MATHEMATICS OF COMPUTATION
Volume 86, Number 304, March 2017, Pages 799–828
http://dx.doi.org/10.1090/mcom/3148
Article electronically published on October 20, 2016

STATISTICAL PROPERTIES OF b-ADIC DIAPHONIES

RAFFAELLO SERI

ABSTRACT. The aim of this paper is to derive the asymptotic statistical properties of a class of discrepancies on the unit hypercube called b-adic diaphonies. They have been introduced to evaluate the equidistribution of quasi-Monte Carlo sequences on the unit hypercube. We consider their properties when applied to a sample of independent and uniformly distributed random points. We show that the limiting distribution of the statistic is an infinite weighted sum of chi-squared random variables, whose weights can be explicitly characterized and computed. We also describe the rate of convergence of the finite-sample distribution to the asymptotic one and show that this is much faster than in the classical Berry-Esséen bound. Then, we consider in detail the approximation of the asymptotic distribution through two truncations of the original infinite weighted sum, and we provide explicit and tight bounds for the truncation error. Numerical results illustrate the findings of the paper, and an empirical example shows the relevance of the results in applications.

1. INTRODUCTION

The b-adic diaphony is a quantitative measure of the irregularity of the distribution of sequences in the d-dimensional unit cube $[0,1)^d$. It is similar to the classical diaphony (introduced in [55]; see also [37, Exercise 5.27, p. 162] or [14, Definition 1.29]), but replaces the trigonometric functions used in it with the b-adic or generalized Walsh (or Chrestenson-Levy) function system. The first instance of the b-adic diaphony was introduced in [33], using the Walsh function system in base 2. It was hence called dyadic diaphony. The general form of the b-adic diaphony was discussed in [28, 29], replacing the original Walsh function system with the generalized Walsh (or Chrestenson-Levy) function system. Generalized versions of the b-adic diaphony were introduced in [24, 25, 27, 32]. The dyadic and b-adic diaphony of special deterministic sequences has been studied in several papers [12, 21–23, 26, 28–31, 36, 42, 45, 52, 53].

The aim of this paper is to investigate instead the statistical properties of the b-adic diaphony of a sequence of N independent uniformly distributed random variables on the d-dimensional unit cube $[0,1)^d$. As the finite-sample statistical properties, i.e. the ones for finite N, are too difficult to derive, we consider the asymptotic properties for $N \to \infty$. The main asymptotic result concerns the derivation of an explicit representation for the asymptotic distribution of the scaled b-adic diaphony,

Received by the editor October 13, 2014 and, in revised form, August 16, 2015.

2010 *Mathematics Subject Classification.* Primary 65D30; Secondary 62E20, 62E17.

Key words and phrases. b-adic diaphony, equidistribution, approximation of distributions, quadratic forms in Gaussian random variables.

The author thanks Kiwi Cave Rafting, Waitomo Caves, New Zealand, for permission to use one of their pictures of the glowworm caves. Financial support through the COFIN Project 2010J3LZEN_006 is gratefully acknowledged. The usual disclaimer applies.

which lends itself to be used for computational purposes. As for other quantitative measures of irregularity of distribution (see, e.g., [4, 5, 38]), this representation involves an infinite weighted sum of chi-square random variables, whose weights can be characterized and computed. We complement this result with a bound on the distance between the finite-sample distribution and the asymptotic one. This result is analogous to the celebrated Berry-Esséen bound describing the convergence rate in the Central Limit Theorem for sums of independent and identically distributed (iid) random variables, but it shows that convergence is much faster for the b-adic diaphony, so that the asymptotic distribution is generally a very good approximation to the finite-sample one. These results allow us to provide a condition characterizing any sequence of point sets whose b-adic diaphony is smaller than the one of a sample of independent and uniformly distributed points with probability converging to 1 when $N \to \infty$.

Then we turn to the computation of the asymptotic distribution. The most natural method to compute this class of distributions is to truncate the infinite sum, thus getting a finite weighted sum of chi-squares. This distribution can be computed with one of the methods available in the literature (see [8, 9, 34, 44, 50], just to name a few). However, the error caused by the truncation of the infinite sequence of weights is not sufficiently investigated in the literature. In the present case, we compare theoretically and empirically two alternative methods of truncation and we provide some easily computable and effective bounds for the truncation error.

The paper is organized as follows. In Section 2, we provide a mathematical introduction to the generalized Walsh or Chrestenson-Levy function system, whose only purpose is to introduce the reader to the concepts that will be used to define the b-adic diaphony (for a more complete introduction, see [13, Appendix A]). The statistical asymptotic properties of this uniformity measure are then derived in Section 3, while the approximation of the asymptotic distribution is dealt with in Section 4. Section 5 contains an application to the equidistribution of glowworms in Waitomo caves, New Zealand, whose main aim is to provide an illustration of the statistical use of b-adic diaphonies and to warn the reader against some pitfalls in their application. Section 6 contains the conclusions, and Section 7 the proofs as well as some auxiliary results.

2. Mathematical preliminaries and definitions

First of all, we set some notation for the following. We will write $\mathbb{N}$ for the positive integers, $\mathbb{N}_0$ for the non-negative integers, $\mathbb{R}$ for the real numbers and $\mathbb{C}$ for the complex numbers. The symbol i denotes the imaginary unit. Vectors are indicated in bold type. Their components are in regular type with an index. As an example, $\mathbf{x} = (x_1, \ldots, x_d)$.

Consider an integer $b \geq 2$. Any non-negative integer k can be represented in base b as:

$$k = \kappa^{(a-1)} b^{a-1} + \cdots + \kappa^{(1)} b + \kappa^{(0)},$$

where $\kappa^{(i)} \in \{0, \ldots, b-1\}$ and $\kappa^{(a-1)} \neq 0$. Every real number $x \in [0, 1)$ has a b-adic representation $x = x^{(1)} b^{-1} + x^{(2)} b^{-2} + \ldots$ where $x^{(i)} \in \{0, \ldots, b-1\}$. If $x = ab^{-g}$, with a and g integers and $0 \leq a < b^g$, then x is said to be a *b-adic rational*. For any b-adic rational $x \neq 0$, there are two b-adic representations, one with the property that $x^{(i)} = 0$ for all i sufficiently large and the other one with the property that $x^{(i)} = b - 1$ for all i sufficiently large. The former is called a *regular*

b-adic representation, and, in the following, we will identify any b-adic rational x with its regular b-adic representation.

We define the *(generalized) Walsh function* ${}_b\mathrm{wal}_k : [0,1) \to \mathbb{C}$ as

$$ {}_b\mathrm{wal}_k\left(x\right) := \exp\left\{2\pi\mathrm{i}\frac{x^{(1)}\kappa^{(0)} + \cdots + x^{(a)}\kappa^{(a-1)}}{b}\right\} $$

for $x = x^{(1)}b^{-1} + x^{(2)}b^{-2} + \ldots$ with $x^{(i)} \in \{0, \ldots, b-1\}$. When the dimension $d \geq 2$, consider $\mathbf{x} = (x_1, \ldots, x_d) \in [0,1)^d$ and $\mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{N}_0^d$. Then, we define ${}_b\mathrm{wal}_{\mathbf{k}} = {}_b\mathrm{wal}_{k_1,\ldots,k_d} : [0,1)^d \to \mathbb{C}$ as

$$ {}_b\mathrm{wal}_{\mathbf{k}}\left(\mathbf{x}\right) = {}_b\mathrm{wal}_{k_1,\ldots,k_d}\left(x_1, \ldots, x_d\right) := \prod_{j=1}^{d} {}_b\mathrm{wal}_{k_j}\left(x_j\right). $$

For any integer $d \geq 1$, the system $\left\{{}_b\mathrm{wal}_{\mathbf{k}}, \mathbf{k} \in \mathbb{N}_0^d\right\}$ is a complete orthonormal system in $L_2\left([0,1)^d\right)$ (see Proposition A.10 in [13]).

The b-adic diaphony is defined as follows.

Definition 2.1. Let $b \geq 2$ be an integer. The *b-adic diaphony* of the point set $\mathcal{P}_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset [0,1)^d$ is defined as

$$ F_{b,N}\left(\mathcal{P}_N\right) := \left(\frac{\sum_{\mathbf{k}\in\mathbb{N}_0^d\setminus\{\mathbf{0}\}} r_b\left(\mathbf{k}\right)\left|\frac{1}{N}\sum_{h=1}^{N} {}_b\mathrm{wal}_{\mathbf{k}}\left(\mathbf{x}_h\right)\right|^2}{(1+b)^d - 1}\right)^{\frac{1}{2}} $$

where for $\mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{N}_0^d$, $r_b\left(\mathbf{k}\right) := \prod_{j=1}^{d} r_b\left(k_j\right)$ and for $k \in \mathbb{N}_0$:

$$ r_b\left(k\right) := \begin{cases} 1 & \text{if } k = 0, \\ b^{-2a} & \text{if } b^a \leq k < b^{a+1} \text{ where } a \in \mathbb{N}_0. \end{cases} $$

Remark 2.2. (i) If $b = 2$, we speak of *dyadic diaphony.*

(ii) Note that $0 \leq F_{b,N}\left(\mathcal{P}_N\right) \leq 1$ for all N.

We define for $\alpha, \beta \in \{0, 1, \ldots, b-1\}$, $\alpha \oplus \beta = \alpha + \beta \pmod{b}$ and

$$ \alpha \ominus \beta = \begin{cases} \alpha - \beta & \text{if } \alpha \geq \beta, \\ b + \alpha - \beta & \text{if } \alpha < \beta. \end{cases} $$

Suppose that $x, y \in [0,1)$ have the representations $x = \sum_{j=1}^{\infty} x^{(j)}b^{-j}$ and $y = \sum_{j=1}^{\infty} y^{(j)}b^{-j}$ in base b. We define the *b-adic sum* and *difference*:

$$ \begin{aligned} x \dot{+} y &= \sum_{j=1}^{\infty}\left(x^{(j)} \oplus y^{(j)}\right)b^{-j}, \\ x \dot{-} y &= \sum_{j=1}^{\infty}\left(x^{(j)} \ominus y^{(j)}\right)b^{-j}. \end{aligned} $$

For $\mathbf{x}, \mathbf{y} \in [0,1)^d$, $\mathbf{x} \dot{+} \mathbf{y} = (x_1 \dot{+} y_1, \ldots, x_d \dot{+} y_d)$ and $\mathbf{x} \dot{-} \mathbf{y} = \left(x_1 \dot{-} y_1, \ldots, x_d \dot{-} y_d\right)$. For an arbitrary real $x \in [0,1)$ with b-adic expansion $x = x^{(g+1)}b^{-g-1} + x^{(g+2)}b^{-g-2} + \ldots$ with $x^{(g+1)} \neq 0$ and $g \geq 0$ an integer, we have

$$ \lfloor \log_b x \rfloor = -g - 1. $$

It is clear that the b-adic diaphony of Definition 2.1 has the form of a V-statistic (see, e.g., [48, p. 174]):[1]

$$F_{b,N}^{2}\left(\mathcal{P}_{N}\right) := \frac{1}{N^{2}}\sum_{h=1}^{N}\sum_{\ell=1}^{N}\frac{\sum_{\mathbf{k}\in\mathbb{N}_{0}^{d}\setminus\{\mathbf{0}\}} r_{b}\left(\mathbf{k}\right) {}_{b}\mathrm{wal}_{\mathbf{k}}\left(\mathbf{x}_{h}\right)\overline{{}_{b}\mathrm{wal}_{\mathbf{k}}\left(\mathbf{x}_{\ell}\right)}}{\left(1+b\right)^{d}-1}. \tag{2.1}$$

However, this formula is not particularly manageable. A more usable formula has been obtained in [28, 29], and we will state it in the following.

Consider the following two conditions:

(C1) $x \dot{+} y$ is not a b-adic rational;
(C2) x and y are b-adic rationals.

Then, let

$$\varphi\left(x\right) = \begin{cases} \left(b+1\right) - \left(b+1\right) b^{1+\lfloor \log_{b} x\rfloor} & \text{if } x\in\left(0,1\right), \\ b+1 & \text{if } x=0, \end{cases}$$

and $\phi : [0,1)^{d} \to \mathbb{R}$:

$$\phi\left(\mathbf{x}\right) = -1 + \prod_{j=1}^{d}\varphi\left(x_{j}\right)$$

for $\mathbf{x} = (x_1, \ldots, x_d)$. Then, for every point set $\mathcal{P}_N$ with generic element $\mathbf{x}_k = (x_{k,1}, \ldots, x_{k,d})$, such that (C1) or (C2) is satisfied with x and y replaced by $x_{h,i}$ and $x_{\ell,j}$ for $1 \le h, \ell \le N$ and $1 \le i, j \le d$, from [28, 29] we have

$$F_{b,N}^{2}\left(\mathcal{P}_{N}\right) := \frac{1}{\left(1+b\right)^{d}-1}\frac{1}{N^{2}}\sum_{h=1}^{N}\sum_{\ell=1}^{N}\phi\left(\mathbf{x}_{h}\dot{-}\mathbf{x}_{\ell}\right). \tag{2.2}$$

Remark 2.3. Note that $\frac{\phi(\mathbf{0})}{(1+b)^{d}-1} = 1$, so that the elements with $h = \ell$ contribute to $F_{b,N}^{2}\left(\mathcal{P}_{N}\right)$ with the constant value $\frac{1}{N}$, independently of the sequence.

