Annals of the Institute of Statistical Mathematics manuscript No.
(will be inserted by the editor)

The Limit Distribution of weighted L^2-Goodness-of-Fit Statistics under fixed Alternatives, with Applications

L. Baringhaus · B. Ebner · N. Henze

Received: date / Accepted: date

Abstract We present a general result on the limit distribution of weighted one- and two-sample L^2-goodness-of-fit test statistics of some hypothesis H_0 under fixed alternatives. Applications include an approximation of the power function of such tests, asymptotic confidence intervals of the distance of an underlying distribution with respect to the distributions under H_0 and an asymptotic equivalence test that is able to validate certain neighborhoods of H_0.

Keywords Goodness-of-fit test · Weighted L^2-statistic · Fixed alternative · Empirical transform · Asymptotic equivalence test

Mathematics Subject Classification (2000) 62G10 · 62G20 · 62G15

1 Introduction

For more than 30 years, numerous goodness-of-fit tests (GOF tests) based on weighted L^2-statistics involving empirical transforms such as the empirical characteristic function (ECF), the empirical Laplace transform (ELT), the empirical moment generating function (EMF), the empirical probability generating function (EGF), the empirical Mellin transform (EMT) and the empirical Hankel transform (EHT) have been proposed for various testing problems. The following list is not exhaustive, but nevertheless shows that such statistics have gained much interest.

Epps and Pulley (1983) considered testing for univariate normality by means of the ECF, and Baringhaus and Henze (1988) and Henze and Zirkler (1990) generalized their approach to the multivariate case. Testing for the Poisson distribution by means of L^2-statistics based on the EGF was studied by Rueda et al. (1991), Baringhaus and Henze (1992) and Gürtler and Henze (2000b). Baringhaus and Henze (1991), Henze (1993), Henze and Meintanis (2002a), Henze and Meintanis (2002b), Henze and Meintanis (2005) and Henze and Meintanis (2010) considered corresponding statistics based on the ELT or the ECF for testing the hypothesis that the underlying distribution is exponential, and Ebner et al. (2012) studied a goodness-of-fit test for the gamma distribution. A weighted L^2-statistic based on

L. Baringhaus
Institut für Mathematische Stochastik, Leibniz Universität Hannover, Welfengarten 1, D-30167 Hannover
E-mail: lbaring@stochastik.uni-hannover.de

B. Ebner and N. Henze
Karlsruher Institut für Technologie (KIT), Institut für Stochastik, Englerstraße 2, D-76131 Karlsruhe
E-mail: Bruno.ebner@kit.edu, norbert.henze@kit.edu

the ELT for testing that the underlying distribution is inverse Gaussian was studied by Henze and Klar (2002), and Gürtler and Henze (2000a) proposed an ECF-based statistic for testing for the Cauchy family. Meintanis (2010) employed the EMF to test for the family of skew-normal distributions, and Meintanis and Tsionas (2010) used the EMF to construct a GOF test for the normal-Laplace distribution. A weighted L^2-statistic using the ECF for testing GOF for normal inverse Gaussian distributions was considered by Fragiadakis et al. (2009), and Meintanis (2008a) employed such statistic based on the EMT to construct tests for generalized exponential laws. Moreover, Iliopoulos and Meintanis (2003) studied L^2-statistics in connection with GOF testing for the Rayleigh distribution, Meintanis (2008b) considered testing for the lognormal family by means of the EMF, and Meintanis (2007) used the ELT for L^2-based tests of fit for bivariate Marshall-Olkin distributions. Meintanis (2004a) and Meintanis (2004b) considered L^2-type statistics for testing GOF for the Laplace and the logistic family of distributions, respectively. More recently, Fermanian (2009) and Genest et al. (2011) employed L^2-statistics for testing for parametric families of copula functions. Last but not least, Alba-Fernández and Jiménez-Gamero (2015) and Novoa-Muñoz and Jiménez-Gamero (2014) studied L^2-type statistics to test for bivariate exponential and bivariate Poisson distributions, respectively.

A weighted L^2-statistic takes the form

$$T_n = n \int_M Z_n^2(t)\, \mu(\mathrm{d}t), \tag{1.1}$$

where M is a Borel subset of $\mathbb{R}^d$, μ is a finite measure on the Borel subsets of M and $Z_n(t) = Z_n(X_1, \ldots, X_n, t)$ is a real-valued measurable function of (not necessarily independent and identically distributed (i.i.d.)) $\mathbb{R}^d$-valued random (column) vectors. For asymptotic theorems, we assume that $X_1, \ldots, X_n$ is the beginning of an infinite sequence $X_1, X_2, \ldots$, where $X_1, X_2, \ldots$ are defined on a common probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Usually, μ is given by a nonnegative weight function w defined on M, and we have $\mu(\mathrm{d}t) = w(t)\, \mathrm{d}t$, where $\mathrm{d}t$ means integration with respect to Lebesgue measure on $\mathbb{R}^d$. The weight function w is often chosen to give T_n a simple expression that is suitable for computations.

Each of the L^2-statistics in the papers listed above has been proposed, in a setting of i.i.d. copies $X_1, X_2, \ldots$ of a random vector X, to test a hypothesis of the type

$$H_0 : \mathbb{P}^X \in \mathcal{Q} = \{Q_\vartheta : \vartheta \in \Theta\}, \tag{1.2}$$

where $\mathbb{P}^X$ denotes the distribution of X, $\mathcal{Q}$ is a family of d-variate distributions indexed by some finite-dimensional parameter ϑ. Weighted L^2-statistics, however, have also been studied for testing nonparametric hypotheses, see e.g. Henze et al. (2003), Ngatchou-Wandji (2009) and Leucht (2012) in the context of testing for reflected symmetry about an unspecified point. The latter paper even relaxed the i.i.d.-assumption.

As an example, we consider the now classical BHEP statistic for testing the hypothesis H_0 that X has some nondegenerate d-variate normal distribution (see, e.g., Henze (2002), Section 6). This statistic is defined as

$$T_n \;=\; n \int_{\mathbb{R}^d} \left| \psi_n(t) - \exp\left(-\frac{\|t\|^2}{2} \right) \right|^2 w(t)\, \mathrm{d}t. \tag{1.3}$$

Here, $\psi_n(t) = n^{-1} \sum_{j=1}^n \exp(\mathrm{i} t^\top Y_{n,j})$ is the ECF of the so-called scaled residuals $Y_{n,j} = S_n^{-1/2}(X_j - \overline{X}_n)$, $j = 1, \ldots, n$, where $\overline{X}_n = n^{-1} \sum_{j=1}^n X_j$ is the sample mean and $S_n^{-1/2}$ denotes the symmetric positive definite square root of the inverse of the sample covariance matrix $S_n = n^{-1}\sum_{j=1}^n (X_j - \overline{X}_n)(X_j - \overline{X}_n)^\top$, which exists almost surely if $n > d$ and the distribution of X is absolutely continuous. Moreover, i stands for the imaginary unit, $\|\cdot\|$ denotes the Euclidean norm, and $x^\top$ is the transpose of a column vector x. With the weight function

$$w(t) = w_\beta(t) = \frac{1}{(2\pi\beta^2)^{d/2}} \exp\left(-\frac{\|t\|^2}{2\beta^2}\right), \tag{1.4}$$

where $\beta > 0$ is fixed, T_n has the simple expression

$$\begin{aligned} T_n = {} & \frac{1}{n}\sum_{j,k=1}^n \exp\left(-\frac{\beta^2}{2}\|Y_{n,j} - Y_{n,k}\|^2\right) \\ & -2(1+\beta^2)^{-d/2}\sum_{j=1}^n \exp\left(-\frac{\beta^2\|Y_{n,j}\|^2}{2(1+\beta^2)}\right) + n(1+2\beta^2)^{-d/2}. \end{aligned}$$

Putting

$$Z_n(t) \ = \ \frac{1}{n}\sum_{j=1}^n \left[\cos(t^\top Y_{n,j}) + \sin(t^\top Y_{n,j}) - \exp\left(-\frac{1}{2}\|t\|^2\right)\right]$$

yields the form (1.1), where μ is the centered d-variate normal distribution with independent components, each having variance β^2.

Theoretical results on weighted L^2-statistics usually involve a nondegenerate limit distribution of T_n under H_0, the limit distribution of T_n under contiguous alternatives to H_0, and a stochastic limit of T_n/n under a fixed alternative distribution. More precisely, one has (writing $\stackrel{\mathbb{P}}{\longrightarrow}$ for convergence in probability)

$$\frac{T_n}{n} \stackrel{\mathbb{P}}{\longrightarrow} \Delta := \int_M z^2(t)\,\mu(\mathrm{d}t) > 0 \tag{1.5}$$

for some measurable function z on M. The latter result is then used to prove the consistency of a GOF test that rejects H_0 for large values of T_n.

With very few exceptions, for example Naito (1997), Bücher and Dette (2010) and Gürtler (2000), there is no stronger result under a fixed alternative. Naito (1997) proved asymptotic normality for weighted L^2-statistics under an i.i.d.-setting for testing for parametric models with regular estimators using the theory of V-statistics with estimated parameters, as given, e.g., in De Wet and Randles (1987). Gürtler (2000) showed in the special case of the BHEP statistic for testing for multivariate normality that, under a fixed alternative distribution satisfying a weak moment condition, we have

$$\sqrt{n}\left(\frac{T_n}{n} - \Delta\right) \stackrel{\mathcal{D}}{\longrightarrow} \mathrm{N}(0, \sigma^2),$$

where $\sigma^2 > 0$ depends on the underlying distribution, and $\stackrel{\mathcal{D}}{\longrightarrow}$ denotes convergence in distribution of random vectors and stochastic processes.

By using a very general Hilbert space approach that carves out the quintessence of asymptotic normality of weighted L^2-statistics (see (2.1)), we will show that, under general conditions, such statistics have centered normal limit distributions. Applications include an approximation of the power function of a GOF test based on T_n, an asymptotic confidence interval for the stochastic limit Δ figuring in (1.5) and an asymptotic 'inverse GOF test' which tests, for a given value $\Delta_0 > 0$, the hypothesis that $\varDelta \geq \varDelta_0$ against the alternative $\varDelta < \varDelta_0$.

The paper is organized as follows. Section 2 contains the main result, the implications of which are given in Section 3. Section 4 discusses some examples. In Section 5 we show that the main result can readily be generalized to the case of two-sample weighted L^2-statistics. Section 6 contains a brief summary. For the sake of readability, the proof of Theorem 2 is deferred to Section 7.

