The Pair Correlation Function of Low-Discrepancy Sequences with Small Stochastic Error Terms

Anja Schmiedt, Christian Weiß

November 21, 2022

Abstract

There is no known example of a low-discrepancy sequence which possesses Poissonian pair correlations. This is in some sense rather surprising, because low-discrepancy sequences always have $\beta$-Poissonian pair correlations for all $0 < \beta < 1$ and are therefore arbitrarily close to having Poissonian pair correlations (which corresponds to the case $\beta = 1$). In this paper, we further elaborate on the closeness of the two notions. We show that Kronecker sequences for badly approximable $\alpha$ with an arbitrary small uniformly distributed stochastic error term generically have $\beta = 1$-Poissonian pair correlations.

1 Introduction

According to a famous theorem of Weyl, [Wey16], a sequence $(x_n)_{n \in \mathbb{N}}$ in $[0,1]$ is uniformly distributed if and only if for all integers $r \in \mathbb{Z} \setminus \{0\}$ it holds that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} e(rx_n) = 0,$$

where $e(\cdot) := \exp(2\pi i \cdot)$. The result established an important link between uniform distribution theory and exponential sums, which are a central tool in analytic number theory. A classical way to quantify the degree of uniformity of $(x_n)_{n \in \mathbb{N}}$ is the discrepancy, which is defined by

$$D_N(x_n) := \sup_{B \subset [0,1]} \left| \frac{1}{N} \# \{(x_i|1 \leq i \leq N) \cap B\} - \lambda(B) \right|,$$

where the supremum is taken over all intervals $B = [a,b) \subset [0,1)$ and $\lambda(\cdot)$ denotes the Lebesgue measure. It is well-known that a sequence $(x_n)_{n \in \mathbb{N}}$ is uniformly distributed if and only if $D_N(x_n) \to 0$ for $N \to \infty$, compare [Nie92].
The so-called Erdős–Turan inequality may be regarded as a quantitative version of Weyl’s theorem and states that for any positive \( m \in \mathbb{N} \) we have

\[
D_N(x_n) \leq \frac{6}{m+1} + \frac{4}{\pi} \sum_{i=1}^{m} \left( \frac{1}{n} - \frac{1}{m+1} \right) \left| \frac{1}{N} \sum_{n=1}^{N} e(hx_n) \right|,
\]

see [KN74], Theorem 2.5, Chapter 2.

If a sequence \((x_n)_{n \in \mathbb{N}} \in [0,1)\) exhibits a discrepancy of order

\[
D_N(x_n) = \mathcal{O}(N^{-1}(\log N)),
\]

where \( \mathcal{O}(\cdot) \) denotes the Landau symbol, then it is called a low-discrepancy sequence. In fact, this is the best possible rate of convergence by the work of Schmidt, [Sch72]. A wide variety of low-discrepancy sequences is known, ranging from more classical ones, like van der Corput sequences and Kronecker sequences, see e.g. [KN74], to more recent ones as the classes of examples in [Car12] or [Wei19].

In this paper, we are mainly interested in Kronecker sequences which are for \( \alpha \in \mathbb{R} \setminus \mathbb{Q} \) defined by \( x_n := \{n\alpha\} \), where \( \{\cdot\} \) denotes the fractional part of a number.

Another concept to quantify the degree of uniformity of a sequence \((x_n)_{n \in \mathbb{N}}\) was introduced by Rudnik and Sarnak in [RS98]. It is based on the pair correlation function defined by

\[
F_N(s) := \frac{1}{N} \# \left\{ 1 \leq k \neq l \leq N : \|x_k - x_l\| \leq \frac{s}{N} \right\},
\]

where \( \|\cdot\| \) is the distance of a number from its nearest integer. The sequence \((x_n)_{n \in \mathbb{N}}\) is said to have Poissonian pair correlations if

\[
\lim_{N \to \infty} F_N(s) = 2s
\]

for all \( s > 0 \). This definition was generalized by Nair and Pollicott in [NP07] to \( \beta \)-Poissonian pair correlations (where the case \( \beta = 1 \) corresponds to the original Poissonian pair correlation property). A sequence \((x_n)_{n \in \mathbb{N}} \in [0,1)\) has \( \beta \)-PPC for \( 0 \leq \beta \leq 1 \) if

\[
\lim_{N \to \infty} F_N^\beta(s) := \lim_{N \to \infty} \frac{1}{N^{2-\beta}} \# \left\{ 1 \leq k \neq l \leq N : \|x_k - x_l\| \leq \frac{s}{N^\beta} \right\} = 2s
\]

for all \( s > 0 \). If a sequence has \( \beta_1 \)-PPC, then it also has \( \beta_2 \)-PPC for all \( 0 \leq \beta_2 < \beta_1 < 1 \) according to [HZ21], Theorem 4. Generalizing results from [GL17] and [ALP18], it was shown in [Ste20] that \( \beta \)-PPC imply uniform distribution for any \( \beta \geq 0 \). Vice versa, a sequence of independent, uniformly distributed random variables \((X_n)_{n \in \mathbb{N}}\) generically has \( \beta \)-PPC for all \( 0 \leq \beta \leq 1 \). In fact, an even stronger statement is known for low-discrepancy sequences.
Theorem 1.1 (Theorem 1.1, [Wei22]). Every low-discrepancy sequence \((x_n)_{n \in \mathbb{N}}\) has \(\beta\)-PPC for all \(0 \leq \beta < 1\).