3. Asymptotic properties

We anticipate some of the results of the paper. As stated above, the b-adic diaphony is a V-statistic. We will show in Theorem 3.1 that it is degenerate and its asymptotic distribution is given by an infinite weighted sum of chi-squared random variables. The reader well acquainted with the literature on degenerate V-statistics will recognize that, since the function system $\mathcal{W}(b)$ is orthonormal, the weights appearing in the infinite sum of chi-squares are the coefficients $\left\{\frac{r_b(\mathbf{k})}{(1+b)^{d}-1}, \mathbf{k}\in\mathbb{N}_0^d\setminus\{\mathbf{0}\}\right\}$. It may seem that a reasonable way to approximate the asymptotic distribution of the b-adic diaphony is to replace the infinite sum with a finite sum whose weights are $\left\{\frac{r_b(\mathbf{k})}{(1+b)^{d}-1}, \mathbf{k}\in\{0,1,\ldots,K\}^{d}\setminus\{\mathbf{0}\}\right\}$. However, as we will see below (see the computational results after Theorem 4.1), this approximation has very poor properties. A much better approximation, whose error is discussed in detail in Theorem 4.1, can be obtained grouping the chi-squared random variables according to the associated weight and truncating the infinite sum.

[1] Here and in the following, we will refer to $F_{b,N}^{2}$, and not to $F_{b,N}$, as the b-adic diaphony.

This requires formulas for the number of times each weight appears in the infinite sum. We provide this result in Theorem 3.1, along with some other relevant but more straightforward facts concerning the asymptotic distribution of the b-adic diaphony.

Theorem 3.1. *Let $\mathcal{P}_N$ be a sample of iid uniform random variables on $[0,1)^d$.*

(i) *When $N \to \infty$, $F^2_{b,N}(\mathcal{P}_N) \xrightarrow{\text{as}} 0$.*

(ii) *When $N \to \infty$,*

$$\limsup_{N\to\infty} \frac{NF^2_{b,N}(\mathcal{P}_N)}{\ln\ln N} = \frac{2}{(1+b)^d - 1}, \quad \text{as.}$$

If b is a prime number, when $N \to \infty$,

$$\liminf_{N\to\infty} \frac{N^2 F^2_{b,N}(\mathcal{P}_N)}{(\ln N)^{d-1}} \geq C_{d,b} > 0, \quad \text{as}$$

where $C_{d,b}$ is a constant depending on d and b but not on the particular sequence $\mathcal{P}_N$.

(iii) *When $N \to \infty$, the following weak convergence result holds:*

$$NF^2_{b,N}(\mathcal{P}_N) \xrightarrow{\mathcal{D}} \sum_{a=0}^{\infty} \lambda_a \chi^2_a(N_a),$$

where $\lambda_a = \frac{1}{(1+b)^d-1} b^{-2a}$ for any $a \geq 0$,

$$N_a = \begin{cases} b^d - 1 & \text{for } a = 0, \\ \sum_{\ell=1}^{\min\{d,a\}} \binom{a-1}{\ell-1}\binom{d}{\ell} b^{a+d-\ell}(b-1)^\ell & \text{for } a > 0, \end{cases}$$

$\chi^2(k)$ denotes a χ^2 random variable with k degrees of freedom, and $\left\{\chi^2_a(N_a); a \in \mathbb{N}_0\right\}$ is a sequence of independent χ^2 random variables.

(iv) *The following uniform bound holds:*

$$\left\| \mathbb{P}\left\{NF^2_{b,N}(\mathcal{P}_N) \leq x\right\} - \mathbb{P}\left\{\sum_{a=0}^{\infty} \lambda_a \chi^2_a(N_a) \leq x\right\} \right\|_\infty \leq \frac{e^{c_0 + c_1 (b+1)^{\frac{d}{2}}} (1.256\,(b+1))^d}{N}$$

where $c_0 > 0$ and $c_1 > 0$ are absolute constants.

Remark 3.2. (a) The values of N_a for $0 \leq a \leq 10$, $1 \leq d \leq 3$ and $2 \leq b \leq 4$ are reported in Table 1. This shows that the increase in N_a with a is much steeper for large d and b.

(b) From [1] (see also [11]), it is easily seen that the result in (ii) provides a worst-case error bound on the integration through the Monte Carlo method.

(c) According to the bound in (iv), the distance between the finite-sample and the asymptotic distributions decreases as N^{-1} and increases as a (more complex) function of b. As we will see below, the decrease in N is of the right order, while the order of the increase with b is probably overly pessimistic.

(d) Suppose that $\mathcal{P}_N$ is a sample of iid not necessarily uniform random variables on $[0,1)^d$. Then, adapting Theorem 4.4 in [17], it is possible to show that, for any fixed $x > 0$, $\lim_{N\to\infty} \mathbb{P}\left\{NF^2_{b,N}(\mathcal{P}_N) > x\right\} = 1$ if and only if the distribution is not uniform.

TABLE 1. Values of N_a for some choices of d and b.

$a\backslash(d,b)$	$(1,2)$	$(1,3)$	$(1,4)$	$(2,2)$	$(2,3)$	$(2,4)$	$(3,2)$	$(3,3)$	$(3,4)$
0	1	2	3	3	8	15	7	26	63
1	2	6	12	8	36	96	24	162	576
2	4	18	48	20	144	528	72	810	4032
3	8	54	192	48	540	2688	200	3618	24768
4	16	162	768	112	1944	13056	528	15066	140544
5	32	486	3072	256	6804	61440	1344	59778	755712
6	64	1458	12288	576	23328	282624	3328	228906	3907584
7	128	4374	49152	1280	78732	1277952	8064	852930	19611648
8	256	13122	196608	2816	262440	5701632	19200	3109914	96141312
9	512	39366	7864320	6144	866052	25165824	45056	11140578	462422016
10	1024	118098	3145728	13312	2834352	110100480	104448	39326634	2189426688

In Figures 1, 2, 3, 4, 5, 6, 7 and 8, we show the difference between the finite-sample and the asymptotic distributions, i.e. $\mathbb{P}\left\{NF_{b,N}^2(\mathcal{P}_N)\leq x\right\}-\mathbb{P}\left\{\sum_{a=0}^{\infty}\lambda_a\chi_a^2(N_a)\leq x\right\}$ as a function of x. The finite-sample distribution has been obtained as the empirical cumulative distribution function based on $18\cdot 10^7$ (for $N=25$), $12\cdot 10^7$ (for $N=50$) and $6\cdot 10^7$ (for $N=100$) replications. The Law of the Iterated Logarithm in [54, p. 268] provides an (asymptotic) upper bound on the error of the empirical cumulative distribution function, given by $1.550632\cdot 10^{-4}$ (for $N=25$), $1.1036536\cdot 10^{-4}$ (for $N=50$) and $0.9044465\cdot 10^{-4}$ (for $N=100$). When $d=1$, the difference exhibits ample waves that are probably due to the fact that the function φ attains only discrete values. When $d=2$, the oscillatory behavior is less pronounced and it disappears for $d=3$. Another interesting feature is the decrease of the maximal value of the difference with N. According to Theorem 3.1 (iv), the decrease should be as N^{-1}: this is indeed the case, even if the phenomenon appears to be more evident in the case $d=2$ since, when $d=1$, the oscillations mask the decrease. On the other hand, the increase of the distance with b predicted by Theorem 3.1 (iv) is real, but the bound seems too pessimistic.

Now we consider a sequence of quasi-random point sets $\mathcal{P}_N^\star$, for $N\in\mathbb{N}$, such that $\lim_{N\to\infty}NF_{b,N}^2(\mathcal{P}_N^\star)=0$ or $F_{b,N}(\mathcal{P}_N^\star)=o\left(N^{-\frac{1}{2}}\right)$. In particular, we show that this condition is necessary and sufficient to have $\lim_{N\to\infty}\mathbb{P}\left\{F_{b,N}(\mathcal{P}_N)\leq F_{b,N}(\mathcal{P}_N^\star)\right\}=0$. This means that the probability that the b-adic diaphony of independent and identically distributed points is lower than the b-adic diaphony of such a sequence becomes negligible as N increases.

The condition $\lim_{N\to\infty}NF_{b,N}^2(\mathcal{P}_N^\star)=0$ is respected by several quasi-random point sets and sequences, such as the Zaremba sequence (for the dyadic diaphony; see [26]), generalized Van der Corput sequences (see [28, 29]), (t,m,s)-nets in base 2 (for the dyadic diaphony; see [12]) and in base b (see [52]), generalized Zaremba nets (see [30, 31, 53]), digital $(0,s)$-sequences over $\mathbb{Z}_2$ (see [42]) and $\mathbb{Z}_b$ (see [21]), digital (t,s)-sequences over $\mathbb{Z}_2$ (see [36]), and digital $(\mathbf{T},s)$-sequences over $\mathbb{Z}_b$ (see [22]). Some numerical examples from which one can compute $NF_{b,N}^2(\mathcal{P}_N^\star)$ can be found in [40].

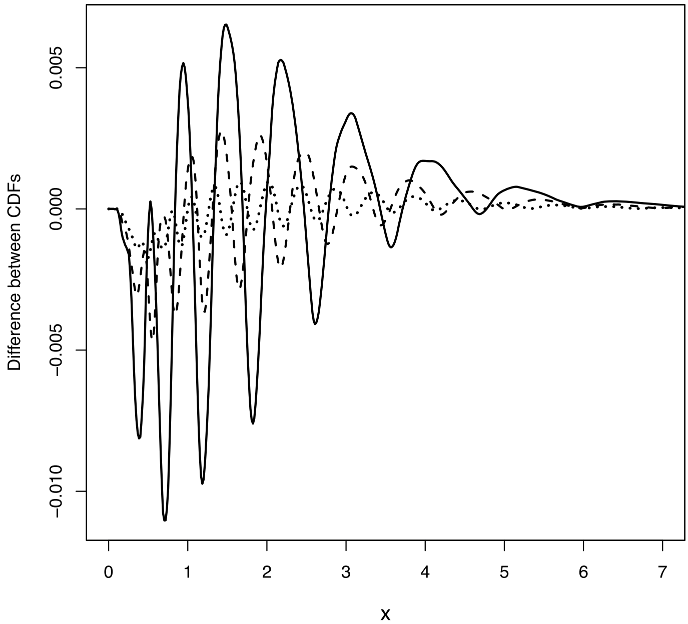

FIGURE 1. Deviation of the finite-sample distribution from the asymptotic one for $b = 2$, $d = 1$ and $N = 25$ (solid line), $N = 50$ (dashed line) and $N = 100$ (dotted line).

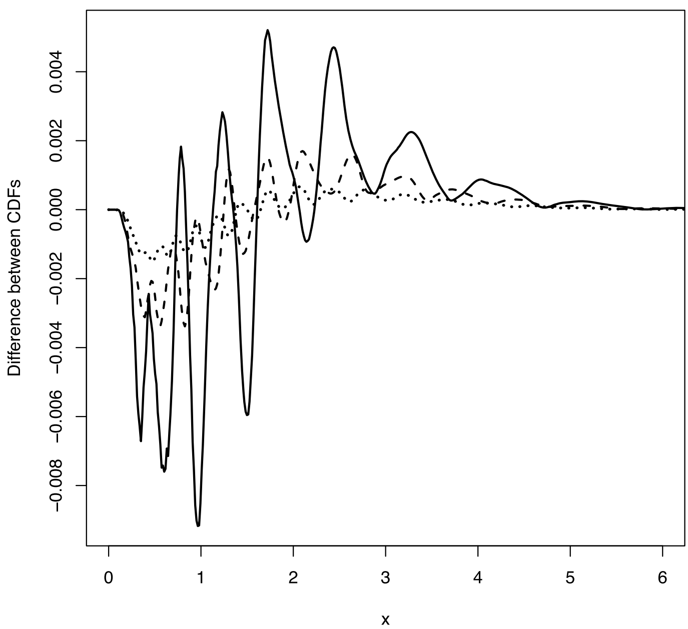

FIGURE 2. Deviation of the finite-sample distribution from the asymptotic one for $b = 3$, $d = 1$ and $N = 25$ (solid line), $N = 50$ (dashed line) and $N = 100$ (dotted line).

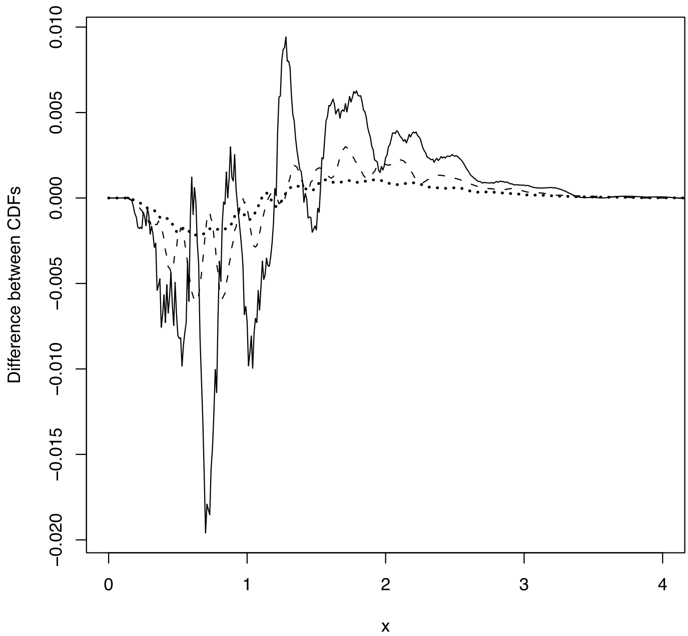

Figure 3. Deviation of the finite-sample distribution from the asymptotic one for $b = 5$, $d = 1$ and $N = 25$ (solid line), $N = 50$ (dashed line) and $N = 100$ (dotted line).

