wishart distributions. *Annals of Statistics*, 39, 3417–3440.
- Henze, N., Meintanis, S. G., Ebner, B. (2012). Goodness-of-fit tests for the gamma distribution based on the empirical Laplace transform. *Communications in Statistics — Theory & Methods*, 41, 1543–1556.
- Herz, C. S. (1955). Bessel functions of matrix argument. *Annals of Mathematics*, 61, 474–523.
- Hochstadt, H. (1973). One-dimensional perturbations of compact operators. *Proceedings of the American Mathematical Society*, 37, 465–467.
- Horn, R. A., Johnson, C. R. (1990). *Matrix analysis*. New York: Cambridge University Press.
- Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48, 419–426.
- James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *The Annals of Mathematical Statistics*, 35, 475–501.
- Jian, B., Vemuri, B. C. (2007). Multi-fiber reconstruction from diffusion MRI using mixture of Wisharts and sparse deconvolution. *Information Processing in Medical Imaging*, 20, 384–395.
- Jian, B., Vemuri, B. C., Özarslan, E., Carney, P. R., Mareci, T. H. (2007). A novel tensor distribution model for the diffusion-weighted MR signal. *NeuroImage*, 37, 164–176.
- Kågström, B. (1977). Bounds and perturbation bounds for the matrix exponential. *BIT Numerical Mathematics*, 17, 39–57.
- Khatri, C. G. (1966). On certain distribution problems based on positive definite quadratic functions in normal vectors. *The Annals of Mathematical Statistics*, 37, 468–479.
- Kim, P. T., Richards, D. St. P. (2011). Deconvolution density estimation on the space of positive definite symmetric matrices. In D. Hunter, et al. (Eds.), *Nonparametric statistics and mixture models*, pp. 58–68. Singapore: World Scientific Press.
- Koev, P., Edelman, A. (2006). The efficient evaluation of the hypergeometric function of a matrix argument. *Mathematics of Computation*, 75, 833–846.
- Kotz, S., Johnson, N. L., Boyd, D. W. (1967). Series representations of distributions of quadratic forms in normal variables. I. Central case. *Annals of Mathematical Statistics*, 38, 823–837.
- Ku, Y. C., Bloomfield, P. (2010). *Generating random Wishart matrices with fractional degrees of freedom in OX*. Preprint: North Carolina State University, Raleigh.
- Le Maître, O. P., Knio, O. M. (2010). *Spectral methods for uncertainty quantification*. New York: Springer.
- Ledoux, M., Talagrand, M. (1991). *Probability in Banach spaces*. New York: Springer.
- Lee, H. N., Schwartzman, A. (2017). Inference for eigenvalues and eigenvectors in exponential families of random symmetric matrices. *Journal of Multivariate Analysis*, 162, 152–171.

- Maass, H. (1971). *Siegel's modular forms and Dirichlet series*. Lecture Notes in Mathematics, Vol 216. New York: Springer.
- Matsui, M., Takemura, A. (2008). Goodness-of-fit tests for symmetric stable distributions—Empirical characteristic function approach. *TEST*, 17, 546–566.
- Matthews, P. M., Arnold, D. L. (2001). Magnetic resonance imaging of multiple sclerosis: New insights linking pathology to clinical evolution. *Current Opinion in Neurology*, 14, 279–287.
- Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. New York: Wiley.
- Neumann-Haefelin, T., Moseley, M. E., Albers, G. W. (2000). New magnetic resonance imaging methods for cerebrovascular disease: Emerging clinical applications. *Annals of Neurology*, 47, 559–570.
- Pomara, N., Crandall, D. T., Choi, S. J., Johnson, G., Lim, K. O. (2001). White matter abnormalities in HIV-1 infection: A diffusion tensor imaging study. *Psychiatry Research: Neuroimaging*, 106, 15–24.
- Richards, D. St. P. (2010). Chapter 35: Functions of Matrix Argument. In F. W. Olver, D. W. Lozier, R. F. Boisvert & C. W. Clark (Eds.), *NIST handbook of mathematical functions*. New York: Cambridge University Press.
- Rosenbloom, M., Sullivan, E. V., Pfefferbaum, A. (2003). Using magnetic resonance imaging and diffusion tensor imaging to assess brain damage in alcoholics. *Alcohol Research and Health*, 27, 146–152.
- Schwartzman, A. (2006). *Random ellipsoids and false discovery Rates: Statistics for diffusion tensor imaging data*, Doctoral dissertation, Stanford University, Palo Alto.
- Schwartzman, A., Dougherty, R. F., Taylor, J. E. (2005). Cross-subject comparison of principal diffusion direction maps. *Magnetic Resonance in Medicine*, 53, 1423–1431.
- Schwartzman, A., Dougherty, R. F., Taylor, J. E. (2008). False discovery rate analysis of brain diffusion direction maps. *The Annals of Applied Statistics*, 2, 153–175.
- Severini, T. A. (2005). *Elements of distribution theory*. New York: Cambridge University Press.
- Shilov, G. E. (1977). *Linear algebra*. New York: Dover.
- Siriteanu, C., Kuriki, S., Richards, D., Takemura, A. (2016). Chi-square mixture representations for the distribution of the scalar Schur complement in a noncentral Wishart matrix. *Statistics and Probability Letters*, 115, 79–87.
- Siriteanu, C., Takemura, A., Kuriki, S., Richards, D., Shin, H. (2015). Schur complement based analysis of MIMO zero-forcing for Rician fading. *IEEE Transactions on Wireless Communications*, 14, 1757–1771.
- Taherizadeh, F. (2009). *Empirical Hankel transform and statistical goodness-of-fit tests for exponential distributions*, Doctoral dissertation, Leibniz Universität Hannover, Hannover.
- Tulino, A. M., Verdú, S. (2004). *Random matrix theory and wireless communications*. Hanover, MA: Now Publishers.
- Wieand, H. S. (1976). A condition under which the Pitman and Bahadur approaches to efficiency coincide. *Annals of Statistics*, 4, 1003–1011.
- Wihler, T. P. (2009). On the Hölder continuity of matrix functions for normal matrices. *Journal of Inequalities in Pure and Applied Mathematics*, 10, 1–5.
- Young, N. (1998). *An introduction to Hilbert space*. New York: Cambridge University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.