However, the above mentioned examples of low-discrepancy all do not have 1-PPC, see [LS20] and [WS22]. More generally, it seems to be challenging to find explicit examples of sequences with 1-PPC, and only few ones are known by now, see e.g. [EBMV15], [LST21].

Theorem 1.1 may however be interpreted that low-discrepancy sequences fail to have 1-PPC as closely as possible. For a class of Kronecker sequences, we further elaborate on this interpretation by showing that an arbitrarily small stochastic distortion of these sequences generically implies 1-PPC, see Theorem 1.2 below.

Recall that a number \(\alpha \in \mathbb{R}\) is called badly approximable if there exists a \(c = c_\alpha > 0\) such that for all \(p \in \mathbb{Z}, q \in \mathbb{Z} \setminus \{0\}\) it holds that

\[
\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^2}
\]

or equivalently

\[
\lim \inf_{q \to \infty} \|q\alpha\| = c > 0.
\]

It is a standard result from analytic number theory that \(\alpha\) is badly approximable if and only if all partial quotients in the continued fraction expansion of \(\alpha\) are bounded.

**Theorem 1.2.** Let \(\alpha \in \mathbb{R}\) be a badly approximable number and let \((X_i)_{i \in \mathbb{N}}\) be a sequence of independent, identically distributed random variables which are uniformly distributed on \([0, 1]\). Furthermore let \(\varepsilon > 0\) be arbitrary. Then the sequence of random variables \((Y_n)_{n \in \mathbb{N}} := (\{n\alpha + \varepsilon X_n\})_{n \in \mathbb{N}}\) generically has Poissonian pair correlations.

In our view, Theorem 1.2 yields new insight in three different ways: First, it adds a new aspect to the interpretation that certain Kronecker sequences (as prominent examples of low-discrepancy sequences) are as close as possible to having Poissonian pair correlations. Second, it constitutes the first non-trivial examples of a pair of two sequences which both do not posses 1-PPC but their sum does.\(^1\) Third, our proof of Theorem 1.2 in Section 3\(^1\) whose structure is inspired by the ones of [Mar07], Theorem 2.3, and [HKL+19], Theorem 1, strongly refines the technique used in the mentioned references. Therefore, it might be the base for further examples of 1-PPC sequences to follow in the future. Indeed, we even conjecture that Theorem 1.2 might hold for any low-discrepancy sequence for the following reason: despite that the proof of Theorem 1.2 relies on the properties of Kronecker sequences and badly approximable numbers, the conjecture seems to be reasonable, because we can show a relevant intermediate step of the proof for arbitrary low-discrepancy sequences. As this result is of independent interest, we state it here separately.

\(^1\)A trivial example would be the following: consider a sequence \((x_n)_{n \in \mathbb{N}}\) with 1-PPC and split it up into two sequences \((y_n)_{n \in \mathbb{N}}, (z_n)_{n \in \mathbb{N}}\), where \(y_n\) is equal to \(x_n\) for even indices and \(z_n\) for odd indices while all other elements of \(y_n\) and \(z_n\) are 0.
Proposition 1.3. Let \((z_n)_{n \in \mathbb{N}}\) be a low-discrepancy sequence, \(\varepsilon > 0\) and \((X_n)_{n \in \mathbb{N}}\) be a sequence of independent random variables which are uniformly distributed on \([0, 1]\). For arbitrary \(s > 0\), let \(F_N(s)\) be the pair correlation function of \((Y_n)_{n \in \mathbb{N}} = (\{z_n + \varepsilon X_n\})_{n \in \mathbb{N}}\). Then the expected value \(\mathbb{E}[F_N(s)]\) converges to \(2s\) for \(N \to \infty\).

The subsequent sections are organized as follows: in Section 2 we collect some rather general results in four lemmas. They are all important for proving Theorem 1.2 but can be formulated separately and are used several times in the remainder of this paper. Afterwards, we give the rather lengthy proof of Theorem 1.2 and of Proposition 1.3 in Section 3.

Acknowledgment. The authors would like to thank Stefan Steinerberger for fruitful discussions on important aspects of this paper.