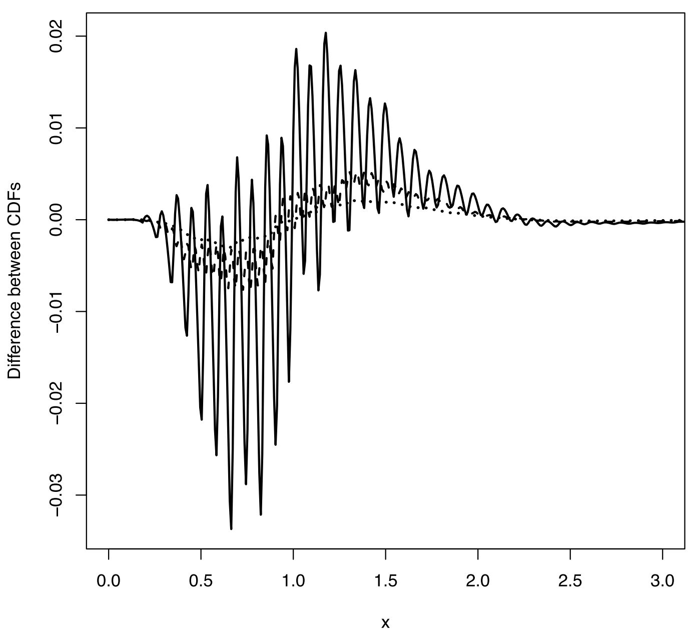

Figure 4. Deviation of the finite-sample distribution from the asymptotic one for $b = 10$, $d = 1$ and $N = 25$ (solid line), $N = 50$ (dashed line) and $N = 100$ (dotted line).

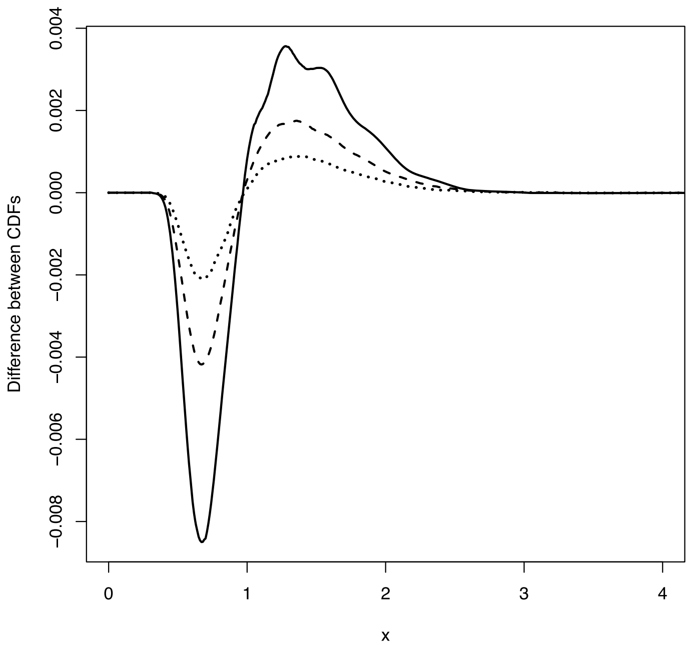

FIGURE 5. Deviation of the finite-sample distribution from the asymptotic one for $b = 2$, $d = 2$ and $N = 25$ (solid line), $N = 50$ (dashed line) and $N = 100$ (dotted line).

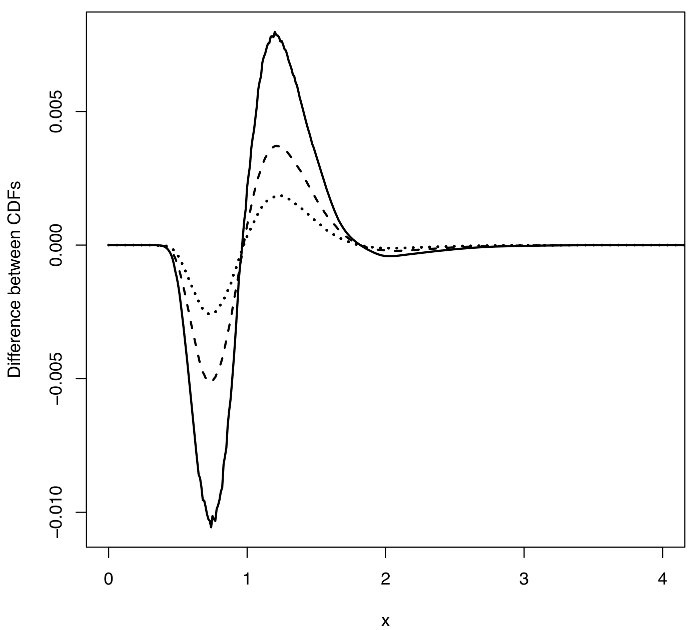

FIGURE 6. Deviation of the finite-sample distribution from the asymptotic one for $b = 3$, $d = 2$ and $N = 25$ (solid line), $N = 50$ (dashed line) and $N = 100$ (dotted line).

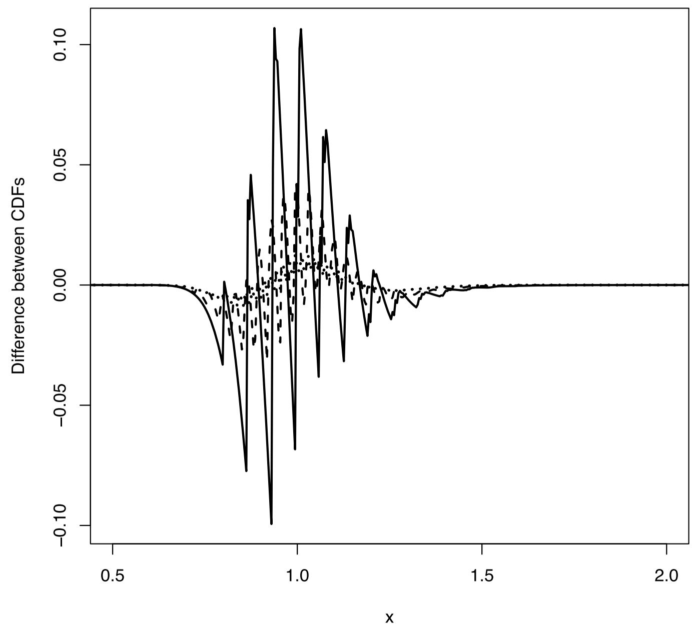

FIGURE 7. Deviation of the finite-sample distribution from the asymptotic one for $b = 10$, $d = 2$ and $N = 25$ (solid line), $N = 50$ (dashed line) and $N = 100$ (dotted line).

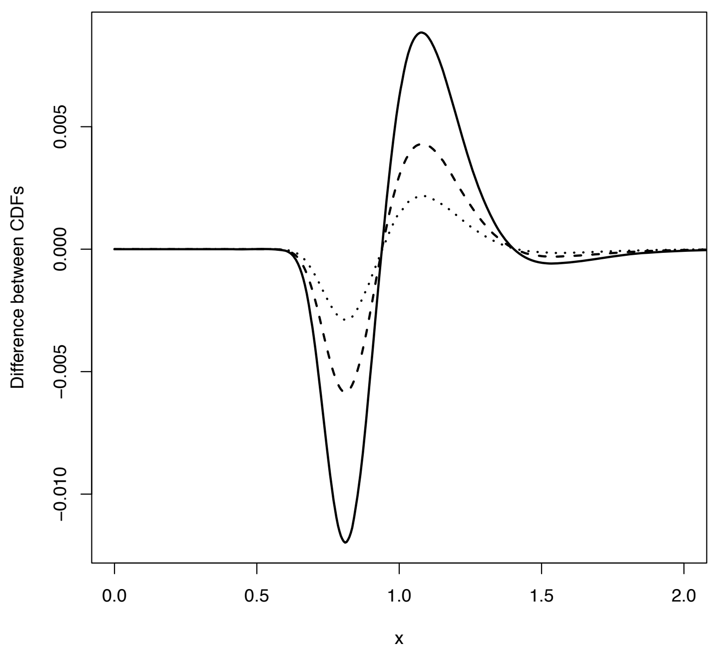

FIGURE 8. Deviation of the finite-sample distribution from the asymptotic one for $b = 2$, $d = 3$ and $N = 25$ (solid line), $N = 50$ (dashed line) and $N = 100$ (dotted line).

Theorem 3.3. *Consider a sequence of point sets* $\{\mathcal{P}_N^\star\}_{N\in\mathbb{N}}$. *The limit*

$$\lim_{N\to\infty} \mathbb{P}\left\{F_{b,N}\left(\mathcal{P}_N\right) \leq F_{b,N}\left(\mathcal{P}_N^\star\right)\right\} = 0$$

holds if and only if

$$\lim_{N\to\infty} NF_{b,N}^2\left(\mathcal{P}_N^\star\right) = 0.$$

Remark 3.4. (a) The result stated above uses convergence in probability. Theorem 1 in [7] (see also Theorem 3.1 (ii)) provides a lower bound holding almost surely, i.e. for almost any sequence of independent and uniformly distributed realizations, when b is prime.

(b) The theorem states what happens for finite fixed d. The behavior of the b-adic diaphony may change when the limit for $d \to \infty$ is taken before the one for $N \to \infty$ (see [12, Section 4]).

4. Approximation of the asymptotic distribution

Consider the following random variables:

$$\begin{aligned} X &= \sum_{a=0}^{\infty} \lambda_a \chi_a^2\left(N_a\right), \\ X_A &= \sum_{a=0}^{A} \lambda_a \chi_a^2\left(N_a\right), \\ X^\star &= X - \mathbb{E}X, \\ X_A^\star &= X_A - \mathbb{E}X_A. \end{aligned}$$

The random variable X_A represents a truncated version of X, suitable for computations, while $X^\star$ and $X_A^\star$ represent respectively two centered versions of the random variables X and X_A. Let $F_Y\left(y\right) := \mathbb{P}\left\{Y \leq y\right\}$ denote the cumulative distribution function of the random variable Y evaluated in y.

Some comments on these approximations are necessary. First of all, note that $X^\star$ is the asymptotic distribution of the U-statistic obtained from the b-adic diaphony removing the elements on the diagonal (i.e. with $h = \ell$ in (2.2)), namely, taking

$$\tilde{F}_{b,N}^2\left(\mathcal{P}_N\right) := \frac{2}{\left(1+b\right)^d - 1} \frac{1}{N^2} \sum_{h=1}^{N} \sum_{\ell=1}^{h-1} \phi\left(\mathbf{x}_h \dot{-} \mathbf{x}_\ell\right).$$

Lemma 7.1 guarantees that restricting the summation to elements with $\ell < h$ is legitimate. Since $\frac{\phi(\mathbf{0})}{(1+b)^d-1} = 1$, the equality $F_{b,N}^2\left(\mathcal{P}_N\right) = \frac{1}{N} + \tilde{F}_{b,N}^2\left(\mathcal{P}_N\right)$ holds, and removing the diagonal elements does not alter the asymptotic properties of the statistic. A second fact is that approximating $X^\star$ through $X_A^\star$ is equivalent to approximating X through $X_A - \mathbb{E}X_A + \mathbb{E}X = X_A + \left(1 - \sum_{a=0}^{A} \lambda_a N_a\right)$. Now, the computation of X_A can be easily performed using the algorithm described in [8] and ported in R from C in [3]; therefore also the computation of $X_A^\star$ is affordable.

The following theorem shows that a considerable gain in accuracy can be obtained using the truncated centered version of the statistic. Indeed, the Kolmogorov distance between the distribution of X and the one of its truncated version X_A decreases as $\left\|F_X - F_{X_A}\right\|_\infty = O\left(b^{-A}A^{d-1}\right)$ when $A \to \infty$, while the same distance for the centered random variables $X^\star$ and $X_A^\star$ decreases as

$\left\|F_{X^\star}-F_{X_A^\star}\right\|_\infty=O\left(b^{-3A}A^{d-1}\right)$. The two bounds show that the rate of decrease is faster when d is smaller and, especially, when b is larger. Moreover, centering yields a much better approximation as the base of the exponent A is b^{-1} in the uncentered case and b^{-3} in the centered case.

Theorem 4.1. *Let $\alpha\in\mathbb{N}_0$ and $\alpha^\star\in\mathbb{N}_0$ be two values of the index a such that $N_\alpha>2$ and $N_{\alpha^\star}>4$. We have*

$$\left\|F_X(x)-F_{X_A}(x)\right\|_\infty \leq \frac{\sqrt{2\sum_{a=A+1}^\infty\lambda_a^2N_a+\left[\sum_{a=A+1}^\infty\lambda_aN_a\right]^2}}{4\pi}\frac{B\left(\frac{1}{2},\frac{N_\alpha-2}{4}\right)}{\sqrt{\frac{\sum_{a=\alpha}^A N_a\lambda_a^2}{N_\alpha}}}$$

where $B(\cdot,\cdot)$ is the beta function, and

$$\left\|F_{X^\star}(x)-F_{X_A^\star}(x)\right\|_\infty \leq \frac{\sum_{a=A+1}^\infty\lambda_a^2N_a}{2\pi\left(\frac{\sum_{a=\alpha^\star}^A N_a\lambda_a^2}{N_{\alpha^\star}}\right)(N_{\alpha^\star}-4)}.$$

The asymptotic behavior of the bounds for large A is

$$\begin{aligned}\left\|F_X(x)-F_{X_A}(x)\right\|_\infty &\lesssim \frac{(b-1)^{d-1}}{(d-1)!\left((1+b)^d-1\right)}\frac{b^{-A}A^{d-1}B\left(\frac{1}{2},\frac{N_\alpha-2}{4}\right)}{4\pi\sqrt{\frac{\sum_{a=\alpha}^A N_a\lambda_a^2}{N_\alpha}}},\\ \left\|F_{X^\star}(x)-F_{X_A^\star}(x)\right\|_\infty &\lesssim \frac{(b-1)^d}{(d-1)!\left((1+b)^d-1\right)^2(b^3-1)}\frac{b^{-3A}A^{d-1}}{2\pi\left(\frac{\sum_{a=\alpha^\star}^A N_a\lambda_a^2}{N_{\alpha^\star}}\right)(N_{\alpha^\star}-4)}.\end{aligned}$$

Remark 4.2. (a) The computations in the following suggest that the bounds are tight.

