Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 06–61

ON THE FINITE SAMPLE BEHAVIOUR OF FIXED BANDWIDTH BICKEL-ROSENBLATT TEST FOR UNIVARIATE AND MULTIVARIATE UNIFORMITY

CARLOS TENREIRO

ABSTRACT: The Bickel-Rosenblatt (BR) goodness-of-fit test with fixed bandwidth was introduced by Fan in 1998 [*Econometric Theory* 14, 604–621, 1998]. Although its asymptotic properties have being studied by several authors, little is known about its finite sample performance. Restricting our attention to the test of uniformity in the d-unit cube for $d \geq 1$, we present in this paper a description of the finite sample behaviour of the BR test as a function of the bandwidth h. For $d = 1$ our analysis is based not only on empirical power results but also on the Bahadur's concept of efficiency. The numerical evaluation of the Bahadur local slopes of the BR test statistic for different values of h for a set of Legendre and trigonometric alternatives give us some additional insight about the role played by the smoothing parameter in the detection of departures from the null hypothesis. For $d > 1$ we develop a Monte-Carlo study based on a set of meta-type uniforme alternative distributions and a rule-of-thumb for the practical choice of the bandwidth is proposed. For both univariate and multivariate cases, comparisons with existing uniformity tests are presented. The BR test reveals an overall good comparative performance, being clearly superior to the considered competitors tests for bivariate data.

KEYWORDS: Kernel distribution function estimation; multistage plug-in bandwidth selection; asymptotic normality.
AMS SUBJECT CLASSIFICATION (2000): 62G05, 62G20.

1. Introduction

Let $X_1, \ldots, X_n, \ldots$ be a sequence of independent and identically distributed d-dimensional absolutely continuous random vectors with unknown density function f. As it has been shown by Fan (1998) a test of the simple hypothesis $H_0 : f = f_0$ against the alternative $H_a : f \neq f_0$, where f_0 is a fixed density function on $\mathbb{R}^d$, can be based on the Bickel-Rosenblatt (BR) statistic with fixed bandwidth.

The classical BR statistic introduced by Bickel and Rosenblatt (1973) is based on the L_2 distance between the kernel density estimator f_n of f introduced by Rosenblatt (1956) and Parzen (1962), and its mathematical

Received December 27, 2006.
This work was partially supported by CMUC (Center for Mathematics, University of Coimbra)/FCT.

expectation under the null hypothesis,

$$I_n^2(h_n) = n \int \{f_n(x) - \mathrm{E}_0 f_n(x)\}^2 dx, \tag{1}$$

where, for $x \in \mathbb{R}^d$,

$$f_n(x) = \frac{1}{n} \sum_{i=1}^{n} K_{h_n}(x - X_i),$$

$K_{h_n}(\cdot) = K(\cdot/h_n)/h_n^d$ with K a kernel on $\mathbb{R}^d$, that is, a bounded and integrable function on $\mathbb{R}^d$, and (h_n), the bandwidth, is a sequence of strictly positive real numbers converging to zero as n goes to infinity (see also Fan, 1994, Gouriéroux and Tenreiro, 2001, and Tenreiro, 2006, for the asymptotic properties of the classical BR test). Analogously to Anderson, Hall and Titterington (1994) that have used kernel density estimators with fixed bandwidth for testing the equality of two multivariate probability density functions, Fan (1998) uses the BR statistic with a constant bandwidth, that is, $h_n = h > 0$, for all $n \in \mathbb{N}$, and shows that $I_n^2(h)$ can be interpreted as a L_2 weighted distance between the empirical characteristic function and the characteristic function implied by the null model with weight function $t \to |\phi_K(th)|^2$. Moreover, he provides an alternative asymptotic approximation for the finite-sample properties of the BR test by showing that the asymptotic distribution of $I_n^2(h)$ is an infinite sum of weighted χ^2 random variables (see also Tenreiro, 2005, 2006).

Although the asymptotic properties of the fixed bandwidth Bickel-Rosenblatt test for a general simple or composite hypothesis are well described in the literature, little is known about its finite sample performance. Previous studies undertaken by Henze and Zirkler (1990), Henze (1997), Henze and Wagner (1997) and Tenreiro (2005) in the case of testing normality indicate that this performance strongly depends on the choice of the bandwidth.

In this paper we explore the empirical properties of the BR statistic with fixed bandwidth to test a univariate or multivariate uniformity hypothesis, that is, we take $f_0 = U$, where U is the density of the uniform density over the d-dimensional unit cube $[0,1]^d$. The choice of this null distribution is mainly motivated by its practical significance. Examples of this practical interest are the assessing of the quality of a pseudo random number generator (see Madras, 2002, pg. 12), and the problem of goodness-of-fit to a given distribution by using the Rosenblatt's (1952) transformation. On the other hand, despite testing uniformity in the unit interval $[0,1]$ has been studied by many authors (see Stephens, 1998, for a review on the subject; see also Marhuenda, Morales and Pardo, 2005, for a recent simulation study comparing several existing univariate uniformity tests), the corresponding multidimensional problem seems to have received less attention in the literature. The exception seems to be the work by Liang, Fang, Hickernell and Li (2001) where multivariate uniformity tests based on several discrepancy criteria that arise in the error analysis of quasi-Monte Carlo methods for evaluating multiple integrals, are considered and compared. Although other easy to evaluate statistics could be used to test multivariate uniformity, like the multivariate versions of the classical Cramér-von Mises and Watson statistics (see Shorack and Wellner, 1986, chapter 5, for the asymptotic behaviour of the classical univariate EDF statistics), no finite sample analysis of these test procedures is, to our knowledge, available in the literature.

The rest of the paper is organized as follows: in section 2 we briefly recall the asymptotic properties of the BR test with fixed bandwidth and some comments are made about the evaluation of the test statistic (1). Although several simple choices for K are possible, in this paper we restrict our attention to the case where K is the standard normal density function. The test of a univariate uniformity hypothesis is considered in section 3. We give numerical evaluations of the principal components and most significant weights of the test statistic as a function of the bandwidth h. Moreover, based on the results of Tenreiro (2005), the Bahadur local slopes are numerically evaluated for different values of h for a set of Legendre and trigonometric alternatives. The simulation study presented in section 2 indicates that the finite sample properties of tests $I^2(h)$ are in good accordance with the theoretical properties based on the Bahadur local slopes. We conclude that for small values of h the BR test is appropriated to detect non-location and high order or high frequency alternatives, whereas for large values of h the test could almost exclusively detect location alternatives. Comparisons with the quadratic EDF tests of Anderson-Darling (1954) and Watson (1961) based on the empirical distribution function, and with the data-driven Neyman's test introduced in Eubank and LaRiccia (1992) are also presented. In Section 4 we consider the test of a multivariate hypothesis of uniformity. We present a simulation study involving a set of meta-type uniform alternative distributions that give us some insight about the finite sample power properties of the multivariate BR test of uniformity as a function of h. A rule-of-thumb for choosing h is proposed and the corresponding BR test is compared with multivariate versions of the Cramér-von Mises and Watson tests and with the multivariate uniformity test introduced by Liang, Fang, Hickernell and Li (2001) based on a symmetric discrepancy χ^2-type statistic. The BR test reveals an overall good comparative performance, being clearly superior to the competitors tests for bivariate data.

2. The fixed bandwidth Bickel-Rosenblatt test

For the sake of completeness we describe in this section the asymptotic behaviour of the test statistic under the null hypothesis and under a fixed alternative distribution. The convergence in distribution and the convergence in probability will be denoted by $\xrightarrow[n\to+\infty]{d}$ and $\xrightarrow[n\to+\infty]{p}$, respectively.

2.1. Asymptotic null distribution and consistency. The asymptotic behaviour of the statistic $I_n^2(h)$, with $h > 0$, for testing a composite null hypothesis was first obtained by Fan (1998). For a simple null hypothesis test it comes easily from the representation of $I_n^2(h)$ as a degenerate V-statistics. When one test the simple hypothesis $H_0 : f = U$, it takes the form

$$I_n^2(h) = \frac{1}{n}\sum_{i,j=1}^{n} Q_h(X_i, X_j), \tag{2}$$

with

$$Q_h(u, v) = W_h(u - v) - W_h \star U(u) - W_h \star U(v) + W_h \star \bar{U} \star U(0), \tag{3}$$

for $u, v \in \mathbb{R}^d$ and $h > 0$, and $W = \bar{K} \star K$, where $\star$ denotes the convolution product and $\bar{\Psi}(u) = \Psi(-u)$. Under the null hypothesis and from the limit distribution of degenerate V-statistics (cf. Theorem 1.2 of Gregory, 1997, and Theorem 4.3.2 of Koroljuk and Borovskich, 1989; see also Fan, 1998, and Tenreiro, 2005), we have

$$I_n^2(h) \xrightarrow[n\to+\infty]{d} I_\infty(h),$$

with

$$I_\infty(h) = \sum_{k=1}^{\infty} \lambda_{k,h} Z_k^2,$$

where $\{Z_k, k \geq 1\}$ are independent and identically distributed standard normal variables, and $\{\lambda_{k,h}, k \geq 1\}$, with $\lambda_{1,h} \geq \lambda_{2,h} \geq \ldots$, denotes the infinite collection of strictly positive eigenvalues of the symmetric positive definite Hilbert-Schmidt operator A_h defined, for $q \in L_2(U)$, by

$$(A_h q)(u) = \langle\, Q_h(u, \cdot), q(\cdot)\, \rangle, \tag{4}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $L_2(U)$. Moreover, under a fixed alternative f with $f \neq U$, we have

$$n^{-1}\, I_n^2(h) \underset{n\to+\infty}{\overset{p}{\longrightarrow}} (2\pi)^d \int |\Phi_f(t) - \Phi_U(t)|^2 |\Phi_K(th)|^2 dt.$$

Therefore, assuming that the Fourier transform Φ_K of K is such that $\{t \in \mathbb{R}^d : \Phi_K(t) = 0\}$ has Lebesgue measure zero, the convergence in probability of the statistic $I_n^2(h)$ to $+\infty$ for a fixed alternative, enable us to conclude that the test associated with the critical regions

$$\{I_n^2(h) > \phi_h^{-1}(1-\alpha)\},$$

for $0 < \alpha < 1$, where ϕ_h is the cdf of the random variable $I_\infty(h)$ given before, is asymptotically of size α and consistent to test H_0 against $H_a : f \neq U$.