2 The main result

To establish our main result, write $\mathcal{B}^d$ for the Borel σ-field of subsets of $\mathbb{R}^d$, and let $\mathcal{H} = L^2(M, \mathcal{B}^d \cap M, \mu)$ be the Hilbert space of (equivalence classes of) square integrable measurable functions on M, equipped with the scalar product $\langle g, h\rangle = \int_M gh\,\mathrm{d}\mu$. Furthermore, let $\|h\|_{L^2}^2 = \int_M h^2\,\mathrm{d}\mu$. We assume that Z_n figuring in (1.1) is a random element of $\mathcal{H}$, which implies $T_n = \|\sqrt{n}Z_n\|_{L^2}^2$. In an i.i.d. setting as stated before (1.2), we typically have $\sqrt{n}Z_n \xrightarrow{\mathcal{D}} Z$ under H_0, where Z is a centered Gaussian element of $\mathcal{H}$. From the Continuous mapping theorem, it thus follows that $T_n \xrightarrow{\mathcal{D}} \|Z\|_{L^2}^2$. The distribution of $\|Z\|_{L^2}^2$ is that of $\sum_{j\geq 1}\lambda_j N_j^2$, where $N_1, N_2, \ldots$ are i.i.d. standard normal random variables, and $\lambda_1, \lambda_2, \ldots$ are the eigenvalues corresponding to eigenfunctions of the integral equation $\lambda f(s) = \int_M K(s,t) f(t)\mu(\mathrm{d}t)$ associated with the covariance kernel K of Z. Under contiguous alternatives to H_0, the limit distribution of T_n is that of $\|Z + c\|_{L^2}^2$, where c is some shift function on M.

To consider the hitherto largely neglected behavior of weighted L^2-statistics under a fixed alternative to H_0, notice that the stochastic limit $\varDelta$ figuring in (1.5) is $\|z\|_{L^2}^2$. We thus have

$$\begin{aligned}\sqrt{n}\left(\frac{T_n}{n} - \varDelta\right) &= \sqrt{n}\left(\|Z_n\|_{L^2}^2 - \|z\|_{L^2}^2\right)\\ &= \sqrt{n}\big\langle Z_n - z, Z_n + z\big\rangle\\ &= \sqrt{n}\big\langle Z_n - z, 2z + Z_n - z\big\rangle\\ &= 2\big\langle \sqrt{n}(Z_n - z), z\big\rangle + \frac{1}{\sqrt{n}}\|\sqrt{n}(Z_n - z)\|_{L^2}^2. \end{aligned} \tag{2.1}$$

This decomposition of $\sqrt{n}(T_n/n - \varDelta)$, which in connection with the Cramér-von Mises statistic has already been used by Chapman (1958), immediately leads to our main result.

Theorem 1. *Let $(T_n)_{n\geq 1}$ be a sequence of weighted L^2-statistic based on (not necessarily i.i.d.) d-dimensional random vectors $X_1, X_2, \ldots$ satisfying* (1.5)*. Putting $W_n(\cdot) = \sqrt{n}(Z_n(\cdot) - z(\cdot))$, suppose further that, as $n \to \infty$,*

$$W_n \xrightarrow{\mathcal{D}} W$$

in $\mathcal{H}$, where W is a centered Gaussian element of $\mathcal{H}$ with covariance kernel $K(s,t) = \mathbb{E}[W(s)W(t)]$. Then

$$\sqrt{n}\left(\frac{T_n}{n} - \varDelta\right) \xrightarrow{\mathcal{D}} \mathrm{N}(0, \sigma^2),$$

where

$$\sigma^2 = 4\int_M\int_M K(s,t)z(s)z(t)\mu(\mathrm{d}s)\mu(\mathrm{d}t). \tag{2.2}$$

Proof. From (2.1) we have

$$\sqrt{n}\left(\frac{T_n}{n} - \Delta\right) \;=\; 2\langle W_n, z\rangle + \frac{1}{\sqrt{n}}\|W_n\|_{L^2}^2.$$

Since $W_n \xrightarrow{\mathcal{D}} W$ and the continuous mapping theorem imply $\langle W_n, z\rangle \xrightarrow{\mathcal{D}} \langle W, z\rangle$ and $\|W_n\|_{L^2}^2 \xrightarrow{\mathcal{D}} \|W\|_{L^2}^2$, it follows that

$$\sqrt{n}\left(\frac{T_n}{n} - \Delta\right) \xrightarrow{\mathcal{D}} 2\langle W, z\rangle.$$

Now, $\langle W, z\rangle$ has a centered normal distribution with variance

$$\begin{aligned}\mathbb{E}\left[\langle W, z\rangle^2\right] &= \mathbb{E}\left[\int_M W(s)z(s)\mu(\mathrm{d}s)\int_M W(t)z(t)\,\mu(\mathrm{d}t)\right]\\ &= \int_M\int_M \mathbb{E}\left[W(s)W(t)\right] z(s)z(t)\mu(\mathrm{d}s)\mu(\mathrm{d}t),\end{aligned}$$

proving the assertion. □

The following corollary is an immediate consequence of Theorem 1 and Sluzky's Lemma.

Corollary 1. *Suppose that, in the setting of Theorem* 1*,* σ^2 *figuring in* (2.2) *is positive. Suppose further that* $\widehat{\sigma}_n^2 = \widehat{\sigma}_n^2(X_1, \ldots, X_n)$ *is a (weakly) consistent sequence of estimators of* σ^2*. Then*

$$\frac{\sqrt{n}}{\widehat{\sigma}_n}\left(\frac{T_n}{n} - \Delta\right) \;\xrightarrow{\mathcal{D}}\; \mathrm{N}(0,1). \tag{2.3}$$

3 Applications

There are some immediate consequences of Theorem 1 and Corollary 1. To this end, consider testing a hypothesis $H_0 : \mathbb{P}^X \in \mathcal{Q} = \{Q_\vartheta : \vartheta \in \Theta\}$ in the setting that $X_1, X_2, \ldots$ are i.i.d. copies of a random vector X. Write F for the distribution function of X under a fixed alternative distribution to H_0. In what follows, assume that the assumptions of Theorem 1 and Corollary 1 hold.

3.1 A confidence interval for Δ

Fix $\alpha \in (0,1)$, and let $u_\alpha = \Phi^{-1}(1-\alpha/2)$ be the $(1-\alpha/2)$-quantile of the standard normal distribution. From (2.3) it follows at once that

$$I_n := \left[\frac{T_n}{n} - \frac{u_\alpha \widehat{\sigma}_n}{\sqrt{n}}, \frac{T_n}{n} + \frac{u_\alpha \widehat{\sigma}_n}{\sqrt{n}}\right] \tag{3.1}$$

is an asymptotic (two-sided) confidence interval at level $1-\alpha$ for Δ, i.e., we have $\lim_{n\to\infty} \mathbb{P}_F(I_n \ni \Delta) = 1 - \alpha$.

Notice that $\Delta = \Delta(F)$ may be regarded as some kind of 'distance' between the true underlying distribution and the distributions in $\mathcal{Q}$, for which $\Delta = 0$. Thus, I_n is an asymptotic confidence interval for this 'distance'.

3.2 Approximation of the power function

As a second application of Theorem 1, we obtain an approximation of the power function of a GOF test that rejects H_0 for large values of T_n. As stated at the beginning Section 2, the limit distribution of T_n under H_0 is the distribution of a sum of weighted independent χ_1^2-variates. When testing for multivariate normality and T_n is affine invariant or when testing for exponentiality and T_n is scale invariant, the distribution of T_n under H_0 does not depend on the value of the 'true' parameter $\vartheta \in \Theta$. In this case, $(1-\alpha)$-quantiles of T_n are obtained by simulation. For the BHEP statistic for testing for multivariate normality, the first three moments of the limit distribution have been established, and a three-parameter lognormal distribution has been fitted to the unknown limit distribution (see Henze and Wagner (1997)). Thus, by the quantiles of the fitted lognormal distribution, alternative critical values are available for large samples. In general, a parametric bootstrap procedure is needed to find critical values in order to carry out the test. Suppose that H_0 is rejected if $T_n > c_n$, and $\lim_{n\to\infty} \mathbb{P}_\vartheta(T_n > c_n) = \alpha$ for each $\vartheta \in \Theta$, where (c_n) is a sequence of constants. If F is the distribution function of an alternative distribution satisfying the assumptions of Theorem 1 and Corollary 1, it follows that the power of the test against this alternative can be approximated by

$$\begin{aligned} \mathbb{P}_F(T_n > c_n) &= \mathbb{P}_F\left(\frac{\sqrt{n}}{\sigma}\left(\frac{T_n}{n} - \Delta\right) > \frac{\sqrt{n}}{\sigma}\left(\frac{c_n}{n} - \Delta\right)\right) \\ &\approx 1 - \Phi\left(\frac{\sqrt{n}}{\sigma}\left(\frac{c_n}{n} - \Delta\right)\right), \end{aligned} \tag{3.2}$$

where Φ is the standard normal distribution function. If a bootstrap procedure is used, let $\widehat{\vartheta}_n = \widehat{\vartheta}_n(X_1, \ldots, X_n)$ be some suitable estimator of $\vartheta \in \Theta$. Denote by $\mathcal{L}(T_n|F)$ the distribution of T_n if $X_1 \sim F$, and by c_n the $(1-\alpha)$-quantile of $\mathcal{L}(T_n|F_{\widehat{\vartheta}_n})$. In typical cases, given $X_1, X_2, \ldots$ with common distribution F, as $n \to \infty$, the weak limit μ_F, say, of $\mathcal{L}(T_n|F_{\widehat{\vartheta}_n})$ exists almost surely, and $c_n \to c$ almost surely, where c is the $(1-\alpha)$-quantile of μ_F. Then, for a given alternative distribution F satisfying the assumptions of Theorem 1 and Corollary 1 we are led to approximate the power ot the test that rejects H_0 if $T_n > c_n$ by

$$\mathbb{P}_F(T_n > c_n) \;\approx\; \mathbb{P}_F(T_n > c) \;\approx\; 1 - \Phi\left(\frac{\sqrt{n}}{\sigma}\left(\frac{c}{n} - \Delta\right)\right).$$

3.3 Neighborhood-of-model validation

A third application of the main results refers to a fundamental drawback inherent in any GOF test. If a level-α-test of H_0 does not lead to a rejection of H_0, the hypothesis H_0 is by no means 'validated' or 'confirmed'. There is probably only not enough evidence to reject it! Suppose on the other hand that we want to tolerate a given 'distance' Δ_0 and consider the 'inverse' testing problem

$$H_{\Delta_0} : \Delta(F) \geq \Delta_0 \;\text{ against }\; K_{\Delta_0} : \Delta(F) < \Delta_0.$$

Here, the dependence of Δ on the underlying distribution has been made explicit.