2 Preparatory Results

In this section, we collect the mainly technical results which we need to prove our main theorem. Since we are in our context not interested in numerically optimal bounds, we do not try to optimize the constants here. This might be up for future research. The first lemma calculates an expected value which will play a central role in the remainder of this paper and establishes the main difference and challenge in comparison to [Mar07] and [HKL+19] due to the involvement of sin-terms.

Lemma 2.1. Let \(r, r' \in \mathbb{Z} \setminus \{0\}\). Furthermore let \((X_n)_{n \in \mathbb{N}}\) be a sequence of independent, identically distributed random variables, which are uniformly distributed on \([0, 1]\). In the case \(r \neq \pm r'\) it follows for \(k, l, m, n \in \mathbb{N}\) with \(k \neq l\) and \(m \neq n\) and any \(\varepsilon > 0\) that

\[
\mathbb{E}[\varepsilon(r(X_k - X_l) + r'\varepsilon(X_m - X_n))]
\]

\[
= \begin{cases} 
\frac{1}{(r+r')\pi\varepsilon}\sin((r + r')\pi\varepsilon)\frac{1}{r\varepsilon}\sin(r\pi\varepsilon)\frac{1}{r'\varepsilon}\sin(r'\pi\varepsilon) & \text{if } k = m, l \neq n \\
\frac{1}{(r+r')\pi\varepsilon}\sin((r + r')\pi\varepsilon)\frac{1}{r\varepsilon}\sin(r\pi\varepsilon)\frac{1}{r'\varepsilon}\sin(r'\pi\varepsilon) & \text{if } k \neq m, l = n \\
\frac{1}{(r-r')\pi\varepsilon}\sin((r - r')\pi\varepsilon)\frac{1}{r\varepsilon}\sin(r\pi\varepsilon)\frac{1}{r'\varepsilon}\sin(r'\pi\varepsilon) & \text{if } k = n, l \neq m \\
\frac{1}{(r-r')\pi\varepsilon}\sin((r - r')\pi\varepsilon)\frac{1}{r\varepsilon}\sin(r\pi\varepsilon)\frac{1}{r'\varepsilon}\sin(r'\pi\varepsilon) & \text{if } k \neq n, l = m \\
\frac{1}{(r+r')\pi\varepsilon}\sin((r + r')\pi\varepsilon)^2 & \text{if } k = m, l = n \\
\frac{1}{(r-r')\pi\varepsilon}\sin((r - r')\pi\varepsilon)^2 & \text{if } k = n, l = m \\
\frac{1}{(r\pi\varepsilon)^2}\sin(r\pi\varepsilon)^2 & \text{else.}
\end{cases}
\]

If \(r = r'\), then all expressions of the form \(\frac{1}{(r-r')\pi\varepsilon}\sin((r - r')\pi\varepsilon)\) need to be replaced by 1 in the formula above. Similarly, if \(r = -r'\), then all expressions of the form \(\frac{1}{(r+r')\pi\varepsilon}\sin((r + r')\pi\varepsilon)\) need to be replaced by 1.
Proof. In the cases \( k \neq m, l \neq n \) and \( k = m, l = n \), the result follows by simple integration and trigonometrical arguments. If \( k = m, l \neq n \), integration leads to
\[
\mathbb{E}[e(\varepsilon(X_k - X_l) + r'\varepsilon(X_m - X_n))] = \\
\frac{\sin((r + r')\pi \varepsilon)}{(r + r')\pi \varepsilon} (\cos((r + r')\pi \varepsilon) + i \sin((r + r')\pi \varepsilon)) \\
\times \frac{\sin(r\pi \varepsilon)}{r\pi \varepsilon} (\cos(r\pi \varepsilon) + i \sin(r\pi \varepsilon)) \\
\times \frac{\sin(r'\pi \varepsilon)}{r'\pi \varepsilon} (\cos(r'\pi \varepsilon) + i \sin(r'\pi \varepsilon))
\]
Using the standard angel sum equations \( \sin(x+y) = \sin(x) \cos(y) + \sin(y) \cos(x) \) and \( \cos(x+y) = \cos(x) \cos(y) - \sin(x) \sin(y) \), we arrive the assertion follows.

The remaining cases follow similarly. \( \square \)

Remark 2.2. If \( r' = 0 \) and \( r \neq 0 \), then the expected value is equal to \( \frac{\sin(r\pi \varepsilon)^2}{(r\pi \varepsilon)^2} \).

There will also appear exponential sums of the Kronecker part of the sequence. This is also the point where we make use of the fact that \( \alpha \) is badly approximable.