2.2. Evaluating the test statistic. From the representation (2) we easily see that $I_n^2(h)$ can be written as

$$I_n^2(h) = -I_n^{2,1}(h) + I_n^{2,2}(h) + W_h(0) + n\, W_h \star \bar{U} \star U(0),$$

where

$$I_n^{2,1}(h) = 2 \sum_{i=1}^{n} W_h \star U(X_i) \tag{5}$$

and

$$I_n^{2,2}(h) = \frac{2}{n} \sum_{1\le i<j\le n} W_h(X_i - X_j). \tag{6}$$

Therefore, the calculation of $I_n^2(h)$ can be easily performed if K is chosen such that the convolutions $W = \bar{K} \star K$ and $W_h \star U$ have close forms. Even in the case where $W_h \star U$ is hard to evaluate, the calculation of (5) can be simplified by using quasi-Monte Carlo methods to approximate the convolution $W_h \star U(u) = \int_{[0,1]^d} W_h(u-x)dx$, by the sample mean $\frac{1}{m}\sum_{j=1}^m W_h(u-x_j)$ over a set of uniformly scattered sample points $\{x_1, \ldots, x_m\} \subset [0,1]^d$ (see Niederreiter, 1992).

If K is a product kernel on $\mathbb{R}^d$, that is, $K(u) = \prod_{i=1}^d k(u_i)$, for $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$, where k is a kernel on $\mathbb{R}$, and the functions $w = \bar{k} \star k$

and $\mathbf{w}(x) = \int_{]-\infty,x]} w(t)dt$ are easy to evaluate, the exact calculation of $I_n^2(h)$ does not involve any problem since

$$W_h(u) = \prod_{i=1}^{d} w_h(u_i)$$

and

$$W_h \star U(u) = \prod_{i=1}^{d} (\mathbf{w}((u_i - 1)/h) - \mathbf{w}(u_i/h)).$$

Several simple choices for k are possible. In the following we restrict our attention to the case where k is the standard normal density function $k(x) = \exp(-x^2/2)/(2\pi)^{1/2}$, for $x \in \mathbb{R}$. However, similar qualitative results can be obtained for other kernels like the standard gamma density function $k(x) = \exp(-x)$, for $x \geq 0$. This is not surprising, because the test statistic depends on k through the convolution $\bar{k} \star k$ which present a quite similar shape for these two quite different kernels.

3. Testing univariate uniformity

In this section we consider the test of a univariate hypothesis of uniformity. In order to get a better understanding of the role played by the smoothing parameter in the detection of departures from the null hypothesis, and to compare the BR tests with existing test procedures, we present in the following an analysis based not only on empirical power results but also on the Bahadur's concept of efficiency (see Nikitin, 1995).

3.1. Local alternatives and Bahadur efficiency. If $\{f(\cdot;\theta) : \theta \in \Theta\}$, where Θ is a nontrivial closed real interval, is a family of probability density functions containing the density U, that is, $U = f(\cdot;\theta_0)$, for some $\theta_0 \in \Theta$, it is natural to compare a set of competitor tests through their Bahadur local exact slopes when $\theta \to \theta_0$. For θ in an appropriate right neighbourhood of the origin, in the following we consider local alternatives of the form

$$f(x;\theta) = 1 + \theta\, Q(x), \tag{7}$$

for $0 \leq x \leq 1$, where Q is a bounded function that belongs to the tangent space $H(U)$ of U defined by $H(U) = \{h \in L_2(U) : \langle\, h, U \,\rangle = 0\}$. From the results of Tenreiro(2005), the Bahadur local exact slope of the BR test $I_n^2(h)$

corresponding to the previous local alternative is given by

$$C_{I_n^2(h)}(f(\cdot;\theta)) = \lambda_{1,h}^{-1}\langle\, W_h \star Q, Q\,\rangle\, \theta^2\,(1 + o(1)),\ \text{when}\ \theta\to 0,$$

where $\lambda_{1,h}$ is the largest eigenvalue of the operator A_h defined by (4).

Moreover, the previous local slope can be written in terms of the infinite collection $\{\lambda_{k,h}, k \geq 1\}$ of strictly positive eigenvalues of A_h and of the principal components $\{q_{k,h}, k \geq 1\}$, which are the orthonormal basis for $H(U)$ corresponding to the previous eigenvalues, that is, for all k and j, $A_h q_{k,h} = \lambda_{k,h} q_{k,h}$, *a.e.* (U) and $\langle\, q_{k,h}, q_{j,h}\,\rangle = \delta_{kj}$, where δ_{kj} is the Kronecker symbol:

$$C_{I_n^2(h)}(f(\cdot;\theta)) = \sum_{k=1}^{\infty} \lambda_{1,h}^{-1}\lambda_{k,h}\ \langle\, q_{k,h}, Q\,\rangle^2\ \theta^2(1 + o(1)),\ \text{when}\ \theta\to 0.$$

From the previous representation, namely from the fact that the eigenvalues $(\lambda_{k,h})$ converge to zero, it is clear that only a finite directions of alternatives effectively contribute to $C_{I_n^2(h)}(f(\cdot;\theta))$. The natural question, that we discuss in the next paragraph, is how rapidly the principal directions loose influence.

3.2. Principal components and most significant weights. As described in the previous paragraph, the Bahadur local slope of $I_n^2(h)$ depends on the weights $(\gamma_{k,h})$, where $\gamma_{k,h} = \lambda_{1,h}^{-1}\lambda_{k,h}$, and on the principal components $(q_{k,h})$. Numerical evaluations of the most significants weights are shown in Table 1 for several values of h. These approximations have been obtained through a quadrature method using Lapack routines (see Anderson *et al.*, 1999). For

	$h = 0.01$	$h = 0.1$	$h = 1.0$	A_n^2	U_n^2
$\gamma_{2,h}$	9.96×10^{-1}	7.40×10^{-1}	2.46×10^{-2}	3.33×10^{-1}	1.00×10^{-0}
$\gamma_{3,h}$	9.89×10^{-1}	4.25×10^{-1}	1.80×10^{-4}	1.67×10^{-1}	2.50×10^{-1}
$\gamma_{4,h}$	9.81×10^{-1}	2.40×10^{-1}	2.12×10^{-6}	1.00×10^{-1}	2.50×10^{-1}
$\gamma_{5,h}$	9.70×10^{-1}	1.07×10^{-1}	3.52×10^{-7}	6.67×10^{-2}	1.11×10^{-1}
$\gamma_{6,h}$	9.59×10^{-1}	4.72×10^{-2}	9.04×10^{-9}	4.76×10^{-2}	1.11×10^{-1}
$\gamma_{7,h}$	9.44×10^{-1}	1.69×10^{-2}	5.19×10^{-13}	3.57×10^{-2}	6.25×10^{-2}
$\gamma_{8,h}$	9.30×10^{-1}	6.04×10^{-3}	2.13×10^{-13}	2.78×10^{-2}	6.25×10^{-2}
$\gamma_{9,h}$	9.12×10^{-1}	1.79×10^{-3}	6.91×10^{-16}	2.22×10^{-2}	4.00×10^{-2}
$\gamma_{10,h}$	8.95×10^{-1}	5.35×10^{-4}	2.60×10^{-16}	1.82×10^{-2}	4.00×10^{-2}
$\gamma_{11,h}$	8.75×10^{-1}	1.35×10^{-4}	1.41×10^{-16}	1.52×10^{-2}	2.78×10^{-2}
$\gamma_{12,h}$	8.55×10^{-1}	3.44×10^{-5}	1.39×10^{-16}	1.28×10^{-2}	2.78×10^{-2}

Table 1: Weights $\gamma_{k,h}$ for $I_n^2(h)$ with K the standard normal density function

