Aequat. Math. 97 (2023), 683–705

0001-9054/23/040683-23
published online May 21, 2023
https://doi.org/10.1007/s00010-023-00958-x

Aequationes Mathematicae

Deviation from equidistance for one-dimensional sequences

Christian Weiss

Abstract. For a finite sequence $(x_i)_{i=1}^N$ in the unit interval, we introduce the gap ratio function which measures the size of the maximal gap length relative to all other gap lengths. This function (asymptotically) captures a lot of information about the degree of uniformity of the sequence. We discuss connections to the theories of dispersion, discrepancy, pair correlation statistics and covering numbers. Furthermore, we explicitly calculate the gap ratio function for some important classes of uniformly distributed sequences.

1. Introduction

A sequence $(x_i)_{i\in\mathbb{N}}$ in $[0,1)$ is called **uniformly distributed** if for any subinterval $[a,b] \subset [0,1)$

$$\lim_{N\to\infty} \left| \frac{\#\{x_1, x_2, \ldots, x_N\} \cap [a,b]}{N} - (b-a) \right| = 0$$

holds. This definition is equivalent to the so-called **star-discrepancy** defined by

$$D_N^*(x_i) := \sup_{b\in[0,1]} \left| \frac{\#\{x_1, x_2, \ldots, x_n\} \cap [0,b)}{N} - \lambda_1([0,b)) \right| \qquad (1)$$

converging to 0 for $N \to \infty$, where $\lambda_1([0,b)$ denotes the one-dimensional Lebesgue measure. The star-discrepancy is therefore often regarded as a quantitative measure for how uniformly distributed a sequence is. A famous result by Schmidt, [26], states that $D_N^*(x_i) \geq c\log(N)N^{-1}$ for any sequence $(x_i)_{i\in\mathbb{N}}$ and infinitely many $N \in \mathbb{N}$. Thereby the best-possible degree of uniformity for one-dimensional sequences was established. On the other hand, if $D_N^*(x_i) \leq \tilde{c}\log(N)N^{-1}$, then $(x_i)_{i\mathbb{N}}$ is called a **low-discrepancy sequence**. Moreover, for a finite set, i.e. $N \in \mathbb{N}$ fixed, it is straightforward to see that $D_N^*(x_i)$ attains its minimal possible value $\frac{1}{2N}$ if and only if $x_i = \frac{2i-1}{2N}$ for $i = 1, \ldots, N$, compare to Lemma 3.2. In other words, the star-discrepancy is

Birkhäuser

minimal if and only if the points are equidistant when they are not regarded as points on the unit interval but on the unit circle, that is $[0, 1)$ with interval ends glued. Due to its importance for e.g. optimization or numerical integration and its connections to many other areas of mathematics, discrepancy theory is nowadays a very active field of research; for more details we refer the reader to one of the many excellent textbooks, e.g. [9, 10, 19, 22]. If instead of intervals $[0, b)$ anchored at zero, we allow arbitrary intervals $[a, b)$ with $0 \leq a < b \leq 1$ for the supremum in (1), then we speak of the **(extreme) discrepancy**.

Another important way of measuring how uniformly distributed a sequence in $[0, 1)$ is, is the **dispersion** defined by

$$\mathrm{disp}(N, x_i) := \sup_{[a,b] \subset [0,1), [a,b) \cap (x_i)_{i=1}^N = \emptyset} b - a.$$

It is a lower bound for the discrepancy. Moreover, a finite sequence is equidistant if and only if $\mathrm{disp}(N, x_i)$ is $\frac{1}{N}$. The dispersion was originally introduced by Hlawka in [16] and later on generalized to a wider class of test sets than intervals by Rothe and Tichy in [24]. In dispersion theory, there are two main strands of research activities, namely the dependence of the dispersion on the number of points, see again e.g. [24], and the dependence on the dimension, see e.g. [25, 28].

In the one-dimensional setting, the dispersion in a sense only encapsulates very limited information about the sequence. To see this, we interpret it in terms of the so-called gaps of $(x_i)_{i=1}^N$: these are the distances between two neighboring elements of $(x_i)_{i=1}^N$ when regarded as elements on the unit circle. If we order the gap lengths by magnitude, i.e.

$$L_{\min,N}(x_i) := L_{1,N}(x_i) \leq L_{2,N}(x_i) \leq \ldots L_{N,N}(x_i) := L_{\max,N}(x_i),$$

then the dispersion corresponds to $L_{\max,N}(x_i)$, the maximal gap length of the finite sequence (while forgetting about all other gaps).

In this paper, we therefore suggest a new approach which uses information about all gaps. For that purpose, we introduce the **gap ratio function** which is for a finite sequence $(x_i)_{i=1}^N$ in $[0, 1)$ and for $0 < \alpha \leq 1$ defined by

$$C_\alpha^N(x_i) := \frac{L_{\max,N}(x_i)}{L_{\alpha N,N}(x_i)}$$

where $L_{\alpha N,N}(x_i) := L_{\lceil \alpha N \rceil,N}(x_i)$ if $\alpha N \notin \mathbb{N}$ and for $\alpha = 0$ as $C_0^N(x_i) := C_{\frac{1}{N}}^N(x_i)$. The gap ratio function is constant, $C_\alpha^N(x_i) \equiv 1$, if and only if all gaps have equal length. In other words, we then have $x_i = \frac{i-1}{N-1} + \beta$ for some $0 < \beta < \frac{1}{N-1}$ up to permutation.

When we consider sequences instead of finite sets, then there might exist gap lengths $L_1, \ldots, L_k$ which appear for all $N > N_0$, i.e. these gaps do not get split up at a later point in time. If there are only finitely many of these gaps with constant length, it is clear from the definition of the gap ratio function that $C_\alpha^N(x_i) \to \infty$ for all $1 > \alpha > 0$. If this situation occurs, we therefore exclude these finitely many gaps from the definition of $C_\alpha^N(x_i)$ for $N > N_0$ and only look at the gap ratio function of the gap lengths $\{L_{1,N}(x_i), L_{2,N}(x_i), \ldots, L_{N,N}(x_i)\} \setminus \{L_1, \ldots, L_k\}$. Why this is a reasonable approach becomes clear in the proof of Theorem 1.3.

In the remainder of this introduction, we will describe several properties of the gap function before we discuss its applications and connections to other concepts from uniform distribution theory. By that we hope to convince the reader that the gap ratio function is a useful tool to consider when analyzing equidistribution properties of one-dimensional sequences. At first, we state a lemma which explains how to get back the gap lengths from the function $C_\alpha^N(x_i)$. This shows that it really contains all the information about the gap lengths.

Lemma 1.1. *Let $(x_i)_{i=1}^N$ be a finite sequence in $[0,1)$ and let $C_\alpha^N(x_i)$ be its gap ratio function. Then*

$$L_{k,N}(x_i) = \frac{1}{C^N_{\frac{k}{N}}(x_i) \cdot \sum_{j=1}^N C^N_{\frac{j}{N}}(x_i)^{-1}}$$

for $k = 1, \ldots, N$.

For a sequence $(x_i)_{i=0}^N$ in $[0,1)$ let $\alpha_0^N := \sup_\alpha \left\{\alpha \,|\, C_\alpha^N(x_i) > C_1^N(x_i)\right\}$ be the share of gaps with non-maximal length. This allows us to derive a necessary condition for a sequence to have a (locally) bounded gap ratio function. As we will see, the latter property has important consequences e.g. for the pair correlation static of $(x_i)_{i\in\mathbb{N}}$

Proposition 1.2. *Let $(x_i)_{i\in\mathbb{N}}$ be a sequence in $[0,1)$ with maximal gap length $L_{\max,N}(x_i)$ for $N \in \mathbb{N}$. For any $M \in \mathbb{N}$ we have $1 - \alpha_0^M \leq \frac{1}{ML_{\max,M}(x_i)}$. If there exists an $\varepsilon > 0$ with $\alpha \geq \frac{1-\frac{1}{ML_{\max,M}(x_i)}}{1-\epsilon}$, then $C_\alpha^M(x_i) \leq \varepsilon^{-1}$.*

The remarkable aspect of Proposition 1.2 is that it only uses information about $L_{\max,N}(x_i)$ and the share of gaps of size at least $L_{\alpha N,N}(x_i)$ to derive information about the local behavior of $C_N^\alpha(x_i)$. The first part of the proposition yields an upper bound for the share of gaps with maximal length. If all gaps have equal length, which happens e.g. for some $M \in \mathbb{N}$ in the case of van der Corput sequences, see Sect. 2, then the bound is trivial, i.e. $1 - \alpha_0^M \leq 1$. However, we then have $C_\alpha^M(x_i) \equiv 1$ for all $\alpha > 0$. Moreover, a situation like that can only occur on rare occasions. In fact, the van der Corput in base 2 is the example where this happens most often and the share of N with this property is still bounded by $\log_2(N)/N$. In addition, for many examples of low-discrepancy sequences, e.g. Kronecker sequences or LS-sequences with $S \geq 1$, see again Sect. 2 for details, the equidistant case does never appear and the proposition always yields a non-trivial upper bound. If the share of gaps with non-maximal length is bounded more strictly than in the *worst* case, then we obtain an upper bound for $C_\alpha^N(x_i)$ which is valid for all N, i.e. $C_\alpha^N(x_i)$ is uniformly bounded in N.