(b) As concerns the choice of α and $\alpha^\star$, in the following computations we have found satisfying to take them as the smallest integers satisfying the inequalities $N_\alpha>2$ and $N_{\alpha^\star}>4$.

The effectiveness of the bounds presented in Theorem 4.1 is shown in Figures 9, 10, 11, 12 and 13. The figures report, on the ordinate, the uniform distance between the cdf of X and the cdf of the approximating random variable, be it X_A, $X^\star$ or $X_A^\star$, as a function of the progressive number of degrees of freedom used, on the abscissa (e.g., $\sum_{a=0}^A N_a$, that is the number of degrees of freedom used in X_A). Therefore, the black dots represent the couples $\left(\sum_{a=0}^A N_a,\left\|F_X-F_{X_A}\right\|_\infty\right)$, the black squares the couples $\left(\sum_{a=0}^A N_a,\left\|F_{X^\star}-F_{X_A^\star}\right\|_\infty\right)$; the corresponding symbols with grey background represent the upper bounds computed in Theorem 4.1. The black curve represents the maximal distance between the distribution of X and that of the approximation $\frac{1}{(1+b)^d-1}\sum_{\mathbf{k}\in\{0,1,\ldots,K\}^d\setminus\{\mathbf{0}\}}r_b(\mathbf{k})\chi^2(1)$. The number of degrees of freedom for this approximation is K^d-1. The curve is represented only for $d>1$, since when $d=1$ it would correspond to a curve joining the black dots. It is evident that the approximation is not very good, being worse than the centered and uncentered truncated sums. The centered truncated sum is by far better than the uncentered one. The bounds are quite good and tight, as each of them decreases as the true distance between the distribution of X and the corresponding approximation.

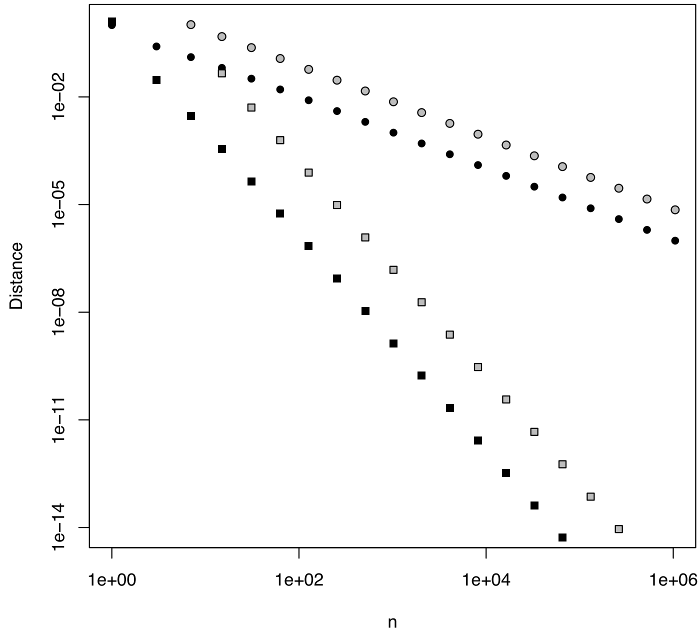

Figure 9. Distances between true and approximated asymptotic distributions for $b = 2$ and $d = 1$.

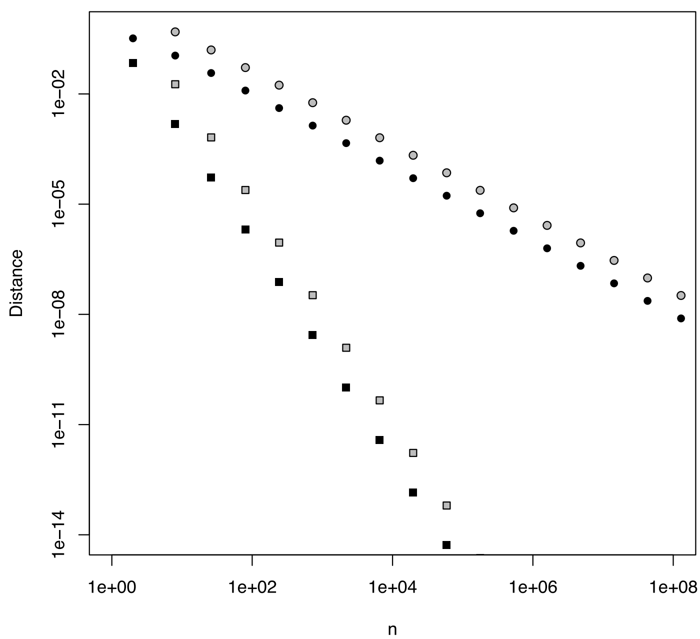

Figure 10. Distances between true and approximated asymptotic distributions for $b = 3$ and $d = 1$.

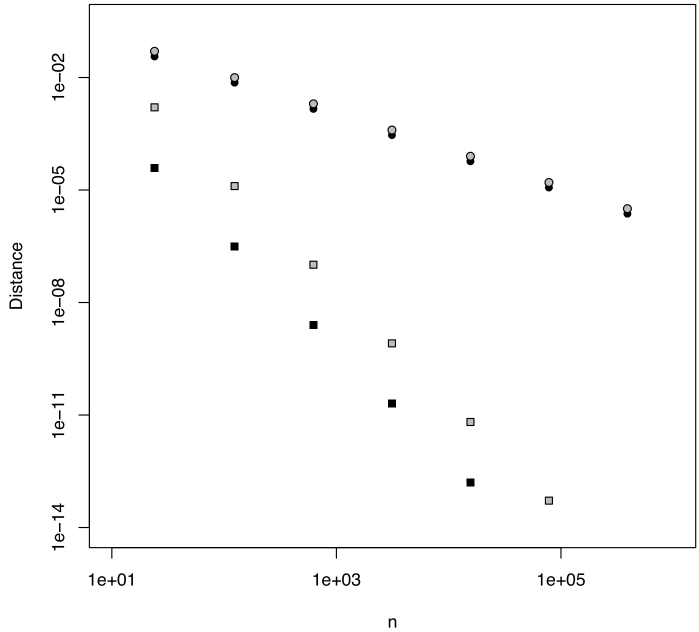

FIGURE 11. Distances between true and approximated asymptotic distributions for $b = 5$ and $d = 1$.

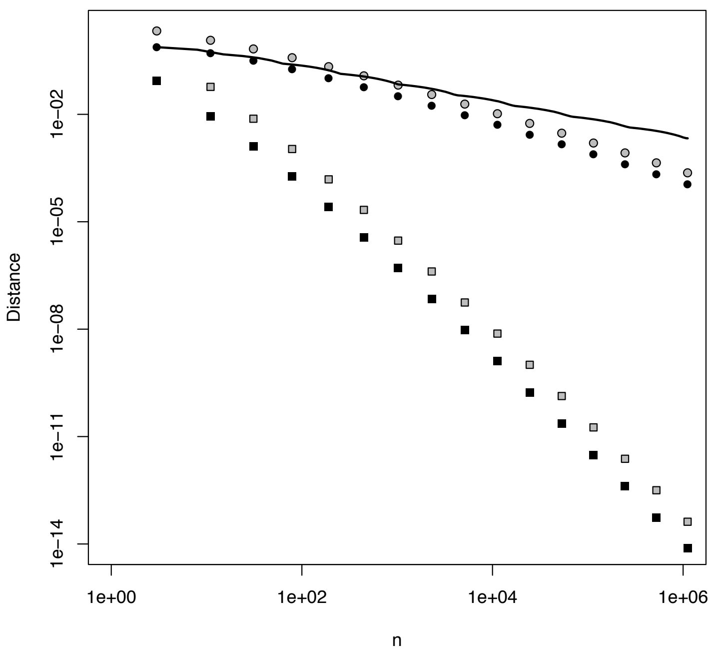

FIGURE 12. Distances between true and approximated asymptotic distributions for $b = 2$ and $d = 2$.

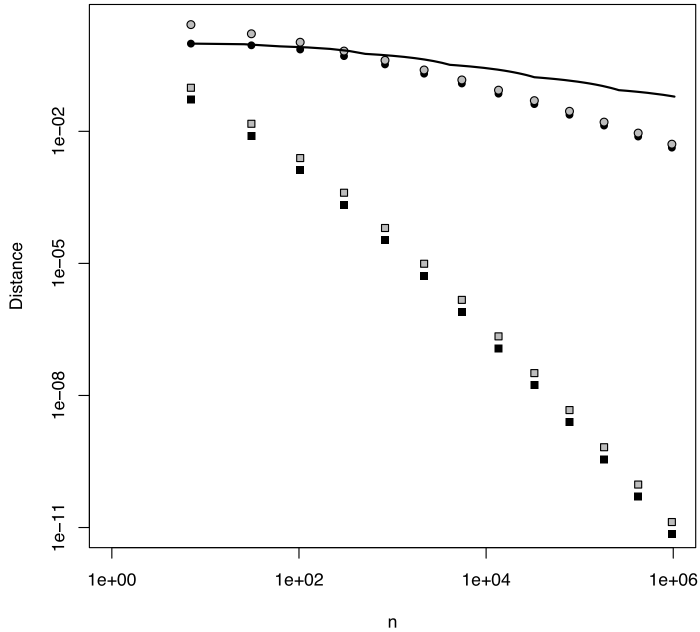

FIGURE 13. Distances between true and approximated asymptotic distributions for $b = 2$ and $d = 3$.

5. An application

In his essay "Glow, Big Glowworm" [18], Stephen Jay Gould described a visit to the Waitomo glowworm (*Arachnocampa luminosa*) caves in New Zealand. This led him to a consideration concerning the equidistribution of sequences of points:

> Here, in utter silence, you glide by boat into a spectacular underground planetarium, an amphitheater lit with thousands of green dots—each the illuminated rear end of a fly larva (not a worm at all). (I was dazzled by the effect because I found it so unlike the heavens. Stars are arrayed in the sky at random with respect to the earth's position. Hence, we view them as clumped into constellations. This may sound paradoxical, but my statement reflects a proper and unappreciated aspect of random distributions. Evenly spaced dots are well ordered for cause. Random arrays always include some clumping, just as we will flip several heads in a row quite often so long as we can make enough tosses—and our sky is not wanting for stars. The glowworms, on the other hand, are spaced more evenly because larvae compete with, and even eat, each other—and each constructs an exclusive territory. The glowworm grotto is an ordered heaven.) [18, pp. 13-14]

The American Nobel Laureate Ed Purcell, a colleague of Gould, performed a little experiment whose result was presented in the postscript of the book version of the essay [19, pp. 265-268].

> Into an array of square cells (144 units on the X-axis and 96 on the Y-axis for a total of 13,824 positions), Purcell placed either "stars"

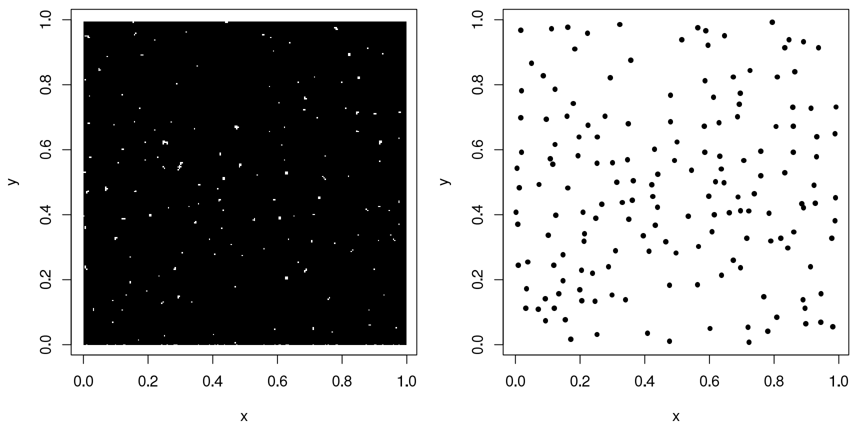

FIGURE 14. Square selections of a picture of the Waitomo caves (see the text for details), corresponding to the 'Real data' point set.

> or "worms" by the following rules of randomness and order [...]. In the stars option, squares are simply occupied at random (a random number generator spits out a figure between 1 and 13824 and the appropriate square is inked in). In the worms option, the same generator spits out a number, but the appropriate square is inked in only if it and all the surrounding squares are unoccupied (just as a worm sets up a zone of inhibition about itself). [19, pp. 267-268]

The two graphs are described in Gould's essay as rectangular with an aspect ratio of 1.5:1; however, in the printed version, the graphs appeared slightly distorted and with a different aspect ratio. In order to evaluate the equidistribution of the sequences, we have rescaled the two scatter plots to the original aspect ratio and we have cropped a square selection as large as possible of both, which is available from the author upon request.

Purcell commented on the appearance as follows:

> What interests me more in the random field of "stars" is the overimposing impression of "features" of one sort or another. It is hard to accept the fact that any perceived feature—be it string, clump, constellation, corridor, curved chain, lacuna—is a totally meaningless accident, having as its only cause the avidity for pattern of my eye and brain! Yet that is perfectly true in this case. [19, p. 268]

The example was also quoted in [43, Chapter 5].

> The one on the left, with the clumps, strands, voids, and filaments (and perhaps, depending on your obsessions, animals, nudes, or Virgin Marys) is the array that was plotted at random, like stars. The one on the right, which seems to be haphazard, is the array whose positions were nudged apart, like glowworms. [43, Chapter 5]

In Figure 14, we reproduce a square selection of a picture, kindly provided to us by Kiwi Cave Rafting, Waitomo Caves, New Zealand, and a scatter plot extracted from it.

In the following, we will use the three scatter plots, i.e. the two appearing in [19, pp. 266-267] (with the 'Stars' point set on the left and the 'Worms' point set on the right) and the one in Figure 14 (displaying the 'Real data' point set) to show the behavior of the b-adic diaphony in some different situations.