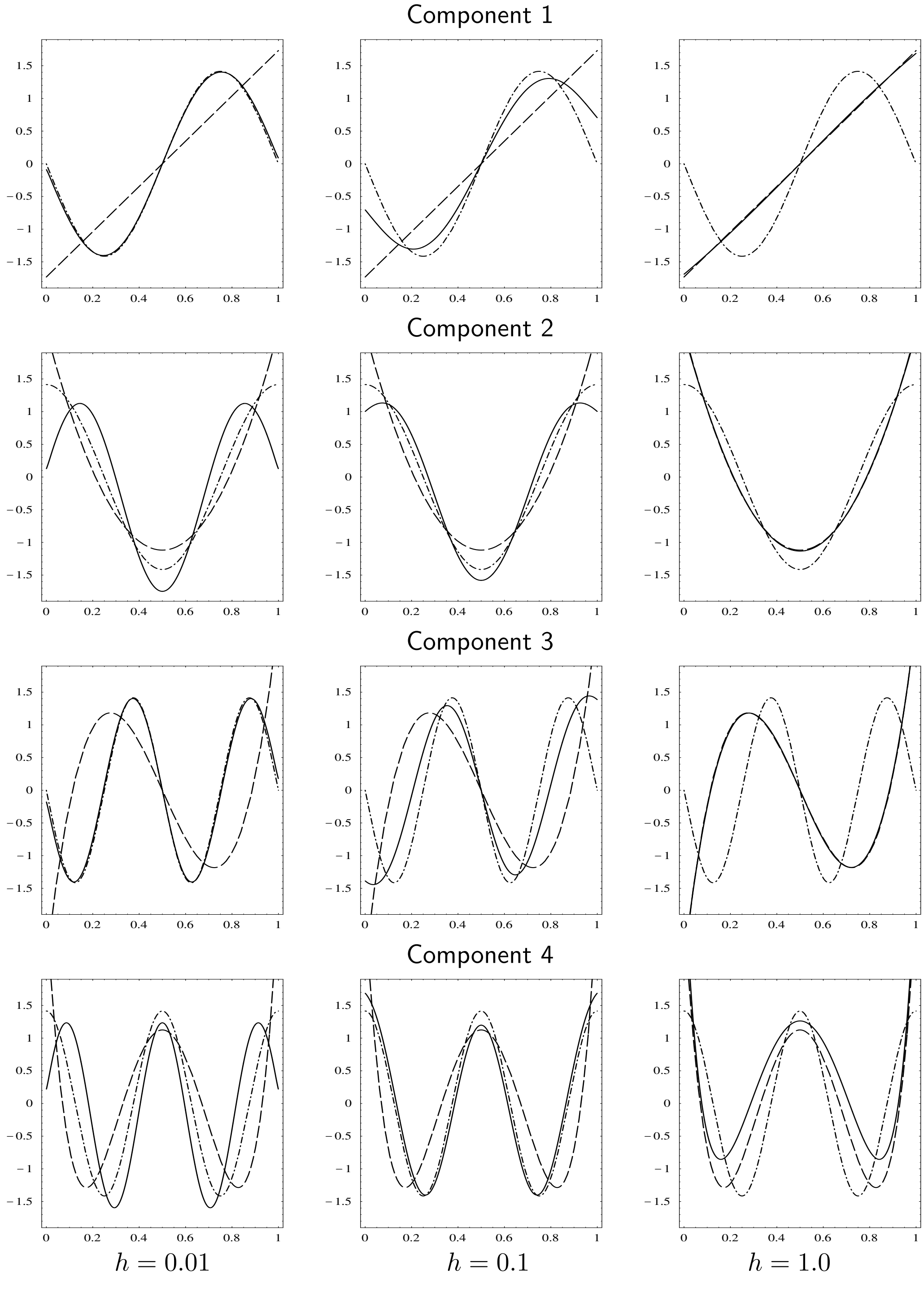

Figure 1: Principal components for: $I_n^2(h)$ – solid line; A_n^2– broken line; U_n^2– broken and dotted line

comparison, we also present the corresponding weights for the well known quadratic EDF tests A_n^2 of Anderson-Darling (1954) and U_n^2 of Watson (1961) based on the empirical distribution function (see also Shorack and Wellner, 1986, chapter 5). These tests are consistent against all alternatives and have shown good performance in testing uniformity in several comparative simulation studies (see Stephens, 1974, Miller and Quesenberry, 1979, and Marhuenda, Morales and Pardo, 2005).

From the values shown in Table 1 and the representation for the Bahadur local slopes given in the previous subsection, we expect that the BR test for small values of h could use information contained in other components different from the first ones. This conclusion is in accordance with the properties of the classical BR test whose asymptotic power function does not depend on the actual direction of the alternative under consideration (see Bickel and Rosenblatt, 1973, Fan, 1994, and Gouriéroux and Tenreiro, 2001). However, for moderate or large values of h, it appears that $I_n^2(h)$ might exclusively use information contained in the first component. See Tenreiro (2005) for similar conclusions in the test of a simple hypothesis of normality. In these last cases the test behaves very much like a parametric test for a one-dimensional alternative whereas in the former cases the test behaves like a well-balanced test for higher-dimensional alternatives.

In Figure 1 we plot the first four principal components of $I_n^2(h)$ for the values of h considered in Table 1. Since the components of the Anderson-Darling test are the Legendre polynomials which arise from the ortogonalization of powers, it is clear that in some sense for all values of h the first four principal components describe deviations in location, scale, skewness and kurtosis, respectively, from the null hypothesis. Therefore, taking into account the previous conclusions, it appears that for small values of h the BR test could be appropriate to detect non-location and high order alternatives, whereas for large values of h the test could, almost exclusively, detect location alternatives. Finally, it is interesting to remark that for very small values of h the first and third principal components of the BR test agree quite well with the corresponding components of the Watson test and for large values of h the first three principal components of the BR test are close to the corresponding components of the Anderson-Darling test.

3.3. Bahadur local exact slopes. In this paragraph the BR test with fixed bandwidth is compared, for different values of h, with the quadratic EDF tests of Anderson-Darling and Watson through their Bahadur exact slopes for two sets of local alternatives of the form (7). In the first set of

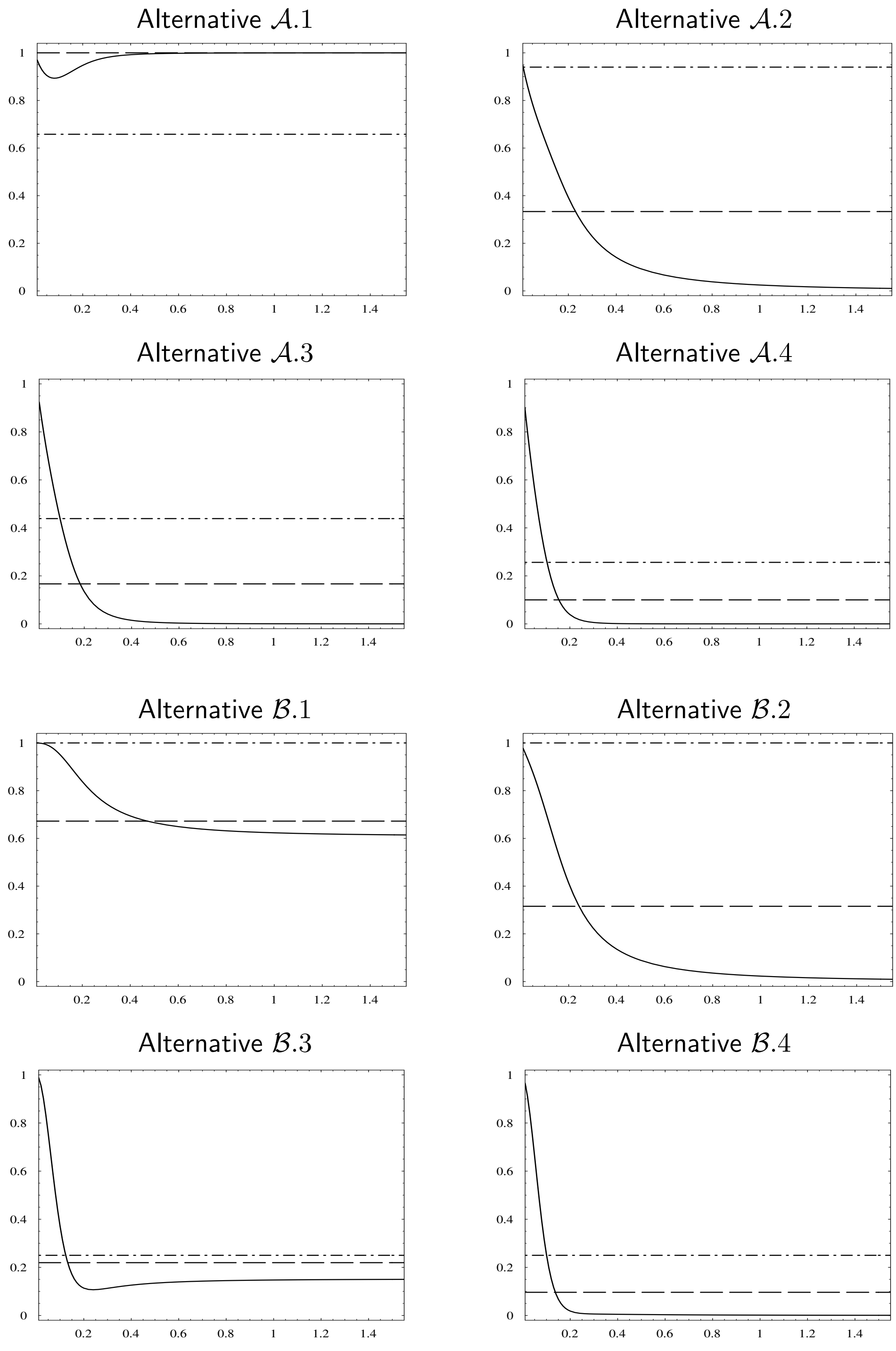

Figure 2: Local indices for: $I_n^2(h)$ – solid line; A_n^2– broken line; U_n^2– broken and dotted line

alternatives that we denote by $(\mathcal{A}.\,j)$, for $j = 1, \ldots, 4$, we take for Q the jth Legendre polynomial defined by:

$$
\begin{aligned}
P_1(x) &= \sqrt{3}\,(2x-1);\\
P_2(x) &= \sqrt{5}\,(6x^2-6x+1);\\
P_3(x) &= \sqrt{7}\,(20x^3-30x^2+12x-1);\\
P_4(x) &= 3\,(70x^4-140x^3+90x^2-20x+1).
\end{aligned}
$$