From (2.3), we obtain the following asymptotic level-α-test of H_{Δ_0} against K_{Δ_0}: This test rejects H_{Δ_0} if

$$\frac{T_n}{n} \;\leq\; \Delta_0 - \frac{\widehat{\sigma}_n}{\sqrt{n}} \Phi^{-1}(1-\alpha).$$

Using (2.3) we have for each $F \in H_{\Delta_0}$

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbb{P}_F\left(\frac{T_n}{n} \le \Delta_0 - \frac{\widehat{\sigma}_n}{\sqrt{n}}\Phi^{-1}(1-\alpha)\right) \\
&= \limsup_{n\to\infty} \mathbb{P}_F\left(\frac{\sqrt{n}}{\widehat{\sigma}_n}\left(\frac{T_n}{n} - \Delta_0\right) \le -\Phi^{-1}(1-\alpha)\right) \\
&\le \alpha.
\end{aligned}$$

Thus, the test has asymptotic level α. Moreover, we have

$$\lim_{n\to\infty} \mathbb{P}_F\left(\frac{T_n}{n} \le \Delta_0 - \frac{\widehat{\sigma}_n}{\sqrt{n}}\Phi^{-1}(1-\alpha)\right) \;=\; \alpha$$

for each F such that $\Delta(F) = \Delta_0$. It is easy to see that

$$\lim_{n\to\infty} \mathbb{P}_F\left(\frac{T_n}{n} \le \Delta_0 - \frac{\widehat{\sigma}_n}{\sqrt{n}}\Phi^{-1}(1-\alpha)\right) = 1$$

if $\Delta(F) < \Delta_0$. Thus, the test is consistent against each alternative.

Notice that this test is in the spirit of bioequivalence testing (see, e.g., Czado et al. (2007), Dette and Munk (2003) or Wellek (2010)), since it aims at validating a certain neighborhood of a hypothesized model.

4 Examples

Example 1. Gürtler (2000) considered the statistics T_n figuring in (1.3) with the weight function w_β given in (1.4), which is the BHEP statistic for testing for multivariate normality. For the sake of simplicity, we assume $d = 1$ in what follows. Then, writing $\varphi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ for the standard normal density,

$$T_{n,\beta} = n\int_{-\infty}^{\infty}\left|\psi_n(t) - \exp\left(-\frac{t^2}{2}\right)\right|^2 \frac{1}{\beta}\varphi\left(\frac{t}{\beta}\right)\,\mathrm{d}t$$

yields the statistic of Epps and Pulley (1983) for testing for normality. Let

$$\Delta_\beta = \int_{-\infty}^{\infty}\left|C(t) - \exp\left(-\frac{t^2}{2}\right)\right|^2 \frac{1}{\beta}\varphi\left(\frac{t}{\beta}\right)\,\mathrm{d}t, \tag{4.1}$$

where $C(t) = \mathbb{E}\left(\cos(tX) + \sin(tX)\right)$ and X is assumed to be standardized, i.e., we have $\mathbb{E}(X) = 0$ and $\mathbb{V}(X) = 1$.

Under the additional condition $\mathbb{E}(X^4) < \infty$, Gürtler (2000) proved that Theorem 1 holds with a centered Gaussian element W of $\mathcal{H} := L^2(\mathbb{R}, \mathcal{B}, w_\beta(t)\mathrm{d}t)$ having covariance kernel

$$\begin{aligned}
K(s,t) = {} & R(t-s) + I(s+t) - C(s)C(t) + tD(t)D'(s) + sD(s)D'(t) + stD(s)D(t) \\
& + \frac{1}{2}\Big\{tC''(s)C'(t) + sC''(t)C'(s)\Big\} + \frac{1}{2}\Big\{tC(s)C'(t) + sC(t)C'(s)\Big\} \\
& + \frac{m_3}{2}st\Big\{D(s)\left(R'(t) + I'(t)\right) + D(t)\left(R'(s) + I'(s)\right)\Big\} \\
& + \frac{1}{4}(m_4 - 1)stC'(s)C'(t).
\end{aligned}$$

Here, $R(t) = \mathbb{E}[\cos(tX)]$, $I(t) = \mathbb{E}[\sin(tX)]$, $C(t) = R(t) + I(t)$, $D(t) = R(t) - I(t)$ and $m_j = \mathbb{E}[X^j]$, $j = 3, 4$. Notice that differentiation can be carried out beneath the expectation operator. For example, we have $R'(t) = -\mathbb{E}[X\sin(tX)]$.

In this case, the function z figuring in (1.5) is

$$z(t) = C(t) - \exp\left(-\frac{t^2}{2}\right),$$

and σ^2 in (2.2) takes the form

$$\sigma^2 = 4\iint K(s,t)\left(C(s) - \exp\left(-\frac{s^2}{2}\right)\right)\left(C(t) - \exp\left(-\frac{t^2}{2}\right)\right)\frac{1}{\beta^2}\varphi\left(\frac{s}{\beta}\right)\varphi\left(\frac{t}{\beta}\right)\mathrm{d}s\mathrm{d}t,$$

where $\int$ is shorthand for $\int_{-\infty}^{\infty}$ (see also Naito (1997), p. 208, for the special case $\beta = 1$).

A consistent estimator $\widehat{\sigma}_n^2$ of σ^2 is obtained if in the above expression $K(s,t)$ is replaced with $K_n(s,t)$ and $C(u)$ with $C_n(u)$, $u \in \{s,t\}$. Here, K_n results from K stated above by replacing R with R_n, I with I_n, C with C_n, D with D_n m_3 with $m_{3,n}$ and m_4 with $m_{4,n}$, where

$$R_n(s) = \frac{1}{n}\sum_{j=1}^{n}\cos(sY_j), \ I_n(s) = \frac{1}{n}\sum_{j=1}^{n}\sin(sY_j),$$

$C_n(s) = R_n(s) + I_n(s)$, $D_n(s) = R_n(s) - I_n(s)$, $m_{l,n} = n^{-1}\sum_{j=1}^{n} Y_j^l$, $l \in \{3,4\}$, and $Y_j = (X_j - \overline{X}_n)/(n^{-1}\sum_{l=1}^{n}(X_l - \overline{X}_n)^2)^{1/2}$, $j = 1, \ldots, n$.

All the resulting integrals may be expressed in terms of

$$\begin{aligned}
J_1(Y_j) &= (1+\beta^2)^{-d/2}\exp\left(-\frac{\beta^2 Y_j^2}{2(1+\beta^2)}\right),\\
J_2(Y_j,Y_k) &= \frac{1}{2}\left[\exp\left(-\frac{\beta^2}{2}(Y_j - Y_k)^2\right) + \exp\left(-\frac{\beta^2}{2}(Y_j + Y_k)^2\right)\right],\\
J_3(Y_j,Y_k) &= \frac{1}{2}\left[\exp\left(-\frac{\beta^2}{2}(Y_j - Y_k)^2\right) - \exp\left(-\frac{\beta^2}{2}(Y_j + Y_k)^2\right)\right],\\
J_4(Y_j) &= \frac{Y_j\beta^2}{(1+\beta^2)^{1+d/2}}\exp\left(-\frac{\beta^2 Y_j^2}{2(1+\beta^2)}\right).
\end{aligned}$$

By tedious calculations one obtains

$$\widehat{\sigma}_n^2 = 4\cdot\left\{K_{1n} + K_{2n}(K_{3n} - K_{2n}) - \frac{1}{4}K_{3n}^2 + K_{5n}^2 + K_{5n}\left(2K_{4n} + K_{6n}\right) + K_{9n}\left(K_{7n} + \frac{1}{4}K_{8n}\right)\right\}.$$

Here,

$$\begin{aligned}
K_{1n} &= \frac{1}{n}\sum_{j=1}^{n}\left\{\left(\frac{1}{n}\sum_{k=1}^{n}J_2(Y_j,Y_k)-J_1(Y_j)\right)^2+\left(\frac{1}{n}\sum_{k=1}^{n}J_3(Y_j,Y_k)\right)^2\right\}\\
&\quad+\frac{2}{n}\sum_{j=1}^{n}\left\{\frac{1}{n}\sum_{k=1}^{n}J_3(Y_j,Y_k)\right\}\left\{\frac{1}{n}\sum_{k=1}^{n}J_2(Y_k,Y_k)-J_1(Y_j)\right\},\\
K_{2n} &= \frac{1}{n^2}\sum_{j,k=1}^{n}\exp\left(-\frac{\beta^2}{2}(Y_j-Y_k)^2\right)-\frac{1}{n}\sum_{j=1}^{n}J_1(Y_j),\\
K_{3n} &= \frac{\beta^2}{n^2}\sum_{j,k=1}^{n}Y_j(Y_k-Y_j)\exp\left(-\frac{\beta^2}{2}(Y_k-Y_j)^2\right)+\frac{1}{n}\sum_{j=1}^{n}Y_jJ_4(Y_j),\\
K_{4n} &= \frac{1}{n}\sum_{j=1}^{n}Y_jJ_1(Y_j)-\frac{1}{n^2}\sum_{j,k=1}^{n}Y_j\exp\left(-\frac{\beta^2}{2}(Y_j-Y_k)^2\right),\\
K_{5n} &= \frac{1}{n}\sum_{j=1}^{n}J_4(Y_j)+\frac{\beta^2}{n^2}\sum_{j,k=1}^{n}(Y_k-Y_j)\exp\left(-\frac{\beta^2}{2}(Y_k-Y_j)^2\right),\\
K_{6n} &= \frac{1}{n^2}\sum_{j,k=1}^{n}Y_jY_k^3J_4(Y_j)+\frac{\beta^2}{n^3}\sum_{i,j,k=1}^{n}Y_jY_k^3(Y_i-Y_j)\exp\left(-\frac{\beta^2}{2}(Y_i-Y_j)^2\right),\\
K_{7n} &= \frac{1}{n}\sum_{j=1}^{n}Y_j^2J_1(Y_j)-\frac{1}{n^2}\sum_{j,k=1}^{n}Y_j^2\exp\left(-\frac{\beta^2}{2}(Y_j-Y_k)^2\right),\\
K_{8n} &= \frac{1}{n^2}\sum_{j,k=1}^{n}Y_jY_k^4J_4(Y_j)+\frac{\beta^2}{n^3}\sum_{i,j,k=1}^{n}Y_jY_k^4(Y_i-Y_j)\exp\left(-\frac{\beta^2}{2}(Y_i-Y_j)^2\right),\\
K_{9n} &= \frac{1}{n}\sum_{j=1}^{n}Y_jJ_4(Y_j)+\frac{\beta^2}{n^2}\sum_{j,k=1}^{n}Y_j(Y_k-Y_j)\exp\left(-\frac{\beta^2}{2}(Y_k-Y_j)^2\right).
\end{aligned}$$

For the following alternative distributions, considered by Gürtler (2000), we take $\beta = 1$ and put $\Delta := \Delta_1$ (see (4.1)). The first alternative to the normal distribution is the uniform distribution $\mathrm{U}[-1/\sqrt{3}, 1/\sqrt{3}]$, the second alternative the Laplace distribution $L(0, \frac{1}{\sqrt{2}})$ with density $f(x) = \exp(-\sqrt{2}|x|)/\sqrt{2}$, $x \in \mathbb{R}$, and the third is a mixture of the normal distributions $\mathrm{N}(\frac{\sqrt{8}}{3}, \frac{1}{9})$ and $\mathrm{N}(-\frac{\sqrt{8}}{3}, \frac{1}{9})$ with equal mixing probabilities, abbreviated by $\mathrm{NMIX}(\pm\frac{\sqrt{8}}{3}, \frac{1}{9})$. Notice that these distributions are standardized. The values of Δ are 0.00647 for $\mathrm{U}[-1/\sqrt{3}, 1/\sqrt{3}]$, 0.00660 for $L(0, \frac{1}{\sqrt{2}})$ and 0.02005 for $\mathrm{NMIX}(\pm\frac{\sqrt{8}}{3}, \frac{1}{9})$.