The assumptions in Proposition 1.2 are in some sense easier to be fulfilled by low-discrepancy sequences: in fact, by applying the extreme discrepancy and the fact that $D_N^*(x_i) \leq D_N(x_i) \leq 2D_N^*(x_i)$ we see that the largest gap may be at most of size $c_N \log(N)/N$ for all N and c_N must lie in the interval $\frac{1}{\log(N)} \leq c_N \leq c$, where $c > 0$ only depends on the sequence. This heuristic argument is supported by our calculations in Sect. 2 and Example 2.1.

Another sufficient condition for the gap ratio function to be bounded on a subsequence $(N_i)_{i\in\mathbb{N}}$ is that the sequence has only three different gap lengths plus that none of the smaller gaps gets split up into two parts of equal size, whenever there are only two (and not three) different gap lengths. As in Proposition 1.2, we can also give an explicit upper bound on $C_\alpha^N(x_i)$ in this case but it is even global here.

Theorem 1.3. *Let the sequence $(x_i)_{i\in\mathbb{N}}$ have the three gap property and denote the occurring gap sizes at step N by $L_1(N) < L_2(N) < L_3(N)$. Moreover if only two gap lengths $L_2(N) < L_3(N)$ exist for some $N \in \mathbb{N}$, assume that $L_2(N+1)/L_1(N+1) \neq 2$. Then there exists a sequence $(N_j)_{j\in\mathbb{N}}$ of natural numbers with $C_\alpha^{N_j}(x_i) \leq 4$ for all $0 \leq \alpha \leq 1$.*

Theorem 1.3 can e.g. be applied in the case of van der Corput and Kronecker sequences which both possess the three gap property as we will explain in more details in Sect. 2. Moreover, the extra condition for the situation when only two different gap lengths occur cannot be ignored as is shown in Example 3.1.

Applications We will now present several applications of the gap ratio function to equidistribution theory. At first, we establish a simple connection to the closest alternative concept, namely dispersion.

Proposition 1.4. *If $C_\alpha^N(x_i) \leq C$ for all $\alpha > 0$ and all $N \in \mathbb{N}$, then there exists a $c > 0$ such that disp$(N, x_i) \leq \frac{c}{N}$ for all $N \in \mathbb{N}$.*

The converse of Proposition 1.4 is not true because there are sequences with $\mathrm{disp}(N, x_i) \leq \frac{c}{N}$ but $C_{\frac{1}{N}}^N(x_i) \to \infty$. To see this take any sequence with $\mathrm{disp}(N, x_i) \leq \frac{C}{N}$, e.g. the van der Corput sequence in base 2, and add elements $y_i = (10^{-i})^{10^{10}}$ for $i = 10^j$. To be more precise, define the first 9 elements of the sequence $(y_i)_{i\in\mathbb{N}}$ to be equal to the first 9 elements of the van der Corput sequence, $y_{10} = (10^{-1})^{10^{10}}$, the following 89 elements by the elements $x_{11}, \ldots, x_{99}$ of the van der Corput sequence, $y_{100} = (10^{-2})^{10^{10}}$ and so on.

Recall that the dispersion is a lower bound for the star-discrepancy. Thus $D_N^*(x_i) \to 0$ implies $\mathrm{disp}(N, x_i) \to 0$ for $N \to \infty$. However, $D_N^*(x_i) \to 0$ does not result in $C_\alpha^N(x_i)$ being bounded. In fact, we will even see in Sect. 2 that

$C_\alpha^N(x_i)$ is generically unbounded for a Monte Carlo sequence drawn from a uniform distribution despite the fact that its discrepancy converges to zero. This shows that a uniformly bounded gap ratio function $C_\alpha^N(x_i)$ is a rather special property of sequences.

As we have already discussed, the relation $C_\alpha^N(x_i) \equiv 1$ holds if and only if all gaps have equal length, i.e. up to permutation $x_i = \frac{i-1}{N-1} + \beta$ for some $0 < \beta < \frac{1}{N-1}$ holds. The latter is the case if and only if $D_N(x_i) = \frac{1}{N-1}$. Furthermore we have

$$D_N^*(x_i) = \frac{1}{2(N-1)} + \left|\beta - \frac{1}{2(N-1)}\right|.$$

We now describe the relation between the gap ratio function and the discrepancy in the general case. From Lemma 1.1 it follows that $L_{k,N}(x_i) < \frac{1}{N}$ if and only if

$$C_{\frac{k}{N}}^N(x_i) > \frac{N}{\sum_{j=1}^N C_{\frac{j}{N}}^N(x_i)^{-1}},$$

compare to Lemma 3.2. The smallest such k is denoted by $N_S^* = N_S$ and indicates the number of *small* gaps while N_L^* is the smallest k with $L_{k,N}(x_i) > \frac{1}{N}$ and $N_L^* = N+1$ if no such k exists. Thus, the number of *large* gaps is $N_L := N + 1 - N_L^*$. Note that $N_L + N_S \neq N$ if and only if there exist gaps of length $\frac{1}{N}$.

Proposition 1.5. *Let $(x_i)_{i=0}^N$ be a finite sequence in $[0,1)$ with gap ratio function $C_\alpha^N(x_i)$ and denote $\sum_{j=1}^N \frac{1}{C_{\frac{j}{N}}^N(x_i)^{-1}}$ by Σ_N. Then*

$$D_N^*(x_i) \leq \frac{1}{2(N+1)} + \max\left(\Sigma_N \sum_{k=1}^{N_S^*} C_{\frac{k}{N}}^N(x_i)^{-1} - \frac{2N_S^*-1}{2(N+1)}, \Sigma_N \sum_{k=1}^{N_L^*} C_{1-\frac{k}{N}}^N(x_i)^{-1} - \frac{2N_L^*-1}{2(N+1)}\right)$$

and

$$D_N^*(x_i) \geq \frac{1}{2}\max\left(\frac{2}{N+1} - \Sigma_N C_{\frac{1}{N}}^N(x_i)^{-1}, \Sigma_N C_{\frac{N}{N}}^N(x_i)^{-1}\right).$$

Both bounds are sharp.

Still, given that a sequence has only a finite number of different gap lengths and that these different gap length are in some sense, which is made precise in [32], equidistributed, better bounds for the star-discrepancy can be derived. For instance, it can even be proved under an extra condition concerning the mixture of gap lengths that certain sequences (van der Corput sequences, Kronecker sequences) have indeed the low-discrepancy property, see again [32].

Another concept from equidistribution theory which has links to the gap ratio function is the pair correlation statistic which is for a sequence $(x_i)_{i\in\mathbb{N}}$ in $[0,1)$ given by

$$F_N(s) := \frac{1}{N}\#\left\{1 \leq l \neq m \leq N:\ \|x_l - x_m\| \leq \frac{s}{N}\right\},$$

where the norm of a point $x \in \mathbb{R}$ is defined by

$$\|x\| := \min(x - \lfloor x \rfloor, 1 - (x - \lfloor x \rfloor)),$$

if $\lfloor x \rfloor$ denotes the floor bracket. In other words $\|\cdot\|$ stands for the distance of a number to the nearest integer. We say that a sequence $(x_i)_{i\in\mathbb{N}}$ has **Poissonian pair correlations**, if

$$\lim_{N\to\infty} F_N(s) = 2s \tag{2}$$

for all $s \geq 0$. Recently, a lot of effort has been made to prove or disprove the Poissonian pair correlation property for classes of sequences, see e.g. [4,11,20,21,34,35] to name only a few. In this paper, we contribute to the discussion by deriving a criterion how to read off from the gap ratio function that a sequence does not possess Poissonian pair correlations.