These four polynomials are the first principal components of Anderson-Darling's test and, as mentioned before, the alternatives $(\mathcal{A}.\,j)$ describe deviations in location, scale, skewness and kurtosis, respectively, from the null hypothesis. The second set of alternatives is based on the first four principal components of Watson's test. These are denoted by $(\mathcal{B}.\,j)$, for $j = 1,\ldots,4$, and Q is one of the trigonometric functions

$$
\begin{aligned}
T_1(x) &= \sqrt{2}\,\sin(2\pi x);\\
T_2(x) &= \sqrt{2}\,\cos(2\pi x);\\
T_3(x) &= \sqrt{2}\,\sin(4\pi x);\\
T_4(x) &= \sqrt{2}\,\cos(4\pi x).
\end{aligned}
$$

Since the Bahadur local exact slopes of the tests we consider take the form $\theta^2(1+o(1))$, up to the multiplication by a constant, when $\theta\to 0$ (see Nikitin, 1995, pp. 73–81, for quadratic EDF tests), for the comparison of such tests it is sufficient to compare the coefficients of θ^2. They are usually called local indices and are plotted in Figure 2 for $h \in [0.01, 1.5]$. We also plot the local indices for the Anderson-Darling and Watson tests.

It is clear from Figure 2 that a moderate or large bandwidth leads to a test with high efficiency for deviations in location whereas a small bandwidth leads to a high efficiency test for other moments alternatives or trigonometric alternatives. However, the gain of efficiency in the location alternative $(\mathcal{A}.1)$ by taking a large value of h implies a severe loss of efficiency in non-location alternatives.

3.4. Some simulation results. To examine the power performance of BR tests for several choices of the bandwidth, and to determine if the previous comparisons based on Bahadur local efficiency reflect the finite sample properties of BR tests, a simulation experiment is undertaken including the BR tests $I_n^2(0.02)$ (small bandwidth), $I_n^2(0.1)$ (medium bandwidth) and $I_n^2(0.5)$ (large bandwidth). These bandwidths have been chosen after some preliminar simulation work with a large set of bandwidths that allowed to identify some systematic behaviour of the sample power of the BR test as a function

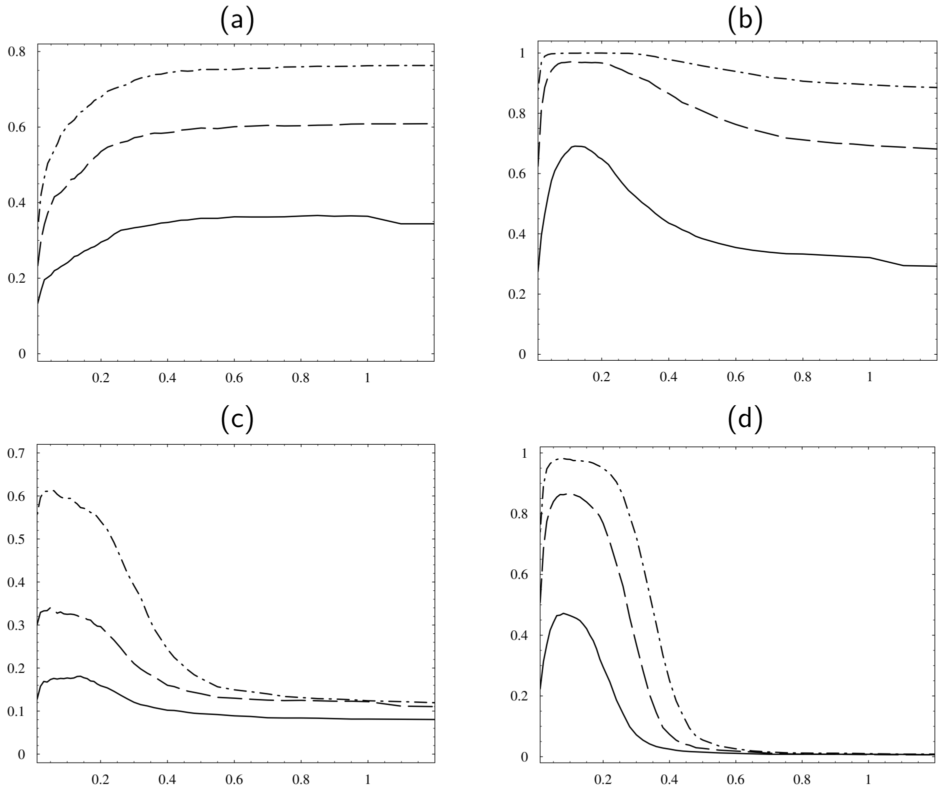

Figure 3: Empirical power of $I_n^2(h)$ as function of h for $n = 20$ (solid line), $n = 40$ (broken line) and $n = 60$ (broken and dotted line) at level 0.05: (a) Beta(0.75,1.125); (b) Beta(1.5,2.25); (c) Beta(0.5,0.5); (d) Beta(2.5,2.5)

of h. In Figure 3, the graphics (a) and (b) show the general behaviour of the power as a function of h for location alternatives and (c) and (d) reveals the general behaviour of the power as a function of h for non-location alternatives. These empirical results are globally in accordance with the theoretical results based on Bahadur local efficiency presented in Figure 2. However, contrary to the Bahadur local efficiency results, a very small bandwidth is not the best choice for h. Also remark that the values of h that maximize the power do not depend significantly on n.

Additionally to the EDF tests A_n^2 and U_n^2 that as before will be use for comparison, we also consider the data-driven Neyman's test Z_{nm} introduced in Eubank and LaRiccia (1992) which is based on the first m principal components of the Cramér-von Mises EDF statistic given, for $j = 1, 2, \ldots$, by $q_j(x) = \cos(\pi j x)$, $0 \leq x \leq 1$, where m is not a fixed integer, but it depends on the observations. The inclusion of this test procedure in our simulation study is motivated by the good empirical power results for the test reported by Eubank and LaRiccia (1992) and Kim (2000).

The power properties of the previous tests are investigated under four sets of alternative distributions:

1) Beta alternatives

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, x^{a-1}(1-x)^{b-1}, \quad 0 \leq x \leq 1,$$

for $a, b > 0$, where Γ is the gamma function.

2) Legendre alternatives of the form

$$f(x) = 1 + \rho\, P_j(x), \quad 0 \leq x \leq 1,$$

for $\rho > 0$ and $j = 1, 2, 3, 4, 5, 6$, where P_j is the jth Legendre polynomial.

3) Cosine alternatives of the form

$$f(x) = 1 + \rho \cos(\pi j x), \quad 0 \leq x \leq 1,$$

for $\rho > 0$ and $j = 1, 2, 3, 4, 6, 8, 10$.

4) Sine alternative with

$$f(x) = 1 + \rho \sin(\pi j x), \quad 0 \leq x \leq 1,$$

for $\rho > 0$ and $j = 2, 4, 8, 12$.

Some of these alternatives have been used in Eubank and LaRiccia (1992) and Kim (2000). The Legendre alternatives permit us to describe deviations to the null hypothesis in location ($j = 1$), scale ($j = 2$), skewness ($j = 3$), kurtosis ($j = 4$) and high moment alternatives ($j = 5, 6$). For the trigonometric alternatives the parameter j controls the frequency of the alternative and allows the analysis of the performance of the tests as a function of the frequency of the alternative. For the trigonometric and Legendre alternatives the parameter ρ determines the L_2 distance of the alternative from the null hypothesis. Several values of ρ where considered but similar qualitative results were observed. A more realistic set of models for the alternative to the null hypothesis of uniformity is given by the beta distributions. The set of values taken for $a, b > 0$ lead to different shape alternatives, and, in particular, to symmetric ($a = b$) and asymmetric ($a \neq b$) alternatives.

a	b	μ	σ^2	n	$I_n^2(0.02)$	$I_n^2(0.1)$	$I_n^2(0.5)$	A_n^2	U_n^2	Z_{mn}
					Asymmetric beta alternatives					
0.5	0.75	0.4	0.107	20	0.35	0.41	0.41	0.60	0.36	0.48
				40	0.62	0.68	0.62	0.82	0.62	0.77
				60	0.82	0.85	0.77	0.94	0.81	0.90
0.75	1.125	0.4	0.083	20	0.17	0.24	0.39	0.37	0.15	0.27
				40	0.30	0.45	0.60	0.62	0.27	0.51
				60	0.42	0.61	0.75	0.78	0.36	0.65
1.0	1.5	0.4	0.069	20	0.15	0.27	0.36	0.32	0.18	0.26
				40	0.25	0.50	0.62	0.61	0.33	0.50
				60	0.39	0.68	0.79	0.79	0.46	0.68
1.5	2.25	0.4	0.051	20	0.23	0.42	0.36	0.34	0.43	0.32
				40	0.49	0.77	0.69	0.76	0.77	0.72
				60	0.75	0.95	0.89	0.95	0.93	0.91
2.0	3.0	0.4	0.040	20	0.39	0.68	0.38	0.46	0.73	0.52
				40	0.81	0.97	0.81	0.94	0.97	0.95
2.5	3.75	0.4	0.033	20	0.56	0.85	0.42	0.59	0.90	0.73
				40	0.95	1.00	0.86	0.99	1.00	1.00
					Symmetric beta alternatives					
0.5	0.5	0.5	0.125	20	0.33	0.33	0.16	0.54	0.44	0.44
				40	0.60	0.59	0.18	0.79	0.74	0.74
				60	0.82	0.81	0.26	0.92	0.91	0.90
0.75	0.75	0.5	0.100	40	0.12	0.13	0.09	0.16	0.18	0.17
				60	0.16	0.18	0.09	0.21	0.25	0.22
				80	0.20	0.25	0.10	0.28	0.34	0.30
1.5	1.5	0.5	0.063	40	0.14	0.19	0.04	0.06	0.28	0.18
				60	0.21	0.28	0.04	0.11	0.41	0.27
				80	0.30	0.41	0.04	0.21	0.54	0.41
2.0	2.0	0.5	0.050	20	0.21	0.29	0.03	0.04	0.39	0.24
				40	0.40	0.57	0.03	0.23	0.74	0.57
				60	0.63	0.80	0.04	0.52	0.91	0.79
2.5	2.5	0.5	0.042	20	0.31	0.46	0.02	0.04	0.39	0.24
				40	0.69	0.87	0.03	0.56	0.95	0.87
3.5	3.5	0.5	0.031	20	0.47	0.67	0.02	0.16	0.82	0.63
				40	0.96	0.99	0.03	0.95	1.00	0.99