Lemma 2.3. For all badly approximable \( \alpha \in \mathbb{R} \) and \( 0 \leq \delta \leq \frac{1}{2} \), there exists \( C_{\alpha, \delta} > 0 \) such that for all \( r, N \in \mathbb{N} \) we have
\[
\left| \sum_{j=1}^{N} e(rj\alpha) \right| \leq C_{\alpha, \delta} r^{1/2-\delta} N^{1/2+\delta}
\]
Proof. Since the left hand side is always \( \leq N \), the inequality holds for \( r \geq N \). In order to prove the assertion for \( r < N \), we show the stronger inequality
\[
\left| \sum_{j=1}^{N} e(rj\alpha) \right| \leq C_{\alpha} r.
\]
We have
\[
\left| \sum_{j=1}^{N} e(rj\alpha) \right| \leq \frac{2}{|1 - \exp(2\pi i r\alpha)|}.
\]
As \( \alpha \) is badly approximable there exists a \( \frac{1}{\sqrt{5}} \geq \hat{C}_{\alpha} > 0 \) with
\[
|r\alpha - p| \geq \hat{C}_{\alpha} \frac{1}{r}
\]
for all \( r \in \mathbb{N} \) and \( p \in \mathbb{Z} \). We may without loss of generality assume that even \( \frac{1}{4} > \hat{C}_{\alpha} \). Therefore,
\[
|1 - \exp(2\pi i r\alpha)| \geq \sin(2\pi \hat{C}_{\alpha}/r) > \hat{C}_{\alpha} \pi \frac{1}{r}
\]
which implies the claim with \( C_{\alpha} = \frac{1}{\pi \hat{C}_{\alpha}} \). \( \square \)
Combining these results, we get the following inequality.

**Lemma 2.4.** For all $0 < \delta \leq \frac{1}{2}$, all $\varepsilon > 0$ and all badly approximable $\alpha \in \mathbb{R}$ we have

$$
\sum_{r \in \mathbb{Z} \setminus \{0\}} \left| \sum_{k=1}^{N} \sum_{l=1}^{N} E[e(r(\varepsilon(X_k - X_l) + k\alpha - l\alpha))] \right| \\
\leq C_{\alpha, \delta, \varepsilon} N^{1+\delta}
$$

for some $C_{\alpha, \delta, \varepsilon} \in \mathbb{R}$ which only depends on $\alpha, \delta$ and $\varepsilon$ but not on $r$.

**Proof.** By Remark 2.2, Lemma 2.3 and the fact that $\sum_{k=1}^{\infty} \frac{1}{\pi^r}$ is a convergent series, we obtain

$$
\sum_{r \in \mathbb{Z} \setminus \{0\}} \left| \sum_{k=1}^{N} \sum_{l=1}^{N} E[e(r(\varepsilon(X_k - X_l) + k\alpha - l\alpha))] \right| \\
\leq \tilde{C}_{\alpha, \delta} \sum_{r=1}^{\infty} \left| \frac{2}{\pi^{2r/\varepsilon^2}} \sin(\pi r \varepsilon)^2 N^{1+\delta} r^{1-\delta} \right| \\
\leq \tilde{C}_{\alpha, \delta, \varepsilon} N^{1+\delta} \sum_{r=1}^{\infty} \left| \frac{\sin(\pi r \varepsilon)}{r^{1+\delta}} \right| \\
\leq \tilde{C}_{\alpha, \delta, \varepsilon} \zeta(1+\delta) N^{1+\delta},
$$

where $\zeta(\cdot)$ denotes the Riemann zeta function. The assertion follows with $C_{\alpha, \delta, \varepsilon} = \tilde{C}_{\alpha, \delta, \varepsilon} \zeta(1+\delta)$. \(\square\)

Finally, for the cases in Lemma 2.1 where two of the variables $X_k, X_l, X_m, X_n$ coincide, we need to use a bound on a certain series which is in our view also of independent interest.

**Lemma 2.5.** For all $r' \in \mathbb{Z} \setminus \{0\}$ and all $0 \leq \sigma < 1$ we have

$$
\sum_{r \in \mathbb{N}, \ r \neq \pm r'} \left| \frac{1}{r} \right|^{\sigma} \frac{1}{|r + r'|^2} \leq (2 + 3\zeta(2)) \frac{1}{|r'|^{2\sigma}}.
$$

**Proof.** We need to show that

$$
\sum_{r \in \mathbb{N}, \ r \neq \pm r'} \left| \frac{r'}{r} \right|^{\sigma} \frac{1}{|r + r'|^2}
$$

is uniformly bounded. At first we split the sum into three parts

$$
\sum_{\substack{r < |r'|/2, \ r \neq r'}} \left| \frac{r'}{r} \right|^{\sigma} \frac{1}{|r + r'|^2} + \sum_{\substack{|r'|/2 \leq r \leq |r'|, \ r \neq r'}} \left| \frac{r'}{r} \right|^{\sigma} \frac{1}{|r + r'|^2} + \sum_{\substack{r \geq |r'|, \ r \neq r'}} \left| \frac{r'}{r} \right|^{\sigma} \frac{1}{|r + r'|^2}
$$

6
The first sum consists of less than $|r'|/2$ terms of size at most $|r'|^α/|r'|/2|^2$ and is thus strictly bounded by 2. The first factor in the second sum is at most $2^2$ and thus the sum can be bounded by $2ζ(2)$. Finally, for the third sum, the factor $|r'|$ is bounded by 1 and therefore this sum can be bounded by $ζ(2)$.