Theorem 1.6. *If there exists a sequence $(N_j)_{j\in\mathbb{N}}$ of natural numbers and $\varepsilon > 0, K > 1$ such that $C^{N_j}_{\alpha_{N_j}}(x_i) \leq K$ for $\alpha_{N_j} \leq 1/\sqrt{N_j(K+\varepsilon)}$, then $(x_i)_{i\in\mathbb{N}}$ does not have Poissonian pair correlations.*

First note that Theorem 1.6 does not hold without any further restrictions on α depending on K and N, because $L_{N,N} = L_{\max,N}$ and therefore $C^N_1(x_i) = 1$ for all sequences x_i. As is shown in Example 3.3, a condition on α like in Theorem 1.6, which allows for a certain number of exceptionally small gap lengths, cannot be completely avoided. From Theorem 1.6 we can immediately deduce the following corollary.

Corollary 1.7. *If there exists an $K > 0$ such that $C^{N_j}_{\alpha}(x_i) \leq K$ for a sequence $(N_j)_{j\in\mathbb{N}} \in \mathbb{N}$ and all $\alpha > 0$, then $(x_i)_{i\in\mathbb{N}}$ does not have Poissonian pair correlations.*

Proof. Given $N_j \in \mathbb{N}$ we choose $\alpha_{N_j} = \frac{1}{\sqrt{N_j(K+1)}}$ and can apply Theorem 1.6.

□

In contrast to Corollary 1.7, Theorem 1.6 allows that $L_{\alpha N,N}(x_i)$ is big for small α, where *small* depends on N. We can again see from Example 3.3 that the relaxed conditions from Theorem 1.6 indeed allow for the construction of examples which are not covered by Corollary 1.7. Moreover, Theorem 1.3 and Theorem 1.6 can be very concretely applied in the case of Kronecker and van der Corput sequences to deduce that they do not have Poissonian pair correlations. This should be compared to the more general result in [20].

Theorem 1.8. (Larcher, Stockinger, [20], Theorem 1) *Let $(x_i)_{i\in\mathbb{N}}$ be a sequence in $[0,1)$ with the following property: There is an $s \in \mathbb{N}$, positive real numbers $K, \gamma \in \mathbb{R}$, and infinitely many $N \in \mathbb{N}$ such that the point set $x_1, \ldots, x_N$ has a subset with $M \geq \gamma N$ elements, denoted by $x_{j_1}, \ldots, x_{j_M}$, which are contained in a set of points with cardinality at most KN having at most s different distances between neighbouring sequence elements, so-called gaps. Then $(x_i)_{i\in\mathbb{N}}$ does not have Poissonian pair correlations.*

Connection to Covering Numbers by Intervals Finally, we shortly discuss connections of the gap ratio function to the concept of covering numbers which is a prevalent topic in the measure-theoretic part of ergodic theory, compare [13]. The notion of covering numbers provides a measure how well disjoint orbits of an ergodic system (X, T, μ) cover X for some optimal subset $B \subset X$, see [7,18]. Since the sets in the general definition may be arbitrarily complicated from a topological point of view, these base sets B are sometimes restricted to (unions of) intervals, see [7,33].

Definition 1.9. Let $X = [0,1)$ be equipped with some measure μ and let T define an ergodic system on X. Then the **covering number by an interval** $F_1(T)$ is the supremum of all $z \in \mathbb{R}$ such that for every $h_0 \in \mathbb{N}$, there exists $h \geq h_0$ and an interval B such that

(i) $B, TB, \ldots, T^{h-1}B$ are disjoint,
(ii) $\mu(\cup_{i=0}^{h-1} T^i B) \geq z$.

If we regard $[0,1)$ as the unit circle with ends glued, $\mathbb{T}^1$, then a different viewpoint of the covering number by an interval for an admissible map T, i.e. $T([a,b]) \subset [T(a), T[b]]$ for all a, b with $T([a,b]) \neq \mathbb{T}^1$, is the following: Fix an element $x_0 \in [0,1)$, a number $N \in \mathbb{N}$ and consider its orbit $(x_i)_{i=0}^N = (T^i(x_0))$. The interval B in Definition 1.9 may then be chosen as the interval $[x_0, x_0+L)$, where $L = L_{1,N}(x_i)$ is the minimal gap length of the orbit. This maximizes the volume of B (given x) while making sure that the orbit of B is disjoint. Thus, $F_1(T)$ is given as $\sup_{x\in[0,1)}(\limsup_{N\to\infty} N \cdot L_{1,N}(x_i))$. An important class of examples of admissible maps are circle rotations.

This simple observation can be used as a definition for the covering by an interval for arbitrary sequences. If we interpret $x = (x_i)_{i=0}^\infty$ as to be given by $x_{i+1} = T(x_i)$ for some (potentially unknown) admissible map T, then the covering number by an interval is

$$F_1(x_i) := \limsup_{N\to\infty} N \cdot L_{1,N}(x_i).$$

The following criterion can now be established.

Proposition 1.10. *Let $(x_i)_{i\in\mathbb{N}}$ be an arbitrary sequence in $[0,1)$. If $F_1(x_i) > 0$, then $C_\alpha^N(x)$ is bounded for all $\alpha > 0$ and $N \in \mathbb{N}$. Therefore, $(x_i)_{i\in\mathbb{N}}$ does not have Poissonian pair correlations.*

Structure of the Paper This article is organized as follows: in Sect. 2 we show how to explicitly calculate $C_\alpha^N(x_i)$ for several classes of uniformly distributed sequences, including van der Corput sequences, Kronecker sequences and Monte Carlo sequences drawn from uniform distribution. These calculations sideline the general theory presented so far. Afterwards we will give proofs of our results in Sect. 3. Furthermore, we present the examples mentioned in the introduction.

2. Calculation of $C_\alpha^N(x_i)$ for certain classes of sequences

In this section, we calculate the gap ratio function $C_\alpha^N(x_i)$ for certain classes of sequences which have a special importance in equidstribution theory. We start by considering three different classes of low-discrepancy sequences, namely van der Corput sequences, Kronecker sequences and LS-sequences. While the first two satisfy the three gap property and the extra condition from Theorem 1.3, the latter does not fall into this class. Afterwards, we analyze the sequence $\{\sqrt{i}\}_{i\in\mathbb{N}}$ without perfect squares, where we denote by $\{\cdot\}$ the fractional part of a real number. On the one hand, the limiting distribution of its gap lengths is known according to [12]. As we will discuss, this yields interesting information about the asymptotic behavior of $C_\alpha^N(x_i)$. On the other hand, this was the first known sequence with Poissonian pair correlations, see [11]. This example will also underline that the Poissonian pair correlation property does not solely rely on the gap ratio function but some extra conditions like in Theorem 1.6 need to be satisfied. Finally, we discuss the asymptotic behavior of $C_\alpha^N(x_i)$ for uniformly distributed sequences. For that purpose we use results from statistics.

Van der Corput Sequences Van der Corput sequences are a classical class of low-discrepancy sequences. They are defined as follows: for an integer $b \geq 2$ the b-ary representation of $n \in \mathbb{N}$ is $n = \sum_{j=0}^{\infty} a_j(n)b^j$ with $a_j(n) \in \mathbb{N}$. The radical-inverse function is defined by $g_b(n) = \sum_{j=0}^{\infty} a_j(n)b^{-j-1}$ for all $n \in \mathbb{N}$. Then, the **van der Corput sequence in base** b is given by $x_i := g_b(i)$. For convenience we add $x_0 = 0$ as the zeroth element of a van der Corput sequence. It follows immediately that van der Corput sequences have at most three different gap lengths, which makes the calculation of $C_\alpha^N(x_i)$ particularly feasible. For a more precise description of the gap structure fix $b \in \mathbb{N}$ and let $N \in \mathbb{N}_{\geq 2}$ with $ab^{n-1} - 1 \leq N < (a+1)b^{n-1} - 1$ for $1 \leq a < b$. Then only the gap lengths

$$\begin{aligned} L_1 &= b^{-n}, \\ L_2 &= b^{-n+1} - ab^{-n}, \\ L_3 &= b^{-n+1} - (a-1)b^{-n} \end{aligned}$$

occur and their corresponding multiplicities are

$$\begin{aligned} N_1 &= N + 1 - b^{n-1}, \\ N_2 &= N + 1 - ab^{n-1}, \\ N_3 &= (a+1)b^{n-1} - 1 - N. \end{aligned}$$

Note that for $a = b - 1$, the lengths L_1 and L_2 coincide and therefore we only have two different gap lengths but the formulae for their multiplicities are still valid. Hence $C_\alpha^N(x_i)$ is for $N = ab^{n-1} + c$ with $1 \le a < b$ and $-1 < c < b^{n-1} - 1$, i.e. $N_3 \neq 0$, given by

$$C_\alpha^N(x_i) = \begin{cases} b - a + 1 & \alpha \le 1 - \frac{b^{n-1}}{N+1} \\ 1 + \frac{1}{b-a} & 1 - \frac{b^{n-1}}{N+1} < \alpha \le 2 - \frac{(a+1)b^{n-1}}{N+1} \\ 1 & \alpha > 2 - \frac{(a+1)b^{n-1}}{N+1}. \end{cases}$$

Similarly for $c = -1$, i.e. $N_3 = 0$, we have

$$C_\alpha^N(x_i) = \begin{cases} b - a & \alpha \le 1 - \frac{b^{n-1}}{N+1} \\ 1 & \alpha > 1 - \frac{b^{n-1}}{N+1}. \end{cases}$$

We see that $C_\alpha^N(x_i)$ is bounded everywhere by b and Corollary 1.7 applies. The limit $\lim_{N\to\infty} C_\alpha^N(x_i)$ however only exists in the trivial case $\alpha = 1$.