Table 2: Empirical power at level 0.05 for beta alternatives

For all the considered alternative distributions, the empirical power results, that we present in Tables 2-3 at level 0.05, were evaluated on the basis of 2000 Monte-Carlo samples of size n, for $n = 20, 40, 60$ or 80. Similar qualitative results were observed for the levels 0.01 and 0.1. The critical values of all the test statistics were found by simulating 10^4 samples from the null distribution.

From Tables 2-3, we conclude that for small values of h the BR test is

ρ	j	μ	σ^2	n	$I_n^2(0.02)$	$I_n^2(0.1)$	$I_n^2(0.5)$	A_n^2	U_n^2	Z_{mn}
					Legendre alternatives					
0.3	1	0.587	0.076	20	0.13	0.22	0.30	0.27	0.15	0.20
				60	0.29	0.53	0.66	0.63	0.37	0.52
				80	0.38	0.66	0.78	0.78	0.49	0.66
0.3	2	0.5	0.105	20	0.12	0.15	0.08	0.14	0.19	0.15
				60	0.28	0.41	0.11	0.29	0.51	0.38
				80	0.37	0.55	0.12	0.42	0.64	0.54
0.3	3	0.5	0.083	20	0.10	0.11	0.06	0.07	0.11	0.10
				60	0.25	0.26	0.06	0.11	0.25	0.28
				80	0.34	0.38	0.06	0.13	0.35	0.40
0.3	4	0.5	0.083	20	0.10	0.09	0.05	0.07	0.08	0.11
				60	0.25	0.16	0.05	0.09	0.13	0.24
				80	0.32	0.22	0.05	0.11	0.17	0.33
0.3	5	0.5	0.083	20	0.10	0.07	0.05	0.06	0.07	0.08
				60	0.23	0.10	0.05	0.07	0.09	0.18
				80	0.32	0.14	0.05	0.08	0.13	0.26
0.3	6	0.5	0.083	20	0.10	0.05	0.06	0.06	0.05	0.08
				60	0.22	0.08	0.05	0.06	0.08	0.17
				80	0.28	0.08	0.05	0.06	0.09	0.23
					Cosine alternatives					
1.0	1	0.297	0.042	20	0.55	0.90	0.95	0.94	0.74	0.89
				40	0.95	1.00	1.00	1.00	0.98	1.00
1.0	2	0.5	0.134	20	0.55	0.85	0.20	0.55	0.89	0.78
				40	0.95	1.00	0.28	0.93	1.00	1.00
1.0	3	0.477	0.083	20	0.53	0.57	0.07	0.16	0.40	0.65
				40	0.94	0.96	0.07	0.43	0.82	0.98
1.0	4	0.5	0.096	20	0.54	0.27	0.08	0.17	0.19	0.60
				40	0.94	0.69	0.06	0.28	0.59	0.96
1.0	6	0.5	0.089	20	0.49	0.07	0.06	0.10	0.10	0.46
				40	0.91	0.10	0.07	0.15	0.18	0.92
1.0	8	0.5	0.086	20	0.43	0.05	0.06	0.08	0.07	0.36
				40	0.87	0.05	0.05	0.10	0.10	0.86
1.0	10	0.5	0.085	20	0.35	0.05	0.05	0.08	0.07	0.29
				40	0.79	0.04	0.05	0.07	0.07	0.80
					Sine alternatives					
0.5	2	0.420	0.078	60	0.44	0.72	0.63	0.61	0.72	0.59
				80	0.57	0.84	0.76	0.77	0.84	0.74
0.5	4	0.460	0.082	60	0.43	0.31	0.18	0.22	0.16	0.40
				80	0.54	0.44	0.21	0.29	0.24	0.52
0.5	6	0.473	0.083	60	0.36	0.09	0.11	0.12	0.08	0.28
				80	0.49	0.10	0.12	0.15	0.09	0.41
0.5	8	0.480	0.083	60	0.33	0.06	0.08	0.09	0.06	0.24
				80	0.44	0.07	0.08	0.10	0.07	0.34
0.5	12	0.487	0.083	60	0.23	0.06	0.07	0.07	0.05	0.18
				80	0.32	0.07	0.07	0.09	0.07	0.26

Table 3: Empirical power at level 0.05 for Legendre and trigonometric alternatives

appropriate to detect non-location, high moment and high frequency alternatives, whereas for large values of h the test exclusively detect location alternatives. For high moment or high frequency alternatives (that is, Legendre and trigonometric alternatives with $j > 4$) the best tests are the Z_{nm} and $I_n^2(0.02)$ tests. Both EDF tests and the BR tests with $h = 0.1, 0.5$ have no power for these alternatives. If we restrict our attention to low frequency alternatives (that is, beta alternatives and Legendre and trigonometric alternatives with $j \leq 3$), the best tests are the Z_{nm} and $I_n^2(0.1)$ tests. Remark that although the superior results obtained by the Anderson-Darling test for pure location alternatives and by the Watson test for scale alternatives, these tests also reveals a poor performance for some of the considered low frequency alternatives.

The results also confirm the good power properties of Z_{nm} test reported by Eubank and LaRiccia (1992) and Kim (2000). This test behaves like a omnibus test which shows a good or reasonable performance against a large range of alternatives. Therefore, it should be used in practice if no information is available about the alternative to the null hypothesis. This omnibus property is also shared by $I_n^2(0.02)$. However, although for high moment and high frequency alternatives the tests Z_{nm} and $I_n^2(0.02)$ have obtained the best power results, for low frequency alternatives Z_{nm} is superior to $I_n^2(0.02)$.

4. Testing multivariate uniformity

The BR test for a multivariate uniformity hypothesis is discussed in this section. In order to describe its finite sample power performance as a function of h, to propose a rule-of-thumb for the practical choice of the bandwidth h, and to compare the corresponding BR test with other existing uniformity tests, we develop in this section a Monte-Carlo study based on a set of meta-type uniform alternative distributions.

A meta-type uniform distribution in $[0,1]^d$ can be seen as the distribution of the random vector $(V_1, \ldots, V_d)$ that is obtained from an absolutely continuous random vector $(X_1, \ldots, X_d)$, by taking $V_i = G_i(X_i)$, where, for $i = 1, \ldots, d$, G_i is the distribution function of X_i (see Fang, Fang and Kotz, 2002, for the idea of meta-distribution). Therefore, these alternatives are appropriate to model the situation where one does not have relevant information about the dependence structure of the alternative to the null distribution but it is known that its support is contained in the unit d-cube and their margins are uniformly distributed in the interval $[0, 1]$.

In the simulation we choose the random vector $(X_1, \ldots, X_d)$ to have one of the following multivariate distributions where $\mu \in \mathbb{R}^d$ and the matrix $\Sigma = [\sigma_{ij}] := \Sigma_\rho$ is chosen as $\sigma_{ii} = 1$ and $\sigma_{ij} = \rho$, with $0 < \rho < 1$, for $1 \le i \ne j \le d$ (see Liang, Fang and Hickernell, 2001, for a similar set of alternative distributions; see also Johnson, 1987, and Kotz, Kozubowski and Podgórski, 2000, for relevant information about these distributions):

1) The multivariate normal distribution, $N_d(\mu, \Sigma)$ with mean μ and covariance matriz Σ.

2) The multivariate t-distribution $T_d(m, \mu, \Sigma)$ with density function

$$g(x) = C|\Sigma|^{-1/2}(1 + m^{-1}(x-\mu)'\Sigma^{-1}(x-\mu))^{-(m+d)/2},$$

with $m > 0$, where, here and in the following, $C > 0$ is a normalizing constant that takes possibly different values in each occurrence.

3) The symmetric Kotz type distribution $K_d(N, \mu, \Sigma)$ with density function given by

$$g(x) = C|\Sigma|^{-1/2}((x-\mu)'\Sigma^{-1}(x-\mu))^{N-1} \exp\left(-\sqrt{(x-\mu)'\Sigma^{-1}(x-\mu)}\right),$$

where $2N + d > 2$.

4) The symmetric Pearson type II distribution $P_d(m, \mu, \Sigma)$ with density function

$$g(x) = C|\Sigma|^{-1/2}(1 - (x-\mu)'\Sigma^{-1}(x-\mu))^m,$$

having support $(x-\mu)'\Sigma^{-1}(x-\mu) \le 1$ and shape parameter $m > -1$.