When applied to a sequence of iid random variables, the b-adic diaphony is able to discriminate between uniformly and non-uniformly distributed random variables. Indeed, Theorem 3.1 shows that if $\mathcal{P}_N$ is a sample of independent uniformly distributed random variables, $F_{b,N}^2(\mathcal{P}_N)$ converges (almost surely) to 0 (item (i)), and $NF_{b,N}^2(\mathcal{P}_N)$ converges in distribution (item (iii)). On the other hand, if $\mathcal{P}_N$ is a sample of independent non-uniformly distributed random variables, Remark 3.2 (d) shows that $NF_{b,N}^2(\mathcal{P}_N)$ diverges. Therefore, under independent sampling, large values of $NF_{b,N}^2(\mathcal{P}_N)$ and correspondingly small p-values witness a departure from uniformity. Therefore, one can use the p-value to perform a statistical test in the Fisher sense (see, e.g., [6, 39]).

The ability of the p-value to discriminate between uniformity and non-uniformity does not necessarily hold without independent sampling. If the observations constitute a quasi-Monte Carlo sequence for which $NF_{b,N}^2(\mathcal{P}_N)$ converges to 0 (see the examples in Section 3), Theorem 3.3 shows that the p-value converges to 1. When the observations are the realization of a repulsive (also called regular or inhibitive) point process, i.e. a point process in which nearby points tend to repel each other, $NF_{b,N}^2(\mathcal{P}_N)$ assumes in general smaller values than in the case of uniform iid points. This also implies that the typical realization of a repulsive point process is associated with a large p-value when the b-adic diaphony is used to test for uniformity.

Now, from a probabilistic point of view, the 'Stars' plot in [19, p. 266] displays a sequence of points independently and uniformly distributed in $[0,1)^2$, namely, the realization of a Poisson process on $[0,1)^2$, while the 'Worms' plot in [19, p. 267] displays the realization of a repulsive point process on $[0,1)^2$, in which repulsion among points leads to a more equidistributed point set. The right plot in Figure 14 shows a real point set that should be distributed according to the repulsive point process described above.

The results of the computations are displayed in Table 2. The first 6 rows contain the scaled values of the b-adic diaphonies $NF_{b,N}^2(\mathcal{P}_N)$, for $b \in \{2,3,4,5\}$, and of the corresponding p-values for the three datasets described above. The p-values have been computed according to the $X_A^\star$ approximation. The last two rows contain the value of A used in the computation of the p-values (the same A has been used for each value in the column) and the error bound in Theorem 4.1. Note that this is only the error associated with the replacement of the original random variable with a truncated version. This differs from the error associated with the computation of the probability of the truncated random variable (see Section VI in [49] for a case in which the computation error is very large), which we have taken equal to 10^{-10} but may be substantially smaller.

Concerning the use of the b-adic diaphony as a figure of merit, it is clear that the 'Worms' point set is more equidistributed than the 'Stars' one, as the b-adic diaphony is consistently smaller for the former than for the latter. This comparison is justified by the fact that the two datasets have almost the same size. If two point sets have different sample sizes and they do not respect the hypothesis of independent sampling from a uniform distribution, then one should not evaluate

TABLE 2. Results of the uniformity test on the three examples.

	N		$b = 2$	$b = 3$	$b = 4$	$b = 5$
'Stars'	744	stat	0.663012	0.6276486	0.8203803	1.086391
		p-value	0.8787808	0.9519859	0.7714008	0.3047332
'Worms'	708	stat	0.54345	0.3312276	0.4404123	0.416133
		p-value	0.9740758	0.9999997	0.9999297	0.999998
Real data	168	stat	0.8500737	0.9666432	0.8986758	0.8689998
		p-value	0.6196428	0.4879706	0.6349837	0.7283115
A			26	15	12	10
Error bound			$7.663 \cdot 10^{-24}$	$3.531 \cdot 10^{-23}$	$6.266 \cdot 10^{-24}$	$1.361 \cdot 10^{-23}$

the degree of uniformity of the two samples by comparing the values of the scaled b-adic diaphonies. The reason is that when $N \to \infty$, $NF_{b,N}^2 (\mathcal{P}_N^\star)$ converges to 0, and therefore smaller values of the scaled b-adic diaphony could simply be due to a larger sample size and not to a better equidistribution.

Now we turn to the interpretation of the p-values. Despite the procedure of sampling the 'Stars' dataset as described in [19] is not exactly compatible with the null hypothesis (because it is based on a discretization of the space), it constitutes a reasonable approximation, by far better than the other two. For the 'Stars' dataset the p-value can therefore be used to perform a test of uniformity that fails to reject the null hypothesis. For the other two datasets, the hypothesis of independence is clearly false, and the asymptotic distribution is different from the one of Theorem 3.1. In this case, the p-value should not be used to compare samples of different size. The reason can be seen using Theorem 3.3: this result shows that for a quasi-Monte Carlo sequence with good equidistribution properties, the p-value converges to 1 when $N \to \infty$. Therefore, higher p-values could be due to larger sample sizes. Nevertheless, the p-value still represents the probability of obtaining by chance (i.e. through a sample of independent and uniformly distributed random variables of the same size as the sample) a value of the test statistic that is higher than the one based on the data. Otherwise stated, it is the probability that a sample of the same size, composed of independent and uniformly distributed random variables, is less equidistributed than the sample on which the diaphony has been computed. As an example, for the 'Worms' dataset, the probability of obtaining a sample of 708 observations that respects the null hypothesis and is less equidistributed than the one under scrutiny according to the 3-adic diaphony is 0.9999997. Despite the non-independence of the points in the dataset, this test provides evidence that the 'Worms' dataset is very well equidistributed.

6. CONCLUSIONS

In this paper, we have considered a class of discrepancies on the unit hypercube called b-adic diaphonies, introduced to test equidistribution of quasi-Monte Carlo sequences on the unit hypercube. In particular, we have derived their asymptotic statistical properties when applied to a sample of uniformly distributed random points. The asymptotic distribution of the statistic is an infinite weighted sum of chi-squared random variables. As this distribution is non-standard, we have illustrated two methods for its computation and provided explicit and tight bounds for the approximation error. We have illustrated the findings of the paper through numerical results and an empirical example.

7. Proofs

In the following we will need the definition of the function $h : [0,1)^d \times [0,1)^d \to \mathbb{R}$:

$$h(\mathbf{x}_h, \mathbf{x}_\ell) := \frac{1}{(1+b)^d - 1} \phi\left(\mathbf{x}_h \dot{-} \mathbf{x}_\ell\right),$$

which in the theory of $V-$statistics is usually called *kernel* (see [48, p. 172]). Moreover, in the following, μ_j, for $j \in \mathbb{N}$, is the j-th element of the sequence in which the eigenvalues λ_a are arranged in decreasing order and repeated a number of times equal to their multiplicity N_a.

Lemma 7.1. *For $\mathbf{x}, \mathbf{y} \in [0,1)^d$ whose coordinates are expressed in regular b-adic representation, the equality $\phi\left(\mathbf{x} \dot{-} \mathbf{y}\right) = \phi\left(\mathbf{y} \dot{-} \mathbf{x}\right)$ holds true.*

Proof of Lemma 7.1. The equality $\phi\left(\mathbf{x} \dot{-} \mathbf{y}\right) = \phi\left(\mathbf{y} \dot{-} \mathbf{x}\right)$ is guaranteed if $\lfloor \log_b (x \dot{-} y) \rfloor = \lfloor \log_b (y \dot{-} x) \rfloor$ for any $x, y \in [0,1)$. Now suppose that $\lfloor \log_b (x \dot{-} y) \rfloor = -g-1$. This means that all coefficients of terms of the form b^{-j} for $1 \leq j \leq g$ in the b-adic expansion of $x \dot{-} y$, i.e. $x^{(j)} \ominus y^{(j)}$, are 0, while the coefficient of b^{-g-1}, i.e. $x^{(g+1)} \ominus y^{(g+1)}$, is not. Therefore, we must just prove that $\alpha \ominus \beta = 0$ implies $\beta \ominus \alpha = 0$, and $\alpha \ominus \beta \neq 0$ implies $\beta \ominus \alpha \neq 0$. Now, by the very definition of $\ominus$, $\alpha \ominus \beta = 0$ can only occur when $\alpha = \beta$, in which case $\beta \ominus \alpha = \beta - \alpha = 0$. Now consider what happens when $\alpha \ominus \beta \neq 0$. If we suppose that $\alpha \ominus \beta \neq 0$ and $\alpha \geq \beta$, we have $\alpha \ominus \beta = \alpha - \beta > 0$; then $\beta \ominus \alpha = b + \beta - \alpha > 0$. At last, if we suppose that $\alpha \ominus \beta \neq 0$ and $\beta > \alpha$, we have $\alpha \ominus \beta = b + \alpha - \beta > 0$; then $\beta \ominus \alpha = \alpha - \beta > 0$. □

Lemma 7.2. *For even $p \in \mathbb{N}$, we have*

$$\begin{aligned}\mathbb{E}\left|h(\mathbf{x}_i, \mathbf{x}_j)\right|^p &= \frac{1}{\left((1+b)^d - 1\right)^p} \sum_{i=0}^{p} \binom{p}{i} (-1)^{p-i} \\ &\quad \cdot (1+b)^{di} \left[(b-1) \sum_{k=0}^{i} \binom{i}{k} \frac{(-b)^k}{b^{k+1}-1}\right]^d.\end{aligned}$$

Remark 7.3. When $p = 2, 4$:

$$\begin{aligned}\mathbb{E}\left|h(\mathbf{x}_i, \mathbf{x}_j)\right|^2 &= \frac{1}{\left((1+b)^d-1\right)^2} \left\{\left[\frac{(b+1)\left(b^2+1\right)}{(b^2+b+1)}\right]^d - 1\right\}, \\ \mathbb{E}\left|h(\mathbf{x}_i, \mathbf{x}_j)\right|^4 &= \frac{1}{\left((1+b)^d-1\right)^4} \left\{6\left[\frac{(b+1)\left(b^2+1\right)}{(b^2+b+1)}\right]^d\right. \\ &\quad -4\left[\frac{(b+1)^2\left(b^4-b^3+3b^2-b+1\right)}{(b^2+1)(b^2+b+1)}\right]^d \\ &\quad \left.+\left[\frac{(b+1)^3\left(b^2-b+1\right)\left(b^6+3b^4+4b^3+3b^2+1\right)}{(b^2+1)(b^2+b+1)(b^4+b^3+b^2+b+1)}\right]^d - 3\right\}.\end{aligned}$$

Proof of Lemma 7.2. Adapting the proofs of Lemma A.12 and Corollary A.13 in [13], it is possible to show that if $\mathbf{x}$ and $\mathbf{y}$ are uniformly distributed on $[0,1)^d$, $\mathbf{x}\dot{-}\mathbf{y}$ is uniformly distributed on $[0,1)^d$, too. This implies that

$$\mathbb{E}\left|\frac{\phi\left(\mathbf{x}\dot{-}\mathbf{y}\right)}{(1+b)^d-1}\right|^p = \mathbb{E}\left|\frac{\phi\left(\mathbf{x}\right)}{(1+b)^d-1}\right|^p .$$

Now, $\mathbb{E}\left|\phi\left(\mathbf{x}\right)\right|^p$ is given by

$$\begin{aligned} & \mathbb{E}\left|-1+(1+b)^d\prod_{j=1}^{d}\left(1-b^{1+\lfloor\log_b x_j\rfloor}\right)\right|^p \\ = & \sum_{i=0}^{p}\binom{p}{i}(-1)^{p-i}(1+b)^{di}\prod_{j=1}^{d}\mathbb{E}\left(1-b^{1+\lfloor\log_b x_j\rfloor}\right)^i \\ = & \sum_{i=0}^{p}\binom{p}{i}(-1)^{p-i}(1+b)^{di}\prod_{j=1}^{d}\left[\sum_{k=0}^{i}\binom{i}{k}(-b)^k\,\mathbb{E}b^{k\lfloor\log_b x_j\rfloor}\right] \end{aligned}$$

where

$$\begin{aligned} \mathbb{E}b^{k\lfloor\log_b x\rfloor} &= \sum_{h=0}^{\infty}\int_{b^{-(h+1)}}^{b^{-h}} b^{k\lfloor\log_b x\rfloor}\mathrm{d}x \\ &= \sum_{h=0}^{\infty}\left(b^{-h}-b^{-(h+1)}\right)b^{k\left\lfloor\log_b b^{-(h+1)}\right\rfloor} \\ &= \sum_{h=0}^{\infty}\left(b^{-h}-b^{-(h+1)}\right)b^{-(h+1)k} \\ &= \left(1-b^{-1}\right)b^{-k}\sum_{h=0}^{\infty}b^{-(k+1)h} = \frac{b-1}{b^{k+1}-1}. \end{aligned}$$

□

Proof of Theorem 3.1. The statistic in Definition 2.1 can be written as a V-statistic. The symmetry of the kernel is established in Lemma 7.1. The fact that this V-statistic is degenerate (see, e.g., [20]) can be seen, e.g., from the intermediate expression (2.1), integrating the kernel with respect to $\mathbf{x}_h$. In general, it is well known (see [48, pp. 193, 196])) that the kernel of a second order V-statistic (and also of a U-statistic) induces an integral operator $\mathcal{A}$ through the formula

$$\mathcal{A}m\left(\mathbf{x}\right) := \int_{[0,1)^d} h\left(\mathbf{y},\mathbf{x}\right)m\left(\mathbf{y}\right)\mathrm{d}\mathbf{y}, \quad \mathbf{x}\in[0,1)^d .$$

The eigenvalues of $\mathcal{A}$ are defined as the values λ_j, for $j=1,2,\ldots$, for which there exists a function h_j such that $\mathcal{A}h_j\left(\mathbf{x}\right)=\lambda_j h_j\left(\mathbf{x}\right)$ for any $\mathbf{x}\in[0,1)^d$ holds. In this case, it is not difficult to identify the eigenvalues as the weights $\frac{1}{(1+b)^d-1}r_b\left(\mathbf{k}\right)$, for $\mathbf{k}\in\mathbb{N}_0^d\backslash\{\mathbf{0}\}$. In the following, we will also repeatedly need the facts that $\mathbb{E}\left|h\left(\mathbf{x}_h,\mathbf{x}_\ell\right)\right|<\infty$ and $\mathbb{E}\left|h\left(\mathbf{x}_h,\mathbf{x}_\ell\right)\right|^2<\infty$: these moment conditions are automatically verified since the kernel is bounded on a bounded domain.