Moreover, we want to present here how to apply Proposition 1.2 to van der Corput sequences.

Example 2.1. Consider the van der Corput sequence $(x_i)_{i\in\mathbb{N}}$ in base $b > 2$. The purpose of this example is to show that $C_\alpha^N(x_i)$ is locally uniformly bounded (in N) for α close to 1 by only using information about multiplicities and the maximal gap length but not (explicitly) about the smaller gap lengths.

Therefore, we compare the actual share of non-maximal gaps

$$2 - (a+1)b^{n-1}/(N+1)$$

to

$$1 - \frac{1}{NL_{\max,N}(x_i)} = 1 - \frac{1}{(N+1)(b^{-n+1} - (a-1)b^{-n})}$$

which holds true for $N = ab^{n-1} + c$ with $1 \le a < b$ and $-1 < c < b^{n-1} - 1$. Let us at first consider the case $a > 1$ which implies $b > 2$. Since the respective quotient is monotonically increasing in $(N+1)$, we have

$$\frac{2 - \frac{(a+1)b^{n-1}}{N+1}}{1 - \frac{1}{(N+1)(b^{-n+1} - (a-1)b^{-n})}} < \frac{2 - \frac{a+1}{a}}{1 - \frac{1}{a-(a-1)/b}}.$$

This expression is again monotonically increasing in a and hence by setting $a = 2$ can be bounded by

$$\frac{1}{1 - \frac{1}{b-1}}.$$

If $a > 1$ and $c = -1$, then the quotient fulfills

$$\frac{1 - \frac{b^{n-1}}{N+1}}{1 - \frac{1}{(N+1)(b^{-n+1} - ab^{-n})}} = \frac{1 - \frac{1}{a}}{1 - \frac{1}{a - a^2/b}} = \frac{1}{1 - \frac{a}{(a-1)(b-a)}}.$$

We remark the trivial inequality $\frac{(b-a)(a-1)}{a} < b - 1$. If $a = 1$ and $c = -1$, then there is only one gap length, which is a trivial case. Hence it remains to consider $a = 1, c > -1$. This simplifies the quotient to

$$\frac{2 - \frac{2b^{n-1}}{N+1}}{1 - \frac{1}{(N+1)b^{-n+1}}} = 2.$$

Summing up, the bound $b - 1$ for $C_\alpha^N(x_i)$ for α close to α_0^N is obtained. This estimate is indeed independent of N and hence we know that $C_\alpha^N(x_i)$ is locally uniformly bounded in N by $b - 1$ for α close to α_0^N.

Kronecker Sequences For $z \in \mathbb{R}$ the corresponding Kronecker sequence is defined by $(x_i)_{i \in \mathbb{N}} = (\{iz\})_{i \in \mathbb{N}}$. The gap ratio function $C_\alpha^N(x_i)$ can be calculated explicitly for each $N \in \mathbb{N}$ with the help of the so-called Three Gap Theorem, originally proved in [27]. The situation is overall reminiscent of van der Corput sequences but only the explicit gap lengths and their multiplicities are different. We formulate the Three Gap Theorem here in a version which is suitable for our context because it explicitly describes gap lengths and their multiplicities, compare to e.g. [1,23,31]. We formulate it here similarly as in [1] (without the point 0) and for that purpose denote the continued fraction expansion of $z \in \mathbb{R} \backslash \mathbb{Q}$ by $[a_0, a_1, a_2, \ldots]$ and denote the corresponding sequence of convergents by $(p_n/q_n)_{n \in \mathbb{N}_0}$. Recall that

$$\begin{aligned} p_{-2} = 0, p_{-1} = 1, \quad p_n &= a_n p_{n-1} + p_{n-2}, n \geq 0, \\ q_{-2} = 1, q_{-1} = 0, \quad q_n &= a_n q_{n-1} + q_{n-2}, n \geq 0. \end{aligned}$$

Theorem 2.2. (Three Gap Theorem) *Let $(iz)_{i \in \mathbb{N}}$ be the Kronecker sequence of $z \in \mathbb{R} \backslash \mathbb{Q}$ and write $N \in \mathbb{N}$ uniquely as*

$$N = cq_n + q_{n-1} + r$$

with $1 \leq c \leq a_{n+1}$ and $0 \leq r < q_n$. Then the gaps between two adjacent terms in the set $\{\{iz\} : 1 \leq i \leq N\}$ that can appear have lengths

$$\begin{aligned} L_1 &= \|q_n z\|, \\ L_2 &= \|q_{n-1} z\| - cL_1, \\ L_3 &= L_1 + L_2, \end{aligned}$$

and their multiplicities are

$$N_1 = N - q_n,$$
$$N_2 = r,$$
$$N_3 = q_n - r.$$

By the theory of continued fractions, it is furthermore well-known that for a convergent q_n the inequalities

$$\frac{1}{(a_{n+1}+2)q_n} \leq \|q_n z\| \leq \frac{1}{a_{n+1}q_n}$$

hold. From Theorem 2.2, we can therefore deduce that $C_\alpha^N(x_i)$ is bounded if and only if the partial quotients are bounded. Still, it is possible to find a subsequence $(N_j)_{j\in\mathbb{N}}$ where $C_\alpha^{N_j}(x_i)$ is bounded by 2, because the dynamics behind the Kronecker sequence is that the smallest gap length gets subtracted from the largest when passing from N to $N+1$. This should be compared to the slightly weaker bound in Theorem 1.3. At the same time, there also exists a subsequence $(N_j')_{j\in\mathbb{N}}$ such that $C_\alpha^{N_j'}$ diverges to ∞ if and only if the sequence of partial quotients is unbounded. For all $0 < \alpha < 1$, the function $C_\alpha^N(x_i)$ never has a limit.

LS-Sequences LS-Sequences are another class of low-discrepancy sequences relatively recently discovered in [6]. As van der Corput and Kronecker sequences, they also have the finite gap property although the structure is a bit more complicated and five different gap lengths may occur. We introduce these sequences very similarly as it is done in [30]. They are based on using partitions as suggested in [17].

Definition 2.3. Let ρ denote a non-trivial partition of $[0,1)$. Then the ρ-**refinement** of a partition π of $[0,1)$, denoted by $\rho\pi$, is defined by subdividing all intervals of maximal length positively homothetically to ρ.

The resulting sequence of partitions is denoted by $\{\rho^n\pi\}_{n\in\mathbb{N}}$.

Definition 2.4. Let $L \in \mathbb{N}, S \in \mathbb{N}_0$ and β be the solution of $L\beta + S\beta^2 = 1$. An LS-**sequence of partitions** $\left\{\rho_{L,S}^n\pi\right\}_{n\in\mathbb{N}}$ is the successive ρ-refinement of the trivial partition $\pi = \{[0,1)\}$ where $\rho_{L,S}$ consists of $L+S$ intervals such that the first L intervals have length β and the successive S intervals have length β^2.