5) The logistic distribution $L_d(\alpha)$ with density function

$$g(x) = C \exp\Big(-\sum_{i=1}^d x_i\Big)\Big(\sum_{i=1}^d \exp(-x_i) + 1\Big)^{-(\alpha+d)},$$

with $\alpha > 0$.

6) The asymmetric Laplace distribution $AL_d(\mu, \Sigma)$ with density function

$$g(x) = C|\Sigma|^{-1/2} \exp(x'\Sigma^{-1}\mu)\left(\frac{x'\Sigma^{-1}x}{2 + \mu'\Sigma^{-1}\mu}\right)^{\nu/2} \times K_\nu\left(\sqrt{(2+\mu'\Sigma^{-1}\mu)(x'\Sigma^{-1}x)}\right),$$

where $\nu = (2-d)/2$ and K_ν is the modified Bessel function of third kind given by $K_\nu(u) = u^\nu \int_0^\infty t^{-\nu-1} \exp(-t - u^2/(4t)) dt/2^{\nu+1}, u > 0$.

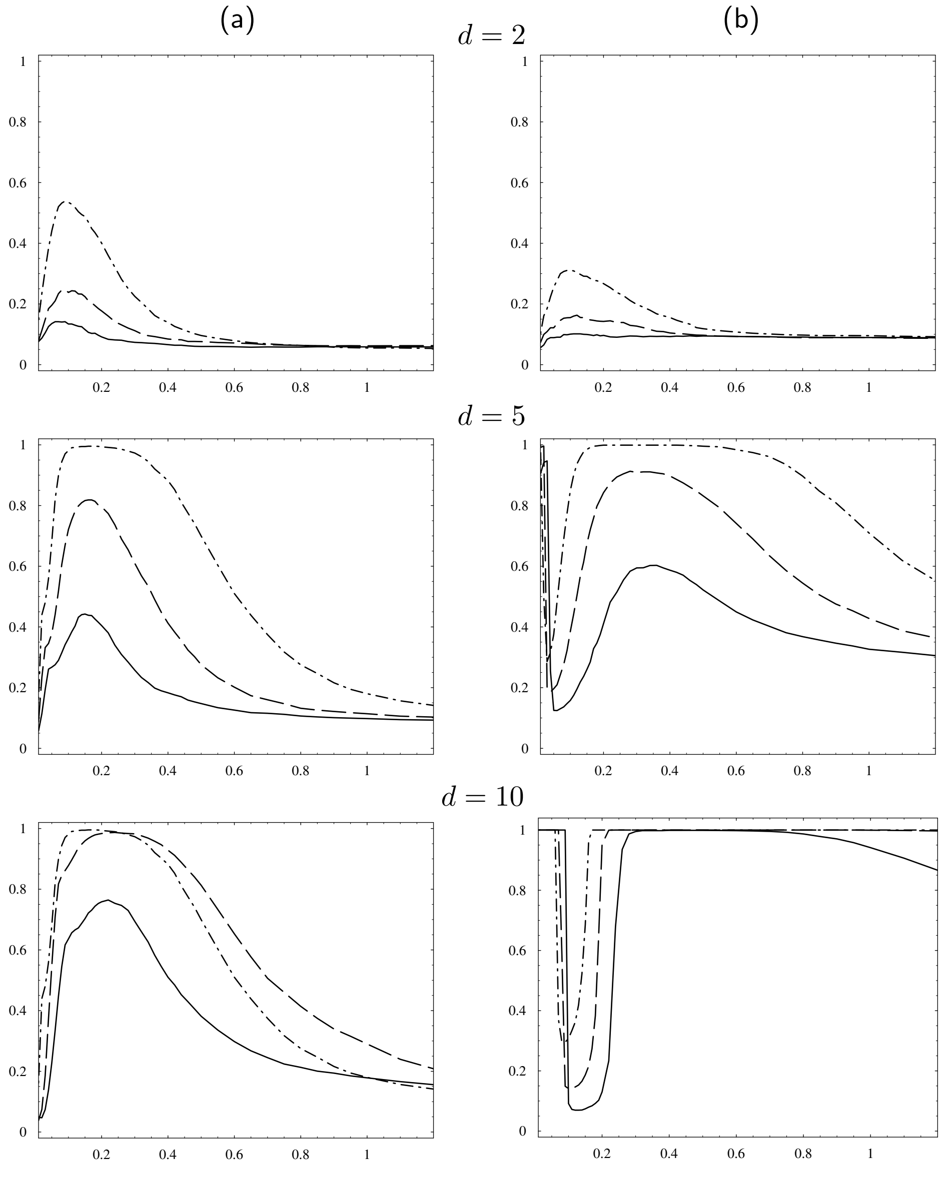

Figure 4: Empirical power of $I_n^2(h)$ as function of h for $n = 10$ (solid line), $n = 20$ (broken line), $n = 40$ (broken and dotted line) and $d = 2, 5, 10$ at level 0.05: (a) $AL_d(0, \Sigma_{0.2})$; (b) $K_d(1, 0, \Sigma_{0.5})$

After some simulation work for some of the previous alternative distributions, we concluded that, as in the univariate case, the value of h that maximize the empirical power of the BR test does not depend significantly on the sample size n but strongly depends on the underlying alternative distribution and data dimension. This is illustrated in Figure 4 where we present two typical behaviours of the power of the BR test as a function of h, for $n = 10, 20, 40$, $d = 2, 5, 10$ and $h \in [0.01, 1.2]$. For the Kotz multivariate distribution (b) it is interesting to remark the large empirical power obtained by the BR test for $d = 5, 10$ when h is very small. This type of behaviour occurs since for small values of hd^{-1} the term $I_n^{2,1}(h)$ given by (5) dominates the term $I_n^{2,2}(h)$ given by (6), and, as a consequence, it determines the behaviour of test statistic $I_n^2(h)$. Therefore, for such values of h the BR test is essentially based on the sum $\frac{1}{n}\sum_{i=1}^n W_h \star U(X_i)$, and detect alternatives f that satisfy $\int W_h \star U(u)U(u)du \neq \int W_h \star U(u)f(u)du$. This is in particular true for U-shaped distributions like meta-uniform distribution based on symmetric Kotz multivariate distribution, because they give significant probability to regions in the neighbourhood of the d-unit cube frontier.

In order to use the BR test in practice and to compare it with some existent multivariate uniformity test procedures, it is essential to propose an easy-to-use rule for choosing the bandwidth h. Since we do not have a particular type of alternative in mind, it is natural to expect that BR test should show a reasonable performance against a large range of alternatives.

With this goal, for data dimensions from $d = 2$ to 10, for each one of the following alternative distributions, and for n taking one of the values $n = 10$ or $n = 20$, we calculate the bandwidth $h_{f,n}$ that maximizes the empirical power over the set $\{0.01, 0.02, \ldots, 1.2\}$ of values of h (when the sample power is maximized for more than one value of h, we take for $h_{f,n}$ the smallest of such values of h). With the exception of the asymmetric Laplace distribution where $\mu = 1 = (1, \ldots, 1)$, we take $\mu = 0$ and $\rho = 0.2, 0.5$ for all the distributions depending on μ and on Σ_ρ, respectively. Moreover, for the Student distribution we take $m = 3$, for the Kotz distribution we consider $N = 1$, for the Pearson distribution we take $m = 3/2$ and the value $\alpha = 0.5$ is considered for the logistic distribution.

The distribution of the bandwidths $h_{f,n}$ for the BR test at level 0.05, is described in Figure 5. A logarithmic regression of the bandwidths $h_{f,n}$ sample medians over the data dimension, leads to the following relation that we will use in the following as rule-of-thumb for the choice of h when the dimension of the data is d:

$$h = 0.09 \ln(d) + 0.036. \tag{8}$$

This rule-of-thumb can also be used for the BR test at levels 0.01 and 0.1, since the distributions of the bandwidths $h_{f,n}$ for these two levels are very similar to that one shown in Figure 5.

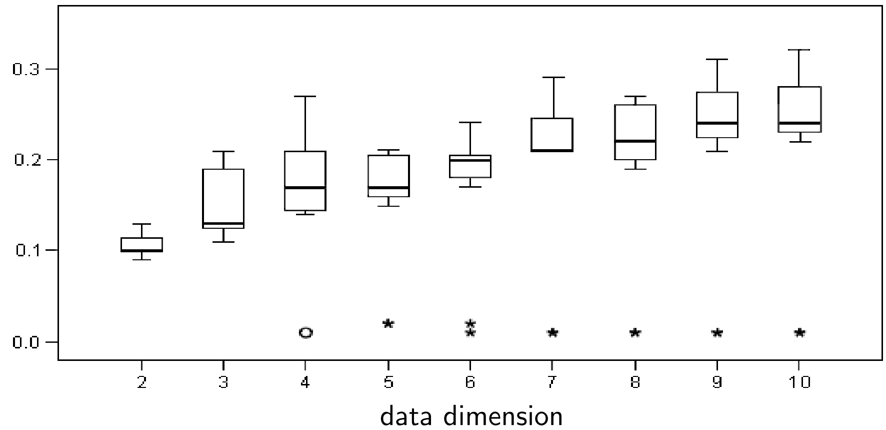

Figure 5: Empirical distribution of $h_{f,n}$ for several values of d at level 0.05

In the following the BR test with h given by (8) is compared with two other easy to evaluate multivariate uniformity tests. They are the multivariate versions of the Cramér-von Mises statistic W_n^2 based on the empirical distribution function and the multivariate uniformity test introduced by Liang, Fang, Hickernell and Li (2001) based on a symmetric discrepancy χ^2-type statistic T_n. Initially the multivariate version of the Watson statistic U_n^2 was also considered but, in general, this test revealed a very poor performance in comparison with the other considered tests. Among the several test statistics considered by Liang, Fang, Hickernell and Li (2001) which are based on quasi-Monte Carlo methods for measuring the discrepancy of points in $[0,1]^d$, the statistic T_n considered in the following has shown the best results. However, contrary to the considered BR tests and the previous EDF tests, the test based on T_n is not asymptotically consistent against all alternative distributions.