(i) When $\mathbb{E}\left|h\left(\mathbf{x}_h,\mathbf{x}_\ell\right)\right|<\infty$, the strong law of large numbers holds for $F_{b,N}^2\left(\mathcal{P}_N\right)$ (see [48, Theorem A, p. 190]) and $F_{b,N}^2\left(\mathcal{P}_N\right)\to 0$, almost surely, if the points are independently uniformly distributed.

(ii) First consider the upper bound. Using the fact that $h\left(\mathbf{0}\right)=1$, we write the b-adic diaphony as

$$F_{b,N}^2\left(\mathcal{P}_N\right)=\frac{2}{N^2}\sum_{h=1}^{N}\sum_{\ell=1}^{h-1}h\left(\mathbf{x}_h,\mathbf{x}_\ell\right)+\frac{1}{N}.$$

From Theorem 2 of [10], if the kernel is square-integrable, then

$$\begin{aligned}
&\limsup_{N\to\infty}\frac{N}{\ln\ln N}F_{b,N}^2\left(\mathcal{P}_N\right)\\
=\;&\limsup_{N\to\infty}\frac{N}{\ln\ln N}\frac{2}{N^2}\sum_{h=1}^{N}\sum_{\ell=1}^{h-1}h\left(\mathbf{x}_h,\mathbf{x}_\ell\right)\\
&+\limsup_{N\to\infty}\frac{N}{\ln\ln N}\frac{1}{N}\\
=\;&2\max_a\left\{\lambda_a\right\}=\frac{2}{\left(1+b\right)^d-1}.
\end{aligned}$$

The same reasoning could be repeated for the lim inf part, but it would yield a 0 limit. A better result can be obtained using Theorem 1 in [7]. From this result, we get

$$F_{b,N}^2\left(\mathcal{P}_N\right)\geq C_{d,b}\frac{\left(\ln N\right)^{d-1}}{N^2}$$

where $C_{d,b}>0$ does not depend on the sequence $\mathcal{P}_N$. From this, the result follows.

(iii) The asymptotic distribution is given by

$$NF_{b,N}^2\left(\mathcal{P}_N\right)\to_{\mathcal{D}}\sum_{\mathbf{k}\in\mathbb{N}_0^d\setminus\{\mathbf{0}\}}\frac{r_b\left(\mathbf{k}\right)}{\left(1+b\right)^d-1}\chi^2\left(1\right),$$

provided $\mathbb{E}\left|h\left(\mathbf{x}_h,\mathbf{x}_\ell\right)\right|^2<\infty$ (see [48, Theorem, p. 194]). The main problem is that this distribution is not suitable for computations. Therefore, we now aggregate the chi-squared random variables according to the weight with which they appear in the infinite weighted sum.

When $d=1$, the equality $r_b\left(k\right)=b^{-2a}$ for $a\in\mathbb{N}_0$ holds for $b^a\left(b-1\right)$ values of k.

When $d>1$, the value $r_b\left(\mathbf{k}\right):=\prod_{j=1}^{d}r_b\left(k_j\right)=b^{-2a}$ can be obtained in several ways. Consider a vector $\mathbf{k}=\left(k_1,k_2,\ldots,k_d\right)$.

- Consider first the case in which $a=0$. Any component j of the vector must have $0\leq k_j<b$. Therefore, there are b^d-1 possible vectors (the subtraction of 1 is due to the removal of the vector with all components equal to 0).
- Any component j of the vector with $0\leq k_j<b$ does not contribute to $r_b\left(\mathbf{k}\right)$. Therefore, we suppose for the moment that ℓ (with $1\leq\ell\leq\min\left\{d,a\right\}$) components $\left(k_1,k_2,\ldots,k_\ell\right)$ are larger than or equal to b. For any ℓ, we can write $a=\sum_{j=1}^{\ell}a_j$, where $1\leq a_j$, $j=1,\ldots,\ell$. Each way of writing a as a sum of ℓ integers (such that order matters) is called an ℓ-composition

of a. For a fixed value of ℓ, there are $\left(\begin{array}{c} a-1 \\ \ell-1 \end{array}\right)$ ℓ-compositions. Let $(a_1, a_2, \ldots, a_\ell)$ be an (ordered) ℓ-composition: this can be obtained for $\prod_{j=1}^{\ell} b^{a_j}(b-1) = b^a (b-1)^\ell$ values of $(k_1, k_2, \ldots, k_\ell)$. Therefore, for any ℓ, we have $\left(\begin{array}{c} a-1 \\ \ell-1 \end{array}\right) b^a (b-1)^\ell$ possible values of $(k_1, k_2, \ldots, k_\ell)$. Note, however, that $d-\ell$ values of $\mathbf{k} = (k_1, k_2, \ldots, k_d)$ have still to be chosen. The number of ways in which the subvector $(k_1, k_2, \ldots, k_\ell)$ can be arranged in $\mathbf{k} = (k_1, k_2, \ldots, k_d)$ is equal to $\left(\begin{array}{c} d \\ \ell \end{array}\right)$; the remaining $d-\ell$ places can be filled with any of the b numbers $\{0, 1, \ldots, b-1\}$, yielding $b^{d-\ell}$ combinations. At last, we have

$$\sum_{\ell=1}^{\min\{d,a\}} \left(\begin{array}{c} a-1 \\ \ell-1 \end{array}\right) b^a (b-1)^\ell \left(\begin{array}{c} d \\ \ell \end{array}\right) b^{d-\ell}$$

vectors $\mathbf{k} = (k_1, k_2, \ldots, k_d)$ yielding $r_b(\mathbf{k}) = \prod_{j=1}^{d} r_b(k_j) = b^{-2a}$.

Summing up, the values $(\lambda_a)_{a\in\mathbb{N}_0}$ and their multiplicities $(N_a)_{a\in\mathbb{N}_0}$ are defined by $\lambda_a = \frac{b^{-2a}}{(1+b)^d-1}$ and

$$N_a = \begin{cases} b^d - 1 & \text{for } a = 0, \\ \sum_{\ell=1}^{\min\{d,a\}} \binom{a-1}{\ell-1}\binom{d}{\ell} b^{a+d-\ell}(b-1)^\ell & \text{for } a > 0. \end{cases}$$

(iv) Define the following moments of the kernel h:

$$\gamma_s = \mathbb{E}\left|h(\mathbf{x}_i, \mathbf{x}_j)\right|^s, \qquad \gamma_{s,r} = \mathbb{E}\left|\mathbb{E}\left[\left|h(\mathbf{x}_i, \mathbf{x}_j)\right|^s \left|\mathbf{x}_i\right.\right]\right|^r . \tag{7.1}$$

From [2, Theorem 1.1] we have

$$\left\| \mathbb{P}\left\{NF_{b,N}^2(\mathcal{P}_N) \le x\right\} - \mathbb{P}\left\{\sum_{a=0}^{\infty} \lambda_a \chi_a^2(N_a) \le x\right\}\right\|_\infty \le \frac{\exp\left\{\frac{c\sqrt{\gamma_2}}{|\mu_{13}|}\right\}}{N}\left(\frac{\gamma_3}{\gamma_2^{\frac{3}{2}}} + \frac{\gamma_{2,2}}{\gamma_2^2}\right)$$

where μ_{13} is defined at the beginning of this section. A simpler formula can be obtained remarking that

$$\frac{\gamma_3}{\gamma_2^{\frac{3}{2}}} + \frac{\gamma_{2,2}}{\gamma_2^2} \le \frac{\gamma_4^{\frac{3}{4}}}{\gamma_2^{\frac{3}{2}}} + \frac{\gamma_4}{\gamma_2^2} \le 2\max\left(\left(\frac{\gamma_4}{\gamma_2^2}\right)^{\frac{3}{4}}, \frac{\gamma_4}{\gamma_2^2}\right) \le 2\frac{\gamma_4}{\gamma_2^2}$$

where the first and the last steps are consequences of the inequalities $\gamma_3^{1/3} \le \gamma_4^{1/4}$, $\gamma_{2,2} \le \gamma_4$ and $\gamma_2^{1/2} \le \gamma_4^{1/4}$ (that are respectively consequences of Liapunov's inequality; see [51, Inequality 4.4, p. 48]). Using Lemma 7.2, we note that

$$\begin{aligned}
\gamma_4 &\le \frac{1}{\left((1+b)^d-1\right)^4}\left[\frac{(b+1)^3\left(b^2-b+1\right)\left(b^6+3b^4+4b^3+3b^2+1\right)}{(b^2+1)(b^2+b+1)(b^4+b^3+b^2+b+1)}\right]^d, \\
\gamma_2 &\le \frac{1}{\left((1+b)^d-1\right)^2}\left[\frac{(b+1)\left(b^2+1\right)}{(b^2+b+1)}\right]^d, \\
\gamma_2^2 &\ge \frac{c'}{\left((1+b)^d-1\right)^4}\left[\frac{(b+1)\left(b^2+1\right)}{(b^2+b+1)}\right]^{2d}
\end{aligned}$$

for a certain constant $c' > 0$ (that can be taken equal to $\frac{64}{225}$). Therefore,

$$\frac{\gamma_3}{\gamma_2^{\frac{3}{2}}} + \frac{\gamma_{2,2}}{\gamma_2^2} \leq \frac{2}{c'} \left[\frac{(b+1)\left(b^2+b+1\right)\left(b^2-b+1\right)\left(b^6+3b^4+4b^3+3b^2+1\right)}{\left(b^2+1\right)^3\left(b^4+b^3+b^2+b+1\right)} \right]^d .$$

The inequalities $\left(b^2+b+1\right)\left(b^2-b+1\right) \leq b^4+b^3+b^2+b+1$ and $b^6+3b^4+4b^3+3b^2+1 \leq 1.256\left(b^2+1\right)^3$ are then used to simplify the result. Concerning μ_{13}, it is different from $\frac{1}{(1+b)^d-1}$ only when $b^d \leq 14$ and, when this inequality is not satisfied, it is larger than this value. Therefore

$$\frac{\sqrt{\gamma_2}}{|\mu_{13}|} \leq \left[\frac{(b+1)\left(b^2+1\right)}{\left(b^2+b+1\right)} \right]^{\frac{d}{2}} \leq (b+1)^{\frac{d}{2}} .$$

□

Lemma 7.4. *For fixed b, d and p and large A, the following asymptotic formulas hold:*

$$\begin{aligned} N_A &\simeq \tfrac{(b-1)^d}{(d-1)!} A^{d-1} b^A, \\ \sum_{a=A+1}^{\infty} N_a \lambda_a^p &\simeq \tfrac{(b-1)^d}{(d-1)!\left((1+b)^d-1\right)^p \left(b^{2p-1}-1\right)} b^{-(2p-1)A} A^{d-1}. \end{aligned}$$

Note that the first formula is exact for $d = 1$.

Remark 7.5. For $p = 1, 2$:

$$\begin{aligned} \sum_{a=A+1}^{\infty} N_a \lambda_a &\simeq \tfrac{(b-1)^{d-1}}{(d-1)!\left((1+b)^d-1\right)} b^{-A} A^{d-1}, \\ \sum_{a=A+1}^{\infty} N_a \lambda_a^2 &\simeq \tfrac{(b-1)^d}{(d-1)!\left((1+b)^d-1\right)^2 \left(b^3-1\right)} b^{-3A} A^{d-1}. \end{aligned}$$

Proof of Lemma 7.4. The asymptotic behavior of N_a for large a is given by

$$\begin{aligned} &\sum_{\ell=1}^{\min\{d,a\}} \binom{a-1}{\ell-1} b^a (b-1)^\ell \binom{d}{\ell} b^{d-\ell} \\ = &\sum_{\ell=1}^{d} \frac{(a-1)!d!}{(\ell-1)!\,(a-\ell)!\ell!\,(\ell-d)!} b^{a+d-\ell} (b-1)^\ell \\ \simeq &\sum_{\ell=1}^{d} \frac{d!}{(\ell-1)!\ell!\,(\ell-d)!} a^{\ell-1} b^{a+d} \left(1-\frac{1}{b}\right)^\ell \\ \simeq &\frac{(b-1)^d}{(d-1)!} a^{d-1} b^a \end{aligned}$$

where the second step derives from 5.11.12 in [41] applied to $(a-1)!/\,(a-\ell)!$, and the third from the fact that the leading term in the sum is the one with $\ell = d$. Therefore, we have

$$\sum_{a=A+1}^{\infty} N_a \lambda_a^p \simeq \tfrac{(b-1)^d}{(d-1)!\left((1+b)^d-1\right)^p} \sum_{a=A+1}^{\infty} a^{d-1} b^{-(2p-1)a}.$$

Clearly, the sum is convergent. A precise estimate of its asymptotic behavior can be obtained noting that the sum is a special instance of the Hurwitz-Lerch zeta function, i.e., the function $\Phi(z,s,a) := \sum_{k=0}^{\infty} \frac{z^k}{(a+k)^s}$, for $1-a \notin \mathbb{N}$, $s \in \mathbb{C}$ and $|z|<1$. This yields $\sum_{a=A+1}^{\infty} N_a \lambda_a^p \simeq \frac{(b-1)^d}{(d-1)!\left((1+b)^d-1\right)^p} b^{-(2p-1)(A+1)} \Phi\left(b^{-(2p-1)}, 1-d, A+1\right)$. The asymptotic behavior of Φ for large values of a can be recovered from Theorem 1 in [16] and gives

$$\begin{aligned}\sum_{a=A+1}^{\infty} N_a \lambda_a^p &\simeq \frac{(b-1)^d}{(d-1)!\left((1+b)^d-1\right)^p} \frac{b^{-(2p-1)(A+1)}}{1-b^{-(2p-1)}} \frac{1}{(A+1)^{1-d}} \\ &\simeq \frac{(b-1)^d}{(d-1)!\left((1+b)^d-1\right)^p (b^{2p-1}-1)} b^{-(2p-1)A} A^{d-1}.\end{aligned}$$

□

Proof of Theorem 3.3. We rewrite the probability appearing in the statement as

$$\begin{aligned}\mathbb{P}\left\{F_{b,N}(\mathcal{P}_N) \leq F_{b,N}(\mathcal{P}_N^\star)\right\} &= \mathbb{P}\left\{NF_{b,N}^2(\mathcal{P}_N) \leq NF_{b,N}^2(\mathcal{P}_N^\star)\right\} \\ &= F_{NF_{b,N}^2}\left(NF_{b,N}^2(\mathcal{P}_N^\star)\right)\end{aligned}$$

where $F_{NF_{b,N}^2}$ is a shortcut for the cumulative distribution function of $NF_{b,N}^2(\mathcal{P}_N)$. Moreover, we take, as in Section 4, $X = \sum_{a=0}^{\infty} \lambda_a \chi_a^2(N_a)$ and we denote by F_X the cumulative distribution function of X. From Proposition 9.1 in [35], the distribution of X possesses a density that we denote f_X. This also implies that F_X is continuous (see [46, Exercise 17 (a), p. 303]).