Table 1 shows the empirical coverage probabilities of the confidence interval (3.1) for Δ, each based on 10 000 replications, for the three alternatives and the sample sizes $n = 20$, $n = 50$, $n = 100$ and $n = 200$. The nominal level is $1 - \alpha = 0.9$.

Obviously, these empirical values are close to the nominal value 0.9 even for small sample sizes.

Table 2 displays the empirical power of the BHEP test for normality, rounded to two decimal places and denoted by MC, against the three alternatives discussed above. The nominal level is 0.9, and each value is based on 10 000 replications. Critical values c_n for T_n have been taken from Henze (1990). The columns denoted by App show the corresponding approximations given by the right-hand side of (3.2). Obviously, the approximation seems to be a lower bound for the true power.

n	$\mathrm{U}[-1/\sqrt{3}, 1/\sqrt{3}]$	$\mathrm{L}(0, \frac{1}{\sqrt{2}})$	$\mathrm{NMIX}(\pm\frac{\sqrt{8}}{3}, \frac{1}{9})$
20	0.91	0.89	0.92
50	0.90	0.87	0.90
100	0.91	0.88	0.90
200	0.90	0.89	0.90

Table 1 Empirical coverage probabilities of I_n for Δ (nominal level 0.9, 10 000 replications)

	$\mathrm{U}[-1/\sqrt{3}, 1/\sqrt{3}]$		$\mathrm{L}(0, \frac{1}{\sqrt{2}})$		$\mathrm{NMIX}(\pm\frac{\sqrt{8}}{3}, \frac{1}{9})$	
	MC	App	MC	App	MC	App
$n = 20$	0.27	0.06	0.35	0.24	0.96	0.86
$n = 50$	0.73	0.58	0.62	0.55	1.0	1.0
$n = 100$	0.98	0.94	0.88	0.78	1.0	1.0
$n = 200$	1.0	0.99	0.99	0.91	1.0	1.0

Table 2 Empirical power and approximation (3.2) against selected alternatives

As a second class of alternatives we consider standard normal distributions that are contaminated by the Laplace distribution $\mathrm{L}(0, 1/\sqrt{2})$. In this case, X has the same distribution as $(1-U)N + UL$, where U, N and L are independent, $\mathbb{P}(U = 1) = \varepsilon = 1 - \mathbb{P}(U = 0)$, $N \sim \mathrm{N}(0, 1)$, and $L \sim L(0, 1/\sqrt{2})$, i.e., X has the density

$$f_\varepsilon(x) \;=\; (1-\varepsilon)\,\frac{1}{\sqrt{2\pi}}\,\exp\left(-\frac{x^2}{2}\right) + \varepsilon\,\frac{1}{\sqrt{2}}\,\exp\left(-\sqrt{2}|x|\right), \quad x \in \mathbb{R},$$

where $0 \le \varepsilon \le 1$. Notice that $\mathbb{E}(X) = 0$ and $\mathbb{V}(X) = 1$.

Since $I(t) = \mathbb{E}[\sin(tX)] = 0$, we have

$$C(t) = \mathbb{E}\left[\cos(tX) + \sin(tX)\right] = (1-\varepsilon)\exp\left(-\frac{t^2}{2}\right) + \varepsilon\,\frac{2}{2+t^2}$$

and

$$z(t) = C(t) - \exp\left(-\frac{t^2}{2}\right) = \varepsilon\left(\frac{2}{2+t^2} - \exp\left(-\frac{t^2}{2}\right)\right).$$

From

$$\int_{-\infty}^{+\infty} \frac{1}{t^2+y^2}\mathrm{e}^{-xt^2}\,\mathrm{d}t = \frac{2\pi}{y}\mathrm{e}^{xy^2}\left(1 - \Phi\left(\sqrt{2xy}\right)\right), \; x > 0, y > 0, \tag{4.2}$$

see, e.g. Magnus et al. (1966), p.350, we obtain

$$\begin{aligned}\Delta &= \varepsilon^2 \int_{-\infty}^{+\infty} \left(\frac{2}{2+t^2} - \mathrm{e}^{-\frac{t^2}{2}}\right)^2 \varphi(t)\,\mathrm{d}t \\ &= \varepsilon^2\left(1 - \sqrt{\pi}\mathrm{e}\left(1 - \Phi\left(\sqrt{2}\right)\right) - 4\sqrt{\pi}\mathrm{e}^2\left(1 - \Phi(2)\right) + 3^{-\frac{1}{2}}\right) \\ &= \varepsilon^2 \cdot 0.006602033.\end{aligned}$$

We calculate σ^2 only for the case $\varepsilon = 1$, i.e., for the Laplace distribution. In this case, the covariance kernel takes the form

$$
\begin{aligned}
K(s,t) = & \frac{1}{1+\frac{1}{2}(s-t)^2} - \frac{1}{1+\frac{1}{2}s^2}\frac{1}{1+\frac{1}{2}t^2} + \frac{\frac{1}{4}s^3t^3 - st}{(1+\frac{1}{2}s^2)^2(1+\frac{1}{2}t^2)^2} \\
& - \frac{1}{2}\left\{\frac{t^2}{(1+\frac{1}{2}t^2)^2}\frac{\frac{3}{2}s^2-1}{(1+\frac{1}{2}s^2)^3} + \frac{s^2}{(1+\frac{1}{2}s^2)^2}\frac{\frac{3}{2}t^2-1}{(1+\frac{1}{2}t^2)^3}\right\} \\
& - \frac{1}{2}\left\{\frac{t^2}{(1+\frac{1}{2}s^2)(1+\frac{1}{2}t^2)^2} + \frac{s^2}{(1+\frac{1}{2}s^2)^2(1+\frac{1}{2}t^2)}\right\} + \frac{5}{4}\frac{s^2t^2}{(1+\frac{1}{2}s^2)^2(1+\frac{1}{2}t^2)^2}.
\end{aligned}
$$

We have

$$
\sigma^2 = 4\left(I_1 - I_2^2 - I_3I_4 - I_3I_2 + \frac{5}{4}I_3^2\right),
$$

where

$$
\begin{aligned}
I_1 &= I_{1,1} - 2I_{1,2} + I_{1,3}, \\
I_2 &= I_{2,1} - I_{2,2}, \\
I_3 &= I_{3,1} - I_{3,2}, \\
I_4 &= I_{4,1} - I_{4,2}.
\end{aligned}
$$

Here, writing $\int$ for $\int_{-\infty}^{\infty}$,

$$
\begin{aligned}
I_{1,1} &= \iint \frac{1}{1+\frac{1}{2}(s-t)^2}\frac{1}{1+\frac{1}{2}s^2}\frac{1}{1+\frac{1}{2}t^2}\varphi(s)\varphi(t)\,\mathrm{d}s\mathrm{d}t, \\
I_{1,2} &= \iint \frac{1}{1+\frac{1}{2}(s-t)^2}\frac{1}{1+\frac{1}{2}s^2}\,\mathrm{e}^{-t^2/2}\,\varphi(s)\varphi(t)\,\mathrm{d}s\mathrm{d}t, \\
I_{1,3} &= \iint \frac{1}{1+\frac{1}{2}(s-t)^2}\mathrm{e}^{-(s^2+t^2)/2}\varphi(s)\varphi(t)\,\mathrm{d}s\mathrm{d}t,
\end{aligned}
$$

$$
\begin{aligned}
I_{2,1} &= \int \left(\frac{1}{1+\frac{1}{2}s^2}\right)^2 \varphi(s)\,\mathrm{d}s = 1 - \sqrt{\pi}\mathrm{e}\left(1 - \Phi\left(\sqrt{2}\right)\right), \\
I_{2,2} &= \int \frac{1}{1+\frac{1}{2}s^2}\,\mathrm{e}^{-s^2/2}\,\varphi(s)\,\mathrm{d}s = 2\sqrt{\pi}\mathrm{e}^2\left(1-\Phi(2)\right), \\
I_{3,1} &= \int \frac{s^2}{(1+\frac{1}{2}s^2)^3}\varphi(s)\,\mathrm{d}s = \frac{1}{2}\left(3 - 7\sqrt{\pi}\mathrm{e}\left(1-\Phi\left(\sqrt{2}\right)\right)\right), \\
I_{3,2} &= \int \frac{s^2}{(1+\frac{1}{2}s^2)^2}\,\mathrm{e}^{-s^2/2}\,\varphi(s)\,\mathrm{d}s = 10\sqrt{\pi}\mathrm{e}^2\left(1-\Phi(2)\right) - 2\sqrt{2}, \\
I_{4,1} &= \int \frac{\frac{3}{2}s^2-1}{(1+\frac{1}{2}s^2)^4}\varphi(s)\,\mathrm{d}s = \frac{25}{12}\sqrt{\pi}\mathrm{e}\left(1-\Phi\left(\sqrt{2}\right)\right) - \frac{13}{12}, \\
I_{4,2} &= \int \frac{\frac{3}{2}s^2-1}{(1+\frac{1}{2}s^2)^3}\,\mathrm{e}^{-s^2/2}\,\varphi(s)\,\mathrm{d}s = \frac{8}{\sqrt{2}} - 20\sqrt{\pi}\mathrm{e}^2\left(1-\Phi(2)\right).
\end{aligned}
$$