In Tables 4-5 we present the empirical power results at level 0.05 of the BR test for h given by (8) and of the above mentioned W_n^2 and T_n tests. These results were obtained on the basis of 2000 Monte-Carlo samples of sizes $n = 10, 20, 40, 60$ or 80. Similar qualitative results were observed for the levels 0.01 and 0.1. For the evaluation of the critical values of all the involved test statistics 10^4 samples from the null distribution were used. The set of alternative distributions we consider includes some of the distributions used to derive the rule-of-thumb (8) and some other meta-uniform distributions based on the Student distributions with $m = 1$ (Cauchy distribution) and $m=5$, on the Kotz distribution with $N=2$, on the Pearson distribution with

	$d=2$				$d=3$				$d=4$			
X-Distribution	n	I_n^2	W_n^2	T_n	n	I_n^2	W_n^2	T_n	n	I_n^2	W_n^2	T_n
$N_d(0,\Sigma_{0.2})$	40	0.10	0.10	0.07	40	0.12	0.20	0.08	40	0.18	0.36	0.13
	60	0.08	0.10	0.06	60	0.16	0.26	0.11	60	0.25	0.46	0.18
	80	0.12	0.11	0.07	80	0.24	0.30	0.12	80	0.35	0.57	0.23
$N_d(0,\Sigma_{0.5})$	40	0.39	0.24	0.19	20	0.41	0.43	0.24	20	0.56	0.67	0.40
	60	0.62	0.31	0.31	40	0.76	0.63	0.55	40	0.91	0.88	0.86
	80	0.79	0.38	0.50	60	0.93	0.77	0.83	60	0.99	0.95	0.99
$T_d(5,0,\Sigma_{0.2})$	60	0.11	0.10	0.07	60	0.24	0.25	0.13	60	0.32	0.45	0.21
	80	0.18	0.11	0.10	80	0.28	0.31	0.14	80	0.44	0.56	0.28
$T_d(5,0,\Sigma_{0.5})$	40	0.45	0.23	0.18	20	0.40	0.39	0.23	10	0.31	0.47	0.22
	60	0.66	0.30	0.30	40	0.80	0.63	0.53	20	0.63	0.66	0.44
	80	0.83	0.38	0.49	60	0.96	0.78	0.83	40	0.94	0.86	0.85
$T_d(1,0,\Sigma_{0.2})$	40	0.28	0.09	0.09	40	0.52	0.20	0.20	40	0.67	0.36	0.32
	60	0.45	0.11	0.11	60	0.75	0.23	0.24	60	0.90	0.45	0.44
	80	0.64	0.12	0.13	80	0.90	0.28	0.29	80	0.98	0.56	0.58
$T_d(1,0,\Sigma_{0.5})$	20	0.34	0.16	0.13	10	0.29	0.26	0.17	10	0.42	0.42	0.29
	40	0.70	0.23	0.20	20	0.62	0.39	0.28	20	0.84	0.60	0.51
	60	0.90	0.29	0.32	40	0.97	0.58	0.58	40	0.99	0.84	0.87
$K_d(2,0,\Sigma_{0.2})$	40	0.52	0.17	0.44	20	0.36	0.21	0.43	20	0.40	0.34	0.52
	60	0.72	0.24	0.60	40	0.70	0.38	0.73	40	0.74	0.54	0.82
	80	0.89	0.37	0.77	60	0.90	0.55	0.92	60	0.92	0.71	0.96
$K_d(2,0,\Sigma_{0.5})$	20	0.39	0.15	0.27	10	0.33	0.27	0.29	10	0.41	0.46	0.44
	40	0.82	0.30	0.59	20	0.70	0.46	0.61	20	0.84	0.68	0.83
	60	0.97	0.45	0.85	40	0.99	0.72	0.97	40	1.00	0.92	1.00
$K_d(1,0,\Sigma_{0.2})$	40	0.32	0.18	0.50	20	0.37	0.27	0.80	10	0.28	0.27	0.84
	60	0.47	0.23	0.65	40	0.73	0.46	0.98	20	0.62	0.42	0.98
	80	0.59	0.28	0.78	60	0.91	0.66	1.00	40	0.95	0.69	1.00
$K_d(1,0,\Sigma_{0.5})$	40	0.62	0.34	0.57	10	0.37	0.40	0.51	10	0.54	0.58	0.82
	60	0.84	0.48	0.77	20	0.68	0.58	0.82	20	0.86	0.79	0.98
	80	0.94	0.60	0.90	40	0.94	0.82	0.99	40	1.00	0.96	1.00
$P_d(0.5,0,\Sigma_{0.2})$	40	0.10	0.09	0.05	40	0.17	0.23	0.09	40	0.21	0.38	0.12
	60	0.13	0.10	0.06	60	0.22	0.28	0.11	60	0.30	0.48	0.18
	80	0.17	0.12	0.07	80	0.29	0.32	0.14	80	0.43	0.59	0.27
$P_d(0.5,0,\Sigma_{0.5})$	40	0.39	0.25	0.19	20	0.37	0.45	0.23	10	0.26	0.50	0.18
	60	0.61	0.33	0.31	40	0.74	0.65	0.55	20	0.54	0.68	0.38
	80	0.80	0.40	0.50	60	0.93	0.80	0.85	40	0.91	0.89	0.84
$L_d(1.0)$	40	0.47	0.24	0.19	20	0.42	0.42	0.24	10	0.31	0.48	0.21
	60	0.70	0.33	0.32	40	0.83	0.61	0.57	20	0.63	0.67	0.43
	80	0.88	0.40	0.50	60	0.97	0.77	0.86	40	0.93	0.87	0.87
$L_d(0.2)$	10	0.57	0.24	0.18	10	0.95	0.57	0.47	10	0.99	0.78	0.85
	20	0.98	0.37	0.40	20	1.00	0.78	0.99	20	1.00	0.94	1.00

Table 4: Empirical power at level 0.05 for multivariate alternatives with $d=2,3,4$

	$d=2$				$d=3$				$d=4$			
X-Distribution	n	I_n^2	W_n^2	T_n	n	I_n^2	W_n^2	T_n	n	I_n^2	W_n^2	T_n
$AL_d(0,\Sigma_{0.2})$	40	0.13	0.09	0.07	40	0.28	0.20	0.12	40	0.43	0.36	0.22
	60	0.21	0.10	0.09	60	0.44	0.24	0.16	60	0.63	0.45	0.29
	80	0.30	0.11	0.09	80	0.59	0.29	0.19	80	0.80	0.56	0.39
$AL_d(0,\Sigma_{0.5})$	40	0.52	0.23	0.20	20	0.48	0.42	0.25	10	0.34	0.44	0.22
	60	0.74	0.30	0.29	40	0.88	0.61	0.55	20	0.72	0.65	0.44
	80	0.89	0.37	0.47	60	0.98	0.76	0.81	40	0.97	0.86	0.83
$AL_d(1,\Sigma_{0.2})$	40	0.66	0.28	0.24	20	0.62	0.48	0.31	10	0.44	0.51	0.27
	60	0.87	0.39	0.42	40	0.96	0.71	0.74	20	0.81	0.75	0.54
	80	0.97	0.45	0.65	60	1.00	0.87	0.94	40	0.99	0.94	0.95
$AL_d(1,\Sigma_{0.5})$	20	0.55	0.26	0.20	10	0.51	0.42	0.23	10	0.72	0.67	0.46
	40	0.94	0.40	0.54	20	0.90	0.62	0.59	20	0.98	0.87	0.89
	60	1.00	0.55	0.87	40	1.00	0.86	0.99	40	1.00	0.98	1.00

Table 4 (cont.): Empirical power at level 0.05 for multivariate alternatives with $d = 2, 3, 4$

$m = 0.5$, on the logistic distribution with $\alpha = 0.2, 1.0$, and on the asymmetric Laplace distribution with $\mu = 0$ (symmetric Laplace distribution).

The empirical power results show that the BR test present excellent comparative properties for bivariate observations. It is clearly more powerful than tests W_n^2 and T_n for the considered set of alternatives. For data dimensions $d = 3$ and $d = 4$, the BR test reveals good performance and, in particular, it is never the worse of the considered test procedures for any one of the considered distributions. For large data dimensions the BR test does not present the same good performance. Although none of the considered tests is uniformly the best test over the considered set of alternative distributions, it seems that the Cramér-von Mises test obtain the best overall results for $d \geq 5$.

References

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Sorensen, D. (1999). *LAPACK users' guide.* Third edition. SIAM.

Anderson, T.W., Darling, D.A. (1954). A test of goodness of fit. *J. Amer. Statist. Assoc.* 49, 765–769.

Anderson, N.H., Hall, P., Titterington, D.M. (1994). Two-sample test statistics for measuring discrepancies between two multivariate probability density functions using kernel-based density estimates. *J. Multivariate Anal.* 50, 41–54.