According to Theorem 3.1 (iv), we have

$$\left|F_{NF_{b,N}^2}\left(NF_{b,N}^2(\mathcal{P}_N^\star)\right) - F_X\left(NF_{b,N}^2(\mathcal{P}_N^\star)\right)\right| \leq \frac{e^{c_0+c_1(b+1)^{\frac{d}{2}}}(1.256(b+1))^d}{N},$$

which can be written as the following two inequalities:

$$0 \leq F_{NF_{b,N}^2}\left(NF_{b,N}^2(\mathcal{P}_N^\star)\right) \leq F_X\left(NF_{b,N}^2(\mathcal{P}_N^\star)\right) + \frac{e^{c_0+c_1(b+1)^{\frac{d}{2}}}(1.256(b+1))^d}{N}, \tag{7.2}$$

$$0 \leq F_X\left(NF_{b,N}^2(\mathcal{P}_N^\star)\right) \leq F_{NF_{b,N}^2}\left(NF_{b,N}^2(\mathcal{P}_N^\star)\right) + \frac{e^{c_0+c_1(b+1)^{\frac{d}{2}}}(1.256(b+1))^d}{N}. \tag{7.3}$$

We start assuming that $\lim_{N\to\infty} NF_{b,N}^2(\mathcal{P}_N^\star) = 0$, and we show, through (7.2), that $\lim_{N\to\infty} \mathbb{P}\left\{F_{b,N}(\mathcal{P}_N) \leq F_{b,N}(\mathcal{P}_N^\star)\right\} = 0$. However, we will not need a precise estimate of $F_X\left(NF_{b,N}^2(\mathcal{P}_N^\star)\right)$ (this could be obtained using the fact that this is a small ball probability). Instead, we use the crude inequality

$$F_X\left(NF_{b,N}^2(\mathcal{P}_N^\star)\right) = \int_0^{NF_{b,N}^2(\mathcal{P}_N^\star)} f_X(x)\,\mathrm{d}x \leq NF_{b,N}^2(\mathcal{P}_N^\star) \max_{x\in\mathbb{R}} f_X(x).$$

From [47, equation (7), p. 72], we get

$$\max_{x\in\mathbb{R}} f_X(x) \leq \frac{1}{2\sqrt{\mu_1\mu_2}}.$$

When $(d,b) \neq (1,2)$, $\mu_1\mu_2 = \left((1+b)^d - 1\right)^{-2}$, while when $(d,b) = (1,2)$, $\mu_1\mu_2 = \left((1+b)^d - 1\right)^{-2} b^{-2}$. Therefore, $\frac{1}{2\sqrt{\mu_1\mu_2}} \leq (1+b)^d - 1$. From (7.2), if we let $N \to \infty$:

$$\begin{aligned}
0 &\leq \lim_{N\to\infty} F_{NF^2_{b,N}}\left(NF^2_{b,N}\left(\mathcal{P}^\star_N\right)\right)\\
&\leq \lim_{N\to\infty}\left((1+b)^d - 1\right) NF^2_{b,N}\left(\mathcal{P}^\star_N\right) + \lim_{N\to\infty} \frac{e^{c_0 + c_1(b+1)^{\frac{d}{2}}}(1.256(b+1))^d}{N} = 0.
\end{aligned}$$

Now we show that $\lim_{N\to\infty} \mathbb{P}\left\{F_{b,N}\left(\mathcal{P}_N\right) \leq F_{b,N}\left(\mathcal{P}^\star_N\right)\right\} = 0$ implies that $\lim_{N\to\infty} NF^2_{b,N}\left(\mathcal{P}^\star_N\right) = 0$. Now, when $N \to \infty$ in (7.3):

$$\begin{aligned}
0 &\leq \lim_{N\to\infty} F_X\left(NF^2_{b,N}\left(\mathcal{P}^\star_N\right)\right)\\
&\leq \lim_{N\to\infty} F_{NF^2_{b,N}}\left(NF^2_{b,N}\left(\mathcal{P}^\star_N\right)\right) + \lim_{N\to\infty} \frac{e^{c_0 + c_1(b+1)^{\frac{d}{2}}}(1.256(b+1))^d}{N}\\
&\leq 0.
\end{aligned}$$

As F_X is continuous, non-decreasing and strictly positive when its argument is strictly positive, $\lim_{N\to\infty} F_X\left(NF^2_{b,N}\left(\mathcal{P}^\star_N\right)\right) = 0$ requires that $\lim_{N\to\infty} NF^2_{b,N}\left(\mathcal{P}^\star_N\right) = 0$. $\square$

Proof of Theorem 4.1. The proof follows the scheme of [5, Theorem 5.1]. Here and in the following, $\psi_X(t) := \mathbb{E}e^{\mathrm{i}tX}$ is the characteristic function of the random variable X.

We start with the uncentered case. Through the Fourier inversion formula for characteristic functions of random variables (see, e.g., [15, p. 33]), it is possible to obtain the following bound:

$$\begin{aligned}
&\left\|F_X(x) - F_{X_A}(x)\right\|_\infty\\
\leq\ & \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{1}{|t|}\left|\psi_X(t) - \psi_{X_A}(t)\right| \mathrm{d}t\\
=\ & \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{1}{|t|}\left|\psi_{X_A}(t)\right| \left|\frac{\psi_X(t)}{\psi_{X_A}(t)} - 1\right| \mathrm{d}t\\
=\ & \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{1}{|t|}\left|\psi_{X_A}(t)\right| \left|\mathbb{E}e^{\mathrm{i}t(X - X_A)} - 1\right| \mathrm{d}t\\
\leq\ & \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{1}{|t|}\left|\psi_{X_A}(t)\right| \mathbb{E}\left|e^{\mathrm{i}t(X - X_A)} - 1\right| \mathrm{d}t\\
\leq\ & \frac{\sqrt{\mathbb{E}\left(X - X_A\right)^2}}{2\pi}\int_{-\infty}^{+\infty} \left|\psi_{X_A}(t)\right| \mathrm{d}t
\end{aligned}$$

where the third step comes from the fact that X_A and $X - X_A$ are independent and the fifth step from the inequalities $\left|e^{\mathrm{i}x} - 1\right| \leq |x|$ and $\mathbb{E}\left|X - X_A\right| \leq \sqrt{\mathbb{E}\left(X - X_A\right)^2}$ (see [51, Lemma 4.2] and [48, p. 197] for a similar majorization). In the centered case, on the other hand, from

$$\begin{aligned}\left|\mathbb{E}e^{\mathrm{i}t(X^\star-X_A^\star)}-1\right| &= \left|\mathbb{E}e^{\mathrm{i}t(X^\star-X_A^\star)}-1-\mathbb{E}\mathrm{i}t\left(X^\star-X_A^\star\right)\right| \\ &\leq \mathbb{E}\left|e^{\mathrm{i}t(X^\star-X_A^\star)}-1-\mathrm{i}t\left(X^\star-X_A^\star\right)\right| \\ &\leq \frac{t^2\mathbb{E}\left|X^\star-X_A^\star\right|^2}{2}\end{aligned}$$

we have

$$\left\|F_{X^\star}(x)-F_{X_A^\star}(x)\right\|_\infty \leq \frac{\mathbb{E}\left|X^\star-X_A^\star\right|^2}{4\pi}\int_{-\infty}^{+\infty}|t|\left|\psi_{X_A}(t)\right|\mathrm{d}t.$$

As concerns ψ_{X_A}, we have the general formula:

$$\begin{aligned}\psi_{X_A}(t) &= \mathbb{E}e^{\mathrm{i}tX_A}=\prod_{a=0}^{A}\mathbb{E}e^{\mathrm{i}t\lambda_a\chi_a^2(N_a)} \\ &= \prod_{a=0}^{A}\left(1-2\mathrm{i}t\lambda_a\right)^{-\frac{N_a}{2}}=\exp\left\{-\sum_{a=0}^{A}\tfrac{N_a}{2}\ln\left(1-2\mathrm{i}t\lambda_a\right)\right\} \\ &= \exp\left\{-\sum_{a=0}^{A}\tfrac{N_a}{4}\ln\left(1+4t^2\lambda_a^2\right)-\mathrm{i}\sum_{a=0}^{A}\tfrac{N_a}{2}\arctan\left(-2t\lambda_a\right)\right\}\end{aligned}$$

where the second equality derives from the multiplication property of characteristic functions (see [51, Proposition 1.1 (g), p. 341]), the third from the characteristic function of a chi-squared random variable (see [51, p. 343]) and the fifth from the formula of the principal branch of the logarithm of a complex number (i.e. $\log(z)=\log\rho+\mathrm{i}\theta$ for $z=\rho e^{\mathrm{i}\theta}$; see [51, p. 352]). As concerns the modulus of ψ_{X_A}, we have

$$\left|\psi_{X_A}(t)\right|=\exp\left\{-\sum_{a=0}^{A}\tfrac{N_a}{4}\ln\left(1+4t^2\lambda_a^2\right)\right\}=\prod_{a=0}^{A}\left(1+4t^2\lambda_a^2\right)^{-\frac{N_a}{4}}.$$

This formula is not directly usable; therefore we consider the following majorizations:

- Consider the uncentered case (namely X_A) and let α be a value of the index a such that $N_\alpha>2$. We have

$$\begin{aligned}\left|\psi_{X_A}(t)\right| &= \frac{1}{\prod_{a=0}^{A}\left(1+4t^2\lambda_a^2\right)^{\frac{N_a}{4}}}\leq\frac{1}{\prod_{a=\alpha}^{A}\left(1+4t^2\lambda_a^2\right)^{\frac{N_a}{4}}} \\ &= \exp\left\{-\sum_{a=\alpha}^{A}\frac{N_a}{4}\ln\left(1+4t^2\lambda_a^2\right)\right\} \\ &\leq \left(1+t^2\frac{4\sum_{a=\alpha}^{A}N_a\lambda_a^2}{N_\alpha}\right)^{-\frac{N_\alpha}{4}};\end{aligned}$$

the result is derived from the following sequence of inequalities:

$$\begin{aligned}
\sum_{a=\alpha}^{A} \frac{N_a}{4} \ln\left(1+4t^2\lambda_a^2\right) &= \frac{N_\alpha}{4} \sum_{a=\alpha}^{A} \frac{N_a}{N_\alpha} \ln\left(1+4t^2\lambda_a^2\right) \\
&= \frac{N_\alpha}{4} \sum_{a=\alpha}^{A} \ln\left(1+4t^2\lambda_a^2\right)^{\frac{N_a}{N_\alpha}} \\
&\geq \frac{N_\alpha}{4} \sum_{a=\alpha}^{A} \ln\left(1+4\frac{N_a}{N_\alpha}t^2\lambda_a^2\right) \\
&= \frac{N_\alpha}{4} \ln \prod_{a=\alpha}^{A} \left(1+4\frac{N_a}{N_\alpha}t^2\lambda_a^2\right) \\
&\geq \frac{N_\alpha}{4} \ln\left(1+t^2\frac{4\sum_{a=\alpha}^{A} N_a\lambda_a^2}{N_\alpha}\right)
\end{aligned}$$

where the third relation holds provided $\frac{N_a}{N_\alpha} \geq 1$. Therefore:

$$\begin{aligned}
\int_{-\infty}^{+\infty} \left|\psi_{X_A}(t)\right| \mathrm{d}t &= 2\int_{0}^{+\infty} \left|\psi_{X_A}(t)\right| \mathrm{d}t \\
&\leq 2\int_{0}^{+\infty} \left(1+t^2\frac{4\sum_{a=\alpha}^{A} N_a\lambda_a^2}{N_\alpha}\right)^{-\frac{N_\alpha}{4}} \mathrm{d}t \\
&= \frac{B\left(\frac{1}{2}, \frac{N_\alpha-2}{4}\right)}{2\sqrt{\frac{\sum_{a=\alpha}^{A} N_a\lambda_a^2}{N_\alpha}}},
\end{aligned}$$

where $B(\cdot,\cdot)$ is the beta function, provided $N_\alpha > 2$ (this is the reason why we replaced 0 with α in the sums). Note that $\alpha = 0$ for every combination (d, b) apart from $b \in \{2, 3\}$ and $d = 1$. It is $\alpha = 1$ for $(d, b) = (1, 3)$, and $\alpha = 2$ for $(d, b) = (1, 2)$.