The values of last six integrals are obtained by equating corresponding partial derivatives of the expressions on the left and the right hand side of (4.2) with respect to x and y. Moreover,

$$
I_{1,1} = \frac{1}{\sqrt{2}}\int \ell^2(w)\mathrm{e}^{-\sqrt{2}|w|}\,\mathrm{d}w,
$$

where

$$\begin{aligned}
\ell(w) &= \int \cos(sw)\frac{1}{1+\frac{1}{2}s^2}\varphi(s)\,\mathrm{d}s \\
&= \int \frac{1}{\sqrt{2}}\left(\int \cos(sw)\cos(sv)\mathrm{e}^{-\sqrt{2}|v|}\,\mathrm{d}v\right)\varphi(s)\,\mathrm{d}s \\
&= \frac{1}{\sqrt{2}}\int\left(\int \cos(sw)\cos(sv)\varphi(s)\,\mathrm{d}s\right)\mathrm{e}^{-\sqrt{2}|v|}\,\mathrm{d}v \\
&= \frac{1}{\sqrt{2}}\int \mathrm{e}^{-(w-v)^2/2}\,\mathrm{e}^{-\sqrt{2}|v|}\,\mathrm{d}v \\
&= \sqrt{\pi}\mathrm{e}((1-\Phi(-w+\sqrt{2}))\exp(-\sqrt{2}w)+\Phi(-w-\sqrt{2})\exp(\sqrt{2}w)),\ \ w\in\mathbb{R}.
\end{aligned}$$

By numerical integration we obtain $I_{1,1} = 0.41736219895$. Moreover,

$$I_{1,2} = \frac{1}{\sqrt{2}}\int\left(\int \frac{\cos(sw)}{1+\frac{1}{2}s^2}\varphi(s)\,\mathrm{d}s\right)\left(\int \cos(tw)\,\mathrm{e}^{-t^2/2}\,\varphi(t)\,\mathrm{d}t\right)\mathrm{e}^{-\sqrt{2}|w|}\,\mathrm{d}w.$$

From $\int \cos(tw)\mathrm{e}^{-t^2/2}\varphi(t)\,\mathrm{d}t = \mathrm{e}^{-w^2/4}/\sqrt{2}$, $w\in\mathbb{R}$, we deduce

$$I_{1,2} = \frac{1}{2}\int \ell(w)\mathrm{e}^{-w^2/4}\,\mathrm{e}^{-\sqrt{2}|w|}\,\mathrm{d}w.$$

By numerical integration we get $I_{1,2} = 0.39748434972$. Finally,

$$I_{1,3} = \frac{1}{\sqrt{2}}\int \frac{1}{2}\mathrm{e}^{-w^2/2}\mathrm{e}^{-\sqrt{2}|w|}\,\mathrm{d}w \ = \ \frac{1}{\sqrt{2}}\int_0^\infty \mathrm{e}^{-w^2/2-\sqrt{2}w}\,\mathrm{d}w = \sqrt{\pi}\mathrm{e}\left(1-\Phi\left(\sqrt{2}\right)\right).$$

Putting pieces together we obtain $\sigma^2 = 0.00231562626$. Based on a sample of size $n = 1000$ from simulated realizations of $X_1,\ldots,X_n$ where $X_1 \sim L(0,1/\sqrt{2})$, we observed the value 0.002848655.

Table 3 shows the empirical coverage probabilities of the confidence interval (3.1) for Δ, each based on 10 000 replications, for several values of ε and the sample sizes $n = 20$, $n = 50$, $n = 100$ and $n = 200$. The nominal level is $1-\alpha = 0.9$. Notice that the values for $\varepsilon = 1$ are also given in Table 1.

	ε			
	0.25	0.5	0.75	1
$n = 20$	0.96	0.97	0.94	0.89
$n = 50$	0.99	0.96	0.91	0.87
$n = 100$	0.99	0.92	0.89	0.88
$n = 200$	0.97	0.91	0.89	0.89

Table 3 Empirical coverage probabilities of I_n for Δ (normal distribution contaminated by Laplace, nominal level 0.9, 10 000 replications)

In contrast with Table 1, the coverage probabilities seem to be much higher than the nominal value, and they seem to increase as ε decreases. Table 4 shows the empirical power of the test for normality based on $T_{n,1}$. The nominal level is 0.9. As was to be expected, the power increases as ε increases.

ε	0.25	0.5	0.75	1.0
$n = 20$	0.14	0.20	0.27	0.35
$n = 50$	0.17	0.29	0.46	0.62
$n = 100$	0.22	0.44	0.69	0.88
$n = 200$	0.28	0.65	0.91	0.99

Table 4 Power Study for the normal distributions contaminated by the Laplace distribution

Example 2. Let $X, X_1, X_2, \ldots$ be i.i.d. positive random variables. Motivated by a characteristic differential equation for the Laplace transform of an exponential distribution, Baringhaus and Henze (1991) considered the test statistic

$$T_{n,a} = n \int_0^\infty Z_n^2(t) \mathrm{e}^{-at} \,\mathrm{d}t \tag{4.3}$$

for testing the hypothesis H_0 that the distribution of X is some exponential distribution. Here,

$$Z_n(t) = \frac{1}{n} \sum_{j=1}^n \mathrm{e}^{-tY_j} \left(1 - (1+t)Y_j\right),$$

$Y_j = X_j / \overline{X}_n$, $j = 1, \ldots, n$, and $a > 0$ is a fixed positive number. Notice that $T_{n,a}$ is scale invariant. Straightforward manipulations of integrals gives the computationally simple form

$$T_{n,a} = \frac{1}{n} \sum_{i,j=1}^n \left[\frac{(1-Y_i)(1-Y_j)}{Y_i + Y_j + a} + \frac{Y_i(Y_j - 1) + Y_j(Y_i - 1)}{(Y_i + Y_j + a)^2} + \frac{2Y_iY_j}{(Y_i + Y_j + a)^3} \right].$$

Under a fixed alternative distribution with finite expectation (which, because of scale invariance, is taken to be 1), we have

$$\frac{T_{n,a}}{n} \xrightarrow{\mathbb{P}} \Delta := \int_0^\infty z^2(t) \mathrm{e}^{-at} \,\mathrm{d}t, \tag{4.4}$$

where

$$z(t) := \mathbb{E}\left[\mathrm{e}^{-tX} \left(1 - (1+t)X\right)\right], \ t \geq 0. \tag{4.5}$$

Notice that Δ depends on a, although this dependence has not been made explicit.

To state the limiting normal distribution of $T_{n,a}$ under a fixed alternative to H_0, let $W_n(t) := \sqrt{n}(Z_n(t) - z(t))$, $t \geq 0$. The process $W_n = (W_n(t), t \geq 0)$ can be regarded as a random element of the Hilbert space $\mathcal{H} := L^2([0,\infty), \mathcal{B}^1 \cap [0,\infty), \mathrm{e}^{-at}\mathrm{d}t)$.

Theorem 2. *If X has mean 1 and positive finite variance τ^2, then*

$$W_n \xrightarrow{\mathcal{D}} W \text{ in } \mathcal{H}.$$

Here, W is a centered Gaussian process on $\mathcal{H}$ having covariance kernel

$$\begin{aligned} K(s,t) &= L(s+t) + (2+s+t)L'(s+t) + (1+s)(1+t)L''(s+t) - z(s)z(t) \\ &\quad + \left(L'(s) + (1+s)L''(s) + z(s)\right) w(t) \\ &\quad + \left(L'(t) + (1+t)L''(t) + z(t)\right) w(s) \\ &\quad + \tau^2 g(s)g(t), \ \ s,t \geq 0, \end{aligned}$$

where $L(t) = \mathbb{E}[\mathrm{e}^{-tX}]$ is the Laplace transform of X and

$$g(t) = \mathbb{E}\left[X\mathrm{e}^{-tX}\left((1+t)(tX-1) - t\right)\right] = (2t+1)L'(t) + t(1+t)L''(t), \ t \geq 0. \tag{4.6}$$

PROOF. The proof of Theorem 2 is given in Section 7.

From Theorem 1 we thus have

$$\sqrt{n}\left(\frac{T_{n,a}}{n} - \Delta\right) \xrightarrow{\mathcal{D}} \mathrm{N}(0, \sigma^2),$$

where

$$\sigma^2 = 4\int_0^\infty \int_0^\infty K(s,t)z(s)z(t)\mathrm{e}^{-a(s+t)}\,\mathrm{d}s\mathrm{d}t. \tag{4.7}$$

Notice that, like Δ, also σ^2 depends on a.

To estimate σ^2, we replace L, L', L'', g, z and τ^2 by their respective empirical counterparts

$$\begin{aligned}
L_n(t) &= \frac{1}{n}\sum_{i=1}^n \mathrm{e}^{-tY_i}, \quad L_n'(t) = -\frac{1}{n}\sum_{i=1}^n Y_i\mathrm{e}^{-tY_i},\ t \geq 0,\\
L_n''(t) &= \frac{1}{n}\sum_{i=1}^n Y_i^2\mathrm{e}^{-tY_i}, \quad z_n(t) = \frac{1}{n}\sum_{i=1}^n (1 - Y_i - tY_i)\mathrm{e}^{-tY_i},\ t \geq 0,\\
g_n(t) &= -\frac{1}{n}\sum_{i=1}^n Y_i\mathrm{e}^{-tY_i}(1 - t(Y_i - 2) - t^2Y_i), \quad \widehat{\tau}_n^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - 1)^2,
\end{aligned}$$

based on the scaled random variables $Y_1, \ldots, Y_n$. Denoting by K_n the resulting estimator of K, the estimator $\widehat{\sigma}_n^2$ of σ^2 is

$$\widehat{\sigma}_n^2 = 4\int_0^\infty \int_0^\infty K_n(s,t)z_n(s)z_n(t)\mathrm{e}^{-a(s+t)}\,\mathrm{d}s\mathrm{d}t.$$

Putting

$$\begin{aligned}
S_{1n} &= \iint L_n(s+t)z_n(s)z_n(t)\mathrm{e}^{-a(s+t)}\,\mathrm{d}s\mathrm{d}t,\\
S_{2n} &= \iint (2+s+t)L_n'(s+t)z_n(s)z_n(t)\mathrm{e}^{-a(s+t)}\,\mathrm{d}s\mathrm{d}t,\\
S_{3n} &= \iint (1+s)(1+t)L_n''(s+t)z_n(s)z_n(t)\mathrm{e}^{-a(s+t)}\,\mathrm{d}s\mathrm{d}t,\\
S_{4n} &= \int z_n^2(t)\mathrm{e}^{-at}\,\mathrm{d}t,\\
S_{5n} &= \int \left(L_n'(t) + (1+t)L_n''(t)\right) z_n(t)\mathrm{e}^{-at}\,\mathrm{d}t,\\
S_{6n} &= \int g_n(t)z_n(t)\mathrm{e}^{-at}\,\mathrm{d}t,
\end{aligned}$$

where $\int$ is shorthand for $\int_0^\infty$, we have

$$\widehat{\sigma}_n^2 \;=\; 4\left\{S_{1n} + S_{2n} + S_{3n} - S_{4n}^2 + 2S_{5n}S_{6n} + 2S_{4n}S_{6n} + \widehat{\tau}_n^2 S_{6n}^2\right\}.$$