3 Sums of non-Poissonian Sequences

In this section we turn to the proof of Theorem 1.2 that uses the intermediate result from Proposition 1.3 that does hold for arbitrary low-discrepancy sequences. The structure of the proof of our main result is essentially similar to the one of Theorem 2.3 in [Mar07] and Theorem 1 in [HKL+19]. However, we face additional technical challenges because the sequence under consideration has a deterministic and a stochastic part with a comparably complicated expected value, see Lemma 2.1.

A main step within the proof of Theorem 1.2 is to apply Chebyshev’s inequality in order to obtain convergence in probability. Therefore, Proposition 1.3 gives an estimate on the expected value of the pair correlation function first. We now come to its proof.

Proof of Proposition 1.3. By definition we have

$$E(F_N(s)) = E\left(\frac{1}{N} \# \left\{ 1 \leq l \neq m \leq N : \|z_l - z_m + ε \cdot (X_l - X_m)\| \leq \frac{sN}{N} \right\} \right).$$

If we regard the expression $ε(X_l - X_m)$ as the difference of two independent, uniformly distributed random variables on $[0, ε]$, then its density is given by

$$f(x) = \frac{1}{ε} \left(1 - \frac{|x|}{ε}\right) I_{[-ε,ε]}(x).$$

For fixed $1 \leq l \leq N$ we consider the expected value of

$$F_N^l(s) := \frac{1}{N} \# \left\{ 1 \leq m \leq N : m \neq l, \|z_l - z_m + ε \cdot (X_l - X_m)\| \leq \frac{s}{N} \right\}.$$ 

Then,

$$E(F_N^l(s)) = \frac{1}{N} \sum_{m \neq l} \int_{z_m - z_l - s/N}^{z_m - z_l + s/N} f(x)dx \int_{z_m - z_l - s/N}^{z_m - z_l + s/N} f(x)dx$$

Note that the corresponding proof in [HKL+19] uses the fact that $ε = 1$ when calculating the expected value. Therefore, the integration bounds only need to be considered modulo 1. This explains the technical difference between their proof and our.
For the moment, we only look at the positive part of the sum
\[
\sum_{m \neq l, \; 0 \leq z_m - z_l \leq \frac{s}{N} + \varepsilon} \int_{z_m - z_l - s/N}^{z_m - z_l + s/N} f(x) \, dx
\]
\[
= -\frac{1}{2} \sum_{m \neq l, \; 0 \leq z_m - z_l \leq \frac{s}{N} + \varepsilon} \left[ \left( 1 - \frac{z_m - z_l + s/N}{\varepsilon} \right)^2 - \left( 1 - \frac{z_m - z_l - s/N}{\varepsilon} \right)^2 \right]
\]
\[
= 2s \cdot \frac{1}{\varepsilon} \sum_{m \neq l, \; 0 \leq z_m - z_l \leq \frac{s}{N} + \varepsilon} \left( 1 - \frac{z_m - z_l}{\varepsilon} \right)
\]
\[
:= I_N
\]

By the Koksma-Hlawka inequality, see e.g. [Nie92], Theorem 2.9, we have
\[
\left| I_N - \int_{z_l}^{z_l + s/N + \varepsilon} \left( 1 - \frac{x - z_l}{\varepsilon} \right) \, dx \right| \leq CD_N(z_n),
\]
with \(C\) being a constant independent of \(N\), or equivalently
\[
\left| I_N - \frac{\varepsilon}{2} \left( 1 - \frac{s^2}{\varepsilon^2 N^2} \right) \right| \leq CD_N(z_n),
\]
which is (very importantly) independent of \(l\). Finally, we consider both the positive and the negative part of the sum and all \(l \in \{1, \ldots N\} \). As \((z_n)_{n \in \mathbb{N}}\) is a low-discrepancy sequence we can hence deduce
\[
\mathbb{E}(F_N(s)) = \frac{2s}{N} \cdot N \cdot \frac{2s}{\varepsilon} \cdot \frac{\varepsilon}{2} \left( 1 - \frac{s^2}{\varepsilon^2 N^2} + O\left( \frac{\log(N)}{N} \right) \right),
\]
which has limit \(2s\) for \(N \to \infty\), as claimed. \(\square\)

All our preparatory results have now paved the way to the proof of Theorem 1.2.