- Consider the centered case and let $\alpha^\star$ be a value of the index a such that $N_{\alpha^\star} > 4$. We have

$$\begin{aligned}
\int_{-\infty}^{+\infty} \left|t\psi_{X_A}(t)\right| \mathrm{d}t &= 2\int_{0}^{+\infty} t\left|\psi_{X_A}(t)\right| \mathrm{d}t \\
&\leq 2\int_{0}^{+\infty} t\left(1+t^2\frac{4\sum_{a=\alpha^\star}^{A} N_a\lambda_a^2}{N_{\alpha^\star}}\right)^{-\frac{N_{\alpha^\star}}{4}} \mathrm{d}t \\
&= \frac{1}{\left(\frac{\sum_{a=\alpha^\star}^{A} N_a\lambda_a^2}{N_{\alpha^\star}}\right)\left(N_{\alpha^\star}-4\right)}.
\end{aligned}$$

Now we turn to $\mathbb{E}\left(X^{\star}-X_A^{\star}\right)^2$ and $\mathbb{E}\left(X-X_A\right)^2$. We have

$$\begin{aligned}
\mathbb{E}\left(X^{\star}-X_A^{\star}\right)^2 &= \mathbb{E}\left(\sum_{a=A+1}^{\infty}\lambda_a\left(\chi_a^2\left(N_a\right)-N_a\right)\right)^2 \\
&= \sum_{a=A+1}^{\infty}\lambda_a^2\mathbb{V}\left(\chi_a^2\left(N_a\right)\right)=2\sum_{a=A+1}^{\infty}\lambda_a^2 N_a, \\
\mathbb{E}\left(X-X_A\right)^2 &= \mathbb{E}\left(\left(X^{\star}-X_A^{\star}\right)+\mathbb{E}\left(X-X_A\right)\right)^2 \\
&= \mathbb{V}\left(X-X_A\right)+\left[\mathbb{E}\left(\sum_{a=A+1}^{\infty}\lambda_a\chi_a^2\left(N_a\right)\right)\right]^2 \\
&= 2\sum_{a=A+1}^{\infty}\lambda_a^2 N_a+\left[\sum_{a=A+1}^{\infty}\lambda_a N_a\right]^2.
\end{aligned}$$

From Lemma 7.4, we get

$$\sqrt{2\sum_{a=A+1}^{\infty}\lambda_a^2 N_a+\left[\sum_{a=A+1}^{\infty}\lambda_a N_a\right]^2}\simeq\frac{(b-1)^{d-1}b^{-A}A^{d-1}}{(d-1)!\left((1+b)^d-1\right)}$$

and

$$\sum_{a=A+1}^{\infty}\lambda_a^2 N_a\simeq\frac{(b-1)^{d}b^{-3A}A^{d-1}}{(d-1)!\left((1+b)^d-1\right)^2\left(b^3-1\right)}.$$

□

References

[1] C. Amstler and P. Zinterhof, *Uniform distribution, discrepancy, and reproducing kernel Hilbert spaces*, J. Complexity **17** (2001), no. 3, 497–515, DOI 10.1006/jcom.2001.0580. MR1851057

[2] V. Bentkus and F. Götze, *Optimal bounds in non-Gaussian limit theorems for U-statistics*, Ann. Probab. **27** (1999), no. 1, 454–521, DOI 10.1214/aop/1022677269. MR1681161

[3] C. Choirat and R. Seri, *The asymptotic distribution of quadratic discrepancies*, Monte Carlo and Quasi-Monte Carlo Methods 2004, Springer, Berlin, 2006, pp. 61–76, DOI 10.1007/3-540-31186-6_5. MR2208702

[4] C. Choirat and R. Seri, *Statistical properties of generalized discrepancies*, Math. Comp. **77** (2008), no. 261, 421–446 (electronic), DOI 10.1090/S0025-5718-07-01839-X. MR2353960

[5] C. Choirat and R. Seri, *Computational aspects of Cui-Freeden statistics for equidistribution on the sphere*, Math. Comp. **82** (2013), no. 284, 2137–2156, DOI 10.1090/S0025-5718-2013-02698-1. MR3073194

[6] R. Christensen, *Testing Fisher, Neyman, Pearson, and Bayes*, Amer. Statist. **59** (2005), no. 2, 121–126, DOI 10.1198/000313005X20871. MR2133558

[7] L. L. Cristea and F. Pillichshammer, *A lower bound for the b-adic diaphony*, Rend. Mat. Appl. (7) **27** (2007), no. 2, 147–153. MR2361027

[8] R. B. Davies, *Numerical inversion of a characteristic function*, Biometrika **60** (1973), 415–417. MR0321152

[9] R. B. Davies, *Statistical algorithms: Algorithm AS 155: The distribution of a linear combination of χ^2 random variables*, Applied Statistics **29** (1980), no. 3, 323–333.

[10] H. Dehling, *Complete convergence of triangular arrays and the law of the iterated logarithm for U-statistics*, Statist. Probab. Lett. **7** (1989), no. 4, 319–321, DOI 10.1016/0167-7152(89)90115-6. MR980708

[11] J. Dick and F. Pillichshammer, *Diaphony, discrepancy, spectral test and worst-case error*, Math. Comput. Simulation **70** (2005), no. 3, 159–171, DOI 10.1016/j.matcom.2005.06.004. MR2176902

[12] J. Dick and F. Pillichshammer, *Dyadic diaphony of digital nets over* $\mathbb{Z}_2$, Monatsh. Math. **145** (2005), no. 4, 285–299, DOI 10.1007/s00605-004-0287-7. MR2162347

[13] J. Dick and F. Pillichshammer, *Digital Nets and Sequences*: *Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press, Cambridge, 2010. MR2683394

[14] M. Drmota and R. F. Tichy, *Sequences, Discrepancies and Applications*, Lecture Notes in Mathematics, vol. 1651, Springer-Verlag, Berlin, 1997. MR1470456

[15] C.-G. Esséen, *Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law*, Acta Math. **77** (1945), 1–125. MR0014626

[16] C. Ferreira and J. L. López, *Asymptotic expansions of the Hurwitz-Lerch zeta function*, J. Math. Anal. Appl. **298** (2004), no. 1, 210–224, DOI 10.1016/j.jmaa.2004.05.040. MR2086542

[17] E. Giné M., *Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms*, Ann. Statist. **3** (1975), no. 6, 1243–1266. MR0388663

[18] S. J. Gould, *Glow, big glowworm*, Natural History **12** (1986), 10–16 (English).

[19] S. J. Gould, *Bully for Brontosaurus: Reflections in Natural History*, W. W. Norton, 1991.

[20] G. G. Gregory, *Large sample theory for U-statistics and tests of fit*, Ann. Statist. **5** (1977), no. 1, 110–123. MR0433669

[21] J. Greslehner, *The b-adic diaphony of digital sequences*, Unif. Distrib. Theory **5** (2010), no. 2, 87–112. MR2781411

[22] J. Greslehner, *The b-adic diaphony of digital* $(\mathbf{T}, s)$*-sequences*, Unif. Distrib. Theory **6** (2011), no. 2, 1–12. MR2904034

[23] V. Grozdanov, *The generalized b-adic diaphony of the net of Zaremba-Halton over finite abelian groups*, C. R. Acad. Bulgare Sci. **57** (2004), no. 6, 43–48. MR2075788

[24] V. Grozdanov, *The weighted b-adic diaphony*, C. R. Acad. Bulgare Sci. **58** (2005), no. 5, 493–498. MR2154330

[25] V. Grozdanov, *The weighted b-adic diaphony*, J. Complexity **22** (2006), no. 4, 490–513, DOI 10.1016/j.jco.2005.11.008. MR2246893

[26] V. S. Grozdanov and E. G. Miteva, *On the dyadic diaphony of the Zaremba sequence*, C. R. Acad. Bulgare Sci. **53** (2000), no. 4, 5–8. MR1780010

[27] V. Grozdanov, E. Nikolova, and S. Stoilova, *Generalized b-adic diaphony*, C. R. Acad. Bulgare Sci. **56** (2003), no. 4, 23–30. MR1980679

[28] V. S. Grozdanov and S. S. Stoilova, *On the theory of b-adic diaphony*, C. R. Acad. Bulgare Sci. **54** (2001), no. 3, 31–34. MR1829550

[29] V. S. Grozdanov and S. S. Stoilova, *The b-adic diaphony* (English, with English and Italian summaries), Rend. Mat. Appl. (7) **22** (2002), 203–221 (2003). MR2041234

[30] V. S. Grozdanov and S. S. Stoilova, *On the irregularities of distribution*, C. R. Acad. Bulgare Sci. **55** (2002), no. 3, 15–18. MR1882673

[31] V. S. Grozdanov and S. S. Stoilova, *On the b-adic diaphony of the Roth net and generalized Zaremba net*, Math. Balkanica (N.S.) **17** (2003), no. 1-2, 103–112. MR2096244

[32] V. Grozdanov and S. Stoilova, *The general diaphony*, C. R. Acad. Bulgare Sci. **57** (2004), no. 1, 13–18. MR2032921

[33] P. Hellekalek and H. Leeb, *Dyadic diaphony*, Acta Arith. **80** (1997), no. 2, 187–196. MR1450924

[34] J. P. Imhof, *Computing the distribution of quadratic forms in normal variables*, Biometrika **48** (1961), 419–426. MR0137199

[35] P. Imkeller, *On exact tails for limiting distributions of U-statistics in the second Gaussian chaos*, Chaos Expansions, Multiple Wiener-Itô Integrals and Their Applications (Guanajuato, 1992), Probab. Stochastics Ser., CRC, Boca Raton, FL, 1994, pp. 179–204. MR1278044

[36] P. Kritzer and F. Pillichshammer, *The weighted dyadic diaphony of digital sequences*, Monte Carlo and Quasi-Monte Carlo Methods 2006, Springer, Berlin, 2008, pp. 549–560, DOI 10.1007/978-3-540-74496-2_32. MR2479245

[37] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Pure and Applied Mathematics, Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974. MR0419394

[38] H. Leeb, *Asymptotic properties of the spectral test, diaphony, and related quantities*, Math. Comp. **71** (2002), no. 237, 297–309 (electronic), DOI 10.1090/S0025-5718-01-01356-4. MR1863001

[39] E. L. Lehmann, *The Fisher, Neyman-Pearson theories of testing hypotheses: one theory or two?*, J. Amer. Statist. Assoc. **88** (1993), no. 424, 1242–1249. MR1245356

[40] I. Lirkov and S. Stoilova, *The b-adic diaphony as a tool to study pseudo-randomness of nets*, Numerical Methods and Applications, Lecture Notes in Comput. Sci., vol. 6046, Springer, Heidelberg, 2011, pp. 68–76, DOI 10.1007/978-3-642-18466-6_7. MR2794677

[41] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (eds.), *NIST Handbook of Mathematical Functions*, U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC; Cambridge University Press, Cambridge, 2010. With 1 CD-ROM (Windows, Macintosh and UNIX). MR2723248

[42] F. Pillichshammer, *Dyadic diaphony of digital sequences* (English, with English and French summaries), J. Théor. Nombres Bordeaux **19** (2007), no. 2, 501–521. MR2394899

[43] S. Pinker, *The Better Angels of Our Nature: The Decline of Violence in History and its Causes*, Penguin Books Limited, 2011.

[44] S. O. Rice, *Distribution of quadratic forms in normal random variables—evaluation by numerical integration*, SIAM J. Sci. Statist. Comput. **1** (1980), no. 4, 438–448, DOI 10.1137/0901032. MR610756

[45] V. Ristovska, *Irregularity of the distribution of the generalized van der Corput sequence*, C. R. Acad. Bulgare Sci. **58** (2005), no. 5, 487–492. MR2154329

[46] H. L. Royden, *Real Analysis*, 3rd ed., Macmillan Publishing Company, New York, 1988. MR1013117

[47] V. V. Sazonov, *Normal Approximation—Some Recent Advances*, Lecture Notes in Mathematics, vol. 879, Springer-Verlag, Berlin-New York, 1981. MR643968

[48] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1980. MR595165

[49] R. Seri, *A tight bound on the distance between a noncentral chi square and a normal distribution*, IEEE Communications Letters **19** (2015), no. 11, 1877–1880.

[50] J. Sheil and I. O'Muircheartaigh, *Statistical algorithms: Algorithm AS 106: The distribution of non-negative quadratic forms in normal variables*, Applied Statistics **26** (1977), no. 1, 92–98.

[51] G. R. Shorack, *Probability for Statisticians*, Springer Texts in Statistics, Springer-Verlag, New York, 2000. MR1762415

[52] S. S. Stoilova, *The b-adic diaphony of an arbitrary* (t,m,s)*-net*, Tech. Report E5-2004-107, Joint Institute for Nuclear Research, Dubna (Russia), 2004.

[53] S. S. Stoilova, *On the b-adic diaphony of the generalized Zaremba net*, Tech. Report E5-2005-96, Joint Institute for Nuclear Research, Dubna (Russia), 2005.

[54] A. W. van der Vaart, *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 3, Cambridge University Press, Cambridge, 1998. MR1652247

[55] P. Zinterhof, *Über einige Abschätzungen bei der Approximation von Funktionen mit Gleichverteilungsmethoden* (German), Österreich. Akad. Wiss. Math.-Naturwiss. Kl. S.-B. II **185** (1976), no. 1-3, 121–132. MR0501760

Dipartimento di Economia, Università degli Studi dell'Insubria, Via Monte Generoso 71, 21100 Varese, Italy – and – Center for Nonlinear and Complex Systems, Università degli Studi dell'Insubria, Via Valleggio 11, 22100 Como, Italy

E-mail address: raffaello.seri@uninsubria.it