Defining

$$T_{ij} \;=\; \frac{1}{Y_i + Y_j + a}, \quad 1 \leq i, j \leq n,$$

tedious, but straightforward calculations yield

$$
\begin{aligned}
S_{1n} &= \frac{1}{n^3} \sum_{i,j,k} T_{ij} T_{ik} (1 - Y_j - Y_j T_{ij})(1 - Y_k - Y_k T_{ik}), \\
S_{2n} &= -\frac{1}{n^3} \sum_{i,j,k} Y_i T_{ij} T_{ik} \Big(2(1 - Y_j - Y_j T_{ij})(1 - Y_k - Y_k T_{ik}) \\
&\qquad + T_{ij}(1 - Y_j - 2Y_j T_{ij})(1 - Y_k - Y_k T_{ik}) + T_{ik}(1 - Y_j - Y_j T_{ij})(1 - Y_k - 2Y_k T_{ik}) \Big), \\
S_{3n} &= \frac{1}{n^3} \sum_{i,j,k} Y_i^2 T_{ij} T_{ik} \Big(1 - Y_j + (1 - 2Y_j) T_{ij} - 2Y_j T_{ij}^2\Big) \Big(1 - Y_k + (1 - 2Y_k) T_{ik} - 2Y_k T_{ik}^2\Big), \\
S_{4n} &= \frac{1}{n^2} \sum_{i,j} T_{ij} \Big((1 - Y_i)(1 - Y_j) - (Y_i + Y_j - 2Y_i Y_j) T_{ij} + 2Y_i Y_j T_{ij}^2 \Big), \\
S_{5n} &= -\frac{1}{n^2} \sum_{i,j} Y_i T_{ij} \Big((1 - Y_i)(1 - Y_j) - (Y_i + Y_j - 2Y_i Y_j) T_{ij} + 2Y_i Y_j T_{ij}^2 \Big), \\
S_{6n} &= \frac{1}{n^2} \sum_{i,j} Y_i T_{ij} \Big(Y_j - 1 + (3Y_j + Y_i - Y_i Y_j - 2) T_{ij} + 2(Y_i + 2Y_j - 2Y_i Y_j) T_{ij}^2 - 6Y_i Y_j T_{ij}^3 \Big).
\end{aligned}
$$

In the sums above, each of the indices runs from 1 to n.

As an alternative to the exponential distribution, we consider the Gamma distribution $\mathrm{G}(\beta, \beta)$ with shape parameter β and scale parameter β, which has the density $f_\beta(x) = \beta^\beta x^{\beta-1} \exp(-\beta x)/\Gamma(\beta)$, $x > 0$, and $f_\beta(x) = 0$, otherwise. The expectation of this distribution is 1, its variance is $\tau^2 = 1/\beta$.

In what follows, the parameter a figuring in (4.3) is chosen to be 1. To calculate Δ defined in (4.4) and σ^2 given in (4.7) in case of the distribution $\mathrm{G}(\beta, \beta)$, notice that, for $t \geq 0$,

$$
L(t) = \left(\frac{\beta}{\beta + t} \right)^\beta, \quad L'(t) = -\left(\frac{\beta}{\beta + t} \right)^{\beta+1}, \quad L''(t) = \frac{\beta + 1}{\beta} \left(\frac{\beta}{\beta + t} \right)^{\beta+2}.
$$

Moreover, $z(t)$ figuring in (4.5) and $g(t)$ defined in (4.6) are given by

$$
z(t) = \frac{(1 - \beta)t}{\beta} \left(\frac{\beta}{\beta + t} \right)^{\beta+1}, \quad g(t) = \frac{t^2(1 - \beta) + \beta(t + 1)}{\beta + t} \left(\frac{\beta}{\beta + t} \right)^{\beta+1},
$$

respectively. Straightforward calculations now yield

$$
\sigma^2 = 4 \left(I_1 + I_2 - \Delta^2 + 2(I_3 + \Delta) I_4 + \frac{I_4^2}{\beta} \right),
$$

where

$$
\begin{aligned}
I_1 &= \int_0^\infty \int_0^\infty \left(L(s+t) + (2+s+t)L'(s+t)\right) z(s)z(t)\mathrm{e}^{-a(s+t)} \,\mathrm{d}s\mathrm{d}t \\
&= \int_0^\infty \left(\int_0^u z(u-t)z(t)\,\mathrm{d}t \right) \left(L(u) + (2+u)L'(u)\right) \mathrm{e}^{-au}\,\mathrm{d}u, \\
I_2 &= \int_0^\infty \int_0^\infty (1+s)(1+t)L''(s+t)z(s)z(t)\mathrm{e}^{-a(s+t)} \,\mathrm{d}s\mathrm{d}t \\
&= \int_0^\infty \left(\int_0^u (1+(u-t))(1+t)z(u-t)z(t)\,\mathrm{d}t \right) L''(u)\mathrm{e}^{-au}\,\mathrm{d}u, \\
I_3 &= \int_0^\infty (L'(t) + (1+t)L''(t))z(t)\mathrm{e}^{-at}\,\mathrm{d}t, \\
I_4 &= \int_0^\infty g(t)z(t)\mathrm{e}^{-at}\,\mathrm{d}t.
\end{aligned}
$$

For $\beta \in \{0.1, 0.25, 0.5, 1.5, 2, 3\}$, the integrals I_1, I_2, I_3, I_4 and $\varDelta = \int_0^\infty z^2(t)\mathrm{e}^{-t}\,\mathrm{d}t$ have been calculated by numerical integration. The corresponding values of $\varDelta$ and σ^2 are shown in Table 5. In addition, Table 5 exhibits realizations of $\frac{1}{n}T_n$ and $\widehat{\sigma}_n^2$, based on a sample of size $n = 1000$ from simulated observations of $X_1, \ldots, X_n$.

β	0.1	0.25	0.5	1.5	2	3
$\varDelta$	0.34770	0.12380	0.02640	0.00547	0.01354	0.02659
$\frac{1}{n}T_n$	0.35816	0.11310	0.02897	0.00928	0.01474	0.02693
σ^2	0.30788	0.09987	0.01516	0.00107	0.00179	0.00192
$\widehat{\sigma}_n^2$	0.32819	0.09182	0.01640	0.00154	0.00176	0.00187

Table 5 Expectation $\varDelta$ and variance σ^2 of the limit distribution in case where the alternative distribution is $\mathrm{G}(\beta, \beta)$, and simulated observations of the estimators $\frac{1}{n}T_n$ and $\widehat{\sigma}_n^2$; $n = 1000$

Table 6 shows the empirical coverage probabilities of the confidence interval (3.1) for $\varDelta$, each based on 10 000 replications, for several values of β and the sample sizes $n = 20$, $n = 50$ and $n = 100$. The nominal confidence level is $1 - \alpha = 0.9$. Likewise, Table 7 displays the empirical power, based on 10 000 replications, of the goodness-of-fit test for exponentiality that rejects the null hypothesis for large values of $T_{n.1}$. The nominal level is 0.1. Critical values have been obtained by simulations based on 100 000 replications. Moreover, Table 7 shows the approximation (3.2) (denoted by App), to the power. The quality of this approximation is quite promising. Notice that, in Tables 6 and 7, there are no entries for the sample size $n = 200$, since the calucations are extremely time consuming due to the fact each of S_{1n}, S_{2n} and S_{3n} is a triple sum.

5 Two-sample weighted L^2-statistics

In the following we show that the results obtained so far are by no means confined to one-sample weighted L^2-statistics, but immediately carry over to two-sample weighted L^2-statistics of the type

$$
T_{m,n} = \frac{mn}{m+n} \int_M Z_{m,n}^2(t)\,\mu(\mathrm{d}t), \tag{5.1}
$$

	β					
	0.1	0.25	0.5	1.5	2	3
$n = 20$	0.83	0.82	0.78	0.81	0.81	0.81
$n = 50$	0.87	0.86	0.84	0.84	0.86	0.86
$n = 100$	0.88	.88	0.87	0.86	0.88	0.88

Table 6 Empirical coverage probabilities of I_n for Δ (distribution $\mathrm{G}(\beta,\beta)$, nominal level 0.9, 10 000 replications)

	$\beta = 0.1$		$\beta = 0.25$		$\beta = 0.5$		$\beta = 1.5$		$\beta = 2$		$\beta = 3$	
	MC	App	MC	App	MC	App	MC	App	MC	App	MC	App
$n = 20$	1.0	1.0	0.99	0.91	0.70	0.71	0.35	0.37	0.70	0.61	0.97	0.95
$n = 50$	1.0	1.0	1.0	0.99	0.96	0.90	0.65	0.59	0.98	0.94	1.0	1.0
$n = 100$	1.0	1.0	1.0	1.0	1.0	0.98	0.90	0.79	1.0	1.0	1.0	1.0

Table 7 Empirical power and approximation (3.2) of the test for exponentiality against selected alternatives from the $\mathrm{G}(\beta,\beta)$-family

where M and μ retain their meanings from Section 1, and $Z_{m,n}(t)$ is based on two samples $X_1,\ldots,X_m,Y_1,\ldots,Y_n$, i.e., $Z_{m,n}(t) = Z_{m,n}(X_1,\ldots,X_m,Y_1,\ldots,Y_n,t)$.

As an example, suppose that the X_i, Y_j are independent d-dimensional random vectors, where $X_1,\ldots,X_m$ are i.i.d. and $Y_1,\ldots,Y_n$ are i.i.d. To test the semiparametric 'location shift' hypothesis H_0 that Y_1 has the same distribution as $X_1 + \mu$ for some unspecified $\mu \in \mathbb{R}^d$, Henze et al. (2005) studied the weighted L^2-statistic (5.1) with $M = \mathbb{R}^d$ and

$$Z_{m,n}(t) = U^{(1)}_{m,n}(t) - U^{(2)}_{m,n}(t),$$

where

$$\begin{aligned}
U^{(1)}_{m,n}(t) &= \frac{1}{m}\sum_{j=1}^{m}\Big\{\cos(t^\top(X_j+\widehat{\mu})) + \sin(t^\top(X_j+\widehat{\mu})) - \Psi(t)\Big\},\\
U^{(2)}_{m,n}(t) &= \frac{1}{n}\sum_{j=1}^{n}\Big\{\cos(t^\top Y_j) + \sin(t^\top Y_j) - \Psi(t)\Big\},\\
\Psi(t) &= \mathbb{E}\left[\cos(t^\top Y_1) + \sin(t^\top Y_1)\right],
\end{aligned}$$

and $\widehat{\mu} = \widehat{\mu}_{m,n}(X_1,\ldots,X_m,Y_1,\ldots,Y_n)$ is some location equivariant regular estimator of μ, e.g., $\widehat{\mu} = \overline{Y}_n - \overline{X}_m$. The measure μ figuring in (5.1) was chosen to be the spherically symmetric d-variate normal distribution $\mathrm{N}_d(0,\beta^2\mathrm{I}_d)$, where $\beta > 0$ is a parameter and I_d is the unit matrix of order d.