Proof of Theorem 1.2: As we want to apply Chebychev’s inequality, we need to calculate the variance of \(F_N(s)\). We closely follow here the approach in [Mar07], although the technical details of our arguments are a lot more involved. Using the Poisson summation formula we write
\[
F_N(s) = \frac{1}{N} \sum_{1 \leq k \neq l \leq N} \sum_{q \in \mathbb{Z}} I \left( \frac{(\varepsilon(X_k - X_l) + \{(k-l)\alpha\} + q)N}{2s} \right)
\]
\[
= \frac{2s}{N^2} \sum_{1 \leq k \neq l \leq N} \sum_{r \in \mathbb{Z}} \mathcal{F}I \left( \frac{2sr \varepsilon}{N} \right) e(r(\varepsilon(X_k - X_l) + \{(k-l)\alpha\}))
\]
\[
= \frac{2s}{N^2} \sum_{1 \leq k \neq l \leq N} \sum_{r \in \mathbb{Z}} \mathcal{F}I \left( \frac{2sr \varepsilon}{N} \right) e(r(\varepsilon(X_k - X_l) + \{(k-l)\alpha\}))
\]
where $\mathcal{F}I(\xi) = \frac{\sin(\xi \xi)}{\pi \xi}$ for $\xi \neq 0$ and $\mathcal{F}I(0) = 1$ is the Fourier transform of the indicator function $I(\cdot)$ of the interval $[-1/2, 1/2]$. Thus, it follows analogously to [HKL+19]

\[
\mathbb{E} \left[ (F_N(s) - \mathbb{E}(F_N(s)))^2 \right] = \frac{(2s)^2}{N^4} \sum_{r, r' \in \mathbb{Z} \setminus \{0\}} \sum_{1 \leq k, l, m, n \leq N, \ k \neq \ell, l \neq m, n \neq n} \mathcal{F}I \left( \frac{2sr \varepsilon}{N} \right) \mathcal{F}I \left( \frac{2sr' \varepsilon}{N} \right) \\
\times \mathbb{E}[e(r\varepsilon(X_k - X_l) + \{k\alpha\} - \{l\alpha\}) + r'(\varepsilon(X_m - X_n) + \{m\alpha\} - \{n\alpha\})]
\]

\[
= \frac{(2s)^2}{N^4} \sum_{r, r' \in \mathbb{Z} \setminus \{0\}} \sum_{1 \leq k, l, m, n \leq N, \ k \neq \ell, l \neq m, n \neq n} \mathcal{F}I \left( \frac{2sr \varepsilon}{N} \right) \mathcal{F}I \left( \frac{2sr' \varepsilon}{N} \right) \\
\times \mathbb{E}[e(r\varepsilon(X_k - X_l) + r'\varepsilon(X_m - X_n))] \\
\times e(r\{\{k\alpha\} - \{l\alpha\}) + r'(\{m\alpha\} - \{n\alpha\})].
\]

If \( \varepsilon \in \mathbb{Z} \) then the proof is finished because \( \{k\alpha + \varepsilon X_k\}_{k \in \mathbb{N}} \) would be uniformly distributed and the proof in [HKL+19] applies. Therefore, we may without loss of generality assume that \( \varepsilon \notin \mathbb{Z} \). We then split up the sum into several parts. According to Lemma 2.1 there are two main cases to distinguish.

1. Case \( r \neq \pm r' \):

For \( r \neq \pm r' \) we split up the sum into the cases which occur in Lemma 2.1. At first, we consider the case \( k = m, l = n \), i.e., the sum

\[
\frac{(2s)^2}{N^4} \sum_{r, r' \in \mathbb{Z} \setminus \{0\}} \sum_{1 \leq k, l, m, n \leq N, \ k \neq \ell, l \neq m, n \neq n} \mathcal{F}I \left( \frac{2sr \varepsilon}{N} \right) \mathcal{F}I \left( \frac{2sr' \varepsilon}{N} \right) \\
\times \mathbb{E}[e(r\varepsilon(X_k - X_l) + r'\varepsilon(X_m - X_n))] \\
\times e(r\{\{k\alpha\} - \{l\alpha\}) + r'(\{m\alpha\} - \{n\alpha\})].
\]