Thus, the partition $\left\{\rho_{L,S}^n\pi\right\}$ after n steps consists of intervals only of length β^n and β^{n+1}. Its total number of intervals is denoted by t_n, the number of intervals of length β^n by l_n and the number of intervals of length β^{n+1} by s_n. Explicit values for these numbers can be deduced from the relations

$$t_n = l_n + s_n, \quad l_n = Ll_{n-1} + Sl_{n-2}, \quad s_n = Ls_{n-1} + s_{n-2}.$$

With $Q = \sqrt{L^2 + 4S}$ this leads to the expressions

$$t_n = \tau_0 \beta^{-n} + \tau_1 (-S\beta)^n, \quad \tau_0 = \frac{L + 2S + Q}{2Q}, \quad \tau_1 = \frac{-L - 2S + Q}{2Q},$$
$$l_n = \lambda_0 \beta^{-n} + \lambda_1 (-S\beta)^n, \quad \lambda_0 = \frac{L + Q}{2Q}, \quad \lambda_1 = \frac{-L + Q}{2Q},$$
$$s_n = \sigma_0 \beta^{-n} + \sigma_1 (-S\beta)^n, \quad \sigma_0 = \frac{2S}{2Q}, \quad \sigma_1 = \frac{-2S}{2Q},$$

see also [2].[1] In other words, the sequence $(\xi_i)_{i=1}^{t_n}$ has s_n gaps of length β^{n+1} and l_n gaps of length β^n.

A specific ordering of the endpoints of the partition yields the LS-sequence of points.

Definition 2.5. Given an LS-sequence of partitions $\left\{\rho_{L,S}^n \pi\right\}_{n \in \mathbb{N}}$, the corresponding LS**-sequence of points** $(\xi^n)_{n \in \mathbb{N}}$ is defined as follows: let $\Lambda_{L,S}^1$ be the first t_1 left endpoints of the partition $\rho_{L,S}\pi$ ordered by magnitude. Given $\Lambda_{L,S}^n = \left\{\xi_1^{(n)}, \dots, \xi_{t_n}^{(n)}\right\}$ an ordering of $\Lambda_{L,S}^{n+1}$ is then inductively defined as

$$\begin{aligned}\Lambda_{L,S}^{n+1} = \{ & \xi_1^{(n)}, \dots, \xi_{t_n}^{(n)}, \\ & \psi_{1,0}^{(n+1)}(\xi_1^{(n)}), \dots, \psi_{1,0}^{(n+1)}(\xi_{l_n}^{(n)}), \dots, \psi_{L,0}^{(n+1)}(\xi_1^{(n)}), \dots, \psi_{L,0}^{(n+1)}(\xi_{l_n}^{(n)}), \\ & \psi_{L,1}^{(n+1)}(\xi_1^{(n)}), \dots, \psi_{L,1}^{(n+1)}(\xi_{l_n}^{(n)}), \dots, \psi_{L,S-1}^{(n+1)}(\xi_1^{(n)}), \dots, \psi_{L,S-1}^{(n+1)}(\xi_{l_n}^{(n)}) \Big\},\end{aligned}$$

where

$$\psi_{i,j}^{(n)}(x) = x + i\beta^n + j\beta^{n+1}, \qquad x \in \mathbb{R}.$$

Theorem 2.6. (Carbone, [6]) *If $L \geq S$, then the corresponding LS-sequence has low-discrepancy.*

For the parameters $S = 0$ and $L = b$, the corresponding LS-sequences are the classical van der Corput sequences. In [29] it was moreover proven that LS-sequences for $S = 1$ coincide with symmetrized Kronecker sequences up to permutation and that neither van der Corput sequences nor Kronecker sequences occur for $S \geq 2$. For $t_n - (S-1)l_{n-1} < N \leq t_{n+1} - (S-1)l_n$, we write N in the unique form $N = t_n - (S-1)l_{n-1} + a l_{n-1} + b l_n + c$, where at first $0 \leq a \leq (S-1)$ is chosen as big as possible and afterwards $0 \leq b \leq L$ is chosen as big as possible, too. This implies that $0 \leq c < S - 1$ if $a < S - 1$

[1] Note that there is a typo in the formulae for σ_0 and σ_1 in the paper [2].

and $0 \le c < L$ otherwise. Then $(\xi_i)_{i=1}^N$ has gaps of lengths

$$\begin{aligned}
L_1 &= \beta^{n+1},\\
L_2 &= \max((S-a)\beta^{n+1}, 0),\\
L_3 &= \max((S-a-1)\beta^{n+1}, 0),\\
L_4 &= \beta^n - (b+1)\beta^{n+1},\\
L_5 &= \beta^n - b\beta^{n+1}.
\end{aligned}$$

We use the short-hand notations $S_a^1 := \max(\min(S-1-a, 1), 0)$ and $S_a^2 := \max(\min(a-(S-2), 1), 0)$ to obtain shorter formulae for the corresponding multiplicities

$$\begin{aligned}
N_1 &= N - t_n - Sl_{n-1} - S_a^1 l_{n-1},\\
N_2 &= S_a^1 \cdot c,\\
N_3 &= S_a^1 \cdot (l_{n-1} - c),\\
N_4 &= S_a^2 \cdot c,\\
N_5 &= S_a^2 \cdot (l_n - c) + S_a^1 \cdot l_n.
\end{aligned}$$

Note that $L_1 = L_2$ for $a = (S-1)$ and that $N_2 = 0$ in this case. Furthermore the relations $N_4 + N_5 = l_n$ and $N_2 + N_3 \in \{0, l_{n-1}\}$ hold. Thus $C_\alpha^N(x_i)$ is uniformly bounded by β and does not converge for any $0 \le \alpha < 1$.

The sequence $\{\sqrt{i}\}$ The sequence $\{\sqrt{i}\}_{i\in\mathbb{N}}$ is uniformly distributed and is one of the few explicitly known examples with Poissonian pair correlations (after the perfect squares have been removed), see [11]. In the paper [12], Elkies and McMullen gave a wonderful description of the limiting distribution of the gaps of $\{\sqrt{i}\} \mod 1$ (with perfect squares included).

Theorem 2.7. (Elkies, McMullen, [12], Theorem 1.1) *The gap distribution for the sequence* $\{\sqrt{i}\}_{i\in\mathbb{N}}$ *is given by a continuous function*

$$F(t) = \begin{cases} 6/\pi^2 & t \in [0, 1/2] \\ F_2(t) & t \in [1/2, 2], \\ F_3(t) & t \in [2, \infty) \end{cases},$$

where the functions $F_2(t), F_3(t)$ *are given in the following way. For* $\frac{1}{2} \le x$, *let* $r = \frac{1}{2x}$ *and define*

$$\psi(r) := \tan^{-1} \frac{2r-1}{\sqrt{4r-1}},\ \frac{1}{2} \le x \le 2, \quad \alpha = \frac{1}{2}(1 - \sqrt{1-4r}),\ 2 \le x.$$

Then

$$F_2(x) = \frac{1}{2\zeta(2)}\left(\frac{2}{3}(4r-1)^{3/2}\psi(r) + (1-6r)\log r + 2r - 1\right),$$

$$F_3(x) = \frac{1}{2\zeta(2)}\left(4(1-4\alpha)(1-\alpha)^2\log(1-\alpha) - 2(1-2\alpha)^3\log(1-2\alpha) - 2\alpha^2\right).$$

That is for any interval $[a,b] \subset [0,\infty)$ we have

$$\frac{\#\{L_{j,N} : L_{j,N} \in [a/N, b/N]\}}{N} \to \int_a^b F(t)\,\mathrm{d}t.$$

The limiting behavior of the α-percentile can be determined by solving the implicit equation

$$\int_a^\infty F(t)\mathrm{d}t = 1-\alpha.$$

Similarly, the maximal gap length can be estimated by solving

$$\int_a^\infty F(t)\mathrm{d}t = \frac{1}{N}.$$

Thereby it is theoretically possible to describe the asymptotic behavior of $C_\alpha^N(x_i)$.

Remark 2.8. For us it is important to note that Theorem 2.7 does not depend on whether perfect squares are removed or not, because their share goes to 0, compare also to [11], but the sequence has Poissonian pair correlations only in the latter case. Therefore, the Poissonian pair correlation property cannot in general be read off from the asymptotic behavior of $C_\alpha^N(x_i)$ but only if certain extra conditions as in Theorem 1.6 are satisfied.

Monte Carlo Sequences It is well-known that Monte Carlo sequences $(X_i)_{i\in\mathbb{N}}$ drawn from a uniform distribution on $[0,1)^d$ with $d \in \mathbb{N}$ generically have Poissonian pair correlations, see [15]. Corollary 1.7 therefore implies that $C_\alpha^N(X_i)$ is generically unbounded for $d=1$. Here, we quantify the limiting behavior of $C_\alpha^N(X_i)$ by giving a lower bound.