Two-sample weighted L^2-statistics have also been employed by Gupta et al. (2004) in the context of testing for affine equivalence of elliptically symmetric distributions and by Baringhaus and Franz (2010), who considered a general class of consistent multivariate rigid motion invariant homogeneity tests. Meintanis (2005) studied permutation tests for homogeneity based on the empirical characteristic function. Moreover, Baringhaus and Kolbe (2015) dealt with two-sample weighted L^2-statistics based on empirical Hankel transforms. All these papers derive the limit distribution under the respective null hypothesis H_0 by showing that, under H_0,

$$\sqrt{\frac{mn}{m+n}}Z_{m,n} \xrightarrow{\mathcal{D}} Z \text{ as } m,n\to\infty$$

in $\mathcal{H} = L^2(M, M \cap \mathcal{B}^d, \mu)$, where Z is some centered Gaussian process on $\mathcal{H}$. In dealing with asymptotics for two-sample problems, sometimes the additional condition

$$\text{there is some } p \in (0,1) \text{ such that } \frac{m}{m+n} \to p \tag{5.2}$$

is imposed.

The following result shows that, under general conditions, two-sample weighted L^2-statistics have a limiting normal distribution.

Theorem 3. *Let $(X_j)_{j\geq 1}$ and $(Y_j)_{j\geq 1}$ be sequences of d-dimensional random vectors on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that, as $m, n \to \infty$,*

$$\frac{m+n}{mn} T_{m,n} \xrightarrow{\mathbb{P}} \Delta := \int_M z^2(t)\, \mu(\mathrm{d}t) > 0 \tag{5.3}$$

for some function $z \in \mathcal{H} := L^2(M, M \cap \mathcal{B}^d, \mu)$. Suppose further that

$$\sqrt{\frac{mn}{m+n}}\, (Z_{m,n} - z) \xrightarrow{\mathcal{D}} W \quad \text{as } m, n \to \infty \tag{5.4}$$

in $\mathcal{H}$, where W is a centered Gaussian process on $\mathcal{H}$ having covariance kernel K. Then

$$\sqrt{\frac{mn}{m+n}} \left(\frac{T_{m,n}}{\frac{mn}{m+n}} - \Delta \right) \xrightarrow{\mathcal{D}} \mathrm{N}(0, \sigma^2),$$

where σ^2 is given in (2.2).

Proof. Let $a_{m,n} = \sqrt{mn/(m+n)}$. Proceeding as in the proof of Theorem 1, we have

$$a_{m,n} \left(\frac{T_{m,n}}{a_{m,n}^2} - \Delta \right) = 2\langle a_{m,n}(Z_{m,n} - z), z \rangle + \frac{1}{a_{m,n}} \| a_{m,n}(Z_{m,n} - z)\|_{L^2}^2.$$

The assertion now follows from (5.4), the continuous mapping theorem and the fact that $a_{m,n} \to \infty$ as $m, n \to \infty$. □

Remark 1. Notice that Theorem 3 is fairly general, since there is no assumption regarding independence or identical distributions among $X_1, X_2, \ldots$ or $Y_1, Y_2, \ldots$. Of course, (5.3) postulates the validity of some sort of large numbers for $T_{m,n}$. The additional condition (5.2), although sometimes imposed, can at second sight often be dispensed with, and it clearly does not restrict the scope of possible applications.

It goes without saying that, in a two-sample setting with independent vectors $X_1, \ldots, X_m, Y_1, \ldots, Y_n$, where the $X_1, \ldots, X_m$ are i.i.d with unknown distribution function F and $Y_1, \ldots, Y_n$ are i.i.d. with unknown distribution function G, Δ figuring in (5.3) will depend on F and G. Then, in the same way as was done in Section 3, Theorem 3 can be used to construct an asymptotic confidence interval for $\Delta(F, G)$ or to establish an inverse goodness-of-fit test of $H_{\Delta_0} : \Delta(F, G) \geq \Delta_0$ versus $K_{\Delta_0} : \Delta(F, G) < \Delta_0$, where Δ_0 is a given positive number, provided that we have a consistent estimator $\widehat{\sigma}_{m,n}^2$ of σ^2.

6 Concluding remarks

We have shown that, under general conditions, weighted one- and two-sample L^2-statistics are asymptotically normally distributed. It is easy to see that the approach also encompasses the multisample case, which was considered by Hušková and Meintanis (2008) in connection with tests for the multivariate k-sample problem based on the ECF. As the examples show, one has to work out the covariance structure of the limiting Gaussian process and the resulting variance of the limiting normal distribution. If the latter can be estimated consistently, this opens the ground for asymptotic confidence intervals for the distance of an underlying distribution with respect to a hypothesized family, and for asymptotic tests that are able to validate neighborhoods of hypothesized models. We hope that this paper will stimulate interest in this important problem, which is in the spirit of bioequivalence testing.

7 Proof of Theorem 2

To derive the limit distribution of W_n, put

$$\widetilde{W}_n(t) := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \Big\{ \mathrm{e}^{-tX_j} - \mathbb{E}\left[\mathrm{e}^{-tX}\right] - (1+t)\left(X_j \mathrm{e}^{-tX_j} - \mathbb{E}\left[X\mathrm{e}^{-tX}\right]\right) \\ - \mathbb{E}\left[X\mathrm{e}^{-tX}\left((1+t)(tX-1)-t\right)\right](X_j-1)\Big\},\ t \geq 0.$$

We first show

$$\|W_n - \widetilde{W}_n\|_{L^2} \xrightarrow{\mathbb{P}} 0. \tag{7.1}$$

To this end, define the processes $A = (A_n(t), t \geq 0)$, $\widetilde{A}_n = (\widetilde{A}_n(t), t \geq 0)$, $B_n = (B_n(t), t \geq 0)$ and $\widetilde{B}_n = (\widetilde{B}_n(t), t \geq 0)$, where

$$\begin{aligned} A_n(t) &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \Big\{ \mathrm{e}^{-tY_j} - \mathbb{E}\left[\mathrm{e}^{-tX}\right] \Big\}, \\ \widetilde{A}_n(t) &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \Big\{ \mathrm{e}^{-tX_j} - \mathbb{E}\left[\mathrm{e}^{-tX}\right] \Big\} + \sqrt{n}(\overline{X}_n - 1)\mathbb{E}\left[tX\mathrm{e}^{-tX}\right], \\ B_n(t) &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \Big\{ (1+t)\left(Y_j \mathrm{e}^{-tY_j} - \mathbb{E}\left[X\mathrm{e}^{-tX}\right]\right) \Big\}, \\ \widetilde{B}_n(t) &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \Big\{ (1+t)\left(X_j \mathrm{e}^{-tX_j} - \mathbb{E}\left[X\mathrm{e}^{-tX}\right]\right) \Big\} \\ &\qquad + \sqrt{n}(\overline{X}_n - 1)\mathbb{E}\left[X\mathrm{e}^{-tX}(1+t)(tX-1)\right]. \end{aligned}$$

Notice that

$$W_n - \widetilde{W}_n = (A_n - \widetilde{A}_n) - (B_n - \widetilde{B}_n). \tag{7.2}$$

A one-term Taylor expansion with integral remainder gives

$$A_n(t) - \widetilde{A}_n(t) = \sqrt{n}(\overline{X}_n - 1)\left(R_{1,n}(t) + R_{2,n}(t) + R_{3,n}(t)\right),\ t \geq 0, \tag{7.3}$$

where

$$R_{1,n}(t) = t\left(\frac{1}{\overline{X}_n} - 1\right)\mathbb{E}\left[X\mathrm{e}^{-tX}\right],$$
$$R_{2,n}(t) = \frac{t}{\overline{X}_n}\left(\frac{1}{n}\sum_{j=1}^n X_j\mathrm{e}^{-tX_j} - \mathbb{E}\left[X\mathrm{e}^{-tX}\right]\right),$$
$$R_{3,n}(t) = \frac{t}{\overline{X}_n}\int_0^\infty\left(\frac{1}{n}\sum_{j=1}^n X_j\left[\mathrm{e}^{-tX_j\left(1+\tau\left(\frac{1}{\overline{X}_n}-1\right)\right)} - \mathrm{e}^{-tX_j}\right]\right)\mathrm{d}\tau.$$

By the law of large numbers, $\int_0^\infty R_{i,n}^2(t)\mathrm{e}^{-at}\,\mathrm{d}t \to 0$ $\mathbb{P}$-a.s. for $i = 1, 2$. Putting

$$\widetilde{R}_{3,n}(t) = \frac{t}{\overline{X}_n}\frac{1}{n}\sum_{j=1}^n X_j\left|\mathrm{e}^{-tX_j\frac{1}{\overline{X}_n}} - \mathrm{e}^{-tX_j}\right|,\ t \geq 0,$$

the relation

$$\sup_{0\leq\tau\leq 1}\left|\mathrm{e}^{-tX_j\left(1+\tau\left(\frac{1}{\overline{X}_n}-1\right)\right)} - \mathrm{e}^{-tX_j}\right| = \left|\mathrm{e}^{-tX_j\frac{1}{\overline{X}_n}} - \mathrm{e}^{-tX_j}\right|$$

yields $0 \leq R_{3,n}(t) \leq \widetilde{R}_{3,n}(t),\ t \geq 0$. Invoking the law of large numbers again, we have $\int_0^\infty \widetilde{R}_{3,n}^2(t)\mathrm{e}^{-at}\,\mathrm{d}t \to 0$ $\mathbb{P}$-a.s. and thus $\int_0^\infty R_{3,n}^2(t)\mathrm{e}^{-at}\,\mathrm{d}t \to 0$ $\mathbb{P}$-a.s. Therefore, (7.3) and $\sqrt{n}(\overline{X}_n - 1) \xrightarrow{\mathcal{D}} \mathrm{N}(0, \sigma^2)$ give $\|A_n - \widetilde{A}_n\|_{L^2} \xrightarrow{\mathbb{P}} 0$. Likewise, $\|B_n - \widetilde{B}_n\|_{L^2} \xrightarrow{\mathbb{P}} 0$. In view of (7.2), this proves (7.1).