	$d=5$				$d=7$				$d=10$			
X-Distribution	n	I_n^2	W_n^2	T_n	n	I_n^2	W_n^2	T_n	n	I_n^2	W_n^2	T_n
$N_d(0,\Sigma_{0.2})$	40	0.24	0.58	0.18	20	0.17	0.57	0.13	20	0.23	0.75	0.22
	60	0.38	0.69	0.26	40	0.34	0.78	0.27	40	0.50	0.93	0.46
	80	0.47	0.79	0.35	60	0.52	0.91	0.44	60	0.71	0.98	0.69
$N_d(0,\Sigma_{0.5})$	10	0.37	0.62	0.28	10	0.49	0.79	0.47	10	0.64	0.89	0.69
	20	0.69	0.83	0.58	20	0.87	0.94	0.85	20	0.96	0.98	0.98
$T_d(5,0,\Sigma_{0.2})$	40	0.28	0.56	0.22	40	0.43	0.77	0.39	20	0.33	0.70	0.40
	60	0.46	0.68	0.32	60	0.67	0.89	0.59	40	0.65	0.91	0.65
	80	0.62	0.76	0.45	80	0.85	0.94	0.75	60	0.86	0.97	0.84
$T_d(5,0,\Sigma_{0.5})$	10	0.38	0.59	0.31	10	0.51	0.77	0.49	10	0.71	0.88	0.76
	20	0.73	0.82	0.59	20	0.90	0.93	0.87	20	0.98	0.98	0.98
$T_d(1,\Sigma_{0.2})$	20	0.43	0.34	0.31	10	0.32	0.35	0.39	10	0.48	0.42	0.60
	40	0.80	0.54	0.50	20	0.61	0.49	0.52	20	0.81	0.61	0.79
	60	0.96	0.67	0.67	40	0.94	0.72	0.78	40	0.99	0.86	0.98
$T_d(1,\Sigma_{0.5})$	10	0.56	0.58	0.41	10	0.69	0.73	0.63	10	0.83	0.83	0.84
	20	0.91	0.78	0.70	20	0.97	0.91	0.91	20	1.00	0.96	0.99
$K_d(2,0,\Sigma_{0.2})$	20	0.36	0.40	0.51	20	0.29	0.58	0.41	10	0.13	0.54	0.20
	40	0.74	0.68	0.84	40	0.61	0.82	0.76	20	0.20	0.74	0.31
	60	0.93	0.85	0.97	60	0.85	0.93	0.95	40	0.43	0.93	0.65
$K_d(2,0,\Sigma_{0.5})$	10	0.47	0.60	0.53	10	0.53	0.78	0.61	10	0.59	0.88	0.68
	20	0.89	0.83	0.91	20	0.94	0.94	0.96	20	0.93	0.98	0.97
$K_d(1,0,\Sigma_{0.2})$	10	0.36	0.36	0.97	10	0.45	0.48	1.00	10	0.85	0.49	1.00
	20	0.80	0.50	1.00	20	0.98	0.64	1.00	20	1.00	0.67	1.00
$K_d(1,0,\Sigma_{0.5})$	10	0.65	0.73	0.95	10	0.74	0.82	1.00	10	0.89	0.87	1.00
	20	0.95	0.89	1.00	20	0.99	0.97	1.00	20	1.00	0.98	1.00
$P_d(0.5,0,\Sigma_{0.2})$	40	0.27	0.60	0.17	40	0.36	0.81	0.27	20	0.24	0.76	0.19
	60	0.41	0.71	0.25	60	0.55	0.92	0.45	40	0.48	0.94	0.41
	80	0.53	0.80	0.37	80	0.73	0.96	0.63	60	0.72	0.98	0.67
$P_d(0.5,0,\Sigma_{0.5})$	10	0.32	0.62	0.24	10	0.46	0.80	0.40	10	0.63	0.88	0.65
	20	0.67	0.83	0.55	20	0.85	0.96	0.84	20	0.95	0.99	0.99
$L_d(1.0)$	10	0.38	0.60	0.29	10	0.54	0.77	0.50	10	0.69	0.85	0.73
	20	0.75	0.81	0.63	20	0.88	0.92	0.88	20	0.97	0.98	0.98
$L_d(0.2)$	10	1.00	0.88	0.99	10	1.00	0.95	1.00	10	1.00	0.98	1.00
$AL_d(0,\Sigma_{0.2})$	40	0.55	0.54	0.31	20	0.43	0.52	0.36	10	0.34	0.49	0.43
	60	0.80	0.69	0.46	40	0.77	0.76	0.58	20	0.59	0.69	0.57
	80	0.91	0.77	0.58	60	0.94	0.89	0.76	40	0.92	0.90	0.85
$AL_d(0,\Sigma_{0.5})$	10	0.45	0.59	0.35	10	0.60	0.74	0.55	10	0.77	0.85	0.78
	20	0.81	0.78	0.64	20	0.94	0.92	0.88	20	0.98	0.97	0.98
$AL_d(1,\Sigma_{0.2})$	10	0.54	0.68	0.38	10	0.70	0.85	0.61	10	0.83	0.93	0.83
	20	0.89	0.89	0.65	20	0.97	0.98	0.95	20	1.00	0.99	1.00
$AL_d(1,\Sigma_{0.5})$	10	0.83	0.80	0.66	10	0.93	0.91	0.89	10	0.98	0.96	0.99
	20	0.99	0.96	0.98	20	1.00	0.99	1.00	20	1.00	1.00	1.00

Table 5: Empirical power at level 0.05 for multivariate alternatives with $d=5,7,10$

Bickel, P.J., Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* 1, 1071–1095.

Eubank, R.L., LaRiccia, V.N. (1992). Asymptotic comparison of Crámer-von Mises and nonparametric function estimation techniques for testing goodness-of-fit. *Ann. Statist.* 20, 2071–2086.

Fan, Y. (1994). Testing the goodness of fit of a parametric density function by kernel method. *Econometric Theory* 10, 316–356.

Fan, Y. (1998). Goodness-of-fit tests based on kernel density estimators with fixed smoothing parameters. *Econometric Theory* 14, 604–621.

Fang, H.-B., Fang, K.-T., Kotz, S. (2002). The meta-elliptical distributions with given marginals. *J. Multivariate Anal.* 82, 1–16.

Gouriéroux, C., Tenreiro, C. (2001). Local power properties of kernel based goodness of fit tests. *J. Multivariate Anal.* 78, 161–190.

Gregory, G.G. (1977). Large sample theory for U-statistics and tests of fit. *Ann. Statist.* 5, 110–123.

Henze, N. (1997). Extreme smoothing and testing for multivariate normality. *Statis. Probab. Lett.* 35, 203–213.

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

Henze, N., Zirkler, B. (1990). A class of invariante consistent tests for multivariate normality. *Comm. Stat. Theory Methods* 19, 3595–3617.

Johnson, M.E. (1987). *Multivariate Statistical Simulation.* Wiley.

Kim, J.-T. (2000). An order selection criterion for testing goodness of fit. *J. Amer. Statist. Assoc.* 95, 829–835.

Koroljuk, V.S., Borovskich, Yu.V. (1989). *Theory of U-Statistics.* Kluwer.

Kotz, S., Kozubowski, T.J., Podgórski, K. (2000). An asymmetric multivariate Laplace distribution. Technical Report No. 367, Department of Statistics and Applied Probability, University of California at Santa Barbara.

Liang, J.-J., Fang, K-T., Hickernell, F.J., Li, R. (2001). Testing multivariate uniformity and its applications. *Math. Comp.* 70, 337–355.

Madras, N. (2002). *Lectures on Monte Carlo Methods.* AMS.

Marhuenda, Y., Morales, D., Pardo, M.C. (2005). A comparison of uniformity tests. *Statistics* 39, 315–328.

Miller, F.L.jun., Quesenberry, C.P. (1979). Power studies of tests for uniformity. II. *Commun. Stat., Simulation Comput. B* 8, 271–290.

Niederreiter, H. (1992). *Random numbers generation and quasi-Monte Carlo methods.* CBMS-NSF Regional Conference Series in Applied Mathematics, 63. SIAM.

Nikitin, Y. (1995). *Asymptotic Efficiency of Nonparametric Tests.* Cambridge University Press.

Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* 33, 1065–1076.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Statist.* 23, 470–472.

Rosenblatt, M. (1956). Remarks on some non-parametric estimates of a density function. *Ann. Math. Statist.* 27, 832–837.

Shorack, G.R., Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics.* Wiley.

Stephens, M.A. (1974). EDF statistics for goodness of fit and some comparisons. *J. Amer. Statist. Assoc.* 69, 730–737.

Stephens, M.A. (1986). Tests for the uniform distribution, In, D'Agostino, R.B., Stephens, M.A., eds. *Goodness-of-Fit Techniques*, 331–366. Marcel Dekker.

Tenreiro, C. (2005). On the role played by the fixed bandwidth in the Bickel-Rosenblatt goodness-of-fit test. *SORT* 29, 201–216.

Tenreiro, C. (2007). On the asymptotic behaviour of location-scale invariant Bickel-Rosenblatt tests. *J. Statist. Plann. Inference* 137, 103–116.

Watson, G.S. (1961). Goodness-of-fit tests on a circle. *Biometrika* 48, 109–114.

Carlos Tenreiro
Departamento de Matemática, Universidade de Coimbra, Apartado 3008, 3001–454 Coimbra, Portugal

E-mail address: tenreiro@mat.uc.pt
URL: http://www.mat.uc.pt/~tenreiro/