We start by describing the general theoretical framework. Suppose $(X_i)_{i=1}^N$ is a finite random sample drawn from a continuous distribution with common density $f(x)$ and cumulative distribution function $F_X(x)$. The so-called *order statistics* $X_1^* \leq X_2^* \leq \ldots \leq X_N^*$ is obtained by ordering the sample in ascending order. Then the density function of X_i^* is given by

$$f_{X_i^*}(x) = \frac{N!}{(i-1)!(N-i!)}F(x)^{i-1}(1-F(x))^{N-i}f(y).$$

If the X_i are drawn from a uniform distribution, it follows that the distribution of the X_i^* is a beta-distribution $\mathrm{Beta}(\alpha,\beta)$ with parameters $\alpha = i$ and $\beta = N-i$. In fact, we are in our situation not interested in the X_i^* themselves but in the gaps $Y_i = X^*_{i+1} - X^*_i$ which are also often called spacings in this context, compare to [8]. Since our analysis is restricted to the uniform distribution on $[0, 1)$ with ends glued, by rotation we may furthermore without loss of generality assume that $X^*_0 = 0, X^*_{N+1} = 1$. Then the Y_i follow a beta distribution with parameters $\alpha = 1$ and $\beta = N$ and the task is to calculate the distribution of the order statistics Y^*_i.

The joint probability distribution of the ordered gaps drawn from uniform distribution is known for a long time, see e.g. [8], Exercise 6.4.4, but the individual distributions of the $L_{j,N}(X_i)$ has only been obtained rather recently in [3].

Proposition 2.9. ([3], Lemma 1) *The distribution of* $L_{j,N}(X_i)$ *is for* $0 \le j \le N$ *given by*

$$P[L_{j,N}(X_i) > x] = \begin{cases} 0 & x(N+2-j) \ge 1 \\ (-1)^{j-1}(N+1)\binom{N}{j-1} \sum_{k=m}^{j} \frac{(-1)^{k-1}}{N+2-k}(1 - x(N+2-k))^n & x \in I_{m,N}, \end{cases}$$

where $I_{m,N} = \left[\frac{1}{N+3-m}, \frac{1}{N+2-m}\right)$ *for* $m = 2, \ldots, N+1$ *and* $I_{1,N} = [0, \frac{1}{N+1})$. *Furthermore* $P[L_{j,N}(X_i) \le x] = 0$ *for* $x \le 0$.

From Proposition 2.9 and the logarithmic growth of the harmonic series, the expected asymptotic growth of the relevant $L_{\cdot,n}(X_i)$ can easily be deduced as

$$E[L_{\max,N}(X_i)] = E[L_{N,N}(X_i)] \sim \frac{\log(N-1)}{N-1},$$
$$E[L_{\alpha N,N}(X_i)] \sim \frac{-\log(1-\alpha)}{N-1}, \quad 0 < \alpha < 1,$$

see [3], Section 3. It follows from the results in [5], Section 2 and Section 3 that for fixed $0 < \alpha < 1$ the correlation and covariance of $L_{\max,N}(X_i)$ and $L_{\alpha N,N}(X_i)$ is negative, and that those of $L_{\max,N}(X_i)$ and $1/L_{\alpha N,N}(X_i)$ are positive for fixed $0 \le \alpha < 1$. Hence we get by using Jensen's inequality

$$\begin{aligned} & E\left[\frac{L_{\max,N}(X_i)}{L_{\alpha N,N}(X_i)}\right] \\ & = E[L_{\max,B}(X_i)] \cdot E\left[\frac{1}{L_{\alpha N,N}(X_i)}\right] + \operatorname{Cov}\left[L_{\max,N}(X_i), \frac{1}{L_{\alpha N,N}(X_i)}\right] \\ & > E[L_{\max,N}(X_i)] \cdot E\left[\frac{1}{L_{\alpha N,N}(X_i)}\right] \\ & > \frac{E[L_{\max,N}(X_i)]}{E[L_{\alpha N,N}(X_i)]} \\ & \sim \frac{\log(N)}{-\log(1-\alpha)}. \end{aligned}$$

Note that the inequality is compatible neither with Corollary 1.7 nor with Theorem 1.6. In fact, it may seem like the lower bound can be significantly improved by avoiding the cancellation of the covariance and the use of Jensen's inequality. This is however not the case, because it is well-known, see e.g. [12], that the limiting distribution of the gap lengths is exponential, i.e.

$$\frac{\#\left\{L_{j,N}(X_i) : L_{j,N}(X_i) \in [a/N, b/N]\right\}}{N} \to \int_a^b e^{-t}\,\mathrm{d}t.$$

Then the deterministic asymptotic behavior

$$\frac{L_{\max,N}(X_i)}{L_{\alpha N,N}(X_i)} \sim \frac{\log(N)}{-\log(1-\alpha)}$$

can be derived by the integral method described in the context of the sequence $\{\sqrt{i}\}$.

3. Proofs of results

In this section, we collect the proofs of the result mentioned in Sect. 1. Moreover, we give examples which show that the extra assumptions made throughout the paper cannot be (completely) avoided.

Although it is rather straightforward, we also include the proof of Lemma 1.1, which explains how to obtain the gap lengths from the gap ratio function.

Proof of Lemma 1.1. The equation

$$\sum_{j=1}^{N} L_{j,N}(x_i) = 1$$

is equivalent to

$$\sum_{j=1}^{N} \frac{L_{\max,N}(x_i)}{C^N_{\frac{j}{N}}(x_i)} = 1.$$

Hence the claim follows because $L_{k,N}(x_i) = \frac{L_{\max,N}(x_i)}{C^N_{\frac{k}{N}}(x_i)}$. □

If the maximal gap lengths and their share satisfy an extra condition, then $C^N_\alpha(x_i)$ is locally bounded for α close to 1 and uniformly in N.

Proof of Proposition 1.2. The inequality

$$(1-\alpha_0^M)NL_{\max,M}(x_i) \leq 1$$

implies the claimed inequality for $1-\alpha_0^M$. Fixing an $\alpha > 0$ and then summing up all gap lengths yields

$$(1-\alpha)ML_{\max,M}(x_i) + \alpha ML_{\alpha M,M}(x_i) \geq 1. \quad (3)$$

Note that equality holds if only two different gap lengths occur and if α equals the share of small length gaps. In any case, (3) yields

$$L_{\alpha M,M}(x_i) \geq \frac{1}{\alpha M} - \frac{(1-\alpha)ML_{\max,M}(x_i)}{\alpha M}.$$

Again we have equality for only two gap lengths. Finally we obtain

$$C_\alpha^M(x_i) = \frac{L_{M,M}(x_i)}{L_{\alpha M,M}(x_i)} \leq \left(1 + \frac{1}{\alpha}\left(\frac{1}{ML_{\max,M}(x_i)} - 1\right)\right)^{-1}.$$

Assuming that $\alpha \geq \frac{1-\frac{1}{ML_{\max,M}}}{1-\varepsilon}$ yields the desired inequality for $C_\alpha^M(x_i)$. □

Next we show the stronger property that $C_\alpha^{N_j}(x_i)$ is globally bounded (in α) by 4 for some subsequence $(N_j)_{j\in\mathbb{N}}$ if there are only three different gap lengths for $(x_i)_{i\in\mathbb{N}}$ and if these gaps satisfy the extra assumption mentioned in Theorem 1.3 regarding the case when only two different gap lengths exist.

Proof of Theorem 1.3. At first recall that we exclude constant gaps from the definition of the gap ratio function. Hence we may without loss of generality assume that all gaps get split up at some point. Let $L_1(N) < L_2(N) < L_3(N)$ be the three gap lengths at step N and $N_1(N), N_2(N), N_3(N)$ be the respective multiplicities. At first we assume $N_3(N) > 1$. If we move from N to $N+1$, then one of the

- a gap of length $L_3(N)$ gets split up into two gaps of length $L_1(N)$,
- or a gap of length $L_3(N)$ gets split up into two gaps of length $L_2(N)$,
- or a gap of length $L_3(N)$ gets split up into one gap of length $L_1(N)$ and one of length $L_2(N)$,
- or a gap of length $L_2(N)$ gets split up into two gaps of length $L_1(N)$.

If the first possibility occurs, then $L_3(N) = 2L_1(N)$ and $C_\alpha^N(x_i)$ is bounded by 2. Hence we may without loss of generality exclude this case.