From the central limit theorem for random elements in the Hilbert space $\mathcal{H}$, see, e.g., Ledoux and Talagrand (2011), there is a centered Gaussian process $W = (W(t), t \geq 0)$, which can be regarded as random element of $\mathcal{H}$, such that $\widetilde{W}_n \xrightarrow{\mathcal{D}} W$ and thus, due to (7.1), also $W_n \xrightarrow{\mathcal{D}} W$.

The covariance function K of W can be expressed in terms of the Laplace transform $L(t) = \mathbb{E}\left[\mathrm{e}^{-tX}\right],\ t \geq 0$, of X. For, noting that $z(t)$ defined in (4.5) takes the form $z(t) = L(t) + (1 + t)L'(t),\ t \geq 0$, the definition of $g(t)$ given in (4.6) gives

$$\begin{aligned}K(s,t) = \mathbb{E}\Big[&\left(\mathrm{e}^{-sX} - (1+s)\mathrm{e}^{-sX} - z(s) - g(s)(X-1)\right)\\ &\left(\mathrm{e}^{-tX} - (1+t)\mathrm{e}^{-tX} - z(t) - g(t)(X-1)\right)\Big],\ s,t \geq 0.\end{aligned}$$

Due to

$$\begin{aligned}&\mathbb{E}\Big[\left(\mathrm{e}^{-sX} - (1+s)X\mathrm{e}^{-sX} - z(s)\right)\left(\mathrm{e}^{-tX} - (1+t)X\mathrm{e}^{-tX} - z(t)\right)\Big]\\ &= L(s+t) + (2+s+t)L'(s+t) + (1+s)(1+t)L''(s+t) - z(s)z(t),\ s,t \geq 0,\end{aligned}$$

and

$$\mathbb{E}\Big[\left(\mathrm{e}^{-sX} - (1+s)X\mathrm{e}^{-sX} - z(s)\right)g(t)(X-1)\Big] = \left(L'(s) + (1+s)L''(s) + z(s)\right)g(t),$$

$s, t \geq 0$, we obtain the representation of $K(s,t)$ given in Theorem 2. □

Acknowledgement: The authors would like to thank Ya. Yu. Nikitin for drawing our attention to the paper of Chapman (1958) and B. Klar for pointing out the reference Naito (1997).

References

Alba-Fernández, V., and Jiménez-Gamero, M.D. (2015). Testing for a class of bivariate exponential distributions. *Int. J. of Comp. Mathem.*, 92, 1733–1754.

Baringhaus, L., and Franz, C. (2010). Rigid motion invariant two-sample tests. *Statist. Sinica*, 20, 1333–1361.

Baringhaus, L., and Henze, N. (1988). An invariant consistent test for multivariate normality. *Metrika*, 13, 269–274.

Baringhaus, L., and Henze, N. (1991). A consistent class of tests for exponentiality based on the empirical Laplace transform. *Ann. Inst. Statist. Mathem.*, 43, 551–564.

Baringhaus, L., and Henze, N. (1992). A goodness-of-fit test for the Poisson distribution based on the empirical generating function. *Statist. & Probab. Lett.*, 13, 269–274.

Baringhaus, L., Gürtler, N., and Henze, N. (2000). Weighted integral test statistics and components of smooth tests of fit. *Aust. N.Z. J. Statist.*, 42, 179–192.

Baringhaus, L., and Kolbe, D. (2015). Two-sample tests based on empirical Hankel transforms. *Statist. Papers*, 56, 597–617.

Bücher, A., and Dette, H. (2010). Some comments on goodness-of-fit tests for the parametric form of the copula based on L^2-distances. *J. Multiv. Anal.*, 101, 749–763.

Chapman, D.G. (1958). A comparative study of several one-sided goodness-of-fit tests. *Ann. Mathem. Statist.*, 29, 655–674.

Czado, C., Freitag, G., and Munk, A. (2007). A nonparametric test for similarity of marginals – with applications to the assessment of bioequivalence. *J. Statist. Plann. Infer.*, 137, 697–711.

Dette, H. and Munk, A. (2003). Some methodological aspects of validation of models in nonparametric regression. *Statist. Neerlandica*, 57, 207–244.

De Wet, T., and Randles, R.H. (1987). On the effect of substituting parameter estimators in limiting χ^2-, U- and V-statistics. *Ann. Statist.*, 15, 398–412.

Ebner, B., Henze, N., and Meintanis, S. (2012). Goodness-of-fit tests for the gamma distribution based on the empirical Laplace transform. *Commun. Statist. - Th. Meth.*, 41, 1543–1556.

Epps, T.W., and Pulley, L.B. (1983). A test for normality based on the empirical characteristic function. *Biometrika*, 70, 723–726.

Fermanian, J. D. (2009). Goodness-of-fit tests for copulas. *J. Multiv. Anal.*, 95, 119–152.

Fragiadakis, K., Karlis, D., and Meintanis, S. (2009). Tests of fit for normal inverse Gaussian distributions. *Statist. Methodology*, 6, 553–564.

Genest, Ch., Kojadinovic, I., Nešlehová, J., and Yan, J. (2011). A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli*, 17, 253–275.

Gürtler, N. (2000). Asymptotic results on the BHEP tests for multivariate normality with fixed and variable smoothing parameter (in German). Doctoral dissertation, University of Karlsruhe.

Gürtler, N., and Henze, N. (2000a). Goodness-of-fit tests for the Cauchy distribution based on the empirical characteristic function. *Ann. Inst. Statist. Mathem.*, 52, 267–286.

Gürtler, N., and Henze, N. (2000b). Recent and classical goodness-of-fit tests for the Poisson distribution. *J. Statist. Plann. Infer.*, 90, 207–225.

Gupta, A.K., Henze, N., Klar, B. (2004). Testing for affine equivalence of elliptically symmetric distributions. *J. Multiv. Anal.*, 88, 222-242.

Henze, N. (1990). An approximation to the limit null distribution of the Epps-Pulley test statistic for normality. *Metrika*, 37, 7–18.

Henze, N. (1993). A flexible class of omnibus tests for exponentiality. *Commun. Statist. - Th. Meth.*, 22, 115–133.

Henze, N. (2002). Invariant tests for multivariate normality: A critical review. *Statist. Papers*, 43, 467–506.

Henze, N., and Klar, B. (2002). Goodness-of-fit tests for the Inverse Gaussian distribution based on the empirical Laplace transform. *Ann. Inst. Statist. Mathem.*, 54, 425–444.

Henze, N., Klar, B., Meintanis, S. (2003). Invariant tests for symmetriy about an unspecified point based on the empirical characteristic function. *J. Multiv. Anal.*, 87, 275–297.

Henze, N., Klar, B., Zhu, L. (2005). Checking the adequacy of the multivariate semiparametric location shift model. *J. Multiv. Anal.*, 93, 238–256.

Henze, N., and Meintanis, S. (2002a). Tests of fit for exponentiality based on the empirical Laplace transform. *Statistics*, 36, 147–161.

Henze, N., and Meintanis, S. (2002b). Goodness-of-fit tests based on a new characterization of the exponential distribution. *Commun. Statist. - Th. Meth.*, 31, 1479–1497.

Henze, N., and Meintanis, S. (2005). Recent and classical tests for exponentiality: A partial review with comparisons. *Metrika*, 61, 29–45.

Henze, N., and Meintanis, S. (2010). A characterization and a class of omnibus tests for the exponential distribution based on the empirical characteristic function. *J. Mathem. Sci.*, 167, 588–595.

Henze, N., and Wagner, T. (1997). A new approach to the BHEP tests for multivariate normality. *J. Multiv. Anal.*, 62, 1–23.

Henze, N., and Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Commun. Statist. - Th. Meth.*, 19, 3595–3617.

Hušková, M., and Meintanis, S. (2008). Tests for the multivariate k-sample problem based on the empirical characteristic function. *J. Nonparam. Statist.*, 20, 263–277.

Iliopoulos, G., and Meintanis, S. (2003). Tests of fit for the Rayleigh distribution based on the empirical Laplace transform. *Ann. Inst. Statist. Mathem.*, 55, 137–151.

Ledoux, M. and Talagrand, M. (2011). *Probability in Banach spaces. Isoperimetry and processes.* Springer-Verlag, Berlin.

Leucht, A. (2012). Characteristic function-based tests under weak dependence. *J. Multiv. Anal.*, 108, 67–89.

Magnus, W., Oberhettinger, F., and Soni, R. (1966). *Formulas and Theorems for the Special Functions of Mathematical Physics.* Springer, Berlin.

Meintanis, S. (2004a). A class of omnibus tests for the Laplace distribution based on the empirical characteristic function. *Commun. Statist. - Th. Meth.*, 33, 925–948.

Meintanis, S. (2004b). Goodness-of-fit tests for the logistic distribution based on empirical transforms. *Sankhyā*, 66, 306–326.

Meintanis, S. (2005). Permutation tests for homogeneity based on the empirical characteristic function. *J. Nonparam. Stat.*, 17, 583–592.

Meintanis, S. (2007). Test of fit for Marshall-Olkin distributions with applications. *J. Statist. Plann. Infer.*, 137, 3954–3963.

Meintanis, S. (2008a). Tests for generalized exponential laws based on the empirical Mellin transform. *J. Statist. Comp. Simul.*, 78, 1077–1085.

Meintanis, S. (2008b). A powerful method of assessing the fit of the lognormal distribution. *Commun. Statist. - Th. Meth.*, 37, 1948–1958.

Meintanis, S. (2010). Testing skew normality via the moment generating function. *Mathem. Meth. Statist.*, 19, 64–72.

Meintanis, S., and Tsionas, E. (2010). Testing for the generalized normal-Laplace distribution with applications. *Comp. Statist. Data Anal.*, 54, 3174–3180.

Naito, K. (1997). On the asymptotic normality of the L_2-distance class of statistics with estimated parameters. *J. Nonparam. Statist.*, 8, 199–214.

Ngatchou-Wandji, J. (2009). Testing for symmetry in multivariate distributions. *Statist. Methodol.*, 6, 230–250.

Novoa-Muñoz, F.N., and Jiménez-Gamero, M.D. (2014). Testing for the bivariate Poisson distribution. *Metrika*, 77, 771–793.

Rueda, R., Pérez-Abreu, V., and O' Reilly, F. (1991). Goodness of fit test for the Poisson distribution based on the probability generating function. *Commun. Statist. - Th. Meth.*, 20, 3093–3110.

Wellek, S. (2010). *Testing statistical hypotheses of equivalence and noninferiority.* CRC Press, Boca Raton.