We now focus the inner sum over \( k = m, l = n \). By Lemma 2.1, we know that the expected value is independent of the explicit values of \( k, l, m \) and \( n \). Therefore, we can apply Lemma 2.1 and bound the part stemming from the two Kronecker sequences by $CN^{3/2} |r|^{1/2}$. If we furthermore write out the functions $\mathcal{F}I(\cdot)$, which are both independent of \( k, l, m \) and \( n \), we obtain by Lemma 2.1 for the inner sum

\[
\left| CN^{3/2} |r|^{1/2} \mathcal{F}I \left( \frac{2sr \varepsilon}{N} \right) \mathcal{F}I \left( \frac{2sr' \varepsilon}{N} \right) \mathbb{E}[e(r\varepsilon(X_k - X_l) + r'\varepsilon(X_k - X_l))] \right|
\]

\[
= \left| CN^{3/2} |r|^{1/2} N^2 \frac{\sin \left( \frac{2\pi sr \varepsilon}{N} \right) \sin \left( \frac{2\pi sr' \varepsilon}{N} \right)}{2\pi^2 sr \varepsilon 2\pi^2 sr' \varepsilon} \frac{\sin \left( (r + r') \varepsilon \right)^2}{\pi^2 \varepsilon^2 (r + r')^2} \right|
\]

\[
\leq CN^{7/2} \frac{1}{|r|^1} \frac{1}{|r|^1} \frac{1}{|r + r'|^2}
\]
We can now apply at first Lemma 2.5 to bound the sum over $r$ by $\frac{C}{|r|^{1/2}}$ and then we can take the sum over $r'$. Including the factor $\frac{(2s)^2}{N^4}$ we achieve a bound of $CN^{-1/2}$ for the complete sum under consideration.

Next we turn to the case $k \neq m, n$ and $l \neq m, n$. Again by Lemma 2.1 the expected value is independent of $k, l, m, n$ and the sum reduces to

$$\frac{(2s)^2}{N^4} \sum_{r, r' \in \mathbb{Z} \setminus \{0\}} \sum_{r \neq \pm r'} \mathcal{F}(\frac{2sr \varepsilon}{N}) \mathcal{F}(\frac{2sr' \varepsilon}{N}) \frac{\sin(r \pi \varepsilon)^2 \sin(r' \pi \varepsilon)^2}{r^2 \varepsilon^2}$$

$$\times e(r(\{k\alpha\} - \{l\alpha\}) + r'(\{m\alpha\} - \{n\alpha\})).$$

In the inner sum, only the parts stemming from the Kronecker sequence depend on $k, l, m$ and $n$ and by Lemma 2.3 the entire product can be bounded by $C \cdot N^3 \cdot |r|^{1/2} |r'|^{1/2}$ because $\#\{k, l, m, n\} = 4$. Therefore,

$$\frac{(2s)^2}{N^4} \sum_{r, r' \in \mathbb{Z} \setminus \{0\}} \sum_{r \neq \pm r'} \mathcal{F}(\frac{2sr \varepsilon}{N}) \mathcal{F}(\frac{2sr' \varepsilon}{N}) \frac{\sin(r \pi \varepsilon)^2 \sin(r' \pi \varepsilon)^2}{r^2 \varepsilon^2}$$

$$\times e(r(\{k\alpha\} - \{l\alpha\}) + r'(\{m\alpha\} - \{n\alpha\}))$$

$$\leq C \cdot \frac{1}{N} \sum_{r, r' \in \mathbb{Z} \setminus \{0\}} \frac{1}{|r|^{3/2}} \frac{1}{|r'|^{3/2}}.$$

$$\leq \tilde{C} \zeta \left(\frac{3}{2}\right)^2 \frac{1}{N}$$

All the remaining cases in Lemma 2.1 can be treated by combining the arguments of the two cases which we discussed here in detail. Summing up, it follows that

$$\frac{(2s)^2}{N^4} \sum_{r, r' \in \mathbb{Z} \setminus \{0\}} \sum_{r \neq \pm r'} \mathcal{F}(\frac{2sr \varepsilon}{N}) \mathcal{F}(\frac{2sr' \varepsilon}{N})$$

$$\times \mathbb{E}[e(r \varepsilon(X_k - X_l) + r' \varepsilon(X_m - X_n))]$$

$$\times e(r(\{k\alpha\} - \{l\alpha\}) + r'(\{m\alpha\} - \{n\alpha\}))$$

$$= \mathcal{O}\left(\frac{1}{N^{1/2}}\right).$$

2. Case $r = \pm r'$:

If $r = r'$ and $k = m, l = n$ or if $r = -r'$ and $k = n, l = m$ respectively, the proof
of the convergence is verbatim the same as in [Mar07] and yields an order of convergence \( O \left( \frac{1}{N} \right) \).

If \( \#\{k, l, m, n\} = 4 \), then the proof from the case \( r \neq \pm r' \) can also be applied almost verbatim with the only difference that the double sum over \( r, r' \) reduces to a single sum. By the convergence of \( \sum_{r \neq 0} \frac{1}{|r|} \), also here an order of convergence \( O \left( \frac{1}{N} \right) \) can be achieved.