If the fourth option occurs, then $L_2(N) = 2L_1(N)$. Thus the fourth option either happens k times in a row until $N_2(N+k) = 0$ or $L_3(N) \leq 4L_1(N)$. Therefore, we assume the first. For $N' = N+k+1$, a gap of length $L_3(N)$ would get split up into one gap of $L_2(N') \geq L_3(N)/2$ and one of $L_1(N)$ or into two of length $L_2(N')$ because we do not allow at this stage that $L_1(N)$ gets split up into two gaps of size $L_1(N)/2$ as that would mean $L_2(N+k+1) = 2L_1(N+k+1)$. In any case we have $C_\alpha^N(x_i) \leq 3$.

Hence we may assume that either only the second *or* the third option occurs since they are mutually exclusive. In any case, we are at some point left with no more gaps of length $L_3(N)$ and we only have gaps of length $L_2(N')$ and gaps of $L_1(N)$. We call this point in time T. If there is only one gap size then the gap ratio function is constant and equal to 1. Therefore a gap of length $L_2(N')$ gets split up

- either into two of $L_1(N)$ and consequently $C_\alpha^N(x_i) \leq 2$,
- or into one of length $L_1(N)$ and one of length $L'(T+1) < L_1(N)$ which implies $C_\alpha^T(x_i) < 2$,
- or into one gap of length $L_1(N)$ and one of length $L'(T+1)$ with $L_2(N+k+1) > L'(T+1) > L_1(N)$

because again we do not allow to split $L_1(N)$ into two gaps of the same size. In the third case we are left with gaps of sizes $L_3(T+1) = L_2(N+k+1) < L_3(N), L_2(T+1) < L_2(N)$ and $L_1(T+1) = L_1(N)$. Thus we are in the same situation as in the beginning of the proof but $C_\alpha^{T+1}(x_i)$ is bounded by a smaller value than $C_\alpha^N(x_i)$. A reduction of length like this can only happen finitely many times and so there must exist an $N^* > T$ with $C_\alpha^{N^*}(x_i) \leq 2$ as desired. This finishes the proof of Theorem 1.3. □

The extra assumption for the situation where only two gap lengths occur cannot be avoided as the following example shows.

Example 3.1. The claim of Theorem 1.3 does not necessarily hold without the extra assumption $L_2(N+1)/L_1(N+1) \neq 2$. To see that assume that $N_3(N) = 1$. Then split up all gaps of length $L_2(N)$ into two of equal length. So we are left with $2(N-1)$ gaps of length $L_2(N)/2$ and one of length $L_3(N)$. Iteratively we can split up each gap of length $L_2(N)/2$ into 2^{M_1-1} gaps of length $L_1(N)/2^{M_1}$ until we at some point diminish the gap of length $L_3(N)$ by $L_1(N)/2^{M_1}$. Afterwards we can again reduce the size of the small gaps by an arbitrary factor 2^{M_2}. By choosing $M_1, M_2, \ldots$ we can make sure that $C_\alpha^N(x_i)$ grows beyond any given upper bound.

By a similar argument as in Lemma 1.1, also the bound on the dispersion from Proposition 1.4 can be established.

Proof of Proposition 1.4. As $C_\alpha^N(x_i) \leq C$ for all $\alpha > 0$, we have $L_{\max,N}(x_i) \leq CL_{\frac{1}{N},N}(x_i)$. Thus $NL_{\max,N}(x_i) \leq CNL_{\frac{1}{N},N}(x_i) \leq C$. This means that the largest gap size goes to 0 with rate of convergence $\frac{1}{N}$ and so does the dispersion. □

The basic relation between $C_\alpha^N(x_i)$ and the star-discrepancy from Proposition 1.5 can be established by using the following well-known formula for explicitly calculating the star-discrepancy of a finite sequence in dimension 1, see e.g. [22], Theorem 2.6.

Lemma 3.2. *If* $0 \leq x_1 \leq x_2 \leq \ldots \leq x_N$, *then*

$$D_N^*(x_i) = \frac{1}{2N} + \max_{1 \leq n \leq N} \left| x_n - \frac{2n-1}{2N} \right|$$

Proof of Proposition 1.5. At first, we establish the upper bound. According to Lemma 3.2, in order to calculate the star-discrepancy of $(x_i)_{i=0}^N$, the expression

$$D(n) := \left| x_n - \frac{2n-1}{2N} \right|$$

needs to be maximized. We have

$$x^*_{N_L} \leq \sum_{j=1}^{N} \frac{1}{C^N_{\frac{j}{N}}(x_i)^{-1}} \sum_{k=1}^{N^*_L} C^N_{1-\frac{k}{N}}(x_i)^{-1}$$

and

$$x^*_{N_S} \geq \sum_{j=1}^{N} \frac{1}{C^N_{\frac{j}{N}}(x_i)^{-1}} \sum_{k=1}^{N^*_S} C^N_{\frac{k}{N}}(x_i)^{-1}$$

by Lemma 1.1. Here equality holds if and only if $x_0 = 0$ and all large gaps respectively all small gaps are subsequent elements of the ordered version of $(x_i)_{i\in\mathbb{N}_0}$. Given only the information from the gap ratio function $C^N_\alpha(x_i)$, then $D(n)$ is thus maximized in one of the two cases just described and the claim follows. Obviously, the bound is sharp in the cases described here.

To see the lower bound, we consider the extreme discrepancy $D_N(x_i)$ instead of the star-discrepancy. According to Lemma 1.1, the first term in the max brackets addresses the fact that two of the N points are the left and right endpoints of the smallest gap, while the second term refers to the fact that no point lies in the interior of the largest gap. The claim for the star discrepancy then follows from $D^*_N(x_i) \geq \frac{1}{2} D_N(x_i)$, see e.g. [22], Proposition 2.4. If all gaps have the same length, then the expression on the right hand side of the inequality is $\frac{1}{2(N+1)}$ which is equal to the star-discrepancy. Therefore also this bound is sharp. □

If the gap ratio function is bounded or only satisfies the weaker assumptions of Theorem 1.6, then the sequence $(x_i)_{i\in\mathbb{N}}$ cannot possess Poissonian pair correlations.

Proof of Theorem 1.6. We may without loss of generality assume that the conditions are fulfilled for all $N \in \mathbb{N}$. From

$$\frac{L_{\max,N}(x_i)}{L_{\alpha_N N,N}(x_i)} \leq K$$

it follows that

$$L_{i,N}(x_i) \leq L_{\max,N}(x_i) \leq K L_{\alpha_N N,N}(x_i) \quad i \leq N.$$

Hence we obtain

$$\begin{aligned} 1 &= \sum_{j=1}^{N\alpha_N} L_{j,N}(x_i) + \sum_{j=N\alpha_N+1}^{N} L_{j,N}(x_i) \\ &\leq N\alpha_N L_{\alpha_N N,N}(x_i) + N(1-\alpha_N)KL_{\alpha_N N,N}(x_i) \\ &= NL_{\alpha_N N,N}(x_i)(\alpha_N + K(1-\alpha_N)) \\ &< NL_{\alpha_N N,N}(x_i)K. \end{aligned}$$

Therefore

$$\frac{1}{N}\#\left\{1 \leq l \neq m \leq N : \|x_l - x_m\| \leq \frac{\frac{1}{K}}{N}\right\} \leq \alpha_N^2 N \leq \frac{1}{K+\varepsilon}$$

and $(x_i)_{i\in\mathbb{N}}$ cannot fulfill (2) for $s \leq \frac{1}{K}$. □

Indeed, the share of exceptionally small gaps which is allowed in Theorem 1.6 does not harm the Poissonian pair correlation property as it becomes clear from the following example.

Example 3.3. We have already seen in Remark 2.8 that the Poissonian pair correlation property does not only depend on the gap ratio function. Here we give another example. Let $(x_i)_{i\in\mathbb{N}}$ be a sequence with Poissonian pair correlations. By applying a rotation we may without loss of generality assume that $x_1 = 0$. Like after the discussion of Proposition 1.4 let y_i be defined by adding elements $y_i = (10^{-i})^{10^{10}}$ for $i = 10^j$ to $(x_i)_{i\in\mathbb{N}}$ and denote the added elements by y_i'. Now fix $s \geq 0$. Since $(x_i)_{i\in\mathbb{N}}$ is uniformly distributed, see e.g. [14], there are $(10^{-i})^{10^{10}}N + o(N)$ elements of $(x_i)_{i\in\mathbb{N}}$ in the interval $[0,(10^{-i})^{10^{10}}]$. It follows that

$$\begin{aligned} &\#\left\{1 \leq l \neq m \leq N \ : \ \|y_l - y_m\| \leq \frac{s}{N}\right\} = \\ &\quad \#\left\{1 \leq l \neq m \leq N - \lfloor \log_{10}(N) \rfloor \ : \ \|x_l - x_m\| \leq \frac{s}{N}\right\} + \\ &\quad \#\left\{1 \leq l \neq m \leq \lfloor \log_{10}(N) \rfloor \ : \ \|y_l' - y_m'\| \leq \frac{s}{N}\right\} + \\ &\quad \#\left\{1 \leq l \leq N - \lfloor \log_{10}(N) \rfloor, 1 \leq m \leq \lfloor \log_{10}(N) \rfloor \ : \ \|x_l - y_m'\| \leq \frac{s}{N}\right\} \\ &\quad = 2s(N - \lfloor \log_{10}(N)) \rfloor \\ &\quad + o(N) + \lfloor \log_{10}(N) \rfloor^2 + o(N) + \lfloor \log_{10}(N) \rfloor^2 + o(N) \\ &\quad = 2sN + o(N). \end{aligned}$$

As a consequence $(y_i)_{i\in\mathbb{N}}$ has Poissonian pair correlations but its gap ratio function is unbounded.

Remark 3.4. Example 2.1 shows the existence of sequences with unbounded $C_\alpha^N(x_i)$ which do not have Poissonian pair correlations, because the sequence therein is not uniformly distributed.

Finally, we consider the covering number by intervals and show that $F_1(T) < 1$ implies that the sequence does not have Poissonian pair correlations. Although, we could almost directly apply Corollary 1.7, we give an independent proof here.

Proof of Proposition 1.10. Let $\alpha > 0$ and $0 < \varepsilon < F_1(x)$ be arbitrary. Choose N_0 such that we may assume $L_{1,N}(x_i) > \frac{F_1(x_i)-\varepsilon}{N}$ without loss of generality for all $N \geq N_0$ instead of a subsequence $(N_j)_{j\in\mathbb{N}}$. Thus $C_\alpha^N(x_i)$ is bounded.

Now let $s < F_1(x_i)$ and $\varepsilon < \frac{F_1(x_i)-s}{2}$ and choose N_0 as above. We obtain

$$\|x_l - x_m\| > \frac{F_1(x_i)-\varepsilon}{N} > \frac{s}{2N}$$

for all $l, m \geq N_0$ and hence $F_N(\frac{s}{2}) = 0$ for all $N > N_0$. This means that $(x_i)_{i\in\mathbb{N}}$ does not have Poissonian pair correlations. □

Funding Information Open Access funding enabled and organized by Projekt DEAL.

Data Availability Statement The author declares that all data supporting the findings of this study are available within the article.

References

[1] Alessandri, P., Berthé, V.: Three distance theorems and combinatorics on words. Enseign. Math. (2) **44**, 103–132 (1998)

[2] Aistleitner, C., Hofer, M., Ziegler, V.: On the uniform distribution modulo 1 of multidimensional ls-sequences. Annali di Matematics Pura ed Applicata **193**, 1329–1344 (2014)

[3] Bairamov, I., Berred, A., Stepanov, A.: Limit results for ordered uniform spacings. Stat. Pap. **51**, 227–240 (2010)

[4] Becher, V., Carton, O., Cunningham, I.: Low discrepancy sequences failing Poissonian pair correlations. Arch. Math. **113**, 169–178 (2019)

[5] Barton, D.E., David, F.N.: Some notes on ordered random variables. J. R. Stat. Soc. B **18**(1), 79–94 (1956)

[6] Carbone, I.: Discrepancy of ls-sequences of partitions and points. Annali di Matematics Pura ed Applicata **191**, 819–844 (2012)
[7] Chekhova, N.: Covering numbers of rotations. Theoret. Comput. Sci. **230**, 97–116 (2000)
[8] David, H.A., Nagaraja, H.N.: Order Statistics. Wiley Series in Probability and Statistics (2003)
[9] Dick, J., Pillichshammer, F.: Digital Nets and Sequences. Cambridge University Press, Cambridge (2010)
[10] Drmota, M., Tichy, R.: Sequences, Discrepancies, and Application. Lecture Notes in Mathematics, vol. 1651. Springer, Berlin (1997)
[11] El-Baz, D., Marklof, J., Vinogradov, I.: The two-point correlation function of the fractional parts of $\sqrt{n}$ is Poisson. Proc. AMS **143**(7), 2815–2828 (2015)
[12] Elkies, N., McMullen, C.: Gaps in $\sqrt{n} \bmod 1$ and ergodic theory. Duke Math. J. **123**(1), 95–139 (2004)
[13] Ferenczi, S.: Systems of finite rank. Colloq. Math. **73**, 35–65 (1997)
[14] Grepstad, S., Larcher, G.: On pair correlation and discrepancy. Arch. Math. **109**, 143–149 (2017)
[15] Hinrichs, A., Kaltenböck, L., Larcher, G., Stockinger, W., Ulrich, M.: On a multi-dimensional Poissonian pair correlation concept and uniform distribution. Monatsh. Math. **190**, 333–352 (2019)
[16] Hlawka, E.: Abschätzung von trigonometrischen Summen mittels diophantischer Approximationen. Österreich. Akad. Wiss. Math.-Naturwiss. Kl. S.-B. **II**(185), 43–50 (1976)
[17] Kakutani, S.: A problem on equidistribution on the unit interval $[0, 1[$. Measure Theory (Proc. Conf. Oberwolfach, 1975), Lecture Notes in Mathematics, 541:369–375, (1975)
[18] King, J.L.: Joining-rank and the structure of finite-rank mixing transformations. J. Analyse Math. **51**, 182–227 (1988)
[19] Kuipers, L., Niederreiter, H.: Uniform Distribution of Sequences. Wiley, New York (1974)
[20] Larcher, G., Stockinger, W.: Some negative results related to Poissonian pair correlation problems. Discrete Math. **343**(2), 555 (2020)
[21] Lutsko, C., Sourmelidis, A., Technau, N.: Pair correlation of the fractional parts of αn^{θ}. to appear in: Journal of the European Mathematical Society, (2021)
[22] Niederreiter, H.: Random Number Generation and Quasi-Monte Carlo Methods. Number 63 in CBMS-NSF Series in Applied Mathematics, SIAM, Philadelphia, (1992)
[23] Polanco, G., Schultz, D., Zaharescu, A.: Continuous distributions arising from the three gap theorem. Int. J. Number Theory **12**(7), 1743–1764 (2016)
[24] Rote, G., Tichy, R.: Quasi-monte Carlo methods and the dispersion of point sequences. Math. Comput. **23**(8–9), 9–23 (1996)
[25] Rudolf, D.: An upper bound of the minimal dispersion via delta covers. In: Contemporary computational mathematics-a celebration of the 80th birthday of Ian Sloan. volume 1, 2, pp. 1099–1108. Springer, Cham (2018)
[26] Schmidt, W.M.: Irregularities of distribution vii. Acta Arith **21**, 45–50 (1972)
[27] Sos, V.: On the distribution mod 1 of the sequence $n\alpha$. Ann. Univ. Sci. Budapest, Eötvös Sect. Math., 1:127–134, (1958)
[28] Ullrich, M., Vybiral, J.: An upper bound on the minimal dispersion. J. Complex. **45**, 120–126 (2018)
[29] Weiß, C.: On the classification of ls-sequences. Uniform Distribution Theory **13**(2), 83–92 (2018)
[30] Weiß, C.: Interval exchange transformations and low-discrepancy. Annali di Matematica **198**, 399–410 (2019)
[31] Weiß, C.: Deducing three gap theorem from rauzy-veech induction. Revista Colombiana de Matematicas **54**, 31–37 (2020)
[32] Weiß, C.: Some connections between discrepancy, finite gap properties and pair correlations. Monatshefte für Mathematik **199**, 909–927 (2022)

[33] Weiß, C.: Systems of rank one, explicit Rokhlin towers, and covering numbers. Arch. Math. **118**(2), 181–193 (2022)
[34] Weiß, C.: Remarks on the pair correlation statistic and lattice point counting. to appear in: Contributions to Algebra and Geometry, (2023)
[35] Weiß, C., Skill, T.: Sequences with almost Poissonian pair correlations. J. Number Theory **236**, 116–127 (2022)

Christian Weiß
Ruhr West University of Applied Sciences: Hochschule Ruhr West
Mühlheim an der Ruhr
Germany
e-mail: christian.weiss@hs-ruhrwest.de

Received: March 7, 2022
Revised: April 18, 2023
Accepted: April 18, 2023