If we consider the case \( k = m, l \neq n \), then the expected value from Lemma 2.1 for \( r = -r' \) takes the form \( \frac{1}{N^{1/2}} \sin(r\pi \varepsilon)^2 \). Applying Lemma 2.3 we thus get an order of convergence \( O \left( \frac{1}{N^{1/2}} \right) \). For all other cases considered in Lemma 2.1, it follows by a similar argument that the sum has order of convergence \( O \left( \frac{1}{N^{1/2}} \right) \).

Combining the two cases \( r \neq r' \) and \( r = \pm r' \) puts us into the position to finally apply Chebyshev’s inequality and get for arbitrary \( \delta > 0 \) that
\[
\mathbb{P} \left( |F_N(s) - 2s| \geq \delta \right) \leq \frac{C}{\delta^2 N^{1/2}}
\]
with \( C \) independent of \( N \), i.e. convergence in probability of \( F_N(s) \) to \( 2s \). In order to get almost sure convergence, the remainder of the proof is analogously to the one of Theorem 1 in [HKL+19] and the presentation of the arguments is therefore omitted here.

References

[ALP18] C. Aistleitner, T. Lachmann, and F. Pausinger. Pair correlations and equidistribution. Journal of Number Theory, 182:206–220, 2018.

[Car12] I. Carbone. Discrepancy of LS-sequences of partitions and points. Annali di Matematics Pura ed Applicata, 191:819–844, 2012.

[EBMV15] D. El-Baz, J. Marklof, and I. Vinogradov. The two-point correlation function of the fractional parts of \( \sqrt{n} \) is Poisson. Proceeding of the AMS, 143 (7):2815–2828, 2015.

[GL17] S. Grepstad and G. Larcher. On pair correlation and discrepancy. Archiv der Mathematik, 109:143–149, 2017.
[HKL⁺19] A. Hinrichs, L. Kaltenböck, G. Larcher, W. Stockinger, and M. Ulrich. On a multi-dimensional Poissonian pair correlation concept and uniform distribution. *Monatshefte für Mathematik*, 190:333–352, 2019.

[HZ21] M. Hauke and A. Zafeiropoulos. Weak poissonian correlations. *arXiv:2112.11813*, 2021.

[KN74] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. John Wiley & Sons, New York, 1974.

[LS20] G. Larcher and W. Stockinger. Some negative results related to Poissonian pair correlation problems. *Discrete Mathematics*, 343(2), 2020.

[LST21] C. Lutsko, A. Sourmelidis, and N. Technau. Pair correlation of the fractional parts of $\alpha n^\theta$. *arXiv:2106.09800*, 2021.

[Mar07] J. Marklof. Distribution modulo one and Ratner’s theorem. In A. Granville and Z. Rudnick, editors, *Equidistribution in Number Theory, An Introduction. NATO Science Series*, volume 237. Springer, Dordrecht, 2007.

[Nie92] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. Number 63 in CBMS-NSF Series in Applied Mathematics, SIAM, Philadelphia, 1992.

[NP07] R. Nair and M. Pollicott. Pair correlations of sequences in higher dimensions. *Israel Journal of Mathematics*, 157:219–238, 2007.

[RS98] Z. Rudnick and P. Sarnak. The pair correlation function of fractional parts of polynomials. *Communication in Mathematical Physics*, 194:61–70, 1998.

[Sch72] W. M. Schmidt. Irregularities of distribution vii. *Acta Arithmetica*, 21:45–50, 1972.

[Ste20] S. Steinerberger. Poissonian pair correlation in higher dimension. *Journal of Number Theory*, 208:47–58, 2020.

[Wei19] C. Weiß. Interval exchange transformations and low-discrepancy. *Annali di Matematica Pura ed Applicata*, 198:399–410, 2019.

[Wei22] C. Weiß. Some connections between discrepancy, finite gap properties and pair correlations. *Monatshefte für Mathematik*, 199:909–927, 2022.

[Wey16] A. Weyl. Ueber die Gleichverteilung von Zahlen mod. Eins. *Mathematische Annalen*, 3:313–352, 1916.
[WS22] C. Weiß and T. Skill. Sequences with almost Poissonian pair correlations. *Journal of Number Theory*, 236:116–127, 2022.

**ROSENHEIM TECHNICAL UNIVERSITY OF APPLIED SCIENCES**, **HOCHSCHUL-STRASSE 1**, D-83024 Rosenheim, **anja.schmiedt@th-rosenheim.de**

**RUHR WEST UNIVERSITY OF APPLIED SCIENCES**, **DUISBURGER STR. 100**, D-45479 Mülheim an der Ruhr, **christian.weiss@hs-ruhrwest.de**